Alternative life histories shape brain gene expression profiles in males of the same population

Nadia Aubin-Horth¹*, Christian R. Landry², Benjamin H. Letcher³,⁴ and Hans A. Hofmann¹

¹Bauer Center for Genomics Research, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA
²Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA
³Silvio O. Conte Anadromous Fish Research Center, US Geological Survey, Biological Resources Division, PO Box 796, One Migratory Way, Turners Falls, MA 01376, USA
⁴Department of Natural Resources Conservation, University of Massachusetts, Amherst, MA 01003, USA

Atlantic salmon (Salmo salar) undergo spectacular marine migrations before homing to spawn in natal rivers. However, males that grow fastest early in life can adopt an alternative ‘sneaker’ tactic by maturing earlier at greatly reduced size without leaving freshwater. While the ultimate evolutionary causes have been well studied, virtually nothing is known about the molecular bases of this developmental plasticity. We investigate the nature and extent of coordinated molecular changes that accompany such a fundamental transformation by comparing the brain transcription profiles of wild mature sneaker males to age-matched immature males (future large anadromous males) and immature females. Of the ca. 3000 genes surveyed, 15% are differentially expressed in the brains of the two male types. These genes are involved in a wide range of processes, including growth, reproduction and neural plasticity. Interestingly, despite the potential for wide variation in gene expression profiles among individuals sampled in nature, consistent patterns of gene expression were found for individuals of the same reproductive tactic. Notably, gene expression patterns in immature males were different both from immature females and sneakers, indicating that delayed maturation and sea migration by immature males, the ‘default’ life cycle, may actually result from an active inhibition of development into a sneaker.

**Keywords:** microarray; gene expression; plasticity; behaviour; reproduction; brain

1. INTRODUCTION

In many species, differences in environmental conditions can result in striking divergences in life history. Atlantic salmon, Salmo salar, exhibit spectacular male alternative reproductive life histories that are the result of developmental plasticity. Males either attain sexual maturity precociously as parr (first reproduction at 1–3 years) to become sneakers (stealing mating) that are 10 times smaller than their migratory conspecifics or migrate out to sea and return at the age of 3–7 years as large smaller than their migratory conspecifics or migrate out to sea and return at the age of 3–7 years as large anadromous animals (Fleming 1998). The quantitative genetic, behavioural and physiological differences between sneakers and immature males have been studied extensively (figure 1; Hutchings & Myers 1994; Aubin-Horth & Dodson 2004). However, little is known regarding the underlying molecular mechanisms of this developmental plasticity. It is generally assumed that the regulation of gene activity is an important component of phenotypic plasticity (West-Eberhard 2003). The advances in functional genomics, particularly the simultaneous monitoring of thousands of genes in individual tissues, can now facilitate a synthesis between mechanistic insights at the molecular, cellular, and physiological levels and an organismal understanding of plasticity within an evolutionary framework.

Here we present the first systematic analysis of the activity of thousands of genes in the brains of these wild animals, the location where environmental demands are integrated with physiological status to sculpt the phenotype (Hofmann 2003; Whitfield et al. 2003). Analysing brain gene expression profiles allows us to identify (i) how many genes vary in expression between sneaker and immature male brains; (ii) what biological processes are implicated and how they relate to known physiological, behavioural and morphological differences between tactics as well as to life-history theory predictions; (iii) genes previously not implicated in the sneaker tactic; and (iv) determine if expression patterns can classify individuals according to their tactic in wild animals. We measured brain gene expression profiles in four sneaker males and eight immature juveniles (four males and four females), collected at spawning time in October 2002 in a tributary of the Connecticut River, Massachusetts, USA, using a microarray constructed from salmonid cDNA libraries (Rise et al. 2004), which allows the monitoring of the transcriptional level of thousands of genes simultaneously. We focused on differences between individuals of the same age (freshwater life before migration), but with different phenotypes, and therefore did not include large anadromous males and females, who are several years older and are protected. Our results illustrate how hypotheses regarding life-history evolution can be addressed using a genomics approach (Stearns & Magwene 2003).

