Conserv Genet
DOI 10.1007/s10592-011-0224-y

RESEARCH ARTICLE

Empirical assessment of software efficiency and accuracy to detect introgression under variable stocking scenarios in brook charr (*Salvelinus fontinalis*)

A. D. Marie · L. Bernatchez · D. Garant

Received: 15 June 2010 / Accepted: 29 April 2011

Abstract Stocking wild populations with domesticated fish is a common practice that promotes variable levels of introgression depending on the stocking intensity. The detection of hybridization and introgression has recently benefited from the application of Bayesian techniques implemented in various software. However, few studies have assessed their efficiency under various scenarios of stocking in the wild. The objective of this study is to assess quantitatively the effects of using two of the most widely distributed software, STRUCTURE and NEWHYBRIDS, on the level of introgression detected in wild brook charr (*Salvelinus fontinalis*) subjected to variable stocking intensities. We first found differences in the efficiency of software assignments based on simulated individuals, with STRUCTURE performing better than NEWHYBRIDS. However, NEWHYBRIDS showed higher assignments accuracy than STRUCTURE for the same sets of individuals. Thus, our results suggest that these software should be used in combination to assess the effects of stocking. Indeed, STRUCTURE is particularly relevant to evaluate the presence of hybrids in wild populations, whereas NEWHYBRIDS might be preferred to accurately assess the number of hybrids present in a sample. When applied to wild populations, STRUCTURE assigned more individuals than NEWHYBRIDS to the wild category. Moreover, the proportions of assigned domestic and hybrid individuals were higher in more intensively stocked lakes, whereas the opposite trend was observed for wild individuals.

Keywords Admixture · Brook charr · Microsatellites · NEWHYBRIDS · Stocking · STRUCTURE

A. D. Marie · D. Garant (✉)
Département de biologie, Université Sherbrooke,
Sherbrooke, QC J1K 2R1, Canada
e-mail: Dany.Garant@USherbrooke.ca

L. Bernatchez
IBIS (Institut de Biologie Intégrative et des Systèmes), Pavillon
Charles-Eugène-Marchand, Université Laval, Québec,
QC G1V 0A6, Canada

Introduction

Introductions and supplementations from exogenous individuals in wild populations are common practices throughout the world. In many cases, introductions are accidental (i.e. escape of domesticated farm fish, reviewed in Weir and Grant 2005). But in other instances, such practices have very diverse objectives such as increasing the size of endangered natural populations (Biebach and Keller 2009) or supplementing populations that are subject to harvesting (Grandjean et al. 2009). Yet, it is now widely recognized that introductions of exogenous individuals could represent a threat for natural populations, for example through hybridization and loss of genetic integrity (Hindar et al. 1991; Ayres et al. 1999; Kidd et al. 2009; Marie et al. 2010). It is therefore imperative to accurately document these effects in wild populations.

In fish, hybridization has been intensively studied at the intraspecific level, and especially in salmonids species (Guyomard 1997; Hansen et al. 2001; Halbisen and Wilson 2009). Salmonids have a considerable economical and recreational value and stocking of natural populations is commonly performed with domesticated exogenous individuals. Hybridization between wild and domestic individuals could result in the modification of the genetic integrity of populations and the loss of local adaptation (Englbrecht et al. 2002; Fraser 2008; Marie et al. 2010;

Published online: 12 May 2011

Bougas et al. 2010; Sauvage et al. 2010), which is a major issue for the management of those populations.

The detection of hybridization and introgression has benefited from the application of microsatellite markers and the development of modern analytical approaches for individual-based multi-locus analyses, such as population assignment and clustering methods (Pritchard et al. 2000; Hansen et al. 2001; Anderson and Thompson 2002; Susnik et al. 2004; Vähä and Primmer 2006). In general, to accurately document hybridization and introgression, a relatively large number of loci (i.e. microsatellite markers) as well as a substantial level of genetic differentiation between hybridizing populations (i.e. F_{ST}) is recommended (reviewed in Vähä and Primmer 2006). Among the various analytical methods available, Bayesian techniques have generally proven to be more efficient than likelihood based approaches (Vähä and Primmer 2006). However, software implementing Bayesian approach, in particular Structure (Pritchard et al. 2000) and NewHybrids (Anderson and Thompson 2002) differ in the method they rely on for assignments. For example, the widely used software Structure assigns the probability of an individual to have a recent ancestry in two or more populations. On the other hand, the software NewHybrids evaluates the probability of one individual belonging to a hybrid or parental classes. Moreover, previous studies have emphasized that the efficiency (defined as the number of correctly identified individuals for a category over the actual number of individuals belonging to that category in the sample) and the accuracy (the number of correctly identified individuals for a category over the actual number of individuals assigned to that category) of these software can differ greatly depending on the population context (see Vähä and Primmer 2006 for details). Indeed, a number of studies compared the global efficiency of methods to detect the level of admixture (Vähä and Primmer 2006; Burgarella et al. 2009; Sanz et al. 2009) as well as the importance of choosing an appropriate threshold probability to discriminate pure versus introgressed individuals (Burgarella et al. 2009). However, these were essentially qualitative in nature, and although Sanz et al. (2009) demonstrated the impacts of variable stocking intensity, the information provided on the stocking history of each lake was binary (e.g. presence/absence) with some degree of uncertainty.

The goal of this study is to complement previous efforts by quantitatively assessing the impacts of employing different assignment software on the level of introgressive hybridization being detected under well documented variable stocking scenarios in the brook charr (*Salvelinus fontinalis*). We have recently shown that stocking may impact on affect the genetic integrity and genetic Structure of brook charr populations depending on its intensity (Marie et al. 2010). We thus quantitatively compared the efficiency and accuracy of the software Structure and NewHybrids in detecting hybrids using simulated genotypes from stocked individuals of domestic and wild origin. We then specifically evaluated the level of admixture within seven lakes with different stocking histories and then compared the effects of different stocking practices and software on the resulting level of hybridization being detected.

Materials and methods

Sampling procedures

Sampling was conducted in the Portneuf wildlife Reserve in Québec, Canada (47°09′ N, 72°17′ W; Fig. 1) in June 2007 and July 2008 (see Marie et al. 2010 for details of sampling procedures). We selected lakes from two categories of stocking intensity: four moderately stocked and three heavily stocked (Fig. 1). A moderately stocked lake underwent stocking in less than 50% of years over the past 15 years (from 1992 to 2007). On average, such type of lake was stocked with 5819 ± 3427 fish (mean ± SD) per year. A heavily stocked lake underwent stocking events in more than 50% of the past 15 years (from 1992 to 2007) with on average 14926 ± 12930 stocked fish per lake per year respectively (see Marie et al. 2010 for details). We also sampled brook charr in four un-stocked lakes (Fig. 1) to determine the wild genetic composition of brook charr in the reserve (see below). Finally, a sample of 51 domesticated brook charr was obtained from the nearby Jacques-Cartier Hatchery (Québec, Canada). Lakes in the Portneuf Reserve have been stocked exclusively using fish from this hatchery, which allowed us to accurately determine the genetic composition of the parental group of domestic origin (see below). Tissue samples (adipose fins) were preserved in 95% ethanol until DNA extraction.

