Gene Expression Divergence and Hybrid Misexpression between Lake Whitefish Species Pairs (*Coregonus* spp. Salmonidae)

S. Renaut, A.W. Nolte,† and L. Bernatchez**

*Département de Biologie, IBIS (Institut de Biologie Intégrative et des Systémes), Université Laval, Québec, Canada; and †Max-Planck-Institute for Evolutionary Biology, Plön, Germany

Genomewide analyses of the transcriptome have confirmed that gene misexpression may underlie reproductive isolation mechanisms in interspecific hybrids. Here, using a 16,006 features cDNA microarray, we compared and contrasted gene expression divergence at two ontogenetic stages in incipient species of normal and dwarf whitefish (*Coregonus clupeaformis*) with that of first generation (normal × dwarf) and second-generation hybrid crosses (backcross: [normal × dwarf] × normal]. Our goal was to identify the main mode of action responsible for gene transcription and to discover key genes misexpressed in hybrids. Very few transcripts (five of 4,950 expressed) differed in mean expression level between parentals and hybrids at the embryonic stage, in contrast to 16-week-old juvenile fish for which 617 out of 5,359 transcripts differed significantly. We also found evidence for more misexpression in backcross hybrids whereby nonadditivity explained a larger fraction of hybrid inheritance patterns in backcross (54%) compared with F1-hybrids (9%). Gene expression in hybrids was more variable than in pure crosses and transgressive patterns of expression were ubiquitous in hybrids. In backcross embryos in particular, the expression of three key developmental genes involved in protein folding and mRNA translation was severely disrupted. Accordingly, gene misexpression in hybrids adds to other factors previously identified as contributing to the reproductive isolation of incipient species of lake whitefish.

Introduction

Under the biological species concept, reproductive isolation arises as a consequence of population divergence, itself driven by natural selection, sexual selection or genetic drift (Coyne and Orr 2004; Bell 2008). The evolution of reproductive isolation is often viewed as a gradual process whereas, over time, more and more barriers will tend to separate lineages and reinforce their divergence (De Queiroz 1998). Frequent gene flow between parental lineages is a characteristic of many early divergence events (Bernatchez 2004; Wu and Ting 2004; Nosil 2008). In these situations, postzygotic isolation may result from the interaction of genetic factors in the parental lineages that, although functional in their normal genetic backgrounds, reduce fitness when recombined in hybrids (Rundle et al. 2000; Burton et al. 2006; Rogers and Bernatchez 2006; Gow et al. 2007). These hybrids can be unfit due to intrinsic factors resulting in increased embryonic mortality or external—environmentally driven—factors, for instance the lack of finding a suitable ecological niche (Schluter 2000).

In nature, rare F1-hybrids, encountering few mates of their kind, may backcross to parental species. This progeny often suffers more problems than first generation hybrids (Barton 2001; Coyne and Orr 2004). Namely, recombination is known to release cryptic genetic variation, resulting in phenotypes that are extreme relative to those of either parental line (Endler 1977; Rieseberg et al. 1999, 2003; Mallet 2007). Ultimately, extreme phenotypes can be lethal or sterile (hybrid breakdown), whereas transgressive segregation refers to phenotypic values in hybrids that extend significantly outside the range defined by the parents (DeVicente and Tanksley 1993; Rockman and Kruglyak 2006). As such, a particular trait in hybrids may, on average, be similar to the parents, yet be transgressive and thus maladapted due to an increased phenotypic variability. Both hybrid breakdown and transgressive segregation may explain the underlying basis of postzygotic isolation in early divergent lineages (Burke and Arnold 2001; Rogers and Bernatchez 2006). The manifestation of such extreme traits supposes nonadditive, epistatic, gene interactions (Dobzhansky 1937; Muller 1942), although transgressive segregation can also be caused solely by the complementary action of genes with additive effects (Rieseberg et al. 1999).

Based on the premise that transcriptional regulation constitutes a major component of the genetic basis for phenotypic evolution (Wray et al. 2003; Wray 2007; but see Hoekstra and Coyne 2007; Carroll 2008), the analysis of gene expression has allowed the identification of many candidate genes underlying phenotypic divergence (Derome et al. 2006; Ranz and Machado 2006; Landry et al. 2007; St-Cyr et al. 2008; Jeukens et al. 2009). Moreover, hybrids have proven extremely valuable to identify cryptically differentiated genetic factors, whereby the combination of divergent regulatory elements into a common genetic background resulted in gene misexpression (reviewed by Landry et al. 2007; Ortiz-Barrientos et al. 2007). For example, Ranz et al. (2004) have shown that the global expression profile of *Drosophila melanogaster* and *Drosophila simulans* are more closely related to each other than to their hybrid progeny. Undeniably, there is a large body of evidence pointing toward largely nonadditive inheritance of gene expression in hybrids (Rockman and Kruglyak 2006), including in fishes (Roberge et al. 2008), which may consequently explain their selective disadvantage. Other studies, however, have reported the predominance of additive patterns of gene expression in F1-hybrids (Hughes et al. 2006; Rottscheidt and Harr 2007), such that the type of inheritance responsible for gene transcription levels in hybrids remains a contentious issue (Rockman and Kruglyak 2006; Moehring et al. 2007). Finally, patterns of gene expression in young species pairs and post-F1 hybrid generations have been little explored, and the underlying transcriptomic basis

Key words: *Coregonus*, hybridization, microarray, reproductive isolation, speciation, transgressive segregation.

E-mail: sebastien.renaut.1@ulaval.ca.

Mol. Biol. Evol. 26(4):925–936. 2009
doi:10.1093/molbev/msp017
Advance Access publication January 27, 2009

Table 1
Origins of Strains and Crosses Used for Gene Expression Analysis

Experimental Group	Lineage	Crosses
D (1)	Témiscouata Lake dwarf	Laboratory strain: single female crossed with five different males
D (2)	Témiscouata Lake dwarf	Wild parental fish, several females crossed with several males
N (1)	Aylmer Lake normal	Laboratory strain: two females crossed with three males
N (2)	Aylmer Lake normal	Laboratory strain: single female crossed with five different males
F1 (1)	Témiscouata Lake dwarf—Aylmer Lake normal	Aylmer Lake female (same as in N1) crossed with five Témiscouata Lake dwarf males (same as D1)
F1 (2)	Témiscouata Lake dwarf—Aylmer Lake normal	Wild-caught Dwarf females (multiple) with three males from Aylmer Lake (same as N2)
BC	F1—-Aylmer Lake normal	F1-hybrid (derived from Aylmer, Témiscouata laboratory strains) female crossed with five males (Aylmer Laboratory strain normal)

NOTE.—Two experimental groups were created for dwarf, normal, and F1, and individuals used for gene expression experiments were composed, in equal part, from the (1) and (2) duplicate families. One experimental group was created for BC. Dwarf and normal families were used in a previous study by Nolte et al. (2009).

of reproductive isolation mechanisms remains largely unknown.

Recent postglacial ecological divergence (12,000–15,000 years ago) of the lake whitefish *Coregonus clupeaformis* (salmonidae) has repeatedly led to the formation of two, benthic and limnetic, whitefish species occurring in sympatry and hereafter referred to as "normal" and "dwarf," respectively (Bernatchez 2004). Extensive experimental work has pointed to differentiation at morphological (Lu and Bernatchez 1999; Rogers et al. 2002; Bernatchez 2004), life history (Bernatchez et al. 1999), physiological (Trudel et al. 2001; Rogers and Bernatchez 2005, 2007) and genetic levels (Bernatchez 2004; Rogers et al. 2007). The recent advent of microarray technology developed for salmonids (von Schalburg et al. 2005) has permitted to identify consistent gene expression divergence between normal and dwarf whitefish in both natural and laboratory settings (Derome et al. 2006, 2008; St-Cyr et al. 2008; Whiteley et al. 2008; Nolte et al. 2009). In particular, these studies, combined with physiological data (Trudel et al. 2001) have shown that, at the juvenile and adult life stages, energy metabolism plays a fundamental role in driving whitefish adaptive divergence. In particular, Nolte et al. (2009) identified a 14-fold increase of differentially expressed genes later (juvenile stage) rather than sooner (embryonic stage) in ontogeny, which may potentially explain the emergence of reproductive isolation as a by-product of adaptive divergence on adult characters. At the same time, the relative lack of divergence in embryos of pure crosses may imply that individual genes are still evolving, while the net gene expression outcome is not altered. Thus Nolte et al. (2009) hypothesized that gene misexpression could manifest as genetic factors segregate in hybrid crosses and particularly so in backcrosses (Nolte et al. 2009). In line with this, both intrinsic (increased hybrid mortality) and extrinsic (transgressive segregation in hatching time) postzygotic isolation mechanisms were shown to be more prevalent and severe in backcross individuals than F1-hybrids (Lu and Bernatchez 1998; Rogers and Bernatchez 2006).

