

Université de Montréal

Épigénétique mitochondriale chez des espèces avec DUI

Par Karim Bouvet

Département de sciences biologiques
Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade M. Sc.
en Sciences biologiques

Septembre, 2018
© Karim Bouvet, 2018

Résumé

La méthylation de l'ADN est l'un des mécanismes épigénétiques les plus étudiés et conséquemment l'un des plus connus. Les chercheurs ont démontré les effets importants qu'elle peut avoir au niveau du phénotype des eucaryotes. La méthylation de l'ADN est très étudiée chez les vertébrés, chez qui elle permet d'inactiver des gènes. Les taux de méthylation sont plus faibles de façon générale chez les invertébrés, mais ces taux sont tout de même non négligeables chez des groupes comme les mollusques. La méthylation de l'ADN mitochondrial a fait l'objet de plusieurs études récemment: les chercheurs commencent à se rendre compte de la présence de méthylation dans le génome mitochondrial. Or, il existe chez les bivalves (mollusques) des espèces qui possèdent deux génomes mitochondriaux différents et qui pourraient potentiellement avoir des fonctions différentes en lien avec la reproduction. Ce système se nomme DUI (*doubly uniparental inheritance*) et il est exclusif aux bivalves. Ces deux génomes différents sont distribués de façon particulière dans les tissus de l'organisme, mais l'expression de chacun des génomes n'est pas la même selon le type de tissu dans lequel ils se trouvent. La méthylation pourrait être l'un des mécanismes permettant de réguler l'expression de ces génomes mitochondriaux et donc être un élément clé pour le fonctionnement du système DUI). C'est pourquoi nous voulions 1) déterminer s'il y a présence de méthylation dans l'ADNmt d'espèces avec DUI et 2) déterminer si les patrons de méthylation observés indiquent rôle fonctionnel, i.e. s'ils peuvent être liés à une régulation de l'expression des génomes mitochondriaux, expliquant ainsi en partie le fonctionnement du DUI.

Mots-clés : Méthylation de l'ADN, Mitochondrie, ADN mitochondrial, Double transmission uniparentale (DUI), Épigénétique, Bivalves, *Venustaconcha ellipsiformis*.

Abstract

Mitochondrial DNA methylation and its role in gene expression regulation remain controversial. In this study, we assessed the presence of mtDNA methylation in the two extremely divergent mt genomes of a naturally heteroplasmic bivalve with doubly uniparental inheritance (DUI) of mtDNA. It was previously shown that the two mt genomes were expressed differently depending on the tissue/sex where they were found. To deal with this potentially deleterious heteroplasmic situation, the existence of different mechanisms in DUI species, such as silencing mechanisms for the M mtDNA at the transcriptional level, were hypothesized. We thus compared methylation patterns across different heteroplasmic tissue samples in male and female individuals of the DUI possessing freshwater mussel *Venustaconcha ellipsiformis*. Overall, our results do not show a clear pattern in the methylation of tissues, although levels were drastically higher in the gonads. We discuss the presence of methylation peaks in putative origins of transcription/replication and propose functional biological significance for the low levels observed consistently across the rest of the genome.

Keywords : DNA methylation, Mitochondria, Mitochondrial DNA, Double Uniparental Inheritance (DUI), Epigenetics, Bivalves, *Venustaconcha ellipsiformis*.

Table des matières

Résumé	ii
Abstract.....	iii
Table des matières	iv
Liste des tableaux.....	vi
Liste des figures	vii
Liste des sigles et abréviations.....	viii
Remerciements	x
1 Introduction	1
1.1 Épigénétique et méthylation de l'ADN	2
1.1.1 Méthylation de l'ADN classique.....	2
1.1.2 Méthylation de l'ADN chez les invertébrés	5
1.1.3 Méthylation de l'ADN mitochondrial	6
1.1.4 Double transmission uniparentale de l'ADNmt (DUI).....	10
1.2 Objectif et hypothèses	13
2 Article: Is mtDNA methylation functionally relevant in animals? Insights from a naturally occurring heteroplasmic system	15
2.1 Introduction.....	16
2.2 Materials and Methods	19

2.2.1	Specimen collection and sample preparation.....	19
2.2.2	DNA extraction and purification.....	20
2.2.3	Chosen samples for sequencing.....	20
2.2.4	Mitochondrial reference sequences.....	21
2.2.5	Bisulfite sequencing and detection of methylation	22
2.3	Results.....	23
2.3.1	General sequencing and mapping information.....	23
2.3.2	Methylation levels in F and M mitochondrial genomes.....	26
2.3.3	Effect of the tissue, sex and mt genome on methylation patterns	28
2.3.4	Methylation in the different cytosine contexts	35
2.4	Discussion.....	36
2.4.1	General sequencing and mapping information.....	36
2.4.2	Mean methylation levels in F (soma and gonad) and M (gonad) mitochondrial genomes...	37
2.4.3	Effect of the tissue, sex and mt genome on methylation patterns	38
2.5	Conclusion.....	42
3	Discussion	44
3.1	Information relative au séquençage	45
3.2	Méthylation de F (soma et gonades mâles) et M (gonades mâles)	48
3.3	Effet du sexe, du tissu et du génome sur les patrons de methylation	50
3.4	Méthylation dans des contextes de cytosines différents.....	54
4	Conclusion et perspectives futures.....	58
	Bibliographie.....	62
	Annexes	xi
	Annexe 1 : tableaux supplémentaires	xii

Liste des tableaux

Table I. General information related to the bisulfite sequencing of <i>V. ellipsiformis</i>	25
Supplementary Table I: Methylation percentage for every gene in the F genome found in male soma	xii
Supplementary Table II: Methylation percentage for every gene in the F genome found in female soma	xiv
Supplementary Table III: Methylation percentage for every gene in the F genome found in male gonads	xvi
Supplementary Table IV: Methylation percentage for every gene in the M genome found in male soma	xviii
Supplementary Table VI: Methylation percentage for every gene in the M found in male gonads	xx
Supplementary Table V: Methylation percentage for every gene in the M genome found in female soma	xxii
Supplementary Table VII: average of the methylation percentages (%) calculated for cytosines in every context (CG, CHG, CHH), for both the top and bottom strands in different samples of <i>V. ellipsiformis</i> and other invertebrates	xxiii

Liste des figures

Figure 1 : Méthylation d'un îlot CpG dans le promoteur d'un gène	4
Figure 2 : Patrons d'hydroxyméthylation des cytosines pour le génome mitochondrial humain....	9
Figure 3 : Schéma représentant le système de double transmission uniparentale (DUI).....	11
Figure 4 : Ratios de l'expression des ADNmt M et F dans les gonades et les tissus somatiques mâles et femelles chez <i>Mytilus galloprovincialis</i>	12
Figure 5: Abondance des transcrits des gènes mitochondriaux en fonction du gène et du génome chez <i>Ruditapes philippinarum</i>	13
Figure 6: Number of reads after trimming for every sample sequenced by shotgun and bisulfite sequencing (WGBS)	24
Fig. 7: Methylation percentage on the template and coding strands of F and M mt genomes present in soma or gonads of <i>V. ellipsiformis</i>	27
Fig. 8: Mean methylation percentage for 100bp windows along the F mt genome found in male soma of <i>V.ellipsiformis</i> for the top and bottom strands	29
Figure 9: Mean methylation percentage for 100bp windows along the F mt genome found in the gills of three males <i>V. ellipsiformis</i>	32
Figure 10: Mean methylation percentage for 100bp windows along the F mt genome found in male and female adductor muscles of <i>V.ellipsiformis</i>	33
Figure 11: Mean methylation percentage for 100bp windows along the F mt genome found in male gonads of <i>V.ellipsiformis</i> for the top and bottom strands of the F and M genomes.....	34

Liste des sigles et abréviations

ADN	Acide Désoxyribonucléique
ADNmt	ADN mitochondrial
ARNm	Acide ribonucléique messenger
CGI	<i>CpG Island</i>
<i>cox</i>	Sous-unité du cytochrome c-oxidase
<i>cytb</i>	Cytochrome b
CpG	Cytosine-phosphate-Guanine
D-loop	<i>Displacement loop</i>
Dnmt	<i>DNA methyltransferase</i>
DUI	<i>Doubly Uniparental Inheritance</i>
GBM	<i>Gene Body Methylation</i>
GSS	<i>Gene Start Site</i>
MeDIP	Methylated DNA Immunoprecipitation
MTS	<i>Mitochondrial Translocation Sequence</i>
PCR	<i>Polymerase Chain reaction</i>
q-PCR	<i>quantitative-PCR</i>
ROS	<i>Reactive Oxygen Species</i>
<i>rrnS</i>	<i>Small Ribosomal RNA</i>
RT-PCR	<i>Reverse-Transcription PCR</i>
SMI	<i>Strict Maternal Inheritance</i>
TET	<i>Ten-eleven translocation methylcytosine dioxygenase</i>
TFAM	<i>Mitochondrial Transcription Factor A</i>

WGBS *Whole Genome Bisulfite Sequencing*

5hmC 5-hydroxymethyl-cytosine

5mC 5-methyl-cytosine

Remerciements

Je tiens avant tout à remercier ma directrice de recherche, Sophie Breton, pour toutes les choses qu'elle a faites pour moi. Je n'aurais pas pu demander mieux comme directrice (et comme laboratoire) et j'en garderai toujours un bon souvenir. Merci de toujours avoir été disponible et de m'avoir permis de découvrir le monde de la recherche. Je suis très reconnaissant pour tous les efforts que tu as faits dans le cadre de ma formation, mais aussi pour les opportunités que j'ai eues ou que j'aurai grâce à toi! Je te souhaite de devenir « famous » mais surtout de garder la même passion pour ce que tu fais, ça fait vraiment plaisir de travailler avec des gens qui aiment ce qu'ils font!

J'aimerais aussi remercier Davide Guerra, qui a été un vrai mentor pour la plus grande partie de ma maîtrise. Je lui dois mes connaissances pour tout ce qui concerne le laboratoire et la bioinformatique. Je lui souhaite de continuer sa carrière de chercheur et de s'épanouir lui aussi ! Merci également à Bernard Angers pour ses précieux conseils et à tous les stagiaires qui m'ont rendu la vie plus facile (surtout Stéphanie et Gabrielle).

Viennent ensuite tous les membres du laboratoire, que je ne considère pas seulement comme des collègues mais véritablement comme de bons amis. Encore une fois, je n'aurais pas pu demander mieux et je vous souhaite beaucoup de succès dans tous vos projets. Stefano et Charlotte, vous êtes un vrai divertissement ! J'espère aussi pouvoir garder contact avec Arthur, Ariane, Philip, Georges et tous les autres. Finalement, merci à ma famille et mes amis pour leur soutien moral, ça n'aurait pas été possible sans vous!

1 Introduction

1.1 Épigénétique et méthylation de l'ADN

1.1.1 Méthylation de l'ADN classique

L'épigénétique peut être définie comme l'ensemble des mécanismes qui agissent à un niveau « au dessus » du génome, c'est-à-dire qu'ils permettent le contrôle de l'expression des gènes, sans pour autant changer la séquence de nucléotides qui compose ces gènes (Holliday, 2006). Ainsi, il est possible de modifier le fonctionnement d'une cellule, voire d'affecter le phénotype complet d'un individu, tout en conservant un génome identique. Ces mécanismes épigénétiques ont une importance énorme notamment durant le développement des eucaryotes (Waddington, 1939), ne serait-ce que pour permettre la différenciation cellulaire, sans laquelle l'organisme ne serait évidemment pas viable (Reik, 2007). Bien qu'il existe plusieurs de ces mécanismes et qu'ils soient tous très importants, celui qui est le plus connu – et probablement celui qui joue le plus grand rôle en ce qui concerne l'expression des gènes – est la méthylation de l'ADN (Reik et al., 2001). Ceci est particulièrement vrai chez les vertébrés (dont l'humain) et c'est sans doute pourquoi la méthylation a été le mécanisme épigénétique le plus étudié. En effet, de très nombreux liens entre la méthylation de l'ADN et des maladies ou des malformations ont été démontrés au cours des dernières années (Robertson, 2005; Wilson et al., 2007).

La méthylation de l'ADN se traduit par la présence d'un groupement méthyle (CH_3) sur un nucléotide. La plupart du temps, chez les animaux, cette méthylation se fait sur une cytosine (Jones, 2012). Toutefois, chez les procaryotes et les plantes, des adénines méthylées ont été observées (Low et al., 2001; Vanyushin et Ashapkin, 2011). Il est important de spécifier ici qu'on parle de méthylation de l'ADN, puisque l'ajout d'un groupement méthyle peut également survenir sur des protéines comme les histones, et ceci représente également un mécanisme

épigénétique. En effet, l'acétylation et la méthylation des histones se produisent sur une partie spécifique des protéines (la queue) et en régulant la condensation de l'ADN (la chromatine), elles permettent indirectement un contrôle de l'expression des gènes (Rice et Allis, 2001). (Hendrich et Tweedie, 2003; Zilberman, 2008).

Plusieurs études ont démontré que la méthylation survient surtout aux cytosines qui sont suivies d'une guanine dans la séquence d'ADN, formant ce qu'on appelle un dinucléotide CpG (cytosine-phosphate-guanine) (Illingworth et Bird, 2009). Puisque l'ADN est composé de deux brins complémentaires, on observe sur le brin opposé un GC, soit l'inverse du CpG. Il s'agit d'une séquence palindromique et chacune des deux cytosines peut être méthylée (Bestor, 1992).

Contrairement aux mutations qui surviennent au niveau du matériel génétique, les mécanismes épigénétiques agissent la plupart du temps de manière coordonnée, parfois en réponse à un stimulus externe provenant de l'environnement (Hajkova et al., 2002 ; Massicotte et al., 2011). C'est en effet là toute l'importance des mécanismes épigénétiques, puisqu'ils permettent d'ajuster le phénotype d'un individu en fonction de l'environnement. La méthylation se produit la plupart du temps à des sites bien précis et suivant un patron bien défini. Une erreur de méthylation à un seul endroit du génome peut entraîner des conséquences graves ; on découvre aujourd'hui que c'est la cause de plusieurs maladies qui affectent les humains, entre autres (ex : Robertson, 2005 ; Nile et al., 2008). Les patrons de méthylation varient selon le type de tissu, voire même selon le type de cellule (Bird, 2002). Cette méthylation varie aussi dans le temps : elle se produit progressivement au cours du développement. Aux premiers stades de développement de la souris, par exemple, on observe des taux de méthylation relativement faibles (environ 30% du niveau habituel au stade adulte) qui augmentent au cours du

développement (Monk et al., 1987). Un phénomène absolument essentiel en ce qui a trait à l'expression des gènes et qui est caractéristique des vertébrés est le regroupement des CpG dans des séquences d'ADN relativement courtes (200 à 3000 nucléotides) situées pour la plupart dans les promoteurs des gènes. Ces courtes séquences d'ADN sont appelées des îlots CpG (CpG islands: CGI) (Illingworth et Bird, 2009). Les CGI fortement concentrés en CpG sont généralement hypométhylés et, à l'inverse, les CGI moins concentrés en CpG sont plutôt hyperméthylés (Elango et Soojin, 2008). Chez les vertébrés, on observe généralement que la méthylation d'un CGI dans le promoteur d'un gène inactive ce dernier (figure 1). Ce système relativement simple rend possible une réaction rapide face à l'environnement et un contrôle local de l'expression des différents gènes sur un même chromosome.

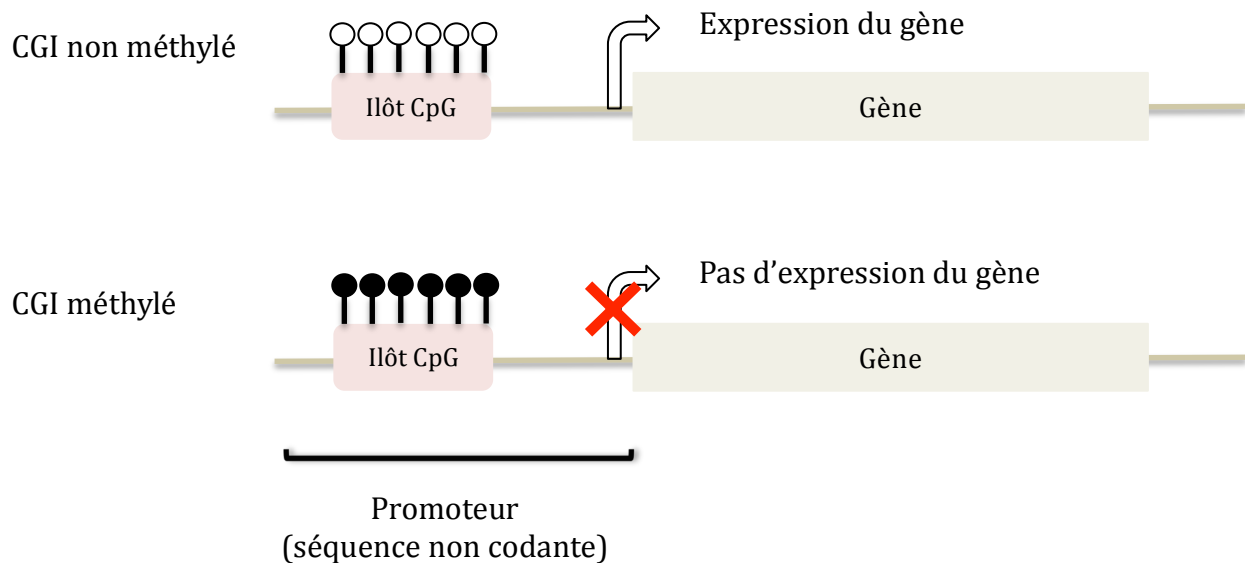


Figure 1 Méthylation d'un îlot CpG dans le promoteur d'un gène, ce qui empêche l'expression de celui-ci (modifiée de Suzuki, 2007)

En haut: Îlot CpG non méthylé, gène exprimé ; En bas: Îlot CpG méthylé, gène non exprimé

Cependant, il existe de la méthylation à d'autres endroits que dans les promoteurs ; on retrouve également des cytosines méthylées dans la région des gènes qui est transcrite (Gene

Body Methylation : GBM) (Zemach et al., 2010). Cette méthylation dans l'unité de transcription a souvent un effet différent de celle qui se fait aux CGI dans les promoteurs : certains sites méthylés dans un gène rendent l'expression de celui-ci plus forte (Jones, 2012). La façon dont cela se produit n'est toutefois pas encore très bien comprise. Une GBM peut avoir pour effet de réprimer l'expression des éléments transposables sans affecter l'élongation de la transcription du gène. Elle peut aussi agir au niveau de l'épissage alternatif en incluant certaines parties de la séquence (exons) qui auraient été rejetées sans méthylation (Jones, 2012).

1.1.2 Méthylation de l'ADN chez les invertébrés

Jusqu'ici, j'ai décrit le patron de méthylation général que l'on observe chez les vertébrés. Cependant, si l'on regarde ce qui se passe du côté des invertébrés, on constate que la situation est non seulement très différente, mais qu'elle est substantiellement plus complexe (Diaz-Freij et al., 2014). Au lieu d'avoir un patron global applicable à tous les groupes d'invertébrés, il existe beaucoup de variation en termes de taux de méthylation, ainsi qu'en ce qui concerne les patrons de méthylation chez différentes espèces. En effet, des taux de méthylation presque nuls ont été observés chez quelques espèces, et des taux beaucoup plus élevés chez d'autres (Sun et al., 2014). Par exemple, *Drosophila melanogaster* et *Caenorhabditis elegans* présentent des taux de méthylation presque nuls, alors que chez les mollusques, on observe des taux allant jusqu'à 20% (Sun et al., 2014). De manière générale, les taux de méthylation chez les groupes d'invertébrés demeurent beaucoup plus faibles que chez les vertébrés, qui ont des taux de méthylation de 70 à 80% des CpG (Olson et Roberts, 2014). De plus, la méthylation chez les invertébrés se fait surtout dans la région intragénique (Rivière, 2014). Il y a donc une GBM relativement forte et peu de CGI dans le promoteur. Il semble que cette méthylation

intragénique permet également le contrôle de l'expression des gènes. Par exemple, Gavery et Roberts (2013) ont établi une corrélation positive entre la méthylation intragénique chez l'huître *Crassostrea gigas* et l'expression des gènes méthylés. De plus, Rivière et al., (2013) et Saint-Carlier et Rivière (2015) ont trouvé un rôle essentiel de la méthylation au cours du développement chez cette même espèce, et ils ont découvert que la méthylation se fait non seulement au niveau intragénique mais aussi dans les promoteurs, à l'instar de ce qui est observé chez les vertébrés. Toujours chez *Crassostrea gigas*, Olson et Roberts (2014) ont calculé un taux d'environ 15% des CpG du génome qui étaient méthylés, ce qui est assez similaire à ce qui a été observé chez *Chlamys farerri*, un pétoncle chez qui Sun et al., (2014) ont observé un taux de méthylation d'environ 20% dans différents tissus.

1.1.3 Méthylation de l'ADN mitochondrial

La plupart des études mentionnées jusqu'ici concernaient exclusivement le génome nucléaire ; relativement peu a été fait pour connaître l'état de la méthylation dans le génome mitochondrial. Ce n'est que récemment que la présence de méthylation dans le génome mitochondrial a été confirmée, et il semblerait qu'elle peut également mener à des problèmes importants si elle n'est pas effectuée correctement (ex : Blanch et al., 2016 ; D'Aquila, 2015 ; Mishra et Kowluru, 2015). Il faut noter que la majorité des travaux ciblent les humains ou d'autres mammifères dans le but d'acquérir des connaissances potentiellement utiles dans le domaine de la santé. Ainsi, il a été démontré que contrairement à l'ADN nucléaire, la méthylation dans l'ADNmt ne se fait pas préférentiellement aux CpG chez les mammifères. En effet, Bellizzi et al., (2013) ont détecté dans l'ADNmt de la méthylation ainsi que de l'hydroxyméthylation sur des cytosines dans des contextes autres que CpG. Cette méthylation se ferait, selon eux, notamment dans la

région D-loop. Cette dernière est la région la plus importante pour le contrôle de la réplication et de la transcription du génome mitochondrial (Sbisà et al., 1997). Il s'agirait donc d'une façon majeure de contrôler la réplication ou l'expression de l'ADNmt.

