

Major disruption of gene expression in hybrids between young sympatric anadromous and resident populations of brook charr (*Salvelinus fontinalis* Mitchill)

JESUS MAVAREZ,^{*,†} CELINE AUDET[‡] & LOUIS BERNATCHEZ^{*}

^{*}Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, Canada

[†]Centro de Ecología, Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela

[‡]Institut des sciences de la mer de Rimouski, Université du Québec à Rimouski, Rimouski, Canada

Keywords:

ecological divergence;
fish;
genomic incompatibilities;
hybridization;
microarrays;
reproductive isolation.

Abstract

Genome-wide analyses of the transcriptome have suggested that male-biased genes are the first targets of genomic incompatibilities (g.i.) in inter-specific hybrids. However, those studies have almost invariably focused on *Drosophila* species that diverged at least 0.9 Ma, and with sterile male hybrids. Here, we use microarrays to analyse patterns of gene expression in very closely related (divergence <12 000 years), sympatric, but ecologically divergent anadromous and resident populations of brook charr (*Salvelinus fontinalis*) and their F₁ hybrids. Our results show a dramatic breakdown of gene expression patterns in hybrids compared with their parental relatives. Several disrupted genes are related to energetic metabolism, immune response, osmoregulation and protection against oxidative stress, and none has sex-biased functions. Besides, pure individuals show no expression differences at most of the genes disrupted in hybrids, which may suggest the operation of some form of stabilizing selection. Taken together, these results both confirm the idea that perturbations of regulatory networks represent a significant source of g.i. and support the suggestion that developmental pathways can diverge through time without any manifest change in the phenotypic outcome. While the role of other evolutionary forces (e.g. genetic drift) cannot be ruled out, this study suggests that ecological selective processes may provide the initial driving force behind disruption of gene expression in inter-specific hybrids.

Introduction

Understanding the causes of reproductive isolation (and hence speciation) remains one of the major tasks faced by evolutionary biologists. Traditionally, reproductive isolation has been classified into pre- and post-zygotic processes, with more emphasis and studies being devoted to the later (Coyne & Orr, 2004). Current comprehension of post-zygotic isolation is based on the general principle that hybrid sterility and/or inviability arises as a consequence of incompatible hybrid allele combinations that are fixed in pure species or populations (Dobzhansky,

1937; Muller, 1942; Orr, 1997; Johnson & Porter, 2000; Orr & Presgraves, 2000; Porter & Johnson, 2002). Thus, alleles may be neutral or positively selected within the genetic background of two different species but interact negatively when brought together in inter-specific hybrid backgrounds. For instance, analyses of coding sequence data have suggested that amino acid differences driven by divergent selection between species may render proteins incapable of interacting correctly in hybrids (e.g. Ting *et al.*, 1998; Rawson & Burton, 2002; Barbash *et al.*, 2003; Brideau *et al.*, 2006). Evidence of such epistatic interactions may also be established through genome-wide analyses of gene expression profiles such as those performed with microarrays or similar techniques (Haerty & Singh, 2006; Wang *et al.*, 2006; Hedgecock *et al.*, 2007; Malone *et al.*, 2007; Moehring *et al.*, 2007; Rottschmidt & Harr, 2007).

Correspondence: Louis Bernatchez, Institut de Biologie Intégrative et des Systèmes (IBIS), Pavillon Charles-Eugène-Marchand, Université Laval, Québec (Québec) G1V 0A6, Canada.
Tel. +1 418 656 3402; fax +1 418 656 7176;
e-mail: louis.bernatchez@bio.ulaval.ca

Detailed analyses of disruption of gene expression profiles in inter-specific hybrid animals have been performed in *Drosophila*, particularly in the *melanogaster* group (Michalak & Noor, 2003; Ranz *et al.*, 2004; Landry *et al.*, 2005; Haerty & Singh, 2006; Moehring *et al.*, 2007), and to a lesser extent in *Xenopus* (Malone *et al.*, 2007) and *Mus* (Rottscheidt & Harr, 2007). These studies have led to two main conclusions: (i) that a significant number of genes are highly under-expressed in hybrids compared with pure species, which supports the hypothesis that divergence in gene-regulation networks is a major contributor to Dobzhansky–Muller incompatibilities (Johnson & Porter, 2000; Porter & Johnson, 2002; Landry *et al.*, 2007; Ortiz-Barrientos *et al.*, 2007) and (ii) that a high proportion of those differentially expressed genes are male-biased (i.e. genes with a higher or exclusive expression in males or the male germ line), supporting the idea that genes with a fast evolutionary rate, such as those associated with the male reproductive function contribute disproportionately to speciation, i.e. faster male theory (Wu *et al.*, 1996; Singh & Kulathinal, 2000; Haerty & Singh, 2006). However, the generality of these conclusions must be considered with care for several reasons. First, comparisons involve species that diverged hundreds of thousands or even millions of years ago such as in *Drosophila*, e.g. *D. simulans*–*D. mauritiana*; 0.93 Ma, *D. simulans*–*D. melanogaster*; 5.1 Ma (Tamura *et al.*, 2004), *Xenopus*, e.g. *X. laevis*–*X. muelleri*; >20 Ma (Evans *et al.*, 2004, 2005), or *Mus*, e.g. *M. musculus* subspecies; 0.3–1.0 Ma (Boursot *et al.*, 1993, 1996), which makes it difficult to determine how much of the incompatibilities observed are indeed related to speciation *per se* or cumulated after the two species had become reproductively isolated. Secondly, these studies almost invariably focus on both male-expressed genes and testes, which in spite of their presumed relevance to speciation obviously represent but a fraction of an organism's transcriptome and therefore cannot be used as a general indication of the importance of gene mis-expression to Dobzhansky–Muller incompatibilities. On the other hand, inter-specific hybrids might suffer from other kind of incompatibilities or reduced fitness not related to sterility, such as inviability, slower growth or poorer performance. Thirdly, all the analyses were performed exclusively in laboratory model organisms and with a manifest predominance of studies in *Drosophila*, which certainly represents a bias. It is therefore clear that the extension of this kind of studies to non-model organisms will be necessary in order to both evaluate the generality of the current trends and to reduce the knowledge gap between the ecological genetics of organisms with and without a strong laboratory genomics toolkit.

The main objective of this study was to compare the transcriptome of whole individuals (i.e. to avoid any bias introduced by focusing on specific tissues or body functions), of two closely related, but ecologically divergent populations of the brook charr (*Salvelinus fontinalis*)

and their F₁ hybrids. As is usual among salmonids (Magnan *et al.*, 2002), *S. fontinalis* is able to exploit a large set of habitats via its impressive variance of phenotypic traits associated with trophic morphology, i.e. limnetic vs. benthic forms (Proulx & Magnan, 2004; Bertrand *et al.*, 2008), or migratory behaviour, i.e. non-migratory forms that remain in freshwater their entire life (resident) vs. seawater migratory forms that return to freshwater to spawn (anadromous) (Lenormand *et al.*, 2004). Migratory strategies are of particular interest because they have motivated a series of studies related to the physiological (Boula *et al.*, 2002; Morinville & Rasmussen, 2003, 2006), morphological (Theriault & Dodson, 2003; Morinville & Rasmussen, 2008), behavioural (Theriault *et al.*, 2007a,b) and genetic (Boula *et al.*, 2002; Perry *et al.*, 2004, 2005) differences between the two forms. Genetic studies in particular have revealed that resident and anadromous brook charr living in sympatry/parapatry in the Laval River, Québec (Canada), show significant population structure ($F_{st} = 0.15$), which suggests some form of reproductive isolation (Boula *et al.*, 2002; Perry *et al.*, 2005). However, our own behavioural observations in the laboratory, together with inferences about the mating patterns in the wild based on parentage analyses of microsatellite data revealed no strong pre-zygotic isolation besides some size homogamy, where anadromous individuals tended to be larger (Theriault *et al.*, 2007a). This indicates that isolation between the two forms at the Laval River is mainly post-zygotic, which raises the hypothesis that reduced hybrid viability and/or performance related to ecological factors may contribute to their reproductive isolation in their natural environment.

Material and methods

Sampling

Resident and anadromous adult brook charr broodstock were collected from the Adam Brook and the Laval River mainstream, respectively, both in the Laval River drainage where spawning sites of both populations are parapatric without physical barriers between them (Fig. 1) (Boula *et al.*, 2002). The individual progenitors used consisted of wild-caught residents and of the 2nd generation of anadromous maintained in the laboratory. Five full-sib families of each type were designed: Anadromous (♀A♂A), resident (♀R♂R) and F₁ hybrid (anadromous female × resident male, ♀A♂R). All 15 families were raised in identical conditions at the Station aquicole de Pointe-au-Père, Université du Québec à Rimouski (Rimouski, Canada), following a 12–12 h light:dark photoperiod regime. Eggs were incubated at 4 °C. From hatching to first-feeding stage (i.e. yolk sac resorption), water temperature was gradually raised (1 °C per week) to reach 8 °C. At this stage, individuals were preserved for further analysis by immediate freezing at –80 °C.

