Université de Montréal

LES IMPACTS DE MITOCHONDRIES ÉTRANGÈRES SUR LE PHÉNOTYPE : UNE ÉTUDE HOLISTIQUE CHEZ LE VENTRE ROUGE DU NORD *CHROSOMUS EOS*

par
Léo Deremiens

Département de sciences biologiques
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
En vue de l'obtention du grade de Ph.D.
En sciences biologiques

03 mai 2016

RÉSUMÉ

Les organismes cybrides souffrent généralement d'une altération de la spécificité des interactions mito-nucléaires, résultant en une détérioration du phénotype. Toutefois, diverses études démontrent que le transfert de mitochondries peut occasionnellement être positif. À l'heure actuelle, de nombreuses questions demeurent quant au degré d'influence de ces transferts sur les différents niveaux d'organisation du phénotype. Afin de répondre à ces questions, les formes sauvages et cybrides du poisson *Chrosomus eos* sont étudiées. Ainsi, le premier volet de ce projet de recherche démontre un impact des mitochondries *Chrosomus neogaeus* à différents niveaux d'organisation du phénotype des poissons *C. eos*, lorsque les formes sauvages et cybrides sont retrouvées en allopatrie. Le deuxième volet de cette thèse révèle, quant à lui, que ces modifications phénotypiques ne sont pas suffisantes pour induire un évènement de spéciation entre les deux biotypes, lorsqu'en sympatrie. De plus, cette étude suggère que la coévolution mito-nucléaire peut ne pas être une condition *sine qua non* à la perpétuation des individus en milieu naturel. Finalement, l'approche holistique considérée dans le troisième volet de cette recherche atteste de l'influence des mitochondries *C. neogaeus* à différents niveaux d'organisation du phénotype de *C. eos*, lorsque les formes sauvages et cybrides sont sympatriques. Cette influence est moins prononcée que celle observée à partir de biotypes allopatriques. Combinés, ces chapitres contribuent à une meilleure compréhension des liens existant entre les mitochondries et le phénotype d'un individu.

Mots-clés : mitochondries, cybrides, phénotype, coévolution mito-nucléaire, *Chrosomus eos*, méthylome, transcriptome, protéome, activité enzymatique, performance de nage, spéciation.

ABSTRACT

Cybrid organisms generally suffer from a disruption of the mito-nuclear interactions specificity, resulting in a phenotype deterioration. However, several studies demonstrate that between-species transfers of mitochondria can occasionally be beneficial. Currently, many questions still remain about how much these transfers can impact the various biological organisation levels of the phenotype. To address these questions, *Chrosomus eos* wild type and cybrids fish are considered. Thus, the first part of this research project demonstrates that *Chrosomus neogaeus* mitochondria impact the *C. eos* phenotype, at various levels of biological organisation, when wild type and cybrids are found in allopatry. On the other hand, the second part of this thesis reveals that these phenotypic modifications cannot trigger a speciation event between the two sympatric biotypes. This study also suggests that mito-nuclear coevolution would not be a *sine qua non* condition to survive and perform in natural environments. Finally, the holistic approach considered in the third part of this research shows that *C. neogaeus* mitochondria influence various levels of *C. eos* phenotype, when wild type and cybrids are sympatric. The magnitude of this influence is less pronounced than the one observed from allopatric biotypes. Combined, all these parts contribute to a better understanding of the links existing between mitochondria and the phenotype of an individual.

Keywords: mitochondria, cybrids, phenotype, mito-nuclear coevolution, *Chrosomus eos*, methylome, transcriptome, proteome, enzymatic activity, swimming performance, speciation.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

LISTE DES SIGLES ET ABRÉVIATIONS

A : Nombre total d'allèles

AC : Cybride Atlantique

ADN / DNA : Acide désoxyribonucléique

ADNc / cDNA : Acide désoxyribonucléique complémentaire

ADNmt / mtDNA : Acide désoxyribonucléique mitochondrial

ADNnc / ncDNA : Acide désoxyribonucléique nucléaire

A_R : Richesse allélique

ARN / RNA : Acide ribonucléique

ARNm / mRNA : Acide ribonucléique messager

ARNr / rRNA : Acide ribonucléique ribosomique

ARNt / tRNA : Acide ribonucléique de transfert

ATP : Adénosine triphosphate

AMOVA : Analyse de variance moléculaire

Brin H : Brin lourd de l'ADNmt

Brin L : Brin léger de l'ADNmt

BSA : Albumine de sérum bovin

C : Complexe

CYT : cytochrome c

CYTB : ubiquinone cytochrome c oxydoréductase

COX : cytochrome c oxydase

DCPIP : Dichlorophénolindophénol

df : Degré de liberté

D-loop: Boucle de déplacement

e^- : Electron

$FADH_2$: forme oxydée de la flavine adénine dinucléotide

F_{CT} : Indice de variation génétique entre les groupes

FQRNT : Fonds Québécois de Recherche en Nature et Technologie

F_{ST} : Indice de variation génétique inter populations

Fstat : Statistiques des indices de variation génétique

GRIL : Groupe de Recherche Interuniversitaire en Limnologie et en environnement aquatique

H^+ : Proton

H_E : Hétérozygotie attendue

H_{EQ} : Hétérozygotie attendue à l'équilibre mutation-dérive

IAM : modèle mutationnel à nombre infini d'allèle

IPG : Gradient de pH immobilisé

IRIC : Institut de Recherche en Immunologie et Cancérologie

K : Nombre de populations

kb : kilobase

KCN : Cyanure de potassium

LDH : Lactate déshydrogénase

m : Taux de migration par génération

M : Taux de migration « mutation-scaled »

m.y. : Million d'années

MC : Cybride Mississipien

MCMC : Chaîne de Markov Monte Carlo

MDE : Equilibre mutation-dérive

Met : Méthionine

mg : Milligramme

mL : Millilitre

mM : Millimolaire

MSAP : Polymorphisme amplifié sensible à la méthylation

NADH : forme oxydée de la nicotinamide adénine dinucléotide

ND : NADH-ubiquinone oxydoréductase

N_e : Taille de la population efficace

nm : Nanomètre

nmol : Nano mole

P-value / *P* : valeur de probabilité

pb / bp : paire de bases

PCR : Réaction en chaîne par polymérase

qRT-PCR : Réaction en chaîne par polymérase quantitative à partir d'ARN rétro-transcrit

R^2 : Coefficient de détermination

RDA : Analyse de redondance

RFLP : Polymorphisme de longueur des fragments de restriction

RU_{crit} : Test de vitesse de nage critique relative

ROS : Espèces réactives de l'oxygène

SDS : Dodécylsulfate de sodium

SMM : Modèle mutationnel par pas

SS : Somme des carrés

SSCP : Polymorphisme de conformation des simples brins

T_i : Dure de la vitesse de nage finale

T_{ii} : Durée du temps d'incrément

TBE : Tris/Borate/EDTA

Thr : Thréonine

TL : Longueur totale du poisson

U_{crit} : Test de vitesse de nage critique

U_i : Vitesse de nage finale

U_{ii} : Incrément de la vitesse de nage

UQ : Ubiquinone

UV/VIS : Ultraviolet-visible

VC : Composantes de la variance

W : Ouest / Watt

WT : Forme sauvage

°C : Degré Celsius

2D : Deux dimensions

μ : Taux de mutation par génération par locus

µg : Microgramme

µM : Micromolaire

ρ0 : sans mitochondrie

ΔΔCt : Méthode de quantification relative de l'expression génique

ε : coefficient d'extinction molaire

Θ : Taille de la population efficace "mutaion-scaled"

%VAR : Pourcentage de la variation

%Vol : Pourcentage volume

~ : Environ

≤ : Inférieur ou égal

REMERCIEMENTS

Mes premiers remerciements sont adressés à mon directeur de recherche, Bernard Angers, sans qui ce projet n'aurait probablement jamais pu aboutir. Sa supervision, ses conseils, son soutien ainsi que la liberté qu'il m'a accordée m'auront permis d'arriver au bout de ce long cheminement que sont les études de troisième cycle. J'aimerais également prendre le temps de remercier mes co-directrices, Hélène Glémet et Annie Angers, qui ont toujours été présentes, disponibles et à l'écoute lorsque j'en avais besoin. Je remercie aussi tout particulièrement les différentes personnes étant passées, à plus ou moins long terme, par le laboratoire d'épigénétique des populations naturelles. Nombreux d'entre vous étant aujourd'hui devenu mes amis, sachez que sans vous rien de tout ceci n'aurait été possible. Merci de m'avoir écouté, soutenu, encouragé dans les moments plus difficiles mais surtout d'avoir partagé les moments de bonheur vécus au cours de ce parcours.

Je tiens à remercier ma famille qui malgré la distance qui nous sépare a toujours été là pour moi. Même si je sais que vous n'avez jamais réellement compris ce sur quoi je travail, merci de m'avoir toujours supporté dans mes choix. Je vous aime. Finalement, j'aimerais remercier mes amis belges ainsi que l'ensemble des personnes rencontrées au sein du département de biologie et en dehors de l'université. Les nombreuses bières, sorties et activités partagées m'auront permis de garder un certain équilibre entre les études/travail et la vie en général.

Exhaustivement, merci Bernard, Hélène, Annie, Inès, Erik, Kevin, Vincent (Chapdelaine), Sandrine, Indiana, Amélie, Bastien, Méloée, Philippe, Logan, Denis, Rachel, Antoine, Thomas

(Grünbaum), Roland, Lucie, Sophie, Yves, JB, Maxime (Leclerc), Emmanuelle, Stéphanie (Massé), Dan, Marie-Pierre, Vincent (Fugère), Amine, Tania, Simone, Jean-Martin, Salim, Alessia, Carl, Catherine, Grégoire, Sarina, Kelly, Stéphanie (Trembley), Kathe, Brian, Pierre-Luc, Adrien, Maryane, Jérôme, Louise, Arthur, William, Charlotte, Yoann, Pierre, Maxime (Amoranitis), Méline, Lola, Brice, Morgan, Louis, Loïc, Romain (Parizel), Maxime (Maréchal), Audrey, Yannick, Thomas (de Marchin), Olivier (Mouchette), Romain (Moens), Adeline, Justine, Fred, Sylvie, Daniel, Francis, Maude, Olivier (Angrignon) et Emilie.

CHAPITRE 1 – INTRODUCTION GÉNÉRALE

1.1 Les mitochondries

Il y a environ 1,5 milliard d'années, l'endocytose d'une α–protéobactérie par un proto-eucaryote mène à l'apparition d'un nouvel organite, les mitochondries (Embley et Martin, 2006 ; Gray et al., 1999 ; Lang et al., 1999). Cet organite d'origine monophylétique et hérité maternellement possède un génome qui lui est propre (ADNmt). Les mitochondries sont délimitées par une double membrane : une membrane interne encadre la matrice mitochondriale et une membrane externe délimite l'espace intermembranaire. Cet organite peut se retrouver sous différentes formes, allant de petites unités individuelles à de larges réseaux formés de mitochondries fusionnées (Chan, 2006). Cette organisation structurale influe sur leurs fonctions métaboliques. Les mitochondries représentent le site principal de production de l'énergie nécessaire à l'activité cellulaire (Magnus et Keizer, 1998). Cette production d'énergie, sous la forme d'adénosine triphosphate (ATP), découle de la phosphorylation oxydative, processus prenant place au sein des complexes de la chaîne respiratoire localisés dans la membrane interne. Bien que cette production soit l'une de leurs fonctions principales, cet organite intervient également dans de nombreux autres processus biologiques fondamentaux, tels que l'homéostasie calcique, la production d'espèces réactives de l'oxygène (ROS), la synthèse de stéroïdes, la thermogenèse, l'apoptose, *etc.* (Green et Reed, 1998 ; Martinou et al., 2000 ; Parekh, 2003 ; Skulachev, 1999).

1.2 Le génome mitochondrial

Structure moléculaire

Le génome mitochondrial est une molécule d'ADN double brin généralement circulaire et de petite taille (15 à 20 kb chez les animaux). Exceptionnellement, une forme linéaire peut être détectée chez quelques organismes tels que certains cnidaires (Bridge et al., 1992 ; Kayal et al.,

2012). Les brins de l'ADNmt sont caractérisés par des densités distinctes (Taanman, 1999) : un brin lourd plus riche en guanine (brin H) et un brin léger plus riche en cytosine (brin L). L'ADNmt est présent en un nombre d'exemplaires par mitochondrie variant de un à plus de 10 copies (Robin et Wong, 1988 ; Satoh et Kuroiwa, 1991). Le nombre de mitochondries par cellule étant lui-même dépendant du type de cellules et de leur activité (e.g. erythrocytes : 0 ; ovocytes de mammifères : 100.000), des milliers voir même des centaines de milliers de copies d'ADNmt peuvent être détectées dans une seule et même cellule (Cummins, 1998 ; Gahan et al., 2016 ; Schon, 2000). A chaque division cellulaire, mitochondries et copies d'ADNmt sont généralement distribuées de façon stochastique au sein des cellules filles.

Suite à leur endocytose, la plupart des gènes issus des α–protéobactéries ont été progressivement transférés au génome nucléaire (Gray et al., 2001). La plus grande partie de l'information génétique nécessaire à la biogenèse et aux fonctions mitochondriales est donc aujourd'hui localisée au sein de l'ADNnc. Bien que la plupart de ces gènes soient d'origine bactérienne, une partie non négligeable des gènes mitochondriaux modernes n'ont jamais été retrouvés chez les procaryotes (Gabaldón et Huynen, 2004 ; Gray et al., 2001 ; Gray, 2012). Si les protéines mitochondriales d'origine bactérienne sont principalement impliquées dans les fonctions mitochondriales (e.g. biosynthèse, production d'énergie, *etc.*), les protéines mitochondriales d'origine eucaryote joueraient essentiellement un rôle dans le transport intermembranaire (Gray et al., 2001).

Chez les animaux, le génome mitochondrial est généralement constitué de 37 gènes (Figure 1.1). Vingt-deux de ces gènes codent pour des ARN de transfert (ARNt) et deux d'entre

eux pour des ARN ribosomaux (ARNr ; 12S et 16S). Les treize gènes restants codent des sous-unités protéiques impliquées dans la formation de trois des quatre complexes de la chaîne respiratoire et de l'ATP synthase (Figure 1.1) : sept sont des sous-unités de la NADH-ubiquinone oxydoréductase (complexe I ; *ND1*-*ND6* et *ND4L*), une sous-unité de l'ubiquinone cytochrome c oxydoréductase (complexe III ; *CYTB*), trois sous-unités de la cytochrome c oxydase (complexe IV ; *COX1*-*COX3*) et deux sous-unités de l'ATP synthase (complexe V ; *ATP6* et *ATP8*). Les ~70 autres sous-unités de ces complexes ainsi que les quatre sous-unités de la succinate déshydrogénase (complexe II) sont encodées par l'ADNnc (Rand et al., 2004). L'ensemble des 37 gènes mitochondriaux possèdent des homologues chez les plantes, champignons et protistes (Gray et al., 1998 ; Levings et Brown, 1989 ; Paquin et al., 1997). A ces gènes sont associés quelques très rares et courtes séquences non-codantes ainsi qu'une région non-codante plus longue appelée la boucle de déplacement ou *D-loop*. Cette *D-loop* s'étend sur environ 700 paires de bases (pb) et contient autant l'origine de réplication du brin H que les origines de transcription des deux brins. La réplication de l'ADNmt est un phénomène fréquent au sein des cellules (Clay Montier et al., 2009), la demi-vie de ces molécules étant relativement courte (i.e. quelques jours ; Kai et al., 2006). Ce processus, tout comme la transcription, est régulé par des protéines d'origine nucléaire (Burton et al., 2013 ; Taanman, 1999).

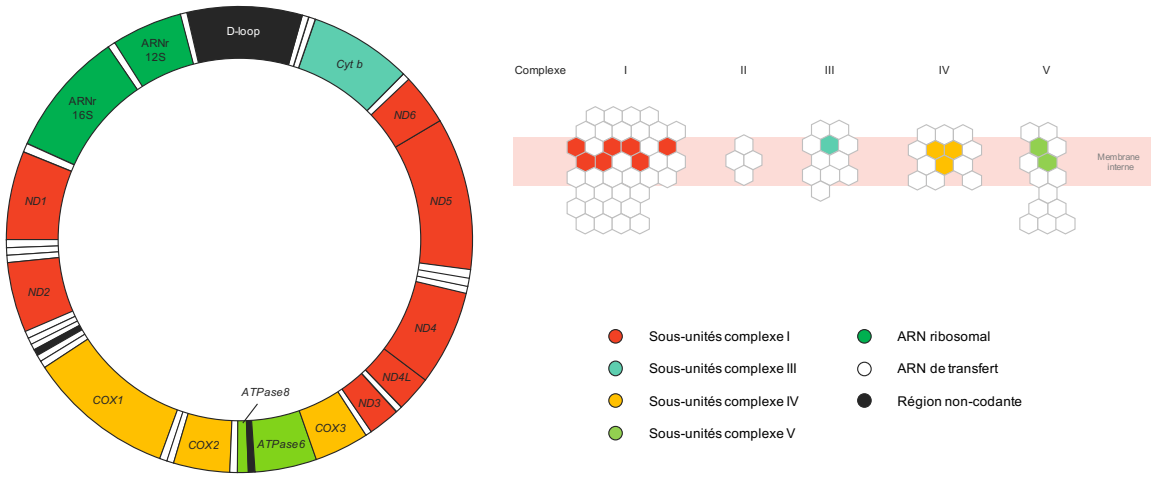

Figure 1.1: Génome mitochondrial humain type codant pour 37 gènes dont 13 encodent les sous-unités protéiques des différents complexes de la chaîne respiratoire mitochondriale. Modifié de Blier et al. (2001).

Héritabilité et transmission

L'ADN des organites cytoplasmiques (mitochondries et chloroplastes) est hérité de façon non-mendélienne. De nombreuses études ont en effet démontré que l'ADNmt est transmis à la descendance via les mitochondries retrouvées au sein des gamètes femelles (Ankel-Simons et Cummins, 1996 ; Birky, 1995 ; Giles et al., 1980). Ce mode de transmission est généralement désigné sous le nom de « transmission uniparentale », et dans ce cas-ci, plus spécifiquement de « transmission uniparentale maternelle ». Ce transfert est un processus commun à un grand nombre de taxons eucaryotes sexués, incluant les algues, les champignons, les bryophytes, les gymnospermes, les plantes supérieures et les animaux (Kuroiwa, 2010).

Selon le mode de transmission uniparentale maternelle, les mitochondries paternelles ne sont généralement pas conservées au sein du zygote. Deux hypothèses ont été proposées pour expliquer ce phénomène. La première de celles-ci repose sur un simple effet de dilution. L'ADNmt paternel étant présent en un nombre de copies beaucoup moins important que celui de l'ADNmt maternel, les mitochondries paternelles seraient difficilement détectables chez les descendants (Gyllensten et al., 1991). Selon la seconde hypothèse, l'ADNmt mâle subirait une digestion active et sélective menant finalement à sa disparition (Sato et Sato, 2013). Récemment, différents mécanismes supportant ce dernier scénario ont pu être mis en évidence chez divers animaux (Al Rawi et al., 2011 ; Ankel-Simons et Cummins, 1996 ; DeLuca et O'Farrell, 2012 ; Nishimura et al., 2006). Chez les mammifères, les mitochondries issues des gamètes mâles sont généralement intégrées au zygote. S'en suit une coexistence temporaire entre les mitochondries des deux sexes avant l'élimination complète et spécifique de la forme mâle (Ankel-simons et Cummins, 1996). Cette élimination se ferait via une ubiquitination des mitochondries mâles lors de la spermatogénèse qui seraient ensuite spécifiquement reconnues dans l'œuf et dégradées par les protéasomes et/ou les lysosomes (Sutovsky et al., 2000). Chez le poisson médaka (*Oryzias latipes*), une dégradation progressive de l'ADNmt paternel commencerait au cours de la spermatogénèse et se terminerait à l'intérieur de l'œuf immédiatement après la fécondation (Nishimura et al., 2006).

Bien que la transmission uniparentale maternelle soit un phénomène conservé chez de nombreuses espèces, le pourquoi de l'élimination des mitochondries paternelles au sein du zygote est encore mal compris. Une raison potentielle serait qu'au cours de la spermatogénèse et de la nage des gamètes mâles vers l'ovule, une quantité importante de ROS serait produite,

induisant de considérables dommages à l'ADNmt (Sato et Sato, 2013). Une autre possibilité serait qu'en raison du haut taux de mutations de ce génome, la transmission uniparentale permettrait de limiter la propagation de mutations délétères (Birky, 1995). Ainsi, l'apparition d'une mutation létale au sein d'une lignée mitochondriale induirait uniquement la disparition de celle-ci et non l'extinction de la population totale. Finalement, ce phénomène permettrait de limiter l'hétéroplasmie mitochondriale (i.e. présence de plus d'un type d'ADNmt/cellule) et d'ainsi éviter une compétition intragénomique pouvant être coûteuse pour l'organisme (Sharpley et al., 2012). Soulignons que ces différentes suppositions représentent des hypothèses mutuellement non-exclusives.

Taux de mutations

L'ADNmt est généralement considéré comme évoluant plus rapidement que l'ADNnc (Lynch, 1996 ; Lynch et al., 2006 ; Scheffler, 2008 ; Wallace et al., 1987). Le haut taux de mutations de ce génome peut être essentiellement attribué à quatre facteurs :

1. Les mitochondries génèrent des ROS pouvant endommager l'ADNmt et induire l'apparition de mutations (Balaban et al., 2005 ; Yakes et Van Houten, 1997). Ce risque de dommage oxydatif est d'autant plus prononcé lors de la réplication ; l'ADNmt étant longuement exposée sous forme simple brin, un état plus sensible aux attaques radicalaires (Kujoth et al., 2005).
2. La fidélité de l'ADN polymérase mitochondriale serait inférieure à celle de son homologue nucléaire (Johnson et Johnson, 2001).
3. Comparées au noyau, les mitochondries possèderaient un système de réparation de l'ADN réduit (Lynch et al., 2006).

4. L'absence de protection par des protéines de type histones, protéines associées à l'ADNnc agissant comme barrière mécanique de protection, participerait également à ce processus (Downs et Jackson, 2003).

Associés à une absence de recombinaison et à un très haut taux de réplication, ces caractéristiques expliqueraient le haut niveau de mutagenèse du génome mitochondrial. De plus, de par son haploïdie et sa transmission uniparentale, l'ADNmt est considéré comme ayant une taille efficace génétique (i.e taille d'une population idéale théorique ayant un degré de dérive génique similaire à la population étudiée) quatre fois plus petite que celle de l'ADNnc (Lynch et al., 2006). Cette particularité le rendrait plus « sensible » à la dérive génétique et lui permettrait d'évoluer plus rapidement.

1.3 La chaîne respiratoire

Les mitochondries représentent le site principal de production de l'énergie nécessaire à l'activité cellulaire (Magnus et Keizer, 1998). Ce processus, appelé phosphorylation oxydative, est assuré par une action conjointe de la chaîne respiratoire et de l'ATP synthase (complexe V), via un transport d'électrons et de protons (Figure 1.2).

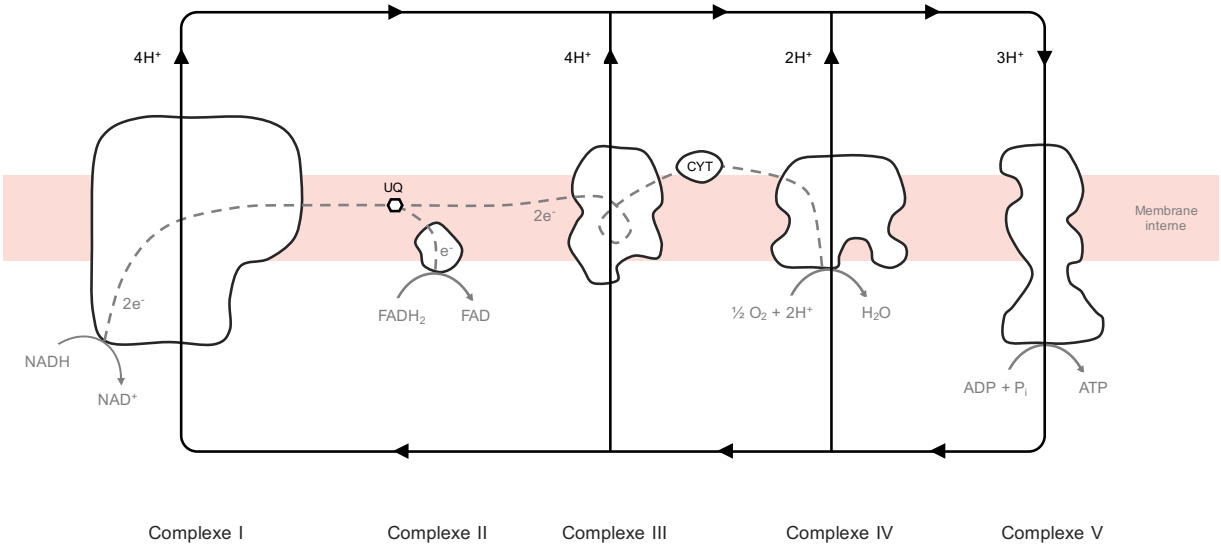

Figure 1.2: Transport d'électrons (e^-) et de protons (H^+) au sein de la chaîne respiratoire mitochondriale. Les e^- provenant du NADH et du $FADH_2$ entre dans la chaîne via les complexes I et II respectivement. Les complexes I, III et IV fonctionnent en tant que pompe à H^+. Le gradient transmembranaire de H^+ formé est consommé par le complexe V pour produire de l'ATP. UQ = ubiquinone ; CYT = cytochrome c. Modifié de Nicholls et Budd (2000).

Les électrons sont transférés depuis le NADH vers l'oxygène par l'intermédiaire des différents complexes de la chaîne respiratoire (i.e. complexes I-IV). Une fois transférés au complexe I, les électrons sont acheminés au complexe III par l'entremise de la forme réduite de la coenzyme Q. Cette molécule, aussi connue sous le nom d'ubiquinone, est une quinone hydrophobe diffusant rapidement au sein de la membrane interne. L'ubiquinone assure également un transfert d'électrons depuis le complexe II vers le complexe III. Les électrons transportés par le complexe II sont issus de la réduction du $FADH_2$, un équivalent réducteur produit lors du cycle de Krebs. Les complexes I et II représentent donc les « portes d'entrées » de la phosphorylation oxydative. Par la suite, la cytochrome c, une protéine soluble de petite taille, transfert les électrons du complexe III au complexe IV. Ces électrons sont finalement

transférés à l'oxygène, dernier accepteur de la chaîne. À ce flux d'électrons est associé un transfert de protons depuis la matrice mitochondriale vers l'espace intermembranaire. Ce pompage, assuré par les complexes I, III et IV, induit l'apparition d'un gradient de protons transmembranaire, de part et d'autre de la membrane interne. Ce gradient est consommé par le complexe V qui va alors pouvoir produire de l'énergie sous forme d'ATP. Il est important de remarquer que pour pouvoir fonctionner de façon optimale, les complexes I, III, IV et V nécessitent une interaction hautement spécifique entre leurs sous-unités protéiques d'origine nucléaire et mitochondriale.

1.4 La coadaptation mito-nucléaire

Au cours de l'évolution, l'accumulation de mutations au sein des génomes peut entrainer une altération de la spécificité des interactions existant entre les sous-unités nucléaires et mitochondriales des complexes de la chaîne respiratoire. Cette altération mène généralement à une diminution plus ou moins forte de l'activité des complexes (Burton et al., 2006 ; Dey et al., 2000 ; Ellison et Burton, 2006 ; Kenyon et Moraes, 1997 ; McKenzie et Trounce, 2000 ; McKenzie et al., 2003 ; Rawson et Burton, 2002 ; Spirek et al., 2000). Ce phénomène a pu être mis en évidence via la construction de cybrides artificiels (ADNnc : espèce 1 ; ADNmt : espèce 2) à partir d'organismes de plus en plus divergents. Selon les études réalisées, plus le niveau de divergence est élevé plus la perturbation des complexes est importante (Dey et al., 2000 ; Kenyon et Moraes, 1997 ; McKenzie et Trounce, 2000 ; McKenzie et al., 2003). A cette perturbation peut également être associée une altération plus ou moins marquée de la valeur adaptative des individus (e.g. viabilité, fécondité ; Ballard et Melvin, 2010 ; Burton et Barreto, 2012 ; Burton et al., 2013 ; Ellison et Burton, 2008a, 2006 ; Gershoni et al., 2009). Afin d'éviter

une baisse d'activité des complexes et une altération de la valeur adaptative, la modification de l'un des génomes peut être contrebalancée par une modification compensatoire sur l'autre génome (Osada et Akashi, 2012 ; Rand et al., 2004). On parle alors de coadaptation.

Le processus de coadaptation ne se limite pas aux interactions protéines-protéines que l'on retrouve au niveau des différents complexes de la chaîne respiratoire. En effet, des interactions protéines-ADN et protéines-ARN sont également à prendre en compte. Lors de la transcription par exemple, l'ARN polymérase, qui est d'origine nucléaire, reconnait une séquence spécifique au niveau de la *D-loop* du génome mitochondrial (Falkenberg et al., 2007 ; Gaspari et al., 2004). Dès lors, le remplacement de l'ARN polymérase par celle d'une autre espèce ou encore la modification de la séquence reconnue par cet enzyme est susceptible d'altérer la transcription mitochondriale (Burton et Barreto, 2012 ; Burton et al., 2013 ; Ellison et Burton, 2008b ; Gaspari et al., 2004). Il y a alors incompatibilité entre ADNmt et protéines nucléaires. Un autre type d'incompatibilité mito-nucléaire se retrouve également au niveau de la traduction (Burton et Barreto, 2012 ; Burton et al., 2013 ; Chou et al., 2010 ; Lee et al., 2008 ; Meiklejohn et al., 2013). Les levures hybrides résultant d'un croisement entre *Saccharomyces cerevisiae* (Sc) et *Saccharomyces bayanus* (Sb) sont incapables de sporuler. Selon Lee et al. (2008), cette stérilité serait essentiellement causée par le gène nucléaire *AEP2*, un gène codant une protéine impliquée dans la traduction du transcrit issu du gène mitochondrial *OLI1* ; lui-même codant une sous-unité du complexe V de la chaîne respiratoire. Dans ce cas, la protéine nucléaire Sb-Aep2 serait incapable d'interagir avec l'ARN issu du gène mitochondrial Sc-*OLI1*.

