

Université de Montréal

Démystifier le lien entre la double transmission uniparentale des mitochondries et la détermination du sexe chez les bivalves

Par
Charlotte Capt

Département de sciences biologiques Faculté des arts et des sciences

Thèse présentée
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en sciences biologiques

Août, 2019

© Charlotte Capt, 2019

Cette thèse intitulée
Démystifier le lien entre la double transmission uniparentale des mitochondries et la détermination du sexe chez les bivalves

Présentée par
Charlotte Capt

A été évaluée par un jury composé des personnes suivantes

Simon Joly
Président-rapporteur

Sophie Breton
Directeur de recherche

Sébastien Renaut
Membre du jury

Richard Cordaux
Examinateur externe

Résumé

Les systèmes sexuels et les mécanismes responsables de la détermination du sexe chez les animaux sont issus de stratégies diverses. Cette incroyable diversité se reflète notamment chez les bivalves, où autant les facteurs génétiques qu'environnementaux y jouent un rôle, avec des espèces utilisant divers modes de reproduction, tels que le gonochorisme ou l'hermaphroditisme simultané ou séquentiel. La découverte la plus notable est celle d'un système de déterminisme sexuel unique qui impliquerait les mitochondries. Spécifiquement, un système de transmission sexe-spécifique de l'ADN mitochondrial, connu sous le nom de DUI (« Double Uniparental Inheritance » ou double transmission uniparentale), serait lié au maintien du gonochorisme chez certaines espèces de bivalves. La DUI implique un ADN mitochondrial qui est transmis de façon maternelle (ADNmt F) aux femelles et aux mâles, et l'autre transmis de façon paternelle (ADNmt M) aux mâles seulement. Les ADNmt F et M chez les espèces à DUI sont caractérisés par des traits uniques, comme une modification du gène *cox2*, ou encore la présence de nouveaux gènes associés à chacun des génomes mitochondriaux (des gènes sexe-spécifiques) qui ont une fonction autre que la production d'énergie contrairement aux autres gènes mitochondriaux typiques. Le lien entre la DUI et la détermination du sexe étant encore flou, trois approches ont été proposées pour aider à le démysterifier, chacune des approches constituant un chapitre de cette thèse.

Les deux premiers chapitres se sont concentrés sur des espèces de moules d'eau douce de l'ordre des Unionida, où une corrélation entre gonochorisme et DUI et hermaphroditisme et SMI (« Strictly Maternally Inheritance » ou transmission strictement maternelle) a été décrite. La première approche consistait à produire une analyse transcriptomique comparative entre les gonades mâles et femelles de deux espèces à DUI gonochoriques, *Venustaconcha ellipsiformis* et *Utterbackia peninsularis* (famille Unionidae), pour mieux comprendre les mécanismes sous-jacents à la détermination du sexe et à la DUI chez ces bivalves. Cette étude a révélé 12 000 gènes orthologues, avec 2 583 gènes différentiellement exprimés chez les deux espèces, dont les gènes *Sry*, *Dmrt1* et *Foxl2* connus pour être des éléments clés dans la détermination du sexe chez les vertébrés et d'autres bivalves. Nos résultats ont aussi été comparés avec d'autres espèces à DUI, notamment avec la palourde marine *Ruditapes philippinarum*, pour identifier des éléments partagés entre des espèces éloignées qui pourraient être responsables de la régulation de la DUI.

Globalement, ces résultats corroborent l'hypothèse selon laquelle un mécanisme d'ubiquitination modifié pourrait être responsable de la rétention de l'ADNmt paternel chez les bivalves mâles. Les analyses ont aussi révélé que la méthylation de l'ADN pourrait être impliquée dans la régulation de la DUI.

Une deuxième analyse transcriptomique comparative a été réalisée afin de discerner les mécanismes sous-jacents à la détermination du sexe et à la DUI, mais cette fois-ci entre l'espèce à DUI gonochorique *U. peninsularis* et l'espèce proche parente à SMI hermaphrodite *U. imbecillis*. Cette étude a permis de supporter l'hypothèse d'une implication des mécanismes d'ubiquitination et de méthylation dans la régulation de la DUI, ainsi que de confirmer un rôle des gènes conservés liés à la détermination du sexe également chez les bivalves hermaphrodites. Nos résultats ont également révélé de nouveaux gènes candidats ayant des rôles potentiels dans la DUI, y compris des nucléases et des facteurs impliqués dans l'autophagie / mitophagie.

Finalement, afin d'identifier des éléments génétiques mitochondriaux qui pourraient faire partie des mécanismes sous-jacents à la DUI et la détermination du sexe chez les bivalves, nous avons séquencé les ADNmt F et M complets de deux nouvelles espèces à DUI de deux familles de l'ordre des Venerida, *Scrobicularia plana* (famille Semelidae) et *Limecola balthica* (famille Tellinidae). En effet, la description complète des ADNmt chez les espèces à DUI a été effectuée chez plusieurs espèces de moules d'eau douce (ordre Unionoida), mais peu d'espèces l'ont été pour les ordres Mytilida et Venerida. Ces études sont essentielles pour retracer des signatures génétiques mitochondrielles partagées par différentes espèces à DUI.

Nos résultats ont révélé les plus grosses différences de taille ($>10\text{kb}$) et de divergence nucléotidique (jusqu'à 50% de divergence) entre les ADNmt M et F, parmi toutes les espèces à DUI. Ces différences de taille sont principalement dues à une immense insertion ($>3.5\text{kb}$) dans la séquence du gène *cox2* du génome mitochondrial M, chez nos deux espèces, un trait précédemment décrit chez les moules d'eau douce. Le gène *cox2* des mâles de *S. plana* est la plus longue séquence à travers le règne animal. Une autre fonctionnalité importante portée par les ADNmt F et M est la présence de nouveaux gènes spécifiques au sexe, comme reportée chez toutes les autres espèces à DUI jusqu'à maintenant. Les résultats combinés de cette thèse soutiennent le partage de plusieurs

éléments génétiques clés entre les espèces à DUI. De plus, un parallèle avec le système CMS (« Cytoplasmic Male Sterility » ou stérilité cytoplasmique mâle) chez les plantes, les seuls autres organismes possédant un déterminisme sexuel qui implique les mitochondries, est proposé pour expliquer le rôle de l'ADNmt dans la détermination du sexe chez les espèces de bivalves à DUI.

Mots-clés : Détermination du sexe, double transmission uniparentale, mitochondrie, transcriptomique, Bivalvia, mitogénomique, gènes ORFans, insertion de gène, *cox2*

Abstract

Sexual systems and sex determining mechanisms described among animals are extraordinarily diverse. This amazing diversity is present in bivalves where both environment and genetic factors occur, leading to, among others, gonochoric and simultaneous or sequential hermaphroditic species. The most impressive discovery is a sex-determining system that would involve mitochondria. Specifically, a unique mitochondrial DNA inheritance system, known as Doubly Uniparental Inheritance (DUI), would be related to the maintenance of gonochorism in some bivalve species. DUI involves two mitochondrial DNA lineages, one that is maternally transmitted (F mtDNA) to females and males, and the other that is transmitted paternally (M mtDNA) to males only. The F and M mtDNAs, in DUI species, are characterized by unique traits, such as a modification of the *cox2* gene, or the presence of new genes associated with each of the mitochondrial genomes (sex-specific genes) that have a function other than energy production, unlike other typical mitochondrial genes. Since the link between DUI and sex determination is still unclear, three approaches have been proposed to help demystify it, with each of the approaches constituting a chapter of this thesis.

The first two chapters focused on freshwater mussel species of the order Unionida, where a correlation between gonochorism and DUI and hermaphroditism and SMI (Strictly Maternally Inheritance) was described. The first approach was to produce a comparative transcriptomic analysis between the male and female gonads of two gonochoric DUI species; *Venustaconcha ellipsiformis* and *Utterbackia peninsularis* (Unionidae family), to better understand the mechanisms underlying sex determination and DUI in these bivalves. This study revealed 12,000 orthologous genes, with 2 583 genes differentially expressed in both species, including *Sry*, *Dmrt1*, and *Foxl2* known to be key sex-determining genes in vertebrates and other bivalve species. Our results were also compared with other DUI species, including the marine clam *Ruditapes philippinarum*, to identify shared elements between distant species that may be responsible for DUI regulation. Overall, these results support the hypothesis that a modified ubiquitination mechanism may be responsible for the retention of paternal mtDNA in male bivalves. The analyzes also revealed that DNA methylation could be involved in DUI regulation.

A second comparative transcriptomic analysis was performed to discern the mechanisms underlying sex determination and DUI between the gonochoric DUI species, *U. peninsularis*, and the closely related SMI hermaphroditic species, *U. imbecillis*. This study supported the hypothesis of an involvement of ubiquitination and methylation mechanisms in DUI regulation, as well as confirmed a role of conserved genes related to sex determination in hermaphroditic bivalves. Our results also revealed novel candidate genes with potential roles in DUI, including nucleases and factors involved in autophagy / mitophagy mechanisms.

Finally, to identify mitochondrial genetic elements that could be part of the mechanisms underlying DUI and sex determination in bivalves, we sequenced the complete F and M mtDNAs of two new DUI species, from two families of the order Venerida; *Scrobicularia plana* (Semelidae family) and *Limecola balthica* (Tellinidae family). The complete description of mtDNAs in DUI species has been carried out for several species of freshwater mussels (Unionoida order), but very few species have been described for the orders Mytilida and Venerida. Such studies are essential for tracing mitochondrial genetic signatures shared by different DUI species.

Our results revealed the largest differences in size (>10kb) and nucleotide divergence (up to 50% divergence) between M and F mtDNAs, among all DUI species. These differences in size are mainly due to a huge insertion (> 3.5kb) in the *cox2* gene of the M mtDNA from both species, a trait previously described in freshwater mussels. The *cox2* gene in *S. plana* males represents the longest *cox2* sequence across the animal kingdom. Another important feature of F and M mtDNAs is the presence of new sex-specific genes, as reported in all other DUI species so far. The combined results of this thesis support the sharing of several key genetic elements among DUI species. In addition, a parallel with the Cytoplasmic Male Sterility (CMS) system in plants, the only other organisms with a sex determination system that involves mitochondria, is proposed to explain the role of mtDNA in sex determination in DUI bivalve species.

Keywords: Sex determination, Doubly Uniparental Inheritance, mitochondria, transcriptomics, Bivalvia, mitogenomics, ORFans genes, gene insertion, *cox2*

Table des matières

1. Introduction	22
1.1 Détermination du sexe	22
1.1.1 Détermination génétique	23
1.1.2 Détermination épigénétique	25
1.1.3 Modèles mixtes et diversité.....	29
1.1.4 Détermination du sexe chez les Mollusques.....	30
1.1.5 Détermination du sexe chez les Bivalves.....	30
1.2 Mitochondries	34
1.2.1 DUI (Doubly Uniparental Inheritance ou Double Transmission Uniparentale).....	36
1.2.1.1 Transmission paternelle des mitochondries chez les bivalves.....	37
1.2.1.2 Distribution taxonomique de la DUI : similarités et différences entre les groupes.....	38
1.2.1.3 Sex-ratio chez les bivalves.....	41
1.2.1.4 Gonochorisme, hermaphrodisme et nouveaux gènes mitochondriaux.....	42
2. Objectif et Hypothèses	45
2.1 Chapitre Un	45
2.2 Chapitre Deux	46
2.3 Chapitre Trois	47
3. Déchiffrer le lien entre la double transmission uniparentale de l'ADNmt et la détermination du sexe chez les bivalves : indices d'une étude transcriptomique comparative	48
3.1 Abstract	50
3.2 Introduction	51
3.3 Materials and Methods	54
3.3.1 Sample collection and RNA isolation.....	54
3.3.2 De novo transcriptome assemblies.....	54
3.3.3 Annotation.....	55
3.3.4 Read mapping, differential expression and functional enrichment analyses.....	56
3.3.5 Blast reciprocal analysis	57
3.4 Results and discussion	57
3.4.1 Sequencing and assemblies of the transcriptomes	57
3.4.2 Sequence annotation	59
3.4.3 Differentially expressed genes (DEG) between male and female gonads in each species	61
3.4.4 Blast reciprocal analysis and enriched gene ontology terms and pathways in male and female gonads	64

3.4.5 Identification of candidate genes involved in sex determination.....	66
3.4.6 Identification of candidate genes involved in DUI.....	71
3.5 Conclusion	73
3.6 Supplementary Material	73
3.6.1 Supplementary tables.....	73
3.6.2 Supplementary figures	74
4. Chapitre 2. Détermination mitochondriale du sexe présumée chez Bivalvia : Informations tirées d'un assemblage hybride de transcriptomes d'espèces de moules d'eau douce.....	76
4.1 Abstract	78
4.2 Introduction.....	79
4.3 Methods	81
4.3.1 Sample Collection and RNA Extraction.....	81
4.3.2 Reads Quality Assessment, De Novo Assemblies and Completeness Appraisal	82
4.3.3 Reads Alignment and Differential Gene Expression	84
4.3.4 Annotations and Gene Ontology Enrichment	85
4.4 Results and Discussion.....	85
4.4.1 Transcriptome Assemblies and Completeness.....	85
4.4.2 Annotation.....	87
4.4.3 Screening of Differentially Expressed and Sex-Specific Genes.....	88
4.4.4 Gene Ontology (GO) and Enrichment Analyses	93
4.4.5 Candidate Genes Involved in Sex Determination.....	95
4.4.6 Candidate Genes Involved in Mitochondrial Inheritance	97
4.5 Conclusion	101
4.6 Supplementary Material	102
4.6.1Supplementary tables.....	102
4.6.2 Supplementary figures	103
5. Chapitre 3. L'analyse mitogénomique comparative de deux bivalves vénéridés utilisant la transmission mitochondriale doublement uniparentale révèle une insertion record dans le gène <i>cox2</i>.....	104
5.1 Abstract	106
5.2 Introduction.....	107
5.3 Materials and methods.....	110
5.3.1 Specimen collection and sample preparation.....	110
5.3.2 Mitochondrial genome sequencing	111
5.3.3 Characterization of mtDNAs and sequence analyses	112
5.4 Results and Discussion.....	114
5.4.1 Main genomic features.....	114
5.4.2 Intraspecific divergences	119

5.4.3 Mitochondrial ORFans.....	123
5.4.4 Insertion in the male <i>cox2</i> gene	125
5.5 Conclusion	127
5.6 Supplementary Material.....	129
5.6.1 Supplementary figures	129
6. Discussion	137
6.1 Aspects nucléaires sous-jacents à la détermination du sexe chez les bivalves	138
6.2 Aspects nucléaires sous-jacents à la DUI	140
6.3 Aspects mitochondriaux sous-jacents à la DUI	143
6.3.1 Une insertion dans le gène <i>cox2</i>	144
6.3.2 Nouveaux ORFs.....	145
6.4 CMS – Stérilité cytoplasmique mâle	148
6.4.1 Liste des facteurs Rf.....	149
7. Conclusions et perspectives	152
8. Références	153

Liste des tableaux

Chapitre 1

Table 1.1. De novo assembly quality statistics of *Utterbackia peninsularis* and *Venustaconcha ellipsiformis* transcriptomes

Table 1.2. Sex-specifically expressed genes in male and female gonads of *Utterbackia peninsularis* and *Venustaconcha ellipsiformis*

Table 1.3. Top ten genes showing the greatest difference in expression in male and female gonads of *Utterbackia peninsularis* and *Venustaconcha ellipsiformis*.

