

Université de Montréal

Impact d'une mitochondrie exogène  
sur le protéome du cybride *Chrosomus eos*

par

Logan Schwartz

Département de sciences biologiques

Faculté des arts et des sciences

Mémoire présenté à la Faculté des arts et des sciences  
en vue de l'obtention du grade de Maître ès Sciences  
en sciences biologiques

© Logan Schwartz, 2014

Université de Montréal  
Faculté des arts et des sciences

Ce mémoire intitulé:  
Impact d'une mitochondrie exogène  
sur le protéome du cybride *Chrosomus eos*

Présenté par :  
Logan Schwartz

a été évalué par un jury composé des personnes suivantes :

Jean-François Pflieger, président-rapporteur

Bernard Angers, directeur de recherche

Annie Angers, codirectrice de recherche

Valérie Legendre-Guillemain, membre du jury

## Résumé

On retrouve dans le complexe *Chrosomus eos-neogaeus* une forme cybride ayant le génome nucléaire de *C. eos* et le génome mitochondrial de *C. neogaeus*. Ce modèle particulier fournit une occasion unique d'étudier l'influence d'une mitochondrie exogène sur le métabolisme et la physiologie d'organismes vivant en milieu naturel, et s'étant donc adaptés à cette situation cellulaire atypique. La mitochondrie jouant un rôle fondamental vital, nous nous attendons à ce que la présence d'une mitochondrie exogène chez la forme cybride ait un impact sur l'expression de son génome et du protéome qui en découle.

L'objectif de ce projet est d'étudier les différences au niveau protéomique entre des individus *C. eos* purs (forme sauvage) et des cybrides provenant d'habitats similaires afin de faire ressortir au maximum les différences dues à la présence de mitochondries *C. neogaeus* chez la forme cybride. Pour ce faire, nous avons comparé les protéomes des formes cybride et sauvage en utilisant l'électrophorèse en deux dimensions. Un sous-groupe de protéines produisant un signal spécifique révélé par l'analyse comparative a été identifié et analysé par spectrométrie de masse (LC/MS).

Les résultats indiquent que la présence de mitochondries *C. neogaeus* chez le cybride influence fortement la régulation génique chez ce dernier. De plus, les protéines identifiées apportent des pistes intéressantes supportant l'hypothèse que la présence de mitochondries *C. neogaeus* chez le cybride rendrait ce biotype plus résistant au froid que la forme sauvage.

**Mots clefs :** cybride, hybride cytoplasmique, protéomique, gel 2D, analyse comparative, *Chrosomus eos-neogaeus*, mitochondrie.

## Abstract

The *Chrosomus eos-neogaeus* genetic complex regroups different forms of hybrids of these two species, among which a cybrid form, that harbours the nuclear genome of *C. eos* and the mitochondria of *C. neogaeus*. This peculiar model is thus a unique opportunity to study the influence of an exogenous mitochondria on the metabolism and cellular physiology in a living animal in the wild, and thus perfectly adapted to this atypical cellular environment. Mitochondria being at the core of fundamental biological processes, we expect that the presence of foreign mitochondria will modify gene expression and the resulting proteome of these fishes.

The overall goal of this master thesis is thus to compare the proteome of pure (wild type) *C. eos* with the cybrid form sampled in similar lakes from the same geographical area so that most differences could be attributed to the different mitochondrial genomes. To achieve this goal, we used two dimensional electrophoresis. We selected a sub-group of proteins that showed the most extreme expression differences and identified these spots by mass spectrometric analyses (LC/MS).

Results demonstrate that *C. neogaeus* mitochondria has a strong influence on gene expression in cybrid. Proteins identified bring new clues supporting the hypothesis that cybrid are more cold tolerant than the wild type biotype.

**Keyword:** cybrid, cytoplasmic hybrid, proteomic, two-dimensional gel electrophoresis, comparative analysis, mitochondria, *Chrosomus eos-neogaeus*.

# Table des matières

Résumé.....	III
Abstract.....	IV
Table des matières.....	V
Liste des figures.....	VIII
Liste des tableaux.....	VIII
Liste des Annexes.....	VIII
Abréviations.....	IX
Remerciements.....	X
<b>1 Introduction.....</b>	<b>1</b>
1.1 Variabilité phénotypique et cause épigénétique.....	1
1.1.1 <i>Le complexe Chrosomus eos-neogaeus</i> .....	1
1.1.2 <i>La plasticité phénotypique</i> .....	2
1.2 Modification de l'expression génique.....	3
1.2.1 <i>Modification du patron de méthylation</i> .....	4
1.2.2 <i>Modifications post-transcriptionnelles</i> .....	5
1.2.3 <i>Modifications post-traductionnelles</i> .....	5
1.3 La mitochondrie.....	7
1.3.1 <i>Origine et caractéristiques de la mitochondrie</i> .....	7
1.3.2 <i>Le complexe d'oxydation phosphorylation</i> .....	8
1.3.3 <i>Les interactions intergénomiques</i> .....	9
1.3.4 <i>Les hybrides cytoplasmiques en général</i> .....	10
1.4 Notre projet.....	12
1.4.1 <i>Modèle d'étude</i> .....	12
1.4.2 <i>Hypothèse et Objectif</i> .....	14
1.4.3 <i>Avantages de l'approche protéomique</i> .....	15
1.4.4 <i>Approche expérimentale</i> .....	16
<b>2 Matériel et méthode.....</b>	<b>17</b>
2.1 Formation de la forme cybride en milieu naturel.....	17
2.2 Prélèvement et génotypage des spécimens.....	18
2.3 Extraction et dosage des protéines.....	20

2.4	Gels 2D.....	21
2.4.1	1ère dimension : réhydratation de la bande et focalisation isoélectrique.....	21
2.4.2	2ème dimension : Gel SDS page.....	21
2.5	Révélation.....	22
2.5.1	Protocole de révélation pour l'analyse comparative des gels .....	22
2.5.2	Protocole de révélation pour le séquençage.....	23
2.6	Analyse des gels 2D .....	23
2.6.1	Comparaison des biotypes.....	23
2.6.2	Diagramme de dispersion.....	26
2.6.3	Sélection des protéines pour la spectrométrie de masse .....	26
2.7	Digestion des bandes et MS/MS.....	27
2.8	Visualisation des résultats de séquençage .....	28
3	Résultats .....	29
3.1	Analyse comparative .....	29
3.1.1	Reproductibilité.....	30
3.1.2	Différence interbiotype .....	31
3.1.3	Diagramme de dispersion.....	32
3.2	Analyse des protéines spécifiques à un biotype .....	34
3.3	Identification des protéines spécifiques.....	37
3.4	Résultat du séquençage .....	38
4	Discussion.....	42
4.1	Reproductibilité des gels 2D .....	43
4.2	Étude comparative du protéome du muscle squelettique .....	44
4.3	Points surexprimés chez un biotype .....	44
4.4	Séquençage : identification des protéines spécifiques à un biotype.....	45
4.4.1	Sous unité alpha de l'énolase1 .....	46
4.4.2	FABP 3.....	47
4.4.3	Succinate déshydrogénase.....	48
4.4.4	Apolipoprotéine AI2 .....	49
4.4.5	Protéines cytosquelettiques .....	50
4.4.6	Triose-phosphate isomérase B .....	51

4.4.7	<i>Guanidinoacétate N-méthyltransférase</i> .....	52
4.5	Protéines spécifiques et thermorégulation.....	54
5	Conclusion.....	57
	Références.....	59
	Annexe 1: Limitations à l'identification des protéines par LC/MS.....	XI

## Liste des figures

Figure 1 : Intégration de l'ensemble des éléments constituant le phénotype .....	3
Figure 2: Complexes constituant le système OXPHOS de la mitochondrie .....	9
Figure 3: Aire de répartition des Cybrides et de la forme sauvage en Amérique du Nord.....	13
Figure 4: Formation des biotypes du complexe <i>Chrosomus eos-neogaeus</i> :.....	18
Figure 5 : Gel à deux dimensions du muscle squelettique de la forme sauvage (A) et de la forme cybride (B).....	30
Figure 6 : Diagramme de Venn représentant la similarité entre le pourcentage volume des points de deux gels.....	32
Figure 7 : Diagrammes de dispersion entre le %Vol des points des gels.....	33
Figure 8 : Localisation des points spécifiques sur les Gels à deux dimensions du muscle squelettique de la forme sauvage (A) et de la forme cybride (B).....	37

## Liste des tableaux

Tableau I : Coordonnées GPS des lacs où les spécimens de la forme cybride et de la forme sauvage ont été prélevés :.....	19
Tableau II: Amorces utilisées pour le génotypage : .....	19
Tableau III : Reproductibilité des gels 2D du protéome du muscle pour la forme sauvage et de la forme cybride : .....	30
Tableau IV : Points exprimés différemment dans le protéome du muscle entre la forme sauvage et de la forme cybride : .....	35
Tableau V : résultats du séquençage des points spécifiques à un biotype ou exprimés différemment : .....	39
Tableau VI : Fonctions et localisations des protéines spécifiques à un biotype ou exprimées différemment : .....	41

## Liste des Annexes

Annexe 1: Limitations à l'identification des protéines par LC/MS.....	XI
---	----



## Abréviations

ApoA	Apolipoprotéine AI2
Eau dd	Eau bidistillée
FABP 3	« <i>Fatty acid binding protein 3</i> »
FDR	« <i>False decoy recovery</i> »
GAMT	Guanidinoacétate N-méthyltransférase
HIF-1	Facteur d'induction hypoxique
HSP	« <i>Heat shock protein</i> »
IPG	Gradient de pH immobilisé
IRIC	Institut de recherche en cancérologie et en immunologie
MLC	Chaîne légère phosphorylable de myosine
MS/MS	Spectrométrie de masse en tandem
SDH	Succinate déshydrogénase
SDS	Dodécylsulfate de sodium
TPI	Triose-phosphate isomérase
%Vol	Pourcentage volume

## Remerciements

Je voudrais remercier mes directeurs Bernard Angers et Annie Angers pour m'avoir donné la chance de vivre cette expérience et pour leur soutien. J'ai appris énormément à votre contact et grâce à votre enseignement. Merci à Hélène Glémet, Jean-François Pflieger et Valérie Legendre-Guillemain d'avoir accepté de prendre de leur temps pour participer à mon comité-conseil et avoir évalué mon mémoire. Merci au FQRNT d'avoir financé ce projet de maîtrise.

J'aimerais aussi remercier mes collègues de travail et avoir une pensée plus particulièrement pour Anne Favier, Denis Beaudet, Florent Renault, Guillaume Desrochers, Jonathan Lord, Leo Deremiens, Maeva Perez, Marc-André Desnoyers et Philippe Lamarre avec qui j'ai partagé les bons moments comme les plus durs survenus au cours de cette maîtrise. C'est grâce à vous et au soutien inconditionnel dont vous avez fait preuve envers moi que j'en suis là aujourd'hui.

Ma dernière pensée est pour ma famille. Aucun mot ne saurait exprimer la dette que j'ai envers vous. Tout ce que j'ai accompli et tout ce que je suis devenu, c'est à vous que je le dois. Pour avoir été à mes côtés chaque jour malgré la distance et les coups durs de la vie, pour tous ces sacrifices que vous avez concédés pour que moi je puisse me réaliser ici, vous resterez à jamais dans mon cœur. Cette réussite personnelle est tout autant la votre.

# 1 Introduction

## 1.1 Variabilité phénotypique et cause épigénétique

### 1.1.1 Le complexe *Chrosomus eos-neogaeus*

Le complexe *Chrosomus eos-neogaeus* (Pisces, Cyprinidae) est issu de plusieurs épisodes d'hybridation ancestraux impliquant deux espèces de cyprinidés d'eau douce, le ventre rouge du nord (*C. eos*) et le ventre citron (*C. neogaeus*). Les différents biotypes du complexe sont ainsi retrouvés vivants en sympatrie ou allopatrie dans la plupart des lacs du nord des États-Unis et du sud du Canada (Scott et Crossman 1973). L'étendue de leur aire de répartition témoigne de la capacité de ces populations à s'adapter à une multitude de conditions environnementales (Massicotte et Angers 2012). Les biotypes du complexe *Chrosomus eos-neogaeus* coexistent tous dans les lacs du sud du Québec et procurent un assortiment d'arrangements génomiques et mitochondriaux uniques en milieu naturel, permettant une étude intégrative de l'impact des différents génomes l'un sur l'autre et sur le phénotype.

On retrouve dans ce complexe une forme de *C. eos* dite « cybride » qui correspond à des individus possédant le génome nucléaire de *C. eos* et mitochondrial de *C. neogaeus*. Bien que la forme cybride soit indissociable visuellement de la forme sauvage (*C. eos* pure), elle possède différentes caractéristiques qui lui sont propres. Les études portant sur ce biotype ont montré que la forme sauvage et la forme cybride sont génétiquement très similaires (Angers , données non publiées), et qu'elles occupent chacune une aire de répartition géographique spécifique à leur forme respective (Mee et Taylor 2012). La capacité de la forme cybride à peupler des niches écologiques inaccessibles à la forme sauvage découlerait de la présence d'une différence au niveau de son patron de méthylation ayant pour conséquence la reprogrammation de son patron d'expression génomique (Angers *et al.*, 2012). L'altération du méthylome induirait ainsi, une variation au niveau du phénotype désigné sous le terme de

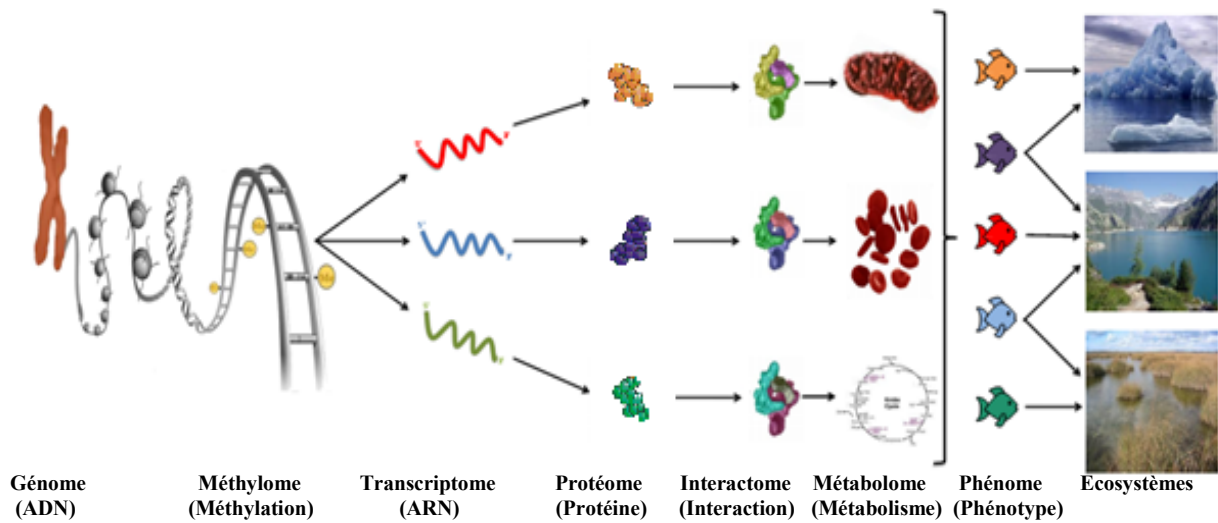
plasticité phénotypique (Roberts et Gavery 2012). Ce mécanisme aurait joué un rôle important au cours de l'évolution en permettant aux espèces de pouvoir s'adapter à un nouvel environnement sans avoir recours à des modifications génomiques (Price *et al.*, 2003).

### **1.1.2 La plasticité phénotypique**

Gustave Flaubert dit un jour : « Le succès est une conséquence et non un but » (Flaubert 1893). Dans un contexte biologique, cette phrase prend tout son sens. L'évolution étant aveugle, le succès d'une espèce à prospérer dans un environnement ne peut être que la conséquence de la capacité d'adaptation de son phénotype face aux différents stress biotiques et abiotiques. La capacité d'adaptation d'un phénotype en réponse à des variations environnementales se définit comme la plasticité phénotypique (Price *et al.*, 2003). Ainsi, la plasticité phénotypique d'une espèce détermine sa capacité à coloniser une aire de répartition géographique et environnementale plus ou moins importante (Queitsch *et al.*, 2000, Massicotte et Angers 2012). Ce point capital représente la frontière départageant les espèces dites endémiques de celles dites généralistes. Le phénotype étant l'intégration de plusieurs niveaux biologiques (figure 1). La variation d'un de ces niveaux aura une conséquence directe sur le phénotype et sa capacité à prospérer dans son environnement (Diz *et al.*, 2012).

Différentes études montrent que bien que la capacité d'une espèce à vivre dans un environnement dépende de son phénotype, celui-ci serait également affecté par l'environnement dans lequel il vit (Lopez *et al.*, 2001, Martinez-Fernandez *et al.*, 2010). Ainsi, l'ensemble des contraintes environnementales imposées à une espèce induit une réponse phénotypique subséquente allant de l'apparition d'une maladie à l'acquisition d'une tolérance accrue à certains environnements (Queitsch *et al.*, 2000). Par exemple, une étude chez la levure a démontré qu'un trait phénotypique apparu en réponse aux variations environnementales pouvait être transmis à la descendance (Halfmann et Lindquist 2010). Ce genre d'études démontre que le phénotype peut être modifié par des mécanismes

transmissibles indépendants de la génétique, désignés sous le terme de mécanismes épigénétiques.



**Figure 1 : Intégration de l'ensemble des éléments constituant le phénotype :**

La capacité d'une espèce à utiliser une niche écologique dépend du phénotype de ses individus. Le phénotype est l'intégration des différents niveaux « omiques » présentés dans la figure 1. La modification de l'un de ces éléments peut mener à une variation du phénotype et avoir un impact sur la capacité de celui-ci à occuper une niche écologique particulière.

Afin de mieux comprendre comment la plasticité phénotypique offre des possibilités complètement distinctes aux deux formes biologiques de *C. eos* malgré qu'elles possèdent le même génome nucléaire, nous allons maintenant passer en revue les différents niveaux biologiques pouvant conduire à la variation phénotypique observée entre la forme cybride et sauvage.

**1.2 Modification de l'expression génique**

Le taux d'expression des gènes peut varier en réponse à plusieurs facteurs, notamment une modification au niveau du patron de méthylation (figure 1), des modifications post-transcriptionnelles ou post-traductionnelles (McGhee et Ginder 1979, ENCODE 2009, Wang *et al.*, 2014). Tandis qu'une modification au niveau du patron de méthylation intervient

directement sur l'expression, les modifications post-transcriptionnelles et post-traductionnelles interviennent respectivement au niveau de l'ARN et des protéines. Comme ces modifications ont la possibilité d'être cumulatives les unes par rapport aux autres, la variation phénotypique qu'elles induisent n'en est que plus importante. La capacité de créer de la diversité sans intervenir sur le contenu du génotype est un outil évolutif formidable. À titre indicatif, le consortium du séquençage du génome humain insiste sur l'importance des mécanismes épigénétiques qui permettraient chez l'Homme de produire un million de protéines différentes à partir de seulement vingt-cinq mille gènes (Human Genome Sequencing 2004). Cet exemple illustre parfaitement le fait qu'une espèce a bien plus de ressources à sa disposition pour répondre aux contraintes de son environnement que le nombre de gènes dont elle dispose. Au-delà de simplement induire de la diversité phénotypique, l'existence de ces mécanismes biologiques est indispensable au bon déroulement de l'embryogenèse et au maintien de la fonctionnalité de chaque type cellulaire durant la vie d'un individu (Li 2002, Wang *et al.*, 2014).

### ***1.2.1 Modification du patron de méthylation***

L'activité transcriptionnelle des gènes peut être modulée par l'ajout ou le retrait de marques de méthylation sur les îlots CpG situés en amont de certains promoteurs. La répartition de l'ensemble des marques de méthylation est spécifique à un type cellulaire et à une espèce. On appelle ce patron le méthylome. Rien qu'avec une simple modification au niveau du méthylome, on peut obtenir différents transcrits d'ARN à partir du même gène (Robertson 2005). Le méthylome est reconnu pour être l'élément déterminant la destinée cellulaire durant l'embryogenèse (Lee *et al.*, 2014). C'est de cette façon que malgré le fait que nos cellules possèdent le même génome, elles n'assurent ou ne produisent ni les mêmes fonctions ni les mêmes organes. On peut considérer le méthylome comme une sorte de bouton de volume disposé sur les gènes modulant leurs taux d'expression.

D'autres mécanismes comme la méthylation et l'acétylation des histones peuvent également réguler l'expression génique. Pour ce faire, ces mécanismes modifient l'état de la chromatine d'une forme condensée à une forme décondensée. La transcription n'étant possible que lorsque la chromatine est décondensée, la méthylation et l'acétylation des histones influencent ainsi l'accessibilité des gènes d'une région aux facteurs de transcription et par conséquent modifient le taux d'expression de ces gènes (Van Lint *et al.*, 1996).

### **1.2.2 Modifications post-transcriptionnelles**

Chez les eucaryotes, l'ARN doit subir une série de modifications dans le but de perdre ses introns et d'acquérir une stabilité. Ces modifications sont reconnues comme des modifications post-transcriptionnelles qui augmentent la diversité des transcrits produits, ainsi que le nombre de fonctions biologiques d'un gène (Black 2003). On distingue deux types de modifications post-transcriptionnelles pouvant modifier le transcrit produit: l'édition de l'ARN et l'épissage alternatif. L'édition de l'ARN correspond à l'addition ou au remplacement de nucléotides par rapport à la séquence d'ADN de départ. L'épissage alternatif permet la production d'une diversité d'ARNm matures chez les eucaryotes. Grâce à l'épissage alternatif, certains gènes sont capables de produire un nombre important d'ARNm différents (Karp 2009). Un exemple extrême est le gène Dscam de la drosophile qui intervient dans la formation des circuits neuronaux et la phagocytose des pathogènes, et est capable à lui seul de produire trente-huit mille ARNm par l'épissage alternatif d'exons variables contenus dans quatre clusters génomiques (Guilherme *et al.*, 2004). Le nombre d'isoformes pour ce gène est tellement important que chaque neurone de la mouche expose sur sa membrane externe un isoforme protéique unique.

