

Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchill) ecotypes

N. DEROME, P. DUCHESNE and L. BERNATCHEZ

Québec Océan, Département de Biologie, Université Laval, Ste-Foy, Québec G1K 7P4, Canada

Abstract

We tested the hypothesis that phenotypic parallelism between dwarf and normal whitefish ecotypes (*Coregonus clupeaformis*, Salmonidae) is accompanied by parallelism in gene transcription. The most striking phenotypic differences between these forms implied energetic metabolism and swimming activity. Therefore, we predicted that genes showing parallel expression should mainly belong to functional groups associated with these phenotypes. Transcriptome profiles were obtained from white muscle by using a 3557 cDNA gene microarray developed for the Atlantic salmon (*Salmo salar*). A total of 1181 genes expressed in both lake populations hybridized on the array. Significant differential expression between ecotypes was detected for 134 (11.3%) and 195 (16.5%) gene clones in Cliff Lake and Indian Pond, respectively. Fifty-one genes (4.3%) showed parallel differential expression between lakes, among which 35 were expressed in opposite directions. Sixteen genes (1.35%) showed true parallelism of transcription, which mainly belonged to energetic metabolism and regulation of muscle contraction functional groups. Variance in expression was significantly reduced for these genes compared to those not showing directionality in parallelism of expression. Candidate genes associated with parallelism in swimming activity and energetic metabolism based on their level and variance in expression were identified. These results add to the growing evidence that parallel phenotypic evolution also involves parallelism at both the genotypic and regulatory level, which may at least partly be associated with genetic constraints. It also provides further evidence for the determinant role of divergent natural selection in driving phenotypic divergence, and perhaps reproductive isolation, in the adaptive radiation of lake whitefish. This study adds to a nascent field employing microarrays as powerful tools for investigating the evolutionary processes of adaptive divergence among natural populations.

Keywords: adaptive radiation, *Coregonus*, gene regulation, microarrays, parallel evolution, speciation

Received 28 May 2005; revision accepted 17 October 2005

Introduction

The independent evolution of adaptively divergent phenotypes among closely related lineages is most likely the consequence of natural selection (Harvey & Pagel 1991). Thus, the study of parallel evolution has contributed importantly towards understanding the role of natural selection in speciation (Harvey & Pagel 1991; Schluter 2000). Parallel evolution has traditionally been documented employing morphology (Huey *et al.* 2000; Rundle & Bronmark 2001) whereas the underlying genetic architecture has been elusive (Remington & Purugganan 2003). Because

selection ultimately acts on the genetic variation underlying trait variation, identifying the genes associated with parallel evolutionary changes among populations is essential to identify candidate genes implicated in adaptive phenotypic divergence. Competing hypotheses have been proposed with respect to parallel genomic evolutionary changes. The first posits that constraints incurred by negative pleiotropic interactions may be considerable, thus the parallel expression of any given trait will likely involve the same genes (Stern 2000). Conversely, in the absence of pleiotropy, the parallel evolution of phenotypes may not necessarily involve the same genes (Schluter *et al.* 2004). While some studies have supported the notion of parallel genotypic evolution (Sucena & Stern 2000; Gompel & Carroll 2003; Colosimo *et al.* 2004; Cresko *et al.* 2004), others have implicated

Correspondence: L. Bernatchez, Fax: 418-656-7176; E-mail: louis.bernatchez@bio.ulaval.ca

differing genomic changes underlying parallel phenotypic evolution (Hoekstra & Nachman 2003; Wittkopp *et al.* 2003). Moreover, these studies involved the analysis of a few candidate genes only, and consequently, tests of these alternate hypotheses have remained inconclusive.

Recent investigations of the simultaneous expression of thousands of genes by means of microarrays revealed that changes in key regulatory genes may have a greater effect than changes in structural genes on phenotypic diversification (Purugganan 1998). Furthermore, variation in gene expression between natural populations has been reported to underlie adaptation (Gibson 2002; Oleksiak *et al.* 2002, 2005; Bochdanovits *et al.* 2003). These studies have cumulatively demonstrated the potential of transcription profiling towards identifying differential gene expression between populations, and thus offer a tremendous opportunity to investigate the genomic basis of parallel, phenotypic divergence under natural environmental conditions. The utility of genome-wide expression studying natural populations allows one to (i) identify the genes implicated in adaptive divergence and (ii) determine the direction of transcriptional divergence using a comparative approach. Moreover, evidence that the same subset of genes reveals parallel, directional changes in expression among independently evolving populations of similar phenotype (up- or down-regulation occurring on the same genes in both allopatric similar phenotypes) would provide strong empirical support that adaptive genomic divergence responds identically to the selective pressures associated with common environmental changes. Moreover, Fisher's theorem predicts that traits under strong selection are expected to display lower variance than traits under weaker selection (Fisher 1930). Since gene expression is a quantitative trait (Cheung & Spielman 2002), it is thus relevant to test whether directional changes show significantly reduced variation compared to nondirectional changes, which would further support the hypothesis that genes subject to directional, parallel changes are under stronger selective and/or genetic constraints. Finally, selection may not only target individual genes, but the whole interacting gene networks, as gene interactions play a central role in development and metabolism (Johnson & Porter 2000; Cork & Purugganan 2004). Therefore, differential regulation of putative candidate genes interacting in the same network may occur between independently evolved, homologous phenotypes.

In this study, we tested the general hypothesis that parallel phenotypic evolution is associated with parallelism in genomic expression by comparing transcription profiles among two isolated, independently evolved population pairs of sympatric 'dwarf' and 'normal' lake whitefish (*Coregonus clupeaformis*, Salmonidae) ecotypes. These reproductively isolated (Lu & Bernatchez 1999) forms diverged 12 000 years ago following the Wisconsinian glaciation,

and demonstrate differential trophic specialization (Lu & Bernatchez 1999; Rogers *et al.* 2002; Campbell & Bernatchez 2004; Rogers & Bernatchez 2005), life history traits (Fenderson 1964), morphology (Fenderson 1964; Bernatchez & Dodson 1990) and behaviour (Bernatchez & Dodson 1990). Thus, they represent populations in the early stages of the speciation process (Lu & Bernatchez 1999). Of particular interest is the observation that ecotypes differ mainly in terms of bioenergetics (Trudel *et al.* 2001) and growth (Rogers & Bernatchez 2005), and that these differences have evolved under divergent selection associated with specialization to limnetic and benthic trophic niches (Rogers & Bernatchez 2005). This study therefore focuses on the white muscle, a tissue highly implicated both in growth and swimming activity (Mommsen 2001; Martinez *et al.* 2004).