Table 1.4. Key sex-determining pathway genes in model organisms and bivalve species.

Table S1.1. Blast reciprocal results.

Table S1.2. Ontology categories significantly enriched in both species.

Table S1.3. GOSeq results for *U. peninsularis*.

Table S1.4. GOSeq results for *V. ellipsiformis*.

Table S1.5. Short description of key sex determining genes in animals.

Chapitre 2

Table 2.1. *De novo* assembly quality and annotation statistics of *Utterbackia imbecillis*, *Utterbackia peninsularis* and the hybrid transcriptomes.

Table 2.2. Sex-specific genes in hermaphrodites, males and females.

Table 2.3. Candidate genes involved in sex determination pathway in freshwater mussels and other bivalves.

Table 2.4. Putative candidate genes involved in mitochondrial inheritance in freshwater mussels.

Table S2.1. Most up-regulated genes based on their logFold-Change (logFC) values in each pairwise-comparison: Male-Female (MF), Hermaphrodite-Female (HF) and Hermaphrodite-Male (HM).

Table S2.2. Hermaphrodite-specific DEGs.

Table S2.3. Male-specific and hermaphrodite-specific DEGs and enriched GO categories related to "ubiquitination pathway", "nucleases/polymerases", "autophagy and mitophagy", and "methylation".

Chapitre 3

Table 3.1. Mitochondrial genome sizes in DUI species exhibiting extreme length differences between M and F types.

Table 3.2. General features of *Scrobicularia plana* and *Limecola balthica* M and F mitochondrial genomes.

Table 3.3. Number of nucleotides at gene boundaries in F and M mitochondrial genomes of *Scrobicularia plana* and *Limecola balthica*.

Table 3.4. Intraspecific divergence (uncorrected *p*-distances) of the 13 mitochondrial protein-coding genes between M and F mtDNAs in DUI species. (a) nucleotides (b) amino acids.

Discussion

Tableau 4. Tableau récapitulatif des gènes différentiellement exprimés attendus pour chaque paire de comparaison (MF, HF, HM), selon chaque catégorie fonctionnelle liée à la transmission mitochondriale.

Liste des figures

Introduction

Figure 1. Modèles théoriques de changements de sexe chez les bivalves, basé sur plusieurs études de changements de sexe, figure issue de Breton et al. (2018). RS : Succès reproducteur. La ligne bleue représente le RS chez les mâles et la ligne rouge celui des femelles. **(a)** Modèle avantageant la taille selon Policansky (1982). Le RS des mâles est indépendant de la taille du corps, tandis qu'il avantage proportionnellement celui des femelles, dû au coût de production des spermatozoïdes qui est moindre comparé à celui des œufs. Après un certain seuil de taille, les femelles ont un RS supérieur à celui des mâles puisqu'elles peuvent produire des œufs plus efficacement que celles plus petites. Puisque le RS des mâles n'a pas été évalué avec l'augmentation de la taille, son RS est dépeint sous forme de plateau pour se concentrer sur le changement de RS lié au changement de sexe. **(b)** Le second modèle illustre le RS lié à la disponibilité en nourriture. Le RS des mâles et des femelles est proportionnel à cette disponibilité et à leur capacité à la consommer. Lorsque les ressources diminuent, le RS des deux sexes diminue, mais celui des femelles chute plus rapidement que celui des mâles puisqu'un œuf demande plus d'énergie de production qu'un spermatozoïde. En-dessous d'un certain seuil d'accessibilité de nourriture, les femelles ont un RS inférieur à celui des mâles, ce qui favorise un changement de sexe de femelle-vers-mâle pour maintenir un meilleur RS chez l'individu.

Figure 2. Modèle de détermination du sexe chez l'huître Pacifique *Crassostrea gigas*. Les mécanismes sous-jacents à la détermination du sexe sont comparés entre organismes modèles selon le résumé de Gamble & Zarkower (2012). Seuls les gènes clés de régulation spécifiques au sexe sont montrés. Les traits noirs indiquent les relations temporelles, et les pointillés indiquent les relations hypothétiques basées sur des données d'expression. Cette figure est issue de Breton et al. (2018), elle-même modifiée de Zhang et al. (2014). Les génotypes FF permettent un changement de sexe (Guo et al. 1998).

Figure 3. Modèle de double transmission uniparentale chez les bivalves. Les mitochondries maternelles sont illustrées par des ronds rouges, tandis que les ronds bleus représentent les mitochondries paternelles. Lors de la fertilisation, l'embryon comprend un mélange des mitochondries maternelles (F) et paternelles (M). Si l'embryon est destiné à devenir une femelle, les mitochondries paternelles seront dégradées pour ne garder que celles de la mère dans tous les tissus de l'individu. À contrario, si l'embryon est destiné à devenir un mâle, les mitochondries paternelles vont s'agréger et migrer dans la cellule qui donnera la lignée germinale donnant ainsi un individu avec des tissus somatiques colonisés par les mitochondries F et des tissus sexuels par les mitochondries M.

Figure 4. Nouveaux gènes associés aux génomes mitochondriaux complets de F (femelle), H (hermaphrodite) et M (mâle). Les génomes M et F sont issus d'espèces de moule d'eau douce. Modifiée à partir de Breton et al. (2009).

Chapitre 1

Figure 1.1. Taxonomic distribution of top BLASTp hits in Uniref90 protein database for *U. peninsularis* contigs (A) and *V. ellipsiformis* contigs (B). Numbers of unique hits in each group are shown.

Figure 1.2. Smear plots of DEG between male and female gonads in *U. peninsularis* (A) and *V. ellipsiformis* (B). Black dots represent genes normally expressed in both tissues. Green dots, DEG revealed by edgeR software; orange dot, DEG revealed by DESeq2. Green dots surrounded by orange represent genes considered as DEG by both software packages. Positive fold-changes represent genes upregulated in males, while negative ones represent genes upregulated in females.

Figure 1.3. Significantly enriched-GO terms involved in molecular function (top), biological process (middle), and cellular component (bottom) in ovary (pink) and testis (blue) of both species. Analyses were performed on DEGs from blast reciprocal analyses between *V. ellipsiformis* and *U. peninsularis*. X-axis shows the number of DEGs contained in each GO category.

Figure S1.1. The results of BUSCO analysis of *U. peninsularis* and *V. ellipsiformis* transcriptomes. (A) Analysis of contigs. (B) Analysis of unigenes. Light-blue: complete and single-copy genes; dark-blue: complete and duplicated genes; yellow: fragmented genes; red: missing genes.

Figure S1.2. Taxonomic distribution of top BLASTp hits in SwissProt protein database for *U. peninsularis* contigs (A) and *V. ellipsiformis* contigs (B). Numbers of unique hits in each group are shown.

Chapitre 2

Figure 2.1. Differential expression analysis of *Utterbackia* spp unigenes. Venn diagram shows differentially expressed genes (DEGs) between sexes: hermaphrodite-female (HF in green), hermaphrodite-male (HM in pink) and male-female (MF in orange) comparisons. The figure shows the number of DEGs unique for one comparison (e.g. 2,877 for MF), the number of DEGs shared between two comparisons (e.g. 5,188 between MF and HM comparisons) and the number of DEGs shared by all comparisons (1,423).

Figure 2.2. Principal Component Analysis (PCA) performed on regularized log transform (rlog) values of the first thousand most expressed genes per pairwise comparison between the three sexes (males, females and hermaphrodites). The four males are shown in blue, the four females in red and the three hermaphrodites in green.

Figure 2.3. Top ten of most significantly enriched Gene Ontology (GO) categories in each comparison; hermaphrodite-male (HM), hermaphrodite-female (HF), and male-female (MF), involved in both biological processes (a) and molecular functions (b). Differentially expressed genes (DEGs) numbers involved in each GO categories are shown by x-axis; female-biased genes are illustrated in pink, male-biased in blue and hermaphrodite-biased in purple. Complete GO terms are written here, *; cellular component assembly involved in morphogenesis, *²; nucleobase-containing compound kinase activity, *³; nucleoside diphosphate kinase activity, *⁴; phosphotransferase activity, phosphate group as acceptor, *⁵; nucleobase-containing compound kinase activity.

Figure S2.1. Quality evaluation of unigenes from *U. imbecillis* and the hybrid transcriptome based on two lineage-specific datasets Benchmarking Universal Single-Copy Orthologs (BUSCO). A. Quality assessment of *U. imbecillis* transcriptome based on Eukaryota and Metazoa datasets, respectively. B. Quality assessment of hybrid transcriptome based on Eukaryota and Metazoa datasets, respectively.

Chapitre 3

Figure 3.1. Gene maps of M and F genomes of *Scrobicularia plana* and *Limecola balthica*. All genes are encoded on the heavy strand, total genome lengths are reported inside their corresponding genome. Gene colors correspond to functional groups (OXPHOS gene families, tRNAs and rRNAs).

Figure 3.2. Individual M-versus-F gene conservation expressed as p-distance and the rates of synonymous and nonsynonymous substitutions within these same mt genes for *Scrobicularia plana* (**A & B, respectively**) and *Limecola balthica* (**C & D, respectively**). Genes are color-coded by gene family.

Figure 3.3. Fcox2 and Mcox2 structural features in *Scrobicularia plana* and *Limecola balthica*. For the sequence alignments of the “heme-patch” regions and Cu_a centers, identical amino acids are indicated by an *.

Figure S3.1. Secondary structures of *trnP* and *trnA* duplications in the M mtDNA of *L. balthica*.

Figure S3.2. Sequence and secondary structure of predicted ORFs found in all unassigned regions longer than 150nt per sex per mitochondrial genome. Only ORFs with at least one predicted TMH were retained.

Figure S3.3. *Scrobicularia plana* F and M COX2 proteins alignment. Identical amino acids are indicated by an *.

Figure S3.4. *Limecola balthica* F and M COX2 proteins alignment. Identical amino acids are indicated by an *.

Liste des sigles et des abréviations

Les termes en anglais sont en caractères italiques, les termes en français sont en caractères droits.

aa: *amino acids* | acides aminés

BLAST: *Basic Local Alignment Search Tool*

BLASTp: *from a protein sequence to a protein sequence*

BLASTx: *from a translated nucleotide sequence to a protein sequence*

tBLASTn: *from a protein sequence to a translated nucleotide sequence*

BUSCO: *Benchmarking Universal Single-Copy Orthologs*

cDNA: *complementary DNA* | ADN complémentaire

CMS: *Cytoplasmic Male Sterility* | Stérilité Cytoplasmique Mâle

DEG: *Differentially Expressed Genes* | Gènes différemment exprimés

DNA, ADN: *Desoxyribonucleic acid* | Acide désoxyribonuléique

DMRT1: *Doublesex / Male-abnormal-3 Related Transcription factor*

DUI: *Doubly Uniparental Inheritance* | Double Transmission Uniparentale

FDR: *False-Discovery Rate*

FPKM: *Fragments per kilobase of exon per million*

F-ORF: *Female open reading frame* | cadre de lecture ouvert femelle

GO: *Gene Ontology* | Catégorie fonctionnelle

H-ORF: *Hermaphrodite open reading frame* | cadre de lecture ouvert hermaphrodite

