Received: 12 April 2017 | Revised: 5 September 2017 | Accepted: 8 September 2017

DOI: 10.1111/mec.14361

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Do genetic drift and accumulation of deleterious mutations preclude adaptation? Empirical investigation using RADseq in a northern lacustrine fish

Charles Perrier[1,2*] | Anne-Laure Ferchaud[1*] | Pascal Sirois[3] | Isabel Thibault[4] | Louis Bernatchez[1]

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

[2]Centre d'Ecologie Fonctionnelle et Evolutive, UMR 5175 Campus CNRS, Université de Montpellier, Montpellier Cedex 5, France

[3]Chaire de Recherche sur les Espèces Aquatiques Exploitées, Laboratoire des Sciences Aquatiques, Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada

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

Correspondence
Charles Perrier, Centre d'Ecologie Fonctionnelle et Evolutive, CEFE UMR 5175, Campus CNRS, Université de Montpellier, Montpellier Cedex 5, France.
Email: charles5perrier@gmail.com

Funding information
MFFP; Canadian Research Chair in Genomics and Conservation of Aquatic Resources; Natural Science and Engineering Research Council of Canada (NSERC)

Abstract

Understanding genomic signatures of divergent selection underlying long-term adaptation in populations located in heterogeneous environments is a key goal in evolutionary biology. In this study, we investigated neutral, adaptive and deleterious genetic variation using 7,192 SNPs in 31 Lake Trout (*Salvelinus namaycush*) populations (n = 673) from Québec, Canada. Average genetic diversity was low, weakly shared among lakes, and positively correlated with lake size, indicating a major role for genetic drift subsequent to lake isolation. Putatively deleterious mutations were on average at lower frequencies than the other SNPs, and their abundance relative to the entire polymorphism in each population was positively correlated with inbreeding, suggesting that the effectiveness of purifying selection was negatively correlated with inbreeding, as predicted from theory. Despite evidence for pronounced genetic drift and inbreeding, several outlier loci were associated with temperature and found in or close to genes with biologically relevant functions notably related to heat stress and immune responses. Outcomes of gene–temperature associations were influenced by the inclusion of the most inbred populations, in which allele frequencies deviated the most from model predictions. This result illustrates challenge in identifying gene–environment associations in cases of high genetic drift and restricted gene flow and suggests limited adaptation in populations experiencing higher inbreeding. We discuss the relevance of these findings for the conservation and management, notably regarding stocking and genetic rescue, of Lake Trout populations and other species inhabiting highly fragmented habitats.

KEYWORDS
adaptation, conservation genetics, fish, genomics, population genetics

1 | INTRODUCTION

Rapid climate change is one of the current main threats to biodiversity, particularly for ectothermic organisms (Deutsch et al., 2008), which are directly affected by the temperature of their environment. Fish are especially threatened by climate warming as they have to cope with both summer temperature rise and the consequential oxygen limitation (Pörtner & Knust, 2007), as well as the disruption of seasonal phenology (Bradshaw & Holzapfel, 2008; Farmer, Marschall, Dabrowski, & Ludsin, 2015). Thermal tolerance is partly genetically based in fish (e.g., in *Gasterosteus aculeatus*, McCairns & Bernatchez, 2012; McCairns, Smith, Sasaki, Bernatchez, & Beheregaray, 2016)

*Joint first authors.

and under directional selection (Leder et al., 2015). Moreover, physiological differences of tolerance have been observed for locally adapted populations along thermal gradients (Somero, 2010), suggesting potential for adaptive response to climate change. Identifying the genomic basis implicated in thermal tolerance is needed to evaluate the adaptive potential and evolutionary constraints of fish species in the face of climate change.

There is evidence for genomic footprints of spatially varying selection along temperature gradients (Gagnaire, Normandeau, Cote, Moller Hansen, & Bernatchez, 2012; Hornoy, Pavy, Gerardi, Beaulieu, & Bousquet, 2015; Kubota et al., 2015; Limborg et al., 2012). However, such evolutionary mechanisms may encounter many constraints, especially for a wide variety of species distributed in isolated populations. First, molecular rates of evolution through mutation may be too slow relative to the rate of climate change (but see DeLong et al., 2016). However, differences in evolution rates are known across genomic regions. In particular, duplicated genes and noncoding sequences near duplicated are known to evolve more rapidly than nonduplicated genes (Conant & Wolfe, 2008). That might confer an advantage to species with a numerous gene duplicates such as species who went through a whole-genome duplication (WGD, Lien et al., 2016; Waples et al., 2016). Second, assuming the presence of sufficient standing genetic variation upon which selection could act, potential for adaptation would largely depend on the respective effects of selection and genetic drift, and on the frequency of advantageous alleles, which are all strongly linked to effective population size (Olson-Manning, Wagner, & Mitchell-Olds, 2012). Crisci, Dean, and Ralph (2016) referred to this optimal combination of selection strength, starting allele frequency and age of the adaptive allele to a "Goldilocks zone" where adaptation is likely to occur. Third, connectivity and gene flow among populations are crucial to allow genetic introgression from better-adapted populations living in warmer environments into less adapted populations (Logan, Duryea, Molnar, Kessler, & Calsbeek, 2016; Somero, 2010) but also to maintain high levels of standing genetic variation. In turn, relatively high gene flow can also hamper local adaptation (Lenormand, 2002), especially at the edges of environmental gradients through continual flow of nonadaptive alleles from centred populations (Polechová & Barton, 2015). Lastly, genetic drift may not only limit the strength of positive selection but also lead to an increased frequency of weakly deleterious, nonsynonymous mutations partly as a consequence of higher linkage disequilibrium and reduced purifying selection (Lynch, Conery, & Bürger, 1995; Reed, Lowe, Briscoe, & Frankham, 2003; Renaut & Rieseberg, 2015; Romiguier et al., 2017). Overall, species structured in small isolated populations may be the most vulnerable in the face of rapid climate change due to highly restricted gene flow and to the relatively high influence of genetic drift that may both lead to reduced standing genetic variation (Brauer, Hammer, & Beheregaray, 2016), accumulation of weakly deleterious nonsynonymous mutations, and decreasing positive selection effect.

One of the major aims in population genomics is to understand how the spatial distribution and the characteristics of populations influence both neutral and adaptive genetic diversity and to identify genetic variants implicated in adaptation. The advances of massive parallel sequencing approaches now allow genotyping of thousands of SNPs (single nucleotide polymorphisms) that can be used to document both neutral and adaptive distribution of genetic diversity among populations distributed in heterogeneous landscapes and through the genome (Manel et al., 2015; Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013). To take advantage of such large data sets to infer footprints of selection and of local adaptation, genome scans using numerous genetic markers have become widely used (Beaumont & Balding, 2004; Excoffier, Hofer, & Foll, 2009b; Foll & Gaggiotti, 2008; Haasl & Payseur, 2015). Methods to test for associations between environmental variables and allele frequencies have also been developed (Günther & Coop, 2013; Frichot & François 2015; Rellstab, Gugerli, Eckert, Hancock, & Holderegger, 2015). However, relatively large numbers of populations coupled with replication tests are needed to ensure outlier analyses reliability (Hand et al., 2016). Using such genotyping approaches coupled with genome scans, several studies investigated gene–temperature associations in several taxa notably including salmonid fishes (Narum, Campbell, Kozfkay, & Meyer, 2010; Bourret, Dionne, Kent, Lien, & Bernatchez, 2013; Matala, Ackerman, Campbell, & Narum, 2014; see Bernatchez, 2016 for a review).

The genetic structure and the candidate genes implicated in adaptation to temperature can then be used for defining evolutionary significant units (Allendorf, Hohenlohe, & Luikart, 2010; Fraser & Bernatchez, 2001), and planning genetic rescue (Whiteley, Fitzpatrick, Funk, & Tallmon, 2015). Small and isolated populations face threats from genetic drift and inbreeding. Genetic rescue (i.e., an increase in population growth owing to the infusion of new alleles) can aid the persistence of small populations and presents a possible temporary solution, albeit contentious, for curtailing the loss of imperilled populations. Before considering conducting genetic rescue, several aspects should be considered (Hedrick & Frederickson 2010, Hedrick & Garcia-Dorado 2016). Some of the major aspects needing to be addressed are as follows: (i) the extent of connectivity among populations (gene flow), (ii) the level of fitness or inbreeding (from molecular data), (iii) the effective population sizes and (iv) the presence of deleterious mutations. The combination of those several aspects would inform on the potential level of genetic load. Moreover, it would be important to detect footprints of selection putatively different among populations (due to local adaptation) to avoid swamping of locally adapted genetic variation.

Lake Trout (*Salvelinus namaycush*, Walbaum 1792) is a freshwater fish native to North America widely distributed in cold freshwaters, from the Laurentian Great Lakes to the south and up to the Arctic, and from the Maritime provinces of Canada to the East all the way west to Alaska. Fish are sexually mature between 7 and 13 years, iteroparous and long living (up to 49 years) (Scott & Crossman, 1998). North American Lake Trout populations have recolonized their current geographic range between 6,000 and 15,000 YBP from several southern refugia (Wilson & Hebert, 1996, 1998). While some rare Lake Trout populations live in rivers (McCracken, Perry, Keefe,

& Ruzzante, 2013), and even more rarely make anadromous migrations (Harris, Moore, McDermid, & Swanson, 2014), most of the populations are lake specialists and have been isolated in lakes of various sizes shaped after the isostatic rebound and glacial scouring during the last glacial retreat (Wilson & Mandrak, 2003). As a result of population isolation with limited or no gene flow between them, pronounced genetic differentiation has been documented among populations and effective population size as well as allelic richness is small compared to other freshwater species (Halbisen & Wilson, 2009; Northrup, Connor, & Taylor, 2010; Valiquette, Perrier, Thibault, & Bernatchez, 2014). Relatively small effective population sizes in this species (e.g., Ne median value of 134 found in Valiquette et al., 2014) are also partly due to the fact that it is a top predator, and therefore expected to maintain smaller census size than prey and forage species. As water temperature is rising in North American lakes, it poses serious challenges to ensure the persistence of Lake Trout, often described as "glacial relicts" (Wilson & Mandrak, 2003). Most of the salmonid species display a genetic basis for differential responses to temperature (Eliason et al. 2011), as revealed notably by common garden experiments (reviewed in Hutchings, 2011), although this has yet to be tested in Lake Trout). Moreover, as for other salmonid species, Lake Trout has experienced at least four WGD events (Dehal & Boore, 2005; Jaillon, Aury, Brunet, Petit, & Stange-Thomann, 2004; Macqueen & Johnston, 2014), so that gene duplication and residual tetrasomic inheritance (from the tetraploid ancestry of salmonids) may have consequences for gene divergence, the mutation rate and therefore into the resilience of salmonids to inbreeding. However, the relatively small effective population sizes and population isolation in Lake Trout may limit their long-term persistence or the replenishment of standing genetic variation essential to evolutionary response to environmental changes.

The general aim of this study was to investigate the extent of neutral genetic structure and particular gene–temperature associations among many Lake Trout populations distributed across a wide geographic and latitudinal range. Specifically, we (i) investigated the distribution of genetic diversity across populations, (ii) assessed the proportion of deleterious mutations, (iii) estimated the influence of several environmental parameters, especially lake size, on the extent of neutral and deleterious genetic diversity and of genetic differentiation in each lake, (iv) searched for gene–temperature associations having significant biological implications for thermal adaptation in Lake Trout and (v) tested whether most inbred populations were likely to exhibit large deviation from predicted adaptive allele frequencies in gene–temperature associations.

