JÉRÔME ST-CYR

ÉTUDE COMPARATIVE DU PROFIL D'EXPRESSION GÉNÉTIQUE ENTRE POPULATIONS DU GRAND CORÉGONE (*COREGONUS CLUPEAFORMIS*) À L'AIDE DE BIOPUCES À ADN

Mémoire présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en biologie
pour l'obtention du grade de Maître ès Sciences (M.Sc.)

DÉPARTEMENT DE BIOLOGIE
FACULTÉ DES SCIENCES ET DE GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

DÉCEMBRE 2007

© Jérôme St-Cyr, 2007

RÉSUMÉ

Un des supports théoriques de la biologie évolutive est la théorie de la radiation adaptative, qui stipule que la diversification des espèces est le fruit de l'action de la sélection naturelle divergente sur les populations, laquelle émerge des caractéristiques biotiques et abiotiques de l'environnement. Il existe plusieurs exemples de cas de radiation adaptative dans la littérature. Or, la majorité des études se penchent sur la variation phénotypique morphologique et sur l'histoire de vie. Ainsi y a-t-il relativement peu de données sur la variation physiologique, tant au niveau de l'organisme qu'au niveau moléculaire. De plus, la caractérisation des bases génétiques de traits phénotypiques complexes s'avère peu aisée et demeure élusive. Dans cette étude, nous avons tiré profit des nouvelles technologies de la génomique pour caractériser les bases génétiques de la variation phénotypique observée chez deux populations divergentes du grand corégone (*Coregonus clupeaformis*), l'écotype limnétique (*nain*) ainsi que l'écotype benthique (*normal*). À l'aide d'une biopuce à ADNc spécifique au Salmonidés, nous avons comparé les profils d'expression génétique des corégones nains et normaux issus de leur milieu naturel ainsi que d'un environnement contrôlé. Le parallélisme observé dans les patrons de transcription génétique entre deux lacs et le milieu contrôlé indique que la sélection naturelle a modulé la régulation de gènes candidats associés à des adaptations phénotypiques distinguant les deux écotypes du corégone. De plus, le compromis observé entre différents traits d'histoire de vie chez les corégones nains et normaux se reflète à travers le profil de transcription pour ces gènes candidats. Enfin, l'analyse de regroupements (*cluster analysis*) entre les patrons de transcription des gènes candidats suggère que les principales fonctions physiologiques impliquées dans la divergence adaptative des corégones nains et normaux sont sous le contrôle d'un seul ou de quelques gènes de régulation.

ABSTRACT

One important theoretical framework in evolutionary biology is the adaptive radiation theory, which states that species diversification is the outcome of diverging natural selection, a selective force stemming from biotic and abiotic characteristics of the environment and acting on populations. Many cases of adaptive radiation have been observed and documented in the literature, focusing mainly on morphological adaptations and on life-history traits. Thus relatively few data are available on the physiological processes of adaptation, both at the organismal and molecular levels. Moreover, characterizing the genetics of phenotypic variation for complex traits is not an easy task and remains elusive in most cases. In this study, we took advantage of technologies of the post-genomic era in order to characterize the genetic bases of adaptive phenotypic variation observed in two diverging populations of the lake whitefish (*Coregonus clupeaformis*), the limnetic (*dwarf*) and the benthic (*normal*) ecotypes. Using a salmonids-specific cDNA microarray, we compared the gene expression profiles of dwarf and normal whitefish individuals sampled from two natural populations and from a controlled environment. Parallelism in regulation patterns observed between two lakes and the control environment indicates that natural selection played a role in modulating the regulation for these candidate genes, in association with phenotypic adaptations distinguishing the two whitefish ecotypes. Also, the observed trade-off between life-history traits in dwarf and normal whitefish is reflected in the expression profiles for these candidate genes. Lastly, cluster analysis performed on transcription profiles for these genes suggests that the major physiological processes involved in the adaptive divergence of the dwarf and normal whitefish are under the control of one or few regulation genes.

Il n'aurait pas été possible d'effectuer tout le travail accompli lors du présent projet de maîtrise, ni d'obtenir des résultats d'une telle qualité au cours de cette étude, sans l'apport précieux de plusieurs acteurs, oeuvrant pour la plupart au sein du laboratoire Bernatchez.

Je voudrais d'abord remercier Louis Bernatchez de m'avoir accueilli dans son équipe en tant que professionnel de recherche. Qui aurait pensé à l'époque que ce nouvel emploi allait susciter chez moi un intérêt si vif qu'il allait précipiter mon retour aux études graduées! Je tiens aussi à remercier Nicolas Derome, qui a supervisé en grande partie les travaux de ce projet de maîtrise, de l'échantillonnage jusqu'à la rédaction, en passant bien sûr par les analyses statistiques.

Je me dois de remercier particulièrement Lucie Papillon, pour son aide précieuse concernant les questions informatiques, ainsi que pour son support constant et ses encouragements. De façon indéniable, l'arrivée à terme ainsi que la réussite de ce projet lui sont dues. Merci aussi à toute l'équipe du labo Bernatchez, que ce soit pour des discussions scientifiques et statistiques, pour l'aide dans le laboratoire, l'aide dans la rédaction anglaise et française, ou simplement pour avoir contribué à rendre ce parcours des plus agréable. Enfin, merci à mes parents et amis, pour leur support et leurs encouragements.

Ce projet visait donc à comparer la régulation des gènes chez deux écotypes sympatriques du grand corégone, et ce au niveau de deux tissus soit le cerveau et le foie. Le chapitre 1 de ce mémoire présentera les résultats obtenus pour le transcriptome du foie. Il s'agit d'un article intitulé: « The transcriptomics of life-history trade-offs in whitefish species pairs (*Coregonus* sp.) » qui sera publié dans la revue *Molecular Ecology* (*sous presse*). Les résultats obtenus pour le cerveau sont en cours d'analyse, et feront vraisemblablement l'objet d'une publication ultérieure.

Le soutien financier pour ce projet a été assuré par la Chaire de recherche du Canada en Génomique et Conservation des Ressources Aquatiques, ainsi que par le groupe interinstitutionnel de recherches océanographiques Québec-Océan.

«Rien n'est permanent, sauf le changement. »

Héraclite

TABLE DES MATIÈRES :

Résumé.....i
Abstract.....ii
Avant-propos.....iii
Liste des tableaux et figures.....vii

Introduction générale.....1

Évolution, radiation adaptative et sélection naturelle.....2
Colonisation post-glaciaire et espèces sympatriques.....3
Le cas du grand corégone.....3
Variations phénotypiques et environnement.....4
Bases génétiques de l'adaptation phénotypique.....5
Biopuces à ADN (*DNA microarrays*).....6
Biopuces à ADN en biologie évolutive.....7
Une biopuce pour les salmonidés.....7
Objectif principal.....8

Chapitre 1 : The transcriptomics of life-history trade-offs between whitefish species pairs (*Coregonus* sp.).....10

Abstract.....11
Introduction.....12
Material and methods.....16
Results.....21
Discussion.....25
Conclusion.....32
Acknowledgments.....33

Conclusion générale……46

Utilisation des biopuces en écologie et en évolution……47
Comparaison entre milieux naturels et contrôlé……48
Régulation des gènes et adaptation……49
Bases génomiques de la divergence adaptative……50
Perspectives futures……51

Bibliographie……54

Annexes……69

LISTE DES TABLEAUX:

Table 1: Number of gene clones showing significant differences in levels of transcription between sympatric dwarf and normal whitefish, and patterns of regulation shared among experimental groups.....34

Table 2: Parallel changes in gene expression between dwarf and normal whitefish from two natural lakes and controlled environmental conditions.....35

Table 3: Parallel changes in gene transcription observed between dwarf and normal whitefish in from two natural lakes only.....37

Table 4: Relative representation of functional groups for genes showing parallel patterns of regulation between sympatric dwarf and normal whitefish 39

Table 5: Proportion of up-regulated genes in dwarf and normal whitefish in regards to their potential involvement in life history trade-offs observed between growth (increased fecundity) vs. survival.....40

LISTE DES FIGURES:

Figure 1: Venn diagram showing the number of gene clones with significant differences in level of transcription between sympatric dwarf and normal whitefish in two natural lakes and a controlled environment.....41

Figure 2: Cluster trees for gene clones showing parallel directional patterns of transcription between sympatric dwarf and normal whitefish from two natural lakes, and from two lakes and a controlled environment.....42

Figure 3: Mean differences in transcription levels for gene clones showing parallel and non-directional patterns of regulation in whitefish from the two natural lakes and from the controlled environment.....45

ANNEXES:

Supplementary Table 1: Genes showing non-directional transcription differences in natural populations.....70

Supplementary Table 2: Genes showing non-directional transcription differences in natural populations and under controlled conditions.....75

INTRODUCTION GÉNÉRALE

Évolution, radiation adaptative et sélection naturelle :

L'étude des processus à l'origine de la diversification des espèces constitue l'un des aspects majeurs de la biologie évolutive. Un des concepts centraux de cette discipline est la théorie de la radiation adaptative (Huxley 1942; Mayr 1942; Lack 1947; Dobzhansky 1951; Simpson 1953; Schluter 2000; Bernatchez 2004), laquelle stipule que la divergence phénotypique est le produit d'une sélection naturelle divergente résultant de l'hétérogénéité de l'environnement et des ressources ainsi que d'interactions compétitives. On parle aussi de spéciation écologique lorsqu'une telle sélection naturelle divergente amène un isolement reproducteur entre les populations touchées (Rundle et al. 2000; Schluter 2001). Un des exemples les mieux connus de la radiation adaptative est celui des pinsons de Darwin dans l'archipel des Galapagos (Lack 1947; Grant & Grant 2002), pour qui l'abondance de ressources et la compétition entre individus ont favorisé la spécialisation de plusieurs espèces pour l'utilisation d'une ressource trophique particulière. Contrairement à la vision classique de la théorie de l'évolution faisant intervenir l'apparition de mutations spontanées dans le patrimoine génétique de populations isolées géographiquement, ce cadre théorique rend compte de l'importance des facteurs écologiques intervenant dans la diversification des formes de vies (Schluter 2000).

Un des enjeux de l'étude des processus évolutifs consiste à démontrer le rôle de la sélection naturelle dans l'apparition de la variation à l'origine des différentes adaptations phénotypiques des organismes à l'environnement (Rieseberg et al. 2002). Un indice du rôle de la sélection se retrouve lorsque des organismes plus ou moins apparentés développent des adaptations semblables de façon indépendante, dans des environnements comparables. On parle ici d'évolution parallèle (Schluter 1996; 2000). La radiation adaptative implique souvent des populations ou des espèces ayant évolué en parallèle dans des environnements similaires (Foster & Baker 2004). Ce phénomène a été documenté pour plusieurs espèces, par exemple chez les épinoches à trois épines (*Gasterosteus aculeatus*, e.g. Rundle et al. 2000), pour lesquels des vagues de colonisation répétées en eau douce ont favorisé l'apparition de variations phénotypiques convergentes dans plusieurs populations, lesquelles variations sont corrélées avec des caractéristiques de l'environnement (McKinnon et al. 2004). Ceci constitue un bon exemple de radiation adaptative, où de nouvelles opportunités écologiques ont amené l'apparition de variations phénotypiques adaptées à l'utilisation de nouvelles ressources (Mayr 1963; Schluter 2000).

Colonisation post-glaciaire et espèces sympatriques :

Ainsi, on retrouve au sein de plusieurs lignées évolutives de poissons des zones tempérées de l'hémisphère nord, des groupes d'espèces ayant divergé relativement récemment à l'échelle des temps géologiques, pour lesquelles les processus de divergence phénotypique et d'isolement reproducteur sont encore à l'œuvre. Ceux-ci constituent de très bons modèles pour l'étude des mécanismes à l'origine de la spéciation et de la diversification des espèces (Schluter 2000; Bernatchez 2004). C'est le cas de certains membres de la famille des Salmonidés, Gasterostéidés et Osméridés de l'hémisphère nord, dont la diversité serait le résultat de l'apparition de nouvelles opportunités écologiques et de la colonisation de nouveaux habitats suite à la fonte des glaciers, il y 10 à 15 milliers d'années (Taylor 1999). Ces évènements de colonisation post-glaciaires sont à l'origine de la formation de complexes de populations, voire d'espèces sympatriques, dont certains se caractérisent par la présence de deux formes ou « écotypes », l'un benthique et l'autre limnétique (Schluter & McPhail 1993). Ces paires d'écotypes sympatriques partagent quatre attributs suggérant l'occurrence d'un processus de sélection écologique : rapide mise en place de mécanismes d'isolement reproducteur, persistance des deux formes malgré un flux génique non négligeable, une ségrégation marquée de leur niche écologique respective ainsi que la viabilité et la fertilité des hybrides (Schluter 1996). De plus, ces systèmes sont intéressants car on y retrouve souvent le phénomène de parallélisme évoqué plus haut, lequel ajoute un solide argument en faveur d'une origine sélective des formes alternatives pour ces écotypes. Finalement, les barrières reproductives n'étant pas complètes entre les formes dans la plupart des cas, ces systèmes représentent de véritables laboratoires naturels pour l'étude des processus évolutifs à leurs stades précoces (Bernatchez 2004; Rogers & Bernatchez 2007).

Le cas du grand corégone :

Le grand corégone (*Coregonus clupeaformis*) est un Poisson téléostéen de la famille des Salmonidés, pour lequel on a pu observé l'apparition de deux sous-espèces ou « écotypes » se distinguant en une forme benthique dite « normale » et une forme limnétique dite « naine » (Fenderson 1964; Bernatchez & Dodson 1990; Pigeon et al. 1997). Ce phénomène a été observé de façon répétée et indépendante : on retrouve ces deux populations en sympatrie dans au moins six lacs du bassin versant de la rivière St-Jean (Nouveau-Brunswick, Canada et Maine; Bernatchez 2004). On a pu démontrer que des variations phénotypiques significatives

et héritables modulées par l'action de la sélection naturelle directionnelle existent entre les deux formes, notamment aux niveaux de traits reliés à l'histoire de vie (Chouinard et al. 1996), à la croissance (Rogers & Bernatchez 2005; 2007), au taux métabolique (Trudel et al. 2001), à l'utilisation des ressources trophiques (Lu & Bernatchez 1999) ainsi qu'aux comportements (Rogers et al. 2002). De plus, la corrélation entre le niveau de différenciation entre corégones nains et normaux et le degré de variation de leur niche respective (Landry et al. 2007) indique une forte association phénotype-environnement chez ces populations. Finalement, il semble qu'il y ait un isolement reproducteur post-zygotique partiel entre les deux formes, mis en évidence par la mesure d'une mortalité plus élevée que la normale chez les hybrides (Lu & Bernatchez 1998; Rogers & Bernatchez 2006). Par des méthodes de génétique quantitative, on a pu cartographier des groupes de liaisons entre marqueurs polymorphiques associés à une réduction du flux génique entre les corégones nains et normaux (Rogers et al. 2001; Campbell & Bernatchez 2004; Rogers et al. 2004), en plus d'estimer le nombre de loci ainsi que de mesurer leurs effets sur la variation phénotypique au niveau de plusieurs traits quantitatifs (QTL : *quantitative trait loci*; pour la croissance et l'activité natatoire; Rogers & Bernatchez 2005; 2007). Les résultats de ces études ont grandement contribué à l'élucidation de l'architecture génomique sous-tendant la divergence phénotypique entre ces deux espèces.

Variations phénotypiques et environnement :

Certains facteurs écologiques à l'origine de la divergence phénotypiques chez le grand corégone ont été documentés (Landry et al. 2007), ainsi que les différences relative à l'histoire de vie et aux caractères morphologiques adaptés à l'environnement et à l'utilisation des ressources (Bernatchez 2004). Traditionnellement, l'étude de la divergence phénotypique s'est attardée à la variation de caractères morphologiques, plus facile à observer et à quantifier (e.g. McKinnon & Rundle 2002; Fjeldsa et al. 2005, Grant & Grant 2006, Diegisser et al. 2007). Or l'adaptation ne se reflète pas toujours dans la forme des organismes (Hoekstra & Coyne 2007). D'autres traits phénotypiques, relativement à la physiologie par exemple, pourraient également être soumis à la sélection naturelle et être à l'origine de divergence entre les populations. D'après les résultats de différentes études, on peut présumer que des populations de poissons peuvent avoir divergé au niveau des différentes fonctions physiologiques dont la modulation leur permet d'augmenter leur succès reproducteur, au niveau des organes des sens (voir Kotrschal et al. 1998) ou des systèmes circulatoire et

respiratoire (McMorrow et al. 1997), du métabolisme énergétique (Trudel et al. 2001) ou de l'apparition de nouveaux comportements (Valdimarsson et al. 2000). D'autres traits phénotypiques relatifs aux stratégies de reproduction et aux traits d'histoire de vie sont aussi sujets à la variation issue de l'action de la sélection naturelle (Roff 1992). L'adaptation des organismes à l'environnement implique donc la variation au niveau d'une multitude de caractères phénotypiques autres que morphologiques.

Bases génétiques de l'adaptation phénotypique :

Il est nécessaire d'obtenir plus de données génétiques et moléculaires concernant la base de la variation phénotypique afin de répondre à des questions cruciales relatives à l'évolution et la diversification du monde vivant (Coyne & Orr 2004). Il existe maintenant plusieurs exemples pour lesquels l'action de la sélection sur un gène précis a été mise en évidence (voir Ford 2002). D'autres études ont été réalisées dans le but de comprendre les bases génétiques de différences phénotypiques (voir Orr 2001; Roff 2007), notamment chez diverses espèces d'insectes (voir Brakefield et al. 2003; Kronforst et al. 2006) et de plantes (voir Ehrenreich & Purugganan 2006). Des travaux ont aussi été effectués chez les poissons, par exemple sur le gène LDH-B (lactate déshydrogénase B) de *Fundulus heteroclitus* (voir Schulte 2001), les gènes d'opsine chez les cichlidés Africains (Terai et al. 2006), ainsi que pour un gène contrôlant le nombre de plaques osseuses chez l'épinoche à trois épines (Colosimo et al. 2005), une espèce pour qui les informations génétiques sont de plus en plus importantes en quantité et en précision (voir Cresko et al. 2007).

La génomique fonctionnelle s'affaire à établir un lien entre les transcrits génétiques, leurs fonctions aux niveaux moléculaire, cellulaire et organismique ainsi que leur effet sur le phénotype (Gilad & Borevitz 2006). Appliquée dans un contexte écologique et évolutif, cette discipline se penche sur l'effet des gènes sur le succès reproducteur des organismes (Feder & Mitchell-Olds 2003; Ouborg & Vriezen 2007). Ainsi il est possible d'évaluer l'effet de l'expression d'un gène précis ou d'un groupe de gènes sur le phénotype d'un organisme, par exemple un processus physiologique. D'autre part, il est possible, à partir d'une fonction physiologique donnée, de concentrer notre attention sur des gènes potentiellement impliqués dans l'élaboration de ce processus physiologique (Feder & Mitchell-Olds 2003). Lorsqu'un processus physiologique est soumis à une contrainte ou force sélective, les gènes qui le sous-tendent sont probablement les cibles de la sélection (Lee & Mitchell-Olds 2006). Il est

maintenant possible grâce aux outils de la génomique (ex. biopuces à ADN) de cribler le génome entier de populations divergentes afin d'identifier les transcrits dont la régulation semble sous contrainte sélective (Whitehead & Crawford 2006a) et donc d'identifier des gènes candidats sous-tendant l'expression de tout trait phénotypique mesurable.

Biopuces à ADN (DNA microarrays) :

Depuis le développement des outils de la génomique, telles que les biopuces à ADN (*DNA microarrays*), il est maintenant possible d'étudier l'expression, la régulation ou la fonction de milliers de transcrits à la fois, et ce à l'échelle du génome (Gracey & Cossins, 2003). Une biopuce à ADNc consiste en une plaque de verre – format lame de microscope – sur laquelle on imprime un nombre *x* de séquences d'ADN dérivées d'*ESTs* (*expressed sequence tags*) obtenus à partir de bibliothèques d'ADN complémentaire (Schena *et al.* 1995; Gibson 2002). Une lame peut contenir jusqu'à quelques dizaines de milliers de séquences, pouvant couvrir en totalité le génome de l'organisme à l'étude – selon l'information disponible et l'effort de séquençage effectué sur cette espèce (Brown & Botstein 1999). Celle-ci est alors utilisée comme sonde lors de l'hybridation avec un échantillon d'ARN, sur lequel une réaction de RT-PCR (transcriptase inverse) a été préalablement effectuée. L'échantillon d'ADNc ainsi obtenu est hybridé avec la biopuce, pour ensuite être marqué avec un fluorophore. On hybride habituellement deux échantillons à la fois sur une même lame, marqués avec deux fluorophores différents, de manière à pouvoir comparer les niveaux d'expression entres eux (Brown & Botstein 1999; Gibson 2002). À l'aide d'un logiciel d'analyse d'images, on obtient les données brutes, qui consistent en une multitude de points fluorescents d'intensités variables, représentant les niveaux d'expression de chacun des transcrits présents pour les deux individus comparés. Les données sont ensuite comparées entre elles et analysées à l'aide de différents tests statistiques (voir Cui & Churchill 2003). Cette technique est de plus en plus utilisée en sciences de la santé, notamment en génomique fonctionnelle pour caractériser l'effet de mutations, pour la découverte de nouveaux gènes et de leur fonction, pour l'élucidation de voies de signalisation moléculaires, en pharmacologie pour caractériser l'effet de nouvelles molécules ou de nouveaux médicaments, dans l'identification de maladies et comme test diagnostic (Hughes & Shoemaker 2001; Butte 2002; Mantripragada et al. 2004). Ses applications diverses en font un outil potentiellement très intéressant pour caractériser la variation au niveau de l'expression génétique en milieu naturel.