* Author for correspondence (naubin-horth@cgr.harvard.edu).

Received 4 March 2005
Accepted 20 April 2005
2. METHODS

(a) Sampling

We collected immature females, immature males and mature sneaker males of Atlantic salmon (*Salmo salar*) of age 1 in the Sawmill River (42° 30' N, 74° 30' W), a tributary of the Connecticut River in Western Massachusetts, USA, in October 2002. All fishes used were collected between 1100 and 1400 of the same day. Electrofishing was used to collect fishes individually. Each fish was immediately euthanized in buffered MS-222 (100 mg l⁻¹). Length and mass were then measured, and the brain was dissected within 5 min after death and transferred to RNAlater storage solution (Ambion), which stops RNase activity. The interval between collecting and tissue storage was always less than 11 min. Sex was then determined by dissection.

(b) Microarray hybridization

Four fishes (biological replicates) were assayed separately per phenotype. Total RNA was extracted from brains according to a standard Trizol protocol (Invitrogen), following tissue homogenization (Tissue Tearor, Biospec Products). All RNA samples were kept separate in subsequent analyses. Four independent reverse transcriptions were performed on total RNA extracted from each brain sample according to a standard 3DNA array 350 protocol (Genisphere). These cDNA samples from one individual were then pooled and aliquoted. cDNA from each fish was used four times (technical replicates, including two dye-swaps) in a loop design comparison, such that an individual of a given group (sneaker male, immature male, female) was directly compared with two individuals of each of the other groups (figure 2), without the need for a reference sample (Churchill 2002; Townsend 2003). Using such a loop design with biological and technical replicates and including dye-swaps for each individual has been shown to substantially increase reliability of microarray results and to reduce the probability of dye bias by means of transitive comparisons (Liang et al. 2003; Townsend 2003). The microarray used was constructed from 18 high complexity salmonid cDNA libraries (described in Rise et al. 2004). We purchased this microarray, called the ‘3700 array’ from the Genomic Research on Atlantic Salmon Project (GRASP) consortium. Physical platform description is available at NCBI GEO GPL966. Clone sequences, library of origin, and annotation are available from the GRASP Web site (http://web.uvic.ca/cbr/grasp/). We used 24 microarrays to compare 48 independent labelling reactions (figure 2). cDNA samples were competitively hybridized on a microarray slide according to a standard 3DNA Array 350 protocol (Genisphere). Arrays were scanned with an Axon 4000B scanner (Axon Instruments) using GenePix 4.0 software (Axon Instruments). Spots were examined individually and flagged as ‘bad’ if irregularities occurred.

(c) Transcription analysis

Raw data (after flags filtering) were imported into R software (v. 1.9; R Development Core Team 2004) and normalized using the Linear Models for Microarray Data package (Llimma v. 1.6.5; Smyth et al. 2003). Background-subtracted mean intensities were normalized using a within array loess normalization. These intensities were used in a Bayesian analysis of gene expression levels (BAgEL v. 3.6 (Townsend & Hartl 2002)). Out of the 7356 spots containing salmonid cDNA, 659 spots could not be reliably analysed because of low hybridization quality in most replicates, such that 6697 spots were used in the BAGEL analysis. We determined that this corresponded to 2917 genes, using the TIGR gene indices for ATLANTIC SALMON v. 1.0 (Quackenbush et al. 2000). This method estimates gene expression level and error variances for each gene by Markov chain Monte Carlo integration of the likelihood function of observed gene expression ratios, and incorporates a prior distribution for the parameters. A non-informative uniform prior distribution is used by default by this method. BAGEL takes advantage of additional information obtained from transitive comparisons of individuals in loop designs experiments (Churchill 2002; Townsend & Hartl 2002; Townsend 2003). For each gene, we obtain a relative gene expression level for each group, 95% credible intervals for gene expression level and a Bayesian posterior probability (p) of significant differences between two groups. Genes found to be significantly highly expressed in sneaker male brains relative to immature male brains (p<0.01) were labelled ‘sneaker biased’ while genes found significantly highly expressed in immature males compared with sneaker males (p<0.01) were labelled ‘immature male biased’.