IRIC: *Institute for Research in Immunology and Cancer Institute* | Institut de Recherche en Immunologie et en Cancérologie

ls-CSD: *locus single-locus complementary sex determination*

Ka: *number of non-synonymous substitutions per non-synonymous sites*

Ks: *number of synonymous substitutions per synonymous sites*

log2FC: *log2 Fold-Change*

M-ORF: *Male open reading frame* | cadre de lecture ouvert mâle

Fcox2, Mcox2: *cox2 gene from mitochondrial genome female or male*

mtDNA, ADNmt: *mitochondrial DNA* | ADN mitochondrial

mRNA: *messenger RNA* | ARN messager

NCBI: *National Centre for Biotechnology Information*

NR/ Nr database: *NCBI non-redundant database*

nt: *nucleotides* | nucléotides

ORF: *Open Reading Frame* | Cadre de lecture ouvert

PCA: *Principal Component Analysis* | Analyse en composante principale

PCG: *Protein-coding genes* | Gènes qui codent pour des protéines

PCR: *Polymerase chain reaction* | Réaction en chaîne de la Polymérase

PPR: Protéines pentatricopeptidiques

RBH: *Reciprocal-best-BLAST-hits*

Rf: *Fertility restorer* | Retaurateur de la fertilité

RNA, ARN: *Ribonucleic acid* | Acide ribonucléique

SMI: *Strictly Maternally Inheritance* | Transmission Strictement Maternelle

SRA: *Sequence Read Archive*

SRY: *Sex-determining Region Y*

TDF: *Testis Determining Factor*

TMH: *Transmembrane domain or helix* | Domaine ou hélice transmembranaire

tRNA: *transfer RNA* | ARN de transfert

Uniprot: *Universal Protein Resource*

À Ann, Jean-Maurice et Maxime

Remerciements

Je tiens à remercier en premier lieu ma directrice de thèse, Sophie Breton. Sophie, merci énormément de m'avoir fait confiance en me donnant l'opportunité de commencer un doctorat dans ton laboratoire, à un moment aussi critique de ma carrière scientifique. Je n'ai jamais su t'exprimer à quel point je te suis reconnaissante, et j'aimerais trouver un terme suffisamment fort pour le manifester. Tu m'as exposée à un système fascinant chez un animal sous-estimé, qui au-delà de son aspect comique, m'a surtout abasourdie quant à la diversité de thèmes qui y sont associés. À travers la transcriptomique, les génomes mitochondriaux, et les systèmes sexuels qui ont façonné ma thèse, j'en ai appris tellement. Cet apprentissage a été d'autant plus bénéfique, grâce à la grande diversité de projets qui règne dans ton laboratoire, et leurs champs d'étude respectifs. Je te remercie également pour ta supervision inestimable; ta patience, tes encouragements, tes conseils, et pour t'être toujours rendue aussi disponible et accessible pour nous aider. C'est ton soutien sans faille qui m'a aidée à prendre confiance dans ce milieu, ainsi que ton enthousiasme, ta passion et tes idées qui ont été un immense plaisir à côtoyer à chacune de nos rencontres, et qui ont été de véritables sources de motivation continues. Merci pour ton ouverture d'esprit, merci de m'avoir toujours poussée et merci pour les nombreuses opportunités offertes durant mon doctorat. Je n'ai pas pu toutes les réaliser, pour différentes raisons, mais je n'oublierai jamais ta générosité. Tous ces points m'ont permis de m'épanouir scientifiquement, d'éveiller sans cesse ma curiosité et mes idées, et rien que pour ça, je te serai éternellement reconnaissante. Au-delà de la science, je voudrais aussi te remercier pour l'incroyable atmosphère que tu as su créer au sein de ton laboratoire, où travail et plaisir ont su s'alterner dans un équilibre parfait, durant ces cinq années.

Je voudrais ensuite remercier très chaleureusement Sébastien Renaut pour son aide précieuse et son accompagnement tout au long de mes analyses transcriptomiques. Tu m'as initiée à un monde captivant, en m'offrant la meilleure formation grâce à tes bonnes et claires explications, ta disponibilité, ton investissement, et tes nombreux conseils d'organisation, même si je n'ai pas toujours su parfaitement appliquer ces derniers. Cette étape a représenté la plus grande partie de mon doctorat, et j'ai été très chanceuse d'apprendre de toi. Au-delà de ça, tu m'as également offert un soutien constant extrêmement apprécié, et j'ai eu beaucoup de plaisir à évoluer à tes côtés, et pouvoir échanger avec toi à ce sujet.

Mon expérience n'aurait pas été la même si je n'avais pas pu fréquenter les personnes aussi formidables que celles qui sont passées par le labo Breton. D'abord un grand merci à Ioana, notre « Saroumaine » au grand coeur, et son aide formidable pour la partie labo de mon projet. Je remercie tous les membres passés ou présents, avec une mention particulière pour Arthur, Davide, Karim, Philip, Georges et notre petit adopté, Rudy. Et le dernier, mais pas le moindre, Stefano. Stefano, merci pour nos discussions très constructives sur nos projets, merci pour nos « drama », merci pour nos pauses, merci pour ton soutien. Je pourrais développer sur des pages, donc je vais résumer ça par « You're the best ». J'en profite pour remercier Maria, avec la même intensité, pour toutes nos conversations, nos rires, ton écoute, tes conseils. Bref, merci à tous pour votre amitié, merci pour votre support et merci pour nos échanges aussi absurdes que sérieux. La « gang » du labo Breton ne serait pas le même sans le labo Dubois, qui ont fait de ma dernière année, une superbe année. Par conséquent un énorme merci à Marie (Number one), Guillaume, Étienne et Victoria.

Je remercie vivement Annie Angers et Simon Joly pour leurs conseils durant mon comité-conseil, qui ont permis à ma thèse de prendre la meilleure direction possible, et à Simon Joly, encore une fois, Sébastien Renaut et Richard Cordaux d'avoir accepté de réviser ma thèse.

Finalement, je tiens à remercier du plus profond de mon cœur, Maxime, Jean-Maurice et Ann pour leur support constant. Je sais que de me voir partir trois années, qui sont devenues quatre, puis cinq, n'a pas été facile pour vous. Surtout l'année passée qui a été très éprouvante. Je vous suis donc d'autant plus reconnaissante d'avoir compris l'importance d'aller « checker le sexe des moules » à l'autre bout de l'Atlantique, de m'avoir écoutée en parler et, pire, de vous y être intéressés. Vous êtes formidables. Sans votre appui moral, votre confiance, vos nombreux dons pascaux, votre amour, nos rires, vos visites, vos encouragements, votre patience à mes crises de panique et mes pleurs, je ne sais pas si j'aurais pu mener cette thèse à son terme. Je voudrais également souligner le support de tous mes amis, avec une reconnaissance très lamesque et très spéciale pour Annie et Nathalie, qui m'ont même fait l'honneur de venir me visiter plusieurs fois. Également un merci particulier à Dimitri qui m'a poussée à entreprendre les démarches nécessaires pour commencer un doctorat, pour son support constant, son aide et ses

encouragements, épaulé ensuite par Sébastien et Alexis. Et surtout, merci à vous trois et à « Jarry » pour m'avoir permis de me vider la tête de la plus belle façon.

1. Introduction

1.1 Détermination du sexe

Depuis des millénaires, le mystère entourant l'évolution du sexe intrigue. Dans les récits bibliques, la notion de genre apparaît au moment où il est dit qu'Ève a été créée à partir d'une des côtes d'Adam. Pendant la majeure partie de l'histoire de l'humanité, les théories sur la détermination du sexe ont dû être formulées en l'absence de connaissance sur les œufs et les spermatozoïdes. Cela a donné lieu à diverses théories farfelues. Par exemple, Pythagore (5^{ième} siècle avant J.-C.) stipulait que seuls les mâles fournissaient les caractéristiques à leur descendance, la mère jouant le simple rôle de fournisseuse de matériel. Anaxagoras (5^{ième} siècle avant J.-C.) pensait que le facteur déterminant le sexe était le testicule dont provenait le sperme : l'embryon était appelé à devenir un mâle s'il avait été conçu avec la semence provenant du testicule droit, tandis que la semence du testicule gauche produisait des femelles. Plusieurs centaines d'années plus tard, Michel Procope-Couteau (1684 – 1753) a même repris un peu cette idée en conseillant aux femmes de s'allonger de manière à ce que la gravité favorise le sexe de l'enfant voulu (Mittwoch, 1985 ; Beukeboom & Perrin, 2014). Suite à l'invention du microscope au courant des années 1670 par le scientifique hollandais Antonie van Leeuwenhoek qui s'en est servi, entre-autre, pour examiner des spermatozoïdes, Oskar Hertwig (1849 – 1922) devient le premier à pouvoir observer une fusion entre les noyaux de l'œuf et du spermatozoïde. À la fin du 19^e siècle, l'effet de l'environnement est lui mis en cause dans la détermination du sexe puisqu'on pensait que le régime alimentaire de la mère, durant ses trois premiers mois de grossesse, définirait le sexe de l'enfant. Au fil des années, les hypothèses se sont affutées avec notamment la découverte des chromosomes sexuels et la contribution claire de la génétique (Beukeboom & Perrin, 2014). Cela dit, l'intérêt pour ces questions est toujours d'actualité aujourd'hui. En effet, quelles sont les forces évolutives responsables du dimorphisme sexuel à travers le règne animal et végétal ? Quels en sont les avantages et, surtout, comment expliquer l'énorme diversité de systèmes sexuels qui existe non seulement entre les embranchements, mais également entre les espèces d'un même genre (Bull, 1983 ; Valenzuela, 2008 ; Wallace et al., 1999) ?

Le sexe se définit par la présence d'événements méiotiques ; un mécanisme spécialisé qui permet la recombinaison entre deux génomes (Beukeboom & Perrin, 2014). L'avantage évolutif de la production de deux sexes permet, entre autres, de contrecarrer les dommages dus aux réactifs oxydatifs sur l'ADN, par recombinaison homologue, ainsi qu'une diversité génétique accrue sur laquelle la sélection naturelle agit (Burt, 2000). Cela favorise, par conséquent, le potentiel adaptatif d'un individu face à son environnement. Pourtant, le coût lié au développement des gamètes est extrêmement élevé (coût énergétique pour passer de la mitose à la méiose) et la recombinaison peut parfois se trouver désavantageuse lorsqu'elle aboutit à briser des combinaisons génétiques bénéfiques à l'organisme (Lehtonen et al., 2012). Malgré les avantages évolutifs évidents de la formation de deux sexes, les coûts semblent plus faciles encore à identifier. Chez plusieurs organismes par exemple, pour se reproduire sexuellement, un individu doit trouver un compagnon volontaire. Ainsi, le mystère entourant l'évolution du sexe est loin d'être résolu mais il semblerait que la grande diversité actuelle des processus sexuels sous-entend qu'une cause évolutionnaire unique dans l'émergence des sexes est exclue (West et al., 1999), et qu'il est conséquemment bon de se pencher sur les différents mécanismes existants pour répondre à ces questions (Beukeboom & Perrin, 2014).

Il est donc important de commencer par définir et différencier les notions de détermination du sexe et de différenciation du sexe. Cette dernière comprend l'ensemble des mécanismes qui sont responsables du développement d'une gonade sexuellement indifférenciée en gonade mâle ou en gonade femelle. La détermination du sexe, quant à elle, comprend les facteurs responsables du déclenchement de cette différenciation (Hayes, 1998). Ces facteurs sont généralement regroupés en deux catégories : détermination génétique du sexe et détermination épigénétique du sexe (Manolakou, Angelopoulou, & Lavranos, 2006).

1.1.1 Détermination génétique

La détermination génétique du sexe représente l'ensemble des facteurs génétiques, tels que les chromosomes sexuels [découverts par Nettie Stevens en 1905 (Brush, 1978)] et leurs (ou autres) gènes associés, qui sont responsables de déclencher la différenciation des gonades. Les chromosomes sexuels sont généralement morphologiquement et génétiquement bien différenciés. Deux principaux systèmes de chromosomes sexuels existent ; XX/XY et ZZ/ZW. Un sexe sera

défini par deux chromosomes dissimilaires (sexe hétérogamétique ; XY ou ZW, qui produit deux types de gamètes haploïdes ; X et Y ou Z et W) et l'autre par deux chromosomes similaires (sexe homogamétique ; XX ou ZZ, qui produit un type de gamète haploïde X ou Z). On retrouve le système de chromosomes sexuels XX/XY principalement chez les mammifères et le système ZZ/ZW chez les oiseaux (Ezaz et al., 2006). Chez les mammifères, les mâles sont le sexe dit hétérogamétique (XY) alors que chez les oiseaux, ce sont les femelles qui sont le sexe hétérogamétique (ZW).

Le principal gène responsable de la détermination du sexe mâle chez les mammifères est *SRY* (Sex-determining Region Y), car il est porté par le chromosome Y (Sinclair et al., 1990). La protéine qu'il code est un facteur de transcription (*TDF* ou *Testis Determining Factor*) qui permet de déclencher la cascade moléculaire à la base de la détermination sexuelle mâle et qui est donc considéré comme le principal régulateur de ce processus chez les mammifères (Sinclair et al., 1990; Koopman, 1991). Une fois la gonade indifférenciée stimulée par l'expression du gène *SRY*, la différenciation en testicules peut débuter, celles-ci vont éventuellement produire des hormones qui vont permettre, d'une part, la production d'organes masculins essentiels au bon développement d'un organisme reproducteur mâle, et, d'autre part, empêcher la formation du canal reproducteur femelle (Hake & O'Connor, 2008; Jakob & Lovell-Badge, 2011). En absence de *SRY*, l'organisme est programmé à se développer en femelle, appuyé par la production de certaines protéines et hormones (œstrogène) qui favoriseront la formation des organes féminins essentiels au succès reproducteur des femelles.

Les facteurs clés déclencheurs de la détermination du sexe tels que *SRY* sont très diversifiés à travers le règne animal, et, même entre deux espèces proches parentes, la détermination du sexe ne sera pas forcément déclenchée par le même facteur (pour une revue ; Bachtrog et al., 2014). Chez les oiseaux et les amphibiens, il semblerait que ce soit *DMRT1* (Doublesex/Male-abnormal-3 Related Transcription factor) qui soit le facteur déclencheur de la détermination du sexe. Ce gène joue également un rôle dans le déterminisme sexuel chez l'homme et la souris, mais il agit en aval comme facteur intermédiaire dans la cascade moléculaire complexe induite par *SRY* (Matson & Zarkower, 2012 ; Picard et al., 2015).

Chez certaines espèces d'invertébrés, dont la drosophile, il existe une situation où le sexe est déterminé par le ratio entre le nombre de chromosomes sexuels (X) et le nombre d'autosomes (A) (Cline & Meyer, 1996 ; Lalli et al., 2003 ; Macdougall et al., 1995; Penalva et al., 2003; Saccone et al., 2002). Les facteurs responsables du développement de la femelle sont portés par les chromosomes X, tandis que ceux liés au développement des mâles sont portés par les autosomes. Dans ce système, le chromosome Y n'est pas impliqué dans la détermination du sexe, mais dans la différentiation du sexe, puisqu'il contient des gènes actifs dans la formation des spermatozoïdes chez les adultes, mais que sa seule présence ne suffit pas à les produire. Normalement, les drosophiles ont un ou deux chromosomes X et deux ensembles d'autosomes. S'il n'y a qu'un seul chromosome X dans une cellule diploïde (1X : 2A), la mouche est de sexe masculin. S'il y a deux chromosomes X dans une cellule diploïde (2X : 2A), la mouche est de sexe féminin. Ainsi, les drosophiles XY sont des mâles fertiles et les XO sont des mâles stériles (Hake & O'Connor, 2008).