Materials and methods

Collection

Ten dwarf and normal sympatric whitefish ecotypes were sampled in June 2003 using gillnets in two separate lakes of the St John River drainage (Maine, USA): Cliff Lake (46°23'51"N, 69°15'05"W) and Indian Pond (46°15'25"N, 69°18'05"W). Gillnets were inspected every 30 min to ensure fish mortality did not occur before tissues were sampled for RNA extractions. Collected white muscle (250–300 mg) was immediately frozen in liquid nitrogen in the field, and stored at –80 °C in the laboratory until RNA extraction. Particular care was given to minimize bias of size difference between contrasted ecotypes because on one hand, metabolic rate is expected to show negative allometry, typical of the mass-specific basal metabolic rate of most animals (Somero & Childress 1980); on the other, glycolytic-enzymes in muscle were observed to increase in larger-size fish (Somero & Childress 1980). In Cliff Lake, dwarf whitefish had a mean fork length of 23.1 cm (SD = 6.15 cm), whereas normals had a mean length of 35.4 cm (SD = 4.9 cm). In Indian Pond, mean length was 22.4 cm (SD = 2.1 cm) for dwarf and 33.0 cm (SD = 5.8 cm) for normal whitefish, respectively. Since metabolic rate has been reported to differ between juvenile and adult fish (Rowan *et al.* 1997), only adult fish were sampled. Because gene expression in white muscle was reported to vary along body axis (Martinez *et al.* 2004), white muscle was systematically sampled from the same body region, immediately below the dorsal fin, for all fish.

Microarrays

Extracted RNA was retrotranscribed following the SuperScript™ II Reverse Transcriptase protocol (Invitrogen Life Technologies). Transcriptome profiles were obtained by using a 3557 cDNA gene microarray developed for the Atlantic salmon (*Salmo salar*) by GRASP (Genomic Research

on Atlantic Salmon Project) (Rise *et al.* 2004), following the Array 50 kit protocol (Genisphere). The transcript levels were quantified using a ScanArray Express scanner (Packard Bioscience). Only genes expressed in both lakes were considered for the data analysis (Tables S1 and S2, Supplementary material).

Functional interpretations of transcriptional patterns should be interpreted cautiously as transcription of a given gene could lead to different protein products (alternative splicing) and ultimately to different functions. Furthermore, cross-hybridization can be a problem for spotted cDNA microarrays because of sequence polymorphisms between strains or paralogous genes that affect the signal for certain genes. However, the 3557 cDNA gene microarray was developed for the Atlantic salmon (*Salmo salar*). Given that the genus *Salmo* and *Coregonus* diverged several million years ago, recently diverged (less than 12 000 bp) whitefish ecotype pairs should be equally affected by cross-hybridization. Also, this study was not conducted in controlled conditions and in principle, differential environmental effects on gene expression should reduce the capacity to detect parallelism of expression between lakes. Yet, we were able to detect highly significant directional parallel changes in expression between ecotypes (see Results). Therefore, the number of genes showing parallelism in levels and directionality of expression should be viewed as conservative.

Statistical analyses

We used a paired design (Churchill 2002) whereby retro-transcribed RNA from an individual of each ecotype of a given lake was randomly chosen to be hybridized. Ten biological replicates were used per analysis. Fluorescence intensity values of duplicated EST (express sequence tags) clone were averaged after correcting for local background and divided the mean intensity for each channel (Draghici 2003). Potential biases in fluorophore intensity variation were minimized by dye swapping (Churchill 2002). We conducted an ANOVA using the R/MAANOVA software (Kerr *et al.* 2000) under a mixed-effect model using a permutation-based *F*-test (*F*₃, with 1000 sample permutations) to detect differentially expressed genes between dwarf and normal ecotypes (Cui & Churchill 2003).

Determination of gene functional groups

Gene loci correspond to EST library annotations using databases from GenBank (see Rise *et al.* 2004 for details). Significantly differentially expressed gene clones were classified into functional groups using the AMIGO browser of Geneontology (www.geneontology.org/), the GO browser of NCBI (www.ncbi.nlm.nih.gov/gquery/gquery.fcgi), the KEGG PATHWAY database (www.genome.jp/kegg/pathway).

html) and completed with references from the literature. Two main functional groups were defined: muscle contraction regulation (MCR) and energetic metabolism (EM), as they are of particular interest given previous observation that dwarf and normal whitefish ecotypes differ mainly in terms of bioenergetics (Trudel *et al.* 2001) and swimming activity (Rogers *et al.* 2002).

P value calculations over observed occurrence of parallel transcriptional changes

For genes showing significant differences in both lakes (L1 and L2), and given some total number of genes *K*, and *Z*₁ or *Z*₂, the number of significant genes in L1 or L2, we computed the *P* value of the observed number of coincidences against the probability of coincidences when assuming stochastic independence between significance in L1 and significance in L2. Using standard notation, the probability function for *X* coincidences can be expressed as:

$$P(X; K, Z_1, Z_2) = \frac{C_X^{Z_1} C_{Z_2-X}^{K-Z_1}}{C_Z^K}$$

However the occurrence of false positives had to be taken into account and so the following corrections were made on parameters *K*, *Z*₁, *Z*₂. Since the significance threshold was set at 0.05 (5%), the observed numbers of significant genes for L1, L2 were corrected to *Z*₁' = *Z*₁ - (0.05 * *K*), *Z*₂' = *Z*₂ - (0.05 * *K*), respectively. The number of coincidences also had to be corrected since false positives may generate false coincidences in three different ways: a pair of false positives, a false positive in L1 paired with a true positive in L2 and, finally, a true positive in L1 paired with a false positive in L2. These three events have respective probabilities: 0.05 * 0.05, 0.05 * (*Z*₂'/*K*) and (*Z*₁'/*K*) * 0.05. Therefore the number of coincidences was corrected to *X*' = *X* - (0.05 * 0.05 + 0.05 * (*Z*₂'/*K*) + (*Z*₁'/*K*) * 0.05) * *K*. The *P* value calculations over coincidences were carried out with the corrected values of the parameters.

Variance in expression profiles

An unpaired one-tailed *t*-test with homoscedasticity (Draghici 2003) was used to test the hypothesis that intrapopulation variance differs between genes showing directional vs. nondirectional parallel changes in expression.