### **1.2.3 Modifications post-traductionnelles**

Les modifications post-traductionnelles correspondent aux modifications affectant les acides aminés formant les protéines. Il est estimé qu'environ 5% du protéome effectuerait plus de

deux cents types de modifications post-traductionnelles (Walsh 2006). Ces modifications sont indispensables au bon déroulement de l'activité cellulaire et ont lieu à différents moments du cycle de vie d'une protéine. Par exemple, les modifications permettant un bon repliement, une bonne stabilité ou simplement impliquées dans le transport de la protéine vont se produire très probablement peu de temps après sa traduction, alors que d'autres modifications visant à modifier l'activité catalytique ou marquer la protéine pour qu'elle puisse être dégradée par le protéasome se produiront plus tard (Glickman et Ciechanover 2002, Mukhopadhyay et Riezman 2007). En plus des modifications induites par des acteurs cellulaires, il faut garder à l'esprit que la population de protéine contenue dans une cellule est très sensible aux variations environnementales.

Une étude, chez la levure *Saccharomyces cerevisiae*, montre que l'altération de la structure tridimensionnelle d'une protéine (forme prion) peut modifier les traits phénotypiques observables tels que la couleur ou la morphologie en altérant la fidélité du mécanisme de terminaison de la traduction. Ainsi, la présence du prion permet à la cellule d'augmenter sa tolérance à des environnements initialement hostiles en révélant une variabilité génétique accessible seulement dans des conditions extrêmes (True et Lindquist 2000). Une autre étude plus extrême a montré que la capacité de la plante *Arabidopsis thaliana* à supporter l'exposition à la chaleur dépend du niveau d'expression de certaines protéines *heat shock* (HSP). L'exposition à 45°C pendant deux heures induit la perte de structure tridimensionnelle d'une grande portion des protéines et, par conséquent, de leurs fonctions biologiques. Alors que les plants normaux récupèrent cependant très bien de ce traitement, la dénaturation des protéines conduit généralement à la mort cellulaire dans les plants où l'expression de HSP 101 est inhibée. La viabilité cellulaire est ainsi remise en question dans la mesure où il ne reste plus suffisamment de protéines HSP101 pour compenser les modifications de la structure tridimensionnelle des protéines dénaturées (Queitsch *et al.*, 2000, Queitsch *et al.*, 2002).



Ces exemples montrent que le phénotype est tout autant sensible à la présence/l'absence d'une protéine qu'aux modifications pouvant altérer sa structure. L'intégration de ces traits phénotypiques aurait joué un rôle important au cours de l'évolution en augmentant le taux de survie de certains phénotypes suite à une variation environnementale brutale (Lindquist 2009).

### 1.3 La mitochondrie

Comme la présence de mitochondrie *C. neogaeus* chez la forme cybride est le seul paramètre cellulaire la distinguant de la forme sauvage. Il est intéressant de comprendre les circonstances ayant permis la cohabitation de ces deux génomes étrangers au sein d'une même cellule, ainsi que les fonctions biologiques qui sont associés à la mitochondrie.

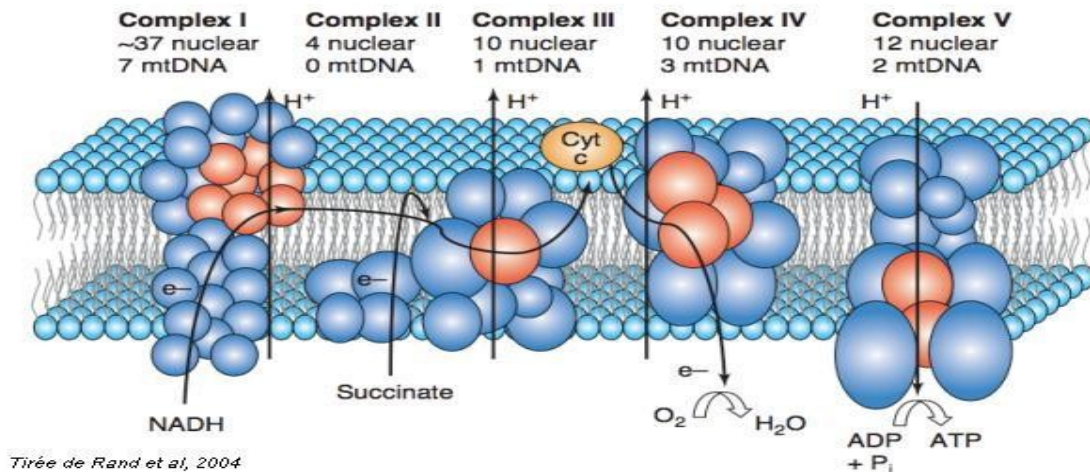
Finalement, nous discuterons de l'importance des interactions intergénomiques existant entre la mitochondrie et le noyau et des conséquences associées à l'altération de ces interactions.

#### *1.3.1 Origine et caractéristiques de la mitochondrie*

Les mitochondries assument principalement le rôle de "centrale énergétique" des cellules eucaryotes car elles contiennent et coordonnent l'activité de la plupart des enzymes métaboliques permettant la formation d'ATP. L'ATP est indispensable à l'exécution de nombreuses réactions métaboliques telles que le transport actif ou la formation de l'acétylcholine (Nachmansohn et Machado 1943, Lodish *et al.*, 2000). Mais le rôle des mitochondries à l'intérieur de la cellule ne se limite pas à la production d'énergie. Elle serait également associée à l'apoptose, l'oxydation des lipides, la glucogenèse, la stéroïdogénèse, la synthèse de l'hème et le cycle de l'urée (Schatz 1995, Dimmer et Scorrano 2006). Mis à part quelques rares exceptions, la mitochondrie est transmise de façon maternelle. La présence de la mitochondrie au sein de nos cellules serait le résultat d'une endosymbiose entre une alpha-protéobactérie et une cellule procaryote ayant eu lieu il y a 2 milliards d'années (Andersson *et al.*, 1998, Lang *et al.*, 1999, Gray *et al.*, 2001).

### 1.3.2 Le complexe d'oxydation phosphorylation

Bien que le génome de la mitochondrie ne code que treize protéines, son rôle reste primordial à la viabilité cellulaire. En effet, les treize protéines encore codées par le génome mitochondrial participent à la formation des complexes I, III, IV et V du système d'oxydation phosphorylation (OXPHOS) située dans la membrane interne de la mitochondrie (figure 2). Ce complexe est d'une importance capitale, puisqu'il est responsable de la production de la majorité de l'ATP cellulaire des cellules eucaryotes. Étant donné que les treize protéines codées par la mitochondrie participent toutes à la formation du système OXPHOS, ce système est un outil d'étude parfait pour visualiser les interactions intergénomiques (McKenzie et Trounce 2000). Dans ce genre d'expériences, la présence de protéines de deux génomes différents implique une coadaptation des protéines mitochondriales provenant du noyau et de la mitochondrie afin de maintenir la production d'ATP et la température interne (Rand *et al.*, 2004, Gershoni *et al.*, 2009). Dans la mesure où les protéines seraient incapables d'interagir efficacement, la baisse de l'activité enzymatique qui en découle pourra conduire à une baisse du « *fitness* » (DiMauro et Schon 2003). Dans le cadre de notre projet, ce système est d'autant plus intéressant qu'il est très probablement altéré à cause de la formation de complexes protéiques chimériques induite par la présence de protéines mitochondriales provenant d'une autre espèce.



**Figure 2: Complexes constituant le système OXPHOS de la mitochondrie :**

Les protéines mitochondriales provenant de la mitochondrie sont représentées en rouge et celles provenant du noyau en bleu. Le complexe II est parfois utilisé comme contrôle dans les expériences intergénomiques du fait de l'absence de protéines mitochondriales. Tirée de Rand *et al.*, (2004).

**1.3.3 Les interactions intergénomiques**

Une étude a estimé que le nombre total de protéines constituant une mitochondrie s'élevait à environ mille cinq cents alors que le génome de la mitochondrie ne code que treize protéines (Lotz *et al.*, 2014). Ces éléments indiquent qu'il y aurait eu au cours de l'évolution une succession de transferts horizontaux. L'existence de ce mécanisme de transfert horizontal a été confirmée à maintes reprises, il a même été démontré que les génomes nucléaires présentent une facilité à recevoir et à perdre de l'ADN provenant d'organelles (Thorsness et Weber 1996, Huang *et al.*, 2003, Timmis *et al.*, 2004, Leister 2005).

Le fait que la majorité des protéines mitochondriales ne soit pas produite par le génome mitochondrial implique la mise en place d'un système de communication entre les génomes nucléaire et mitochondrial afin de combler les besoins mitochondriaux et maintenir l'activité mitochondriale. De cette façon, la mitochondrie peut tout de même avoir un contrôle indirect sur l'expression des protéines mitochondriales codées par le génome nucléaire, dont elle a besoin pour se reproduire ou maintenir/augmenter certaines voies métaboliques pour répondre

efficacement à un stimulus. Étant donné que les treize protéines mitochondriales codées par le génome mitochondrial rentrent dans la formation du système OXPHOS, la mitochondrie doit emprunter d'autres voies rétrogrades pour communiquer avec le génome nucléaire s'appuyant entre autres (mais non exclusivement) sur l'UPR (« *unfolded protein response*») ou un relâchement cytosolique de  $\text{Ca}^{2+}$  (Amuthan *et al.*, 2002, Ryan et Hoogenraad 2007, Haynes et Ron 2010). Un tel système résultant de la coévolution de deux génomes contraints à coopérer est relativement fragile dans la mesure où une légère altération du génome mitochondrial où la présence de mitochondries d'une autre espèce peut conduire à l'effondrement de l'activité mitochondriale (McKenzie et Trounce 2000, Mochel et Haller 2011).

Parallèlement, certaines altérations peuvent également induire un effet bénéfique, être sélectionnées et transmises aux générations suivantes. Ainsi, certaines mutations au niveau du génome mitochondrial auraient pour conséquence de diminuer la capacité de couplage des protéines au niveau des complexes chimériques du système OXPHOS, cette baisse du couplage interprotéiques induit une augmentation du relâchement de proton associée à un dégagement de chaleur par la mitochondrie (Rousset *et al.*, 2004). Il n'est pas surprenant de constater que ces mutations mitochondriales ont parfois été sélectionnées dans certaines populations de l'arctique (Ruiz-Pesini *et al.*, 2004). Cet exemple démontre que la balance au niveau de la quantité des protons relâchés et la quantité d'ATP produite semble avoir été un paramètre important permettant l'adaptation d'une espèce à son environnement (Speakman *et al.*, 2004).

#### **1.3.4 Les hybrides cytoplasmiques en général**

Le meilleur moyen connu à ce jour pour étudier les interactions intergénomiques est l'utilisation de cybrides artificiels. De telles cellules peuvent être obtenues par la fusion cytoplasmique de cellules dépourvues d'ADN mitochondrial avec des cellules énucléées

provenant d'une autre espèce (McKenzie *et al.*, 2003). Avec un tel système, McKenzie *et al.*, (2003) ont démontré que l'activité du système OXPHOS chez des cybrides artificiels est inversement proportionnelle à la divergence évolutive entre le génome nucléaire et mitochondrial des cellules cybrides. À partir d'un certain degré de divergence, l'activité du complexe est tellement diminuée que la majorité de la production d'ATP provient de la fermentation. Ce faisant, le milieu cellulaire s'acidifie suite à l'accumulation d'acide lactique (sous-produit de la fermentation) et la viabilité cellulaire de ces cellules est alors réduite. Bien que l'activité du complexe OXPHOS dans les cellules cybrides obtenues en laboratoire soit compromise, il est peu probable qu'en milieu naturel les formes cybrides soient aux prises avec une chaîne respiratoire déficiente. À contrario, Mee et Taylor, (2012) ont suggéré qu'au contraire de nouveaux traits phénotypiques avantageux pourraient découler de la présence d'une mitochondrie provenant d'une espèce étrangère.

Un excellent exemple est celui de l'omble de fontaine *Salvelinus fontinalis*. Chez cette espèce, on retrouve des formes cybrides issues de l'introgression de mitochondries provenant de *Salvelinus alpinus* (Glémet *et al.*, 1998). L'introgression mitochondriale fait référence au transfert interspécifique du génome mitochondrial découlant d'une série de rétrocroisement entre un hybride et l'une de ses espèces parentales. Sur le plan géographique, ces deux espèces vivent en allopatrie. *S. alpinus* est retrouvé dans les eaux arctiques et subarctiques alors que *S. fontinalis* est retrouvé dans des environnements plus chauds. L'existence d'une forme cybride chez cette espèce implique un événement d'hybridation ayant eu lieu à l'époque où les espèces parentales ont vécu en sympatrie. Les cybrides sont apparus suite au croisement répété d'hybrides *S. fontinalis* x *S. alpinus* avec *S. fontinalis*, diluant progressivement le génome nucléaire, mais conservant la mitochondrie *S. alpinus* intacte. En laboratoire, l'étude des performances respiratoires de ces poissons a montré qu'à 6°C, la forme cybride de *S. fontinalis* avait une activité métabolique plus forte que les individus

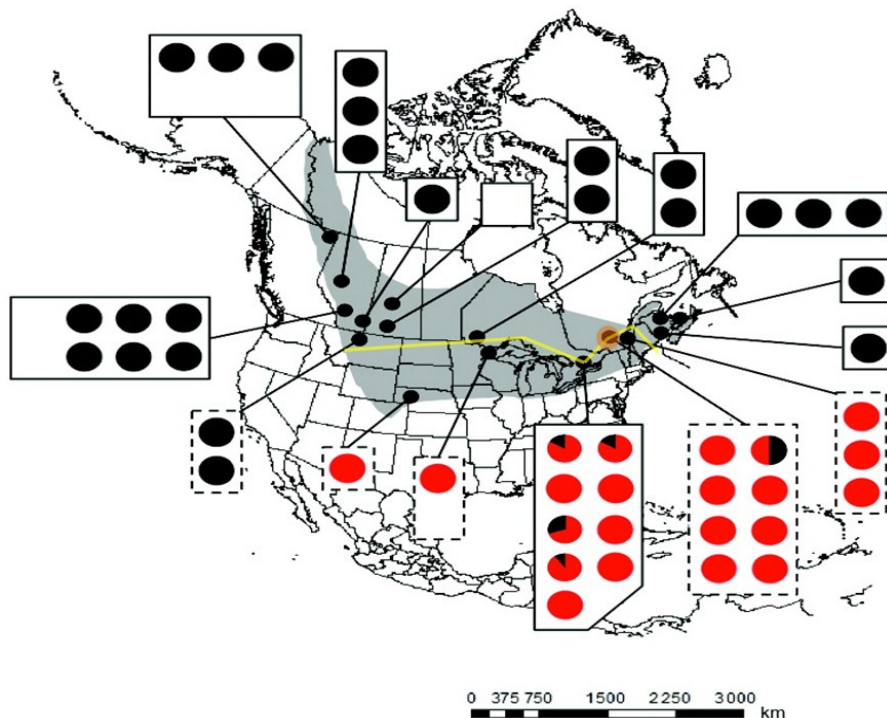
*S. fontinalis* sauvages (Glémet *et al.*, 1998). Les individus appartenant à la forme cybride de *S. alpinus* possèdent des mitochondries *S. alpinus* et sont génétiquement très proches de la forme sauvage bien que certains gènes nucléaires de *S. alpinus* aient pu être conservés. La différence métabolique observable entre le cybride et l'espèce sauvage serait donc principalement due à la présence de la mitochondrie exogène de *S. alpinus*. Cette particularité cellulaire aurait ainsi conféré à la forme cybride de la possibilité de coloniser des niches environnementales inaccessibles à l'espèce sauvage.

Le fait qu'il existe des formes cybrides produites de façon naturelle suggère que ces individus ne souffrent pas d'altération de l'activité enzymatique du système OXPHOS contrairement aux cybrides artificiels de McKenzie *et al.*, (2003). Cet élément soulève l'existence d'une coévolution entre le génome nucléaire et mitochondrial permettant le maintien de la viabilité cellulaire chez la forme cybride, tout en atténuant les effets néfastes liés à l'origine divergente des deux composantes cellulaires (McKenzie *et al.*, 2003, Rand *et al.*, 2004).

## 1.4 Notre projet

### 1.4.1 Modèle d'étude

Dans le cadre de notre projet, nous tirons avantage de la présence d'une forme cybride formée naturellement au sein du complexe *Chrosomus eos-neogaeus*. On retrouve donc chez ce biotype, l'ADN nucléaire de *C. eos* et l'ADN mitochondrial d'une lignée de *C. neogaeus* originaire d'un refuge glaciaire du Mississippi (Goddard et Schultz 1993, Binet et Angers 2005, Mee et Taylor 2012).



**Figure 3: Aire de répartition des Cybrides et de la forme sauvage en Amérique du Nord :** La localisation de la forme cybride est représentée par des cercles noirs et celle de la forme sauvage par des cercles rouges. La ligne jaune représente la frontière entre les aires de répartition des deux biotypes. Le point orange sur la carte correspond à notre zone d'échantillonnage. Adaptée de Mee et Taylor (2012), Scott et Crossman (1979), Angers et Schlosser (2007).

L'aire de répartition du cybride et de la forme sauvage se chevauche au niveau du sud du Québec et des Laurentides, où nous effectuons nos prélèvements (point orange, figure 3). Cependant, ces deux biotypes vivent globalement en allopatrie avec une aire de répartition spécifique à chaque biotype. Tandis que la forme sauvage est retrouvée dans la partie sud, le cybride est concentré dans la partie nordique de l'aire de répartition géographique (Angers et Schlosser 2007, Mee et Taylor 2012). Les résultats amassés sur la forme cybride de *C. eos* suggèrent que la mitochondrie des cybrides leur permet d'être mieux adaptés aux environnements nordiques et qu'elle leur aurait fourni un avantage pour la colonisation de ces environnements par rapport à l'espèce sauvage. La mitochondrie jouant un rôle déterminant

au niveau de la balance énergétique des cellules, l'étude comparative des cybrides avec l'espèce parentale fournit une occasion unique de mesurer l'impact du génome de cette organelle sur le protéome des individus.

#### ***1.4.2 Hypothèse et Objectif***

Bien qu'il existe une vaste quantité d'études épigénétiques, très peu ont étudié l'impact de l'environnement sur le phénotype à l'échelle cellulaire. Notre laboratoire a été à ma connaissance le premier à démontrer que les composantes cellulaires représentaient un microenvironnement tout autant capable d'affecter le phénotype que l'environnement écologique. L'article d'Angers *et al.*, (2012) a déjà démontré que la présence de la mitochondrie exogène chez la forme cybride du complexe était associée à une modification de son méthylome. Cette altération méthylomique affecterait le protéome du cybride, induisant une différence interbiotype de 8,7% pour le protéome du muscle (Angers *et al.*, 2012). La reprogrammation du méthylome observable chez la forme cybride du complexe *Chrosomus eos-neogaeus* aurait pour conséquence de modifier autant le taux d'expression, que les gènes exprimés. De cette façon, une portion des protéines de chacun des biotypes leur serait spécifique et les protéines communes seraient régulées différemment pour chaque biotype. Le but de cette reprogrammation serait probablement d'accommoder la mitochondrie étrangère en lui permettant d'interagir avec les protéines d'une espèce différente pour pouvoir maintenir une activité cellulaire viable.

Nos objectifs dans le cadre de cette maîtrise étaient dans un premier temps de confirmer la différence interbiotype au niveau protéomique, puis de faire une analyse quantitative des protéines entre les deux biotypes, tout en tentant d'identifier certaines des protéines dont l'expression serait spécifique ou modifiée de façon importante chez un biotype.



### 1.4.3 Avantages de l'approche protéomique

Bien que beaucoup plus complexe à réaliser, l'étude des protéomes présente de nombreux avantages sur les études génomiques ou transcriptomiques lorsque l'on s'intéresse à la variabilité phénotypique et à l'écologie d'une population (Karr 2008). L'expression des gènes n'aboutissant pas au niveau transcriptomique, la compréhension globale des mécanismes et des acteurs biologiques participant à la détermination du phénotype implique l'étude de tous les niveaux biologiques jusqu'à leur produit final : le protéome (Feder et Walser 2005). C'est pour cette raison que de nombreuses études insistent sur l'intérêt d'avoir une approche expérimentale combinant la génomique, la transcriptomique et la protéomique lors d'études phénotypiques (Biron *et al.*, 2005, Aubin-Horth et Renn 2009). Le principal intérêt est d'obtenir des informations globales nous permettant de visualiser autant les causes que les conséquences de la variation phénotypique.

Par exemple, une étude chez une souche de *Pseudomonas fluorescens* a montrée à partir d'une approche génomique que la substitution d'une seule paire de bases dans le génome a permis à une souche de coloniser une nouvelle niche écologique inaccessible à l'espèce ancestrale (Knight *et al.* 2006). En plus d'analyser le génome, ces auteurs ont aussi étudié le protéome de ces bactéries et ont pu montrer que cette souche exprime différemment 46 protéines par comparaison avec la souche sauvage (Knight *et al.*, 2006). Bien que cette différence découle de la substitution d'une paire de bases dans le génome, cet exemple démontre bien que l'information récoltée avec l'approche protéomique est plus informative que l'approche génomique pour comprendre les mécanismes et les acteurs biologiques induisant l'acquisition du phénotype.