The present study anchors itself on a previous analysis of gene expression data of pure normal and dwarf whitefishes at two developing stages (embryonic and juvenile, Nolte et al. 2009). Here, the goal was to document the main mode of action responsible for gene transcription in hybrids and to identify genes misexpressed relative to dwarf and normal whitefish. More specifically, under the assumption that similar genes involved in the adaptive divergence of these species are also responsible for driving their reproductive isolation, and based on previous comparisons between pure forms (Nolte et al. 2009), we predicted an excess of hybrid misexpression at the juvenile compared with the embryonic stage. Secondly, we also predicted more evidence of hybrid misexpression in second generation (backcross) compared with first generation hybrids.

Material and Methods

Strains, Crosses, and Fish Maintenance

Details regarding strain origin, crosses, and fish maintenance are provided in Nolte et al. (2009). Briefly, eggs were obtained from outbred laboratory strains (normal whitefish originating from Aylmer Lake (45°54′N, 71°20′W), dwarf whitefish originating from Témiscouata Lake (47°41′N, 68°47′W), as detailed in Lu and Bernatchez 1998) at the Laboratoire de Recherche en Sciences Aquatiques (LARSA, Université Laval, Quebec, Canada). We also used wild dwarf whitefish caught in Témiscouata Lake in October 2006. In order to reduce family-specific effects, we used crosses that were composed of several parents, depending on the availability of mature fish (table 1). The group D-1 was created using one laboratory strain dwarf female crossed to five different dwarf males, all originating from Témiscouata Lake. D-2 was created by crossing wild-caught dwarf whitefish from the same lake using multiple females and multiple males. Two groups of Normal whitefish (N-1 and N-2) were created from one and five as well as two and three, females and males of the laboratory strain of normal whitefishes from Aylmer Lake. F1-hybrids (F1-1) were generated between the females of group N-1 and males of Group D-1. Likewise, another group of F1-hybrids (F1-2) was created among the parents of the second pair of parental groups (D-2 and N-2). Finally, an independent group of backcross (BC) was obtained using an F1-hybrid female generated in the laboratory in a previous study (Rogers et al. 2007) crossed to five normal whitefish (laboratory strain). As such, the backcrosses are composed of a 75% normal and 25% dwarf genetic background.

Sampling

We sampled embryos during the beginning of the segmentation period. For our experiments, we chose embryos that had formed approximately 20 segments in the detached portion of their tail, which was also curved at an angle of approximately 30°. Furthermore, in this stage, the optic primordium begins to hollow, thus initiating the formation of the eye lens. Viable eggs with well-formed embryos were individually selected, preserved in RNA later (Ambion, Austin, TX) and frozen at −20 °C for storage.

All hatched larvae were transferred to tanks, and we sampled juvenile fish at an age of 16 weeks (May 2007), when these reached a weight of approximately 860 mg (500–1190 mg). Young immature fish chosen for gene expression analysis were well developed and in good general shape. Sampling was done in the morning following an 18-h fast. Fish were then killed with a blow, kept on ice (no longer than 20 min), homogenized in Trizol Reagent (Invitrogen, Carlsbad, CA) using a polytron homogenizer and stored at −80 °C prior to RNA extraction.

Choice of Samples and Analysis of Gene Expression

Total RNA was extracted using the Trizol Reagent protocol. For the embryo experiment, a pool of five whole embryos preserved in RNA later was homogenized using a bead mill (Qiagen, Germantown, MD). RNA pooling is a common practice when quantity of material is limiting, and inference for most genes is not affected by this (Kendziorski et al. 2005). For the juvenile fish experiment, a single whole juvenile fish was used. Crude total RNA was further cleaned by ultrafiltration using microcon (Millipore, Billerica, MA) spin columns (embryo experiment) or a combination of a lithium chloride precipitation (1 vol. 5 M LiCl, precipitation at −20 °C for 2 h, centrifugation at 16,000 × g at 4 °C for 30 min, 70% ethanol wash) and subsequent ultrafiltration (juvenile experiment). Total RNA was quantified and quality checked using Experion RNA StdSens Analysis Kit (Bio-Rad, Hercules, CA). Total RNA was stored in pure water supplemented with Superase-In RNase Inhibitor (Ambion) at −80 °C.

Gene expression analysis was performed using the 16K (v2.0) Salmon cDNA microarray provided by the cGRASP consortium (von Schalburg et al. 2005). Following the vendor's protocol, Genisphere (Hatfield, PA) 3DNA Array Detection Array 350 Kit (Cy3/Alex647) was used in the embryo experiment because it requires less starting material (we used 3–5 μg of RNA). Genisphere 3DNA Array Detection Array 50 Kit (Cy3/Cy5) was used in the juvenile experiment, where we used 15–20 μg of starting RNA. Reverse transcription reactions were performed using Superscript II Kit (Invitrogen). Microarrays were scanned using a ScanArray Express scanner (Packard Bioscience, Wellesley, MA).

Samples of dwarf, normal, F1-hybrid, and backcross were hybridized in a loop design, involving eight biological replicates, and dye swap performed between each replicate (fig. 1). In this way, technical replication for dwarf and normal samples was 3-fold, whereas F1-hybrid and backcross samples were each involved in two pairwise comparisons.

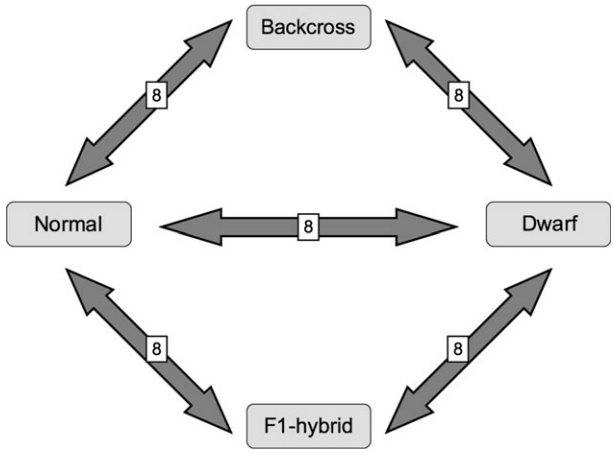

FIG. 1.—Microarray design with the four experimental groups. Each double-headed arrow represents one microarray slide such that one complete loop corresponds to five slides. Each loop was replicated eight times for a total of 40 slides for both the embryo and juvenile data sets.

As a result, we obtained a final set of 40 experiments for each, embryo and juvenile, data sets.

We used an ANOVA-based approach using the R package Rmaanova (v1.4.1) (Kerr et al. 2000) to identify transcripts differentially expressed. The mixed model used included the following terms as fixed sources of variance: Group (normal, dwarf, F1-hybrid, or backcross) and Dye (Fluorescent dye). Sample (biological sample) and Array (individual microarray) were included as random sources of variance. Statistical testing for overall divergence in gene expression is based on an F test (1,000 permutations, Fs test option). We corrected for multiple testing using a False Discovery Rate (FDR) cutoff value of 0.05, as implemented in Rmaanova. To test for specific pairwise differences between groups (N, D, F1, and BC), we used the contrast option (*t*-test) implemented in Rmaanova (1,000 permutations, Fs test option, FDR cutoff value of 0.05). In order to remain conservative in interpreting the number of significant features, transcripts with less than 2% sequence divergence were also compressed into a single transcript.