Biopuces à ADN en biologie évolutive :

Bien qu'encore rarissime jusqu'à tout récemment, l'utilisation de la technologie des biopuces en biologie évolutive s'est avérée très profitable, tel que l'avait entrevue certains chercheurs (voir Singh 2003). Une idée lancée il y a plus d'une trentaine d'années (Britten & Davidson 1969; King & Wilson 1975) veut que l'adaptation des organismes passe principalement par la régulation des gènes. Il semble en effet que la régulation de l'expression génétique soit un des facteurs 1) ayant le plus de sensibilité en termes de réponse à l'environnement, et 2) étant à l'origine de l'apparition de la variation phénotypique dans bon nombre de cas. Certains auteurs (e.g. Cheung & Spielmann 2002) ont également proposé de considérer le niveau d'expression d'un gène comme un trait phénotypique quantifiable. Finalement, plusieurs études ciblant des gènes associés à une variation phénotypique induite par l'environnement se basent sur la régulation différentielle de ces gènes candidats, plutôt que sur la variation génétique au niveau de la séquence codante en tant que telle (voir Kammenga et al. 2007). Cette approche est d'autant plus pertinente dans le cas d'organismes non modèles tels que le grand corégone, pour lesquelles il existe peu de données de séquences génomiques (Gracey & Cossins 2003; Rogers & Bernatchez 2007). Malgré ces obstacles potentiels, il est de plus en plus aisé de développer des biopuces, avec un minimum d'informations et de matériel, et ce même sans avoir accès à la séquence de génomes complets (Lee & Mitchell-Olds 2006).

Une biopuce pour les salmonidés :

Un groupe de chercheurs de l'Université de Victoria (cGRASP : Consortium for Genomic Research on All Salmons Project, Victoria, C.-B.) ont récemment développé une base de données EST ainsi qu'une biopuce à ADNc spécifique aux salmonidés (Rise et al. 2004; von Schalburg et al. 2005) contenant 16 006 gènes. Une description complète du contenu de la biopuce, incluant une liste de l'identité des clones représentés, des renseignements sur leur provenance ainsi que leur annotation fonctionnelle est disponible à l'adresse suivante : http://web.uvic.ca/cbr/grasp. Cette puce s'avère être un outil puissant pour l'étude de l'expression génétique chez le grand corégone, puisque des tests d'hybridations interspécifiques ont donné des résultats très satisfaisants, laissant croire en une grande similarité entre les ADNc orthologues du saumon atlantique (*Salmo salar*), de la truite arc-en-ciel (*Oncorhynchus mykiss*) et du grand corégone (*Coregonus clupeaformis*) (Rise et al. 2004). L'hybridation hétérologue peut donner des résultats significatifs pour des espèces

apparentées dont le temps de divergence peut aller jusqu'à 200 Ma (Renn et al. 2004; Rise et al. 2004; 2006).

Derome et collaborateurs (2006) ont publié les premiers résultats relatifs à la régulation différentielle de l'expression génétique entre les deux écotypes sympatriques du grand corégone. Cette étude conduite sur le transcriptome du muscle blanc d'individus nains et normaux de corégones provenant de deux lacs du bassin de la rivière Saint-John (Cliff Lake & Indian Pond, Maine USA) démontre l'origine sélective, à travers le parallélisme retrouvé dans les patrons de transcription génétiques, de la régulation de certains gènes potentiellement impliqués dans l'adaptation de la forme naine à un mode de vie limnétique, par le biais d'une augmentation de l'efficacité de l'activité musculaire et du métabolisme énergétique basal (Derome et al. 2006).

Objectif principal :

Ce projet de maîtrise vise la caractérisation des bases génomiques de la divergence phénotypique entre deux écotypes sympatriques du grand corégone. Il s'agit d'évaluer l'action potentielle de la sélection naturelle divergente sur le profil d'expression des gènes au niveau de deux tissus, soit le foie et le cerveau, chez des individus nains et normaux issus de deux populations naturelles ainsi que d'un milieu contrôlé. Le foie joue un rôle central dans plusieurs processus physiologiques, notamment en relation avec le métabolisme énergétique et la croissance (Tortora & Grabowski 1993), deux traits phénotypiques montrant une divergence marquée entre les deux formes. Des études précédentes ont démontré que ces variations ont une base génétique héritable (Rogers & Bernatchez 2007). Le cerveau est évidemment à la base du comportement. Il interprète tous les signaux provenant des différents organes des sens. De même, plusieurs organes, glandes, membres, et incidemment, plusieurs fonctions physiologiques sont sous le contrôle du cerveau (Tortora & Grabowski 1993; Kotrschal et al. 1998). Les corégones nains et normaux montrent des différences significatives et héritables au niveau de certains comportements liés à la prédation ainsi qu'à la cueillette de nourriture, soit dans la fréquence d'accélérations subites, la fréquence de changements de direction et dans le choix de la profondeur de nage (Rogers et al. 2002).

Cette étude permettra l'identification de gènes candidats (Feder & Mitchell-Olds 2003) dont la régulation est à la base de l'adaptation chez le corégone nain à sa niche limnétique.

L'hypothèse de base propose que le parallélisme observé dans l'adaptation phénotypique se reflète dans le profil d'expression génétique, en milieu naturel et en milieu contrôlé. Parallèlement, le volet génomique fonctionnelle de cette étude pourrait permettre d'identifier d'autres fonctions physiologiques évoluant sous l'influence de la sélection naturelle divergente chez le corégone, par le biais de l'analyse des fonctions occupées par ces gènes candidats ainsi que de leur patrons de regroupements à l'intérieur de réseaux co-régulatoires. Finalement, cette étude devrait contribuer à l'avancement des connaissances en ce qui a trait au mécanisme moléculaire du compromis (*trade-off*) observé entre deux traits reliés à l'histoire de vie, notamment le taux de croissance et la fécondité, chez les corégones nains et normaux.

Le chapitre 1 du présent mémoire présente les résultats obtenus pour le foie par le biais d'un article qui est présentement sous presse dans la revue *Molecular Ecology*. Les résultats obtenus pour le cerveau feront l'objet d'une publication ultérieure, en collaboration avec d'autres membres du laboratoire (Whiteley et al., *en préparation*).

CHAPITRE 1

THE TRANSCRIPTOMICS OF LIFE-HISTORY TRADE-OFFS IN WHITEFISH SPECIES PAIRS (*COREGONUS* SP.)

Abstract

Despite the progress achieved in elucidating the ecological mechanisms of adaptive radiation, there has been little focus on documenting the extent of adaptive differentiation in physiological functions during this process. Moreover, a thorough understanding of the genomic basis underlying phenotypic adaptive divergence is still in its infancy. One important evolutionary process for which causal genetic mechanisms are largely unknown pertains to life-history trade-offs. We analysed patterns of gene transcription in liver tissue of sympatric dwarf and normal whitefish from two natural lakes, as well as from populations reared in controlled environments, using a 16,006 gene cDNA microarray in order to: i) document the extent of physiological adaptive divergence between sympatric dwarf and normal species pairs, and ii) explore the molecular mechanisms of differential life history trade-offs between growth and survival potentially involved in their adaptive divergence. In the two natural lakes, 6.45% of significantly transcribed genes showed regulation either in parallel fashion (2.39%) or in different directions (4.06%). Among genes showing parallelism in regulation patterns, we observed a higher proportion of over-expressed genes in dwarf relative to normal whitefish (70.6%). Patterns observed in controlled conditions were also generally congruent with those observed in natural populations. Dwarf whitefish consistently showed significant over-expression of genes potentially associated with survival though enhanced activity (energy metabolism, iron homeostasis, lipid metabolism, detoxification), whereas more genes associated with growth (protein synthesis, cell cycle, cell-growth) were generally down-regulated in dwarf relative to normal whitefish. Overall, parallelism in patterns of gene transcription, as well as patterns of inter-individual variation across controlled and natural environments, provide strong indirect evidence for the role of selection in the evolution of differential regulation of genes involving a vast array of potentially adaptive physiological processes between dwarf and normal whitefish. Our results also provide a first mechanistic, genomic basis for the observed trade-off in life-history traits distinguishing dwarf and normal whitefish species pairs, wherein enhanced survival via more active swimming, necessary for increased foraging and predator avoidance, engages energetic costs that translate into slower growth rate and reduced fecundity in dwarf relative to normal whitefish.

Introduction

Unravelling processes that underlie population divergence and speciation is a crucial step towards elucidating the origin and maintenance of biodiversity (Howard and Berlocher 1998; Coyne and Orr 2004). Over the last decades, much progress has been achieved in elucidating the ecological mechanisms responsible for phenotypic divergence, particularly in the context of an adaptive radiation (Schluter 2000). In contrast, there has been less focus on several issues that are also relevant to a full understanding of how divergent populations adapt to different environments under the effect of natural selection, and ultimately evolve into biological species. First, the vast majority of studies on adaptive radiation have focussed on divergence in "external phenotypes", putting most emphasis on adaptations in morphological (see Gray and McKinnon 2006; Grant and Grant 2006) and life-history traits (reviewed in Schluter 2000, but see also Roff 2002). Comparatively, there has been little focus on elucidating the extent of adaptive differentiation in physiological functions during the process of population adaptation and divergence (but see Whitehead and Crawford 2005; Schulte 2007). Yet, the paramount importance of physiological adaptations in coping with different environments has been amply documented by comparative physiologists (e.g. see Willmer et al. 2000), albeit very seldom in the context of adaptive radiation *per se*. This was recently pointed out by Hoekstra and Coyne (2007): « ...it must be the case that many major evolutionary innovations and transitions involved changes that were not reflected in body forms ». Secondly, a thorough understanding of the genomic basis underlying phenotypic adaptive divergence is still in its infancy (MacCallum and Hill 2006; Mekel-Bobrov and Lahn 2006) although remarkable progress is being made in a few organisms, such as the threespine stickleback *Gasterosteus aculeatus* (Colosimo et al. 2005; Cresko et al. 2007), cichlid fishes (Terai et al. 2006), Darwin's finches (Schneider 2006), fruit flies (Matzkin et al. 2006, Laayouni et al. 2007), sunflowers (Edelist et al. 2006), monkeyflower (Streisfeld and Kohn 2005) and forest trees (Gonzàlez-Martinez et al. 2006).

One important evolutionary process for which causal genetic mechanisms is largely unknown pertains to life-history trade-offs, defined as negatively correlated responses to selection (Roff 2007). As expressed by Stearns and Magwene (2003), such trade-offs still represent "...black boxes located within theories that are much more explicit about mechanisms at the level of whole organisms that they are about mechanisms inside organisms". Many species that originated from an adaptive radiation differ in life-history traits potentially involving trade-

offs. In fish for instance, there are many cases of sympatric and parapatric occurrence of species pairs that strikingly differ in life-history traits. These generally involve a limnetic form, characterized by slower growth, shorter developmental time to reach maturity and shorter lifespan, and an alternative benthic form (Taylor 1999). In these species, fecundity increase with body size is achieved through a longer period of growth. On the other hand, the probability of surviving to reproduction will decrease with increased time-to-maturity (Roff 1992). For instance, Rennie et al. (2005) demonstrated that higher activity costs, accounting for a major proportion of fish energy budgets, traded-off against observed growth rate in wild populations of yellow perch (*Perca flavescens*). Although not a species pair, a parallel can nonetheless by drawn between this study on perch and the divergence between limnetic-benthic species pairs, which apparently involves differential trade-offs between fecundity and survival mediated through body size and developmental time. In other species however, different trade-offs may be in question, and these processes might involve different physiological, behavioral, morphological or ecological factors (Roff 2007).

It has been proposed for more than 30 years that changes in gene regulation may play a crucial role in driving rapid evolutionary changes under the effect of selection (Britten and Davidson 1969; King and Wilson 1975; Schulte 2001; Saetre et al. 2004). This hypothesis was first supported by the pioneering work of Powers and colleagues on the physiological adaptation of *Fundulus heteroclitus* to different thermal environments (reviewed in Powers and Schulte 1998). Recently, Stearns and Magwene (2003) proposed that in a genomic context, "trade-offs could be perceived as antagonistic pleiotropy representing conflicts between whole-organism function over the whole-genome patterns of gene expression which can be described by considering gene expression pattern in response to two physiological challenges" (e.g. growth vs. survival). Thus, studies that examine how quantitative variation in gene expression relates to phenotypic and quantitative variation in life-history traits would be of considerable interest (Roff 2007).

The development of microarray technologies, allowing the simultaneous detection of expression modulations at thousands of genes offers a powerful means of assessing the importance of evolutionary change in gene regulation involved in population divergence and adaptation (e.g. Olesiak et al. 2002, 2005; Bochdanovits et al. 2003; Singh 2003; Ranz and Machado 2005; Brodsky et al. 2005; Gilad and Borevitz 2006; Matzkin et al. 2006; Franchini and Egli 2006; Gonzàlez-Martinez et al. 2006; Street et al. 2006). Not only can gene

transcription be considered a phenotypic trait (Gibson 2002; Gracey and Cossins 2003), but microarray studies also represent a unique way of documenting evolutionary change at specific genes occupying distinct molecular functions, relative to biological processes involved in morphological as well as in physiological adaptations to natural environments (Ouborg and Vriezen 2007). Furthermore, surveying simultaneous transcriptomic activity of thousands of genes offers the opportunity to conduct cluster analyses on their expression profiles which in turn allows testing for functional groupings of gene transcripts, thus giving new insights on the architecture of regulation pathways underlying phenotypic divergence (Eisen et al. 1998). This has been little explored in the specific context of an adaptive radiation. Moreover, microarray technology potentially represents a powerful means for exploring the molecular mechanisms of trade-offs and searching for genes that affect life-history traits in opposite ways (antagonistic pleiotropy). This has been investigated in a few systems to date (Oakeshott et al. 2003; Bochdanovits and de Jong 2004), although not in vertebrates to our knowledge.

The lake whitefish species complex (*Coregonus* sp., Salmonidae) has contributed in many ways to the understanding of the genetic basis of evolutionary change in the course of an adaptive radiation (Bernatchez 2004). Following ice-cap retreat after the Wisconsin glaciations 15,000 years ago, recurrent and independent sympatric divergence of two reproductively isolated whitefish phenotypic forms occurred in many lakes, so-called "normal" and "dwarf" ecotypes, adapted to occupying the benthic and limnetic niches, respectively (Pigeon et al. 1997, Lu and Bernatchez 1999). Previous studies revealed that divergent natural selection led to the evolution of many phenotypic differences between them including morphological, behavioral, ecological and life-history traits that most likely represent adaptations towards exploiting distinct trophic resources (Fenderson 1964; Lu and Bernatchez 1999; Bernatchez et al. 1999; Rogers, Gagnon and Bernatchez 2002). Linkage mapping has been used to document the number and effects of quantitative trait loci (QTL) involved in controlling the expression of these adaptive traits (Rogers et al. 2007), and genome scans performed in natural populations provided evidence that directional selection is maintaining genetic divergence between sympatric dwarf and normal whitefish by restricting gene flow at these QTLs (Rogers and Bernatchez 2005, 2007). Recently, microarray studies provided further evidence for the role of directional selection in maintaining divergence between these sympatric ecotypes at the transcription level, mainly at genes involved in

swimming activity and energy metabolism in the white muscle tissue (Derome et al. 2006; Derome and Bernatchez 2006).

Several lines of evidence indicate that the divergence observed in whitefish species pairs might involve differential trade-offs between fecundity and survival. Dwarf whitefish mature as early as one year old, seldom exceed 20 cm in length and 100 g in weight and rarely live more than five years, whereas normal whitefish mature at an older age (greater than three years old), commonly exceed 40 cm and 1000 g, and can live up to 20 years (Fenderson 1964). Experimental work on swimming behaviour performed in a controlled environment showed that dwarf whitefish are more active swimmers (Rogers et al. 2002; Rogers and Bernatchez 2007). Moreover, higher metabolic rate (partly associated with the cost of swimming activity) and lower bioenergetic conversion efficiency (defined as growth rate/consumption rate ratio) are associated with slower growth and younger age at sexual maturity in dwarf whitefish (Trudel et al. 2001). These observations raise the hypothesis that the adaptive divergence and the evolution of distinct life-history strategies between dwarf and normal whitefish might involve differential trade-offs between fecundity and survival that is mediated through the higher energetic cost of occupying the limnetic relative to the benthic trophic niche. Thus, higher metabolic rate, more active swimming and reduced bioenergetic conversion efficiency may constrain available energy for growth and reproduction at older ages in dwarf whitefish.

In this study, we analysed patterns of gene transcription in liver tissue using cDNA microarrays in order to: i) further document the extent of physiological adaptive divergence between dwarf and normal whitefish, and ii) explore the molecular mechanisms of differential life-history trade-offs potentially involved in the adaptive divergence of the two species. Rise et al. (2006) reviewed the multiples functions of the liver in growth regulation in salmonids, suggesting that this organ likely plays a central role in a large array of physiological processes for which dwarf and normal whitefish are known to show heritable divergence, namely energy metabolism (affecting survival via swimming activity devoted to foraging and predator avoidance) and protein synthesis (affecting growth and fecundity), and likely constitutes a major target of directional selection driving adaptive divergence between them. Moreover, the higher food consumption rate of dwarf relative to normal whitefish (Trudel et al. 2001) should also impact genes involved in blood filtration and waste management functions of the liver (Tortora and Grabowski 1993). More specifically, we predicted that

differential trade-offs involving growth (and correlated fecundity) vs. survival between dwarf and normal whitefish should translate into differential patterns of gene transcription, with dwarf whitefish predominantly showing over-expression of genes associated with energy metabolism and under-expression at genes involved in protein synthesis and other functions assocated with growth in normal whitefish.

Material and methods

Sample collection

Sympatric dwarf and normal whitefish were collected in both Cliff Lake (46°23'51") and Indian Pond (46°15'25"), located in the Allagash basin (St. John River drainage), Maine, USA. Fish were sampled using gill nets during the growth season in late June 2003. Nets were pulled every 30 min, ensuring the fish were still alive prior to tissue collection, thus avoiding RNA degradation. Ten adult individuals (mean fork length for dwarf =23.1 cm SD=6.15 cm (Cliff Lake) ; 22.4 cm SD=2.1 cm (Indian Pond) ; normal =35.4 cm SD=4.9 cm (Cliff Lake) ; 33.0 cm SD=5.8 cm (Indian Pond)) were randomly collected in each population. Fish were euthanized with 0.001% Eugenol solution (Sigma-Aldrich) immediately prior to tissue collection. Liver tissue samples were frozen immediately in liquid nitrogen, and stored at -80°C. Fish used in this study are the same individuals as those used for the transcriptomics analysis of white muscle by Derome et al. (2006).

Dwarf and normal whitefish reared in a controlled environment were originally sampled in 1998 in Témiscouata Lake and Aylmer Lake (Québec) respectively, and reared in captivity at the LARSA facilities (Laboratoire Régional des Sciences Aquatiques, Université Laval, Québec). Family crosses were made in 2001 to generate pure F1 breeds. Both normal and dwarf lines were kept at all times under the same environmental conditions (water temperature, photoperiod, diet). In September 2004, 12 individuals of comparable developmental stage for each population were randomly selected, measured (mean fork length =dwarf : 22.9 cm SD=1.8 cm; normal : 28.9 cm SD=2.2 cm) then euthanized with a 0.001% Eugenol solution just prior to tissue extraction. Liver tissue samples were immediately frozen on dry-ice and stored at -80oC.

Microarray experiments

Total RNA extracts were obtained from 48 fish, including eight normal and eight dwarf individuals from Cliff Lake, Indian Pond and the controlled environment according to the TriZOL® Reagent extraction Protocol (Gibco BRL, Carlsbad, CA) as described in Roberge et al. (2006). RNA integrity was verified with a 2100 Bioanalyzer (Agilent, Palto Alto, CA). A total of 10 ug of RNA were retrotranscribed to cDNA using the Superscript-II protocol (Invitrogen Life Technologies, Carlsbad, CA) and probed on 16,006 gene cDNA microarrays developed for the Atlantic salmon (*Salmo salar*) by the cGRASP (consortium for Genomic Research on All Salmon Project ; Rise et al. 2004 ; von Schalburg et al. 2005) following the Array 50 kit protocol (Genisphere, Hatfield, PA ; adapted as described in Roberge et al. 2006 and Derome et al. 2006). Efficiency of hybridization on the cDNA array was measured for lake whitefish (see Rise et al. 2007) and is comparable in both normal and dwarf whitefish (Derome and Bernatchez 2006). Two samples (one dwarf and one normal) were differentially labelled by fluorescence (Cy3 and Alexa647) and probed simultaneously on each array. Dye swaps were performed to minimise biases inherent to the dyes uneven fluorescence intensity (Churchill 2002). Gene expression levels were quantified using a ScanArray Express scanner (Packard Bioscience, Wellesley, MA) and the QuantArray software (Perkin Elmer, Wellesley, MA).

Statistical analysis

A total of 3,842 genes had significant levels of transcription (fluorescence intensity higher than the mean of controls plus twice their standard deviation; Roberge et al. 2006) for both dwarf and normal whitefish in all three experimental groups and were considered for further analyses. This percentage of significantly expressed genes is comparable with previous studies conducted on whitefish with the salmon cDNA array (Derome et al. 2006; Derome & Bernatchez 2006) and is also consistent with the calculated efficiency of heterologous hybridization conducted on the salmon cDNA array with various salmonid species, including the lake whitefish (Rise et al. 2007). Transcription data were corrected for intensity-related bias using a regional R-LOWESS algorithm and analysed by ANOVA using the R/MAANOVA software package (Kerr et al. 2000) under a mixed-effect model (sample and dye: fixed, array: random) using a permutation-based *F*-test (*F3*, 1000 sample permutations) in order to detect significant differences (P-val <0.05) in level of gene transcription between dwarf and normal whitefish (Cui and Churchill 2003). For each experimental group, eight biological replicates were used to insure sufficient statistical strength in the estimation of

biological variance, leaving 6 degrees of freedom (df =8 biological samples – 2 treatments). The number of degrees of freedom for adequate microarray experimental designs should be 5 or more according to Churchill (2002). Changes in transcription levels were calculated as a D/N ratio using un-logged R-LOWESS normalized transcription data, where the mean transcription in dwarf individuals (D) was divided by mean transcription in normal individuals (N) for each gene. Genes showing at least 5% differences in transcription levels (D/N <0.95 or >1.05) and with a *P*-value <0.05 were considered differentially expressed and were kept for subsequent analyses.

The amount of false discoveries was accounted for by comparing samples from independent experimental groups (two lakes and a controlled environment group) for transcripts sharing patterns of regulation between dwarf and normal whitefish. Patterns of gene regulation were of two types: i) "parallel genes" applies to genes with D/N ratios either <0.95 or >1.05 in all compared groups, that is, either between the two lakes (when comparing natural populations only, see Results) or between the two lakes and the controlled environment (when comparing natural populations with a control group); ii) "non-directional genes" applies to significantly expressed genes but regulated in opposed directions between two lakes when comparing natural populations only (see Results), or regulated in parallel in two of the three groups and in the opposed direction in the third experimental group when comparing natural populations with a control group. The probability to find the same false positive gene is thus decreased with the increasing number of independent experimental settings (see Cui and Churchill 2003).