Même si d'un point de vue évolutif, le génome détermine la capacité d'une espèce à occuper une niche écologique, c'est la sélection naturelle qui agit sur le phénotype découlant de l'expression de ce génome et qui permet aux individus de surmonter les différents stress

biotiques et abiotiques (Martinez-Fernandez *et al.*, 2010, Mee et Taylor 2012). Le fait que le protéome varie de façon importante en réponse aux stimuli environnementaux démontre qu'il est plus flexible que le génome dont il découle et peut permettre une adaptation indépendante des changements génomiques (Knight *et al.*, 2006). De plus, il a été souvent observé que l'abondance des ARNm est plutôt faiblement corrélée avec celle des protéines qu'ils codent, ce qui suggère encore une fois que l'étude du protéome permet de mieux comprendre les phénotypes observés (Gygi *et al.*, 1999, Nie *et al.*, 2006, Rossignol *et al.*, 2009, Zhong *et al.*, 2009, Makowsky *et al.*, 2011). Dans ce contexte, l'approche protéomique est justifiée par le fait qu'elle permet autant de visualiser les informations génétiques issues des processus évolutifs que les conséquences des mécanismes épigénétiques (Danchin *et al.*, 2011).

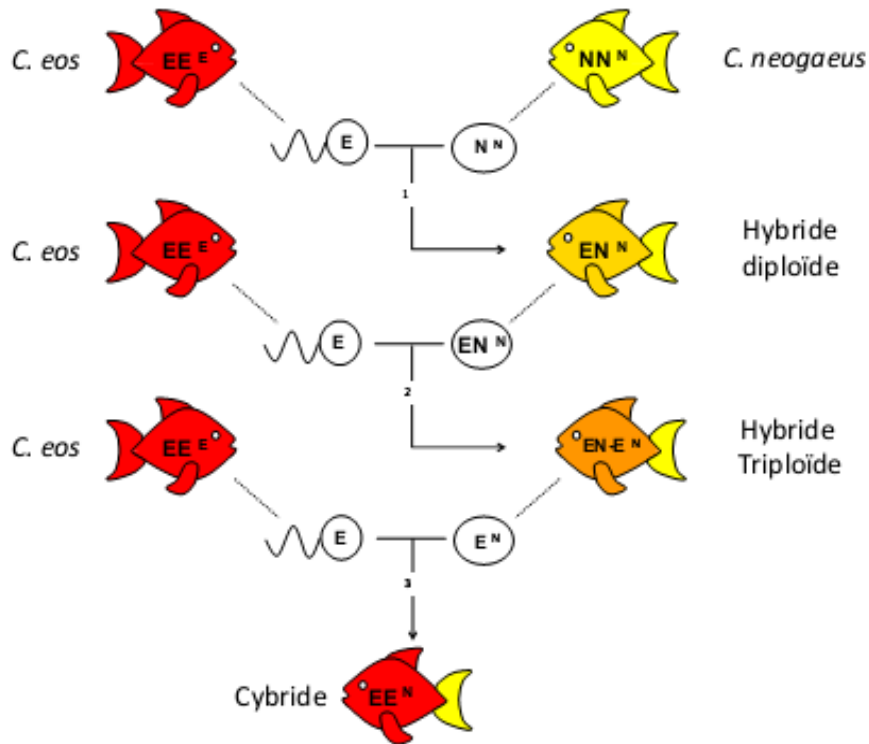
#### ***1.4.4 Approche expérimentale***

Pour mener à bien notre projet, nous avons tout d'abord extrait les protéines contenues dans les muscles latéraux des poissons. Nous avons ensuite produit des gels 2D afin d'isoler les protéines musculaires selon la technique de Kenrick et Margolis (1970). Les protéines ressortant de cette analyse comparative comme étant spécifiques ou surexprimées chez un biotype ont été extraites et envoyées à séquencer. Le séquençage des protéines a été effectué via l'utilisation combinée de la chromatographie en phase liquide et de la spectrométrie de masse (LC/MS). La comparaison des ratios masse/charge à une banque de données a permis de détecter les peptides présents dans l'échantillon. Ceux-ci ont ensuite été comparés à une banque de données transcriptomique partielle pour identifier la où les protéines présentes dans le point envoyé initialement.

## 2 Matériel et méthode

### 2.1 Formation de la forme cybride en milieu naturel

La formation de la forme cybride du complexe *C. eos-neogaeus* repose sur une série de croisements découlant de l'hybridation des deux espèces parentales (figure 4). Le croisement d'un mâle *C. eos* et d'une femelle *C. neogaeus* produit des femelles se reproduisant par gynogenèse, lorsque le développement de leurs œufs diploïdes (non-réduits) est stimulé par un spermatozoïde d'une ou l'autre espèce parentale (Goddard *et al.*, 1998). Il peut arriver cependant que durant ce processus, le génome du spermatozoïde d'un mâle *C. eos* soit intégré à l'ovule, produisant des hybrides triploïdes (*C. eos-neogaeus* x *C. eos*). Lors de la gamétogenèse des hybrides triploïdes, le chromosome n'ayant pas d'homologue est dégradé, et les cellules diploïdes forment des ovules haploïdes (Dawley *et al.*, 1987, Goddard *et al.*, 1998). Les ovules d'une telle femelle contiendront donc le génome nucléaire de *C. eos* et des mitochondries de type *C. neogaeus*. Fertilisés par un mâle *C. eos*, ils engendreront un hybride cytoplasmique (cybride) possédant le génome nucléaire de *C. eos* et le génome mitochondrial de *C. neogaeus* (Goddard et Schultz 1993).



**Figure 4: Formation des biotypes du complexe *Chrosomus eos-neogaeus*:**

E et N réfèrent respectivement aux génomes nucléaires de *C. eos* (rouge) et de *C. neogaeus* (jaune). Les exposants <sup>E</sup> et <sup>N</sup> représente l'identité du génome mitochondrial de *C. eos* (rouge) et de *C. neogaeus* (jaune). Tirée d'Angers *et al.*, (2012).

## 2.2 Prélèvement et génotypage des spécimens

La région utilisée pour l'échantillonnage de nos spécimens est représentée à la figure 3 sous la forme d'un point orange. Les spécimens expérimentaux ont été capturés durant la période de reproduction des étés 2010, 2011 et 2012 sur les bords des lacs Beaver, Desjardins et Triton de la région des Laurentides (tableau I). De l'eugénol a servi à anesthésier les spécimens. Ceux-ci ont ensuite été euthanasiés puis placés sur la glace sèche afin d'abaisser leur température corporelle et d'inhiber ainsi les protéases, puis stockés à -80°C jusqu'à leur prochaine utilisation.

**Tableau I : Coordonnées GPS des lacs où les spécimens de la forme cybride et de la forme sauvage ont été prélevés :**

<u>Lac :</u>	<u>Latitude :</u>	<u>Longitude :</u>
Triton (forme sauvage)	46°,59',14"	74°, 00',30 "
Beaver (forme cybride)	45°, 55', 29.3"	74° ,3' ,52.3"
Desjardins (forme cybride)	45°, 54', 58.7"	74°, 4', 24.9"

Afin de discriminer les biotypes, un génotypage a été effectué à partir d'une portion de la nageoire caudale prélevée sur chaque individu. En premier lieu, l'ADN est extrait à partir d'un protocole standard au phénol/chloroforme (Sambrook *et al.*, 1989). On amplifie ensuite l'ADN par PCR en utilisant les marqueurs moléculaires mitochondriaux et nucléaires développés par Binet et Angers (2005) (tableau II).

**Tableau II: Amorces utilisées pour le génotypage :**

Amorces	Séquence (5'-3')
Zpeg amorce sens	CCTACGTCAAGCTAT
Zpeg amorce anti-sens	AGCGTTGATTGAGAGAATCC
D-loop : Cr1-eos amorce sens	CCTAGATATGGCTCCCAACAG
D-loop : Cr3c-eos amorce anti-sens	ATAATGCGGCTAACCC
D-loop : Cr1-neogaeus amorce sens	CTCAACTTTTTCTTGACATAC
D-loop : CR3c-neogaeus amorce anti-sens	ATAATGCGGATGGCTAACGG

On utilise ensuite une enzyme de restriction HinfI sur les amplifications faites à partir des amorces mitochondriales. Finalement, on fait migrer les extraits de l'amplification sur gel d'agarose afin de pouvoir identifier l'origine du génome nucléaire et mitochondrial des spécimens. À savoir que les individus présentant le génome nucléaire (220 paires de bases avec les amorces Zpeg) et mitochondrial (2 bandes avec les amorces D-loop) de *C. eos* sont assimilés à la forme sauvage, et ceux possédant le génome nucléaire (220 paires de bases avec les amorces Zpeg) de *C. eos* et mitochondrial de *C. neogaeus* (1 bande avec les amorces D-loop) sont reconnus comme la forme cybride.

### 2.3 Extraction et dosage des protéines

Des échantillons congelés de muscle squelettique ont été broyés. Un volume de 1:5 de solution d'extraction (urée 8M, triton X100 4%) est ajouté. Le tout est agité 20 minutes à 4°C, puis centrifugé 5 minutes à 1000 g. Le surnageant est ensuite transvasé, puis centrifugé à 70 000 g pendant 15 minutes. Enfin, on mélange les extraits protéiques de quatre individus par biotype et par gel afin de minimiser les différences individuelles pouvant avoir un impact sur les analyses des gels 2D. La concentration en protéine de chaque échantillon est déterminée grâce au dosage de Bradford (1976). Ainsi on commence par créer une courbe étalon à partir d'une série de triplicata de BSA (« *Bovin serum albumin* ») dont les quantités protéiques sont connues (0 µg ; 0,3 µg ; 0,6 µg ; 0,9 µg ; 1,5 µg ; 1,8 µg ; 2,4µg et 5µg). Les échantillons à tester sont dilués en triplicata de 1 µL dans deux concentrations différentes (1/10 et 1/100), puis un volume de 250 µL de réactif de Bradford est ajouté. Au contact des protéines, le réactif vire à un bleu dont l'intensité reflète la quantité de protéines dans l'échantillon. Le tout est finalement placé dans un spectrophotomètre « µquant » (Bio-tech Instruments Inc.), puis analysé avec le programme « Kc Junior » (Fisher Scientific). La concentration en protéines peut alors être estimée en rapportant la densité optique de chaque échantillon à la concentration en protéines correspondante dans la courbe étalon.

## 2.4 Gels 2D

### 2.4.1 1ère dimension : réhydratation de la bande et focalisation isoélectrique

Un volume de 250 µL de solution de réhydratation « DeStreak » contenant 15 µL de tampon IPG pH 4-7 (Amersham Bioscience) est constitué afin d'avoir une concentration en protéine de 1 µg/µL. La solution est par la suite placée dans une cuvette de porcelaine. Une bande de gel « *Immobiline Dry Strip* » de 13cm possédant un gradient de pH de 4 à 7 est déposée dans la cuvette (GE 17-6001-13) ainsi qu'une couche d'huile minérale d'environ 1ml. On procède alors à la focalisation isoélectrique dans un appareil « *Ettan IPG phorIEF System* » (Amersham pharmaciabiotec). La bande est d'abord réhydratée pendant 15 heures à 30 Volts afin de favoriser l'entrée des protéines de haut poids moléculaire (Goerg *et al.*, 2004). La focalisation débute automatiquement après la réhydratation de la bande en suivant la série d'étapes suivantes : une heure à 500 Volts, une heure à 1000 Volts, et enfin cinq heures à 8000 Volts.

### 2.4.2 2ème dimension : Gel SDS page

Après la focalisation isoélectrique, on soumet les bandes à deux étapes d'équilibration afin de briser les ponts disulfures et d'empêcher leur reformation. La première étape consiste à faire agiter pendant 20 minutes la bande avec 10 ml de solution d'équilibration (50mM de Tris-HCl pH 8.5, 6 M d'urée, 30% du volume en poids de glycérol, 2% de SDS, 0,002% de bleu de bromophénol, H<sub>2</sub>O) additionnés de 100 mg de dithiothréitol (Fisher Scientific). Pour la seconde étape, on place la bande dans 10 ml de solution d'équilibration et 250 mg d'iodoacétamide (Sigma-Aldrich) et on fait agiter de nouveau l'échantillon durant 20 minutes. On place ensuite la bande au sommet d'un gel dénaturant de 12,5% de polyacrylamide (18cm x 16cm). La bande est fixée dans l'espace entre les vitres avec une solution scellante (100ml de tampon de migration SDS, 0,5% d'agarose, 0,002% bleu de bromophénol). Après avoir

disposé le montage dans la cuve à électrophorèse, la migration est lancée pour recevoir 900 Vh.

## 2.5 Révélation

Dans le cadre de notre projet, 2 protocoles de révélation ont été utilisés. Le premier surexpose les gels à chaque solution et produit des gels ayant un contraste prononcé. Cependant, le fait d'avoir une meilleure visibilité engendre une contrepartie importante. La surexposition entraîne une modification chimique des protéines les rendant impropres à l'analyse par spectrométrie de masse. Le deuxième protocole possède des étapes d'exposition plus courtes afin de minimiser la surexposition au prix d'un contraste beaucoup moins important. On a donc utilisé le premier protocole pour procéder à l'analyse comparative des gels et le second pour prélever les protéines analysées par MS/MS.

### *2.5.1 Protocole de révélation pour l'analyse comparative des gels*

Une fois la migration du gel 2D terminée, le gel est plongé durant deux heures dans une solution de fixation (150 ml d'eau dd, 75 ml d'éthanol, 25 ml d'acide acétique glaciale), puis deux heures dans une solution de sensibilisation (75 ml d'éthanol, 10 ml de thiosulfate de sodium à 5% de w/v, 17 g d'acétate de sodium, eau dd). Le gel est ensuite rincé cinq fois d'affilée dans une eau dd pendant quinze minutes afin de retirer l'excédant de thiosulfate de sodium. Puis, il est plongé deux heures dans une solution de nitrate d'argent (250 ml d'eau dd, 0,625 g de nitrate d'argent). On rince ensuite le gel deux fois dans de l'eau dd durant une minute, puis il est placé dans une solution de développement (6,25 g de carbonate de sodium, 250 ml d'eau dd) jusqu'à l'obtention du contraste désiré. Finalement, on change la solution de développement pour une solution-stop (3,65 g d'acide éthylène diamine tétraacétique, 250 ml d'eau dd) dans laquelle on laisse le gel deux heures pour arrêter la révélation. Une fois la révélation terminée, le gel est numérisé et entreposé à 4°C dans une solution d'acide acétique (1%).



### 2.5.2 Protocole de révélation pour le séquençage

Le gel est plongé dans une solution de fixation 2 heures (50% de méthanol, 12% d'acide acétique, 0.05% de formaldéhyde). Il est lavé trois fois 20 minutes avec une solution d'éthanol (35%), puis placé dans une solution de sensibilisation (0.02% de thiosulfate de sodium) durant 2 minutes. On le rince ensuite trois fois durant 5 minutes avec de l'eau dd. Le gel est plongé dans une solution de coloration (0.2% de nitrate d'argent, 0.075% de formaldéhyde) pendant 20 minutes, puis rincée deux fois durant 1 minute dans de l'eau dd pour retirer l'excès de nitrate d'argent. On le met ensuite dans une solution de développement (6% de carbonate de sodium, 0.05% de formaldéhyde, 0.0004% de thiosulfate de sodium) jusqu'à ce que le maximum de points devienne visible. Normalement, cette étape n'excède pas 5 minutes. On rince ensuite les gels avec une solution-stop (50% de méthanol, 12% d'acide acétique) pendant 5 minutes nécessaires à la neutralisation de la solution de développement. Le gel est entreposé à 4°C dans une solution d'acide acétique (1%) jusqu'à l'extraction des points.

## 2.6 Analyse des gels 2D

### 2.6.1 Comparaison des biotypes

Afin d'avoir un nombre de répliqua suffisant, les trois gels provenant de la forme sauvage et les trois gels provenant de la forme cybride possédant la meilleure résolution et le plus grand nombre de points ont été comparés avec le programme Mélanie 7 (Genebio). Ce programme a la particularité de digitaliser une image en pixels caractérisés par une coordonnée x, y ainsi qu'un niveau d'intensité. À cause d'un temps de révélation variable, les gels présentaient un bruit de fond différent. Par conséquent, les paramètres permettant la détection automatique des points ont été sélectionnés de façon à être spécifiques pour chacun des gels. Le seuil de détection a été sélectionné afin d'être supérieur au bruit de fond. Cette manœuvre a également eu pour conséquence de ne pas prendre en compte certains points dont l'intensité était

inférieure ou similaire à celle du bruit de fond. Dans ces cas, une transposition de la zone de détection a été effectuée à partir du point correspondant d'un autre gel appartenant au même biotype.

Dans le but d'assigner une correspondance entre les points des différents gels, 5 points non ambigus, de petite taille et présents sur l'ensemble des gels ont été désignés comme référence. À partir de ces points, le programme a pu aligner les gels ensemble et associer les points ayant les mêmes coordonnées automatiquement à travers les différents gels. Une contre-vérification a été effectuée pour corriger toute erreur d'assignation potentielle au niveau des correspondances entre les points pouvant avoir des conséquences sur les analyses sous-jacentes. Le poids moléculaire a été déterminé en utilisant une échelle moléculaire (Bio-Rad) dans un puits en bordure d'un gel 2D de sorte à ne pas empiéter sur la migration de celui-ci. Le point isoélectrique a été déterminé en rapportant le pourcentage d'une longueur de gel donnée à la courbe linéaire fournie par le fabricant (Amersham pharmaciabiotech). À partir de ces informations, le poids moléculaire et le point isoélectrique ont été automatiquement extrapolés par le programme pour l'ensemble des points de tous les gels.

Une analyse qualitative (absence/présence) a été effectuée, comptabilisant les points trouvant une correspondance dans les autres gels d'un même biotype (intrabiotype) ou dans les gels de l'autre biotype (interbiotype). L'analyse qualitative intrabiotype a été effectuée pour nous permettre de déterminer la reproductibilité des résultats de nos gels 2D au sein d'un même groupe biologique. De cette façon, nous pouvons nous assurer que toute différence détectée au niveau de l'analyse qualitative interbiotype n'est pas la conséquence d'artéfacts techniques, ou de la variabilité interindividuelle. Les points ne trouvant aucune correspondance dans les gels de la forme cybride ou de la forme sauvage n'ont pas été pris en compte dans l'analyse qualitative interbiotype. Dans l'analyse qualitative interbiotype, les points considérés comme

spécifiques à un biotype correspondent à des points constamment retrouvés dans tous les gels d'un biotype alors qu'ils sont absents dans tous les gels de l'autre biotype.

Une analyse quantitative (taux d'expression) a également été effectuée en s'appuyant sur le pourcentage volume (%Vol), une valeur normalisée correspondant au pourcentage de volume que représente un point par rapport à la somme des volumes de l'ensemble des points du même gel. L'intérêt de normaliser nos valeurs est de nous permettre de comparer l'intensité des points d'un gel à l'autre malgré la présence d'une variation technique. Cette variation peut être causée par une différence de quantité de protéines chargées ou du temps de révélation. Ainsi, si un gel est plus contrasté qu'un autre, le volume de ses points sera proportionnellement plus important et ne permettra pas une comparaison directe entre les différents gels. Cependant, dans la mesure où le taux d'expression des protéines est comparable pour un point donné, le pourcentage volume associé sera similaire entre les gels (Byrjalsen *et al.*, 1999, Tsuji *et al.*, 1999).

La formule permettant le calcul du pourcentage volume est présentée ci-dessous :

$$\%Vol = \frac{Vol}{\sum_{S=1}^n Vol_S} \times 100$$

Vol S est le volume total des points dans un gel contenant n points

Un test de Student a été effectué à partir du %Vol des points présents sur l'ensemble des gels (n=6) dans le cadre de l'analyse quantitative. Pour cette analyse, les points non distincts ou présentant des signes de saturation ont été retirés de l'analyse afin de ne pas biaiser les résultats. Les résultats ont été filtrés pour ne conserver parmi les points considérés différents que les correspondances dont le %Vol est au moins deux fois supérieures entre les deux biotypes afin de limiter les faux positifs. Les points ne trouvant qu'une correspondance intrabiotype ou nettement surexprimés (3 points vs 3 points avec une valeur-p < 0.05) ont été

comptabilisés comme spécifiques à leur biotype. Les points retrouvés chez les deux biotypes sont comptabilisés comme étant communs s'ils ne présentent pas de différence significative en termes de %Vol.

### **2.6.2 Diagramme de dispersion**

En plus d'une analyse comparative, une analyse quantitative a été effectuée sur les gels 2D à partir du programme Mélanie 7 (Genebio). Le pourcentage volume (%Vol) de l'ensemble des points présentant une correspondance entre deux gels ont été comparés avec un diagramme de dispersion. Le diagramme de dispersion cherche à créer la courbe de tendance linéaire entre la valeur du % Vol des points d'un gel (variable x) et le %Vol des points correspondant dans le gel de référence (variable y). La courbe de tendance linéaire est définie par l'équation suivante :

$$y = ax + b$$

Où a représente le coefficient directeur et b l'ordonnée à l'origine.