La détermination sexuelle peut également se faire de façon génétique sans passer par des chromosomes sexuels spécialisés (Manolakou et al., 2006). C'est le cas chez quelques insectes Hymenoptères, par exemple, qui présentent un mode de déterminisme sexuel régulé par le niveau de ploïdie et d'hétérozygotie. Un œuf non-fertilisé induira le développement de l'individu en mâle haploïde, et un œuf fertilisé en femelle diploïde (Beukeboom & Perrin, 2014). C'est la présence de certains allèles pour un unique locus, ls-CSD (« locus single-locus complementary sex determination »), qui sera responsable de déterminer le sexe, en fonction de l'hétérozygotie de ce gène ; un hemi-zygote mène à la formation d'un mâle, et un hétérozygote à celle d'une femelle (Beukeboom et al., 2000 ; Beukeboom & Van De Zande, 2010 ; Beye et al., 2003 ; Sanchez, 2008). L'homozygotie chez les individus diploïdes mène à la formation de mâles généralement stériles et avec une viabilité réduite (Wilgenburg et al., 2006).

1.1.2 Détermination épigénétique

À l'opposé de la détermination génétique du sexe, lorsque le génome est identique entre les mâles et les femelles, il existe des facteurs extérieurs à ce génome, qui lui permettront de produire autant un sexe que l'autre. Parmi ces facteurs, on retrouve principalement l'effet de l'environnement, qui joue un rôle dans la détermination sexuelle chez un grand nombre d'espèces animales, tels que la température, la nourriture, certains composés chimiques, les phéromones, la photopériode, etc

(Beukeboom & Perrin, 2014). Jusqu'à présent, le facteur le plus étudié est la température. Son influence sur la détermination du sexe a été initialement découverte chez les lézards (Charnier, 1966) et les tortues (Pieau, 1971), pour finalement se retrouver être le grand responsable de la détermination du sexe chez de nombreux autres reptiles (Lang & Andrews, 1994 ; Mitchell et al., 2006 ; Warner & Shine, 2005) et poissons (Ospina-Alvarez & Piferrer, 2008). Dans ces situations, le sexe de l'embryon sera défini selon la température à laquelle il se développera. Une étape développementale dite thermosensible, qui a généralement lieu tôt dans l'embryogénèse ou même au stade d'œuf (Pieau & Dorizzi, 1981), représente l'étape cruciale dans la détermination du sexe. Le sexe de l'embryon sera dicté par les conditions thermiques de son habitat auxquelles il sera confronté lors de cette étape-clé de son développement (Perrin & Beukeboom, 2014). Il peut exister un (certains lézards et les Sphenodon) ou deux (certaines tortues et crocodiles) seuils de températures qui induisent une masculinisation ou une féminisation (Pieau, 1995). Par exemple, les œufs de l'alligator américain, *Trachemys scripta*, vont produire des mâles tant que la température externe ne dépassera pas les 30 degrés pendant le stade de vie critique propice à la différentiation et au développement des gonades chez cette espèce. Dépassé cette température, les œufs incubés donneront naissance à des femelles. Pour l'alligator américain, *Alligator mississippiensis*, il existe deux seuils critiques. Toutes les températures plus basses que 30 ou plus hautes que 35 degrés Celsius induiront la production de mâles, tandis que les températures intermédiaires entre les deux permettront la production de femelles (Gilbert, 2000).

La détermination du sexe thermo-dépendante est induite par une enzyme, l'aromatase, qui permet la transformation chimique de l'hormone sexuelle mâle, la testostérone, en hormone sexuelle femelle, l'oestrogène (Pieau & Dorizzi, 2004). Son activité va dépendre de la température afin d'induire la production des hormones nécessaires au développement d'un sexe ou de l'autre (Desvages et al., 1993). La régulation de l'expression de l'aromatase peut parfois être induite par des événements de méthylation, un processus épigénétique qui permet de réguler l'expression des gènes dans chaque cellule. En effet, la température agit sur la méthylation du promoteur de l'aromatase ou encore de certains promoteurs de gènes impliqués dans la gamétogenèse, comme *SOX9* chez certains alligators (Parrott et al., 2014), et va ainsi permettre une différence dans leurs taux d'expression et, conséquemment, une différenciation des gonades (Navarro-Martin et al., 2011). Par exemple, lorsque l'aromatase est fortement exprimée, la testostérone est transformée en oestrogène et l'individu devient femelle, et vice versa.

La détermination épigénétique du sexe implique donc des facteurs extérieurs au génome, c'est l'expression différentielle des gènes issus d'un même génome qui est responsable du déterminisme sexuel. Outre les facteurs environnementaux, des facteurs sociaux, parasites ou encore maternels peuvent être impliqués dans la détermination épigénétique du sexe. Par exemple, un des facteurs maternels responsable d'un système de détermination du sexe particulièrement bien étudié chez les insectes, plus précisément chez l'ordre des isopodes comprenant les crustacés terrestres, est la présence d'une endobactéries symbiotique, *Wolbachia*. Chez le coléoptère *Armadillidium vulgare*, dont la détermination du sexe est initialement dépendante des chromosomes sexuels ZW (Cordaux & Gilbert, 2017), *Wolbachia* est capable de bloquer la production d'une hormone androgénique afin de forcer le développement en femelles chez les individus que le portent, peu importe les chromosomes sexuels présents. Puisque ces bactéries envahissent le cytoplasme des œufs de leur hôte, nous sommes faces à un contenu cytoplasmique qui prime sur le contenu chromosomique en matière de différentiation sexuelle, ce qui signifie que la détermination du sexe chez les hôtes, porteurs de cette bactéries, a été substituée pour un déterminisme cytoplasmique du sexe (Perlman et al., 2015; Cordaux & Gilbert, 2017). Presque la moitié des insectes arthropodes sont infectés par cette bactéries et, par conséquent, ont subi grand nombre de transferts génétiques horizontaux de sa part (Zug et al., 2012 ; Weinert et al., 2015). Bien que la plupart de ces transferts semblent neutres pour l'hôte qui les porte, un facteur s'est fait remarquer chez *Armadillidium vulgare* qui vient rajouter une diversité supplémentaire aux systèmes sexuels présents. Il s'agit d'un agent féminisant, nommé *f*, qui féminise lui aussi la descendance et qui provient du génome de *Wolbachia* (Legrand & Juchault, 1984 ; Leclercq et al., 2016). La présence d'un autre facteur est soupçonnée chez les mâles, le facteur *M*, qui permettrait de contrer le facteur *f* afin de restaurer le phénotype mâle. L'effet de ce facteur domine celle du facteur *f* et du chromosome W. En revanche la présence féminisante de *Wolbachia* prend le dessus sur le facteur M (Juchault & Legrand, 1976 ; Rigaud & Juchault, 1993). Ces systèmes ont parfois donné naissance à des individus intersexes, plus communément à cause de *Wolbachia* (Legrand & Juchault, 1969). Globalement, puisque ces bactéries envahissent le cytoplasme des œufs de leur hôte, couplé au transfert horizontal du facteur *f*, elles altèrent de manière significative le sex-ratio en le biaisant vers les femelles afin de leur assurer une transmission quasi systématique (Cordaux & Gilbert, 2017).

Les différents facteurs impliqués dans la détermination épigénétique du sexe peuvent conduire à la formation d'espèces gonochoriques (deux sexes séparés) ou hermaphrodites, pour lesquels la différenciation des gonades mâles et femelles se fait au sein du même individu. Si cette différenciation se produit au même moment, on parle d'hermaphrodisme simultané, qui est probablement favorisé dans un environnement qui limite la possibilité de trouver un partenaire (Eppley & Jesson, 2008). Si cette différenciation est séparée dans le temps (par exemple ; un individu qui commence sa vie en tant que mâle puis, après un stimuli extérieur défini, la termine en tant que femelle), on va parler d'hermaphrodisme séquentiel (Kroon et al., 2005). Les causes évolutives de ces systèmes peuvent dépendre de la probabilité de trouver un partenaire sexuel, la sessilité des individus, la densité des populations, l'avantage conféré à un sexe selon les disponibilités en nourriture, la photopériode, le nombre de partenaires, certaines phéromones (Beukeboom & Perrin, 2014), ou encore même la taille ou l'âge des individus (Zimmerman, 1991). En fonction de ces conditions, si cela confère un véritable avantage à la valeur adaptative de l'individu, et/ou de l'espèce, étant donné son lourd coût physiologique (Policansky, 1982 ; Warner, 1984), un individu peut même changer plusieurs fois de sexe (Perrin & Beukeboom, 2014).

1.1.3 Modèles mixtes et diversité

Les catégories décrites précédemment, représentent chacune l'extrême d'un continuum dans lequel grand nombre de systèmes de déterminisme du sexe trouvés dans la nature sont issus d'un mixte entre déterminisme purement génétique et déterminisme purement épigénétique. La grenouille *Rana temporaria*, illustre bien ce cas de figure puisqu'elle possède un système de chromosomes sexuels XX/XY mais est également sensible à l'effet de la température. Il arrive donc que des individus XY développent des ovaires et des individus XX des testicules résultant en une divergence entre le sexe génotypique et le sexe phénotypique (Dournon et al., 1990 ; Matsuba et al., 2008). Ces évènements sont connus sous le terme « d'inversion sexuelle », traduit de l'anglais « sex reversal ». En 2014, Ashman et collègues ont créé trois banques de données qui regroupent le maximum d'informations collectées sur les systèmes sexuels décrits dans la littérature pour les plantes (11 038 données à partir de 77 clades d'angiospermes), les vertébrés (2 145 données dont 705 poissons, 173 amphibiens, 593 reptiles, 195 oiseaux et 479 mammifères) et les invertébrés (11 566 données à partir de 45 archives) (Ashman et al., 2014). Cette dernière catégorie contient uniquement des informations concernant les groupes taxonomiques des arthropodes et ne contient aucune information à propos des autres groupes d'invertébrés. Les auteurs de cette banque de données se sont concentrés sur l'ensemble de la littérature faite à ce jour pour les différents groupes taxonomiques, en focalisant principalement sur les études actualisées et vérifiées, et sur les espèces les plus étudiées jusqu'à maintenant. Il y a donc beaucoup d'espèces, tant végétales qu'animales, pour lesquelles le déterminisme du sexe est largement méconnu ou reste encore entièrement à découvrir. Ces informations sont cruciales afin de dresser un tableau plus complet de l'ensemble des systèmes de déterminisme sexuel et pour mieux comprendre les forces évolutives responsables de leur immense diversité en nature (The Tree of Sex Consortium, 2014). Chez les invertébrés en particulier, un groupe est totalement absent de l'étude de Ashman et collègues (2014). Il s'agit des mollusques, et c'est ce groupe qui va nous intéresser pour le présent projet de recherche, qui constituent le deuxième groupe taxonomique d'invertébrés le plus diversifié après celui des insectes (Myers et al., 2015). Malgré ceci, et malgré l'existence de plusieurs espèces d'importance économique au sein de ce groupe animal, les connaissances sur les facteurs responsables de la détermination du sexe chez les mollusques sont peu nombreuses.

1.1.4 Détermination du sexe chez les Mollusques

Parmi les quelques études sur la détermination du sexe chez les mollusques, il est principalement mention de déterminisme purement génétique ou de déterminisme purement épigénétique. Chez certaines espèces de Gastéropodes, par exemple, la détermination du sexe est définie par des chromosomes sexuels dont la diversité d'une espèce à l'autre est aussi importante qu'entre les mammifères, les oiseaux et les insectes, puisqu'on y retrouve trois systèmes de chromosome sexuels (XX/XY, ZZ/ZW et XX/XY/X0) (Ahmed & Sparks, 1967 ; Thiriot-Quiévreux, 2003), parmi la panoplie complexe de systèmes de chromosomes sexuels existants (Blackmon et al., 2017). Toujours chez les Gastéropodes, le sexe peut aussi être déterminé par la température ou encore par des facteurs sociaux (Collin, 2007), comme chez *Crepidula fornicata* dont le sexe est déterminé selon celui des individus voisins au sein de la colonie. La larve commencera généralement sa vie en tant que mâle, pour une brève période, puis se développera en femelle, au contact des nouveaux individus venus rejoindre la colonie, afin d'être fécondée à son tour. Ce stade de développement induit la production des organes féminins et la régression du pénis. Le dernier individu arrivé peut garder son rôle de mâle jusqu'à 6 ans, probablement en fonction des phéromones produites par les femelles environnantes (Collin, 1995 ; Fretter & Graham, 1981).

L'effet d'autres facteurs environnementaux tels que la pollution ou les hormones sexuelles sur le développement des gonades a également été observé chez certains Gastéropodes ainsi que chez certains Bivalves (Wang & Croll, 2004). Il existe aussi, chez ces deux mêmes groupes, de l'hermaphrodisme simultané et séquentiel (Heller, 1993). Les mécanismes de détermination du sexe chez les autres groupes de mollusques, composés majoritairement d'espèces gonochoriques quant à eux (i.e., Scaphopodes, Céphalopodes, Monoplacophores, Polyplacophores, Caudofovéates, et les Solenogastres qui eux sont hermaphrodites), semblent peu connus (Collin, 2007).