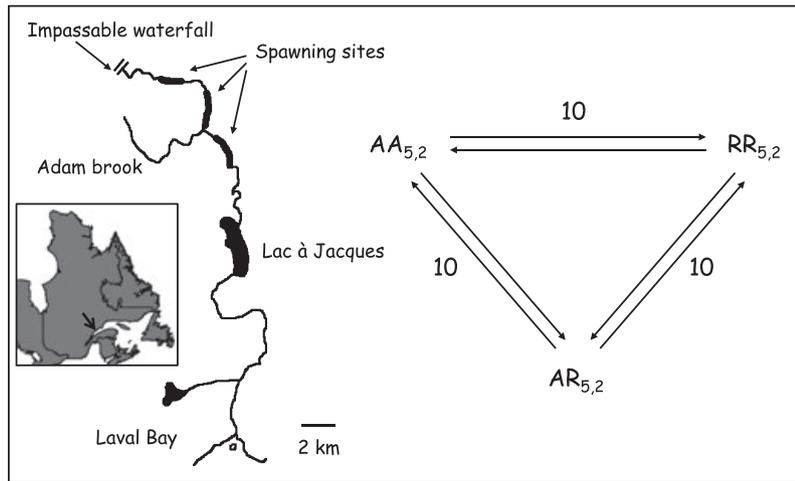


Fig. 1 Map of Québec (insert box) indicating the location of the Laval River (left). Anadromous brook charr were collected close to their spawning sites in the upper Laval River, while resident fish were collected in the Adam brook. Analysis of microarray data (right) follows a loop design with 10 loops using five families of each type (anadromous AA, resident RR and hybrids AR) and two individuals per family.

Experimental design

Transcription profiles of 30 fish (two individuals for each of the fifteen families) were hybridized on the same number of cDNA microarray slides using a loop-design (Fig. 1). Thus, there were three kinds of slides depending on the comparisons assayed: AA vs. RR, AR vs. AA and AR vs. RR. Each individual was used twice; once labelled with Cy3 and once with Alexa 647 (i.e. dye swap). This design combines the relatively high statistical power of the simultaneous and direct comparison of the three types of fish with the relatively reduced cost of a small number of slides (Draghici, 2003). The microarray slides used were fabricated by the consortium for Genomic Research on All Salmon Project (cGRASP) and comprise 16 006 genes obtained from over 175 salmonid cDNA libraries derived from a large set of tissues and developmental stages of *Salmo salar* and *Oncorhynchus mykiss* (von Schalburg *et al.*, 2005). Detailed information on the genes printed in the microarrays can be found at the cGRASP website (<http://web.uvic.ca/grasp/array.html>).

RNA extraction, labelling and cDNA hybridization

Whole frozen individuals (~0.1 g) were homogenized using 1 mL of TRIzol® reagent (Amersham, Amersham, UK) on a Tissue-Lyser (Qiagen, Valencia, USA) following the manufacturer's protocol. Briefly, 200 µL of chloroform (Sigma-Aldrich, St Louise, USA) were added to each homogenate, mixed thoroughly and centrifuged for 15 min at 4 °C and 12 000 g. The aqueous layer was separated into a new tube containing 1 mL of isopropanol (Sigma-Aldrich) and stored overnight at -80 °C. The day after, the samples were centrifuged for 1 h at 4 °C and 12 000 g, and the isopropanol discarded. The remaining pellets were washed in 70 mL of ethanol, dried for 15 min at room temperature and resuspended in 200 µL of nuclease-free water and 1 µL of RNAase

inhibitor (Ambion, Austin, USA). In order to exclude any co-precipitant that might disturb retro-transcription reactions (see below), each RNA extraction was further cleaned with Microcon YM30 columns (Amicon, Billerica, USA), following the manufacturer's protocol. RNA integrity and concentration was independently verified on an Experion (Bio-Rad, Hercules, USA) and a Gene-Quant (Pharmacia, Uppsala, Sweden). For each sample, 15 µg of RNA were retro-transcribed and labelled using the 3DNA Array 50 kit (Genisphere, Hatfield, USA), Superscript II retro-transcriptase (Invitrogen, Carlsbad, USA), Alexa 647/Cyanine 3 dyes (Genisphere) and microarray hybridization chambers (Corning, Lowell, USA), following the modifications of the manufacturers' protocols available at the cGRASP website. Briefly, 15 µg of RNA were reverse-transcribed using oligo (dT) primers with a 5' unique sequence overhang for the labelling reaction. Microarray slides were prepared for hybridization by washing twice in 0.1% SDS for 5 min, five times in milliQ water for 1 min and dried by centrifugation at 512 g for 5 min. The cDNA were hybridized to the slides in formamide-based buffer (25% formamide, 4× SSC, 0.5% SDS, 2× Denhardt's solution) with competitor DNA (LNA dT bloker, Genisphere and human COT-1 DNA, Sigma) at 51 °C for 16 h, using a microarray hybridization chamber (Corning) in a water bath. The arrays were washed once at 45 °C for 5 min (2× SSC, 0.1% SDS), twice at room temperature for 3 min (2× SSC, 0.1% SDS), twice at room temperature for 3 min (1× SSC) and twice at room temperature for 3 min (0.1% SSC), and dried by centrifugation as described before. The Cy3 and Alexa 647 dyes (3DNA capture reagent, Genisphere) were then hybridized to the complementary sequences in the dT oligos linked to the bound cDNA on the microarray slides, using the same hybridization reagents/protocols as before, for 2 h at 51 °C. The arrays were then washed and dried as previously described. Hybridization signals were detected with a ScanArray scanner (Packard BioScience, Meriden, USA), while spots were

located and quantified with QuantArray 3.0 (Perkin Elmer, Waltham, USA).

Data analysis

Local background was removed from the dataset by keeping only values higher than the average intensity of empty spots plus twice their standard deviation in both channels (Cy3 and Alexa 647), which left 6561 gene clones that were considered significantly expressed in all three experimental groups for subsequent analyses. All remaining algorithms and analyses were performed in the R/MAANOVA package (<http://www.jax.org/staff/churchill/labsite/software/Rmaanov/>), unless stated otherwise. Missing values were then imputed using the K-nearest neighbours (15 neighbours). This data set was then \log_2 -transformed, corrected for intensity-linked distortions by using a regional LOWESS algorithm and fitted to a mixed model ANOVA for microarray data (Kerr *et al.*, 2000, 2002). The model considers as random the terms 'Array' (slides 1 to 30) and 'Sample' (ind. 1 to 30), and as fixed the terms 'Dye' (Cy3 and Alexa 647) and 'Type' (AA, AR, RR). A permutation-based *F*-test ($F_{2,9} \geq 4.26$, $P \leq 0.05$), (Cui *et al.*, 2005) was performed and restricted maximum likelihood was used to solve the mixed model equations in R/MAANOVA (*F*-values, with 1000 sample ID permutations). In order to cope with the false positives problem due to the multiple testing typical of microarray experiments (Draghici, 2003), we applied a false discovery rate procedure (FDR, Benjamini & Hochberg, 1995; Reiner *et al.*, 2003) to the ANOVA permutation-based *P*-values and only selected values with an FDR lower than 5%. Pairwise comparison tests ($t_{16} \geq 2.58$, $P \leq 0.01$, RR vs. AA, AR vs. AA and AR vs. RR) were performed on the subset of FDR-significant genes using a test-wise $\alpha = 0.01$, which translates into a rather conservative overall type-I error per gene of 0.029. For each significant gene, we calculated the \log_2 of the average fold-change, defined as the ratio of expression values in the two groups considered in the contrast, e.g. $\log_2(\text{RR}/\text{AA})$. Hereafter, fold-change refers to the \log_2 values.

Functional classification of differences in gene expression

For each significant gene, we used the BLAST-gathered gene annotations available on the cGRASP website in order to obtain biological process information from the Gene Ontology (GO) database (<http://www.geneontology.org>). Gene and protein accession numbers were converted into Unigene cluster codes using the DAVID web-based conversion tool (<http://david.abcc.ncifcrf.gov/conversion.jsp>). These were used as input for PANTHER 2.2 (<http://www.pantherdb.org/>) (Thomas *et al.*, 2003) in order to evaluate the differences in GO categories between all features that were significantly

expressed ($n = 6561$, see Results) and those that were significantly different between groups (using Fisher exact tests).

Results

The ANOVA model revealed that a high proportion (1255/6561 $\approx 20\%$) of the features analysed show levels of expression that were significantly different in at least one of the three groups, which is four times higher than the number of false positives expected by chance alone ($n = 328$). After FDR correction, the number of significant genes still remains considerably high ($n = 313$).