Bayesian Analysis of Gene Expression

Relative level of expression was also estimated for each transcript using a Bayesian approach (BAGEL V.2, Townsend and Hartl 2002). Normalized ratio data were implemented in the software (using the default parameters), and the most probable relative gene expression values were then calculated per transcript, per group (N, D, F1, or BC) and per replicate. The average relative gene expression and variance estimate were then calculated from the eight replicates. Previous work has shown that this Bayesian method strongly corroborates results obtained through analysis of variance-based methods, yet provided a simpler interpretation of relative gene expression and variance (Meiklejohn and Townsend 2005). In our juvenile data set, we verified that, for all comparisons, fold changes calculated from the relative gene expression values obtained through the Bayesian method and fold changes provided by Rmaanova were

highly correlated ($r^2 > 0.97$, data not shown). Similar correlation values were obtained in the embryo data set, but the low number of transcripts differentially expressed limited the validity of the approach.

Variance Estimation in Gene Expression

Test of homogeneity of variance (Bartlett 1937) between the groups (N, D, F1, and BC) was performed to identify general patterns of variability and transgressivity in hybrids compared with parental species, using the relative gene expression values calculated for each replicate. Histograms of the variance estimates were drawn for each group from all the transcripts that showed heterogeneity of variance (at $P < 0.05$, Bartlett test). Nonparametric Wilcoxon rank sum test were then used to compare the groups (Bauer 1972).

Classification of Gene Expression Inheritance Patterns

In order to categorize different types of inheritance, we analyzed the distribution of dominance effects (*d/a* ratio distribution, Gibson et al. 2004; Hughes et al. 2006, Rottscheidt and Harr 2007) to decipher between additivity, dominance, or nonadditivity for genes differentially expressed between parentals. For F1-hybrids, $d = \mu_{\text{F1-hybrid}} - ((\mu_{\text{normal}} + \mu_{\text{dwarf}})/2)$, $a = (\mu_{\text{normal}} - \mu_{\text{dwarf}})/2$, where μ are relative gene expression values. In backcross, formulas were modified to take into account the 75% normal, 25% dwarf hybrid background, $d = \mu_{\text{backcrosss}} - ([\mu_{\text{normal}} + (\mu_{\text{dwarf}} + \mu_{\text{normal}})/2]/2)$ and $a = [(\mu_{\text{normal}} - (\mu_{\text{dwarf}} + \mu_{\text{normal}})/2]/2$.

Additivity: Transcript whose hybrid gene expression value corresponds to the midvalue of the parents. A *d/a* ratio of 0 corresponds to perfect within-locus additivity (i.e., $d = 0$). We then set up an arbitrary range of -0.5 to $+0.5$ to include transcripts showing patterns of gene expression resembling additivity, rather than dominance.

Dominance: Transcript whose hybrid gene expression value resembles more closely one parent than another. A *d/a* ratio $= -1$ or $+1$ corresponds to complete dominance. In order to also include transcripts showing near complete dominance, we applied a *d/a* ratio threshold of $+0.5$ to $+1.5$ (normal dominance) or -0.5 to -1.5 (dwarf dominance).

Nonadditivity: Transcript whose hybrid gene expression is lower or higher than both parents. A *d/a* ratio greater than $+1.5$ or smaller than -1.5 corresponds to nonadditivity. These genes were further classified as under/overdominant if the hybrid has an expression lower/higher than the mean of both parents. We also included in this category transcripts not significantly differentiated between the parents, but for which hybrid expression was significantly different from the parental values, either by being under or overexpressed.

Transgressivity: As defined by Brem and Kruglyak (2005) and Rockman and Kruglyak (2006), transgressivity corresponds to a transcript whose level of expression in segregating hybrids does not necessarily differ in average from midparent values but whose variance extends outside the range of both parents values. Firstly, the maximum range of values defined for the parents was calculated as mean expression plus two standard deviation ($\mu_{\text{parent}} + 2\sigma_{\text{parent}}$) of the highest parent (either normal or dwarf) minus ($\mu_{\text{parent}} - 2\sigma_{\text{parent}}$) of the lowest parent. Similarly, the range of values for each hybrid group was calculated as ($\mu_{\text{F1-hybrid}} + 2\sigma_{\text{F1-hybrid}}$) $-$ ($\mu_{\text{F1-hybrid}} - 2\sigma_{\text{F1-hybrid}}$) and ($\mu_{\text{BC}} + 2\sigma_{\text{BC}}$) $-$ ($\mu_{\text{BC}} - 2\sigma_{\text{BC}}$). Transcripts were qualified as transgressive if the hybrid had a range of values greater than the maximum range of the parents. The total number of transgressive transcripts (as defined above) for each hybrid group was then calculated. We estimated the total number of false positive transgressive transcripts expected by permuting for each transcript separately, the identity of each of the four groups and then calculating a new total number of transgressive transcripts for hybrids as aforementioned (4 groups $\times$ 8 replicate loops = 32 columns permuted using the function "sample" in R). *P* value was then estimated as the number of times, over 1,000 permutations, that the total number of false positive transgressive transcripts was greater than the observed number of transgressive transcripts. Lastly, we examined the function and identity of the transcripts that showed the highest transgressivity values in hybrids (range of hybrid, either F1-hybrid or backcross, minus maximum range of parents >0.5). This cutoff value, although arbitrary, permits to identify the most severely transgressive transcripts, which one would predict to be more likely biologically relevant compared with nearly transgressive transcripts. *P* values for those transcripts were also estimated as the number of times a randomized data set would produce a hybrid transcript more transgressive than the real calculated value (1,000 permutations).

Results

This study primarily focuses on hybrid gene expression relative to that reported recently between pure parentals by Nolte et al. (2009). As such, results obtained for pure parental comparisons are not treated in detail but reported for the sake of comparison only. In brief, Nolte et al. (2009) found 33 and 502 transcripts differentially expressed in the embryo and juveniles, respectively. Their numbers are slightly different than the ones we obtained for the same comparisons, that is, 5 and 543 transcripts, respectively. This discrepancy is likely due to the overall different number of technical replicates in both studies. Yet it does not influence any of the earlier conclusions reported by these authors, namely, that the number of significant transcripts between normal and dwarf is much larger at the juvenile than the embryonic stage.

Gene Expression Differentiation among Groups

The number of transcripts (Expressed Sequence Tag clones spotted on the microarray) for which we obtained gene expression data of sufficient quality for subsequent analyses was 4,950 and 5,359 for the embryos and for the juvenile data set, respectively. After an FDR correction (0.05) and a compression of replicate spots, we identified five transcripts (0.1% of all transcripts expressed) and 573 transcripts (12%) as differentially expressed in the

Table 2
Number of Transcripts Identified in the ANOVA (FDR Corrected Permutation *P* Value < 0.05, Fs Test Option, 1,000 Permutations) and the Subsequent *t*-Tests between the Different Experimental Groups for Embryos and Juveniles (FDR Corrected Permutation *P* Value < 0.05, Fs Test Option, 1,000 Permutations)

		N	D	F1
Embryos				
ANOVA	5			
Backcross		0	5	2
Normal			5	0
Dwarf				1
Juveniles				
ANOVA	573			
Backcross		343	161	177
Normal			501	94
Dwarf				403

NOTE.—N = Normal, D = Dwarf, and F1 = F1-hybrid.