The SAS software (The SAS Institute, Inc., Cary NC, USA) was used to test for normality of distribution (Shapiro-Wilk) for variance in transcription levels, and for non-parametric tests on mean variance of gene transcription, in parallel versus non-directional genes. Additional statistical analyses regarding proportions of functional groups represented in the parallel gene samples were carried out using the R Commander software package (http://www.r-project.org,http://socserv.socsci.mcmaster.ca/jfox/Misc/Rcmdr/), where a Chi-Squared (X^2) test (exact binomial with 1 df) was used to estimate the probability of over-representation for each functional group, given their respective representation among the 3,842 significantly transcribed genes. To avoid bias caused by redundancy of several repeated ESTs, DNA sequences of multiple clones of the same genes were considered as one single gene for every repeated gene with significant difference in transcription levels. An exception was the

(CB491826) glyceraldehydes-3-phosphate dehydrogenase clone that failed to align with the other 6 GADPH spots (see Table 2) using the BioEdit Sequence Alignment Editor software (Hall 1999), and was considered to be a second independent gene in this case. Similarly, multiple clones of the same gene were considered as one single gene when estimating the probability of over-representation of up-regulated genes found in each whitefish species in the relevant functional groups (X^2-test, exact binomial with 1 df), to test the hypothesis of a genomic basis underlying the phenotypic trade-offs in life-history traits.

Cluster analysis

Normal and dwarf whitefish individuals were initially profiled using transcription datas for genes showing parallel patterns of transcription in the two lakes, and subsequently in all three experimental groups. Cluster trees were obtained by using the average-linking method (Sokal and Michener 1958, also described in Eisen et al. 1998) with pair-wise distances estimated from the Pearson correlation coefficient (Qu and Xu 2006) using the GeneSight software (BioDiscovery). Given that bias introduced by differential array fluorescence could potentially cause individuals paired on the same array to systematically cluster together, we used a conservative approach whereby unlogged R-LOWESS data were further normalized by dividing expression values from one individual by that of the other and vice-versa (ex: array #1 : D1 = dwarf #1 expression values / normal #1 expression values; N1 = normal #1 expression values / dwarf #1 expression values ; for each gene). In this way, gene expression values in the cluster trees do not represent absolute, but rather relative gene transcription levels. This does not affect consistancy in terms of directionality of regulation, nor does it affect the meaning of the clustering of genes or samples together. Moreover, repeated clones of the same gene were considered independent and were all included in this analysis, as listed in Tables 2 & 3.

Determination of gene functional groups

Spots on the cDNA array correspond to genes derived from EST library annotations using databases from GeneBank (described in Rise et al. 2004; updates available on the cGRASP webpage: http://woodstock.ceh.uvic.ca/estproj/index.cgi). Gene clones showing parallel patterns of regulation between the three experimental groups or between the two natural populations (Cliff Lake and Indian Pond) were classified into 12 functional groups using information provided by the cGRASP website (http://web.uvic.ca/cbr/grasp/array.html), the SwissProt/TrEMBL database (http://ca.expasy.org/sprot/), the NCBI browser

(http://www.ncbi.nlm.nih.gov/), the KEGG Pathway database (www.genome.jp/kegg/pathway.html), the SOURCE Database (http://genome-www5.stanford.edu:80/cgi-bin/source/sourceSearch), the EMBL Bioinformatic harvester (http://harvester.embl.de/) and completed with references from the literature. Functional groups (see Table 2 and 3) each include various biological processes (functions as annotated by cGRASP are in parenthesis) and are defined as follows : 1) blood and transport (plasma retinol-binding, transport, antifreeze protein); 2) cell cycle regulation (intracellular signaling cascade, protein amino acid phosphorylation, response to oxidative stress, regulation of GTPase activity, phosphoenolpyruvate-dependent sugar phosphotransferase system, nucleotide metabolism, apoptosis signaling); 3) cell structure (cytoskeleton organization and biogenesis); 4) detoxification pathways (detoxification of xenobiotic agents and metabolic by-products); 5) energy metabolism (electron transport, glycolysis, tricarboxylic acid cycle intermediate metabolism, tricarboxylic acid cycle, malate metabolism, nucleotide-sugar metabolism, digestion, proteolysis and peptidolysis, tyrosine metabolism, tyrosine catabolism, L-phenylalanine catabolism); 6) immunity (endocytosis, transport, antigen-binding, pathogen recognition, T-cell protease); 7) germ-line formation (transport, sexual reproduction, ovulation (sensu Mammalia), progesterone metabolism); 8) iron homeostasis (iron ion homeostasis, iron ion transport, negative regulation of cell proliferation); 9) lipid metabolism (transport, lipid transport, lipoprotein metabolism, regulation of cholesterol absorption, bile acid metabolism, lipid biosynthesis, steroid biosynthesis, prostaglandin metabolism, prostaglandin biosynthesis, regulation of circadian sleep/wake cycle); 10) muscle contraction (regulation of muscle contraction, muscle development, hydrogen transport); 11) protein catabolism (proteolysis and peptidolysis, protein metabolism, protein modification, ubiquitin-dependent protein catabolism); and 12) protein synthesis (protein biosynthesis, electron transport, protein folding, ribosome biogenesis).

Sequences from ESTs previously matched with unknown gene loci were submitted to the BlastN and BlastX browsers from NCBI. Significant results were returned for only one unidentified, parallel gene spot (CK 990291) which partly matched the beta actin mRNA sequence ($\geq$80% homology) for three species (*Dicentrarchus labrax, Epinephelus coioides* and *Onchorhynchus mykiss*). Accordingly, this gene clone was included in the cell structure functional group in the two lakes comparison (Table 3).

Results

Differences in gene transcription by group

The number of differences in gene transcription levels between dwarf and normal whitefish varied among lakes and the control group at LARSA, but was significantly higher than expected by chance alone in all cases (expected: 5%; X^2–test *P*-val $<2.2e^{-16}$) (Table 1). Fish from Cliff Lake showed the highest number of significant differences (*P*-val $<2.2e^{-16}$), with 793 gene clones showing at least a 5% increase or decrease in transcription level in dwarf vs. normal whitefish, whereas fish from Indian Pond and the control group had 502 and 540 differentially expressed gene clones, respectively. Thus, the proportion of genes out of 3,842 that were significantly expressed showing differences in transcription level between dwarf and normal whitefish for each experimental group was 21% for Cliff Lake, 13% for Indian Pond, and 14% or the control group.

Parallel patterns of gene regulation between dwarf and normal whitefish

Table 1 provides details regarding the number of genes showing parallel patterns of regulation between dwarf and normal whitefish. When comparing groups for shared genes under regulation, we expected by chance alone to find 0.21 x 0.13 x 0.14 =0.0038 (0.38 %) of genes showing patterns of regulation between dwarf and normal whitefish in all three experimental groups (Cliff Lake, Indian Pond and control at LARSA). Here, 1.98% of significantly expressed genes (76 gene clones) showed patterns of regulation between dwarf and normal whitefish in all three groups, which was significantly higher than expected by chance alone (X^2–test *P*-val $<2.2e^{-16}$). Of those 76 gene clones, 34 showed parallel patterns of regulation, with six being down regulated in dwarf compared to normal whitefish, with an average 23.50% decrease in gene transcription level (min: 12.5%; max: 47.0%), whereas 28 were up-regulated in dwarf compared to normal samples, with an average 39.70% increase in gene transcription level (range: 10.1% to 152.3 %). The other 42 genes (1.09%) showed non-directional patterns of regulation between dwarf and normal whitefish among all three experimental groups (Table 1).

When comparing whitefish from the two natural lakes only, we expected to find a proportion of 0.21 x 0.13 =0.027 (2.7%) of all significantly expressed genes showing regulation in both groups. We found that 6.45% of gene clones (248) showed patterns of regulation between dwarf and normal whitefish in both lakes, which was also significantly higher than expected

by chance alone (X^2–test P-val $< 2.2e^{-16}$). A total of 92 genes (2.39%) showed parallel patterns of transcription, which included the aforementioned 34 parallel directional plus 58 others. Of these 92 genes, 65 were up-regulated in dwarf compared to normal whitefish, with an average 29.52% increase of gene transcription level (range 9.84% to 177%) and 27 were down-regulated in dwarf compared to normal whitefish samples, with an average 18.7% decrease in transcription level (range: 5.42% to 65.62%). The other 156 genes showed non-directional patterns of regulation between dwarf and normal whitefish. The total number of genes with significant differences of transcription between dwarf and normal whitefish in each group, and shared among experimental groups are illustrated in Fig. 1. Details regarding the directionality of regulation are provided in Table 1. When considering each experimental group separately, proportions of up- and down-regulated genes between dwarf and normal whitefish were not different from a 50/50 distribution (X^2–test on proportions: P-val = 0.3269), whereas for genes with parallel patterns of regulation, the proportion of up-regulated genes in dwarf whitefish was significantly higher than 50% in all comparisons (P-val $<$ 0.001).

Fig. 2 presents cluster trees obtained for 92 genes showing parallel patterns of regulation among the two natural lakes (2a), as well as for 34 genes showing parallel patterns of regulation between dwarf and normal whitefish among all three groups (2b). The cluster analysis performed on fish from the two natural lakes perfectly grouped dwarf and normal whitefish separately, whereas in the three group comparison, two major sample clusters stood out: one comprising a majority of normal individuals (91% normal; 20 normal and 2 dwarf whitefish), and the other being mostly composed of dwarf whitefish (84.6% dwarf; 22 dwarf and 4 normal samples). In contrast, cluster analysis performed on a data set including genes showing both parallel and non-directional patterns of regulation between dwarf and normal whitefish were less discriminant, revealing four clusters with proportions of: 20% dwarf / 80% normal, 25% dwarf / 75% normal, 85% dwarf / 15% normal, and 80% dwarf / 20% normal (data not shown). Then, a cluster analysis performed using only non-directional gene data from the three experimental groups failed to reveal any meaningful clustering (that is grouping either by form or origin), with three clusters in proportions varying between 45% dwarf / 55% normal to 50% / 50% (data not shown).

Functional groups associated with phenotypic divergence

Significant grouping patterns also appeared in both gene clusters, where a clear distinction emerged between genes that were up-regulated in dwarf from those that were up-regulated in normal whitefish (Fig. 2a and 2b). Moreover, different clones of the same genes tended to cluster together most of the time (e.g. brain protein 44-like protein clones, Fig. 2a; GAPDH clones, liver carboxylesterase 22 clones, malate dehydrogenase clones, anionic trypsin II clones, Fig. 2b). Genes with related functions also tended to cluster together, as illustrated by the five main gene clusters defined in Fig. 2a. These consistent clustering patterns provide support for the reliability of the normalization method used to compare transcription profiles obtained from different experimental replicates together. Cluster 1 consisted of a group of 35 genes that were up-regulated in dwarf whitefish, 11 (31.4%) of which belonged to the energy metabolism group making it the best represented functional group in this cluster. Three out of four lipid metabolism genes also up-regulated in the dwarf whitefish grouped in cluster 2 as well as all three clones of malate dehydrogenase from the energy metabolism group. Cluster 3, 4 and 5 grouped genes up-regulated in the normal whitefish. Cluster 3 comprised nine gene clones, four of which pertained to protein synthesis and two to protein degradation functions. Cluster 4 contained seven genes from four different yet related functional groups (cell cycle, lipid metabolism, cell structure and energy metabolism; see Discussion for details); cluster 5 comprised six genes, three of which were classified in the immunity functions group, with another gene assigned to the blood and transport functional group (alpha-fetoprotein precursor) potentially involved in immune processes as well, by analogy of functions (see Discussion). Two additional small gene clusters included three proteases (elastase 2 and two anionic trypsin II gene clones) that were classified in the energy metabolism group, as well as three brain protein 44-like protein gene clones (cell cycle regulation) grouped with perforin and serum amyloid P-component precursor gene clones from the immunity functions group.

Table 2 provides full names and functional classifications of gene clones with parallel patterns of regulation among all three experimental groups, whereas Table 3 provides similar information for the remaining 58 gene clones showing parallel patterns of transcription in the two natural lakes only. A complete list of genes showing non-directional regulation is provided in the online Supplementary Table 1.

We compared the proportions of different functional groups represented in the data set, under the hypothesis that a significant fraction of genes under differential regulation should be

functionally related to the phenotypic divergence observed between dwarf and normal whitefish. After regrouping repeated gene clones when required (see Material and Methods), all but one of the 12 functional groups of genes showing parallel patterns of transcription among the two natural lakes were over-represented in the data set compared to proportions expected by chance (X^2 p-val <0.05 except for the 'muscle contraction' functional group p-val >0.05; Table 4). When including comparison with the controlled environment condition, 5 out of 8 functional groups present in nature and in the control group were also significantly over-represented in the data set compared to expectation by chance, based on their respective representation in the group of 3, 842 significantly expressed gene clones (Table 4).

Transcriptomic basis of trade-offs in life-history traits

A Chi-squared test (X^2) was conducted on the proportions of up-regulated genes to test if any of the different functional groups were significantly over-represented in the dwarf or normal whitefish, under the assumption that up-regulation of relevant transcripts would underly the observed trade-offs in life-history traits between fecundity and survival via growth and activity (Table 5). Significant up-regulation for a majority of genes involved in functions potentially associated with enhanced activity and survival, namely energy metabolism, lipid metabolism, iron homeostasis and detoxification, was observed in the dwarf whitefish, both in natural populations and all three groups comparisons (P-val <0.05; Table 5). Although less clear, a trend of up-regulation was observed for growth genes (protein synthesis, DNA synthesis, cell cycle) in the normal whitefish, but statistical significance was obtained solely for the protein synthesis group when comparing both natural populations and the control environment (P-val <0.0001; Table 5). Although not directly involved in growth-survival trade-off but possibly reflecting earlier sexual maturation in dwarf whitefish (see Discussion), up-regulation of genes involved in germ-line formation was also observed in dwarf whitefish (P-val <0.0001; Table 5), both in natural populations or when compared with the controlled conditions.

Differential variance in gene transcription between parallel and non-direcitonal genes

In natural populations, the variance in transcription for genes showing parallel patterns of regulation for both lakes (mean variance $=3.31$) was significantly lower than the variance observed for non-directional genes (mean variance $=5.78$) (one sided Wilcoxon test, P-val $=$ 0.0021). However, this reduced variance in parallel genes was not statistically significant when including genes from the control group. When compared with fish held in controlled

conditions, variance in transcription levels was significantly lower in fish from the two lakes than in fish from the controlled environment for every pattern of regulation (up-regulated in dwarf: mean =4.27 in lakes vs. 10.17 in control *P*-val = 0.0028 ; up-regulated in normal : 0.55 in lakes vs. 2.10 in control *P*-val =0.0339 ; parallel (up and down-regulated) : 3.62 in lakes vs. 8.75 in control *P*-val =0.0012 ; non-directional : 2.10 in lakes vs. 13.07 in control *P*-val <0.0001 ; differently expressed in lakes mean =2.79 vs. differently expressed in control mean =11.11 *P*-val <0.0001 ; one sided Wilcoxon test). Moreover, in genes showing parallel differences, the mean increase (D/N; %) or decrease (1 – D/N; %) in transcription levels in dwarf individuals was significantly lower in natural populations than in the controlled environment, with an average decrease of 19.5 % for down-regulated genes in Cliff Lake and Indian Pond vs. 31.6 % in the control (*P*-val = 0.0083, one-tailed *t*-test), and an average increase of 32.6 % for up-regulated genes in Cliff Lake and Indian Pond vs. 54% in the control (*P*-val =0.0014, one-tailed *t*-test) (Fig. 3a). The more pronounced differences in level of transcription between dwarf and normal whitefish in the controlled environment relative to those from natural lakes held true for both parallel genes (up-regulated and down-regulated taken together) as well as for non-directional genes (Fig. 3b).

Discussion

Patterns of gene regulation and parallel evolution of dwarf and normal whitefish

The first objective of this study was to investigate patterns of gene transcription in liver tissue using cDNA microarrays in order to further document the extent of physiological adaptive divergence between dwarf and normal whitefish. To this end, this study adds to that of Derome et al. (2006) performed on white muscle of the same individuals in several ways. First, this study was based on the analysis of gene transcription by means of a 16,006 gene cDNA microarray (von Schalburg et al. 2005), which allowed coverage of a much broader set of genes and functions relative to the 3,557 gene cDNA array that was available for the previous study. Of the 16, 006 ESTs sequences printed on the array, a total of 3,842 had significant levels of transcription in both dwarf and normal whitefish samples. Such a seemingly low number of significantly transcribed genes is expected, since this experiment consists in heterologous hybridization (see Renn et al. 2004) and since the chip includes transcripts derived from various tissues and developmental stages whereas we analysed one tissue at one life history stage only (von Schalburg et al. 2005).

Second, the use of the same individuals studied by Derome et al. (2006) allowed a rigourous comparison of the extent of parallelism in gene transcription of different tissues between dwarf and normal whitefish, and as such, to evaluate the possible role of directional selection in driving the evolution of transcription level of genes expressed in both tissues. The set of genes that we identified here, and for which transcription profiles are likely to have evolved under directional selection, adds to those identified in our previous study on white muscle tissue performed on dwarf and normal whitefish from the same natural populations (Derome et al. 2006). Overall, both studies yielded very comparable results, although the proportion of differentially expressed genes between dwarf and normal whitefish was higher in liver relative to the white muscle tissue. This is consistent with the fact that the fractional rate of protein synthesis in the liver is 40 times that measured in white muscle (Mommsen 2001). When considering only the two natural lakes in this study, 6.45% of all significantly transcribed genes showed differences in transcription levels in the liver (248 gene clones), of which 2.39% showed parallel patterns of regulation (92 gene clones). For white muscle, and under the same statistical criteria, 4.3% (51 gene clones) of all significantly transcribed genes showed differences in level of transcription, with 1.35% (16 gene clones) showing parallel regulation between sympatric whitefish ecotypes in both lakes (Derome et al. 2006). Also, when considering parallel genes in both studies, we observed a higher proportion of over-expressed genes in dwarf relative to normal whitefish (70.6% of 92 genes) similar to what was observed for white muscle (62.3% of 16 genes). Moreover, the level of fold change for parallel genes between dwarf and normal whitefish was comparable for both tissues; we observed an average 29.5% increase of transcription level (range: 9.8-177%) for over-expressed genes in liver of dwarf whitefish, whereas Derome et al. (2006) reported an average 47% increase of transcription level (range 10-235%) for over-expressed genes in white muscle of dwarf whitefish. In both tissues the level of expression in up-regulated genes in normal whitefish was lower than that observed in up-regulated genes in dwarf whitefish, being 18.7% for liver and 17% in white muscle.

Third, this study added the analysis of dwarf and normal whitefish maintained in identical, controlled environmental conditions. Even though inter-individual and inter-populational variation in gene expression has been reported to be heritable in fish (Whitehead and Crawford 2006, Roberge et al. 2007), differential regulation of genes is also certainly modulated by environmental variance, that may in turn influence phenotypic variation in response to different environmental conditions (Bochdanovits et al. 2003, West-Eberhart

2005, Cossins et al. 2006, Alvarez and Nicieza 2006). To our knowledge, however, there is no other published comparison of gene transcription analysed by microarrays between populations from both natural and controlled environments. Here, such comparison provided strong support to our interpretation regarding the predominant role of a genetic rather than environmental control on differential patterns of transcription observed at numerous genes between dwarf and normal whitefish in their natural environments. Thus, 76 gene clones (among which 34 were regulated in parallel) representing the same functional groups were differentially transcribed between dwarf and normal whitefish reared in controlled environment and sampled from natural lakes. This number of shared genes was much higher than expected by chance alone. Admittedly, there were also a number of genes that did not show parallelism in transcription between controlled and natural populations. However, this was anticipated since the number of genes remaining differentially expressed, in parallel or in a non-directional fashion across different environments is expected to decrease with the number of groups being compared. Also, dwarf and normal whitefish reared in controlled conditions originated from different populations than those from the natural lakes we studied. Consequently, it was expected that these would also show population-specific transcription profiles at some genes (Derome et al. 2006). Finally, fish maintained in controlled conditions are most certainly exposed to different selective constraints than those found in natural environments. Despite all these differences, the fact that we detected a high number of genes under parallel regulation in fish from both natural and controlled environments provides very strong support for a predominant genetic control for the differential patterns of transcription shown by these genes between dwarf and normal whitefish. These results also provide strong indirect evidence for the role of directional selection in shaping these patterns in nature, since the independent evolution of divergent phenotypes among closely related lineages, as evidenced in previous studies of lake whitefish (e.g. Pigeon et al.1997; Lu et al. 2001; Bernatchez 2004), is unlikely to be random, and is generally interpreted as being the consequence of natural selection (Schluter 2000).

Further evidence for the role of selection in shaping differential patterns of transcription comes from the comparison of variance between parallel and non-directional genes. Stabilising selection within population should exert a constraining action on the phenotypic response, which in the present case, is predicted to translate into reduced inter-individual variance of transcription for genes that are most likely to be under strong selective constraints (Fisher 1930). Here, we observed that the inter-individual variance of transcription for parallel

genes was almost half that measured for non-directionnal genes (see Results). This suggests that genes with parallel transcription profiles experience stronger selective pressure than non-directional ones. Thus, our results suggest both stabilising selection within population and directional selection between populations are acting synergistically, therefore enhancing differential patterns of transcription across species pairs, while reducing transcription variance within each population for genes most likely to be under selective constraints (Cheung and Spielman 2002; Gilad, Oshlack and Rifkin 2006).

That reduced inter-individual transcription variance may truly reflect the consequences of selective constraints acting on directional parallel genes in nature is further supported, although indirectly, by the comparison with transcription variance observed in controlled conditions. Fish reared in controlled conditions for three generations have certainly been less subject to selective pressures, both during early life history stages (as evidenced by high survival rate; Lu and Bernatchez 1999), and throughout their life cycle vis-à-vis pressures associated with competition for resources (since always fed *ad libidum*), diseases (no mortality associated with pathogens has been recorded at LARSA; L.B. personal observation), and predation. Theory would predict that these relaxed conditions should allow for a higher range in phenotypic responses that could be measured by an increased variance in inter-individual mean gene expression levels from controlled vs. natural environments (Fisher 1930). Our results supported this prediction; despite the fact that dwarf and normal whitefish lines maintained in the laboratory were initiated with a maximum of 40 breeders each (Lu and Bernatchez 1999), inter-individual variance observed for parallel directional genes was more than twice higher in controlled relative to natural conditions (8.75 vs. 3.62, $P = 0.0012$).