2 | METHODS

2.1 | Sampling and study system

A total of 673 individuals sampled in 31 lakes from 2003 to 2013 in Québec, Canada, were successfully genotyped (Figure 1, Table 1). The map of the sampled lakes was drawn using the R-package GGMAP (Kahle & Wickham, 2013) with R 3.1.0 (R Core Team, 2016). The median number of individuals per lake was 19 (Table 1). Fish were sampled by technicians of the Ministère des Forêts, de la Faune et des Parcs du Québec (MFFP) using gillnets or by anglers using fishing rods. Adipose fin clips were preserved in 95% EtOH. Populations were chosen posterior to sampling according to their geographic location, the absence of stocking (Valiquette et al., 2014) 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 Trout most likely colonized these lakes from the single Atlantic refugium (Wilson & Hebert, 1996, 1998; Wilson & Mandrak, 2003). The size of the lakes sampled ranged from 36 to 207181 hectares, with a median value of 575 hectares (Table 1). Latitude ranged from 46.004 to 54.546 and longitude from −79.038 to −67.871. The average distance between lakes was 480 km, with a maximum of 984 km. Records of water temperature were unavailable for most of the lakes, and so records of air temperature through Quebec were used to estimate ten-year averages (2004–2014) of minimum, maximum, mean minimum, mean maximum and mean air temperature for each lake (Régnière, Saint-Amant, & Béchard, 2014). Indeed, growth rate in Lake Trout has been shown to be linearly correlated with air temperature (Black, Biela von, Zimmerman, & Brown, 2013). Moreover, in remote locations where the climate data not available (like in the Arctic), Lake Trout otoliths are used as indicators of past climate patterns (Torvinen 2017). Annual average temperature ranged from −4.98 to 4.13°C (Figure 1, Table 1).

2.2 | Molecular analyses

Genomic DNA was extracted from tissues using a modified version of a salt extraction protocol (Aljanabi & Martinez, 1997). An RNAse A (Qiagen) treatment was applied following the manufacturer's recommendation. DNA quality was checked using agarose gel electrophoresis. Quantity of DNA was evaluated using a NanoDrop spectrophotometer (Thermo Scientific) and then using Quant-iT Picogreen dsDNA Assay Kit (Invitrogen). Genotyping-by-sequencing libraries were prepared following a modified version of the two-enzyme GBS protocol, using *PstI* and *MspI* restriction enzymes (Poland, Brown, Sorrells, & Jannink, 2012). Forty-eight individuals were barcoded and pooled per library, and 96 barcode sequences of four to eight nucleotides were used. Single-end 100-bp length sequencing on Illumina HiSeq2000 platform was conducted at the Genome Quebec Innovation Centre (McGill University, Montreal, Canada). Real-time PCR was used to quantify DNA and adjust libraries loading.

2.3 | Bioinformatics

The program FASTQC 0.11.1 (Andrews, 2010) was used to inspect the quality of the sequence reads. Stacks pipeline version 1.21 (Catchen, Amores, Hohenlohe, Cresko, & Postlehwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) was used to demultiplex

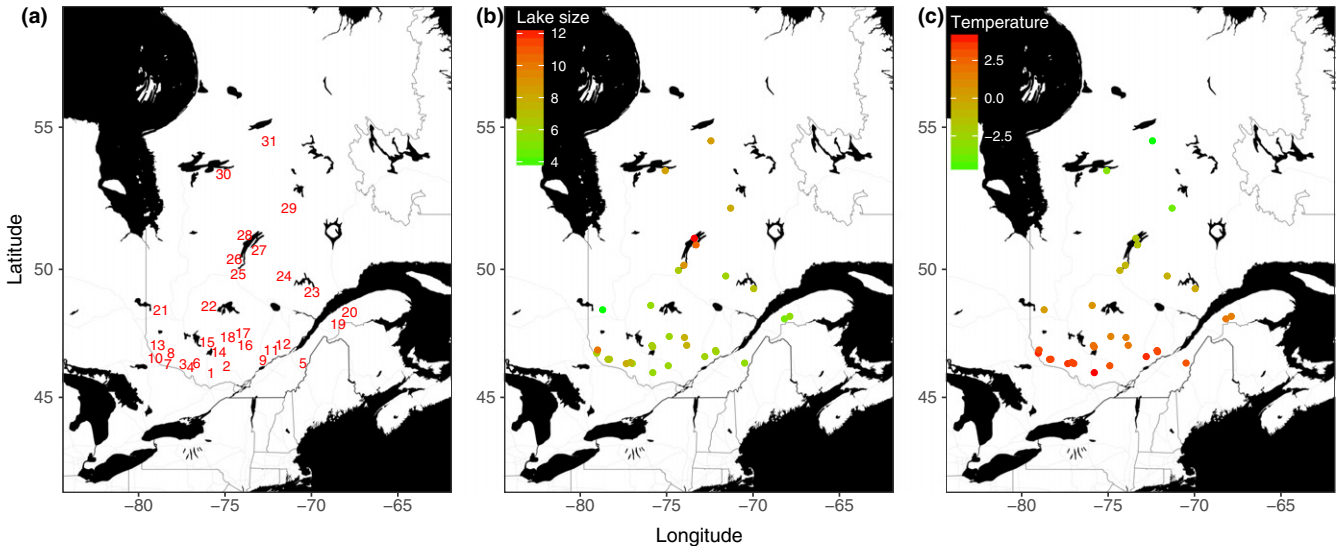

FIGURE 1 (a) Geographic location of the 31 lakes studied. (b) Log(lake size). (c) Average daily air temperature [Colour figure can be viewed at wileyonlinelibrary.com]

reads, discover loci and call genotypes. The libraries were demultiplexed using the *process_radtags* program. Reads were filtered for quality and ambiguous barcodes, allowing one mismatch in individual tags. Reads were trimmed to 80 bp as read quality tends to be lower at later positions and due to the occasional presence (~3%) of Illumina adaptors or fragments of adaptors at the end of the read (Pujolar et al., 2013). Stacks were created using *ustacks* allowing a maximum of two nucleotide mismatches ($M = 2$) among primary reads and four nucleotide mismatches ($n = 4$) among secondary reads. No polymorphisms were called if they were only present in the secondary reads. *M* and *n* values were identified as an optimum threshold according to the method developed by Ilut, Nydam, and Hare (2014). Given the potentially large proportion of private polymorphism (see Results), all individuals were used to form the catalog of loci, using *cstacks* with $n = 3$. The minimum stacks depth was set to five ($m = 5$). Sets of stacks were searched against the catalog using *sstacks*. *rxstacks* was used to correct genotypes, and *cstacks* and *sstacks* were run again with the correction. The stacks' module POPULATIONS (Stacks version 1.34) was then used to call genotypes. Several quality-filtering steps were implemented to remove polymorphisms with weak Stacks' genotype likelihood, with too few individuals or populations genotyped, deviating too much from Hardy–Weinberg equilibrium (due to genotyping errors or to the merging of paralogous loci), or being too rare. To be retained by POPULATIONS, a SNP had to be bi-allelic; genotyped in a least 60% of the individuals of a population for at least 24 of the 31 populations, had a minimum read depth of five and have a minimum genotype likelihood of −10. We kept only SNPs having a minimum average read depth of six across the entire data set. We retained SNPs with a minor allele frequency (MAF) >0.05 in at least one population. Lastly, we retained SNPs having an average heterozygosity inferior or equal to 0.50 and maximum heterozygosity inferior to 0.75; and an average F_{IS} between −0.05 and 0.3, a maximum F_{IS} superior to −0.05 and minimum F_{IS} between −0.03 and 0.03. Filtration and conversion of the VCF file was performed using R 3.1.0, PYTHON 2.7.6 (van Rossum & de Boer, 1991), VCFTOOLS 0.1.11 (Danecek et al., 2011), PLINK 1.07 (Purcell et al., 2007), and PGDSPIDER 2.0.7.2 (Lischer & Excoffier, 2012).

2.4 | Population diversity

We described genomic diversity within and among populations using several genomic indices. Nucleotide diversity "Π all loci" was estimated within each individual and population, on all sequenced nucleotides including variant and nonvariant sites, with the function *summary_haplotypes* in the stackr, R-package (Gosselin & Bernatchez, 2016). Nucleotide diversity among the loci listed in the catalog, "Π catalogued loci," was estimated using the POPULATIONS module of stacks. The proportion of SNPs genotyped and the ratio of polymorphic SNPs were reported for each population. In order to illustrate the limited proportion of polymorphisms shared among populations, we calculated the average number of polymorphic SNPs in 1000 random subsampling of 1–30 populations. Average inbreeding per population was estimated using the method of moments (Ritland, 1996) implemented in VCFTOOLS 0.1.11. Effective population size was estimated in NeEstimator 2.0.1 (Do et al., 2013) using the LDNe algorithm (Waples & Do, 2008) with a threshold of 0.05 as the lowest allele frequency considered. F_{st} among all population pairs was estimated using Genodive 3.0 (Meirmans & van Tienderen, 2004), which uses an analysis of molecular variance, performed between each pair of populations (Excoffier, Smouse, & Quattro, 1992; Michalakis & Excoffier, 1996). We reported an average value of pairwise F_{ST} for each population to document its average differentiation from all other populations, potentially reflecting the extent of genetic drift. The minor allele frequency spectra (mAFS) were plotted for each population in order to provide a visual illustration of the extent to which populations exhibit an L-shaped distribution as expected at a mutation–drift equilibrium (Nei & Li, 1976). Deviations from an

TABLE 1 Population numbering and corresponding lake name, sample size, latitude, longitude, lake size, and maximum, average and minimum air temperatures

PopID	Name	Sample size	Latitude	Longitude	Lake size	MaxT°C	MeanT°C	MinT°C
1	Vert	17	46.00367	−75.79911	300	33.64	4.13	−34.5
2	Desert	24	46.280601	−74.903241	323	32.91	3.08	−35.97
3	StPatrice	19	46.365556	−77.334444	2,841	33.65	3.63	−34.28
4	Antoine	17	46.37	−76.985556	437	33.52	3.33	−34.87
5	Etchemin	18	46.38861	−70.4875	247	31.04	3.06	−32.28
6	Lynch	23	46.411389	−77.095278	1,632	33.6	3.39	−34.78
7	Maganasipi	43	46.53417	−78.38972	939	32.83	3.26	−34.14
8	Caugnawana	22	46.53972	−78.3075	747	32.92	3.36	−34.18
9	Piles	18	46.64667	−72.79417	392	32.83	3.83	−34.71
10	Tee	22	46.78417	−79.03833	493	32.89	3.87	−31.72
11	Long	21	46.83778	−72.13833	306	32.88	3.38	−35.68
12	Montauban	16	46.88472	−72.16917	451	32.92	3.23	−35.82
13	Kipawa	20	46.90611	−78.99333	25,661	32.92	3.28	−34.08
14	Marguerite	22	47.02861	−75.80333	624	32.78	2.34	−38.06
15	Bondy	20	47.08389	−75.85222	526	32.46	2.18	−38.33
16	Devenyns	47	47.09028	−73.83722	2,141	32.1	1.9	−37.62
17	Mondanac	47	47.39917	−73.96528	2,328	31.99	1.56	−38.55
18	Turnbull	23	47.43889	−74.84667	342	31.55	1.37	−39.52
19	Cote	17	48.12222	−68.18139	121	31.37	1.92	−33.74
20	Chasseurs	14	48.22028	−67.87056	237	30.94	1.61	−33.9
21	Haie	20	48.47306	−78.70083	36	31.29	0.31	−41.02
22	Terrasses	13	48.63917	−75.91972	268	31.12	0.69	−39
23	SaultAuCochons	19	49.27556	−69.96167	1,002	33.44	−0.26	−41.98
24	Dulain	16	49.751744	−71.577779	376	30.6	−0.68	−41.99
25	Bourbeau	19	49.96	−74.317222	575	30.53	−0.77	−38.87
26	Wacanichi	15	50.15111	−73.99972	8,078	30.69	−0.91	−39.63
27	Albanel	15	50.9025	−73.295	40,083	30.14	−1.6	−41.23
28	Mistassini	38	51.141436	−73.396109	207,181	30.07	−1.78	−41.53
29	Pluto	13	52.22278	−71.29611	3,000	27.89	−3.88	−43.56
30	Guyer	18	53.531389	−75.080556	4,625	27.93	−3.34	−41.17
31	Oeufs	17	54.545278	−72.434444	5,110	25.97	−4.98	−43.17

L-shaped distribution would indicate a population bottleneck (Luikart, Allendorf, Cornuet, & Sherwin, 1998).

