Molecular Ecology (2006) **15**, 9–20 doi: 10.1111/j.1365-294X.2005.02807.x

FAST TRACK

Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon

CHRISTIAN ROBERGE,* SIGURD EINUM,† HELGA GUDERLEY* and LOUIS BERNATCHEZ*
**Département de biologie, Université Laval, Québec, Canada G1K 7P4 and †Norwegian Institute for Nature Research, Trondheim, Norway*

Abstract

Farmed salmon strains have been selected to improve growth rates as well as other traits of commercial interest but the 2 million farmed salmon escaping annually may enhance the risk of extinction of wild populations through genetic and ecological interactions. Here, we compare the transcription profiles of 3557 genes in the progeny of farmed and wild Atlantic salmon from Norway and Canada grown in controlled conditions, and demonstrate that five to seven generations of artificial selection led to heritable changes in gene transcription profiles, the average magnitude of the differences being 25% and 18% for at least 1.4% and 1.7% of the expressed genes in juvenile salmon from Norway and Canada, respectively. Moreover, genes showing significant transcription profile differences in both farmed strains (16%) all exhibited parallel changes. These findings, along with the identification of several genes whose expression profiles were modified through artificial selection, provide new insights into the molecular basis of parallel evolution, and suggest how gene flow from farmed escapees may affect the genetic integrity of wild populations.

Keywords: farmed salmon escapees, genome-wide transcription profiles, parallel evolution, recent evolution, salmon breeding, salmon conservation

Received 14 June 2005; revision received 28 September 2005; accepted 20 October 2005

Introduction

Consumers and environmentalists are increasingly concerned about the potential health and environmental consequences of producing genetically modified organisms. Yet, some of the criticisms of transgenic plants and animals may also apply to artificially selected breeds. In animals, strains selected for high production efficiency may be more susceptible to behavioural, physiological and immunological problems than wild populations (Rauw *et al*. 1998, 2000; Van der Waaij 2004). Also, many of the genetic changes accumulated in breeding strains are likely to be unknown.

Selective breeding of Atlantic salmon (*Salmo salar*) was initiated in Norway in 1971 and is now intensively practised in Chile, the United Kingdom, the United States, and Canada. In Norway, artificial selection was first limited to improving growth rate, but currently also targets traits such as age at sexual maturity, bacteria resistance, fat content and flesh colour. The selection response on each of these traits has been approximately 10% per generation (Gjoen & Bentsen 1997). However, phenotypic changes not specifically selected for have also been observed in Norwegian farmed salmon, including increased fat content in flesh (Rye & Gjerde 1996), poorer performance in the wild (Fleming *et al*. 2000; McGinnity *et al*. 2003) morphological and behavioural changes (Fleming & Einum 1997), increased growth hormone (GH) and insulin-like growth factor (IGF) levels (Fleming *et al*. 2002), as well as a higher feeding rate and food conversion efficiency (Thodesen *et al*. 1999). In Canada, the first breeding programme was initiated during the late 1980s, and targeted growth rate, age at sexual maturity, and more recently, bacteria resistance (Friars *et al*. 1995; O'Flynn *et al*. 1999).

The last decade has seen the worldwide production of farmed Atlantic salmon outstrip that of fisheries (FAO 2004), and the problem of farmed salmon escapees has taken alarming proportions, with about 2 million farmed salmon escaping annually from their sea cages while the natural populations in the North Atlantic are estimated to comprise approximately 4 million individuals (McGinnity *et al*. 2003). Escaped farmed salmon are thought to greatly enhance the wild populations' risk of extinction through genetic and ecological interactions (Fleming *et al*. 2000;

Correspondence: Christian Roberge, Fax: 1-418-656-7176; E-mail: christian.roberge@giroq.ulaval.ca

McGinnity *et al*. 2003), as well as by spreading diseases (Naylor *et al*. 2005). Knowledge of evolutionary changes induced by salmon breeding is crucial for assessing the consequences of genetic interactions between wild and escaped farmed salmon (Bentsen 1991). Namely, it is increasingly acknowledged that evolutionary changes may strongly depend on alterations in gene regulation (Wilson 1976), and the microarray technology offers the possibility of a genome-wide scan for gene transcription differences between farmed and wild salmon. Yet, the oldest salmon breeding programme has only operated for eight generations and no microarray study in any species has assessed the accumulation of heritable expression differences following so few generations of divergence. Recent studies in salmonids suggest that evolution may be rapid for phenotypic traits under strong selection (Hendry *et al*. 2000; Koskinen *et al*. 2003), but the genomic basis for such rapid changes is unknown.

Due to the similarities in controlled selective regimes, breeding strains also represent a powerful system for studies on parallel evolution. Parallel phenotypic evolution is considered one of the most eloquent manifestations of the role of selection in driving adaptive change (Harvey & Pagel 1991). Yet the genetic bases of parallel evolution are poorly understood (Schluter *et al*. 2004). Genetic changes implicated in natural phenotypic variation and evolution of a given trait often involve a disproportionate use of the same genes in independent parallel instances of phenotypic evolution instead of affecting genes randomly picked from the pool of all genes influencing this trait (Remington & Purugganan 2003; Schluter *et al*. 2004). This could either be due to the similarly biased production of genetic variation in close relatives (Haldane 1932; Schluter *et al*. 2004) or to a limited number of genes being actually important in phenotypic variation and evolution for a given trait (Stern 2000; Remington & Purugganan 2003). To our knowledge, very few published studies have specifically investigated parallelism in gene transcription (Ferea *et al*. 1999; Cooper *et al*. 2003).

Here, we used a 3557-gene cDNA array (Rise *et al*. 2004) to compare levels of gene transcription in the progeny of Atlantic salmon from the Norwegian National breeding programme first brood line (DOM1 or 'population 1' *sensu* Gjedrem *et al*. 1991) and its main population of origin (river Namsen, 64°27′0″N, 11°28′0″E), and from the Canadian breeding programme (strain 84JC, St John River, from the Atlantic Salmon Broodstock Development Program in St Andrews, New Brunswick) and its population of origin (St John River, 50°17′0″N, 64°20′0″W). Salmon of wild and farmed origin were reared under identical conditions within each country and transcription profiles of fish from the farmed strains and their wild counterparts were contrasted on 23 microarrays (Canada, 13 involving 26 fry; Norway, 10 involving 20 fry).

Materials and methods

Fish crosses

Twenty wild and 20 farmed genitors were used to produce 10 full-sibs pure wild families and 10 pure farmed families in both Norway and Canada, for a total of 40 families. Norwegian genitors were from the seventh generation of the Norwegian breeding programme first brood line and Canadian ones from the fifth generation of brood strain 84JC. Fertilized eggs of both types were kept under identical controlled conditions within systems [24 h darkness; yet temperatures were slightly different between countries: Norway, mean temperature (range) 3.9 °C (0.8–11.1); Canada, mean water temperatures 0.5–1 °C from January to March, gradually rising in April to 4 °C, and from 4 °C to 9 °C by the first of June]. Because juvenile characteristics play a determinant role in the expression of life history traits at later stages (Metcalfe *et al*. 1989), this study focuses on young of the year (fry stage). Sexually undifferentiated fry were sampled at the yolk-sac resorption stage, before exogenous feeding. The number of families used is relatively high compared to current standards for microarray experiments (see for instance Mackay *et al*. 2005), minimizing biases associated with individual parental effects, either by randomized sampling (Canada) or by using the same number of fry from each family (Norway). For the Canadian fish, however, some of the families may have been over-represented (and others under-represented) in the random sample. Genitors of wild origin were either grown from the egg in controlled conditions (Norway) or caught in the wild and kept in controlled conditions for several months until spawning (Canada), allowing us to control for environmental effects. The rearing of the parents in a controlled environment is also expected to considerably reduce environmental maternal effects (where offspring from different females have different characteristics due to their mothers having lived in different environments). Hence, differences in gene transcription related to maternal effects are expected to be mainly due to genetically based maternal effects, which are heritable and may respond to selection. These may thus be an important source of evolutionary changes among offspring of wild and farmed stocks, at least during the early stages, when genetic maternal effects have most influence in salmonids (Perry *et al*. 2005).