(d) Functional annotation

Each significantly differentially regulated gene between sneaker and immature males, as determined by the Bayesian analysis, was assigned to a biological process category following its annotation by the GRASP EST database (http://web.uvic.ca/cbr/grasp/) and TIGR gene indices. Genes with no sequence, that had a BLAST hit to unannotated sequence or that were of unknown function in the brain were removed from analysis. Seventeen biological processes were associated to differences in the brain of sneaker and immature males. Each gene found to be significantly upregulated in one male type compared with the other male type had an associated biological process. Therefore, genes related to a biological process can be sneaker biased, while other genes classified in that same biological process can be immature male biased. We hypothesize that if a biological process is generally independent of the alternative phenotypes examined here, genes associated with that process would probably be represented at random (equally) in both sneaker-biased and immature male-biased gene lists. If non-random distributions of genes that are linked to a biological process are found between the sneaker-biased gene list and the immature male-biased gene list, then we hypothesized that this biological process is statistically associated with, and is probably important for, the development of the phenotype. For each biological process category, the proportion of sneaker-biased and immature male-biased genes was calculated. An exact binomial test (R software v. 1.9) was performed to detect statistically significant over-representation towards one reproductive tactic in the number of genes in this category.

(e) Individual variation in transcription profiles

We asked whether the overall gene expression profiles are shared among representatives of the same phenotype despite the individual variation probably present in animals collected in the natural environment. To this end, we used unsupervised hierarchical clustering to investigate how expression profiles of individual fish group. Using the Bayesian analysis of gene expression levels (BAgEL), we first estimated gene expression levels for each of the twelve individuals separately.
In the cluster analysis, we then used the expression levels for each individual of the 1017 genes that had shown significant differential expression in the first Bayesian analysis comparing phenotypes. We created 1000 gene expression datasets of the same size as the original by re-sampling gene expression levels with replacement in the individual gene expression dataset (bootstrap; Sokal & Rohlf 1995). Pearson correlation coefficients ($r$) were calculated between log$_2$-transformed brain gene expressions of all individuals for each of these 1000 datasets and the resulting dissimilarity indices ($1-r$) were used in a clustering analysis of individuals (successive agglomerative clustering using an average-linkage method, UPGMA method, Neighbour package, PHYLIP v. 3.6 (Felsenstein 2004)). The 1000 clusters obtained were used to construct a consensus tree (Consense package, PHYLIP v. 3.6), which was visualized using TREEVIEW (v. 1.6.6; Page 1996).

(f) Quantitative real-time PCR
The same individuals used in the microarrays were used in the quantitative real-time PCR validation experiments except for one individual (immature male) for which there was no RNA left. RNA samples were first treated with DNase I following manufacturer protocol (Invitrogen). An aliquot of these RNA samples was used in a Ribogreen assay (Invitrogen) to quantify RNA concentration of each sample based on fluorescence. This allows using exactly the same amount of starting RNA for each sample in later reverse transcription reactions and alleviates the need for inclusion of ‘house keeping genes’ (Hashimoto et al. 2004). Reverse transcription was performed in duplicate according to standard superscript II protocol (Invitrogen) using a primer solution containing poly dTand random hexamers (Invitrogen). cDNA from two replicates was then pooled. 5 ng of cDNA was used in a quantitative real-time PCR assay using a standard SYBR-GREEN RTPCR manufacturer’s protocol (Qiagen) in an Opticon thermocycler (MJ research). Each fish was tested in three replicates for a given gene. Primers were designed using Primer3 and AmpLyzer (1.0) software. Primer sequences were: Prolactin: 5-CGCCCCACTCTACACACCC-3 and 5-AGCAGGACATGAGGAAGTGG-3; Pro-opiomelanocortin (POMC): 5-GAGCAGTGTGTTCTGACTGC-3 and 5-GGAGGCTGGAAGTGG-3; Somatostatin: 5-AGCCAAGGACTGCCCTG-3 and 5-ACACGCAGGCCTCTTTGG-3.