Gene expression differences between pure forms

Of these 313 genes, only 18 showed significant differences in expression in the pure (RR vs. AA) forms comparison (Fig. 2). This value represents just 0.27% of the total number of genes analysed and 5.75% of the number of significant genes after FDR. The average fold-change in this comparison, i.e. \log_2 of the ratio RR/AA, was 0.301 (SD = 0.152), which also appeared relatively small when compared with the other contrasts (see below). More genes were under-expressed in resident relative to anadromous fish (72.22% vs. 27.78%), although this difference was marginally not significant ($Z = 1.89$, $P = 0.059$). Similarly, differences in fold-change between genes down-regulated (0.314, SD 0.174) and up-regulated (0.269, SD 0.067) were not significant ($t_{16} = 0.79$, $P = 0.222$).

Given that the low number of genes used raised some doubts about the statistical power of the aforementioned tests, we performed another *t*-test for this pairwise comparison using an $\alpha = 0.05$ ($t_{16} \geq 1.75$, $P \leq 0.05$), which resulted in a list of 31 significant genes. This adjustment has a double justification. First, an $\alpha = 0.01$ can indeed be too conservative when differences in gene expression are small. Secondly, none of the statistical tests assayed, i.e. the proportion of down- and up-regulated genes (72.22% and 27.28% vs. 64.52% and 35.48%), the average fold-change for all genes (0.301 vs. 0.288) or the average fold-change for down- and up-regulated genes (0.314 and 0.269 vs. 0.306 and 0.256), appear to be significantly different between the RR vs. AA ($\alpha = 0.01$) and RR vs. AA ($\alpha = 0.05$), respectively. Therefore, the lack of significant differences observed between RR and AA fishes seems biologically meaningful and not the consequence of low statistical power.

There is reliable and detailed Gene Ontology information of biological processes for 14 of the significant genes in the RR vs. AA comparison (Table 1a). These comprised many genes involved in electron transport and energy metabolism (i.e. aldehyde dehydrogenase 9 family, NADH-ubiquinone oxidoreductase chain 1, NADH-ubiquinone dehydrogenase flavoprotein 3 and pyruvate

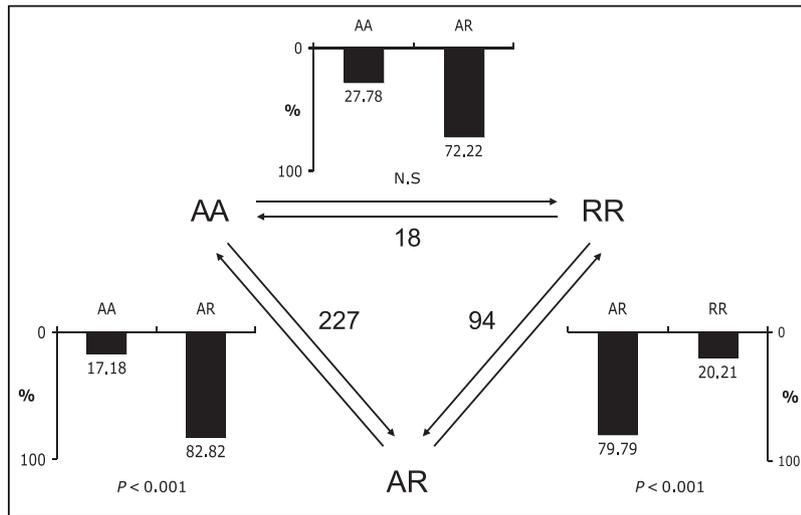


Fig. 2 Number of genes with significant differences in expression levels in each of the three *t*-test comparisons; RR vs. AA, AR vs. AA and AR vs. RR. Bar graphs show the percentage of under-expressed genes of each fish type per comparison (e.g. 17.18% of the 227 significant genes in the AR vs. AA comparison are under-expressed in the AA individuals). *P*-values correspond to *Z*-tests for proportions.

kinase) and several transport/carrier proteins, including a sodium chloride co-transporter associated to re-absorption of this salt (Simon *et al.*, 1996).

Gene expression differences in F₁ hybrids vs. pure forms

In contrast to the pure form comparison, a very high number of genes were differentially expressed in the hybrid/pure forms comparisons (Fig. 2). In the case of AR vs. AA fish, 227 genes showed a significant difference of expression, which represents 3.46% of the total of genes analysed and 72.52% of the set of significant genes after FDR. On the other hand, 94 genes were differentially expressed between AR vs. RR fish, which represents 1.43% of the total number of genes and 30.03% of the FDR significant set. These differences between hybrid and pure fish are better appreciated when contrasted to the comparisons between pure forms described above (Fig. 2). Thus, a five-fold increase in the number of differences (94 vs. 18) was observed when the AR vs. AA comparison is contrasted to the RR vs. AA. Moreover, an even higher 12-fold increase (227 vs. 18) was observed when the AR vs. RR comparison is contrasted to the RR vs. AA.

Moreover, there was a clear pattern of under-expression in the AR F₁ hybrids. Thus, 82.8% (188/227) of the differentially expressed genes in the AR vs. AA comparison were down-regulated in the AR individuals ($Z = 9.89$, $P < 0.0001$). A similar pattern was also observed in the AR vs. RR comparison, where 79.79% (75/94) of the significant genes show a lower level of expression in the AR individuals ($Z = 5.78$, $P < 0.0001$). These percentages of down-regulated genes in the AR individuals were not significantly different between the AR vs. AA and the AR vs. RR comparisons (Fisher exact test, $P = 0.101$).

Analyses of fold-change expression differences (i.e. fold-changes) also revealed some interesting patterns (Fig. 3). Namely, the average (SD) fold-changes observed in our comparisons were: 0.301 (0.152) in RR/AA, 0.334 (0.238) in AR/AA and 0.499 (0.267) in AR/RR. Pairwise *t*-tests indicated that fold-changes in the RR/AA contrast were significantly lower from those observed in the AR/RR ($t_{110} = 3.05$, $P = 0.001$), and lower but not significantly different from those in the AR/AA ($t_{243} = 0.57$, $P = 0.28$) comparison, suggesting that differences in expression were particularly strong in comparisons involving hybrids. Moreover, inspection of fold-change data also showed the existence of a striking asymmetry between down- and up-regulated genes *within* a single comparison. In both the AR/AA and AR/RR cases, fold-change differences between down- and up-regulated genes were highly significant (*t*-tests, both $P < 0.0001$). Thus, not only the majority of genes in those comparisons were down-regulated in hybrids but also the average level of expression for genes down-regulated in hybrids was significantly much lower than for genes down-regulated in pure forms. In contrast, there was no significant difference between down- and up-regulated genes in the RR/AA comparison. In sum, these results clearly indicate that hybridization between anadromous and resident brook charr was accompanied by a massive breakdown of gene regulation in F₁ individuals in the form of a majority of transcripts exhibiting low levels of expression relative to pure forms.

Functional analysis of gene disruption in hybrids

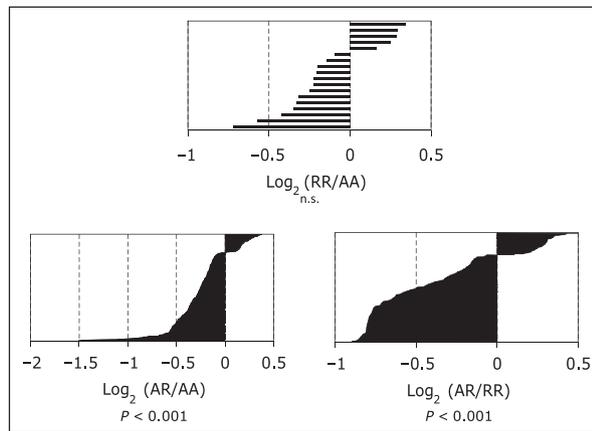
Functional analyses revealed that disrupted genes in F₁ hybrids represented 23 different biological processes (all *P*-values in this section were obtained through Fisher's exact test) (Fig. 4). Several categories associated with energy metabolism (i.e. oxidative phosphorylation,

Table 1 (a) Significant genes with GO information in the Resident (RR) vs. Anadromous (AA) brook charr comparison. cGRASP ID corresponds to EST clone identification (full information available in the cGRASP website). GO biological functions are defined by the Gene Ontology database (<http://www.geneontology.org>) and the protein knowledge database at UNIPROT (<http://beta.uniprot.org>). Fold-change represents the \log_2 of the average RR/AA ratio (positive and negative values are genes over- and under-expressed in RR, respectively) and *P*-value (FDR) represents the FDR-corrected (1000 permutations) *P*-value of the ANOVA test (see text for details). (b) Significant genes in the hybrids vs. pure forms comparisons with a \log_2 fold-change (AR/AA or AR/RR) higher than 0.5 (which represents an absolute difference of at least 41%). All listed genes are under-expressed in hybrids.