RNA extraction, labelling and cDNA hybridization

Following sampling, all experiments were performed by the same person (C.R.), in the same laboratory (Pavillon Marchand, Université Laval, Québec). Juveniles that had been frozen in liquid nitrogen were homogenized individually in 4 mL TRIzol® Reagent using a Polytron homogenizer, and total RNA was extracted in four separate

assays (to limit the error associated with the variability in the RNA extraction efficiency). Briefly, 200 µL chloroform (Sigma) was added to each millilitre of fish homogenate in TRIzol. After mixing and centrifuging (12 000 *g*, 4 °C, 15 min), the aqueous layers were transferred in new tubes and 1 mL isopropanol (Sigma) was added. The samples were then stored overnight at –80 °C. The following day, the samples were centrifuged 1 h (12 000 *g*, 4 °C) and the isopropanol discarded. The pellets were washed with 1 mL 70% ethanol, dried for 15 min at room temperature, resuspended in 40 µL non-DEPC treated nuclease-free water (Ambion) and spiked with 1 µL RNAse inhibitor (Ambion). RNA integrity was verified with a 2100 Bio-Analyzer (Agilent). For each sample, 15 µg of the pooled RNA of the four separate extractions was then retro-transcribed and labelled using Genisphere 3DNA Array 50 kit, Invitrogen's Superscript II retro-transcriptase and Cy3 and Alexa 647 dyes (Genisphere). The detailed protocol of the retro-transcription, labelling and hybridization procedures can be found at http://web.uvic.ca/cbr/grasp/(Genisphere Array 50 Protocol). Briefly, 15 µg total RNA were reverse-transcribed by using special oligo d(T) primers with 5′ unique sequence overhangs for the labelling reactions. Microarrays were prepared for hybridization by washing twice for 5 min in 0.1% SDS, washing five times for 1 min in MilliQ H_2O, immersing 3 min in 95 °C MilliQ H_2O, and drying by centrifugation (5 min at 800 RCF in 50 mL conical tubes). The cDNA was hybridized to the microarrays in a formamide-based buffer (25% formamide, 4X SSC, 0.5% SDS, 2X Denhardt's solution) with competitor DNA [LNA dT bloker (Genisphere), human COT-1 DNA (Sigma)] for 16 h at 51 °C in a humidified hybridization oven. The arrays were washed once for 5 min at 45 °C (2X SSC, 0.1% SDS), twice for 3 min in 2X SSC, 0.1% SDS at room temperature (RT), twice for 3 min in 1X SSC at RT, two times for 3 min in 0.1X SSC at RT, and dried by centrifugation. The Cy3 and Alexa 647 fluorescent dye attached to DNA dendrimer probes (3DNA capture reagent, Genisphere) were then hybridized to the bound cDNA on the microarray using the same hybridization solution as earlier; the 3DNA capture reagents bound to their complementary cDNA capture sequences on the oligo d(T) primers. This second hybridization was done for 2 h at 51 °C in a humidified hybridization oven. The arrays were then washed and dried as before.

Transcription profiles of fish from the farmed strains and their wild counterparts were contrasted on 23 microarrays (Canada, 13 involving 26 fry; Norway, 10 involving 20 fry). Since our main interest was in detecting changes driven by artificial selection, one farmed and one wild juvenile of the paired populations were compared directly on each microarray. This simple design maximizes the statistical power for detection of such differences for a given number of microarrays. Dye-sample coupling was balanced between fry of wild and domestic origins in both systems (dye flip on the biological replicates). The cDNA microarrays used here were obtained through the Genomic Research on Atlantic Salmon Project (GRASP, available from Ben F. Koop, bkoop@uvic.ca) and comprises 3557 clones from 18 high-complexity salmonid cDNA libraries, each printed as double, side-by-side spots on Telechem Superamine slides [6440 Atlantic salmon and 916 rainbow trout (*Oncorhynchus mykiss*) elements or spots; Rise *et al*. 2004]. Detailed information on the clones printed on the microarray can be found at http://web.uvic.ca/cbr/grasp/. Rise *et al*. (2004) provided information on the relative representation of diverse molecular functions in several cDNA libraries. However, they did not provide such information for the sequences printed on the cDNA microarray. Using the Web-based tool FatiGO (Al Shahrour *et al*. 2004) as well as information available on the GRASP website, we calculated the proportions of genes on the array involved in several biological processes (Fig. S1, Supplementary materials). Although calculated with all the information available, the results must be interpreted cautiously since gene ontology (GO) information regarding biological process at level 4 was only available for 18% (631 genes) of the sequences on the array.

Signal detection, data preparation and statistical analysis

Signal was detected using a ScanArray scanner from Packard BioScience. Spots were located and quantified with the QUANTARRAY software, using the adaptive circle spot quantification method and keeping the mean value of intensity for each spot. Local background and the data from bad spots were removed. Missing data were then imputed using the K-nearest neighbours imputer in SAM (Tusher *et al*. 2001) (15 neighbours). Signals from pairs of neighbouring replicate spots (PNRS) were averaged and data were normalized by dividing by the channel mean. To assess differences between wild and farmed salmon, data from Canada and Norway were analysed separately using a mixed model of ANOVA (Wolfinger *et al*. 2001) and the R/MAANOVA package (Kerr *et al*. 2000, 2002). Genes with low intensity data (mean intensity smaller than the mean intensity of the empty spot controls plus twice its standard deviation in both channels) were removed from the analysis, leaving 3058 and 2552 detected clones for the Norwegian and Canadian experiments, respectively. Under the assumption that all null hypotheses are true, 31 and 26 false positives at the $P < 0.01$ significance threshold are expected by chance alone. Prepared data were corrected for intensity-linked distortion using a regional LOWESS algorithm and fitted to a mixed model of variance using the R/MAANOVA package (www.jax.org/staff/churchill/labsite/software/Rmaanova/index.html). The model included the 'array' term as a random term and the 'sample type' (farmed

or wild) and 'dye' terms as fixed terms. A permutation-based *F*-test (F2, with 1000 sample ID permutations) was then performed, and restricted maximum likelihood was used to solve the mixed model equations. Variance homoscedasticity was tested prior to each ANOVA using the *F*-test function in the Excel spreadsheet software.

While our main interest was in detecting changes in farmed vs. wild salmon, we were also interested in using our data to assess for differences in transcription levels between salmon from Norway (river Namsen) and Canada (St John River). To do this, normalized intensity data obtained previously on separate arrays for wild individuals of both origins were analysed together using an ANOVA model with 'dye' and 'country' as fixed terms. The 'array' term had to be removed from this model because experimental variance associated with the array term would have been confounded with that linked to the 'country' term. This simpler model fits the data without notable bias, as verified by plotting the residuals vs. the fitted value for each gene in R/MAANOVA (not shown). Since the array-linked experimental variance is considered as random (Draghici 2003), removing the 'array' term from our model is not expected to bias the conclusions of the analysis. Yet, it is not ideal since the array-linked experimental variance will further inflate random error, making the test less powerful. Hence, while our experimental design was well suited for our main question of interest, this led to use a less powerful test to assess for differences between fish from Norway (river Namsen) and Canada (St John River).