Transcriptomic basis of trade-offs in life-history traits

Our second objective was to investigate how differential patterns of transcription between dwarf and normal whitefish could inform on the molecular mechanisms of life-history trade-offs potentially involved in the adaptive divergence of these species pairs. Based on their strikingly different life-history strategies, we predicted that differential trade-offs involving growth (and correlated fecundity) vs. survival between dwarf and normal whitefish should translate into predictable differences in patterns of gene transcription. Our results followed the general trend predicted under the hypothesis of trade-offs involving i) survival via overall enhanced standard metabolism (for instance associated with increased feeding rate; Trudel et al. 2001) and swimming activity to occupy the limnetic niche in dwarf relative to normal

whitefish; and ii) increased growth rate (and associated increased fecundity) in normal relative to dwarf whitefish. Dwarf whitefish consistently showed significant over-expression of genes potentially associated with survival through increased activity (including genes involved in energy metabolism, iron homeostasis, lipid metabolism and detoxification; Table 5). Thus, when comparing normal and dwarf whitefish from natural and controlled environment, 8 out of the 9 energy metabolism genes were overexpressed in parallel in dwarf relative to normal whitefish. Genes involved in lipid metabolism and detoxification mechanisms were also generally overexpressed in dwarf whitefish, indicating enhanced metabolic activity. It is noteworthy that we also observed significant over-expression of genes involved in germ-line formation in dwarf relative to normal whitefish, despite the fact that fish of both species were of comparable age (data not shown). This is consistent with the observation that dwarf whitefish become sexually mature at an earlier age. In contrary, genes more likely to be associated with growth, such as those involved in protein synthesis, cell growth and cell cycle promoting-genes were generally down-regulated in dwarf relative to normal whitefish. In the natural vs. controlled environment comparison, genes involved in various stages of protein synthesis pathways were overexpressed in parallel in normal whitefish (Table 5), thus corroborating previous observations on enhanced growth rate for this species (Rogers and Bernatchez 2005).

In the following sections, we discuss possible (hypothetical) functions of specific candidate genes that are the most likely to be involved in the differential life-history trade-offs distinguishing dwarf and normal whitefish species pairs. In doing so, one should bear in mind that a thorough appraisal of such functional links will require detailed mechanistic studies.

Dwarf whitefish are known to display more active swimming behaviors (Rogers et al. 2002) owing to their limnetic foraging habits and predator avoidance (Kahilainen and Lehtonen 2002). Accordingly, when comparing dwarf and normal whitefish populations from natural and control populations, we found over-expression of genes potentially associated with enhanced swimming activity in dwarf whitefish. Thus, relative to the normal, dwarf whitefish were prone to overexpression of genes playing key roles in energy producing mechanisms such as glycolysis (GDPH, aldolase A and B) and the TCA cycle (malate dehydrogenase), as well as genes facilitating transport of macromolecules such as lipids towards cellular target compartments such as AFABP (Glatz et al. 2003). We also observed in dwarf whitefish parallel up-regulation for three genes related to other biological processes involved in the

energy metabolism network (fructose-biphosphate aldolase B, malate dehydrogenase, GDP-L-fucose synthetase), as well as for three genes involved in lipid metabolism and transport (beta-2-glycoprotein I, gastrotropin, thyroid hormone-inducible hepatic protein). Conversely, three other lipid metabolism genes (apolipoprotein A-I-1, A-IV and CII) involved in lipid absorption, as well as one gene involved in the oxydative phosporylation of fatty acids (NADH-ubiquinone oxidoreductase chain 1) were down-regulated in dwarf whitefish, which would be consistent with an enhanced exportation of lipidic products towards energy demanding systems such as skeletal muscles. This hypothesis of enhanced lipid export activity in dwarf as opposed to normal whitefish liver tissue is supported by cluster analyses of transcription profiles for fish from natural populations (cluster 4, Fig. 2a). Thus, the aforementioned genes involved in lipid absorption and oxidative phosphorylation were grouped with genes potentially involved in vacuole formation and endocytosis (polyposis locus protein 1 homolog) as well as cytoskeleton-mediated vacuolar transport (Unknown CK990291, a beta actin homolog). Moreover, the clustering of these genes suggests that they are part of a common co-regulation network (Quackenbush et al. 2003).

Up-regulation of genes involved in detoxification pathways (glutathione-S-transferase and carboxylesterase 22, aldehyde dehydrogenase 9 family member A1-like 1, plasma glutathione peroxidase) observed in dwarf whitefish could also potentially be associated with higher metabolic activity and more active swimming behaviour. Thus, the higher energetic demands and metabolic activity would tend to generate a proportionally larger amount of metabolic by-products and wastes that would need to be detoxified in order to maintain proper muscle activity and avoid complications caused by oxydative stress (Claiborne 1998). For example, glutathione-S-transferase is known to play a role in protecting the plasma membrane against oxydating agents (Bioinformatic Harvester: http://harvester.embl.de/harvester/P464/P46439.htm). Another gene of interest also up-regulated in dwarf whitefish is ARF GTPase-activating protein GIT2, which is known for its possible involvment in DNA repair and apoptosis (Premont et al. 2000). Cells subjected to increased metabolic activity would tend to be faced with a greater production of metabolic by-products and free reactive oxygen species via an augmented respiratory activity in the mitochondria, which could potentially be DNA damaging (Mommsen 1998).

Dwarf whitefish are also characterized by a higher food consumption rate, perhaps to compensate for their less efficieint food conversion efficiency, itself linked to higher energy

demands for basal metabolism and/or swimming activity (Trudel et al. 2001). In turn, higher consumption rate should translate in differential transcription profiles between dwarf and normal whitefish for genes involved in catabolic functions. Accordingly, we observed parallel up-regulation in dwarf whitefish of three genes involved in the breakdown of proteins (trypsin II, homogentisate 1,2-oxygenase, elastase 2), a catabolic activity associated with lysosomes and digestive functions.

Other parallel up-regulated genes in dwarf whitefish could also be associated with adaptations to the limnetic niche. For instance, the heme oxygenase and two ferritin heavy chain gene clones, here classified in the iron homeostasis network (Bioinformatic Harvester : http://marvester.embl.de/marvester/O702/O70252.htm), could provide a more efficient way of ensuring oxygen transport by expressing higher levels of heme cleavage enzymes and keeping more iron ions readily available for oxygen transport (Rise et al. 2006). This would be consistent with the overall higher energetic demands of enhanced swimming and foraging activity in dwarf compared to normal whitefish. In addition to its potential role in respiratory functions, heme oxygenase is also associated with oxydative stress and inflamatory response (Stanford et al. 2003), consistent with enhanced activity and stress in the dwarf whitefish.

Two additional genes potentially associated with enhanced swimming activity were up-regulated in parallel in dwarf whitefish: troponin I, which is involved in the regulation of skeletal muscle contraction (UniProtKB/Swiss-Prot : http://ca.expasy.org/uniprot/P13412), and the prostaglandin-H2 D-isomerase involved in the hormonal-dependent regulation of muscle contraction. This latter gene could play a role in a more efficient blood circulation of nutrients and energy molecules associated with increased metabolic demand, and is also known for its role as scavenger for harmful hydrophopic molecules (UniProtKB/Swiss-Prot : http://ca.expasy.org/uniprot/O09114 ; http://ca.expasy.org/uniprot/P22057).

Four genes involved in protein synthesis (*O.mykiss* calreticulin and 60S ribosomal protein L9, protein disulfide-isomerase A4 precursor and peptidyl-prolyl cis-trans isomerase B precursor) were up-regulated in normal whitefish, which would be consistent with their higher growth rate relative to dwarf whitefish (Rogers and Bernatchez 2005). The level of DNA synthesis may also be indicative of cell-cycle progression, and can therefore be associated with growth rate at the cellular and molecular level. Thus, two genes involved in nucleoside phosphate metabolism (nucleoside diphosphate kinases A and B) were up-regulated in normal whitefish,

suggesting a higher DNA synthesis attributable to genome replication and cell division, consistent with an overall higher growth rate in normal whitefish. Up-regulation of three clones of the 'brain protein 44' gene in normal whitefish, a gene whose putative functions pertain to central nervous system development and differentiation via apoptosis regulation (UniProtKB/Swiss-Prot : http://ca.expasy.org/uniprot/BR44L_RAT), is also consistent with higher growth rate and developmental activity in this species. Finally, up-regulation of the polyposis locus protein 1 gene (UniProtKB/Swiss-Prot : http://ca.expasy.org/uniprot/Q60870) in normal whitefish, which is a putative membrane protein involved in cell signaling, membrane trafficking and indirectly in cellular division and cell cycle regulation, could hypothetically be associated with a higher responsiveness of cells to various growth signal. The role of the polyposis locus protein 1 gene, a human TB2 homolog, has been studied in yeast (*YOP1*) by Calero et al. (2001) who demonstrated that overexpression of *YOP1* resulted in an accumulation of internal cell membrane and a block in membrane traffic. This accumulation of internal membrane accompanied by a reduction in cellular membrane surface and restriction in membrane traffic may be associated with early stages of mitosis (Boucrot and Kirchhausen 2007), and is consistent with an overall enhanced cell cycle activity and cellular division associated with higher growth rate observed in the normal whitefish. Finally, some other genes did not fit the pattern predicted by the trade-off hypothesis. For instance, expression profiles of two cell cycle regulation genes (ARF GTPase-activating protein, Table 2; XP8, Table 3) were not consistent with the hypothesis of accelerated cellular growth and replication expected in normal whitefish in association with their higher growth rate.

Conclusion

Parallelism in gene transcription, as well as patterns of inter-individual variation across controlled and natural environments, provide strong indirect evidence for the role of selection in the evolution of differential regulation of genes that involve a vast array of potentially adaptive physiological processes between dwarf and normal whitefish, including energy metabolism, protein synthesis, cell cycle regulation, cell proliferation and differenciation, detoxification pathways, transport in the plasma, cellular intake of various nutrients and molecules, and immunity. Moreover, cluster analyses reveal significant grouping of genes with similar or related functions together, suggesting that clustered genes belong to common pathways that could share regulational effectors (Eisen et al. 1998; Ihmels et al. 2005). The fact that a relatively small number of clusters were observed also raises the hypothesis that the

transcription regulation of these genes could be under the control of one or few genes with major pleiotropic effects on the genome. The nature of changes in DNA sequences underlying such differences between dwarf and normal whitefish remains elusive at this point. However, even if the actual coding sequence of functional genes was differentiated between dwarf and normal whitefish ecotypes, the overexpression we observed in dwarf for most of the parallel changes would suggest that their corresponding alleles would be functionally more similar together, compared to those found in the normal populations. This, in turn, would further support the evolution of adaptive divergence at the genetic level between dwarf and normal whitefish. Notwithstanding the apparently contradictory results observed for a few genes, our results also provide a first mechanistic genomic basis for the observed trade-off in life-history traits distinguishing dwarf and normal whitefish species pairs, wherein enhanced survival via a more active swimming, necessary for increased foraging and predator avoidance, engages energetic costs that translate into slower growth rate and reduced fecundity in dwarf relative to normal whitefish.

Acknowledments:

The authors are grateful to D. Basley (Maine Department of Inland Fisheries and Wildlife), R. St-Laurent and D. Campbell for their invaluable help in fish sampling, S. Higgins and the LARSA staff, and B.F. Koop and W.S. Davidson at cGRASP for their collaboration in this project. Special thanks also go out to C. Roberge for his insights on cluster analyses, J. Jeukens for RNA integrity analysis, L. Papillon for her precious help in graphical editing and S. McCairns for constructive comments on ealier versions of the manuscript. We are also grateful to J. Slate and two anonymous referees for their helpful comments on the previous version of this manuscript. This research was financially supported by the Natural Sciences and Engineering Research Council of Canada to L.B. (Discovery Grant and EWR Steacie supplement). The supporting agency had no role in planning or any other aspect of this study.

Table 1. Total number of gene clones showing significant differences in level of transcription between sympatric dwarf (D) and normal (N) whitefish in each experimental group, as well as detailed number of gene clones with parallel (D/N >1 : up-regulated in dwarf ; D/N <1 : up-regulated in normal) and non-directional patterns of transcription in comparisons among experimental groups.

		Genes shared between two groups:				Three groups:
Groups:	Total	Regulation pattern:	Cliff Lake	Indian Pond	Control	Cliff Lake and Indian Pond
Cliff Lake	793	D/N > 1	354	65	84	
		D/N < 1	439	27	53	
		non-directional	-	156	50	
Indian Pond	502	D/N > 1		244	64	
		D/N < 1		258	36	
		non-directional		-	61	
Control	540	D/N > 1			298	28
		D/N < 1			242	6
		non-directional			-	42

Table 2. Parallel changes in gene expression between dwarf and normal whitefish from two natural lakes and controlled environmental conditions

Functional group: (1)	EST clone number and gene name: (2)	Transcription ratio: (3)					
		Cliff Lake	Indian Pond	Control	*P*-CL (4)	*P*-IP (4)	*P*-ctl (4)
Cell cycle regulation	CB517934 [GO] [Q9JLQ2] ARF GTPase-activating protein GIT2	1.37	1.36	1.32	0.0038	0.0012	0.0149
	CB492176 [GO] [Q60870] Polyposis locus protein 1 homolog	0.86	0.87	0.71	0.0220	0.0322	0.0146
Detoxification	CA057214 [GO] [Q64176] Liver carboxylesterase 22 precursor	1.35	1.23	2.24	0.0019	0.0051	0.0001
	CB496876 [GO] [Q64176] Liver carboxylesterase 22 precursor	1.33	1.17	2.16	0.0030	0.0174	0.0009
	CB496493 [GO] [P48774] Glutathione S-transferase Mu 5	1.13	1.63	1.39	0.0258	0.0001	0.0415
	CB497579 [GO] [P48774] Glutathione S-transferase Mu 5	1.10	1.16	1.73	0.0347	0.0182	0.0081
Energy metabolism	CB516178 [GO] [P07146] Anionic trypsin II precursor	1.55	2.52	2.25	0.0096	0.0000	0.0020
	CB515463 [GO] [P05208] Elastase 2 precursor	1.23	2.24	1.56	0.0208	0.0001	0.0097
	CA045033 [GO] [P07146] Anionic trypsin II precursor	1.24	2.14	1.45	0.0271	0.0001	0.0017
	CA062911 [GO] [P05064] Fructose-bisphosphate aldolase A	1.21	1.18	1.28	0.0151	0.0085	0.0235
	CB502483 [GO] [Q91Y97] Fructose-bisphosphate aldolase B	1.19	1.42	1.36	0.0094	0.0007	0.0205
	CB497681 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.13	1.46	1.71	0.0459	0.0027	0.0067
	CA768062 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.16	1.30	1.61	0.0080	0.0028	0.0019
	CB493574 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.23	1.18	1.47	0.0081	0.0101	0.0033
	BU965756 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.31	1.23	1.32	0.0076	0.0107	0.0128
	CB498361 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.23	1.28	1.33	0.0072	0.0032	0.0205
	CB514460 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.26	1.21	1.33	0.0073	0.0084	0.0223
	CB491826 [GO] [P16858] Glyceraldehyde-3-phosphate dehydrogenase	1.18	1.11	1.39	0.0345	0.0421	0.0112
	CA055883 [GO] [O09173] Homogentisate 1,2-dioxygenase	1.28	1.33	1.41	0.0051	0.0026	0.0044
	CB493498 [GO] [P14152] Malate dehydrogenase, cytoplasmic	1.34	1.20	1.47	0.0013	0.0151	0.0422
	CB518115 [GO] [P14152] Malate dehydrogenase, cytoplasmic	1.25	1.10	1.63	0.0074	0.0084	0.0195
	CA062141 [GO] [P43023] Cytochrome c oxidase polypeptide VIa-heart	0.87	0.78	0.72	0.0339	0.0036	0.0100
Germ-line formation	CA062348 [GO] [P51658] Estradiol 17-beta-dehydrogenase 2	1.12	1.20	1.28	0.0256	0.0146	0.0341
	CB496948 [GO] [O09114] Prostaglandin-H2 D-isomerase precursor	1.20	1.13	1.71	0.0132	0.0348	0.0012
Immunity	CA037686 [NR] [AAM73701] C1q-like adipose specific protein	1.40	1.13	1.21	0.0011	0.0456	0.0270
Iron homeostasis	CB515893 [NR] [O73688] Heme oxygenase	1.83	1.10	1.13	0.0001	0.0266	0.0103
Lipid metabolism	CK990220 [GO] [P04117] Adipocyte Fatty acid-binding protein	1.28	1.11	1.70	0.0032	0.0186	0.0021
Protein synthesis	CA063352 [GO] [P24369] Peptidyl-prolyl cis-trans isomerase B precursor	0.84	0.76	0.53	0.0244	0.0040	0.0030
	CA048973 [GO] [P08003] Protein disulfide-isomerase A4 precursor	0.75	0.76	0.69	0.0036	0.0065	0.0042
Unknown	CA063623 Unknown	1.83	1.42	1.22	<0.0001	0.0005	0.0211
	CB509889 Unknown	1.23	1.24	1.83	0.0190	0.0056	0.0072
	CB498458 Unknown	1.28	1.20	1.60	0.0072	0.0180	0.0393
	CK990521 Unknown	0.79	0.75	0.66	0.0047	0.0024	0.0016
	CA045102 Unknown	0.80	0.83	0.79	0.0073	0.0124	0.0175

(1) : Functional groups were defined as described in the Material and Methods section;
(2) : Each EST clone number represents a single EST sequence;
(3) : Transcription ratio corresponds to mean dwarf expression level divided by mean normal expression level (D/N);
(4) : *P*-CL, *P*-IP and *P*-ctl correspond to permutated *P* values for Cliff Lake, Indian Pond and the control conditions, respectively, (ANOVA, *F*3 test, 1000 permutations);

Table 3. Parallel changes in gene transcription observed between dwarf and normal whitefish from two natural lakes only.

Functional group:(1)	Gene clone number and name:(2)	Transcription ratio: (3)			
		Cliff Lake	Indian Pond	*P*-CL(4)	*P*-IP(4)
Blood and transport	CB501248 [GO] [Q00724] Plasma retinol-binding protein precursor	1.41	1.23	0.0015	0.0098
	CB509509 [GO] [Q00724] Plasma retinol-binding protein precursor	1.34	1.12	0.0017	0.0429
	CB498410 [NR] [AAT95404] Cod1AFP antifreeze protein	1.34	1.15	0.0030	0.0162
	CB493984 [GO] [P02772] Alpha-fetoprotein precursor	0.87	0.89	0.0211	0.0354
Cell cycle regulation	CA043387 [NR] [BAB33387] XP8	1.33	1.15	0.0022	0.0325
	CB510934 [GO] [P15532] Nucleoside diphosphate kinase A	0.85	0.82	0.0254	0.0109
	CK991305 [GO] [Q01768] Nucleoside diphosphate kinase B	0.84	0.83	0.0120	0.0098
	CA059018 [GO] [P63030] Brain protein 44-like protein	0.83	0.78	0.0114	0.0035
	CB496707 [GO] [P63030] Brain protein 44-like protein	0.84	0.87	0.0156	0.0255
	CA059976 [GO] [P63030] Brain protein 44-like protein	0.88	0.88	0.0366	0.0324
Cell structure	CA052792 [GO] [P05784] Keratin, type I cytoskeletal 18	1.15	1.13	0.0304	0.0319
	CK990291 UNKNOWN (possible beta actin) (5)	0.87	0.87	0.0361	0.0165
Detoxification	CB494429 [NR] [AAH45932] Aldehyde dehydrogenase 9 family, member A1 like 1	1.14	1.18	0.0245	0.0065
	CA052877 [GO] [P46412] Plasma glutathione peroxidase precursor	1.19	1.15	0.0248	0.0203
Energy metabolism	CA062426 [GO] [Q91Y97] Fructose-bisphosphate aldolase B	1.16	1.48	0.0040	0.0003
	CB497834 [GO] [P14152] Malate dehydrogenase, cytoplasmic	1.33	1.27	0.0015	0.0041
	CB505763 [GO] [P23591] GDP-L-fucose synthetase	1.17	1.12	0.0355	0.0393
	CN442510 [GO] [P18929] NADH-ubiquinone oxidoreductase chain 1	0.87	0.90	0.0404	0.0326
Germ-line formation	CA061998 [GO] [P23298] Protein kinase C, eta type	1.40	1.15	0.0009	0.0122
Immunity	CA040697 [NT] [AY071854] Oncorhynchus mykiss clone GC71 MHC class I antigen	1.18	1.26	0.0285	0.0050
	CB502941 [NR] [AAS89353] perforin	0.45	0.34	<0.0001	<0.0001
	CB496842 [GO] [Q91ZW7] CD209 antigen-like protein E	0.64	0.86	0.0001	0.0181
	CB497396 [GO] [P12246] Serum amyloid P-component precursor	0.89	0.78	0.0460	0.0100
	CA037647 [GO] [P07724] Serum albumin precursor	0.82	0.85	0.0299	0.0208
	CB504246 [NR] [AAU21486] fucolectin	0.86	0.89	0.0204	0.0410
Iron homeostasis	CA056696 [GO] [P09528] Ferritin heavy chain	1.17	1.48	0.0277	0.0004
	CA768211 [GO] [P09528] Ferritin heavy chain	1.11	1.16	0.0372	0.0227
	CA056544 [GO] [Q921I1] Serotransferrin precursor	0.95	0.80	0.0381	0.0053
Lipid metabolism	CK990702 [GO] [Q62264] Thyroid hormone-inducible hepatic protein	1.98	1.60	<0.0001	<0.0001
	CK990953 [GO] [P51162] Gastrotropin	1.27	1.26	0.0061	0.0014
	CA038031 [GO] [Q01339] Beta-2-glycoprotein I precursor	1.18	1.10	0.0241	0.0487
	CB497259 [NR] [AAB96972] apolipoprotein A-I-1	0.83	0.84	0.0135	0.0374
	CB510585 [GO] [P06728] Apolipoprotein A-IV precursor	0.86	0.84	0.0298	0.0259
	CB510956 [NR] [AAG11410] apolipoprotein CII	0.84	0.86	0.0280	0.0229

Table 3. (continued from previous page)

Functional group:(1)	Gene clone number and name:(2)	Expression ratio:(3)			
		Cliff Lake	Indian Pond	*P*-CL(4)	*P*-IP(4)
Muscle contraction	CB497373 [GO] [P13412] Troponin I, fast skeletal muscle	1.12	1.25	0.0247	0.0031
Protein degradation	CK990562 [GO] [P06797] Cathepsin L precursor	1.16	1.26	0.0204	0.0062
	CB510468 [GO] [P62991] Ubiquitin	1.17	1.18	0.0423	0.0104
	CB511632 [GO] [Q60692] Proteasome subunit beta type 6 precursor	0.78	0.80	0.0052	0.0052
	CA053701 [GO] [Q9Z2U0] Proteasome subunit alpha type 7	0.89	0.85	0.0371	0.0160
Protein synthesis	CB504359 [GO] [Q9CQE3] 28S ribosomal protein S17	1.21	1.45	0.0124	0.0005
	CB500108 [NR] [AAM34649] 60S ribosomal protein L35	1.29	1.21	0.0022	0.0074
	CA050584 [NR] [AAQ97817] nascent-polypeptide-associated complex α-polypeptide	1.13	1.15	0.0417	0.0362
	CK991302 [NT] [AY372389] Oncorhynchus mykiss calreticulin	0.67	0.74	0.0005	0.0030
	CB496532 [GO] [P51410] 60S ribosomal protein L9	0.78	0.88	0.0032	0.0374
Unknown	CA040785 UNKNOWN	2.77	1.26	<0.0001	0.0191
	CA038900 UNKNOWN	1.28	1.63	0.0013	0.0001
	CB492348 UNKNOWN	1.42	1.36	0.0030	0.0035
	CK990518 UNKNOWN	1.48	1.17	0.0006	0.0126
	CA052504 UNKNOWN	1.42	1.21	0.0007	0.0047
	CA045203 UNKNOWN	1.23	1.29	0.0126	0.0152
	CA046404 UNKNOWN	1.30	1.16	0.0044	0.0215
	CB511601 UNKNOWN	1.19	1.25	0.0224	0.0100
	CK990765 UNKNOWN	1.24	1.21	0.0110	0.0055
	CK990939 UNKNOWN	1.31	1.13	0.0028	0.0496
	CK990930 UNKNOWN	1.25	1.17	0.0049	0.0417
	CB502618 UNKNOWN	1.24	1.12	0.0071	0.0331
	CA037110 UNKNOWN	1.21	1.16	0.0471	0.0156
	CA051731 UNKNOWN	1.13	1.17	0.0374	0.0131

(1) : Functional groups were defined as described in Material and methods;
(2) : Each EST clone number represents a single EST sequence;
(3) : Transcription ratio corresponds to mean dwarf expression level divided by mean normal expression level (D/N);
(4) : P-CL and P-IP correspond to permutated P values for Cliff Lake and Indian pond (ANOVA, F3 test, 1000 permutations);
(5) : Blastn results partly match actin beta mRNA sequences for *D. labrax, E. coioides* and *O. mykiss* ;

Table 4. Relative representation of functional groups for genes showing parallel patterns of regulation between sympatric dwarf and normal whitefish from two natural lakes (left), and from the two natural lakes plus the control group (right).