La qualité de la courbe de tendance linéaire est définie par son coefficient de détermination ( $R^2$ ). Ce dernier peut varier dans un intervalle compris entre 0 et 1, où une valeur se rapprochant de 1 indique une bonne corrélation. Le gel de muscle utilisé comme référence dans les diagrammes de dispersion pour les comparaisons intrabiotypes (forme sauvage/forme sauvage) et interbiotypes (forme sauvage/forme cybride) est le gel « Forme sauvage1 » (espèce sauvage). Le gel « Forme cybride 1 » (forme cybride) a été utilisé comme gel de référence pour les comparaisons intrabiotypes (forme cybride /forme cybride).

### **2.6.3 Sélection des protéines pour la spectrométrie de masse**

Après avoir terminé l'analyse comparative sur les six gels, nous avons identifié les points spécifiques à chaque biotype. Sur la base des critères énoncés dans la section précédente, une protéine est considérée comme spécifique à un biotype lorsqu'elle est uniquement retrouvée

dans ce biotype ou lorsqu'elle présente une différence d'abondance significative et que le ratio de cette différence est supérieur à 2:1 en faveur de ce biotype (valeur-p < 0,05 et  $\Delta\%$  Vol > 2:1). Afin de maximiser les chances d'identification par spectrométrie de masse, les protéines spécifiques à chaque biotype ont été isolées à partir de nouveaux gels révélés selon le protocole spécifique au séquençage. De cette façon, 35 échantillons de points ressortant de l'analyse comparative comme étant spécifiques ont été extraits et envoyés à séquencer à l'institut de recherche en cancérologie et immunologie (IRIC) de l'Université de Montréal. Sur ces 35 échantillons, 11 correspondent à des points uniquement retrouvés chez la forme sauvage et 13 à des points uniquement retrouvés chez la forme cybride. Les 11 autres échantillons correspondent à des points présentant une différence d'abondance, soit 4 en faveur de la forme sauvage et 7 en faveur de la forme cybride. Certains points spécifiques ont été envoyés à plusieurs reprises parmi ces 35 échantillons en raison d'un nombre important d'échecs (voir annexe 1).

## **2.7 Digestion des bandes et MS/MS**

Après leur réception à l'IRIC, les points ont été découpés et placés dans une solution contenant de l'acétonitrile (50%), du bicarbonate d'ammonium (50mM) et de Tris 2-carboxyéthyl phosphine (10mM). Le tout a été agité pendant une heure à 37°C. Du chloroacétamide (55mM) a été ensuite ajouté pour permettre l'alkylation, les extraits sont alors agités durant une autre heure à 37°C. Afin de permettre la digestion, 1 $\mu$ g de trypsine est ajouté aux échantillons. Après un temps d'attente de 8 heures à 37°C, on procède à l'extraction des peptides avec une solution d'acétonitrile (90%). Les extraits sont ensuite séchés, puis resolubilisés dans une solution d'acétonitrile (5%) et d'acide formique (0,2%). On les charge ensuite sur des colonnes C18 faites maison (de 0.3mm par 5mm) reliée directement à une soupape de commutation qui assure leur passage et leur séparation vers une colonne en phase inverse. Cette seconde colonne est constituée d'un gradient d'acétonitrile

(10-60%) et d'acide formique (0.2%) permettant le passage des extraits selon un débit de 600nl/min vers un Nano LC-2D (Eksigent) couplé à un élite LTQ-Orbitrap (Thermo Fisher Scientific). Chaque spectre MS possédant une résolution d'au moins 60.000 est suivi de 12 spectres MS/MS. Les 12 spectres présentant le ratio d'ions à charge multiple le plus important ont été sélectionnés pour être séquencés par MS/MS. Le tandem MS a été effectué en analysant les collisions induites par la dissociation au niveau du piège à ions linéaire. Les données ont été analysées avec le programme Mascot 2.4 (Matrix Science). Les peptides détectés ont été comparés à la banque de données protéomique du NCBI et également une banque de données issue de l'assemblage du transcriptome des différents biotypes du complexe *Chrosomus eos-neogaeus* (B. Angers ; données non publiées). Le niveau de tolérance concernant les précurseurs et les fragments a été réglé respectivement à 15ppm et 0.5 Da. Les modifications post-traductionnelles variables sélectionnées dans Mascot sont la carbamidomethyl, l'oxydation, la désamination, la méthylation, le diméthyl et le triméthyl.

## 2.8 Visualisation des résultats de séquençage

Les résultats du séquençage sont transférés dans le logiciel Scaffold 4 (Proteome software) afin d'être visualisés. On règle le nombre de peptides minimums à deux par protéine et le filtre des peptides et des protéines à 1% de la FDR (*false decoy recovery*) afin de ne considérer que les peptides spécifiques les plus fiables dans l'analyse. Dans un premier temps, on retire la contamination représentée par les protéines présentes dans plusieurs échantillons correspondant généralement à des protéines liées au cytosquelette, telles que la myosine ou l'actine très présente dans les cellules musculaires. On associe ensuite la séquence protéique se situant en première place comme la protéine isolée dans le point. Le poids moléculaire et le point isoélectrique de la protéine de cette séquence sont ensuite calculés et comparés à celui des points envoyés afin de filtrer au besoin la contamination propre à l'échantillon. Pour se faire, nous avons utilisé l'application « Compute pI/Mw tool » du site ExPASy en utilisant

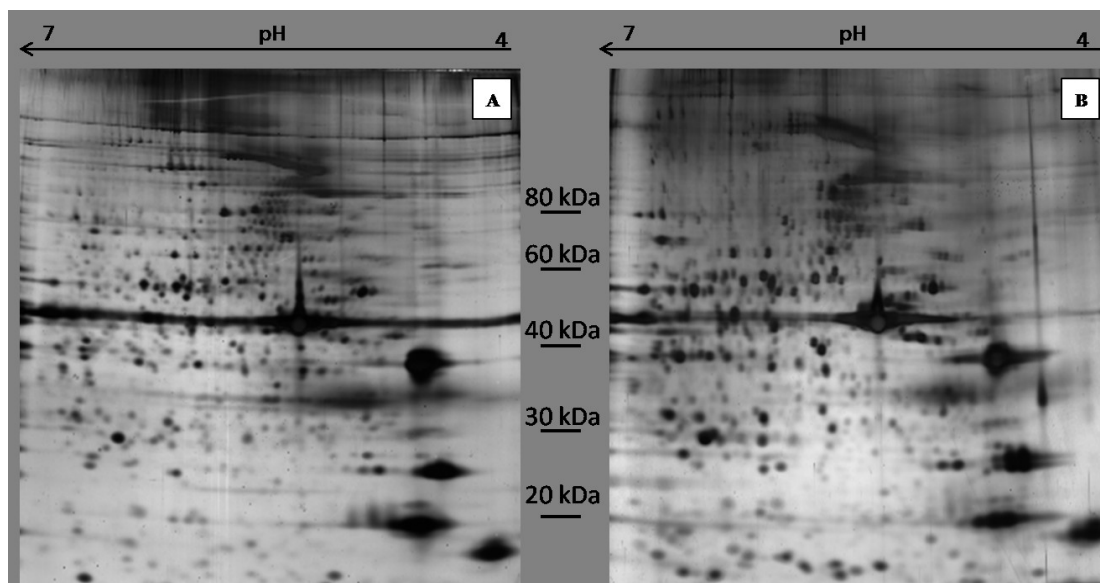
comme résolution la moyenne (Wilkins *et al.*, 1999). Le numéro Gi de la protéine ayant la plus forte homologie dans la banque de données du NCBI a été annoté pour chaque protéine. Le pourcentage de couverture ainsi que le nombre de spectres et de peptides uniques ont été comptabilisés pour chaque protéine.

### **3 Résultats**

#### **3.1 Analyse comparative**

Afin d'étudier les différences existant au niveau protéomique entre nos deux biotypes, nous avons réalisé une vingtaine de gels 2D. Nous avons sélectionné pour chaque biotype les trois gels présentant la meilleure résolution et le plus grand nombre de points. L'utilisation de gels 2D nous permet de produire un patron protéomique spécifique à chaque biotype qui peut être utilisé à des fins de comparaison dans l'analyse comparative. Un gel représentatif pour chaque biotype est représenté à la figure 5.

On peut observer que de façon générale, si les protéines les plus abondantes montrent un patron de migration très similaire et facile à établir, plusieurs protéines semblent produire des patrons de migration très différents entre les biotypes (figure 5). Pour analyser ces patrons de façon systématique, nous avons eu recours au programme de reconnaissance d'image Melanie 7.0, qui a permis d'établir la reproductibilité des gels et de quantifier la variation protéomique entre les deux biotypes.



**Figure 5 : Gel à deux dimensions du muscle squelettique de la forme sauvage (A) et de la forme cybride (B) :** Le pH de la première dimension et le poids moléculaire de la seconde dimension sont indiqués respectivement sur le dessus et au centre des gels.

### 3.1.1 Reproductibilité

Afin de s'assurer que les différences observées ne sont pas induites par une variation technique ou interindividuelle nous procédons à une analyse intrabiotype qui nous permettra de vérifier ces paramètres et ultimement la reproductibilité (tableau III).

**Tableau III : Reproductibilité des gels 2D du protéome du muscle pour la forme sauvage et de la forme cybride :**

Biotype	Nombre de points trouvant une correspondance intrabiotype	Nombre de points ne trouvant pas de correspondance intrabiotype	Reproductibilité (%)
Forme sauvage	414	15*	96,4
Forme cybride	414	12*	97,1

\* Points retirés des étapes subséquentes de l'analyse.



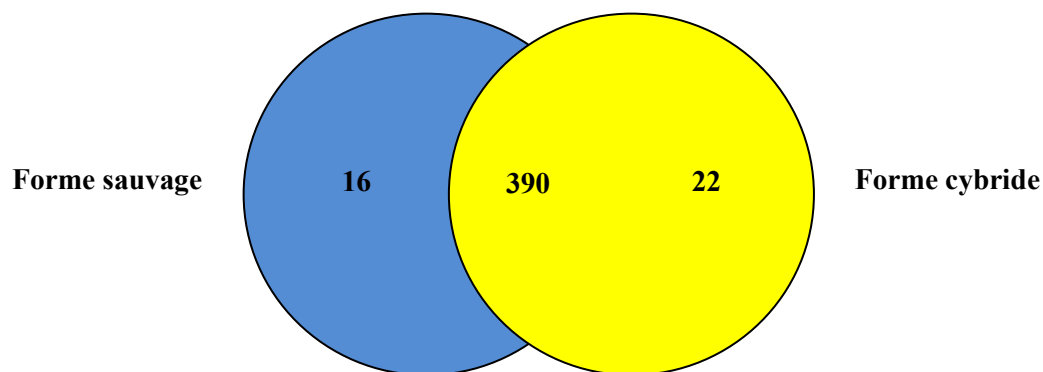
L'analyse comparative réalisée à partir des gels 2D a permis de détecter 414 correspondances intrabiotypes pour chaque biotype. Sur les 414 points communs au sein des gels d'un même biotype, 15 points chez la forme sauvage et 12 points chez la forme cybride ne trouvent pas de correspondance dans les autres gels de leur biotype. La reproductibilité obtenue est similaire entre les deux biotypes avec 96,4% pour la forme sauvage et 97,1% chez la forme cybride (Tableau III). En moyenne, quatre à cinq points par gel ne trouvaient pas de correspondance dans les autres gels produits à partir du même biotype. Dans le cadre de l'analyse qualitative interbiotype, les points ne trouvant aucune correspondance avec aucun autre gel (forme cybride et forme sauvage) n'ont pas été comptabilisés dans l'analyse comparative. En calculant l'inverse du pourcentage correspondant à la reproductibilité pour chaque biotype, on obtient la variation intrabiotype. Celle-ci correspond à la différence protéomique qualitative existant entre différents individus appartenant au même biotype. La variation intrabiotype est donc de 3,6% pour la forme sauvage et de 2,9% pour la forme cybride.

### *3.1.2 Différence interbiotype*

L'analyse qualitative intrabiotype précédente nous a permis de nous assurer que la reproductibilité était suffisante pour pouvoir considérer les différences interbiotype observées comme significatives. L'analyse précédente nous a également permis d'isoler les points ne trouvant aucune correspondance afin de les retirer de l'analyse qualitative interbiotype. Les différences interbiotypes sont représentées sous forme de diagramme de Venn dans la figure 6.

Sur les 428 correspondances totales (interbiotypes et intrabiotypes), 390 points sont communs aux deux biotypes et ne présentent pas de différence d'expression significative. Les protéines présentes seulement chez un biotype ou montrant une variation moyenne d'expression supérieure à deux fois le %Vol entre les deux biotypes sont considérés comme spécifiques à ce biotype (valeur-p < 0,05 et  $\Delta\% \text{ Vol} > 2:1$ ). Ainsi, 38 points sont spécifiques ou

significativement surexprimés chez un biotype ou l'autre. Parmi ces 38 points, on retrouve 7 points surexprimés et 9 points uniques pour la forme sauvage contre 13 points surexprimés et 9 points uniques pour la forme cybride. La différence interbiotype représente 8,9% des correspondances totales dans le protéome du muscle (figure 6).

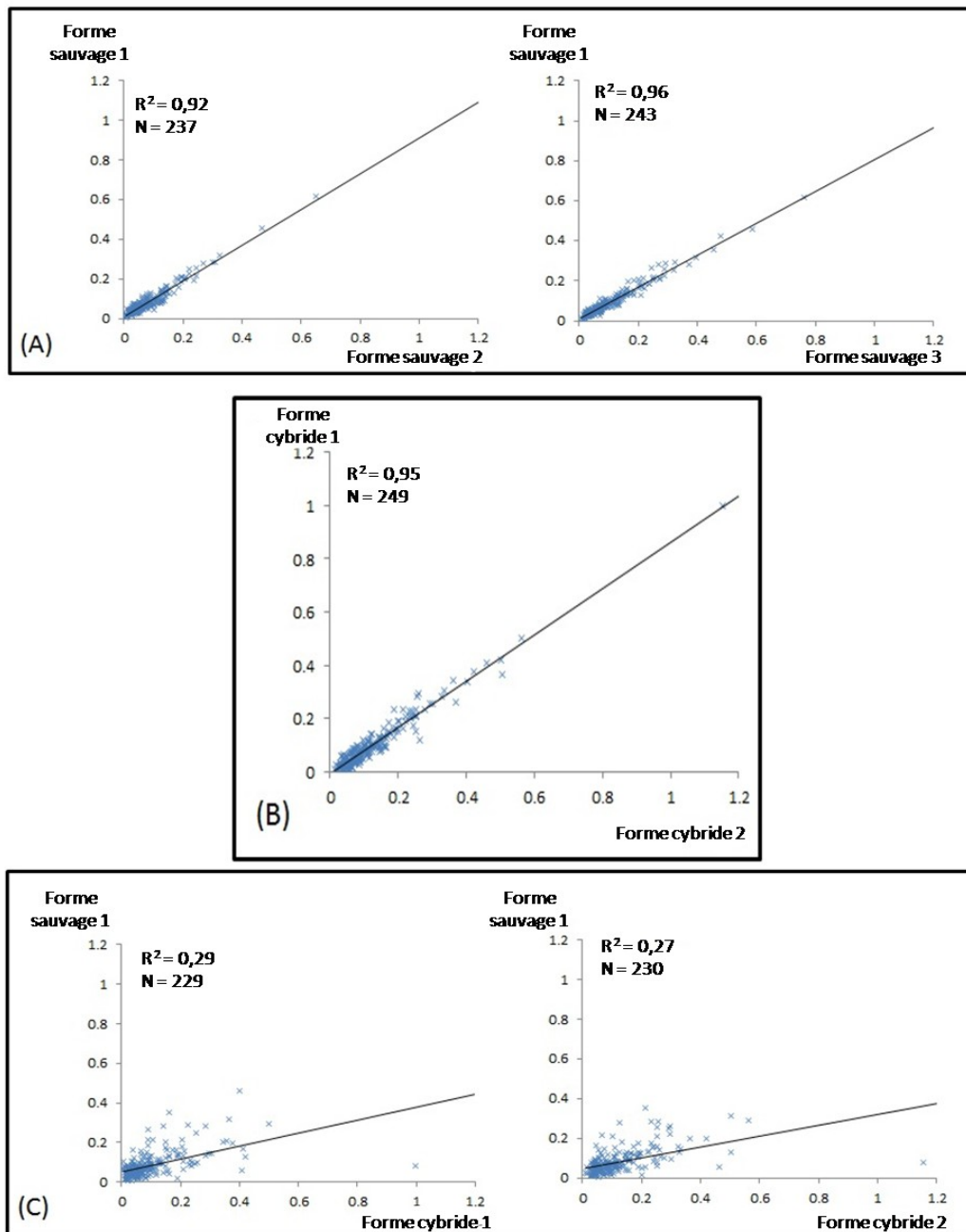


**Figure 6 : Diagramme de Venn représentant la similarité entre le pourcentage volume des points de deux gels :** Le nombre de points spécifiques à chaque biotype est présenté dans la section bleue pour la forme sauvage et jaune pour la forme cybride. Le nombre de points communs est présenté dans la section centrale verte. (N=428)

### ***3.1.3 Diagramme de dispersion***

Pour effectuer une analyse au niveau de l'abondance de protéines entre les biotypes nous avons utilisé des diagrammes de dispersion. Dans cette analyse, le %Vol de chaque point d'un gel est comparé avec le %Vol des points correspondants sur un autre gel. Une régression linéaire permet de calculer la ressemblance entre les gels et le coefficient de détermination ( $R^2$ ). Ainsi, le diagramme de dispersion comparant deux gels identiques montrera une pente de 1 et un  $R^2$  de 1 puisque les %Vol seront identiques pour chaque point. Pour évaluer la variabilité au niveau de l'abondance en protéines existant entre les deux gels produits à partir d'échantillons appartenant au même biotype, nous avons produit les diagrammes de dispersion en comparant l'ensemble des gels deux à deux. Nous avons ainsi montré que l'abondance protéique est similaire et qu'il n'y a pas de variation interindividuelle ou technique importante

au sein d'un même biotype ( $R^2 = 0,92$  et  $0,96$  pour la forme sauvage et  $R^2 = 0,95$  pour la forme cybride; figure 7A et B).



**Figure 7 : Diagrammes de dispersion entre le %Vol des points des gels :**

Intrabiotype forme sauvage (A), intrabiotype forme cybride (B) et interbiotype forme sauvage/forme cybride (C). Le nom des gels comparés est indiqué à l'extrémité des axes. Les valeurs sur les axes sont en %Vol. Le N correspond au nombre de points communs aux deux gels analysés. La pente est tracée pour chaque diagramme de dispersion.

En raison d'un nombre important de points saturés, le gel « forme cybride 3 » a dû être retiré de cette analyse. Au vu de la similarité entre les coefficients de détermination des diagrammes de dispersion intrabiotype pour les formes sauvage et cybride, on peut considérer que les gels sont reproductibles lorsque l'on considère l'estimation du niveau d'expression.

Nous avons ensuite effectué une comparaison de l'abondance en protéine entre les deux biotypes. Cette fois, le coefficient de détermination est de 0,29 pour le diagramme de dispersion « forme sauvage 1/ forme cybride 1 » et de 0,27 pour le diagramme de dispersion « forme sauvage 1/ forme cybride2 » (figure 7C). Ces résultats démontrent que les patrons d'expression de la forme sauvage et de la forme cybride sont très différents l'un de l'autre lorsqu'on tient compte de l'abondance relative des protéines.

### **3.2 Analyse des protéines spécifiques à un biotype**

Sur les 38 points spécifiques retrouvés dans l'analyse interbiotypes, 20 points proviennent de protéines présentant une différence d'expression entre les deux biotypes (valeur-p < 0,05 et  $\Delta\% \text{ Vol} > 2:1$ ) (tableau IV).

Sur ces 20 points, 13 sont surexprimés en faveur de la forme cybride et 7 en faveur de la forme sauvage. Les points présentant la plus grande différence d'abondance entre les biotypes sont généralement surexprimés chez la forme cybride. Le tableau IV présente la valeur-p et la moyenne du %Vol en plus de la différence d'expression interbiotype. La figure 8 présente la distribution des différences d'expression des protéines spécifiques entre les deux biotypes.

**Tableau IV : Points exprimés différemment dans le protéome du muscle entre la forme sauvage et de la forme cybride :**

Échantillon <sup>1</sup>	Biotype <sup>2</sup>	VALEUR-P	Moyenne	Moyenne	Différence d'expression
			% Vol Forme sauvage	% Vol Forme cybride	
D1*	Forme cybride	0,007	0,087	0,951	10,9
D2*	Forme sauvage	0,012	0,091	0,012	7,8
D3	Forme cybride	0,036	0,017	0,131	7,6
D4	Forme cybride	0,008	0,023	0,149	6,5
D5	Forme cybride	0,032	0,041	0,180	4,3
D6	Forme cybride	2,70E-05	0,036	0,154	4,3
D7	Forme cybride	0,036	0,021	0,078	3,7
D8	Forme sauvage	0,009	0,202	0,055	3,7
D9	Forme sauvage	0,034	0,098	0,028	3,5
D10*	Forme sauvage	0,015	0,160	0,046	3,5

<sup>1</sup> Le « D » de la colonne « échantillon » fait référence aux points spécifiques surexprimés.

<sup>2</sup> La colonne biotype représente le biotype pour lequel le point est surexprimé.

\* Échantillons identifiés par séquençage (tableau V et VI).