La découverte de méthyltransférases à l'intérieur des mitochondries fut un des premiers arguments de taille confirmant la présence de méthylation dans les mitochondries (il s'agit des protéines responsables de la méthylation de l'ADN). Chestnut et al., (2011) ont démontré la présence intra-mitochondriale de Dnmt3a, qui sert à la méthylation *de novo*, c'est-à-dire l'ajout d'un groupement méthyle sur un site où il n'y avait pas du tout de méthylation, tandis que Shock et al., (2011), ont rapporté qu'ils avaient décelé la présence de Dnmt1, qui sert plutôt à la méthylation « de maintenance » : elle ajoute un groupement méthyle sur un site déjà hémiméthylé (Hermann et al., 2004). Cette hémiméthylation se produit lorsqu'il y a réplication de l'ADN, par exemple, et le nouveau brin doit lui aussi être méthylé. Shock et al., (2011) ont prouvé la pertinence de leurs résultats en décrivant la translocation de la protéine vers la mitochondrie. L'ARNm contient une séquence signal ciblant la mitochondrie (mitochondrial translocation sequence : MTS) en amont du site d'initiation de la traduction. Par ailleurs, Dnmt3b (une autre méthyltransférase *de novo*) ainsi que des protéines de la famille TET, dont le rôle est la déméthylation, ont aussi été localisées à l'intérieur des mitochondries (Bellizzi et al., 2013).

Tiedemann et al., (2014) ont démontré que dans le génome nucléaire, une quatrième méthyltransférase, Dnmt3L, favorisait la méthylation aux CpG plutôt qu'aux cytosines non-CpG. En effet, ils ont produit des souris « knock out » pour cette protéine et ils ont observé de la méthylation *de novo* effectuée par Dnmt3a et 3b qui n'était pas dirigée aux CpG,

contrairement aux souris « contrôles » pour qui le Dnmt3L fonctionnait correctement. On n'a pas trouvé jusqu'à maintenant cette méthyltransférase dans les mitochondries et cela pourrait peut-être expliquer pourquoi la méthylation ne se fait pas préférentiellement aux CpG au sein de celles-ci (van der Wijst et Rots 2015).

De plus, la distribution des différentes méthyltransférases dans les mitochondries, de même que la méthylation, varient selon le type de tissu ou de cellule (Wong et al., 2013). En effet, Wong et al., (2013) ont déterminé que les mitochondries des tissus « excitables », c'est-à-dire le cœur, les muscles squelettiques et le système nerveux, étaient plus propices à être méthylés. La Dnmt3a est plus couramment observée à l'intérieur des mitochondries de ces tissus, tandis que la Dnmt1 est accolée à la membrane externe de l'organelle, sans être présente à l'intérieur. D'autres auteurs sont parvenus à des résultats similaires concernant la méthylation dans différents tissus ou pour un même tissu dans le temps. Par exemple, Liu et al., (2016) ont trouvé des différences entre les patrons de méthylation de l'ADNmt pour des cellules du sang et de la salive. La méthylation se fait à des endroits précis sur des CpG dans les gènes et on peut constater une variation importante d'un tissu à l'autre. La variation entre les patrons est encore plus évidente pour l'hydroxyméthylation (Li et Liu, 2011; Nestor et al., 2011), 5hmC, dont le rôle est encore flou, mais qui agit probablement de façon similaire à la méthylation. Le passage de 5mC à 5hmC serait régulé à l'aide d'une TET (Ghosh et al., 2016). L'hydroxyméthylation se ferait donc aussi de manière non aléatoire et serait associée au type de cellule (figure 2).

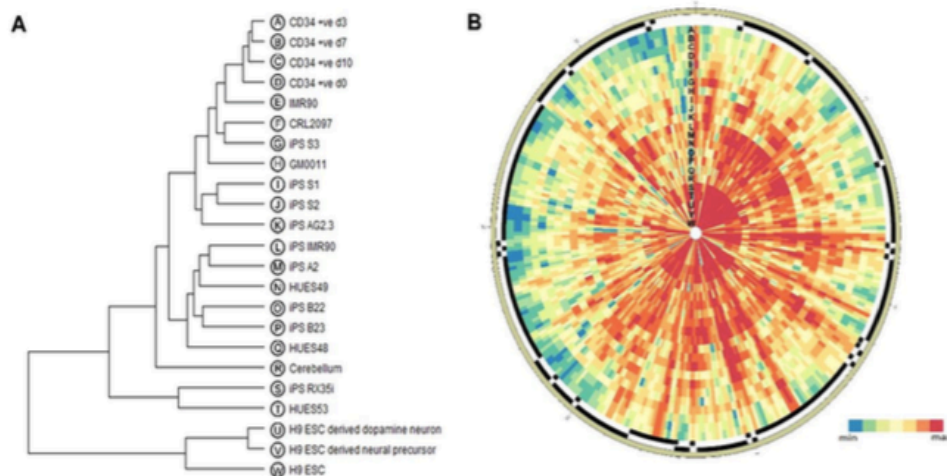


Figure 2 Patrons d'hydroxyméthylation des cytosines pour le génome mitochondrial humain (Ghosh et al., 2016)

La densité en hydroxyméthylcytosines a été mesurée pour des séquences de 50 paires de bases
 A) L'arbre présente le lien entre chacun des 23 types cellulaires étudiés (l'arbre a été réalisé avec R, en fonction de la similarité des cellules) B) Le génome mitochondrial (circulaire) complet est présenté avec les types cellulaires classés de haut en bas, en suivant le rayon du cercle. Les couleurs représentent la densité d'hydroxyméthylation (rouge étant le plus élevé).

Il y a évidemment un lien à faire ici avec cette méthylation (ou hydroxyméthylation) et l'expression du génome mitochondrial. En effet, cette méthylation différentielle, puisqu'elle est non aléatoire et coordonnée de la sorte, a sans doute un rôle particulier dans la cellule. Le rôle le plus probable serait l'expression du génome mitochondrial complet ou de certains gènes mitochondriaux en particulier. Il faut alors porter attention aux emplacements où survient la méthylation dans le génome afin de comprendre comment l'expression du génome pourrait être affectée. C'est ce qu'ont fait Bellizzi et al., (2013), Ghosh et al., (2014) et Bianchessi et al., (2016) pour la région D-loop et pour les « *Genes Start Sites* » (GSS) qui pourraient respectivement permettre de contrôler la transcription de l'ADNmt et de réprimer localement des gènes particuliers. De la méthylation a été trouvée à la fois dans la D-loop et dans les GSS, ce qui renforce l'idée d'un contrôle de l'expression de l'ADNmt. De plus, une augmentation du

niveau des sites déméthylés dans la D-loop de cellules tumorales a été fortement associée à une expression accrue du gène mitochondrial *nad2* et du nombre de copies d'ADNmt (Feng et al., 2012). Cependant, le lien entre la méthylation de l'ADNmt et le contrôle de l'expression des gènes mitochondriaux reste à être clairement démontré.

1.1.4 Double transmission uniparentale de l'ADNmt (DUI)

La règle générale chez les animaux concernant la transmission de l'ADNmt est la SMI (Strict Maternal Inheritance), c'est-à-dire la transmission du génome mitochondrial de la mère seulement à ses enfants (Hayashi et al., 1978). Or, il existe chez les bivalves plusieurs espèces qui ont un système de transmission du génome mitochondrial unique qui défie cette règle générale (l'existence de ce système a été confirmée chez 105 espèces pour l'instant, Gusman et al., 2017). Il s'agit du système DUI (Doubly Uniparental Inheritance), qui à l'opposé de la SMI, permet également la transmission du génome mitochondrial du père. Cependant, ceci se fait selon des conditions particulières : l'ADNmt du père est transmis par les gamètes (spermatozoïdes) et il est généralement localisé dans les gonades de la descendance mâle, tandis que les tissus somatiques des mâles comme des femelles, en plus des gonades femelles, contiennent plutôt l'ADNmt qui leur a été transmis par la mère (Breton et al., 2007). Ceci fait en sorte que les mâles sont hétéroplasmiques (ils possèdent plus d'un type d'ADNmt), tandis que les femelles, en principe, sont homoplasmiques (elles ne possèdent qu'un type d'ADNmt) (figure 3). Cela fait aussi en sorte que les deux génomes mitochondriaux (mâle et femelle, M et F) ont évolué indépendamment l'un de l'autre depuis que ce système a été mis en place et c'est la raison pour laquelle on observe une grande divergence génétique entre ces deux génomes (Doucet-Beaupré et al., 2010).

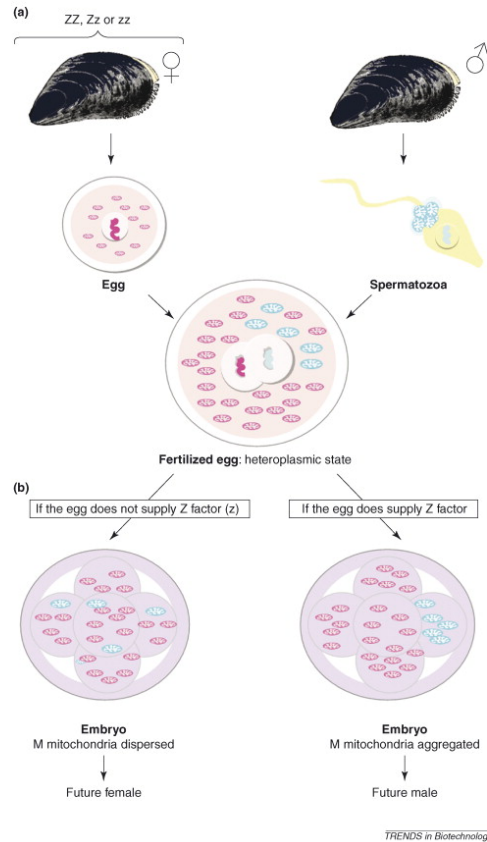


Figure 3 Schéma représentant le système de double transmission uniparentale (DUI) (Breton et al., 2007)

Transmission de l'ADNmt F par les oeufs des femelles; transmission des ADNmt M par les spermatozoïdes des mâles

Ségrégation ou élimination des mitochondries dans le zygotes; formation d'une femelle (en bas à gauche) ou d'un mâle (en bas à droite)

Chez certaines espèces comme les moules bleues du genre *Mytilus* spp. ainsi que la palourde marine *Ruditapes philippinarum*, l'ADNmt M a été trouvé dans certains tissus somatiques mâles et femelles (Garrido-Ramos et al., 1998 ; Ghiselli et al., 2011). Obata et al., (2011), se sont penchés sur l'expression de chacun des génomes dans les différents tissus des mâles et des femelles chez l'espèce *Mytilus galloprovincialis*. Ils ont observé une expression de l'ADNmt F dans les tissus somatiques ainsi que dans les gonades mâles et femelles (figure 4). L'ADNmt M, même lorsque présent dans les tissus somatiques, était seulement exprimé dans les gonades mâles, de manière plus forte durant la spermatogénèse. Les chercheurs suggèrent donc qu'il existe un mécanisme de contrôle de la transcription de l'ADNmt qui permettrait une expression différentielle des génomes dans les différents tissus. Dans le cas du génome M présent dans les tissus somatiques, il pourrait s'agir d'une répression de sa transcription, mais

ceci reste à être clairement démontré puisque Milani et al. (2014) ont observé une faible transcription du génome M dans les tissus somatiques des mâles chez l'espèce *Ruditapes philippinarum*. Dans leur étude, Milani et al., (2014) ont également noté une absence de corrélation entre le taux de transcription de l'ADNmt et la quantité d'ADNmt. Les auteurs suggèrent l'existence de mécanismes complexes de régulation de la transcription mitochondriale chez les espèces à DUI. Cette absence de corrélation entre le taux de transcription de l'ADNmt et la quantité d'ADNmt a aussi été observée chez l'espèce de moule d'eau douce *Venustaconcha ellipsiformis* (Capt et al., 2018). Chez *V. ellipsiformis*, tout comme chez *R. philippinarum*, l'abondance des transcrits des gènes mitochondriaux varie en fonction du gène et du génome (figure 5) (Ghiselli et al., 2013 ; Capt et al., 2018). Il serait pertinent de vérifier si cette abondance peut être régulée, entre autres, par l'épigénétique.

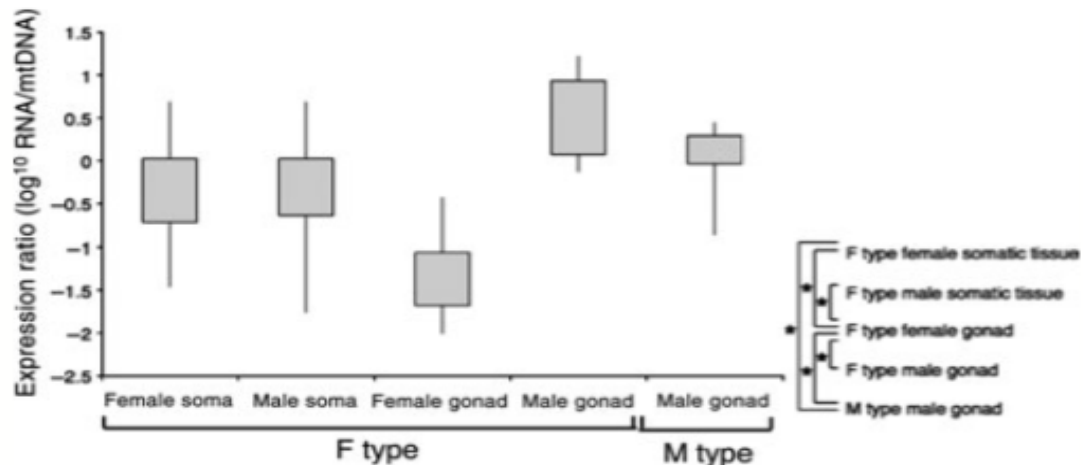


Figure 4 Ratios de l'expression des ADNmt M et F dans les gonades et les tissus somatiques mâles et femelles chez *Mytilus galloprovincialis* (Obata et al., 2011)

On peut constater des taux d'expressions variables pour l'ADNmt F en fonction du tissu. Ce taux est particulièrement bas dans les gonades femelles. Étonnamment, le taux d'expression dans les gonades mâles est plus élevé. L'expression de l'ADNmt M n'a pas été détectée ailleurs que dans les gonades mâles.

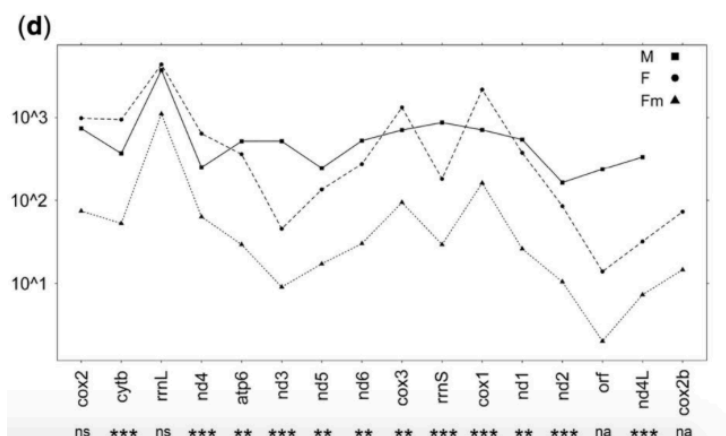


Figure 5: Abondance des transcrits des gènes mitochondriaux en fonction du gène et du génome chez *Ruditapes philippinarum* (Ghiselli et al., 2013)
Le nombre de transcrits varie en fonction du gène et du génome. M: transcription du génome M chez les mâles; F: transcription du génome F chez les femelles; Fm: transcription du génome F chez les mâles

1.2 Objectif et hypothèses

L'objectif de mon projet de maîtrise était de déterminer, chez une espèce de bivalve à DUI, les patrons de méthylation des ADN mitochondriaux F et M en fonction des différents tissus où ils sont trouvés (tissus hétéroplasmiques ou non). Ce projet s'inscrit dans un autre, plus grand, et dont le but est de déterminer si la méthylation de l'ADNmt a une fonction biologique précise, en permettant par exemple la régulation de l'expression des génomes mitochondriaux. Il s'agirait alors de comprendre comment cette régulation se produit et pourquoi elle pourrait être utile, notamment pour les espèces à DUI. Le projet portait sur la moule d'eau douce *Venustaconcha ellipsiformis*, dont les ADNmt F et M complets avaient été caractérisés, et chez qui une expression différentielle des gènes mitochondriaux F et M avait été observée (Capt et al., 2018). Chez cette espèce, la présence de l'ADNmt M a été observée non seulement au niveau de la gonade mâle mais également dans les tissus somatiques d'individus mâles et

femelles (Breton et al., 2017). *V. ellipsiformis* représente un modèle intéressant puisqu'elle constitue la première espèce de moules d'eau douce pour laquelle la présence de l'ADNmt M a été observée dans les tissus somatiques (jusqu'à maintenant l'ADNmt M avait été répertorié seulement au niveau des gonades mâles chez les moules d'eau douce) (Zouros 2013 ; Breton et al., 2017). Cette situation offre donc la possibilité d'étudier les mécanismes épigénétiques de contrôle de la transcription des ADNmt F et M au sein de tissus homoplasmiques et hétéroplasmiques. L'objectif du projet avait pour hypothèse nulle que les patrons de méthylation seraient similaires pour les deux types d'ADNmt dans tous les tissus. Une hypothèse alternative était que les ADNmt F et M présenteraient des patrons de méthylation différents à l'intérieur d'un même tissu et/ou entre les différents tissus. Nous avons également pour objectif de vérifier si l'abondance des transcrits des gènes était corrélée au taux de méthylation.

Nous étions confiants en ce qui concerne la présence de méthylation dans l'ADNmt puisque certains patrons avaient déjà été identifiés à plusieurs reprises dans les mitochondries de mammifères (ex : Ghosh et al., 2014 ; Wong et al., 2013). Contrairement aux autres espèces animales qui n'ont qu'un seul type d'ADNmt à réguler, les bivalves à DUI ont à gérer deux types de mitochondries et deux haplotypes mitochondriaux hautement divergents. Nous avons supposé que la méthylation de l'ADNmt pouvait être un mécanisme clé permettant de réguler le système DUI chez les bivalves. En comparant les patrons de méthylation, nous nous attendions à déterminer quelles étaient les différences qui existaient entre les tissus. Ces résultats, combinés avec des analyses d'expression génique par PCR quantitative, nous ont permis d'établir un lien entre la méthylation et l'expression des gènes mitochondriaux.

2 Article: Is mtDNA methylation functionally relevant in animals? Insights from a naturally occurring heteroplasmic system

Karim Bouvet^{1,*}, Davide Guerra¹, Bernard Angers¹, Sophie Breton^{1,*}

¹ Department of Biological Sciences, Université de Montréal, Montréal, QC, Canada, H2V 2S9

* Authors for correspondence: karim.bouvet-hasab.alla@umontreal.ca; s.breton@umontreal.ca

En préparation pour *Frontiers in Genetics*



2.1 Introduction

Although DNA methylation has been recognized for decades as an important epigenetic mechanism in the nucleus (Jones, 2012), its presence and role in the mitochondria have been a matter of debate since the 1970s (reviewed in van der Wijst and Rots, 2015; Castegna et al., 2015; Bacaliani et al., 2017; D'Aquila et al., 2017; Matilainen et al., 2017; Mposhi et al., 2017). On the one hand, some recent studies suggest that cytosine methylation is virtually absent or very rare in the human and mouse mitochondrial DNA (mtDNA) (e.g. Liu et al., 2016; Mechta et al., 2017; Matsuda et al., 2018; Owa et al., 2018), thus questioning its functional relevance. On the other hand, a growing body of literature suggest that methylation adds a potential epigenetic layer of regulation control of mtDNA replication and transcription (D'Aquila et al., 2017 and reviews aforementioned). Also, recent studies using human cultured cells have reported significant correlation between mtDNA epigenetic marks, environmental exposure and peculiar phenotypes, aging and diseases (D'Aquila et al., 2017 and reviews aforementioned). Coupled to this is the discovery of a DNA methyltransferase targeted to and functioning in human and mouse mitochondria, where it modifies mtDNA methylation levels and presumably influences mitochondrial transcription (Chestnut et al., 2011; Shock et al., 2011; Wong et al., 2013; Saini et al., 2017). However, although mtDNA methylation has been detected by different methods over the years (van der Wijst and Rots, 2015; Mposhi et al., 2017 and reviews aforementioned), its physiological role(s) still remain unclear. Due to the difference in structure and gene organisation between mitochondrial and nuclear genomes, we can expect a different mode of action in each case. For example, methylation in the nucleus, even though it was proven to regulate gene expression when present in gene bodies (e.g. Rivière, 2014), is generally described as concentrated in CpG islands (CGI) in gene promoters where it represses

transcription (Gardiner-Garden and Frommer, 1987; Jones, 2012). Unlike the nuclear DNA, which contains at least one promoter region per gene, the human mtDNA contains only three promoter regions located within or in the vicinity of a region known as the mitochondrial displacement loop (D-loop), or mtDNA non-coding region, that produce polycistronic transcripts (Clayton, 2000; Mposhi et al., 2017). Also, besides CpG methylation, it has been shown that mtDNA exhibits a peculiar non-CpG methylation pattern (CpC, CpA and CpT) (Mposhi et al., 2017). To date, most articles published on mtDNA methylation focused on this “control region” in human and mouse mtDNAs (e.g. Bellizzi et al., 2013; Gao et al., 2015; Zheng et al., 2015; Armstrong et al., 2016; Tong et al., 2017), and very little is known regarding mtDNA methylation in other animals, including invertebrates (e.g. Li et al., 2017).