cGRASP ID	Gene	GO biological function	Log ₂ fold-change	<i>P</i> -value (FDR)
(a)				
CA060078	Pyruvate kinase isozyme M2	Energy metabolism. Glycolysis	-0.294	0.028
CB516580	Kelch-like 6	Protein binding	-0.285	0.037
CB507338	Metalloproteinase inhibitor 2 precursor (TIMP-2)	Protein binding. Negative regulation of cell proliferation	-0.254	0.038
CK990779	Beta-2-microglobulin precursor	MHC-related immune response	-0.166	0.031
CB497043	NADH dehydrogenase (ubiquinone) flavoprotein 3	Electron transport, respiratory chain	0.145	<0.001
CB493934	Aldehyde dehydrogenase 9 family, member A1	Oxireductase activity, respiratory chain	0.203	<0.001
CK990220	Fatty acid binding protein, adipocyte (AFABP)	Lipid binding	0.223	0.038
CB505664	Nucleolar RNA helicase II (DEAD-box protein 21)	Nucleic acid binding	0.251	0.017
CK991112	Annexin A2	Protein binding. Skeletal development and angiogenesis	0.320	0.044
CA058762	Cell proliferation-inducing gene 17	Transmembrane protein	0.331	0.006
CA046928	NADH-ubiquinone oxidoreductase chain 1	Electron transport, respiratory chain	0.352	0.001
CA051446	Solute carrier family 12 member 3	Ion transport. Sodium chloride cotransporter activity	0.427	<0.001
CB492794	Guanosine-3',5'-bis(diphosphate)	Nucleic acid binding	0.571	0.028
CA042416	<i>Salmo salar</i> zonadhesin-like gene	Protein binding	0.722	0.037
(b)				
CB506103	Apolipoprotein A-IV	Lipid transport. Antioxidant activity	1.507	0.043
CB505753	Apolipoprotein A-IV precursor	Lipid transport. Antioxidant activity	1.436	0.023
CB497640	Lysosomal acid lipase 1 cholesteryl ester hydrolase	Lipid degradation. Hydrolase	1.298	0.009
CK990548	Fatty acid binding protein-1, liver	Lipid transport and metabolism. Ligase	1.106	0.028
CB498566	Procollagen, type X alpha 1	Structural protein. Skeletal development	1.003	0.028
CB510634	Immune-related Hdd11	Immune response in <i>Hyphantria cunea</i> (fall webworm)	0.926	0.042
CB489981	Tissue inhibitor of metalloprotein	Protein binding. Negative regulation of cell proliferation	0.893	0.013
CA052615	Lipase A, lysosomal acid, cholesterol esterase	Lipid degradation. Hydrolase	0.886	0.028
CX984314	Leucocyte cell-derived chemotaxin precursor	Chemotaxis. Regulation of chondrocyte proliferation	0.861	0.006
CB503780	Ferritin heavy chain (H)	Iron storage. Oxyreductase	0.849	0.012
CB491393	Ribosomal protein S13	Ribonucleoprotein	0.828	0.011
CA037206	Ferritin middle subunit (M)	Iron storage. Oxyreductase	0.790	0.014
CB509723	Glutathione peroxidase	Oxyreductase. Protects from oxidative stress	0.789	0.027
CB504359	Mitochondrial ribosomal protein S17	Ribonucleoprotein	0.762	0.012
CB503310	GDP dissociation inhibitor	Signal transduction. GTPase activation	0.748	0.023
CB516580	Kelch-like 6	Protein binding	0.709	0.012
CB509797	Apolipoprotein B (including AG(X) antigen)	Lipid transport. Cholesterol metabolism	0.682	0.041
CB492648	Apolipoprotein B 100	Lipid transport. Cholesterol metabolism	0.679	0.043
CA057016	NF-kappaB repressing factor	Transcription regulation	0.666	0.015
CA056693	T-complex protein 1, gamma subunit	Molecular chaperone	0.632	0.012
CA058205	Gamma interferon-inducible thiol reductase	Oxyreductase. MHC class II-antigen processing	0.627	0.008
CA056647	Iron(III)-Zinc(II) purple acid phosphatase	Iron and zinc binding. Hydrolase	0.614	0.006
CB500588	Collagen, type X alpha 1	Structural protein. Skeletal development	0.608	0.028
CB503763	S100 calcium binding protein A1	Protein binding. Calcium and zinc binding	0.583	0.034
CA054909	Annexin A1	Protein binding. Negative regulation of cell proliferation	0.572	0.015
CB510782	Nonhistone chromosomal protein H6 (histone T)	Antibiotic. Antimicrobial	0.563	0.027
CB497693	Beta thymosin	Protein binding. Organization of cytoskeleton	0.559	0.041

Table 1 (Continued)

cGRASP ID	Gene	GO biological function	Log ₂ fold-change	P-value (FDR)
CA056568	Lipopolysaccharide-induced TNF factor	Transcription regulation	0.553	0.008
CK990545	Beta-2 microglobulin	MHC-related immune response	0.553	0.036
CA055287	Procollagen, type I alpha 2	Structural protein. Transmembrane signal transduction	0.551	0.034
CA041451	Thioredoxin (ATL derived factor)	Electron transport. Cell redox homeostasis	0.548	0.027
CA057408	Annexin A2	Protein binding. Skeletal development and angiogenesis	0.527	0.019
CK990715	Syntaxin binding protein	Protein binding. Transport	0.522	0.034

**Fig. 3** Log₂ of fold-change for genes with significant differences in expression levels in each of the three comparisons analysed; RR/AA, AR/AA and AR/RR.

$P = 0.0073$; electron transport, $P = 0.0370$), lipid metabolism ($P = 0.0027$) and nucleic acid metabolism ($P = 0.0004$) were significantly different from random expectations. Table 1b shows the genes with a log₂ fold-change higher than 0.5 (\approx fold difference of 40%) between hybrids and pure forms. The commonest biological functions among the genes with the most severe differences in fold-chance between hybrids and pure forms are related to lipid transport and metabolism (e.g. apolipoprotein A-IV, apolipoprotein B100, lysosomal acid lipase-1 cholesteryl ester hydrolase, fatty acid binding protein-1), immune response (e.g. immune-related Hdd11, beta-2 microglobulin, gamma interferon-inducible thiol reductase), as well as defence against oxidative stress (e.g. glutathione peroxidase, ferritin).

Discussion

Breakdown of expression in hybrids

Our results revealed that when compared to their parental anadromous and resident pure forms of *S. fontinalis*, F₁ hybrid individuals showed a significantly higher proportion of differentially expressed genes, with a clear tendency for an excess of under-expression

(i.e. \sim 80% of mis-expressed genes in hybrids are under-expressed). Moreover, in hybrids, average under-expression was significantly stronger than average over-expression. These results are in agreement with similar studies of genome-wide gene expression in other taxonomic groups, such as *Drosophila* (Meiklejohn *et al.*, 2003; Michalak & Noor, 2003; Ranz *et al.*, 2004; Landry *et al.*, 2005; Haerty & Singh, 2006; Moehring *et al.*, 2007), *Xenopus* (Malone *et al.*, 2007), *Arabidopsis* (Wang *et al.*, 2006), and *Coregonus* (Renaut *et al.*, 2009) in which the transcription levels of tens to hundreds of genes in hybrid individuals showed values that fall outside the range observed in the parental forms, with a significantly higher proportion of under-expression among deregulated genes in inter-specific hybrids. This pattern cannot be explained by a reduced size of juvenile F₁ hybrids in relation to their parental relatives since there are no significant differences in size at this age for any of the three fish types considered (Lefrant, 2006). We also carefully measured the concentration of extracted RNA in order to ensure that the same quantity was used for all the experiments. Another possible explanation for the discrepancy observed between hybrid and pure groups could relate to an overall stress because of the negative interaction between two different genomes in the same cells (McClintock, 1984; Comai *et al.*, 2003; Landry *et al.*, 2007). However, we believe that this is implausible here. Anadromous and resident forms are very closely related, and their postglacial divergence puts an upper limit of approximately 12 000 years since they shared a common ancestor (Castric & Bernatchez, 2003). This appears too short for the two genomes to be considered as 'overall divergent', as also illustrated by the fact that most genes that were significantly expressed did not differ between pure forms.