2.5 | Proportion of nonsynonymous polymorphisms and of deleterious mutations

To identify putative deleterious mutations, all defined loci (80 bp long) were first used in a BLAST query against the Rainbow Trout (*Oncorhynchus mykiss*) transcriptome (Berthelot et al., 2014) using blastx. All hits that had higher than 25 amino acid similarity and more than 90% similarity between the query read and the transcriptome sequence were retained. Both variants of each SNP in each query read were used with BLAST against the transcriptome, and translation results were compared pairwise. For each locus, results were only kept if the lowest e-value hit for both variants was the same (i.e., same protein name and length), and these were used to extract the protein sequence and identifier for the following step. Finally, PROVEAN was used (protein variation effect analyser; Choi, Sims, Murphy, Miller, & Chan, 2012) to predict the deleterious effect of nonsynonymous mutations with the default deleterious threshold value (−2.5). This program uses a versatile alignment-based score to predict the damaging effects of variations not limited to single amino substitutions but also in-frame insertions, deletions and multiple amino acid substitutions. This alignment-based score measures the change in sequence similarity of a query sequence to a protein sequence homolog before and after the introduction of an amino acid variation to the query sequence.

The average proportion of deleterious mutations (defined, for a given population and considering only the loci genotyped in that

population, as the number of loci showing a deleterious mutation over the number of loci identified as harbouring a deleterious mutation across all the populations) was compared to the proportion of the rest of the polymorphic SNPs between populations using *t* tests in R. Similarly, the average frequency of deleterious mutations was compared with the average minor allele frequency of the rest of the SNP, using *t* test in R. Finally, we estimated the ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs, per population. The MAF filtering obviously impacts the discovery and frequency of both putatively neutral and deleterious polymorphisms, but we do not expect that changing the MAF thresholds would affect these comparisons and subsequent relationships.

2.6 | Genetic differentiation among populations and assignment

The distribution of F_{ST} among all population pairs estimated using Genodive 3.0 was illustrated using a histogram. Genetic structure among populations was investigated with ADMIXTURE 1.23 program (Alexander, Novembre, & Lange, 2009) for *K* ranging from 2 to 40. A neighbour-joining tree based on Nei genetic distances among populations was built using the R-packages GGDENDRO (de Vries & Ripley, 2013), POPPR (Kamvar, Tabima, & Grünwald, 2014) and GGPLOT2 (Wickham, 2009). A DAPC implemented in the ADEGENET 2.0.1 R-package (Jombart & Ahmed, 2011) was used to estimate the genetic structure among populations. Following the optim.a.score function, we used 63 PCs and 4 DAs. The most likely number of genetic clusters among populations was inferred using the function *find.clusters* in ADEGENET 2.0.1 with a maximum of 40 clusters. The percentage of individuals assigned to their population of origin was estimated, based on the discriminant functions used in the DAPC, using as the input the original data set and then a data set in which all individuals were randomized throughout all populations.

2.7 | Relationships between environmental variables, population genetic diversity, population differentiation and proportion of deleterious mutations

Potential effects of lake size, geographic distance between lakes, latitude and longitude on parameters of genetic diversity and genetic differentiation were investigated using linear models in R. Correlations were determined between log(lake size) and three indices: Π catalogued loci, inbreeding and average pairwise F_{ST} per population. Isolation-by-distance signals were tested using a correlation between $F_{ST}/(1\text{-}F_{ST})$ and geographic distance and log(geographic distance) (Rousset, 1997). Effect of spatial heterogeneity in effective population sizes (SHNe) was investigated by testing a correlation between $F_{ST}/(1\text{-}F_{ST})$ and *di* of log(lake size), where *di* corresponds ((lake size 1+ lake size 2)/(lake size 1 × lake size 2)) and refers to the "*distance based on the inverse*" (Prunier, Dubut, Chikhi, & Blanchet, 2017). Using linear models in R, correlations between either average Π catalogued loci or average pairwise F_{ST} were also determined against latitude, and longitude. Correlations were also calculated between the proportion of deleterious mutations and average log(lake size) and inbreeding. Finally, the ratio of the proportion of deleterious mutations over the proportion of polymorphic sites among all the SNPs used in the study was correlated with the average log(lake size) and inbreeding.

2.8 | Detecting gene–environment associations

To identify SNPs for which the allele frequencies were correlated with the several temperatures recorded, logistic regressions were performed between the frequency of the minor allele and temperature (minimum temperature, mean minimum temperature, mean temperature, maximum temperature, and mean maximum temperature) using the GLM function implemented in R with a binomial variance and logit link. A linear model was not used as in cases of strong directional selection and absence of gene flow, and this can possibly lead to the fixation of alleles in the different environments, and therefore, one might rather expect a sigmoidal or a logistic response of allele frequencies along the environmental gradient (Rellstab et al., 2015). Because signal significance may sometimes be largely influenced by only one of the populations, each GLM was ran 100 times, randomly removing one of the populations at each iteration. Significant associations with environmental variables were considered when the SNPs for which more than 95% of the models had *p*-values $\leq$.05.

To test whether most inbred populations were likely to exhibit larger deviation from predicted adaptive allele frequencies in gene–temperature associations than less inbred ones, we investigated the relationship between average inbreeding per population and average absolute values of the GLMs' residuals for outlier loci in each population.

To detect the occurrence of false positives and false negatives in the outlier tests, the GLMs were repeated after excluding the five populations with the highest average inbreeding coefficients together with the highest absolute values of residuals over the GLMs for outlier loci. The rationale for removing these populations was that large average inbreeding coefficients are likely associated with pronounced genetic drift that may counteract the effect of selection. Such an enhanced effect of drift could thus reduce the associations between allele frequencies at catalogued loci and environment variables in these populations, resulting in false negative *via* a decrease of the fit of GLMs. Alternatively, the enhanced effect of drift could artificially create false positives by enhancing allele frequencies differences that randomly correlates with an environmental variable.

2.9 | Gene ontology

Nucleotide sequences containing deleterious mutations or SNPs for which allele frequencies were putatively correlated with temperature (as revealed using GLMs on 31 and 26 populations) were used in a BLAST query using *blastn* on the NCBI website (Altschul, Gish, Miller, Myers, & Lipman, 1990), considering salmonids sequences (and not only *Oncorhynchus mykiss*, to increase our gene ontology results). Only hits with an *E*-value $<10^{-10}$ were retained. For each

sequence, the hit with the smallest *E*-value was retained. If a sequence had several hits with similar top *E*-values, all hits were kept. Annotations were then recorded and discussed.

3 | RESULTS

3.1 | Population diversity, inbreeding and effective population size

The total number of reads retained was 2,315,769,589 with an average of 3,440,965 reads per individual. A total of 137,906 loci were collected in the catalog after *rxstacks* corrections. A total of 42,683 SNPs were outputted by POPULATIONS, 454 SNPs were removed as they had a minimum average read depth below 6, 27,224 SNPs were removed because they had too small MAF, 7,813 SNPs were removed as they did not meet heterozygosity or Fis criteria such that 7,192 SNPs in 5,738 loci were kept for population genomics analyses. Per population, the median value of SNPs genotyped was 6,832 (95%) and the median depth coverage was 11. The median value of Π all loci was 0.00021, and varied approximately fourfold among populations, ranging from 0.000125 (pop 24) to 0.000391 (pop 29) (Table 2). The median value of Π catalogued loci was 0.0279, ranging from 0.0125 (pop 24) to 0.0411 (pop 29) (Table 2). The cumulative number of polymorphic SNPs increased with the number of populations considered (Appendix S1). Depending on the population, 12%–56% of the SNPs were polymorphic, with a median value of 27% (Table 2; Appendix S1). Inbreeding was generally pronounced and ranged from 0.23 to 0.84 depending on the population, for a median of 0.62 (Table 2; Appendix S1). This inbreeding index was highly correlated to Π of catalogued loci (r^2 = .88, *p*-value: 6.57e-15). F therefore largely reflected the reduction in genetic diversity in each population. One explanation for overall high values of F is that we applied the test to the entire data set of SNP without filtering for MAF to remove weakly polymorphic loci in each population. Ne estimates were generally small, ranging from 1 to 255, with a median of 70 (Table 2). However, five "infinite" Ne values were estimated, most likely reflecting problems of trying to estimate Ne from linkage disequilibrium in these weakly polymorphic populations while using in the end relatively small numbers of bi-allelic markers. Allele frequency spectra (mAFS) varied markedly among populations (Appendix S1). Less than a third of the populations (2, 7, 11, 14, 16, 17, 19, 21, 24, 28 and 29) exhibited L-shaped mAFS consistent with historically large and stable effective population sizes. L-shaped distributions but with much fewer rare alleles were observed in four populations (6, 10, 26 and 27), suggesting recent past demographic events differing from large and stable effective population sizes. In the remaining populations, mAFS were drastically distorted, indicating recent bottlenecks or expansion.

3.2 | Proportion of nonsynonymous polymorphism and detection of deleterious mutations

Among the 5,736 genotyped loci (7,192 SNPs), 410 SNPs had significant BLAST results against the Rainbow Trout transcriptome and were retained to assess synonymy. Synonymous substitutions were identified for 203 SNPs and nonsynonymous for 207. Among the 207 nonsynonymous mutations, 68 were predicted to be neutral and 124 (60%) were putatively deleterious. The 13 remaining nonsynonymous mutations could not be attributed to one or the other category. The proportion of deleterious mutations varied from 0.10 to 0.43 among populations (Table 2; Appendix S1). On average, a smaller proportion of populations harboured polymorphism for deleterious sites compared to the entire SNP data set (t = −4.45, p = 1.38E-05, Appendix S1). Similarly, the average frequency of deleterious mutations was smaller than that of the MAFs for the entire data set of SNP (t = −5.42, p = 3162E-07, Appendix S1). These results suggest that deleterious mutations are more likely to be lost than putatively neutral polymorphisms but that there are still relatively common across all populations.

3.3 | Genetic differentiation among populations and assignment

All analyses confirmed a general pattern of genetic structure reflecting a situation of highly constrained gene flow and pronounced genetic drift. Pairwise population differentiation (F_{ST}) ranged from 0.03 to 0.78 (median = 0.46) (Appendix S1) whereas mean pairwise F_{ST} per population ranged from 0.25 to 0.61 (median = 0.46) (Table 2). Bayesian clustering analyses with the ADMIXTURE software for K = 25, 26, 31 showed that most populations could be assigned to a unique genetic cluster (Figure 2a). However, two population pairs were consistently assigned to the same genetic clusters (10 and 13 and 26 and 27) in accordance with their geographic proximities and possible physical connectivity (Valiquette et al., 2014). Additionally, two populations (seven and 17) were assigned to more than one genetic cluster, indicating putative within-lake genetic differentiation. The population tree illustrated the relatively large differentiation among populations but no clear geographic clustering (Figure 2b) except for lakes that clustered together in the admixture results. The DAPC (Appendix S1) confirmed the pronounced genetic structure among populations and the overall lack of geographic patterns, reflecting again the likely effect of pronounced genetic drift. The most likely number of genetic clusters inferred using the find.clusters function was 30. In accordance with the generally pronounced differentiation, the percentage of correctly assigned individuals to their population of origin was very high, with a mean value of 98.8% (Appendix S1), and much higher than the average of 11.6% obtained from a randomized data set.