To better understand the purpose of mtDNA methylation and the mechanism by which it could regulate gene expression, it is crucial to compare different mitochondrial contexts such as different taxa, different tissues or different development time (see Wong et al., 2013; Ghosh et al., 2016; Liu et al., 2016; Devall et al., 2017 for a few examples in mammals). One particularly interesting model to study mtDNA methylation is the system of doubly uniparental inheritance (DUI) of mtDNA in bivalve molluscs, a naturally heteroplasmic system where two highly divergent mitochondrial genomes coexist inside the same organism (Breton et al., 2007). Specifically, DUI is a “mother-to-daughter & son” and “father-to-son” mtDNA inheritance system where females are usually homoplasmic for the maternally-transmitted F mtDNA and males are usually heteroplasmic and possess the F mtDNA in their soma and the paternally-transmitted M mtDNA in their gametes (Breton et al., 2007). These two sex-specific mt genomes can be extremely divergent (up to 41% DNA divergence; Doucet-Beaupré et al., 2010). This is very impressive considering that heteroplasmy in other species (in humans for example), even when it is the result of a few mutated bases, is known to cause severe diseases

and even death (Wallace 2016; Picard et al., 2016). In freshwater mussels, which are the animal models of the present study, mitochondrial genes are encoded on both strands and each gender-associated mtDNA possesses a novel mt-encoded protein gene (Breton et al., 2009; Breton et al., 2011), i.e. in addition to the typical set of 13 protein-coding genes usually found in animal mtDNAs (Breton et al., 2014). These genes, called *f-orf* in female-transmitted mtDNA and *m-orf* in male-transmitted mtDNA, have hypothesized roles in the maintenance of DUI and sex determination in bivalves (Breton et al., 2011; Mitchell et al., 2016), but this remains to be clearly demonstrated.

Although the general DUI portrait is depicted as male gametes carrying only the M mt genome whereas all other tissues (males and females) are exempt of it, many exceptions have been reported in marine and freshwater DUI species, showing the presence and expression of the M mtDNA in male and female somatic tissues (reviewed in Zouros, 2013; Breton et al., 2017). Overall, these studies showed that although both F and M mtDNAs can be present in male and female somatic and gonadal tissues, they are co-expressed more often in males than in females, and Breton et al. (2017) speculated that this could be the general scenario for all DUI species. Because of the potentially deleterious nature of heteroplasmy, Breton et al. (2017) also suggested the existence of different mechanisms in DUI species to deal with this possibly harmful situation, such as silencing mechanisms for the M mtDNA at the transcriptional, post-transcriptional and/or post-translational levels. One hypothesis is that mtDNA methylation could very well be one mechanism underlying this differential expression between sexes (or among tissues). Our aim is thus to determine if methylation is present in the F and M mtDNAs of DUI bivalves, and to see if there is a change in methylation pattern depending on the context in which a haplotype is found.

For this study, we analyzed mitochondrial methylation patterns using whole genome bisulfite sequencing (WGBS) of heteroplasmic somatic and gonadal tissues of adult individuals of the freshwater mussel species *Venustaconcha ellipsiformis*, a bivalve with DUI (Breton et al., 2009), as a first step to see if methylation could represent a mechanism regulating mitochondrial expression in DUI species. We assessed the presence of methylation in both M and F mt genomes, then compared the patterns between the two. We also looked for variation or similarities among tissue samples for a same haplotype (different tissues and different sex). Finally, we discuss the distribution of methylation marks for both F and M mt genomes.

2.2 Materials and Methods

2.2.1 Specimen collection and sample preparation

A total of twelve adult specimens (6 males and 6 females) of *Venustaconcha ellipsiformis* were collected from Straight River (Minnesota, USA; Lat 44.006509, Long -93.290899) in July 2014. Animals were shipped alive to the Université de Montréal, where they were dissected and sexed by microscopic examination of gonad smear. Gonadal and somatic tissues (gills, foot, mantles and adductor muscles) were immediately frozen at -80°C for preservation (Breton et al. (2017)). Four other mollusc species plus one brachiopod were also obtained: one chiton *Chaetopleura apiculata* sent from Gulf Specimen Marine Laboratories, Inc. (Panacea, FL, USA), one brachiopod *Glottidia pyramidata* sent by the University of Hawaii, one slug *Arion fuscus* from the south-east of Quebec, and one male blue mussel *Mytilus edulis* bought from a grocery store. Only the foot was used for the chiton and the slug, and only gonads were used for the blue mussel and the brachiopod.

2.2.2 DNA extraction and purification

Total DNA was extracted from every sample (for *V. ellipsiformis*: 5 tissues/specimen except for 2 females for which only 4 tissues were obtained, for a total of 58 DNA extractions), plus the other 5 mollusc and brachiopod samples. Extractions were performed with the MasterPure Complete DNA and RNA Purification Kit (Epicentre). Quality and quantity of the extracted DNA were respectively verified with a 1% agarose gel and with a Biodrop μ LITE spectrophotometer.

2.2.3 Chosen samples for sequencing

As described in Breton et al., (2017) about the detection of the M mtDNA in gonadal and somatic tissues in males and females *V. ellipsiformis*, 30 of the 58 samples were found to be heteroplasmic (i.e. both F and M mitochondrial genomes were detected). Since our objective was to determine if mtDNA methylation could act as a way of regulating the expression of the two mitochondrial genomes present in a same tissue, we selected heteroplasmic tissues for this study. We also chose our samples in function of the expression of the M genome in a tissue (i.e. the F genome was expressed in all tissues, but not the M genome; Breton et al., 2017). Specifically for this project, we chose 12 samples, 11 of them originating from male individuals (3 gills, 2 adductor muscles, 2 gonads, 3 mantles and one foot) and one from a female (one adductor muscle). The M genome was expressed in all samples where it was present, except for one male gill and one male mantle.

2.2.4 Mitochondrial reference sequences

To obtain mtDNA reference sequences, total DNA from one *V. ellipsiformis* male gonad (to obtain the M genome) and one male foot (to obtain the F genome) was used to construct DNA libraries that were paired-end sequenced (2 x 100 bp) using an Illumina HiSeq platform (G  nome Qu  bec Innovation Centre, Montr  al, QC, Canada). Trimmomatic 0.32 (Bolger et al., 2014) was used to trim reads, which were then assembled with PEAR 0.9.6 (Zhang et al., 2013), a free tool available on the Galaxy platform (Afgan et al., 2018). Quality of the reads was checked at every step with FastQC 0.11.5 (Andrews, 2010). The two complete mt genomes were assembled using MITObim 1.8 (Hahn et al. 2013), by giving as baits (input sequences), complete F and M *V. ellipsiformis* mt genome sequences from a different population previously published (Breton et al., 2009; Genbank accession numbers FJ809753 and FJ809752 for the F and M mtDNAs, respectively). The new mt genome sequences, which are both 99.6% similar to the ones from Breton et al. (2009), are available in GenBank (accession numbers XXXXXX and UUUUUU for the F and M mtDNA, respectively). This very low level of polymorphism at the species level indicates that using one F and one M reference mt genomes for all samples in the present study is correct. For the chiton, the slug and the brachiopod mt genomes, a *de novo* assembly was necessary; the protocol used was the same, except that the bait was not a complete genome (*cox1* sequences were used; see Guerra et al, 2018 for more details). For *Mytilus edulis* the process was the same as for the two *V. ellipsiformis* mt genomes. Having only a male gonad of *M. edulis*, which contained mainly the M genome, we were able to assemble this genome but not the F. Just like *V. ellipsiformis*, a complete M genome was used as a bait (Genbank accession number AY823624).

2.2.5 Bisulfite sequencing and detection of methylation

Whole genome bisulfite sequencing (WGBS) was performed at Génome Québec Innovation Centre (Montréal, QC, Canada) using EpiGenome™ Methyl-Seq Kit (Epicentre, Illumina) according to the manufacturer's protocol. Libraries were subjected to 150 bp paired-end sequencing using Illumina MiSeq instrument. As for mtDNA sequence references, reads were trimmed with Trimmomatic 0.32 and assembled with PEAR 0.9.6, checking their quality at every step with FastQC 0.11.5. Preprocessed reads were mapped onto a mitochondrial reference sequence using Bismark Bisulfite Mapper v.0.17.0 (Krueger and Adrews, 2011). The program operates in three separate steps: first, a genome reference conversion is required so that the bisulfite converted reads map to the converted genome. Then, the alignment step calls out the methylation status of every cytosine – this is done using Bowtie (version 2.2.4; Langmead and Salzberg, 2012). Finally, Bismark methylation extractor, which is part of the Bismark Bisulfite Mapper v.0.17.0, organises the information in an understandable fashion, enabling further analyses. This allowed to determine the exact number of methylated and unmethylated cytosines at a precise position of the genome and to calculate the methylation proportion for every position of the genome. This was done for both coding and template strands separately (and also for both + and - strands for non-coding regions). Strand-biased DNA methylation has previously been reported (e.g. Singal and vanWert, 2001; Luo et al., 2003; Zhong et al., 2014), and could possibly account for transcriptional differences among genes. For example, since the coding strand does not have an active role in RNA polymerisation, it has been suggested that the methylation state of the template strand would more likely have an impact on transcription rates, with an unmethylated and well transcribed template strand contributing to higher levels of mRNAs (and vice versa) (Stöger et al., 2011).

2.3 Results

2.3.1 General sequencing and mapping information

The first thing we wanted to look at was the effect of bisulfite treatment on sequencing quality. Figure 6 shows a comparison of the total number of reads after trimming (i.e. only those having a good quality) for both shotgun and bisulfite sequencing, in samples from four mollusc and one brachiopod species. Our results show that the number of good quality reads is not necessarily higher in shotgun sequencing, suggesting that the bisulfite treatment does not affect the quality of the sequencing.

Information related to the bisulfite sequencing of *V. ellipsiformis* has been compiled in Table 1. Overall, between 37.32 and 50.96 millions methylated reads were assessed for the quality, and between 20.94 and 43.60 high quality reads were obtained. A total of 52344 reads mapped to the F mt genome in the only female somatic tissue, covering 100% of the genome, an average of 22650 reads mapped to the F mtDNA in male somatic tissues, covering >99.99% of the genome, and an average of 900 reads mapped to the F mt genome in male gonads (with ~99.62% genome length coverage). The number of reads that mapped to the M mt genome in heteroplasmic somatic tissues was very low with an average of 140 in male soma (with ~56.97% genome length coverage) and only 50 in the female adductor muscle (with ~48.26% genome length coverage). In male soma, the lowest coverage was observed in a mantle (0.15X coverage) and the highest (6.42X coverage) also in a mantle, which interestingly corresponded to an heteroplasmic sample in which the M mtDNA was not expressed (Table 1; Breton et al., 2017). Conversely, an average of 9900 reads mapped to the M mt genome in male gonads, covering 100% of the M genome.

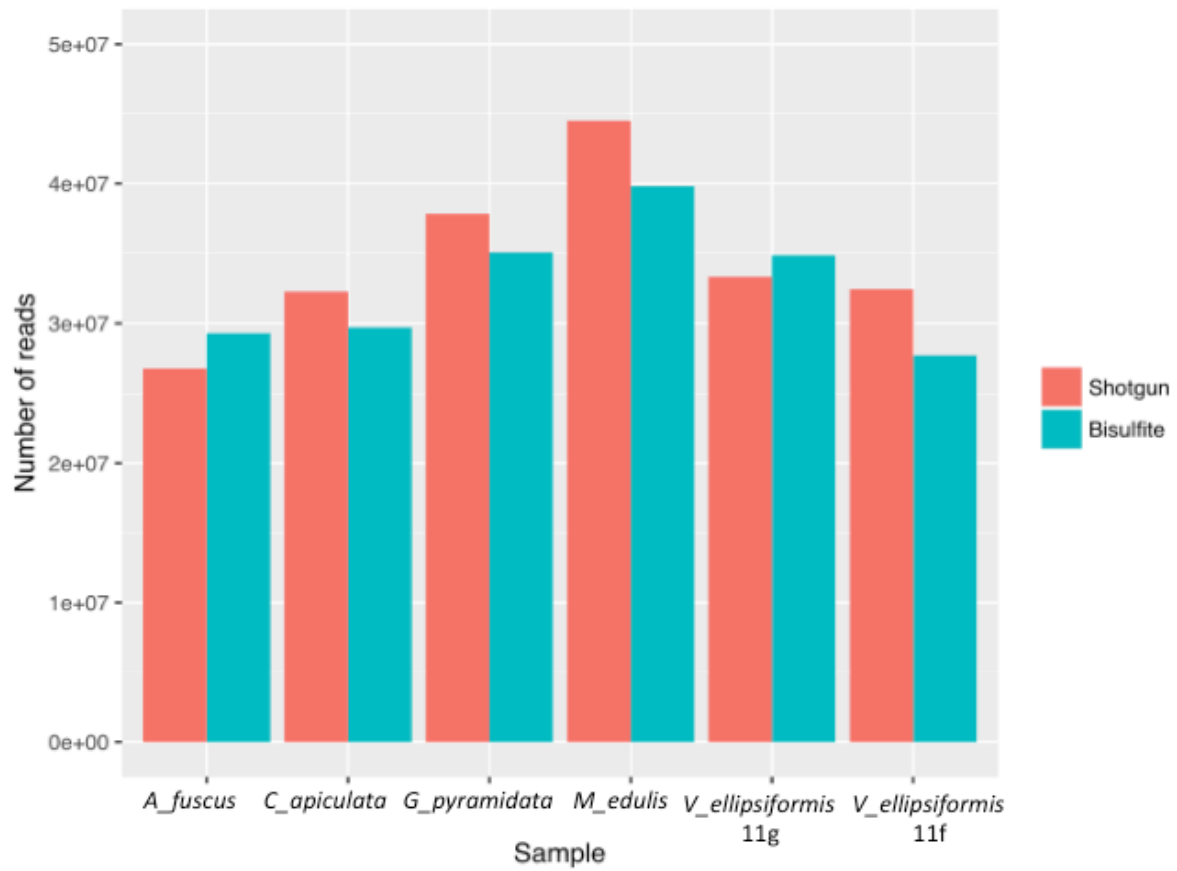


Figure 6: Number of reads after trimming for every sample sequenced by shotgun and bisulfite sequencing (WGBS). g: gonad; f: foot

Table I. General information related to the bisulfite sequencing of *V. ellipsiformis*

Sample	Total number of reads		Genome	Number of mapped (mt) reads	Mean coverage	% of mt genome covered by reads
	Before trimming	After trimming				
Male 1gi	47364469	39493691	1giF	28016	508.54	100%
			1giM	36	0.60	34.09%
Female 2a.m	44785340	37361856	2a.mF	52344	949.07	100%
			2a.mM	50	0.84	48.26%
Male 7gi	37319305	20940573	7giF	5865	105.90	100%
			7giM	32	0.55	34.02%
Male 7a.m	46701589	31323410	7a.mF	40065	726.04	100%
			7a.mM	177	3.00	83.49%
Male 8gi*	47542010	36271327	8giF	28957	525.91	99.98%
			8giM	10	0.17	8.51%
Male 8go	50963114	43603876	8goF	1126	20.28	99.94%
			8goM	8822	147.94	100%
Male 8m	47530581	38441535	8mF	27256	491.40	100%
			8mM	347	5.80	97.97%
Male 9m*	44136487	32270133	9mF	20501	371.92	100%
			9mM	382	6.42	98.20%
Male 11go	42511007	34859399	11goF	673	12.43	99.29%
			11goM	10952	187.84	100%
Male 11m	43776634	29261884	11mF	17092	309.67	100%
			11mM	9	0.15	11.71%
Male 11a.m	39020969	31646878	11a.mF	9862	178.93	99.94%
			11a.mM	140	2.39	78.59%
Male 11f	39625754	27698522	11fF	26399	488.54	100%
			11fM	105	1.79	66.11%

NOTE – a.m: adductor muscle; f: foot; gi: gills; go: gonad; m: mantle; *: heteroplasmic samples in which the M mtDNA was not expressed

Genome	Average number of mapped reads	Mean coverage	% of mt genome covered by reads
F in male soma	22668	411.87	99.99%
F in female soma	52344	949.07	100%
F in male gonads	900	16.35	99.62%
M in male soma	138	2.32	56.97%
M in female soma	50	0.84	48.26%
M in male gonads	9887	167.89	100.00%

2.3.2 Methylation levels in F and M mitochondrial genomes

Figure 7 shows the quantitative results of methylated cytosines averaged (%) for the template and coding strands, for the F mtDNA in male soma, female soma and male gonads, as well as for the M mtDNA in male gonads. Because of the very low coverage obtained for the M mtDNA in somatic tissues, methylation levels for the M mtDNA in male and female soma are not shown in Figure 7. Supplementary Tables 1-6 present average methylation levels for the M and F mtDNAs in male and female somatic tissues and male gonads for every protein-coding gene, tRNA, rRNA, and total intergenic regions for both template and coding strands. Supplementary File 1 presents our complete data for each sample. Overall, our results show that (i) methylation levels are relatively low and oscillate around 1% and they apply to all cytosines, regardless of their context (CpG, CpA, CpC and CpT), (ii) methylation levels in the F mtDNA – whether it is found in male or female soma – are similar for both coding and template strands (around 1%), and (iii) methylation levels for coding and template strands seem different between F and M mtDNAs in male gonads: for the coding strand, the most methylated protein-coding gene for the F genome is *cox3* (4.49% and also some tRNAs obtained >5%) and the least methylated genes are *f-orf* and *nad6* (0%). For the M genome, the most methylated protein-coding gene is also *cox3* (1.78%), but the least methylated gene is *nad2* (0.44%). For the template strand, the most methylated protein-coding gene in the F genome is *cytb* (1.46%) and the least methylated genes are *atp8* and *nad4l* (0.14% et 0.27%, respectively), whereas in the M genome, the most methylated protein-coding gene is *m-orf* (2.74%) and the least methylated is *nad6* (0.53%).

Interestingly, while there seem to be no difference for total protein-coding gene regions, our result suggest that total rRNA are a bit more methylated on the template strand for

the F mtDNA in male gonads whereas these differences do not really appear for the M mtDNA in the same tissue (Supplementary Tables 1-6).