Alternatively, our results concur to the growing body of evidence, both theoretical (Johnson & Porter, 2000; Porter & Johnson, 2002; Landry *et al.*, 2005) and empirical (Landry *et al.*, 2007; Ortiz-Barrientos *et al.*, 2007; Renaut *et al.*, 2009) supporting the idea that incompatibilities in hybrids are in part due to epistatic interactions among elements of regulatory transcriptional networks. Thus, the necessary co-adaptation of regulatory elements within 'species' could be broken in hybrids, probably due

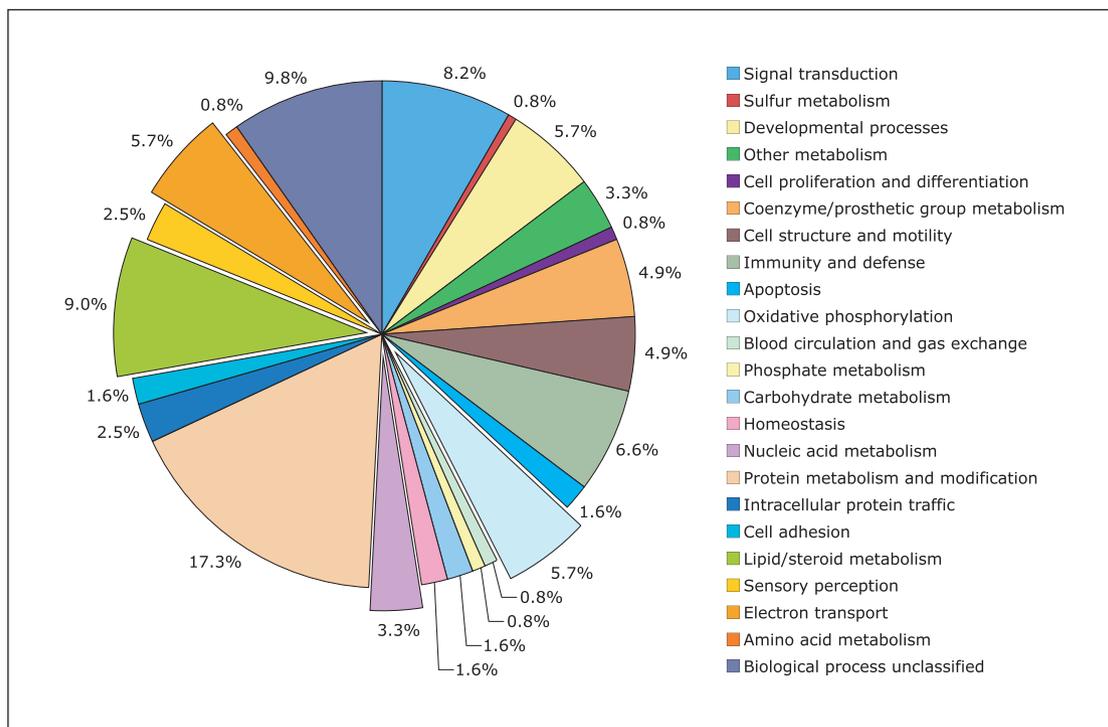


Fig. 4 Pie chart representing the relative frequencies of each of the 23 GO biological functions observed among genes with significant differences in expression levels in the ANOVA analysis. Detached pie sections represent GO functions whose frequencies among deregulated transcripts in hybrids are significantly different from expectations given the gene composition of the cGRASP array.

to failures in the interaction between the transcription factors of one species and their binding sites in the regulatory modules or enhancers of the other species. Although current direct evidence in support of this idea is scarce, it seems highly plausible given our knowledge of the regulation of transcription dynamics (see reviews and models in Arnosti, 2003; Levine & Tjian, 2003; Wray *et al.*, 2003; Landry *et al.*, 2007; Segal *et al.*, 2008). For instance, Inga *et al.* (2005) have shown that the trans-activation capacity of several mutants of a transcription factor required for organogenesis of the human heart varies from full activity to complete loss of function, depending on the amino-acid substitutions in their binding domains and the sequence of target elements in the DNA. On the other hand, a single lysine to methionine substitution in the binding domain of a transcription factor controlling anthocyanin biosynthesis in the beefsteak plant (*Perilla frutescens*) can increase its trans-activation capacity by 50-fold, but an alanine to aspartic acid substitution in a neighboring residue completely eliminates it (Pattanaik *et al.*, 2008). It is therefore reasonable to expect that if two populations or species fixed for different transcriptional interactions come into contact and hybridize, the resulting F₁ hybrids can suffer from these kinds of deregulations and will show a certain number of mis-expressed genes when compared to their parents. Arguably, the actual number of interactions

causing incompatibilities could be much smaller than the observed number of mis-expressions, especially out of microarray experiments, given the elevated levels of pleiotropy associated with transcription regulation (Gibson, 1996; Wray *et al.*, 2003). For example, a single transcription factor can target response elements (RE) in many genes, such as has been described in yeast (Brem *et al.*, 2002) or as in the case of human transcription factor p53, known to control the regulation of expression of over 50 genes with different biological functions through direct binding with several REs (Levine, 1997). Moreover, eQTL mapping analyses have revealed the existence of single transcriptional ‘hotspots’ associated with the regulation of tens of genes across the genome (reviewed in Gibson & Weir, 2005). This has also recently been shown to occur in the whitefish *Coregonus clupeaformis*, a salmonid fish that is a close relative of *S. fontinalis* (Derome *et al.*, 2008; Whiteley *et al.*, 2008).

Stabilizing selection in pure forms

In contrast to comparisons with hybrids, our results indicate that patterns of gene expression in pure anadromous and resident *S. fontinalis* are very similar. Only 18 transcripts were expressed differentially in this comparison, which represents about only 6% of the number of genes with significant expression differences in hybrids.

The fact that there may be hybrid incompatibilities at genes that are not different between parental forms poses an apparent paradox. However, regulatory pathways can diverge through time between different species without any obvious change in the final phenotype (True & Haag, 2001; Lynch, 2007). Indeed, there is ample theoretical (Stone & Wray, 2001; Segal *et al.*, 2008) and empirical evidence (Ludwig *et al.*, 2000, 2005; True & Haag, 2001; Shaw *et al.*, 2002; Romano & Wray, 2003; Wray *et al.*, 2003) suggesting that within-lineage co-evolution of transcription factors and their binding sites can provide a powerful mechanism explaining functional conservation between pure species but disruptions in hybrids. For example, Ludwig *et al.* (2000, 2005) have shown that in spite of a >80% divergence in their binding sites, orthologous *eve* stripe 2 enhancers from several *Drosophila* species have identical functional activity. However, the introduction in *D. melanogaster* of chimerical enhancers combining the 5' and 3' ends from this species and *D. pseudoobscura* no longer drives proper expression of the gene, which indicates that compensatory changes in the *cis*- and *trans*-regulation have occurred since the divergence of the two species. Similarly, Shaw *et al.* (2002) have found a high degree of conservation in the functional development of the Hunchback protein across several dipteran taxa, yet the *cis*-regulatory elements in their *hunchback* (*hb*) enhancers have undergone major changes affecting particularly the binding sites of the transcription factor Bicoid. Several experiments have shown that the reliability of the Bicoid–*hb* interaction has been achieved through coevolution of lineage specific Bicoid-binding domains and their respective *hb cis*-regulatory sites. These studies, and several others involving more distantly related species (e.g. Romano & Wray, 2003), suggest that some form of stabilizing selection operates to preserve phenotype (i.e. gene expression) fidelity in diverging lineages. Moreover, genome-wide analyses of gene expression indicate that stabilizing selection is an important force behind the evolution of gene expression (Denver *et al.*, 2005; Lemos *et al.*, 2005; Gilad *et al.*, 2006).

Functional analyses of differences in gene expression

Functional analyses of the transcripts studied in this work must be taken with care given that only a fraction of the genes considered here have reliable Gene Ontology entries and in most cases the biological process is inferred by comparison with the human, mouse and even *Drosophila* species, which reflects the lack of fish information in the bioinformatics databases. However when available, functional analyses provide useful insights into some of the ecological and genetic processes associated with divergence in anadromous and resident brook charrs and with the performance of their F₁ hybrids.