1.5 Les cybrides

Formation

Les cybrides sont des organismes dont l'ADNnc et l'ADNmt proviennent d'espèces différentes ou de populations divergentes. Lorsque créés *in vitro*, ces organismes résultent le plus souvent d'une fusion entre des cellules énucléées et des cellules dépourvues de mitochondries (cellules ρ0 ; Blier et al., 2001 ; Kenyon et Moraes, 1997 ; McKenzie et Trounce, 2000 ; McKenzie et al., 2003). Que ce soit en laboratoire ou en milieu naturel, les cybrides peuvent également être formés par introgression (Bernatchez et al., 1995 ; Blier et al., 2001 ; Boratynski et al., 2011 ; Edmands, 2002 ; Mila et al., 2011 ; Nagao et al., 1998). L'introgression mitochondriale se fait par hybridation entre deux espèces divergentes suivie de rétrocroisements successifs entre les hybrides femelles et l'espèce parentale paternelle (Edmands, 2002). Ainsi, bien que l'ADNnc maternel soit peu à peu remplacé par l'ADNnc paternel, l'ADNmt maternel, quant à lui, est conservé au cours des générations successives. En milieu naturel, les hybrides polyploïdes peuvent dans certains cas produire des gamètes femelles possédant l'ADNnc d'une espèce et l'ADNmt d'une autre espèce. Lorsque fécondés par des spermatozoïdes possédant un ADNnc identique, ces gamètes donnent naissance à des cybrides « instantanés » (Angers et al., 2012 ; Binet et Angers, 2005).

Contrairement aux cybrides obtenus artificiellement, les cybrides naturels doivent s'ajuster à leurs mitochondries étrangères de manière à pouvoir survivre et se reproduire dans leur environnement. L'existence de tels organismes représente alors une opportunité unique pour évaluer les effets de mitochondries étrangères sur le phénotype et l'écologie des individus.

Impacts phénotypiques

De l'épigénome à la spéciation

Le phénotype se définit comme une intégration des interactions existant entre les différents génomes, modulables en fonction de l'environnement. L'introduction de mitochondries étrangères induit généralement une modification de ce phénotype, à différents niveaux d'organisation biologique (e.g. épigénome, transcriptome, protéome, métabolisme, physiologie, *etc.*). Le premier niveau d'organisation du phénotype est l'épigénome, ensemble des marques épigénétiques régulant l'expression des gènes d'un individu en l'absence de changement de la séquence d'ADN (Angers et al., 2010 ; Razin et Riggs, 1980). Les modifications épigénétiques peuvent être regroupées en deux catégories : les processus engendrant une modification de la structure de la chromatine et ceux induisant une modification de l'ADN (sans changement de la séquence nucléotidique). Dans le premier cas, la désacétylation et la méthylation des histones induisent une compaction de la chromatine (hétérochromatine) résultant en une répression de la transcription (Rice et Allis, 2001). Dans le second cas, l'expression des gènes est réprimée via une méthylation de l'ADN qui empêche la liaison des facteurs de transcription (Boyes et Bird, 1991). Ces processus épigénétiques, qui agissent de concert afin de réguler l'expression génique, sont influencés par l'environnement. Un exemple particulièrement intéressant est celui des abeilles (*Apis mellifera*). Chez ces animaux, ouvrières stériles et reines fertiles sont produites à partir de larves génétiquement identiques (Tautz, 2008). Pour pouvoir se développer en reines, ces larves doivent être nourries à partir de gelée royale, une solution hypométhylante de l'ADN menant à l'expression de gènes normalement réprimés chez les ouvrières (Chittka et Chittka, 2010 ; Kucharski et al., 2008). Ainsi, dans ce système, la détermination du phénotype est fonction non pas de la génétique mais d'un signal environnemental induisant un remaniement

des marques épigénétiques et de l'expression génique.

Suite à l'intégration de mitochondries étrangères, une altération de la spécificité des interactions entre sous-unités nucléaires et mitochondriales des complexes de la chaîne respiratoire peut avoir lieu. Lors de la phosphorylation oxydative, ce phénomène induit généralement un ralentissement du flux d'électrons, menant à un état réduit des complexes, et finalement à une production accrue de ROS (Barja, 2007 ; Moreno-Loshuertos et al., 2006). Bien que ces molécules soient connues pour leurs effets délétères sur les lipides, protéines et nucléotides, certaines études démontrent aujourd'hui leur implication dans la modification de l'expression génique via, notamment, un impact sur les marques épigénétiques (Afanas'ev, 2015 ; Apel et Hirt, 2004 ; Lim et al., 2008 ; Wu et Ni, 2015). L'introduction de mitochondries étrangères, qui correspond à une modification de l'environnement intracellulaire, est donc susceptible d'induire une reprogrammation de l'expression de l'ADNnc par l'intermédiaire d'un remaniement de l'épigénome, possiblement déclenché par une surproduction de ROS.

Les modifications épigénétiques induites par la présence de mitochondries étrangères entraînent généralement un changement du transcriptome, du protéome et possiblement d'autres niveaux d'organisation supérieure du phénotype (e.g. voies métaboliques, physiologie, *etc.* ; Angers et al., 2012 ; He et al., 2010 ; Leitch et Leitch, 2008). Bien que dépendant de l'épigénome, le transcriptome et le protéome peuvent également être influencés par des modifications post-transcriptionnelles et post-traductionnelles, respectivement.

Ultimement, l'ensemble des modifications phénotypiques observées pourrait aller jusqu'à induire une réduction du flux génétique entre la forme cybride et son homologue sauvage, menant éventuellement à un évènement de spéciation entre les deux biotypes (Ellison et Burton, 2008a ; Gershoni et al., 2009). Afin d'illustrer ce phénomène, considérons le cas de l'introgression naturelle de mitochondries provenant de la paruline à croupion jaune (*Setophaga coronata*) chez la paruline d'Audubon (*Setophaga auduboni* ; Brelsford et al., 2011 ; Mila et al., 2011). Chez ces oiseaux, une altération du comportement migrateur saisonnier est observée ; passant de résident chez la forme sauvage à migrateur chez les cybrides (Toews et Brelsford, 2012). Selon Toews et al. (2014) cette modification pourrait être expliquée par la plus grande efficacité des mitochondries de la paruline à croupion jaune, qui seraient dès lors mieux adaptée à la demande énergétique associée aux migrations de longue distance. L'ensemble de ces changements phénotypiques pourraient induire une réduction du flux génétique entre les deux biotypes. Afin de vérifier cette hypothèse, des analyses de génétique des populations, et plus précisément du taux de migration entre populations, sont nécessaires.

Un impact positif ?

Malgré le processus de coadaptation existant entre l'ADNnc et l'ADNmt, de nombreux transferts inter-espèces de l'ADNmt surviennent naturellement chez les animaux et les plantes (voir Chapitre 2 ; Arnold, 2012 ; Boratynski et al., 2011 ; Du et al., 2011 ; Ellison et Burton, 2006 ; Glemet et al., 1998 ; Senjo et al., 1999 ; Toews et al., 2014). Tel que mentionné précédemment, de tels échanges affectent généralement la spécificité des interactions mito-nucléaires ainsi que la valeur adaptative des individus (Arnqvist et al., 2010 ; Ellison et Burton, 2006 ; Kenyon et Moraes, 1997 ; McKenzie et al., 2003). Cependant, diverses études

démontrent que l'introduction de mitochondries étrangères peut également être neutre (Blier et al., 2006 ; Boratynski et al., 2011 ; McKenzie et al., 2003) voir même avoir des effets bénéfiques (voir Chapitre 2 ; Pichaud et al., 2012 ; Toews et al., 2014).

Tout comme l'étude de Toews et al. (2014) décrite ici plus haut, l'étude de Pichaud et al. (2012) témoigne des effets bénéfiques que peut avoir l'intégration de mitochondries étrangères sur le phénotype. Dans cette étude, l'introgression artificielle de mitochondries provenant de *Drosophila simulans si*II chez *Drosophila simulans si*III est réalisée. Selon les auteurs, les cybrides ainsi produits sont caractérisés par des chaînes respiratoires présentant une activité catalytique supérieure à celle retrouvée chez les individus *si*III mais similaire à celle mesurée chez la forme *si*II. Un tel changement pourrait donner divers avantages à la forme introgressée comparée à la forme sauvage *si*III en terme d'intensité de l'activité aérobie, d'endurance et/ou de diminution de la production de ROS (Pichaud et al., 2012, 2011).

1.6 Le modèle biologique

Le modèle biologique utilisé pour cette étude est constitué de la forme sauvage et de la forme cybride du poisson ventre rouge du Nord, *Chrosomus eos* (Figure 1.3). Dans cette section, le complexe dont sont issus ces deux biotypes sera décrit et les caractéristiques qui en font un modèle d'intérêt seront détaillées.

Figure 1.3: Représentation photo des deux espèces parentales, *C. eos* sauvage et *C. neogaeus*, ainsi que des *C. eos* cybrides.

Le complexe Chrosomus eos-neogaeus

Le complexe *Chrosomus eos-neogaeus* (Pisces, Cyprinidae ; Figure 1.4) est issu d'hybridations entre les mâles *C. eos* et les femelles *C. neogaeus* (Figure 1.3 et 1.4). Ces événements d'hybridation sont à l'origine de lignées hybrides clonales se reproduisant de façon asexuée par gynogenèse (Dawley et al., 1987 ; Goddard et al., 1998). Ce mode de reproduction implique la formation d'œufs diploïdes dont le développement est stimulé par un spermatozoïde de l'une ou l'autre des espèces parentales. Occasionnellement, le matériel génétique mâle d'individus *C. eos* peut être intégré à l'œuf, créant ainsi des hybrides triploïdes de type *C. eos-neogaeus x eos* (Dawley et al., 1987 ; Goddard et al., 1998). Lors de la gamétogenèse, chez ces individus, l'assortiment des chromosomes non homologues *C. neogaeus* est éliminé au cours de la méiose,

menant à des gamètes haploïdes possédant un génome nucléaire de type *C. eos* et des mitochondries de type *C. neogaeus*. La fécondation de ces œufs par un mâle *C. eos* produit des individus cybrides, ayant un ADNnc de type *C. eos* et un ADNmt de type *C. neogaeus* (Binet et Angers, 2005).

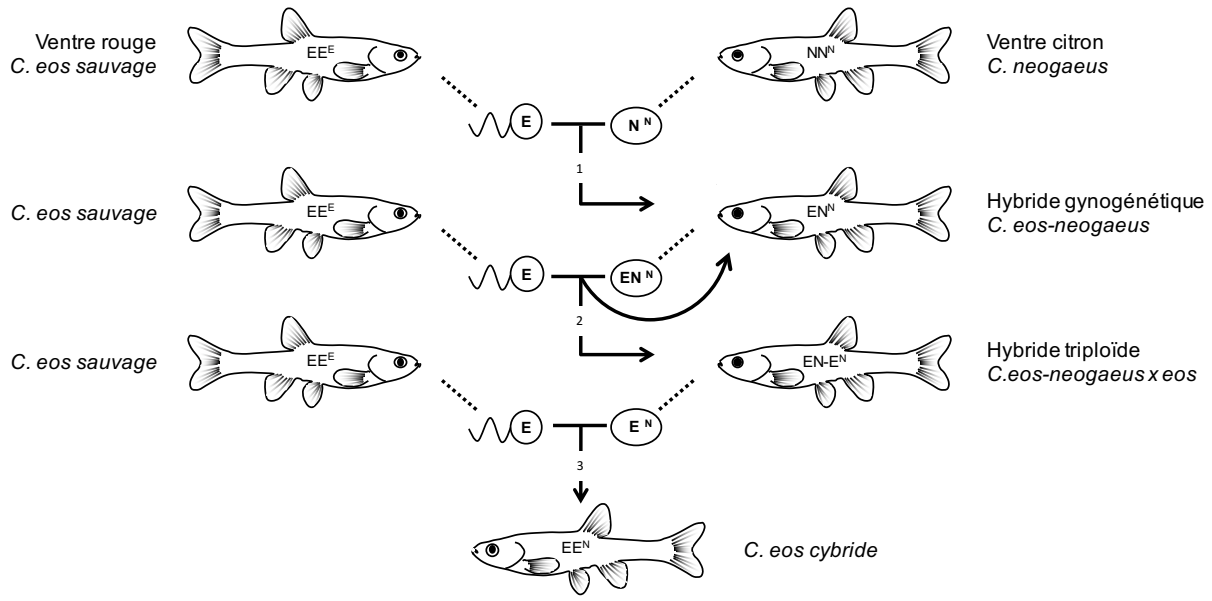

Figure 1.4: Mécanisme attendu menant à la formation des individus *C. eos* cybrides. 1– Les hybrides gynogénétiques résultent d'hybridations entre les femelles *C. neogaeus* et les mâles *C. eos*. 2– Les hybrides triploïdes apparaissent lorsque l'ADNnc du spermatozoïde de *C. eos* est intégré à l'œuf diploïde *C. eos-neogaeus*. 3– Les cybrides sont formés lorsque les œufs haploïdes issus des hybrides triploïdes sont fertilisés par un spermatozoïde *C. eos*. E et N réfèrent à l'ADNnc de *C. eos* et *C. neogaeus* respectivement ; les exposants représentent l'ADNmt. Modifié de Deremiens et al. (2015).

Répartition géographique

Les membres du complexe *C. eos-neogaeus* sont retrouvés en sympatrie ou en allopatrie dans la plupart des lacs du nord de l'Amérique du Nord (Scott et Crossman, 1973). Selon Mee et

Taylor (2012), les cybrides ainsi que *C. neogaeus* occuperaient une position dominante par rapport aux individus *C. eos* sauvages dans la portion nord de la répartition du complexe. Cette dominance suggère que les mitochondries *C. neogaeus* pourraient être mieux adaptées au froid que les mitochondries *C. eos*. Bien qu'ayant des répartitions géographiques distinctes, la forme sauvage et les cybrides sont tous deux retrouvés dans les lacs des Laurentides, région située à la frontière de leur répartition présumée respective.

Influence de l'ADNmt C. neogaeus

En se basant sur des études précédentes comparant une région du gène mitochondrial *cox1* (Angers et Schlosser, 2007) et en considérant le polymorphisme de longueur des fragments de restriction (RFLP) sur l'ADNmt complet (Goddard et al., 1989), une différence de 9,6% est mesurée entre l'ADNmt de *C. neogaeus* et de *C. eos*. Assumant une divergence de 1,3% par million d'années (Bermingham et al., 1997), on peut estimer à sept millions d'années la séparation entre la forme sauvage et la forme cybride. Pour un même niveau de divergence chez des souris cybrides, McKenzie et al. (2003) ont détecté une augmentation de la concentration en lactate au sein de fibroblastes, reflet d'une altération subtile du métabolisme. Lors d'une étude comparative des formes *C. eos* sauvages et cybrides en allopatrie, une variation des profils de méthylation (i.e. méthylome) a été observée entre les biotypes, et ce pour chacun des tissus étudiés (Angers et al., 2012). Dans la même étude, une analyse du protéome révèle que 7,5 à 8,7% des 405 protéines analysées sont différentes entre les deux formes. Sur base de ces observations, nous pouvons constater que la présence de mitochondries étrangères peut, à elle seule, induire une large reprogrammation de l'expression des gènes nucléaires.

Intérêts du modèle biologique

Au Québec, dans la région des Laurentides, formes sauvages et cybrides du poisson *C. eos* se côtoient, permettant ainsi une comparaison de ces dernières sous des contraintes environnementales et évolutives similaires. En se référant à l'étude de McKenzie et al. (2003) sur les souris, les sept millions d'années de divergence estimés entre la forme sauvage et les cybrides pourraient induire des modifications métaboliques. De plus, dans leur étude, Angers et al. (2012), démontrent un impact direct des mitochondries *C. neogaeus* sur le méthylome et le protéome *C. eos*. Ainsi, la forme sauvage et la forme cybride de *C. eos* nous procure l'unique opportunité d'étudier les effets de mitochondries étrangères à différents niveaux d'organisation biologique du phénotype, à partir d'organismes ayant évolué en milieu naturel.

1.7 Objectifs et hypothèses

L'objectif premier de cette thèse est d'évaluer l'impact de mitochondries étrangères à différents niveaux d'organisation du phénotype, lorsque les formes sauvages et cybrides de *C. eos* sont retrouvées en allopatrie ou en sympatrie. Notre hypothèse est que les mitochondries étrangères *C. neogaeus* déclenchent l'apparition de modifications phénotypiques avantageuses chez *C. eos*, permettant à ces cybrides d'utiliser de nouvelles niches écologiques et expliquant leur répartition géographique particulière.

Notre second objectif est de déterminer l'influence des modifications phénotypiques engendrées par la présence de mitochondries étrangères sur le flux génétique existant entre les formes *C. eos* sauvages et cybrides, quand sympatriques. Notre hypothèse est que lorsque ces

formes sont en sympatrie, une réduction du flux génétique est observée, signe d'une divergence progressive des biotypes.

1.8 Structure de la thèse

Dans le premier volet de cette thèse, les répercussions de mitochondries étrangères sur le phénotype ont été étudiées à différents niveaux d'organisation biologique. A cette fin, la forme sauvage et la forme cybride du ventre rouge (*Chrosomus eos*) ont été utilisées comme modèle biologique. Généralement retrouvés en allopatrie, ces biotypes ont été échantillonnés dans différents lacs de la région des Laurentides (Québec, Canada). Afin de tenir compte de l'influence des effets environnementaux sur nos analyses, les individus de chaque biotype ont été échantillonnés dans plusieurs lacs. Dans cette étude, les niveaux d'organisation du phénotype investigués sont : l'activité d'enzymes associés au métabolisme, le protéome ainsi que la performance de nage de ces poissons. Nous avons également retracé l'origine potentielle de la différence d'activité observée pour l'un des enzymes analysés. Ces analyses ont permis de mieux comprendre l'impact de mitochondries allospécifiques chez des organismes évoluant en milieu naturel.

Suite aux modifications phénotypiques détectées dans le cadre du volet précédent, l'objectif principal du second volet de cette thèse était de déterminer l'étendue du flux génétique entre des individus cybrides et leur homologue sauvage, lorsque retrouvés en sympatrie. Les analyses de migration effectuées révèlent un important flux génétiques entre ces biotypes, menant à une homogénéisation génétique de ceux-ci. Par ailleurs, un tel flux de gènes pourrait entraver la mise en place d'allèles adaptatifs, impliquant une absence de coadaptation entre

l'ADNnc et l'ADNmt des organismes cybrides, lorsque ce biotype est retrouvé en sympatrie avec la forme sauvage. Contrairement au premier volet, ces individus ont été échantillonnés au sein d'un seul et même lac de la région des Laurentides (Québec, Canada). De façon surprenante, une troisième forme de *C. eos* y a été découverte. Ce lac renferme ainsi la forme sauvage et deux formes cybrides divergentes issues des refuges glaciaires Mississippien et Atlantique.

Le troisième et dernier volet de ce projet avait pour but d'évaluer l'influence de mitochondries étrangères sur le phénotype, à partir d'organismes retrouvés en sympatrie. A cette fin, nous avons eu recourt à la forme sauvage ainsi qu'aux cybrides Mississippiens et Atlantiques de *C. eos*. Cette étude holistique a débuté par une analyse comparative de l'expression du génome nucléaire (i.e. approche épigénétique et transcriptomique). Par la suite, des niveaux d'organisation biologique de plus hautes complexités ont été investigués à l'aide d'analyses protéomiques et enzymatiques. Il est intéressant de remarquer que, par l'évolution divergente de leurs mitochondries, les cybrides Mississippiens et Atlantiques pourraient être caractérisés par des modifications phénotypiques distinctes. Par ailleurs, soulignons que contrairement à l'utilisation d'organismes retrouvés en allopatrie, l'utilisation d'organismes sympatriques permet une élimination sans équivoque des effets environnementaux et nucléaires sur nos analyses. De manière générale, ce volet démontre que les modifications phénotypiques observées entre les formes sauvages et cybrides sympatriques sont moins prononcées que celles mesurées à partir de biotypes allopatriques.

CHAPITRE 2 – ARTICLE 1

Interactions between nuclear genes and a foreign mitochondrial genome in the redbelly dace *Chrosomus eos*

Léo Deremiens, Logan Schwartz, Annie Angers, Hélène Glémet, Bernard Angers

Publié par Comparative Biochemistry and Physiology, Part B (2015), **189**: 80-86.

Résumé

Étant donné le processus de coévolution survenant entre les génomes nucléaires et mitochondriaux, les effets d'une hybridation introgressive sont encore mal compris. Dans cette étude, nous profitons de la cooccurrence de deux biotypes constitués d'un même génome nucléaire (*Chrosomus eos*) mais de mitochondries provenant d'espèces différentes (forme sauvage : *C. eos* ; cybrides : *C. neogaeus*) afin de déterminer l'étendue des changements phénotypiques liés à la divergence des génomes mitochondriaux. Ces modifications ont été estimées via une analyse de l'expression de certains gènes cibles, de l'activité d'enzymes associés au métabolisme, du protéome et de l'activité de nage des poissons. Nos données démontrent une augmentation significative de l'activité du complexe IV chez les cybrides, lorsque comparée à celle de la forme sauvage. Cette différence pourrait être le résultat de la modification d'un acide aminé au niveau de la sous-unité protéique *COX3* et/ou des importantes modifications observées au niveau protéomique. Finalement, nous montrons que les cybrides présentent une performance de nage plus élevée que celle de la forme sauvage. En définitive, nos résultats démontrent que l'absence de coévolution pour une période d'environ sept millions d'années entre les génomes nucléaires et mitochondriaux ne semble pas nécessairement néfaste, mais qu'au contraire, ce phénomène pourrait avoir des répercussions bénéfiques. En effet, la capture de mitochondries étrangères permettrait de contourner les processus de sélection s'appliquant à la coévolution génomique, induisant ainsi une rapide accumulation de nouvelles mutations chez les cybrides *C. eos*.

Abstract

Given the coevolution process occurring between nuclear and mitochondrial genomes, the effects of introgressive hybridization remain puzzling. In this study, we take advantage of the natural co-occurrence of two biotypes bearing a similar nuclear genome (*Chrosomus eos*) but harboring mitochondria from different species (wild type: *C. eos*; cybrids: *C. neogaeus*) to determine the extent of phenotype changes linked to divergence in the mitochondrial genome. Changes were assessed through differences in gene expression, enzymatic activity, proteomic and swimming activity. Our data demonstrate that complex IV activity was significantly higher in cybrids that in wild type. This difference could result from one variable amino acid on the *COX3* mitochondrial subunit and/or from a tremendous change in the proteome. We also show that cybrids present a higher swimming performance than wild type. Ultimately, our results demonstrate that the absence of coevolution for a period of approximately seven million years between nuclear and mitochondrial genomes does not appear to be necessarily deleterious but could even have beneficial effects. Indeed, the capture of foreign mitochondria could be an efficient way to circumvent the selection process of genomic coevolution, allowing the rapid accumulation of new mutations in *C. eos* cybrids.

2.1 Introduction

Mitochondria play a vital role in cell metabolism. Although the mitochondrial genome (mtDNA) of animals only encodes a few proteins, they are critically positioned in the energy producing complexes of the respiratory chain (Andersson et al., 1998; Garesse and Vallejo, 2001). Of the thirteen mtDNA-encoded proteins, seven are subunits of NADH-ubiquinone oxidoreductase (complex I; *ND1*-*ND6* and *ND4L*), one is a subunit of ubiquinone-cytochrome c oxidoreductase (complex III; *CYTB*), three are subunits of cytochrome c oxidase (complex IV; *COX1*-*COX3*, forming the functional core of the enzyme) and two are subunits of ATP synthase (complex V; *ATP6* and *ATP8*). All the components of complex II (succinate dehydrogenase) are nuclear encoded. Consequently, all other complexes require highly specific interactions between nuclear and mitochondrial subunits to function effectively, suggesting that a strong coevolution of genomes must occur (Blier et al., 2001; Rand et al., 2004).

Studies of cybrids with nuclear DNA (ncDNA) and mtDNA derived from different species provide support for the coevolution hypothesis (James and Ballard, 2003; Kenyon and Moraes, 1997; McKenzie et al., 2003; McKenzie and Trounce, 2000). Artificial cybrids obtained by the experimental introduction of mtDNA from progressively more divergent species into *Mus musculus domesticus* mtDNA-less (ρ0) cells revealed that cell viability and mitochondrial metabolism decreased as the genetic distance between mtDNA and ncDNA increased (McKenzie et al., 2003).

While most studies involve cybrids created *in vitro* (Cannon et al., 2011; Kenyon and Moraes, 1997; McKenzie et al., 2003; McKenzie et al., 2004) or obtained by directed

backcrossing (Ellison and Burton, 2006; Harrison and Burton, 2006; Pichaud et al., 2012; Sackton et al., 2003), some cybrid organisms spontaneously appear in natural environments (Angers et al., 2012; Boratynski et al., 2011; Toews and Brelsford, 2012). Unlike artificially obtained cybrids, natural cybrids must deal with their foreign mitochondria to survive and reproduce in their environment. Moreover, interspecific transfer of mitochondria may induce a modification of phenotype and confer adaptability to new environmental conditions (Glémet, 1997; Toews et al., 2014). Indeed, even though the absence of coevolution between genomes is expected to decrease individual performances, several studies demonstrate that this can sometimes present advantageous effects (Pichaud et al., 2012; Toews et al., 2014). A particularly interesting example is the introgression of the Myrtle warbler (*Setophaga coronata*) mitochondria in Audubon's warbler (*Setophaga audubonì*) (Brelsford et al., 2011; Mila et al., 2011). For the latter birds, mitochondrial introgression permanently alters their seasonal migratory behavior so that residents become migratory (Toews and Brelsford, 2012). According to Toews et al. (2014), this phenomenon results from mitochondria that are more efficient, and better adapted for the energetic demands of long-distance migration.

A spontaneous naturally occurring cybrid derived from the redbelly dace *Chrosomus eos* (Cyprinidae, Pisces) is found in the freshwater lakes of North America. This cybrid originates from the crossfertilization of *C. eos-neogaeus x eos* triploid females by *C. eos* males (Figure 2.1), effectively producing fish bearing diploid *C. eos* nuclei and *C. neogaeus* mitochondria (Angers et al., 2012; Angers and Schlosser, 2007; Binet and Angers, 2005; Dawley et al., 1987; Goddard et al., 1998). Based on previous estimates from comparisons of the *cox1* gene region (Angers and Schlosser, 2007) and of restriction fragment length polymorphism (RFLP) of the

complete mitochondrial genome (Goddard et al., 1989), *Chrosomus neogaeus* mtDNA differs by 9.6% compared to *C. eos*. This implies that species have been separated for approximately seven million years. In addition, methylome as well as proteome are strongly different between allopatric *C. eos* wild type and cybrids (Angers et al., 2012). Given these important divergences, modification of key mitochondrial function in cybrids, such as ATP production, might be expected. According to Mee and Taylor (2012), *C. eos* wild type are found mainly in the southern part of the distribution area whereas cybrids, as well as *C. neogaeus*, occur predominately in the north. This peculiar distribution suggests that cybrids might be able to support colder conditions than wild type. In the southern part of Quebec (Canada), a region where their distribution overlaps, wild type and cybrids are found in close proximity (Angers and Schlosser, 2007), thereby allowing direct comparison of the different biotypes under similar environmental and evolutionary constraints.

This study aims at assessing the influence of foreign mitochondria on *C. eos* phenotype. To do so, we first compared the enzymatic activity of wild type and cybrids where *C. eos* nuclear genome is in association with *C. eos* and *C. neogaeus* mtDNA respectively. We measured the maximal activity of metabolic enzymes involved in aerobic metabolism (complex I, II, II+III and IV of the mitochondrial respiratory chain) and anaerobic metabolism (lactate dehydrogenase) in white muscle. Afterwards, the source responsible for variation in enzymatic activity was investigated by analyses of gene expression, gene sequences and proteome reprogramming. To this end, qRT-PCR, sequencing and two-dimensional gel electrophoresis were respectively used. Finally, we assess whether the *C. neogaeus* mitochondria of cybrids could impact function at a higher level of biological organization by determining individual

swimming performance (U_{crit}).

Since *C. eos* cybrids are able to deal with their foreign mitochondria and because the absence of coevolution between mitochondrial and nuclear genomes can sometimes present beneficial effects, we first predict that respiratory chain complex I, complex III and complex IV have higher activities in *C. eos* cybrids than in wild type. Secondly, we hypothesize that higher activities of complexes I, III and IV in cybrids improve the efficiency of mitochondrial metabolism, which in turn should improve the swimming capacity (U_{crit}) of these fish.

2.2 Material and method

The biological model

The cybrids used in this study occur naturally and are derived from the *Chrosomus eos-neogaeus* hybrids (Dawley et al., 1987; Goddard et al., 1998). Hybridization between *C. eos* males and *C. neogaeus* females results in the *C. eos-neogaeus* diploid hybrids that reproduce clonally via gynogenesis (Binet and Angers, 2005; Goddard et al., 1998). Occasionally, the nuclear genome of the *C. eos* sperm can be incorporated into the egg's diploid genome leading to the formation of triploid hybrids *C. eos-neogaeus x eos*. During gametogenesis of triploid hybrids, the unmatched set of chromosomes is discarded and meiosis occurs on the *C. eos* diploid genome, producing haploid gametes (Goddard et al., 1998). Fertilization of such eggs by a *C. eos* male results in the formation of individuals with a pure *C. eos* nuclear diploid genome but a female inherited *C. neogaeus* mitochondrial genome (Binet and Angers, 2005; Figure 2.1).