3.4 | Relationship between environmental variables, population genetic diversity, population differentiation and proportion of deleterious mutations

A positive correlation was identified between average Π over all loci and log(lake size), (r^2 = .46, p = 2.65e-05, Figure 3a), as well as a negative correlation between average inbreeding and log(lake size),

TABLE 2 Genetic characteristics for each sample. Proportion of genotyped SNP; Proportion of polymorphic SNP, Pi over catalogued loci; Pi over all loci found in the catalog; average inbreeding coefficient; effective population size (Ne); Average F_{ST} of the considered population with the other 30 populations (F_{ST}); proportion of deleterious mutations and ratio of deleterious mutation over ratio of polymorphic SNP. Best demographic scenario inferred using dadi for each population. Time since bottleneck or expansion was reported for populations in which the best scenario was either bottleneck or expansion. Ancestral population size was reported for all populations

Pop ID	Genotyping ratio	Ratio of polymorphic SNP	Pi selected loci	Pi all loci	Inbreeding	Ne	F_{ST}	Ratio of deleterious mutations	Ratio of deleterious mutations over ratio of polymorphic SNP
1	0.93	0.19	1.91E-02	1.85E-04	0.66	94.5	0.50	0.10	0.52
2	0.92	0.30	2.28E-02	2.25E-04	0.55	249.4	0.43	0.25	0.85
3	0.92	0.24	2.11E-02	2.14E-04	0.62	INF	0.45	0.20	0.82
4	0.93	0.21	2.07E-02	2.00E-04	0.67	70.9	0.51	0.21	1.02
5	0.98	0.31	3.26E-02	3.01E-04	0.45	14.3	0.40	0.22	0.72
6	0.96	0.31	2.58E-02	2.39E-04	0.47	81.1	0.42	0.20	0.66
7	0.99	0.27	1.75E-02	1.64E-04	0.68	25.8	0.52	0.27	1.00
8	0.75	0.21	1.53E-02	1.66E-04	0.72	142.1	0.51	0.22	1.02
9	0.92	0.25	2.22E-02	2.20E-04	0.62	INF	0.49	0.22	0.85
10	0.94	0.35	2.80E-02	2.69E-04	0.41	32.3	0.38	0.26	0.74
11	0.98	0.22	1.90E-02	1.78E-04	0.64	73.4	0.51	0.14	0.63
12	0.82	0.19	1.67E-02	1.71E-04	0.72	INF	0.51	0.14	0.74
13	0.75	0.32	2.62E-02	2.67E-04	0.53	273.1	0.38	0.22	0.68
14	0.97	0.32	2.75E-02	2.56E-04	0.44	254.5	0.39	0.24	0.75
15	0.95	0.23	1.98E-02	1.90E-04	0.64	75.9	0.50	0.21	0.91
16	0.97	0.34	2.11E-02	2.02E-04	0.59	88.6	0.46	0.43	1.26
17	0.97	0.31	2.21E-02	2.12E-04	0.68	2.3	0.44	0.26	0.84
18	0.85	0.23	1.69E-02	1.72E-04	0.72	41.3	0.53	0.19	0.82
19	0.95	0.25	2.13E-02	1.99E-04	0.84	0.4	0.61	0.22	0.87
20	0.82	0.16	1.88E-02	1.85E-04	0.76	183.8	0.58	0.20	1.28
21	0.94	0.21	1.76E-02	1.70E-04	0.82	0.5	0.60	0.16	0.75
22	0.72	0.20	2.17E-02	2.18E-04	0.69	40.5	0.53	0.15	0.78
23	0.89	0.13	1.25E-02	1.25E-04	0.81	27	0.59	0.17	1.34
24	0.96	0.12	1.35E-02	1.29E-04	0.83	50.7	0.60	0.21	1.73
25	0.98	0.27	2.84E-02	2.60E-04	0.43	60.8	0.45	0.16	0.59
26	0.98	0.40	4.00E-02	3.64E-04	0.23	INF	0.32	0.23	0.56
27	0.98	0.40	3.82E-02	3.49E-04	0.30	69	0.32	0.20	0.51
28	0.98	0.56	4.11E-02	3.91E-04	0.24	205.5	0.25	0.39	0.70
29	0.98	0.30	3.18E-02	2.92E-04	0.41	109.8	0.39	0.21	0.70
30	0.95	0.39	3.62E-02	3.39E-04	0.38	37.9	0.37	0.29	0.75
31	0.90	0.34	3.20E-02	3.10E-04	0.38	INF	0.38	0.16	0.46

(r^2 = .50, p = 9.62e-06, Figure 3b). A positive correlation was identified between the proportion of polymorphic SNPs per population and log(lake size) (r^2 = .57, p = 1.03E-06, Figure 3c). Similarly, a positive correlation was identified between the proportion of deleterious mutations per population and log(lake size) (r^2 = .22, p = 8.00E-03, Figure 3d) and a negative correlation with inbreeding (r^2 = .11, p = 7.00E-02, Figure 3e), suggesting a main effect of drift in driving both neutral and deleterious diversity. Nevertheless, the ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs per population was not significantly correlated with log (lake size) (r^2 = .08, p = 1.30E-01, Figure 3f) but was positively correlated with inbreeding (r^2 = .37, p = 2.61E-04, Figure 3g), suggesting that deleterious mutations are lost faster in less inbred lakes than in more inbred ones. Isolation by distance (IBD) was extremely small (Figure 3h, r^2 = .00 or 0.02, p = 0.34 or 4.03e-06, when considering either geographic distance or log(geographic distance) as explanatory variable, respectively), reflecting a highly constrained contemporary gene flow. In turn, pairwise F_{ST} between populations were positively correlated with di of log(lake size) (Figure 3i, r^2 = .50, p < 2.2e-16). Moreover, a negative correlation was also observed between log(lake size) and average pairwise F_{ST} per population (r^2 = .63, p = 9.74e-08, Figure 3j), highlighting a pronounced

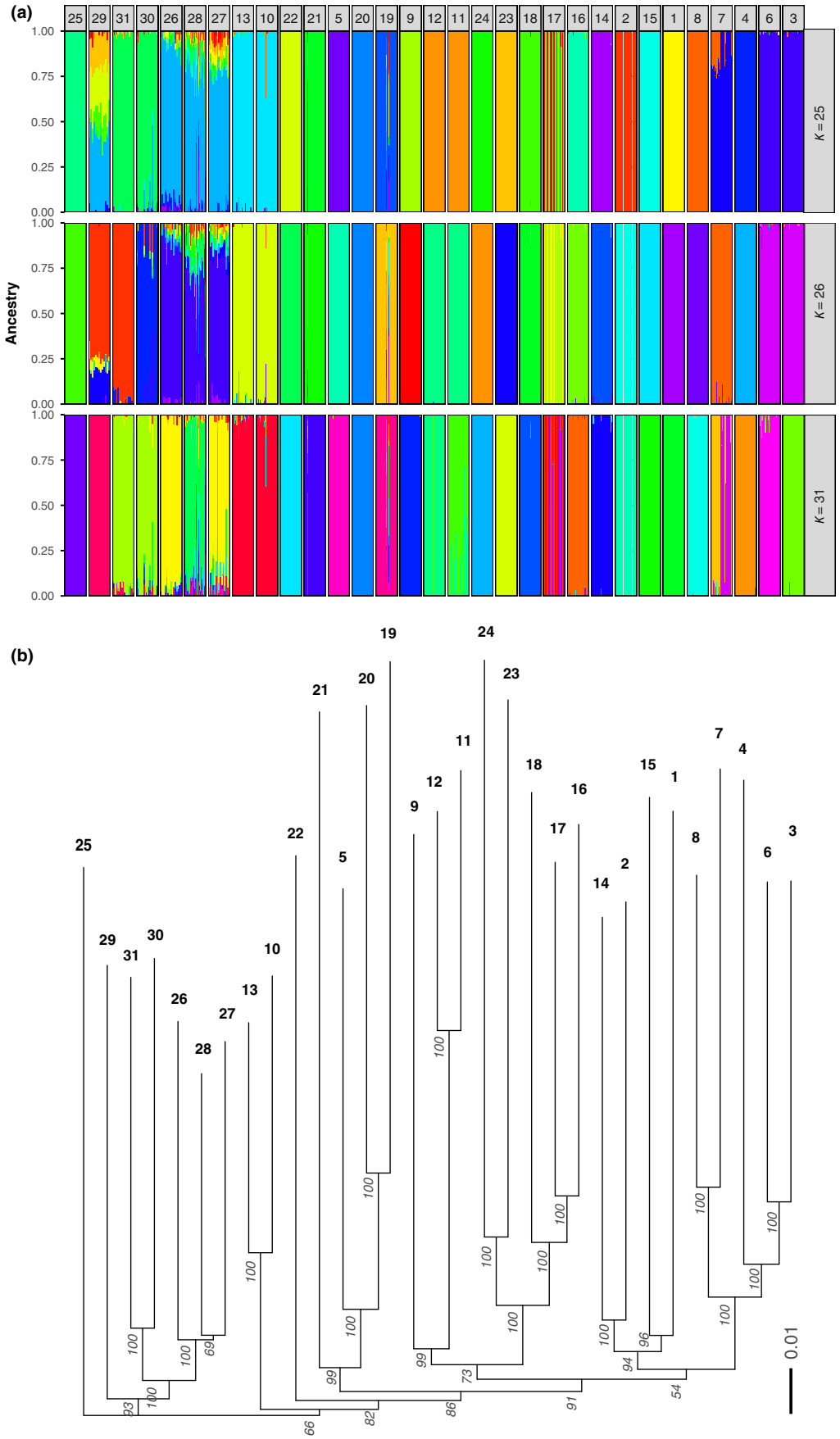

FIGURE 2 (a) Admixture results for k = 25, 26, 31 (populations ordered as in the tree below). (b) Neighbour-joining tree of Nei genetic distances between populations [Colour figure can be viewed at wileyonlinelibrary.com]

genetic drift negatively correlated with lake size, a proxy of effective population size. No significant correlation was found between longitude and either Π, inbreeding and F_{ST} (r^2 = 1.77E-05, p = 9.82E-01; r^2 = 1.27E-02, p = 5.46E-01; r^2 = .04, p = 3.14E-01) Π, inbreeding and F_{ST} were weakly influenced by latitude (r^2 = .31, p = 1.21E-03; r^2 = .21, p = 9.51E-03; r^2 = .13, p = 4.74E-02).

3.5 | Markers under positive selection

Using GLMs on the 31 populations, we found that among the 7,192 SNPs, 44 SNPs (in 42 loci) showed significant correlation between frequency of the minor allele and temperature indices (Figure 4—upper panel; Appendix S1). Twenty-three SNPs (in 22 loci) were associated with minimum temperature, 21 with mean minimum temperature, 21 with mean temperature, 26 with mean maximum temperature, 14 with maximum temperature and two were shared among all of the five temperature indices (data not shown).

A positive significant correlation was found between the average value of the residuals from all putatively outlier SNPs in the GLMs with all of the five temperatures and the average individual inbreeding coefficient (r^2 = .26 p = 3.52e-03, Figure 5). This result suggests that the most inbred populations are those that are most likely to depart from a predicted association between "adaptive" allele frequencies and temperature, and therefore being less adapted to the local thermal regime.

Populations 20, 21, 22, 24 and 25, which exhibited the highest levels of inbreeding and absolute values of residuals in GLM models for outlier SNPs (Figure 5) were then removed from the data set before applying again GLM association tests. For the 26 remaining populations, 23 of the 44 formerly identified SNPs were identified as associated with temperature (Figure 4—lower panel). No additional SNPs were found associated in this new analysis.

3.6 | Gene ontology

Among the 42 loci containing the 44 SNPs putatively associated with temperature identified using GLMs, 21 loci had significant hits referring to usable annotated sequences in other salmonids (Atlantic Salmon (*Salmo salar*), Lake Whitefish (*Coregonus clupeaformis*), Chinook Salmon (*Oncorhynchus tshawytscha*), Whitespotted Charr (*Salvelinus leucomaenis*), and Rainbow Trout (*Oncorhynchus mykiss*)). Nine of these 42 loci corresponded to putative synonymous mutations, six were putatively nonsynonymous including a single deleterious mutation, and the remaining outliers were not determinable in terms of types of mutations (Appendix S2). When excluding putative false positives and false negatives among those outlier loci (i.e., detected outliers only for GLM31 in Appendix S2), 22 loci had significant hits. Five of these corresponded to synonymous mutations, three were nonsynonymous in which none was detected as deleterious. DNA sequences identified by the analysis included genes representing multiple biological functions including genes coding for protein or transcription factors mainly related to immune system genes, but also a variety of other processes including notably response to heat shock (three hits for *hsp*70), sex determination, growth factor and oogenesis (Appendix S2).