DNA extraction and microsatellite analyses

DNA was extracted from fin clips (5 mm^2) using the salting out method of Aljanabi and Martinez (1997). A total of 866 brook charr from the reserve as well as all individuals from the hatchery were genotyped using nine microsatellite loci: sfoC129, sfoC113, sfoC88, sfoB52, sfoD75, sfoC24, sfoC86, SCO218, sfoD100 as detailed in Marie et al. (2010). PCR products were visualized using an AB 3130 capillary DNA sequencer (Applied Biosystems) and allele size was established using the software Genemapper version 4.0 (Applied Biosystems).

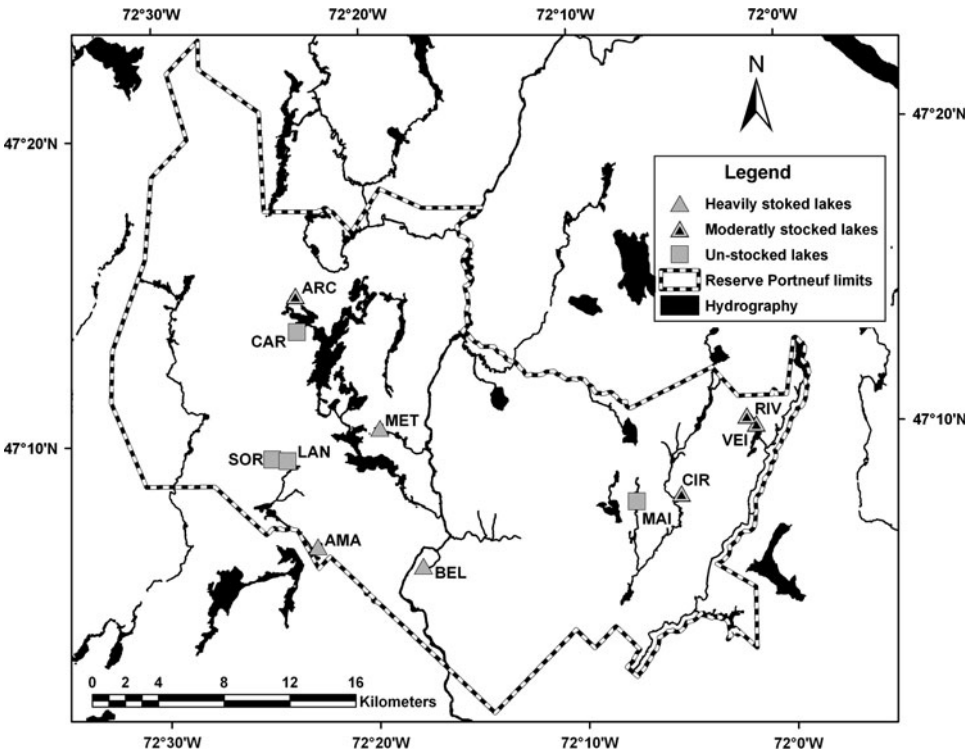

Fig. 1 Geographical locations of lakes within the Portneuf wildlife Reserve in Québec, Canada. *BEL* Belles-de-Jour Lake, *AMA* Amanites Lake, *MET* Methot Lake, *RIV* Rivard Lake, *VEI* Veillette Lake, *ARC* Arcand Lake, *CIR* Circulaire Lake, *CAR* Caribou Lake, *LAN* Langoumois Lake, *SOR* Sorbier Lake, *MAI* Main de Fer Lake

Selection of wild and domestic individuals for assignment tests

First, we selected wild and domestic individuals for each lake independently with both STRUCTURE (henceforth ST) and NEWHYBRIDS (NH) to simulate individuals of different background (Fig. 2). Namely, each run of each software included individuals from the stocked lake of interest (both years pooled) and individuals of the potential parental sources (Fig. 2). We considered wild individuals from four un-stocked lakes ($N = 316$) of the Portneuf Reserve and domestic individuals from the Jacques-Cartier Hatchery ($N = 51$) as potential parental sources.

We estimated the individual admixture proportions (q-values) and their posterior probability intervals for each individual in each stocked lake using ST. The number of clusters for each lake was assessed using the ad hoc statistic ΔK (see Evanno et al. 2005), which revealed the presence of three clusters. Using a $K = 3$ allowed us to obtain a clear discrimination between wild, potential hybrids and domestic individuals (at least one group for each individual category) and thus to properly select individuals belonging to domestic and wild category for our simulations. In each analysis, we used the admixture model with correlated allele frequency with 250,000 steps of the Markov-Chain preceded by a burnin-period of 100,000 steps.

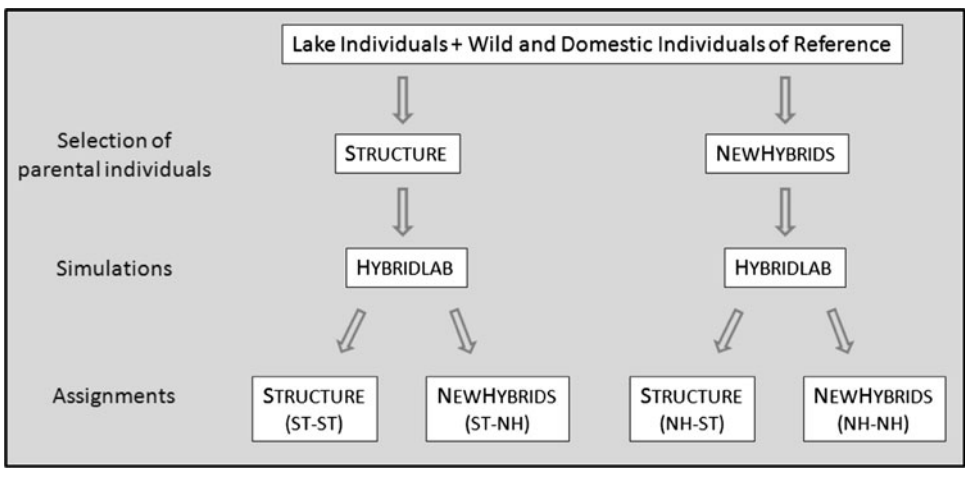

Fig. 2 Summary of the different steps used for individual assignments. See text for a detailed explanation. *ST* STRUCTURE, *NH* NEWHYBRIDS, *ST–ST* selection of parental individuals with ST and assignments with ST, *ST–NH* selection of parental individuals with ST and assignments with NH, *NH–ST* selection of parental individuals with NH and assignments with ST, *NH–NH* selection of parental individuals with NH and assignments with NH

We also performed the assessment of individual probability of belonging to a given group using NH. We ran the analysis by specifying the genotype frequency category of wild and domestic individuals (i.e. wild and domestic individuals each forming a group). Each run of the Markov-Chain consisted in a burnin-period of 100,000 iterations followed by 250,000 iterations and no prior species information was assumed. Individuals belonging to a category with a threshold $\geq 70\%$ were considered correctly assigned (following Gunnell et al. 2008 and Gagnaire et al. 2009).