Table 3
Number of Transcripts Identified as Additive, Dominant, NonAdditive, or Transgressive

	F1-Hybrid	Backcross
Additive		
Embryos	1 (60%)	2 (40%)
Juveniles	217 (44%)	98 (20%)
Dominant		
Embryos	4 (40%)	3 (60%)
Juveniles	237 (47%)	133 (27%)
Nonadditive		
Embryos	0 (0%)	0 (0%)
Juveniles	47 (9%)	269 (54%)
Overdominant	19 (3%)	100 (20%)
Underdominant	28 (6%)	169 (34%)
Overexpressed	5 (7%)	15 (20%)
Underexpressed	3 (4%)	3 (4%)
Transgressive		
Embryos	1,306 (26%)	2,622 (53%)
Juveniles	2,097 (39%)	2,316 (43%)

NOTE.—See Material and Methods for criteria defining the categories. Percentage relate to the total number of transcripts differentiated between the parents (5 in embryo data set, 501 in juvenile data set), except for over and underexpressed transcripts where percentage refers to the total number of genes specifically differentiated in the juvenile hybrids (75), and for transgressive transcripts, where percentages refer to the total number of transcripts significantly expressed in the embryo (4,950) and juvenile (5,359) data set.

embryo and juvenile data sets, respectively among all groups compared. In the embryo data set, most of that difference was due to normal–dwarf comparison, but the small number of differentiated transcripts hampered the interpretation of this trend (table 2). In the juvenile data set, most of the observed differences were in the normal–dwarf comparison because 501 of the 573 transcripts (87%) were differentially expressed between the parents. For comparisons involving hybrids, there was much less differentiation between F1 and normal (94 genes or 16% of all transcripts differentially expressed) compared with differences between F1 and dwarf (403 genes or 70%). In contrast, 161 (28%) transcripts differed in the backcross–dwarf comparison compared with 343 (60%) in the backcross–normal comparison. Finally, 177 (31%) transcripts significantly differed between F1 and BC hybrids (see tables 2 and supplementary table 1, Supplementary Material online, for all the transcript IDs, fold changes, and relative expression).

Estimates of Gene Expression Variability

In the embryo data set, 799 transcripts (16% of all transcripts expressed) showed heterogeneity of variance in expression between groups (Bartlett test, $P < 0.05$). Of those, backcross hybrids showed the highest mean and median values of variance, followed by dwarf, F1-hybrids, and normal. All comparisons were significant (Wilcoxon test, $P < 0.0001$, fig. 2). In juvenile fish, 656 transcripts (13%) showed significant heterogeneity of variance among groups (Bartlett test, $P < 0.05$). On average, F1-hybrid and backcross were more variable than both normal and dwarf (Wilcoxon test, $P < 0.0001$) but not different from one another (Wilcoxon test, $P = 0.7$) (fig. 3).

Type of Inheritance Observed in Hybrids

It is noteworthy that the same transcripts generally showed different patterns of expression in F1 and backcross hybrids such that there was actually little correlation between *d*/*a* values for specific transcripts in F1 and in backcross hybrids ($r^2 = 0.1$ for juveniles). Secondly, the distribution of dominance effect was skewed toward pure normal whitefish (mean = 0.38, $P < 0.0001$, *t*-test) in the F1-hybrid distribution. This reflects the fact that, as stated above, patterns of expression observed in F1-hybrids were generally more similar to normal than dwarf whitefish. Conversely, backcross *d*/*a* ratio distribution displayed skewness toward dwarf (mean = −1.27, $P < 0.0001$, *t*-test) (fig. 4).

Embryos

Because the identification of inheritance patterns relies mostly on differentially expressed transcripts, the very small number of differentiated transcripts hampered the use of this approach in the embryos. On the other hand, we identified a large fraction of transcripts showing evidence of transgressivity both in F1-hybrids (1,306, or 26% of all transcripts expressed) and backcross (2,622, 53%; table 3). The number of false positive transgressive transcripts expected based on 1,000 permutation was 1,452 (29%), and consequently there were significantly more transgressive transcripts than expected by chance in backcross ($P = 0.01$) but not F1-hybrids ($P = 0.55$). Nine transcripts showed very high transgressivity in backcross hybrids (range of backcross − maximum range of parents >0.5). These transcripts, not differentially expressed in any comparisons, are involved in protein folding, mRNA translation, signal transduction, germ-line formation, and endocytosis (table 4). Moreover, five of those nine transcripts closely match with three different homologs ("Immunoglobulin binding protein" [protein folding], E-value: 1e−157; "translation elongation factor alpha 1" [mRNA translation], E-value: 1e−15 and "40S ribosomal protein

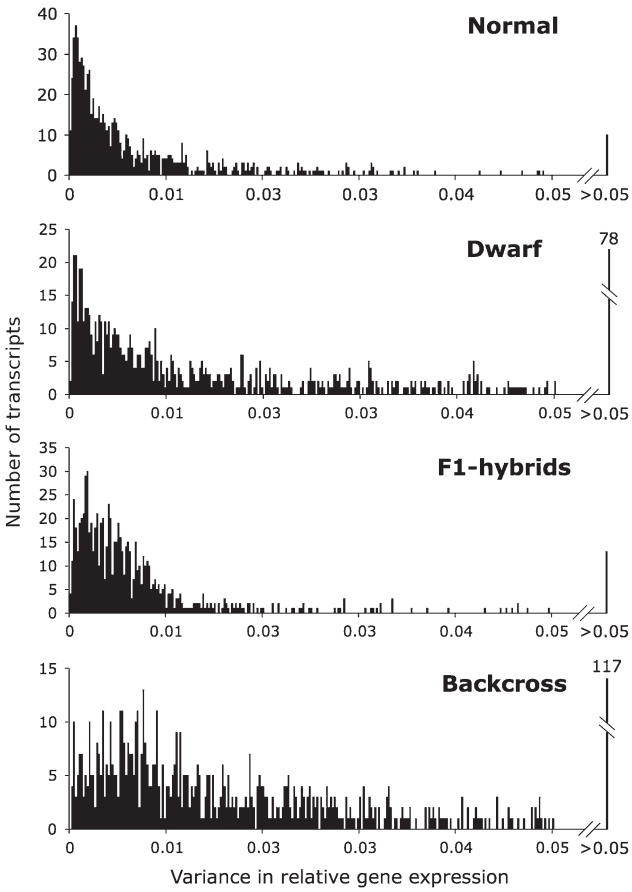

FIG. 2.—Frequency distribution of variance of relative gene expression (embryo data set) for the four groups for genes showing heterogeneity of variance according to a Bartlett test ($P < 0.05$). Backcross shows the greatest mean and median of variance (mean = 0.026, median = 0.014), followed by dwarf (mean = 0.018, median = 0.0089), F1-hybrids (mean = 0.0075, median = 0.0046), and normal (mean = 0.0081, median = 0.0035). All pairwise comparisons are significant (Wilcoxon rank test, $P < 0.0001$).

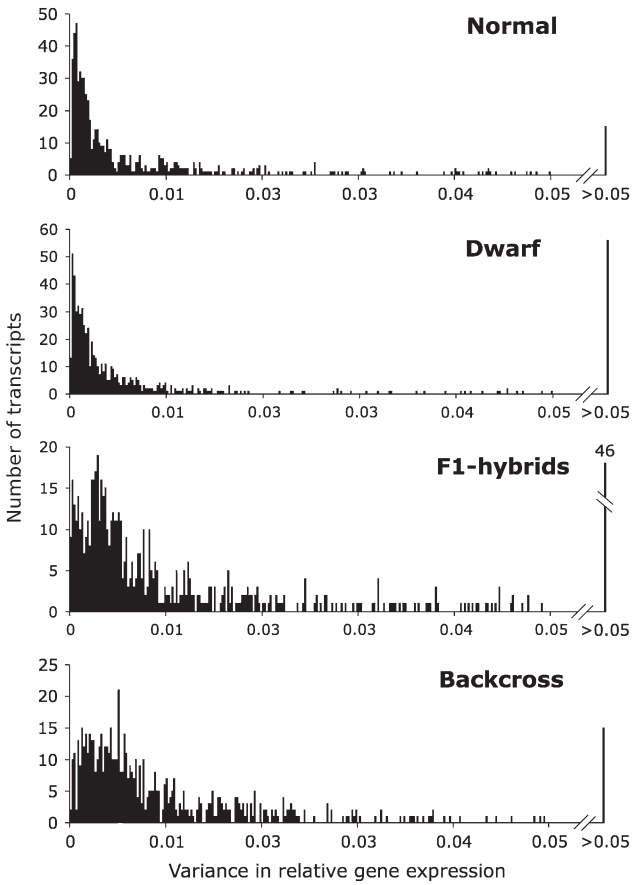

FIG. 3.—Frequency distribution of variance of relative gene expression (juvenile data set) for the four groups for genes showing heterogeneity of variance according to a Bartlett test ($P < 0.05$). Backcross (mean = 0.012, median = 0.0059) and F1-hybrids (mean = 0.014, median = 0.0053) show comparable and greatest mean and median of variance followed by dwarf (mean = 0.017, median = 0.0024) and normal (mean = 0.008, median = 0.025). BC–F1-hybrid comparison and N–D comparison are not significant (Wilcoxon rank test, $P = 0.76$ and $P = 0.86$, respectively), whereas all other comparisons are significant (Wilcoxon rank test, $P < 0.0001$).