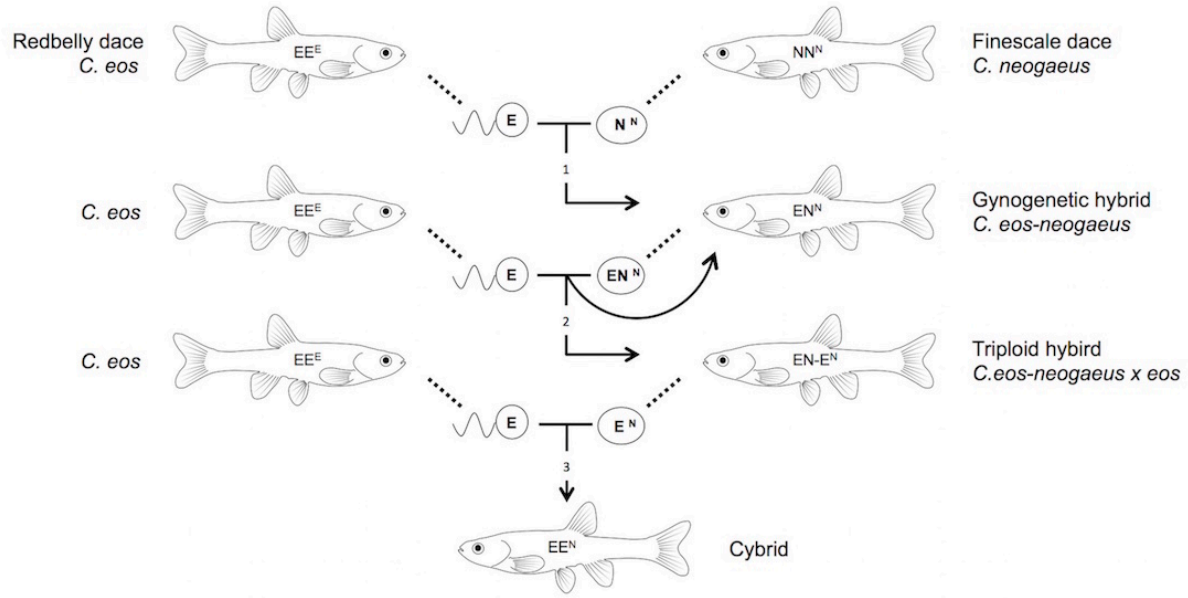

Figure 2.1: Expected mechanism leading to the formation of natural cybrids. 1- Gynogenetic hybrids resulting from hybridization between *Chrosomus neogaeus* females and *C. eos* males. 2- Triploid hybrids arise when nuclear genome of *C. eos* sperm is incorporated into diploid eggs of *C. eos-neogaeus*. 3- Cybrids are formed when haploid eggs of triploid hybrids are fertilized by *C. eos* sperm. E and N refer to the nuclear genome of *C. eos* and *C. neogaeus* respectively, superscripts represent mitochondrial genome.

Sampling and biotype identification

Fish were sampled in the Laurentian region (Quebec, Canada) in six lakes located in close proximity of each other (Table 2.1). All fish were captured at the same period of the year, in the summer when lake temperature was around 24°C. Except for individuals used for swimming performance (U_{crit}) tests, all fish were euthanized in the field by eugenol overdose and immediately placed on dry ice. They were then brought back to the laboratory and stored at -80°C for a few months to two years; depending on the sampled year (Table 2.2).

Table 2.1: Characteristics of the sampled lakes. Geographic localization of the sites and biotypes identified in the lakes.

Lake	Geographic coordinates	Biotypes
Saad	45.913333, -74.050833	Wild type
Morin	45.873832, -74.140385	Wild type
Triton	45.987786, -74.008194	Wild type
Beaver	45.924722, -74.064167	Cybrids
Plat	45.913333, -74.050833	Cybrids
Desjardins	45.916111, -74.073056	Cybrids

We identified *C. eos* biotypes by assessing their genome composition as described by Binet and Angers (2005). Briefly, a first step, based on the nuclear genome composition of individuals, enabled the distinction between each parental species and hybrids. Afterward, a second step performed on the mitochondrial genome, allowed the discrimination of wild type from cybrids.

Enzymatic activity analysis

A survey of enzyme activity was performed on three to nine individuals from two lakes for each biotype (Table 2.2). The entire white muscle was sampled from one side of these fish and red muscle, found in very small amounts, was removed. The white muscle samples were then powdered using a dry-ice cooled mortar and pestle. Each muscle sample was weighed (80-110 mg) and homogenized as described by Spinazzi et al. (2012). All manipulations were performed on ice.

Table 2.2: Origin, sampled year and number of individuals analysed for the enzymatic activity, the gene expression (qRT-PCR), the proteome survey and the critical swimming speed (U_{crit}).

Lake	Sampled year	Enzymatic activity							qRT-PCR	Proteome	U_{crit}
		CI (24°C)	CII (24°C)	CII+III (24°C)	CIV (24°C)	CIV (17°C)	CIV (10°C)	LDH (24°C)			
Saad	2011[a] – 2013[b]	7	7	9	9	5	5	6			13
Morin	2012	3	4	4	7	5	5	6	5		
Triton	2011									8	
Beaver	2012[a, c] – 2013[b]	7	7	7	9	7	7	8	4		9
Plat	2012	5	5	5	5	3	7	4			
Desjardins	2011									8	

*Not*e. CI, CII, CII+III, CIV and LDH refer to complex I, complex II, complex II+III, complex IV and lactate dehydrogenase respectively.

[a] Fish sampled for the enzymatic activity analysis.

[b] Fish sampled for the critical swimming speed analysis.

[c] Fish sampled for the gene expression analysis.

Complex I, complex II, complex II+III, complex IV and lactate dehydrogenase (LDH) were measured in wild type and cybrids using a UV/VIS spectrophotometer (Shimadzu, UV-1800) equipped with a thermoelectrically temperature controlled cell holder (CPS-240A). All assays were run at least in duplicate and performed at 24°C, a temperature close to the water temperature where fish were sampled. Complex IV activity was also assessed at 17°C and 10°C. Specific activities were expressed as nmol per minute per mg of total proteins. Conditions for the enzymatic assays were as follow:

(i) Complex I: 50 mM potassium phosphate (pH 7.5), 100 μM NADH, 60 μM ubiquinone-1, 3 mg mL^{-1} bovine serum albumin (BSA), 300 μM KCN. Activity was measured by following the disappearance of NADH at 340 nm and calculated using an extinction coefficient (ε_{340}) of 6.2 mM^{-1} cm^{-1} (Spinazzi et al., 2012). To determine the rotenone-insensitive activity, duplicate reactions with 10 μM rotenone were performed.

(ii) Complex II: 25 mM potassium phosphate (pH 7.5), 20 mM succinate, 80 μM 2,6-dichlorophenolindophenol (DCPIP), 50 μM decylubiquinone, 1mg mL^{-1} BSA, 300 μM KCN. Activity was assayed by monitoring the decrease of absorbance due to the reduction of DCPIP at 600 nm (ε_{600} = 19.1 mM^{-1} cm^{-1}) (Spinazzi et al., 2012).

(iii) Complex II+III: 20 mM potassium phosphate (pH 7.5), 10 mM succinate, 50 μM oxidized cytochrome c, 300 μM KCN. Activity was measured by following the increase of absorbance caused by the reduction of oxidized cytochrome c at 550 nm (ε_{550} = 18.5 mM^{-1} cm^{-1}) (Spinazzi et al., 2012).

(iv) Complex IV: 50 mM potassium phosphate (pH 7.0), 60 μM reduced cytochrome c. Cytochrome c was reduced by the addition of a few grains of sodium dithionite.

Excess dithionite was eliminated by bubbling air into the solution (Thibault et al., 1997). Activity was monitored by following the decrease of absorbance owing to the oxidation of reduced cytochrome c at 550 nm (ε_{550} = 18.5 mM^{-1} cm^{-1}) (Spinazzi et al., 2012).

(v) LDH: 100 mM potassium phosphate (pH 7.0), 160 µM NADH, 400 µM pyruvate. Activity was measured by following the decrease of absorbance due to the oxidation of NADH at 340 nm (ε_{340} = 6.2 mM^{-1} cm^{-1}) (Thibault et al., 1997).

Total soluble protein content of each homogenate was determined by the Bradford method (Bradford, 1976) using bovine serum albumin (BSA) as a standard.

Gene expression analysis

To trace back the origin of variation observed in the enzymatic activity of biotypes, we analysed the expression level of genes encoding the complex IV mitochondrial subunits (COX1, COX2 and COX3) and the complex IV nuclear subunits (COX5b, COX6b and COX7a1). The nuclear gene of β-actin was used as endogenous control. This assay was conducted on five wild type individuals and four cybrids, which were sampled from two lakes (Table 2.2). Total RNA was extracted from 50-100 mg of white muscle from -80°C conserved wild type and cybrids using TRIzol Reagent (Invitrogen Life Technologies). The muscle was manually homogenized using a small pestle (Fisher Scientific). RNA was then further purified and DNase treated using a RNeasy Fibrous Tissue Mini Kit (Qiagen). Total RNA was quantified using a BioDrop µLITE. First strand cDNA was synthesized from 1 µg of total RNA using a QuantiTect Reverse Transciption Kit (Qiagen). mRNA expression for genes of interest was at least measured in duplicate with a PikoReal 96 Real-Time PCR System (Thermo Scientific) using DyNamo Flash

SYBR Green qPCR Kit (Thermo Scientific). The relative quantification method (ΔΔCt; Livak and Schmittgen, 2001) was used to estimate the relative amount of targets in each sample. Specific primers were designed using Primer3Plus software (Table A1). Specificity of all primer pairs was tested using PCR amplification: electrophoresis showed a single band for each sample, having a length consistent with that expected. Additionally, a SYBR green dissociation curve was run to ensure a single fluorescence peak at the appropriate temperature.

Sequencing

To further understand the source responsible for variation in the observed enzymatic activity of biotypes, the mitochondrial genes *cox1*, *cox2* and *cox3* were sequenced with a Beckman Coulter CEQ 2000XL sequencer. Sequences alignment and mtDNA-encoded protein comparisons were realized with MEGA 5.2.2 software (Tamura et al., 2011).

Proteome reprogramming analysis

A quantitative survey of the proteome was also performed to help determine the source responsible for modification in observed enzymatic activity. Protein homogenates were prepared from white muscle, as described by Angers et al. (2012). Total protein (250 μg) was separated using two-dimensional gel electrophoresis (Klose, 1975; O'Farrell, 1975). Isoelectric focusing was performed using IPG strips (pH 4-7; GE Healthcare). Molecular weight separation was carried out on a SDS polyacrylamide gel (12.5%). 2D-gels were stained with silver nitrate coloration (Heukeshoven and Dernick, 1988). This assay was conducted on eight wild type individuals and eight cybrids, which were sampled from two lakes (Table 2.2). For each 2D-gel, four individuals belonging to the same biotype were pooled. Automatic spot detection was then

realized for each scan using MELANIE software version 7.05 (Swiss Institute of Bioinformatics). Only spots with good resolution and contrast were selected. To compare the protein expression within and between *C. eos* biotypes, the percentage volume (%Vol) was used. The %Vol of a spot refers to the ratio between its volume and that of all spots found in the gel.

Critical swimming tests (U_{crit})

The critical swimming speed was assessed to determine whether the *C. neogaeus* mitochondria of cybrids could impact function at higher levels of biological organization than enzymatic activity: organism *vs*. macromolecular level. This assay was conducted on thirteen wild type individuals and nine cybrids which were sampled from two lakes and maintained under identical laboratory conditions for one year at 17°C (Table 2.2). These fish were different from the ones used for the enzymatic activity assay. Critical swimming tests were carried out in modified Blazka-type swim tunnels (3.37 ± 0.03L) similar to those described by Beamish et al. (1989). Except for the relationship between pump voltage and water velocity which was determined daily instead of weekly, swim tunnel conditions and their utilization were as described by Rouleau et al. (2010).

Since fish were acclimatized to 17˚C, all swimming tests were realized at that temperature. Before being tested, fish were deprived of food for approximately 48h. A single fish was transferred to each swim tunnel and left to habituate for 45 min at a water velocity of 10 cm s^{-1}. Following habituation, flow rate was increased by 5 cm s^{-1} every 10 min until the fish could no longer hold its position against the current (Koumoundouros et al., 2002; Leris et al., 2013). Critical swimming speed (U_{crit}) was calculated using the following formula:

$$U_{crit} = U_i + U_{ii} \times (T_i/T_{ii})$$

where U_i is the final speed (cm s^{-1}) completed, U_{ii} is the speed increment (cm s^{-1}), T_i is the time spent at the final speed and T_{ii} the time increment (s).

On average, tested wild type and cybrid individuals were of similar size (Student *t*-test, $P > 0.1$). Nevertheless, individual size affects swimming performance (Beamish, 1978). Therefore, the relative critical swimming speed (RU_{crit}), which is the ratio between U_{crit} and total fish length (TL) (Beamish, 1978), was calculated for every fish.

Statistical analysis

All statistical analyses were performed using SPSS Statistics 22 software. A Student's *t*-test was used to evaluate differences between groups. The significance level was set at $P < 0.05$.

2.3 Results

Biotype identification

Genetic identification confirmed the presence of wild type individuals in lake Saad, Morin and Triton whereas cybrids were present in lake Beaver, Plat and Desjardins (Table 2.1). As reported in Angers et al. (2012), the two biotypes occurred in allopatry, where a single biotype was found in each lake.

Enzymatic activity

To account for possible environmental effects due to conditions specific to each lake, we first compared the activity of complex I, II, II+III and IV, as well as that of LDH, for each biotype,

since biotypes originated from different lakes (Table 2.2). No significant differences were found for any of the biotypes (Student's *t*-test, $P > 0.1$) suggesting that lake condition had no detectable effect on enzyme activity. Wild type and cybrid individuals from different lakes were then grouped in the subsequent analyses. Since cell macromolecular concentrations are known to be able to modify enzyme activities (Garner and Burg, 1994; Hochachka and Somero, 2002), total soluble protein content of wild type and cybrid homogenates were also compared. No significant difference was found between the two biotypes (Student's *t*-test, $P = 0.786$).

To assess the influence of foreign mitochondria on *C. eos* metabolism, maximal enzyme activity in wild type and cybrids was compared (Table 2.3). At 24°C, a significant increase of complex IV activity was observed in cybrids when compared to wild type (Student's *t*-test, $P < 0.005$) whereas no significant differences were detected for complex I, complex II and complex II+III activities (Student's *t*-test, $P > 0.3$). Finally, no significant difference in the mean activity of LDH was observed between wild type and cybrids (Student's *t*-test, $P > 0.4$).

Table 2.3: White muscle enzyme activities in *C. eos* wild type and cybrids. Enzyme activities are expressed in nanomoles per minute per milligram of total proteins (Mean ± standard deviation).

Enzymes	Temperature (°C)	Wild type	Cybrids	*P*-value
Complex I	24	0.0105 ± 0.0068	0.0134 ± 0.0083	0.381
Complex II	24	0.0142 ± 0.0053	0.0150 ± 0.0080	0.779
Complex II + III	24	0.0084 ± 0.0020	0.0088 ± 0.0019	0.596
Complex IV	24	0.0452 ± 0.0237	0.0736 ± 0.0268	0.004
	17	0.0409 ± 0.0252	0.0482 ± 0.0230	0.510
	10	0.0330 ± 0.0199	0.0346 ± 0.0180	0.849
LDH	24	2.4242 ± 0.5709	2.1689 ± 0.8846	0.410

As the overall geographical distribution of both biotypes hints towards different temperature sensitivities, we also measured the complex IV activity at lower temperatures. Interestingly, for complex IV activity in wild type individuals, a gradual decrease at lower temperatures was observed, whereas for cybrids, activity decreased sharply at 17 and 10˚C, attaining a similar activity to that of wild type (Student's *t*-test, $P > 0.5$). Complex IV activity of cybrid fish thus appears more sensitive to temperature, being significantly higher at 24°C as compared to wild type, but comparable at colder temperatures (Table 2.3).

Gene expression

In order to assess whether differential expression of complex IV subunit encoding genes could be responsible for the enzymatic activity variation observed at 24°C, qRT-PCR experiments were performed to determine the amount of mRNA for COX1, COX2, COX3 (encoded on

mtDNA) and COX5b, COX6b, COX7a1 (from ncDNA). Our results revealed no difference in gene expression between wild type and cybrid individuals for any of the complex IV subunit genes (Student's *t*-test, $P > 0.7$; Figure Annex (A) 1).

Sequence analysis

We sequenced the genes coding for the mitochondrial subunits of complex IV; *cox1* and *cox3* were completely sequenced while 94.4% of *cox2* sequence was retrieved. All sequences have been deposited in Genbank (accession numbers: KP866220, KP866221, KP866222). Out of the 1551 bp of *cox1* gene, 121 sites were variable between the *C. eos* mtDNA of wild type and the *C. neogaeus* mtDNA of cybrids. This variability represents a 7.8% difference. Regarding *cox2* and *cox3*, a difference of 6.6% (43 variable sites/652 bp) and 8.1% (64 variable sites/784 bp) was detected respectively. These results are slightly below the previously estimated 9.6% (Angers and Schlosser, 2007; Goddard et al., 1989).

Amongst the analysed genes, no non-synonymous mutation was detected in *cox1*, only one in *cox2* and two in *cox3*. Every non-synonymous mutation was conservative (i.e. similar properties) except for one in *COX3*, where Thr41 in wild type was replaced by a Met in cybrids. This modification induces the incorporation of a nonpolar amino acid instead of a polar amino acid in the protein.

Proteome comparative analysis

To compare protein expression between the biotypes, we analysed the %Vol of a subsample of 162 unambiguous protein spots present in both wild type and cybrid 2D-gels. Comparisons

within biotypes revealed a strong correlation between wild type gels (R^2 = 0.87), as well as between cybrid gels (R^2 = 0.95; Figure 2.2A and B). Approximately 10% of the assessed proteins showed a more than 2-fold difference in expression for both wild type (9.26%) and cybrids (11,73%). These results are in sharp contrast with gel comparisons between biotypes (wild type 1 *vs.* cybrid 1, wild type 2 *vs.* cybrid 2, wild type 1 *vs.* cybrid 2, wild type 2 *vs.* cybrid 1), which demonstrated a very low correlation (R^2 < 0.20; Figure 2.2C and D) and where 29.01 to 37.04% of the assessed proteins showed a more than 2-fold difference in expression.

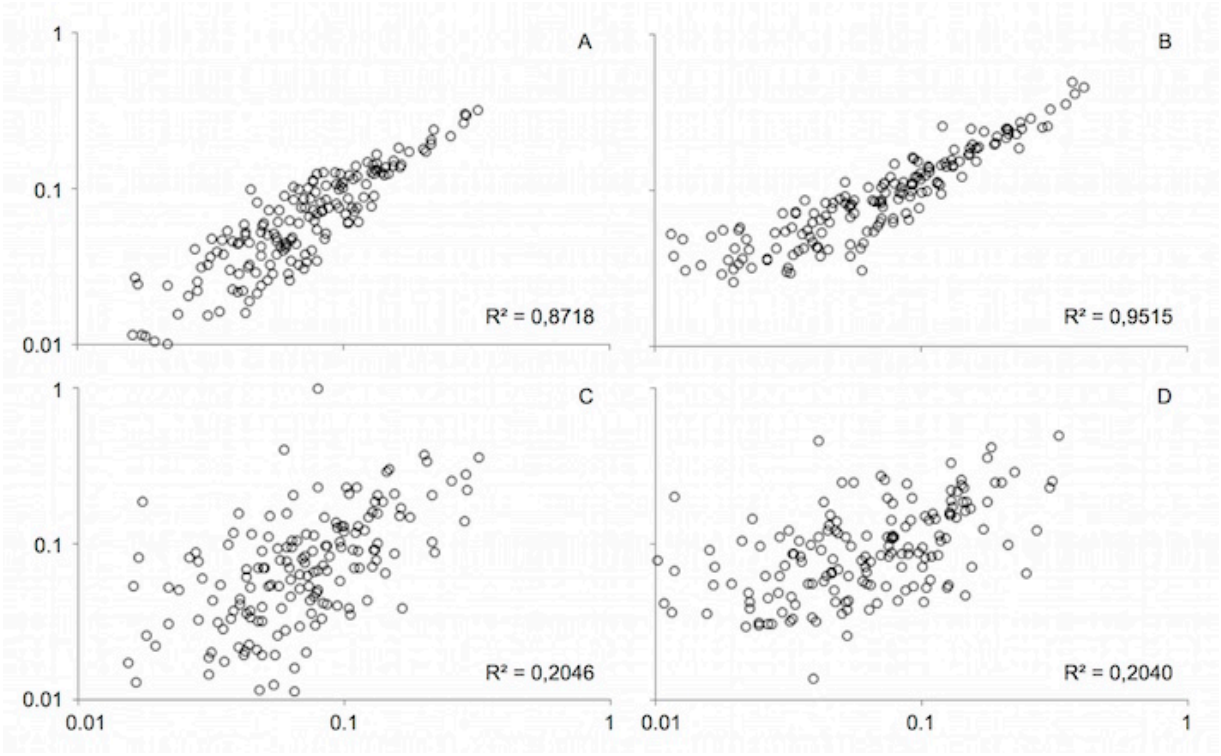

Figure 2.2: Scatter plots representing protein expression variations within and between wild type and cybrids. Gel comparisons within biotypes: wild type 1 *vs.* wild type 2 (A), cybrid 1 *vs.* cybrid 2 (B). Gel comparisons between biotypes: wild type 1 *vs.* cybrid 1 (C), wild type 2 *vs.* cybrid 2 (D). Axes units are expressed in %Vol.

Critical swimming speed (U_{crit})

Critical swimming tests revealed that cybrids presented a significant increase in their relative maximum sustainable swimming speed (RU_{crit}) when compared to wild type at 17°C (Student's *t*-test, $P < 0.0005$). Cybrids reached an average relative sustainable swimming speed of 5.60 ± 0.71 TL s^{-1} compared to 4.20 ± 0.74 TL s^{-1} for wild type.

2.4 Discussion

To better understand the influence of foreign mitochondria on phenotype, we compared key metabolic enzyme activities of *C. eos* and *C. neogaeus* mitochondria when associated with a *C. eos* nuclear genome and then investigated the potential source responsible for the observed variability. Because a modification in the respiratory chain complex is expected to modify metabolism efficiency and possibly impact function at higher levels of biological organization (e.g. organism level), we finally assessed whether *C. neogaeus* mitochondria could influence individual swimming performance of cybrids.

At the macromolecular level, our data suggest that only complex IV activity is modified, with a significant enzymatic activity increase in cybrids compared to wild type at 24°C. This complex is probably one of the most studied complexes of the mitochondrial respiratory chain. It catalyzes electron transfer from cytochrome c to molecular oxygen, thereby reducing the latter to water (Nicholls and Ferguson, 2003). It comprises ten ncDNA-encoded subunits and three mtDNA-encoded subunits (Kadenbach, 1983; Tsukihara et al., 1996). The latter subunits form the functional core of the enzyme and interact with the nuclear-encoded cytochrome c (Tsukihara et al., 1996).

Cytochrome c oxidase is thought to exert a major control on mitochondrial respiration (Arnold, 2012). Indeed, its activity has been widely used as a proxy of oxidative phosphorylation capacity in many ectotherm species (Guderley and Seebacher, 2011; Lannig et al., 2003; Mark et al., 2012; Williard and Harden, 2011). However, complex IV activity is expected to have little effect on the maximal respiration rate of mitochondria *per se* but would rather act on respiration regulation through impact upon redox state of the respiratory chain (Blier and Lemieux, 2001). Indeed, due to its higher activity, complexes of this chain would mostly be in an oxidized state and, consequently, ensure a sharp thermodynamic gradient in the respiratory chain under a wide temperature range (Blier and Lemieux, 2001). In doing so, the complex IV excess capacity could allow a decrease of reactive oxygen species (ROS) production and proton leak which would reduce oxidative stress and energy waste, respectively (Blier et al., 2014; Kadenbach et al., 2010). If cybrids are better adapted to cold environments, an increase in water temperature could effectively provoke an overproduction of respiratory chain substrates (Blier et al., 2014). This excess could ultimately override the maximal functional capacity of complex IV, which might lead to a higher respiratory chain reduction state and thus, higher ROS production and proton leak. This hypothesis could explain why cybrids are not found at low latitudes where they would be maladapted to the warmer environment.

As previously described by Duggan et al. (2011), the increase of cytochrome c oxidase activity is not always correlated with a coordinated rise in expression of its thirteen subunit-encoding genes. Indeed, it has been shown that some of them can be upregulated whereas others remain at the same expression level (Duggan et al., 2011). In that case, rate-determining subunits might be those where changes in mRNA level best parallel complex IV activity. In the present

study, six of the thirteen complex IV subunit-encoding genes were considered (COX 1, 2, 3, 5b, 6b and 7a1) and no expression variation between wild type and cybrids was recorded. According to these results, we can assume that the modification of cytochrome c oxidase activity would not be caused by a change in expression of its subunit-encoding genes.

After sequencing mtDNA protein-encoding genes, only one non-conservative mutation was detected on the *COX3* subunit. According to preliminary analyses based on bovine (*Bos taurus*) complex IV crystal structure (PDB ID 2occ), *COX3* variable residue could possibly be in contact with the nuclear-encoded subunit *COX7a1* (distance ≤ 4 Å). According to Melvin et al. (2008), an amino acid change is considered likely to affect complex IV activity if its property is changed and/or if the site is in close contact with another subunit. Indeed, substitutions able to modify interactions between nuclear and mitochondrial subunits at positions of structural importance to enzyme function are known to modify enzymatic activity (Barrientos et al., 2000; Breton et al., 2009; Melvin et al., 2008). Therefore, the change at residue 41 on the *COX3* mitochondrial subunit could potentially explain the difference of complex IV activity observed between wild type and cybrids.

The quantitative proteome comparison showed small differences within biotypes but a tremendous one between them. This quantitative difference in expression is consistent with a previous qualitative analysis that showed a major difference between the two biotypes (Angers et al. 2012). Indeed, almost 40% of the analysed proteins showed a more than 2-fold difference of expression between wild type and cybrids. Together with results from Angers et al. (2012) based on the presence/absence of expression (not considered here), our results reveal a massive

difference between biotypes. In another proteomic study, Lopez et al. (2001) discovered that 48.9% of 92 analysed proteins showed a significant difference of expression between mussels *Mytillus galloprovincialis* living in totally different physical (temperature, desiccation and turbulence) and ecological (food availability and predation) conditions. This indicates that the major proteomic modifications triggered by a drastic change of habitat can almost be equaled by cellular environment changes caused by the presence of foreign mitochondria. Moreover, we can speculate that such a large reprogramming effect could eventually modify the expression of genes other than those analysed in this study, leading to the observed modification of cytochrome c oxidase activity (e.g. expression of complex IV nuclear subunit isoforms).

The introgression of foreign mitochondria might impact whole organism performance traits, as exemplified in the seasonal migratory behavior of Audubon's warbler (*Setophaga auduboni*) (Toews et al., 2014). In the present study, the critical swimming speed analysis revealed that cybrids present a higher swimming performance than wild type. Although our results were achieved at 17°C instead of 24°C, we can assume that the observed difference would be maintain or even increased at higher temperatures. Here, we make the assumption that the higher swimming capacity of cybrids results from more efficient mitochondria caused by the modification of complex IV activity and/or by the observed proteome reprogramming. To test this hypothesis, an analysis of mitochondrial respiration in *C. eos* wild type and cybrids, would be ideal.

As previously mentioned, the enzymatic activity survey was performed on individuals captured in different lakes where either wild type or cybrids were found. However, because of

sampling limitations, we were not able to do so for gene expression and proteome reprogramming analyses. Although this approach is a great way to test for the effect of lakes within biotype, it is important to note that there may still exist an effect of lakes between wild type and cybrids, which we cannot assess. This effect, which includes many biotic and abiotic factors, could be responsible of the proteome and complex IV variations detected between the two biotypes. Indeed, numerous studies already have demonstrated how the environment can influence the phenotype (Blier and Guderley, 1993; Blier and Lemieux, 2001; Garner and Burg, 1994; Hochachka and Somero, 2002; Lucassen et al., 2003; Pelster et al., 2003). Moreover, since both biotypes were sampled from distinct populations, nuclear genome differences could also potentially influence our data. However, because the estimated divergence time for these populations is quite recent (≤ 10.000 years; Angers and Schlosser, 2007), this impact is likely limited.

In conclusion, this comparative study of *Chrosomus eos* wild type and cybrids provides insights into the effects of foreign mitochondria on phenotype, at various levels of biological organization. Our results demonstrate that the presence of foreign mitochondria can severely modify enzymatic activity, as well as the proteome and that these changes can ultimately impact functional traits such as swimming capacity. As predicted, our results reveal that the absence of coevolution for a period of approximately seven million years between mitochondrial and nuclear genomes is not necessarily deleterious but could even be beneficial. Indeed, the capture of foreign mitochondria could be an efficient way to circumvent the selection process of genomic coevolution, allowing the rapid accumulation of new mutations in *C. eos* cybrids.

2.5 Acknowledgments

The authors are grateful to Sophie Breton and Pierre Magnan for providing the necessary equipment for qRT-PCR and critical swimming speed assays, respectively. We also thank the members of B. Angers laboratory for their invaluable assistance. LD was partly supported by a PhD scholarship from the GRIL. This research was supported by a team grant from the *Fonds Québécois pour la Recherche sur la Nature et les Technologies* (FQRNT) to BA and AA.

CHAPITRE 3 – ARTICLE 2

"*Ménage à trois*" among sympatric wild type and cybrids of the red belly dace *Chrosomus eos*

Léo Deremiens, Roland Vergilino, Bernard Angers

Soumis à la revue BMC Evolutionary Biology

Résumé

De nombreuses études ont démontré l'influence de l'introgression mitochondriale sur le phénotype. Toutefois, nos connaissances restent limitées concernant l'impact de mitochondries allospécifiques sur la coadaptation intergénomique, ainsi que sur la différenciation génétique entre les individus cybrides et leur homologue sauvage en milieu naturel. Parmi les nombreux biotypes retrouvés au sein du complexe *Chrosomus eos-neogaeus*, un cybride apparu spontanément en milieu naturel, formé d'un ADNnc *C. eos* et d'un ADNmt *C. neogaeus*, peut être distingué. Dans cette étude, nous avons mesuré le niveau de différenciation existant entre trois formes différentes de *C. eos*, retrouvées en sympatrie dans la région des Laurentides : la forme sauvage et deux formes cybrides divergentes issues des refuges glaciaires Mississippien et Atlantique. Les trois formes étaient uniformément distribuées parmi les sites échantillonnés. Bien qu'une différenciation génétique faible mais significative ait été détectée entre ces sites, les analyses de migration révélaient un important flux génétique entre ces formes, menant à une homogénéisation génétique de celles-ci. De manière générale, ces résultats démontrent l'absence de barrières reproductives entre les formes sauvages et cybrides de *C. eos*, lorsqu'en sympatrie. Un tel flux de gènes pourrait entraver la mise en place de nouveaux allèles localement adaptatifs, suggérant que l'ADNnc *C. eos* et l'ADNmt *C. neogaeus* des cybrides n'ont jamais eu la chance de s'adapter l'un à l'autre, lorsque ce biotype est retrouvé en sympatrie avec la forme sauvage.