Results

Differential expression associated with phenotypic divergence

A total of 1181 genes expressed in both lake populations hybridized on the 3557 cDNA chip. Significant differential

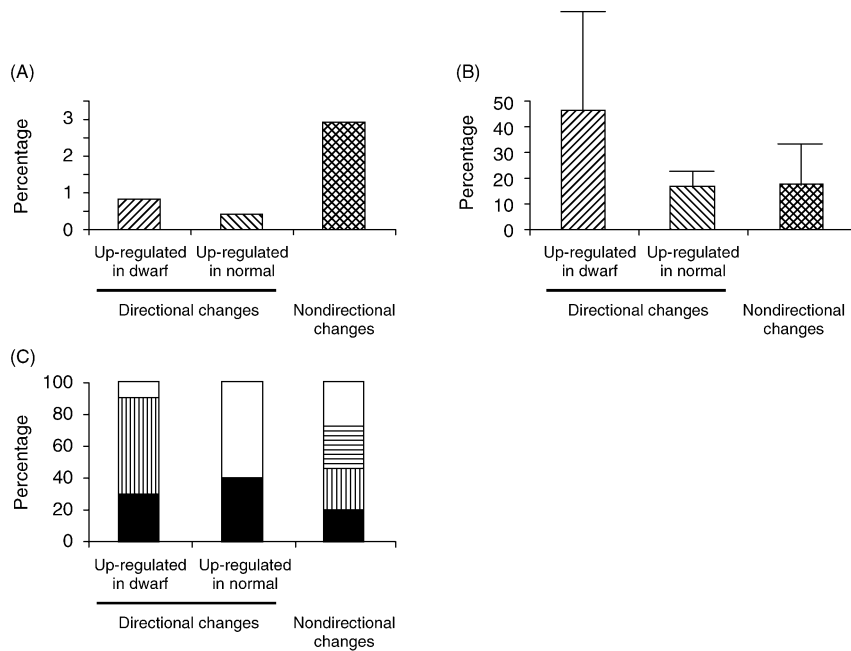


Fig. 1 (A) Percentage of genes that are differentially expressed in parallel between dwarf and normal ecotypes. (B) Mean differential gene expression (%) between dwarf and normal ecotypes. (C) Breakdown (%) of the main functional groups associated with directional and nondirectional parallel differential expression between dwarf and normal ecotypes. Error bars are standard deviations. Directional change refers to overexpression in the same ecotype in the two lake populations, whereas nondirectional change refers to overexpression in different ecotypes between the two lakes. Black, energetic metabolism; vertically striped, muscle contraction regulation; horizontally striped, protein synthesis; and white, other functional groups.

expression between dwarf and normal ecotypes was detected for 134 (11.3%) and 195 (16.5%) gene clones in Cliff Lake and Indian Pond, respectively. Of these, 51 gene clones (4.3%) showed parallel differential expression between lakes (Table 1). Such a level of parallelism is highly unlikely by chance alone ($P = 0.166 \cdot 10^{-17}$, see methods). However, 35 of the common genes (3% of 1181) were differentially expressed in opposite directions (nondirectional) between the two lake populations (Fig. 1A). On average, the magnitude of differential expression was 21% (SD = 0.23) higher for the 15 nondirectional gene clones up-regulated in the dwarf ecotypes within Cliff Lake and 16% (SD = 0.08) higher for the 20 gene clones up-regulated in the dwarf ecotypes of Indian Pond. The remaining 16 gene clones (1.35% of 1181) showed parallel pattern of transcription in both lake populations (Table 1). Among these, 10 genes [EST clone 'nwh⁷14[P]' which assigned to gamma-crystallin gene was spotted in two separate duplicate pairs, see Table 1 and Rise *et al.* (2004)] were overexpressed in the dwarf ecotype, while five were overexpressed in the normal ecotype (Fig. 1A). The level of differential expression between dwarf and normal ecotypes was different when considering these genes: on average, the level of up-regulation was 47% (SD = 0.39) higher for the 10 genes up-regulated in the dwarf ecotype, but only 17% (SD = 0.07) for the five up-regulated genes in the normal ecotype (Fig. 1B). Moreover, the variance in expression for directional parallel genes (mean = 2.42) was significantly lower than that observed for nondirectional parallel genes (mean = 22.08) ($P = 0.006$, one-tailed *t*-test). Finally, the number of genes that showed lake-specific differential expression between forms was significantly higher ($\chi^2 = 9.93$, $P = 0.002$) in Indian Pond

($n = 144$) than in Cliff Lake ($n = 83$), the latter being marginally significantly different from the number of the 59 expected false positives ($\chi^2 = 0.383$, $P = 0.05$) (table of all significant genes available upon request to the authors).

Functional groups: directional vs. nondirectional parallelism in expression

Among genes showing directional parallel changes (Table 1), up-regulation was twice more common in the dwarf ecotype than in the normal, and mainly involved both energetic metabolism (3 distinct EST clones out of 10) and muscle contraction regulation genes (six distinct EST clones out of 10). Gamma-crystallin was up-regulated by 109% in Cliff Lake and 54% in Indian Lake; ATP synthase subunit c was up-regulated by 10.4% in Cliff Lake and 17% in Indian Pond. Up-regulation for six parvalbumin variants varied between 14% and 135%. Up-regulated genes in the normal ecotype concerned energetic metabolism, iron binding and tissular differentiation. Nondirectional parallel changes mainly involved functional groups associated with glycolysis/ATP synthesis (20%), muscle contraction regulation (26%) and protein synthesis (26%) (Table 2, Fig. 1C).

Population-specific differences of expression

Since the significance threshold was set at 0.05 (5%), the observed numbers of significant genes for lake-specific changes were 83 and 144 for Cliff Lake and Indian Pond, respectively (tables available upon request to the authors). The main functional groups involved in both lake were protein synthesis (30.1% for Cliff Lake, 36.8% for Indian

Table 1 Parallel directional changes in gene expression between dwarf and normal ecotypes

Function*	EST clon†	Gene‡	Fold change§			
			Cliff Lake	Indian Pond	P-CL¶	P-IP¶
MCR	Nwh ²⁰ 43 ^[P]	= (AF538283) parvalbumin beta	1.22	1.14	0.0010	0.0395
	Nwh ¹⁰ 49 ^[P]	= (AF538283) parvalbumin beta	1.28	1.15	0.0004	0.0223
	Nwh ¹⁸ 44 ^[P]	= (AF538283) parvalbumin beta	1.20	1.24	0.0017	0.0155
	Nwh ¹¹ 85 ^[P]	= (AF538283) parvalbumin beta	1.42	1.37	0.0005	0.0090
	Nwh ⁶ 27 [^]	= (AF538283) parvalbumin beta	2.35	1.55	0.0000	0.0018
EM	Nwh ¹¹ 21 ^[P]	= (AF538283) parvalbumin beta	2.13	1.82	0.0000	0.0007
	Nwh ⁷ 14 ^[P]	[JC2355] gamma-crystallin††	2.16	1.59	0.0000	0.0004
	Nwh ⁷ 14 ^[P]	[JC2355] gamma-crystallin	2.02	1.49	0.0000	0.0007
	rbha ³ 93 ^[P]	[AAG16310] enolase-1	1.18	1.01	0.0121	0.0291
	Rgb ⁵¹⁵ 50 ^[P]	[P07926] ATP synthase (c)	1.10	1.17	0.0247	0.0343
OF	Nwh ⁹ 90 [^]	unknown	1.36	1.33	0.0002	0.0219
IB	Rgb ⁵²⁶ 361 ^[P]	[AAB34576] ferritin M	0.86	0.70	0.0049	0.0012
EM	Rgb ⁵⁰⁷ 246 ^[N]	[BI468053] = triose P isomerase	0.81	0.86	0.0063	0.0205
TD	Nwh ¹³ 82 ^[P]	[AAB82747] GAPDH	0.91	0.87	0.0272	0.0276
	Rblb ³ 48 ^[P]	[AAC99813] SPARC	0.90	0.90	0.0231	0.0204
	rgb ⁵²³ 219 ^[P]	[AAC99813] SPARC	0.90	0.91	0.0246	0.0381