s11" [mRNA translation], E-value: 3e−93) identified as essential for early embryonic development of *Danio rerio* (Amsterdam et al. 2004) (table 4).

Juvenile Stage

For the 501 transcripts that were differently expressed between the pure forms at the juvenile stage, 217 (44%) and 98 (20%) had a relative expression resembling additivity for F1 and backcross hybrids, respectively. A total of 237 transcripts (47%) had an inheritance pattern close to dominance in F1-hybrids compared with 133 (27%) in backcross. A total of 47 (9%) and 269 (54%) of all transcripts showed level of expression that fell outside the mean of the parents (nonadditive) for F1-hybrid and backcross, respectively ($|d/a|$ ratio > 1.5; table 3). Twenty-eight and 169 transcripts were underdominant compared with 19 and 100 overdominant in F1-hybrid and backcross, respectively. Seventy-five transcripts were not differentiated between the parents (supplementary table 1, Supplementary Material online), yet still identified in the ANOVA; most (49) of those significantly differentiated from only one parent yet never from both. In juvenile F1-hybrid, three (underexpressed) and five (overexpressed) transcripts had a mean level of expression falling significantly outside the parental range. Under the same criteria, 3 and 15 transcripts were, respectively, under and overexpressed in the backcross hybrids.

We detected a large fraction of transcripts showing evidence of transgressivity both in F1-hybrids (2,097, or 39% of all transcripts expressed) and backcross (2,316, 43%). The number of false positive transgressive transcripts expected based on a randomized data set was 1,605 (33%), and thus, there were more transgressive transcripts than expected by chance, yet this was nonsignificant in F1 (P value = 0.16, based on 1,000 permutations as described in Material and Methods) and backcross hybrids ($P = 0.09$), respectively. Eighteen and three transcripts showed highly significant transgressivity in F1 and backcross hybrids, respectively. These transcripts, not differentially expressed in any comparisons, belong to muscle contraction, energy metabolism, lipid metabolism, protein degradation, and transport functional categories (table 4).

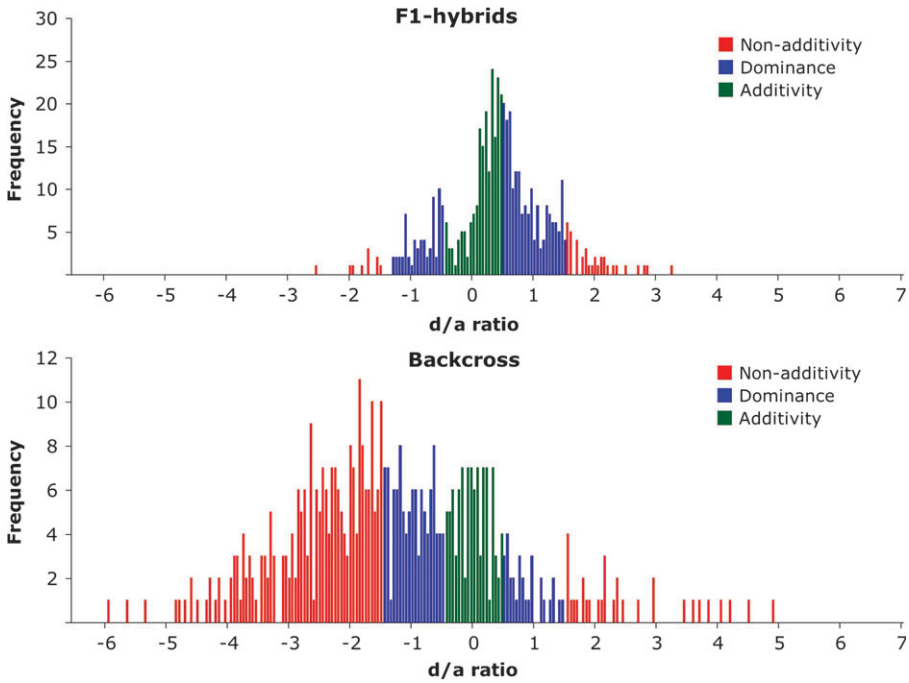

FIG. 4.—Distribution of dominance effects (*d*/*a* ratio) for F1-hybrids and backcross for transcripts significantly different between the parents (501 juvenile transcripts). A |*d*/*a*| ratio between 0.5 and 1.5 is considered as dominant, >1.5 nonadditive, <0.5 additive (see Material and Methods). Positive values imply that hybrids are more closely related to the normal parent whereas negative values imply that hybrids are more closely related to the dwarf parent.

Discussion

The main objective of this study was to document patterns of gene expression divergence in first (normal × dwarf) and second-generation hybrid crosses (backcross: [normal × dwarf] × normal], and compare them with pure normal and dwarf parental forms, at both embryonic and juvenile ontogenetic stages. More specifically, under the assumption that similar genes involved in the adaptive divergence of these species are also responsible for driving their reproductive isolation (Nolte et al. 2009), we predicted more evidence of misexpression at the juvenile compared with the embryonic stage. In general, our results supported this prediction as we observed that few genes differed in average expression in hybrids compared with parentals at the embryonic stage, whereas many more did so at the juvenile stage. Secondly, we predicted more evidence of hybrid misexpression in backcross hybrids and the fact that nonadditivity was more prevalent in backcross compared with F1-hybrids supported this prediction. Lastly, extreme transgressivity of several key developmental genes was observed in backcross embryos. This emphasizes that, at the transcriptomic level, intrinsic hybrid misexpression may also play a role in explaining reproductive isolation of dwarf and normal whitefish. Below, we discuss the potential implications of those results, also considering the limitations of the data.

Patterns of Inheritance in Hybrids

Strikingly, patterns of inheritance were quite distinct between backcross and F1-hybrids, and *d*/*a* values were not correlated between them. Under an additive model of inheritance, we would also have expected F1-hybrid to be the midvalue of their parents and backcross to be closer to the normal phenotype (with which they share 75% of their genome). This was not the case, as more genes differentiated F1-hybrids to dwarf whereas, and to a lesser extent, more genes differentiated backcross to normal, a result also exemplified by the asymmetry in the direction of dominance in both F1-hybrids (normal dominance) and backcross (dwarf dominance). This idiosyncratic result cannot be ignored, yet is difficult to interpret beyond the fact that, as it has been emphasized many times before, gene expression is a complex phenotype, whose behavior is hard to predict and whose inheritance often does not follow simple Mendelian rules (Rockman and Kruglyak 2006). Indeed, this asymmetry of gene expression divergence toward one parent is apparently quite common in F1-hybrids event though the underlying mechanistic reasons responsible for this trend are poorly understood (*Drosophila*: Ranz et al. 2004; Gibson et al. 2004; *Mus*: Rottscheidt and Harr 2007; and *Salvelinus*: Mavárez et al. submitted).