Pure forms

We observed differences in the expression level of genes whose biological functions are mainly associated with electron transport and energy metabolism such as aldehyde dehydrogenase 9 family, NADH-ubiquinone oxidoreductase chain 1, NADH-ubiquinone dehydrogenase flavoprotein 3 and pyruvate kinase. It is noteworthy that anadromous and resident brook charr fish exhibit significant bioenergetic differences not only at the adult stage, *i.e.* during smoltification and migration in anadromous fish (Boula *et al.*, 2002), but also during early life history stages. For instance, Morinville & Rasmussen (2003) have demonstrated that juvenile future-migrant anadromous charr have consumption rates 1.4 higher than same age resident charr but a smaller ratio of growth to consumption, which indicates that they have higher metabolic costs. Moreover, anadromous forms, both juvenile and post-migratory, use freshwater habitats with faster currents inducing higher bioenergetic costs associated with swimming (Morinville & Rasmussen, 2006). Interestingly, similar differences in the expression of metabolism-related genes have been previously reported in studies of the transcriptome of recently (12 000 years bp) diverged salmonid taxa, such young sympatric species of *Coregonus* (Derome & Bernatchez, 2006; Derome *et al.*, 2006; Nolte *et al.*, 2009), and anadromous/resident forms in *Salmo* (Giger *et al.*, 2008). Moreover, in *Fundulus heteroclitus*, about 80% of the individual variation in cardiac metabolism is explained by variation in the expression of metabolism-associated genes, implying a functional link between both (Oleksiak *et al.*, 2005). Our results suggest therefore that divergence between anadromous and resident brook charr was accompanied by limited but significant transcriptional changes at a small number of genes mainly involved in energy metabolism. Admittedly, however, rigorously testing whether such transcriptional divergence has been driven by directional selection rather than mere genetic drift will require the analysis of multiple population pairs (e.g. St-Cyr *et al.*, 2008).

Hybrids

Disrupted genes in hybrids belong to a large set (23) of different biological processes, each representing a small fraction of the ensemble. For example, the two most abundant categories, protein metabolism/modification and lipid/steroid metabolism represent only 17.2% and 9.0% of all disrupted genes, respectively. This suggests that genomic incompatibilities can target a diverse array of biological functions even at an early stage of the divergence process. Both observations provide therefore valuable insights into a literature on gene expression and speciation in animals entirely dominated by studies on species pairs with divergences dated to >0.5 Ma and in

which reproductive genes appear almost as the only relevant biological function (Michalak & Noor, 2003; Ranz *et al.*, 2004; Landry *et al.*, 2005; Haerty & Singh, 2006; Moehring *et al.*, 2007).

Secondly, the most severely under-expressed genes in hybrids were related to electron transport, immune defence, lipid transport/metabolism and protection against oxidative stress. The last category is interesting from an ecological standpoint because brook charr tend to live in oxygen-hypersaturated waters in which the risk of oxidative stress can be particularly high (Ritola *et al.*, 2002; Martinez-Alvarez *et al.*, 2005a). Oxidative stress occurs when oxygen itself and reactive oxygen species (ROS, i.e. highly reactive molecular byproducts of metabolism), attack (oxidize) biomolecules such as proteins, DNA, steroids and unsaturated lipids in cell membranes. The consequences of oxidative stress reach a high variety of dysfunctions, diseases and syndromes (see review in Davies, 2000). Consequently, all aerobic organisms have developed powerful antioxidant defences, including glutathione peroxidase (GPX) activity – the principal enzyme that converts the highly reactive ROS H_2O_2 into H_2O . Interestingly, compared with both pure forms, hybrid brook charr show an important reduction in the expression levels of the gene encoding for this enzyme, which could have a negative impact in their antioxidant performance. For example, rainbow trout (*Oncorhynchus mykiss*) exposed to oxygen-hypersaturated water respond with a rapid increase in GPX levels, but depletion of the enzyme entails oxidative damage in gills and probably other tissues (Ritola *et al.*, 2002). Additional support for a reduced defence against oxidative stress in brook charr hybrids comes from the observation of low levels of expression of the iron-sequestering protein ferritin. Although iron is essential to vertebrate life via its role in oxygen binding/release and in numerous critical cell processes, its excess is potentially harmful as it catalyses the formation of highly reactive ROS molecules (Torti & Torti, 2002). There is ample experimental evidence confirming that ferritin plays a central role in maintaining a fine intracellular iron balance and that levels of this enzyme are positively correlated with a reduction in the accumulation of ROS and oxidative stress (Orino *et al.*, 2001; Torti & Torti, 2002). Finally, we also observed significant low levels of expression for a high number of lipid/fatty acid transporters in hybrids. Fatty acids are essential structural elements of the cell membrane. Modifications in the fatty acid composition will bring changes in membrane rigidity and on its transport function, which could then affect many physiological processes associated to trans-membrane exchanges such as thermal tolerance and osmo-iono-regulation processes. Indeed, fatty acids have been shown to play an essential role in osmoregulation in salmon (Sheridan *et al.*, 1985), trout (Li & Yamada, 2002) and sturgeons (Martinez-Alvarez *et al.*, 2005b). Here, low expression level of these genes

suggests a possible reduction of osmoregulatory performance of hybrids, especially if they develop an anadromous strategy later. In summary, functional analyses of gene expression raise the hypothesis that hybrids between anadromous and resident forms of *S. fontinalis* might suffer from reduced performance at some key physiological processes which could cause outbreeding depression, as reported in other salmonids (Lu & Bernatchez, 1998; Rogers & Bernatchez, 2006; Renaut *et al.*, 2009).

To conclude, in this study we documented hybrid dysfunctions by identifying gene transcripts that appear deregulated in F_1 hybrids between two closely related but ecologically divergent forms of a freshwater fish. We propose that our results are concordant with the Dobzhansky–Muller model in that incompatibilities in hybrids seem caused by the highly modified expression of genes that otherwise show no incompatibilities in pure forms (Dobzhansky, 1937; Muller, 1942; Orr & Presgraves, 2000; Porter & Johnson, 2002). This in turn, indicates that some form of stabilizing selection for gene expression is at action in the anadromous and resident fish. Our results thus provide further evidence that ecological divergence in fish is associated with evolutionary change of transcription levels, likely driven by natural selection (Oleksiak *et al.*, 2002; Derome & Bernatchez, 2006; Derome *et al.*, 2006; Giger *et al.*, 2006, 2008; St-Cyr, Derome & Bernatchez, 2006).

Acknowledgments

We are grateful to Simon C. Tétreault, Lise Michaud, Dominique Lavallée, and Nathalie Morin for their invaluable help in fish sampling and fish rearing. We thank Bérénice Bougas, Fabien Lamaze, Arne Nolte, Eric Normandeau, Sébastien Renaut and Christian Roberge for their help with laboratory techniques and data analysis. We are grateful to Nicolas Derome, Arne Nolte, Eric Normandeau, Sébastien Renaut, and two anonymous referees for their constructive comments on an earlier version of the manuscript. We also acknowledge the kind collaboration of B.F. Koop and W.S. Davidson, both leaders of the cGRASP. This research was financially supported by the Canadian Research Chair in Genomics and Conservation of Aquatic Resources to L.B. (Discovery Grant Program). J.M. was financially supported by a postdoctorate fellowship from Réseau Aquaculture Québec.

References

- Arnosti, D.N. 2003. Analysis and function of transcriptional regulatory elements: insights from *Drosophila*. *Annu. Rev. Entomol.* **48**: 579–602.
- Barbash, D.A., Siino, D.F., Tarone, A.M. & Roote, J. 2003. A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**: 5302–5307.