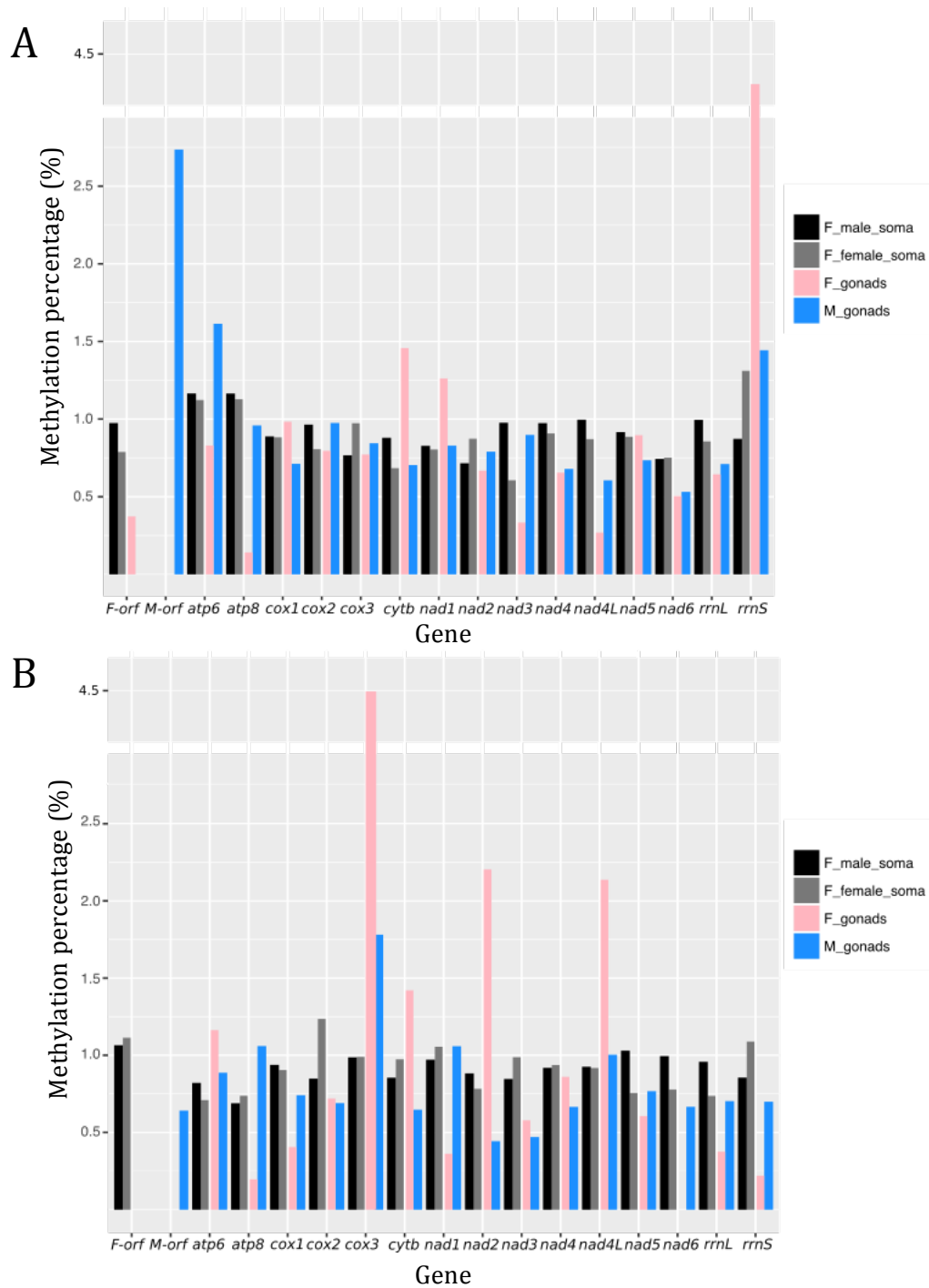
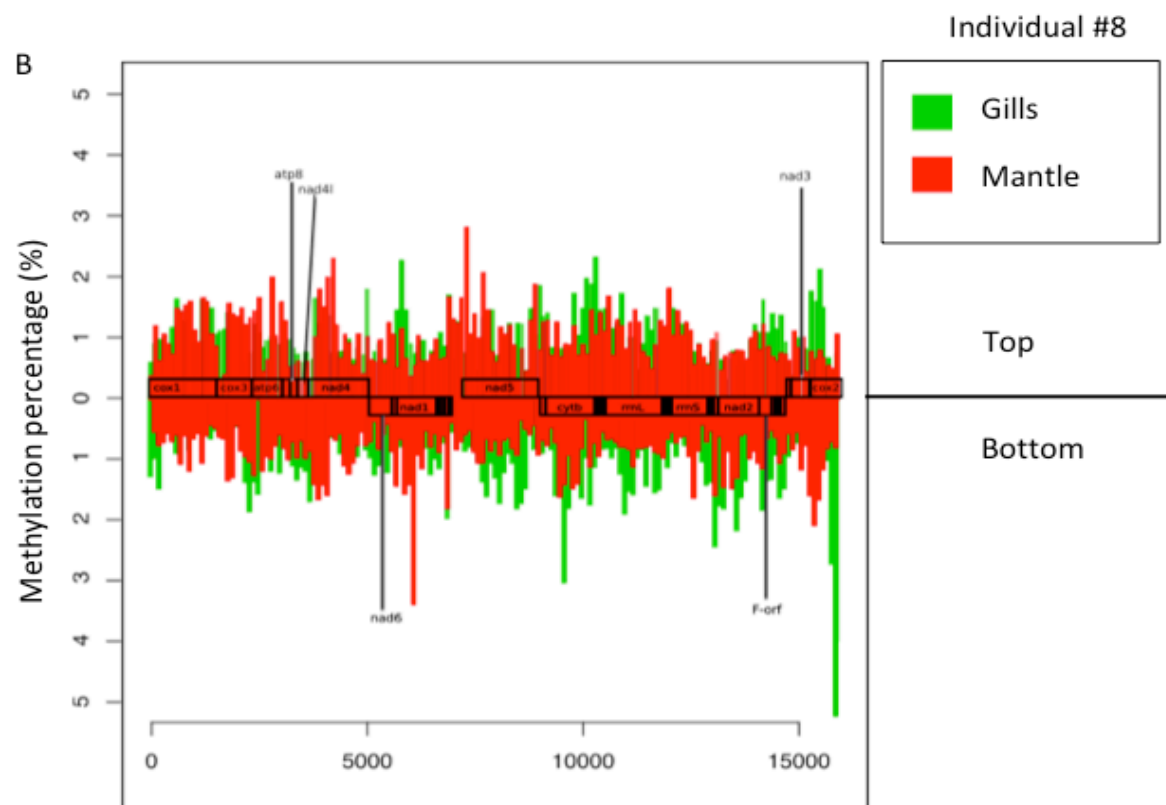
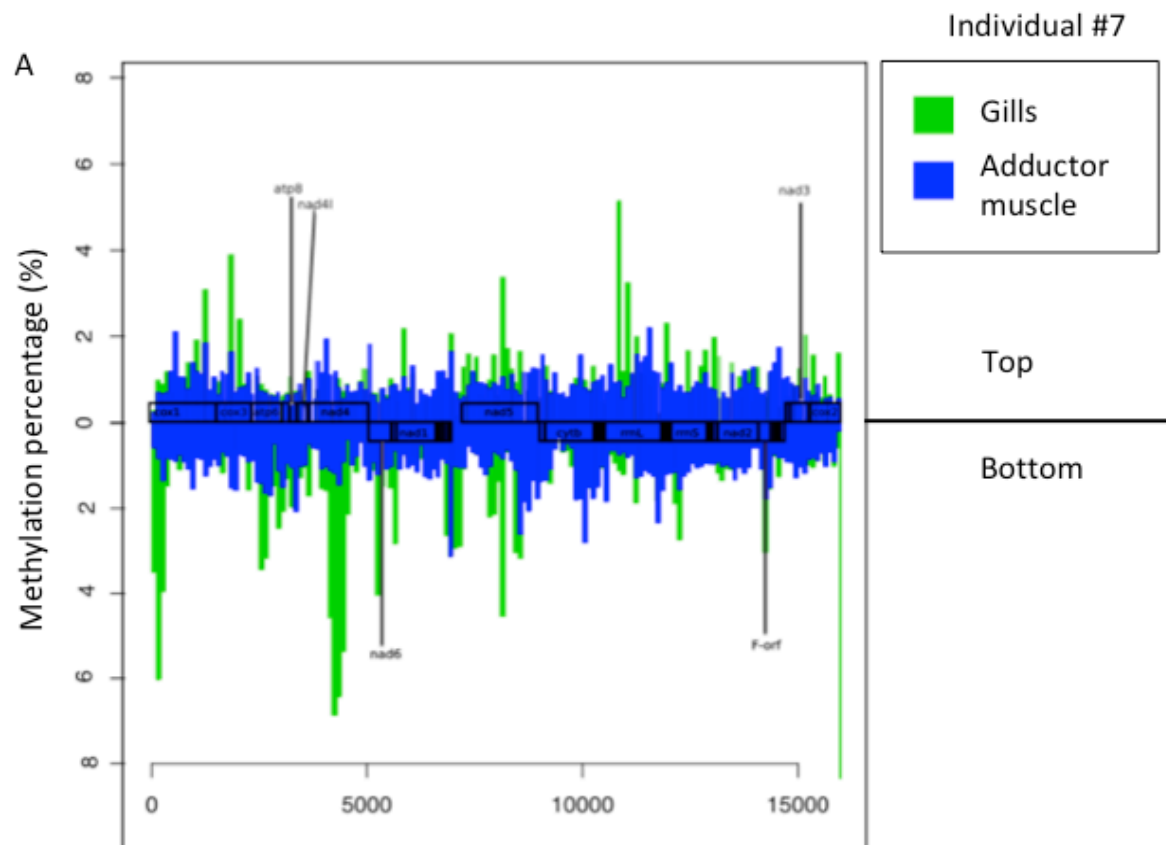


Figure 7: Methylation percentage on the (A) template and (B) coding strands of F and M mt genomes present in soma or gonads of *V. ellipsiformis*

2.3.3 Effect of the tissue, sex and mt genome on methylation patterns

2.3.3.1 Methylation patterns of the F mtDNA in different tissues within individuals

Methylation levels for the well-covered F mtDNA were averaged across 100bp non-overlapping windows for each sample. This allowed for a comparison of the tissues originating from a same individual, but also for a comparison of the same tissues across individuals. There is no striking difference in the methylation patterns of the F mtDNA in different tissues within individuals or across different individuals (Fig. 8); the only peaks higher than 5% being located at the extremities of the genome and probably the result of a bias. In fact, mtDNA is a circular genome that we cut arbitrarily for the purpose of these analyses. For this reason, reads do not map as efficiently at the extremities and we encounter such a bias. Other than that, it is possible to see some peaks inside genes but not necessarily located in the same genes and following a consistent pattern. In other words, it is difficult to identify a peak present in two different tissues of the same individual, but it is equally difficult to identify one that is shared for a tissue in different animals (see also Fig. 9). Globally, our results indicate a lower methylation percentage for the male foot in general, which does not exceed 1%, when compared to all other samples. It is worth noting that this sample was sent for sequencing separately.



are distinctively more methylated than the rest, as it is the case for the gills of individual #7 for example. Again, the most prominent peaks at the extremities are probably the result of a bias, and it would be best not to focus on these.

The male and female adductor muscles do not appear to be very differently methylated, when compared to the gills samples: we can even see that the highest peak is shared between samples, no matter the sex. This peak is located in the same position as the one found in the gonad of individual #8 (Fig. 11a), i.e. in one putative origin of replication/transcription (Breton et al. 2009). Other smaller peaks also seem to be shared, like the ones at the end of *nad5* on the template strand. The recurrence of peaks in the non-coding region following *nad5* could represent a pattern that would not necessarily be related to the individual, the sex or to a certain extent, the type of tissue.

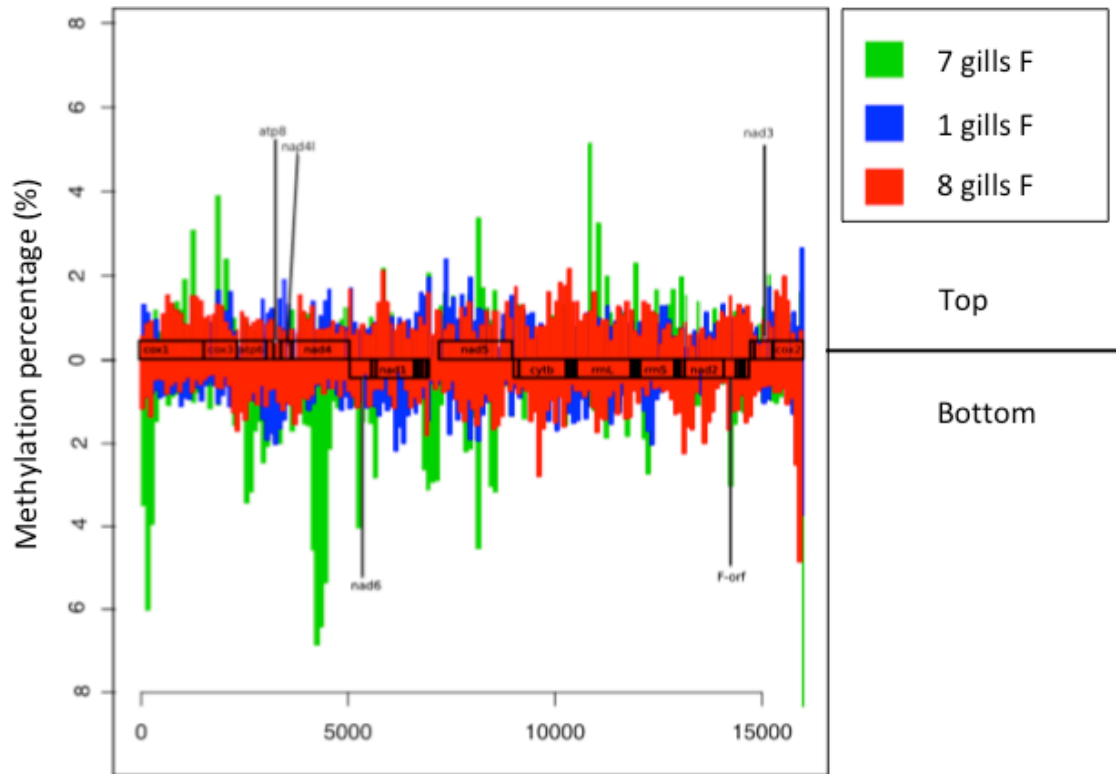


Figure 9: Mean methylation percentage for 100bp windows along the F mt genome found in the gills of three males *V. ellipsiformis*

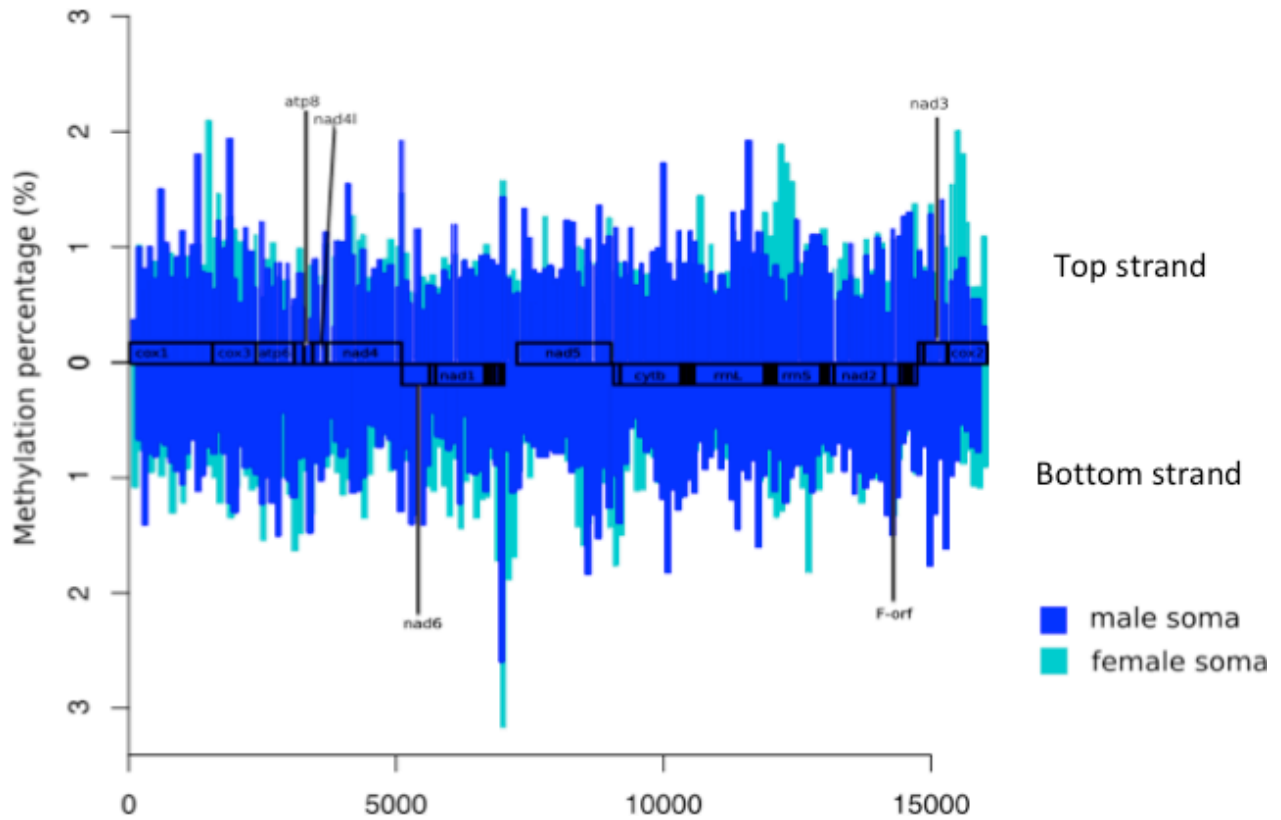


Figure 10: Mean methylation percentage for 100bp windows along the F mt genome found in male and female adductor muscles of *V.ellipsiformis*

2.3.3.3 Methylation patterns of the F and M mt genomes present in male gonads

Interestingly, methylation reaches much higher levels for the F mtDNA in male gonads. High methylation levels can also be observed in the M mt genome in male gonads (Fig. 11). This tissue does indeed stand out in terms of methylation percentages, yet patterns vary between individuals and it is not possible to point to a specific region of the genome that is consistently more methylated, i.e. similarly to what was observed in the soma, we see high methylation peaks in the F mt genome for both individuals, but they are not shared between the two. We would like to point out that two of the three peaks appearing on the bottom strand of individual # 8 have a very interesting location: as mentioned above, these parts of the genome, where there

is a switch in the coding direction, are suspected origins of replication/transcription (Breton et al. 2009). Finally, no similarity in methylation patterns was observed between the two genomes present in the gonad of the same individual.

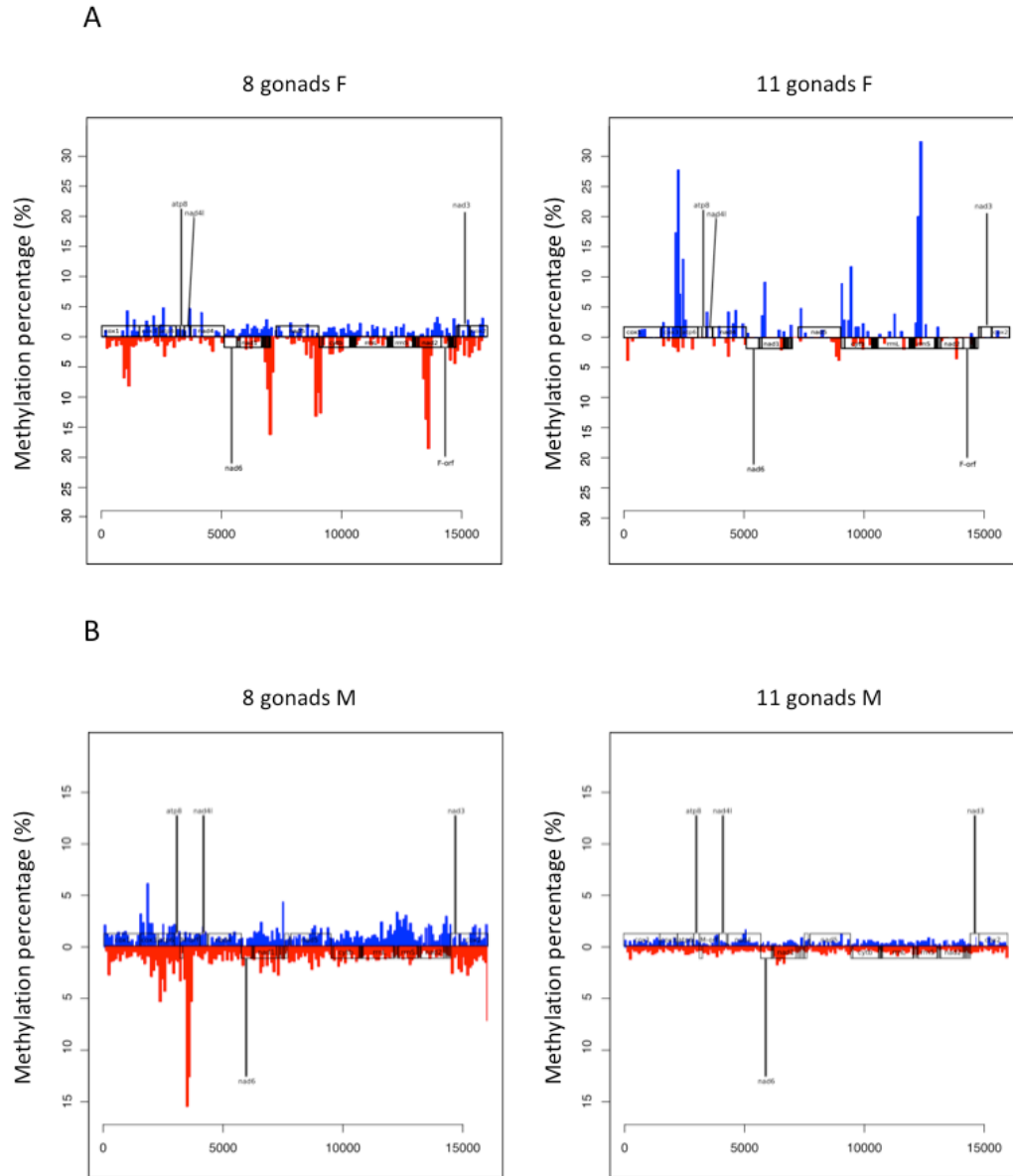


Figure 11: Mean methylation percentage for 100bp windows along the F mt genome found in male gonads of *V.ellipsiformis* for the top (blue) and bottom (red) strands of the F (A) and M (B) genomes.

2.3.4 Methylation in the different cytosine contexts



In order to determine whether methylation could have a role in regulating the expression/the replication of the mt genome, we tried to see if it was randomly distributed along the genome, or if on the contrary, it was directed to cytosines in a certain context in particular (CG, CHG or CHH). We calculated the methylation percentage of cytosines in every of these contexts, for every sample (Supplementary Table 7). We also compared methylation on the two strands since using general percentages. Given that it is still uncertain how methylation acts in the mt genome, it is a fair assumption that it could be directed specifically to one of the strands. It is worth mentioning that although both strands have the same length, the bottom one has more cytosines than the top one (69% vs. 31%) in *V. ellipsiformis*. These proportions seem to vary from one species to another, however. For example, the mt genome of the chiton *C. apiculata* has 52% of its cytosines on the bottom strand and 48% on the top strand.

Our results indicate that, for most samples, there is no apparent difference in the methylation percentage of each strand. It is very different in gonads, however. In fact, they exhibit remarkably high peaks on either strands, but these peaks cause a very high variance as well. It is thus not possible to determine statistically if such high peaks in the gonads represent strand specificity in the methylation process. From what we can see in Figure 11a, it would be surprising if one of the two strands was preferably methylated in the gonads.

2.4 Discussion

2.4.1 General sequencing and mapping information

Tens to hundreds of mitochondria can be found in most (bivalve) cells and each of these mitochondria generally holds 2-8 copies of mtDNA, at least in humans (Legros et al., 2004; Ghiselli et al., 2011); for this reason, there is usually no need for a great sequencing depth in order to ensure a sufficient coverage of the mtDNA (Rasmussen and Noor, 2009). In fact, we managed to reach a coverage of more than 700X in one of the samples for the F mt genome (see Table 1) with a sequencing depth of only 10X. However, DUI species are unique in the sense that they exhibit two highly divergent mt genomes that can be both present and expressed in male and female somatic and gonadal tissues (Zouros, 2013; Breton et al., 2017). This situation was recently observed in the freshwater mussels *Venustaconcha ellipsiformis* and *Utterbackia peninsularis* using PCR and RT-PCR (Breton et al., 2017). Although this method proved to be a good way of localizing the M mt genome, it did not provide a quantitative insight on its abundance in each tissue. In the marine mussels *Mytilus* spp., the M mtDNA is usually present at extremely low concentration in male and female somatic tissues whereas in the marine clam *Ruditapes philippinarum*, male somatic tissues can contain more M mtDNA than the F mtDNA and heteroplasmic female somatic tissues appear to be very rare (Ghiselli et al., 2011; Zouros, 2013). Our results for the freshwater mussel species *V. ellipsiformis* suggest that female and male somatic tissues have low M mtDNA levels. For example, while a mean coverage of >400X was obtained for the F mt genome in male somatic tissues, the average was only of 2.3X for the M mtDNA, which was not fully covered in most tissues (below 50%). Due to this low coverage, detailed methylation analyses could not be made for the M mtDNA in somatic tissues. That said, if methylation does indeed regulate mtDNA copy number, with high

methylation levels of the replication control region(s) leading to decreased mtDNA copy number (e.g. Gao et al., 2015; Tong et al., 2017), we could expect higher methylation of the control region(s) of the M mtDNA compared to the F mtDNA in somatic tissues (see Breton et al., 2009 for a detailed analysis of mtDNA control regions in *V. ellipsiformis*). Further experiments with a greater sequencing depth or using methylated DNA capture-based methods like MeDIP (methylated DNA immunoprecipitation) will be needed to test this hypothesis.