Échantillon <sup>1</sup>	Biotype <sup>2</sup>	VALEUR- P	Moyenne % Vol Forme sauvage	Moyenne % Vol Forme cybride	Différence d'expression
D11	Forme cybride	0,012	0,022	0,074	3,3
D12	Forme cybride	0,008	0,018	0,058	3,3
D13	Forme cybride	0,045	0,088	0,253	2,9
D14	Forme cybride	0,002	0,074	0,199	2,7
D15	Forme cybride	0,041	0,031	0,080	2,6
D16	Forme sauvage	0,016	0,169	0,073	2,3
D17	Forme cybride	0,023	0,050	0,114	2,3
D18	Forme cybride	0,001	0,046	0,105	2,3
D19*	Forme sauvage	0,013	0,182	0,082	2,2
D20*	Forme sauvage	0,047	0,231	0,109	2,1

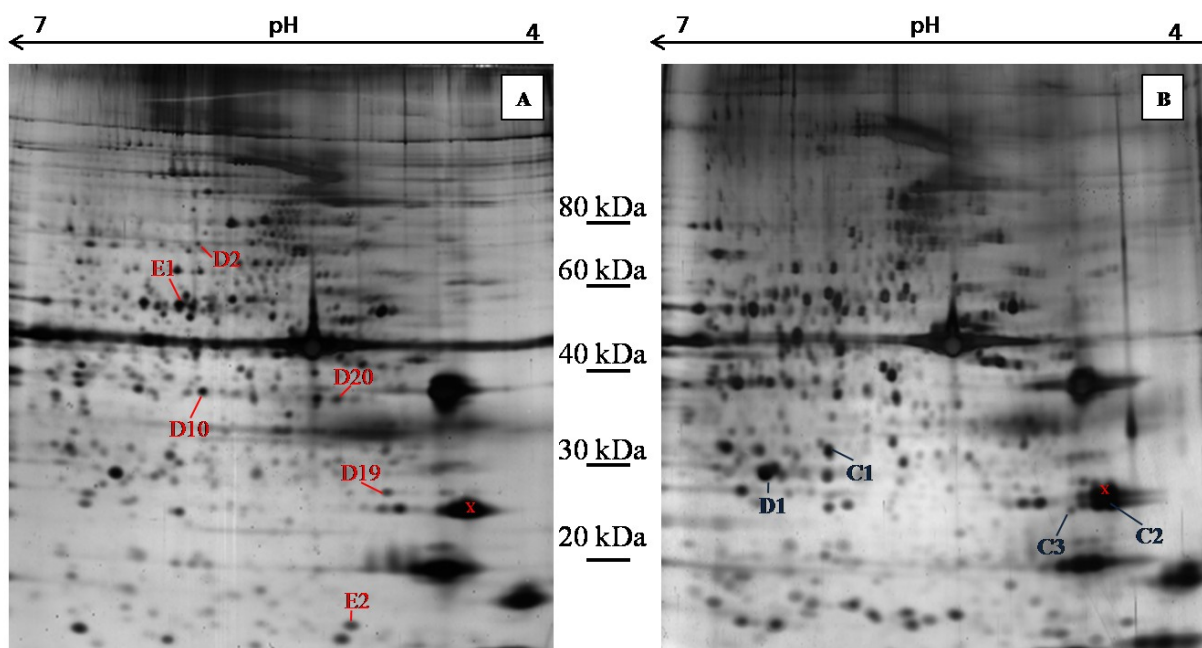
<sup>1</sup> Le « D » de la colonne « échantillon » fait référence aux points spécifiques surexprimés.

<sup>2</sup> La colonne biotype représente le biotype pour lequel le point est surexprimé.

\* Échantillons identifiés par séquençage (tableau V et VI).

### 3.3 Identification des protéines spécifiques

Au cours de ce projet, 35 échantillons ont été envoyés au service de protéomique de l'IRIC et soumis à une analyse par spectrométrie de masse pour tenter d'identifier les protéines qu'ils contenaient. De ces 35 échantillons, seulement 10 ont pu être identifiés avec certitude, à cause de diverses difficultés techniques (présentées dans l'annexe 1), et surtout à cause de l'inexistence de données génétiques détaillées pour cette espèce. Une banque transcriptomique partielle n'a été disponible que vers la fin du projet, et a rendu possible les identifications présentées. La position respective sur les gels 2D de ces protéines est indiquée à la Figure 8.



**Figure 8 : Localisation des points spécifiques sur les Gels à deux dimensions du muscle squelettique de la forme sauvage (A) et de la forme cybride (B) :** Les gels présentent les points spécifiques qui ont été identifiés par spectrométrie de masse. Les points différemment exprimés entre les biotypes sont identifiés par un D suivi d'un chiffre allant de 1 à 20 (tableau IV pour plus d'information). Les points uniquement retrouvés chez un biotype sont identifiés par la lettre « E » pour la forme sauvage et « C » pour la forme cybride (tableau V pour plus d'information). Le pH de la première dimension et le poids moléculaire de la seconde dimension sont indiqués respectivement sur le dessus et à droite du gel. Le X est un repère visuel pour le lecteur (texte section 3.3).

On retrouve 6 points spécifiques chez la forme sauvage et 4 points spécifiques chez la forme cybride. Les points D1 et D2 sont les points présentant la plus grande différence d'expression

entre les biotypes (tableau IV). La localisation adjacente du point unique à la forme cybride C2 et de la tache au-dessus (x, figure 8) semble montrer que C2 et ce point correspondent à la même protéine (hypothèse vérifiée par spectrométrie de masse). Le fait que ces points aient une masse moléculaire différente dans nos gels 2D démontre que la différence protéomique observée entre ces points est très probablement associée à une modification post-traductionnelle.

### 3.4 Résultat du séquençage

L'ensemble des informations collectées lors de l'analyse par spectrométrie de masse est présenté dans le tableau V pour chacune des protéines spécifiques identifiées. Les dix échantillons séquencés correspondent à 9 protéines spécifiques.

Le tableau V présente l'identité des neuf protéines spécifiques. Pour valider l'identification, nous avons comparé les poids moléculaires et points isoélectriques théoriques des protéines identifiées avec les positions desquels les points analysés par spectrométrie de masse ont été extraits du gel. Dans tous les cas, sauf pour la TPI B, les positions observées étaient raisonnablement proches des valeurs théoriques. L'identification de la TPI B dans un échantillon isolé d'un autre gel à la position correspondante nous amène cependant à considérer qu'il ne s'agit pas d'une erreur de séquençage. Le pourcentage de couverture représenté par les peptides uniques détectés ainsi que le nombre de peptides et de spectres uniques ont été indiqués pour chaque protéine afin de s'assurer que les protéines identifiées sont présentes en quantité non négligeable dans leur échantillon respectif. Avec ces informations, on remarque que, mise à part la myozénine 1, les protéines sont présentes en quantité non négligeable. La myozénine 1 a tout de même été considérée en raison de sa spécificité dans l'échantillon D20 par rapport aux autres échantillons envoyés à séquencer en même temps. De plus, la myozénine était en première position derrière les protéines cytosquelettiques contaminant la majorité de nos échantillons (annexe 1).



L'identification des protéines spécifiques nous a permis de pouvoir découvrir la fonction et la localisation cellulaire associée à chaque protéine dans la littérature scientifique. Ces résultats sont présentés dans le Tableau VI.

**Tableau V : résultats du séquençage des points spécifiques à un biotype ou exprimés différemment :**

Échantillon <sup>1</sup>	Identité <sup>2</sup>	MM/pI théorique	MM/pI échantillon	Gi accession <sup>3</sup>	Spectre / peptide unique <sup>4</sup>	Couverture (%)
E1	Enolase 1 sous unité alpha	47 kDa / 6,16	46 kDa / 6,18	47229720	14/10	23
E2	Protéine de liaison des acides gras (FABP 3)	14,8 kDa / 5,71	16 kDa / 5,2	23308625	2/2	2,52
D1 <sup>C</sup>	Triose-phosphate isomérase B	7.61 kDa / 26.8	6,44/26	47271422	6/5	21,3
D2 <sup>E</sup>	Succinate deshydrogénase sous unité A	72 kDa / 6,2	71 kDa / 6,13	28374211	2/2	2,7
D10 <sup>E</sup>	Alpha tropo myosine	32,7 kDa / 4,7	35 kDa / 5,34	42476235	2/2	1,28
D19 <sup>E</sup>	Apolipo protéine AI 2 précurseur	30,6 kDa / 5,25	26 kDa / 5,05	206598062	7/4	12
D20 <sup>E</sup>	Myozenine 1	33 kDa / 5,62	34 kDa / 6,12	45387767	1/1	0,62
C1	Guanidioacétate N-méthyltransférase	26,5 kDa / 5,8	29 kDa / 6,16	157743330	3/2	6,4
C2	Chaîne légère phosphorylable de myosine	18,9 kDa / 4,65	22 kDa / 4,6	533528662	7/6	36,2
C3	Chaîne légère phosphorylable de myosine	18,9 kDa / 4,66	22 kDa / 4,8	533528663	4/4	24,5

<sup>1</sup> Les lettres de la colonne échantillon représentent le biotype dont l'échantillon est issu, soit (E) pour les protéines identifiées la forme sauvage, (C) pour la forme cybride et (D) pour les protéines exprimées différemment. <sup>E</sup> et <sup>C</sup> en indice des points exprimés différemment réfèrent au biotype surexprimant le point.

<sup>2</sup> Les protéines ont été comparées à une banque de données faite à partir d'un assemblage du transcriptome des différents biotypes du complexe *Chrosomus eos-neogaeus*.

<sup>3</sup> Le numéro d'accèsion Gi renvoie à la protéine ayant la plus forte homologie avec la séquence d'acides aminés issus de notre banque de données.

<sup>4</sup>Les valeurs de la colonne spectre et peptide unique correspondent au nombre d'acides aminés et de spectres détectés avec un seuil de FDR à 1% pour les protéines et les peptides.

Sur les dix protéines spécifiques identifiées, quatre ont des fonctions rapportées associées au métabolisme mitochondrial, deux sont associées à la glycolyse, quatre sont des protéines participant à la formation du cytosquelette et 5 sont associées à une réponse à l'acclimatation au froid (tableau VI). De façon plus descriptive, la sous-unité alpha de l'énolase identifiée chez la forme sauvage est localisée dans la mitochondrie et participe à la glycolyse et au transport d'ARNt<sup>lys</sup> vers la mitochondrie. FABP 3, identifiée chez la forme sauvage est localisée dans la mitochondrie et participe à la bêta-oxydation des acides gras et joue un rôle au niveau de la thermorégulation. La TPI B, surexprimée chez la forme cybride, est localisée dans le cytosol. Elle participe à la glycolyse et est associée à la production de glycérol. La sous-unité alpha de la SDH surexprimée chez la forme sauvage est localisée dans la mitochondrie et le cytosol. Elle participe au cycle de Krebs et à la formation du complexe II du système OXPHOS. L'alpha-tropomyosine surexprimée chez la forme sauvage est retrouvée dans le cytosol et participe à la formation du cytosquelette. Le précurseur de l'ApoA surexprimé chez la forme sauvage est localisé dans le cytosol. Cette protéine participe à l'importation de protéines mitochondriales encodées par le génome nucléaire vers la mitochondrie. La myozénine 1 surexprimée chez la forme sauvage est localisée dans le cytosol et participe à la formation du cytosquelette. La GAMT identifiée chez la forme cybride et localisée dans le cytosol participe à la formation de la créatine. Deux points identifiés chez la forme cybride et localisés dans le cytosol correspondent à la MLC, qui participe à la formation du cytosquelette en permettant la contraction cellulaire.

**Tableau VI : Fonctions et localisations des protéines spécifiques à un biotype ou exprimées différemment :**

Échantillon <sup>1</sup>	Identité <sup>2</sup>	Fonction	Localisation	Référence
E1	Enolase 1, sous unité alpha	Participe à la glycolyse et au transport d'ARNt vers la mitochondrie	Mitochondrie : espace inter-membranaire et membrane externe	Brandina <i>et al.</i> , (2006); Entelis <i>et al.</i> , (2006)
E2	Protéine de liaison des acides gras (FABP 3)	Alimente la réaction d'oxydation mitochondriale en acide gras, thermorégulation.	Mitochondrie : membrane externe	Fisher <i>et al.</i> , (2006) Vergnes <i>et al.</i> , (2011)
D1 <sup>C</sup>	Triose-phosphate isomérase B	Glycolyse, associée à la production de glycérol	Cytoplasme	Karp, (2009) Compagno <i>et al.</i> , (1996)
D2 <sup>E</sup>	Succinate deshydrogénase sous unité A	Participe à la formation de la chaîne respiratoire et au cycle de Krebs	Mitochondrie : complexe II de la chaîne respiratoire	Oyedotun et Lemire, (2004)
D10 <sup>E</sup>	Alpha tropo myosine	Formation du cytosquelette	Cytoplasme: filament d'actine	Mannherz et Goody, (1976)
D19 <sup>E</sup>	Apolipo protéine ai2 precursor	Transport de Lipide, transport de protéine mitochondriale probable	Cytoplasme: gouttelette de lipide	Zhang <i>et al.</i> , (2011)
D20 <sup>E</sup>	Myozénine 1	Formation du cytosquelette	Cytoplasme: filament d'actine	Takada <i>et al.</i> , (2001)
C1	Guanidinoacétate méthyltransférase	Formation de la créatine	Cytoplasme	Ha et Bhagavan (2011)
C2	Chaîne légère phosphorylable de myosine	Formation du cytosquelette	Cytoplasme: filament d'actine	Szczesna <i>et al.</i> , (2002)
C3	Chaîne légère phosphorylable de myosine	Formation du cytosquelette	Cytoplasme: filament d'actine	Szczesna <i>et al.</i> , (2002)

<sup>1</sup> Les lettres de la colonne échantillon représentent le biotype dont l'échantillon est issu, soit (E) pour la forme sauvage, (C) pour la forme cybride et (D) pour les protéines exprimées différenciellement. Les exposants <sup>E</sup> et <sup>C</sup> en indice des points exprimés différenciellement réfèrent au biotype surexprimant le point.

<sup>2</sup> Les protéines ont été comparées à une banque de données faite à partir d'un assemblage du transcriptome des différents biotypes du complexe *Chrosomus eos-neogaeus*.

## 4 Discussion

Dans ce projet de maîtrise, nous avons tiré avantage de la présence d'une forme cybride dans le complexe *Chrosomus eos-neogaeus* pour étudier l'impact de la mitochondrie sur le protéome. Pour ce faire, nous avons effectué une analyse comparative et quantitative partielle sur le protéome du muscle. Le choix d'utiliser des tissus musculaires pour cette analyse découle du fait que le protéome du muscle a déjà été étudié chez de nombreuses espèces (Egginton et Sidell 1989, Guderley 1990, Guderley 2004, Anttila *et al.*, 2006, Fischer *et al.*, 2006, Hayley *et al.*, 2011, Angers *et al.*, 2012). Nous disposons ainsi d'une base pour analyser certains de nos résultats. Ce tissu est également très abondant et facile à isoler, même chez des poissons de petite taille tels que les *Chrosomus*.

L'utilisation de gels 2D a été préférée à d'autres techniques pour leur capacité à séparer efficacement les protéines d'un échantillon. La migration des protéines sur un gel 2D produit un patron protéomique spécifique qui peut ensuite être utilisé comme référence lors de comparaisons avec d'autres groupes biologiques. Cette approche peut aussi permettre de visualiser les protéines ayant subi des modifications post-traductionnelles ou présentes sous différentes isoformes. L'utilisation de 3 gels 2D par biotype, comprenant chacun l'extrait protéique de 4 individus par gel nous a permis d'obtenir des coefficients de détermination quasi identiques et très proches de 1 à l'intérieur d'un même biotype (figure 7 A et B). Ce résultat démontre que nous disposons d'une reproductibilité satisfaisante en conformité avec

d'autres études utilisant des gels 2D (López *et al.*, 2002, Rodriguez-Piñeiro *et al.*, 2004, Hunt *et al.*, 2005).

L'analyse comparative de gels 2D étant une activité extrêmement chronophage, nous n'avons effectué l'analyse comparative que sur les trois meilleurs gels pour les formes sauvage et cybride. Nous disposons toutefois d'une large banque de données constituée de l'ensemble des gels effectués. Nous avons ainsi été en mesure de confirmer que les points ressortant de l'analyse comparative comme étant spécifiques à un biotype étaient bien présents dans tous les autres gels 2D qui n'ont pas été utilisés pour ce biotype. Cette approche nous permet d'affirmer que les résultats obtenus lors de ce projet sont représentatifs de la différence interbiotype et intrabiotype existant entre la forme sauvage et la forme cybride.

Puisque notre objectif principal était de mesurer l'impact d'une mitochondrie étrangère sur le protéome du muscle, il était important de minimiser au maximum les variations dues à d'autres facteurs. Comme les formes sauvage et cybride sont allopatriques, il n'a pas été possible de capturer les spécimens des différents biotypes dans le même lac. Nous avons donc tenté de réduire au maximum la variabilité environnementale en sélectionnant des lacs les plus proches possible les uns des autres. Ainsi les trois lacs utilisés pour l'échantillonnage sont distants de moins de 9,4 km. Bien que nous ne disposions pas de données physico-chimiques détaillées, ces lacs présentent des superficies et des profondeurs similaires, ce qui suggère que ces différents environnements sont globalement très similaires.

#### 4.1 Reproductibilité des gels 2D

Si nous avons obtenu un résultat comparable à celui rapporté par Angers *et al.*, (2012) en ce qui a trait à la variabilité interbiotype, la variabilité intrabiotype que nous avons obtenue est légèrement supérieure, soit respectivement 3,6% et 2,9% pour la forme sauvage et la forme cybride, contre en moyenne 1,1% et 0,4% rapportés dans l'étude antérieure. Il faut cependant réaliser que cet écart ne représente qu'environ 4,5 points par gel dans notre étude contre 3

points par gel dans l'article, sur un total de respectivement 414 et 405 points dans notre étude et dans l'article. Les résultats obtenus sont donc tout à fait comparables, et la reproductibilité des gels est bien démontrée par cette comparaison intrabiotype. De plus, les coefficients de détermination obtenus en comparant les différents gels provenant d'un même biotype sont très élevés (0,94 en moyenne), ce qui atteste que l'abondance relative des protéines au sein d'un même groupe reste très similaire d'un gel à l'autre. La reproductibilité de la technique semble donc raisonnablement établie et la technique est fiable pour comparer les biotypes (figure 7 et tableau III).

#### 4.2 Étude comparative du protéome du muscle squelettique

Les résultats de notre projet nous ont permis de confirmer qu'il existe bien une différence significative entre les protéomes de la forme sauvage et de la forme cybride. La différence interbiotype qualitative, c'est-à-dire en termes de présence/absence des protéines sur les gels, est similaire à celle de 8,1% trouvée par Angers *et al.*, (2012). L'analyse de la différence d'abondance des protéines à l'aide des coefficients de détermination calculés selon les %Vol de tous les points communs aux deux biotypes montre qu'il existe une différence interbiotype beaucoup plus importante que la différence intrabiotype. Sachant que a) la forme sauvage et la forme cybride sont semblables génétiquement (Angers *et al.*, 2012) et b) que l'analyse du génome mitochondrial de l'hybride (*C. eos x C. neogaeus*), de la forme parentale *C. neogaeus* et de la forme cybride n'a démontré aucune différence (Angers et Schlosser 2007), nos résultats soutiennent que la présence de mitochondrie *C. neogaeus* chez la forme cybride a un impact majeur sur la régulation génique et sur le protéome qui en découle.

#### 4.3 Points surexprimés chez un biotype

Puisque le nombre d'échantillons est relativement petit, nous n'avons considéré que les points présents sur l'ensemble des gels avec une valeur-p < 0,05 et dont la différence d'expression était d'au minimum 2:1 entre les biotypes. L'analyse statistique effectuée sur l'ensemble des

points nous a ainsi permis d'identifier 20 points répondant à ces critères, répertoriés dans le Tableau IV. De ces 20 points, 13 points sont surexprimés chez la forme cybride et 7 points chez la forme sauvage. Comme l'expression interbiotype de ces points semble varier de façon importante entre les biotypes, les protéines correspondant à ces points pourraient potentiellement être à la base de la différence de « *fitness* » observée entre les deux biotypes. L'identification de ces protéines nous permettrait donc de mieux comprendre qu'elles sont les voies moléculaires permettant à la forme cybride de s'adapter à la présence de la mitochondrie *C. neogaeus* et à son environnement potentiellement plus nordique.

#### 4.4 Séquençage : identification des protéines spécifiques à un biotype

Sur les trente-cinq échantillons séquencés, seulement dix protéines ont pu être identifiées avec certitude à cause des sérieuses limitations techniques décrites à l'annexe 1. Soit, six pour la forme sauvage et quatre pour la forme cybride. Les dix points spécifiques identifiés correspondent à neuf protéines différentes. De ces neuf protéines, six sont spécifiques à la forme sauvage et trois sont spécifiques à la forme cybride. Sur les neuf protéines spécifiques, quatre protéines sont de par leur fonction ou leur localisation associées à la mitochondrie, trois protéines sont associées au cytosquelette et deux protéines sont associées à la glycolyse. Finalement, sur les neuf protéines spécifiques, cinq présentent un patron d'expression associé dans la littérature scientifique à une réponse à l'acclimatation au froid.

Ceci dit, il faut bien garder à l'esprit que bien que nous puissions observer des points clairement surexprimés dans un biotype par rapport à l'autre, il est difficile de conclure que la protéine correspondante est forcément surexprimée. En effet, l'expression d'isoformes différents ou des modifications post-traductionnelles affectant le point isoélectrique ou la masse moléculaire d'une protéine dans un biotype feraient en sorte qu'on ne retrouve pas son équivalent chez l'autre biotype, bien que la protéine ait été présente dans les deux échantillons au départ. Toutefois, des isoformes différents ou des modifications post-traductionnelles étant

susceptibles de modifier significativement le destin d'une protéine ou l'activité d'une enzyme, les différences au niveau du gel 2D sont fort probablement significatives, même si elles ne représentent pas nécessairement des différences d'expression proprement dites.