Results

At a significance threshold for permutation corrected *P* values of $P < 0.01$, 68 and 74 pairs of neighbouring replicate spots (PNRS) on the array represented genes with significantly different transcription levels in the progeny of wild and farmed populations from Canada and Norway, respectively (Tables S1 and S2, Supplementary materials). After correcting for the expected number of false positives, this represented 1.4% and 1.7% of all detected PNRS for Norwegian and Canadian juveniles, respectively. These values represent conservative estimates of the proportion of genes for which the transcription level was modified by five to seven generations of domestication, which may include the effects of artificial selection, founder effects, genetic drift and inadvertent selection due to artificial rearing or to the selection of traits correlated to the traits of interest.

Potential caveats of this study include the difference in genetic diversity between farmed and wild salmon as well as the use of imputed data. First, a reduced genetic diversity is expected in farmed vs. wild salmon and was documented for Norwegian fish (e.g. Norris *et al*. 1999). This difference could be a problem when using ANOVA. Thus, *F*-tests for variance homoscedasticity in gene transcription profiles between farmed and wild salmon were performed for each detected gene on the prepared data before performing the ANOVA. In both the Canadian and Norwegian systems, less significant genes than expected by chance alone were observed (at $P < 0.05$, 13 genes showed significant differences over 2552 tests in Canada and 20 genes over 3058 tests in Norway). These were considered as false positives. Only one of them (clone SK1-0197, similar to Beta actin 1) was present in the lists of genes showing significant differences between farmed and wild salmon (Norwegian system). Hence, reduced genetic diversity in farmed populations does not ward against the use of ANOVA in this study. Second, most microarray data analysis methods require that all the data be present. Missing data must therefore be imputed [here by using the K-nearest neighbours imputer in SAM (Tusher *et al*. 2001)], which could influence the results. In this study, only a very small subset of the data was missing and had to be imputed (0.08% for the Canadian system and 0.12% for the Norwegian one). Only two of the genes for which a significant difference in transcription profile was observed between farmed and wild salmon had received an imputed value; both of these were in the Canadian system (nwh^18^87^nucleoside diphosphate kinase and pitl^504^374^ZipA). In these two cases, one value over 26 (13 arrays, 2 side-by-side replicates averaged) was imputed. As for nucleotide diphosphate kinase, two other non-imputed clones showed equivalent results. Thus, data imputation did not have much influence on our results.

Functional classification of the differentially expressed genes

Changes in transcription profiles were observed in different functional classes (Fig. 1). The proportion of significant changes for genes implicated in protein synthesis and muscle function was similar in Norway and Canada. In contrast, the proportion of changes for genes implicated in energy metabolism and metal ion sequestration was greater in the Canadian than in the Norwegian salmon, whereas the opposite pattern was observed for genes coding for molecular chaperones, blood transport proteins and cellular matrix proteins. A Fisher exact test for equality of proportions (Fisher & van Belle 1993; Draghici 2003) was applied to these data. Only the over-representation of genes implicated in regulation of cellular division and growth in the Canadian vs. Norwegian salmon significant genes was significant at $P > 0.05$ (Fig. 1). Tools allowing more reliable functional analysis of microarray data, such as DAVID/EASE (Dennis *et al*. 2003; Hosack *et al*. 2003) or ONTO-EXPRESS (Khatri *et al*. 2002) are presently poorly adapted to the analysis of data from nonmodel species custom cDNA microarrays for several reasons, the first one being the difficulty to get meaningful functional annotations for all the clones on such arrays. Hence, we

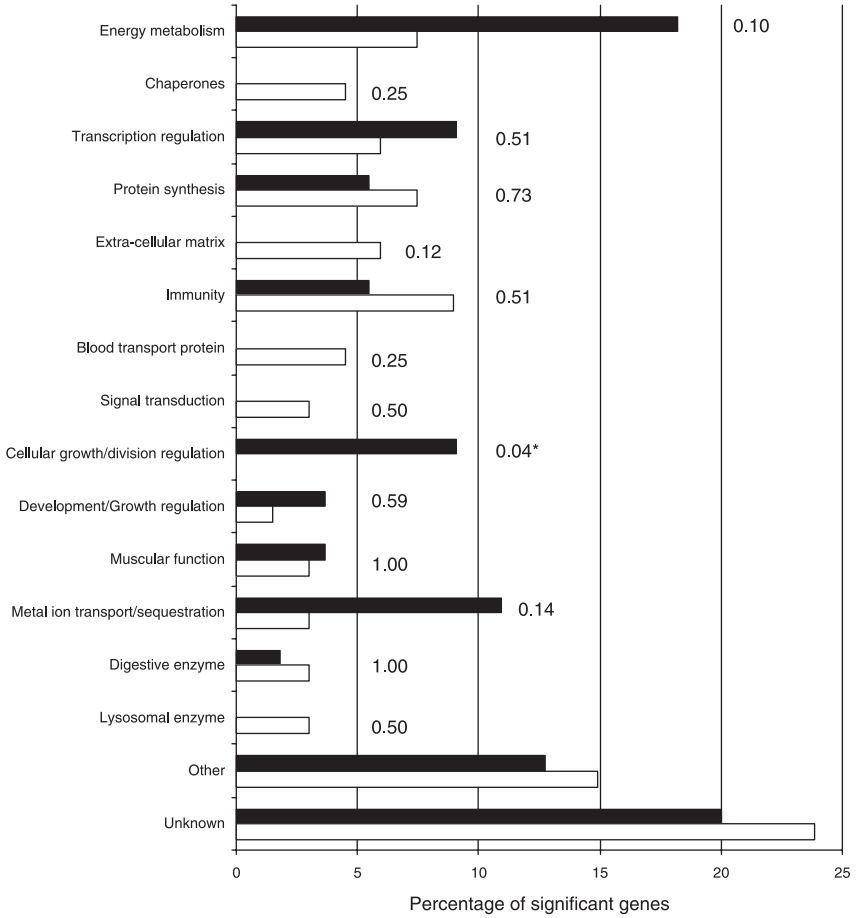


Fig. 1 Distribution of the differentially expressed genes in various functional classes. We show the percentage of all genes showing significant expression differences ($P < 0.01$, permutation-corrected P values) between farmed and wild salmon originating from the St John River (black) and from the river Namsen (white) in each of 16 functional classes defined from the literature. P values from Fisher's exact bilateral test for equality of proportions (Fisher & van Belle 1993; Draghici 2003) were calculated in the R environment (R Development Core Team, 2002) to assess for the significance of the differences in proportion between the Canadian and Norwegian systems observed among the diverse categories and added to the figure.

were only able to get GO information for a subset (< 18%) of the sequences on the array, using BLAST-gathered information available on the GRASP website. Protein accession numbers were converted to Unigen clusters numbers using DAVID Web-based conversion tool and these were imputed in EASE to assess for differences in GO representation between the array and each of the two sets of significant genes (using Fisher exact test). The results obtained partly reflect those presented in Fig. 1. When compared to the array, the 'extracellular', 'extracellular space', 'extracellular matrix' and 'protein binding' categories were all significantly over-represented in the Norwegian significant genes list (P value: 0.007, 0.018, 0.023, 0.045, respectively), while the 'regulation of cell proliferation', 'regulation of cellular process' and 'regulation of biological process' categories were significantly over-represented in the Canadian significant genes list (P value: 0.033, 0.039, 0.039, respectively). For the Canadian significant genes list, it is interesting to note that the 'transition metal ion binding' and the 'iron binding' categories also had low P values (0.051, 0.091, respectively). However, none of these changes would still be significant following a Bonferroni correction and the partial coverage of the array content appeals for a wary consideration of these results.