The Transcriptomic Basis of Ecological (Extrinsic) Reproductive Isolation Factors

Previous gene expression studies (Derome et al. 2006; St-Cyr et al. 2008; Nolte et al. 2009) combined with physiological data (Trudel et al. 2001) have shown that changes in the expression of metabolic genes are largely responsible for the physiological adaptation to distinct whitefish benthic (normal) and limnetic (dwarf) niches. Notably, a suite of six key metabolic genes (glyceraldehyde-3-phosphate

Table 4
Range of Relative Gene Expression Level (in %) of Highly Transgressive Transcripts for Both Embryo and Juvenile Data Sets

Transcript ID	Maximum Range (Parents)	Range (Backcross)	Range (F1-Hybrids)	Gene Product	Functional Group
Embryos					
CB485951	[36–181]	[0–233]***	[50–157]	Heat shock cognate 70 kDa protein[a]	Protein folding
CK991158	[46–165]	[17–214]**	[72–136]	Heat shock cognate 70 kDa protein[a]	Protein folding
CB516765	[29–164]	[15–219]**	[39–190]	Fish-egg lectin	Lipopolysaccharide-binding protein
CA060826	[38–179]	[9–226]*	[66–148]	Elongation factor 1 alpha[b]	Translation
CK990889	[48–155]	[28–197]**	[76–139]	Guanine nucleotide-binding protein	Signal transduction
CA051954	[56–150]	[34–198]**	[78–126]	Protein kinase C	Germ line formation
CB502683	[39–174]	[16–224]**	[74–135]	Heat shock cognate 70 kDa protein[a]	Protein folding
CB502825	[69–125]	[45–151]**	[65–156]	Asialoglycoprotein receptor 2	Endocytosis
CN442505	[58–123]	[37–153]***	[72–156]	40S ribosomal protein S11[c]	Translation
Juveniles					
CB507670	[78–123][a]	[74–128]	[49–145]***	Collagen alpha-2(I) chain precursor	Muscle contraction
CA064346	[59–126]	[97–157]	[24–142]**	Proproteinase E precursor	Protein degradation
CA053777	[76–117]	[76–120]	[67–159]***	SJCHGC04882	Unknown
CK990215	[18–178]	[14–156]	[4–216]*	Unknown	Unknown
CA043836	[67–120]	[62–123]	[63–168]**	Phosducin-like protein 3	Other
CA037858	[70–135]	[58–144]	[30–147]***	Unknown	Unknown
CB488336	[94–117]	[78–122]	[52–126]***	Collagen alpha-1(I) chain precursor	Muscle contraction
CB510992	[76–138]	[65–126]	[33–148]***	Apolipoprotein A-I precursor	Lipid metabolism
CB504468	[68–132]	[87–136]	[29–147]***	Elastase-1	Muscle contraction
CA038612	[60–146]	[70–148]	[12–151]**	Serotransferrin-2 precursor	Transport
CB500533	[84–115]	[65–141]	[41–137]**	Collagen alpha-1(I) chain precursor	Muscle contraction
CB507066	[80–116]	[74–128]	[41–143]***	Collagen alpha-1(I) chain precursor	Muscle contraction
BU965755	[86–123]	[73–131]	[45–138]***	Coiled-coil domain	Other
CA038358	[86–121]	[76–128]	[45–137]**	Proteasome subunit alpha type 2	Protein degradation
CB496771	[58–148]	[65–148]	[7–159]**	Serotransferrin precursor	Transport
CA043815	[67–144]	[67–124]	[21–163]***	Unknown	Unknown
CB492384	[82–111]	[55–139]***	[71–147]	Creatine kinase B-type	Energy metabolism
CB497013, CB496702	[47–131]	[47–200]**	[29–175]**	Myosin heavy chain	Muscle contraction

NOTE.—See Materials and Methods for criteria defining the categories. Values below 100% imply underexpression compared with the average of all four groups, whereas values over 100% imply overexpression compared with the average of all four groups. Transcripts for which the gene expression range of hybrid minus the maximum range of the parents >50: ***P value < 0.01, **P value < 0.05, *P value < 0.1 (based on 1,000 permutations as described in Materials and Methods). Gene product and functional groups are based on the latest annotation file provided by cGRASP (May 2008).

Corresponding homologs and knockdown phenotypes in *Danio rerio* study (Amsterdam et al. 2004):

[a] Immunoglobulin-binding protein (BlastN, E value: 1e−157). Knockdown phenotype: Day 1: pinched midbrain/hindbrain boundary. Day 2: small head and eyes, inflated hindbrain ventricle, thin body. Day 5: very small head and eyes, thin body with underdeveloped liver/gut.

[b] Eukaryotic translation elongation factor 1 alpha (BlastN, e value: 1e−15). Knockdown phenotype: Day 3: small head and eyes, rounder yolk. Day 5: increasingly necrotic.

[c] 40S ribosomal protein s11 (BlastN, E value: 3e−93) -Day 1: pinched midbrain/hindbrain boundary. Day 2: small head and eyes, inflated hindbrain ventricle, thin body. Day 5: small head and eyes, thin or necrotic body, round gray yolk, and underdeveloped liver/gut.

dehydrogenase, Fructose–bisphosphate aldolase A, Beta-enolase, Trypsin-1 precursor, Cytochrome c oxidase polypeptide VIa, and Nucleoside diphosphate kinase) was identified as consistently divergent between normal and dwarf whitefish (Nolte et al. 2009). We may then hypothesize that misexpression for those metabolic genes could contribute, to an atypical physiological phenotype and to an inferior, ecologically maladapted, individual. Here, we found that, in F1-hybrids juveniles, those genes mostly showed an intermediate pattern of expression (supplementary table 1, Supplementary Material online) and no transgressivity compared with parents. In backcross hybrids, two of those genes (G3PDH, FBPA A) showed additivity of expression, the rest being slightly nonadditive, whereas none revealed transgressivity.

According to the ecological theory of adaptive radiation, intermediate hybrid phenotypes may be selected against if no suitable ecological niche for them exists in nature (Schluter 2000). Recent work in sticklebacks (Gow et al. 2007) and cichlids (van der Sluijs et al. 2008) has shown such environment driven natural selection may be key in explaining incipient population divergence. As such, reproductive isolation of lake whitefish could be at least partly seen as a by-product of divergent selection acting

on metabolic genes. Admittedly, a clear demonstration of the association between the expression of such genes and phenotypic variation between dwarf and normal whitefish is lacking, and we are currently conducting quantitative trait loci, expression quantitative trait loci, and gene mapping studies toward this end (Renaut S, Nolte AW, Bernatchez L, unpublished data).

A Possible Role of Transgressivity in Reproductive Isolation

Gene expression in hybrids was generally more variable than parental, both at embryonic and juvenile stages. One might argue that the patterns of variance observed are confounded by a different number of families used in each treatment. Although this cannot be entirely ruled out, it is noteworthy that backcross individuals, who showed the highest level of variance in gene expression, consisted of a single female crossed to five males, relative to all other treatments, which consisted of many half-sib families. This family effect probably also explains the relatively large variance of dwarf whitefish, which comprised both half-sib families and many wild-caught natural families. It then seems unlikely that this factor would explain the general increased variance observed in the backcross group. Alternatively, recombination can release hidden variation and generate transgressive phenotypes (Rieseberg et al. 1999, 2003; Mallet 2007). Populations are known to accumulate cryptic variation only revealed under certain genotypic or environmental conditions (Le Rouzic and Carlborg 2007). Recently, Landry et al. (2007) have illustrated how in hybrids, the regulation of coevolved *cis* regulatory regions and *trans* transcription factors could be disrupted and lead to increased phenotypic novelties in hybrids. This may explain why many whitefish genes showed increased variance in expression and transgressivity in hybrids and yet were not differentially expressed between parentals. In fact, all the highly transgressive transcripts presented in table 4 were not differentially expressed in any comparisons.