We selected source individuals for genotypes simulations based on the recommendations from two previous studies. Nielsen et al. (2003) recommended using between 30 and 50 individuals for each simulation to reduce biases. Also, Vähä and Primmer (2006) suggested a q-value threshold ≤ 0.1 (estimated using ST) to define parental populations and to obtain both a good efficiency and accuracy to detect hybrids in each lake. For analyses using ST, we thus kept 36 domestic fish from the hatchery and 9 wild individuals from each of our four un-stocked lakes ($N = 36$) and used q-values ≤ 0.05 as a threshold to be more stringent in our choice of individuals representing parental populations. To be consistent in our analysis, we also selected 36 individuals of each parental category to run the analyses with NH. The selection of individuals in this case was performed according to their probability of belonging to a given category (i.e. between 98% and 100% for wild and domestic individuals, data not shown). F_{ST} between groups were calculated with FSTAT (version 2.9.3.2; Goudet 1995) using the 36 individuals selected for each parental simulation of each software. Moreover, we calculated the F_{ST} for all individuals with a q-values threshold of 0.05, estimated using ST and compared it (Wilcoxon matched pair test) with the F_{ST} of the 36 individuals selected for each parental population to assess the consequence of using different selection methods.

Simulations of different categories of individuals

To assess the relative performance of each software to detect hybrids, we simulated individuals using HYBRIDLAB 1.0 (Nielsen et al. 2001). Six groups of individuals (pure wild, pure domestic, F_1 hybrids, F_2 hybrids and reciprocal backcrosses: F1*wild parents and F1*domestic parents) were simulated for each lake using the 36 individuals selected with ST and NH (Fig. 2). We performed three sets of simulations to compare the global efficiency of each software with different proportions of hybrids. First, we generated group 1800 genotypes for each parental group (SWI for simulated wild individuals and SDI for simulated domestic individuals) and 100 genotypes of each of the four hybrid groups (F1 and F2 hybrids, F1*wild parents (BWP) and F1*domestic parents (BDP) backcrosses), where hybrids represented 10% of all simulated individuals. Then, we generated groups that comprised higher proportions (40 and 67%) of hybrids to assess the effect of having variable hybrids proportion in samples on the software assignments capability. Samples including 40 and 67% of hybrids were represented by the simulation of 1000 genotypes for each parental group. And, 333 and 1000 genotypes for each hybrid group were respectively simulated for samples with 40 and 67% of hybrids.

Admixture analyses from simulated individuals

Admixture analyses were carried out with each software using the same settings as above. However, here we used a q-values threshold of 0.1 with ST which allowed us to define hybrid individuals as those having q-values >0.1 and <0.9. For NH, we used a threshold ≥ 0.7 to consider the correctly assigned hybrids to their respective category. We were especially interested in assessing the effect of software used for the selection of individuals and then admixture analyses on the resulting difference in hybrid detection capability (for example comparing the performance of a ST–ST combination to a NH–ST combination—see Fig. 2). We first applied a t-test to compare the hybrid detection capability between samples with 10, 40 and 67% of hybrids. We then compared the efficiency and the accuracy of ST and NH, both from the standpoint of selecting individuals for simulations and for quantifying admixture, as well for assessing the effect of stocking intensity using simulated individual assignments. We used an ANOVA (GenStat version 12; VSN intl.) with software (ST–ST, ST–NH, NH–ST, NH–NH), stocking intensity (MS or HS), category of individuals (domestic, wild or hybrid), and their interactions as factors to assess differences in both efficiency and accuracy. Post-hoc Tukey tests were performed when factors were deemed significant in the ANOVA.

Admixture analyses using wild-caught samples

Finally, we used ST and NH to assess the proportions of admixture in the seven lakes sampled. We used the same setting as above for each software assuming a $K = 2$ with ST and specifying the genotype frequency category of domestic individuals with NH. We also included domestic individuals (from the hatchery) in the analyses when assigning individuals of each lake to a given category. We used an ANOVA to assess the effects of the software, stocking intensity and their interactions on the assignment of each category of individuals (wild, hybrid and domestic).

Results

Pairwise F_{ST} between wild and domestic individuals selected for simulations

Structure

The mean level of genetic differentiation (F_{ST}) among wild (the four un-stocked lakes) and domestic individuals was 0.220 ± 0.032, and the average F_{ST} among un-stocked lakes was 0.199 ± 0.109 (see also Marie et al. 2010). The average F_{ST} value between the 36 most extreme selected (based on *q*-values) wild and domestic individuals was 0.244 ± 0.006 ($P < 0.05$; Table 1). In comparison, the average F_{ST} value between individuals with a *q*-value ≤ 0.05 (from 267 to 276 wild individuals depending on lake) and *q*-value ≥0.95 (from 37 to 45 domestic individuals depending on the lake) was reduced to 0.193 ± 0.004 ($P < 0.05$; Table 1). This difference was significant ($Z = 2.37$, $P = 0.018$).

NewHybrids

The average F_{ST} value between wild and domestic individuals selected was 0.233 ± 0.036 ($P < 0.05$; Table 1). A total of 95.6% of wild individuals (302 ± 1.2 individuals) and 98.3% of domestic individuals (50 ± 1.1 individuals) were defined as belonging respectively to the wild and domestic category (data not shown). This level of differentiation was similar to that obtained using 36 individuals with ST ($Z = 0.25$, $P = 0.80$). However, this F_{ST} value was greater than the level of differentiation obtained using individuals with a *q*-value ≤0.05 in ST ($Z = 2.20$, $P = 0.028$).

Assignments of simulated individuals

Individuals selected using Structure

Using ST, over 99% of SWI and SDI were on average respectively assigned to their category when the sample comprised 10% of hybrids (Table 2a). Individuals that were incorrectly assigned as SWI or SDI had respectively *q*-values >0.1 and *q*-values <0.9. F1 hybrids were always correctly assigned as being hybrids (99.4% ± 0.1 on average; data not shown). In contrast, the other types of hybrids overlapped with the parental categories (higher overlap for backcrosses with an average of 29.6% ± 2.3; data not shown). Compared to samples including 40 and 67% of hybrids, we found a significant difference for assignments of SDI ($P = 0.003$ and $P < 0.001$ respectively for 40 and 67% of hybrids), which was lower in samples with more hybrids (Table 2b and c). Moreover, assignments of SDI were significantly lower ($P < 0.001$) in samples including 67% of hybrids than those including 40% of hybrids (Table 2b and c). Finally, significant differences appeared between hybrids that overlapped with parental categories, with lower overlap in samples including 67% of hybrids than samples with 40 and 10% of hybrids ($P = 0.026$; data not shown). There was also a lower overlap in samples including 40% than samples with 10% of hybrids ($P = 0.047$; data not shown).