1.1.5 Détermination du sexe chez les Bivalves

Estimé à 20,000-30,000 espèces vivantes (www.bivatol.org), le groupe des Bivalves constitue le second groupe le plus diversifié chez les mollusques, après les Gastéropodes. Plusieurs espèces de Bivalves, telles que l'huître, le pétoncle, les palourdes et les moules, sont économiquement importantes. Très peu d'études ont été effectuées concernant les mécanismes de détermination du

sex chez ce groupe animal. La plupart des systèmes sexuels recensés chez ce groupe taxonomique englobe des facteurs épigénétiques (Breton et al., 2018 pour une revue de littérature – ce chapitre de livre, pour lequel je suis co-auteure, est en annexe). Les espèces d'eau douce sont sujettes plus fréquemment à l'hermaphrodisme (Morton, 1991) bien que l'hermaphrodisme ait également été rapporté dans certains groupes marins (Collin, 2013). La découverte d'individus hermaphrodites dans des espèces supposément gonochoriques doit néanmoins être traitée avec précaution car les éléments chimiques liés à la pollution sont connus pour interagir avec le développement sexuel de ces espèces aquatiques et fausser le sexage (Ciocan, 2012). Parmi les espèces hermaphrodites, on retrouve des hermaphrodites simultanés, notamment chez beaucoup d'espèces de pétoncles (famille Pectinidae ; Shumway & Parsons, 2011). Des cas d'hermaphrodisme séquentiel ont aussi été révélés, tant protandriques (l'individu débute sa vie en exprimant son phénotype mâle, puis change pour celui de la femelle) que protogyniques (l'individu exprime son phénotype femelle avant de changer pour un phénotype mâle ; Oldfield, 1961). Le premier est illustré par le pétoncle *Chlamys varia* (Lubet, 1959 ; Lucas, 1965) dont la transition de mâle à femelle est corrélée avec l'âge (Burnell, 1995). En effet, le changement aurait lieu lors de la deuxième saison de croissance, probablement stimulé par l'accumulation de produits énergétiques pendant l'hiver, puisque la production de femelles, et donc d'œufs, est plus demandante énergétiquement que celle des spermatozoïdes. D'autres cas associent ce changement mâle-vers-femelle à la taille (Ghiselin, 1969 ; Charnov, 1982 ; Policansky, 1982 ; Breton et al., 2018). Même si la protogynie est plus rare, elle a été décrite chez certaines espèces, notamment du genre *Pecten*, où le changement femelle-vers-mâle découle de l'absence de ressources nutritionnelles (Coe, 1943). Ces exemples sont illustrés et expliqués dans la figure 1.

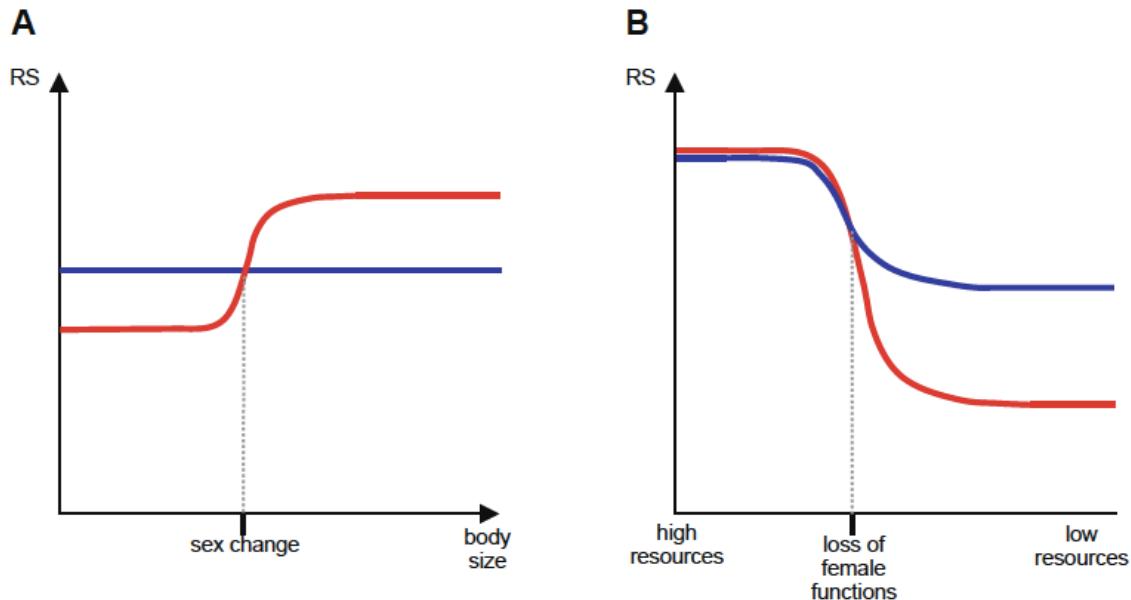


Figure 1. Modèles théoriques de changements de sexe chez les bivalves, basé sur plusieurs études de changements de sexe, figure issue de Breton et al. (2018). RS : Succès reproducteur. La ligne bleue représente le RS chez les mâles et la ligne rouge celui des femelles. **(a)** Modèle avantageant la taille selon Policansky (1982). Le RS des mâles est indépendant de la taille du corps, tandis qu'il avantage proportionnellement celui des femelles, dû au coût de production des spermatozoïdes qui est moindre comparé à celui des œufs. Après un certain seuil de taille, les femelles ont un RS supérieur à celui des mâles puisqu'elles peuvent produire des œufs plus efficacement que celles plus petites. Puisque le RS des mâles n'a pas été évalué avec l'augmentation de la taille, son RS est dépeint sous forme de plateau pour se concentrer sur le changement de RS lié au changement de sexe. **(b)** Le second modèle illustre le RS lié à la disponibilité en nourriture. Le RS des mâles et des femelles est proportionnel à cette disponibilité et à leur capacité à la consommer. Lorsque les ressources diminuent, le RS des deux sexes diminue, mais celui des femelles chute plus rapidement que celui des mâles puisqu'un œuf demande plus d'énergie de production qu'un spermatozoïde. En-dessous d'un certain seuil d'accèsibilité de nourriture, les femelles ont un RS inférieur à celui des mâles, ce qui favorise un changement de sexe de femelle-vers-mâle pour maintenir un meilleur RS chez l'individu.

Par ailleurs, plusieurs facteurs environnementaux sont aussi responsables de la détermination du sexe chez les bivalves, tels que les stéroïdes exogènes, la température, la disponibilité en ressource et les polluants (tous décrits dans Breton et al., 2018).

Concernant la détermination génétique, il semble que toutes les espèces de bivalves étudiées jusqu'à maintenant, à l'exception de deux mais ceci reste à être confirmé en microscopie, ne détiennent pas de paire de chromosome sexuels (Borsig & Thiriot-Quiévreux, 1990 ; Breton et al., 2018; Giese, 2012). Malgré cette absence de chromosomes spécialisés, plusieurs facteurs génétiques ont été identifiés comme étant impliqués dans la détermination du sexe (Breton et al., 2018). Par exemple, l'analyses transcriptomique des gonades de plusieurs espèces de bivalves a permis de mettre en lumière les facteurs impliqués dans la différentiation des sexes mâles et

femelles. Les mécanismes génétiques de la détermination du sexe ont été étudiés plus en détail chez les huîtres que chez tout autre groupe de bivalves. Parmi les gènes connus chez les animaux en général pour agir sur la détermination du sexe et / ou au début de la différenciation gonadique, il existe *Oyvlg* (*Crassostrea gigas* ; Fabiou et al., 2004, 2009), un homologue de *vasa*, qui est un gène impliqué dans le développement des cellules germinales primordiales et la différenciation sexuelle précoce chez les eucaryotes ; *Cg-DMI* (*C. gigas* ; Naimi et al., 2009a, b), un homologue de *Dmrt1* discuté plus haut ; *Cg-SoxE*, un homologue de *Sox9* qui est exprimé lorsque le sexe n'est pas encore distingué chez *C. gigas* (Santerre et al., 2014) ; ainsi que *CgFoxL2*, un gène important dans le déterminisme sexuel femelle, *CgSoxH* (ou *Sry-like*) et *CgDsx* (*C. gigas* ; Zhang et al., 2014), ces deux derniers gènes étant liés au déclenchement du développement masculin.

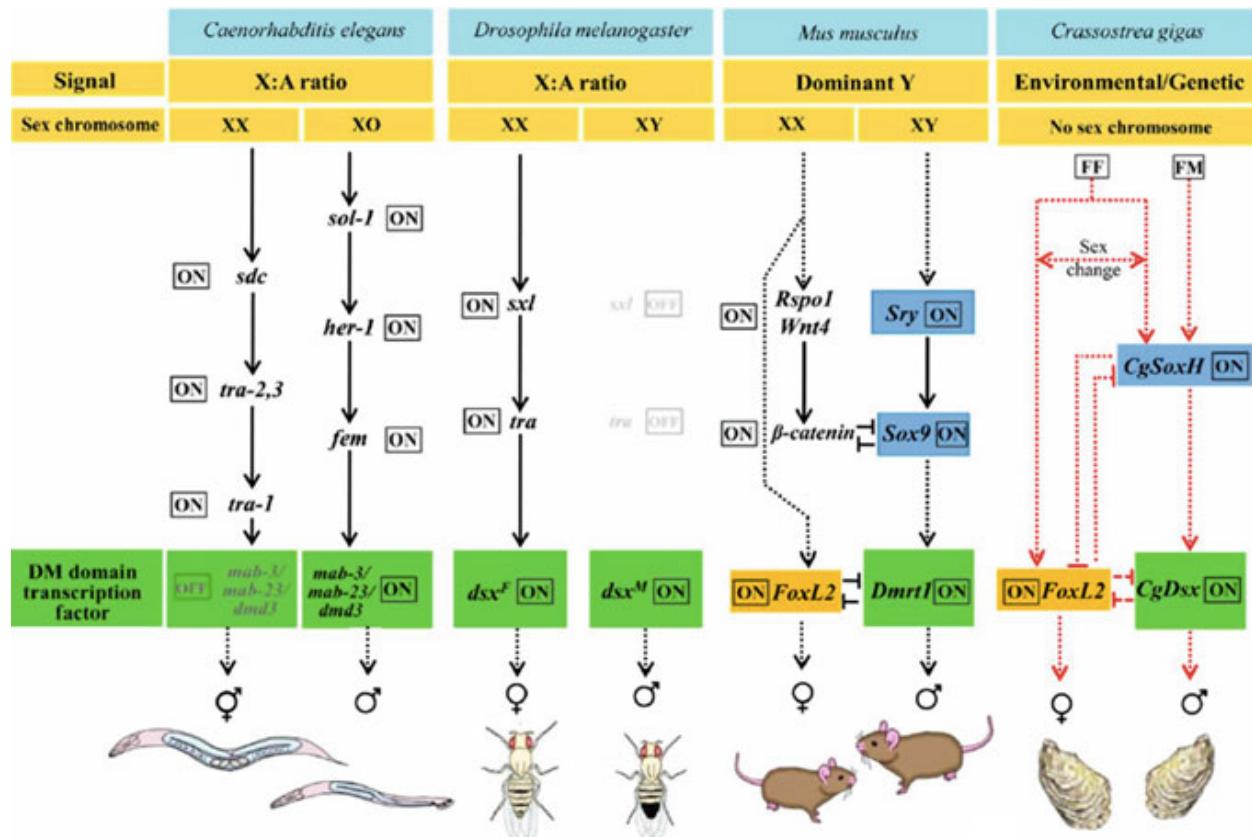


Figure 2. Modèle de détermination du sexe chez l'huître Pacifique *Crassostrea gigas*. Les mécanismes sous-jacents à la détermination du sexe sont comparés entre organismes modèles selon le résumé de Gamble & Zarkower (2012). Seuls les gènes clés de régulation spécifiques au sexe sont montrés. Les traits noirs indiquent les relations temporelles, et les pointillés indiquent les relations hypothétiques basées sur des données d'expression. Cette figure est issue de Breton et al. (2018), elle-même modifiée de Zhang et al. (2014). Les génotypes FF permettent un changement de sexe (Guo et al. 1998).

À la lumière de ces résultats, Zhang et al. (2014) ont proposé un modèle pour la détermination du sexe chez l'huître creuse *Crassostrea gigas* (Fig. 2). Cette espèce se caractérise par un hermaphrodisme protandrique, et l'initiation de la détermination du sexe impliquerait à la fois des facteurs génétiques et environnementaux (Zhang et al., 2014). En se basant sur l'homologie de séquences et les fonctions déduites des données transcriptomiques, Zhang et al. (2014) ont émis l'hypothèse que *CgSoxH*, un gène de type *Sry* strictement exprimé dans les gonades mâles, pourrait jouer un rôle de premier plan dans la voie de détermination du sexe chez *C. gigas*. *CgSoxH* activerait directement ou indirectement *CgDsx*, un gène à domaine DM comme par exemple *Dmrt1* qui a été identifié comme commutateur maître pour le développement testiculaire dans tous les métazoaires étudiés jusqu'à présent. *CgSoxH* et *CgDsx* interagiraient avec, ou inhiberaient, *CgFoxL2*, qui est habituellement spécifiquement exprimé dans les ovaires. L'interaction possible entre ces gènes favorisant les mâles et les femelles est compatible avec l'interaction rapportée entre *Sry*, *Sox9*, *Dmrt1* et *FoxL2* chez les mammifères (Veitia, 2010; Zhang et al., 2014) (Fig. 2). Ce modèle fourni par Zhang et al. (2014) encouragera certainement de nouvelles recherches sur les voies de détermination du sexe chez les mollusques et autres invertébrés.

Dans le cadre de mon projet de doctorat, je m'intéresse à un système de détermination du sexe hypothétique chez les bivalves qui impliquerait, pour la première fois dans le monde animal, les mitochondries. En effet, chez certaines espèces de bivalves, il existe un mode de transmission des mitochondries particulier qui suggère fortement un lien avec la détermination du sexe (Breton et al., 2011). Démystifier ce lien est l'objectif principal de ma thèse.