2.4.2 Mean methylation levels in F (soma and gonad) and M (gonad) mitochondrial genomes

Our whole mt genome bisulfite sequencing (WMGBS) analysis revealed that methylation occurs on both strands, in all regions (intergenic and within genes), and is applied to all cytosines, regardless of their context (CpG, CpA, CpC, CpT). These results are in line with previous studies (Shock et al., 2011; Bellizzi et al., 2013; Ghosh et al., 2014; Blanch et al., 2016; Bianchessi et al., 2016). It is still uncertain, however, whether there is a strand that is more important than the other for methylation in the mtDNA.

Our results also showed that mean methylation levels oscillate around 1% for every gene. These results are also in line with previous WMGBS studies (Liu et al., 2016; Mechta et al., 2017; Matsuda et al., 2018; Owa et al., 2018), which reported mtDNA methylation levels ranging from ~0.5% to 2%. These low methylation levels have led some authors to question the existence and functional significance of mtDNA methylation (e.g. Mechta et al., 2017; Matsuda et al., 2018). However, strong evidence such as the discovery of a mitochondria-targeted DNA methyltransferase in human and the correlations of methylation levels with diseases such as cancer and diabetes convincingly supports the existence of mtDNA methylation (see Mposhi et

al., 2017 for a review). In the present study, a certain variability in cytosine methylation percentages has been observed, with values ranging from 0 to 10% at each site (Supplementary File 1), which is probably related to the heterogeneity of the sample that is composed of a mixture of tissue cells and/or of several copies of the mtDNA per mitochondrion/cell that may have heterogeneous methylation profiles. Other studies also observed such heterogeneity (e.g. Bartelli et al., 2018). As suggested by Mposhi et al. (2017), this may point to a possible physiological threshold above which methylation has a noticeable effect on mitochondrial function and/or gene expression. For example, as for mtDNA mutations, single mtDNA methylated cytosines at a frequency of 5-10% could have an influence on the mitochondrial phenotype (Wallace 2016; Picard et al., 2016). As discussed below, we were not able to identify a clear pattern associated to an individual, a tissue type, or to one sex. In fact, it seems that there are only few inter-sex differences within a tissue type (the adductor muscle), and that the greatest differences are observed between F and M mitochondrial genomes functioning in the same tissue type (the male gonad) (Fig. 7-11). These results are interesting considering that they could point to a functional relevance of mtDNA methylation in species with ~~doubly~~ uniparental inheritance of mtDNA. They should be taken lightly, however, as some biases could also explain this differential methylation in the gonads.

2.4.3 Effect of the tissue, sex and mt genome on methylation patterns

Although overall F mtDNA methylation levels were low in all tissues, our results revealed a particularly low level in the foot with respect to other tissues. This variation could be explained by the fact that the foot sample and another gonad sample from the same individual were sent for sequencing at a different time than the rest of the samples. The gonad sample also presented

lower methylation levels globally, but other tissues from the same animal were comparable to the rest of the samples. Thus, there is more likely a potential bias related to sequencing rather than an intra-individual pattern. When analyzing differentially methylated regions among tissues, we could detect peaks that were not necessarily shared by all tissues, supporting the notion of tissue specificity of mtDNA methylation as reported in previous studies (Bellizzi et al., 2013; Devall et al., 2017). The observation that inter-sex differences within a tissue type seemed lower than intra-sex differences across tissue types further reinforce the notion that methylation is likely tissue specific and associated with different cellular adaptations/functions in the mussel body.

The most impressive results of this study are those involving the F and M mtDNAs in male gonads. Overall, the distribution of methylation had the following features: (1) for the M mtDNA, one important methylation peak reaching more than 15% was observed on the template strand in the *m-orf* gene of one sample whereas the methylation level in other genes and intergenic regions was relatively similar (around 1%) (Figure 9 and Supplementary Table 6), and (2) for the F mtDNA three distinctive peaks were observed for the bottom strand of one sample: in the two regions flanking *nad5* and in the *nad2* genes, respectively (Figure 11). For the other sample, no peaks were found on the bottom strand. It is possible to see that the regions flanking *nad5* are also slightly more methylated, but the most noticeable are the two tall peaks located on the top strand (in *cox3/atp6* and in the *rrnS*).

To date, two possible functions have been proposed for mtDNA methylation: the regulation of mtDNA replication and expression (see Mposhi et al., 2017; van der Wijst et al., 2017; Bartelli et al., 2018 for reviews of the molecular mechanisms possibly involved). As mentioned above, previous studies have shown that mtDNA hypomethylation of the mitochondrial D-loop in human cells is associated with increased mtDNA copy number (e.g.

Gao et al., 2015; Tong et al., 2017; van der Wijst et al., 2017). Although the cellular consequences of this increase in mtDNA copy number are still poorly known, some studies suggest that it could influence mitochondrial gene expression (e.g. Gao et al., 2015; Mposhi et al., 2017). For example, Gao et al. (2015) observed that de-methylation of the D-loop region was associated with an elevated mtDNA copy number and an increased *nad2* expression in colorectal cancer. It is also possible that D-loop methylation influences mitochondrial gene expression not only indirectly by modulating mtDNA copy number, but also directly via the regulation of the mitochondrial transcription factor A (TFAM) binding to promoter region (van der Wijst and Rots, 2015; Mposhi et al., 2017).

A peculiar aspect of our data is that two highly methylated regions in the F mtDNA of the male gonad corresponded to the two intergenic regions flanking the *nad5* gene, suspected to control the replication and/or expression of the mt genome in *V. ellipsiformis* (Breton et al., 2009). These two peaks were not observed in the M mtDNA, which instead showed an important methylation peak in the male genome-specific *m-orf* gene, suggesting that regulation of methylation could be genome-specific in species with DUI, with intergenic regions possibly more targeted in the F mtDNA for regulating whole mt genome expression/replication whereas methylation of gene bodies in the M mtDNA would be the main mechanism for regulating mtDNA expression. It is also possible that methylation patterns differ during development, or with different physiological or environmental conditions (e.g a reverse situation could be observed in juvenile mussels compared to adults with methylation in gene bodies in the F mtDNA instead of in the M). Alternatively, our results could reflect different distribution of DNA binding proteins on the F vs. M mtDNA, such as the TFAM, with some regions being more or less accessible to methyltransferases (Rebelo et al., 2009; Bartelli et al., 2018).

In both F and M mtDNAs, methylation occurred in CpG and non-CpG contexts in gene bodies. At the nuclear level, methylation in gene bodies has also been reported and is positively correlated with transcription (Jones, 2012). At the mitochondrial level, the situation seems controversial. For example, one study reported that an increase in the methylation of the mtDNA-encoded *nad6* gene was correlated with a decrease in its mRNA abundance and protein expression (Pirola et al., 2012). By contrast, Bartelli et al. (2018) suggested that the observed decrease in the overall mtDNA methylation profile of a *C. albicans* strain was possibly associated with the transcriptional profile of this strain in adaptation to hypoxia, which is characterized by a decreased expression of the genes involved in mitochondrial oxidative phosphorylation that is oxygen dependent. The same authors (and others) also suggested that higher levels of methylation in mtDNA may reflect a less compacted mtDNA that is more accessible to methyltransferases, and also to the transcriptional machinery, thus leading to increased gene activation/expression (Bartelli et al., 2018; see also van der Wijst and Rots, 2015).

Because methylation seems to occur equally on CpG and non-CpG cytosines in the mtDNA, it is possible that the methyltransferases targeting mitochondria follow a slightly different procedure than those found in the nucleus. Also, the molecular mechanisms recognizing methylated cytosines might not require both strands to be methylated, as it is the case for CpG islands in the nucleus. For example, Dnmt1 is more likely to apply methylation on cytosines at hemimethylated CpG sites (Viltakis et al., 2005). If it happened to be absent from the mitochondria or if the isoform sent towards the mitochondria varied from the ones in the nucleus, this could lead to having one strand more methylated than the other. Moreover, it has previously been proposed that methylation on the template strand would be the most likely to bear biological significance for controlling mitochondrial expression (Stöger et al., 2011).

It is not clear yet how mtDNA methylation in gene bodies may be linked to mitochondrial gene expression given the polycistronic nature of mtDNA transcription (Clayton 2000). For instance, in the nuclear DNA, it has been reported that methylation in gene bodies (within exons) may influence gene expression by affecting gene splicing (Maor et al., 2015). In most animal mtDNAs, which are devoid of introns, it could be possible that the polycistronic mitochondrial mRNA precursor undergoes unknown post-transcriptional modifications to influence gene expression (Mposhi et al. 2017).

Our results also support the notion of genome-specificity for the regulation of mitochondrial methylation and expression in species with DUI. For example, one recent study using mitochondria-specific methylation-induction experiments showed that mtDNA gene expression remained unchanged upon induction of CpG mtDNA methylation in mitochondrial promoters, whereas induction of C-methylation in the GpC context decreased mtDNA gene expression (van der Wijst et al., 2017). Mitochondrial promoters have not been clearly identified in *V. ellipsiformis* (Breton et al., 2009), but it is tempting to speculate that different context of cytosine methylation (CpG vs GpC) could affect the F and M mtDNAs in male gonads. Although our observations point towards functional relevance of cytosine methylation in bivalves with DUI, future studies will be needed to confirm if it is indeed the case.

2.5 Conclusion

This preliminary study suggests functional biological relevance of mtDNA methylation in a bivalve species with doubly uniparental inheritance of mtDNA. This is based on the observation that methylation levels in the male gonads for both genomes are drastically higher than in other tissues. Indeed, no difference was found in the methylation of distinct somatic tissues from individuals of the same sex. There was perhaps even less variation in the methylation patterns

of different sexes for a specific tissue type (the adductor muscle). However, not only do the F and M mitochondrial genomes found in the male gonads appeared to be more methylated than in other tissues, the regions of these genomes with the highest methylation peaks have a potential biological significance. In fact, these could be strategic regions allowing for a regulation of the expression/replication of the whole genome in one case (the F mtDNA^{AA}) and the expression of a genome-specific gene in the other case (*m-orf* in the M mtDNA).. We hypothesized that the regulation of mitochondrial methylation and expression in species with DUI could be genome-specific, although more studies will be needed to determine the exact process of mt gene regulation in these species. MtDNA methylation could be an important mechanism involved in regulating the DUI system.

Author Contributions

KB, DG and SB contributed to data analysis. KB, DG, BA, and SB contributed to the study design. KB and DG contributed to data acquisition. KB, DG, BA, and SB contributed to data interpretation and manuscript drafting and approved the final version of the manuscript.

Funding

KB received a scholarship from the Groupe de Recherche Interuniversitaire en Limnologie et Environnement Aquatique (GRIL), Hydro-Québec and Banque Nationale. This work was supported by the National Sciences and Engineering Research Council (NSERC) (grant no RGPIN/435656-2013 to SB and 238600-2013 to BA).

Acknowledgments

We would like to thank Bernard Sietman for providing *Venustaconcha ellipsiformis* samples.

3 Discussion


L'objectif principal de ce projet de maîtrise était de déterminer si les organismes à DUI arboraient de la méthylation au niveau de l'ADN mitochondrial. Étant donné que ces espèces possèdent deux génomes mitochondriaux extrêmement divergents, parfois présents dans un même tissu chez les mâles et les femelles, mais pas nécessairement tous les deux exprimés, nous avons émis l'hypothèse que ces ADNmt pourraient être régulés via la méthylation. Nous avons donc exploré la méthylation des ADNmt F et M entre différents tissus d'un même sexe, entre les sexes, et à l'intérieur d'un même tissu. Nous avons également vérifié si la méthylation était faite de façon égale sur les deux brins du génome et dans les différents contextes dans lesquels les cyosines peuvent se trouver.

3.1 Information relative au séquençage

Du fait que la technologie de séquençage à haut débit (*HiSeq*) devient plus accessible, il peut être intéressant de privilégier ce type de méthode pour séquencer le génome mitochondrial complet, par rapport aux méthodes traditionnelles qui demandent plus de temps et de manipulations expérimentales en laboratoire (ex. *long range PCR*). Lorsqu'on utilise le séquençage à haut débit, il est important de prendre en considération la profondeur de séquençage : celle-ci donne une approximation de la couverture de chaque base séquencée (Sims et al., 2014). Or, puisqu'il y a généralement de multiples copies du génome mitochondrial dans chaque cellule, il n'est pas nécessaire d'avoir une grande profondeur de séquençage afin d'obtenir une couverture satisfaisante de ce génome ; ce n'est pas le cas du génome nucléaire qui est généralement présent en seulement une ou deux copies par cellule. Dans cette étude, nous avons eu recours à une profondeur de séquençage de 10X, ce qui est considéré comme relativement faible mais qui nous a permis d'obtenir une couverture du génome mitochondrial F de 500X en moyenne dans les tissus somatiques, et avec un maximum de plus de 700X pour un

des échantillons. Toutefois, à cause de la distribution très inégale des deux génomes mitochondriaux dans les organismes à DUI, il n'a pas été possible de faire pareil pour le deuxième génome mitochondrial (M). En effet, la faible couverture pour ce génome dans les tissus somatiques a fait en sorte que nous n'avons pas pu analyser ces données : il manquait de l'information pour certaines parties du génome dans la plupart des tissus, et il n'y avait qu'une partie des échantillons qui présentaient des séquences couvrant les autres parties du génome. Il n'était donc pas possible de faire des moyennes et les pourcentages de méthylation étaient fortement biaisés.

Pour remédier à ce problème, il serait possible de réutiliser le même protocole en changeant toutefois la profondeur de séquençage. D'autres méthodes pourraient également nous renseigner sur la présence de méthylation dans le génome M, sans pour autant avoir le même degré de précision que le séquençage-bisulfite (voir Kurdyukov et Bullock, 2016, pour un résumé des méthodes). Plusieurs d'entre elles impliquent par exemple des enzymes de restriction et permettent de cibler certaines régions du génome plus susceptibles d'être méthylées ou d'avoir un rôle dans l'expression/la réplication du génome. Ces régions pourraient alors être multipliées par PCR ou qPCR, et ensuite séquencées ou analysées sur gel pour un aperçu général de l'état de méthylation. Il faut être prudent toutefois, car plusieurs de ces méthodes sont conçues pour l'analyse des cytosines dans un contexte CpG et d'après nos résultats, ce n'est pas le seul type de méthylation présent dans l'ADNmt. L'une des méthodes qui ne pose pas ce problème et qui serait appropriée dans un cas comme celui-ci est l'immunoprécipitation combinée au séquençage (MeDIP-seq). Cette méthode évite effectivement les biais associés aux enzymes de restriction en faisant plutôt appel à la sonication pour couper l'ADN en fragments (Taiwo et al., 2012).

Alternativement, il pourrait aussi être intéressant d'essayer de déterminer les patrons de méthylation du génome M dans les tissus somatiques de d'autres espèces. En effet, il existe plusieurs autres espèces chez qui le génome M n'est pas limité aux gonades (Breton et al., 2017; Dalziel et Stewart, 2002; Garrido-Ramos et al., 1998 ; Ghiselli et al., 2011). De plus, les quantités varient fortement dans les tissus somatiques en fonction de l'espèce et d'un sexe à l'autre chez une même espèce. Pour plusieurs espèces du genre *Mytilus*, par exemple, le génome M est présent en très faible quantité dans les tissus somatiques mâles et femelles (Zouros, 2013) (c'est un peu la même situation que nous observons ici chez *V. ellipsiformis*). Cependant, dans les tissus somatiques mâles de *R. philippinarum*, le génome M est présent en quantité parfois plus importante que le F, alors que c'est l'inverse dans les tissus somatiques femelles (Ghiselli et al., 2011). Il semble également  que le génome M soit transcrit à des niveaux différents en fonction du tissu chez cette espèce. Il serait donc intéressant d'étudier les patrons de méthylation du génome M dans les tissus somatiques mâles de *R. philippinarum* à l'aide du séquençage-bisulfite, puis de tenter de corrélérer ces patrons à l'expression des gènes.

Il est à noter que le traitement bisulfite ne semble pas affecter l'efficacité du séquençage : en observant la figure 6, on constate que le nombre de lectures (*reads*) obtenues pour le séquençage shotgun est tout à fait comparable au nombre obtenu pour le séquençage bisulfite. Afin d'éviter un biais, nous n'avons pas utilisé toutes les lectures pour effectuer cette comparaison, mais seulement celles qui passaient le contrôle de qualité après une première étape de « *trimming* ». Le traitement bisulfite ne semble donc pas dégrader l'ADN pour les différents échantillons d'invertébrés testés.

3.2 Méthylation de F (soma et gonades mâles) et M (gonades mâles)

L'une des choses qui ressort clairement de nos analyses est le taux de méthylation, qui concorde avec les résultats de plusieurs études (utilisant le traitement bisulfite) quant au taux de méthylation dans l'ADNmt chez l'humain (0-2%). Ces faibles taux ont poussé certains auteurs à proposer que la méthylation de l'ADNmt serait inexistante : Owa et al. (2018) ainsi que Matsuda et al. (2018) expliquent que des taux de 0-2% seraient probablement dus aux limites de la détection par séquençage-bisulfite, tandis que Mechta et al., (2017) parlent de «bruit de fond ». Ces derniers ont même observé une corrélation entre la couverture de séquençage et le taux de méthylation. Cela pourrait peut-être expliquer nos résultats pour le génome F dans la gonade mâle, qui présente un taux de méthylation relativement élevé (jusqu'à 15%) et dont la couverture est beaucoup plus faible dans ce tissu par rapport aux tissus somatiques. Ceci dit, même si quelques chercheurs prétendent que la méthylation de l'ADNmt est probablement inexistante et « non-fonctionnelle », beaucoup d'études suggèrent le contraire (Liu et al., 2016; Ghosh et al., 2014, 2016; Devall et al., 2017; Wong et al., 2013). Par exemple, Liu et al., (2016), relèvent une différence en ce qui concerne les profils de méthylation entre les échantillons de sang et de salive analysés. Devall et al. (2017) ont également observé que les patrons de méthylation de l'ADNmt étaient associés au type de tissu : ils ont pu regrouper les échantillons de sang, de cortex et de cervelet en se basant seulement sur les régions différemment méthylées du génome mitochondrial. Wong et al. (2013), quant à eux, ont trouvé que les méthyltransférases Dnmt3a étaient présentes surtout dans certains types de tissus (muscle squelettique, système nerveux central et testicules) et presque absents dans d'autres. Ils ont lié la présence de Dnmt3a à de la méthylation dans ces tissus et ils ont même observé une méthylation ainsi qu'une quantité de Dnmt3a anormales dans ces mêmes tissus pour des

individus atteints de sclérose latérale amyotrophique (SLA). La méthylation de l'ADNmt pourrait donc avoir un rôle spécifique associé au tissu, ou encore être liée à l'âge ou au sexe comme le proposent D'aquila et al., (2015). Par ailleurs, une méthylation de l'ADNmt anormale semble être liée à diverses maladies (maladies cardiaques, diabète, cancer, etc.; voir Mposhi et al. (2017) pour une revue des articles publiés). Si l'on ajoute à cela les découvertes de Shock et al. (2011) et Chestnut et al., (2011) concernant la translocation des méthyltransférases spécifiquement vers les mitochondries, on doit envisager que non seulement la méthylation de l'ADNmt pourrait exister, mais qu'elle a possiblement un rôle fonctionnel.

Nos résultats sont mitigés : d'une part, nous n'observons pas de patrons clairs associés à chaque type de tissu, à chacun des génomes ou à chacun des sexes. D'autre part, nous constatons la présence de taux élevés de méthylation dans un tissu en particulier : les gonades mâles. Or, c'est une région où il serait attendu qu'une régulation plus importante de chacun des deux génomes soit nécessaire au fonctionnement du système DUI chez cette espèce. On pourrait donc aisément lier la méthylation de l'ADNmt à un rôle fonctionnel dans ce cas. Il faut considérer que des variations importantes dans le patron de méthylation (du génome F dans les gonades mâles, par exemple) pourraient être expliquées par le fait que les tissus échantillonnés sont hétérogènes en ce sens qu'ils sont composés de différents types cellulaires pouvant avoir des fonctions très distinctes. Il est donc possible que seule une partie des cellules dans les gonades mâles portent le génome F ou bien que ce dernier soit utile (et exprimé) dans seulement une partie des cellules (voir Obata et al., 2011, figure 7).

Nous proposons aussi que la méthylation pourrait servir de mécanisme permettant de contrôler la réplication ou l'expression des copies mutées de l'ADNmt et d'éviter les maladies



citées en exemple plus haut. En effet, si ~~ces maladies~~ apparaissent seulement à partir d'un certain seuil (un pourcentage des copies de l'ADNmt qui portent des mutations), il est nécessaire pour la cellule de faire en sorte que ce seuil physiologique ne soit pas franchi (Latorre-Pellicer et al., 2016; Picard et al., 2016; Elliott et al., 2008). La méthylation de l'ADNmt ciblée pour certaines parties du génome pourrait 1) empêcher la transcription de certains gènes en particulier, 2) empêcher la transcription du génome complet par la méthylation d'un ou plusieurs sites à travers le génome, 3) empêcher non seulement la transcription mais aussi la réplication de la molécule mutée. Les copies méthylées pourraient être « réparées » puis déméthylées, ou tout simplement détruites. Dans tous les cas, la proportion des copies d'ADNmt méthylées ne devrait pas être élevée. Ce type de mécanisme pourrait donc expliquer le faible pourcentage de méthylation que nous avons observé et qui revient constamment dans la littérature. De plus, il est possible que ce seuil soit spécifique au tissu, puisque chacun est associé à des fonctions énergétiques, entre autres (Iacobazzi et al., 2013).