Among the 124 deleterious mutations, 101 loci had significant hits against the best annotated transcriptome in salmonids (Rainbow Trout, Berthelot et al., 2014). Among relevant biological functions represented among those hits, 11 genes refer to immune response, five to neurogenesis, one to memory and learning ability, one to eye morphogenesis and one to heat shock response (Appendix S2).

4 | DISCUSSION

Using genetic inferences based on genotyping-by-sequencing data, we investigated the extent of neutral genetic structure, gene–temperature associations and the extent of deleterious variation among 31 Lake Trout populations distributed across a broad geographic and latitudinal range from Québec, Canada. Genetic diversity was generally low, differentiation among lakes was very pronounced and the extent of both genetic diversity and differentiation were largely influenced by lake size, with smaller diversity and higher divergence in smaller lakes, indicating a major role for genetic drift in shaping the distribution of genetic diversity. These results therefore suggest decreased standing genetic variation and possibly adaptive potential for a changing environment in the smaller populations. Nevertheless, gene–temperature associations were found for biologically relevant functions notably related to heat stress and immune responses. Some outcomes of gene–temperature associations were influenced by the inclusion of populations with the highest inbreeding coefficients, illustrating challenges in finding gene–environment associations in cases of high genetic drift and weak gene flow, as well as suggesting limitations for adaptations in small populations. Moreover, approximately equal numbers of nonsynonymous and synonymous mutations were identified among genes that could be annotated, and 60% of the nonsynonymous mutations were predicted to be deleterious. Although the proportion of deleterious mutations was negatively correlated with inbreeding, these deleterious mutations were on average at smaller frequencies than the rest of the SNPs within a population, and their proportion in each population relative to the proportion of all polymorphic loci were positively correlated with inbreeding. Both of these results suggest a positive effect of population size on the efficacy of purifying selection. Although the large genetic differentiation among populations and the evidence for local adaptation calls for a management at the population level, reduced genetic diversity and potentially reduced adaptive potential in smaller populations may indicate that genetic rescue could be beneficial based on the identification of adaptive polymorphisms in the context of rapid environmental changes. Therefore, we discuss the relevance of these findings for the conservation and management of Lake Trout populations in particular, but also for any highly fragmented species, notably regarding supplementation and/or genetic rescue.

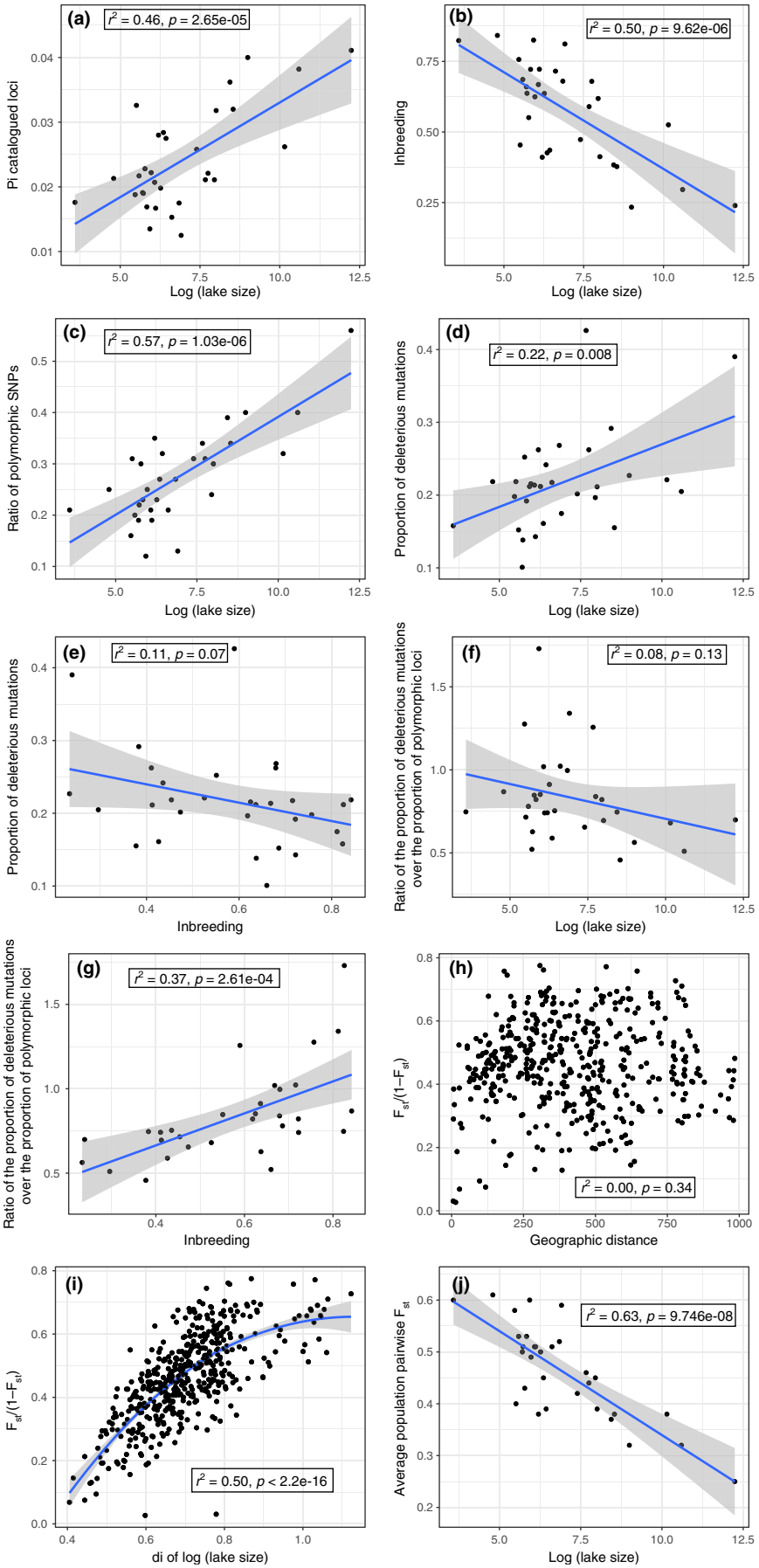

FIGURE 3 Relationships among genomic indices and environmental parameters. Relationship between (a) nucleotide diversity at catalogued loci and log(lake size), (b) inbreeding and log(lake size), (c) ratio of polymorphic SNPs and log (lake size), (d) proportion of deleterious mutations and log(lake size), (e) proportion of deleterious mutations and inbreeding, (f) ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs and log(lake size), (g) ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs and inbreeding, (h) F_{ST} / (1-F_{ST}) and geographic distance,(i) F_{ST} /(1-F_{ST}) and *di* of log(lake size), (j) Average population pairwise F_{ST} and log(lake size). R^2 and associated *p*-values are given for the linear regressions [Colour figure can be viewed at wileyonlinelibrary.com]

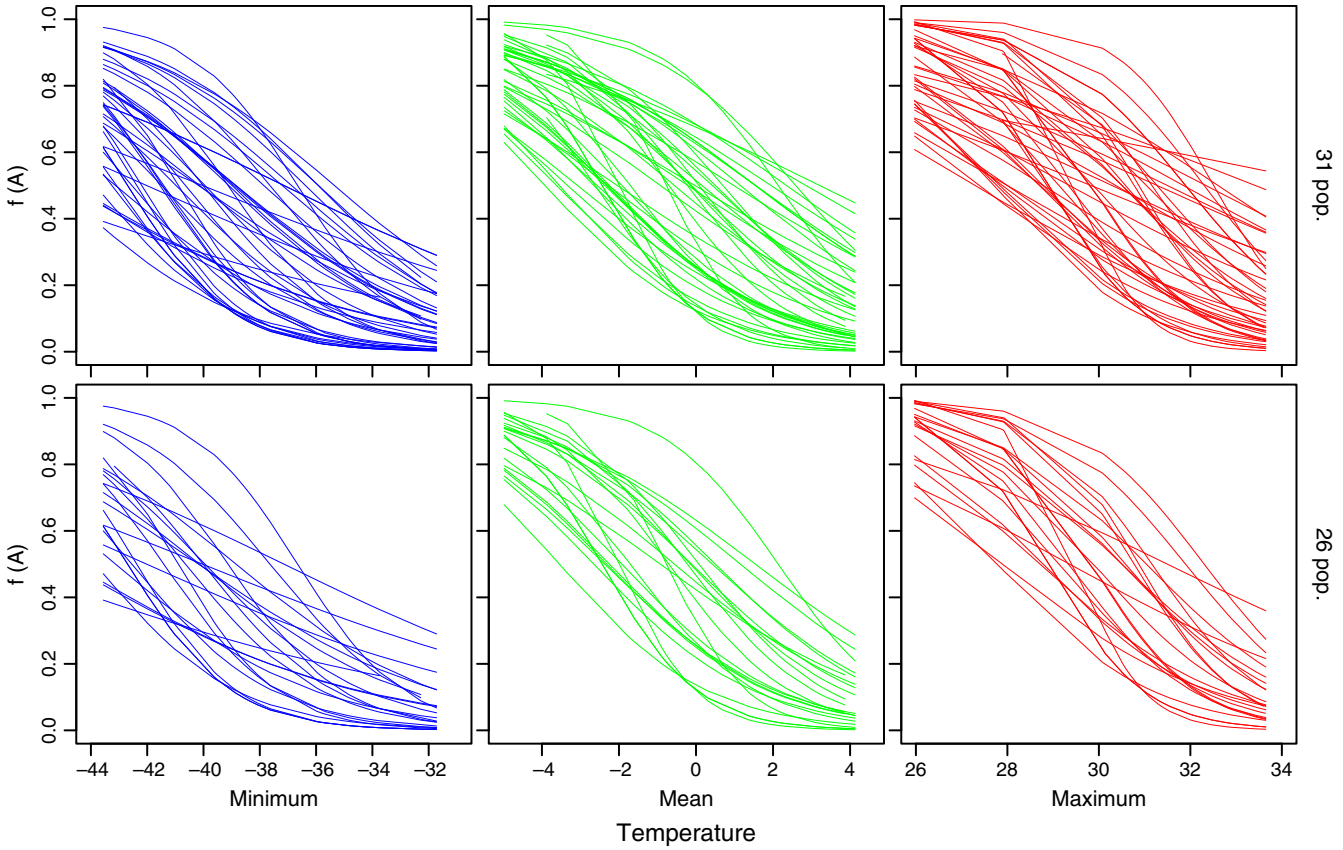

FIGURE 4 Fitted allele frequencies as function of minimum, mean and maximum temperature for the 44 and 23 outlier loci found using GLMs based on all of the 31 populations (upper panel) or 26 less inbred populations (lower panel) [Colour figure can be viewed at wileyonlinelibrary.com]

4.1 | Dramatically reduced genetic diversity and accumulation of deleterious mutations in Lake Trout populations

Low genetic diversity is in line with the fact that Lake Trout has colonized present-day northern freshwater ecosystems following the last glacial retreat and since then have been predominantly isolated in lakes formed by glacial scoring and isostatic rebound (Wilson & Mandrak, 2003). Such demographic history may have frequently implicated founder effects and bottlenecks, most probably resulting in enhanced genetic drift and reduction in genetic diversity. Similarly to several previous studies (Halbisen & Wilson, 2009; Northrup et al., 2010; Valiquette et al., 2014), we found that genetic diversity was highly positively correlated with lake size, indicating that lake size may modulate effective population sizes and consequently the extent of genetic drift and ultimately genetic diversity. As also reported by Valiquette et al. (2014), we found a very weak signal of isolation by distance, suggesting the absence of a significant effect of gene flow among lakes in shaping the distribution of genetic diversity among these populations. This absence of isolation by distance was also supported by the almost absence of an effect of latitude and longitude on neutral genetic diversity and differentiation. This is coherent with the proposed patterns of rapid recolonization of the Northern Hemisphere, and subsequent isolation and by the almost null contemporary migration among lakes of this species (Wilson & Mandrak, 2003).