- Benjamini, Y. & Hochberg, Y. 1995. Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B – (Methodological)* **57**: 289–300.
- Bertrand, M., Marcogliese, D.J. & Magnan, P. 2008. Trophic polymorphism in brook charr revealed by diet, parasites and morphometrics. *J. Fish Biol.* **72**: 555–572.
- Boula, D., Castric, V., Bernatchez, L. & Audet, C. 2002. Physiological, endocrine, and genetic bases of anadromy in the brook charr, *Salvelinus fontinalis*, of the Laval River (Quebec, Canada). *Environ. Biol. Fish.* **64**: 229–242.
- Boursot, P., Auffray, J.-C., Britton-Davidian, J. & Bonhomme, F. 1993. The evolution of house mice. *Annu. Rev. Ecol. Syst.* **24**: 119–152.
- Boursot, P., Din, W., Anand, R., Darviche, D., Dod, B., VonDeimling, F., Talwar, G.P. & Bonhomme, F. 1996. Origin and radiation of the house mouse: Mitochondrial DNA phylogeny. *J. Evol. Biol.* **9**: 391–415.
- Brem, R.B., Yvert, G., Clinton, R. & Kruglyak, L. 2002. Genetic dissection of transcriptional regulation in budding yeast. *Science* **296**: 752–755.
- Brideau, N.J., Flores, H.A., Wang, J., Maheshwari, S., Wang, X. & Barbash, D.A. 2006. Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science* **314**: 1292–1295.
- Castric, V. & Bernatchez, L. 2003. The rise and fall of isolation by distance in the anadromous brook charr (*Salvelinus fontinalis* Mitchell). *Genetics* **163**: 983–996.
- Comai, L., Madlung, A., Josefsson, C. & Tyagi, A. 2003. Do the different parental ‘heteromes’ cause genomic shock in newly formed allopolyploids? *Phil. Trans. R. Soc. Lond. B* **358**: 1149–1155.
- Coyne, J.A. & Orr, H.A. 2004. *Speciation*. Sinauer Associates, Sunderland, MA, USA.
- Cui, X.G., Hwang, J.T.G., Qiu, J., Blades, N.J. & Churchill, G.A. 2005. Improved statistical tests for differential gene expression by shrinking variance components estimates. *Biostatistics* **6**: 59–75.
- Davies, K.J.A. 2000. Oxidative stress, antioxidant defenses, and damage removal, repair, and replacement systems. *IUBMB Life* **50**: 279–289.
- Denver, D.R., Morris, K., Streebman, J.T., Kim, S.K., Lynch, M. & Thomas, W.K. 2005. The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nat. Genet.* **37**: 544–548.
- Derome, N. & Bernatchez, L. 2006. The transcriptomics of ecological convergence between 2 limnetic coregonine fishes (Salmonidae). *Mol. Biol. Evol.* **23**: 2370–2378.
- Derome, N., Duchesne, P. & Bernatchez, L. 2006. Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Mol. Ecol.* **15**: 1239–1249.
- Derome, N., Bougas, B., Rogers, S.M., Whiteley, A., Labbe, A., Laroche, J. & Bernatchez, L. 2008. Pervasive sex-linked effects on transcription regulation as revealed by eQTL mapping in lake whitefish species pairs (*Coregonus* sp, Salmonidae). *Genetics* **179**: 1903–1917.
- Dobzhansky, T. 1937. *Genetics and the Origin of Species*. Columbia University Press, New York.
- Draghici, S. 2003. *Data Analysis Tools for Microarrays*. Chapman and Hall/CRC Press, Boca Raton, FL, USA.
- Evans, B.J., Kelley, D.B., Tinsley, R.C., Melnick, D.J. & Cannatella, D.C. 2004. A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. *Mol. Phylogenet. Evol.* **33**: 197–213.
- Evans, B.J., Kelley, D.B., Melnick, D.J. & Cannatella, D.C. 2005. Evolution of RAG-1 in polyploid clawed frogs. *Mol. Biol. Evol.* **22**: 1193–1207.
- Gibson, G. 1996. Epistasis and pleiotropy as natural properties of transcriptional regulation. *Theor. Popul. Biol.* **49**: 58–89.
- Gibson, G. & Weir, B. 2005. The quantitative genetics of transcription. *Trends Genet.* **21**: 616–623.
- Giger, T., Excoffier, L., Day, P.J.R., Champigneulle, A., Hansen, M.M., Powell, R. & Largiader, C.R. 2006. Life history shapes gene expression in salmonids. *Curr. Biol.* **16**: R281–R282.
- Giger, T., Excoffier, L., Amstutz, U., Day, P.J.R., Champigneulle, A., Hansen, M.M., Kelso, J. & Largiader, C.R. 2008. Population transcriptomics of life-history variation in the genus *Salmo*. *Mol. Ecol.* **17**: 3095–3108.
- Gilad, Y., Oshlack, A. & Rifkin, S.A. 2006. Natural selection on gene expression. *Trends Genet.* **22**: 456–461.
- Haerty, W. & Singh, R.S. 2006. Gene regulation divergence is a major contributor to the evolution of Dobzhansky-Muller incompatibilities between species of *Drosophila*. *Mol. Biol. Evol.* **23**: 1707–1714.
- Hedgecock, D., Lin, J., DeCola, S., Haudenschild, C.D., Meyer, E., Manahan, D.T. & Bowen, B. 2007. Transcriptomic analysis of growth heterosis in larval Pacific oysters (*Crassostrea gigas*). *Proc. Natl Acad. Sci. USA* **104**: 2313–2318.
- Inga, A., Reamon-Buettner, S.M., Borlak, J. & Resnick, M.A. 2005. Functional dissection of sequence-specific NKX2-5 DNA binding domain mutations associated with human heart septation defects using a yeast-based system. *Hum. Mol. Genet.* **14**: 1965–1975.
- Johnson, N.A. & Porter, A.H. 2000. Rapid speciation via parallel, directional selection on regulatory genetic pathways. *J. Theor. Biol.* **205**: 527–542.
- Kerr, M.K., Martin, M. & Churchill, G.A. 2000. Analysis of variance for gene expression microarray data. *J. Comput. Biol.* **7**: 819–837.
- Kerr, M.K., Afshari, C.A., Bennett, L., Bushel, P., Martinez, J., Walker, N.J. & Churchill, G.A. 2002. Statistical analysis of a gene expression microarray experiment with replication. *Stat. Sin.* **12**: 203–217.
- Landry, C.R., Wittkopp, P.J., Taubes, C.H., Ranz, J.M., Clark, A.G. & Hartl, D.L. 2005. Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. *Genetics* **171**: 1813–1822.
- Landry, C.R., Hartl, D.L. & Ranz, J.M. 2007. Genome clashes in hybrids: insights from gene expression. *Heredity* **99**: 483–493.
- Lefrant, A. 2006. *Resistance des ombles (Salvelinus fontinalis et Salvelinus alpinus) aux infections bactériennes opportunistes en milieu marin*, M.sc. Thesis. Université du Québec à Rimouski, Rimouski.
- Lemos, B., Meiklejohn, C.D., Caceres, M. & Hartl, D.L. 2005. Rates of divergence in gene expression profiles of primates, mice, and flies: Stabilizing selection and variability among functional categories. *Evolution* **59**: 126–137.
- Lenormand, S., Dodson, J.J. & Menard, A. 2004. Seasonal and ontogenetic patterns in the migration of anadromous brook charr (*Salvelinus fontinalis*). *Can. J. Fish. Aquat. Sci.* **61**: 54–67.
- Levine, A.J. 1997. p53, the cellular gatekeeper for growth and division. *Cell* **88**: 323–331.
- Levine, M. & Tjian, R. 2003. Transcription regulation and animal diversity. *Nature* **424**: 147–151.