Functional groups: (1)	Cliff Lake vs. Indian Pond		Cliff Lake vs. Indian Pond vs. control	
	Proportion(%): (2)	*P*-val: (3)	Proportion(%):	*P*-val:
Blood and transport	3.9	0.03079	-	-
Cellc cycle regulation	7.8	3.324e-05	8.0	0.0007494
Cell structure	2.6	0.03173	-	-
Detoxification	5.2	0.005477	8.0	0.00543
Energy metabolism	14.3	6.895e-05	36.0	8.764e-08
Germ-line formation	3.9	0.0007837	8.0	0.0009456
Immunity	9.1	0.002373	4.0	0.08622
Iron homeostasis	3.9	0.02045	4.0	0.1504
Lipid metabolism	9.1	1.699e-06	4.0	0.1778
Muscle contraction	1.3	0.1295	-	-
Protein catabolism	5.2	0.004873	-	-
Protein synthesis	9.1	0.02215	8.0	0.0143

(1) : Functional groups are described in Material and Methods ;
(2) : Proportion was calculated by dividing the number of genes in each functional group by the total number of non-redundant parallel genes x 100 (Two lakes : n =77 ; two lakes and control : n =25) ;
(3) : *P*-val from X^2–test on proportions ;

Table 5. Proportion of up-regulated genes in dwarf and normal whitefish in regards to their potential involvement in life history trade-offs observed between growth (increased fecundity) vs. survival. a) Cliff Lake vs. Indian Pond; b) Cliff Lake vs. Indian Pond vs. control group.

a.)	Dwarf		Normal	
Functions potentially associated to survival	% of up-regulated genes	*P*-val(4)	% of up-regulated genes	*P*-val
Energy metabolism	81.8	0.03271	-	
Iron homeostasis	66.7	0.5	-	
Lipid metabolism	57.1	0.5	-	
Detoxification	100	<0.0001	-	
All energy (1)	76.0	0.007317	-	
Germ-line formation (2)	100	<0.0001	-	
Functions potentially associated to growth (increased fecundity)				
Cell cycle & cell growth	-		66.7	0.3438
Protein synthesis	-		57.1	0.5
All growth (3)	-		61.5	0.2905

b.)	Dwarf		Normal	
Functions potentially associated to survival	% of up-regulated genes	*P*-val	% of up-regulated genes	*P*-val
Energy metabolism	88.9	0.01953	-	
Iron homeostasis	100	<0.0001	-	
Lipid metabolism	100	<0.0001	-	
Detoxification	100	<0.0001	-	
All energy	92.3	0.001709	-	
Germ-line formation	100	<0.0001	-	
Functions potentially associated to growth (increased fecundity)				
Cell cycle & cell growth	-		50	0.75
Protein synthesis	-		100	<0.0001
All growth	-		75	0.3125

1. All energy includes the energy metabolism, iron homeostasis, lipid metabolism and detoxification functional groups taken together;
2. This functional group does not relate to survival or growth but was differentially over-expressed as a whole in dwarf whitefish.
3. All growth includes the cell cycle/cell growth and protein synthesis functional groups taken together;
4. *P*-val from X^2–test on proportions;

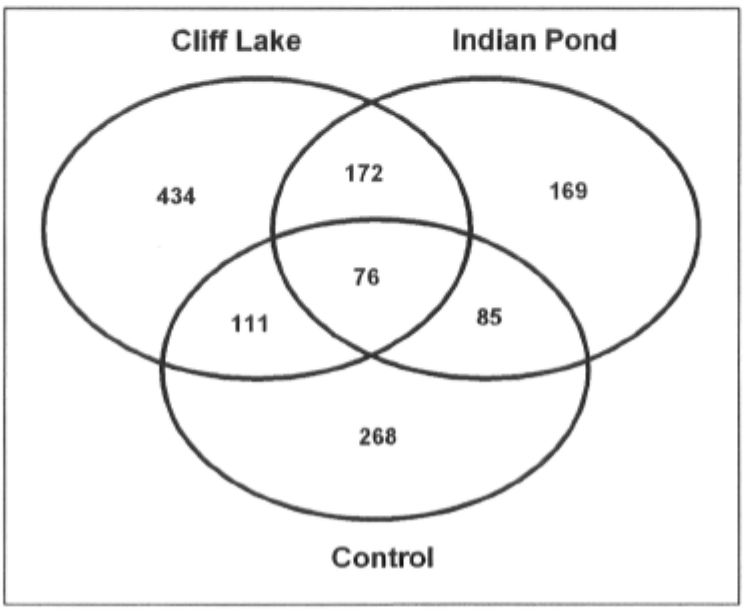
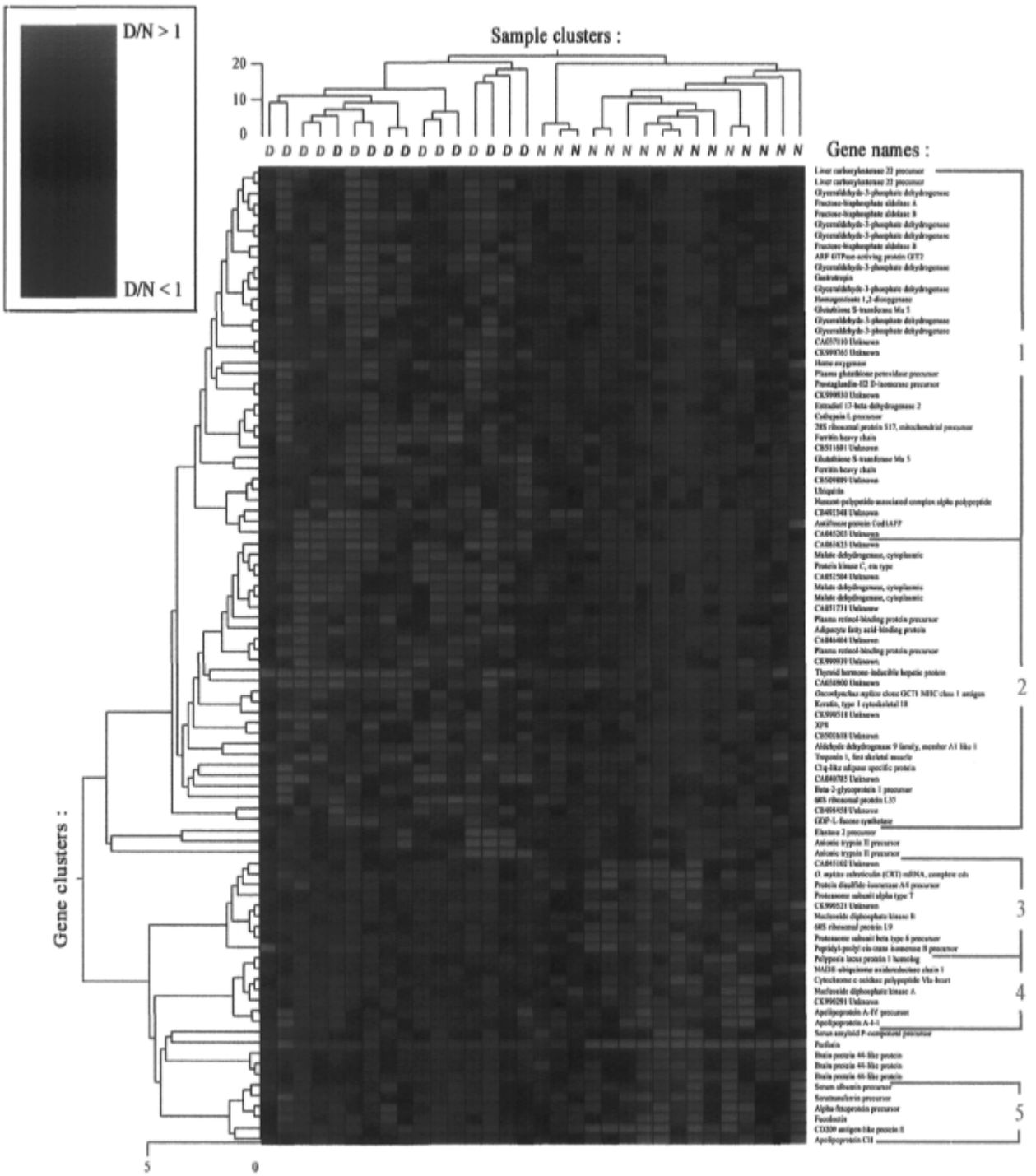


Figure 1. Venn diagram showing the number of gene clones with significant differences in level of transcription between sympatric dwarf and normal whitefish in two natural lakes (Cliff Lake and Indian Pond) and controlled environment. Numbers are presented for each separate group, as well as for the number of genes shared between any of the two or all three groups. Details on directionality of transcription are presented in Table 1.

a.)


D/N > 1
D/N < 1
Sample clusters :
20
10
0
D D D D D D D D D D D D D D D N N N N N N N N N N N N N N N N
Gene names :
Gene clusters :
1
2
3
4
5
5
0

b.)

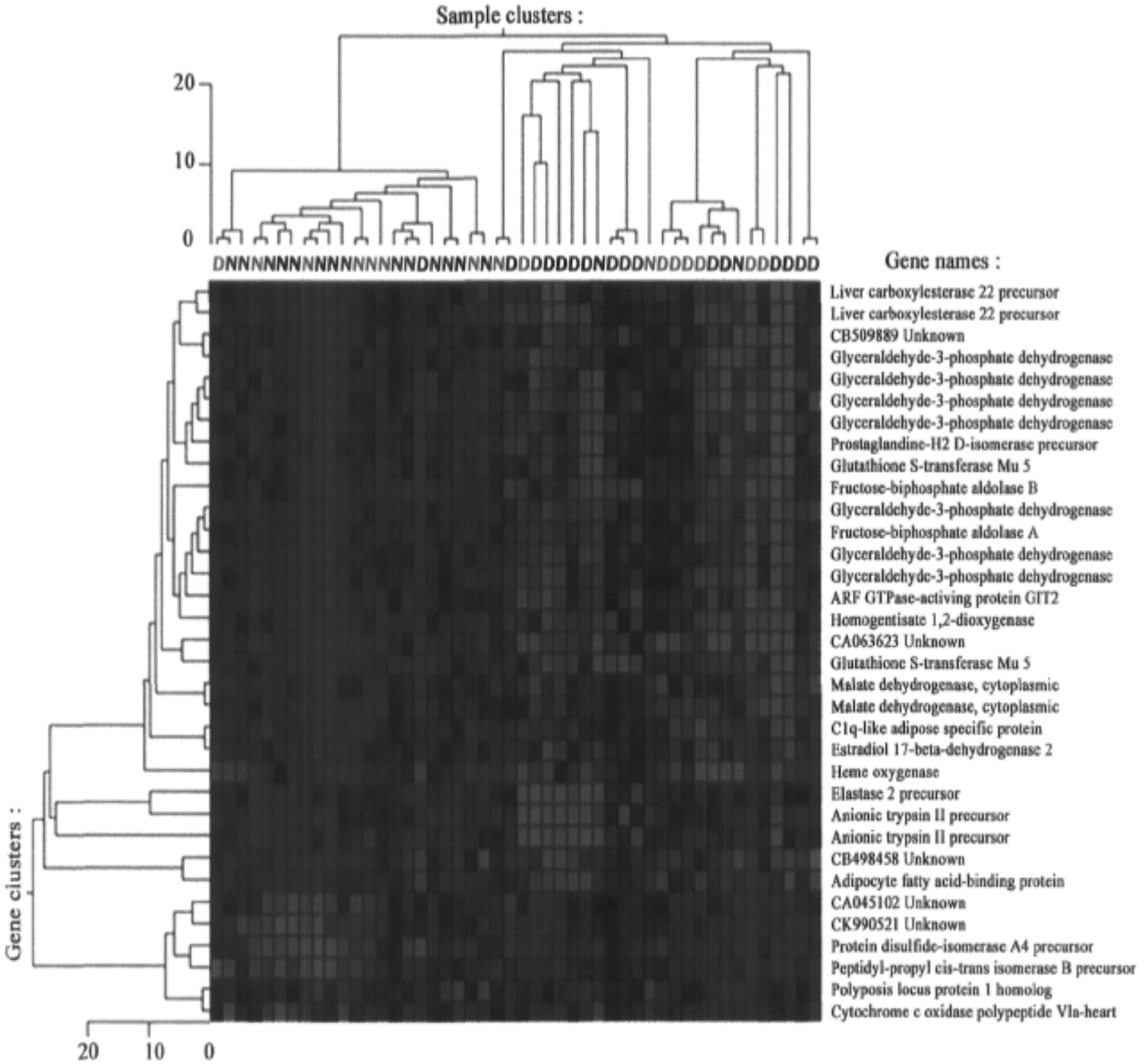
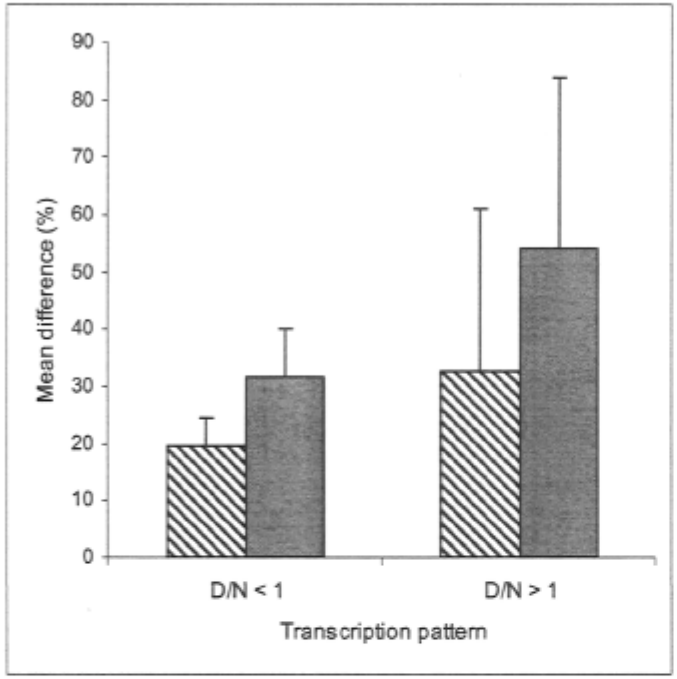

Sample clusters :
20
10
0
DNNNNNNNNNNNNNNDNNNNNNDDDDDDDNDDDNDDDDDDNDDDDDD
Gene names :
Liver carboxylesterase 22 precursor
Liver carboxylesterase 22 precursor
CB509889 Unknown
Glyceraldehyde-3-phosphate dehydrogenase
Glyceraldehyde-3-phosphate dehydrogenase
Glyceraldehyde-3-phosphate dehydrogenase
Glyceraldehyde-3-phosphate dehydrogenase
Prostaglandine-H2 D-isomerase precursor
Glutathione S-transferase Mu 5
Fructose-biphosphate aldolase B
Glyceraldehyde-3-phosphate dehydrogenase
Fructose-biphosphate aldolase A
Glyceraldehyde-3-phosphate dehydrogenase
Glyceraldehyde-3-phosphate dehydrogenase
ARF GTPase-activing protein GIT2
Homogentisate 1,2-dioxygenase
CA063623 Unknown
Glutathione S-transferase Mu 5
Malate dehydrogenase, cytoplasmic
Malate dehydrogenase, cytoplasmic
C1q-like adipose specific protein
Estradiol 17-beta-dehydrogenase 2
Heme oxygenase
Elastase 2 precursor
Anionic trypsin II precursor
Anionic trypsin II precursor
CB498458 Unknown
Adipocyte fatty acid-binding protein
CA045102 Unknown
CK990521 Unknown
Protein disulfide-isomerase A4 precursor
Peptidyl-propyl cis-trans isomerase B precursor
Polyposis locus protein 1 homolog
Cytochrome c oxidase polypeptide VIa-heart
Gene clusters :
20
10
0

Figure 2. a) Cluster tree for 92 gene clones showing parallel patterns of regulation between sympatric dwarf and normal whitefish from the two natural lakes; b) Cluster tree for 34 gene clones showing parallel patterns of transcription between sympatric dwarf and normal whitefish from two natural lakes and controlled conditions. Gene names are listed on the right of each tree. The dendrogram on top (Sample cluster) groups individuals based on similar patterns of transcription for genes with an average D/N value indicating parallel patterns of transcription in compared groups. Individuals are designated by the letters N and D for 'normal' and 'dwarf' whitefish (*red* : Cliff Lake ; *green* : Indian Pond ; *grey* : control). The dendrogram on the left (Gene cluster) groups genes with similar patterns of expression between individuals, where five broad functional clusters were defined (see Results). After normalization (see Material and Methods), expression values shown in red (up-regulated in dwarf on average) are relatively higher in the experimental pairwise comparison and values shown in green (up-regulated in normal on average) are relatively lower in the experimental pairwise comparison. Hierarchical clustering analysis were performed using the average linking method (described in Eisen et al. 1998) with pairwise distances calculated by Pearson correlation coefficients (Qu and Xu 2006).

a.)



b.)

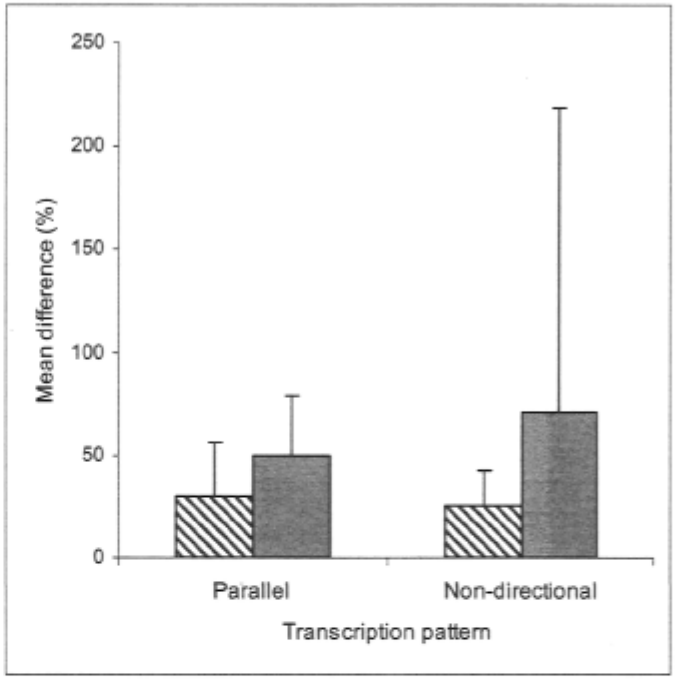


Figure 3. a.) Mean differences (in percent) in transcription levels for gene clones showing parallel patterns of regulation (D/N >1 : up-regulation in dwarf ; D/N <1 : up-regulated in normal) in whitefish from the two natural lakes (striped bars) and from the controlled environment (plain bars) ; b) mean differences (%) in transcription levels for gene clones showing parallel (left) or non-directional (right) patterns of regulation in fish from the two natural lakes (striped bars) or from the controlled environment (plain bars). Differences in expression were calculated using expression value in the normal ecotype as the reference; *t*-bars represent standard deviation for transcription values.