3. RESULTS AND DISCUSSION
Of the 2917 genes included in the analysis, 432 genes or 15% were significantly differentially expressed in the brain of mature sneaker males compared with immature males ($p<0.01$; Electronic Appendix, supplementary table S1).

There is no other study to date that has examined brain gene expression profiles of plastic phenotypes in a vertebrate, although it has been found that even higher proportions of the genome (up to 40%) are differentially regulated in two honeybee behavioural phenotypes (Whitfield et al. 2003). For simplicity, genes found to be significantly upregulated in sneaker males relative to immature male brains are called ‘sneaker biased’, while genes found significantly upregulated in immature males compared with sneaker males are called ‘immature male biased’. As the genes represented on the array were randomly selected with regard to the tactics studied, we infer that roughly 15% of the genome varies in gene expression in the brain between the two male tactics. Of course, it is probable that some differences in genes expressed only in a few cells and/or at very low levels may be masked by examining whole brains. This would result in false negatives (e.g. a gene is found to be not differentially expressed between the two male types when in fact it is). For most of the differentially expressed genes, the differences between the two male types were very small (less than twofold) but highly significant ($p<0.01$); 85% of sneaker-biased genes and 94% of immature male-biased genes. Therefore, large changes in phenotype were accompanied mostly by subtle changes in transcription levels when considering the whole brain, demonstrating that significant but subtle variation and/or localized to specific brain regions could be captured using our approach.

According to life-history theory, trade-offs between reproduction and growth are expected to result in differential allocation of resources in divergent phenotypes (Stearns 1992). However, the mechanisms underlying trade-offs have rarely been examined and are usually treated as a black box (Stearns & Magwene 2003). We now know that developmental switches are largely determined by changes in gene expression (Carroll et al. 2001; Abouheif & Wray 2002). We therefore expect developmental trade-offs to be detectable at the transcriptional level. We hypothesize that this reproduction/growth trade-off should be reflected at the molecular level within specific biological processes. We test this prediction by classifying genes according to their biological function. We grouped the differentially expressed genes into 17 biological process categories following functional annotations (Rise et al. 2004) and tested whether these processes were preferentially activated in one phenotype relative to the other (Electronic Appendix, supplementary table S2; figure 3). Here, we find that the diverse molecular and cellular processes that are differentially activated in the two phenotypes are consistent with trade-offs between reproduction and growth. We find a significant disparity in the number of sneaker-biased and immature male-biased genes for many processes (exact binomial test; table 1): genes relevant in reproduction ($p<0.01$), feeding, reproduction and associated processes ($p<0.01$), growth hormone related, and neuronal signalling and neural plasticity ($p<0.001$), were sneaker biased significantly more often. Conversely, genes involved in transcription regulation, protein synthesis ($p<0.0001$), protein folding and maturation (chaperones), and neurodegeneration ($p=0.03$), were mostly upregulated in immature males, consistent with resource allocation towards somatic growth. In contrast, genes relating to blood, immune function, protein degradation, cellular metabolism and development were sneaker biased and immature biased in comparable proportions (table 1).

Finding reproduction-associated genes to be almost exclusively sneaker biased (table 1), although not extremely surprising, is a validation of our approach. For example, gonadotropins have previously been shown to be upregulated in sneaker males (Antonopoulou et al. 1999). Other biological processes were also in broad agreement with earlier findings that focused on candidate genes and hormones in Atlantic salmon and other species; growth hormone-related genes were found upregulated in sneaker males, which is consistent with findings that growth
hormone, notwithstanding its name, is elevated in plasma during sexual maturation of anadromous individuals of this species (Bjornsson et al. 1994). The circadian clock gene Bmal1, a basic helix-loop-helix transcription factor, was upregulated in sneakers, which is intriguing since in salmonids photoperiod and maturation are linked (Amano et al. 1995).