1.2 Mitochondries

Les mitochondries sont des organelles présentes dans presque toutes les cellules animales et végétales, et possèdent leur propre génome, l'ADN mitochondrial (ADNmt). Depuis son séquençage complet pour la première fois chez l'humain (Anderson et al., 1981), cet ADN est décrit, chez les espèces animales, de forme circulaire d'environ 16kb et contenant 37 gènes. Treize d'entre eux codent pour des protéines qui vont participer à la formation des complexes enzymatiques du système de phosphorylation oxydative qui est responsable de produire l'énergie cellulaire. Les 24 gènes restants codent pour des ARN de transferts (22 ARNt) et des ARN ribosomaux (2 ARNr) nécessaires à la traduction des 13 protéines (Scheffler, 2007). Les protéines

codées par le génome mitochondrial interagissent avec celles issues du génome nucléaire résultant d'une co-évolution entre les deux génomes afin d'assumer les fonctions mitochondrielles (Bar-Yaacov et al., 2012).

Pendant plusieurs décennies, une participation à la production d'énergie était la seule fonction que l'on attribuait aux protéines codées par l'ADN mitochondrial. Cependant, grâce au développement des domaines -omiques, plusieurs études ont permis de constater que la fonction et le contenu génétique des mitochondries est bien moins fixe qu'initialement prétendu. En effet, de nouveaux gènes ont été découverts dans le génome mitochondrial de certaines méduses (Kayal et al., 2012), et même chez l'humain [Humanin et Gau (Capt et al., 2016 – cet article est en annexe) ; et MOTSc (Lee et al., 2015)]. En plus du contenu, les fonctions assumées par le génome mitochondrial et les mitochondries ont aussi été revues et il a été démontré qu'outre la production d'ATP, les mitochondries sont impliquées dans plusieurs processus tels que la thermogenèse, le transport de calcium, la signalisation cellulaire, l'apoptose ou encore le vieillissement (Brookes et al., 2002 ; Chabi et al., 2008 ; Pollack & Leeuwenburgh, 2001), et les produits du génome mitochondrial prennent part à certains de ces processus (ex. Humanin est une protéine anti-apoptotique ; Capt et al., 2016).

Un autre critère propre aux mitochondries, est leur mode de transmission. En effet, malgré quelques rares exceptions, principalement chez les champignons et les plantes, les mitochondries sont transmises par un seul parent, qui est généralement la mère (voir Breton & Stewart, 2015 pour une revue). On parle alors de transmission strictement maternelle ou SMI (Strict Maternal Inheritance) (Hayashi et al., 1978). Ce mode de transmission permet de maintenir un état homoplasmique, c'est-à-dire des copies de l'ADNmt génétiquement identiques dans toutes les cellules (Clayton, 1991). Cette homoplasmie aurait hypothétiquement évolué afin de maintenir des interactions mito-nucléaires optimales pour favoriser les fonctions des mitochondries (Bar-Yaacov et al., 2012).

1.2.1 DUI (*Doubly Uniparental Inheritance ou Double Transmission Uniparentale*)

Pourtant, en 1994, une exception à cette transmission strictement maternelle a été décrite chez un groupe de bivalves ; les moules marines du genre *Mytilus* (Zouros et al., 1994). Ce système unique de transmission mitochondriale est caractérisé par la présence de deux génomes mitochondriaux : un qui est transmis de façon maternelle (ADNmt F) aux femelles et aux mâles, et l'autre transmis de façon paternelle (ADNmt M) aux mâles seulement (Breton et al., 2007 ; Passamonti & Ghiselli, 2009). L'ADNmt F est donc retrouvé dans les tissus somatiques des mâles et des femelles et dans les tissus gonadiques des femelles, tandis que l'ADNmt M est seulement retrouvé dans les spermatozoïdes des mâles (Breton et al., 2007 ; Passamonti & Ghiselli, 2009). On parle alors de double transmission uniparentale (DUI, de l'anglais « Double Uniparental Inheritance »). Ce système garde un état d'homoplasmie chez les femelles, tandis qu'il induit un état d'hétéroplasmie chez les mâles avec le génome mitochondrial paternel dans les spermatozoïdes, et le génome mitochondrial issu des œufs dans le soma (Fig. 3). La divergence de séquence entre ces deux génomes est très grande et peut atteindre jusqu'à 52% chez certaines moules d'eau douce (Doucet-Beaupré et al., 2010).

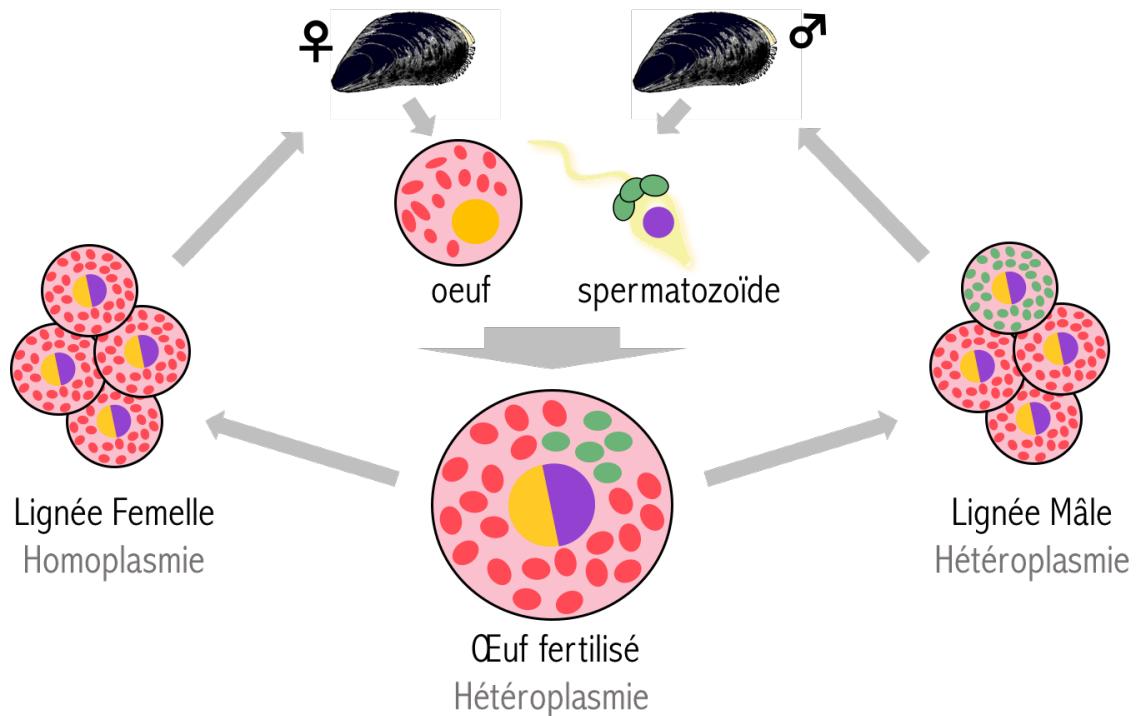


Figure 3. Modèle de double transmission uniparentale chez les bivalves. Les mitochondries maternelles sont illustrées par des ronds rouges, tandis que les ronds bleus représentent les mitochondries paternelles. Lors de la fertilisation, l'embryon comprend un mélange des mitochondries maternelles (F) et paternelles (M). Si l'embryon est destiné à devenir une femelle, les mitochondries paternelles seront dégradées pour ne garder que celles de la mère dans tous les tissus de l'individu. À contrario, si l'embryon est destiné à devenir un mâle, les mitochondries paternelles vont s'agrégner et migrer dans la cellule qui donnera la lignée germinale donnant ainsi un individu avec des tissus somatiques colonisés par les mitochondries F et des tissus sexuels par les mitochondries M.

1.2.1.1 Transmission paternelle des mitochondries chez les bivalves

Le phénomène de double transmission uniparentale de l'ADNmt chez les bivalves est illustré à la Figure 3. Lors de la fertilisation chez les bivalves avec la DUI, le spermatozoïde transmet seulement 5 mitochondries à l'embryon contre des centaines de mitochondries dans l'oeuf contenant des milliers de copies du génome mitochondrial maternel (Kyriakou et al., 2010). Il a été observé que cette faible quantité d'ADNmt M est capable de se maintenir lors du développement de l'embryon mâle, grâce à un phénomène d'agrégation-dispersion (Cao et al., 2004 ; Milani et al., 2012). Chez les mâles, ces 5 mitochondries s'assemblent et se maintiennent dans la même cellule du blastomère pour former la future lignée germinale de l'individu. Chez les femelles, il a été proposé soit une élimination de ces mitochondries issues du sperme par des processus d'ubiquitination, comme c'est le cas chez les mammifères (Sutovsky et al., 2000), soit une

dispersion de celles-ci et donc une dilution importante de l'ADNmt M présent dans les différents tissus rendue possible par la quantité nettement supérieure de l'ADNmt F (Cao et al., 2004 ; Milani et al., 2012). Une modification du mécanisme d'ubiquitination pourrait être la cause de la rétention des mitochondries paternelles chez les embryons mâles (Kenchington et al., 2002). Fait intéressant, il a été démontré chez la drosophile que la machinerie d'ubiquitination peut participer à la régulation du dimorphisme sexuel (Bayerer et al., 2005).

1.2.1.2 Distribution taxonomique de la DUI : similarités et différences entre les groupes

En tant qu'exception majeure à la SMI, la question de l'origine de la DUI présente un grand intérêt dans le domaine de la biologie mitochondriale, ce phénomène complexe entretenant de nombreux liens avec d'autres questions biologiques clés, telles que la détermination du sexe (Breton et al., 2011, 2014; Passamonti et Ghiselli, 2009; Zouros, 2013) et les conflits génomiques (Milani et al., 2016 ; Passamonti et Ghiselli, 2009). Le système DUI a été recensé chez plus d'une centaine d'espèces de bivalves, réparties dans 4 ordres et douze familles (toutes résumées dans Gusman et al., 2016). Comme on retrouve la DUI chez des espèces très éloignées, un débat portant sur son origine est toujours d'actualité. Une origine unique, qui serait survenue lors de la radiation d'Eulamellibranchiata (ou proche de l'origine de la classe moderne des Bivalvia; Boyle & Etter, 2013) suivie de sa perte dans certaines lignées, a longtemps semblé l'hypothèse la plus parcimonieuse (Boyle & Etter, 2013; Doucet-Beaupré et al., 2010 ; Hoeh et al., 1997; Gusman et al., 2016). Zouros (2013) a fait valoir, dans l'hypothèse où une telle perte impliquerait la perte de la lignée de l'ADNmt M, que soit les ADNmt F devraient se regrouper avec des ADNmt d'espèces non-DUI, ce qui n'est pas le cas, soit un événement de masculinisation serait requis à la base de chaque clade DUI, ce qui ne semble pas l'explication la plus parcimonieuse (Plazzi et al., 2019). En fait, il existe de nettes différences entre les systèmes de DUI, et les ADNmt liés au sexe présentent des caractéristiques différentes selon les espèces (Plazzi, 2015; Zouros, 2013). De nouvelles espèces à DUI continuent d'être découvertes parmi les bivalves, et leur distribution demeure éparsillée (Gusman et al., 2016; Plazzi, 2015). Cependant, la détection de la DUI étant particulièrement sujette aux faux négatifs, de nombreuses espèces à DUI ont peut-être été négligées (Theologidis et al., 2008; Zouros, 2013). Récemment, plusieurs preuves suggèrent une origine multiple de la DUI (Milani et al., 2013; Milani et al., 2016; Zouros, 2013). Milani et collègues (2013, 2016) ont suggéré que des événements d'endogénisation de gènes viraux dans

les génomes mitochondriaux pourraient être liés à l'origine de la DUI, un mécanisme qui expliquerait la distribution dispersée de la DUI parmi les bivalves. Pour maintenir la DUI, une machinerie cellulaire complexe est nécessaire et, bien que des similarités évidentes soient retrouvées entre espèces à DUI, elles présentent également beaucoup de différences, ce qui génère une difficulté à instaurer une généralité dans ce système. Parmi ces différences, on peut citer des dissimilarités au niveau de la distribution et l'expression des haplotypes mitochondriaux au sein des différents types tissulaires. De manière générale, les femelles sont homoplasmiques pour l'ADNmt F et les mâles hétéoplasmiques (Breton et al., 2007 ; Passamonti & Ghiselli, 2009 ; Zouros, 2013 pour des revues). Il existe pourtant des cas où des traces du génome mitochondrial M ont été détectées dans les tissus somatiques mâles et femelles chez les moules marines Mytilidés et les palourdes marines Vénéridés (Dalziel & Stewart, 2002 ; Passamonti & Scali, 2001). Quelques hypothèses pour expliquer ces résultats ont été proposées (Batista et al., 2010 ; Cao et al., 2004 ; Garrido-Ramos et al., 1998 ; Kyriakou et al., 2010 ; Obata et al., 2006; 2011), notamment un défaut dans le mécanisme d'agrégation des mitochondries M lors de la formation de la lignée germinale chez les mâles ou encore un défaut dans le mécanisme d'élimination des mitochondries M chez les femelles, mais toutes ces hypothèses restent encore à être vérifiées. La présence de ces deux génomes a également été investiguée dans les tissus gamétiques. Le génome F est absent des spermatozoïdes chez les Mytiloida (Venetis et al., 2006) et chez les Veneroida (Ghiselli et al., 2011). Ghiselli et collègues (2011) ont d'ailleurs également prouvé qu'il n'y avait pas trace du génome M dans les œufs chez les Veneroida. Les tissus gamétiques des femelles chez les Mytiloida, en revanche, semblent subir une régulation moins stricte sur la présence de l'haplotype mitochondrial M (Dalziel & Stewart, 2002). La distribution des haplotypes mitochondriaux mâles et femelles diffère en deux autres points, entre les Mytiloida et les Veneroida ; premièrement la présence du génome M dans les tissus somatiques des mâles dépasse largement celle du génome F chez les Veneroida, contrairement aux Mytiloida, et deuxièmement, le génome M est présent en beaucoup plus faible quantité dans les tissus somatiques des femelles chez les Veneroida par rapport aux Mytiloida (Ghiselli et al., 2011). Ces différences relèvent un mécanisme actif responsable de la distribution au sein des tissus somatiques. La répartition tissulaire des génomes F et M a moins été étudiée chez les moules d'eau douce, mais elle a été examinée chez trois espèces ; *Unio crassus* (Mioduchowska et al., 2016), *Utterbackia peninsularis* et *Venustaconcha ellipsiformis* (Breton et al., 2017). Malgré une

discordance entre les résultats, la deuxième étude démontre une présence et expression de l'ADNmt M dans tous les tissus somatiques et gonadiques, à des degrés variables, et dont la répartition est plus sporadique dans les tissus des femelles, ce qui rejoint les résultats rapportés chez *Mytilus* spp. et *Ruditapes philippinarum* (Breton et al., 2017). Une attribution inégale de la présence des génomes entre sexes et tissus est intuitivement liée à leurs tâches. En effet, dans le système DUI, les génomes mitochondriaux des femelles fonctionnent dans les gonades femelles et les tissus somatiques des deux sexes, tandis que les génomes mitochondriaux paternels semblent jouer un rôle uniquement dans les gonades mâles et, à moindre mesure, dans le soma des mâles (Stewart et al., 1996 ; Breton et al., 2017), ce qui conduit à une division assymétrique du travail. Une activité moindre est liée à une pression de sélection plus souple pour le génome paternel : en effet, celui-ci évolue plus vite que l'ADNmt maternel (Stewart et al., 1996 ; Zouros, 2013).