CONCLUSION GÉNÉRALE

Utilisation des biopuces en écologie et en évolution :

L'utilisation de technologies de la génomique telles que les biopuces à ADN (*microarrays*) pour l'étude du phénomène de la radiation adaptative semble très prometteuse, comme il est possible de le constater par les résultats obtenus dans la présente étude. En résumé, la comparaison des profils de transcription des deux formes sympatriques du grand corégone, provenant de deux lacs ainsi que d'un milieu contrôlé, nous a permis de mettre en évidence plusieurs dizaines de gènes candidats, potentiellement impliqués dans la divergence phénotypique entre les corégones nains et normaux, et ce au niveau de plusieurs processus physiologiques. Les patrons de variation interindividuels, ainsi que les patrons d'embranchements observés à l'intérieur de l'arbre de regroupements (*cluster tree*) pour ces gènes en nature (Fig. 2a), suggèrent que ces gènes candidats font partie de cinq grands réseaux de corégulation, potentiellement sous contrôle d'un seul ou de quelques gènes de régulation. Par contre, cette hypothèse constitue une explication parmi plusieurs possibles et devrait faire l'objet d'expérimentations plus approfondies pour être validée.

Plusieurs éléments nous permettent d'affirmer que ces gènes sont importants dans la divergence adaptative du corégone nain. Premièrement, la présence de patrons de régulation en parallèle dans plusieurs systèmes indépendants (deux lacs ou populations naturelles ainsi qu'une population en milieu contrôlé) constitue un argument majeur en faveur du rôle de la sélection naturelle directionnelle dans l'apparition et le maintien de ces différences de régulation entre les deux formes. Deuxièmement, la présence d'une majorité de gènes surexprimés dans la forme naine par rapport à la forme normale est cohérente avec l'hypothèse voulant que la forme naine ait émergé à partir de la forme normale pour coloniser une niche écologique disponible (Bernatchez 2004). Il semble en effet que la régulation des gènes dans le cas de l'apparition de nouvelles adaptations corresponde dans certains cas à une augmentation de l'expression, ce qui est observé par exemple dans le cerveau humain comparativement à celui du chimpanzé (Preuss et al. 2004). Ainsi, la surexpression de certains gènes aurait permis l'augmentation du métabolisme nécessaire à l'exploitation de la niche écologique limnétique par les corégones nains. Troisièmement, les groupes fonctionnels représentés par les gènes ayant des patrons de régulation parallèles sont tous cohérents avec les adaptations observées au niveau de plusieurs traits phénotypiques distinguant les deux formes du grand corégone (comportements, activité natatoire, métabolisme énergétique, morphologie, histoire de vie, taux de croissance). Quatrièmement, la variance du niveau de

transcription des gènes ayant des patrons parallèles de régulation est réduite par rapport à celle des gènes non-directionnels en nature, ce qui plaide en faveur du rôle de la sélection naturelle stabilisatrice sur le maintien de ces patrons d'expression. Il apparaît donc que la sélection naturelle agisse de deux façons sur le profil d'expression. D'abord d'une façon directionnelle, en favorisant une divergence dans les patrons de transcription génétique entre les corégones nains et normaux. Ensuite en stabilisant l'expression des gènes importants pour l'adaptation optimale des deux populations à leur niche écologique respective.

Comparaison entre milieux naturels et contrôlé :

Une des forces majeures de cette étude réside dans la comparaison de la variation dans les patrons de transcription génétique entre des populations naturelles et des populations maintenues en milieu contrôlé. Ceci constitue à notre connaissance une première dans le domaine de la génomique écologique et évolutive. En effet, la transcription des gènes est sujette à la variation environnementale, ce qui ne peut manquer d'introduire une multitude de sources de variation dans les niveaux de transcription, autant du point de vue interindividuel qu'interpopulationnel (Whitehead & Crawford 2006a; Kammenga et al. 2007). La comparaison avec des populations issues d'un milieu contrôlé permet d'éliminer la majeure partie de ces sources de variation (Whitehead & Crawford 2006b). Ceci nous permet d'affirmer que les différences de régulation entre nains et normaux retrouvées dans deux lacs ainsi qu'en milieu contrôlé sont héritables, c'est-à-dire qu'elles ont une base génétique et ne sont pas seulement dues à des différences dans les conditions environnementales prévalant dans les différentes niches, ni seulement à des phénomènes épigénétiques attribuables à l'effet de l'environnement sur les stades précoces du développement (Kammenga et al. 2007). De plus, ces patrons de régulation ayant été observés en parallèle dans trois milieux indépendants écarte la possibilité qu'ils ne soient le fait de phénomènes stochastiques de variation (e.g. dérive génétique; Khaitovich et al. 2004) et que ces observations soient fortuites (Schluter 1996). En définitive, le parallélisme observé dans les patrons de régulation pour ces gènes candidats suggère fortement un rôle pour la sélection naturelle dans la divergence au niveau de ces traits phénotypiques d'expression.

Régulation des gènes et adaptation :

L'idée que la régulation des gènes soit un mécanisme important dans l'apparition d'adaptations nouvelles chez les espèces a été lancée il y a de cela plusieurs décennies déjà (Britten & Davidson 1969; King & Wilson 1975). En effet, l'accumulation d'observations et de résultats depuis les dernières années ajoute à la solidité d'une telle hypothèse. Le cas du grand corégone ne fait pas exception. En gardant à l'esprit que la forme naine soit apparue à la suite du retrait des glaciers aux termes de la dernière période de glaciation, il y a environ 12 000 à 15 000 ans, et en considérant le nombre de variations qui différencient la forme naine de la forme normale au niveau du transcriptome – comparé à la vitesse d'accumulation normale des mutations dans le génome (ou horloge moléculaire) – il semble en effet que la régulation de l'expression des gènes ait contribué de façon très importante à la divergence et à l'adaptation phénotypique du corégone nain vers l'exploitation de la niche écologique limnétique. Ceci est en accord avec l'hypothèse voulant qu'une mutation ayant des conséquences sur l'expression de plusieurs gènes constitue un mécanisme de divergence adaptative beaucoup plus rapide et fréquent que l'accumulation de mutations dans la séquence codante de ces gènes (Gibson & Weir 2005).

Cette étude enrichit les approches plus traditionnelles mettant en jeu des traits morphologiques, en rendant possible l'accès à plusieurs traits phénotypiques moins apparents, comme les patrons de régulation génétique et les processus physiologiques aux niveaux cellulaire et moléculaire. En effet, il est possible de relier de façon assez convaincante plusieurs gènes spécifiques, ou groupes de gènes, à des fonctions physiologiques variées, lesquelles seraient importantes dans l'adaptation de la forme naine à un mode de vie limnétique. Par exemple, la surexpression chez le corégone nain de 17 gènes impliqués dans le métabolisme énergétique est cohérente avec une activité natatoire accrue, en lien avec un mode de nutrition limnétique et à des comportements d'évitement des prédateurs. Cette activité musculaire accrue entraîne une augmentation de la production de métabolites secondaires et d'agents potentiellement dommageables pour l'intégrité du génome, ce qui est cohérent avec la surexpression, toujours chez le nain, de quatre gènes impliqués dans différents processus de détoxification. Il en va ainsi pour plusieurs autres groupes fonctionnels représentés dans les données d'expression (*Voir à ce sujet la section* Discussion *du Chapitre 1 du présent mémoire*).

Bases génomiques de la divergence adaptative :

De plus, il semble que le compromis (*trade-off*) observé au niveau de traits d'histoire de vie entre les deux formes, à savoir une augmentation de la fécondité via un taux de croissance et une longévité accrue chez la forme normale *versus* une augmentation de l'activité métabolique nécessaire à la survie en milieu limnétique chez la forme naine, se reflète dans les différences du profil d'expression génétique entre les corégones nains et normaux. Nous avons pu vérifier la prédiction que des gènes associés à une augmentation de l'activité métabolique de base – autant au niveau de processus énergétiques qu'au niveau du métabolisme secondaire, ainsi que de la détoxification qui y est associée - seraient surexprimés chez le corégone nain, tandis que des gènes associés à la croissance (synthèse protéique, synthèse d'ADN, cycle cellulaire) seraient en contre partie surexprimés chez le corégone normal. Par conséquent, cette étude vient contribuer à combler un manque au niveau des données moléculaires et mécanistiques relativement à la théorie de l'histoire de vie, domaine de recherche en écologie où peu de résultats de cette nature ont été publiés jusqu'à présent (Roff 2007).

En lien avec ce compromis énergétique entre deux stratégies d'histoire de vie, les profils d'expression génétiques peuvent, en plus d'offrir un support moléculaire aux divergences phénotypiques observées, expliquer en partie le déficit énergétique apparent observé chez les corégones nains. En effet, même une très faible augmentation de la synthèse de macromolécules, telle l'ARNm, entraîne un coût énergétique pour la cellule (Wagner 2007). Considérant que les nains nagent plus activement (Rogers et al. 2002) et ont une moins bonne efficacité de conversion énergétique des ressources trophiques par rapport aux corégones normaux (Trudel et al. 2001), une augmentation de la synthèse d'ARN ne pourrait qu'accentuer le déficit général de ressources énergétiques nécessaires à des processus physiologiques très exigeants au niveau énergétique, tel que la croissance par exemple. Cette hypothèse est aussi cohérente avec une durée de vie moyenne plus courte, attribuable en partie à un métabolisme basal généralement plus élevé chez le corégone nain. Par contre, la prédation demeure sans doute le facteur ayant la plus grande incidence sur la durée de vie chez les corégones nains (Kahilainen & Lehtonen 2002).

En terme de succès reproducteur, il serait donc profitable pour le corégone nain d'atteindre la maturité sexuelle plus rapidement, ce qui est généralement observé (Chouinard et al. 1996).

Le profil d'expression génétique va dans ce sens pour au moins trois gènes, ayant des patrons de régulation parallèles en nature (2) et en milieu contrôlé (1), impliqués dans la maturation des gonades ainsi que dans la différenciation des cellules sexuelles, lesquels gènes sont surexprimés chez le corégone nain.

Tout au long de cette étude, il a surtout été question de gènes dont le patron de régulation entre corégones nains et normaux montre un parallélisme dans la différence de niveau de transcription et dans la direction de la régulation. Les annexes 1 et 2 fournissent la liste des gènes montrant un parallélisme dans les différences de niveau de transcription entre les groupes, mais de façon non-directionnelle. Bien qu'il n'en ait pas été question dans la présente analyse fonctionnelle, ces gènes pourraient constituer des candidats à l'adaptation phénotypique des corégones nains et normaux. En effet, il est fait état dans la littérature de la possibilité d'emprunter différents chemins moléculaires pour arriver à un même phénotype (Johnson & Porter 2000 ; Derome et al. 2006). De plus, les traits phénotypiques complexes faisant intervenir des réseaux complexes de gènes s'influençant les uns les autres, il est fort probable qu'une fraction significative des gènes listés en annexe soient impliqués dans les principaux réseaux régulatoires mis en évidence par la Figure 2a, ou à tout le moins qu'une proportion significative de ces gènes non-directionnels soient associés à un des 12 groupes fonctionnels représentés dans les données relatives aux gènes candidats, listés dans les tableaux 2 et 3 du chapitre 1 du présent mémoire.

Perspectives futures:

Il est fréquent de valider les résultats obtenus à l'aide de biopuces à ADN par des méthodes complémentaires de mesures de la transcription génétique (Draghici et al. 2006). La méthode la plus couramment utilisée est la PCR quantitative (ou RT-PCR pour *real time PCR* ou qPCR pour *quantitative PCR*) qui consiste à quantifier la quantité d'ARN d'un transcrit particulier par une réaction de PCR couplée à l'utilisation d'une sonde marquée d'un fluorophore (Heid et al. 1996). Cette méthode s'avère très efficace et donne des résultats réplicables dans la majorité des cas (Dallas et al. 2005). Dans cette perspective, les niveaux d'expression de deux gènes candidats obtenus lors de ce projet de maîtrise (anionic trypsin II et carboxylesterase 22; voir Chapitre 1, Tableau 2) ont été quantifiés par PCR quantitative pour ces mêmes populations avec des résultats consistants (Jeukens J, *communication personnelle*). Des études

similaires ont aussi été conduites sur le gène de la parvalbumine-beta, gène candidat surexprimé dans le muscle blanc chez le corégone nain (Derome et al. 2006) et chez le cisco (*C. artedi*), un compétiteur écologique du corégone nain (Derome & Bernatchez 2006), avec des résultats consistants pour la majorité des populations étudiées (Jeukens J, *communication personnelle*).

Il serait aussi possible d'obtenir des données de transcription génétiques par l'utilisation de biopuces à oligos (e.g. Affymetrix; Lee et al. 2003), une méthode beaucoup plus précise notamment à cause de la spécificité des séquences, mais nécessitant une connaissance de la séquence du génome pour l'organisme à l'étude, ce qui n'est pas le cas du grand corégone.

Parallèlement, nous avons aussi procédé à la cartographie des traits quantitatifs d'expression (eQTL), méthode qui consiste à identifier des régions du génome dont la variation est corrélée avec la variation des niveaux de transcription génétique (Wayne & McIntyre 2002), ce qui pourra nous renseigner davantage sur l'architecture génomique des différences de régulation adaptative entre les corégones nains et normaux (Rockman & Kruglyak 2007; Rogers et Bernatchez 2007; Derome et al. *soumis*).

Accessoirement, l'analyse des réseaux de corégulation par le regroupement de gènes selon leurs profils de transcription interindividuels offre plusieurs possibilités en termes de génomique fonctionnelle. Par exemple, l'arbre obtenu avec 92 gènes montrant des patrons de régulations parallèles chez les corégones issus de deux lacs (Chapitre 1, Fig. 2a) suggère que les principaux groupes fonctionnels divergents entre les deux populations sont régulés par un seul ou quelques gènes, donnant ainsi une première indication de l'architecture génétique de l'adaptation phénotypique au niveau de la régulation de l'expression des gènes chez le corégone. De plus, par la règle du « coupable par association » (*guilt by association :* Quackenbush et al. 2003), ces analyses indiquent que des gènes regroupés dans des branches communes, en plus de faire partie d'un même réseau de régulation, partagent des fonctions soit complémentaires ou semblables et participent au même processus à l'échelle moléculaire, cellulaire ou physiologique. Ceci pourrait s'avérer fort intéressant pour la caractérisation de transcrits encore inconnus, pour lesquels les données font défaut. L'exemple du clone « Unknown 90992 », dont la séquence présente des homologies d'environ 80% avec l'ARNm de l'actine chez trois autres Téléostéens est frappant : il se groupe avec des gènes dont la fonction indique qu'ils pourraient être impliqués dans l'absorption et le transport vacuolaire

des lipides vers la mitochondrie, où ces lipides seraient utilisés comme source d'ATP. À cet égard, l'actine est une composante du cytosquelette, potentiellement impliquée dans le transport intracellulaire de vacuoles plasmiques (voir Reggiori et al. 2005). De plus, l'hypothèse d'une augmentation de l'absorption et de la phosphorylation oxydative des lipides dans le foie chez le corégone normal, par opposition à un besoin accru de transport de ces molécules énergétiques vers des cibles plus actives comme le muscle blanc chez le corégone nain, est tout à fait cohérente avec une augmentation de l'activité natatoire chez ce dernier. Il serait intéressant de vérifier la réplicabilité de cette tendance dans les patrons de regroupement des autres gènes dont l'identité est encore inconnue (19 au total; Chapitre 1, tableaux 2 et 3), à mesure que les données d'annotations fonctionnelles et de séquences homologues seront rendues disponibles pour ces transcrits.

Finalement, il est certain que le lien établi par la présente étude entre les gènes candidats et leur rôle dans l'adaptation phénotypique vers l'utilisation d'une niche écologique spécifique, que ce soit chez le corégone nain ou normal, demeure hypothétique. En effet, des études plus poussées sur les effets des gènes sur le phénotype devraient être conduites (e.g. *knock-out*, mutagenèse dirigée) dans le but d'établir un lien de cause à effet valable entre la régulation différentielle de ces transcrits et la variation phénotypique, lien qui malgré la cohérence observée entre l'analyse fonctionnelle et la divergence phénotypique, demeure de nature corrélationnelle (Kammenga et al. 2007).

BIBLIOGRAPHIE :

Alvarez D, Nicieza AG. 2006. Factors determining tadpole vulnerability to predators: can prior experience compensate for suboptimal shape? Evol. Ecol. 20:523-534.

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

Bernatchez L, Chouinard A, Lu G. 1999. Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, Coregonus sp., as a case study. Biol. J. Linn. Soc. 68:173-194.

Bernatchez L, Dodson JJ. 1990. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) revealed by mitochondrial DNA restriction analysis. Evolution 44: 1263-1271.

Bochdanovits Z, de Jong G. 2004. Antagonistic pleiotropy for life-history traits at the gene expression level. Proc. R. Soc. Lond. B (Suppl.) 271:S75-S78.

Bochdanovits Z, van der Klis H, de Jong G. 2003. Covariation of larval gene expression and body size in natural populations of Drosophila melanogaster. Mol. Biol. Evol. 20:1760-1766.

Boucrot E, Kirchhausen T. 2007. Endosomal recycling controls plasma membrane area during mitosis. Proc. Nat. Acad. Sci. USA 104: 7939-7944.

Brakefield PM, French V, Zwaan BJ. 2003. Development and the genetics of evolutionary change within insect species. Annu. Rev. Ecol. Evol. Syst. 34: 633-660.

Britten RJ, Davidson EH. 1969. Gene regulation for higher cells: a theory. Science 165: 349-357.

Brodsky LI, Jacob-Hirsch J, Avivi A, et al. (11 co-authors). 2005. Evolutionary regulation of the blind subterranean mole rat, Spalax, revealed by genome-wide gene expression. Proc. Natl. Acad. Sci. USA 102:17047-17052.

Brown PO, Botstein D. 1999. Exploring the new world of the genome with DNA microarrays. Nature Genet. Suppl. 21: 33-37.

Butte A. The use and analysis of microarray data. Nature Rev. Drug Disc. 1: 951-960.

Calero M, Whittaker GR, Collins RN. 2001. Yop1, the yeast homolog of the polyposis locus protein 1, interacts with Yip1 and negatively regulates cell growth. J. Biol. Chem. 276: 12100-12112.

Campbell D, Bernatchez L. 2004. Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol. Biol. Evol. 21: 945-956.

Cheung VG, Spielman RS. 2002. The genetics of variation in gene expression. Nature Genet. Suppl. 32:522-525.

Chouinard A, Pigeon D, Bernatchez L. 1996. Lack of specialisation in trophic morphology between genetically differentiated dwarf and normal forms of lake whitefish (Coregonus clupeaformis Mitchill) in Lac de l'Est, Quebec. Can. J. Zool. 74:1989-1998.

Churchill GA. 2002. Fundamentals of experimental design for cDNA microarrays. Nature Genet. Suppl. 32:490-495.

Claiborne JB. 1998. Acid-base regulation. In: Evans DH, editor. The physiology of fishes (Second edition). CRC Press. Boca Raton, FL. p. 177-198.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal Jr.G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. Science 303:1928-1933.

Cossins A, Fraser J, Hughes M, Gracey A. 2006. Post-genomic approaches to understanding the mechanisms of environmentally induced phenotypic plasticity. J. Eperim. Biol. 209:2328-2336.

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

Cresko WA, McGuigan KL, Phillips PC, Postlethwait JH. 2007. Studies of threespine stickleback developmental evolution: progress and promise. Genetica 129:105-126.

Cui X, Churchill GA. 2003. Statistical tests for differential expression in cDNA microarray experiments. Gen. Biol. 4:210.

Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, Freitas JR, Boag JM, Cummings AJ, Kees UR. 2005. BMC Genomics 6: Art. No. 59.

Derome N, Bernatchez L. 2006. The transcriptomics of ecological convergence between 2 limnetic coregonine fishes (Salmonidae). Mol. Biol. Evol. 23:2370-2378.

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

Derome N, Bougas B, Rogers S, Whiteley A, Labbe A, Laroche J, Bernatchez L. 2007. Pervasive sex-linked effects on transcription regulation as revealed by eQTLmapping in lake whitefish species pairs (Coregonus sp, Salmonidae). *(article soumis).*

Diegesser T, Seitz A, Johannesen J. 2007. Morphological adaptation in host races of *Tephritis conura.* Entomol. Exp. Applic. 122: 155–164.

Dobzansky T. 1951. Genetics and the origin of species. Columbia University Press, New York.

Edelist C, Lexer C, Dillmann C, Sicard D, Rieseberg LH. 2006. Microsatellite signature of ecological selection for salt tolerance in a wild sunflower hybrid species, Helianthus paradoxus. Mol. Ecol. 15:4623–4634.

Ehrenreich IM, Purugganan MD. 2006. The molecular genetic basis of plant adaptation. Am. J. Bot. 93: 952-962.

Eisen MB, Spellman PT, Brown PO, Botstein D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA 95:14863-14868.

Feder ME, Mitchell-Olds T. 2003. Evolutionary and ecological functional genomics. Nature Rev. Gen. 4 : 649-655.

Fenderson OC. 1964. Evidence of subpopulations of lake whitefish, *Coregonus clupeaformis*, involving a dwarf form. Trans. A. Fish. Soc. 93:77–94.

Fisher RA. 1930. The genetical theory of natural selection. Oxford University Press, Oxford.

Fjeldsa J, Irestedt M, Eicson PGP. 2005. Molecular data reveal some major adaptational shifts in the early evolution of the most diverse avian family, the Furnariidae. J. Ornithol. 146: 1-13.

Ford MJ. 2002. Applications of selective neutrality tests to molecular ecology. Mol. Ecol. 11: 1245-1262.

Foster SA, Baker JA. 2004. Evolution in parallel: new insights from a classic system. Trends Ecol. Evol. 19:546-549.

Franchini AG, Egli T. 2006. Global gene expression in Escherichia coli K-12 during short-term and long-term adaptation to glucose-limited continuous culture conditions. Microbiology 152: 2111-2127.

Gibson G. 2002. Microarrays in ecology and evolution: a preview. Mol. Ecol. 11:17-24.

Gibson G, Weir B. 2005. The quantitative genetics of transcription. Trends Genet. 21:616-623.

Gilad Y, Borevitz J. 2006. Using DNA microarrays to study natural variation. Curr. Op. Gen. Dev. 16:553-558.

Gilad Y, Oshlack A, Rifkin SA. 2006. Natural selection on gene expression. Trends Genet. 22:456-461.