With 10% of hybrids present in the sample, on average over 99% of SWI and SDI were correctly assigned to their

Table 1 Pairwise genetic differentiation (F_{ST}) between wild and domestic individuals for each lake obtained using Structure (for the first 36 individuals selected and for all individuals with a *q*-values ≥0.95) and NewHybrids (for the first 36 individuals selected)

Lake	Structure		NewHybrids
	The first 36 individuals	Individuals with a *q*-value ≥ 0.95	The first 36 individuals
Heavily stocked lakes			
Belles-de-Jour	0.254	0.202	0.257
Amanites	0.246	0.192	0.242
Methot	0.237	0.190	0.243
Average	0.246	0.194	0.247
Standard error	0.009	0.007	0.008
Moderately stocked lakes			
Rivard	0.247	0.190	0.249
Veillette	0.242	0.191	0.153
Arcand	0.243	0.192	0.241
Circulaire	0.240	0.195	0.246
Average	0.243	0.192	0.223
Standard error	0.003	0.002	0.046
Global average	0.244	0.193	0.233
Standard error	0.006	0.004	0.036

The averages and the standard errors are given for all lakes

Table 2 Percentage of simulated wild, hybrid and domestic individuals that were assigned respectively to the wild, hybrid and domestic category with STRUCTURE (ST) and NEWHYBRIDS (NH), using individuals selected with either STRUCTURE or NEWHYBRIDS for samples including (a) 10% of hybrids, (b) 40% of hybrids and (c) 67% of hybrids

Lake	Individuals selected with STRUCTURE						Individuals selected with NEWHYBRIDS					
	SWI		Hybrids		SDI		SWI		Hybrids		SDI	
	ST	NH	ST	NH	ST	NH	ST	NH	ST	NH	ST	NH
(a)												
Heavily stocked lakes												
Belles-de-Jour	100.0	100.0	85.3	66.5	99.6	99.6	99.9	100.0	84.0	59.5	99.8	99.9
Amanites	100.0	100.0	87.0	55.3	99.4	99.8	100.0	100.0	80.8	60.5	99.4	99.4
Methot	100.0	100.0	82.8	63.8	99.7	99.4	100.0	100.0	86.3	65.8	99.4	99.4
Average	100.0	100.0	85.0	61.9	99.6	99.6	100.0	100.0	83.7	61.9	99.5	99.6
Standard error	0.0	0.0	2.1	5.8	0.2	0.2	0.1	0.0	2.8	3.4	0.2	0.3
Moderately stocked lakes												
Rivard	100.0	100.0	83.3	58.0	99.7	99.7	100.0	100.0	84.0	54.5	99.6	99.7
Veillette	100.0	99.9	85.0	56.3	99.6	99.7	96.7	98.7	63.3	21.8	97.6	98.9
Arcand	100.0	100.0	86.8	58.0	99.7	99.8	100.0	100.0	87.0	64.0	99.8	99.7
Circulaire	99.6	99.5	74.3	40.0	99.2	99.3	100.0	100.0	85.0	67.3	99.6	99.7
Average	99.9	99.9	82.4	53.1	99.6	99.6	99.2	99.7	79.8	51.9	99.2	99.5
Standard error	0.2	0.2	5.6	8.8	0.2	0.2	1.7	0.7	11.1	20.8	1.0	0.4
Global average	99.9	99.9	83.5	56.8	99.6	99.6	99.5	99.8	81.5	56.2	99.3	99.5
Standard error	0.2	0.2	4.4	8.5	0.2	0.2	1.2	0.5	8.3	15.8	0.8	0.3
(b)												
Heavily stocked lakes												
Belles-de-Jour	100.0	100.0	89.4	69.7	98.3	96.2	99.8	99.7	91.0	71.5	97.9	96.7
Amanites	100.0	100.0	88.3	60.3	97.5	98.6	99.9	99.7	89.8	63.5	96.9	98.3
Methot	100.0	99.9	89.0	64.0	97.4	98.0	99.8	99.9	89.4	64.5	97.9	97.0
Average	100.0	100.0	88.9	64.7	97.7	97.6	99.8	99.8	90.1	66.5	97.6	97.3
Standard error	0.0	0.1	0.6	4.7	0.5	1.2	0.1	0.1	0.8	4.4	0.6	0.9
Moderately stocked lakes												
Rivard	100.0	99.8	89.3	64.4	97.0	94.9	99.6	99.5	89.8	60.2	98.4	98.9
Veillette	100.0	99.9	88.0	55.3	97.3	98.5	82.9	97.5	86.9	30.3	83.2	97.0
Arcand	99.8	99.5	89.8	62.6	98.0	98.0	99.7	99.8	89.4	61.1	96.9	97.9
Circulaire	95.8	97.9	85.4	38.7	93.9	98.1	100.0	100.0	88.1	64.0	98.6	98.4
Average	98.9	99.3	88.1	55.3	96.6	97.4	95.6	99.2	88.6	53.9	94.3	98.1
Standard error	2.1	0.9	2.0	11.7	1.8	1.7	8.4	1.2	1.3	15.8	7.4	0.8
Global average	99.4	99.6	88.5	59.3	97.1	97.5	97.4	99.4	89.2	59.3	95.7	97.7
Standard error	1.6	0.8	1.5	10.1	1.5	1.4	6.4	0.9	1.3	13.3	5.5	0.8
(c)												
Heavily stocked lakes												
Belles-de-Jour	100.0	99.9	93.8	64.8	89.6	93.2	100.0	99.9	93.7	64.9	89.9	93.2
Amanites	99.4	99.6	93.5	63.5	90.6	94.5	99.0	99.2	93.8	64.8	91.1	94.7
Methot	99.5	99.7	94.0	63.4	88.6	94.6	99.4	99.6	93.4	67.7	90.6	93.8
Average	99.6	99.7	93.8	63.9	89.6	94.1	99.5	99.6	93.6	65.8	90.5	93.9
Standard error	0.3	0.2	0.3	0.8	1.0	0.8	0.5	0.4	0.2	1.6	0.6	0.8
Moderately stocked lakes												
Rivard	99.6	99.4	92.8	61.1	92.8	94.7	97.5	98.7	93.8	61.9	89.2	94.4
Veillette	98.5	98.8	93.6	57.0	89.4	92.6	61.5	98.8	95.9	57.0	62.7	92.6
Arcand	99.0	99.1	93.8	65.2	92.7	94.7	99.8	99.8	93.0	67.8	92.0	94.4
Circulaire	82.2	97.9	92.9	39.4	79.1	95.7	82.4	97.9	92.9	39.3	79.0	95.7

Table 2 continued

Lake	Individuals selected with STRUCTURE						Individuals selected with NEWHYBRIDS					
	SWI		Hybrids		SDI		SWI		Hybrids		SDI	
	ST	NH	ST	NH	ST	NH	ST	NH	ST	NH	ST	NH
Average	94.8	98.8	93.3	55.7	88.5	94.4	85.3	98.8	93.9	56.5	80.7	94.3
Standard error	8.4	0.6	0.5	11.4	6.5	1.3	17.6	0.8	1.4	12.3	13.3	1.3
Global average	96.9	99.2	93.5	59.2	89.0	94.3	91.4	99.1	93.8	60.5	84.9	94.1
Standard error	6.5	0.7	0.5	9.2	4.6	1.0	14.6	0.7	1.0	10.1	10.7	1.0