- Li, H. & Yamada, H. 2002. Changes of the fatty acid composition in smolts of masu salmon (*Oncorhynchus masou*), associated with desmoltification and sea-water transfer. *Comp. Biochem. Physiol. A* **103**: 221–226.
- Lu, G. & Bernatchez, L. 1998. Experimental evidence for reduced hybrid viability between dwarf and normal ecotypes of lake whitefish (*Coregonus clupeaformis* Mitchill). *Proc. R. Soc. Lond. Ser. B* **265**: 1025–1030.
- Ludwig, M.Z., Bergman, C., Patel, N.H. & Kreitman, M. 2000. Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* **403**: 564–567.
- Ludwig, M.Z., Palsson, A., Alekseeva, E., Bergman, C.M., Nathan, J. & Kreitman, M. 2005. Functional evolution of a cis-regulatory module. *PLoS Biol.* **3**: 588–598.
- Lynch, M. 2007. *The Origins of Genome Architecture*. Sinauer Associates, Sunderland, MA, USA.
- Magnan, P., Audet, C., Glemet, H., Legault, M., Rodriguez, M.A. & Taylor, E.B. 2002. Developments in the ecology, evolution, and behaviour of the charrs, genus *Salvelinus*: relevance for their management and conservation. *Environ. Biol. Fish* **64**: 9–14.
- Malone, J.H., Chrzanowski, T.H. & Michalak, P. 2007. Sterility and gene expression in hybrid males of *Xenopus laevis* and *X. muelleri*. *PLoS ONE* **2**: e781.
- Martinez-Alvarez, R.M., Morales, A.E. & Sanz, A. 2005a. Antioxidant defenses in fish: biotic and abiotic factors. *Rev. Fish Biol. Fisheries* **15**: 75–88.
- Martinez-Alvarez, R.M., Sanz, A., Garcia-Gallego, M., Domezain, A., Domezain, J., Carmona, R., Ostos-Garrido, M.D. & Morales, A.E. 2005b. Adaptive branchial mechanisms in the sturgeon *Acipenser naccarii* during acclimation to saltwater. *Comp. Biochem. Physiol. A* **141**: 183–190.
- McClintock, B. 1984. The significance of responses of the genome to challenge. *Science* **226**: 792–801.
- Meiklejohn, C.D., Parsch, J., Ranz, J.M. & Hartl, D.L. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc. Natl Acad. Sci. USA* **100**: 9894–9899.
- Michalak, P. & Noor, M.A. 2003. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. *Mol. Biol. Evol.* **20**: 1070–1076.
- Moehring, A.J., Teeter, K.C. & Noor, M.A. 2007. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. *Mol. Biol. Evol.* **24**: 137–145.
- Morinville, G.R. & Rasmussen, J.B. 2003. Early juvenile bioenergetic differences between anadromous and resident brook charr (*Salvelinus fontinalis*). *Can. J. Fish. Aquat. Sci.* **60**: 401–410.
- Morinville, G.R. & Rasmussen, J.B. 2006. Does life-history variability in salmonids affect habitat use by juveniles? A comparison among streams open and closed to anadromy. *J. Anim. Ecol.* **75**: 693–704.
- Morinville, G.R. & Rasmussen, J.B. 2008. Distinguishing between juvenile anadromous and resident brook charr (*Salvelinus fontinalis*) using morphology. *Environ. Biol. Fish* **81**: 171–184.
- Muller, H.J. 1942. Isolating mechanisms, evolution, and temperature. *Biol. Symp.* **6**: 71–125.
- Nolte, A.W., Renaut, S. & Bernatchez, L. 2009. Divergence in gene regulation at young life history stages of whitefish (*Coregonus* sp.) and the emergence of genomic isolation. *BMC Evol. Biol.* **9**: 59.
- Oleksiak, M.F., Churchill, G.A. & Crawford, D.L. 2002. Variation in gene expression within and among natural populations. *Nat. Genet.* **32**: 261–266.
- Oleksiak, M.F., Roach, J.L. & Crawford, D.L. 2005. Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nat. Genet.* **37**: 67–72.
- Orino, K., Lehman, L., Tsuji, Y., Ayaki, H., Torti, S.V. & Torti, F.M. 2001. Ferritin and the response to oxidative stress. *Biochem. J.* **357**: 241–247.
- Orr, H.A. 1997. Haldane's rule. *Annu. Rev. Ecol. Syst.* **28**: 195–218.
- Orr, H.A. & Presgraves, D.C. 2000. Speciation by postzygotic isolation: forces, genes and molecules. *BioEssays* **22**: 1085–1094.
- Ortiz-Barrientos, D., Counterman, B.A. & Noor, M.A.F. 2007. Gene expression divergence and the origin of hybrid dysfunctions. *Genetica* **129**: 71–81.
- Pattanaik, S., Xie, C.H. & Yuan, L. 2008. The interaction domains of the plant Myc-like bHLH transcription factors can regulate the transactivation strength. *Planta* **227**: 707–715.
- Perry, G.M., Audet, C., Laplatte, B. & Bernatchez, L. 2004. Shifting patterns in genetic control at the embryo-alevin boundary in brook charr. *Evolution* **58**: 2002–2012.
- Perry, G.M., Audet, C. & Bernatchez, L. 2005. Maternal genetic effects on adaptive divergence between anadromous and resident brook charr during early life history. *J. Evol. Biol.* **18**: 1348–1361.
- Porter, A.H. & Johnson, N.A. 2002. Speciation despite gene flow when developmental pathways evolve. *Evolution* **56**: 2103–2111.
- Proulx, R. & Magnan, P. 2004. Contribution of phenotypic plasticity and heredity to the trophic polymorphism of lacustrine brook charr (*Salvelinus fontinalis* M.). *Evol. Ecol. Res.* **6**: 503–522.
- Ranz, J.M., Namgyal, K., Gibson, G. & Hartl, D.L. 2004. Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. *Genome Res.* **14**: 373–379.
- Rawson, P.D. & Burton, R.S. 2002. Functional coadaptation between cytochrome c and cytochrome c oxidase within allopatric populations of a marine copepod. *Proc. Natl Acad. Sci. USA* **99**: 12955–12958.
- Reiner, A., Yekutieli, D. & Benjamini, Y. 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* **19**: 368–375.
- Renaut, S., Nolte, A.W. & Bernatchez, L. 2009. Transcriptomic investigation of post-zygotic isolation in lake whitefish (*Coregonus clupeaformis*). *Mol. Biol. Evol.* **26**: 925–936.
- Ritola, O., Livingstone, D.R., Peters, L.D. & Lindstrom-Seppa, P. 2002. Antioxidant processes are affected in juvenile rainbow trout (*Oncorhynchus mykiss*) exposed to ozone and oxygen-supersaturated water. *Aquaculture* **210**: 1–19.
- Rogers, S.M.R. & Bernatchez, L. 2006. The genetic basis of intrinsic and extrinsic postzygotic reproductive isolation jointly promoting speciation in the lake whitefish species complex (*Coregonus clupeaformis*). *J. Evol. Biol.* **19**: 1979–1994.
- Romano, L.A. & Wray, G.A. 2003. Conservation of Endo16 expression in sea urchins despite evolutionary divergence in both cis and trans-acting components of transcriptional regulation. *Development* **130**: 4187–4199.
- Rottscheldt, R. & Harr, B. 2007. Extensive additivity of gene expression differentiates subspecies of the house mouse. *Genetics* **177**: 1553–1567.

- von Schalburg, K.R., Rise, M.L., Cooper, G.A., Brown, G.D., Gibbs, A.R., Nelson, C.C., Davidson, W.S. & Koop, B.F. 2005. Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray. *BMC Genomics* **6**: 126.
- Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U. & Gaul, U. 2008. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature* **451**: 535–U531.
- Shaw, P.J., Wratten, N.S., McGregor, A.P. & Dover, G.A. 2002. Coevolution in bicoid-dependent promoters and the inception of regulatory incompatibilities among species of higher Diptera. *Evol. Dev.* **4**: 265–277.
- Sheridan, M.A., Woo, N.Y.S. & Bern, H.A. 1985. Changes in the rates of glycogenesis, glycogenolysis, lipogenesis, and lipolysis in selected tissues of the coho salmon (*Oncorhynchus kisutch*) associated with parr-smolt transformation. *J. Exp. Zool.* **236**: 35–44.
- Simon, D.B., Nelson-Williams, C., Bia, M.J., Ellison, D., Karet, F.E., Molina, A.M., Vaara, I., Iwata, F., Cushner, H.M., Koolen, M., Gainza, F.J., Gitelman, H.J. & Lifton, R.P. 1996. Gitelman's variant of Bartter's syndrome, inherited hypokalaemic alkalosis, is caused by mutations in the thiazide-sensitive Na-Cl cotransporter. *Nat. Genet.* **12**: 24–30.
- Singh, R.S. & Kulathinal, R.J. 2000. Sex gene pool evolution and speciation: A new paradigm. *Genes Genet. Syst.* **75**: 119–130.
- St-Cyr, J., Derome, N. & Bernatchez, L. 2008. The transcriptomics of life-history trade-offs in whitefish species pairs (*Coregonus* sp.). *Mol. Ecol.* **17**: 1850–1870.
- Stone, J.R. & Wray, G.A. 2001. Rapid evolution of cis-regulatory sequences via local point mutations. *Mol. Biol. Evol.* **18**: 1764–1770.
- Tamura, K., Subramanian, S. & Kumar, S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol. Biol. Evol.* **21**: 36–44.
- Theriault, V. & Dodson, J.J. 2003. Body size and the adoption of a migratory tactic in brook charr. *J. Fish Biol.* **63**: 1144–1159.
- Theriault, V., Bernatchez, L. & Dodson, J.J. 2007a. Mating system and individual reproductive success of sympatric anadromous and resident brook charr, *Salvelinus fontinalis*, under natural conditions. *Behav. Ecol. Sociobiol.* **62**: 51–65.
- Theriault, V., Garant, D., Bernatchez, L. & Dodson, J.J. 2007b. Heritability of life-history tactics and genetic correlation with body size in a natural population of brook charr (*Salvelinus fontinalis*). *J. Evol. Biol.* **20**: 2266–2277.
- Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A. & Narechania, A. 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* **13**: 2199–2141.
- Ting, C.T., Tsaur, S.C., Wu, M.L. & Wu, C.I. 1998. A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**: 1501–1504.
- Torti, F.M. & Torti, S.V. 2002. Regulation of ferritin genes and protein. *Blood* **99**: 3505–3516.
- True, J.R. & Haag, E.S. 2001. Developmental system drift and flexibility in evolutionary trajectories. *Evol. Dev.* **3**: 109–119.
- Wang, J.L., Tian, L., Lee, H.S., Wei, N.E., Jiang, H.M., Watson, B., Madlung, A., Osborn, T.C., Doerge, R.W., Comai, L. & Chen, Z.J. 2006. Genomewide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics* **172**: 507–517.
- Whiteley, A.R., Derome, N., Rogers, S.M., St-Cyr, J., Laroche, J., Labbe, A., Nolte, A., Renaut, S., Jeukens, J. & Bernatchez, L. 2008. The Phenomics and Expression Quantitative Trait Locus Mapping of Brain Transcriptomes Regulating Adaptive Divergence in Lake Whitefish Species Pairs (*Coregonus* sp.). *Genetics* **180**: 147–164.
- Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V. & Romano, L.A. 2003. The evolution of transcriptional regulation in eukaryotes. *Mol. Biol. Evol.* **20**: 1377–1419.
- Wu, C.I., Johnson, N.A. & Palopoli, M.F. 1996. Haldane's rule and its legacy: Why are there so many sterile males? *Trends Ecol. Evol.* **11**: 281–284.

Received 27 March 2009; revised 7 May 2009; accepted 8 May 2009