Abstract

Numerous studies have demonstrated an influence of mitochondrial introgression on phenotype. However, little is known about the impact of allospecific mitochondria on the intergenomic coadaptation as well as on the genetic differentiation between wild type and the introgressed forms in natural environments. Among the many biotypes of the *Chrosomus eos-neogaeus* fish complex, a spontaneous natural cybrid *C. eos* with a *C. neogaeus* mtDNA can be distinguished. In this study, we measured the differentiation of three different sympatric forms of *C. eos* found in a Laurentian lake: the wild type and two divergent cybrids from the Mississippian and the Atlantic glacial refuges. The three forms were uniformly distributed across the sampling sites. A low but significant genetic differentiation was detected among sites within the lake. On the other hand, migration estimates showed an important gene flow among the forms, leading to a lack of significant genetic differentiation. Overall, these results demonstrate the absence of reproductive barrier between wild type and cybrids in sympatry. Such continuous gene flows could potentially prevent the establishment of locally adaptive alleles, indicating that the *C. eos* ncDNA and the *C. neogaeus* mtDNA of cybrids do not adapt to each other when living in sympatry with wild types.

3.1 Introduction

Mitochondria are essential organelles to cellular activity and individual survival. They are involved in many cellular processes such as energetic metabolism, apoptosis or signaling pathways. In animals, the mitochondrial genome (mtDNA) is maternally inherited (Gray et al., 2001), rarely recombine (Lynch, 1996), and undergo higher rates of nucleotide substitution compared with the nuclear DNA (ncDNA) (Moritz et al., 1987; Pesole et al., 1999). This genome only encodes a few proteins, but they are critically positioned in the energy producing complexes of the respiratory chain (Andersson et al., 1998; Garesse and Vallejo, 2001). These complexes require highly specific interactions between nuclear and mitochondrial subunits to be fully operational, implying a strong coevolution between genomes across time (Blier et al., 2001; Burton et al., 2013; Osada and Akashi, 2012; Rand et al., 2004).

Despite this coevolution process, between-species transfers of mtDNA are frequently observed in both animals and plants (Arnold, 2012; Boratynski et al., 2011; Deremiens et al., 2015; Du et al., 2011; Glemet et al., 1998; Senjo et al., 1999; Toews and Brelsford, 2012; Zielinski et al., 2013). Such transfers are expected to be deleterious when mito-nuclear interactions specificity is altered (Blier et al., 2001; Ellison and Burton, 2006; Kenyon and Moraes, 1997; McKenzie et al., 2003; Rand et al., 2004), but could also be neutral (Blier et al., 2006; Boratynski et al., 2011; McKenzie et al., 2003) or even induce beneficial effects (Deremiens et al., 2015; Pichaud et al., 2012; Toews et al., 2014). Whether deleterious or beneficial, all these mitochondrial introgressions lead to modifications of the phenotype, at various levels of biological organisation (Angers et al., 2012; Cannon et al., 2011; Deremiens et al., 2015; McKenzie et al., 2003; Pichaud et al., 2012; Toews et al., 2014). As mitochondrial

introgression generally trigger phenotypic changes, foreign mitochondria could go as far as to induce genetic differentiation and speciation between wild type and introgressed forms in natural environments (Ellison and Burton, 2008a; Gershoni et al., 2009). A common condition to the differentiation process in speciation event and the coadaptation of the mitochondrial and the nuclear genomes of introgressed organisms is the reduction of gene flow between the different forms (Mallet et al., 2009). Therefore, processes leading to speciation would also allow the ncDNA and the mtDNA of cybrids to progressively adapt to each other.

Among the many biotypes of the *Chrosomus eos-neogaeus* fish complex (Cyprinidae, Pisces), a spontaneous natural cybrid *C. eos* with a *C. neogaeus* mtDNA can be distinguished (Dawley et al., 1987; Goddard et al., 1998). These cybrids originate from triploid hybrids *C. eos-neogaeus x eos*. It is assumed that, during gametogenesis of such triploid individuals, the unmatched set of chromosomes is discarded and meiosis solely occurs on the *C. eos* diploid genome, producing haploid eggs (Goddard et al., 1998). When fertilized by a *C. eos* male, these eggs give birth to individuals with a pure *C. eos* nuclear diploid genome but a female inherited *C. neogaeus* mitochondrial genome (Angers and Schlosser, 2007; Angers et al., 2012; Binet and Angers, 2005; Deremiens et al., 2015). As does *C. eos* wild type, these cybrids reproduce sexually.

C. eos wild type and cybrids live in the freshwater lakes of North America. Based on previous estimates from comparison of restriction fragment length polymorphism (RFLP) over the complete mitochondrial genome (Goddard et al., 1989), of the partial *cox1* gene region (Angers and Schlosser, 2007), and of the *cox1*, *cox2* and *cox3* genes (Deremiens et al., 2015),

C. neogaeus mtDNA differs by 7.6 to 9.6% from *C. eos* mtDNA. Additionally, it has been demonstrated that methylome, enzymatic activity, proteome as well as swimming performance are strongly different between allopatric wild type and cybrids (Angers et al., 2012; Deremiens et al., 2015). These important phenotypic differences reflect divergent evolutionary trajectories that may result in the allopatric formation of distinct species.

Recently, both wild type and cybrids have been found in sympatry (Mee and Taylor, 2012). According to the metabolic and physiologic differences observed in allopatry, coexistence of both forms provides an ideal system to investigate the effects of mitochondrial introgression on evolutionary processes, such as genomes coadaptation and sympatric speciation. During an ongoing speciation process, a reduction of gene flow between sympatric organisms is expected (Mallet et al., 2009) at more or less long term (Feder et al., 2012; Gagnaire et al., 2013). However, no data about the gene flow between sympatric wild type and cybrids is available.

This study aims at assessing the extent of gene flow between sympatric *C. eos* wild type and cybrids. Our first assumption is that phenotypic differences resulting from the allospecific mitochondria may lead to a reduction of gene flow between sympatric wild type and cybrids. In addition to this differentiation process, we hypothesize that low gene flow may lead to the establishment of locally adaptive alleles in cybrids, which allows their ncDNA and mtDNA to progressively adapt to each other.

3.2 Materials and methods

Sampling

In the summer 2004, 282 fish were sampled at four different sites in Desjardins Lake, southern Quebec, Canada (45°54'58.00"N, 74°04'23.00"W; Figure 3.1). From these fish, a small piece of caudal fin was withdrawn and preserved in 95% ethanol. Total DNA was extracted by proteinase K digestion followed by phenol-chloroform purification and ethanol precipitation (Sambrook et al., 1989). DNA samples were conserved at -20°C for approximately ten years, until analyses.

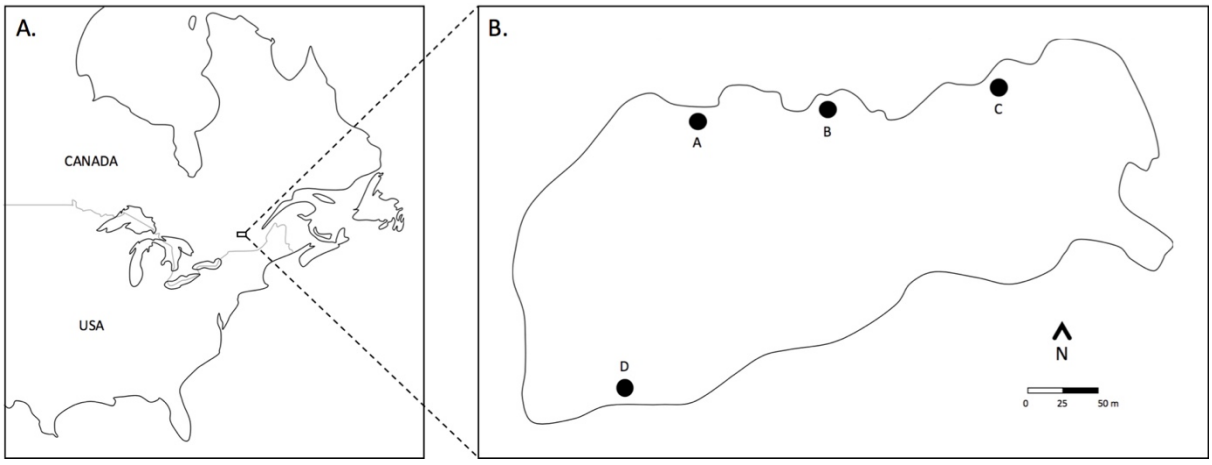

Figure 3.1: Location of the studied site in North America (A) and map of Desjardins Lake (B). Solid circles (A, B, C, D) correspond to the sampled sites.

Biotype identification and genotyping

Individuals sampled from Desjardins Lake were previously genetically identified and genotyped by Massicotte et al. (2008). Briefly, an intron of the mesoderm specific transcript gene (PEG/MST1) was amplified to differentiate *C. eos* from hybrids (Binet and Angers, 2005). Subsequently, *C. eos* individuals were genotyped with five polymorphic microsatellite markers:

Pho-1 and *Pho-2* (Binet and Angers, 2005) and *Pho-4*, *Pho-60* and *Pho-61* (Angers and Schlosser, 2007). PCR conditions and polymorphism detection were as described by Binet and Angers (2005).

Along with *C. eos-neogaeus* hybrids, Desjardins Lake harbors *C. eos* individuals consisting of a mixture of wild type and cybrids. To assess *C. eos* mitochondrial haplotype, a small fragment (200 bp) of the mitochondrial-encoded gene *nd3* was amplified using the following primers: ND3_cybF (5'-CCCAGGGAAAGATAATGAACT) and ND3_cybR (5'-GAGAATTGCTACGAGGAAGA). This *nd3* gene fragment allows the differentiation of both wild type and cybrids on a single polyacrylamide gel using the single-strand conformation polymorphism (SSCP) method (Orita et al., 1989). Amplifications were carried out with the following conditions: one cycle of denaturation at 92°C for 30 sec, and 45 cycles of denaturation at 92°C for 30 sec, annealing at 52°C, 54°C for 20 sec, and extension at 68°C for 30 sec, with a final extension at 68°C for 2 min. PCR products were separated on an 8% non-denaturing polyacrylamide gel (acrylamide/bisacrylamide 37.5:1) for 12h at 15W and 4°C in TBE 1X. To confirm the identification as wild type or cybrids, a fragment overlapping the *nd3*, tRNA Arg and *nd4l* genes as well as a fragment of the *cox1* mitochondrial-encoded gene were amplified and sequenced (*nd3-nd4L* primers: 5'-CCCAGGGAAAGATAATGAACT and 5'-TGGCTACTAGGAGTGCAAG; *cox1* primers: see Mee and Taylor (2012)).

Population genetics analyses

Genetic diversity was estimated for each sampled site, haplotype and haplotype/site in Desjardins Lake through the total number of alleles (A), the allelic richness (A_R; El Mousadik

and Petit, 1996) and Nei's gene diversity (H_E; Nei, 1987) using the program FSTAT (version 2.9.3; Goudet, 2001). The difference in genetic diversity among sites, haplotypes and haplotypes/site was tested using a Friedman test on A_R and H_E values. All statistical analyses were performed using the program SPSS Statistics 22 (SPSS Inc., Chicago, IL, USA). Allelic differentiation was tested among haplotypes/site and sites/haplotype with the exact G test by means of a Markov chain Monte Carlo (MCMC) variation using the program GENEPOP (version 4.2; Raymond and Rousset, 1995). Linkage disequilibrium between each nuclear locus and between each nuclear locus and mitochondrial haplotype was tested for each site and significance was assessed by 10,000 iterations using GENEPOP (version 4.2; Raymond and Rousset, 1995). Sequential Bonferroni corrections were applied to all multiple comparisons (Holm, 1979; Rice, 1989).

The expected heterozygosity at mutation-drift equilibrium (H_{EQ}) was assessed under the infinite allele model (IAM) using the program BOTTLENECK (version 1.2.02; Piry et al., 1999). To test for mutation-drift equilibrium, the Wilcoxon test (one tail for heterozygosity excess) was applied under the same model and using the same program. An alternative hypothesis explaining the results obtained with BOTTLENECK was tested using the simulation program EASYPOP (version 2.0.1; Balloux, 2001). We simulated a recent admixture of wild type and cybrids by generating data sets for two populations of 10,000 diploid individuals genotyped at 50 microsatellite loci. The stepwise mutation model (SMM) was selected, 20 possible allelic states were chosen and the mutation rate was set to $\mu = 10^{-4}$. 10,000 generations were generated whose the first 9,950 were used to obtain populations at mutation-drift equilibrium (no migration). To simulate a recent admixture, the last 50 generations were

followed using a migration rate of one. 20 replicates were computed. These replicates were analysed with the BOTTLENECK program.

The genetic organisation of *C. eos* individuals within Desjardins Lake was assessed using hierarchical analyses of molecular variance (AMOVA) conducted with the program Arlequin (version 3.5; Excoffier et al., 2005). Individuals were organized according to their sampling site (groups = sites; populations = haplotypes) as a spatial structure was previously detected as significant in this lake (Massicotte et al., 2008). We also tested whether reproductive isolation exists between wild type and cybrids by grouping individuals by haplotype (groups = haplotypes; populations = sites). All analyses were carried out with 10,000 permutations based on distance matrices.

Genetic structure of the *C. eos* population within Desjardins Lake was also investigated using the Bayesian clustering program STRUCTURE (version 2.3.2; Falush et al., 2003; Pritchard et al., 2000). Based on genotype data, this program clusters individuals into *K* populations and allows the identification of migrants and admixed individuals. An admixture model with correlated frequencies was used, the *K* values were set from one to twelve, and three runs were performed for each value of *K*. Each run included a burn-in period of 1,000,000 steps followed by 2,000,000 MCMC repetitions. Moreover, to corroborate and extend STRUCTURE analysis, another Bayesian analysis using a greedy stochastic optimization algorithm was achieved with the program BAPS (version 6.0; Corander and Marttinen, 2006; Corander et al., 2008). This program not only explores the number of clusters that best explain the population structure but it also allows to compare the posterior probabilities of various clustering scenarios.

Regarding how many clusters best represent the population of Desjardins Lake, ten runs for each value of *K* (set from one to twelve) were conducted. On the other hand, the tested clustering scenarios relied on the haplotypes and the sampled sites of the *C. eos* fish (Table A2). The prior probability for each scenario followed a uniform distribution considering that each scenario was equally possible.

The coalescence-based program MIGRATE-N (version 3.6; Beerli, 2006) was used to test for and estimate the direction and magnitude of gene flow between wild type and cybrids. MIGRATE-N allowed to assess the mutation-scaled migration rates $M = m/\mu$, where m is the immigration rate per generation among populations and μ the mutation rate per generation per locus, as well as the mutation-scaled effective population size $\Theta = 4N_e\mu$, where N_e is the effective population size. Migration rates were estimated from individuals coming from the local site D, where each biotype is the most abundant (Table 3.1). For each run, Bayesian analysis was selected, the Brownian microsatellite model was set and the initial genealogies were started from a random tree. A full migration matrix model was adopted and initial Θ and migration rate values were generated using the default F_{ST} calculation. Uniform priors were selected for these two parameters. All analyses were run with one long chain, a sampling increment of 1,000, a burn-in of 3,000, a sampling of 100,000, 10 replicates and a static heating with four temperatures (1, 1.5, 3 and 1,000,000).

Table 3.1: Abundance, diversity indices (mean ± standard deviation) and estimation of mutation-drift equilibrium (MDE) under the infinite allele model (IAM) for each haplotype/site, each haplotype and each site in Desjardins Lake. Significant *P*-values of MDE probability tests are indicated in bold.

Population	n (%)	A	A_R	H_E	H_{EQ}	MDE probability test
A - MC	53 (70.67)	9.6 ± 3.5	6.0 ± 1.6	0.785 ± 0.076	0.698 ± 0.090	**0,016**
A - AC	9 (12.00)	6.4 ± 1.5	6.4 ± 1.4	0.813 ± 0.050	0.767 ± 0.069	**0,016**
A - WT	13 (17.33)	7.0 ± 2.4	6.3 ± 1.7	0.776 ± 0.084	0.737 ± 0.128	0,312
B - MC	51 (68.93)	10.6 ± 3.7	6.5 ± 1.6	0.792 ± 0.082	0.726 ± 0.103	**0,031**
B - AC	12 (12.21)	6.8 ± 1.6	6.2 ± 1.1	0.782 ± 0.065	0.755 ± 0.070	0,078
B - WT	11 (14.86)	5.8 ± 2.7	5.4 ± 1.9	0.755 ± 0.094	0.667 ± 0.157	**0,016**
C - MC	35 (57.38)	8.2 ± 2.2	5.9 ± 1.3	0.783 ± 0.054	0.692 ± 0.092	**0,016**
C - AC	14 (22.95)	6.6 ± 2.6	5.6 ± 1.6	0.770 ± 0.055	0.699 ± 0.144	0,313
C - WT	12 (19.67)	6.6 ± 2.1	6.0 ± 1.6	0.790 ± 0.088	0.722 ± 0.120	**0,016**
D - MC	41 (56.95)	10.0 ± 3.8	6.7 ± 1.5	0.816 ± 0.067	0.723 ± 0.110	**0,016**
D - AC	16 (22.22)	6.8 ± 2.3	6.0 ± 1.6	0.786 ± 0.100	0.699 ± 0.116	**0,016**
D- WT	15 (20.83)	6.6 ± 1.5	5.8 ± 1.0	0.786 ± 0.070	0.718 ± 0.071	**0,031**
MC	180 (63,84)	13 ± 5.1	12.9 ± 5.0	0.796 ± 0.067	0.695 ± 0.125	**0,016**
AC	51 (18,08)	9.6 ± 2.9	9.5 ± 2.8	0.786 ± 0.074	0.706 ± 0.091	**0,016**
WT	51 (18,08)	9.8 ± 4.0	9.62 ± 3.7	0.772 ± 0.079	0.699 ± 0.123	**0,031**
A	75	11.0 ± 4.0	10.6 ± 3.7	0.785 ± 0.073	0.711 ± 0.104	**0,016**
B	74	10.8 ± 4.1	10.4 ± 3.7	0.786 ± 0.078	0.703 ± 0.107	**0,016**
C	61	9.6 ± 2.9	9.6 ± 2.9	0.777 ± 0.057	0.690 ± 0.109	**0,016**
D	72	10.6 ± 4.2	10.2 ± 3.7	0.802 ± 0.075	0.698 ± 0.115	**0,016**
Total	282	13.8 ± 5.5	13.7 ± 5.4	0.790 ± 0.069	0.688 ± 0.133	**0,016**

Note. n – number of individuals; % - haplotype percentage; A – number of alleles; A_R – Allelic richness; H_E – observed gene diversity; H_{EQ} – estimated gene diversity under MDE (IAM); A, B, C, D – studied sites; MC – Mississippian Cybrid; AC – Atlantic Cybrid; WT – Wild Type.

3.3 Results

Biotype identification

The SSCP screening using *nd3* mitochondrial gene fragment confirms the results of Angers and Schlosser (2007) stating that *Chrosomus eos* wild type and cybrids are both found in Desjardins Lake. Surprisingly, SSCP analyses also reveal the presence of a third haplotype, in addition to the ones expected for the wild type and the cybrid. When compared to Genbank sequences, *nd3-nd4L* and *cox1* sequences amplified from one of these biotypes (Genbank accession numbers: KU556702 to KU556705) are attributed to *C. eos* while the sequences of the two other (Genbank accession numbers: KU556700, KU556701, KU556703, KU556704) belong to *C. neogaeus*. This indicates the presence of the wild type *C. eos* and two distinct cybrid haplotypes in Desjardins Lake. The *cox1* sequences obtained from the two different cybrids diverge by 13 nucleotides over 1,137 base pairs and are closely related to the haplotypes A (Genbank accession number: KU556701) and F (Genbank accession number: KU556700) of Angers and Schlosser (2007) and Mee and Taylor (2012) studies. The haplotype A is found in numerous drainage basins across North America and is suspected to originate from the Mississippian refuge while the haplotype F is only found along the North American East coast and is suspected to originate from the Atlantic refuge (Goddard et al., 1989; Mee and Taylor, 2012). According to their probable origin, the cybrid haplotypes are henceforth named Mississippian and Atlantic cybrids.

When *nd3-nd4l* and *cox1* sequences are combined, 20 variable sites over 1,667 bp can be found between the two cybrids. This variability represents a difference of 1.20%. When wild type and Mississippian cybrids are compared, a difference of 8.38% is detected (132 variable sites/1,575 bp). The comparison of wild type and Atlantic cybrids lead to a difference of 8.44%

(133 variable sites/1,575 bp). The last two results are slightly below the 9.6% of Goddard et al. (1989) and Angers and Schlosser (2007) but are similar to what was found by Deremiens et al. (2015).

Analysis of the 282 individuals reveals that the three haplotypes are found at each sampling site of Desjardins Lake. The relative abundance of each haplotype vary across these sites (Table 3.1) with an average of 63.84%, 18.08% and 18.08% for the Mississippian, Atlantic and wild type *C. eos* haplotypes respectively.

Population genetics analyses

Diversity indices, differentiation tests and mutation-drift equilibrium

As shown in Table 3.1, the microsatellite loci were extremely variable and exhibited 13.8 ± 5.5 alleles over the 282 studied individuals. The allelic richness (A_R) varied from 9.6 ± 2.9 to 10.6 ± 3.7 across sites and from 9.5 ± 2.8 to 12.9 ± 5.0 across haplotypes (Table 3.1). No significant linkage disequilibrium between nuclear loci (Fisher's method, $P > 0.1982$) or between nuclear loci and mitochondrial haplotypes (Fisher's method, $P > 0.1567$) was detected. When diversity indices were compared, no significant difference was found among sites or haplotypes/site but a significant difference was detected among haplotypes for A_R (Friedman test, $P = 0.022$) and H_E (Friedman test, $P = 0.007$), with a higher diversity observed in Mississippian cybrids (Table 3.1).

Differentiation tests performed among haplotypes within site revealed a single significant difference between wild type and Mississippian cybrids from site A (Table 3.2). When a given

haplotype is compared across sites, several combinations involving the Mississippian cybrids were significantly different (Table 3.2).

Of the twenty tested groups and subgroups of Desjardins Lake population (sites, haplotypes, haplotypes/site and the whole population), seventeen displayed a significant departure from mutation-drift equilibrium ($P < 0.05$; Table 3.1). In most cases, the estimated H_E were higher than the H_{EQ} values (Table 3.1), reflecting a H_E excess.

Such H_E excess is usually explained by a population bottleneck effect (Luikart and Cornuet, 1998). However, given the different origins expected for wild type and cybrids forms, an alternative hypothesis of admixture was tested using simulations. The H_E and H_{EQ} values obtained with simulations (Table A3) were similar to the ones measured (Table A4), suggesting that admixture events that occurred in Desjardins Lake may mimic demographic fluctuations.

Table 3.2: Differentiation tests among haplotypes/site and sites/haplotype. Significant *P*-values are indicated in bold.

Population pair		Chi^2	df	*P*
A. Haplotypes/site				
Site A	MC *vs* AC	6.26	10	0.7928
	MC *vs* WT	21.52	10	**0.0178**
	AC *vs* WT	9.56	10	0.4797
Site B	MC *vs* AC	6.49	10	0.7722
	MC *vs* WT	12.01	10	0.2844
	AC *vs* WT	11.05	10	0.3537
Site C	MC *vs* AC	8.62	10	0.5680
	MC *vs* WT	7.04	10	0.7221
	AC *vs* WT	7.83	10	0.6458
Site D	MC *vs* AC	9.16	10	0.5169
	MC *vs* WT	5.65	10	0.8435
	AC *vs* WT	4.84	10	0.9015
B. Sites/haplotype				
MC	A *vs* B	15.83	10	0.1047
	A *vs* C	20.53	10	**0.0247**
	B *vs* C	11.22	10	0.3408
	A *vs* D	27.30	10	**0.0023**
	B *vs* D	32.07	10	**0.0004**
	C *vs* D	20.15	10	**0.0279**
AC	A *vs* B	10.61	10	0.3887
	A *vs* C	8.99	10	0.5335
	B *vs* C	11.03	10	0.3549
	A *vs* D	4.91	10	0.8974
	B *vs* D	14.39	10	0.1558
	C *vs* D	17.20	10	0.0701
WT	A *vs* B	11.61	10	0.3123
	A *vs* C	8.45	10	0.5850
	B *vs* C	8.34	10	0.5958
	A *vs* D	8.21	10	0.6088
	B *vs* D	6.49	10	0.7726
	C *vs* D	4.66	10	0.9128

Note. WT - Wild Type; MC - Mississippian Cybrids; AC - Atlantic Cybrids; df – degree of freedom; *P* = *P*-values.

Genetic organisation of *C. eos*

To assess whether the sampled individuals clustered according to their haplotype ($K = 3$), their site of origin ($K = 4$) or both ($K = 12$), STRUCTURE analyses were carried out with K values set from one to twelve. In all runs, the highest likelihood was obtained for $K = 1$ (Figure 3.2A). Figure 3.2B also illustrates the absence of structure among the sampled fish. This result was partially confirmed by BAPS, as four clustering scenarios were equally probable and displayed the highest likelihood (log likelihood = -5269, Table A2): (*i*) $K = 1$, no partition across the population, (*ii*) $K = 2$, site C is differentiated from the other sites, (*iii*), $K = 2$, site D is differentiated from the other sites, (*iv*) $K = 2$, sites C + D are differentiated from sites A + B. None of the clustering scenarios involving partition among haplotypes presented a high likelihood value, the most likely being the one with wild type differentiated from the Mississippian + Atlantic haplotypes (log likelihood = -5338, Table A2). Therefore, haplotypes across sites do not explain the partitioning of the population.

Hierarchical AMOVA also support these results and provides no significant F_{CT} when individuals are organised by haplotypes (Table 3.3). On the other hand, a low but significant genetic differentiation was detected among sites ($F_{CT} \leq 0.00515$, $P < 0.05$; Table 3.3) while the "haplotype within site" component is not significant ($P > 0.7561$). This suggests that *C. eos* individuals from Desjardins Lake are spatially organized but do not preferentially reproduce according to their haplotype.

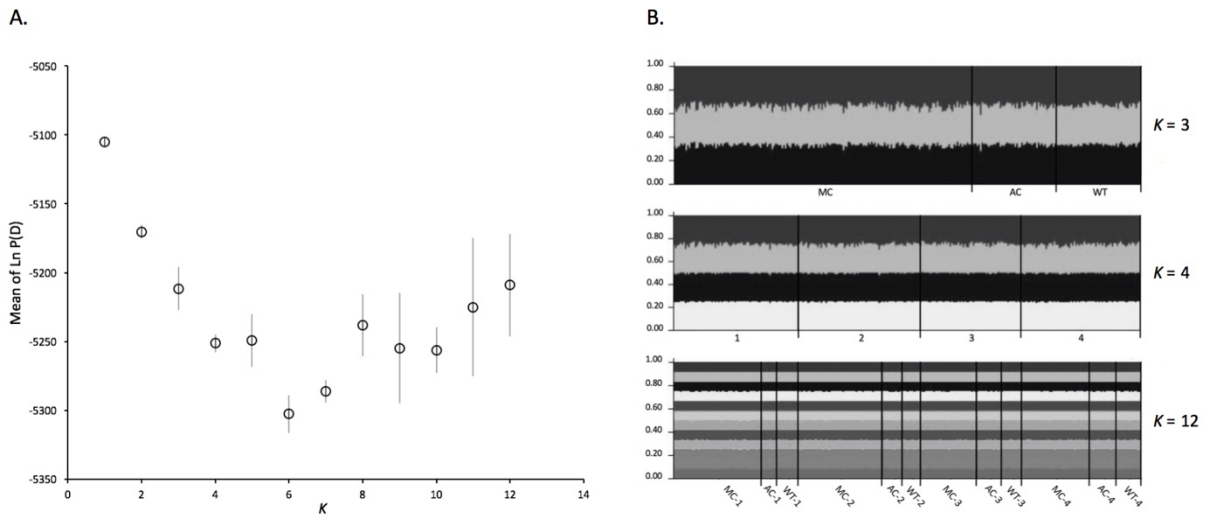

Figure 3.2: Estimated population structure of wild type (WT), Mississippian cybrids (MC) and Atlantic cybrids (AC) sampled in four sites (1, 2, 3, 4) of Desjardins Lake using STRUCTURE. (A) Mean of the estimated probability of the data given K = 1 to 12. (B) Bar plot where each individual is represented by a thin vertical line, which is partitioned into K segments that represent its estimated population group membership fractions. Black lines separate individuals according to their haplotype (K = 3), their sampled site (K = 4) or both (K = 12).

Migration rates

Estimation of N_e and m among wild type, Mississippian and Atlantic cybrids were performed using Bayesian inference and the Brownian motion mutation model. All the estimates of the mutation-scaled effective population Θ were scaled using a microsatellite mutation rate of 5.56 $\times 10^{-4}$ per locus per generation to calculate N_e (Bradic et al., 2012; Yue et al., 2007). The mean Θ obtained for wild type, Mississippian and Atlantic cybrids were 1.20, 7.95 and 1.21 respectively (Table 3.3). These values correspond to an effective population size of ~540 wild types, ~3,576 Mississippian cybrids and ~544 Atlantic cybrids (Table 3.4). The haplotype proportion obtained from these values (i.e. 11.59% wild type, 76.74% Mississippian cybrids and 11.67% Atlantic cybrids) is roughly of the same order as the one observed in the sampled

individuals (i.e. 18.08% wild type, 63.84% Mississippian cybrids and 18.08% Atlantic cybrids). Regarding the migration rates, *M* varied from 8.64 to 26.89 with the highest values attributed to "from Mississippian to Atlantic cybrids" and "from Atlantic cybrids to wild type" (Table 3.5). On the other hand, *m* varied from 2.60 to 20.90 and its highest values were detected for "from Atlantic to Mississippian cybrids" and "from wild type to Mississippian cybrids" (Table 3.5). Although *M* and *m* values reveal a high gene flow, it is important to mention that, because of their large confidence interval, gene flow magnitude differences among *C. eos* forms are not reliable.

Table 3.3: Analyses of molecular variance (AMOVA) of the four haplotype combinations obtained from wild type, Mississippian and Atlantic cybrids when groups = haplotypes (A) and when groups = sites (B). Significant *P*-values are indicated in bold.