The means and standard errors are given for all lakes. *SWI* simulated wild individuals, *SDI* simulated domestic individuals

respective category using NH (Table 2a). However, the percentage of correct assignment of hybrids to their respective category was generally smaller (from 10% for BDP to 76% for F1; data not shown). Although the global assignment of hybrids was relatively low (59.2% ± 9.2; Table 2a), almost every F1 hybrids were assigned to the hybrids categories, whereas some individuals of other hybrid categories overlapped with the parental classes (data not shown). Compared to the assignments comprising 40 and 67% of hybrids, assignments of SDI ($P = 0.007$ and $P < 0.001$ respectively for 40 and 67% of hybrids) and BDP ($P = 0.005$ for 67% of hybrids) were significantly different (correct assignment of SDI lower in samples with 40 and 67% of hybrids, Table 2b and c; and opposite trend for BDP with 67% of hybrids, data not shown). Moreover, assignments of SDI differed between samples including 40 and 67% of hybrids ($P = 0.003$), with lower assignments at 67% of hybrids (Table 2b and c). Moreover, the proportion of BDP overlapping with the parental categories was significantly lower in samples with 67% of hybrids than in samples with 10% of hybrids ($P = 0.004$). Similarly, the proportion of BDP and F2 hybrids overlapping with the parental categories in samples including 40% of hybrids were significantly lower than in samples with 10% of hybrids ($P = 0.011$ and $P = 0.021$, respectively, data not shown).

Individuals selected using NEWHYBRIDS

Using ST, on average over 99% of SWI and SDI were assigned to their parental category (Table 2a). Some of the incorrectly assigned hybrids also overlapped with the parental categories (higher overlap for backcrosses with an average of 31.7% ± 1.3; data not shown). Compared to the samples including 40 and 67% of hybrids, the *t*-tests revealed significant differences between methods for assignments of SDI ($P = 0.005$ and $P = 0.009$ respectively for 40 and 67% of hybrids), with on average fewer correctly assigned individuals in samples including 40 and 67% of hybrids (95.7% ± 5.5 and 84.9% ± 10.7 respectively; Table 2b and c) than in 10% samples (99.3 ± 0.8). Moreover, results showed significant differences between hybrid individuals overlapping with the parental categories (lower proportion of overlap in samples including 67% of hybrids; $P < 0.001$; data not shown). Similarly, the proportion of hybrids overlapping with the parental categories was significantly lower in samples with 40% of hybrids than in samples with 10% ($P < 0.001$; data not shown).

With NH, over 99% of SWI and SDI on average were correctly assigned to their respective category, whereas the proportion of hybrids assigned to the hybrid category was lower (60.5% ± 10.1; Table 2a). Simulated hybrids were again assigned in lower proportion to their respective class (from 18% for BDP to 76% for F1; data not shown). Moreover, some hybrid individuals overlapped with the parental classes (from 2% for F1 to 10% for backcrosses; data not shown). Our comparisons with samples including 40% and 67% of hybrids showed significant differences between assignments for the SDI ($P = 0.001$ and $P < 0.001$), which were lower in samples with 40% and 67% of hybrids (97.7% ± 0.8; Table 2b and 94.1% ± 1.0; Table 2c, respectively).

Comparisons of assignment methods

As our results revealed that the clear identification of F2 hybrids and backcrosses was problematic (misidentification in their respective category and overlap with the parental populations), we pooled hybrids (F1, F2 and backcrosses assigned to the hybrid categories) in a same hybrid category for subsequent analyses.

Efficiency

The results of analyses including 10% of hybrids showed a significant difference in the assignment efficiency between methods according to the category of individuals ($P < 0.001$; Table 3b). The Tukey test showed that the proportions of assigned wild and domestic individuals (on average 99.8% ± 0.7 and 99.5% ± 0.4 respectively) were

Table 3 Effects of software, stocking intensity, individuals category (wild, hybrid or domestic) and their interaction on the efficiency (number of correctly assigned individuals for a category over the actual number of individuals of that category in the sample) and accuracy (number of correctly identified individuals for a category over the actual number of individuals assigned to that category) of assignment of simulated individuals for samples including (a) 10% of hybrids, (b) 40% of hybrids and (c) 67% of hybrids

	F-statistic	d.f.	*P*-value
(a)			
Efficiency			
Software	7.13	3	< 0.001
Individuals category	117.08	2	< 0.001
Software * Individuals category	16.01	6	< 0.001
Accuracy			
Software	2.81	3	0.045
(b)			
Efficiency			
Software	6.83	3	< 0.001
Individuals category	63.42	2	< 0.001
Software * Individuals category	26.15	6	< 0.001
Accuracy			
Software	5.28	3	0.002
Individuals category	24.97	2	< 0.001
Stocking intensity	8.06	1	0.006
(c)			
Efficiency			
Software	2.79	3	0.046
Individuals category	20.63	2	< 0.001
Software * Individuals category	28.26	6	< 0.001
Accuracy			
Software	12.51	3	< 0.001
Individuals category	35.07	2	< 0.001
Stocking intensity	11.89	1	< 0.001

Results from the final model including significant terms are presented (from ANOVAs with the corresponding *F*-statistic)

higher than hybrid individuals (on average 69.5% ± 16.3; $P < 0.001$). Moreover, our analyses showed a significant effect of software on assignment efficiency ($P < 0.001$; Table 3). Indeed, post-hoc analyses revealed significantly higher efficiency in four cases: ST–ST > ST–NH (on average 94.3% ± 8.2 and 85.5% ± 21.2 respectively; $P < 0.001$); ST–ST > NH–NH (on average 94.3% ± 8.2 and 85.2% ± 22.7 respectively; $P < 0.001$); NH–ST > ST–NH (on average 93.4% ± 9.8 and 85.5% ± 21.2 respectively; $P < 0.001$); NH–ST > NH–NH (on average 93.4% ± 9.8 and 85.2% ± 22.7 respectively; $P < 0.001$). Others comparisons did not show significant differences ($P \geq 0.96$). Our results thus suggest that ST is more effective than NH at quantifying admixture when a small proportion of hybrids (10%) are present in the sample. A significant interaction between software and the category of individuals was also revealed ($P < 0.001$). More specifically, using ST to assign individuals, the proportions of assigned hybrids (on average 82.5% ± 6.4) was higher ($P < 0.001$) than using NH (on average 56.5% ± 12.2), whereas no difference between software appeared for the assignments of wild and domestic individuals (all $P \geq 0.99$). Similar results concerning the assignment efficiency were obtained with 40% and 67% of hybrids in the sample (data not shown). Thus, our results generally suggest that the software performance is driven by the hybrid individuals.