3.3 Effet du sexe, du tissu et du génome sur les patrons de méthylation

Bien que la méthylation soit retrouvée partout à travers le génome, nous n'avons pas été en mesure d'identifier un patron spécifique à chaque type de tissu ou au sexe. Les patrons de méthylation semblent être aussi différents pour un même type de tissu, entre individus de sexes différents, qu'entre différents tissus d'individus du même sexe. Dans certains cas, on observe quand même des pics de méthylation qui reviennent d'un échantillon à l'autre : dans le muscle adducteur, par exemple, on observe un pic dans la région précédant *nad5* aussi bien chez les individus mâles et femelles. Cependant, on n'observe pas exactement les mêmes patrons entre le muscle adducteur et les autres tissus somatiques mâles. La seule grande différence que l'on

observe entre les tissus est le taux beaucoup plus élevé dans la gonade mâle, et ce pour les deux génomes. Des études récentes comme celles de, Wong et al. (2013), Liu et al. (2016), ou encore Devall et al. (2017), ont aussi comparé les patrons de méthylation de l'ADNmt entre différents tissus, voire même différents types cellulaires de mêmes individus. Ces études réalisées chez l'humain ou chez d'autres mammifères ont pour conclusion que les patrons de méthylation inter-individus et intra-tissus sont plus similaires que ceux inter-tissus et intra-individu. En plus des exemples décrits plus haut, on peut citer Bellizi et al. (2013), qui ont comparé des échantillons de sang et de cellules de culture, provenant à la fois d'humains et de souris. Pour les deux modèles, les résultats étaient similaires : il y avait une spécificité du patron qui était associée au tissu. Effectivement, les cellules sanguines et HeLa étaient plus méthylées que les fibroblastes et les ostéosarcomes pour l'humain; et la même situation a été observée chez la souris. De façon surprenante, même l'ADNmt des plantes est méthylé de façon spécifique à l'espèce, au tissu, et même à l'âge (Vanyushin, 2006).

Étant donné qu'une des fonctions principales des mitochondries est la production d'énergie, on peut supposer que la méthylation de l'ADNmt pourrait agir comme un contrôle indirect de la production d'énergie en affectant le nombre de transcrits des gènes mitochondriaux. Aussi, comme on sait que les différents tissus/types cellulaires ont des besoins en énergie différents (Fernández-Vizarra et al., 2011; Betinazzi et al., en préparation), il est raisonnable de penser que la méthylation différentielle et spécifique au tissu peut être liée à la production d'énergie. Nous pouvons supposer que les gonades, qui doivent produire des spermatozoïdes, ont des besoins en énergie différents des autres tissus, et que ce besoin en énergie peut varier en fonction de l'année, de la maturité sexuelle de l'individu, etc. Cela pourrait expliquer pourquoi la gonade semble être le tissu le plus différemment méthylé. Dans

le même ordre d'idées, Wong et al. (2013) avancent que les tissus où ils ont retrouvé la *dnmt3a* associée à la méthylation étaient des tissus « excitables ». Reprenons l'idée de seuil physiologique où la méthylation de l'ADNmt a pour rôle d'empêcher les copies mutées de se multiplier ou d'être exprimées. On peut alors imaginer que les tissus où l'activité métabolique est plus élevée seraient plus susceptibles de produire des ROS endommageant l'ADNmt (Newsholme et al., 2007), ce qui entraînerait une augmentation du nombre de copies méthylées dans ces tissus.

Une chose intéressante à constater dans les figures 10 et 11 est que le pic que l'on retrouve pour le génome F dans le muscle adducteur (mâle et femelle) avant *nad5*, est également présent dans le génome F des gonades mâles. Malgré la grande variation entre les tissus en termes de patrons de méthylation, c'est un pic qui revient et qui pourrait avoir une signification biologique importante. En effet, cette région exempte de gènes avait été suggérée par Breton et al. (2009), comme étant une origine de réplication ou une région contrôle de la transcription pour le génome F chez les moules d'eau douce. Dans la gonade mâle, on observe également un autre pic à l'autre extrémité de *nad5*, là où il y a également un changement dans l'orientation des gènes (le brin codant change). Là encore, c'est une région potentiellement intéressante d'un point de vue biologique puisque la méthylation pourrait se faire préférentiellement aux origines de transcription du brin lourd (H) et du brin léger (L), si elles sont bel et bien situés à ces endroits chez cette espèce et pour ce génome. Jusqu'à présent, il n'est pas clair si la méthylation de l'ADNmt agirait plutôt en empêchant l'expression ou en bloquant la réplication du génome (Mposhi et al., 2017; van der Wijst et al., 2017; Bartelli et al., 2018). Pourtant, plusieurs études établissent une corrélation entre la méthylation, notamment dans la région D-loop, et le nombre de copies du génome mitochondrial (Gao et al., 2015; Tong

et al., 2017; van der Wijst et al., 2017). Par exemple, Gao et al. (2015) ont remarqué que le nombre de copies d'ADNmt était plus élevé dans les tissus cancéreux, mais aussi que ce nombre de copies plus élevé allait de pair avec une dé-méthylation de la région D-loop, ainsi qu'une expression plus prononcée de *nad2*. De plus, Tong et al. (2015) ont constaté que la dé-méthylation de seulement quelques cytosines présentes dans la D-loop était suffisante pour voir une augmentation du nombre de copies d'ADNmt. Ceci suppose donc qu'une action très locale de la méthylation pourrait avoir un grand impact sur les processus métaboliques de la cellule. Cette régulation à partir de seulement quelques cytosines méthylées dans la région de contrôle pourrait faire intervenir le « Facteur de Transcription Mitochondrial A » (TFAM) selon van der Wijst et Rots (2015); Mposhi et al. (2017). En effet, TFAM est absolument essentiel à l'initiation de la transcription mitochondriale chez l'humain (Shi et al., 2012) et la méthylation de l'ADNmt pourrait empêcher l'interaction entre ce dernier et TFAM (van der Wijst et Rots, 2015). Les TFAM sont des molécules qui peuvent être modifiées post-traduction, un peu comme les histones, et la régulation épigénétique pourrait donc être beaucoup plus complexe que prévu, en faisant intervenir au moins deux mécanismes coordonnés par le noyau. La méthylation pourrait faire en sorte que l'ADNmt devienne plus dense, ou bien des modifications apportées aux TFAM pourraient faire en sorte que celle-ci ne se lie plus correctement à l'ADNmt. Il y aurait ensuite des répercussions sur l'activité métabolique et le noyau devrait donc ajuster cette activité par un mécanisme ou un autre. Les résultats que nous avons obtenus pourraient donc s'expliquer par la présence de TFAM rendant plus ou moins accessibles certaines parties du génome aux méthyltransférases (Rebelo et al., 2009; Bartelli et al., 2018).

Pour les raisons que nous avons données précédemment, il n'a pas été possible d'aller en profondeur de la même façon pour notre analyse du génome M dans les tissus somatiques. Nous

retrouvons tout de même dans la gonade mâle un pic de presque 8% à l'intérieur du gène *m-orf*. Ce dernier est unique au génome M et sa fonction n'est pas encore totalement définie. Toutefois, Milani et al. (2013), Mitchell et al. (2016) ont proposé des fonctions possibles pour les protéines codées à partir des gènes additionnels uniques à ces espèces (*m-orf* et *f-orf*). La protéine associée à *m-orf* est très probablement liée à la détermination du sexe ou aux spermatozoïdes puisque ceux-ci transportent exclusivement le génome M et que ce gène codant pour une protéine est l'une des différences majeures entre les deux génomes. En effet, l'une des hypothèses avancées par ces auteurs est qu'elle servirait potentiellement à l'agrégation des mitochondries portant le génome M durant les premiers stades de vie (chez les embryons) ; ces protéines interagissent peut-être avec le cytosquelette dans le but de déplacer les mitochondries pour les regrouper ou les agréger (Milani et al., 2013; Mitchell et al., 2016). Les fonctions spécifiques associées à chacun des génomes mitochondriaux pourraient fournir une autre explication pour la méthylation de l'ADNmt spécifique au tissu. En effet, un contrôle de l'expression des gènes mitochondriaux pourrait mener à l'expression différentielle de ces protéines en fonction du tissu. La méthylation de l'ADNmt pourrait donc, par l'intermédiaire des protéines, être un mécanisme permettant la transmission exclusive des ADNmt de type M ou encore la ségrégation de celui-ci dans ce qui deviendra la gonade mâle.

3.4 Méthylation dans des contextes de cytosines différents

Un autre aspect important de nos résultats est que la méthylation se fait généralement de façon plus ou moins uniforme à travers le génome, pas seulement dans un contexte de dinucléotides CpG, et que la méthylation se fait sur les deux brins du génome, même s'il n'est pas tout à fait clair si des effets biologiques pourraient être associés à chacun d'entre eux. En effet, il est possible que la méthylation soit présente sur les deux brins, mais qu'elle

empêche/augmente la transcription d'un gène seulement lorsqu'elle est sur le brin matrice ou à l'inverse, seulement sur le brin codant. Étant donné que dans le génome mt, les deux brins codent pour des gènes, les notions de « brin codant » et de « brin matrice » sont relatives aux gènes. Stöger et al. (2011) suggèrent que le brin matrice aurait plus d'importance pour contrôler l'expression du génome mitochondrial. Pour leur part, Guo et al. (2014) ont étudié la distribution de la méthylation dans l'ADN nucléaire de cellules pluripotentes chez l'humain. Ils ont déterminé que celle-ci se faisait toujours de façon symétrique lorsqu'il s'agissait de CpG, tandis qu'elle était asymétrique et spécifique au brin lorsque les cytosines étaient dans un contexte non-CpG (dans des introns et des éléments répétés). D'après nos résultats, il ne semble pas y avoir de préférence ni pour un contexte en particulier, ni pour l'un des deux brins de l'ADNmt, et ce chez plusieurs espèces de mollusques/invertébrés. La proportion de cytosines varie beaucoup entre les deux brins chez certaines espèces (ex : 69% vs 31% chez *V. ellipsiformis*), mais les taux de méthylation par cytosine sont généralement similaires sur chacun d'eux.

Guo et al. (2014) ne sont pas les seuls à s'être intéressés à la méthylation non-CpG; en fait, plusieurs études dans l'ADN nucléaire des mammifères ont rapporté la présence de méthylation fonctionnelle des cytosines dans des contextes « non-CpG », même si elle est relativement rare (Barrès et al., 2009 ; Pinney, 2014; Sharma et al., 2016). Tomizawa et al. (2011) ont même observé de la méthylation non-CpG dans les oocytes, sans pouvoir en détecter dans les spermatozoïdes. Ces résultats supposent que la méthylation de l'ADN qui ne cible pas particulièrement les CpG pourrait donc agir dans certaines situations précises et être contrôlée en fonction du tissu et/ou du sexe. En ce qui concerne la méthylation non-CpG de l'ADNmt, quelques chercheurs mentionnent en avoir détecté (Bellizi et al., 2013; van der Wijst et al.,

2017), ce qui appuie nos résultats. Ceci est assez intéressant considérant la quantité restreinte d'études portant sur la méthylation de l'ADNmt. Par exemple, van der Wijst et al. (2017) proposent que ce serait la méthylation des GpC qui serait plus importante en ce qui concerne la régulation de l'ADNmt. Bellizi et al. (2013), quant à eux, se sont concentrés sur la région de contrôle chez l'humain (D-loop). Ils ont même détecté la présence de différents types de méthyltransférases (Dnmt1, Dnmt3a et Dnmt3b) dans les mitochondries. Il faut savoir que la Dnmt1 sert plutôt à apposer un groupement méthyle là où il y a déjà une cytosine méthylée sur le brin complémentaire (lorsque le dinucléotide CpG est hémiméthylé) tandis que les Dnmt3a et b sont plutôt associées à la méthylation *de novo*. En inactivant ces trois méthyltransférases, Bellizi et al. (2013) n'ont constaté une réduction que dans la méthylation des CpG et non pas dans celle des non-CpG. Parallèlement à cette étude, Tiedemann et al. (2014) ont créé des *knock down* pour Dnmt3L, puis ont réalisé que la cible principale de méthylation de Dnmt3a et Dnmt3b passait alors de CpG à non-CpG. Tout ceci soulève plusieurs questions : est-il possible que d'autres types de méthyltransférases soient présents dans les mitochondries? Y a-t-il une combinaison de méthyltransférases qui agissent de façon complémentaire pour effectuer la méthylation dans l'ADNmt? Aussi, est-il possible que cette combinaison soit différente d'un tissu à l'autre, par exemple? En effet, Wong et al. (2013) ont observé que la Dnmt3a avait une distribution qui variait en fonction du tissu, tandis que Saini et al. (2017) décrivent la présence d'une méthyltransférase à l'intérieur des mitochondries qui varie à travers le temps chez des cellules en culture. De plus, la présence de cette méthyltransférase est directement liée à la méthylation de l'ADNmt. Ces auteurs proposent donc qu'il y aurait une auto-régulation de l'expression de l'ADNmt via les méthyltransférases, en acheminant dans ce cas-ci l'isoforme 3 de la Dnmt1 vers l'intérieur des mitochondries pour répondre à des pressions environnementales (Saini et al., 2017, figure 7). Le fait qu'il existe différents types de

méthyltransférases et différents isoformes de chacune d'entre elles, combiné au fait que la distribution spatiale et temporelle de ces méthyltransférases peut varier en fonction de la situation, pourrait expliquer en partie nos résultats, et tous ceux qui témoignent de patrons de méthylation spécifiques au tissu.

Il se peut donc qu'une ou plusieurs méthyltransférases présentes dans les mitochondries ne fasse pas de discrimination entre les cytosines CpG et non-CpG. Étant donné la nature du génome mitochondrial et l'absence de promoteur pour chaque gène, ce type de méthylation non spécifique aux CpG (généralement regroupés en îlots dans les promoteurs) pourrait être avantageuse dans une configuration comme celle-ci. Puisque ce type de méthylation ne semble pas limité à la région de contrôle, il se pourrait qu'il permette une régulation plus précise (gène par gène), sans forcément agir sur le génome complet.

Il y a bien sûr dans le génome nucléaire de la méthylation intragénique; elle semble même être liée à l'expression des gènes chez les bivalves (Gavery et Roberts, 2013). Cependant, plusieurs suggèrent que cette méthylation serait liée à l'épissage des exons. Or, ces derniers ne sont pas présents dans le génome mitochondrial et il serait donc étonnant que la méthylation intragénique agisse de la même façon dans les mitochondries étant donné la nature polycistronique de leur transcription (Clayton, 2000). Cependant, nous avons trouvé de la méthylation intragénique et d'autres auteurs en ont également rapporté la présence dans l'ADNmt (Mposhi et al., 2017). La méthylation intragénique a même été corrélée à la progression de maladies comme la stéatose hépatique non alcoolique et il n'est pas clair comment elle pourrait affecter l'expression des gènes (réduire ou augmenter l'expression) (Mposhi et al., 2017). Ces auteurs suggèrent que des modifications à l'ARN messager encore

inconnues pour l'instant pourraient survenir après la transcription et que malgré la nature polycistronique de cette transcription, il pourrait y avoir un contrôle de l'expression des gènes causé par la méthylation. Cela ressemblerait donc en quelque sorte à ce qui a été proposé dans l'ADN nucléaire concernant la méthylation qui permettrait l'épissage des exons, à la différence peut-être que ce serait ici certains gènes qui ne seraient pas exprimés au lieu d'exons.

4 Conclusion et perspectives futures

Nos résultats suggèrent que la méthylation est présente dans les ADNmt des bivalves d'eau douce à DUI, mais à des niveaux relativement faibles (quoique comparables à ceux qui ont déjà été observés dans les mitochondries). Ces taux et patrons de méthylation ne semblent pas être associés à l'individu : peu de similarités sont observables entre les tissus pour un même individu... De plus, les taux de méthylation varient peu en fonction du tissu dans lequel le génome se trouve, à l'exception des gonades mâles, où ils sont beaucoup plus élevés que dans les autres tissus. Il est cependant difficile de dire à partir de nos résultats s'il existe véritablement un patron spécifique au type de tissu. En effet, les taux de méthylation dans les gonades mâles sont plus élevés qu'ailleurs, mais les patrons ne sont pas constants d'un individu à l'autre pour un même génome mt. Une chose intéressante, cependant, est que les deux génomes présents dans les gonades mâles ont des pics de méthylation situés à des endroits différents, des endroits notamment stratégiques pour la régulation de la réplication/transcription du génome ou encore pour le fonctionnement du DUI. Ces pics semblent donc indiquer à la fois une régulation globale (dans la présumée région de contrôle) mais aussi possiblement une régulation plus précise en agissant directement au niveau des gènes (par exemple *m-orf*). Un

autre aspect de cette étude était d'étudier la distribution de la méthylation à travers les deux brins du génome mt et à travers les différents contextes dans lesquels se trouvent les cytosines. Encore une fois, il ne semblait pas y avoir de préférence pour l'un ou l'autre des brins, ou encore pour un contexte de cytosine particulier. Une étude approfondie serait cependant nécessaire afin de savoir si la méthylation affecte réellement l'expression et si le brin matrice est le seul impliqué dans cette régulation de l'expression.

Nos résultats soulèvent plusieurs questions concernant la nature de la méthylation dans l'ADNmt, notamment sur la fonction de la méthylation non-CpG, sur le rôle que la méthylation peut avoir dans l'expression ou la réplication du génome, ainsi que sur le rôle différent associé à la méthylation de chaque brin. Il se peut qu'un ou plusieurs mécanismes permettent le contrôle du nombre de copies de l'ADNmt/nombre de mitochondries, et que ces mécanismes contrôlent aussi l'expression différentielle des génomes M et F en fonction du tissu. Par exemple, les mécanismes responsables de cette régulation pourraient cibler directement la mitochondrie ou bien agir au niveau de l'ADN nucléaire et indirectement affecter la réplication mitochondriale. C'est ce que proposent certains auteurs qui ont trouvé des corrélations entre la méthylation de gènes nucléaires et la densité mitochondriale (Barrès et al., 2009) ou encore avec le nombre de copies du génome mt en fonction du tissu (Kelly et al., 2012). Il sera intéressant, dans des études futures, de considérer la méthylation des gènes nucléaires liés aux mitochondries, ainsi que d'autres mécanismes épigénétiques susceptibles d'affecter l'expression des gènes mitochondriaux. En effet, plusieurs voies restent à explorer, mais les avancées de toutes sortes en lien avec le système de DUI pourraient enligner ces futures études. Il est à noter que Renaut et al. (2018) ont récemment séquencé le génome complet de *Venustaconcha ellipsiformis* et ceci pourra faciliter l'analyse du méthylome pour cette espèce. De plus, on pourrait en apprendre

davantage sur les fonctions associées à chaque type mitochondrial de par la caractérisation des protéines issues des gènes mitochondriaux spécifiques (et uniques) à chaque haplotype (Ouimet et al., en préparation). Une fonction spécifique à chacun des génomes pourrait évidemment supporter le fait d'avoir une distribution et une expression aussi particulière à travers les tissus. Finalement, plusieurs biais expérimentaux potentiels sont à considérer et d'autres études faisant intervenir d'autres techniques permettraient de confirmer nos résultats. Par exemple, en ce qui concerne le séquençage-bisulfite, certains auteurs suggèrent que des pics de méthylation seraient plus susceptibles d'être observés dans des régions du génome où il y aurait des structures secondaires formées par l'ADN et qu'ils seraient le résultat d'une conversion inefficace lors du traitement bisulfite (Mechta et al., 2017). Il faudrait vérifier si c'est le cas pour les génomes que nous avons utilisés. D'autres méthodes impliquant des enzymes de restriction censées éviter un biais associé à la structure du génome (Liu et al., 2016) pourraient permettre cette vérification. Même si la découverte de structures secondaires à certains endroits du génome menait à infirmer nos résultats et que les pics de méthylation étaient causés par des faux-positifs, il serait intéressant d'observer pour quelle raison des structures se forment et si elles ont tout de même un rôle lié à l'expression des gènes. Des méthodes comme MeDIP pourraient aussi permettre l'étude du génome complet sans toutefois recourir au traitement bisulfite et pourraient confirmer la méthylation fonctionnelle du génome mitochondrial chez les espèces à DUI.

Il serait évidemment très pertinent de voir s'il existe un lien entre la méthylation de l'ADNmt et son expression. Pour ce faire, il faudrait disposer de données de méthylation comme celles que nous avons utilisées (préféremment avec une bonne profondeur de séquençage) et de données de transcriptomique. Idéalement, il faudrait baser ces expériences sur les mêmes individus et sur les mêmes tissus. Il serait aussi intéressant d'étudier l'expression des

gènes pour un même génome comportant différents patrons de méthylation. On pourrait ainsi comparer différents tissus, ou encore induire soi-même la méthylation de l'ADNmt à l'aide de méthyltransférases dirigées vers les mitochondries (Saini et al., 2017). Il n'est pas encore clair comment les patrons de méthylation changent à travers le temps et si les facteurs environnementaux ont une influence sur ceux-ci. Si c'est le cas, il serait donc intéressant de comparer les taux d'expression à différents stades de vie aux taux de méthylation à ces mêmes stades.

Il ne faut pas non plus rejeter la possibilité que la méthylation de l'ADNmt pourrait avoir plusieurs rôles et qu'elle ne soit pas uniquement associée à l'expression des gènes. C'est le cas chez les procaryotes (Colot et Rossignol, 1999), mais aussi dans le génome nucléaire, où la méthylation sert possiblement à empêcher une multiplication d'éléments répétés qui se seraient insérés dans le génome (ex. Liu et Schmid 1993). Il faut préciser ici que cette répression de la mobilité des éléments répétés passe par la répression de leur transcription. Il a même été observé que certains éléments transposables pouvaient être la cible de méthyltransférases pour la méthylation *de novo* (Arnaud et al., 1999). Ce serait donc une façon de protéger le génome. Bien que les éléments transposables soient généralement absents du génome mitochondrial soumis à une transmission maternelle (uniparentale), ils sont fréquemment présents et actifs dans les génomes mitochondriaux de levures qui ont une transmission biparentale (Bestor, 2003). Selon cet auteur, la méthylation des éléments transposables pourrait avoir comme fonction de protéger les génomes sexuels, tandis que son rôle serait moins important dans les organismes asexués. Étant donné le lien possible entre les génomes mitochondriaux des espèces à DUI et la détermination du sexe, cette protection via la méthylation de l'ADN pourrait également exister dans l'ADNmt de ces espèces.