Le séquençage des génomes F et M complets chez plusieurs espèces de bivalves a permis d'établir une liste d'éléments similaires et divergents interspécifiques, mais aussi intra-spécifiques. Un premier élément notable est la découverte de nouveaux gènes dans les ADNmt M et F (Breton et al., 2014; Plazzi et al., 2016), qui semblent corrélés au phénomène de la DUI (Breton et al., 2009, 2011), et qui pourraient provenir d'événements de duplication de gènes (Mitchell et al., 2016), plutôt que de transferts horizontaux de gènes viraux. Cependant, les similarités de séquence ou de structure sont rarement retrouvées entre les ORFans d'espèces de bivalves, même entre espèces apparentées (Milani et al., 2013; Mitchell et al., 2016; Plazzi et al., 2016). Cette particularité est décrite de manière plus détaillée dans la section 2.1.4. Un deuxième élément remarquable concerne le gène *cox2* du génome M, qui contient une extension ou une insertion importante chez plusieurs espèces DUI étudiées jusqu'à présent (seules les espèces du genre *Mytilus* ne possèdent pas de modification du gène *cox2* : Bettinazzi et al., 2016). Chez les espèces de moule d'eau douce (ordre Unionoida) par exemple, l'extension du gène *cox2* se trouve à l'extrémité 3' et sa taille est approximée à 550 pb (Doucet-Beaupré et al., 2010). Sa transcription et traduction ont été prouvées, et les protéines issues de ces processus ont été localisées tant dans la membrane mitochondriale interne qu'externe (Chakrabarti et al., 2007, 2006), lui prêtant ainsi un rôle dans le marquage mitochondrial qui permettrait aux mitochondries paternelles d'échapper à leur élimination dans les embryons mâles (Chapman et al., 2008). Chez les vénéridés, une insertion d'environ 300

nucléotides dans le gène *cox2* du génome M a été rapportée chez l'espèce *Meretrix lamarckii* (Bettinazzi et al., 2016). En ce qui concerne les protéines COX2 codées par le génome F, elles ont toutes été décrites de taille similaire aux autres animaux à l'exception de *Ruditapes philippinarum*, de la famille des vénéridés, qui possède une copie dupliquée du gènes *cox2*, à l'extrémité 3' (Passamonti et al., 2011). Une autre particularité réside chez *Venerupis philippinarum*, dont le génome nucléaire est régulé par des petites séquences d'ARN non-codantes, codées par le génome mitochondrial, qui pourraient être connectées à sa détermination du sexe et au maintien de la DUI (Pozzi et al., 2017). La découverte de séquences spécifiques retrouvées dans la région contrôle des mytilidés qui ont mené à des événements de masculinisation chez ceux-ci (Zouros, 2000; Burzyński, et al., 2003; Venetis et al., 2007; Zouros, 2013) achève de supposer qu'un élément unique, ou une portion de l'ADNmt, suffit à l'origine de la DUI et est propre à chaque système. Selon une étude de Plazzi et Passamonti (2019), le peu de mutations retrouvées en commun, à travers les groupes taxonomiques de bivalves à DUI, et les événements indépendants de sélection directionnelle à la base de chaque système DUI, soutient une origine multiple de celui-ci.

Ces études suggèrent fortement que les espèces à DUI ont emprunté des chemins évolutifs différents, mais que certaines caractéristiques semblent plus partagées que d'autres, notamment la présence d'ORFans dans les génomes mitochondriaux et une modification du gène *cox2*. L'analyse de génomes mitochondriaux d'espèces supplémentaires est essentielle pour dresser un portrait plus clair des scénarios évolutifs de la DUI.

1.2.1.3 Sex-ratio chez les bivalves

Toute espèce possédant des chromosomes sexuels de type XX/XY ou ZZ/ZW aura un sex-ratio attendu de 1 : 1 (ex. 50% de chance durant la méiose d'obtenir un gamète portant le chromosome X ou le chromosome Y). Un sex-ratio biaisé donne donc une indication sur les forces qui agissent sur la détermination du sexe et est extrêmement utile pour identifier les facteurs potentiellement responsables, principalement chez les espèces dont on ne connaît pas les composantes génétiques sous-jacentes (Bull, 1983). C'est ce qui a été étudié par Saavedra et collègues (1997) et par Kenchington et collègues (2002, 2009) chez les moules marines du genre *Mytilus*. Le croisement de femelles avec plusieurs mâles a permis de constater que la sex-ratio chez la progéniture était toujours identique chez les familles issues de la même mère et ce, indépendamment du père avec

lequel on la croisait. À l'inverse, un même mâle croisé avec différentes femelles pouvait résulter en des sex-ratios différents. Au final, une femelle *Mytilus* se retrouve soit (i) à toujours produire des familles avec un sex-ratio fortement biaisé pour les femelles ($>90\%$ de femelles produites), (ii) à toujours produire des familles avec un sex-ratio fortement biaisé pour les mâles ($>90\%$ de mâles produits), ou (iii) à produire des familles avec un sex-ratio 1 : 1.

Plusieurs modèles de détermination du sexe ont été proposés suite à ces études. Les différentes hypothèses semblent s'accorder sur l'effet d'un facteur nucléaire clé d'origine maternelle (Cogswell et al., 2006 ; Kenchington et al., 2002, 2009 ; Saavedra et al., 1997), et la détermination sexuelle serait polygénique (Kenchington et al., 2002), supposant donc la participation d'autres facteurs génétiques dans ce processus. Étant donné le patron mitochondrial spécifique au sexe chez *Mytilus*, il a été suggéré que certains de ces facteurs génétiques pourraient être d'origine mitochondriale (Yusa et al., 2013). En effet, la corrélation absolue entre la présence de l'ADNmt M dans la lignée germinale et le sexe mâle suggère fortement un lien entre le système DUI et la détermination du sexe (Breton et al., 2011 ; Yusa et al., 2013).

1.2.1.4 Gonochorisme, hermaphrodisme et nouveaux gènes mitochondriaux

Une autre découverte supportant le lien entre le système DUI et la détermination du sexe a été faite chez les moules d'eau douce (Unionoida). Ce groupe comprend des espèces proches parentes qui sont hermaphrodites et gonochoriques, et une corrélation absolue a été observée entre gonochorisme et présence de la DUI, ainsi qu'entre hermaphrodisme et une transmission strictement maternelle (donc avec une perte de la DUI et de l'ADNmt M) (Breton et al., 2011). Autre fait intéressant, Breton et collègues (2009, 2011) ont découvert chez ces espèces, un nouveau gène associé à chacun des génomes mitochondriaux (Fig. 4). Ces nouveaux gènes, qui ont été nommé F-orf pour celui recensé dans le génome mitochondrial F, et M-orf pour le génome mitochondrial M, n'ont pas d'homologie avec aucun autre gène décrit et leur fonction exacte demeure pour le moment inconnue (Breton et al., 2009, 2011). Cependant, des analyses *in silico* suggèrent fortement qu'ils pourraient être impliqués dans les mécanismes sous-jacents à la DUI ; les protéines F-ORF sont soupçonnées d'interagir avec les acides nucléiques, d'adhérer aux membranes et avoir un rôle dans les voies de signalisation, tandis que les protéines M-ORF ont été associées à une interaction avec le cytosquelette, et à un rôle dans les processus d'ubiquitination et

d'apoptose (Milani et al., 2013 ; Mitchell et al., 2016 ; Guerra et al., 2017). De plus, des analyses biochimiques ont démontré que ces protéines sont localisées non seulement dans les mitochondries, mais aussi hors des mitochondries, dans la membrane nucléoplasmique ainsi que dans le nucléoplasme, ce qui suggère une fonction autre que la production d'énergie (Breton et al., 2011 ; Milani et al., 2013, 2014). Finalement, la séquence du gène F-orf retrouvée chez les espèces hermaphrodites possédant la SMI (nommé ici H-orf), a subi des macromutations importantes qui rendent la protéine H-ORF très divergente de la protéine F-ORF des espèces gonochoriques proches parentes, suggérant une perte ou une modification de sa fonction chez les espèces hermaphrodites (Breton et al., 2011). Breton et collègues (2011) ont émis l'hypothèse que ces nouveaux gènes associés au maintien des sexes séparés constituaient les éléments génétiques mitochondriaux impliqués dans la détermination du sexe chez les moules d'eau douce avec la DUI. Mais ceci reste à être clairement démontré.

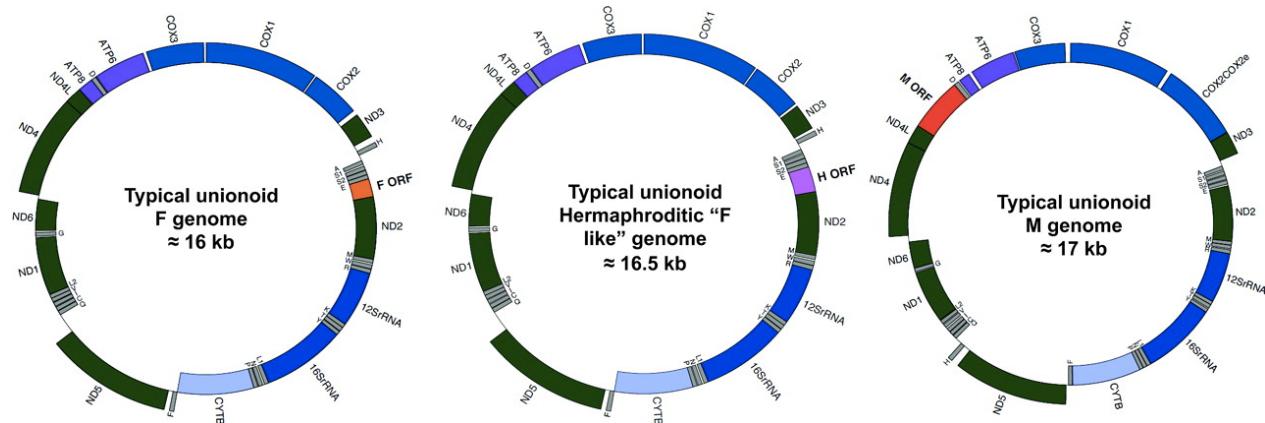


Figure 4. Nouveaux gènes associés aux génomes mitochondriaux complets de F (femelle), H (hermaphrodite) et M (mâle). Les génomes M et F sont issus d'espèces de moule d'eau douce. Tirée à partir de Breton et al., 2009.

Ghiselli et collègues (2012) ont réalisé une première étude approfondie du lien entre le système DUI et la détermination du sexe chez l'espèce *Ruditapes philippinarum* (Veneroida). Partant du principe que pour qu'un même génome donne un dimorphisme sexuel il faut qu'il y ait une expression différentielle génétique entre mâles et femelles, ils ont analysé les transcriptomes des gonades de plusieurs individus mâles et femelles de cette espèce afin de mettre en lumière les gènes qui pourraient être responsables de cette différenciation sexuelle. Ils ont également comparé les transcriptomes d'individus provenant de familles à fort biais pour les femelles ou pour les mâles, et ont découvert un lien entre l'expression différentielle des gènes liés à l'ubiquitination et les biais

de sex-ratio. Comme cité plus haut, l'élimination des mitochondries paternelles chez les embryons de mammifères étant prise en charge par des processus d'ubiquitination (Sutovsky et al., 1999), une modification de ce processus chez les espèces de bivalves ayant la DUI pourrait expliquer le maintien exceptionnel du génome paternel chez les embryons mâles (Ghiselli et al., 2012). Outre un rôle dans la rétention des mitochondries paternelles chez les mâles, des facteurs d'ubiquitination dans les œufs pourraient également avoir un rôle dans l'activation de l'expression de gènes liés au sexe durant l'embryogénèse, et le processus de différenciation sexuelle se voudrait multifactoriel et quantitatif, avec un développement mâle requérant une certaine quantité de transcrits afin de dépasser un certain seuil pour induire leur effet masculinisant (Ghiselli et al., 2012). Selon de récentes études bioinformatiques et biochimiques, le nouveau gène mitochondrial spécifique à l'ADNmt M pourrait faire partie de ces facteurs masculinisant (Milani et al., 2013 ; Milani et al., 2014 ; Mitchell et al., 2016).

En résumé, les bivalves qui possèdent la DUI, en plus de bouleverser le dogme de l'héritabilité uniparentale des mitochondries, contiennent des gènes mitochondriaux supplémentaires avec des nouvelles fonctions biologiques. Il a été supposé que ce système DUI avec ces nouveaux gènes aurait une fonction liée à la détermination du sexe (Breton et al., 2011). Si cela est vrai, la DUI serait le premier système de détermination du sexe qui impliquerait des protéines codées par l'ADNmt chez les animaux, et ceci expliquerait pourquoi ce système persiste chez les bivalves depuis plus de 400 millions d'années. Cependant les mécanismes sous-jacents à la DUI et la détermination du sexe sont encore méconnus. C'est l'étude de cet aspect qui est proposé dans ce projet de recherche.