Genes with multiple significant PNRS

The specific design of the microarrays we used generated additional information (Table 1), which further supported the results obtained for several of the significant genes (Tables S1 and S2). Namely, two types of replicates were present according to their relative positions on the array: neighbouring replicates for each spot on the chip (PNRS, which were averaged), and additional non-neighbour replicates, often representing different cDNA clones of the same mRNA, for a limited proportion of all the cDNAs represented on the array (which were analysed separately without averaging). Hence, within the lists of significant genes, 12 genes were represented more than once. For instance, the metallothionein A-coding gene was represented by six separate pairs of spots, which all showed significant differences in transcription level in the same direction at $P < 0.01$. Such results substantially reduce the probability that these genes represent false positives due to technical artefacts.

Parallel changes

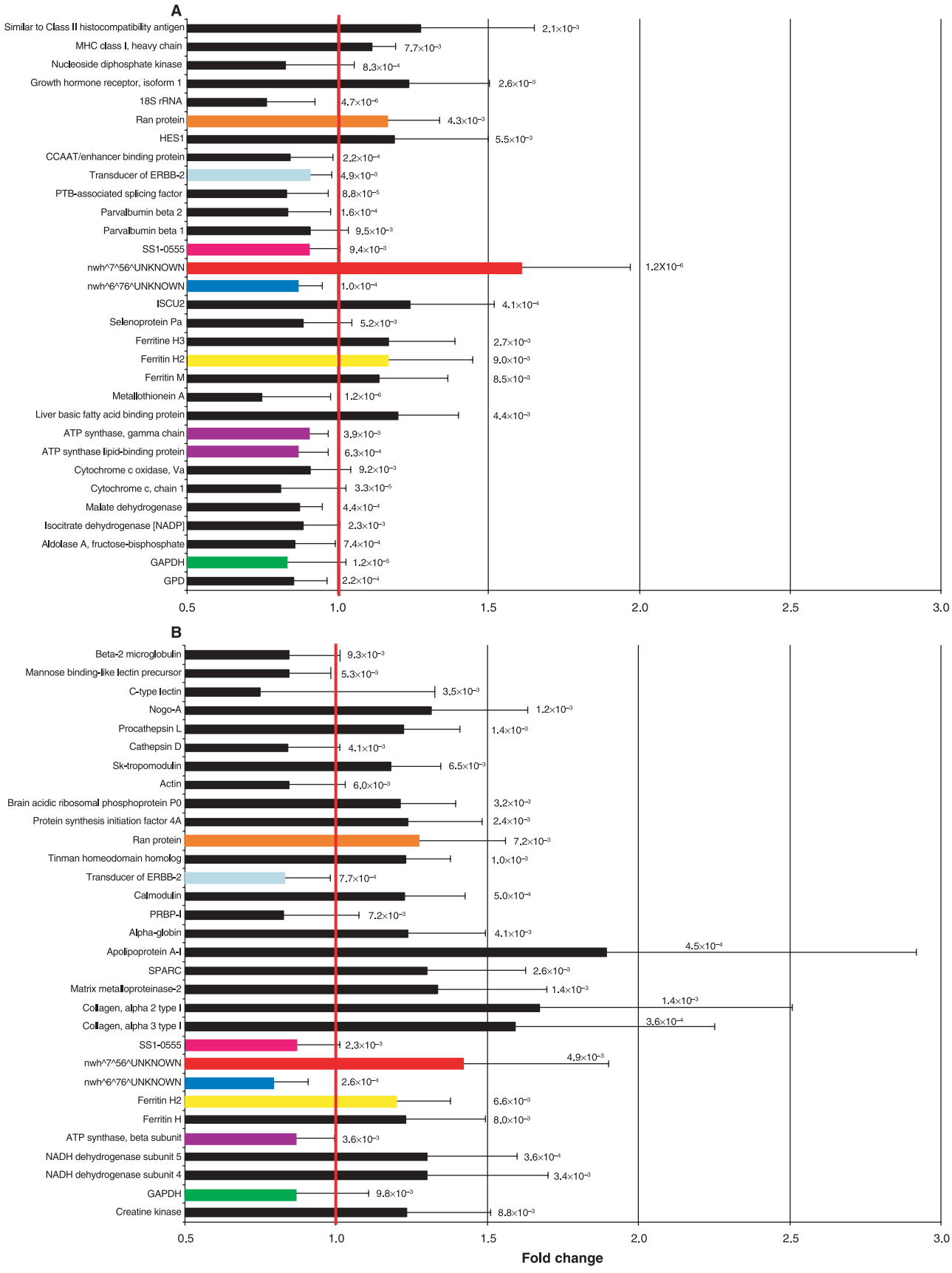
Many changes in gene transcription levels observed here were strain specific. Yet, our results also illustrate at least

Table 1 Genes for which multiple non-neighbour PNRS showed significant transcription level differences between the progeny of wild and domestic salmon. For each PNRS of genes with multiple significant PNRS at the $P < 0.01$ threshold, the average fold change in farmed relative to wild individuals from 13 (Norwegian) or 10 (Canadian) arrays and the permutation corrected (1000 permutations) P value from the F-test on the mixed model of ANOVA is given

Origin	Genes with multiple significant PNRS	Average fold change	P value
Canada	Cytochrome *c*	0.8	3.3×10^{-5}
		0.8	1.9×10^{-4}
		0.9	6.7×10^{-3}
	Metallothionein A	0.7	1.2×10^{-6}
		0.8	1.1×10^{-4}
		0.8	1.1×10^{-4}
		0.8	1.4×10^{-4}
		0.7	2.8×10^{-6}
		0.9	1.8×10^{-3}
	GAPDH	0.9	4.4×10^{-4}
		0.8	1.2×10^{-5}
		0.9	2.8×10^{-3}
		0.9	8.4×10^{-4}
	Nucleoside diphosphate kinase	0.9	8.3×10^{-4}
		0.9	1.2×10^{-3}
		0.9	6.7×10^{-3}
	HES1	1.2	9.2×10^{-3}
		1.2	5.5×10^{-3}
	Brain protein (44-like)	0.9	2.1×10^{-3}
		0.9	4.0×10^{-3}
Norway	NADH dehydrogenase subunit 5	1.3	3.6×10^{-4}
		1.3	4.1×10^{-3}
	Calmodulin	1.3	5.0×10^{-4}
		1.3	6.7×10^{-3}
		1.2	6.0×10^{-3}
		1.2	9.0×10^{-3}
	Clone SS1-0555	0.8	2.3×10^{-3}
		0.9	9.9×10^{-3}
	Protein synthesis initiation factor 4A	1.2	2.4×10^{-3}
		1.2	6.7×10^{-3}
	Brain acidic ribosomal phosphoprotein P0	1.2	3.2×10^{-3}
		1.2	3.5×10^{-3}
	α-globin	1.3	4.1×10^{-3}
		1.2	5.4×10^{-3}

four distinct ways by which directional selection could act on gene expression to result in parallel phenotypic evolution. First, equivalent changes could occur in the same genes in populations submitted to similar selective regimes. Hence, seven genes showed significant expression differences ($P < 0.01$) between farmed and wild salmon in both Norway and Canada, and all changed in a parallel direction (Fig. 2). The probability that seven genes show parallel changes by chance alone is $P = 7.8 \times 10^{-3}$, as evaluated using the binomial law with $\pi = 0.5$ (which is equivalent to the probability of at least seven successes over seven iterations in a sign test with $\pi = 0.5$). Given that some of the parallel genes were down-regulated and others up-regulated in the farmed strains, that they belong to very different functional groups and that no relation was found between them either through clustering analysis or with the PATHWAYASSIST software (Ariadne Genomics), we consider that the assumption of independence between the genes required for this test is reasonable. These seven genes represent 16% of the expected number of true positives in each population. It is noteworthy that three of these genes had unknown functions. A second level of