Bibliographie

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., ... & Guerler, A. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic acids research*, 46(W1), W537-W544.
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.
- Armstrong, D. A., Green, B. B., Blair, B. A., Guerin, D. J., Litzky, J. F., Chavan, N. R., ... & Marsit, C. J. (2016). Maternal smoking during pregnancy is associated with mitochondrial DNA methylation. *Environmental epigenetics*, 2(3), dvw020.
- Arnaud, P., Goubely, C., Péliissier, T., & Deragon, J. M. (2000). SINE retroposons can be used in vivo as nucleation centers for de novo methylation. *Molecular and cellular biology*, 20(10), 3434-3441.
- Bacalini, M. G., D'Aquila, P., Marasco, E., Nardini, C., Montesanto, A., Franceschi, C., ... & Bellizzi, D. (2017). The methylation of nuclear and mitochondrial DNA in ageing phenotypes and longevity. *Mechanisms of ageing and development*, 165, 156-161.
- Barrès, R., Osler, M. E., Yan, J., Rune, A., Fritz, T., Caidahl, K., ... & Zierath, J. R. (2009). Non-CpG methylation of the PGC-1 α promoter through DNMT3B controls mitochondrial density. *Cell metabolism*, 10(3), 189-198.
- Bartelli, T. F., Bruno, D. C., & Briones, M. R. (2018). Evidence for mitochondrial genome methylation in the yeast *Candida albicans*: a potential novel epigenetic mechanism affecting adaptation and pathogenicity?. *Frontiers in genetics*, 9, 166.
- Bellizzi, D., D'Aquila, P., Scafone, T., Giordano, M., Riso, V., Riccio, A., & Passarino, G. (2013). The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. *DNA research*, 20(6), 537-547.
- Bestor, T. H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *The EMBO Journal*, 11(7), 2611.
- Bestor, T. H. (2003). Cytosine methylation mediates sexual conflict. *Trends in Genetics*, 19(4), 185-190.
- Bianchessi, V., Vinci, M. C., Nigro, P., Rizzi, V., Farina, F., Capogrossi, M. C., ... & Lauri, A. (2016). Methylation profiling by bisulfite sequencing analysis of the mtDNA Non-Coding Region in replicative and senescent Endothelial Cells. *Mitochondrion*, 27, 40-47.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & development*, 16(1), 6-21.

- Blanch, M., Mosquera, J. L., Ansoleaga, B., Ferrer, I., & Barrachina, M. (2016). Altered Mitochondrial DNA Methylation Pattern in Alzheimer Disease–Related Pathology and in Parkinson Disease. *The American journal of pathology*, 186(2), 385-397.
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.
- Breton, S., Beaupré, H. D., Stewart, D. T., Hoeh, W. R., & Blier, P. U. (2007). The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough?. *Trends in Genetics*, 23(9), 465-474.
- Breton, S., Beaupré, H. D., Stewart, D. T., Piontkivska, H., Karmakar, M., Bogan, A. E., ... & Hoeh, W. R. (2009). Comparative mitochondrial genomics of freshwater mussels (Bivalvia: Unionoida) with doubly uniparental inheritance of mtDNA: gender-specific open reading frames and putative origins of replication. *Genetics*, 183(4), 1575-1589.
- Breton, S., Ghiselli, F., Passamonti, M., Milani, L., Stewart, D. T., & Hoeh, W. R. (2011). Evidence for a fourteenth mtDNA-encoded protein in the female-transmitted mtDNA of marine mussels (Bivalvia: Mytilidae). *PLoS One*, 6(4), e19365.
- Breton, S., Milani, L., Ghiselli, F., Guerra, D., Stewart, D. T., & Passamonti, M. (2014). A resourceful genome: updating the functional repertoire and evolutionary role of animal mitochondrial DNAs. *Trends in genetics*, 30(12), 555-564.
- Breton, S., Bouvet, K., Auclair, G., Ghazal, S., Sietman, B. E., Johnson, N., ... & Guerra, D. (2017). The extremely divergent maternally-and paternally-transmitted mitochondrial genomes are co-expressed in somatic tissues of two freshwater mussel species with doubly uniparental inheritance of mtDNA. *PloS one*, 12(8), e0183529.
- Capt, C., Renaut, S., Ghiselli, F., Milani, L., Johnson, N. A., Sietman, B. E., ... & Breton, S. (2018). Deciphering the link between doubly uniparental inheritance of mtDNA and sex determination in bivalves: clues from comparative transcriptomics. *Genome biology and evolution*, 10(2), 577-590.
- Castegna, A., Iacobazzi, V., & Infantino, V. (2015). The mitochondrial side of epigenetics. *Physiological genomics*, 47(8), 299-307.
- Chestnut, B. A., Chang, Q., Price, A., Lesuisse, C., Wong, M., & Martin, L. J. (2011). Epigenetic regulation of motor neuron cell death through DNA methylation. *The Journal of Neuroscience*, 31(46), 16619-16636.
- Clayton, D. A. (2000). Transcription and replication of mitochondrial DNA. *Human Reproduction*, 15(suppl_2), 11-17.
- Colot, V., & Rossignol, J. L. (1999). Eukaryotic DNA methylation as an evolutionary device. *Bioessays*, 21(5), 402-411.

- Dalziel, A. C., & Stewart, D. T. (2002). Tissue-specific expression of male-transmitted mitochondrial DNA and its implications for rates of molecular evolution in *Mytilus* mussels (Bivalvia: Mytilidae). *Genome*, 45(2), 348-355.
- D'Aquila, P., Giordano, M., Montesanto, A., De Rango, F., Passarino, G., & Bellizzi, D. (2015). Age-and gender-related pattern of methylation in the MT-RNR1 gene. *Epigenomics*, 7(5), 707-716.
- D'Aquila, P., Montesanto, A., Guarasci, F., Passarino, G., & Bellizzi, D. (2017). Mitochondrial genome and epigenome: two sides of the same coin. *Front. Biosci.(Landmark Ed)*, 22, 888-908.
- Devall, M., Smith, R. G., Jeffries, A., Hannon, E., Davies, M. N., Schalkwyk, L., ... & Lunnon, K. (2017). Regional differences in mitochondrial DNA methylation in human post-mortem brain tissue. *Clinical epigenetics*, 9(1), 47.
- Díaz-Freije, E., Gestal, C., Castellanos-Martínez, S., & Morán, P. (2014). The role of DNA methylation on *Octopus vulgaris* development and their perspectives. *Frontiers in physiology*, 5, 62.
- Doucet-Beaupré, H., Breton, S., Chapman, E. G., Blier, P. U., Bogan, A. E., Stewart, D. T., & Hoeh, W. R. (2010). Mitochondrial phylogenomics of the Bivalvia (Mollusca): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. *BMC evolutionary biology*, 10(1), 50.
- Elango, N., & Soojin, V. Y. (2008). DNA methylation and structural and functional bimodality of vertebrate promoters. *Molecular biology and evolution*, 25(8), 1602-1608.
- Elliott, H. R., Samuels, D. C., Eden, J. A., Relton, C. L., & Chinnery, P. F. (2008). Pathogenic mitochondrial DNA mutations are common in the general population. *The American journal of human genetics*, 83(2), 254-260.
- Feng, S., Xiong, L., Ji, Z., Cheng, W., & Yang, H. (2012). Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer. *Molecular medicine reports*, 6(1), 125-130.
- Fernández-Vizarra, E., Enríquez, J. A., Pérez-Martos, A., Montoya, J., & Fernández-Silva, P. (2011). Tissue-specific differences in mitochondrial activity and biogenesis. *Mitochondrion*, 11(1), 207-213.
- Gao, J., Wen, S., Zhou, H., & Feng, S. (2015). De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. *Molecular medicine reports*, 12(5), 7033-7038.
- Gardiner-Garden, M., & Frommer, M. (1987). CpG islands in vertebrate genomes. *Journal of molecular biology*, 196(2), 261-282.

- Garrido-Ramos, M. A., Stewart, D. T., Sutherland, B. W., & Zouros, E. (1998). The distribution of male-transmitted and female-transmitted mitochondrial DNA types in somatic tissues of blue mussels: implications for the operation of doubly uniparental inheritance of mitochondrial DNA. *Genome*, 41(6), 818-824.
- Gavery, M. R., & Roberts, S. B. (2013). Predominant intragenic methylation is associated with gene expression characteristics in a bivalve mollusc. *PeerJ*, 1, e215.
- Gavery, M. R., & Roberts, S. B. (2014). A context dependent role for DNA methylation in bivalves. *Briefings in functional genomics*, elt054.
- Ghiselli, F., Milani, L., & Passamonti, M. (2011) Strict sex-specific mtDNA segregation in the germ line of the DUI species *Venerupis philippinarum* (Bivalvia: Veneridae). *Molecular biology and evolution* 28(2), 949-961.
- Ghiselli, F., Milani, L., Guerra, D., Chang, P. L., Breton, S., Nuzhdin, S. V., & Passamonti, M. (2013). Structure, transcription, and variability of metazoan mitochondrial genome: perspectives from an unusual mitochondrial inheritance system. *Genome biology and evolution*, 5(8), 1535-1554.
- Ghosh, S., Sengupta, S., & Scaria, V. (2014). Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. *Mitochondrion*, 18, 58-62.
- Ghosh, S., Sengupta, S., & Scaria, V. (2016). Hydroxymethyl cytosine marks in the human mitochondrial genome are dynamic in nature. *Mitochondrion*, 27, 25-31.
- Guo JU, Su Y, Shin JH, et al. (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. *Nature neuroscience*, 17(2), 215-222. doi:10.1038/nn.3607.
- Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic acids research*, 41(13), e129-e129.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., ... & Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of development*, 117(1), 15-23.
- Hayashi, J. I., Yonekawa, H., Gotoh, O., Watanabe, J., & Tagashira, Y. (1978). Strictly maternal inheritance of rat mitochondrial DNA. *Biochemical and biophysical research communications*, 83(3), 1032-1038.
- Hendrich, B., & Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *TRENDS in Genetics*, 19(5), 269-277.

- Hermann, A., Goyal, R., & Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *Journal of Biological Chemistry*, 279(46), 48350-48359.
- Holliday, R. (2006). Epigenetics: a historical overview. *Epigenetics*, 1(2), 76-80.
- Iacobazzi, V., Castegna, A., Infantino, V., & Andria, G. (2013). Mitochondrial DNA methylation as a next-generation biomarker and diagnostic tool. *Molecular genetics and metabolism*, 110(1), 25-34.
- Illingworth, R. S., & Bird, A. P. (2009). CpG islands—‘a rough guide’. *FEBS letters*, 583(11), 1713-1720.
- Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7), 484-492.
- Kelly, R. D., Mahmud, A., McKenzie, M., Trounce, I. A., & St John, J. C. (2012). Mitochondrial DNA copy number is regulated in a tissue specific manner by DNA methylation of the nuclear-encoded DNA polymerase gamma A. *Nucleic acids research*, 40(20), 10124-10138.
- Krueger, F., & Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *bioinformatics*, 27(11), 1571-1572.
- Krueger, F., Kreck, B., Franke, A., & Andrews, S. R. (2012). DNA methylome analysis using short bisulfite sequencing data. *Nature methods*, 9(2), 145-151.
- Kurdyukov, S., & Bullock, M. (2016). DNA methylation analysis: choosing the right method. *Biology*, 5(1), 3.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), 357.
- Latorre-Pellicer, A., Moreno-Loshuertos, R., Lechuga-Vieco, A. V., Sánchez-Cabo, F., Torroja, C., Acín-Pérez, R., ... & Bernad-Miana, M. L. (2016). Mitochondrial and nuclear DNA matching shapes metabolism and healthy ageing. *Nature*, 535(7613), 561.
- Legros, F., Malka, F., Frachon, P., Lombès, A., & Rojo, M. (2004). Organization and dynamics of human mitochondrial DNA. *Journal of cell science*, 117(13), 2653-2662.
- Lewis, J. B., & Riebel, P. N. (1984). The effect of substrate on burrowing in freshwater mussels (Unionidae). *Canadian Journal of Zoology*, 62(10), 2023-2025.
- Li, W., & Liu, M. (2011). Distribution of 5-hydroxymethylcytosine in different human tissues. *Journal of nucleic acids*, 2011.

- Liu, B., Du, Q., Chen, L., Fu, G., Li, S., Fu, L., ... & Bin, C. (2016). CpG methylation patterns of human mitochondrial DNA. *Scientific reports*, 6, 23421.
- Liu, W. M., & Schmid, C. W. (1993). Proposed roles for DNA methylation in Alu transcriptional repression and mutational inactivation. *Nucleic acids research*, 21(6), 1351-1359.
- Low, D. A., Weyand, N. J., & Mahan, M. J. (2001). Roles of DNA adenine methylation in regulating bacterial gene expression and virulence. *Infection and Immunity*, 69(12), 7197-7204.
- Luo, S., & Preuss, D. (2003). Strand-biased DNA methylation associated with centromeric regions in Arabidopsis. *Proceedings of the National Academy of Sciences*, 100(19), 11133-11138.
- Mandrioli, M. (2007). A new synthesis in epigenetics: towards a unified function of DNA methylation from invertebrates to vertebrates. *Cellular and molecular life sciences*, 64(19-20), 2522-2524.
- Maor, G. L., Yearim, A., & Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. *Trends in Genetics*, 31(5), 274-280.
- Matilainen, O., Quirós, P. M., & Auwerx, J. (2017). Mitochondria and epigenetics—crosstalk in homeostasis and stress. *Trends in cell biology*, 27(6), 453-463.
- Matsuda, S., Yasukawa, T., Sakaguchi, Y., Ichiyanagi, K., Unoki, M., Gotoh, K., ... & Kang, D. (2018). Accurate estimation of 5-methylcytosine in mammalian mitochondrial DNA. *Scientific reports*, 8(1), 5801.
- Massicotte, R., Whitelaw, E., & Angers, B. (2011). DNA methylation: a source of random variation in natural populations. *Epigenetics*, 6(4), 421-427.
- Mechta, M., Ingerslev, L. R., Fabre, O., Picard, M., & Barrès, R. (2017). Evidence suggesting absence of mitochondrial DNA methylation. *Frontiers in genetics*, 8, 166.
- Milani, L., Ghiselli, F., Maurizii, M. G., Nuzhdin, S. V., & Passamonti, M. (2014). Paternally transmitted mitochondria express a new gene of potential viral origin. *Genome biology and evolution*, 6(2), 391-405.
- Mishra, M., & Kowluru, R. A. (2015). Epigenetic Modification of Mitochondrial DNA in the Development of Diabetic Retinopathy. Methylation of mtDNA in Diabetic Retinopathy. *Investigative ophthalmology & visual science*, 56(9), 5133-5142.
- Mitchell, A., Guerra, D., Stewart, D., & Breton, S. (2016). In silico analyses of mitochondrial ORFans in freshwater mussels (Bivalvia: Unionoida) provide a framework for future studies of their origin and function. *BMC genomics*, 17(1), 597.

- Monk, Marilyn Boubelik, Michael., & Lehnert, Sigrid. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, 99(3), 371-382.
- Mposhi, A., Van der Wijst, M. G., Faber, K. N., & Rots, M. G. (2017). Regulation of mitochondrial gene expression, the epigenetic enigma. *Front. Biosci.(Landmark Ed)*, 22, 1099-1113.
- Nestor, C. E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., ... & Meehan, R. (2011). Tissue-type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome research*, gr-126417.
- Newsholme, P., Haber, E. P., Hirabara, S. M., Rebelato, E. L. O., Procopio, J., Morgan, D., ... & Curi, R. (2007). Diabetes associated cell stress and dysfunction: role of mitochondrial and non - mitochondrial ROS production and activity. *The Journal of physiology*, 583(1), 9-24.
- Nile, C. J., Read, R. C., Akil, M., Duff, G. W., & Wilson, A. G. (2008). Methylation status of a single CpG site in the IL6 promoter is related to IL6 messenger RNA levels and rheumatoid arthritis. *Arthritis & Rheumatism*, 58(9), 2686-2693.
- Obata, M., Sano, N., & Komaru, A. (2011). Different transcriptional ratios of male and female transmitted mitochondrial DNA and tissue - specific expression patterns in the blue mussel, *Mytilus galloprovincialis*. *Development, growth & differentiation*, 53(7), 878-886.
- Olson, C. E., & Roberts, S. B. (2014). Genome-wide profiling of DNA methylation and gene expression in *Crassostrea gigas* male gametes. *Front Physiol*, 5, 224.
- Owa, C., Poulin, M., Yan, L., & Shioda, T. (2018). Technical adequacy of bisulfite sequencing and pyrosequencing for detection of mitochondrial DNA methylation: Sources and avoidance of false-positive detection. *PloS one*, 13(2), e0192722.
- Picard, M., Wallace, D. C., & Burelle, Y. (2016). The rise of mitochondria in medicine. *Mitochondrion*, 30, 105-116.
- Pinney, S. E. (2014). Mammalian non-CpG methylation: stem cells and beyond. *Biology*, 3(4), 739-751.
- Pirola, C. J., Gianotti, T. F., Burgueño, A. L., Rey-Funes, M., Loidl, C. F., Mallardi, P., ... & Sookoian, S. (2012). Epigenetic modification of liver mitochondrial DNA is associated with histological severity of nonalcoholic fatty liver disease. *Gut*, gutjnl-2012.
- Rasmussen, D. A., & Noor, M. A. (2009). What can you do with 0.1× genome coverage? A case study based on a genome survey of the scuttle fly *Megaselia scalaris* (Phoridae). *Bmc Genomics*, 10(1), 382.

- Renaut, S., Guerra, D., Hoeh, W., Stewart, D., Bogan, A., Ghiselli, F., ... & Breton, S. (2018). Hybrid de novo assembly of the draft genome of the freshwater mussel *Venustaconcha ellipsiformis* (Bivalvia: Unionida). *bioRxiv*, 265157.
- Rebelo, A. P., Williams, S. L., & Moraes, C. T. (2009). *In vivo* methylation of mtDNA reveals the dynamics of protein–mtDNA interactions. *Nucleic acids research*, 37(20), 6701-6715.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, 447(7143), 425-432.
- Reik, W., Dean, W., & Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science*, 293(5532), 1089-1093.
- Rice, J. C., & Allis, C. D. (2001). Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Current opinion in cell biology*, 13(3), 263-273.
- Rivière, G. (2014). Epigenetic features in the oyster *Crassostrea gigas* suggestive of functionally relevant promoter DNA methylation in invertebrates. *Frontiers in physiology*, 5, 129.
- Rivière, G., Wu, G. C., Fellous, A., Goux, D., Sourdain, P., & Favrel, P. (2013). DNA methylation is crucial for the early development in the oyster *C. gigas*. *Marine biotechnology*, 15(6), 739-753.
- Robertson, K. D. (2005). DNA methylation and human disease. *Nature Reviews Genetics*, 6(8), 597-610.
- Saini, S. K., Mangalhara, K. C., Prakasam, G., & Bamezai, R. N. K. (2017). DNA Methyltransferase1 (DNMT1) Isoform3 methylates mitochondrial genome and modulates its biology. *Scientific reports*, 7(1), 1525.
- Saint-Carlier, E., & Riviere, G. (2015). Regulation of Hox orthologues in the oyster *Crassostrea gigas* evidences a functional role for promoter DNA methylation in an invertebrate. *FEBS letters*, 589(13), 1459-1466.
- Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., & Saccone, C. (1997). Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene*, 205(1), 125-140.
- Sharma, G., Sowpati, D. T., Singh, P., Khan, M. Z., Ganji, R., Upadhyay, S., ... & Khosla, S. (2016). Genome-wide non-CpG methylation of the host genome during *M. tuberculosis* infection. *Scientific reports*, 6, 25006.
- Shi, Y., Dierckx, A., Wanrooij, P. H., Wanrooij, S., Larsson, N. G., Wilhelmsson, L. M., ... & Gustafsson, C. M. (2012). Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proceedings of the National Academy of Sciences*, 201119738.

- Shock, L. S., Thakkar, P. V., Peterson, E. J., Moran, R. G., & Taylor, S. M. (2011). DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proceedings of the National Academy of Sciences*, 108(9), 3630-3635.
- Sims, D., Sudbery, I., Illott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nature Reviews Genetics*, 15(2), 121.
- Singal, R., van Wert, J., & Bashambu, M. (2001). Cytosine methylation represses glutathione S-transferase P1 (GSTP1) gene expression in human prostate cancer cells. *Cancer research*, 61(12), 4820-4826.
- Stöger, R., Genereux, D. P., Hagerman, R. J., Hagerman, P. J., Tassone, F., & Laird, C. D. (2011). Testing the FMR1 promoter for mosaicism in DNA methylation among CpG sites, strands, and cells in FMR1-expressing males with fragile X syndrome. *PLoS One*, 6(8), e23648.
- Sun, Y., Hou, R., Fu, X., Sun, C., Wang, S., Wang, C., ... & Bao, Z. (2014). Genome-wide analysis of DNA methylation in five tissues of zhikong scallop, *Chlamys farreri*. *PloS one*, 9(1), e86232.
- Taiwo, O., Wilson, G. A., Morris, T., Seisenberger, S., Reik, W., Pearce, D., ... & Butcher, L. M. (2012). Methylome analysis using MeDIP-seq with low DNA concentrations. *Nature protocols*, 7(4), 617.
- Tiedemann, R. L., Putiri, E. L., Lee, J. H., Hlady, R. A., Kashiwagi, K., Ordog, T., ... & Robertson, K. D. (2014). Acute depletion redefines the division of labor among DNA methyltransferases in methylating the human genome. *Cell reports*, 9(4), 1554-1566.
- Tomizawa, S. I., Kobayashi, H., Watanabe, T., Andrews, S., Hata, K., Kelsey, G., & Sasaki, H. (2011). Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development*, dev-061416.
- Tong, H., Zhang, L., Gao, J., Wen, S., Zhou, H., & Feng, S. (2017). Methylation of mitochondrial DNA displacement loop region regulates mitochondrial copy number in colorectal cancer. *Molecular medicine reports*, 16(4), 5347-5353.
- van der Wijst, M. G., & Rots, M. G. (2015). Mitochondrial epigenetics: an overlooked layer of regulation?. *Trends in Genetics*, 31(7), 353-356.
- van der Wijst, M. G., van Tilburg, A. Y., Ruiters, M. H., & Rots, M. G. (2017). Experimental mitochondria-targeted DNA methylation identifies GpC methylation, not CpG methylation, as potential regulator of mitochondrial gene expression. *Scientific Reports*, 7(1), 177.