Transgressive segregation may in some cases create fitter phenotypes (e.g., hybrid species resulting from selected "hopeful monsters", Barton 2001; Mallet 2007). Conversely, it also underlies postzygotic isolation mechanisms such that transgressive hybrids often suffer a highly reduced survival (Barton 2001; Coyne and Orr 2004). We propose that the overall patterns of transgressivity we observed, including misexpression of several key developmental genes, may contribute to abnormal hybrid development and increased embryonic mortality identified by Lu and Bernatchez (1998) and also Rogers et al. 2007 as a plausible postzygotic reproductive isolation mechanism. Namely, five of the nine transcripts identified as highly transgressive in embryos and involved in protein folding and mRNA translation are especially good candidates because knockdown mutants in *D. rerio* for those genes are known to show visible embryonic defect and almost invariably die prior to, or early after, hatching (Amsterdam et al. 2004). This proportion (five of nine or 54%) was also significantly higher ($P < 0.001$, one tailed Fisher's Exact test) than the actual proportion of transcripts in the whole embryo data set that matched to essential genes identified in *D. rerio* (257 of 4,950 or 5%). In fact, the abnormal phenotypes described in the *D. rerio* study closely match our own observation that a large fraction of backcross eggs (35%) started to show visible defects (asymmetric axial body plan, small eyes, heart not beating, deformed tail) 15 days after our sampling and eventually die prior to hatching (Renaut S, unpublished data). A previous study on reproductive isolation in lake whitefish also showed that a large fraction of the backcross progeny died around the same developmental time (Rogers and Bernatchez 2006). Of course, there is a leap between linking a knockdown mutation completely obliterating a gene product (as it is the case in the *D. rerio* study) and a simple increase in biological variation. Yet, it is noteworthy that these key developmental genes are the most transgressive of 4,950 surveyed in the embryo data set and were generally significantly more transgressive than expected by chance. Moreover, it is plausible that hybrid genetic combinations creating even greater misexpression for those genes may have caused early lethality (prior to our sampling) and thus may have reduced our ability to pick out such abnormal phenotypes. Consequently, the increased patterns of variance observed in hybrid embryos are likely to be conservative estimates.

Gene Expression Studies of Speciation

Most gene expression differentiation we observed was between normal and dwarf parental forms rather than between hybrids. These results contrast with many recent gene expression–speciation studies that have identified pervasive nonadditive patterns of gene expression in first generation hybrids (see recent reviews by Landry et al. 2007; Ortiz-Barrientos et al. 2007). Namely, our study brings new lights into gene expression studies of speciation for two main reasons. Firstly, the bulk of the work, done mostly in *Drosophila*, particularly in the *melanogaster* group (Michalak and Noor 2003; Ranz et al. 2004; Landry et al. 2005; Haerty and Singh 2006; Moehring et al. 2007), and to a lesser extent in *Xenopus* (Malone et al. 2007) and *Mus* (Rottscheidt and Harr 2007) involves biological species that have diverged millions of years ago (e.g., *D. simulans–Drosophila mauritiana*: 0.93 Ma, *D. simulans–D. melanogaster*: 5.1 Ma (Tamura et al. 2004), *Xenopus laevis–Xenopus muelleri:* >20 Ma (Evans et al. 2004), *Mus musculus* subspecies: 0.3–1.0 Ma (Boursot et al. 1996)]. Because genetic incompatibilities continue to accumulate over time even after complete reproductive isolation has been established, this complicates the identification of the loci that initially led to the divergence event (Mallet 2006). In these studies, most hybrids are known to be poorly fit, sterile, or simply inviable (Ranz et al. 2004) rendering it difficult to disentangle whether gene expression misexpression is the cause rather than the consequence of hybrid inviability. In contrast, in young diverging lineages such as whitefish species pairs that diverged 12,000–15,000 years ago (Bernatchez 2004), much less genetic divergence is expected due to frequent gene flow or recent common ancestry. The effects of hybridization may then be subtler

and the genetic changes identified more likely to be involved in the very early steps of reproductive isolation. Secondly, even though the effect of early reproductive barriers may be more important in later hybrid generations (Barton 2001), most recent studies have focused on first generation hybrids. Clearly, we identified different genes and patterns of inheritance in first and second-generation hybrids. Moreover, backcross hybrids have also revealed increased nonadditivity as well as transgressive expression of essential developmental genes. To our knowledge, no gene expression studies of speciation in natural systems have previously used genomewide gene expression data to compare the expression profile of second-generation hybrids with that of parental lineages.

Conclusion

We have attempted to determine the role of gene expression divergence in the development of reproductive isolation of recently evolved lineages of lake whitefish. We demonstrated that F1 and backcross hybrids showed different patterns of expression and that gene misexpression at both embryonic and juvenile stages might take different forms (intermediacy, nonadditivity and transgressivity of expression). We then identified candidate genes whose role in driving reproductive isolation will have to be further confirmed. As pointed out in a recent review, the identification of candidate genes does not constitute an end in itself, but rather the beginning of a new set of evolutionary relevant questions (Stinchcombe and Hoesktra 2007). Ultimately, by combining information obtained through gene expression studies such as this one, to QTL mapping, gene sequencing, mapping and genome scan data, we will be able to better answer questions regarding the underlying genetic architecture of adaptive traits and expression phenotypes, the role of standing genetic variation and de novo mutations in driving the emergence of those traits, and the role of natural selection and/or drift in maintaining this divergence.

Supplementary Material

Supplementary table 1 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We would like to thank the Ministère des Ressources Naturelles et de la Faune du Québec (MRNF) for issuing scientific collecting permits; J.-C. Therrien and S. Higgins (LARSA) for their assistance in spawning and rearing whitefish; J. Mavárez, E. Normandeau, and C. Roberge for many insightful suggestions and help with the analysis. This project was funded by a Natural Science and Engineering Research Council of Canada (NSERC) postgraduate scholarship to S.R., NSERC Discovery grant and Canadian Research Chair to L.B., and a postdoctoral research stipend from the German Research Foundation to A.N. This paper is a contribution to the research program of Québec Océan.

Literature Cited

Amsterdam A, Nissen RM, Sun Z, Swindell EC, Farrington S, Hopkins N. 2004. Identification of 315 genes essential for early zebrafish development. Proc Natl Acad Sci USA. 101:12792–12797.

Bartlett MS. 1937. Properties of sufficiency and statistical tests. Proc R Soc Lond B Biol Sci. 160:268–282.

Barton NH. 2001. The role of hybridization in evolution. Mol Ecol. 10:551–568.

Bauer DF. 1972. Constructing confidence sets using rank statistics. J Am Stat Assoc. 67:687–690.

Bell G. 2008. Selection: the mechanisms of evolution, 2nd ed.. New York: Oxford University Press.

Bernatchez L. 2004. Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (Salmoniformes). In: Hendry AP, Stearns SC, editors. Evolution illuminated: salmon and their relatives. Oxford: Oxford University Press. p. 175–207.

Bernatchez L, Chouinard A, Lu GQ. 1999. Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. Bio J Lin Soc. 68: 173–194.

Boursot P, Din W, Anand R, Darviche D, Dod B, VonDeimling F, Talwar GP, Bonhomme F. 1996. Origin and radiation of the house mouse: mitochondrial DNA phylogeny. J Evol Biol. 9:391–415.

Brem RB, Kruglyak L. 2005. The landscape of genetic complexity across 5,700 gene expression traits in yeast. Proc Natl Acad Sci USA. 102:1572–1577.

Burke JM, Arnold ML. 2001. Genetics and the fitness of hybrids. Ann Rev Genet. 35:31–52.

Burton RS, Ellison CK, Harrison JS. 2006. The sorry state of F2 hybrids: consequences of rapid mitochondrial DNA evolution in allopatric populations. Am Nat. 168:S14–S24.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell. 134:25–36.

Coyne JA, Orr HA. 2004. Speciation. Sunderland (MA): Sinauer Associates, Inc.

De Queiroz K. 1998. The general lineage concept of species, species criteria, and the process of speciation. In: Howard DJ, Berlocher SH, editors. Endless forms species and speciation. New York: Oxford University Press. p. 57–75.