Accuracy

In samples with 10% of hybrids, the assignment accuracy differed only between software ($P = 0.045$; Table 3) and the only post-hoc significant ($P = 0.035$) comparison was found between ST-NH (on average 99.1% ± 1.1) and NH-ST (on average 96.3% ± 6.0). Contrastingly, in samples including 40% and 67% of hybrids, the accuracy of assignments differed significantly between individual categories (both $P < 0.001$; Table 3). For both 40% and 67% of hybrids, the accuracy was significantly higher for hybrids (98.0% ± 4.4 and 94.7% ± 5.8) than wild individuals (94.2% ± 5.4 and 87.0% ± 10.2; $P = 0.006$ and $P < 0.001$) and domestic individuals (90.2% ± 3.9 and 80.6% ± 7.1; $P < 0.001$ and $P < 0.001$). Moreover, the accuracy of wild individuals was also higher than domestic individuals ($P = 0.004$ and $P = 0.001$ respectively for 40% and 67% of hybrids). Our results revealed also significant differences between software ($P = 0.002$ and $P < 0.001$ respectively for 40% and 67% of hybrids; Table 3). For samples including 40 and 67% of hybrids respectively, the accuracy of assignments was lower for NH–ST (on average 93.4% ± 7.1 and 81.8% ± 10.3, respectively) than NH–NH (on average, 95.7% ± 5.5, $P = 0.031$ and 91.5% ± 9.4, $P < 0.001$, respectively) and ST–NH (on average, 96.1 ± 4.4, $P = 0.013$ and 91.5% ± 9.1, $P < 0.001$, respectively). For samples with 67% of hybrids, the accuracy of assignments were significantly higher for NH–NH (on average 91.5% ± 9.4) than ST–ST (84.9% ± 6.4; $P = 0.011$), as well as higher for ST–NH (on average 91.5% ± 9.1) than ST–ST (84.9% ± 6.4; $P = 0.011$). Other comparisons between software, for samples with 40% and 67% of hybrids, did not differ significantly (all $P \geq 0.14$). Thus, our analyses suggest that NH accuracy is greater than ST when assigning individuals. Moreover, the accuracy of assignment also increased significantly with stocking intensity ($P = 0.006$ and $P < 0.001$ respectively for 40 and 67% of hybrids; Table 3). For the highly stocked lakes, the software accuracy was of 95.6% ± 3.6 and 90.2% ± 6.5 while it dropped to 93.1% ± 6.5 and 85.4% ± 11.2 for moderately

stocked lakes, respectively for samples including 40 and 67% of hybrids.

Assignments using wild-caught samples

ANOVA results showed a significant difference between software in terms of number of wild individuals assigned (see Table 4). The number of wild individuals assigned was significantly higher when using ST (74.2% ± 27.8) than with NH (57.2% ± 42.9) (Table 5). As expected, the proportion of wild individuals assigned decreased significantly with stocking intensity ($P < 0.001$; Table 4). Using ST, the moderately stocked lakes exhibited a proportion of 96.4% ± 3.0 of wild fish whereas the heavily stocked lakes included 44.6% ± 3.5 of wild fish (Table 5). With NH, wild fish represented 89.4% ± 6.8 of individuals in the moderately stocked lakes and 14.4% ± 24.9 in the heavily stocked lakes (Table 5). Conversely, the number domestic and hybrid individuals increased significantly with the stocking intensity ($P < 0.001$ and $P = 0.010$ respectively for the domestic and hybrid individuals; Table 4). The moderately stocked lakes showed no domestic individuals and 5.7% ± 4.6 hybrid individuals on average (Table 5). The heavily stocked lakes exhibited on average 38.0% ± 15.0 domestic individuals and 27.3% ± 19.5 hybrid individuals (Table 5).

Discussion

This study aimed to quantitatively assess the impacts of employing different assignment software on the level of introgressive hybridization being detected under variable stocking practices in the brook charr. Our analyses revealed that the software STRUCTURE (ST) has a higher efficiency in assigning individuals than NEWHYBRIDS (NH), both when samples contained a low (10%), intermediate (40%) or high (67%) proportion of hybrids. This difference in efficiency was mainly related to the much higher assignment of hybrids individuals when using ST than with NH. However, our results also suggested that NH was more accurate than ST in assigning individuals, especially when the proportion of hybrids in the sample was high. Interestingly, under this same scenario, the assignment accuracy increased with the stocking intensity in the sample. Finally, when applied to wild populations, ST assigned more individuals than NH to the wild category. As expected, the proportion of individuals assigned to the domestic and hybrid categories increased with the stocking intensity, whereas the opposite trend was observed for wild individuals.

Table 4 Effects of software, stocking intensity and their interaction on categorical assignment of wild, hybrids and domestic individuals sampled from lakes in the Portneuf Reserve Quebec, Canada

	F-statistic	d.f.	*P*-value
Wild individuals			
Software	5.82	1	0.035
Stocking intensity	80.00	1	< 0.001
Software * Stocking intensity	3.22	1	0.103
Hybrid individuals			
Software	1.46	1	0.252
Stocking intensity	9.38	1	0.010
Software * Stocking intensity	0.47	1	0.507
Domestic individuals			
Software	0.23	1	0.638
Stocking intensity	52.92	1	< 0.001
Software * Stocking intensity	0.29	1	0.601

Results from ANOVAs are presented with the corresponding *F*-statistic

Selection of individuals for simulations

Our results revealed the importance of choosing a strict threshold when using ST in order to meet the criteria suggested by Vähä and Primmer (2006). Indeed, these authors recommended a minimal level of genetic divergence (F_{ST}) between parental populations of 0.21 and 0.12 using respectively 12 and 21 microsatellite loci to detect hybrids. Here, with nine loci and using only individuals with q-value <0.05, we achieved a F_{ST} of 0.24 using ST and 0.23 using NH. Our results thus suggest that even with the typical number of loci (6–10) used in previous studies (see Sanz et al. 2009), sufficient resolution could be achieved when performing admixture analyses.

Detection of simulated hybrids

Analyses performed with 10, 40 or 67% of simulated hybrids showed that both software could adequately detect F1 hybrids, but that the power of detection of the subsequent generations (F2 hybrids and backcrosses) was much more limited. With each software, F2 and backcrosses hybrids overlapped with the parental populations and the proportion of overlap was higher with lower proportion of hybrids in the sample (see also Vähä and Primmer 2006). In general, the overlap was higher for backcrosses, which can be potentially attributed to the repeated backcrosses of admixed individuals with individuals of parental populations (Oliveira et al. 2008). Similar results were found in others studies (Vähä and Primmer 2006; Oliveira et al. 2008; reviewed in Randi 2008; Burgarella et al. 2009; Sanz et al. 2009), and it is likely that the power in the detection of F2 and backcross hybrids could be raised by increasing the number of loci,

Table 5 Number of assigned individuals and categorical assignments of individuals (wild, hybrid and domestic individuals in percentage) of each lake obtained with STRUCTURE (q-values $\geq$0.9 or $\leq$0.1) and NEWHYBRIDS (threshold of 0.7)