Structure tested	SS	VC	%VAR	Fstat	*P*
A. Groups = haplotypes – Populations = sites					
Groups: WT *vs.* MC *vs.* AC - Populations: A *vs.* B *vs.* C *vs.* D					
Among groups	3.488	-0.00163	-0.08273	-0.00083	0.88356
Among populations within groups	17.866	0.00021	0.01081	0.00011	0.42535
Within populations	1034.637	1.97569	100.07192	-0.00072	0.55812
Groups: WT *vs.* [MC+AC] - Populations: A *vs.* B vs C *vs.* D					
Among groups	2.132	0.00091	0.04589	0.00046	0.59436
Among populations within groups	12.125	0.00063	0.03187	0.00032	0.33228
Within populations	1041.050	1.97445	99.92224	0.00078	0.38079
Groups: AC *vs.* [WT+MC] - Populations: A *vs.* B *vs.* C *vs.* D					
Among groups	1.324	-0.00500	-0.25347	-0.00253	0.99257
Among populations within groups	13.325	0.00372	0.18848	0.00188	0.17317
Within populations	1040.658	1.97366	100.06499	-0.00065	0.29307
Groups: MC *vs.* [WT+AC] - Populations: A *vs.* B *vs.* C *vs.* D					
Among groups	1.684	-0.00184	-0.09316	-0.00093	0.83337
Among populations within groups	12.953	0.00280	0.14185	0.00142	0.19545
Within populations	1040.670	1.97367	99.95131	0.00049	0.29119
B. Groups = sites – Populations = haplotypes					
Groups: A *vs.* B *vs.* C *vs.* D - Populations: WT *vs.* MC *vs.* AC					
Among groups	7.673	0.00790	0.39970	0.00400	**0.01931**
Among populations within groups	13.634	-0.00760	-0.38432	-0.00386	0.89921
Within populations	1033.993	1.97599	99.98462	0.00015	0.57871
Groups: A *vs.* B *vs.* C *vs.* D - Populations: WT *vs.* [MC+AC]					
Among groups	7.646	0.01017	0.51471	0.00515	**0.04891**
Among populations within groups	6.612	-0.00833	-0.42131	-0.00423	0.86960
Within populations	1041.050	1.97445	99.90660	0.00093	0.39079
Groups: A *vs.* B *vs.* C *vs.* D - Populations: AC *vs.* [WT+MC]					
Among groups	7.646	0.00814	0.41212	0.00412	**0.03644**
Among populations within groups	7.003	-0.00551	-0.27889	-0.00280	0.76455
Within populations	1040.658	1.97366	99.86676	0.00133	0.28099
Groups: A *vs.* B *vs.* C *vs.* D - Populations: MC *vs.* [WT+AC]					
Among groups	7.646	0.00621	0.31434	0.00314	**0.0302**
Among populations within groups	6.991	-0.00358	-0.18113	-0.00182	0.75614
Within populations	1040.670	1.97367	99.86680	0.00133	0.29257

Note. WT - Wild Type; MC - Mississippian Cybrids; AC - Atlantic Cybrids; A, B, C, D - Local sites; SS - Sum of squares; VC - Variance components; %VAR - Percentage of variation; Fstat = F-statistics; *P* = *P*-values.

Table 3.4: Estimates of mutation-scaled effective population size (Θ) and effective population size (N_e) of wild type (WT), Mississippian (MC) and Atlantic cybrids (AC) in site D of Desjardins Lake.

Parameter		WT	MC	AC
Θ				
	2.5%	0.13	4.80	0.00
	97.5%	3.13	14.27	3.00
	Median	1.30	6.77	1.30
	Mean	1.20	7.95	1.21
N_e				
	2.5%	59.95	2158.27	0.00
	97.5%	1408.87	6414.87	1348.92
	Median	584.53	3042.57	584.53
	Mean	540.34	3576.15	543.60

Table 3.5: Mutation-scaled migration rates (*M*) and migration rates per generation (*m)* among wild type (WT), Mississippian (MC) and Atlantic cybrids (AC) in site D of Desjardins Lake.

Parameter	AC-MC	WT-MC	MC-AC	WT-AC	MC-WT	AC-WT
M						
0.025	0.00	0.00	2.00	0.00	0.00	0.00
0.975	28.00	28.67	48.67	30.00	27.33	48.00
Median	13.00	13.67	25.67	14.33	12.33	21.00
Mean	9.43	10.51	26.89	11.83	8.64	22.66
m						
0.025	0.00	0.00	0.60	0.00	0.00	0.00
0.975	55.67	57.00	14.71	9.07	8.21	14.42
Median	25.85	27.17	7.76	4.33	3.71	6.31
Mean	18.74	20.90	8.13	3.57	2.60	6.81

3.4 Discussion

An unsuspected finding of the present study is that Desjardins Lake harbors three different forms of *Chrosomus eos*: the wild type and two different cybrids harboring the mitochondrial genome of *C. neogaeus* from the Mississippian and the Atlantic glacial refuges. Based on *nd3-nd4l* and *cox1* gene sequences, Mississippian and Atlantic *C. neogaeus* mtDNA of cybrids are characterised by a divergence of 1.20%, while a difference of 8.38% and 8.44% is found between wild type and Mississippian cybrids and between wild type and Atlantic cybrids respectively. Assuming a 1.3% divergence per million years based on the *nd2* and *atp6* genes (Bermingham et al., 1997), these results imply that mitochondria of the two cybrids have diverged for approximately one million years while *C. eos* and *C. neogaeus* mtDNA would have diverged for about seven million years.

The three *C. eos* forms are present at each of the studied sites, with the Mississippian cybrid being the most abundant. Because of an extensive gene flow, the differentiation analyses and the performed AMOVA could not detect any genetic differentiation among these forms. However, overall sites, Mississippian cybrids were genetically more diversified than the wild type and Atlantic cybrids; their A_R and H_E values being the highest. This difference in diversity suggests that wild type and cybrids do not reproduce panmictically as these estimators are then expected to provide the very same values among forms.

On the other hand, a significant but low genetic differentiation characterizes fish from different sites. Similar results were obtained when multiple scenarios were compared using the Bayesian clustering program BAPS. This is in agreement with results previously obtained by

Massicotte et al. (2008), who showed through Hardy-Weinberg equilibrium probability test, allelic differentiation tests and F_{ST} analyses, that the whole lake is not organized as a unique panmictic reproductive unit. The average F_{ST} estimated between site (0.00530) by Massicotte et al. (2008) and our estimation of average F_{CT} (0.00407) are very close; the difference may be accounted for by the higher sample size in the present study. Such low levels of differentiation are generally characteristic of recent intralacustrine population divergence of fish (Barluenga et al., 2006; Strange and Stepien, 2007; Varnavskaya et al., 1994) due in the present case to *C. eos* homing behavior (Massicotte et al., 2008).

In Desjardins Lake, when studying the genetic diversity of each site, haplotype or haplotype/site, a significant departure from mutation-drift equilibrium with a heterozygosity excess is recorded for almost every tested population (Table 3.1). The presence of a unique hybrid, characterized by a haplotype different from the one of both cybrids (Angers and Schlosser, 2007), added to the complete absence of *C. neogaeus* in this lake, rules out the *in situ* formation of the introgressed *C. eos* forms. Colonization of cybrids from Mississippian and Atlantic glacial refuges, in addition to wild type, results in an admixture among these three differentiated groups. Considering a constant effective population size, such admixture may have resulted in the coexistence of alleles with relatively high and similar frequencies and the loss of the rare ones, as expected following a bottleneck effect (Garza and Williamson, 2001). This may have led to the heterozygosity excess observed. While we do not exclude founding events that are generally associated to postglacial colonization (Hewitt, 2000, 1996), admixture seems to be responsible, at least partially, of the observed disequilibrium.

The magnitude of gene flow measured among the three *C. eos* haplotypes appears to be relatively important. These groups were likely differentiated as they originated from at least two distinct glacial refuges. The current lack of genetic differentiation among these forms indicates admixture and extensive homogenization of their original gene pools. Such rates of gene flow demonstrate that despite the many phenotypic modifications triggered by their foreign mitochondria, cybrids remain able to reproduce with their wild type counterpart. Even more fascinating, high gene flows are known to prevent local adaptation, this continuous exchange hindering the establishment of locally adaptive alleles (Feder et al., 2012). According to this effect, our results reveal that the *C. eos* ncDNA and the *C. neogaeus* mtDNA of cybrids could likely not adapt to each other when both wild type and cybrid forms coexist in sympatry.

Alternatively, it has been shown that high gene flow can be accompanied by speciation (i.e. speciation-with-gene-flow) (Gagnaire et al., 2013; Nadeau et al., 2012; Nosil and Feder, 2011; Via, 2012). Gene flow may be sufficient to homogenize most of the nuclear genome while divergent selection, acting on a few regions, is strong enough to create genomic islands characterised by a low effective migration rate compared to the gross mutation rate between populations (Feder et al., 2012). Accordingly, regions under selection such as genes interacting with the mitochondrial genome could be characterised by a lower effective gene flow and would allow the mtDNA and the ncDNA of cybrids to progressively adapt to each other.

The foreign mitochondria of *C. eos* cybrids is known to induce numerous phenotypic modifications. In their study, Angers et al. (2012) demonstrate that the *C. neogaeus* mitochondria are able to impact the ncDNA methylation profile as well as to qualitatively

change the proteome (presence/absence) of this biotype. Moreover, according to Deremiens et al. (2015), cybrids also present a higher mitochondrial respiratory chain complex IV activity, a quantitative change in their proteome and a higher swimming capacity. Given all these modifications, taking place at various levels of biological organisation, we can hypothesize that wild type and cybrids generally colonize different kind of environments and thus explain why the two biotypes are usually found in allopatry (Angers et al., 2012; Deremiens et al., 2015; Mee and Taylor, 2012). The sympatry observed in Desjardins Lake could be the result of environmental conditions allowing wild type and cybrids to live and reproduce indiscriminately with each other. This would explain why AMOVA and Bayesian clustering analyses failed in detecting a significant genetic differentiation or cluster between the two biotypes respectively, in Desjardins Lake (Table 3.3, Figure 3.2A and B).

To conclude, acquisition of allospecific mitochondria is an efficient way to circumvent the selection process of genomic coevolution, resulting in multiple and possibly beneficial mutations in the mtDNA of cybrids. Surprisingly, an important gene flow exists among the sympatric wild type, Mississippian and Atlantic cybrids. This migration rate could prevent differentiation at the nuclear level and coevolution between nuclear and different mitochondrial genomes. While counterintuitive, this study reveals that intergenomic coevolution may not be a *sine qua non* condition to survive and perform in natural environments.

3.5 Acknowledgments

The authors are grateful to J. Shapiro for providing the necessary equipment for migration rate analyses. We also thank the members of B. Angers laboratory for their invaluable assistance. LD was partly supported by a PhD scholarship from the GRIL. This research was supported by a team grant from the *Fonds Québécois pour la Recherche sur la Nature et les Technologies* (FQRNT) to BA.

CHAPITRE 4 – ARTICLE 3

To what extent can foreign mitochondria impact the phenotype, a holistic study of the red belly dace *Chrosomus eos*

Léo Deremiens, Vincent Chapdelaine, Logan Schwartz, Annie Angers, Hélène Glémet, Bernard Angers

Résumé

Le transfert d'ADNmt entre espèces affecte généralement les fonctions mitochondriales ainsi que la valeur adaptative des individus. Toutefois, diverses études démontrent que de tels transferts peuvent également être bénéfiques. Bien que les introgressions mitochondriales soient connues pour induire des modifications phénotypiques, de nombreuses questions persistent encore à ce jour concernant l'étendue de cette influence sur les différents niveaux d'organisation du phénotype. Le complexe *Chrosomus eos* représente un système unique pour répondre à ces questions. Au sein de ce complexe, trois formes sympatriques différentes peuvent être distinguées dans un même lac : la forme sauvage et deux formes cybrides arborant l'ADNmt de *Chrosomus neogaeus* issus des refuges glaciaires Mississippien et Atlantique. Dans cette étude, nous évaluons les variations épigénétiques, transcriptomiques, protéomiques ainsi que l'activité du complexe IV de la chaîne respiratoire mitochondriale parmi les trois formes détectées. Tel qu'escompter, l'amplitude de ces variations était plus prononcée entre la forme sauvage et les cybrides qu'entre les cybrides Mississippiens et Atlantiques ; ces deux derniers étant caractérisés par des ADNmt moins divergents. Qui plus est, lorsque comparée à une étude similaire mais considérant des formes allopatriques, notre approche holistique révèle que les modifications phénotypiques détectées entre des biotypes sympatriques étaient moins importantes. Ce contraste pourrait provenir d'effets environnementaux et nucléaires dérivant de l'utilisation de poissons allopatriques. De façon générale, nos résultats démontrent que les mitochondries allospécifiques peuvent modifier le phénotype de *C. eos*, à différents niveaux d'organisation, lorsque la forme sauvage et les cybrides sont retrouvés en sympatrie.

Abstract

Between-species transfers of mtDNA usually have deleterious effects on mitochondrial function and individual fitness. However, various studies demonstrate that such transfers can also be beneficial. Although mitochondrial introgressions are known to induce phenotypic changes, many questions remain regarding how much these events can impact the phenotype, at various levels. The *Chrosomus eos* complex is a unique system allowing to address these questions. In this complex, we find three distinct forms living in sympatry in a single lake: a wild type and two different cybrids harboring the mtDNA of *Chrosomus neogaeus* derived from the Mississippian and the Atlantic glacial refuges. Here, we assessed epigenetic, transcriptomic, proteomic and enzymatic activity changes among these three forms. According to our results, each of these phenotypic levels displayed variations. As expected, because of the higher mtDNA divergence detected between wild type and cybrids, the magnitude of this variation was more important between these two than between Mississippian and Atlantic cybrids. Additionally, this holistic survey demonstrated that the phenotypic modifications observed between the sympatric wild type and cybrids were less pronounced than the ones detected from allopatric *C. eos* biotypes, analysed in a previous study. This discrepancy might be the result of environmental and nuclear effects introduced by the use of allopatric fish. Ultimately, our results show that allospecific mitochondria are able to impact *C. eos* phenotype, at various biological levels, when wild type and cybrids are found in sympatry.

4.1 Introduction

Mitochondrial energy production, which affects every aspect of cellular and individual fitness, requires the interaction of two genomes. The nuclear genome (ncDNA) encodes approximately 1500 proteins that are imported to mitochondria where they constitute the majority of the organelle proteome (Calvo et al., 2006). The mitochondrial genome (mtDNA), on the other hand, generally encodes for 22 transfer RNAs, 2 ribosomal RNAs and 13 protein subunits. The 13 proteins are all involved in the formation of most of the mitochondrial respiratory chain complexes (Blier et al., 2001; Rand et al., 2004). The ~74 other subunits of these complexes are coded by the ncDNA (Rand et al., 2004). Specific interactions between the products of both genomes are required at many levels; including mtDNA replication, repair, transcription, translation and assembly of the respiratory chain complexes (Burton et al., 2013; Ellison and Burton, 2010). The critical importance of maintaining the specificity of these interactions suggests the necessity of a strong coevolution between both genomes (Ballard and Melvin, 2010; Blier et al., 2001; Burton et al., 2013; Rand et al., 2004; Wolff et al., 2014).

Despite this strong constraint, an increasing number of natural organisms characterised by a mtDNA and a ncDNA derived from divergent species have been discovered (Boratynski et al., 2011; Deremiens et al., 2015; Du et al., 2011; Ellison and Burton, 2006; Glemet et al., 1998; Senjo et al., 1999; Toews et al., 2014). These cybrids may suffer from a disruption of the mito-nuclear interaction specificity, resulting in an alteration of the mitochondrial function and individual fitness (Arnqvist et al., 2010; Ellison and Burton, 2006; Kenyon and Moraes, 1997; McKenzie et al., 2003). Nevertheless, various studies demonstrate that between-species transfers of mtDNA can also be neutral (Blier et al., 2006; Boratynski et al., 2011; McKenzie et

al., 2003) or even have beneficial effects (Deremiens et al., 2015; Pichaud et al., 2012; Toews et al., 2014). As an example, the artificial introgression of the fly *Drosophila simulans si*II mitochondria into a *Drosophila simulans si*III nuclear background induces an increase of the respiratory chain catalytic capacity, eventually leading to a better fitness of the individuals (Pichaud et al., 2012). Although these mitochondrial introgressions are known to induce phenotypic changes, many questions still remain about how much allospecific mitochondria can impact the various phenotypic levels (e.g. epigenome, transcriptome, proteome, *etc.*).

The *Chrosomus eos-neogaeus* fish complex (Cyprinidae, Pisces) is a particularly interesting model to address these questions. Among the various biotypes found in this complex, a spontaneous cybrid bearing the ncDNA of *Chrosomus eos* and the mtDNA of *Chrosomus neogeaus* can be distinguished (Dawley et al., 1987; Goddard et al., 1998). These organisms originate from the crossfertilization of *C. eos-neogaeus x eos* triploid hybrid females by *C. eos* males. During the gametogenesis of the triploid individuals, it is assumed that the unmatched set of chromosomes is discarded and meiosis solely occurs on the *C. eos* diploid genome, producing haploid eggs (Goddard et al., 1998). When fertilized by a *C. eos* male, these eggs give birth to individuals with a pure *C. eos* nuclear diploid genome and a female inherited *C. neogaeus* mitochondrial genome (Angers and Schlosser, 2007; Angers et al., 2012; Binet and Angers, 2005; Deremiens et al., 2015). As does *C. eos* wild type, these cybrids reproduce sexually.

Although generally allopatric, *C. eos* wild type and cybrids have been found in sympatry (see Chapter 3; Mee and Taylor, 2012). The third chapter of this thesis revealed the existence of

three different forms of *C. eos* in a same lake of the Laurentian region (Quebec, Canada): the wild type and two different cybrids harboring the mitochondrial genome of *C. neogeaus* from the Mississippian and the Atlantic glacial refuges. Based on *nd3-nd4l* and *cox1* gene sequences, Mississippian and Atlantic *C. neogaeus* mtDNA of cybrids are characterised by a divergence of 1.20%, while a difference of 8.38% and 8.44% is found between wild type and Mississippian cybrids and between wild type and Atlantic cybrids, respectively. Assuming a 1.3% divergence per million years (Bermingham et al., 1997), these results imply that the mitochondria, harbored by the two cybrids, have diverged for approximately one million years while *C. eos* and *C. neogaeus* mtDNA would have diverged for about seven million years. This difference represents a unique opportunity to study the constraints of mito-nuclear genome interactions at different time scales, in naturally occurring organisms.

According to Angers et al. (2012) and Deremiens et al. (2015), allopatric *C. eos* wild type and cybrids present methylome, enzymatic activity, proteome and swimming performance differences. However, because allopatric individuals were used, these studies could not entirely account for possible environmental and ncDNA effects (Deremiens et al., 2015). The discovery of *C. eos* forms that do not show notable nuclear divergences (see Chapter 3), provide an ideal system to address these issues and perform a holistic study of allospecific mitochondria impact upon the various biological organization levels of the phenotype.

This study aims at assessing the influence of allospecific mitochondria on *C. eos* nuclear genome considering various levels of biological integration. To address this, wild type, Mississippian and Atlantic cybrid fish living in sympatry were compared through a holistic

approach. We first analysed the epigenetic and transcriptomic variations associated with the nuclear gene expression. Afterwards, higher organisation levels were assessed through proteomic and enzymatic analyses. Because cell viability and mitochondrial metabolism decrease as the genetic distance between mtDNA and ncDNA increases (Bayona-Bafaluy et al., 2005; McKenzie et al., 2003), we expected a greater divergence between *C. eos* wild type and cybrids than between Mississippian and Atlantic cybrids. Moreover, since no significant difference is measured among the nuclear genomes of the sympatric *C. eos* forms (see Chapter 3), and because environmental effects are minimized compared to our previous study (see Chapter 2), the phenotypic modifications detected among sympatric fish are expected to be less pronounced than the ones measured between allopatric wild type and cybrids.

4.2 Materials and methods

Sampling and genotyping

The sympatric forms of *C. eos* were sampled in Desjardins Lake, southern Quebec, Canada (45°54'58.00"N, 74°04'23.00"W). Regarding the epigenetic and transcriptomic surveys, individuals were also sampled in Beaver Lake (45°55'29.6"N, 74°03'52.5"W) and Saad Lake (45°54'53.6"N, 74°01'40.2"W) respectively. These fish were used as outgroup to compare the magnitude of environmental/nuclear effects to mitochondrial effects. The number and the origin of wild type, Mississippian and Atlantic cybrids used in each analysis are found in Table 4.1.

Sampled individuals were genetically identified according to Binet and Angers (2005). Briefly, total DNA extractions were carried out by proteinase K digestion followed by phenol-chloroform purification and ethanol precipitation (Sambrook et al., 1989). An intron of the

mesoderm specific transcript gene (PEG/MST1) was amplified to differentiate *C. eos* from hybrids. Afterwards, the amplification of a cytochrome c oxidase subunit I gene segment allowed the discrimination of wild type from cybrids. To differentiate Mississippian cybrids from Atlantic cybrids, an additional step was conducted: a segment of the mitochondrial *nd5* gene was amplified (5' to 3' primer sequences: GTGACGGTCCCCCTAATTGG and TGGCATCCTTTGAGAAGAACC) and then digested with the restriction enzyme *NcoI*. The Mississippian haplotype was characterised by a single 559 bp fragment whereas the Atlantic form had two 129 and 428 bp fragments

Table 4.1: Number and origin of the *C. eos* wild type (WT), Missisippian (MC) and Atlantic cybrids (AC) used for the epigenetic, the transcriptomic, the proteomic and the enzymatic activity analyses.

Lake	*C. eos* form	Epigenetic	Transcriptomic	Proteomic	Enzymatic activity
Desjardins	WT	7	2	4	5
	MC	8	2	4	5
	AC	8	2	4	7
Beaver	MC	8			
Saad	WT		2		

Epigenetic

A survey of the DNA methylation variation was performed on female fish sampled on the same day, at the same time. DNA from the caudal fin, was used for the analysis. The epigenetic variation was investigated by assessing the methylation polymorphism at CCGG motifs, using a methylation-sensitive amplified polymorphism analysis (MSAP; Xiong et al., 1999). Fragments displaying a methylation polymorphism with the methylation sensitive treatment (*HpaII*) were considered only when the non-sensitive (*MspI*) treatment exhibited a different

banding pattern across individuals. Selective amplifications were performed using ten primer combinations (HpaII-CAA; HpaII-CAC; HpaII-CAG; HpaII-CAT; HpaII-CCA; HpaII-CCC; HpaII-CCG; HpaII-CCT; HpaII-CGT; HpaII-CTG). Amplification products were separated on a 6% denaturing polycarylamide gel (19:1 acrylamide:bis-acrylamide) and visualized using silver nitrate staining (Bassam et al., 1991).

Transcriptomic

A transcriptomic survey was performed to assess the magnitude of *C. neogaeus* mitochondria impact on the *C. eos* nuclear genome expression. Female fish were euthanized in the field by eugenol overdose, immediately placed on dry ice, brought back to the laboratory and then stored at -80°C. Total RNA was extracted from the white muscle of these fish, as described by Deremiens et al. (2015). To prepare samples for RNA sequencing on the Illumina platform, mRNA was separated from the extracted total RNA and then retro transcribed into cDNA using a Truseq Stranded mRNA sample prep kit. cDNA libraries were constructed with the same kit. Ultimately, samples were barcoded, multiplexed and sequenced (paired-end, 100bp) in a single lane on the Illumina HiSeq2000 platform at IRIC (Université de Montréal, Quebec, Canada).

De novo assembly of the transcriptome was conducted in *Trinity* (version 2.1.1; Grabherr et al., 2011). This was performed with reads previously filtered for poly-A tails using the default parameters of the program Trimmomatic. Following assembly, contigs from the obtained transcriptome (~192,000) were clustered at 90% identity with CD-HIT-EST using the CD-HIT program. Subsequently, filtered reads were aligned on the assembled reference transcriptome using Bowtie2 (Langmead and Salzberg, 2012). Only contigs with one count per million reads

for at least two individuals were considered. Transcripts of the final assembly were identified using a BLASTx search against a NCBI non-redundant (NR) sub-database only composed of Cypriniform fishes (TaxID 7952). Transcripts assigned to the mitochondrial genome were removed from all analyses (E-value < 10^{-20}).

Differences in gene expression for all transcripts were analysed with pairwise comparisons using the edgeR package (Robinson et al., 2010). With this package, *P*-values were adjusted using the Benjamini and Hochberg (1995) method for false discovery rate control. The significance level was set at $P_{adj} < 0.05$. Ultimately, differentially expressed genes were identified using a BLASTx search against the whole NR database or against the NR sub-database described earlier; only matches with an E-value below 10^{-5} were considered. The whole NR database was used for "among groups within site" comparisons whereas the NR sub-database was used for "between sites within group" comparisons.

Proteomic

A quantitative proteomic analysis was achieved to assess if protein expression differences could emerge among the sympatric wild type, Mississippian and Atlantic cybrids. Protein homogenates were prepared as described by Angers et al. (2012) and total protein separation and revelation were performed according to Deremiens et al. (2015). For each 2D-gel, two individuals belonging to the same haplotype were pooled. Gels were analysed using the MELANIE program (version 7.05, Swiss Institute of Bioinformatics). Only spots with good resolution and contrast were selected. To improve the analysis resolution, spots missing in a gel were manually duplicated at the same (x,y) coordinates. Protein expression among the *C. eos*

forms was compared using the percentage volume (%Vol). The %Vol of a spot is a normalized value referring to the ratio between its own volume and the total volume of all the spots in the gel.

Enzymatic activity

According to Deremiens et al. (2015), the cytochrome c oxidase (complex IV) was the only mitochondrial respiratory chain complex displaying a significant difference in enzymatic activity between allopatric wild type and cybrids. To assess whether this difference was maintained among the three sympatric *C. eos* forms, the entire white muscle was sampled from one side of these fish and red muscle, found in very small amount, was removed. Homogenates were prepared as described by Deremiens et al. (2015). Complex IV activity was measured using a UV/VIS spectrophotometer (Shimadzu, UV-1800) equipped with a thermoelectrically temperature-controlled cell holder (CPS-240A). All assays were performed in duplicate at 24°C. Conditions for the cytochrome c oxidase enzymatic assay were as described by Deremiens et al. (2015). Total soluble protein content of each homogenate was determined by the Bradford method (Bradford, 1976) using bovine serum albumin (BSA) as standard.

Statistical analyses

Redundancy analyses (RDA) were performed to compare the epigenetic, transcriptomic and proteomic profiles among the wild type, Mississippian and Atlantic cybrids from Desjardins Lake. For the first two surveys, Mississippian cybrids from Beaver Lake and wild types from Saad Lake were respectively added to the analyses. The epigenetic, transcriptomic and proteomic profiles assessments were carried out using the presence-absence matrix of MSAP

bands, a matrix of the read count per million of each transcript and a matrix of the %Vol of each protein, respectively. Each of these matrices were combined to a matrix grouping of individuals according to their mitochondrial haplotype or the lake of origin. The significance of the results was tested with 999 permutations. Regarding enzymatic analyses, because of the very small sample data sets, non-parametric Mann-Whitney *U*-tests were applied. All statistical analyses were achieved with the R program (version 3.2.1), using the vegan package (Oksanen et al., 2007). The significance level was set at $P < 0.05$.

4.3 Results

Epigenetic

Out of a total of 367 loci screened for epigenetic changes, 128 (34.9%) displayed variation. Among the variable loci, 32 (25%) methylated bands differentiated individuals from Desjardins Lake and Beaver Lake; more precisely 23 of these bands were found in every individual from Desjardins Lake and 9 of them characterised every fish from Beaver Lake.

The first RDA, which considered the Mississippian cybrids from Desjardins Lake and Beaver Lake, revealed a difference between sites. This "between sites within group" difference was significant (ANOVA, $P = 0.002$). In this case, the first axis of the RDA explained 50.4% of the variation. The second RDA, which considered the three *C. eos* forms from Desjardins Lake, demonstrated a difference among groups. This “among groups within site” difference was significant between the wild type and both cybrids (ANOVA, $P < 0.015$) but not between Mississippian and Atlantic cybrids (ANOVA, $P = 0.566$). Here, 12.7% of the variance of the epigenetic profiles was explained by the first two RDA axes. The third RDA (Figure 4.1),

combining the three *C. eos* forms from Desjardins Lake with Mississippian cybrids from Beaver Lake (ANOVA , $P = 0.003$), showed that the difference observed "between sites within group" was much more important than the one measured "among groups within site". The first two axes of this RDA explained 19.1% of the variance.

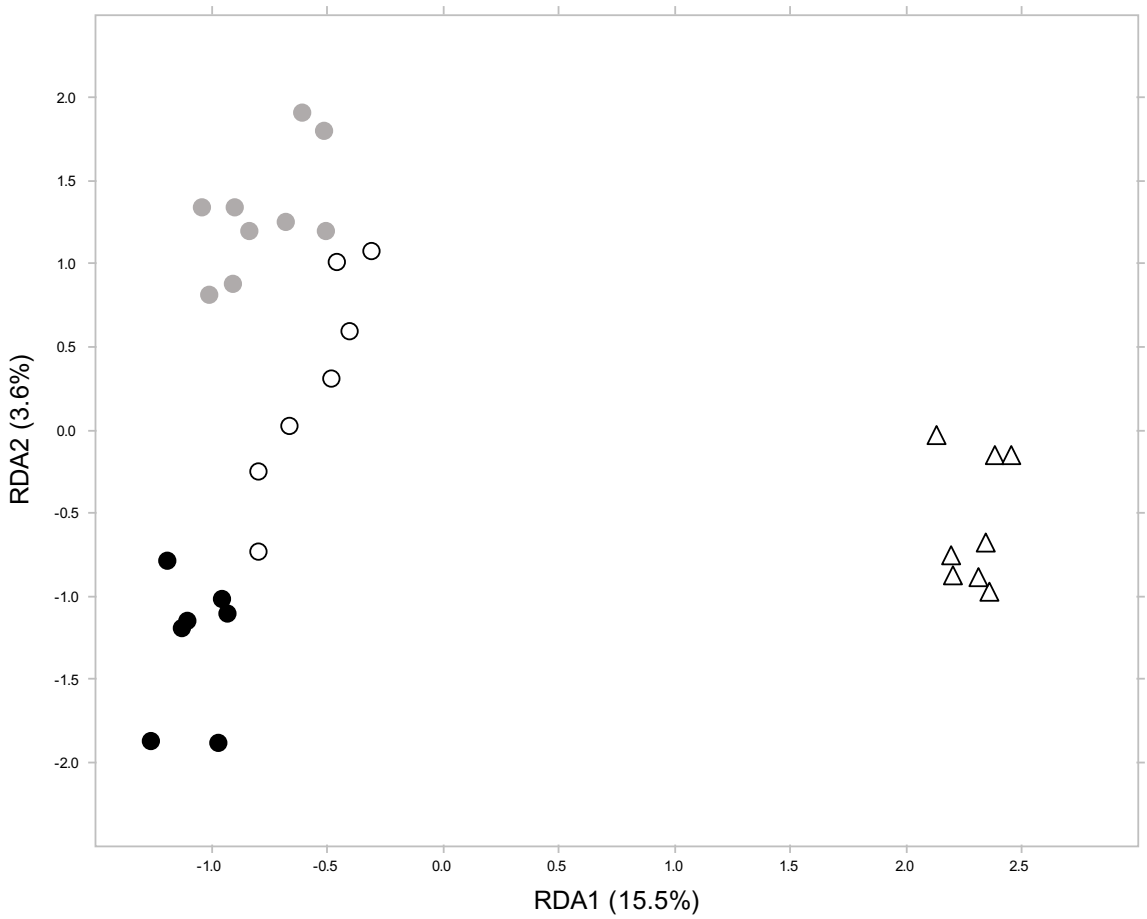

Figure 4.1: RDA of the different epigenetic profiles obtained from the caudal fin of the *C. eos* wild type, Mississippian and Atlantic cybrids. Circles and triangles refer to individuals from Desjardins Lake and Beaver Lake respectively. Black circles correspond to the wild type, grey circles to Atlantic cybrids and blank circles and triangles to Mississippian cybrids.