The level of deleterious mutations in these Lake Trout populations is in line with theoretical studies on such processes in populations located in an expansion front (Agrawal & Whitlock, 2012; Balick et al., 2014; Peischl et al., 2016) and in small or domesticated populations (Caballero, Bravo, & Wang, 2017; Renaut & Rieseberg, 2015). Most of the deleterious mutations are common to several lakes, illustrating that they have a common origin preceding the isolation of populations in different lakes. Indeed, the expansion load model predicts that private de novo mutations arising during the phase of population expansion itself are often more deleterious than rare genetic variants already present in the source population (Peischl, Dupanloup, Kirkpatrick, & Excoffier, 2013). Moreover, linkage disequilibria between desirable beneficial and unwanted deleterious mutations may hinder the ability of selection to efficiently fix beneficial mutations, while weeding out deleterious ones. This occurs because of Hill–Robertson effect (Hill & Robertson, 1966); selection acts on the net effect of both beneficial and deleterious mutations for a given genotype (Hill & Robertson, 1966). A corollary of this effect is that selection against deleterious mutations will be less effective in regions of the genome with reduced levels of recombination (Charlesworth, Morgan, & Charlesworth, 1993), leading to a predicted enrichment in deleterious mutations in these regions. Moreover, as a salmonids species, Lake Trout went through a relatively recent WGD (Macqueen & Johnston, 2014) that might have contributed to the observed accumulation of deleterious mutations

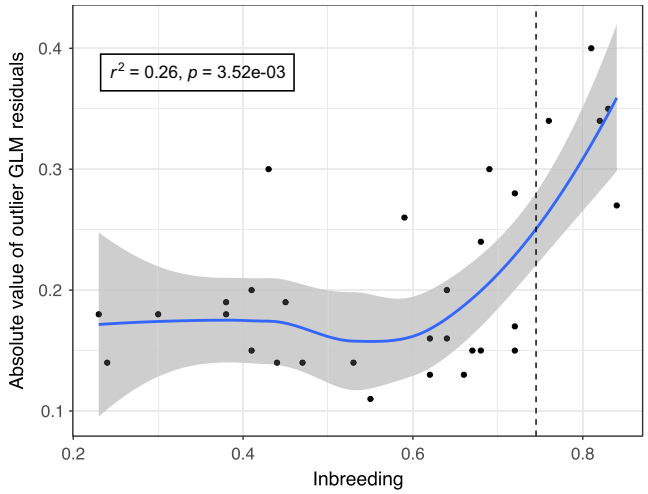

FIGURE 5 Average values of the residuals from all putatively outlier SNPs in the GLMs with temperature, in relation to the average individual inbreeding values, for each population. The dashed vertical line separates the populations kept and removed from subsequent genome scans analyses, based on their inbreeding value [Colour figure can be viewed at wileyonlinelibrary.com]

caused by relaxed purifying selection on the becoming nonfunctionalized copy of the genes. Therefore, the gene duplication and residuals tetrasomic inheritance might have also consequences for the resilience of salmonids to inbreeding. Further, limited effective population size in small populations would increase the intensity of linkage disequilibrium, thus enhancing the segregation of deleterious alleles by hitchhiking. Concordantly, in the current study, the proportion of deleterious mutations was negatively correlated with inbreeding and lake size, suggesting that drift and inbreeding (mainly caused by limited effective size) are partly responsible for the retention of deleterious mutations, as for common SNPs (Whitlock & Bürger, 2004). However, deleterious mutations were on average at lower frequencies than the rest of the SNPs, suggesting that the efficacy of purifying selection increased with population size (Lynch et al., 1995; Reed et al., 2003). Moreover, the proportion of deleterious mutations in each population, relative to the level of the entire polymorphism characterized, was positively correlated with inbreeding, again suggesting a positive influence of population size and a negative influence of inbreeding on the effectiveness of purifying selection. However, various gene functions notably related to immune system and HSP were found among deleterious mutations, suggesting that some populations may possess reduced putatively adaptive standing genetic variation that may ultimately result in decreased local adaptation and adaptive potential.

4.2 | Gene–temperature association

Theory predicts that the power to detect adaptation is maximized at intermediate values of selection strength (starting allele frequency and age of adaptive alleles (Crisci et al., 2016). In our study, detecting adaptation could have been facilitated by the relatively recent isolation of populations and the potentially strong selection differential across the temperature gradient. However, the genetic drift that may have occurred after the founding and the isolation of the lakes and subsequently could have both facilitated and limited adaptation, and thus affects the potential to detect adaptive signals. On the one hand, because of genetic drift, adaptation is expected to produce soft selective sweeps, where multiple alleles at the same locus rise in frequency simultaneously, facilitating adaptation and their detection in genome scans (Messer & Petrov, 2013). On the other hand, genetic drift could also reduce the potential for adaptation if its strength is high relative to selection, and if adaptive standing genetic variation is reduced either through fixation or deleterious mutation accumulation, in particular in smaller populations (*as discussed in the next section*). Here, we identified several outlier loci that may be implicated in response to selection by temperature, concordantly with other studies conducted on marine and catadromous organisms (Atlantic herring (*Clupea harengus*): Limborg et al., 2012; Guo et al., 2015; European Eel (*Anguilla anguilla)*: Pujolar et al., 2014; American Eel (*A. rostrata*): Gagnaire et al., 2012), anadromous species (Atlantic salmon (*Salmo salar*): Bourret et al., 2013; Moore et al., 2014; Chinook Salmon (*Oncorhynchus tshawtscha*): Hecht, Matala, Hess, & Narum, 2015), as well as freshwater fish (Rainbow Trout (*O. mykiss*): Hand et al., 2016). Of particular interest is the study of Bay & Palumbi, 2014; genotyped over 15,000 SNPs in corals (*Acropora hyacinthus*) within a natural temperature range, and suggested that natural populations harbour a reservoir of alleles pre-adapted to high-temperatures, suggesting evolutionary potential to respond to future climate change. Our empirical population genomic study on small isolated populations strengthens the argument that the potential for response to selection at small population sizes might be more extensive than previously assumed in evolutionary and conservation biology, and may be constrained only in the most extreme cases of small population sizes. Indeed, in a meta-analysis, Wood et al. (2016) investigated the link between natural selection, quantitative genetic variation and population size across a large number of populations and species in nature and found that heritability does not decrease with population size unless it is extremely low (Ne <100). Moreover, Kavanagh, Haugen, Gregersen, Jernvall, and Vøllestad (2010) provided evidence that natural selection is sufficiently powerful for temperate lake fish populations to adapt to novel temperature regimes within few generations, even under conditions with low genetic variation and under an influence of gene flow.

Furthermore, gene duplication may provide raw material for functional innovation (Conant & Wolfe, 2008) and rapid evolution (Kostka, Hahn, & Pollard, 2010). In Atlantic Salmon, Yasuike et al. (2010) indicated the duplication of the IgH (Immunoglobulin heavy chain) locus significantly contributes to the increased diversity of the antibody repertoire, as compared with the single IgH locus in other vertebrates. Furthermore, recent studies provided new insights into the architecture of duplicated loci in *Salvelinus* (Sutherland et al., 2016; Nugent, Easton, Norman, Ferguson, & Danzmann, 2017). All salmonids species share substantial regions of duplication in the distal ends of seven or eight pairs of chromosomes (Lien et al., 2016;

Sutherland et al., 2016; Waples et al. 2016) and several studies in other taxa (mainly human) revealed that putative adaptively important gene families can accumulate in the distal ends of chromosomes (Bethke et al., 2006; Fan et al., 2008; Linardopoulou et al., 2005). Unfortunately, there is currently a lack of hard empirical data to support the direct link between duplication and fitness, particularly in salmonids.

Admittedly, it is not possible here to rule out the possibility that some of outlier SNPs identified here resulted from allele surfing (i.e., process that enables mutation to establish at expansion range and reach high frequencies (Edmonds, Lillie, & Cavalli-Sforza, 2004). In turn, we identified putatively false-positives outlier SNPs detected only when including all populations but not when populations with the largest average inbreeding coefficient were removed. We hypothesize that these outliers resulted from random shifts in allele frequencies in the most inbred populations in which genetic drift was more pronounced. Meanwhile, while leaving out the most inbred populations, we found several new potential outliers, suggesting that including populations in which the action of selection may be limited due to higher drift may limit the associations between allele frequencies at selected loci and environment variables (Blanquart, Gandon, & Nuismer, 2012; Crisci et al., 2016).

4.3 | Biological functions of outliers associated with temperature

Several of the biological functions identified for the outlier loci could be relevant in the context of adaptation to temperature and in particular HSP and immune system genes. Among the three outlier SNPs found in a gene coding for *heat shock protein HSP70*, two were nonsynonymous mutations and one was detected as an outlier either based on 31 or 26 populations and consequently was a top candidate for temperature adaptation. If this SNP has the potential to modify conformation of the HSP70, or the gene expression of *hsp70*, it could have implications for adaptation to temperature. Heat shock protein activities are known to be involved in responses to various abiotic stressors, including temperature variations and elevated levels of toxic substances. In most experiments, an increase in HSP70 expression is the most prominent response at the effect of heat shock on HSPs (Iwama, Thomas, Forsyth, & Vijayan, 1998). For example, increased levels of HSP70 were observed in various tissues of Cutthroat Trout (*O. clarki*) exposed to heat shock (Mazur, 1996). Such HSP70 increases after heat shock were also observed in Rainbow Trout (Currie & Tufts, 1997). Furthermore, sequence variation in HSP70 genes within and among *Oncorhynchus* species have been detected, notably in cis-regulatory regions, and may influence HSP70 transcription rate and performance (Narum & Campbell, 2010). Hemmer-Hansen, Nielsen, Frydenberg, and Loeschcke (2007) documented pronounced differential gene expression in *heat shock cognate 70* (a member of the HSP70 gene family) between geographically close Baltic sea and North sea populations of a marine fish, highlighting adaptive divergence and local adaptation. For the three HSP70 outlier SNPs, most of the populations which present the highest residuals values from GLM tests (i.e., potentially maladapted for the hsp70 variation) correspond to small lakes with high inbreeding estimates (population 19, 20, 21, 23, 24), and the northernmost population (31) presents the highest residual value for one of those three SNP (SNP182512). Based on these results, we suggest that future works should aim to investigate the genetic polymorphism in HSP70 and other HSPs in these Lake Trout populations along a temperature gradient, possibly coupled with the estimation of the related expression levels and performance of the HSPs, that could further unravel the genetic bases of adaptation to temperature by HSPs in these Lake Trout populations.

More than a third of annotations obtained were related to immune system genes., notably including MHC (major histocompatibility complex) class 1/2, IgH (Immunoglobulin) locus A/B, TCR-alpha/delta/gamma, interferon alpha 1-like, and CRFB (corticotrophin-releasing factor-binding). This result may not be surprising given the relatively large diversity at these genes, and given that they are under strong selective pressures. In particular, MHC class 2 B locus is known to be highly polymorphic in Lake Trout (Dorschner, Duris, Bronte, Curtis, & Phillips, 2000). In Atlantic salmon, Dionne, Miller, Dodson, Caron, and Bernatchez (2007) showed that MHC class 2 allelic diversity increases with temperature, suggesting adaptation to pathogen load that is correlated with temperature (Dionne, Miller, Dodson, & Bernatchez, 2009).