Fig. 2 *(Opposition)* Average fold change and standard deviation for a selection of genes differentially expressed between the progeny of farmed and wild salmon originating from (A) the St John River, Canada (mean of 13 microarrays), and (B) the river Namsen, Norway (mean of 10 microarrays). Permutation-corrected P values are given for each gene. For genes with multiple PNRS, only the smallest P value observed is given. A value above one (bold line) for a given gene represents overexpression in farmed relative to wild salmon. Colours indicate parallel changes between Canadian and Norwegian salmon, either in genes coding for different subunits of the same functional entity (ATP synthase) or in the same gene (the seven other instances). The genes in this figure were chosen either for their high level of significance, their particular contextual interest, or because they showed parallel changes in expression in both systems.


A
Similar to Class II histocompatibility antigen
MHC class I, heavy chain
Nucleoside diphosphate kinase
Growth hormone receptor, isoform 1
18S rRNA
Ran protein
HES1
CCAAT/enhancer binding protein
Transducer of ERBB-2
PTB-associated splicing factor
Parvalbumin beta 2
Parvalbumin beta 1
SS1-0555
nwh^7^56^UNKNOWN
nwh^6^76^UNKNOWN
ISCU2
Selenoprotein Pa
Ferritine H3
Ferritin H2
Ferritin M
Metallothionein A
Liver basic fatty acid binding protein
ATP synthase, gamma chain
ATP synthase lipid-binding protein
Cytochrome c oxidase, Va
Cytochrome c, chain 1
Malate dehydrogenase
Isocitrate dehydrogenase [NADP]
Aldolase A, fructose-bisphosphate
GAPDH
GPD
2.1×10⁻³
7.7×10⁻³
8.3×10⁻⁴
2.6×10⁻³
4.7×10⁻⁶
4.3×10⁻³
5.5×10⁻³
2.2×10⁻⁴
4.9×10⁻³
8.8×10⁻⁵
1.6×10⁻⁴
9.5×10⁻³
9.4×10⁻³
1.2X10⁻⁶
1.0×10⁻⁴
4.1×10⁻⁴
5.2×10⁻³
2.7×10⁻³
9.0×10⁻³
8.5×10⁻³
1.2×10⁻⁶
4.4×10⁻³
3.9×10⁻³
6.3×10⁻⁴
9.2×10⁻³
3.3×10⁻⁵
4.4×10⁻⁴
2.3×10⁻³
7.4×10⁻⁴
1.2×10⁻⁵
2.2×10⁻⁴
0.5
1.0
1.5
2.0
2.5
3.0
B
Beta-2 microglobulin
Mannose binding-like lectin precursor
C-type lectin
Nogo-A
Procathepsin L
Cathepsin D
Sk-tropomodulin
Actin
Brain acidic ribosomal phosphoprotein P0
Protein synthesis initiation factor 4A
Ran protein
Tinman homeodomain homolog
Transducer of ERBB-2
Calmodulin
PRBP-I
Alpha-globin
Apolipoprotein A-I
SPARC
Matrix metalloproteinase-2
Collagen, alpha 2 type I
Collagen, alpha 3 type I
SS1-0555
nwh^7^56^UNKNOWN
nwh^6^76^UNKNOWN
Ferritin H2
Ferritin H
ATP synthase, beta subunit
NADH dehydrogenase subunit 5
NADH dehydrogenase subunit 4
GAPDH
Creatine kinase
9.3×10⁻³
5.3×10⁻³
3.5×10⁻³
1.2×10⁻³
1.4×10⁻³
4.1×10⁻³
6.5×10⁻³
6.0×10⁻³
3.2×10⁻³
2.4×10⁻³
7.2×10⁻³
1.0×10⁻³
7.7×10⁻⁴
5.0×10⁻⁴
7.2×10⁻³
4.1×10⁻³
4.5×10⁻⁴
2.6×10⁻³
1.4×10⁻³
1.4×10⁻³
3.6×10⁻⁴
2.3×10⁻³
4.9×10⁻³
2.6×10⁻⁴
6.6×10⁻³
8.0×10⁻³
3.6×10⁻³
3.6×10⁻⁴
3.4×10⁻³
9.8×10⁻³
8.8×10⁻³
0.5
1.0
1.5
2.0
2.5
3.0
Fold change

parallelism was exemplified by the parallel changes observed in the expression of ATP synthase coding genes (Fig. 2). Here, parallel change did not affect the same gene but genes coding for different subunits of the same functional entity, ATP synthase. Third, equivalent changes in different protein-coding genes within a same pathway were observed. Thus, Canadian farmed salmon showed 23% overexpression of the growth hormone (GH) receptor gene, which corroborates the increased levels of GH reported in Norwegian farmed salmon by Fleming *et al*. (2002). However, this study did not confirm their observation since the transcription level of the GH-coding genes was not significantly different between Norwegian farmed and wild salmon. Many explanations could account for this: lack of statistical power, different life stages used (Fleming *et al*. 2002: 1 and 2+ parr; here: alevins before first feeding) and more importantly, the fact that plasmatic levels of GH are regulated post-transcriptionally by several mechanisms (see, for instance, Muller *et al*. 1999). Fourth, equivalent changes in gene transcription profiles could also occur in different pathways affecting a same biological function. Hence, many genes showing significant changes were in different molecular pathways but belonged to similar functional classes, including energy metabolism, transcription regulation, protein synthesis, immunity, muscle function and digestion, in both farmed strains, suggesting the importance of these functions in the response to the similar selection regimes (Fig. 1).

Differences between Norwegian and Canadian wild salmon populations

To assess for a potential source of the strain-specific differences in transcription profiles caused by domestication, we investigated transcription profile differences between the wild fry from river Namsen (Norway) and those from St John River (Canada). Of 3147 detected PNRS (mean intensity over that of the empty spot controls plus twice its standard deviation) in the grouped normalized data from 12 wild Canadian and 10 wild Norwegian fry, 944 (30%) exhibited significantly different variances in the detected signals between countries at $P < 0.05$. For the remaining 2203 genes, data were fitted to an ANOVA model and a permutation-based *F*-test was performed. While 22 false positives are expected by chance at the $P < 0.01$ threshold, 171 PNRS exhibited significant differences between North American and European salmon (Table S3, Supplementary materials). After subtracting the expected number of false positives, this represented 6.8% of the detected genes. The average fold change in transcription profiles within the set of significant genes for $P < 0.01$ was 63%. Several of the continent-specific changes induced by domestication involved the same genes for which transcription differed between wild salmon from Norway and Canada. Hence, of the 39 and 41 genes (expected number of true positives) for which transcription profiles were modified differentially by domestication in Norway and Canada, respectively (Tables S1 and S2), 17 (43%) and 13 (32%) also differed significantly between continents (Table 2). Of these, genes transcribed at higher levels in wild Norwegian salmon were generally either down-regulated in farmed Norwegian salmon (e.g. mannose binding lectin, lysozyme G) or up-regulated in farmed Canadian salmon (e.g. ependymin). Reciprocally, genes transcribed at higher levels in wild Canadian salmon were generally either up-regulated in farmed Norwegian salmon (e.g. Collagen, apolipoprotein A-I, NADH dehydrogenase) or down-regulated in Canadian farmed salmon (e.g. nucleoside diphosphate kinase, parvalbumin beta 1 and 2, malate dehydrogenase; Table 2).