*Functional classes are defined using AMIGO browser of Geneontology (www.geneontology.org/), the go browser of NCBI (www.ncbi.nlm.nih.gov/gquery/gquery.fcgi), the KEGG PATHWAY database (www.genome.jp/kegg/pathway.html) and completed with references from the literature. EM, energetic metabolism; IB, iron binding; MCR, muscle contraction regulation; OF, other functions; TD, tissular differentiation.

†Each EST clone number corresponds to a single EST sequence.

‡Gene loci correspond to gene names assignation of each library's EST using databases from GenBank see Rise *et al.* (2004) for details.

‘=’ means unknown EST sequences resubmitted in this study to GenBank databases. Locus accession number are in brackets.

§Fold change correspond to a ratio of mean expression values of dwarf to normal ecotypes: > 1 implies up-regulated in dwarf; < 1 implies up-regulated in normal.

¶P-CL and P-IP correspond to permuted P values for Cliff Lake and Indian Pond, respectively (ANOVA, F3 test with 1000 permutations).

††EST clone ‘nwh⁷14^[P]’ which assigned to gamma-crystallin gene was spotted in two separate duplicate pairs. See Rise *et al.* (2004).

Pond), energetic metabolism (16.9% for Cliff Lake, 9.7% for Indian Pond), and muscle contraction (8.4% for Cliff Lake, 7.6% for Indian Pond). The main difference between two lakes was overexpression of several independent EST clones of Ferritin H gene in Indian Pond normal phenotype. In Cliff Lake, several EST clones matching to the same gene were overexpressed in dwarf ecotype: four other parvalbumin variants, two cytochrome c oxidase polypeptide VIa precursor (lna⁸10[^] and plnb⁵⁰⁴159[^]) and two cytochrome c oxidase subunit II (shc⁵⁰³299[^] and pitl⁵⁰⁴173[^]).

Discussion

Our main objective was to test the general hypothesis that phenotypic parallelism between dwarf and normal whitefish ecotypes is accompanied by parallelism in gene expression, which would further support the role of divergent selection in driving population divergence and ultimately reproductive isolation between whitefish ecotypes. Given that the most striking phenotypic differences previously reported between these forms implied energetic metabolism and

swimming activity (Trudel *et al.* 2001; Rogers *et al.* 2002), we predicted that genes showing the clearest evidence of parallel expression should belong to functional groups associated with these phenotypes. Genes showing evidence of parallelism are more likely to be under directional selective constraints, which should in turn result in reduced genetic variance. Therefore, such genes should also exhibit lower variance in expression. Previous studies also showed that both metabolism (Trudel *et al.* 2001) and swimming activity (Rogers *et al.* 2002) are higher in dwarf relative to normal phenotypes, genes associated with these traits should be up-regulated in the dwarf whitefish. The results obtained supported these predictions since we observed directional parallelism for 16 genes (1.35% of the 1181 genes expressed in both lakes) and these mainly belong to energetic metabolism and regulation of muscle contraction functional groups (Fig. 1C). Moreover, variance in expression was significantly reduced for these genes compared to those not showing directionality in parallelism of expression. Finally, we observed a 47% increase of mean expression for these genes in dwarf relative to normal ecotypes, which contrasted with a mean difference of 18%

Table 2 Parallel nondirectional changes in gene expression between dwarf and normal ecotypes

Function*	EST clon†	Gene‡	Fold change§				
			Cliff Lake	Indian Pond	P-CL¶	P-IP¶	
MCR	nwh^1^27^ [P]	= (SSPRVB1) parvalbumin beta	1.20	0.86	0.0004	0.0196	
	nwh^19^62^ [P]	= (SSPRVB1) parvalbumin beta	1.19	0.83	0.0027	0.0091	
	nwh^11^86^ [P]	= (SSPRVB1) parvalbumin beta	1.19	0.82	0.0047	0.0050	
	nwh^10^26^ [P]	= (SSPRVB1) parvalbumin beta	1.15	0.87	0.0071	0.0139	
	nwh^20^13^ [P]	= (SSPRVB2) parvalbumin beta	1.28	0.42	0.0019	0.0000	
	nwh^8^70^ [N]	[AF330142] actin	1.11	0.76	0.0156	0.0112	
	rb1b^3^3^ [N]	[L25609] tropomyosin	1.11	0.88	0.0259	0.0373	
	rbhb^2^80^ [P]	[NP_003794] myomesin	0.86	1.15	0.0096	0.0214	
	nwh^11^82^ [P]	[AAC96094] creatine kinase	0.97	1.24	0.0257	0.0054	
	EM	ela^2^71^ [P]	[AAF61384] cytochrome <i>c</i> oxidase (III)	1.15	0.90	0.0302	0.0210
rgb^525^81^ [P]		[AAF61384] cytochrome <i>c</i> oxidase (III)	1.20	0.78	0.0040	0.0010	
lna^7^82^ [P]		[T09959] cytochrome <i>c</i> reductase	1.09	0.81	0.0250	0.0065	
lna^13^27^ [N]		[AY024367] glycerol ₃ PDH	1.08	0.86	0.0445	0.0261	
plnb^504^321^ [P]		[BAA88482] alpha 1 enolase	0.81	1.21	0.0013	0.0061	
nwh^19^56^ [P]		[NP_036627] aldolase	0.84	1.17	0.0028	0.0088	
lna^2^51^ [N]		[AF360980] GAPDH	0.85	1.20	0.0044	0.0083	
PS		nwh^6^94^ [P]	[JC1235] transcription factor BTF3a	0.87	1.12	0.0116	0.0415
	rgb^526^260^ [P]	[AAK09383] ribosomal protein S3a	0.88	1.25	0.0131	0.0080	
	pha^501^321^ [P]	[P27952] ribosomal protein S2	0.86	1.14	0.0142	0.0073	
	rgb^514^102^ [P]	[NP_065625] ribosomal protein S14	0.90	1.09	0.0342	0.0338	
	rgb^523^34^ [P]	[AAD26692] ribosomal protein S13	0.92	1.14	0.0349	0.0418	
	rgb^534^122^ [P]	[XP_049354] = ribosomal protein L9	0.91	1.11	0.0221	0.0354	
	lna^17^71^ [P]	[1909362 A] ribosomal protein L27	0.94	1.28	0.0430	0.0051	
	rgb^522^123^ [N]	[BG934793] = ribosomal protein L37A	0.84	1.23	0.0064	0.0078	
	rgb^531^17^ [P]	[BAB27309] = ribosomal protein L13	0.87	1.16	0.0184	0.0121	
	lna^13^72^ [N]	[BG935658] = (AF503957) ribosomal protein	0.90	1.39	0.0202	0.0102	
	OF	rbha^4^69^ [N]	[BI468155] = (AF452171) lysozyme type II	0.95	1.07	0.0210	0.0489
		rgb^510^358^ [P]	[NP_035159] laminin receptor	0.91	1.14	0.0154	0.0273
plnb^506^166^ [P]		[BAA88568] ubiquitin	0.87	1.26	0.0157	0.0037	
lna^16^10^ [P]		[BAA88568] ubiquitin	0.92	1.33	0.0396	0.0036	
nwh^3^95^ [P]		[AAG12164] ornithine aminotransferase	1.32	0.80	0.0008	0.0282	
rb1b^2^45^ [P]		[AAG30028] prostaglandin D synthase	1.14	0.83	0.0115	0.0296	
rbhb^3^59^ [P]		[AAF73200] methyltransferase	1.06	0.98	0.0167	0.0433	
plnb^503^306^ [P]		[CAC19682] NTT4	1.11	0.97	0.0324	0.0427	
rb1b^2^22^ [P]		[BAA91208] unknown protein	0.92	1.12	0.0410	0.0321	