- Vilkaitis, G., Suetake, I., Klimašauskas, S., & Tajima, S. (2005). Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *Journal of Biological Chemistry*, 280(1), 64-72.
- Vanyushin, B. F. (2006). DNA methylation in plants. In DNA methylation: basic mechanisms (pp. 67-122). Springer, Berlin, Heidelberg.
- Vanyushin, B. F., & Ashapkin, V. V. (2011). DNA methylation in higher plants: past, present and future. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1809(8), 360-368.
- Waddington CH. (1939). Introduction to Modern Genetics. London: Allen and Unwin
- Wallace, D. C. (2016) Mitochondrial DNA in evolution and disease. *Nature*, 535(7613), 498-500.
- Wilson, A. S., Power, B. E., & Molloy, P. L. (2007). DNA hypomethylation and human diseases. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1775(1), 138-162.
- Wong, M., Gertz, B., Chestnut, B. A., & Martin, L. J. (2013). Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. *Frontiers in Cellular Neuroscience*, 7, 279.
- Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, 328(5980), 916-919.
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2013). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30(5), 614-620.
- Zheng, L. D., Linarelli, L. E., Liu, L., Wall, S. S., Greenawald, M. H., Seidel, R. W., ... & Cheng, Z. (2015). Insulin resistance is associated with epigenetic and genetic regulation of mitochondrial DNA in obese humans. *Clinical epigenetics*, 7(1), 60.
- Zhong, X., Du, J., Hale, C. J., Gallego-Bartolome, J., Feng, S., Vashisht, A. A., ... & Jacobsen, S. E. (2014). Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. *Cell*, 157(5), 1050-1060.
- Zilberman, D. (2008). The evolving functions of DNA methylation. *Current opinion in plant biology*, 11(5), 554-559.
- Zouros, E. (2013). Biparental inheritance through uniparental transmission: the doubly uniparental inheritance (DUI) of mitochondrial DNA. *Evolutionary Biology*, 40(1), 1-31.

Annexes

Annexe 1 : tableaux supplémentaires

Supplementary Table I: Methylation percentage for every gene in the F genome found in male soma

gene	start	end	Gene length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1542	1541	+	38.7	0.937503441	0.888523165
cox3	1574	2353	779	+	40	0.98635889	0.766644777
atp6	2398	3105	707	+	38.7	0.820868257	1.165242228
trnD	3114	3178	64	+	24.6	0.981324477	1.082084912
atp8	3179	3388	209	+	35.2	0.689059133	1.164820991
nad4L	3392	3688	296	+	36	0.925762324	0.995716041
nad4	3681	5027	1346	+	39	0.918464368	0.973936138
nad6	5081	5569	488	-	35	0.994793997	0.743691151
trnG	5593	5654	61	-	41.9	0.940332509	0.782281008
nad1	5673	6575	902	-	37.4	0.971551485	0.827652574
trnL	6576	6638	62	-	44.4	0.889744542	0.903763945
trnV	6639	6701	62	-	31.7	1.091615746	1.041303041
trnI	6712	6776	64	-	46.2	0.786572176	0.967058417
trnC	6788	6851	63	-	37.5	0.877062709	0.842099368
trnQ	6858	6925	67	-	38.2	1.106329412	0.849098066
nad5	7273	8961	1688	+	37.8	1.029635956	0.91606909
trnF	9030	9092	62	-	36.5	0.621658146	1.009373622
cytb	9093	10248	1155	-	37.7	0.854817836	0.879431911
trnP	10250	10314	64	-	26.2	1.366051608	1.403174433
trnN	10326	10391	65	-	39.4	0.903854387	0.84215181
trnL	10417	10482	65	-	30.3	0.95134128	0.993289198
rrnL	10483	11796	1313	-	36.2	0.957687403	0.994867018
trnY	11797	11858	61	-	46.8	0.782513241	0.895928233
trnT	11866	11932	66	-	37.3	0.866573873	1.0322626
trnK	11934	11996	62	-	42.9	0.711801496	1.070685728
rrnS	11997	12861	864	-	39.1	0.855541	0.872509682
trnR	12862	12927	65	-	36.4	0.748499803	0.715023236
trnW	12932	12997	65	-	25.8	0.73873756	0.761206263
trnM	13022	13085	63	-	35.9	0.690276041	0.813889326
nad2	13086	14051	965	-	37.3	0.882805768	0.715566291
F-orf	14052	14321	269	-	31.5	1.065126922	0.974402672
trnE	14335	14400	65	-	42.4	0.79846088	0.90786137
trnS	14403	14472	69	-	42.9	0.740840768	1.028808744
trnS	14478	14542	64	-	36.9	0.656426269	0.977140578
trnA	14569	14632	63	-	31.3	1.305988864	0.979397554

trnH	14721	14786	65	+	27.3	0.764847089	1.121847502
nad3	14878	15234	356	+	38.7	0.8463468	0.976415743
cox2	15305	15988	683	+	40.9	0.848583565	0.96448585

Total protein coding					38.04	0.927104711	0.902131968
Total rRNA					37.4	0.917147903	0.946306275
Total tRNA					36.5	0.919974058	1.00122087
Total intergenic					32.7	0.951318238*	1.056920471*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table II: Methylation percentage for every gene in the F genome found in female soma

gene	start	end	Gene length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1542	1541	+	38.7	0.904556494	0.882564505
cox3	1574	2353	779	+	40	0.989677878	0.973295554
atp6	2398	3105	707	+	38.7	0.709408185	1.122701352
trnD	3114	3178	64	+	24.6	1.038705951	1.627904162
atp8	3179	3388	209	+	35.2	0.737157781	1.127413796
nad4L	3392	3688	296	+	36	0.917248336	0.870804718
nad4	3681	5027	1346	+	39	0.936437359	0.907734002
nad6	5081	5569	488	-	35	0.777675026	0.75186777
trnG	5593	5654	61	-	41.9	1.190475686	0.702663186
nad1	5673	6575	902	-	37.4	1.054869575	0.80386762
trnL	6576	6638	62	-	44.4	1.368659962	1.015045273
trnV	6639	6701	62	-	31.7	0.907864418	0.77190727
trnI	6712	6776	64	-	46.2	0.736339058	0.610764816
trnC	6788	6851	63	-	37.5	1.530942985	0.698074663
trnQ	6858	6925	67	-	38.2	1.588444118	0.640648692
nad5	7273	8961	1688	+	37.8	0.754720658	0.885931122
trnF	9030	9092	62	-	36.5	2.068491532	0.761903591
cytb	9093	10248	1155	-	37.7	0.973962397	0.683996357
trnP	10250	10314	64	-	26.2	1.69878567	0.878741192
trnN	10326	10391	65	-	39.4	1.002424823	0.741292404
trnL	10417	10482	65	-	30.3	0.685564043	0.903091846
rrnL	10483	11796	1313	-	36.2	0.736333384	0.856103298
trnY	11797	11858	61	-	46.8	0.857081461	1.256525984
trnT	11866	11932	66	-	37.3	1.073026936	1.273882611
trnK	11934	11996	62	-	42.9	0.996780955	0.996232036
rrnS	11997	12861	864	-	39.1	1.088724314	1.310587309
trnR	12862	12927	65	-	36.4	1.01485674	1.13423448
trnW	12932	12997	65	-	25.8	0.837683442	0.988653623
trnM	13022	13085	63	-	35.9	1.168777447	0.607401873
nad2	13086	14051	965	-	37.3	0.782823418	0.873167481
F-orf	14052	14321	269	-	31.5	1.113590284	0.788008023
trnE	14335	14400	65	-	42.4	0.669572796	0.545027959
trnS	14403	14472	69	-	42.9	0.611976799	0.801674484
trnS	14478	14542	64	-	36.9	0.663655158	0.630190826
trnA	14569	14632	63	-	31.3	0.788981662	1.307357778

trnH	14721	14786	65	+	27.3	0.838540665	0.627695359
nad3	14878	15234	356	+	38.7	0.987376709	0.606047133
cox2	15305	15988	683	+	40.9	1.235480631	0.805623903

Total protein coding					38.04	0.907647659	0.863072638
Total rRNA					37.4	0.87618904	1.036477292
Total tRNA					36.5	1.108200605	0.928649296
Total intergenic					32.7	0.991512732*	2.011494253*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table III: Methylation percentage for every gene in the F genome found in male gonads

gene	start	end	Gene length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1542	1541	+	38.7	0.405213718	0.982988189
cox3	1574	2353	779	+	40	4.494268078	0.771932647
atp6	2398	3105	707	+	38.7	1.162989699	0.829228243
trnD	3114	3178	64	+	24.6	0	0
atp8	3179	3388	209	+	35.2	0.196078431	0.141242938
nad4L	3392	3688	296	+	36	2.136948529	0.268817204
nad4	3681	5027	1346	+	39	0.860417576	0.655006859
nad6	5081	5569	488	-	35	0	0.504050799
trnG	5593	5654	61	-	41.9		0
nad1	5673	6575	902	-	37.4	0.361721612	1.262162751
trnL	6576	6638	62	-	44.4	1.785714286	0.433298319
trnV	6639	6701	62	-	31.7	0	1.83982684
trnI	6712	6776	64	-	46.2	0.925925926	1.041666667
trnC	6788	6851	63	-	37.5		0.5
trnQ	6858	6925	67	-	38.2	5.224358974	2.008547009
nad5	7273	8961	1688	+	37.8	0.606226039	0.896054314
trnF	9030	9092	62	-	36.5	5.654761905	4.444444444
cytb	9093	10248	1155	-	37.7	1.420521945	1.456771855
trnP	10250	10314	64	-	26.2	1.388888889	0.657894737
trnN	10326	10391	65	-	39.4	0.555555556	0.627705628
trnL	10417	10482	65	-	30.3	0	0.37037037
rrnL	10483	11796	1313	-	36.2	0.375287662	0.644229647
trnY	11797	11858	61	-	46.8	0	0
trnT	11866	11932	66	-	37.3	0.357142857	0
trnK	11934	11996	62	-	42.9	0	0
rrnS	11997	12861	864	-	39.1	0.219690489	4.205246914
trnR	12862	12927	65	-	36.4	0.384615385	0.324675325
trnW	12932	12997	65	-	25.8	0.555555556	0.833333333
trnM	13022	13085	63	-	35.9	0	0
nad2	13086	14051	965	-	37.3	2.203821819	0.667136441
F-orf	14052	14321	269	-	31.5	0	0.373298199
trnE	14335	14400	65	-	42.4	0.892857143	0.793650794
trnS	14403	14472	69	-	42.9	0.446428571	0.681818182
trnS	14478	14542	64	-	36.9	1.923076923	0
trnA	14569	14632	63	-	31.3	0.961538462	0.649350649

trnH	14721	14786	65	+	27.3	0	6.944444444
nad3	14878	15234	356	+	38.7	0.578703704	0.333333333
cox2	15305	15988	683	+	40.9	0.719451004	0.795420079

Total protein coding					38.04	1.106264846	0.850212125
Total rRNA					37.4	0.31353481	2.057513486
Total tRNA					36.5	1.007279602	1.059128915
Total intergenic					32.7	0.957661753*	2.500526759*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table IV: Methylation percentage for every gene in the M genome found in male soma

Gene	start	stop	Length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1537	1537	+	37.7	0.350877193	0.882089017
cox3	1622	2395	774	+	38.1	0.630630631	1.269482152
atp6	2413	3108	696	+	37.5	0.786111111	0.37037037
atp8	3163	3336	174	+	42.5	9.811594203	0
trnD	3359	3432	74	-	41.9	1.388888889	0.641025641
M-orf	3433	4278	846	+	34.5	0.806252681	1.323529412
nad4L	4268	4564	297	+	36.7	1.313131313	0.651340996
nad4	4685	6061	1377	+	36.4	0.832121653	0.792310167
nad6	6118	6597	480	-	35	0.744047619	0.627893519
trnG	6600	6663	64	-	40.6	1.19047619	1.851851852
nad1	6680	7582	903	-	36.3	4.406344985	0.699827735
trnL	7595	7660	66	-	48.5	0	0.621118012
trnV	7677	7743	67	-	34.3	1.923076923	0
trnI	7771	7838	68	-	41.2	3.846153846	0
trnC	7845	7912	68	-	32.4	2.083333333	0
trnQ	7924	7993	70	-	30	0	0.378787879
trnH	8113	8184	72	+	36.1	0.277777778	0
nad5	8371	10056	1686	+	35.6	0.749055178	0.518538913
trnF	10090	10159	70	-	32.9	0	0.833333333
cytb	10167	11333	1167	-	37.4	0.416666667	0.994770162
trnP	11334	11399	66	-	30.3	0	0
trnN	11403	11467	65	-	46.2	0	1.794871795
trnL	11470	11540	71	-	31	0	0.621118012
rrnL	11541	12871	1331	-	38.2	0.438869992	0.5
trnY	12872	12933	62	-	45.2	0	8.333333333
trnT	12952	13019	68	-	51.5	0	7.083333333
trnK	13021	13083	63	-	41.3	0	1.515151515
rrnS	13084	13938	855	-	37.9	0.596679081	1.041871921
trnR	13939	14006	68	-	33.8	0	0
trnW	14023	14086	64	-	34.4	0	0
trnM	14090	14152	63	-	31.7	0	2.083333333
nad2	14151	15107	957	-	33.3	0.865112994	0.761469222
trnE	15133	15201	69	-	34.8	0	1.133333333
trnS	15215	15284	70	-	44.3	0	0
trnS	15300	15367	68	-	41.2	0	0.641025641

trnA	15389	15455	67	-	40.3	1.851851852	0
nad3	15549	15902	354	+	37.3	2.777777778	0.641025641
cox2	15908	17140	1233	+	36.3	0.734767025	1.04151019

Total protein coding	36.4	1.145911544	0.817932133
Total rRNA	38.1	0.500593126	0.711939841
Total tRNA	38.3	0.576597175	1.213344813
Total intergenic	33	0.349726776*	0.4539801*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table V: Methylation percentage for every gene in the M genome found in female soma

Gene	start	stop	Length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1537	1537	+	37.7	0	0.523560209
cox3	1622	2395	774	+	38.1	0	0
atp6	2413	3108	696	+	37.5	0	0
atp8	3163	3336	174	+	42.5	0	0
trnD	3359	3432	74	-	41.9	0	0
M-orf	3433	4278	846	+	34.5	0	0
nad4L	4268	4564	297	+	36.7	0	0
nad4	4685	6061	1377	+	36.4	0	0
nad6	6118	6597	480	-	35	0	0
trnG	6600	6663	64	-	40.6	0	0
nad1	6680	7582	903	-	36.3	0	0
trnL	7595	7660	66	-	48.5	0	0
trnV	7677	7743	67	-	34.3	0	0
trnI	7771	7838	68	-	41.2	0	0
trnC	7845	7912	68	-	32.4	0	0
trnQ	7924	7993	70	-	30	0	0
trnH	8113	8184	72	+	36.1	0	0
nad5	8371	10056	1686	+	35.6	0	0
trnF	10090	10159	70	-	32.9	0	0
cytb	10167	11333	1167	-	37.4	0	1.869158879
trnP	11334	11399	66	-	30.3	0	0
trnN	11403	11467	65	-	46.2	0	0
trnL	11470	11540	71	-	31	0	0
rrnL	11541	12871	1331	-	38.2	0	0
trnY	12872	12933	62	-	45.2	0	14.28571429
trnT	12952	13019	68	-	51.5	0	0
trnK	13021	13083	63	-	41.3	0	0
rrnS	13084	13938	855	-	37.9	0	0
trnR	13939	14006	68	-	33.8	0	0
trnW	14023	14086	64	-	34.4	0	0
trnM	14090	14152	63	-	31.7	0	0
nad2	14151	15107	957	-	33.3	0	0
trnE	15133	15201	69	-	34.8	0	0
trnS	15215	15284	70	-	44.3	0	0
trnS	15300	15367	68	-	41.2	6.666666667	0
trnA	15389	15455	67	-	40.3	0	0
nad3	15549	15902	354	+	37.3	0	0

cox2	15908	17140	1233	+	36.3	0	0
------	-------	-------	------	---	------	---	---

Total protein coding	36.4	0	0.239245289
Total rRNA	38.1	0	0
Total tRNA	38.3	0.305686671	0
Total intergenic	33	0*	0*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table VI: Methylation percentage for every gene in the M found in male gonads

Gene	start	stop	Length	coding strand	%GC	%mC in coding strand	%mC in template strand
cox1	1	1537	1537	+	37.7	0.741368899	0.711685471
cox3	1622	2395	774	+	38.1	1.780883617	0.844495171
atp6	2413	3108	696	+	37.5	0.886736215	1.613933222
atp8	3163	3336	174	+	42.5	1.059659287	0.959119214
trnD	3359	3432	74	-	41.9	0.807799026	0.430677422
M-orf	3433	4278	846	+	34.5	0.641382659	2.735791082
nad4L	4268	4564	297	+	36.7	1.003496095	0.605450754
nad4	4685	6061	1377	+	36.4	0.665609179	0.679153178
nad6	6118	6597	480	-	35	0.666259925	0.531201269
trnG	6600	6663	64	-	40.6	0.524339802	0.571634946
nad1	6680	7582	903	-	36.3	1.058536787	0.828805303
trnL	7595	7660	66	-	48.5	1.037777788	0.521461268
trnV	7677	7743	67	-	34.3	0.945045646	0.712969584
trnI	7771	7838	68	-	41.2	1.216177012	0.897252593
trnC	7845	7912	68	-	32.4	1.313099296	0.880838124
trnQ	7924	7993	70	-	30	0.676829943	0.549259922
trnH	8113	8184	72	+	36.1	0.577773719	0.643712696
nad5	8371	10056	1686	+	35.6	0.76749656	0.734275565
trnF	10090	10159	70	-	32.9	0.55393483	0.675793179
cytb	10167	11333	1167	-	37.4	0.647556387	0.704213597
trnP	11334	11399	66	-	30.3	0.697044142	1.038909878
trnN	11403	11467	65	-	46.2	0.316634346	0.576598543
trnL	11470	11540	71	-	31	1.037777788	0.521461268
rrnL	11541	12871	1331	-	38.2	0.702958055	0.710464634
trnY	12872	12933	62	-	45.2	0.681329737	0.857529823
trnT	12952	13019	68	-	51.5	0.399739315	0.775844092
trnK	13021	13083	63	-	41.3	0.613725951	0.807391475
rrnS	13084	13938	855	-	37.9	0.69928523	1.443346473
trnR	13939	14006	68	-	33.8	0.832526777	0.400212406
trnW	14023	14086	64	-	34.4	0.382038271	0.391733062
trnM	14090	14152	63	-	31.7	0.567558421	0.390399387
nad2	14151	15107	957	-	33.3	0.443409787	0.790176117
trnE	15133	15201	69	-	34.8	0.51252609	0.974532701
trnS	15215	15284	70	-	44.3	0.404909516	1.404839501
trnS	15300	15367	68	-	41.2	0.561283418	1.70590544
trnA	15389	15455	67	-	40.3	0.197649573	0.558231612

nad3	15549	15902	354	+	37.3	0.470109227	0.898018852
cox2	15908	17140	1233	+	36.3	0.69007978	0.974935402

Total protein coding	36.4	0.788688417	0.945962855
Total rRNA	38.1	0.70152152	0.997113294
Total tRNA	38.3	0.677804673	0.741319703
Total intergenic	33	0.881008082*	0.755524742*

*For intergenic regions, there is no coding or template strands. The percentage associated with the coding strand represents methylation on the + strand and the percentage associated with the template strand represents methylation on the – strand.

Supplementary Table VII: average of the methylation percentages (%) calculated for cytosines in every context (CG, CHG, CHH), for both the top and bottom strands in different samples of *V. ellipsiformis* and other invertebrates.

Sample	Top			Average	Bottom			Average
	CG	CHG	CHH		CG	CHG	CHH	
V. ell1 gi F	0.9167	0.9302	1.0092	0.9819	0.9827	1.0451	0.9857	0.9912
V. ell7 gi F	0.8348	0.9844	0.9455	0.9378	0.9331	1.1794	1.2930	1.2600
V. ell8 gi F	1.0074	0.9742	1.0370	1.0212	1.1060	1.0097	1.0846	1.0787
V. ell2 a.m. F	0.8490	0.8606	0.9109	0.8932	0.8487	0.9397	0.9392	0.9337
V. ell7 a.m. F	1.0794	0.9679	1.0508	1.0391	1.2070	1.1397	1.1323	1.1376
V. ell11 a.m. F	0.7851	0.7600	0.7083	0.7283	0.6804	0.7385	0.7563	0.7500
V. ell8 m F	1.0950	1.0012	0.9977	1.0114	0.8604	0.7957	0.8654	0.8584
V. ell9 m F	0.7842	0.7652	0.8868	0.8502	0.8612	0.9603	0.9497	0.9453
V. ell11 m F	1.0969	1.1907	1.2272	1.2028	1.1178	1.0671	1.1363	1.1285
V. ell11 f F	0.3970	0.3918	0.3808	0.3850	0.3942	0.3268	0.3428	0.3445
V. ell8 go F	0.9679	0.8254	0.9115	0.9030	1.4681	1.5969	1.5487	1.5483
V. ell11 go F	1.6204	1.6374	1.5298	1.5622	0.0870	0.4432	0.4000	0.3848
V. ell8 go M	1.2821	1.2795	1.2478	1.2583	1.3627	1.1995	1.3618	1.3464
V. ell11 go M	0.4608	0.4139	0.3997	0.4106	0.4011	0.5385	0.4013	0.4144
<i>A. fuscus</i>	0.4033	0.4028	0.3750	0.3817	0.4026	0.3728	0.3813	0.3822
<i>C. apiculata</i>	0.4253	0.4501	0.4051	0.4134	0.3943	0.4166	0.4083	0.4083
<i>G. pyramidata</i>	0.4855	0.4524	0.4225	0.4371	0.4879	0.4412	0.4367	0.4433
<i>M. edulis</i>	0.4708	0.4347	0.4232	0.4318	0.3941	0.3850	0.4339	0.4243

Legend : Samples are grouped by tissue types, represented in different colours. V.ell: *Venustaconcha ellipsiformis* ; gi: gills ; a.m: adductor muscle ; m: mantle ; f: foot ; go: gonad