Discussion

Elucidating the nature of changes in gene expression levels in farmed salmon may contribute to better assess the possible consequences of genetic interactions between wild and escaped farmed salmon. It is also of direct relevance to the understanding of contemporary evolution driven by human alterations. Here, we demonstrated that only five to seven generations of domestication led to significant changes in gene expression, the average magnitude of the observed differences being approximately 20% for at least 1.4 and 1.7% of the expressed genes at the juvenile stage. As evidenced by the parallel changes observed, some of these changes were generated by directional artificial selection for traits of interest for commercial purposes. In addition, other factors including founder effects, genetic drift and inadvertent selection due to artificial rearing or to the selection of traits correlated to the traits of interest may have contributed to the observed differences.

Changes in the transcription levels of genes involved in energy metabolism are especially interesting since they may influence the most important traits targeted by artificial selection: growth rate and age at sexual maturity. With the exception of the gene coding for NADH dehydrogenase and creatine kinase in Norwegian salmon, genes coding for enzymes involved in energy metabolism, including glycolysis, citric acid cycle and oxidative phosphorylation, were all under-transcribed in farmed salmon of both strains (Fig. 2). Decreased expression of these genes could lower the functional capacity in these pathways or may reflect reduced turnover of their components. The increased expression of creatine kinase is not in contradiction with a diminution in metabolic capacities since this enzyme acts downstream of ATP production and catalyses both ATP and creatine formation from phosphocreatine and ATP and the reverse reaction (Voet & Voet 2004). Under the assumption that over-transcription of the creatine

Table 2 Relation between country-specific changes caused by domestication and between-country differences in transcription profiles. Genes for which both a country-specific domestication changes and between-country differences in gene transcription profiles were observed at the $P < 0.01$ significance threshold are presented with their average fold change between and within countries. A value superior to 1 in the 'between countries' column represents an over-transcription in the Canadian compared to Norwegian salmon. A value above 1 in the country-specific column represents an over-transcription in the domestic population compared to its wild population of origin. Asterisks indicate cases where the between-country differences seem partly counter-balanced by domestication-linked changes

Gene	Between countries	Norway	Canada	
HMG-1	2.29	1.27		*
Malate dehydrogenase 2, NAD (mitochondrial)	1.85		0.87	*
Parvalbumin beta 1	2.57		0.91	*
Ferritin, middle subunit	3.96		1.14	
Salmo salar cDNA clone SK1-0455	1.69		1.16	
Mannose binding-like lectin	0.57	0.84		*
Salmo salar cDNA clone SK1-0473 similar to Beta actin 1	2.15	1.64		*
Salmo salar cDNA clone SS1-0134	1.73	0.83		
Salmo salar cDNA clone SS1-0466 similar to Ependymin	0.68		1.23	*
Salmo salar cDNA clone lna^9^95^UNKNOWN	0.66	0.85		*
Nucleoside diphosphate kinase	1.67		0.86–0.90	*
Parvalbumin beta 1	1.89		0.91	*
NADH dehydrogenase subunit 4	1.41	1.30		*
Actin	2.23	0.84		
Fatty acid binding protein	1.60		1.16	
Glyceraldehyde 3-phosphate dehydrogenase	1.41	0.83	0.86	
Collagen alpha 3(I)	1.69	1.59		*
LYSOZYME G	0.48	0.85		*
Salmo salar cDNA clone SNN-1095	0.55	0.81		*
APOLIPOPROTEIN A-I-1 PRECURSOR	1.45	1.89		*
Creatine kinase (EC 2.7.3.2)	1.93	1.23		*
Parvalbumin beta 2	1.52		0.83	*
Ferritin middle subunit	1.32		1.14	
P0 acidic ribosomal phophoprotein	1.71	1.21–1.22		*
Collagen alpha 2(I) chain precursor	1.93	1.67		*
Secreted protein, acidic, rich in cysteine (SPARC)	1.53	1.30		*
Beta-2 microglobulin	1.23	0.84		
Salmo salar cDNA clone lna^9^26^UNKNOWN	0.69		0.89	
Ferritin H-2	1.42	1.20	1.16	
High choriolytic enzyme 1 precursor	1.49		1.20	
Selenoprotein precursor	1.28		0.89	*
Lysozyme	0.56	0.85		*

kinase gene translates into increased enzymatic activity, this could represent a more efficient ATP buffering by phosphocreatine in the Norwegian farmed salmon relative to wild ones. Alternatively, heritable increase in the relative mass of skeletal muscle, with its moderate aerobic capacity and high levels of creatine kinase, could underlie the patterns observed in farmed salmon. In accordance with the increased fat levels in the flesh of farmed salmon (Rye & Gjerde 1996), the gene for Apolipoprotein A-I, a constituent of both chylomicrons and HDL, was strongly over-transcribed in the farmed progeny from Norway, and the transcription level of the gene for fatty acid binding protein was increased in Canadian farmed salmon. Altogether, these patterns are consistent with a reduction in basal metabolic rate and an increased metabolic efficiency of farmed juvenile salmon favouring allocation of resources towards growth and fat deposition. Knowing that resistance to infection has been selected both in Norway and Canada farmed salmon, it is also noteworthy that two genes coding for MHC antigens were over-transcribed in farmed salmon from Canada (Fig. 2).

We found evidence for parallelism in transcriptome evolution at several distinct levels in this study. Thus, up to 16% of all transcription profile changes between farmed populations and their wild populations of origin occurred for the same genes and in parallel for both farmed populations. In these cases, there is therefore strong evidence for the role of selection in driving phenotypic changes. Assuming that a large number of genes contribute to the expression of the natural variation in polygenic traits under selection (e.g. growth rate and age at sexual maturity), parallel phenotypic evolution should rarely involve

the same genes (Schluter *et al*. 2004). Three major non-mutually exclusive explanations were proposed to account for changes in the same genes in two independent lineages having evolved in parallel: (i) the similarly biased production of genetic variation in close relatives (Haldane 1932; Schluter *et al*. 2004); (ii) a limited number of genes being actually important in phenotypic variation and evolution for a given trait (Stern 2000; Remington & Purugganan 2003); or (iii) developmental constraints shared between closely related groups (e.g. West-Eberhard 2003). It is noteworthy that three of the parallel genes we observed have unknown functions: their parallel expression changes in both breeding strains suggest they might affect growth rate or age at sexual maturity, giving a first glimpse of their potential functions. Yet, we also observed parallel changes in the expression of genes coding for different subunits of the ATP synthase (Fig. 2), equivalent changes for different genes within pathways and the representation of common functional classes in the significant changes. These levels of parallelism not necessarily involving the same genes, as well as the fact that most observed changes are strain specific, suggest that selection also acted on genes for which genetic variation was not biased in a similar way in Norwegian and Canadian salmon. Accordingly, significant differences in both gene expression variance and mean between the wild Norwegian and Canadian salmon were observed, which could modify their responses to similar selection regimes. Also, 43% (Canada) and 32% (Norway) of the genes for which transcription profiles differed between farmed and wild fish also differed significantly between the wild populations of river Namsen (Norway) and St John River (Canada). Of these, genes transcribed at higher levels in wild Norwegian salmon were generally either down-regulated in farmed Norwegian salmon (e.g. mannose binding lectin, lysozyme G) or up-regulated in farmed Canadian salmon (e.g. ependymin). Reciprocally, genes transcribed at higher levels in wild Canadian salmon were generally either up-regulated in farmed Norwegian salmon (e.g. Collagen, apolipoprotein A-I, NADH dehydrogenase) or down-regulated in Canadian farmed salmon (e.g. nucleoside diphosphate kinase, parvalbumin beta 1 and 2, malate dehydrogenase; Table 2). This illustrates how pre-existing differences in gene expression profiles can modify the way by which two related populations respond to selection. It also suggests that genetic or developmental constraints can limit the contribution of a given gene to the evolution of a phenotype under directional selection.