* , †, ‡, § and ¶: see Table 1.

in levels of expression for other functional groups, or genes not showing directional parallelism of expression.

We also identified potential candidate genes associated with parallelism in swimming activity and energetic metabolism based on their level and variance in expression. One such candidate is the gamma-crystallin gene. The beta/gamma-crystallin superfamily has evolved from pre-existing proteins, and is closely related and highly homologous to lactate dehydrogenase (Jones *et al.* 1999). Moreover, it has been reported that vertebrate enzyme-crystallins have retained enzymatic activities (Piatigorsky 2003). Therefore, it appears more plausible that transcripts that hybridized to the gamma-crystallin EST were lactate dehydrogenase transcripts rather than genes with lens crystallin functions

expressed in muscles. Lactate dehydrogenase plays an important role in glycolysis and swimming activity, and is correlated with glycolytic capacity such that a more active use of fast muscle fibres is expected to translate into higher expression of glycolytic enzymes (Martinez *et al.* 2004). Indeed, higher activity of lactate dehydrogenase was reported in the cisco (*Coregonus artedii*) relative to normal lake whitefish (Guderley *et al.* 1986), the former being a pelagic whitefish species characterized by high swimming activity (Bernatchez & Dodson 1985). Similarly, over-expression of proteins which are highly homologous to gamma-crystallins could also be associated with a genetically determined, increased swimming activity in dwarf whitefish (Rogers *et al.* 2002). A second potential candidate

gene was the mitochondrial ATP synthase, a key enzyme in oxidative phosphorylation, which produces energy for the cell. The *c* subunits form oligomeric rings in the membrane domain of ATP synthase, and are believed to play a central role in providing the energy for the synthesis of ATP (Medeiros & Jennings 2002). Thus, higher expression of ATP synthase observed in both dwarf populations is consistent with the more elevated energy demands associated with increased swimming activity in this ecotype.

The parvalbumin genes showed the most pronounced pattern of differential expression as a group, with six variants being overexpressed in both dwarf populations, and exhibiting significantly reduced variance in expression (mean = 1.15) relative to nondirectional parvalbumin variants (mean = 30.26) ($P = 9 \cdot 10^{-5}$, unpaired one-tailed *t*-test) (Fig. 3). In addition, a resubmission of the 11 distinct parvalbumin EST clones to sequence databases showed that the six variants overexpressed in both dwarf populations matched to one parvalbumin gene (AF538283) with 87–98% of sequence homology. It thus appears likely that the six parvalbumin transcripts that were overexpressed in the dwarf ecotype and hybridized to the six parvalbumin EST clones may belong to the same gene region. Among the five nondirectional parvalbumin variants, four showed little sequence homology with (AF538283), but rather matched to (SSPRVB1) with 98–99% of sequence homology. Although one-third of the sequence of the other nondirectional parvalbumin variant aligned with (AF538283), its entire sequence length matched to (SSPRVB2) with 98% of sequence homology. In addition, directional and nondirectional parvalbumin EST sequences form two clusters (directional vs. nondirectional). These clusters do not align over the first 400 sites, but do align over the last 193 sites, where they are differentiated by 25 fixed substitutions (data not shown). Together, it appears very likely that four of five nondirectional parvalbumin transcripts would belong to the same gene, whereas the fifth would belong to another parvalbumin variant, all of them being distinct from the directional parvalbumin (AF538283) locus. Parvalbumins are intracellular calcium-binding proteins that play a major role in the regulation of muscle contraction regulation (Rall 1996). Namely, they accelerate muscle relaxation between brief contractions by calcium sequestration, such that a higher concentration of parvalbumins is expected to translate into more muscle contractions per unit of time (Rall 1996). Moreover, these genes have been shown to increase the hyperexcitability of fast muscle in mutant mice (Schleef *et al.* 1994). Thus, parallelism in overexpression of one parvalbumin variant is also congruent with the higher swimming activity observed in dwarf whitefish.

Altogether, the above results support the hypothesis that parallel phenotypic evolution in energetic metabolism and swimming activity at least partly involves directional

changes in expression at the same genes. This, in turn, brings further support to the hypothesis that directional selection has played a determinant role in the rapid adaptive radiation of whitefish in recent evolutionary times. Parallel evolution and reduced variance at the genome level may also partly result from the same genetic constraints being shared by closely related species or populations (Haldane 1932). Genetic constraints depend both on gene position in metabolic networks, as well as network functions themselves (Johnson & Porter 2000; Cork & Purugganan 2004). For instance, genes acting at branch points of metabolic pathways are potentially more constrained because of their more important impact on pleiotropic interactions (Cork & Purugganan 2004). Conversely, more external genes are expected to be less constrained since they have one or few functions, and similar genes are often available for doing the same task (Cork & Purugganan 2004). Thus, under the hypothesis that genes showing directional parallelism are the most likely targets of selection, the genetic constraint hypothesis would predict that such genes should also be more likely to play a central role in their metabolic network. Our results partly support this hypothesis. For instance, parvalbumins are at the basis of gene interactions controlling muscle contraction regulation (Rall 1996). Similarly, the putative LDH-like crystalline enzyme that hybridized on the gamma-crystallin EST would have a central role in muscle energetics, as LDH activity was reported to be an indicator for glycolytic ATP needs in fast-twitch (i.e. white) muscle (Odell *et al.* 2003).