Transcriptomic

Our transcriptomic inquiries initially yielded 216 million paired-end fragments (2 × 100 bp), averaging 54 million reads of high quality per sample: phred33+ > 30 for 92%, with the exception of one Atlantic cybrid individual at 90.46%. In our final assembly of 30,671 transcripts, 21,140 (68.9 %) found an homology. The quality of that assembly proved to be high; with an N50 of 2948 bases.

Once the transcriptomes were assembled, transcriptomic profiles were compared. The first RDA, which considered the wild types from Desjardins Lake and Saad Lake, demonstrated a difference between sites but which was not significant (ANOVA, $P = 0.333$). The first axis of this RDA explained 14.2% of the variation. The second RDA, which considered the sympatric wild type, Mississippian and Atlantic cybrids from Desjardins Lake, revealed a trend separating these forms but which was not significant (ANOVA, $P > 0.3$). In this case, 54.4% of the transcriptomic profiles was explained by the first two RDA axes. The lack of significance observed for these two RDA probably originates from the small number of individuals analysed for each *C. eos* forms. The last RDA (Figure 4.2), which considered the three *C. eos* forms from Desjardins Lake along with the wild type from Saad Lake (ANOVA, $P = 0.009$), showed that the difference observed "between sites within group" was more pronounced than the one measured "among groups within site". The first two axes of the performed RDA explained 82.4% of the variance.

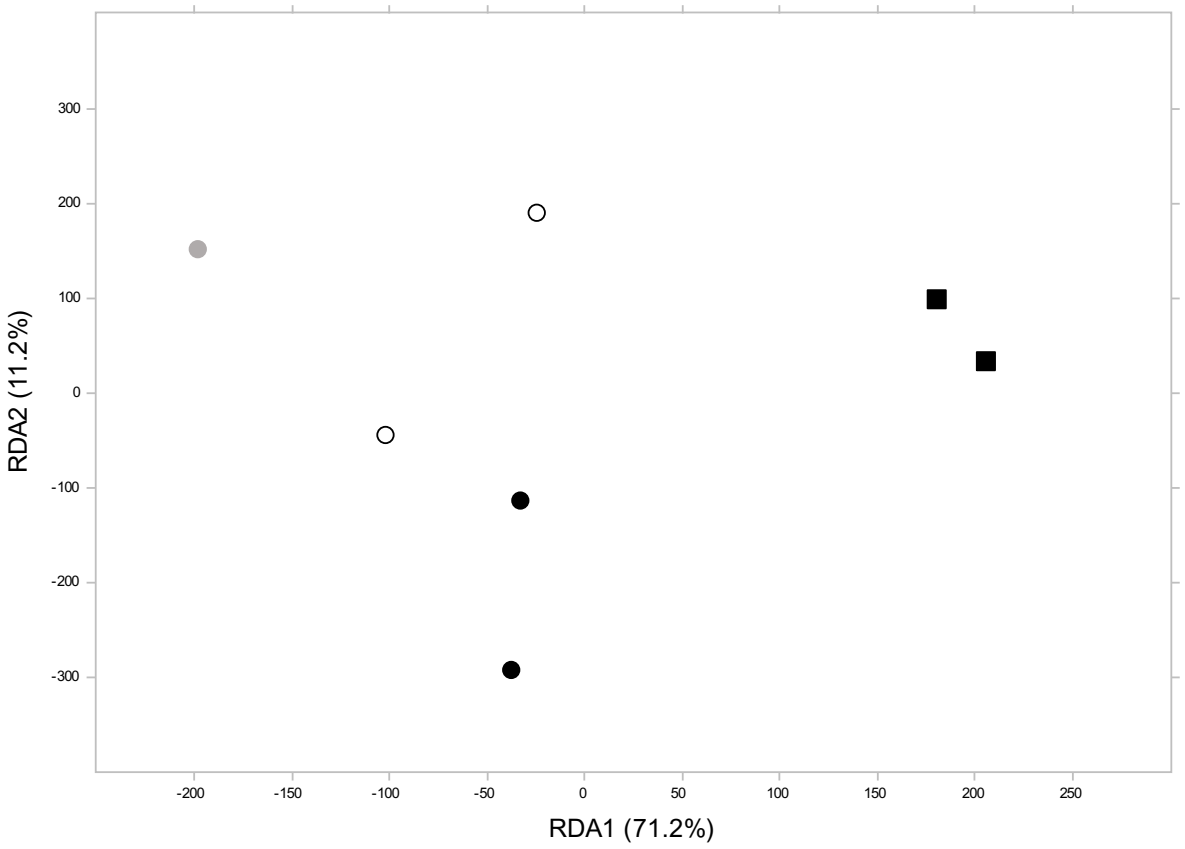

Figure 4.2: RDA based on the read count per million of each transcript detected in the *C. eos* wild type, Mississippian and Atlantic cybrids. Circles and squares refer to individuals from Desjardins Lake and Saad Lake respectively. Black circles and squares correspond to the wild type, the grey circle to an Atlantic cybrid and blank circles to Mississippian cybrids.

Transcript expression variations were compared "between sites within group", but also "among groups within site". Out of the 30,671 transcripts, 2,392 (7.80%) were considered as differentially expressed between the wild types from Desjardins Lake and Saad Lake (Figure 4.3A). On the other hand, when the wild type and Mississippian cybrids from Desjardins Lake were compared, only 104/30,671 (0.34%) transcripts were characterised as differentially expressed (Figure 4.3B). Based on the whole NR database, 72 of these mRNAs were

homologous to 63 proteins. Among those 63 proteins, 31 were characterised (Table A5) while 32 had unknown functions; the latter were not considered in this study. Among the caracterised ones, three were ncDNA-encoded proteins associated to mitochondria. When "between sites within group" and "among groups within site" results were compared, 53/104 transcripts were common to both analyses. These mRNAs corresponded to 34 of the 63 homologies previously detected, among which 17 were characterised. Accordingly, the 51/104 differentially expressed transcripts left could be attributed to foreign mitochondria effects alone. These mRNAs corresponded to 29 of the 63 detected homologies, among which 14 were characterised (Table A5).

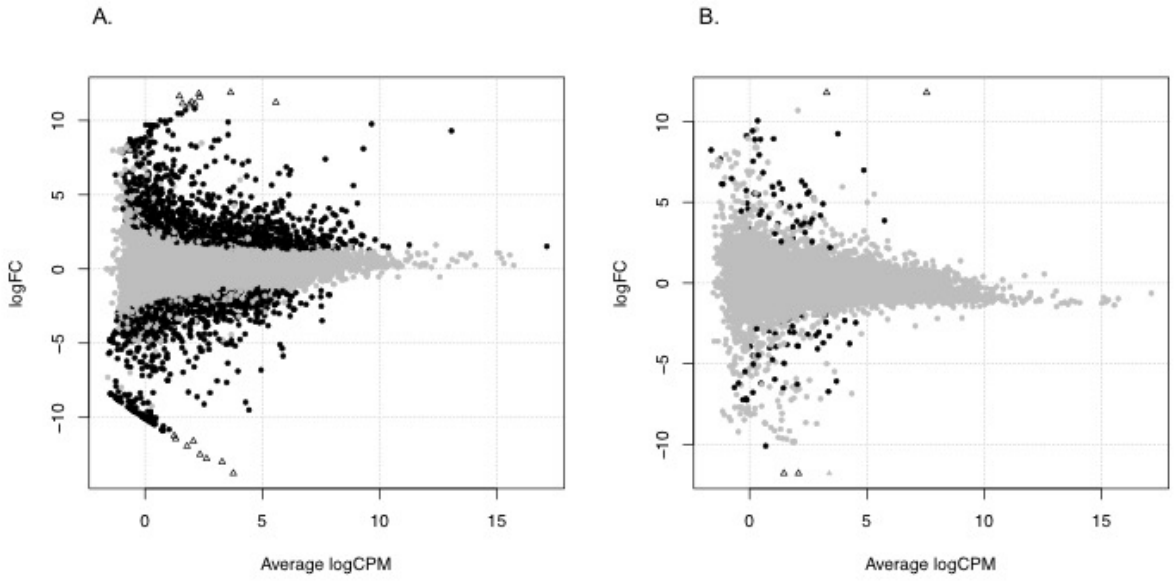

Figure 4.3: Smear plot visualization of differential expression analyses of *C. eos* wild type and Mississippian cybrid transcriptomes. (A) Transcriptome comparisons between wild types from different lakes (Desjardins and Saad). (B) Transcriptome comparisons between wild type and Mississippian cybrids from the same lake (Desjardins). Black dots represent the mRNAs considered as differentially expressed ($P_{adj} < 0.05$). Triangles refer to transcripts that could not be represented in the logFC interval. CPM = counts per million; FC = fold change.

Proteomic

To infer mitochondria influence on the protein expression of the three *C. eos* forms, the %Vol of a subsample of 132 unambiguous protein spots was analysed. Although the RDA revealed a difference among the wild type, Mississippian and Atlantic cybrids from Desjardins Lake (variance explained by the first two axes = 66.2%), this difference was not statistically significant ($P = 0.067$)

Despite this result, inter-form comparisons indicated protein expression variations (Figure 4.4). Considering an expression ratio difference greater than or equal to 2×, the proportion of differentially expressed proteins within a given *C. eos* form was low; varying from 3.8 to 12%. When inter-form comparisons were tested, a higher percentage of differentially expressed proteins was detected: (*i*) wild type *versus* Atlantic cybrids, 15.1 to 23.5%; (*ii*) wild type *versus* Mississippian cybrids, 27.3 to 31.8%; (*iii*) Atlantic cybrids *versus* Mississippian cybrids, 13.6 to 20.4%.

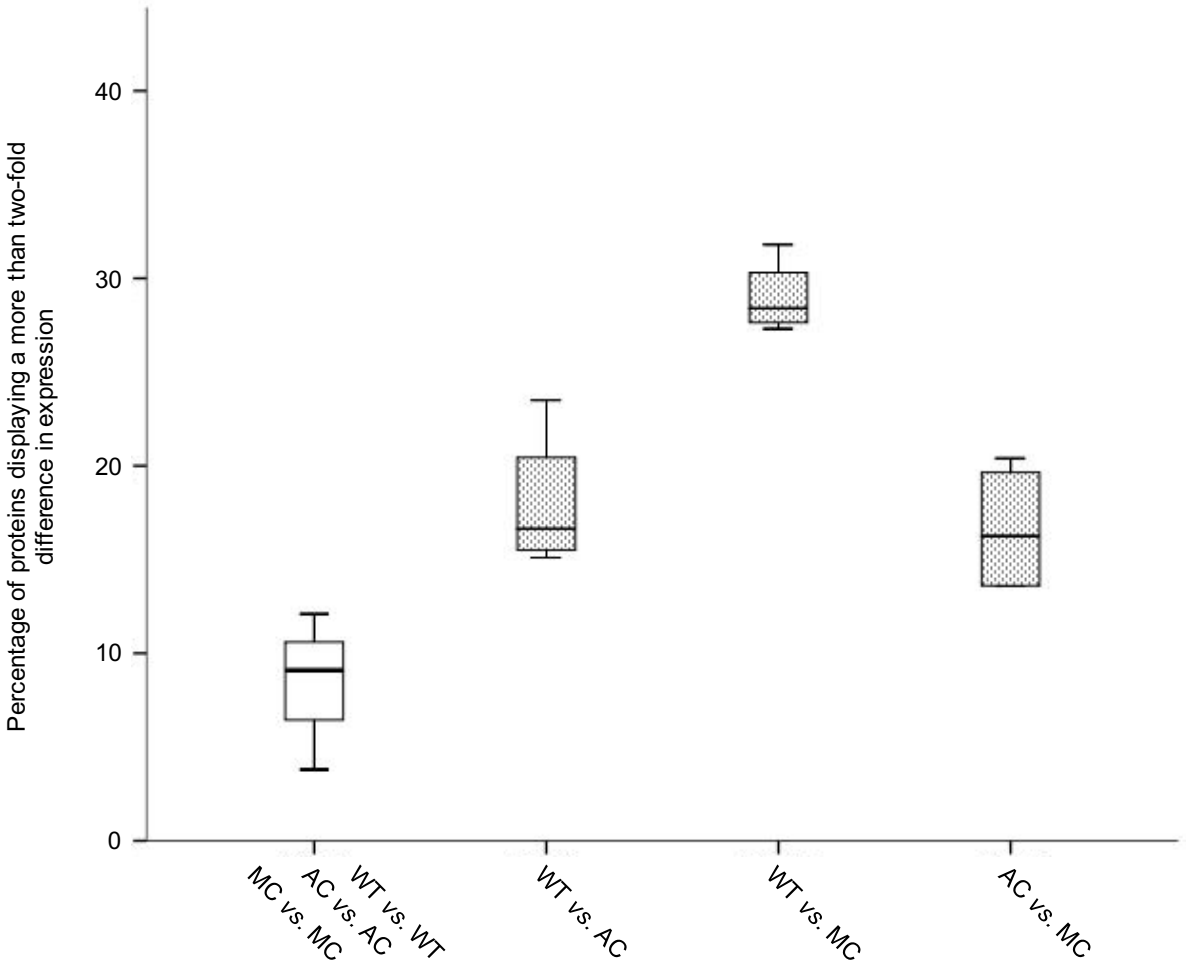

Figure 4.4: Protein expression comparisons in *C. eos* wild type and cybrids. Among the 132 studied proteins, only the ones presenting a more than two-fold difference in expression were considered. The comparisons were performed with two protein gels for each *C. eos* form. The blank boxplot represents the intra-form comparisons while the dotted boxplots represent the inter-form comparisons. Intra-from comparisons were pooled. WT = wild type; AC = Atlantic cybrid; MC = Mississippian cybrid.

Enzymatic activity

A significant difference of complex IV activity was observed in both Mississippian and Atlantic cybrids when compared to their sympatric wild type counterparts (Figure 4.5; Mann-Whitney, $P < 0.02$). These cybrids exhibited a nearly, twice higher activity than the wild type. No significant difference was detected between the two cybrid forms (Mann-Whitney, $P = 0.861$).

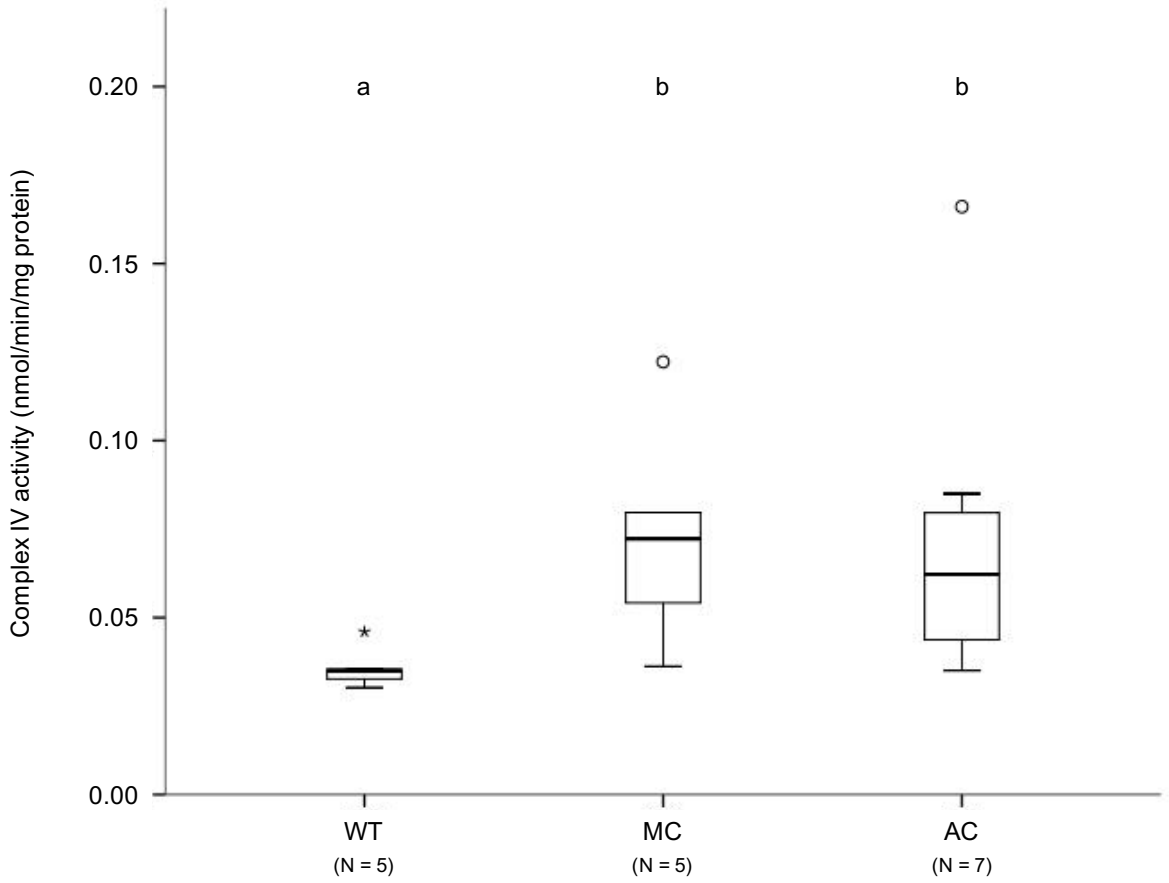

Figure 4.5: Complex IV activity in the white muscle of *C. eos* wild type (WT), Mississippian (MC) and Atlantic (AC) cybrids. Significance was set as $P < 0.05$. Letters denote differences between *C. eos* forms: a is significantly different from b.

4.4 Discussion

As previously described, Mississippian and Atlantic cybrids are characterised by a 1.20% divergence of their mtDNA while a difference of 8.38% and 8.44% is detected between the wild type and Mississippian cybrids and between the wild type and Atlantic cybrids, respectively (see Chapter 3). While a divergence of 1.20% would not trigger a wide phenotypic modulation, the 8.38-8.44% difference observed between the wild type and the two cybrids could potentially do so. To test this hypothesis, a holistic approach, going from nuclear genome expression to proteome and enzymatic comparisons, was considered.

In the present study, the comparison of the methylation profiles among the wild type, Mississippian and Atlantic cybrids revealed that *C. neogaeus* mitochondria are able to trigger a reprogramming of the *C. eos* nuclear genome expression. However, this comparison showed that the environment has a stronger effect on DNA methylation than allospecific mitochondria. Presently, the link between mitochondria and epigenetic changes is not yet fully understood. One of the possible processes could imply the production of reactive oxygen species (ROS; Afanas'ev, 2015; Wu and Ni, 2015), a by-product of the mitochondrial respiration (Bleier and Drose, 2013; Murphy, 2009). More precisely, mitochondrial respiration involves the transfer of electrons through a succession of redox centers, from an electron donor (NADH) to a terminal acceptor (O_2). Following the integration of foreign mitochondria, a reduction of this electron flux is generally observed due to mismatches between ncDNA- and mtDNA-encoded subunits of the respiratory chain complexes. Ultimately, this phenomenon leads to a more highly reduced state of these enzymes, which finally induces a higher production of ROS (Barja, 2007; Gusdon et al., 2007; Moreno-Loshuertos et al., 2006). These molecules would impact epigenetic profiles

through both DNA methylation modifications and changes of histone acetylation/deacetylation state (Afanas'ev, 2015; Wu and Ni, 2015). Contrarily to the expected effects of foreign mitochondria integration on respiratory chain complex activity, this study revealed a significant increase of complex IV activity in cybrids. This higher activity would allow complexes I, II and III to be in a more highly oxidized state (Blier and Lemieux, 2001), thereby inducing a lower production of ROS compared to *C. eos* wild type, which would eventually modify cybrid epigenetic profiles.

Although transcriptomic profiles of the sympatric wild type and Mississippian cybrids were not significantly different, our results demonstrated a few changes between them. These limited changes might be related to the detected epigenetic variations and/or could derive from post-transcriptional modifications. As revealed by the epigenetic analysis, environmental effects were more pronounced than those of allospecific mitochondria. Although most of the 104 differentially expressed mRNAs detected between the wild type and Mississippian cybrids were mainly unidentified or associated to uncharacterised functions, three of them were homologous to ncDNA-encoded mitochondrial proteins. Among those three, the "1,25-dihydroxyvitamin D(3) 24-hydroxylase" was found in both "between sites within group" and "between groups within site" comparisons. On the other hand, the "glutaminase kidney isoform" and the "pyruvate dehydrogenase kinase isozyme 3" were only found in the "between groups within site" comparison. These two transcript expression modifications can therefore be attributed to foreign mitochondria effects alone. While the "glutaminase kidney isoform" is an enzyme associated to the maintenance of acid-base homeostasis, the "pyruvate dehydrogenase kinase isozyme 3" regulates the pyruvate dehydrogenase complex activity, an enzyme directly involved

in glucose metabolism and aerobic respiration. According to their function, discrete metabolic differences could thus happen between the wild type and Mississippian cybrids. These mRNAs, therefore, reveal that phenotypic modifications might occur at a biological organization level not yet investigated, the metabolome.

As shown in the transcriptomic survey, proteomic profiles of the sympatric *C. eos* forms were not significantly differentiated but some protein expression variations were observed among them. These variations might result from the detected transcriptomic changes but could also come from post-translational modifications. Although apparently high, the magnitude of differentially expressed proteins was actually less pronounced when intra-form comparisons were considered. One particularly interesting feature of this analysis come from a similar number of differentially expressed proteins between Mississippian and Atlantic cybrids (i.e. 13.6-20.4%) and between the wild type and Atlantic cybrids (i.e. 15.1-23.5%). This similarity suggests that the 1.20% mtDNA divergence measured between the two cybrids can impact the proteomic phenotype as much as the 8.44% mtDNA difference detected between the wild type and Atlantic cybrids. In the present study, 15.1 to 31.8% of the assessed proteins showed a more than two-fold difference in expression between sympatric wild type and cybrids. On the other hand, according to Deremiens et al. (2015), allopatric wild type and cybrids were characterised by a proteomic difference ranging from 29.0 to 37.0%. These higher values would mainly be caused by environmental and possibly nuclear effects resulting from the use of allopatric biotypes. Ultimately, it is important to mention that part of the inter-form variations could come from proteins encoded by the mtDNA itself. Therefore, the magnitude of allospecific

mitochondria influence on protein expression could be less pronounced than what is actually measured.

As mentioned earlier, our comparative analysis demonstrated that *C. neogaeus* mitochondria are able to increase the complex IV activity of *C. eos*. According to Deremiens et al. (2015), similar results were obtained with allopatric *C. eos* wild type and cybrids, and could result from one amino acid change on the *COX3* mitochondrial subunit. This enzyme being of importance in the respiratory chain regulation (Arnold, 2012; Blier and Lemieux, 2001), its activity change might impact the aerobic metabolism. Moreover, since the sympatric *C. eos* forms do not present notable nuclear differentiation (see Chapter 3) and because no transcript involved in the mitochondrial respiration *per se* was differentially expressed among the wild type, Mississippian and Atlantic cybrids, the modification of complex IV activity can be assumed to only come from the allospecific mitochondria. Altogether, these observations suggest that mitochondria can occasionally confer specific functional differences, which might ultimately impact mitochondrial metabolism.

According to Mee and Taylor (2012), the wild type form is mainly found in the southern part of the *C. eos* distribution area whereas cybrids, as well as *C. neogaeus*, predominantely occur in the north. This distribution suggests that cybrids can support colder conditions than wild type. In the Laurentian region, where their distributions overlaps, wild type and cybrids are usually found in allopatry and sometimes in sympatry (see Chapter 3; Mee and Taylor, 2012). To explain this particularity, one assumption is that under certain environmental conditions, the phenotypic differences observed between wild type and cybrids are such that

they promote the establishment of one of the two biotypes. However, under less common conditions, these phenotypic variations would be less pronounced and too low to allow the prevalence of any biotypes. On a larger geographical scale, in Laurentian lakes, the influence of *C. neogaeus* mitochondria on *C. eos* ncDNA expression would not be enough to prevent wild type and cybrids to live together or next to each other. On the other hand, further north, in cold environments, one hypothesis is that the nuclear reprogramming induced by these foreign mitochondria is more pronounced; inducing the expression of genes usually repressed in wild type, eventually leading to cybrids better adapted to cold environments and therefore explaining the peculiar distribution of this biotype. If the higher complex IV activity in cybrids is extrapolated to the metabolic activity, we can also assume that these fish would be metabolically more active than wild types at low temperature; advantage promoting their dominance in cold environments. Both hypotheses are mutually non-exclusive.

To conclude, the holistic approach applied in this study enabled to demonstrate that allospecific mitochondria are able to impact *C. eos* phenotype, at various biological levels, when wild type and cybrids are found in sympatry. As expected, the phenotypic differences detected between these biotypes were more important than the ones measured between Mississippian and Atlantic cybrids. Additionally, this survey revealed that the magnitude of foreign mitochondria influence was more pronounced between allopatric wild type and cybrids than between sympatric wild type and cybrids. This discrepancy might be the result of environmental and nuclear effects resulting from the use of allopatric fishes. Ultimately, all the results obtained through this holistic research significantly contribute to the disentangling of the links existing between mitochondria and the different organisation levels of the phenotype.

Furthermore, this approach participates of the age-old question of how genotypic variations affect the phenotype of an organism.

4.5 Acknowledgments

The authors are grateful to L. Baillon, K. Duclos, C. Leung and E. L'Heureux for their precious assistance. LD was partly supported by a PhD scholarship from the GRIL. This research was supported by a generous contribution from S. Lumbroso and by a team grant from the *Fonds Québécois pour la Recherche sur la Nature et les Technologies* (FQRNT) to BA and AA.

CHAPITRE 5 : DISCUSSION GÉNÉRALE

L'objectif principal de cette thèse de doctorat était d'évaluer l'influence de mitochondries étrangères à différents niveaux d'organisation du phénotype, lorsque les poissons *C. eos* sauvages et cybrides sont retrouvés en allopatrie ou en sympatrie. Les sujets traités de même que les résultats obtenus nous ont permis d'approfondir les liens existant entre l'ADNmt et l'ADNnc, ainsi que de mieux comprendre les conséquences de changements génétiques sur les variations phénotypiques. À l'heure actuelle, il apparait clair que les études cherchant à estimer les impacts du génotype sur le phénotype nécessitent la considération des différents niveaux d'organisation caractérisant ce dernier ; depuis l'épigénome jusqu'à l'individu.

5.1 *C. eos*, un modèle biologique irrésistible

Lors de l'étude des impacts engendrés par la présence de mitochondries étrangères sur le phénotype, des cybrides artificiels, créés *in vitro* ou par rétrocroisement, sont généralement utilisés (Ellison et Burton, 2006 ; Kenyon et Moraes, 1997 ; McKenzie and Trounce, 2000 ; McKenzie et al., 2003 ; Pichaud et al., 2012). Contrairement à la majorité de ces recherches, notre étude propose l'utilisation d'un organisme cybride apparu spontanément en milieu naturel.

Bien qu'on en découvre de plus en plus d'exemples, le nombre d'organismes cybrides recensés en milieu naturel reste limité (Boratynski et al., 2011 ; Deremiens et al., 2015 ; Du et al., 2011; Ellison et Burton, 2006 ; Glemet et al., 1998 ; Senjo et al., 1999 ; Toews et Brelsford, 2012 ; Toews et al., 2014). Parmi ces organismes, seuls certains d'entre eux ont été étudiés afin d'évaluer les effets de mitochondries étrangères sur le phénotype (Boratynski et al., 2011 ; Deremiens et al., 2015 ; Ellison et Burton, 2006 ; Glemet et al., 1998 ; Toews et al., 2014). Le plus souvent non-intégratives, ces études ne considèrent généralement pas les effets

environnementaux ou nucléaires et sont basées sur l'utilisation de cybrides résultant de phénomènes d'hybridations suivis de rétrocroisements successifs (Boratynski et al., 2011 ; Ellison et Burton, 2006 ; Glemet et al., 1998 ; Toews et al., 2014). Chez de tels organismes, l'introgression mitochondriale est régulièrement accompagnée d'introgressions au niveau de l'ADNnc.

Dans le cadre de cette thèse, les poissons *C. eos* sauvages et cybrides sont utilisés. Résultant d'un phénomène d'hybridation suivi d'une augmentation de ploïdie, ces organismes sont produits *de novo*, sans introgression nucléaire. Seules les formes sauvages et cybrides en allopatrie ont été considérées dans le premier volet de notre étude. Afin de tenir compte des effets environnementaux, chacun de ces biotypes a été échantillonné dans plusieurs lacs de la même région. Les effets nucléaires, quant à eux, seraient limités en conséquence d'une divergence supposée relativement récente entre les deux formes (Angers et al., 2012). Dans les volets deux et trois, la sympatrie des formes sauvages et cybrides permet l'élimination des effets environnementaux. De plus, le deuxième volet de la présente étude révèle, par des analyses de différenciation génétique, d'organisation génétique et de taux de migration, une absence de divergence nucléaire entre les biotypes, supprimant ainsi les effets potentiels de ce génome. Lorsque réunies, l'ensemble de ces caractéristiques rendent compte de l'opportunité exceptionnelle qu'offre le ventre rouge du Nord *C. eos* pour étudier l'influence de mitochondries étrangères sur les différents niveaux d'organisation du phénotype. À l'heure actuelle, l'approche holistique considérée dans cette thèse représente une des études intégratives les plus complètes en regard de l'influence du génotype sur le phénotype.