4.4 | Implications for the definition of management units and genetic rescue

An important aim of conservation genetics is to help defining evolutionary significant units (ESU) and management units (MU) (Crandall, Bininda-Emonds, Mace, & Wayne, 2000; Funk, McKay, Hohenlohe, & Allendorf, 2012; Moritz, 1994). While all lakes studied here most likely belong to the same glacial lineage (Wilson & Mandrak, 2003) and should therefore be considered as a single ESU, the results from this study (notably the high differentiation between lakes), in line with Valiquette et al. (2014), confirm that almost all the studied Lake Trout populations living in isolated lakes can be considered as divergent populations and should be therefore treated as different MUs. Moreover, no clear geographical organization of the genetic variation was found, meaning that the geographic closeness among populations could not be used as an indicator of genetic similarity among populations. In cases where populations are weakly differentiated at neutral markers and where ESUs and MUs could omit local adaptations, outlier loci can be used to define a higher number of groups of fewer populations each (Funk et al., 2012). In contrast, in species with small isolated and genetically differentiated populations (as for Lake Trout), ESUs and MUs based on neutral markers may be too numerous as almost every population (and even subpopulations) could be considered as unique. For such species, outlier loci should be used to regroup populations according to allele frequencies (Funk et al., 2012). Some lakes even harbour sympatric, genetically differentiated populations (lakes 1, 4, 7, 16 and 17), some of them corresponding to distinct planktonic and piscivorous ecotypes (Bernatchez,

Laporte, Perrier, Sirois, & Bernatchez, 2016). Overall, the management and conservation of Lake Trout populations would benefit from more consideration of both genetic and phenotypic variations, both occurring at a relatively small geographic scale, which is smaller than the one currently generally considered for management.

Although we found outlier loci potentially implicated in adaptation to temperature, bottlenecks and genetic drift may have caused a reduction in local adaptation due to the loss of genetic diversity or higher proportions of deleterious mutations, particularly in small isolated populations, as suggested by higher residual values in our GLM models in less diverse populations. Importantly, fast environmental changes (notably climatic) may disrupt local adaptations, especially in the Lake Trout context where migration is impossible between most of the lakes and fish have to face the environmental changes to survive. Such fast environmental changes may more especially affect smaller populations already showing reduced diversity, both neutral and putatively adaptive. Implementing artificial gene flow between these smaller, genetically impoverished and isolated Lake Trout populations may reduce the mismatches between organisms and rapidly human-altered environments (Carroll et al., 2014; Frankham, 2015; Waller, 2015). Several studies have now revealed successful persistence of small populations by an increase in population growth owing to the introduction of new alleles in both experimental translocations (Fitzpatrick et al., 2016) or wild populations (Madsen, Shine, Olsson, & Wittzell, 1999; Johnson et al., 2010; Harrisson et al., 2016; and see Frankham, 2015; Waller, 2015; Whiteley et al., 2015 for reviews). The populations showing the highest genetic diversities and the lowest inbreeding (26, 27, 28, 29 and 31) should be the primary sources of such genetic rescue (Frankham, 2015). In doing so however, this is important to avoid disruption of potential adaptation to other aspects of the environment as well as other negative impacts associated with fish translocation (i.e., diseases propagation). While the last point is difficult to control for, the first point may be solved by translocating relatively few individuals in order to keep the frequencies of potentially detrimental alleles low, and thus favour their purge by local selection.

ACKNOWLEDGEMENTS

We thank all 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. We are grateful to Clément Rougeux for assistance in the library preparation, to Éric Normandeau for bioinformatics assistance, to Thierry Gosselin for his thoughtful comments while analysing the data, to Julien Prunier for help in extracting air temperatures, and to Ben Sutherland for comments on the manuscript. We are also grateful to AE M.M. Hansen and three anonymous referees for their very useful and constructive comments on a previous version of this manuscript. This research was funded by the MFFP, the Canadian Research Chair in Genomics and Conservation of Aquatic Resources, as well as by a Strategic Project Grant from the Natural Science and Engineering Research Council of Canada (NSERC) to LB and PS.

DATA ACCESSIBILITY

Individual raw sequences are available at the Sequence Read Archive (SRA) (Study Accession no. SRP115105). The VCF file MEC-17-0449.vcf used for population genomic analyses is available on Dryad Digital Repository https://doi.org/10.5061/dryad.2b8f1. In this VCF file, individual names contain population number, population name, and individual number. Environmental and geographic data are in the tables of the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

AUTHOR CONTRIBUTION

L.B., P.S., I.T. and C.P. conceived the study. C.P. contributed to laboratory work and conducted bioinformatic analyses. C.P. and A.-L.F. achieved statistical analyses and wrote the manuscript. All co-authors critically revised and contributed to edit the manuscript and approved the final version to be published.

ORCID


Charles Perrier http://orcid.org/0000-0001-5820-9374

REFERENCES

Agrawal, A. F., & Whitlock, M. C. (2012). Mutation load: The fitness of individuals in populations where deleterious alleles are abundant. *Annual Review of Ecology, Evolution, and Systematics*, *43*, 115–135.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

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

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*, 697–709.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*, 403–410.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Balick, D. J., Do, R., Cassa, C. A., Reich, D., & Sunyaev, S. R. (2014). Dominance of deleterious alleles controls the response to a population bottleneck (G Coop, Ed.). *PLoS Genetics*, *11*, e1005436-23.

Bay, R. A., & Palumbi, S. R. (2014). Multilocus adaptation associated with heat resistance in reef-building corals. *Current Biology*, *24*(24), 2952–2956.

Beaumont, M. A., & Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, *13*, 969–980.

Bernatchez, L. (2016). On the maintenance of genetic variation and adaptation to environmental change: Considerations from population genomics in fishes. *Journal of Fish Biology*, *89*, 2519–2556.

Bernatchez, S., Laporte, M., Perrier, C., Sirois, P., & Bernatchez, L. (2016). Investigating genomic and phenotypic parallelism between piscivorous and planktivorous Lake Trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses. *Molecular Ecology*, *25*, 4773–4792.

Berthelot, C., Brunet, F. E. D. E. R., Chalopin, D., Juanchich, A., Bernard, M., Noel, B., … Guiguen, Y. (2014). The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications*, *5*, 1–10.

Bethke, L. L., Zilversmit, M., Nielsen, K., Daily, J., Volkman, S. K., Ndiaye, D., … Wirth, D. F. (2006). Duplication, gene conversion, and genetic diversity in the species-specific acyl-CoA synthetase gene family of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology*, *150*, 10–24.

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.

Blanquart, F., Gandon, S., & Nuismer, S. L. (2012). The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, *25*, 1351–1363.

Bourret, V., Dionne, M., Kent, M. P., Lien, S., & Bernatchez, L. (2013). Landscape genomics in Atlantic salmon (*Salmo salar*): Searching for gene-environment interactions driving local adaptation. *Evolution*, *67*, 3469–3487.

Bradshaw, W. E., & Holzapfel, C. M. (2008). Genetic response to rapid climate change: It's seasonal timing that matters. *Molecular Ecology*, *17*, 157–166.

Brauer, C. J., Hammer, M. P., & Beheregaray, L. B. (2016). Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology*, *25*, 5093–5113.

Caballero, A., Bravo, I., & Wang, J. (2017). Inbreeding load and purging: Implications for the short-term survival and the conservation management of small populations. *Nature Publishing Group*, *118*, 177–185.

Carroll, S. P., Jorgensen, P. S., Kinnison, M. T., Bergstrom, C. T., Denison, R. F., Gluckman, P., … Tabashnik, B. E. (2014). Applying evolutionary biology to address global challenges. *Science*, *346*(6207), 1245993.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlehwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *Genes Genomes Genetics*, *1*(3), 171–182.

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.

Charlesworth, B., Morgan, M. T., & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics*, *134*, 1289–1303.

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.

Conant, G. C., & Wolfe, K. H. (2008). Turning a hobby into a job: How duplicated genes find new functions. *Nature Reviews Genetics*, *9*, 938–950.

Crandall, K. A., Bininda-Emonds, O. R. P., Mace, G. M., & Wayne, R. K. (2000). Considering evolutionary processes in conservation biology. *Trends in Ecology and Evolution*, *15*(7), 290–295.

Crisci, J. L., Dean, M. D., & Ralph, P. (2016). Adaptation in isolated populations: When does it happen and when can we tell? *Molecular Ecology*, *25*, 3901–3911.

Currie, S., & Tufts, B. (1997). Synthesis of stress protein 70 (Hsp70) in rainbow trout (*Oncorhynchus mykiss*) red blood cells. *Journal of Experimental Biology*, *200*, 607–614.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … McVean, G. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158.

Dehal, P., & Boore, J. L. (2005). Two rounds of whole genome duplication in the ancestral vertebrate (P Holland, Ed.). *PLOS Biology*, *3*, e314–e319.

DeLong, J. P., Forbes, V. E., Galic, N., Gibert, J. P., Laport, R. G., Phillips, J. S., & Vavra, J. M. (2016). How fast is fast? Eco-evolutionary dynamics and rates of change in populations and phenotypes. *Ecology and Evolution*, *6*, 573–581.

Deutsch, C. A., Tewksbury, J. J., Huey, R. B., Sheldon, K. S., Ghalambor, C. K., Haak, D. C., & Martin, P. R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences*, *105*, 6668–6672.

Dionne, M., Miller, K. M., Dodson, J. J., & Bernatchez, L. (2009). MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. *Philosophical Transactions of the Royal Society of London*, *364*, 1555–1565.

Dionne, M., Miller, K. M., Dodson, J. J., Caron, F., & Bernatchez, L. (2007). Clinal variation in mhc diversity with temperature: Evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. *Evolution*, *61*, 2154–2164.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2013). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214.

Dorschner, M. O., Duris, T., Bronte, C. R., Curtis, M. B., & Phillips, R. B. (2000). High levels of MHC class II allelic diversity in lake trout from Lake Superior. *Journal of Heredity*, *91*(5), 359–363.

Edmonds, C. A., Lillie, A. S., & Cavalli-Sforza, L. L. (2004). Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences*, *101*, 975–979.

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.

Excoffier, L., Hofer, T., & Foll, M. (2009b). Detecting loci under selection in a hierarchically structured population. *Heredity*, *103*, 285–298.

Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*(2), 479–491.

Fan, C., Zhang, Y., Yu, Y., Rounsley, S., Long, M., & Wing, R. A. (2008). The subtelomere of *Oryza sativa* chromosome 3 short arm as a hot bed of new gene origination in rice. *Molecular Plant*, *1*, 839–850.

Farmer, T. M., Marschall, E. A., Dabrowski, K., & Ludsin, S. A. (2015). Short winters threaten temperate fish populations. *Nature Communications*, *6*, 1–10.

Fitzpatrick, S. W., Gerberich, J. C., Angeloni, L. M., Bailey, L. L., Broder, E. D., Torres-Dowdall, J., … Chris Funk, W. (2016). Gene flow from an adaptively divergent source causes rescue through genetic and demographic factors in two wild populations of Trinidadian guppies. *Evolutionary Applications*, *9*, 879–891.

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.

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

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, D. J., & Bernatchez, L. (2001). Adaptive evolutionary conservation: Towards a unified concept for defining conservation units. *Molecular Ecology*, *10*, 2741–2752.

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, *27*(9), 489–496.

Gagnaire, P.-A., Normandeau, E., Cote, C., Moller Hansen, M., & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the panmictic American eel (*Anguilla rostrata*). *Genetics*, *190*, 725–736.

Gosselin, T., & Bernatchez, L. (2016). stackr: GBS/RAD Data Exploration, Manipulation and Visualization using R. R package version 0.3.1. Retrieved from https://github.com/thierrygosselin/stackr. https://doi.org/10.5281/zenodo.53832

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

Guo, B. C., DeFaveri, J., Sotelo, G., Nair, A., & Merila, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, *13*, 19. https://doi.org/10.1186/s12915-015-0130-8

Haasl, R. J., & Payseur, B. A. (2015). Fifteen years of genomewide scans for selection: Trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology*, *25*, 5–23.

Halbisen, M. A., & Wilson, C. C. (2009). Variable introgression from supplemental stocking in Southern Ontario populations of Lake Trout. *Transactions of the American Fisheries Society*, *138*, 699–719.

Hand, B. K., Muhlfeld, C. C., Wade, A. A., Kovach, R. P., Whited, D. C., Narum, S. R., … Stanford, J. A. (2016). Climate variables explain neutral and adaptive variation within salmonid metapopulations: The importance of replication in landscape genetics. *Molecular Ecology*, *25*, 689–705.

Harris, L. N., Moore, J. S., McDermid, C. G., & Swanson, H. K. (2014). Long-distance anadromous migration in a fresh water specialist: The Lake Trout (*Salvelinus namaycush*). *The Canadian Field-Naturalist*, *128*(3), 260–264.