Lake	Number of individuals	Categorical assignment					
		Wild individuals		Hybrids individuals		Domestic individuals	
		ST	NH	ST	NH	ST	NH
Heavily stocked lakes							
Belles-de-Jour	81	43.2	0	9.9	44.4	46.9	53.1
Amanites	88	42.0	43.2	13.6	6.8	44.3	44.3
Methot	101	48.5	0	37.6	51.5	13.8	25.7
Average	90	44.6	14.4	20.4	34.2	35.0	41.0
Standard error		3.5	24.9	15.0	24.0	18.4	14.0
Moderately stocked lakes							
Rivard	46	97.8	84.8	2.2	13	0	0
Veillette	101	99.0	99.0	1.0	0	0	0
Arcand	56	96.4	89.3	3.6	8.9	0	0
Circulaire	77	92.2	84.4	7.8	9.1	0	0
Average	70	96.4	89.4	3.7	7.8	0	0
Standard error		3.0	6.8	3.0	5.5	0	0
Global average	78.6	74.2	57.2	10.8	19.1	15.0	17.6
Standard error		27.8	42.9	12.6	20.2	21.5	23.4

The mean percentage of assignments and standard error are given for the heavily and moderately stocked lakes

given the genetically proximity of the parental populations, as suggested by several studies (Vähä and Primmer 2006; Oliveira et al. 2008; Randi 2008; Burgarella et al. 2009). However, it is noteworthy that Albert et al. (2006) were not able to tell apart F2 from BC interspecific hybrids between American (*Anguilla rostrata*) and European ell (*A. anguilla*) even when using 373 AFLP markers. Moreover, the extent of misclassification of hybrids may be also reduced by using a more severe threshold to increase the accuracy at the expense of efficiency (Vähä and Primmer 2006; Burgarella et al. 2009).

Efficiency and accuracy of software

Our quantitative assessment of software efficiency revealed significant differences among software combinations. More specifically, we found that ST showed a higher efficiency than NH, especially when assigning hybrid simulated individuals, both under scenarios of low, intermediate and high proportion of hybrids. Burgarella et al. (2009) previously showed that when a low proportion of hybrids were present in their sample (2%), the efficiency of ST was higher than NH. Also, Vähä and Primmer (2006) showed that both software possessed a similar efficiency when the proportion of hybrids was around 10% in the sample but that NH efficiency decreased more rapidly than ST when the proportion of hybrids was smaller (1%). Thus, our results, in combination with previous published evidences, suggest that ST should generally show a greater efficiency and be less prone to fluctuations in the number of hybrids present in the sample. We also found that NH was more accurate than ST to assign simulated individuals, as supported previously (Vähä and Primmer 2006; Burgarella et al. 2009). Differences in the performance of software could be mainly explained by their respective modelling approaches rather than by their algorithms, which are equivalent. The main difference stands in the underlying assumption of each software with respect to the presence of hybrids in the populations. Indeed, NH models evaluate directly the posterior probability that each individual belongs to a parental or hybrid class, because it assumes that hybridization is occurring in the populations studied. On the other hand, the ST models evaluate the posterior probability of an individual of belonging to a population, without assumptions regarding the presence of hybridization.

Thus, ultimately the choice of the software to detect hybrids will clearly depend on the main objective of the study. The high efficiency of ST should be appealing for conservation studies aiming to assess the presence of hybrids in wild populations, especially given the difficulty to predict a priori the proportion of expected hybrids in such samples (see Marie et al. 2010, for example). Alternatively, the use NH should be favored in studies where hybridization is known to occur and where the main aim is to accurately assess the number of hybrids present in a

subset of individuals (see Adams et al. 2007, for example). Obviously, the availability of parental baseline genetic information should increase the efficiency and accuracy of software as estimations of allele frequencies are done without errors in such cases.

Effect of stocking intensity on the efficiency and the accuracy of software

We found no effect of stocking intensity on the software's assignment efficiency. Previous studies showed qualitatively that the efficiency of software may be influenced by the proportion of hybrids in the sample but in equivocal fashion. For example, Vähä and Primmer (2006) showed that the efficiency of software decreased when fewer hybrids were present. At the opposite, Sanz et al. (2009) showed that the efficiency of the assignment method was reduced with greater levels of introgression in the samples. Our analyses, however, revealed a significant effect of stocking intensity on the assignment accuracy in samples including 40 and 67% of hybrids. The accuracy was higher in heavily stocked lakes than in moderately stocked lakes, suggesting that a greater resolution is achieved with more hybridization in the sample.

Our study represents the first quantitative assessment of the assignment accuracy with different proportions of hybrids in samples and thus further studies are required to conclude on the generality of our findings. Yet, our conclusions are somewhat different from those reached by Vähä and Primmer (2006) and Sanz et al. (2009). Several technical reasons could explain these differences and here we emphasize only the most obvious ones. First, the three studies differ in terms of proportions of simulated hybrids. In their study, Sanz et al. (2009) used only one proportion of simulated hybrids (about 29%) whereas Vähä and Primmer (2006) investigated a range of hybrid proportion varying between 1% versus 10% that was smaller than the range we covered here (10% vs. 40% vs. 67%). Second, our study is the only one that pooled the second generation (or more) hybrids for the analyses, given the problems related to discriminating hybrid status beyond the F2 generation. Finally, the accuracy of stocking intensity differs among studies. Sanz et al. (2009) considered the stocking intensity as being binary (presence or absence), whereas we used different level of stocking (moderately or highly stocked).

Assessment of hybridization in wild populations

Our admixture analyses revealed that the assigned proportions of wild, hybrid and domestic individuals were significantly influenced by the stocking intensity at a given location. More specifically, and as expected, our results showed that the proportion of wild individuals was significantly higher in moderately stocked lakes than in heavily stocked lakes, whereas the opposite trend appeared for the domestic and hybrid individuals. Thus, these results confirm the genetic impact of the stocking practices on wild populations, with a greater potential of introgressive hybridization in heavily stocked lakes rather than moderately stocked lakes. Similar conclusions were reported by Marie et al. (2010), who showed that the mean individual admixture of brook charr in lakes of two wildlife reserves in Quebec (Canada) increased significantly with the number of stocking events performed in these lakes. Hansen and Mensberg (2009) also showed that rivers that were more intensely stocked with brown trout (*Salmo trutta*) showed higher levels of introgression.

Based on our results from simulated individuals, we expected that ST would perform better than NH in terms of assignment efficiency since stocking in our lakes result in the presence of hybrids in our sample. Yet, the only difference in individual's assignment was found for wild individuals for which the number of individuals assigned was higher with ST than NH. This was especially evident in heavily stocked lakes where ST assigned on average three times as many wild individuals than NH (Table 5). Such difference, reflecting a possible trade-off between efficiency and accuracy, is a little worrying as it could greatly impact interpretations on the consequences of stocking. Indeed, if ST was the software used to conduct admixture analyses, one could conclude that wild individuals are still common under conditions of intense stocking. Yet, if the analyses were performed with NH only, the main conclusion would be totally opposite: that intense stocking reduce the number of wild individuals detected. Thus, in order to efficiently understand the effect of stocking it might be necessary to use the two software in combination: using first ST to detect the presence of hybrids and then NH to assess the number of hybrids. Other studies also showed differences between simulations and real case scenarios. Sanz et al. (2009), for example, compared the efficiency of four Bayesian assignment methods to detect the level of admixture in stocked populations of brown trout in Spain. From simulated individuals, the authors showed that any combination of markers and methods gave qualitatively similar conclusions, whereas ST seemed to be the best choice to detect admixture in wild populations. Thus, even though simulations can help to choose the software with the best performances, our study and that of Sanz et al. (2009) suggest that results obtained under real contexts can sometimes provide inconsistent results.