5.2 *C. eos* sauvages et cybrides en allopatrie

Le premier volet de cette thèse a permis de mettre en évidence les répercussions de mitochondries étrangères à différents niveaux d'organisation biologique du phénotype, lorsque les formes sauvages et cybrides de *C. eos* sont retrouvées en allopatrie. Dans cette étude, le premier niveau d'organisation du phénotype analysé était l'activité d'enzymes clés du métabolisme aérobie et anaérobie. Ainsi, une augmentation de l'activité du complexe IV de la chaîne respiratoire mitochondriale a été mise en évidence chez les individus cybrides, à 24°C. Bien que l'activité de cet enzyme n'était pas significativement différente entre les deux biotypes à 17 et 10°C, l'activité moyenne de ce complexe était tout de même supérieure chez la forme introgressée. Cette augmentation pourrait avoir été causée par la modification de l'acide aminé en position 41 de la sous-unité mitochondriale *COX3* et/ou par d'importantes modifications protéomiques. Selon l'étude de ce deuxième niveau d'organisation, 29.01 à 37.04% des protéines analysées démontraient une variation d'expression supérieure ou égale à un facteur deux, entre les deux biotypes. Le troisième et dernier niveau d'organisation du phénotype considéré dans ce volet était la vitesse de nage critique (U_{crit}) des poissons. Selon nos résultats, les cybrides présentaient une performance de nage supérieure à celle de la forme sauvage à 17°C.

Contrairement à ce qui est généralement attendu, l'ensemble des résultats obtenus démontrent qu'une absence de coévolution entre l'ADNnc et l'ADNmt n'induit pas nécessairement une détérioration du phénotype, mais qu'au contraire, ce phénomène pourrait effectivement avoir des répercussions bénéfiques. Afin de mieux comprendre ce phénomène, rappelons que l'ADNmt est communément considéré comme évoluant plus rapidement que l'ADNnc (Lynch, 1996 ; Lynch et al., 2006 ; Scheffler, 2008 ; Wallace et al., 1987). Selon cette

particularité, afin d'éviter une détérioration du phénotype, les changements apparaissant au niveau de ce génome sont généralement contrebalancés au niveau nucléaire (Osada et Akashi, 2012). Bien qu'efficace dans certains cas, ce processus pourrait limiter l'éventail de combinaisons possibles entre les deux génomes, restreignant ainsi l'apparition de nouvelles interactions potentiellement bénéfiques. Dès lors, l'originalité de notre étude est qu'elle suggère la possibilité que la capture de mitochondries étrangères permettrait de contourner les processus de sélection s'appliquant à la coévolution mito-nucléaire, induisant ainsi l'apparition de nouvelles combinaisons génomiques, potentiellement avantageuses, chez les cybrides *C. eos*.

5.3 *C. eos* sauvages et cybrides en sympatrie

Ménage à trois et échanges génétiques

Les formes sauvages et cybrides du poisson *C. eos* sont généralement retrouvées en allopatrie dans la région des Laurentides (Québec, Canada) (Angers and Schlosser, 2007 ; Angers et al., 2012). De nombreuses variations phénotypiques ont été observées entre ces biotypes ; que ce soit dans le premier volet de cette thèse (protéome, activité enzymatique et performance de nage) ou dans l'étude de Angers et al. (2012) (méthylome et protéome). À la suite de récents échantillonnages, il s'avère que ces deux biotypes peuvent occasionnellement se retrouver en sympatrie. Cette répartition offre l'opportunité unique de déterminer si l'étendue des modifications précédemment détectées en allopatrie peut aller jusqu'à déclencher un processus de différenciation génétique entre les formes sauvages et cybrides lorsqu'en sympatrie.

De façon surprenante, le deuxième volet de cette thèse a révélé l'existence en sympatrie de trois formes différentes de *C. eos* : la forme sauvage et deux formes cybrides arborant

l'ADNmt de *C. neogaeus* issus des refuges glaciaires Mississippien et Atlantique. Sur base des séquences obtenues à partir des gènes *nd3-nd4l* et *cox1*, les ADNmt des cybrides Mississippiens et Atlantiques divergeraient depuis environ un million d'année alors que ceux de *C. eos* et *C. neogaeus* évolueraient indépendamment depuis approximativement sept millions d'années.

Le second volet de cette thèse avait pour objectif principal de caractériser l'étendue du flux génétique existant parmi les formes sauvages, les cybrides Mississippiens et les cybrides Atlantiques. Dans le lac Desjardins, chacune de ces formes a été retrouvée dans chacun des quatre sites échantillonnés. Quel que soit le site considéré, les cybrides Mississipiens étaient invariablement prédominants. Au sein d'un seul et même site, l'ampleur du flux de gènes parmi les différentes formes de *C. eos* était telle qu'aucune différenciation génétique n'a pu être mise en évidence. En revanche, lorsque l'ensemble des sites étaient combinés, les cybrides Mississippiens présentaient une diversification génétique (A_R et H_E) légèrement plus élevée que celle des autres formes ; notons ici que les valeurs obtenues pour A_R et H_E sont corrigées pour la taille de l'échantillon (Goudet, 1995 ; Leberg, 2002). Ainsi, ces derniers résultats suggèrent que bien que génétiquement non différenciées, les formes sauvages et cybrides ne se reproduisent pas de façon panmictique.

En définitive, notre analyse démontre que l'ensemble des modifications phénotypiques précédemment observées entre les formes sauvages et cybrides ne sont pas suffisantes pour induire un phénomène de spéciation entre ces dernières. De plus, sachant qu'un important flux génétique entrave la mise en place de nouveaux allèles localement adaptatifs, nos résultats suggèrent que l'ADNnc *C. eos* et l'ADNmt *C. neogaeus* n'ont jamais eu la chance de s'adapter

l'un à l'autre lorsque la forme sauvage et les cybrides sont retrouvés en sympatrie. Selon nos résultats, ces ADNmt évolueraient indépendamment depuis approximativement sept millions d'années. Pour un niveau de divergence similaire, McKenzie et al. (2003) observaient une légère altération du métabolisme de souris cybrides. Contrairement à ces observations, les résultats obtenus dans le premier volet de ce projet démontraient une influence positive des mitochondries étrangères sur le phénotype (i.e. augmentation de l'activité du complexe IV et de la performance de nage). Ces effets bénéfiques pourraient provenir de l'exploration de nouvelles combinaisons génomiques qui, selon les résultats de ce volet, ne nécessiteraient que peu ou pas de compensation nucléaire depuis l'introgression des mitochondries *C. neogaeus*. Remarquons que dans ce cas de figure, les formes sauvages et cybrides sont considérées comme étant en sympatrie depuis la formation de ces derniers. En fin de compte, bien que contre-intuitifs, les premiers et seconds volets de cette thèse suggèrent que la coévolution intergénomique peut ne pas être une condition *sine qua non* à la survie, la reproduction et la perpétuation des individus en milieu naturel.

Ménage à trois et modifications phénotypiques

Lorsque combinés, l'étude menée par Angers et al. (2012) et les résultats obtenus dans le premier volet de cette thèse révèlent une modification du méthylome, du protéome, de l'activité d'enzymes associés au métabolisme et de la performance de nage entre les *C. eos* sauvages et les cybrides allopatriques. Ces études ne peuvent totalement écarter l'influence des effets environnementaux et nucléaires sur les mesures effectuées, même si elles en tiennent compte. Dans le second volet de cette recherche, trois formes de *C. eos* caractérisées par des ADNnc semblables ont été découvertes en sympatrie. Ces différentes formes procurent un système idéal

permettant de résorber les points faibles rencontrés lors de nos précédentes analyses. De plus, le temps de divergence estimé entre la forme sauvage et les cybrides (~7 millions d'années), et entre les cybrides Mississippiens et Atlantiques (~1 million d'années) représentent l'unique opportunité de mieux comprendre les contraintes associées aux interactions mito-nucléaires à différentes échelles temporelles, à partir d'organismes issus d'un environnement naturel.

L'objectif du troisième volet était d'évaluer l'influence de mitochondries étrangères à différents niveaux d'organisation du phénotype, à partir d'organismes sauvages et cybrides retrouvés en sympatrie. Cette approche holistique consiste en une analyse comparative de l'expression de l'ADNnc (i.e. méthylome et transcriptome), suivie d'une analyse de niveaux d'organisation biologique de plus hautes complexités (i.e. protéome et activité enzymatique). Selon nos résultats, les formes sauvages et cybrides de *C. eos* présentent des profils de méthylation distincts ainsi qu'un certain nombre de transcrits et protéines différentiellement exprimés. Parmi les ARNm identifiés, trois sont associés à la mitochondrie ; la modification d'expression de deux d'entre eux peut directement être attribuée à la présence des mitochondries *C. neogaeus* et a le potentiel d'induire des changements métaboliques. Que ce soit pour les analyses épigénétiques ou transcriptomiques, l'influence de l'environnement est plus importante que celle des mitochondries étrangères. Finalement, tel que détecté dans le premier volet, une augmentation significative de l'activité du complexe IV est mesurée chez les individus cybrides. D'un autre côté, lorsque les cybrides Mississippiens et Atlantiques sont comparés, les différences observées sont faibles au niveau épigénétique, indisponibles au niveau transcriptomique et absentes au niveau enzymatique. Au niveau protéomique, le niveau de différences détectées entre les deux formes est légèrement inférieur à celui mesuré entre les

formes sauvages et cybrides. Ainsi, tel qu'attendu sur base du degré de divergence existant entre les ADNmt des trois formes de *C. eos*, les variations phénotypiques mesurées entre les deux cybrides sont généralement moins marquées que les différences mesurées entre la forme sauvage et les cybrides.

En définitive, l'approche holistique employée dans le troisième volet de cette recherche démontre l'existence d'un impact des mitochondries étrangères sur le méthylome, le transcriptome, le protéome et l'activité du complexe IV de la chaîne respiratoire mitochondriale, lorsque les formes sauvages et cybrides sont retrouvées en sympatrie. Cette influence, bien que présente, semble généralement moins prononcée que celle observée à partir d'organismes allopatriques. La plus forte variation détectée chez ces derniers proviendrait vraisemblablement d'effets environnementaux et possiblement d'effets nucléaires. En effet, bien que la méthode d'échantillonnage appliquée dans le premier volet de cette recherche soit capable de tenir compte de l'effet des lacs pour chacun des biotypes, cette dernière ne permet pas de tester les effets de l'environnement entre ceux-ci. Les effets nucléaires, quant à eux, seraient réduits suite à une divergence supposée relativement récente entre les deux biotypes (Angers et al., 2012). Finalement, à la suite d'analyses de différentiation génétique, d'organisation génétique et de taux de migration, les formes sauvages et cybrides de *C. eos* apparaissent indifférenciées au niveau nucléaire. Néanmoins, une reproduction non panmictique est détectée entre les deux biotypes. Ainsi, si l'ensemble des variations phénotypiques décelées au cours de ce projet de recherche semblent insuffisantes pour induire un phénomène de spéciation, nous pouvons imaginer que ces dernières pourraient entrainer un décalage partiel de la période de reproduction des formes sauvages et cybrides. Ce faisant, bien qu'un échange de gènes serait maintenu entre les deux

biotypes, ces derniers se reproduiraient préférentiellement avec des individus appartenant à leur propre type ; constituant ainsi une population se reproduisant de façon non panmictique.

5.4 Répartition géographique

Lorsque l'on considère la répartition géographique de *C. eos*, les cybrides, ainsi que *C. neogaeus*, seraient majoritairement présents dans la portion nord de cette répartition alors que la forme sauvage serait principalement retrouvée dans la portion sud (Mee et Taylor, 2012). Cette observation suggère que les cybrides sont capables de tolérer de plus froides températures que la forme sauvage.

Dans les Laurentides, région située à la frontière de la répartition des deux biotypes, les formes sauvages et cybrides sont généralement retrouvées en allopatrie et occasionnellement en sympatrie. Afin d'expliquer cette particularité, une supposition serait que sous certaines conditions environnementales les différences phénotypiques observées entre la forme sauvage et les cybrides seraient telles qu'elles favoriseraient l'établissement d'un des deux biotypes, expliquant ainsi leur allopatrie. Toutefois, sous d'autres conditions, plus rares, les variations phénotypiques mesurées seraient moins prononcées et ne permettraient la prévalence d'aucune des deux formes, expliquant alors leur sympatrie.

A plus large échelle géographique, dans les Laurentides, l'influence des mitochondries *C. neogaeus* sur l'expression de l'ADNnc *C. eos* ne serait pas suffisante pour empêcher les formes sauvages et cybrides de se côtoyer (i.e. allopatrie et sympatrie). Par contre, à de plus hautes latitudes, dans des environnements plus froids, la reprogrammation nucléaire induite par ces

mitochondries serait plus prononcée ; induisant l'expression d'un plus grand nombre de gènes normalement réprimés chez la forme sauvage, modifiant plus fortement le phénotype des cybrides et expliquant finalement la dominance de ce biotype dans ce type d'environnement. Parmi les gènes impliqués, certains pourraient induire une augmentation de la production de phospholipides à acides gras insaturés afin de maintenir la fluidité des membranes plasmiques et d'augmenter l'activité de certains enzymes membranaires (Guderley, 2004 ; Hochachka et Somero, 2002 ; Ohtsu et al., 1998 ; Podrabsky et Somero, 2004) ; favoriser la production d' « heat shock proteins », molécules chaperonnes ayant pour fonction d'assister d'autres protéines dans leur repliement (Doucet et al., 2009 ; Hochachka et Somero, 2002 ; Podrabsky et Somero, 2004) ; ou encore déclencher la synthèse de cryoprotectants (glycérol, sorbitol, inositol, *etc.*), permettant d'abaisser le point de congélation des cellules et de stabiliser les protéines (Hochachka et Somero, 2002 ; Storey et Storey, 1991). La différence d'activité mesurée au niveau du complexe IV pourrait également influer sur la prépondérance des cybrides dans des environnements plus froids. En effet, bien que l'activité de ce complexe ne soit pas significativement différente entre la forme sauvage et les cybrides à 17 et 10°C, l'activité moyenne de cet enzyme reste supérieure chez la forme introgressée. Si l'activité du complexe IV est extrapolée à l'activité métabolique, nous pouvons émettre l'hypothèse que les cybrides sont métaboliquement plus actif lors d'un abaissement de la température, avantage favorisant leur dominance à de plus hautes latitudes. Cette hypothèse semble être en partie appuyée par les résultats obtenus lors des analyses de performance de nage puisque l'endurance accrue des cybrides à 17°C implique un métabolisme plus performant à cette température.

5.5 Perspectives

Les résultats présentés dans cette thèse de doctorat mènent à une multitude de points méritant d'être explorés et/ou éclaircis. Entre autres, il serait particulièrement intéressant d'approfondir les répercussions que peuvent avoir les modifications phénotypiques détectées dans ce projet à un niveau d'organisation non-exploré, le métabolome. En effet, suite aux variations transcriptomiques et protéomiques précédemment observées, certaines voies métaboliques pourraient avoir été modifiées. De plus, l'augmentation de l'activité du complexe IV de la chaîne respiratoire chez les cybrides pourrait influer sur la respiration mitochondriale et par conséquent sur le métabolisme aérobie. Les cybrides semblant tolérer de plus froides températures que la forme sauvage, il serait également pertinent de comparer les modifications transcriptomiques et métaboliques engendrées suite à une acclimatation d'environ un an dans des conditions contrôlées, à de basses températures. Les données ainsi récoltées permettraient non seulement de déterminer si l'influence de l'ADNmt *C. neogaeus* augmente lorsque la température diminue, mais aussi de tester l'hypothèse selon laquelle les cybrides possèdent un métabolisme plus performant lorsque retrouvés dans des environnements froids. Finalement, il serait intéressant d'analyser la performance de nage des formes sauvages et cybrides lorsque retrouvées en sympatrie, à l'aide de tunnels de nage. En effet, dans le cadre du premier volet de cette thèse, l'analyse de ce niveau d'organisation avait été réalisée à partir d'individus allopatriques maintenus en laboratoire dans des conditions contrôlées identiques pour une période d'un an. Si cette méthode limitait l'influence des effets environnementaux, celle-ci ne pouvait écarter l'influence des effets nucléaires. L'analyse ici proposée permettrait l'élimination définitive de ces effets et d'ainsi confirmer l'influence de mitochondries étrangères à un niveau d'organisation de haute complexité du phénotype.

Pour conclure, l'influence des variations génotypiques sur le phénotype d'un organisme reste à l'heure actuelle l'une des questions à laquelle de nombreuses études tentent encore de répondre. Les résultats obtenus dans le cadre de ce projet de thèse participent à une meilleure compréhension de ce phénomène en établissant un lien entre les mitochondries et les différents niveaux d'organisation du phénotype, et possiblement entre les mitochondries et la structure des populations. À notre connaissance, l'approche holistique abordée dans cette recherche représente l'une des études intégratives les plus complètes jamais réalisée à partir d'organismes retrouvés en milieu naturel.

RÉFÉRENCES

Afanas'ev, I., 2015. Mechanisms of superoxide signaling in epigenetic processes: relation to aging and cancer. Aging Dis. 6, 216–227.

Al Rawi, S., Louvet-Vallee, S., Djeddi, A., Sachse, M., Culetto, E., Hajjar, C., Boyd, L., Legouis, R., Galy, V., 2011. Postfertilization autophagy of sperm organelles prevents paternal mitochondrial DNA transmission. Science 334, 1144–1147.

Andersson, S.G., Zomorodipour, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., Kurland, C.G., 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396, 133–140.

Angers, B., Castonguay, E., Massicotte, R., 2010. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. Mol Ecol 19, 1283–1295.

Angers, B., Dallaire, A., Vervaet, S., Vallières, F., Angers, A., 2012. The influence of mitochondria in epigenetics revealed through naturally occuring fish cybrids. Curr. Zool. 58, 138–145.

Angers, B., Schlosser, I.J., 2007. The origin of unisexual hybrids. Mol. Ecol. 16, 4562–4571.

Ankel-Simons, F., Cummins, J.M., 1996. Misconceptions about mitochondria and mammalian fertilization: implications for theories on human evolution. Proc. Natl. Acad. Sci. U. S. A. 93, 13859–13863.

Apel, K., Hirt, H., 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55, 373–399.

Arnold, S., 2012. The power of life - Cytochrome c oxidase takes center stage in metabolic control, cell signalling and survival. Mitochondrion 12, 46–56.

Arnqvist, G., Dowling, D.K., Eady, P., Gay, L., Tregenza, T., Tuda, M., Hosken, D.J., 2010. Genetic architecture of metabolic rate: environment specific epistasis between mitochondrial and nuclear genes in an insect. Evolution 64, 3354–3363.

Balaban, R.S., Nemoto, S., Finkel, T., 2005. Mitochondria, Oxidants, and Aging. Cell 120, 483–495.

Ballard, J.W., Melvin, R.G., 2010. Linking the mitochondrial genotype to the organismal phenotype. Mol Ecol 19, 1523–1539.

Balloux, F., 2001. EASYPOP (version 1.7): a computer program for population genetics simulations. J. Hered. 92, 301–302.

Barja, G., 2007. Mitochondrial oxygen consumption and reactive oxygen species production are independently modulated: implications for aging studies. Rejuvenation Res. 10, 215–224.

Barluenga, M., Sanetra, M., Meyer, A., 2006. Genetic admixture of burbot (Teleostei: *Lota lota*) in Lake Constance from two European glacial refugia. Mol. Ecol. 15, 3583–3600.

Bassam, B.J., Caetano-Anolles, G., Gresshoff, P.M., 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. Anal. Biochem. 196, 80–83.

Bayona-Bafaluy, M.P., Müller, S., Moraes, C.T., 2005. Fast Adaptive Coevolution of Nuclear and Mitochondrial Subunits of ATP Synthetase in Orangutan. Mol. Biol. Evol. 22, 716–724.

Beerli, P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22, 341–345.

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and

Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B 57, 289–300.

Bermingham, E., McCafferty, S.S., Martin, A.P., 1997. Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. Mol. Syst. Fishes 7, 431–452.

Bernatchez, L., Glémet, H., Wilson, C.C., Danzmann, R.G., 1995. Introgression and fixation of Arctic char (Salvelinus alpinus) mitochondrial genome in an allopatric population of brook trout (*Salvelinus fontinalis*). Can. J. Fish. Aquat. Sci. 52, 179–185.

Binet, M.C., Angers, B., 2005. Genetic identification of members of the *Phoxinus eos-neogaeus* hybrid complex. J. Fish Biol. 67, 1169–1177.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc. Natl. Acad. Sci. U. S. A. 92, 11331–11338.

Bleier, L., Drose, S., 2013. Superoxide generation by complex III: from mechanistic rationales to functional consequences. Biochim. Biophys. Acta 1827, 1320–1331.

Blier, P.U., Breton, S., Desrosiers, V., Lemieux, H., 2006. Functional conservatism in mitochondrial evolution: insight from hybridization of arctic and brook charrs. J. Exp. Zool. 306, 425–432.

Blier, P.U., Dufresne, F., Burton, R.S., 2001. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. Trends Genet. 17, 400–406.

Blier, P.U., Lemieux, H., 2001. The impact of the thermal sensitivity of cytochrome c oxidase on the respiration rate of Arctic charr red muscle mitochondria. J. Comp. Physiol. B. 171, 247–253.

Blier, P.U., Lemieux, H., Pichaud, N., 2014. Holding our breath in our modern world: will mitochondria keep the pace with climate changes? Can. J. Zool. 92, 591–601.

Blier, P.U., Guderley, H.E., 1993. Effects of pH and temeprature on the kinetics of pyruvate

oxidation by muscle mitochondria from rainbo trout (*Oncorhynchus mykiss*). Physiol. Zool. 66, 474-489.

Boratynski, Z., Alves, P.C., Berto, S., Koskela, E., Mappes, T., Melo-Ferreira, J., 2011. Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics? BMC Evol. Biol. 11, 355.

Boyes, J., Bird, A., 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. Cell 64, 1123–1134.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Bradic, M., Beerli, P., García-de León, F.J., Esquivel-Bobadilla, S., Borowsky, R.L., 2012. Gene flow and population structure in the Mexican blind cavefish complex (*Astyanax mexicanus*). BMC Evol. Biol. 12, 1–17.

Brelsford, A., Mila, B., Irwin, D.E., 2011. Hybrid origin of Audubon's warbler. Mol. Ecol. 20, 2380–2389.

Bridge, D., Cunningham, C.W., Schierwater, B., DeSalle, R., Buss, L.W., 1992. Class-level relationships in the phylum Cnidaria: evidence from mitochondrial genome structure. Proc. Natl. Acad. Sci. U. S. A. 89, 8750–8753.

Burton, R.S., Barreto, F.S., 2012. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities? Mol Ecol 21, 4942–4957.

Burton, R.S., Ellison, C.K., Harrison, J.S., 2006. The sorry state of F2 hybrids: consequences of rapid mitochondrial DNA evolution in allopatric populations. Am Nat 168 Suppl, S14–24.

Burton, R.S., Pereira, R.J., Barreto, F.S., 2013. Cytonuclear Genomic Interactions and Hybrid

Breakdown. Annu. Rev. Ecol. Evol. Syst. 44, 14.1–14.22.

Calvo, S., Jain, M., Xie, X., Sheth, S.A., Chang, B., Goldberger, O.A., Spinazzola, A., Zeviani, M., Carr, S.A., Mootha, V.K., 2006. Systematic identification of human mitochondrial disease genes through integrative genomics. Nat. Genet. 38, 576–582.

Cannon, M. V, Dunn, D.A., Irwin, M.H., Brooks, A.I., Bartol, F.F., Trounce, I.A., Pinkert, C.A., 2011. Xenomitochondrial mice: investigation into mitochondrial compensatory mechanisms. Mitochondrion 11, 33–39.

Chan, D.C., 2006. Mitochondrial fusion and fission in mammals. Annu. Rev. Cell Dev. Biol. 22, 79–99.

Chittka, A., Chittka, L., 2010. Epigenetics of Royalty. PLoS Biol 8, 1–4.

Chou, J.-Y., Hung, Y.-S., Lin, K.-H., Lee, H.-Y., Leu, J.-Y., 2010. Multiple molecular mechanisms cause reproductive isolation between three yeast species. PLoS Biol 8, e1000432.

Clay Montier, L.L., Deng, J.J., Bai, Y., 2009. Number matters: control of mammalian mitochondrial DNA copy number. J. Genet. Genomics 36, 125–131.

Corander, J., Marttinen, P., 2006. Bayesian identification of admixture events using multilocus molecular markers. Mol. Ecol. 15, 2833–2843.

Corander, J., Marttinen, P., Sirén, J., Tang, J., 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9, 539.

Cummins, J., 1998. Mitochondrial DNA in mammalian reproduction. Rev. Reprod. 3, 172–182.

Dawley, R.M., Schultz, R.J., Goddard, K.A., 1987. Clonal reproduction and polyploidy in unisexual hybrids of *Phoxinus eos* and *Phoxinus neogaeus* (Pisces; Cyprinidae). Copeia 1987, 275–283.

DeLuca, S.Z., O'Farrell, P.H., 2012. Barriers to male transmission of mitochondrial DNA in sperm development. Dev. Cell 22, 660–668.

Deremiens, L., Schwartz, L., Angers, A., Glemet, H., Angers, B., 2015. Interactions between nuclear genes and a foreign mitochondrial genome in the redbelly dace *Chrosomus eos*. Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 189, 80–86.

Dey, R., Barrientos, A., Moraes, C.T., 2000. Functional constraints of nuclear-mitochondrial DNA interactions in xenomitochondrial rodent cell lines. J Biol Chem 275, 31520–31527.

Doucet, D., Walker, V.K., Qin, W., 2009. The bugs that came in from the cold: molecular adaptations to low temperatures in insects. Cell. Mol. Life Sci. 66, 1404–1418.

Downs, J.A., Jackson, S.P., 2003. Cancer: protective packaging for DNA. Nature.

Du, F.K., Peng, X.L., Liu, J.Q., Lascoux, M., Hu, F.S., Petit, R.J., 2011. Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. New Phytol. 192, 1024–1033.

Edmands, S., 2002. Does parental divergence predict reproductive compatibility? Trends Ecol. & Evol. 17, 520–527.

El Mousadik, A., Petit, R.J., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa (L.)* Skeels] endemic to Morocco. Theor. Appl. Genet. 92, 832–839.

Ellison, C.K., Burton, R.S., 2006. Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. Evolution (N. Y). 60, 1382–1391.

Ellison, C.K., Burton, R.S., 2008a. Interpopulation hybrid breakdown maps to the mitochondrial genome. Evolution (N. Y). 62, 631–638.

Ellison, C.K., Burton, R.S., 2008b. Genotype-dependent variation of mitochondrial

transcriptional profiles in interpopulation hybrids. Proc Natl Acad Sci U S A 105, 15831–15836.

Ellison, C.K., Burton, R.S., 2010. Cytonuclear conflict in interpopulation hybrids: the role of RNA polymerase in mtDNA transcription and replication. J. Evol. Biol. 23, 528–538.

Embley, T.M., Martin, W., 2006. Eukaryotic evolution, changes and challenges. Nature 440, 623–630.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol. Bioinform. Online.

Falkenberg, M., Larsson, N.G., Gustafsson, C.M., 2007. DNA replication and transcription in mammalian mitochondria. Annu Rev Biochem 76, 679–699.

Falush, D., Stephens, M., Pritchard, J.K., 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164, 1567–1587.

Feder, J.L., Egan, S.P., Nosil, P., 2012. The genomics of speciation-with-gene-flow. Trends Genet. 28, 342–350.

Gabaldón, T., Huynen, M.A., 2004. Shaping the mitochondrial proteome. Biochim. Biophys. Acta - Bioenerg. 1659, 212–220.

Gagnaire, P.-A., Pavey, S.A., Normandeau, E., Bernatchez, L., 2013. The genetic architecture of reproductive isolation during speciation-with-gene-flow in Lake Whitefish species pairs assessed by RAD sequencing. Evolution (N. Y). 67, 2483–2497.

Gahan, M.E., Miller, F., Lewin, S.R., Cherry, C.L., Hoy, J.F., Mijch, A., Rosenfeldt, F., Wesselingh, S.L., 2016. Quantification of mitochondrial DNA in peripheral blood mononuclear cells and subcutaneous fat using real-time polymerase chain reaction. J. Clin.

Virol. 22, 241–247.

Garesse, R., Vallejo, C.G., 2001. Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. Gene 263, 1–16.

Garner, M.M., Burg, M.B., 1994. Macromolecular crowding and confinement in cells exposed to hypertonicity. AM. J. Physiol. 266, C877-C892.

Garza, J.C., Williamson, E.G., 2001. Detection of reduction in population size using data from microsatellite loci. Mol. Ecol. 10, 305–318.

Gaspari, M., Falkenberg, M., Larsson, N.G., Gustafsson, C.M., 2004. The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. EMBO J 23, 4606–4614.

Gershoni, M., Templeton, A.R., Mishmar, D., 2009. Mitochondrial bioenergetics as a major motive force of speciation. Bioessays 31, 642–650.

Giles, R.E., Blanc, H., Cann, H.M., Wallace, D.C., 1980. Maternal inheritance of human mitochondrial DNA. Proc. Natl. Acad. Sci. U. S. A. 77, 6715–6719.

Glemet, H., Blier, P., Bernatchez, L., 1998. Geographical extent of Arctic char (*Salvelinus alpinus*) mtDNA introgression in brook char populations (*S. fontinalis*) from eastern Quebec, Canada. Mol. Ecol. 7, 1655–1662.

Goddard, K.A., Dawley, R.M., Dowling, T.E., 1989. Origin and genetic relationships of diploid, triploid, and diploid-triploid mosaic biotypes in the Phoxinus eos-neogaeus unisexual complex, in: Dawley, R.M., Bogart, J.P. (Eds.), Evolution and Ecology of Unisexual Vertebrates. New-York State Education Department, Albany, pp. 269–280.

Goddard, K.A., Megwinoff, O., Wessner, L.L., Giaimo, F., 1998. Confirmation of gynogenesis in *Phoxinus eos-neogaeus* (Pisces: Cyprinidae). J. Hered. 89, 151–157.

Goudet, J., 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. J.Hered. 86, 485–486.

Goudet, J., 2001. FSTAT: a program to estimate and test gene diversities and fixation indices.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotech 29, 644–652.

Gray, M.W., Lang, B.F., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., Brossard, N., Delage, E., Littlejohn, T.G., Plante, I., Rioux, P., Saint-Louis, D., Zhu, Y., Burger, G., 1998. Genome structure and gene content in protist mitochondrial DNAs. Nucleic Acids Res. 26, 865–878.

Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. Science. 283, 1476–1481.

Gray, M., Burger, G., Lang, B.F., 2001. The origin and early evolution of mitochondria. Genome Biol. 2, 1018.1–1018.5.

Gray, M.W., 2012. Mitochondrial evolution. Cold Spring Harb Perspect Biol 4, a011403.

Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. Science. 283, 1476–1481.

Green, D.R., Reed, J.C., 1998. Mitochondria and apoptosis. Science 281, 1309–1312.

Guderley, H., 2004. Metabolic response to low temperature in fish muscle. Biol. Rev. Camb. Philos. Soc. 79, 409–427.

Gusdon, A.M., Votyakova, T. V, Reynolds, I.J., Mathews, C.E., 2007. Nuclear and mitochondrial interaction involving mt-Nd2 leads to increased mitochondrial reactive oxygen species production. J. Biol. Chem. 282, 5171–5179.

Gyllensten, U., Wharton, D., Josefsson, A., Wilson, A.C., 1991. Paternal inheritance of mitochondrial DNA in mice. Nature 352, 255–257.

He, G., Zhu, X., Elling, A.A., Chen, L., Wang, X., Guo, L., Liang, M., He, H., Zhang, H., Chen, F., Qi, Y., Chen, R., Deng, X.-W., 2010. Global epigenetic and transcriptional trends among two rice subspecies and their reciprocal hybrids. Plant Cell.

Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. 58, 247–276.

Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.

Hochachka, P., Somero, G., 2002. Biochemical adaptation: mechanism and process in physiological evolution. Oxford University Press, Oxford.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.

Johnson, A.A., Johnson, K.A., 2001. Exonuclease proofreading by human mitochondrial DNA polymerase. J Biol Chem 276, 38097–38107.

Kai, Y., Takamatsu, C., Tokuda, K., Okamoto, M., Irita, K., Takahashi, S., 2006. Rapid and random turnover of mitochondrial DNA in rat hepatocytes of primary culture. Mitochondrion 6, 299–304.

Kayal, E., Bentlage, B., Collins, A.G., Kayal, M., Pirro, S., Lavrov, D. V, 2012. Evolution of linear mitochondrial genomes in medusozoan cnidarians. Genome Biol. Evol. 4, 1–12.

Kenyon, L., Moraes, C.T., 1997. Expanding the functional human mitochondrial DNA database by the establishment of primate xenomitochondrial cybrids. Proc. Natl. Acad. Sci. U. S. A. 94, 9131–9135.

Kucharski, R., Maleszka, J., Foret, S., Maleszka, R., 2008. Nutritional control of reproductive

status in honeybees via DNA methylation. Science. 319, 1827–1830.

Kuroiwa, T., 2010. Review of cytological studies on cellular and molecular mechanisms of uniparental (maternal or paternal) inheritance of plastid and mitochondrial genomes induced by active digestion of organelle nuclei (nucleoids). J. Plant Res. 123, 207–230.

Lang, B.F., Gray, M.W., Burger, G., 1999. Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33, 351–397.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Leberg, P.L., 2002. Estimating allelic richness: effects of sample size and bottlenecks. Mol. Ecol. 11, 2445–2449.

Lee, H.-Y., Chou, J.-Y., Cheong, L., Chang, N.-H., Yang, S.-Y., Leu, J.-Y., 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell 135, 1065–1073.

Leitch, A.R., Leitch, I.J., 2008. Genomic plasticity and the diversity of polyploid plants. Science 320, 481–483.

Levings, C.S., Brown, G.G., 1989. Molecular biology of plant mitochondria. Cell 56, 171–179.

Lim, S.-O., Gu, J.-M., Kim, M.S., Kim, H.-S., Park, Y.N., Park, C.K., Cho, J.W., Park, Y.M., Jung, G., 2008. Epigenetic changes induced by reactive oxygen species in hepatocellular carcinoma: methylation of the E-cadherin promoter. Gastroenterology 135, 2128–2140.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-ΔΔCt) Method. Methods 25, 402–408.

Lucassen, M., Schmidt, A., Eckerle, L.G., Pörtner, H.O., 2003. Mitochondrial proliferation in the permanent vs. temporary cold: enzyme activities and mRNA levels in Antarctic and

temperate zoarcid fish. Am. J. Physiol.-Reg. I. 285, R1410–R1420.

Luikart, G., Cornuet, J., 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conserv. Biol. 12, 228–237.

Lynch, M., 1996. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Mol. Biol. Evol. 13, 209–220.

Lynch, M., Koskella, B., Schaack, S., 2006. Mutation pressure and the evolution of organelle genomic architecture. Science. 311, 1727–1730.

Magnus, G., Keizer, J., 1998. Model of beta-cell mitochondrial calcium handling and electrical activity. II. Mitochondrial variables. Am. J. Physiol. 274, C1174–84.

Mallet, J., Meyer, A., Nosil, P., Feder, J.L., 2009. Space, sympatry and speciation. J. Evol. Biol. 22, 2332–2341.

Martinou, J.C., Desagher, S., Antonsson, B., 2000. Cytochrome c release from mitochondria: all or nothing. Nat. Cell Biol.

Massicotte, R., Magnan, P., Angers, B., 2008. Intralacustrine site fidelity and nonrandom mating in the littoral-spawning northern redbelly dace (*Phoxinus eos*). Can. J. Fish. Aquat. Sci. 65, 2016–2025.

McKenzie, M., Chiotis, M., Pinkert, C.A., Trounce, I.A., 2003. Functional respiratory chain analyses in murid xenomitochondrial cybrids expose coevolutionary constraints of cytochrome b and nuclear subunits of complex III. Mol. Biol. Evol. 20, 1117–1124.

McKenzie, M., Trounce, I., 2000. Expression of *Rattus norvegicus* mtDNA in *Mus musculus* cells results in multiple respiratory chain defects. J. Biol. Chem. 275, 31514–31519.

Mee, J.A., Taylor, E.B., 2012. The cybrid invasion: widespread postglacial dispersal by *Phoxinus* (Pisces: Cyprinidae) cytoplasmic hybrids. Can. J. Zool. 90, 577–584.

Meiklejohn, C.D., Holmbeck, M.A., Siddiq, M.A., Abt, D.N., Rand, D.M., Montooth, K.L., 2013. An incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in *Drosophila*. PLoS Genet 9, e1003238.

Mila, B., Toews, D.P.L., Smith, T.B., Wayne, R.K., 2011. A cryptic contact zone between divergent mitochondrial DNA lineages in southwestern North America supports past introgressive hybridization in the yellow-rumped warbler complex (Aves: Dendroica coronata). Biol. J. Linn. Soc. 103, 696–706.

Moreno-Loshuertos, R., Acin-Perez, R., Fernandez-Silva, P., Movilla, N., Perez-Martos, A., Rodriguez de Cordoba, S., Gallardo, M.E., Enriquez, J.A., 2006. Differences in reactive oxygen species production explain the phenotypes associated with common mouse mitochondrial DNA variants. Nat Genet 38, 1261–1268.

Moritz, C., Dowling, T.E., Brown, W.M., 1987. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu. Rev. Ecol. Syst. 18, 269–292.

Murphy, M.P., 2009. How mitochondria produce reactive oxygen species. Biochem. J. 417, 1–13.

Nadeau, N.J., Whibley, A., Jones, R.T., Davey, J.W., Dasmahapatra, K.K., Baxter, S.W., Quail, M.A., Joron, M., Ffrench-Constant, R.H., Blaxter, M.L., Mallet, J., Jiggins, C.D., 2012. Genomic islands of divergence in hybridizing *Heliconius* butterflies identified by large-scale targeted sequencing. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367, 343–353.

Nagao, Y., Totsuka, Y., Atomi, Y., Kaneda, H., Lindahl, K.F., Imai, H., Yonekawa, H., 1998. Decreased physical performance of congenic mice with mismatch between the nuclear and the mitochondrial genome. Genes Genet Syst 73, 21–27.

Nei, M., 1987. Molecular evolutionary genetics. Columbia university press, New York.

Nicholls, D.G., Budd, S.L., 2000. Mitochondria and neuronal survival. Physiol. Rev. 80, 315–360.

Nishimura, Y., Yoshinari, T., Naruse, K., Yamada, T., Sumi, K., Mitani, H., Higashiyama, T., Kuroiwa, T., 2006. Active digestion of sperm mitochondrial DNA in single living sperm revealed by optical tweezers. Proc. Natl. Acad. Sci. U. S. A. 103, 1382–1387.

Nosil, P., Feder, J.L., 2011. Genomic divergence during speciation: causes and consequences. Philos. Trans. R. Soc. London B Biol. Sci. 367, 332–342.

Ohtsu, T., Kimura, M.T., Katagiri, C.,. 1998. How *Drosophila* species acquire cold tolerance. Eur. J. Biochem. 252, 608–611.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M., 2007. The vegan package. Community Ecol. Packag. 10.

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., Sekiya, T., 1989. Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. Proc. Natl. Acad. Sci. U. S. A. 86, 2766–2770.

Osada, N., Akashi, H., 2012. Mitochondrial-nuclear interactions and accelerated compensatory evolution: evidence from the primate cytochrome c oxidase complex. Mol Biol Evol 29, 337–346.

Paquin, B., Laforest, M.-J., Forget, L., Roewer, I., Wang, Z., Longcore, J., Lang, F.B., 1997. The fungal mitochondrial genome project: evolution of fungal mitochondrial genomes and their gene expression. Curr. Genet. 31, 380–395.

Parekh, A.B., 2003. Mitochondrial regulation of intracellular Ca^{2+} signaling: more than just simple Ca^{2+} buffers. News Physiol. Sci. an Int. J. Physiol. Prod. jointly by Int. Union Physiol. Sci. Am. Physiol. Soc. 18, 252–256.

Pelster, B., Sänger, A.M., Siegele, M., Schwerte, T., 2003. Influence of swim training on cardiac activity, tissue capillarization, and mitochondrial density in muscle tissue of zebrafish larvae. Am. J. Physiol.-Reg. I. 285(2), R339–R347.

Pesole, G., Gissi, C., De Chirico, A., Saccone, C., 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. J. Mol. Evol. 48, 427–434.

Pichaud, N., Ballard, J.W., Tanguay, R.M., Blier, P.U., 2011. Thermal sensitivity of mitochondrial functions in permeabilized muscle fibers from two populations of *Drosophila simulans* with divergent mitotypes. Am J Physiol Regul Integr Comp Physiol 301, R48–59.

Pichaud, N., Ballard, J.W.O., Tanguay, R.M., Blier, P.U., 2012. Naturally occuring mitochondrial DNA haplotypes exhibit metabolic differences: insight intoo functional properties of mitochondria. Evolution (N. Y). 66, 3189–3197.

Piry, S., Luikart, G., Cornuet, J.-M., 1999. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. J. Hered. 90, 502–503.

Podrabsky, J.E., Somero, G.N., 2004. Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. J. Exp. Biol. 207, 2237–2254.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

Rand, D.M., Haney, R.A., Fry, A.J., 2004. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol. Evol. 19, 645–653.

Rawson, P.D., Burton, R.S., 2002. Functional coadaptation between cytochrome c and

cytochrome c oxidase within allopatric populations of a marine copepod. Proc Natl Acad Sci U S A 99, 12955–12958.

Raymond, M., Rousset, F., 1995. Genepop: population genetics software for exact tests and ecumenicism. J. Hered. 86.

Razin, A., Riggs, A.D., 1980. DNA methylation and gene function. Science. 210, 604–610.

Rice, J.C., Allis, C.D., 2001. Histone methylation versus histone acetylation: new insights into epigenetic regulation. Curr. Opin. Cell Biol. 13, 263–273.

Rice, W.R., 1989. Analyzing Tables of Statistical Tests. Evolution 43, 223–225.

Robin, E.D., Wong, R., 1988. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J. Cell. Physiol. 136, 507–513.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Sato, M., Sato, K., 2013. Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. Biochim. Biophys. Acta 1833, 1979–1984.

Satoh, M., Kuroiwa, T., 1991. Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. Exp. Cell Res. 196, 137–140.

Scheffler, I.E., 2008. Mitochondria, 2nd edn. ed. Wiley-Liss Publishers, Hoboken, NJ.

Schon, E.A., 2000. Mitochondrial genetics and disease. Trends Biochem. Sci. 25, 555–560.

Scott, W.B., Crossman, E.J., 1973. Freshwater fishes of Canada. Edited by O. d. r. s. l. p. d. Canada. Vol. Bulletin 184. Ottawa.

Senjo, M., Kimura, K., Watano, Y., Ueda, K., Shimizu, T., 1999. Extensive Mitochondrial Introgression from *Pinus pumila* to *P. parviflora var. pentaphylla* (Pinaceae). J. Plant Res. 112, 97–105.

Sharpley, M.S., Marciniak, C., Eckel-Mahan, K., McManus, M., Crimi, M., Waymire, K., Lin, C.S., Masubuchi, S., Friend, N., Koike, M., Chalkia, D., MacGregor, G., Sassone-Corsi, P., Wallace, D.C., 2012. Heteroplasmy of mouse mtDNA is genetically unstable and results in altered behavior and cognition. Cell 151, 333–343.

Skulachev, V.P., 1999. Mitochondrial physiology and pathology; concepts of programmed death of organelles, cells and organisms. Mol. Aspects Med. 20, 139–184.

Spirek, M., Horvath, A., Piskur, J., Sulo, P., 2000. Functional co-operation between the nuclei of *Saccharomyces cerevisiae* and mitochondria from other yeast species. Curr Genet 38, 202–207.

Storey, K., Storey, J., 1991. Biochemistry of cryoprotectants in insects at low temperature. R. Lee, Jr. and D. Denlinger, editors. Springer US. 64-93.

Strange, R.M., Stepien, C.A., 2007. Genetic divergence and connectivity among river and reef spawning groups of walleye (*Sander vitreus vitreus*) in Lake Erie. Can. J. Fish. Aquat. Sci. 64, 437–448.

Sutovsky, P., Moreno, R.D., Ramalho-Santos, J., Dominko, T., Simerly, C., Schatten, G., 2000. Ubiquitinated sperm mitochondria, selective proteolysis, and the regulation of mitochondrial inheritance in mammalian embryos. Biol. Reprod. 63, 582–590.

Taanman, J.W., 1999. The mitochondrial genome: structure, transcription, translation and replication. Biochim. Biophys. Acta 1410, 103–123.

Tautz, J., 2008. The buzz about bees: biology of a superorganism. Springer Berlin Heidelberg,

Berlin, Heidelberg, pp. 140–154.

Toews, D.P., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907–3930.

Toews, D.P., Mandic, M., Richards, J.G., Irwin, D.E., 2014. Migration, mitochondria, and the yellow-rumped warbler. Evolution (N. Y). 68, 241–255.

Varnavskaya, N. V, Wood, C.C., Everett, R.J., Wilmot, R.L., Varnavsky, V.S., Midanaya, V. V, Quinn, T.P., 1994. Genetic Differentiation of Subpopulations of Sockeye Salmon (*Oncorhynchus nerka*) Within Lakes of Alaska, British Columbia, and Kamchatka, Russia. Can. J. Fish. Aquat. Sci. 51, 147–157.

Via, S., 2012. Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367, 451–460.

Wallace, D.C., Ye, J., Neckelmann, S.N., Singh, G., Webster, K.A., Greenberg, B.D., 1987. Sequence analysis of cDNAs for the human and bovine ATP synthase β subunit: mitochondrial DNA genes sustain seventeen times more mutations. Curr Genet 12, 81–90. doi:10.1007/bf00434661

Wolff, J.N., Ladoukakis, E.D., Enriquez, J.A., Dowling, D.K., 2014. Mitonuclear interactions: evolutionary consequences over multiple biological scales. Philos. Trans. R. Soc. London B Biol. Sci. 369: 20130443.

Wu, Q., Ni, X., 2015. ROS-mediated DNA methylation pattern alterations in carcinogenesis. Curr. Drug Targets 16, 13–19.

Xiong, L.Z., Xu, C.G., Saghai Maroof, M.A., Zhang, Q., 1999. Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. Mol. Gen. Genet. 261, 439–446.

Yakes, F.M., Van Houten, B., 1997. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. Proc Natl Acad Sci U S A 94, 514–519.

Yue, G.H., David, L., Orban, L., 2007. Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio L.*). Genetica 129, 329–331.

Zielinski, P., Nadachowska-Brzyska, K., Wielstra, B., Szkotak, R., Covaciu-Marcov, S.D., Cogalniceanu, D., Babik, W., 2013. No evidence for nuclear introgression despite complete mtDNA replacement in the Carpathian newt (*Lissotriton montandoni*). Mol. Ecol. 22, 1884–1903.

ANNEXES

Table A1: Primer sequences used for qRT-PCR analysis of mRNA expression in *C. eos* wild type and cybrids.

SYBR target	Primer 3'-5'	Product (bp)
COX1_F	AACAATTTTCTCSCTTCACC	108
COX1_R	GTTTGGTATTGRGAAATGGC	
COX2_F	GACACCAATGATACTGAAGC	109
COX2_R	TCTAGGAGCCGAAATTGACC	
COX3_F	CCGTTTGAAGTYCCTCTCC	100
COX3_R	CCTGTTTTCGTTCGCCTTC	
COX5b_F	TACACTGGATCCAGAGATGACC	112
COX5b_R	CGTGAAGCCAGAACCAAACG	
COX6b_F	GCGCTTAAAAGTCTTCGAGAGG	133
COX6b_R	GAAAACGGGCATCGAATGG	
COX7a1_F	GCACAACAGTCCGGCAAATG	133
COX7a1_R	GTGATGGCCATGGTCAAACG	
β-actin_F	TACAATGAGCTGCGTGTTGC	117
β-actin_R	GGTGTTGAAGGTCTCGAACATG	

Table A2: Tested clustering scenarios (*a* to *s*) based on the haplotypes and the sampled sites of Desjardins Lake using the program BAPS. The highest log likelihood are indicated in bold.

Group	Scenario																		
	a	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*	*j*	*k*	*l*	*m*	*n*	*o*	*p*	*q*	*r*	*s*
A - MC	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A - AC	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2
A- WT	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	2	2	1
B - MC	4	1	2	2	1	1	1	2	2	2	2	1	2	2	2	1	1	1	1
B - AC	5	1	2	2	1	1	1	2	2	2	2	1	2	2	2	2	1	2	2
B- WT	6	1	2	2	1	1	1	2	2	2	2	1	2	2	2	3	2	2	1
C - MC	7	1	3	2	2	1	2	1	1	2	2	2	1	3	3	1	1	1	1
C - AC	8	1	3	2	2	1	2	1	1	2	2	2	1	3	3	2	1	2	2
C - WT	9	1	3	2	2	1	2	1	1	2	2	2	1	3	3	3	2	2	1
D - MC	10	1	4	2	2	2	1	1	2	1	3	3	3	1	2	1	1	1	1
D - AC	11	1	4	2	2	2	1	1	2	1	3	3	3	1	2	2	1	2	2
D- WT	12	1	4	2	2	2	1	1	2	1	3	3	3	1	2	3	2	2	1
Log likelihood	-5559	**-5269**	-5352	-5352	**-5269**	**-5269**	-5352	**-5269**	-5352	-5352	-5352	-5352	-5338	-5369	-5352	-5352	-5338	-5436	-5417

Note. MC – Mississippian Cybrid; AC – Atlantic Cybrid; WT – Wild Type; A, B, C, D - Local sites.

Table A3: Measure of H_E, H_{EQ}, H_E-H_{EQ} (mean ± standard deviation) and estimation of the mutation-drift equilibrium (MDE) for two simulated populations (EASYPOP). Among the 50 simulated loci, only 20 are represented.

Locus	Simulated population A			Simulated population B		
	H_E	H_{EQ}	H_E - H_{EQ}	H_E	H_{EQ}	H_E - H_{EQ}
1	0.729 ± 0.105	0.617 ± 0.078	0.112 ± 0.090	0.721 ± 0.091	0.616 ± 0.057	0.104 ± 0.076
2	0.763 ± 0.077	0.624 ± 0.081	0.139 ± 0.062	0.729 ± 0.099	0.615 ± 0.080	0.114 ± 0.058
3	0.755 ± 0.112	0.636 ± 0.084	0.120 ± 0.075	0.739 ± 0.080	0.615 ± 0.088	0.124 ± 0.066
4	0.743 ± 0.097	0.630 ± 0.053	0.113 ± 0.089	0.682 ± 0.143	0.574 ± 0.107	0.109 ± 0.115
5	0.732 ± 0.096	0.624 ± 0.066	0.108 ± 0.066	0.729 ± 0.130	0.621 ± 0.074	0.109 ± 0.093
6	0.748 ± 0.072	0.614 ± 0.065	0.135 ± 0.070	0.726 ± 0.128	0.603 ± 0.098	0.123 ± 0.102
7	0.796 ± 0.048	0.652 ± 0.076	0.144 ± 0.057	0.746 ± 0.101	0.627 ± 0.097	0.119 ± 0.061
8	0.743 ± 0.068	0.624 ± 0.064	0.119 ± 0.059	0.765 ± 0.063	0.619 ± 0.087	0.146 ± 0.069
9	0.780 ± 0.051	0.642 ± 0.075	0.138 ± 0.059	0.769 ± 0.078	0.644 ± 0.078	0.126 ± 0.062
10	0.781 ± 0.093	0.666 ± 0.054	0.115 ± 0.060	0.705 ± 0.122	0.611 ± 0.093	0.094 ± 0.073
11	0.760 ± 0.082	0.610 ± 0.109	0.150 ± 0.059	0.715 ± 0.107	0.609 ± 0.092	0.106 ± 0.078
12	0.739 ± 0.096	0.653 ± 0.068	0.088 ± 0.088	0.750 ± 0.071	0.637 ± 0.068	0.113 ± 0.061
13	0.754 ± 0.061	0.627 ± 0.042	0.127 ± 0.049	0.770 ± 0.092	0.650 ± 0.072	0.119 ± 0.075
14	0.766 ± 0.082	0.639 ± 0.066	0.127 ± 0.064	0.770 ± 0.088	0.648 ± 0.056	0.122 ± 0.057
15	0.721 ± 0.089	0.627 ± 0.057	0.093 ± 0.092	0.709 ± 0.106	0.610 ± 0.078	0.099 ± 0.081
16	0.762 ± 0.089	0.617 ± 0.105	0.145 ± 0.060	0.778 ± 0.091	0.629 ± 0.090	0.149 ± 0.036
17	0.775 ± 0.062	0.632 ± 0.060	0.144 ± 0.052	0.730 ± 0.089	0.603 ± 0.095	0.127 ± 0.075
18	0.727 ± 0.102	0.600 ± 0.090	0.127 ± 0.070	0.785 ± 0.046	0.651 ± 0.058	0.133 ± 0.061
19	0.740 ± 0.110	0.618 ± 0.093	0.122 ± 0.064	0.686 ± 0.146	0.620 ± 0.083	0.067 ± 0.122
20	0.776 ± 0.079	0.645 ± 0.064	0.131 ± 0.078	0.734 ± 0.090	0.617 ± 0.047	0.117 ± 0.074
MDE probability test		< 0.0005			< 0.0005	

Table A4: Measure of H_E, H_{EQ}, H_E-H_{EQ} and estimation of the mutation-drift equilibrium (MDE) for wild type (WD), Mississippian (MC) and Atlantic cybrids (AC) in Desjardins Lake.

Locus	WT			MC			AC		
	H_E	H_{EQ}	H_E - H_{EQ}	H_E	H_{EQ}	H_E - H_{EQ}	H_E	H_{EQ}	H_E - H_{EQ}
Pho1	0,830	0,730	0,100	0,841	0,787	0,054	0,839	0,777	0,062
Pho2	0,868	0,841	0,027	0,880	0,812	0,068	0,876	0,802	0,074
Pho4	0,701	0,502	0,199	0,737	0,495	0,242	0,729	0,572	0,157
Pho60	0,772	0,711	0,061	0,798	0,706	0,092	0,786	0,708	0,078
Pho61	0,689	0,711	-0,022	0,724	0,676	0,048	0,698	0,671	0,027
MDE probability test		< 0,05			< 0,05			< 0,05	

Table A5: The significantly differentially expressed and characterised transcripts in white muscle, when *C eos* wild type and Mississippian cybrids from Desjardins Lake are compared. Transcripts associated to foreign mitochondria effects (14) alone are indicated by an asterisk. mRNAs homologous to proteins attributed to mitochondria (3) are in bold. Only matches with an E-value below 10^{-5} were considered. Significance was set as $P_{adj} < 0.05$.

Transcript	Protein ID	Description	E-value
1	XP_691804.2	Predicted: Transmembrane glycoprotein NMB isoform X2*	0
2	NP_998105.2	Ligand of numb-protein X 2b	0
3	XP_007241229.1	Predicted: Zinc finger BED domain-containing protein	0
4	NP_001018510.1	Guanine deaminase*	0
5	XP_007250496.1	Predicted: Carboxypeptidase Q	0
6	XP_007259625.1	**Predicted: Pyruvate dehydrogenase (acetyl-transferring) kinase isozyme 3***	0
7	Q08CS6.2	DBH-like monooxygenase protein 2 homolog	0
8	XP_010792626.1	Predicted: Zinc finger protein 862-like*	0
9	XP_003200076.2	Predicted: Integrin alpha-E-like*	0
10	XP_001346132.4	Predicted: Interferon-induced very large GTPase 1-like*	0
11	NP_956704.1	Solute carrier family 16, member 9a	0
12	XP_005465275.1	Predicted: Protein NYNRIN-like	0
13	KKF10883.1	Retrovirus-related Pol polyprotein from transposon 297	0
14	XP_005167957.1	**Predicted: Glutaminase kidney isoform***	6E-176
15	XP_002664750.1	Predicted: Sphingosine-1-phosphate phosphatase 2	8E-152
16	XP_009297040.1	Predicted: Kin of IRRE like b isoform X1*	1E-124
17	KRX16276.1	Transposon Ty3-G Gag-Pol polyprotein*	1E-108
18	XP_009291085.1	Predicted: Afadin-like*	1E-100
19	NP_997932.1	Transducin beta-like protein 2 precursor*	1E-89
20	XP_695841.5	Predicted: G protein-coupled receptor kinase 5-like isoform X2	6E-87
21	XP_007242179.1	Predicted: Dihydropyrimidinase*	5E-83
22	XP_005167479.1	Predicted: nectin-1 isoform X2*	1E-75
23	XP_002663014.2	Predicted: Macrophage mannose receptor 1	2E-69
24	KKF13591.1	Zinc finger protein 235	1E-64
25	NP_001082927.1	**1,25-dihydroxyvitamin D(3) 24-hydroxylase**	9E-62
26	XP_014914825.1	Predicted: Endogenous retrovirus group K member 19 Pol protein-like	2E-61
27	CBN81309.1	Transposable element Tcb1 transposase	3E-51
28	XP_014065326.1	Predicted: Fibroblast growth factor receptor 3-like isoform X1	1E-28
29	AKL71657.1	Hepcidin	8E-27
30	NP_001035458.1	Ankyrin repeat and SOCS box protein 16*	8E-21
31	XP_011492157.1	Predicted: RNA-directed DNA polymerase from mobile element jockey-like	2E-18

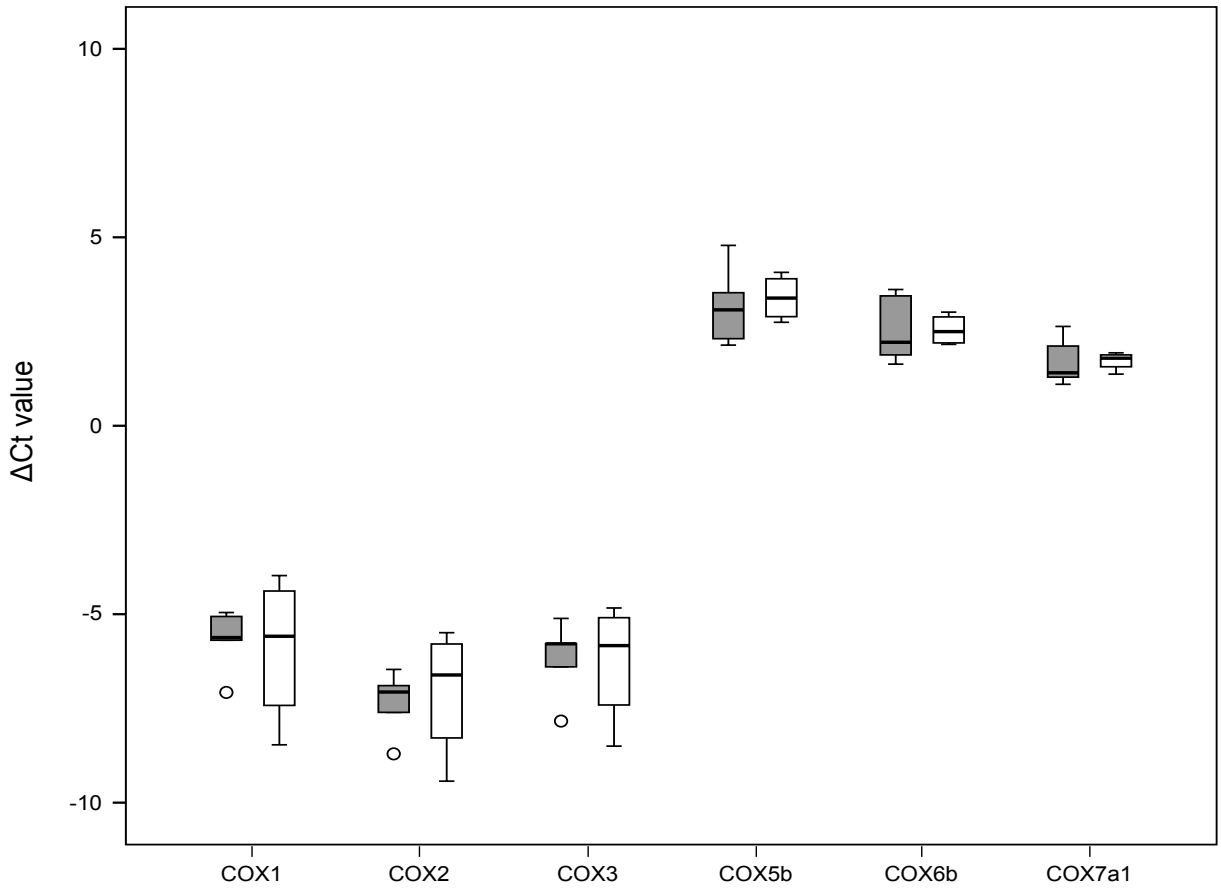

Figure A1: Box plot of the expression level of genes encoding complex IV nuclear and mitochondrial subunits in wild type (grey) and cybrids (white). ΔCt values of COX1, COX2, COX3 COX5b, COX6b and COX7a1 represent the difference between the Ct of the target gene and the housekeeping gene β-actin. Refer to Table 2.2 for sample size.