Quoi qu'il en soit, il ressort clairement de l'analyse par spectrométrie de masse que sur les neuf protéines spécifiques identifiées, sept sont des protéines cytosoliques et trois sont des protéines mitochondriales. Ce premier constat démontre que malgré que la différence entre nos deux biotypes soit la présence de mitochondrie *C. neogaeus* chez la forme cybride, une minorité de protéines affectées sont des protéines mitochondriales contrairement à ce quoi on aurait pu s'attendre. Nous discutons brièvement ci-après des fonctions physiologiques associées à chacune des protéines identifiées lors du séquençage.

#### ***4.4.1 Sous unité alpha de l'énolase1***

Le premier point présent sur les gels de la forme sauvage et non détecté chez la forme cybride (point E1 sur le Tableau V) correspond à la sous-unité alpha de l'énolase 1. L'énolase participe à l'avant-dernière étape de la glycolyse en convertissant le 2-phosphoglycérate en phosphoenolpyruvate et jouerait également un rôle au niveau du transport d'ARNt jusqu'à la mitochondrie (Berg *et al.*, 2002, Entelis *et al.*, 2006). La sous-unité alpha correspond à l'un des trois isoformes de l'énolase ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) et est retrouvée dans pratiquement tous les tissus où elle forme des complexes protéiques homodimériques ou hétérodimériques, tels que  $\alpha\alpha$ ,  $\alpha\beta$ ,  $\beta\beta$ ,  $\alpha\gamma$  ou  $\gamma\gamma$  (Pancholi 2001). On retrouve ces complexes majoritairement dans le cytosol, mais également dans les mitochondries (Brandina *et al.*, 2006). L'expression de l'énolase 1 (alpha) serait induite en condition hypoxique par le facteur d'induction hypoxique HIF-1 (Semenza *et al.*, 1996, Gracey *et al.*, 2001). Chez les mammifères, l'énolase peut subir de nombreuses modifications post-traductionnelles, notamment l'acétylation, la méthylation et la phosphorylation (Cooper *et al.*, 1984, Bottalico *et al.*, 1993). Cependant, on ignore si ces modifications modifient son activité catalytique. L'alpha émolase est surexprimée dans de



nombreux types de tumeurs, ce qui est associé avec le changement métabolique vers la glycolyse aérobie, ou effet Warburg (Warburg 1930), observé durant le développement des tumeurs cancéreuses (Capello *et al.*, 2011). Les modifications post-traductionnelles sont également modifiées dans les tissus cancéreux (Zhou *et al.*, 2010). Qu'un point spécifique à un des deux biotypes soit associé à l'énolase alpha laisse donc penser que les deux biotypes adoptent des stratégies métaboliques différentes. La présence d'un point unique chez la forme cybride correspondant à cette protéine indique qu'il existe potentiellement des différences importantes entre les deux formes au niveau glycolytique et/ou mitochondrial.

#### ***4.4.2 FABP 3***

Un second point uniquement retrouvé chez la forme sauvage correspond à la protéine FABP 3. FABP 3 est une protéine reconnue pour alimenter le cycle de Krebs en acétyl-CoA et la réaction d'oxydation en acide gras, mais aussi pour son implication dans la thermorégulation (Martinez *et al.*, 2003, Fischer *et al.*, 2006, Karp 2009, Vergnes *et al.*, 2011, Zhang *et al.*, 2011 b). La bêta-oxydation est responsable de dégrader les acides gras dans la cellule. Cette réaction métabolique produit un relâchement d'électrons qui est réacheminé vers la chaîne respiratoire mitochondriale pour produire de l'énergie sous la forme d'ATP. L'oxydation des acides gras produit aussi de l'acétyl-CoA qui est réutilisé dans les voies métaboliques du cycle de Krebs, de la gluconéogenèse et de la cétogenèse (Karp 2009). En ce qui concerne la thermorégulation, le gène de FABP3 serait 10 fois plus exprimé dans le tissu adipeux brun de rat après une exposition à 4°C durant 48h (Yamamoto *et al.*, 2011). Il existe différentes isoformes et modifications post-traductionnelles potentielles pour cette protéine, cependant il n'y a pas d'altération de la fonction connue à ce jour lorsque la protéine est modifiée (Jones *et al.*, 1988, Matarese *et al.*, 1989, Nielsen et Spener 1993, Schroeder *et al.*, 1998). Les souris doubles K.O pour le gène FABP3 de l'étude de Vergnes *et al.*, (2011) présentent une importante sensibilité au froid malgré l'induction de gènes liés à la voie de la bêta-oxydation

et l'hydrolyse des lipides présents dans les adipocytes bruns. La différence d'expression entre les formes cybrides et sauvage de FABP3 pourrait peut-être être liée à une telle réponse au froid, dans la mesure où la forme sauvage se retrouve dans l'extrême nord de son aire de distribution dans notre zone d'échantillonnage, en opposition à la forme cybride qui elle se trouve dans le sud de son aire de distribution (figure 3).

#### 4.4.3 Succinate déshydrogénase

La sous-unité A de la SDH correspond à un point 7,8 fois surexprimé chez la forme sauvage. Elle participe avec trois autres sous-unités à la formation de la SDH qui constitue le complexe II (succinate-coenzyme Q réductases) du complexe OXPHOS dont nous avons déjà parlé dans l'introduction (figure 2). En dehors de la chaîne respiratoire, la SDH serait aussi impliquée dans le cycle de Krebs (Zerbetto *et al.*, 1997, Oyedotun et Lemire 2004). Ces deux voies métaboliques sont intimement liées puisque c'est le cycle de Krebs qui alimente la chaîne respiratoire en NADH et FADH (Berg *et al.*, 2002). Les électrons relâchés par ces molécules fournissent l'énergie nécessaire à la constitution d'un gradient électrochimique dans la membrane interne mitochondriale permettant à l'ATP synthase de transformer de l'ADP en ATP. Dans la voie métabolique du cycle de Krebs, la SDH transforme le succinate en fumarate. L'implication du cycle de Krebs ne se limite pas seulement à alimenter la chaîne respiratoire en coenzyme oxydée. Elle permet également la formation de plusieurs composés réutilisés par la suite dans d'autres voies métaboliques telles que la synthèse du glutamate (l'alpha-cétoglutarate), la synthèse du cholestérol (citrates), ou la gluconéogenèse (oxaloacétate) (Berg *et al.*, 2002). Des études ont montré que l'expression de la SDH augmente après une exposition d'un an à 18°C chez *Danio rerio* et suite à un exercice physique important chez *Salmo salar* L. (Malek *et al.*, 2004, Anttila *et al.*, 2006).

Dans nos résultats, la sous-unité A de la SDH est en deuxième position dans le tableau IV avec la plus grande différence d'expression observable entre les deux biotypes. Le fait que

cette protéine fasse partie du le complexe II du complexe OXPHOS entièrement codé par le génome nucléaire est intéressant pour plusieurs raisons. Il s'agit du seul complexe de la chaîne respiratoire dont l'activité n'est pas affectée dans des hybrides cytoplasmiques murins, donc phylogénétiquement plus éloignés que nos deux biotypes (McKenzie *et al.*, 2003). Contrairement à cette étude, nos résultats suggèrent que les modifications biologiques induites par la présence d'une mitochondrie exogène ne se limitent pas seulement aux complexes protéiques encodés par le génome nucléaire et mitochondrial. Sa surexpression chez la forme sauvage suggère que le complexe II de la chaîne respiratoire est potentiellement autant sensible à la présence d'une mitochondrie exogène que les autres complexes. Une sous-expression de la SDH chez la forme cybride pourrait aussi avoir des conséquences importantes au niveau de l'expression d'autres gènes nucléaires. En effet, d'après une étude de Cervera *et al.*, (2009), la diminution de l'expression de la SDH dans des cellules de mammifères entraîne l'augmentation de la méthylation des histones H3.

#### **4.4.4 Apolipoprotéine AI2**

Bien que la différence du taux d'expression chez la forme sauvage ne soit pas excessive avec un taux de 2,2 fois. Le précurseur de l'ApoA est exprimé plus abondamment chez la forme sauvage que chez la forme cybride (D19 dans le tableau IV). Les apolipoprotéines participent au transport des molécules/protéines hydrophobes. Lorsqu'une apolipoprotéine interagit avec une gouttelette lipidique, elle forme une lipoprotéine. Une étude a démontré que les lipoprotéines contiendraient une grande quantité de protéines mitochondriales intervenant dans le cycle de Krebs (Zhang *et al.*, 2011 b). Une autre étude a démontré que la séquence de récepteurs de HDL (lipoprotéine de haute densité) est très similaire à la séquence de la chaîne bêta de l'ATP synthase. Le fait que ces récepteurs soit reconnus pour avoir une forte affinité pour l'ApoA et que l'immunofluorescence montre une colocalisation de l'ApoA et de la chaîne bêta de l'ATP synthase démontre qu'il y a une interaction entre ces deux protéines

(Martinez *et al.*, 2003). Si cette protéine assume un rôle d'importation de protéines mitochondriales liées au cycle de Krebs (Zhang *et al.*, 2011 b), la différence physiologique pourrait bien être plus importante que sa différence d'expression.

#### ***4.4.5 Protéines cytosquelettiques***

Les deux dernières protéines surexprimées chez la forme sauvage sont la myozénine et l' $\alpha$ -tropomyosine. Il existerait au moins trois isoformes de tropomyosine et de myozénine dans les cellules de muscles squelettiques (Zhang *et al.*, 2011 a). Il est possible que ces protéines soient régulées différemment entre nos biotypes, ce qui pourrait induire une différence au niveau du ratio de chaque isoforme participant à la formation du cytosquelette. Un élément qui pourrait soutenir cette hypothèse est le fait que l'impact de la température sur la stabilité de la tropomyosine varierait dépendamment de l'isoforme (Huang et Ochiai 2005, Hayley *et al.*, 2011). La tropomyosine est toujours retrouvée couplée aux filaments d'actine dans le cytoplasme (Mannherz et Goody 1976). Dans les muscles striés, la position de la tropomyosine sur le filament d'actine est dépendante de la conformation de la troponine, elle même modulable en présence de  $Ca^{2+}$  (Dominguez 2011).

La myozénine également nommée FATZ, est une protéine cytosquelettique participant à la formation des lignes Z qui séparent les sarcomères des myofibres (Wang *et al.*, 2005, Stout *et al.*, 2008). Elle interagit avec l' $\alpha$ -actinine et la  $\gamma$ -filamine et est majoritairement exprimée dans les muscles squelettiques (Takada *et al.*, 2001). En raison d'une quantité de peptides uniques ainsi qu'un pourcentage de couverture faible, nous ne pouvons rejeter la possibilité que l'échantillon D20 ait été mal identifié. Dans la mesure où la myozénine 1 s'est placée en première position après les protéines cytosquelettiques contaminantes retrouvées en grande quantité dans la majorité des échantillons séquencés (annexe 1) nous avons tout de même pris la décision de la présenter dans nos résultats.

On retrouve chez la forme cybride deux points spécifiques correspondant à la MLC reconnue pour leur association au cytosquelette et plus particulièrement, à la contraction musculaire. Ces points sont uniquement retrouvés chez la forme cybride. La différence existant entre le point isoélectrique et le poids moléculaire avec le point commun aux deux biotypes identifiés par un « x » sur la figure 8 et correspondant également à la même protéine montre que la forme cybride présente une forme modifiée de cette protéine. La littérature associe les modifications post-traductionnelles à une modification de l'activité de cette protéine, par conséquent ce résultat suggère qu'il existe une différence entre nos deux biotypes au niveau des propriétés contractiles du muscle (Szczesna *et al.*, 2002).

Il est cependant possible que la différence observée pour nos deux points chez la forme cybride soit due à la présence d'isoformes spécifiques, puisque plusieurs isoformes ont été identifiés pour la MLC. Selon une étude, les isoformes exprimées pour la MLC varieraient en réponse à l'activité musculaire (Wahrmann *et al.*, 2001). Par conséquent même si les points C2 et C3 correspondaient à des isoformes différents que ceux retrouvés chez la forme sauvage, ils impliqueraient tout de même une différence au niveau de l'activité physique entre les biotypes.

#### **4.4.6 Triose-phosphate isomérase B**

Le point correspondant à la TPI est celui qui montre la plus grande différence d'abondance dans l'analyse quantitative de nos gels 2D avec un ratio supérieur à 10 :1 en faveur de la forme cybride (D1, tableau IV). La TPI B participe à la formation du dimère enzymatique formant la TPI. Cette enzyme intervient au niveau de la 5<sup>ème</sup> étape de la glycolyse et catalyse la réaction permettant la transformation du dihydroxyacétone phosphate en D-glycéraldéhyde-3-phosphate (Karp 2009). Certaines études ont montré que l'activité enzymatique de la TPI est tellement efficace qu'elle est reconnue comme étant une enzyme parfaite (Albery et Knowles 1976). Ces enzymes possèdent une vitesse de réaction qui n'est limitée que par la

vitesse de diffusion des produits et des substrats autour de leur site catalytique (Berg *et al.*, 2002). La différence des points isoélectriques observée entre la protéine théorique et la protéine échantillonnée semble soulever la présence de modifications post-traductionnelles pour la TPI. Certaines études ont montré que la TPI peut dans certaines conditions être modifiée post-traductionnellement, les conséquences de ces modifications chez la TPI demeurent cependant méconnues (Skala-Rubinson *et al.*, 1976, Tomonaga *et al.*, 2004).

Au niveau de la fonction de la TPI, une étude sur des souches mutantes de *Saccharomyces cerevisiae* ayant une déficience de la TPI montre que cette enzyme est associée à la production de glycérol (Compagno *et al.*, 1996). Les résultats montrent qu'une baisse de l'activité enzymatique de la TPI induit une hausse de la production de glycérol par rapport aux souches non mutées. Dans notre contexte, la différence d'expression observée entre les deux formes pour cette enzyme pourrait être un élément révélateur d'une production de glycérol accrue chez la forme sauvage.

#### 4.4.7 Guanidinoacétate N-méthyltransférase

Le premier point unique à la forme cybride a été identifié comme étant la guanidinoacétate N-méthyltransférase (GAMT). La GAMT est une protéine participant à la formation de la créatine (Moussard 2006). La majorité des études sur la GAMT portent sur les conséquences phénotypiques associées à une déficience de cette protéine conduisant à un déficit en biosynthèse de la créatine dont les symptômes sont variables. On retrouve peu d'études dans la littérature scientifique sur la GAMT qui ne soit pas liées à un déficit en biosynthèse de la créatine. À notre connaissance il n'y a pas de référence dans la littérature à des modifications post-traductionnelles de la GAMT. Les informations dont nous disposons sur la GAMT indiquent qu'elle serait majoritairement exprimée dans le foie et la rate chez d'autres poissons tels que *Cyprinus carpio* et *Ictiobus cyprinella* (Van Pilsum *et al.*, 1972). La biosynthèse de la créatine se déroule en deux étapes successives, et la GAMT assure la seconde d'entre elles.

Elle consiste en la méthylation du guanidinoacétate et son produit est la créatine (Wyss et Kaddurah-Daouk 2000). La créatine kinase phosphoryle ensuite la créatine en phosphocréatine, une source énergétique mobilisable pour régénérer la réserve d'ATP cellulaire nécessaire à la contraction cellulaire (Bittl et Ingwall 1985). La créatine kinase existe sous quatre formes cellulaires dont deux d'entre elles sont localisées dans la mitochondrie (Schlegel *et al.*, 1988, Schlattner *et al.*, 1998).

La phosphocréatine formée dans la mitochondrie est ainsi rapatriée dans le cytosol selon un système de navette de transport de la phosphocréatine (Bessman et Carpenter 1985). Cette navette permettrait de maintenir le niveau de phosphocréatine cytosolique nécessaire à la régénération de l'ATP. C'est pour cette raison que la GAMT et le métabolisme de la créatine qui en découle ont été longtemps associés à l'activité physique. Cependant, une étude récente de Lygate *et al.*, (2013) montre que l'absence de créatine chez des souris K.O pour la GAMT n'induit pas une baisse des capacités physiques maximales, ce qui démontre de l'existence de mécanismes cellulaires alternatifs pour régénérer l'ATP. Par conséquent, on ne peut conclure que la présence unique du point associé à la GAMT dans les gels 2D de la forme cybride serait la résultante d'une différence au niveau des capacités physiques des biotypes.

Une autre explication potentielle qui pourrait justifier la différence d'expression de la GAMT est liée aux propriétés d'antioxydant de la créatine. Une étude de Lawler *et al.*, (2002) montre que la créatine est capable de retirer des dérivés réactifs de l'oxygène tels que, l'acide 2,2-azobis-3-éthylbenzthiazoline-6-sulfonique (ABTS), les ions superoxyde ( $O_2^{-2}$ ) et les ions peroxynitrite ( $ONOO^-$ ). Dans la mesure où la mitochondrie étrangère chez la forme cybride induit des altérations physiologiques similaires à ce qui est observé chez les cybrides artificiels de McKenzie *et al.*, (2003), et que ces altérations impliquent une augmentation du stress oxydatif. La différence d'expression au niveau de la GAMT pourrait théoriquement être liée à une production accrue de créatine cellulaire afin de limiter le stress oxydatif.

Pour conclure sur les conséquences découlant de la différence d'expression interbiotype pour la GAMT. La diminution du stress oxydatif résultant d'une augmentation de la production de créatine chez la forme cybride serait un argument plus soutenu par la littérature scientifique que l'augmentation des capacités ou de la masse musculaire.

#### 4.5 Protéines spécifiques et thermorégulation

Comme la présence d'une mitochondrie exogène chez la forme cybride semble modifier sa tolérance au froid (Mee et Taylor 2012), on pourrait s'attendre à ce que les différences protéomiques observées reflètent cette différence. Mise en relation avec le fait que la zone d'échantillonnage que nous avons utilisée pour ce projet a la particularité de se situer à la frontière entre les aires de répartition spécifiques à chaque biotype (point orange, figure 3). Nous nous retrouvons à comparer dans nos gels des individus prélevés (sur la même période) de la forme sauvage provenant de la partie septentrionale de leur aire de répartition, alors que les individus de la forme cybride proviennent de la partie méridionale de leur aire de répartition. Cette distinction semble avoir eu une conséquence inattendue au niveau de l'apparition d'un patron d'expression protéique correspondant à une réponse à l'acclimatation au froid.

La triose-phosphate isomérase B correspond à la protéine présentant la plus grande différence d'expression entre les biotypes (D1, tableau IV). Cette protéine est sous-exprimée chez la forme sauvage avec un ratio supérieur à 10 :1 en faveur de la forme cybride. Nous avons discuté précédemment de son rôle dans la production glycérol dans les souches de *S. cerevisiae* de Compagno *et al.*, (1996) et de la possibilité au vu de la différence d'expression existant pour cette protéine que la forme sauvage ait une production de glycérol plus importante que la forme cybride. Une étude a montré qu'en réponse à une exposition de 72h à 4°C, les cellules de foie de l'éperlan d'Amérique (*Osmerus mordax*) démontrent une



accumulation de glycérol (Hall *et al.*, 2011). Dans ce contexte, il est possible que l'espèce sauvage produise davantage de glycérol en réponse à l'acclimatation au froid.

La SDH fait partie des protéines présentant la plus grande différence d'expression entre les biotypes (D2, tableau IV). En dehors du fait que la surexpression de cette protéine chez *Danio rerio* est observée en réponse à une période d'acclimatation au froid (Malek *et al.*, 2004), elle participe comme nous en avons déjà parlé dans l'introduction à la formation du complexe II du système OXPHOS. Par conséquent sa surexpression pourrait également indiquer la forme sauvage présente une augmentation de la densité des mitochondries ou de l'activité enzymatique des protéines mitochondriales préalablement reportée dans d'autres études comme des adaptations face à l'acclimatation au froid (Egginton et Sidell 1989, Sängner 1993).

Dans d'autres études, une augmentation de l'oxydation des lipides et de la quantité de gouttelettes lipidiques dans les fibres musculaires du bar rayé (*Morone saxatilis*) est observée lorsque ces poissons sont acclimatés à 5°C (Jones et Sidell 1982, Egginton et Sidell 1989). Ces modifications pourraient potentiellement être associées à l'apolipoprotéine AI2 et FABP3. L'augmentation de l'expression de FABP 3 au niveau transcriptomique et protéomique observée suite à une période d'acclimatation à 4°C et son rôle dans la thermorégulation supporterait notre hypothèse (Londrville et Sidell 1996, Vergnes *et al.*, 2011).

Concernant l'énolase, à ce jour aucune étude n'a démontré que l'acclimatation au froid augmentait l'abondance de cette protéine dans le muscle squelettique. Nous avons déjà indiqué que le gène encodant pour cette protéine est induit par le facteur d'induction hypoxique HIF-1 (Semenza *et al.*, 1996). Il a été démontré que la sous-unité alpha de la protéine HIF-1 était surexprimée à faible température (Rissanen *et al.*, 2006). Cela suggère donc que l'énolase pourrait également être indirectement surexprimée en réponse à l'acclimatation au froid.