Substantial genetic (based on neutral markers) and heritable phenotypic differences were observed between Atlantic salmon from Europe and North America (Ståhl 1987; Phillips & Hartley 1988; Verspoor & McCarthy 1997). Accordingly, we observed differences both in gene expression variance (30% of the detected genes had non-homogenous variances) and means (average change of 63%, for 6.8% of the detected genes with homoscedastic variances) between the progeny of wild individuals from Norway and Canada. These differences could be causally linked with some of the strain-specific differences in the domestic strains (Table 2). Strain-specific changes between wild and domestic salmon may also be associated with differences in the selection regimes applied in Norway and Canada.

In addition, this study provided a first identification of many structural genes for which average transcription levels differ significantly between European and North American salmon (Table S3). Among the most interesting of these are fish allergens (type I Collagen and parvalbumines), key genes of energy metabolism (malate dehydrogenase, NADH dehydrogenase, creatine kinase, GAPDH) and several ribosomal protein coding genes (Table S3). However, these results must be interpreted cautiously since they were obtained from only two wild populations. Clearly, studies involving more populations from both continents will be needed to confirm whether the observed differences truly reflect general intercontinental differences of expression at the early fry stage.

Finally, our results also exemplify how founder effects, genetic drift or inadvertent selection may have caused nondesirable evolutionary changes in farmed salmon (McGinnity *et al*. 2003). Hence, previous studies revealed a positive correlation between growth rate and fat content in Atlantic salmon (Quinton *et al*. 2005) and showed that farmed salmon have a higher fat content and eat more frequently (Thodesen *et al*. 1999). Our results provided strong evidence for a 21% underexpression (average of six PNRS) of the metallothionein (MT) A gene in Canadian farmed salmon. Interestingly, it has been shown that MT-null mouse are obese, as a result of a higher food intake and of an abnormal energy balance (Beattie *et al*. 1998), which points towards a role for MT in the regulation of energy balance. MT is also a key factor in adaptation to heavy-metal environments (Posthuma & Vanstraalen 1993). Therefore, introgression of this reduced MT expression from farmed into wild populations could reduce their resistance to environmental pollutants. Admittedly, the phenotypic importance of the transcription level changes observed in this study must be interpreted cautiously since the link between transcription level and phenotypic expression is subjected to other levels of regulation. Studying the actual impact of the transcription profile changes observed on phenotype variation and fitness will therefore represent a necessary and exciting challenge in future studies.

Acknowledgements

This work was supported by a research grant from Science and Engineering Research Canada (NSERC strategic programme) to

L.B., H.G. and J.J. Dodson, a post-graduate scholarship from the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) to C.R., and funding from the Research Council of Norway to S.E. This study is a contribution to the research programmes of CIRSA (Centre Inter-universitaire de Recherche sur le Saumon Atlantique) and Québec-Océan. We thank Trevor Goff and the staff from the Mactaquac Fish Hatchery, Department of Fisheries and Oceans (DFO) Canada and the staff at the NINA Research Station, Ims, Norway, for their assistance in rearing fish. We are also grateful to Brian Glebe (DFO) for providing fish from the Canadian farmed strain, to Sean Rogers, Jon Slate and three anonymous referees for their valuable input on an earlier version of the manuscript, as well as to Hélène Crépeau and Gary Churchill for their comments on statistical issues.

Supplementary materials

The supplementary material is available from http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC2807/MEC2807sm.htm

Fig. S1 Distribution of 631 genes on the array with GO annotation at level 4 among different biological processes. Details of how GO annotations were linked to the salmon sequences are provided on the GRASP website (http://web.uvic.ca/cbr/grasp/geneIDmethods.v2.txt). Briefly, the BLASTX online tool was used to find annotated polypeptidic sequences similar to those translated from the sequences printed on the array. The calculation of the proportion of GO annotation in each biological process was performed using the Web-based tool FatiGO (Al-Shahrour *et al.* 2004), using protein identifiers as inputs. Biological processes for which the observed proportion was under 2% were grouped in the category 'other'.

Table S1 PNRS representing genes differentially expressed between the progeny of farmed and wild Canadian salmons. PNRS are classified in various biological/cellular functional groups. The average fold change in gene expression in farmed salmon compared to wild individuals (data from 13 arrays) and its standard deviation is given for each PNRS, as well as *P* values obtained from the *F*-test on the mixed model of ANOVA. Genes for which several PNRS were significant are in bold. Colours identify genes or functional entities showing parallel changes in Canada and Norway. Clone identification follows notation from the GRASP. GenBank Accession numbers refer to the GenBank sequence the most similar to the amplified cDNA portion spotted on the array

Table S2 PNRS representing genes differentially expressed between the progeny of farmed and wild Norwegian salmons. PNRS are classified in various biological/cellular functional groups. The average fold change in gene expression in farmed salmon compared to wild individuals (data from 10 arrays) and its standard deviation is given for each PNRS, as well as *P* values obtained from the *F*-test on the mixed model of ANOVA. Genes for which several PNRS were significant are in bold. Colours identify genes or functional entities showing parallel changes in Canada and Norway. Clone identification follows notation from the GRASP. GenBank Accession numbers refer to the GenBank sequence the most similar to the amplified cDNA portion spotted on the array

Table S3 PNRS representing the 171 genes differentially expressed between the progeny of wild salmon from Norway (Namsen River) and Canada (St John River) at the $P < 0.01$ threshold on permutation-based *P* values. The average fold change in gene expression in Canadian compared to Norwegian wild individuals (data from 22 individuals) is given for each PNRS, as well as *P* values obtained from the *F*-test on the ANOVA model. GenBank Accession numbers refer to the GenBank sequence the most similar to the amplified cDNA portion spotted on the array. Clone identification follows notation from the GRASP

References

Al Shahrour F, Diaz-Uriarte R, Dopazo J (2004) FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics*, **20**, 578–580.

Beattie JH, Wood AM, Newman AM *et al.* (1998) Obesity and hyperleptinemia in metallothionein (-I and -II) null mice. *Proceedings of the National Academy of Sciences, USA*, **95**, 358–363.

Bentsen HB (1991) Quantitative genetics and management of wild populations. *Aquaculture*, **98**, 263–266.

Cooper TF, Rozen DE, Lenski RE (2003) Parallel changes in gene expression after 20,000 generations of evolution in *Escherichia coli*. *Proceedings of the National Academy of Sciences, USA*, **100**, 1072–1077.

Dennis G, Sherman BT, Hosack DA *et al.* (2003) DAVID: database for annotation, visualization, and integrated discovery. *Genome Biology*, **4**.

Draghici S (2003) *Data Analysis Tools for DNA Microarrays*. CRC Press, London.