Other biological processes yielded some surprising insights, as many genes were in categories not previously implicated in this type of developmental divergence, such as neural plasticity and neural signalling. For example, genes involved in synaptic function and plasticity like neuronal pentraxin (a secreted immediate early gene product induced by synaptic activity (Reti et al. 2002)), synaptotagmin (triggers calcium-induced fusion of pre-synaptic vesicles with the plasma membrane (Zamponi 2003)), MHC class I proteins (activity-dependent synaptic remodelling (Huh et al. 2000)) and ependymin (hormonally induced glycoprotein involved in learning and memory formation in fishes (Schmidt 1995)) were all highly expressed in sneaker males compared with immature males. In addition, several genes involved in the synthesis of the gaseous neurotransmitter nitric oxide were upregulated in sneakers, which may be indicative of its role in the regulation of neuropeptides (Hawkes & Kar 2004). These differences in gene expression may reflect the cognitive demands of sneaking behaviour. Clearly, behavioural and neurobiological studies that explore some of these genes in the context of learning and cognition are needed.

Among genes relevant for feeding, reproduction and associated processes, somatolactin was sneaker biased, which confirms previous work (figure 1). The same was found for another pituitary hormone, prolactin, which has been implicated in salmon reproductive function and osmoregulation (Onuma et al. 2003). Another intriguing gene upregulated in sneaker males codes for POMC, a master regulator of feeding, reproduction, and a variety of other physiological processes (Raffin-Sanson et al. 2003). These microarray results have been validated using quantitative real-time PCR (figure 4). The growth category included insulin-like growth factor 2 receptor (IGF-2-R) upregulated in sneakers, which facilitates neuronal proliferation and differentiation and has been implicated in neuroendocrine regulation (Hawkes & Kar 2004). Genes involved in neurodegeneration were over-represented in immature males. Examples include β-1 integrin, which is involved in neural proliferation, differentiation, and apoptosis (Van der Flier & Sonnenberg 2001) and Brp44l, an apoptosis-regulating basic protein. Many immature male-typical genes are involved in transcription regulation (including histone H3), protein synthesis (ribosomal proteins), protein folding and maturation (chaperones) and protein degradation (proteasome-related genes), possibly indicating the increased somatic (and neural) growth in those animals compared with sneakers in the fall at the time of capture (figure 1).

Because considerable variation may exist (based on studies in other species and tissues (Oleksiak et al. 2002)),
we determined whether the broad range of environmental and genetic factors involved in an individual’s developmental trajectory, particularly in wild animals, result in variable gene expression profiles among individuals of the same phenotype. If individuals would cluster independently of phenotype, then this would reflect the null hypothesis of no correlation between expression profiles at the individual level. However, in this study, individual expression profiles clustered mostly according to their macroscopic phenotypes (e.g. sneaker males, immature males, immature females) in a correlation-based consensus tree, which groups individual expression profiles according to their similarity (figure 5). This indicates that despite the potential for large environmental effects in natural habitats, specific cellular and physiological processes are consistently activated in a phenotype. In the face
Table 1. Number of genes significantly upregulated in sneaker male brains and immature male brains for 17 biological process categories and significance level of exact binomial statistical test. (The five biological process categories with significant bias towards one male type in the number of genes represented in that category (in bold) were still all significantly biased after applying the Dunn-Sidak method of sequential correction of alpha level of significance (Sokal & Rohlf 1995).)