Glatz JFC, Schaap FG, Binas B, Bonen A, van der Vusse GJ, Luiken JJFP. 2003. Cytoplasmic fatty acid-binding protein facilitates fatty acid utilization by skeletal muscle. Acta Physiol. Scand. 178:367-371.

Gonzàlez-Martinez SC, Krustovsky KV, Neale DB. 2006. Forest-tree population genomics and adaptive evolution. New Phytol. 170:227-238.

Gracey AY, Cossins AR. 2003. Application of microarray technology in environmental and comparative physiology. Ann. Rev. Physiol. 65:231-259.

Grant PR, Grant R. 2006. Evolution of character displacement in Darwin's finches. Science 313:224-226.

Gray SM, McKinnon JS. 2006. Linking color polymorphism maintenance and speciation. Trends Ecol. Evol. 22:71-79.

Hall, TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, 41, 95-98

Heid CA, Stevens J, Livak KJ, Williams PM. 1996. Real time quantitative PCR. Genome Res. 6: 986-994.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. Evolution 61:995-1016.

Howard DJ, Berlocher SH. 1998. Endless Forms - Species and Speciation. Oxford University Press, Oxford.

Hughes TR, Shoemaker DD. 2001. DNA microarrays for expression profiling. Curr. Op. Chem. Biol. 5: 21-25.

Huxley, JS. 1942. Evolution, the modern synthesis. Allen and Unwin, London.

Ihmels J, Bergmann S, Berman J, Barkai N. 2005. Comparative gene expression analysis by a differential clustering approach: application to the Candida albicans transcription program. PLoS Genet. 1:0380-0393.

Kahilainen K, Lehtonen H. 2002. Brown trout (Salmo trutta L.) and Arctic charr (Salvelinus alpinus (L.)) as predators on three sympatric whitefish (Coregonus lavaretus (L.)) forms in the subarctic Lake Muddusjärvi. Ecol. Freshw. Fish 11:158-167.

Kammenga JE, Herman MA, Ouborg JN, Johnson L, Breitling R. 2007. Microarray challenges in ecology. Trends Ecol. Evol. 22:273-279.

Kerr MK, Martin M, Churchill GA. 2000. Analysis of variance for gene expression microarray data. J. Comput. Biol. 7:819-837.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Pääbo S. 2004. A neutral model of transcriptome evolution. Plos Biol. 2:682-689.

King MC, Wilson AC. 1975. Evolution at two levels in Humans and Chimpanzees. Science 188:107-116.

Kotrschal K, van Staaden MJ, Huber R. 1998. Fish brains: evolution and environmental relationships. Rev. Fish. Biol. Fisher. 8: 373-408.

Kronforst MR, Kapan DD, Gilbert LE. 2006. Parallel genetic architecture of parallel adaptive radiations in mimetic Helioconius butterfly. Genetics 174: 535-539.

Laayouni H, Garcia-Franco F, Chàchez-Sandoval BE, Trotta V, Beltran S, Corominas M, Santos M. 2007. Thermal evolution of gene expression profiles in Drosophila subobscura. BMC Evol. Biol. 7: Art. No. 42.

Lack D. 1947. Darwin's Finches. Cambridge University Press, Cambridge.

Landry L, Vincent WF, Bernatchez L. 2007. Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. J. Evol. Biol. 20: 971-984.

Lee JK, Bussey KJ, Gwadry FG et al. (14 co-authors). 2003. Comparing cDNA and oligonucleotide array data: concordance of gene expression across platforms for the NCI-60 cancer cells. Genome Biol. 4:R82.

Lee CE, Mitchell-Olds T. 2006. Preface to the special issue: ecological and evolutionary genomics of populations in nature. Mol. Ecol. 15:1193-1196.

Lu G, Basley DJ, Bernatchez L. 2001. Contrasting patterns of mitochondrial DNA and Microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. Molecular Eclogy, 10, 965-985.

Lu G, Bernatchez L. 1998. Experimental evidence for reduced hybrid viability between dwarf and normal ecotypes of lake whitefish (*Coregonus clupeaformis* Mitchill). Proc. R. Soc. Lond. B 265: 1025-1030.

Lu G, Bernatchez L. 1999. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (Coregonus clupeaformis): support for the ecological speciation hypothesis. Evolution 53:1491-1505.

MacCallum C, Hill E. 2006. Being positive about selection. PLoS Biol. 4:0293-0295.

Mantripragada KK, Buckley PG, de Stahl TD, Dumanski JP. 2004. Genomic microarrays in the spotlight. Trends Genet. 20: 87-94.

Matzkin LM, Watts TD, Bitler BG, Machado CA, Markow TA. 2006. Functional genomics of cactus host shifts in Drosophila mojavensis. Mol. Ecol. 15:4635-4643.

Mayr, E. 1942. Systematics and the origin of species. Dover Publications, New York.

--- 1963. Animal Species and Evolution. Harvard University Press, Cambridge.

McKinnon JS, Rundle HD. 2002. Speciation in nature: the threespine stickleback model systems. Trends Ecol. Evol. 17:480-488.

McKinnon JS, Mori S, Blackman BK, David L, Kingsley DM, Jamieson L, Chou J, Schluter D. 2004. Evidence for ecology's role in speciation. Nature 429:294-298.

McMorrow T, Wagner A, Harte T, Gannon F. 1997. Sequence analysis and tissue expression of a non-Bohr beta-globin cDNA from Atlantic salmon. Gene 189: 183-188.

Mekel-Bobrov N, Lahn BT. 2006. What makes us human: revisiting an age-old question in the genomic era. J. Biomed. Disc. Coll. 1:18.

Mommsen TP. 1998. Growth and metabolism. In: Evans DH, editor. The physiology of fishes (Second edition). CRC Press. Boca Raton, FL. p. 65-97.

Mommsen TP. 2001. Paradigms of growth in fish. Comp. Biochem. Physiol. pt.B 129:207-219.

Oakeshott JG, Horne I, Sutherland TD, Russell RJ. 2003. The genomics of insecticide resistance. Genome Biol. 4:202.

Olesiak MF, Churchill GA, Crawford DL. 2002. Variation in gene expression within and among natural populations. Nature Genet. 32:261-266.

Olesiak MF, Roach JL, Crawford DL. 2005. Natural variation in cardiac metabolism and gene expression in Fundulus heteroclitus. Nature Genet. 37:67–72.

Orr HA. 2001. The genetics of species differences. Trends Ecol. Evol. 16: 343-350.

Ouborg NJ, Vriezen WH. 2007. An ecologist's guide to ecogenomics. Journ. Ecol. 95:8-16.

Pigeon D, Chouinard A, Bernatchez L. 1997. Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, salmonidae). Evolution 51:196-205.

Powers DA, Schulte PM. 1998. Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: a multidisciplinary approach to address the million-year saga of a small fish. J. Exp. Zool. 282:71-94.

Premont RT, Claing A, Vitale N, Perry SJ, Lefkowitz RJ. 2000. The GIT family of ADP-ribosylation factor GTPase-activating proteins. J. Biol. Chem. 275:22373-2238.

Preuss TM, Caceres M, Oldham MC, Geschwind DH. 2004. Human brain evolution: insights from microarrays. Nature 5:850-860.

Quackenbush J. 2003. Microarrays - Guilt by association. Science 302: 240-241.

Qu Y, Xu S. 2006. Quantitative trait associated microarray gene expression data analysis. Mol. Biol. Evol. 23:1558-1573.

Ranz JM, Machado CA. 2005. Uncovering evolutionary paterns of gene expression using microarrays. Trends Ecol. Evol. 21: 29-37.

Reggiori F, Monastyrska I, Shintani T., Klionsky DJ. 2005. The actin cytoskeleton is required for selective types of autophagy, but not non-specific autophagy, in the yeast *Saccharomyces cerevisiae*. Mol. Biol. Cell 16:5843-5856.

Renn SCP, Aubin-Horth N, Hofmann HA. 2004. Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. BMC Genomics 5: Art. No. 42.

Rennie MD, Collins NC, Shuter BJ, Rajotte JW, Couture P. 2005. A comparison of methods for estimating activity costs of wild fish populations : more active fish observed to grow slower. Can. J. Fish. Aquat. Sci. 62:767-780.

Rieseberg LH, Widmer A, Arntz AM, Burke JM. 2002. Directional selection is the primary cause of phenotypic diversification. PNAS 99:12242-12245.

Rise ML, Douglas SE, Sakhrani D, Williams J, Ewart KV, Rise M, Davidson WS, Koop BF, Devlin RH. 2006. Multiple microarray platforms utilized for hepatic gene expression profiling of GH transgenic coho salmon with and without ration restriction. J. Mol. Endocr. 37:259-282.

Rise ML, von Schalburg KR, Brown GG, et al. (24 co-authors). 2004. Development and application of a salmonid EST database and cDNA microarray: data mining and interspecific hybridization characteristics. Gen. Research 14:478-490.

Rise ML, von Schalburg KR, Cooper GA, Koop BF. 2007. Salmonid DNA microarrays and other tools for functional genomics research. In: Aquaculture genome technologies. (ed. Liu Z), pp. 369-411. Blackwell Publishing Ltd, Oxford.

Roberge C, Einum S, Guderley H, Bernatchez L. 2006. Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. Mol. Ecol. 15:9-20.

Roberge C, Guderley H, Bernatchez L. 2007. Genome-wide identification of genes under selection: gene transcription Qst scan in diverging Atlantic salmon subpopulations. Genetics *(in press)*.

Rockman MV, Kruglyak L. 2007. Genetics of global gene expression. Nature Rev. Genet. 7:862-872.

Roff DA. 1992. The evolution of life histories: theory and analysis. Chapman and Hall, New York.

--- 2002. Life history evolution. Sinauer Associates, Inc., Sunderland, MA.

--- 2007. Contributions of genomics to life-history theory. Nat. Rev. Genet. 8:116-125.

--- 2007. A centennial celebration for quantitative genetics. Evolution 61: 1017-1032.

Rogers SM, Bernatchez L. 2005. Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (Coregonus clupeaformis). Mol. Ecol. 14: 351-361.

Rogers SM, Bernatchez L. 2006. The genetic basis of intrinsic and extrinsic postzygotic reproductive isolation jointly promoting speciation in the lake whitefish sepcies complex (*Coregonus Clupeaformis*). J. Evol. Biol. 19: 1979-1994.

Rogers SM, Bernatchez L. 2007. The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. Mol. Biol. Evol. 24:1423-1438.

Rogers SM, Campbell D, Baird SJE, Danzmann RG, Bernatchez L. 2001. Combining the analyses of introgressive hybridisation and linkage mapping to investigate the genetic architecture of population divergence in the lake whitefish (*Coregonus clupeaformis*, Mitchill). Genetica 111: 25-41.

Rogers SM, Gagnon V, Bernatchez L. 2002. Genetically based phenotype-environment association for swimming behavior in lake whitefish (*Coregonus clupeaformis* Mitchill). Evolution 56:2322-2329.

Rogers SM, Isabel N, Bernatchez L. 2007. Linkage maps of the dwarf and normal lake whitefish (*Coregonus clupeaformis*) species complex and their hybrids reveal the genetic architecture of population divergence. Genetics 175:375-398.

Rogers MR, Marchand M-H, Bernatchez L. 2004. Isolation, characterization, and cross-salmonid amplification of 31 microsatellite loci in the lake whitefish (*Coregonus clupeaformis*, Mitchill). Mol. Ecol. Notes. 4: 89-92.

Rundle HD, Boughman JW, Schluter D. 2000. Natural speciation and parallel speciation in sympatric sticklebacks. Science 287:306-308.

Saetre P, Lindgerg J, Leonard JA, Olsson K, Pettersson U, Ellegren H, Bergström TF, Vilà C, Jazin E. 2004. From wild wolf to domestic dog: gene expression changes in the brain. Mol. Brain Res. 126:198-206.

Schena M, Shalon D, Davis RW, Brown PO. 1995. Quantitative monitoring of gene-expression patterns with complementary-DNA microarray. Science 270: 467-470.

Schluter D. 1996. Ecological speciation in postglacial fishes. Philo. Trans. Roy. Soc. Lond. Biol. Sc. 351:807-814.

--- 2000. The ecology of adaptive radiation. Oxford University Press, Oxford.

--- 2001. Ecology and the origin of species. Trends Ecol. Evol. 16:372-380.

Schluter D, McPhail JD. 1993. Character displacement and replicate adaptive radiation. Trends Ecol. Evol. 8: 197-200.

Schneider RA. 2006. How to tweak a beak: molecular techniques for studying the evolution of size and shape in Darwin's finches and other birds. BioEssays 29:1-6.

Schulte PM. 2001. Environmental adaptations as windows on molecular evolution. Comp. Biochem. Physiol. pt.B 128:597-611.

Schulte PM. 2007. Responses to environmental stressors in an estuarine fish: interacting stressors and the impact of local adaptation. J. Thermal Biol. 32:152-161.

Simpson GG. 1953. The major features of evolution. Columbia Univiversity Press, New York.

Singh RS. 2003. Darwin to DNA, molecules to morphology: the end of classical population genetics and the road ahead. Genome 46:938-942.

Sokal RR, Michener CD. 1958. A statistical method for evaluation systematic relationships. Univ. Of Kans. Sci. Bull. 38:1409-1438.

Somero GN, Childress JJ. 1980. A violation of the metabolism-size scaling paradigm – activities of glycolytic-enzymes in muscle increase in larger-size fish. Physiol. Zool. 53:322–337.

Stanford SJ, Walters MJ, Hislop AA, Haworth SG, Evans TW, Mann BE, Motterlini R, Mitchell JA. 2003. Heme oxygenase is expressed in human pulmonary artery smooth muscle where carbon monoxide has an anti-proliferative role. Euro. J. Pharma. 473:135-141.

Stearns SC, Magwene P. 2003. The naturalist in a world of genomics. Am. Nat. 161: 171-180.

Street NR, Skogström O, Sjödin A, et al. (8 co-authors). 2006. The genetics and genomics of the drought response in Populus. The Plant Journal, 48, 321-341.

Streisfeld MA, Kohn JR. 2005. Contrasting patterns of floral and molecular variation across a cline in Mimulus aurantiacus. Evolution 59:2548-2559.

Taylor EB. 1999. Species pairs of north temperate freshwater fishes: evolution, taxonomy and conservation. Rev. Fish Biol. Fisher. 9:299-324.

Terai Y, Seehausen O, Sasaki T, et al. (14 co-authors). 2006. Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. PLoS Biol. 4:2244-2251.

Tortora GJ, Grabowski SR. 1993. Principales of anatomy and physiology. Seventeenth edition. Biological Textbooks, Inc., A and P Textbooks, Inc.

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*). Can. J. Fish. Aquat. Sci. 58:394-405.

Valdimarsson SK, Metcalfe NB, Skulason S. 2000. Experimental demonstration of differences in sheltering behaviour between Icelandic populations of Atlantic salmon (*Salmo salar*) and Arctic char (*Salvelinus alpinus*). Can. J. Fisher. Aqua. Sc. 57: 719-724.

von Schalburg KR, Rise ML, Cooper GA, Brown GD, Gibbs AR, Nelson CC, Davidson WS, Koop BF. 2005. Fish and chips: various methodologies demonstrate utility of a 16,006-gene salmonid microarray. BMC Genomics 6:126.

Wagner A. 2007. Energy costs constrain the evolution of gene expression. J. Exp. Zool. 308B:322–324.

Wayne ML, McIntyre LM. 2002. Combining mapping and arraying: an approach to candidate gene identification. PNAS 99: 14903-14906.

Wen C, Zhang Z, Ma W, Xu M, Wen Z, Peng J. 2005. Genome-wide identification of female-enriched genes in zebrafish. Dev. Dynamics. 232:171-179.

West-Eberhart MJ. 2005. Developmental plasticity and the origin of species differences. Proc. Natl. Acad. Sci. USA (Suppl.1) 102:6543-6549.

Whitehead A, Crawford DL. 2005. Variation in tissue-specific gene expression among natural populations. Genome Biol. 6:R13.

---. 2006a. Variation within and among species in gene expression: raw material for evolution. Mol. Ecol. 15:1197-1211.

---. 2006b. Neutral and adaptive variation in gene expression. PNAS 103: 5425-5430.

Whiteley AR, Derome N, Rogers SM, St-Cyr J, Labbe A, Bernatchez L. The phenomics of brain transcriptomes regulating adaptive divergence in lake whitefish species pairs from eQTL mapping (*en préparation*).

Willmer P, Stone G, Johnston I. 2000. Environmental physiology of animals. Blackwell Science, Oxford.

ANNEXES

Supplementary Table 1 : Genes showing non-directional transcription differences in natural populations

Gene clone number and name: (1)	Transcription ratio: (2) Cliff Lake	Indian Pond	*P*-CL (3)	*P*-IP (3)
CA061100 [GO] [Q9DCD0] 6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44).	0,8535814	1,2914683	0,0164475	0,0027954
CB497730 [NR] [AAW82445] 14 kDa apolipoprotein [Carassius auratus gibelio]	1,2929791	0,9201138	0,0041108	0,0401685
CB498079 [NR] [AAW82445] 14 kDa apolipoprotein [Carassius auratus gibelio]	1,188727	0,8646246	0,0142863	0,0087909
CB516377 [GO] [P25444] 40S ribosomal protein S2 (S4) (LLRep3 protein).	0,8588715	1,1154307	0,0333704	0,0495013
CA058850 [GO] [P62908] 40S ribosomal protein S3.	0,8901225	1,176212	0,0292778	0,0287414
CB497613 [GO] [P97351] 40S ribosomal protein S3a.	0,7403824	1,1268996	0,001624	0,0336292
CB505295 [GO] [P97351] 40S ribosomal protein S3a.	0,8328921	1,1541075	0,0121621	0,0216174
CB504457 [GO] [P62082] 40S ribosomal protein S7.	0,8386863	1,125131	0,0269551	0,0326026
CB496418 [NR] [AAG22824] 40S ribosomal protein S11 [Salmo salar]	0,801525	1,1171157	0,0059836	0,0379088
CB493119 [GO] [P62301] 40S ribosomal protein S13.	0,9012127	1,1345675	0,0484058	0,0345283
CB495065 [GO] [P62843] 40S ribosomal protein S15 (RIG protein).	0,878733	1,1479677	0,0250722	0,0358114
CK990478 [GO] [P63276] 40S ribosomal protein S17.	1,1931524	0,8873098	0,0173195	0,0363664
CB496672 [GO] [P47955] 60S acidic ribosomal protein P1.	0,7565745	1,1280028	0,0019158	0,0282979
CA052925 [GO] [P27659] 60S ribosomal protein L3 (J1 protein).	0,8886622	1,1971573	0,0412818	0,0098988
CK991334 [GO] [P14148] 60S ribosomal protein L7.	0,8830969	1,1461774	0,048202	0,0278451
CK990919 [GO] [P46223] 60S ribosomal protein L7a.	0,8126682	1,1258878	0,0075031	0,0066837
CB500821 [GO] [P62918] 60S ribosomal protein L8.	1,1382444	0,8765825	0,0391771	0,0226549
CB515441 [GO] [P62717] 60S ribosomal protein L18a.	0,8905062	1,2408083	0,0318945	0,0141427
CA039207 [GO] [P62717] 60S ribosomal protein L18a.	0,8403178	1,1648097	0,0178435	0,0198108
CB496920 [GO] [P62717] 60S ribosomal protein L18a.	0,7856628	1,1911484	0,0060451	0,0143243
CA051876 [GO] [P84099] 60S ribosomal protein L19.	0,8939194	1,2320078	0,038476	0,0337738
CA046247 [GO] [O09167] 60S ribosomal protein L21.	0,8181574	1,1298187	0,0089069	0,0303198
CB516607 [GO] [P67984] 60S ribosomal protein L22 (Heparin binding protein HBp15).	0,7334133	1,2387565	0,0014012	0,004641
CA042605 [GO] [P61255] 60S ribosomal protein L26 (Silica-induced gene 20 protein) (SIG-20).	0,8121284	1,2211351	0,0066016	0,0134875
CB501837 [GO] [P41105] 60S ribosomal protein L28.	1,3965847	0,8069692	0,0015924	0,0075466
CK990714 [GO] [P62900] 60S ribosomal protein L31.	0,8543322	1,1483425	0,0332954	0,0338281
CB492180 [GO] [P62900] 60S ribosomal protein L31.	0,808849	1,266591	0,008249	0,0072207
CB500526 [GO] [Q9D823] 60S ribosomal protein L37.	0,893476	1,0953247	0,0331469	0,0155827
CB488850 [GO] [Q9D823] 60S ribosomal protein L37.	0,827271	1,1771245	0,013897	0,0151888
CB497286 [GO] [Q9D823] 60S ribosomal protein L37.	0,9006862	1,1721616	0,0354954	0,0171317
CB493464 [GO] [Q07456] AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha- trypsin inhibitor light chain (ITI-LC)	1,127925	0,8964802	0,0332167	0,0433745
CB497703 [NR] [CAA57449] apolipoprotein B [Salmo salar]	1,2608731	0,8333783	0,010492	0,0174915
CA050658 [GO] [P05202] Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1) (Transaminase A)	0,869375	1,1815384	0,0296298	0,0153204

Supplementary Table 1 : (continued from previous page)