FAO (2004) *Fisheries Department statistical databases and software: FISHSTAT PLUS*. Available from www.fao.org/fi/statist/FISOFT/FISHPLUS.asp.

Ferea TL, Botstein D, Brown PO, Rosenzweig RF (1999) Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proceedings of the National Academy of Sciences, USA*, **96**, 9721–9726.

Fisher LD, van Belle G (1993) *Biostatistics: A Methodology for Health Sciences*. John Wiley & Sons, New York.

Fleming IA, Einum S (1997) Experimental tests of genetic divergence of farmed from wild Atlantic salmon due to domestication. *ICES Journal of Marine Science*, **54**, 1051–1063.

Fleming IA, Hindar K, Mjolnerod IB *et al.* (2000) Lifetime success and interactions of farm salmon invading a native population. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **267**, 1517–1523.

Fleming IA, Agustsson T, Finstad B, Johnsson JI, Bjornsson BT (2002) Effects of domestication on growth physiology and endocrinology of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, **59**, 1323–1330.

Friars GW, Bailey JK, Oflynn FM (1995) Applications of selection for multiple traits in cage-reared Atlantic salmon (*Salmo salar*). *Aquaculture*, **137**, 213–217.

Gjedrem T, Gjoen HM, Gjerde B (1991) Genetic-origin of Norwegian farmed Atlantic salmon. *Aquaculture*, **98**, 41–50.

Gjoen HM, Bentsen HB (1997) Past, present, and future of genetic improvement in salmon aquaculture. *ICES Journal of Marine Science*, **54**, 1009–1014.

Haldane JBS (1932) *The Causes of Evolution*. Longmans, London.

Harvey PH, Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford, UK.

Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, **290**, 516–518.

Hosack DA, Dennis G, Sherman BT, Lane HC, Lempicki RA (2003) Identifying biological themes within lists of genes with EASE. *Genome Biology*, **4**, R70. doi: 10.1186/gb-2003-4-10-r70.

Kerr MK, Martin M, Churchill GA (2000) Analysis of variance for gene expression microarray data. *Journal of Computational Biology*, **7**, 819–837.

Kerr MK, Afshari CA, Bennett L *et al.* (2002) Statistical analysis of a gene expression microarray experiment with replication. *Statistica Sinica*, **12**, 203–217.

Khatri P, Draghici S, Ostermeier GC, Krawetz SA (2002) Profiling gene expression using ONTO-EXPRESS. *Genomics*, **79**, 266–270.

Koskinen MT, Haugen TO, Primmer CR (2003) Corrigendum: Contemporary fisherian life-history evolution in small salmonid populations (vol. 419, p. 826, 2002). *Nature*, **421**, 656.

Mackay TFC, Heinsohn SL, Lyman RF, Moehring AJ, Morgan TJ, Rollmann SM (2005) Genetics and genomics of *Drosophila* mating behavior. *Proceedings of the National Academy of Sciences, USA*, **102**, 6622–6629.

McGinnity P, Prodohl P, Ferguson K *et al.* (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **270**, 2443–2450.

Metcalfe NB, Huntingford FA, Graham WD, Thorpe JE (1989) Early social-status and the development of life-history strategies in Atlantic salmon. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **236**, 7–19.

Muller EE, Locatelli V, Cocchi D (1999) Neuroendocrine control of growth hormone secretion. *Physiological Reviews*, **79**, 511–607.

Naylor R, Hindar K, Fleming IA *et al.* (2005) Fugitive salmon: assessing the risks of escaped fish from net-pen aquaculture. *Bioscience*, **55**, 427–437.

Norris AT, Bradley DG, Cunningham EP (1999) Microsatellite genetic variation between and within farmed and wild Atlantic salmon (*Salmo salar*) populations. *Aquaculture*, **180**, 247–264.

O'Flynn FM, Bailey JK, Friars GW (1999) Responses to two generations of index selection in Atlantic salmon (*Salmo salar*). *Aquaculture*, **173**, 143–147.

Perry GML, Audet C, Bernatchez L (2005) Maternal genetic effects on adaptive divergence between anadromous and resident brook charr during early life history. *Journal of Evolutionary Biology*, **18**, 1348–1361.

Phillips RB, Hartley SE (1988) Fluorescent banding-patterns of the chromosomes of the genus *Salmo*. *Genome*, **30**, 193–197.

Posthuma L, Vanstraalen NM (1993) Heavy-metal adaptation in terrestrial invertebrates — a review of occurrence, genetics, physiology and ecological consequences. *Comparative Biochemistry and Physiology Part C: Pharmacology Toxicology and Endocrinology*, **106**, 11–38.

Quinton CD, McMillan I, Glebe B (2005) Development of an Atlantic salmon Salmo salar genetic improvement program: genetic parameters of harvest body weight and carcass quality traits estimated with animal models. *Aquaculture*, **247**, 211–217.

Rauw WM, Kanis E, Noordhuizen-Stassen EN, Grommers FJ (1998) Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Production Science*, **56**, 15–33.

Rauw WM, Luiting P, Bakken M *et al.* (2000) Behavioural differences in non-reproductive adult females in a long-term selection experiment for litter size in mice. *Applied Animal Behaviour Science*, **66**, 249–262.

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.

Rye M, Gjerde B (1996) Phenotypic and genetic parameters of body composition traits and flesh colour in Atlantic salmon, *Salmo salar* L. *Aquaculture Research*, **27**, 121–133.

Schluter D, Clifford EA, Nemethy M, McKinnon JS (2004) Parallel evolution and inheritance of quantitative traits. *American Naturalist*, **163**, 809–822.

Ståhl G (1987) Genetic population structure of Atlantic salmon. In: *Population Genetics and Fishery Management* (eds Ryman N, Utter F), pp. 121–140. University of Washington Press, Seattle.

Stern DL (2000) Perspective: evolutionary developmental biology and the problem of variation. *Evolution*, **54**, 1079–1091.

Thodesen J, Grisdale-Helland B, Helland SJ, Gjerde B (1999) Feed intake, growth and feed utilization of offspring from wild and selected Atlantic salmon (*Salmo salar*). *Aquaculture*, **180**, 237–246.

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences, USA*, **98**, 5116–5121.

Van der Waaij EH (2004) A resource allocation model describing consequences of artificial selection under metabolic stress. *Journal of Animal Science*, **82**, 973–981.

Verspoor E, McCarthy E (1997) Genetic divergence at the NAD+-dependent malic enzyme locus in Atlantic salmon from Europe and North America. *Journal of Fish Biology*, **51**, 155–163.

Voet D, Voet JG (2004) *Biochemistry*. John Wiley & Sons, New York.

West-Eberhard MJ (2003) *Developmental Plasticity and Evolution*. Oxford University Press, New York.

Wilson AC (1976) Gene regulation in evolution. In: *Molecular Evolution* (ed. Ayala FJ), pp. 225–234. Sinauer Associates, Sunderland, Massachusetts.

Wolfinger RD, Gibson G, Wolfinger ED *et al.* (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *Journal of Computational Biology*, **8**, 625–637.

This study is part of C. Roberge's doctoral research in L. Bernatchez's laboratory, which aims to study the genetic basis of recent evolution in Atlantic salmon, both in nature and through artificial selection. S. Einum is interested in various aspects of salmon evolution and conservation, including assessing the effects of interactions between wild and farmed salmon. This study also overlaps H. Guderley's interest in biochemical adaptation to changes in abiotic and biotic factors. The major interests of L. Bernatchez are in the understanding of the patterns and processes of molecular and organismal evolution, as well as their significance to conservation.