<table>
<thead>
<tr>
<th>biological process</th>
<th>sneaker males</th>
<th>immature males</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>circadian rhythm</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>growth hormone related</td>
<td>3</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>reproduction/others</td>
<td>10</td>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>reproduction</td>
<td>10</td>
<td>1</td>
<td>0.01</td>
</tr>
<tr>
<td>neural plasticity</td>
<td>19</td>
<td>3</td>
<td>0.0009</td>
</tr>
<tr>
<td>neural signalling</td>
<td>12</td>
<td>5</td>
<td>0.14</td>
</tr>
<tr>
<td>blood</td>
<td>6</td>
<td>3</td>
<td>0.51</td>
</tr>
<tr>
<td>immune function</td>
<td>6</td>
<td>3</td>
<td>0.51</td>
</tr>
<tr>
<td>development</td>
<td>21</td>
<td>19</td>
<td>0.87</td>
</tr>
<tr>
<td>growth</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>metabolism and energy</td>
<td>19</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>production</td>
<td>6</td>
<td>8</td>
<td>0.79</td>
</tr>
<tr>
<td>neuroprotection</td>
<td>9</td>
<td>13</td>
<td>0.52</td>
</tr>
<tr>
<td>protein degradation</td>
<td>4</td>
<td>7</td>
<td>0.55</td>
</tr>
<tr>
<td>protein folding and maturation</td>
<td>3</td>
<td>7</td>
<td>0.34</td>
</tr>
<tr>
<td>protein synthesis</td>
<td>5</td>
<td>48</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>neurodegeneration</td>
<td>0</td>
<td>6</td>
<td>0.03</td>
</tr>
</tbody>
</table>

of the continuous variation found in many internal and external traits relevant to this kind of developmental plasticity (Roff 1996; Nijhout 2003; West-Eberhard 2003), the developmental trajectories at both the molecular and physiological levels nonetheless produce two discrete and canalsized alternative phenotypes instead of a continuous spectrum.

Surprisingly, the cluster analysis showed that brain gene expression profiles of sneaker males were more similar to those of immature females than immature males. Although evidence at the molecular and cellular level is limited, this suggests that early maturation as a sneaker could be the default developmental pathway in salmon, with males that fail to reach the threshold size to develop into sneakers, the immature males of this study, maturation component of the divergence between these two male types. Delayed maturation and future sea migration, as exhibited by immature males, is often seen as the ‘default’ life-history tactic. However, our study suggests it may actually result from actively inhibiting the developmental trajectory towards becoming a sneaker male.

The goal of our study was also to lay the foundation for future expression profiling studies to determine whether the same molecular networks (or ‘modules’) are used across species in males showing the (plastic or genetically determined) sneaker tactic (Hartwell et al. 1999). Fishes
Figure 5. Cluster analysis of similarity of individual gene expression profiles of mature sneaker males 1–4, immature males 1–4 and immature females 1–4. Distance matrix based on Pearson correlation. Numbers at branch nodes in the consensus tree represent the percentage of trees that showed this clustering configuration of individuals when those trees were built using 1000 bootstrapped datasets (see §2).

are ideally suited for this comparative approach, as numerous species are known to exhibit such alternative male reproductive tactics (Taborsky 2001; Knapp 2003). Additionally, it has recently been shown that the same microarray platform can be used to study even relatively distantly related teleost species (Renn et al. 2004; Aubin-Horth et al. 2005). Combined with the existing research into the ultimate causes of phenotypic plasticity and behavioural ecology, functional genomics paves the way for similar analyses in a wide variety of organisms in order to uncover the molecular and cellular determinants and constraints underlying evolutionary trajectories of these complex plastic traits.

We thank Kitty Griswold, Sarah Annis and Keith Nislow for their assistance during sampling in the field and Yehoda Martei for laboratory assistance. We are grateful to Scott Edwards, Laura Garwin, Susan Renn, Brian Trainor and two anonymous reviewers for comments on earlier versions of this manuscript. N.A.H. was supported by a postdoctoral fellowship from Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) and an NSERC (Natural Science and Engineering Research Council of Canada) postdoctoral fellowship. C.R.L. is a Frank Knox Memorial Fellow at Harvard University and is supported by NSERC, Desjardins foundation and FQRNT graduate fellowships. Funding of this project was provided by the Olin Foundation Fellowship from the Atlantic Salmon Federation to N.A.H. and the Bauer Center for Genomics Research to H.A.H.

REFERENCES
Felsenstein, J. 2004 PHYLP (Phylogeny Inference Package) version 3.6b. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.