An alternative, nonexclusive hypothesis is that differences in expression between more vs. less constrained genes could also reflect the sequential chronology of adaptive genomic response. Orr (2000) hypothesized that large effect mutations are selected first, and because they often have negative pleiotropic interactions, small effect compensatory mutations are recruited secondarily in order to canalize adaptive response (Gibson & Wagner 2000). Under this framework, directional parallel genes, also showing the most pronounced pattern of differential expression, may have been recruited first during the process of population divergence, later accompanied by recruitment of different (i.e. those showing no directional parallelism) genes to compensate for possible deleterious effects of major ones. Thus, transcriptional regulation changes would have occurred first on genes of larger effects such as parvalbumin beta, ATP-synthase and LDH-like crystalline enzyme, followed by compensatory mutations modulating expression of genes in the same interacting networks (e.g. actin, tropomyosin, myomesin and creatine kinase in muscle contraction; glycolytic enzymes such as aldolase, triose phosphate isomerase, GAPDH, enolase, as well as cytochrome *c* oxidases in the ATP synthesis network). Further indication for the role of genetic constraints on the association between gene expression and phenotypic divergence

comes from the results obtained for the 35 genes that showed significant changes in expression in both lakes but in opposite directions. It is plausible that genetic bias resulting in parallelism of metabolic networks associated with differential phenotypic expression does not necessarily involve similar changes in expression of the same genes, but rather different combinations of up- and down-regulations that may result in similar phenotypic expression. Thus, parallel, nonrandom, transcriptional changes were observed in both lake populations, suggesting they too may have played an important, functional role in the phenotypic divergence of lake whitefish ecotypes. The three main functional groups of these nondirectional parallel genes were glycolysis/ATP synthesis, muscle contraction regulation and protein synthesis. Interestingly, these three functional groups presented an amount of differential expression, although opposed, that was not significantly different between the two lake population pairs (glycolysis/ATP synthesis: $P = 0.48$; muscle contraction regulation: $P = 0.26$; and protein synthesis: $P = 0.17$). This raises the hypothesis that different balances between up- and down-regulated genes in a given pathway could potentially lead to the same large-scale effect (Johnson & Porter 2000). The most striking case concerns glycolysis/ATP synthesis genes (Fig. 2), wherein nondirectional, parallel changes typically belong to the two stages of the glucose catabolism pathway: in one stage, aldolase, glyceraldehyde 3-phosphate dehydrogenase and enolase, operate in the glycolysis pathway leading to the formation of pyruvate (stage 1); in the other, cytochrome *c* oxidases, cytochrome *c* reductase and glycerol 3 phosphate dehydrogenase are involved in the tricarboxylic acid cycle, which occurs within the mitochondria (stage 2). Thus, when stage 1 genes are up-regulated, stage 2 genes are down-regulated. Conversely, when stage 2 genes are up-regulated, genes from stage 1 are down-regulated. Therefore, these two different settings of gene regulation could potentially lead to comparable amounts of ATP production in different populations of the same ecotype. Similarly, such a phenomenon might also have occurred in muscle contraction regulation, wherein parvalbumins, actin, myomesin, tropomyosins and creatine kinase interact in the same network (Fig. 3). It is also noteworthy that two different EST clones of both glyceraldehyde 3-phosphate dehydrogenase and enolase genes showed dissimilar results. This is not surprising as there is a nearly 50% sequence divergence between glyceraldehyde 3-phosphate dehydrogenase ESTs, and that enolase ESTs differed by two insertions/deletions of 3 and 35 nucleotides, respectively. This raises the hypothesis that the different EST of these two genes may have distinct functions. Altogether, these observations corroborate the hypothesis that different regulation settings in a given pathway or interaction network could lead to a similar expression of physiological phenotypes (Johnson & Porter 2000).

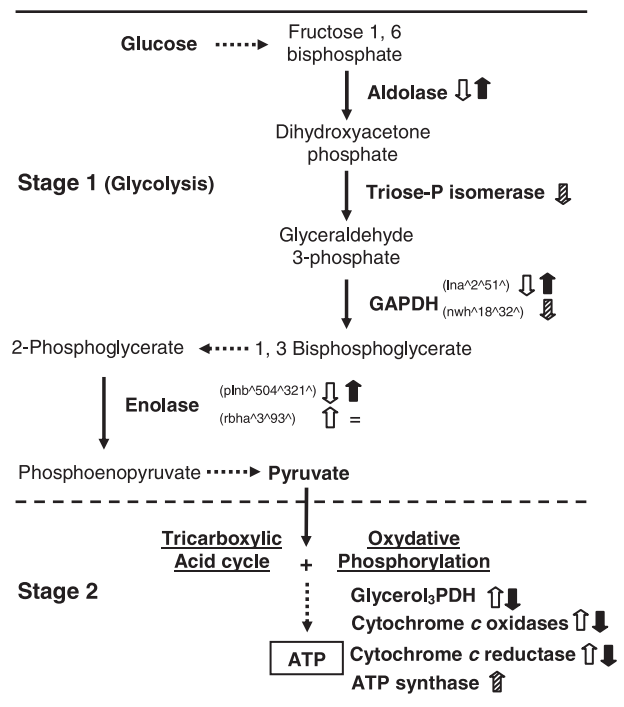


Fig. 2 Differential gene expression in energetic metabolism network (simplified). Arrows represent gene regulation changes: diagonally striped for directional parallel changes, white for Cliff Lake population (left) and black for Indian Pond population (right). '=' means no differential expression. EST clones are indicated in parentheses when different (see Tables 1 and 2).

Differences of transcription regulation between dwarf and normal could potentially be partly affected by size differences. For instance, glycolytic enzyme activities in white muscle are positively correlated with body size (Somero & Childress 1990). However, no systematic bias in levels of expression between the two forms was detected. It thus seems unlikely that allometry significantly affected differences of transcriptional patterns observed.