Harrisson, K. A., Pavlova, A., Gonçalves da Silva, A., Rose, R., Bull, J. K., Lancaster, M. L., … Sunnucks, P. (2016). Scope for genetic rescue of an endangered subspecies though re-establishing natural gene flow with another subspecies. *Molecular Ecology*, *25*, 1242–1258.

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.

Hedrick, P. W., & Fredrickson, R. (2010). Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. *Conservation Genetics*, *11*, 615–626.

Hemmer-Hansen, J., Nielsen, E. E., Frydenberg, J., & Loeschcke, V. (2007). Adaptive divergence in a high gene flow environment: Hsc70 variation in the European flounder (*Platichthys flesus* L.). *Heredity*, *99*, 592–600.

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetic Research*, *8*, 269–294.

Hornoy, B., Pavy, N., Gerardi, S., Beaulieu, J., & Bousquet, J. (2015). Genetic adaptation to climate in white spruce involves small to moderate allele frequency shifts in functionally diverse genes. *Genome Biology and Evolution*, *7*(12), 3629–3285.

Hutchings, J. A. (2011). Old wine in new bottles: Reaction norms in salmonid fishes. *Heredity*, *106*, 421–437.

Ilut, D. C., Nydam, M. L., & Hare, M. P. (2014). Defining loci in restriction-based reduced representation genomic data from nonmodel species: Sources of bias and diagnostics for optimal clustering. *BioMed Research International*, *675158*.

Iwama, G. K., Thomas, P. T., Forsyth, R. B., & Vijayan, M. M. (1998). Heat shock expression in fish. *Reviews in Fish Biology and Fisheries*, *8*, 35–56.

Jaillon, O., Aury, J.-M., Brunet, F., Petit, J.-L., & Stange-Thomann, N. (2004). Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature*, *432*, 1–12.

Johnson, W. E., Onorato, D. P., Roelke, M. E., Land, E. D., Cunningham, M., Belden, R. C., … Howard, J. (2010). Genetic restoration of the Florida panther. *Science*, *329*, 1641–1645.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

Kahle, D., & Wickham, H. (2013). ggmap: Spatial visualization with ggplot2. *The R Journal*, *5*, 144–161.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281. https://doi.org/10.7717/peerj.281

Kavanagh, K. D., Haugen, T. O., Gregersen, F., Jernvall, J., & Vøllestad, L. A. (2010). Contemporary temperature-driven divergence in a Nordic freshwater fish under conditions commonly thought to hinder adaptation. *BMC Evolutionary Biology*, *10*(1), 1.

Kostka, D., Hahn, M. W., & Pollard, K. S. (2010). Noncoding sequences near duplicated genes evolve rapidly. *Genome Biology and Evolution*, *2*, 518–533.

Kubota, S., Iwasaki, T., Hanada, K., Nagano, A. J., Fujiyama, A., Toyoda, A., … Morinaga, S. I. (2015). A genome scan for genes underlying microgeographic-scale local adaptation in a wild *Arabidopsis* species (R Mauricio, Ed.). *PLoS Genetics*, *11*, e1005361-26.

Leder, E. H., McCairns, R. S., Leinonen, T., Cano, J. M., Viitaniemi, H. M., Nikinmaa, M., … Merilä, J. (2015). The evolution and adaptive potential of transcriptional variation in sticklebacks—Signatures of selection and widespread heritability. *Molecular Biology and Evolution*, *32*, 674–689.

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, *17*, 183–189.

Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., … Grammes, F. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature Publishing Group*, *533*, 200–205.

Limborg, M. T., Helyar, S. J., De Bruyn, M., Taylor, M. I., Nielsen, E. E., Ogden, R. O. B., … Bekkevold, D. (2012). Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, *21*, 3686–3703.

Linardopoulou, E. V., Williams, E. M., Fan, Y., Friedman, C., Young, J. M., & Trask, B. J. (2005). Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature*, *437*, 90–100.

Lischer, H., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299.

Logan, M., Duryea, M., Molnar, O., Kessler, B., & Calsbeek, R. (2016). Spatial variation in climate mediates gene flow across an island archipelago. *Evolution*, *70–10*, 2395–2403.

Luikart, G., Allendorf, F., Cornuet, J. M., & Sherwin, W. B. (1998). Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity*, *89*(3), 238–247.

Lynch, M., Conery, J., & Bürger, R. (1995). Mutation accumulation and the extinction of small populations. *American Naturalist*, *146*, 489–518.

Macqueen, D. J., & Johnston, I. A. (2014). A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132881.

Madsen, T., Shine, R., Olsson, M., & Wittzell, H. (1999). Restoration of an inbred adder population. *Nature*, *402*, 34–35.

Manel, S., Perrier, C., Pratlong, M., Abi-Rached, L., Paganini, J., Pontarotti, P., & Aurelle, D. (2015). Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, *25*, 170–184.

Matala, A. P., Ackerman, M. W., Campbell, M. R., & Narum, S. R. (2014). Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, *7*, 682–701.

Mazur, C. F. (1996). The heat shock protein response and physiological stress in aquatic organisms.

McCairns, R. J. S., & Bernatchez, L. (2012). Plasticity and heritability of morphological variation within and between parapatric stickleback demes. *Journal of Evolutionary Biology*, *25*, 1097–1112.

McCairns, J. S., Smith, S., Sasaki, M., Bernatchez, L., & Beheregaray, L. B. (2016). The adaptive potential of subtropical rainbowfish in the face of climate change: Heritability and heritable plasticity for the expression of candidate genes. *Evolutionary Applications*, *9*, 531–545.

McCracken, G. R., Perry, R., Keefe, D., & Ruzzante, D. E. (2013). Hierarchical population structure and genetic diversity of lake trout (*Salvelinus namaycush*) in a dendritic system in Northern Labrador. *Freshwater Biology*, *58*(9), 1903–1917.

Meirmans, P. G., & van Tienderen, P. H. (2004). genotype and genodive: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, *4*, 792–794.

Messer, P. W., & Petrov, D. A. (2013). Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology and Evolution*, *28*, 659–669.

Michalakis, Y., & Excoffier, L. (1996). A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, *142*(3), 1061–1064.

Moore, J.-S., Bourret, V., Dionne, M., Bradbury, I., O'Reilly, P., Kent, M., … Bernatchez, L. (2014). Conservation genomics of anadromous Atlantic salmon across its North American range: Outlier loci identify the same patterns of population structure as neutral loci. *Molecular Ecology*, *23*(23), 5680–5697.

Moritz, C. (1994). Defining "evolutionarily significant units" for conservation. *Trends in Ecology and Evolution*, *9*(10), 373–375.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*, 2841–2847.

Narum, S. R., & Campbell, N. R. (2010). Sequence divergence of heat shock genes within and among 3 Oncorhynchids. *Journal of Heredity*, *101*, 107–112.

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.

Nei, M., & Li, W.-H. (1976). The transient distribution of allele frequencies under mutation pressure. *Genetics Research*, *28*, 205–214.

Northrup, S., Connor, M., & Taylor, E. B. (2010). Population structure of lake trout (*Salvelinus namaycush*) in a large glacial-fed lake inferred from microsatellite DNA and morphological analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*, 1171–1186.

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.

Olson-Manning, C. F., Wagner, M. R., & Mitchell-Olds, T. (2012). Adaptive evolution: Evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, *13*, 867–877.

Peischl, S., Dupanloup, I., Foucal, A., Jomphe, M., Bruat, V., Grenier, J.-C., … Excoffier, L. (2016). Relaxed selection during a recent human expansion. *bioRxiv*, 064691. https://doi.org/10.1101/064691

Peischl, S., Dupanloup, I., Kirkpatrick, M., & Excoffier, L. (2013). On the accumulation of deleterious mutations during range expansions. *Molecular Ecology*, *22*, 5972–5982.

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach (T Yin, Ed.). *PLoS ONE*, *7*, e32253-8.

Polechová, J., & Barton, N. H. (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 6401–6406.

Pörtner, H. O., & Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science*, *315*, 95–97.

Prunier, J. G., Dubut, V., Chikhi, L., & Blanchet, S. (2017). Contribution of spatial heterogeneity in effective population sizes to the variance in pairwise measures of genetic differentiation. *Methods in Ecology and Evolution*, 1–12.

Pujolar, J. M., Jacobsen, M. W., Als, T. D., Frydenberg, J., Munch, K., Jónsson, B., … Hansen, M. M. (2014). Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*, *23*(10), 2514–2528.

Pujolar, J. M., Jacobsen, M. W., Frydenberg, J., Als, T. D., Larsen, P. F., Maes, G. E., … Hansen, M. M. (2013). A resource of genome-wide single-nucleotide polymorphisms generated by RAD tag sequencing in the critically endangered European eel. *Molecular Ecology Resources*, *13*, 706–714.

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

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Reed, D. H., Lowe, E. H., Briscoe, D. A., & Frankham, R. (2003). Inbreeding and extinction: Effects of rate of inbreeding. *Conservation Genetics*, *4*, 405–410.

Régnière, J., Saint-Amant, R., & Béchard, A. (2014). BioSIM 10 – User's manual. Nat. Resour. Can., Can. For. Serv., Laurentian For. Cent., Québec (Quebec). Inf. Rep. LAU-X-137E.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*, 4348–4370.

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.

Ritland, K. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetical Research*, *67*, 175–185.

Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., … Loire, E. (2017). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, *515*, 261–263.

van Rossum, G., & de Boer, J. (1991). Interactively testing remote servers using the Python programming language. *CWi Quarterly*, *4*(4), 283–303.

Rousset, F. (1997). Genetic differentiation. *Genetics*, *145*, 1219–1228.

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

Somero, G. N. (2010). The physiology of climate change: How potentials for acclimatization and genetic adaptation will determine 'winners' and "losers". *Journal of Experimental Biology*, *213*, 912–920.

Sutherland, B. J., 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*(12), 3600–3617.

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

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.

de Vries, A., & Ripley, B. D. (2013). Ggdendro: tools for extracting dendrogram and tree diagram plot data for use with ggplot. R package version 0.1-12, Retrieved from http://CRAN.R-project.org/package=ggdendro.

Waller, D. M. (2015). Genetic rescue: A safe or risky bet? *Molecular Ecology*, *24*, 2595–2597.

Waples, R. S., & Do, C. (2008). ldne: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*, 753–756.

Waples, R. K., Seeb, L. W., & Seeb, J. E. (2016). Linkage mapping with paralogs exposes regions of residual tetrasomic inheritance in chum salmon (*Oncorhynchus keta*). *Molecular Ecology Resources*, *16*, 17–28.

Whiteley, A. R., Fitzpatrick, S. W., Funk, W. C., & Tallmon, D. A. (2015). Genetic rescue to the rescue. *Trends in Ecology and Evolution*, *30*, 1–8.

Whitlock, M. C., & Bürger, R. (2004). The good, the bad, and the ugly: Fixation of new mutations in small populations. In R. Ferriere, U. Dieckmann, & D. Couvet (Eds.), *Evolutionary conservation biology: The genetics of rarity*. Cambridge: Cambridge University Press.

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

Wilson, C. C., & Hebert, P. D. (1996). Phylogeographic origins of lake trout (*Salvelinus namaycush*) in eastern North America. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*, 2764–2775.

Wilson, C. C., & Hebert, P. D. (1998). Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*, 1010–1024.

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

Wood, J. L. A., Yates, M. C., & Fraser, D. J. (2016). Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evolutionary Applications*, *9*, 640–657.

Yasuike, M., De Boer, J., von Schalburg, K. R., Cooper, G. A., McKinnel, L., Messmer, A., . . . & Koop, B. F. (2010). Evolution of duplicated IgH loci in Atlantic salmon, *Salmo salar*. *BMC Genomics*, *11*, 486.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Perrier C, Ferchaud A-L, Sirois P, Thibault I, Bernatchez L. Do genetic drift and accumulation of deleterious mutations preclude adaptation? Empirical investigation using RADseq in a northern lacustrine fish. *Mol Ecol*. 2017;26:6317–6335. https://doi.org/10.1111/mec.14361