Gene clone number and name: (1)	Transcription ratio: (2) Cliff Lake	Indian Pond	*P*-CL (3)	*P*-IP (3)
CA060011 [GO] [O54962] Barrier-to-autointegration factor (Breakpoint cluster region protein 1) (LAP2-binding protein 1).	1,310232	0,7576417	0,0139012	0,009815
BU965641 [GO] [Q05793] Basement membrane-specific heparan sulfate proteoglycan core protein precursor (HSPG) (Perlecan) (PLC).	1,326111	0,7066325	0,0061965	0,0015124
CN442516 [GO] [P01887] Beta-2-microglobulin precursor.	1,1472285	0,7289341	0,0481265	0,0010244
CB505897 [GO] [P01887] Beta-2-microglobulin precursor.	1,1409032	0,8422559	0,0460289	0,0112979
CA051169 [NR] [NP_001012498] betaine-homocysteine methyltransferase [Danio rerio]	0,8686761	1,2530698	0,0165629	0,006828
CA060239 [GO] [O89110] Caspase-8 precursor (EC 3.4.22.-) (CASP-8).	1,2684049	0,7268547	0,005845	0,0012881
CA039257 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41) (Complement component 1, r subcomponent).	1,8668636	0,7530046	3,60E-05	0,0022026
CB510638 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41) (Complement component 1, r subcomponent).	1,2471935	0,788319	0,0030601	0,0091906
CB510300 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41) (Complement component 1, r subcomponent).	1,5652076	0,8516593	7,74E-05	0,042111
CB514355 [GO] [P01027] Complement C3 precursor (HSE-MSF) [Contains: C3a anaphylatoxin].	1,324352	0,7503689	0,0018762	0,0014085
CB493612 [GO] [P00417] Cytochrome c oxidase subunit 3 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III).	1,1334904	0,8229974	0,0477583	0,0134202
CK991041 [GO] [P00417] Cytochrome c oxidase subunit 3 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III).	1,1456101	0,8342477	0,0112632	0,0156922
CK990846 [GO] [P00417] Cytochrome c oxidase subunit 3 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III).	1,2372222	0,8867563	0,0103508	0,0157774
CA060609 [NT] [BC045512] Danio rerio ubiquitin-conjugating enzyme E2G 2, mRNA (cDNA clone MGC:55974 IMAGE:3819548)	1,3368066	0,8086747	0,0034803	0,0087938
CA039356 [GO] [Q9JIQ3] Diablo homolog, mitochondrial precursor (Second mitochondria-derived activator of caspase)	1,6028523	0,636988	7,97E-05	9,07E-05
CB492655 [GO] [P20067] DNA-binding protein inhibitor ID-1 (ID).	0,8613388	1,1573218	0,0238904	0,0217362
CB493567 [GO] [P20067] DNA-binding protein inhibitor ID-1 (ID).	0,8212071	1,1354374	0,0081182	0,0260727
CB496981 [GO] [Q24320] DNA-directed RNA polymerases I, II, and III 14.4 kDa polypeptide (EC 2.7.7.6) (RPB6).	1,6892372	0,8086576	0,0004838	0,0108693
CA054908 [NR] [AAH56554] E74-like factor 2 (ets domain transcription factor) [Danio rerio]	1,1403265	0,8518737	0,027521	0,0212761
CA052512 [GO] [Q9D8N0] Elongation factor 1-gamma (EF-1-gamma) (eEF-1B gamma).	0,8673114	1,1474487	0,0454309	0,0236924
CB510664 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,6367748	1,403316	0,0014497	0,0004896
CK991218 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,7133696	1,5270449	0,0019522	0,0002086
CK990756 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,8115752	1,307805	0,0085015	0,0017739
CB516896 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,8477653	1,2012827	0,0277657	0,0117616
CK990586 [GO] [P09528] Ferritin heavy chain (Ferritin H subunit).	1,591863	0,7964209	0,0001198	0,0068001
CA051332 [GO] [P09528] Ferritin heavy chain (Ferritin H subunit).	1,3575705	0,8772595	0,0009518	0,0267936
CB496788 [NR] [AAQ91198] glutathione S-transferase [Micropterus salmoides]	0,875763	1,1483632	0,0379052	0,0288589
CB496919 [GO] [P13707] Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic (EC 1.1.1.8) (GPD-C) (GPDH-C).	1,1641061	0,7352983	0,0337665	0,0018939
CA055402 [GO] [P68040] Guanine nucleotide-binding protein beta subunit 2-like 1 (Receptor of activated protein kinase C 1) (RACK1)	0,8567416	1,1437714	0,0215257	0,0419467
CA047661 [GO] [P62874] Guanine nucleotide-binding protein G(I)/G(S)/G(T) beta subunit 1 (Transducin beta chain 1).	0,7740878	1,1212062	0,0054109	0,0492544
CK990501 [GO] [Q80SZ7] Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-5 subunit.	1,6226548	0,8541217	0,0008249	0,0295149
CK991158 [GO] [P63017] Heat shock cognate 71 kDa protein (Heat shock 70 kDa protein 8).	0,9101976	1,3447582	0,0492489	0,0010898

Supplementary Table 1 : (continued from previous page)

Gene clone number and name: (1)	Transcription ratio: (2) Cliff Lake	Indian Pond	*P*-CL (3)	*P*-IP (3)
CA049318 [GO] [P01942] Hemoglobin alpha chain.	1,5982282	0,7203088	0,0003311	0,0006639
CA064277 [GO] [P01942] Hemoglobin alpha chain.	1,2515691	0,7214393	0,0105702	0,0012481
CA058603 [GO] [P01942] Hemoglobin alpha chain.	1,3683543	0,7101354	0,0005109	0,0004077
CB492263 [GO] [P01942] Hemoglobin alpha chain.	1,4021603	0,9282511	0,0008027	0,0362984
CA061084 [GO] [P01942] Hemoglobin alpha chain.	1,3247554	0,747976	0,0065487	0,0042899
CB500796 [GO] [P01942] Hemoglobin alpha chain.	1,2440367	0,8589644	0,0123512	0,018987
CA058361 [GO] [P01942] Hemoglobin alpha chain.	1,7614824	0,8508082	3,60E-05	0,0106739
CA770045 [GO] [P01942] Hemoglobin alpha chain.	1,5910308	0,7998175	0,0001748	0,0062926
CA049305 [GO] [P01942] Hemoglobin alpha chain.	1,623414	0,7645175	0,000321	0,0041856
CA049294 [GO] [P01942] Hemoglobin alpha chain.	1,3774295	0,762419	0,0060412	0,0033207
CB498419 [GO] [P01942] Hemoglobin alpha chain.	1,3566779	0,7460976	0,0028654	0,0017871
CA770045 [GO] [P01942] Hemoglobin alpha chain.	1,790727	0,7717601	3,02E-05	0,0009132
CB492226 [GO] [P01942] Hemoglobin alpha chain.	1,4330296	0,7216406	0,0016083	0,0019804
CK991045 [GO] [P01942] Hemoglobin alpha chain.	1,3907317	0,7709938	0,0040723	0,0042757
CB496604 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,7577922	0,8401846	5,13E-05	0,0017818
CA047133 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,5729979	0,8350049	0,0002986	0,0149645
CK990563 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,5673625	0,5732896	0,0002347	0,000157
CA049300 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,2894709	0,7303951	0,0327449	0,002342
CB492133 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,5014561	0,7022301	0,0006058	0,000767
CB494536 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,5364343	0,7056865	0,0007154	0,0010148
CB498575 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,7952463	0,6781499	3,60E-05	0,0004291
CA041804 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,4078289	0,8576216	0,0030067	0,010035
CB498416 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,4495517	0,6262628	0,0008798	0,0003812
CB498665 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,6576605	0,8086488	7,92E-05	0,0027864
CA058206 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,3200056	0,8566068	0,0047421	0,0190774
CB497335 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,7260998	0,8696993	0,0008066	0,0027135
CB497723 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,7038172	0,8695153	0,0001907	0,0089738
CB500031 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,2244601	0,6542699	0,0200729	0,0010604
CB497309 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,590527	0,6132346	0,0003361	0,0002843
CK990457 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,2475128	0,8875095	0,0159199	0,028051

Supplementary Table 1 : (continued from previous page)

Gene clone number and name: (1)	Transcription ratio: (2) Cliff Lake	Indian Pond	*P*-CL (3)	*P*-IP (3)
BU965636 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,4742553	0,7080418	0,0016974	0,0024294
CB510387 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,4898227	0,7079391	0,0004854	0,001125
CB509758 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,7395007	0,6758907	0,0003095	0,0008486
CA036686 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,2232728	0,8963907	0,0142769	0,0379732
CB501013 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,7968274	0,7559762	3,67E-05	0,0004929
CK991073 [GO] [P06467] Hemoglobin zeta chain (Alpha-like embryonic globin chain x).	1,1992464	0,8104988	0,0225064	0,0065825
CA039501 [NR] [AAO85553] hepcidin [Salmo salar]	0,7901909	1,2798691	0,0047695	0,0027221
CK991068 [NR] [AAO85553] hepcidin [Salmo salar]	0,524843	1,8395147	6,64E-05	2,19E-05
CA051011 [NT] [BC025274] Homo sapiens ARP1 actin-related protein 1 homolog B, centractin beta (yeast), mRNA	1,776363	0,717199	3,05E-05	0,000926
CB509856 [NR] [AAQ55174] inhibitor of differentiation 1C; Id 6-2 [Oncorhynchus mykiss]	0,8698884	1,1184126	0,0264864	0,0397959
CA039272 [GO] [Q61704] Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3) (Inter-alpha-inhibitor heavy chain 3)	1,1910952	0,6777096	0,0079061	0,0005268
CA037372 [GO] [Q61704] Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3) (Inter-alpha-inhibitor heavy chain 3)	1,2064306	0,5823486	0,0102179	0,0001998
CB497601 [GO] [O08677] Kininogen precursor [Contains: Bradykinin].	1,1684249	0,8530293	0,0294313	0,0262993
CB494268 [GO] [O08677] Kininogen precursor [Contains: Bradykinin].	1,2305951	0,8303762	0,0058908	0,0170806
CA043151 [NR] [P01044] Kininogen 1 precursor (Kininogen I) (Thiol proteinase inhibitor)	1,1751697	0,7417702	0,0148381	0,0028131
CA051373 [NT] [AY071855] Oncorhynchus mykiss clone GC41 MHC class I antigen (Onmy-U41p) pseudogene, partial sequence	1,2073905	0,8417605	0,0065031	0,0120989
CA050896 [NT] [AJ250180] Oncorhynchus mykiss mRNA for beta thymosin	1,1356041	0,7512258	0,0348297	0,0039837
CB509775 [GO] [P19221] Prothrombin precursor (EC 3.4.21.5).	1,2023084	0,8303019	0,0179034	0,0096624
CB493471 [GO] [P19221] Prothrombin precursor (EC 3.4.21.5).	1,1450311	0,8504191	0,0427137	0,0265312
CA041430 [GO] [P26350] Prothymosin alpha.	1,1532836	0,7946179	0,0359743	0,0064425
CA037810 [NR] [CAH57705] putative 14 kDa apolipoprotein [Platichthys flesus]	1,2516324	0,8523353	0,0022193	0,0104928
CA038051 [NR] [CAH57705] putative 14 kDa apolipoprotein [Platichthys flesus]	1,3521329	0,7968231	0,0030976	0,0074946
CA054395 [GO] [O35704] Serine palmitoyltransferase 1 (EC 2.3.1.50) (Long chain base biosynthesis protein 1) (LCB 1)	0,7859322	1,1289384	0,0036874	0,0466179
CB510002 [GO] [P12246] Serum amyloid P-component precursor (SAP).	1,2033678	0,6632703	0,0091557	0,0002394
CB497429 [GO] [Q9CYP7] Sestrin 3.	0,7558547	1,2909638	0,0023136	0,0025296
CK990717 [NT] [X97287] S.salar gene encoding beta-globin, clone 4	1,7320212	0,8064659	0,0004784	0,0001678
CA042931 [GO] [P12613] T-complex protein 1, alpha subunit (TCP-1-alpha) (CCT-alpha).	0,813687	1,1899367	0,0084207	0,0108516
CB493361 [NR] [AAM21198] toxin-1 [Oncorhynchus mykiss]	1,1828529	0,8563697	0,0260224	0,0229925
CA060308 [GO] [Q93092] Transaldolase (EC 2.2.1.2).	0,8611856	1,2924804	0,0233815	0,0043073
CA063762 [GO] [Q64152] Transcription factor BTF3 (RNA polymerase B transcription factor 3).	0,8792952	1,139654	0,0407578	0,0300774
CB516494 [GO] [Q9EPF4] Transforming growth factor-beta-inducible early growth response protein	1,8153147	0,7183712	7,45E-05	0,0015403
CA048021 [GO] [P62991] Ubiquitin.	0,8634597	1,1354999	0,0216661	0,0411776
CA053740 [GO] [Q9R0P9] Ubiquitin carboxyl-terminal hydrolase isozyme L1 (EC 3.4.19.12) (UCH- L1) (Ubiquitin thiolesterase L1)	1,7408542	0,7497077	8,31E-05	0,0011276

Supplementary Table 1 : (continued from previous page)

Gene clone number and name: (1)	Transcription ratio: (2) Cliff Lake	Indian Pond	*P*-CL (3)	*P*-IP (3)
CA037818 UNKNOWN	1,3851864	0,7424035	0,0024879	0,0036927
CA043177 UNKNOWN	1,1484189	0,852223	0,0425894	0,0161317
CA046756 UNKNOWN	0,8542244	1,1289523	0,0142331	0,0260278
CA047023 UNKNOWN	1,1133419	0,8695654	0,0319703	0,024425
CA050489 UNKNOWN	0,8640453	1,1235199	0,0388551	0,0349785
CA053773 UNKNOWN	1,5911765	0,6572318	0,0003609	0,0001344
CA055233 UNKNOWN	0,8515824	1,1849392	0,017209	0,013032
CA059099 UNKNOWN	0,8719047	1,0807217	0,0345809	0,0396422
CA060279 UNKNOWN	1,4487186	0,6196508	0,0008277	0,0003311
CA060896 UNKNOWN	1,8993484	0,8423187	4,06E-05	0,0026786
CA061786 UNKNOWN	1,422862	0,7585834	0,0030335	0,0023701
CA062447 UNKNOWN	0,8895198	1,1737141	0,0412932	0,0106645
CB494013 UNKNOWN	1,1490451	0,8271586	0,0372955	0,0124177
CB494761 UNKNOWN	1,5961612	0,7190478	0,0004497	0,0019775
CB496908 UNKNOWN	1,1666538	0,8815144	0,0327089	0,0443757
CB497819 UNKNOWN	1,1412675	0,8414706	0,0150916	0,0314466
CB499949 UNKNOWN	1,1816692	0,8448579	0,0155259	0,0168141
CB499950 UNKNOWN	0,8464576	1,111588	0,0130273	0,0350522
CB499966 UNKNOWN	1,3242245	0,857042	0,0024249	0,0116694
CB500536 UNKNOWN	0,8497397	1,2233914	0,0178885	0,0065386
CB500553 UNKNOWN	1,168808	0,8770597	0,019414	0,0350332
CB501180 UNKNOWN	1,2061536	0,8720134	0,0193843	0,0213774
CB512159 UNKNOWN	1,1460818	0,8140583	0,0332284	0,0103404
CK990629 UNKNOWN	0,8441009	1,1555147	0,0242765	0,0235451
CK990889 UNKNOWN	0,8574793	1,4221237	0,0147047	0,000728
CK991008 UNKNOWN	1,5780597	0,7486218	0,0002183	0,0045003
CK991021 UNKNOWN	1,698344	0,7288652	0,0001725	0,0008568
CK991159 UNKNOWN	0,8790365	1,1459921	0,0289371	0,0225621

(1): Each EST clone number represents a single EST sequence;
(2): Transcription ratios correspond to mean dwarf expression level divided by mean normal expression level (D/N);
(3): P-values for Cliff Lake and Indian pond obtained from ANOVA, F3 test, 1000 permutations;

Supplementary Table 2 : Genes showing non-directional transcription differences in natural populations and under controlled conditions

Gene clone number and name: (1)	Transcription ratio: (2)					
	Cliff Lake	Indian Pond	Control	*P*-CL (3)	*P*-IP (3)	*P*-ctl (3)
CB496418 [NR] [AAG22824] 40S ribosomal protein S11 [Salmo salar]	0,801525	1,1171157	0,7320699	0,0059836	0,0379088	0,0108505
CB496672 [GO] [P47955] 60S acidic ribosomal protein P1.	0,7565745	1,1280028	0,7963657	0,0019158	0,0282979	0,0252264
CB493464 [GO] [Q07456] AMBP protein precursor [Contains: Alpha-1-microglobulin; Inter-alpha- trypsin inhibitor light chain]	1,127925	0,8964802	1,4024043	0,0332167	0,0433745	0,0044177
CA060011 [GO] [O54962] Barrier-to-autointegration factor (Breakpoint cluster region protein 1) (LAP2-binding protein 1).	1,310232	0,7576417	1,4000171	0,0139012	0,009815	0,0227974
CN442516 [GO] [P01887] Beta-2-microglobulin precursor.	1,1472285	0,7289341	1,3154655	0,0481265	0,0010244	0,0417846
CA051169 [NR] [NP_001012498] betaine-homocysteine methyltransferase [Danio rerio]	0,8686761	1,2530698	1,3670548	0,0165629	0,006828	0,0173023
CA039257 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41)	1,8668636	0,7530046	1,4450816	3,60E-05	0,0022026	0,0014463
CB510638 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41)	1,2471935	0,788319	1,4426516	0,0030601	0,0091906	0,0022834
CB510300 [GO] [Q8CG16] Complement C1r subcomponent precursor (EC 3.4.21.41)	1,5652076	0,8516593	1,5360994	7,74E-05	0,042111	0,0061199
CB493612 [GO] [P00417] Cytochrome c oxidase subunit 3 (EC 1.9.3.1) (Cytochrome c oxidase polypeptide III).	1,1334904	0,8229974	0,7883161	0,0477583	0,0134202	0,028707
CA039356 [GO] [Q9JIQ3] Diablo homolog, mitochondrial precursor (Second mitochondria-derived activator of caspase)	1,6028523	0,636988	5,1455899	7,97E-05	9,07E-05	4,95E-06
CB510664 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,6367748	1,403316	0,5738228	0,0014497	0,0004896	0,0003327
CK991218 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,7133696	1,5270449	0,605616	0,0019522	0,0002086	0,000903
CB516896 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,8477653	1,2012827	0,6654061	0,0277657	0,0117616	0,0035887
CK990756 [GO] [P51880] Fatty acid-binding protein, brain (B-FABP) (Brain lipid-binding protein) (BLBP).	0,8115752	1,307805	0,6254347	0,0085015	0,0017739	0,0011404
CB504246 [NR] [AAU21486] fucolectin [Fundulus heteroclitus]	0,8575269	0,8897919	1,2456885	0,0204173	0,0409973	0,0145531
CB505763 [GO] [P23591] GDP-L-fucose synthetase (EC 1.1.1.271) (FX protein) (Red cell NADP(H)- binding protein)	1,1711345	1,1174343	0,7915311	0,0355045	0,039267	0,0322678
CB496788 [NR] [AAQ91198] glutathione S-transferase [Micropterus salmoides]	0,875763	1,1483632	1,8562218	0,0379052	0,0288589	0,0009049
CB496919 [GO] [P13707] Glycerol-3-phosphate dehydrogenase [NAD+], cytoplasmic (EC 1.1.1.8) (GPD-C) (GPDH-C)	1,1641061	0,7352983	0,6786473	0,0337665	0,0018939	0,0040426
CK990501 [GO] [Q80SZ7] Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-5 subunit.	1,6226548	0,8541217	0,7161969	0,0008249	0,0295149	0,0053724
CA061084 [GO] [P01942] Hemoglobin alpha chain.	1,3247554	0,747976	0,815165	0,0065487	0,0042899	0,0373036
CB500796 [GO] [P01942] Hemoglobin alpha chain.	1,2440367	0,8589844	0,8052556	0,0123512	0,018987	0,0411896
CB492263 [GO] [P01942] Hemoglobin alpha chain.	1,4021603	0,9282511	0,7759731	0,0008027	0,0362984	0,0470762
CK990870 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,3955736	0,9643801	0,8126101	0,0009794	0,0216354	0,0288842
CA047133 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,5729979	0,8350049	0,7631862	0,0002986	0,0149645	0,0252798
CA049300 [GO] [P02089] Hemoglobin beta-2 chain (B2) (Hemoglobin beta-minor chain).	1,2894709	0,7303951	0,3685128	0,0327449	0,002342	0,0021338
CA036686 [GO] [P02104] Hemoglobin epsilon-Y2 chain.	1,2232728	0,8963907	0,8314166	0,0142769	0,0379732	0,0312866
CA039272 [GO] [Q61704] Inter-alpha-trypsin inhibitor heavy chain H3 precursor (ITI heavy chain H3)	1,1910952	0,6777096	1,487788	0,0079061	0,0005268	0,009275

Supplementary Table 2 : (continued from previous page)

Gene clone number and name: (1)	Transcription ratio: (2)					
	Cliff Lake	Indian Pond	Control	*P*-CL (3)	*P*-IP (3)	*P*-ctl (3)
CB502941 [NR] [AAS89353] perforin [Ctenopharyngodon idella]	0,4480266	0,3438191	9,7884408	1,82E-06	0	1,30E-06
CB511632 [GO] [Q60692] Proteasome subunit beta type 6 precursor (EC 3.4.25.1)	0,7825467	0,7996623	1,4450844	0,0051686	0,0051947	0,0081211
CA061998 [GO] [P23298] Protein kinase C, eta type (EC 2.7.1.-) (nPKC-eta) (PKC-L).	1,3952612	1,1503632	0,8994678	0,0008955	0,0122111	0,0300543
CB510002 [GO] [P12246] Serum amyloid P-component precursor (SAP).	1,2033678	0,6632703	1,4208429	0,0091557	0,0002394	0,0053567
CB493361 [NR] [AAM21198] toxin-1 [Oncorhynchus mykiss]	1,1828529	0,8563697	1,1249195	0,0260224	0,0229925	0,0236702
CB497373 [GO] [P13412] Troponin I, fast skeletal muscle (Troponin I, fast-twitch isoform).	1,1235968	1,2460832	0,5039127	0,0247319	0,0031124	0,0001363
CA062447 UNKNOWN	0,8895198	1,1737141	1,377127	0,0412932	0,0106645	0,0130305
CA060896 UNKNOWN	1,8993484	0,8423187	4,1959425	4,06E-05	0,0026786	2,45E-05
CB497819 UNKNOWN	1,1412675	0,8414706	0,8471984	0,0150916	0,0314466	0,0405213
CB494761 UNKNOWN	1,5961612	0,7190478	1,6032603	0,0004497	0,0019775	0,005702
CA043177 UNKNOWN	1,1484189	0,852223	1,1790041	0,0425894	0,0161317	0,0417046
CA059099 UNKNOWN	0,8719047	1,0807217	1,4713014	0,0345809	0,0396422	0,0065
CK990291 UNKNOWN	0,8719434	0,8730443	1,3398755	0,0360657	0,0165103	0,0425772
CA046404 UNKNOWN	1,3015541	1,1597577	0,795569	0,0044248	0,0214824	0,0376016

(1): Each EST clone number represents a single EST sequence;
(2): Transcription ratios correspond to mean dwarf expression level divided by mean normal expression level (D/N);
(3): P-values for Cliff Lake, Indian Pond and control obtained from ANOVA, F3 test, 1000 permutations;