To date, very few studies have investigated parallelism in gene expression in the context of adaptive radiation or other rapid evolutionary processes in natural environments (but see Cooper *et al.* 2003; Zhong *et al.* 2004). In a study on *Escherichia coli* performed in a controlled experiment, the proportion of directional, parallel changes observed in the Cooper *et al.* (2003) study was very similar to that we observed in this study (1.40% vs. 1.35%). Our results are also congruent with previous studies that addressed patterns of parallel evolution at the genome level in whitefish. For instance, Campbell & Bernatchez (2004) compared the pattern of genetic differentiation obtained using 440 amplified fragment length polymorphism (AFLP) loci with that expected under neutrality in four sympatric pairs of lake whitefish ecotypes. Among loci that were most likely under directional selection, only six (1.3%) exhibited

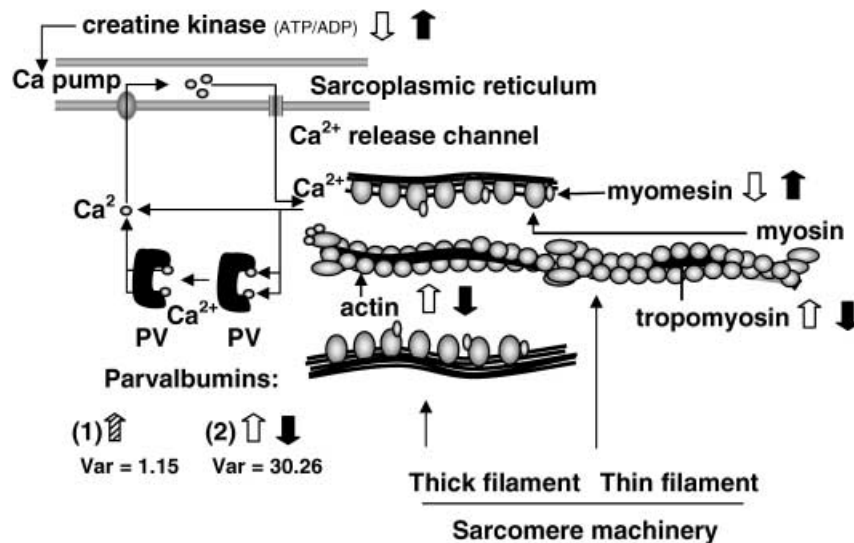


Fig. 3 Differential gene expression in muscle contraction regulation network. During relaxation phase, Ca^{2+} is bound to parvalbumin (PV). Ca^{2+} returns to the sarcoplasmic reticulum (SR) via a Ca^{2+} pump activated by ATP modulated by creatine kinase. This establishes the next contraction phase wherein Ca^{2+} is released from the SR, thereby removing contraction inhibition of tropomyosin. This allows interaction between actin and myosin, leading to muscle contraction. Arrows represent gene regulation changes: diagonally striped for directional parallel changes, white for Cliff Lake population (left) and black for Indian Pond population (right). EST clones are indicated in parentheses when different (see Tables 1 and 2). (1): nwh^{20^43}[P], nwh^{10^49}[P], nwh^{18^44}[P], nwh^{11^85}[P], nwh^{6^27} matched to (AF538283) parvalbumin beta variant, up-regulated in both dwarf populations. (2): nwh^{1^27}[P], nwh^{19^62}[P], nwh^{11^86}[P] and nwh^{10^26}[P] matched to (SSPRVB1) parvalbumin beta variant, and nwh^{20^13}[P] matched to (SSPRVB2) parvalbumin beta variant, both of them are up-regulated in Cliff Lake dwarf population and down-regulated in Indian Pond dwarf population. 'Var' indicates mean variance of gene expression among dwarf and normal ecotypes in both lake populations.

parallel patterns of divergence, indicating that only a small proportion of scored AFLP loci might be linked to genes implicated in the adaptive divergence of lake whitefish. More recently, Rogers & Bernatchez (2005) integrated adaptive quantitative trait loci (QTL) mapping and genome scans among the same four sympatric pairs to test the hypothesis that differentiation between dwarf and normal ecotypes at growth-associated QTLs was maintained by directional selection. Among the 27 growth-related QTLs that were analysed, only two exhibited parallel reductions of gene flow. Altogether, these studies are consistent with the hypothesis that adaptive phenotypic divergence between closely related populations is accompanied by relatively few genetic changes, either in the form of structural or regulatory genes (Purugganan 1998).

To conclude, this study adds to the growing evidence that parallel phenotypic evolution also involves parallelism at both the genotypic and regulatory level, which may at least partly be associated with genetic constraints as envisioned by Haldane (1932). It also provides further evidence for the determinant role of divergent natural selection in driving phenotypic divergence, and perhaps reproductive isolation in the adaptive radiation of lake whitefish. Finally, this study adds to a nascent field employing microarrays as powerful tools to investigate the evolutionary processes of adaptive divergence among natural populations.

Acknowledgements

We are grateful to D. Basley, Maine Department of Inland Fisheries and Wildlife, D. Campbell and R. St Laurent for their invaluable help in fish sampling, as well as to Christian Roberge for microarray techniques training. Thank you also to S. Rogers, S. McCairns, W. Cresko and H. Guderley for many valuable discussions and comments on earlier versions of the manuscript. We also acknowledge the GRASP project, B.F. Koop and W.S. Davidson. This research was financially supported by the Natural Sciences and Engineering research Council of Canada to L.B. (EWR Steacie supplement, grant no. 25610-2002). The supporting agency had no role in planning or any other aspect of this study.

Supplementary material

The supplementary material is available from <http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC2804/MEC2804sm.htm>

Table S1 and Table S2

Column A: 'ID' corresponds to the spot number on the GRASP 3557 gene chip (see Rise *et al.* (2004) for details).

Column B: 'Name' includes EST clone number (Each EST clone number corresponds to a single EST sequence), Genbank accession number of the best blast query, and its gene name (see Rise *et al.* (2004) for details).

Column C: 'Pvalperm' corresponds to permuted *P*-values for Cliff Lake and Indian Pond respectively (ANOVA, *F*3 test with 1000 permutations).

Column D: 'Fold change' corresponds to a ratio of mean expression values of dwarf to normal ecotypes: > 1 implies up-regulated in dwarf; < 1 implies up-regulated in normal.