Au niveau transcriptomique, l'alpha-tropomyosine et la myozénine ne sont pas surexprimées dans le muscle en réponse au froid (Gracey *et al.*, 2004). Plusieurs études ont cependant montré que les propriétés contractiles du muscle de *Cyprinus carpio* et *Myoxocephalus scorpius* sont modifiées en réponse à l'acclimatation au froid (Johnston *et al.*, 1990, Temple et Johnston 1998, Temple *et al.*, 2000). Cela suggère que la différence d'abondance observée entre nos deux biotypes pour l'alpha-tropomyosine et la myozénine découlerait d'une différence au niveau de la traduction ou de la dégradation des protéines.

Concernant la MLC, nous avons déjà évoqué que les isoformes exprimés varient en réponse à l'activité musculaire dans l'étude de Wahrmann *et al.*, (2001). Une autre étude montre que la vitesse de nage est dépendante de la température de l'eau chez des saumons coho (*Oncorhynchus kisutch*) et qu'on détecte la vitesse de nage la plus faible à 3°C (Griffiths et Alderdice 1972, Randall et Brauner 1991). Par conséquent, il est possible que la différence d'activité musculaire observée entre les biotypes soit la conséquence d'une différence de sensibilité au froid entre les biotypes. Maintenant, si la différence d'expression au niveau de la MLC est dû à des modifications post-traductionnelles, alors on sait par défaut qu'on a une différence au niveau de l'activité de cette protéine et que cette différence à des répercussions sur les propriétés contractiles musculaires entre les biotypes.

Comme les individus de la forme sauvage ne sont pas retrouvés dans la partie nordique de l'aire de répartition de *C. eos*, il se peut que la latitude à laquelle nous prélevons les individus de la forme sauvage corresponde à leur limite physiologique face à la tolérance au froid. Si cette hypothèse s'avérait vraie, il serait logique que les individus appartenant à la forme sauvage et vivant dans notre zone d'échantillonnage présentent une série d'adaptation physiologique leur permettant de s'acclimater à des températures plus basses contrairement à des individus de la forme cybride vivant dans les lacs voisins de notre zone de prélèvement. De ce fait, il serait cohérent de retrouver dans nos résultats une série de modifications

protéomiques résultant d'une adaptation physiologique face à l'acclimatation au froid chez la forme sauvage. Une hypothèse permettant d'expliquer la réponse protéomique associée à l'acclimatation au froid observée chez la forme sauvage serait basée sur le fait ce biotype serait plus sensible au froid que la forme cybride.

## **5 Conclusion**

Lors de ce projet, nous avons voulu découvrir la nature des protéines impliquées dans la différence interbiotype au niveau du protéome du muscle entre la forme sauvage et la forme cybride du complexe *Chrosomus eos-neogaeus* observées par Angers *et al.*, (2012). Dans cette optique nous avons fait une analyse comparative sur gels 2D entre nos deux biotypes, en augmentant significativement l'effort d'échantillonnage et en ajoutant un volet quantitatif ainsi qu'une analyse par spectrométrie de masse de certaines protéines. Ainsi, nous avons pu confirmer la différence interbiotype qualitative obtenue dans l'article d'Angers *et al.*, (2012) et l'existence d'une différence quantitative interbiotype.

Le fait que la majorité des protéines spécifiques identifiées ne soit pas des protéines mitochondriales soulève le fait que la mitochondrie étrangère chez la forme cybride induit une reprogrammation de l'expression génique dans le génome nucléaire. La différence de stratégie physiologique employée par chaque biotype au niveau de la glycolyse, de la formation du cytosquelette et de la réponse au froid est évocatrice de l'impact de la mitochondrie sur ces voies métaboliques. Nous avons également montré que pour un sous-groupe de protéines, le patron d'expression protéomique est associé dans la littérature à une réponse à l'acclimatation au froid chez la forme sauvage. Les lacs des deux biotypes étant proches, cela suggérerait qu'à des conditions climatiques semblables seule la forme sauvage présente une modification de l'expression de son protéome caractéristique d'une réponse à l'acclimatation au froid. Cela soulève l'hypothèse que la forme sauvage serait plus sensible que la forme cybride au froid.

Nous pourrions vérifier cette hypothèse en comparant le protéome des individus de la forme sauvage de notre zone d'échantillonnage avec des individus de la forme sauvage vivant dans la partie méridionale de ce biotype. Si le patron protéomique des individus méridionaux se rapproche du patron protéomique de la forme cybride dans notre zone d'échantillonnage, cela nous permettrait de confirmer notre hypothèse.

Les acteurs cellulaires permettant à la mitochondrie *C. neogaeus* de modifier la régulation génique chez la forme cybride restent encore inconnus. Une hypothèse au vu de nos résultats serait que la présence de la mitochondrie exogène chez la forme cybride entraîne une baisse de couplage entre les complexes chimériques du système OXPHOS. L'altération du système OXPHOS entraîne une augmentation de la quantité des protons relâchés et de la chaleur produite par la mitochondrie. Le relâchement de protons par la chaîne respiratoire étant associé au stress oxydatif, il est fort probable qu'il y aurait une augmentation de la production de dérivé réactif de l'oxygène. La différence d'expression observée au niveau de la GAMT et de la créatine qu'elle synthétise permettrait dans ce contexte de diminuer les effets néfastes associés au stress oxydatif chez la forme cybride. De plus, nous avons vu que certaines protéines sous-exprimées chez la forme cybride pouvaient entraîner des modifications épigénétiques, mais il est probable que d'autres mécanismes interviennent en amont. Ces mécanismes se feraient probablement au niveau de la communication intergénomique entre la mitochondrie et le noyau. De nouvelles études seront donc nécessaires pour avoir une meilleure compréhension des mécanismes responsables de la différence de phénotype entre la forme sauvage et la forme cybride.

## Références

- Albery, W. J. and J. R. Knowles (1976). "Free-energy profile of the reaction catalyzed by triosephosphate isomerase." Biochemistry 15(25): 5627-5631.
- Amuthan, G., G. Biswas, H. K. Ananadatheerthavarada, C. Vijayasarathy, H. M. Shephard and N. G. Avadhani (2002). "Mitochondrial stress-induced calcium signaling, phenotypic changes and invasive behavior in human lung carcinoma A549 cells." Oncogene 21(51): 7839-7849.
- Andersson, S. G., A. Zomorodipour, J. O. Andersson, T. Sicheritz-Ponten, U. C. Alsmark, R. M. Podowski, A. K. Naslund, A. S. Eriksson, H. H. Winkler and C. G. Kurland (1998). "The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria." Nature 396(6707): 133-140.
- Angers, B., Dallaire A, Vervaeet S and Angers A (2012). "The influence of mitochondria in epigenetics revealed through naturally occurring fish cybrids." Current Zoology 58: 138-145.
- Angers, B. and I. J. Schlosser (2007). "The origin of *Phoxinus eos-neogaeus* unisexual hybrids." Molecular Ecology 16(21): 4562-4571.
- Anttila, K., S. Mänttari and M. Järvilehto (2006). "Effects of different training protocols on Ca<sup>2+</sup> handling and oxidative capacity in skeletal muscle of Atlantic salmon (*Salmo salar* L.)." Journal of Experimental Biology 209(15): 2971-2978.
- Aubin-Horth, N. and S. C. Renn (2009). "Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity." Molecular Ecology 18(18): 3763-3780.
- Barrett, J., P. M. Brophy and J. V. Hamilton (2005). "Analysing proteomic data." International Journal of Parasitology. 35: 543.
- Berg, J. M., J. L. Tymoczko and L. Stryer (2002). Biochemistry, Fifth Edition, W.H. Freeman.
- Bessman, S. P. and C. L. Carpenter (1985). "The creatine-creatine phosphate energy shuttle." Annual Review of Biochemistry 54: 831-862.
- Binet, M. C. and B. Angers (2005). "Genetic identification of members of the *Phoxinus eos-neogaeus* hybrid complex." Journal of Fish Biology 67(4): 1169-1177.
- Biron, D. G., L. Marché, F. Ponton, H. D. Loxdale, N. Galéotti, L. Renault, C. Joly and F. Thomas (2005). "Behavioural manipulation in a grasshopper harbouring hairworm: a proteomics approach." Proceedings of the Royal Society B: Biological Sciences 272(1577): 2117-2126.
- Bittl, J. A. and J. S. Ingwall (1985). "Reaction rates of creatine kinase and ATP synthesis in the isolated rat heart. A 31P NMR magnetization transfer study." Journal of Biological Chemistry 260(6): 3512-3517.

- Black, D. L. (2003). "Mechanisms of alternative pre-messenger RNA splicing." Annual Review of Biochemistry 72: 291-336.
- Bottalico, L. A., N. C. Kendrick, A. Keller, Y. Li and I. Tabas (1993). "Cholesteryl ester loading of mouse peritoneal macrophages is associated with changes in the expression or modification of specific cellular proteins, including increase in an alpha-enolase isoform." Arterioscler Thromb 13(2): 264-275.
- Brandina, I., J. Graham, C. Lemaitre-Guillier, N. Entelis, I. Krasheninnikov, L. Sweetlove, I. Tarassov and R. P. Martin (2006). "Enolase takes part in a macromolecular complex associated to mitochondria in yeast." Biochimica et Biophysica Acta 1757(9-10): 1217-1228.
- Byrjalsen, I., P. Mose Larsen, S. Fey, L. Nilas, M. Larsen and C. Christiansen (1999). "Two-dimensional gel analysis of human endometrial proteins: characterization of proteins with increased expression in hyperplasia and adenocarcinoma." Molecular Human Reproduction 5: 748 - 756.
- Capello, M., S. Ferri-Borgogno, P. Cappello and F. Novelli (2011). "alpha-Enolase: a promising therapeutic and diagnostic tumor target." The FEBS journal 278(7): 1064-1074.
- Compagno, C., F. Boschi and B. M. Ranzi (1996). "Glycerol production in a triose phosphate isomerase deficient mutant of *Saccharomyces cerevisiae*." Biotechnology Progress 12(5): 591-595.
- Cooper, J. A., F. S. Esch, S. S. Taylor and T. Hunter (1984). "Phosphorylation sites in enolase and lactate dehydrogenase utilized by tyrosine protein kinases in vivo and in vitro." Journal of Biological Chemistry 259(12): 7835-7841.
- Danchin, E., A. Charmantier, F. A. Champagne, A. Mesoudi, B. Pujol and S. Blanchet (2011). "Beyond DNA: integrating inclusive inheritance into an extended theory of evolution." Nature reviews. Genetics 12(7): 475-486.
- Dawley, R., Schultz R and Goddard K (1987). "Clonal Reproduction and Polyploidy in Unisexual Hybrids of *Phoxinus eos* and *Phoxinus neogaeus* (Pisces; Cyprinidae)." Copeia 1987(2): 275-283.
- DiMauro, S. and E. A. Schon (2003). "Mitochondrial respiratory-chain diseases." The New England journal of medicine 348(26): 2656-2668.
- Dimmer, K. S. and L. Scorrano (2006). "(De)constructing mitochondria: what for?" Physiology (Bethesda) 21: 233-241.
- Diz, A. P., M. Martinez-Fernandez and E. Rolan-Alvarez (2012). "Proteomics in evolutionary ecology: linking the genotype with the phenotype." Molecular Ecology 21(5): 1060-1080.
- Dominguez, R. (2011). "Tropomyosin: The Gatekeeper's View of the Actin Filament Revealed." Biophysical Journal 100(4): 797-798.

- Egginton, S. and B. D. Sidell (1989). "Thermal acclimation induces adaptive changes in subcellular structure of fish skeletal muscle." American Journal of Physiology 256(1 Pt 2): R1-9.
- ENCODE (2009). "Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs." Nature 457(7232): 1028-1032.
- Feder, M. E. and J. C. Walser (2005). "The biological limitations of transcriptomics in elucidating stress and stress responses." Journal of Evolutionary Biology 18(4): 901-910.
- Fischer, H., Gustafsson T, Sundberg C J, Norrbom J, Ekman M, Johansson O and Jansson E (2006). "Fatty acid binding protein 4 in human skeletal muscle." Biochemical and Biophysical Research Communications 346(1): 125-130.
- Flaubert, G. (1893). Correspondance, Ed.
- Gershoni, M., A. R. Templeton and D. Mishmar (2009). "Mitochondrial bioenergetics as a major motive force of speciation." Bioessays 31(6): 642-650.
- Gibert, J. M., F. Peronnet and C. Schlotterer (2007). "Phenotypic plasticity in *Drosophila* pigmentation caused by temperature sensitivity of a chromatin regulator network." PLoS Genetics 3(2): e30.
- Glémet, H., P. Blier and L. Bernatchez (1998). "Geographical extent of Arctic char (*Salvelinus alpinus*) mtDNA introgression in brook char populations (*S. fontinalis*) from eastern Québec, Canada." Molecular Ecology 7(12): 1655-1662.
- Glickman, M. H. and A. Ciechanover (2002). "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." Physiological Reviews 82(2): 373-428.
- Goddard, K., O. Megwinoff, L. Wessner and F. Giaimo (1998). "Confirmation of gynogenesis in *Phoxinus eos-neogaeus* (Pisces: Cyprinidae)." Journal of Heredity 89(2): 151-157.
- Goddard, K. and R. Schultz (1993). Aclonal reproduction by polyploid members of the clonal hybrid species *Phoxinus eos-neogaeus* (Cyrinidae). Lawrence, KS, ETATS-UNIS, American Society of Ichthyologists and Herpetologists.
- Goerg, A., Weiss W and Dunn M J (2004). "Current two-dimensional electrophoresis technology for proteomics." Proteomics 4(12): 3665-3685.
- Gracey, A., E. Fraser, W. Li, Y. Fang, R. Taylor, J. Rogers, A. Brass and A. Cossins (2004). "Coping with cold: An integrative, multitissue analysis of the transcriptome of a poikilothermic vertebrate." Proceedings of the National Academy of Sciences of the United States of America 101(48): 16970-16975.
- Gracey, A., Troll J V and Somero G N (2001). "Hypoxia-induced gene expression profiling in the euryoxic fish *Gillichthys mirabilis*." Proceedings of the National Academy of Sciences 98(4): 1993-1998.

- Gray, M. W., Burger G and Lang BF (2001). "The origin and early evolution of mitochondria." Genome Biology 2(6): 1-5.
- Griffiths, J. S. and D. F. Alderdice (1972). "Effects of Acclimation and Acute Temperature Experience on the Swimming Speed of Juvenile Coho Salmon." Journal of the Fisheries Research Board of Canada 29(3): 251-264.
- Guderley, H. (1990). "Functional significance of metabolic responses to thermal acclimation in fish muscle." American Journal of Physiology 259(2 Pt 2): R245-252.
- Guderley, H. (2004). "Metabolic responses to low temperature in fish muscle." Biological reviews of the Cambridge Philosophical Society 79(2): 409-427.
- Guilherme, N., Jacob Z, Mark D and Andrew C (2004). "Stochastic yet biased expression of multiple Dscam splice variants by individual cells." Nature Genetics 36(3): 240-246.
- Gygi, S. P., Y. Rochon, B. R. Franza and R. Aebersold (1999). "Correlation between Protein and mRNA Abundance in Yeast." Molecular and Cellular Biology 19(3): 1720-1730.
- Halfmann, R. and S. Lindquist (2010). "Epigenetics in the extreme: prions and the inheritance of environmentally acquired traits." Science 330(6004): 629-632.
- Hall, J. R., K. A. Clow, M. L. Rise and W. R. Driedzic (2011). "Identification and validation of differentially expressed transcripts in a hepatocyte model of cold-induced glycerol production in rainbow smelt (*Osmerus mordax*)." American Journal of Physiology - Regulatory, Integrative and Comparative Physiology 301(4): R995-r1010.
- Hayley, M., T. Chevaldina and D. H. Heeley (2011). "Cold adaptation of tropomyosin." Biochemistry 50(30): 6559-6566.
- Haynes, C. M. and D. Ron (2010). "The mitochondrial UPR - protecting organelle protein homeostasis." J Cell Sci 123(Pt 22): 3849-3855.
- Huang, M. C., M. A. Ayliffe and J. N. Timmis (2003). "Direct measurement of the transfer rate of chloroplast DNA into the nucleus." Nature 422(6927): 72-76.
- Huang, M. C. and Y. Ochiai (2005). "Fish fast skeletal muscle tropomyosins show species-specific thermal stability." Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 141(4): 461-471.
- Human Genome Sequencing, C. (2004). "Finishing the euchromatic sequence of the human genome." Nature 431(7011): 931-945.
- Hunt, S. M., M. R. Thomas, L. T. Sebastian, S. K. Pedersen, R. L. Harcourt, A. J. Sloane and M. R. Wilkins (2005). "Optimal replication and the importance of experimental design for gel-based quantitative proteomics." Journal of Proteome Research 4: 809.



- Johnston, I. A., J. D. Fleming and T. Crockford (1990). "Thermal acclimation and muscle contractile properties in cyprinid fish." American Journal of Physiology 259(2 Pt 2): R231-236.
- Jones, P. D., A. Carne, N. M. Bass and M. R. Grigor (1988). "Isolation and characterization of fatty acid binding proteins from mammary tissue of lactating rats." The Biochemical journal 251(3): 919-925.
- Jones, P. D. and B. Sidell (1982). "Metabolic responses of striped bass (*Morone saxatilis*) to temperature acclimation. II. Alterations in metabolic carbon sources and distributions of fiber types in locomotory muscle." Journal of Experimental Zoology 219(2): 163-171.
- Karp, G. (2009). Cell and Molecular Biology: Concepts and Experiments, John Wiley & Sons.
- Karr, T. L. (2008). "Application of proteomics to ecology and population biology." Heredity (Edinb) 100(2): 200-206.
- Knight, C. G., N. Zitzmann, S. Prabhakar, R. Antrobus, R. Dwek, H. Hebestreit and P. B. Rainey (2006). "Unraveling adaptive evolution: how a single point mutation affects the protein coregulation network." Nature reviews. Genetics 38(9): 1015-1022.
- Lang, B. F., M. W. Gray and G. Burger (1999). "Mitochondrial genome evolution and the origin of eukaryotes." Annual review of Genetics 33: 351-397.
- Lee, H. J., T. A. Hore and W. Reik (2014). "Reprogramming the Methylome: Erasing Memory and Creating Diversity." Cell Stem Cell 14(6): 710-719.
- Leister, D. (2005). "Origin, evolution and genetic effects of nuclear insertions of organelle DNA." Trends in Genetics 21(12): 655-663.
- Li, E. (2002). "Chromatin modification and epigenetic reprogramming in mammalian development." Nature reviews. Genetics 3(9): 662-673.
- Lindquist, S. (2009). "Protein folding sculpting evolutionary change." Cold Spring Harb Symp Quant Biol 74: 103-108.
- Lodish, H., Berk A, Zipursky SL, Matsudaira P, Baltimore D and D. J. (2000). Molecular Cell Biology, New York: W. H. Freeman: Section 15.15.
- Londraville, R. L. and B. D. Sidell (1996). "Cold acclimation increases fatty acid-binding protein concentration in aerobic muscle of striped bass, *Morone saxatilis*." Journal of Experimental Zoology 275(1): 36-44.
- Lopez, J., E. Mosquera, J. Fuentes, A. Marina, J. Vazquez and G. Alvarez (2001). "Two-dimensional gel electrophoresis of *Mytilus galloprovincialis*. Differences in protein expression between intertidal and cultured mussels." Marine Ecology Progress Series 224: 149 - 156.

- López, J. L., A. Marina, G. Alvarez and J. Vazquez (2002). "A proteomic approach to the study of the marine mussels, *Mytilus edulis* and *Mytilus galloprovincialis*." Proteomics 2: 1658.
- Lotz, C., A. J. Lin, C. M. Black, J. Zhang, E. Lau, N. Deng, Y. Wang, N. C. Zong, J. H. Choi, T. Xu, D. A. Liem, P. Korge, J. N. Weiss, H. Hermjakob, J. R. Yates, 3rd, R. Apweiler and P. Ping (2014). "Characterization, design, and function of the mitochondrial proteome: from organs to organisms." Journal of Proteome Research 13(2): 433-446.
- Makowsky, R., Pajewski N M, Klimentidis Y C, Vazquez A I, Duarte C W, Allison D B and de los Campos G (2011). "Beyond Missing Heritability: Prediction of Complex Traits." PLoS Genetics 7(4): e1002051.
- Malek, R. L., Sajadi H, Abraham J, Grundy M A and Gerhard G S (2004). "The effects of temperature reduction on gene expression and oxidative stress in skeletal muscle from adult zebrafish." Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 138(3): 363-373.
- Mannherz, H. G. and R. S. Goody (1976). "Proteins of Contractile Systems." Annual Review of Biochemistry 45(1): 427-466.
- Martinez-Fernandez, M., M. P. de la Cadena and E. Rolan-Alvarez (2010). "The role of phenotypic plasticity on the proteome differences between two sympatric marine snail ecotypes adapted to distinct micro-habitats." BMC evolutionary biology 10: 65.
- Martínez-Fernández, M., A. M. Rodríguez-Piñeiro, E. Oliveira, M. Páez de la Cadena and E. Rolán-Alvarez (2008). "Proteomic Comparison between Two Marine Snail Ecotypes Reveals Details about the Biochemistry of Adaptation." Journal of Proteome Research 7(11): 4926-4934.
- Martinez, L. O., Jacquet S, Esteve J P, Rolland C, Cabezon E, Champagne E, Pineau T, Georgeaud V, Walker J E, Terce F, Collet X, Perret B and Barbaras R (2003). "Ectopic [beta]-chain of ATP synthase is an apolipoprotein A-I receptor in hepatic HDL endocytosis." Nature 421(6918): 75-79.
- Massicotte, R. and B. Angers (2012). "General-purpose genotype or how epigenetics extend the flexibility of a genotype." Genetics Research International 2012: 317175.
- Matarese, V., Stone R L, Waggoner D W and Bernlohr D A (1989). "Intracellular fatty acid trafficking and the role of cytosolic lipid binding proteins." Progress in Lipid Research 28(4): 245-272.
- McGhee, J. D. and G. D. Ginder (1979). "Specific DNA methylation sites in the vicinity of the chicken [beta]-globin genes." Nature 280(5721): 419-420.
- McKenzie, M., M. Chiotis, C. A. Pinkert and I. A. Trounce (2003). "Functional respiratory chain analyses in murid xenomitochondrial cybrids expose coevolutionary constraints of cytochrome b and nuclear subunits of complex III." Molecular Biology and Evolution 20(7): 1117-1124.