Conclusion

When aiming at assessing the level of introgression between genetically related parental populations of the

same species (e.g. wild and domestic populations), the choice of an appropriate threshold to select individuals for simulations appeared to be of importance to respect the recommendations of Vähä and Primmer (2006) and should thus be tested a priori. Our results confirmed previously published evidences that a high number of microsatellite loci is required in order to go beyond the detection of F1 hybrids (Vähä and Primmer 2006; Oliveira et al. 2008; Burgarella et al. 2009; Sanz et al. 2009). Analyses of efficiency suggested that ST should be used when performing admixture analyses independently of the expected proportion of hybrids in the sample. Moreover, it appeared that NH had a higher accuracy, which increased with stocking intensity. Our results also confirmed the genetic impact of the stocking practices on wild populations, with a greater introgressive hybridization in lakes subjected to intensive stocking. Yet, the differences detected among software in terms of number of wild individuals assigned suggest that simulations and real case scenarios might provide contrasted results.

Acknowledgments We are grateful to Fabien Lamaze, Bruno Mayot, Marc-André Poulin, Andréanne Lessard and Mélissa Lieutenant-Gosselin for their help with field sampling and laboratory analyses. We also thank Associate editor Michael Schwartz and two anonymous reviewers for their helpful comments on a previous version of the manuscript. We acknowledge the important contributions of the Ministère des Ressources Naturelles et de la Faune du Québec (MRNF), the Société des Établissement de Plein-Air du Québec (SÉPAQ) and the Réseau Aquaculture Québec (RAQ) to the project. This research was financially supported by a strategic project grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to LB and DG.

References

Adams JR, Lucash C, Schutte L, Waits LP (2007) Locating hybrid individuals in the red wolf (*Canis rufus*) experimental population area using a spatially targeted sampling strategy and faecal DNA genotyping. Mol Ecol 16:1823–1834

Albert V, Jónsson B, Bernatchez L (2006) Natural hybrids in Atlantic eels (*Anguilla anguilla, A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. Mol Ecol 15:1903–1916

Aljanabi SM, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucl Acids Res 25:4692–4693

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229

Ayres DR, Garcia-Rossi D, Davis HG, Strong DR (1999) Extent and degree of hybridization between exotic (*Spartina alterniflora*) and native (*S. foliosa*) cordgrass (Poaceae) in California, USA determined by random amplified polymorphic DNA (RAPDs). Mol Ecol 8:1179–1186

Biebach I, Keller LF (2009) A strong genetic footprint of the re-introduction history of Alpine ibex (*Capra ibex ibex*). Mol Ecol 18:5046–5058

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

Burgarella C, Lorenzo Z, Jabbour-Zahab R, Lumaret R, Guichoux E, Petit RJ, Soto A, Gil L (2009) Detection of hybrids in nature: application to oaks (*Quercus suber* and *Q. ilex*). Heredity 102: 442–452

Englbrecht CC, Schliewen U, Tautz D (2002) The impact of stocking on the genetic integrity of Arctic charr (*Salvelinus*) populations from the Alpine region. Mol Ecol 11:1017–1027

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620

Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. Evol Appl 1:535–586

Gagnaire PA, Albert V, Jonsson B, Bernatchez L (2009) Natural selection influences AFLP intraspecific genetic variability and introgression patterns in Atlantic eels. Mol Ecol 18:1678–1691

Goudet J (1995) FSTAT (Version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Grandjean F, Verne S, Cherbonnel C, Richard A (2009) Fine-scale genetic structure of Atlantic salmon (*Salmo salar*) using microsatellite markers: effects of restocking and natural recolonization. Fresh Biol 54:417–433

Gunnell K, Tada MK, Hawthorne FA, Keeley ER, Ptacek MB (2008) Geographic patterns of introgressive hybridization between native Yellowstone cutthroat trout (*Oncorhynchus clarkia bouvieri*) and introduced rainbow trout (*O. mykiss*) in the South Fork of the Snake River watershed, Idaho. Conserv Genet 9:49–64

Guyomard R (1997) Consequences of fish introduction in the absence of reproductive isolation: interest and limits of the experimental approach. Bull Fr Pêche Piscic 344–345:301–308

Halbisen MA, Wilson CC (2009) Variable introgression from supplemental stocking in Southern Ontario populations of lake trout. Trans Am Fish Soc 137:699–719

Hansen MM, Mensberg KLD (2009) Admixture analysis of stocked brown trout populations using mapped microsatellite DNA markers: indigenous trout persist in introgressed populations. Biol Lett 5:656–659

Hansen MM, Nielsen EE, Bekkevold D, Mensberg KLD (2001) Admixture analysis and stocking impact assessment in brown trout (*Salmo trutta*), estimated with incomplete baseline data. Can J Fish Aquat Sci 58:1853–1860

Hindar K, Ryman N, Utter F (1991) Genetic effects of cultured fish on natural fish populations. Can J Fish Aquat Sci 48:945–957

Kidd AG, Bowman J, Lesbarrères D, Schulte-Hostedde AI (2009) Hybridization between escaped domestic and wild American mink (*Neovison vison*). Mol Ecol 18:1175–1186

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

Ministère des Ressources Naturelles et de la Faune du Québec (2008) Lignes directrices sur les ensemencements de poissons. Secteur Faune Québec, Direction de l'expertise sur la faune et ses habitats. Québec. 41p

Nielsen EE, Hansen MM, Bach L (2001) Looking for a needle in a haystack: discovery of indigenous salmon in heavily stocked populations. Conserv Genet 2:219–232

Nielsen EE, Hansen MM, Ruzzante DE, Meldrup D, Gronkjaer P (2003) Evidence of a hybrid-zone in Atlantic cod (*Gadus morhua*) in the Baltic and the Danish Belt Sea revealed by individual admixture analysis. Mol Ecol 12:1497–1508

Oliveira R, Godinho R, Randi E, Alves PC (2008) Hybridization versus conservation: are domestic cats threatening the genetic

integrity of wildcats (*Felis silvestris silvestris*) in Iberian Peninsula? Phil Trans R Soc B 363:2953–2961

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. Mol Ecol 17:285–293

Sanz N, Araguas RM, Fernandez R, Vera M, Garcia-Marin JL (2009) Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. Conserv Genet 10:225–236

Sauvage C, Derôme N, Normandeau É, St-Cyr J, Audet C, Bernatchez L (2010) Fast transcriptional responses to domestication in the brook charr *Salvelinus fontinalis*. Genetics 185:105–112

Susnik S, Berrebi P, Dovc P, Hansen MM, Snoj A (2004) Genetic introgression between wild and stocked salmonids and the prospects for using molecular markers in population rehabilitation: the case of the Adriatic grayling (*Thymallus thymallus* L. 1785). Heredity 93:273–282

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol Ecol 15:63–72

Weir LK, Grant JWA (2005) Effects of aquaculture on wild fish populations: a synthesis of data. Environ Rev 13:145–168