References

- Bernatchez L, Dodson JJ (1985) Influence of temperature and current speed on the swimming capacity of lake whitefish (*Coregonus clupeaformis*) and cisco (*C. artedii*). *Canadian Journal of Fishes and Aquatic Sciences*, **42**, 1522–1529.
- Bernatchez L, Dodson JJ (1990) Allopatric origin of sympatric populations of Lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. *Evolution*, **44**, 1263–1271.
- Bochdanovits Z, van der Klis H, de Jong G (2003) Covariation of larval gene expression and adult body size in natural populations of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **20**, 1760–1766.
- Campbell D, Bernatchez L (2004) Genomic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Molecular Biology and Evolution*, **21**, 945–956.
- Cheung VG, Spielman RS (2002) The genetics of variation in gene expression. *Nature Genetics*, **32**, 522–525.
- Churchill GA (2002) Fundamentals of experimental design for cDNA microarrays. *Nature Genetics*, **32**, 490–495.
- Colosimo PF, Peichel CL, Nereng K *et al.* (2004) The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology*, **2**, e109.
- Cooper TF, Rozen DE, Lenski RE (2003) Parallel changes in gene expression after 20000 generations of evolution in *Escherichia coli*. *Proceedings of the National Academy of Sciences, USA*, **100**, 1072–1077.
- Cork JM, Purugganan MD (2004) The evolution of molecular genetic pathways and networks. *BioEssays*, **26**, 479–484.
- Cresko WA, Amores A, Wilson C *et al.* (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences, USA*, **101**, 6050–6055.
- Cui X, Churchill GA (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology*, **4**, 210.
- Draghici S (2003) *Data Analysis Tools for DNA Microarrays*. Chapman & Hall, London.
- Fenderson OC (1964) Evidence of subpopulations of lake whitefish, *Coregonus clupeaformis*, involving a dwarfed form. *Transactions of the American Fisheries Society*, **93**, 77–94.
- Fisher RA (1930) *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
- Gibson G (2002) Microarrays in ecology and evolution: a preview. *Molecular Ecology*, **11**, 17–24.
- Gibson G, Wagner G (2000) Canalization in evolutionary genetics: a stabilizing theory? *BioEssays*, **22**, 372–380.
- Gompel N, Carroll SB (2003) Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. *Nature*, **424**, 931–935.
- Guderley H, Blier P, Richard L (1986) Metabolic changes during the reproductive migration of two sympatric coregonines, *Coregonus artedii* and *Coregonus clupeaformis*. *Canadian Journal of Fishes and Aquatic Sciences*, **43**, 1859–1865.
- Haldane JBS (1932) The time of action of genes, and its bearing on some evolutionary problems. *American Midland Naturalist*, **66**, 5–24.
- Harvey PH, Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.
- Hoekstra HE, Nachman MW (2003) Different genes underlie adaptive melanism in different populations of rock pocket mice. *Molecular Ecology*, **12**, 1185–1194.
- Huey RB, Gilchrist GW, Carlson ML, Berrigan D, Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science*, **287**, 308–309.
- Johnson NA, Porter AH (2000) Rapid speciation via parallel, directional selection on regulatory genetic pathways. *Journal of Theoretical Biology*, **205**, 527–542.
- Jones SE, Jomary C, Grist J, Makwana J, Neal MJ (1999) Retinal expression of gamma-crystallins in the mouse. *Investigative Ophthalmology and Visual Sciences*, **40**, 3017–3020.
- Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. *Journal of Computational Biology*, **7**, 819–837.
- Lu GQ, 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.
- Martinez M, Bedard M, Dutil JD, Guderley H (2004) Does condition of Atlantic cod (*Gadus morhua*) have a greater impact upon swimming performance at U-crit or sprint speeds? *Journal of Experimental Biology*, **207**, 2979–2990.
- Medeiros DM, Jennings D (2002) Role of copper in mitochondrial biogenesis via interaction with ATP synthase and cytochrome c oxidase. *Journal of Bioenergetics and Biomembranes*, **34**, 389–395.
- Mommsen TP (2001) Paradigms of growth in fish. *Comparative Biochemistry and Physiology Part B*, **129**, 207–219.
- Odell JP, Chappell MA, Dickson KA (2003) Morphological and enzymatic correlates of aerobic and burst performance in different populations of Trinidadian guppies *Poecilia reticulata*. *Journal of Experimental Biology*, **206**, 3707–3718.
- Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. *Nature Genetics*, **32**, 261–266.
- Oleksiak MF, Roach JL, Crawford DL (2005) Natural variation in cardiac metabolism and gene expression in *Fundulus heteroclitus*. *Nature Genetics*, **37**, 67–72.
- Orr HA (2000) Adaptation and the cost of complexity. *Evolution*, **54**, 13–20.
- Piatigorsky J (2003) Crystallin genes: specialization by changes in gene regulation may precede gene duplication. *Journal of Structural and Functional Genomics*, **3**, 131–137.
- Purugganan M (1998) The molecular evolution of development. *BioEssays*, **20**, 700–711.
- Rall JA (1996) Role of parvalbumin in skeletal muscle relaxation. *News in Physiological Sciences*, **11**, 249–255.
- Remington DL, Purugganan MD (2003) Candidate genes, quantitative trait loci, and functional trait evolution in plants. *International Journal of Plant Sciences*, **164**, S7–S20.
- Rise ML, von Schalburg KR, Brown GD *et al.* (2004) Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. *Genome Research*, **14**, 478–490.
- 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*). *Molecular Ecology*, **14**, 351–361.
- 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.
- Rowan DJ, Ramussen JB, Chant L (1997) A bioenergetics approach to modeling seasonal patterns in the bioaccumulation of radiocesium. In: *Freshwater and Estuarine Radioecology*. Elsevier Science Publishers, Amsterdam.
- Rundle SD, Bronmark C (2001) Inter- and intraspecific trait compensation of defence mechanisms in freshwater snails. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 1463–1468.
- Schleef M, Zuhlke C, Schoffl F, Jockusch H (1994) Subtractive cDNA cloning as a tool to analyse secondary effects of a muscle disease. Characterization of affected genes in the myotonic ADR mouse. *Neuromuscular Disorders*, **4**, 205–217.
- Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, New York.
- Schluter D, Clifford EA, Nemethy M, McKinnon JS (2004) Parallel evolution and inheritance of quantitative traits. *American Naturalist*, **163**, 809–822.
- Somero GN, Childress JJ (1980) A violation of the metabolism-size scaling paradigm — activities of glycolytic-enzymes in muscle increase in larger-size fish. *Physiological Zoology*, **53**, 322–337.
- Somero GN, Childress JJ (1990) Scaling of ATP-supplying enzymes, myofibrillar proteins and buffering capacity in fish muscle. *Journal of Experimental Biology*, **149**, 319–333.
- Stern D (2000) Perspective: evolutionary developmental biology and the problem of variation. *Evolution*, **54**, 1079–1091.
- Sucena E, Stern DL (2000) Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven baby*. *Proceedings of the National Academy of Sciences, USA*, **97**, 4530–4534.
- Trudel M, Tremblay A, Schetagne R, Rasmussen JB (2001) Why are dwarf fish so small? An energetic analysis of polymorphism in lake whitefish (*Coregonus clupeaformis*). *Canadian Journal of Fishes and Aquatic Sciences*, **58**, 394–405.
- Wittkopp PJ, Carroll SB, Kopp A (2003) Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends in Genetics*, **19**, 495–504.
- Zhong SB, Khodursky A, Oykhuzen DE, Dean AM (2004) Evolutionary genomics of ecological specialization. *Proceedings of the National Academy of Sciences, USA*, **101**, 11719–11724.

This study is in continuation with our long-term research program on the comprehension of genotype-environment interactions involved in the adaptive radiation of the lake whitefish. Nicolas Derome is a postdoctoral fellow interested in the molecular response of genotype-environment interactions that lead to adaptive divergence. Louis Bernatchez's research focuses on understanding the patterns and processes of molecular and organismal evolution. Pierre Duchesne is a mathematician and member of Bernatchez's laboratory.