- McKenzie, M. and I. Trounce (2000). "Expression of *Rattus norvegicus* mtDNA in *Mus musculus* cells results in multiple respiratory chain defects." Journal of Biological Chemistry 275(40): 31514-31519.
- Mee, J. A. and E. B. Taylor (2012). "The cybrid invasion: widespread postglacial dispersal by *Phoxinus* (Pisces: Cyprinidae) cytoplasmic hybrids." Canadian Journal of Zoology 90(5): 577-584.
- Mochel, F. and R. G. Haller (2011). "Energy deficit in Huntington disease: why it matters." The Journal of Clinical Investigation 121(2): 493-499.
- Moussard, C. (2006). Biochimie structurale et métabolique, De Boeck Supérieur.
- Mukhopadhyay, D. and H. Riezman (2007). "Proteasome-independent functions of ubiquitin in endocytosis and signaling." Science 315(5809): 201-205.
- Nachmansohn, D. and A. L. Machado (1943). "The formation of acetylcholine, a new enzyme: "Choline acetylase"." Journal of Neurophysiology 6(5): 397-403.
- Nie, L., G. Wu and W. Zhang (2006). "Correlation between mRNA and protein abundance in *Desulfovibrio vulgaris*: A multiple regression to identify sources of variations." Biochemical and Biophysical Research Communications 339(2): 603-610.
- Nielsen, S. U. and F. Spener (1993). "Fatty acid-binding protein from rat heart is phosphorylated on Tyr19 in response to insulin stimulation." Journal of Lipid Research 34(8): 1355-1366.
- O'Connell, K. and K. Ohlendieck (2009). "Proteomic DIGE analysis of the mitochondria-enriched fraction from aged rat skeletal muscle." Proteomics 9(24): 5509-5524.
- Oyedotun, K. S. and B. D. Lemire (2004). "The Quaternary Structure of the *Saccharomyces cerevisiae* Succinate Dehydrogenase: HOMOLOGY MODELING, COFACTOR DOCKING, AND MOLECULAR DYNAMICS SIMULATION STUDIES." Journal of Biological Chemistry 279(10): 9424-9431.
- Pancholi, V. (2001). "Multifunctional alpha-enolase: its role in diseases." Cellular and Molecular Life Sciences 58(7): 902-920.
- Price, T. D., A. Qvarnstrom and D. E. Irwin (2003). "The role of phenotypic plasticity in driving genetic evolution." Proceedings. Biological Sciences 270(1523): 1433-1440.
- Queitsch, C., S. W. Hong, E. Vierling and S. Lindquist (2000). "Heat shock protein 101 plays a crucial role in thermotolerance in *Arabidopsis*." The Plant Cell 12(4): 479-492.
- Queitsch, C., T. A. Sangster and S. Lindquist (2002). "Hsp90 as a capacitor of phenotypic variation." Nature 417(6889): 618-624.

- Rand, D. M., R. A. Haney and A. J. Fry (2004). "Cytonuclear coevolution: the genomics of cooperation." Trends in Ecology & Evolution 19(12): 645-653.
- Randall, D. and C. Brauner (1991). "Effects of Environmental Factors on Exercise in Fish." Journal of Experimental Biology 160(1): 113-126.
- Richert, S., S. Luche, M. Chevallet, A. Van Dorselaer, E. Leize-Wagner and T. Rabilloud (2004). "About the mechanism of interference of silver staining with peptide mass spectrometry." Proteomics 4(4): 909-916.
- Rissanen, E., H. K. Tranberg, J. Sollid, G. E. Nilsson and M. Nikinmaa (2006). "Temperature regulates hypoxia-inducible factor-1 (HIF-1) in a poikilothermic vertebrate, crucian carp (*Carassius carassius*)." Journal of Experimental Biology 209(Pt 6): 994-1003.
- Roberts, S. B. and M. R. Gavery (2012). "Is There a Relationship between DNA Methylation and Phenotypic Plasticity in Invertebrates?" Frontiers in Physiology 2: 116.
- Robertson, K. D. (2005). "DNA methylation and human disease." Nature reviews. Genetics 6(8): 597-610.
- Rodriguez-Ortega, M. J., B. E. Grosvik, A. Rodriguez-Ariza, A. Goksoyr and J. Lopez-Barea (2003). "Changes in protein expression profiles in bivalve molluscs (*Chamaelea gallina*) exposed to four model environmental pollutants." Proteomics 3: 1535.
- Rodriguez-Piñeiro, A. M., D. Ayude, F. J. Rodriguez-Berrocal and M. Paez de la Cadena (2004). "Concanavalin A chromatography coupled to two-dimensional gel electrophoresis improves protein expression studies of the serum proteome." Journal of Chromatography B: Biomedical Sciences and Applications; 803: 337.
- Rossignol, T., D. Kobi, L. Jacquet-Gutfreund and B. Blondin (2009). "The proteome of a wine yeast strain during fermentation, correlation with the transcriptome." Journal of applied microbiology 107(1): 47-55.
- Rousset, S., Alves-Guerra M C, Mozo J, Miroux B, Cassard-Doulcier A M, Bouillaud F and Ricquier D (2004). "The Biology of Mitochondrial Uncoupling Proteins." Diabetes 53(suppl 1): S130-S135.
- Ruiz-Pesini, E., Mishmar D, Brandon M, Procaccio V and Wallace D C (2004). "Effects of Purifying and Adaptive Selection on Regional Variation in Human mtDNA." Science 303(5655): 223-226.
- Ryan, M. T. and N. J. Hoogenraad (2007). "Mitochondrial-nuclear communications." Annu Rev Biochem 76: 701-722.
- Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular cloning : a laboratory manual. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press.

- Sanger, A. M. (1993). "Limits to the acclimation of fish muscle." Reviews in Fish Biology and Fisheries 3(1): 1-15.
- Schatz, G. (1995). "Mitochondria: beyond oxidative phosphorylation." Biochimica et Biophysica Acta 1271(1): 123-126.
- Schlattner, U., Forstner M, Eder M, Stachowiak O, Fritz-Wolf K and Wallimann T (1998). "Functional aspects of the X-ray structure of mitochondrial creatine kinase: A molecular physiology approach." Molecular and Cellular Biochemistry 184(1-2): 125-140.
- Schlegel, J., B. Zurbriggen, G. Wegmann, M. Wyss, H. M. Eppenberger and T. Wallimann (1988). "Native mitochondrial creatine kinase forms octameric structures. I. Isolation of two interconvertible mitochondrial creatine kinase forms, dimeric and octameric mitochondrial creatine kinase: characterization, localization, and structure-function relationships." Journal of Biological Chemistry 263(32): 16942-16953.
- Schroeder, F., C. A. Jolly, T. H. Cho and A. Frolov (1998). "Fatty acid binding protein isoforms: structure and function." Chemistry and Physics of Lipids 92(1): 1-25.
- Scott W.B and Crossman E.J (1973). "Freshwater Fishes of Canada. Bulletin of the Fisheries." Research Board of Canada 184: 1-1092.
- Semenza, G. L., Jiang BH, Leung S W, Passantino R, Concordet J P, Maire P and Giallongo A (1996). "Hypoxia Response Elements in the Aldolase A, Enolase 1, and Lactate Dehydrogenase A Gene Promoters Contain Essential Binding Sites for Hypoxia-inducible Factor 1." Journal of Biological Chemistry 271(51): 32529-32537.
- Skala-Rubinson, H., M. Vibert and J. Dreyfus (1976). "Electrophoretic modifications of three enzymes in extracts of human and bovine lens. Posttranslational "aging" of lens enzymes." Clinica Chimica Acta 70(3): 385-390.
- Speakman, J. R., D. A. Talbot, C. Selman, S. Snart, J. S. McLaren, P. Redman, E. Krol, D. M. Jackson, M. S. Johnson and M. D. Brand (2004). "Uncoupled and surviving: individual mice with high metabolism have greater mitochondrial uncoupling and live longer." Aging Cell 3(3): 87-95.
- Stout, A. L., J. Wang, J. M. Sanger and J. W. Sanger (2008). "Tracking changes in Z-band organization during myofibrillogenesis with FRET imaging." Cell Motility and the Cytoskeleton 65(5): 353-367.
- Szczesna, D., J. Zhao, M. Jones, G. Zhi, J. Stull and J. D. Potter (2002). "Phosphorylation of the regulatory light chains of myosin affects Ca<sup>2+</sup> sensitivity of skeletal muscle contraction." Journal of Applied Physiology 92(4): 1661-1670.
- Takada, F, Woude D L V, Tong H Q, Thompson T G, Watkins S C, Kunkel L M and Beggs A H (2001). "Myozenin: An  $\alpha$ -actinin- and  $\gamma$ -filamin-binding protein of skeletal muscle Z lines." Proceedings of the National Academy of Sciences 98(4): 1595-1600.

- Temple GK and Johnston I. (1998). "Testing hypotheses concerning the phenotypic plasticity of escape performance in fish of the family Cottidae." Journal of Experimental Biology 201(3): 317-331.
- Temple GK, Wakeling J. M. and Johnston I. A. (2000). "Seasonal changes in fast-starts in the short-horn sculpin: integration of swimming behaviour and muscle performance." Journal of Fish Biology 56(6): 1435-1449.
- Thorsness, P. E. and E. R. Weber (1996). "Escape and migration of nucleic acids between chloroplasts, mitochondria, and the nucleus." International review of Cytology 165: 207-234.
- Timmis, J. N., M. A. Ayliffe, C. Y. Huang and W. Martin (2004). "Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes." Nature reviews. Genetics 5(2): 123-135.
- Tomonaga, T., Matsushita K, Yamaguchi S, Oh-Ishi M, Kodera Y, Maeda T, Shimada H, Ochiai T and Nomura F (2004). "Identification of Altered Protein Expression and Post-Translational Modifications in Primary Colorectal Cancer by Using Agarose Two-Dimensional Gel Electrophoresis." Clinical Cancer Research 10(6): 2007-2014.
- True, H. L. and S. L. Lindquist (2000). "A yeast prion provides a mechanism for genetic variation and phenotypic diversity." Nature 407(6803): 477-483.
- Tsuji T, Shimohama S, Kamiya S, Sazuka T and O. O (1999). "Analysis of brain proteins in Alzheimer's disease using high-resolution two-dimensional gel electrophoresis." Journal of the Neurological Sciences 166: 100–106.
- Van Lint, C., S. Emiliani and E. Verdin (1996). "The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation." Gene expression 5(4-5): 245-253.
- Van Pilsum, J. F., G. C. Stephens and D. Taylor (1972). "Distribution of creatine, guanidinoacetate and the enzymes for their biosynthesis in the animal kingdom. Implications for phylogeny." The Biochemical Journal 126(2): 325-345.
- Vergnes, L., R. Chin, S. G. Young and K. Reue (2011). "Heart-type fatty acid-binding protein is essential for efficient brown adipose tissue fatty acid oxidation and cold tolerance." Journal of Biological Chemistry 286(1): 380-390.
- Verhoeckx, K. C., M. Gaspari, S. Bijlsma, J. van der Greef, R. F. Witkamp, R. P. Doornbos and R. J. Rodenburg (2005). "In search of secreted protein biomarkers for the anti-inflammatory effect of beta2-adrenergic receptor agonists: application of DIGE technology in combination with multivariate and univariate data analysis tools." Journal of Proteome Research 4: 2015.
- Wahrmann, J. P., R. Winand and M. Rieu (2001). "Plasticity of skeletal myosin in endurance-trained rats (I). A quantitative study." European Journal of Applied Physiology 84(5): 367-372.

- Walsh, C. (2006). Posttranslational Modification of Proteins: Expanding Nature's Inventory, Roberts and Company Publishers.
- Wang, N. Shaner, B. Mittal, Q. Zhou, J. Chen, J. Sanger and J. Sanger (2005). "Dynamics of Z-band based proteins in developing skeletal muscle cells." Cell Motility and the Cytoskeleton 61(1): 34-48.
- Wang, Y.-C., S. E. Peterson and J. F. Loring (2014). "Protein post-translational modifications and regulation of pluripotency in human stem cells." Cell Research 24(2): 143-160.
- Warburg, O (1930). "Warburg report on the metabolism of tumors." Journal of Chemical Education 7(1): 179.
- Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel and D. F. Hochstrasser (1999). "Protein identification and analysis tools in the ExPASy server." Methods Mol Biol 112: 531-552
- Wyss, M. and R. Kaddurah-Daouk (2000). "Creatine and creatinine metabolism." Physiological Reviews 80(3): 1107-1213.
- Yamamoto, T., A. Yamamoto, M. Watanabe, M. Kataoka, H. Terada and Y. Shinohara (2011). "Quantitative evaluation of the effects of cold exposure of rats on the expression levels of ten FABP isoforms in brown adipose tissue." Biotechnology Letters 33(2): 237-242.
- Zerbetto, E., L. Vergani and F. Dabbeni-Sala (1997). "Quantification of muscle mitochondrial oxidative phosphorylation enzymes via histochemical staining of blue native polyacrylamide gels." Electrophoresis 18(11): 2059-2064.
- Zhang , G., W. Chu, S. Hu, T. Meng, L. Pan, R. Zhou, Z. Liu and J. Zhang (2011 a). "Identification and analysis of muscle-related protein isoforms expressed in the white muscle of the mandarin fish (*Siniperca chuatsi*)." Marine Biotechnology 13(2): 151-162.
- Zhang, H., Y. Wang, J. Li, J. Yu, J. Pu, L. Li, H. Zhang, S. Zhang, G. Peng, F. Yang and P. Liu (2011 b). "Proteome of skeletal muscle lipid droplet reveals association with mitochondria and apolipoprotein a-I." Journal of Proteome Research 10(10): 4757-4768.
- Zhong, S., S. P. Miller, D. E. Dykhuizen and A. M. Dean (2009). "Transcription, translation, and the evolution of specialists and generalists." Molecular Biology and Evolution 26(12): 2661-2678.
- Zhou, W., M. Capello, C. Fredolini, L. Piemonti, L. A. Liotta, F. Novelli and E. F. Petricoin (2010). "Mass spectrometry analysis of the post-translational modifications of alpha-enolase from pancreatic ductal adenocarcinoma cells." Journal of proteome research 9(6): 2929-2936.

## Annexe 1: Limitations à l'identification des protéines par LC/MS

Malgré les d'efforts que nous avons mis pour identifier les protéines spécifiques, nous n'avons pas été en mesure d'identifier une majorité des points analysés par spectrométrie de masse. De ce fait nous ne pouvons être certains des facteurs impliqués. Deux éléments méthodologiques majeurs, reconnus également par d'autres études (Lopez *et al.*, 2001, Martínez-Fernández *et al.*, 2008) peuvent expliquer en partie cet échec. Il s'agit de la quantité de protéine et l'absence de banque de données génomique pour notre modèle d'étude.

Dans le premier cas, la différence de peptides détectés que nous avons observée durant ce projet entre la révélation ayant des temps d'expositions plus courtes et plus longues démontre que la révélation a un impact sur la quantité de protéine disponible pour le séquençage. En général, plus les temps d'exposition ont été longs, moins on a identifié de peptides dans les échantillons. Cela peut s'expliquer par le fait que la coloration au nitrate d'argent est reconnue pour induire une modification des protéines conduisant à une baisse du nombre de peptides séquencés (Richert *et al.*, 2004). Étant donné que la sélection des points à séquencer a été faite à partir des gels révélés avec la méthode ayant des temps d'expositions prolongés, il est difficile d'avoir une idée concrète de la quantité de protéine contenue dans ces points à cause de la surexposition. Par conséquent, la quantité en protéine de certains points envoyés au séquençage aurait pu être surévaluée, alors qu'en réalité elle était insuffisante pour permettre leur identification. Cette hypothèse est confortée par le fait que les protéines identifiées avec succès ont été extraites de gels soumis à un protocole de révélation abrégé. L'absence de protéine se démarquant des contaminants dans les échantillons initiaux démontre qu'une faible quantité de protéine a été soumise à la base et qu'une faible quantité de protéine est associée à un faible signal (Verhoeckx *et al.*, 2005).



Dans beaucoup d'échantillons, l'abondance de contaminant par rapport à la faible quantité de protéines isolées aurait empêché l'identification de ces protéines. Dans la majorité des cas, les protéines contaminant nos échantillons avaient le meilleur score et étaient associées à des protéines participant à la formation du cytosquelette. Étant donné que l'on en retrouve dans la poussière, il est très difficile d'empêcher cette contamination. De plus, comme ces protéines ont une séquence peptidique très conservée entre les espèces, il est très probable que ces peptides trouvent une correspondance avec la banque de données utilisée.

Face à un faible taux d'identification avec cette méthode, nous avons tenté de contourner ce problème en soumettant directement les extraits totaux protéomiques de chaque biotype à séquencer. Les résultats n'ont vraiment pas été plus concluants, la raison principale étant sûrement liée à la complexité de ces échantillons. Le fait de ne pas appliquer d'étape de séparation comme dans le cas dans les gels 2D augmente significativement la quantité de protéines liées au cytosquelette dans l'échantillon masquant le signal des autres protéines plus faiblement exprimées. Avec le recul, la solution aurait été de fractionner l'échantillon de tissus autant que possible afin d'en diminuer la complexité. Il faudrait fractionner le protéome de la cellule en sous-groupe : les protéines cytoplasmiques et le mitoprotéome (membrane et/ou matrice). Le fait de produire ce fractionnement aurait pu nous permettre dans le cas de l'étude du mitoprotéome de nous affranchir d'une grande partie des protéines cytosquelettiques qui posaient problème avec cette approche (O'Connell et Ohlendieck 2009).

L'autre élément majeur ayant nui à l'identification des échantillons fut l'absence de banque de données génomiques pour notre modèle d'étude. D'autres études protéomiques ont également rencontré des problèmes à ce niveau démontrant l'importance de ce point (Lopez *et al.*, 2001, Rodriguez-Ortega *et al.*, 2003, Barrett *et al.*, 2005). Par conséquent, nous n'avons eu d'autre choix au début de notre projet où nous n'avions pas encore la banque de données transcriptomique d'utiliser la banque de données des téléostéens de NCBI. Au départ, nous

pensions utiliser plus précisément la banque de données du poisson-zèbre parce que c'est l'organisme le plus proche génétiquement de notre modèle d'étude pour lequel nous disposions d'une banque de données complète. Mais contrairement à cette préconception, il semblerait qu'il y ait trop de divergence au niveau protéomique entre notre modèle d'étude et le poisson-zèbre pour permettre une identification à partir des rares peptides dont nous disposions. Étant donné que nous n'avons reçu la banque de données transcriptomique qu'au début de l'année 2014, nous avons dû explorer d'autres approches pour tenter d'exploiter les données brutes du séquençage. Nous avons d'abord essayé d'utiliser une approche MS/MS en comparant directement l'ensemble des spectres d'un échantillon pour trouver une correspondance avec la banque de données de spectre de « the Global Proteome Machine » ([www.theGPM.org](http://www.theGPM.org)). Pour les mêmes raisons que celle que nous venons d'énoncer, cette approche n'a pas fonctionné non plus.

Nous nous sommes par la suite tournés vers une nouvelle analyse *de novo* de nos données brutes avec le programme Peaks (Bioinformatics Solution). L'avantage de ce programme est qu'il tente de déterminer directement à partir des spectres les séquences d'acides aminés correspondants sans avoir recours à une banque de données de spectre de masse. Il utilise ensuite ces données pour trouver une homologie dans la banque de données de NCBI en tenant compte de la présence d'une éventuelle divergence au niveau de la séquence d'acides aminés. Cette approche a permis d'augmenter le score de certaines protéines en leur réattribuant des peptides ignorés jusque là dans les analyses précédentes. À savoir que jusqu'à l'analyse *de novo* les peptides ne correspondant pas à 100% à une séquence de la banque de données n'étaient pas pris en compte dans l'analyse. Mais même avec cette nouvelle analyse, les protéines semblant être spécifiques à un échantillon se retrouvaient tout de même dans les 10 premières places ce qui ne nous a pas permis d'augmenter significativement notre niveau de confiance envers les identifications bioinformatiques. Ces échantillons ont été depuis

réanalysés à partir de la banque de données transcriptomique partielle de notre modèle d'étude. Les résultats obtenus ont été tellement différents que cela démontre qu'il est crucial de disposer d'une banque de données propre à l'organisme étudié dans le cadre d'un projet de protéomique similaire au nôtre. Les protéines que nous soupçonnions spécifiques de l'échantillon avec l'analyse *de novo* se sont placées dans le top 3 des protéines ayant le score le plus grand (en règle générale après l'actine et la myosine) et nous avons finalement pu déterminer avec certitude l'identité des protéines isolées. Le fait que tous les échantillons analysés à partir de la banque de données transcriptomiques n'aient pas donné un résultat n'implique pas qu'il n'y avait pas du tout ou pas suffisamment de peptides de protéines isolés à l'intérieur du point. La banque de données transcriptomiques est toujours incomplète, ce qui expliquerait pourquoi certains peptides n'y ont pas trouvé de correspondance.