Derome N, Bougas B, Rogers SM, 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.

Derome N, Duchesne P, Bernatchez L. 2006. Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes. Mol Ecol. 15:1239–1249.

DeVicente MC, Tanksley SD. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics. 134:585–596.

Dobzhansky T. 1937. Genetics and the origin of species. New York: Columbia University Press.

Endler JA. 1977. Geographic variation, speciation, and clines. Princeton (NJ): Princeton University Press.

Evans BJ, Kelley DB, Tinsley RC, Melnick DJ, Cannatella DC. 2004. A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. Mol Phyl Evol. 33:197–213.

Gibson G, Riley-Berger R, Harshman L, Kopp A, Vacha S, Nuzhdin S, Wayne M. 2004. Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*. Genetics. 167:1791–1799.

Gow JL, Peichel CL, Taylor EB. 2007. Ecological selection against hybrids in natural populations of sympatric threespine sticklebacks. J Evol Biol. 20:2173–2180.
Haerty W, Singh RS. 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.
Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution. 61:995–1016.
Hughes KA, Ayroles JF, Reedy MM, Drnevich JM, Rowe KC, Ruedi EA, Cáceres CE, Paige KN. 2006. Segregating variation in the transcriptome: cis regulation and additivity of effects. Genetics. 173:1347–1355.
Jeukens J, Bittner D, Knudsen R, Bernatchez L. 2009. Candidate genes and adaptive radiation: insights from transcriptional adaptation to the limnetic niche among coregonine fishes (*Coregonus* spp., *Salmonidae*). Mol Biol Evol. 26:155–166.
Kendziorski C, Irizarry RA, Chen KS, Haag JD, Gould MN. 2005. On the utility of pooling biological samples in microarray experiments. Proc Natl Acad Sci USA. 102: 4252–4257.
Kerr MK, Martin M, Churchill GA. 2000. Analysis of variance for gene expression microarray data. J Comp Biol. 7:819–837.
Landry CR, Hartl DL, Ranz J. 2007. Genome clashes in hybrids: insights from gene expression. Heredity. 99:483–493.
Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. 2005. Compensatory cis–trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. Genetics. 171:1813–1822.
Le Rouzic A, Carlborg O. 2007. Evolutionary potential of hidden genetic variation. Trends Ecol Evol. 23:33–37.
Lu G, Bernatchez L. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. Evolution. 53:1491–1505.
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 B Biol Sci. 265:1025–1030.
Mallet J. 2007. Hybrid speciation. Nature. 446:279–283.
Mallet J. 2006. What does Drosophila genetics tell us about speciation? Trends Ecol Evol. 21:386–393.
Malone JH, Chrzanowski TH, Michalak P. 2007. Sterility and gene expression in hybrid males of *Xenopus laevis* and *X. muelleri*. PLoS ONE. 2:e781.
Mavárez J, Audet C, Bernatchez L. Submitted. Ecological divergence drives disruption of gene expression in hybrids between two sympatric populations of Brook Charr (*Salvelinus fontinalis*). Mol Biol Evol.
Meiklejohn CD, Townsend JP. 2005. A Bayesian method for analysing spotted microarray data. Brief Bioinform. 6: 318–330.
Michalak P, Noor MA. 2003. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. Mol Biol Evol. 20:1070–1076.
Moehring AJ, Teeter KC, Noor MA. 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.
Muller HJ. 1942. Isolating mechanisms, evolution, and temperature. Biol Symp. 6:71–125.
Nolte AW, Renaut S, Bernatchez L. 2009. Regulatory divergence of life history stages of whitefish and the emergence of postzygotic isolation. BMC Evol Biol.
Nosil P. 2008. Speciation with gene flow could be common. Mol Ecol. 17:2103–2106.
Ortiz-Barrientos D, Counterman BA, Noor MAF. 2007. Gene expression divergence and the origin of hybrid dysfunctions. Genetica. 29:71–81.
Ranz JM, Machado C. 2006. Uncovering evolutionary patterns of gene expression using microarrays. Trends Ecol Evol. 21:29–37.
Ranz JM, Namgyal K, Gibson G, Hartl D. 2004. Anomalies in the expression profile of interspecific hybrids of Drosophila melanogaster and Drosophila simulans. Genome Res. 14:373–379.
Rieseberg LH, Archer MA, Wayne RK. 1999. Transgressive segregation, adaptation and speciation. Heredity. 83:363–372.
Rieseberg LH, Raymond O, Rosenthal DM, Lai Z, Livingstone K, Nakazato T, Durphy JL, Schwarzbach AE, Donovan LA, Lexer C. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. Science. 301:1211–1216.
Roberge C, Normandeau E, Einum S, Guderley H, Bernatchez L. 2008. Genetic consequences of interbreeding between farmed and wild Atlantic salmon: insights from the transcriptome. Mol Ecol. 17:314–324.
Rockman M, Kruglyak L. 2006. Genetics of global gene expression. Nat Rev Genet. 7:862–872.
Rogers SM, Bernatchez L. 2005. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). Mol Ecol. 14:351–361.
Rogers SM, Bernatchez L. 2006. The genetic basis of intrinsic and extrinsic post-zygotic reproductive isolation jointly promoting speciation in the lake whitefish species complex (*Coregonus clupeaformis*). J Evol Biol. 19:1979–1994.
Rogers SM, Bernatchez L. 2007. The genetic architecture of ecological speciation and the association with signatures of selection in Natural Lake Whitefish (*Coregonus* sp. Salmonidae) species pairs. Mol Biol Evol. 24:1423–1438.
Rogers SM, Gagnon V, Bernatchez L. 2002. Genetically based phenotype–environment association for swimming behavior in lake whitefish ecotypes (*Coregonus clupeaformis*, Mitchill). Evolution. 56:2322–2329.
Rogers SM, Isabel N, Bernatchez L. 2007. Linkage maps of the dwarf and normal lake Whitefish (*Coregonus clupeaformis*) species complex and their hybrids Reveal the genetic architecture of population divergence. Genetics. 175: 375–398.
Rottscheidt R, Harr B. 2007. Extensive additivity of gene expression differentiates subspecies of the house mouse. Genetics. 177:1553–1567.
Rundle HD, Nagel L, Boughman JW, Schluter D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. Science. 287:306–308.
Schluter D. 2000. The ecology of adaptive radiation. New York: Oxford University Press.
St-Cyr J, Derome N, Bernatchez L. 2008. The transcriptomics of life-history trade-offs between whitefish species pairs (*Coregonus* sp.). Mol Ecol. 17:1850–1870.
Stinchcombe JR, Hoesktra HE. 2007. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity. 100:158–170.
Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. Mol Biol Evol. 21:36–44.
Townsend JP, Hartl DL. 2002. Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. Genome Biol. 3:RESEARCH0071-0071.16.
Trudel M, Tremblay A, Schetagne R, Rasmussen J. 2001. Why are dwarf fish so small? An energetic analysis of polymorphism

in lake whitefish (*Coregonus clupeaformis*). Can J Fish Aquat Sci. 58:394–405.
van der Sluijs I, van Dooren TJM, Seehausen O, van Alphen JJM. 2008. A test of fitness consequences of hybridization in sibling species of Lake Victoria cichlid fish. J Evol Biol. 21:480–491.
von Schalburg KR, Rise ML, Cooper GA, Brown GD, Gibbs AR, Nelson CC, Davidson WS, Koop BF. 2005. Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray. BMC Genomics. 15:126.
Whiteley AR, Derome N, Rogers SM, St-Cyr J, Nolte AW, Renaut S, Jeukens J, Laroche J, Labbe A, 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 GA. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet. 8:206–216.
Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. 2003. The Evolution of transcriptional regulation in eukaryotes. Mol Biol Evol. 20: 1377–1419.
Wu CI, Ting CT. 2004. Genes and speciation. Nat Rev Genet. 5:114–122.

Koichiro Tamura, Associate Editor

Accepted January 23, 2009