2. Objectif et Hypothèses

L'objectif principal de cette thèse vise à mieux comprendre les mécanismes sous-jacents à la DUI et à la détermination du sexe chez les espèces à DUI. Pour ce faire, j'ai décidé de diviser cet objectif principal en deux objectifs spécifiques. Dans un premier temps, j'ai choisi d'étudier les transcriptomes de gonades matures d'espèces de moules d'eau douce (Unionoida) afin de mettre en lumière certains facteurs, nucléaires et mitochondriaux, qui pourraient être impliqués dans la détermination du sexe et la transmission mitochondriale. Étant donné qu'il existe des espèces proches parentes hermaphrodites (sans DUI) et gonochoriques (avec DUI) au sein d'un même genre, cela fait de ce groupe le candidat idéal pour tenter de répondre à ce premier objectif spécifique qui représente les deux premiers chapitres de la thèse. Dans un deuxième temps, j'ai voulu me concentrer sur l'aspect purement mitochondrial, afin de dresser un portrait plus spécifique des traits propres à la DUI, en caractérisant les génomes mitochondriaux de deux nouvelles espèces à DUI qui ont été recensées dans deux nouvelles familles de bivalves ; *Scrobicularia plana* (famille des Semelidae) et *Limecola balthica* (famille Tellenidae). Ce deuxième objectif spécifique représente le troisième chapitre de la thèse. Les trois chapitres de la thèse sont présentés dans les trois sous-sections suivantes.

2.1 Chapitre Un

Le premier chapitre a porté sur la comparaison des profils d'expression des gènes des tissus gonadiques, de femelles et de mâles, des espèces DUI *Venustaconcha ellipsiformis* et *Utterbackia peninsularis* grâce à une approche transcriptomique. Étant donné qu'aucun chromosome sexuel n'a été découvert dans ce groupe de mollusques, cette étude a permis de mieux comprendre les composantes génétiques (regroupées sous le terme de « facteurs génétiques », incluant ainsi également les gènes) qui agissent sur la différentiation phénotypique entre mâles et femelles, pourtant issus d'un même génome nucléaire. Les gènes qui s'expriment différemment entre les mâles et les femelles ont été identifiés comme des candidats potentiels impliqués dans la détermination du sexe ou la différentiation des gonades, ou encore la double transmission uniparentale de l'ADNmt. Ma première prédition était que les gènes qui ont une expression différente entre gonades mâles et femelles sont liés au succès reproducteur des mâles (liés à la

motilité des spermatozoïdes, par exemple) et des femelles, ainsi que des gènes connus pour leur rôle dans le déterminisme du sexe, chez d'autres espèces. Puis, pour comprendre comment la rétention des mitochondries paternelles chez les espèces à DUI est possible, à l'instar de l'étude réalisée par Ghiselli et collègues (2012), je m'attendais à trouver des gènes associés à des processus d'ubiquitination, étant donné son rôle dans la transmission mitochondriale chez les mammifères (Sutovsky *et al.* 2000), en soupçonnant une modification de ces mécanismes chez les mâles par rapport aux femelles. Afin d'éviter un biais d'expression dû à des facteurs environnementaux ou propres aux espèces, un blast réciproque a été réalisé entre les deux espèces pour maximiser la pertinence du set de gènes sélectionnés restant, dont la fonction devrait principalement être liée à la détermination/ différenciation sexuelle, et/ou aux mécanismes impliqués dans le processus de la DUI. Ce premier chapitre a permis de mettre en avant plusieurs facteurs liés au mécanisme d'ubiquitination, supportant l'hypothèse de son implication dans la régulation de la DUI. L'étude a également permis de constater que des mécanismes de méthylation pourraient être impliqués, suggérant une autre catégorie fonctionnelle à investiguer concernant la régulation de la DUI.

2.2 Chapitre Deux

Le deuxième chapitre a porté sur le profil d'expression des gènes des tissus gonadiques de l'espèce hermaphrodite *Utterbackia imbecillis* à SMI (= sans DUI). Le but étant de comparer ces résultats avec ceux de l'espèce proche parente gonochorique à DUI, *U. peninsularis*, afin de souligner les facteurs génétiques propres à l'hermaphrodisme vs. gonochorisme, mais aussi SMI vs. DUI. J'ai produit un assemblage comprenant toutes les séquences provenant de tous les individus, qualifié d'« hypride » puisqu'il est le résultat de deux espèces, qui a ensuite été utilisé comme référence pour pouvoir identifier les gènes différentiellement exprimés pour chaque sexe (mâle M, femelle F, hermaphrodite H). Des comparaisons MF, HF et HM ont été effectuées afin d'identifier les gènes candidats impliqués dans la détermination du sexe et / ou la transmission mitochondriale. À la lumière d'informations provenant d'études précédentes (Ghiselli *et al.*, 2012 ; Milani *et al.*, 2013 ; Capt *et al.*, 2018 ; Punzi *et al.*, 2018), une attention particulière a été accordée aux catégories d'ontologies géniques suivantes : reproduction, ubiquitination, nucléases, autophagie / mitophagie, pour leur rôle dans l'élimination des mitochondries et/ou de l'ADNmt paternel, et la méthylation, pour confirmer les résultats de mon premier chapitre. Je prédisais une

expression différentielle de ces gènes chez les mâles comparé aux femelles et/ou aux hermaphrodites, selon la nature du mécanisme. Mes résultats ont révélé l'existence de gènes avec des rôles conservés liés à la détermination du sexe chez les moules d'eau douce gonochoriques et hermaphrodites, et ont montré que les étapes de la détermination du sexe sont conservées chez des espèces de bivalves éloignées. Mes résultats ont également suggéré la participation potentielle de plusieurs catégories d'ontologies géniques dans l'élimination des mitochondries paternelles chez l'espèce SMI, notamment les endonucléases et les processus d'autophagie/mitophagie et de méthylation.

2.3 Chapitre Trois

Enfin, le dernier chapitre a été consacré à la caractérisation d'éléments génétiques mitochondriaux, associé aux ADNmt spécifiques aux mâles (M) et aux femelles (F), de deux nouvelles espèces à DUI de l'ordre des Venerida ; *Scrobicularia plana* (Gusman et al., 2016) et *Limecola balthica* (Pante et al., 2017). Les génomes complets ont été séquencés et annotés afin de comparer les similitudes et différences intra-spécifiques et inter-spécifiques. Je m'attendais à retrouver au moins une des caractéristiques propres à la DUI, listées dans l'introduction, avec une emphase particulière sur les nouveaux ORFs spécifiques aux sexes, étant donné leur présence et fonctionnalité dans l'ADNmt F et M de toutes les espèces à DUI (Breton et al., 2011 ; Ghiselli et al., 2012 ; Milani et al., 2013 ; Ouimet et al. en préparation). Étant donné les grandes variations structurales recensées dans le gène *cox2* du génome mitochondrial M de beaucoup d'espèces à DUI, je m'attendais également à en retrouver dans les ADNmt M de mes espèces. Nos résultats ont validé l'existence de modifications dans le gène *Mcox2*, avec la détection des plus longues insertions jamais recensées dans le règne animal pour ce gène. De plus, un nouveau gène spécifique aux mâles semble également présent dans l'ADNmt M de nos espèces. Cette étude a permis d'élargir nos connaissances sur les caractéristiques de la DUI propres à chaque groupe taxonomique afin d'avoir une vision plus précise des éléments partagés et divergents chez les espèces à DUI.

Chapitre 1.

***Déchiffrer le lien entre la double transmission
uniparentale de l'ADNmt et la détermination du
sexe chez les bivalves : indices d'une étude
transcriptomique comparative.***

Deciphering the link between doubly uniparental inheritance of mtDNA and sex determination in bivalves: clues from comparative transcriptomics

Charlotte Capt^{1,*}, Sébastien Renaud^{1,2}, Fabrizio Ghiselli³, Liliana Milani³, Nathan Johnson⁴, Bernard Sietman⁵, Donald T. Stewart⁶, Sophie Breton^{1,*}

¹ Department of Biological Sciences, Université de Montréal, Montréal, QC, Canada

² Centre de la Science de la Biodiversité du Québec, Université de Montréal, Montréal, QC, Canada

³ Dipartimento di Scienze Biologiche, Geologiche ed Ambientali, University of Bologna, Bologna, Italy

⁴ Wetland and Aquatic Research Center, Gainesville, FL, USA

⁵ Center for Aquatic Mollusk Programs, Minnesota Department of Natural Resources, Lake City, MN, USA

⁶ Department of Biology, Acadia University, Wolfville, NS B4P 2R6 Canada

Published in Genome Biology and Evolution 10:577-590.
doi: 10.1093/gbe/evy019.

** Minor edits to the published text have been made.*

Author Contributions

CC, SR, and SB designed experiments. NJ conducted sampling. CC performed laboratory experiments. CC and SR analyzed data. CC, SR, NJ, DTS and SB wrote the manuscript. SB conceived and managed the project. All authors read and approved the final version of the manuscript.

Abstract

Bivalves exhibit an astonishing diversity of sexual systems and sex-determining mechanisms. They can be gonochoric, hermaphrodites or androgenetic, with both genetic and environmental factors known to determine or influence sex. One unique sex-determining system involving the mitochondrial genome has also been hypothesized in bivalves with doubly uniparental inheritance (DUI) of mtDNA. However, the link between DUI and sex determination remains obscure. In this study, we performed a comparative gonad transcriptomics analysis for two DUI-possessing freshwater mussel species to better understand the mechanisms underlying sex determination and DUI in these bivalves. We used a BLAST reciprocal analysis to identify orthologs between *Venustaconcha ellipsiformis* and *Utterbackia peninsularis* and compared our results with previously published sex-specific bivalve transcriptomes to identify conserved sex-determining genes. We also compared our data with other DUI species to identify candidate genes possibly involved in the regulation of DUI. A total of ~12,000 orthologous relationships was found, with 2,583 genes differentially expressed in both species. Among these genes, key sex-determining factors previously reported in vertebrates and also in bivalves (e.g. *Sry*, *Dmrt1*, *Foxl2*) were identified, suggesting that some steps of the sex determination pathway may be deeply conserved in metazoans. Our results also supported the hypothesis that a modified ubiquitination mechanism could be responsible for the retention of the paternal mtDNA in male bivalves, and revealed that DNA methylation could also be involved in the regulation of DUI. Globally, our results suggest that sets of genes associated with sex determination and DUI are similar in distantly-related DUI species.

Key words: sex determination, mitochondrial DNA, doubly uniparental inheritance, comparative transcriptomics, Bivalvia

Introduction

Bivalves show extreme diversity in their reproductive systems, ranging from functional (simultaneous) hermaphroditism, alternative sexuality (sequential hermaphroditism), to strict gonochorism (i.e. species that exist as separate males and females). Both genetic and environmental factors appear to determine sex in most bivalve species studied to date (reviewed in Breton et al., 2018; Chávez-Villalba et al., 2011; Coe, 1943). However, because sex determination has been studied in great detail only in oysters, the genetic and environmental factors that control sexual diversification in bivalves are still poorly known. An important reason for studying sex determination in bivalves relates to their biology and ecology: with ~25,000 living species (www.bivatol.org), they are sufficiently diverse to provide a rich source of material to better understand the evolution of sex and sex determination in general, and to provide unique examples of sex-determining mechanisms, including the only possible example of a sex-determining system involving the mitochondrial genome in animals (e.g. Breton et al., 2011a).

A variety of approaches have been undertaken to elucidate the mode of sex determination in bivalves. Globally, these studies demonstrated an absence of heteromorphic sex chromosomes, an effect of several environmental factors on sex-ratio (i.e. temperature, food availability and pollutants, such as heavy metals organochlorines and exogenous steroids), and a polygenic architecture of sex determination (Breton et al., 2018). Genes homologous to sex-determining pathway genes in other animals have been identified in several bivalve species. For example, homologs of *Sry* (sex determining region-y)-box 30 (*Sox 30*) and *Dmrt1* (doublesex and mab-3 related transcription factor 1), known for their role in male sex determination in nematodes, fruit flies and vertebrates (Kopp, 2012; Wallis, Waters, & Graves, 2008), have been found in the marine clam *Ruditapes philippinarum*, and in some species of scallop and oyster (Ghiselli et al., 2012; Y. Li et al., 2016; Llera-Herrera et al., 2013; Teaniniuraitemoana et al., 2014; Yu et al., 2011; N. Zhang, Xu, & Guo, 2014). Several other genes known to act in sex determination in other animals or during early gonadal differentiation have been identified in bivalves, such as *Oyvlg* and *vasph*, i.e. homologues of *vasa*, which is a gene involved in primordial germ cell development and early sex differentiation in eukaryotes; *Cg-SoxE*, a homologue of *Sox9*, and *β-catenin*, which are respectively expressed when sex is still not distinguishable (or in mature females and vitellogenesis

oocytes in *Crassostrea gigas* [Santerre et al., 2014]); and *CgFoxL2*, a homologue of *FoxL2*, a key gene involved in ovarian determination in vertebrates (Fabioux et al., 2004; Milani et al., 2011; Milani et al., 2015; Uhlenhaut & Treier, 2006; N. Zhang et al., 2014).

An unconventional sex-determining mechanism involving doubly uniparental inheritance (DUI) of mitochondrial DNA has been hypothesized in bivalves (e.g. Breton et al., 2011a; Milani, Ghiselli, & Passamonti, 2016). This unique system of mitochondrial transmission in the animal kingdom has been discovered in >100 species from different bivalve orders (Mytiloida, Unionoida, Veneroida and Nuculanoida; Gusman et al., 2016). DUI involves two distinct mitochondrial genomes that are transmitted in a sex-specific way: the female-type, or F-type, transmitted through eggs and usually found in both male and female somatic tissues and in female gonadic tissues, and the male type, or M-type, transmitted through sperm and usually found in male gametes (reviewed in Breton et al., 2007a; Passamonti & Ghiselli, 2009; Zouros, 2013). In addition to the typical set of 13 mtDNA-encoded proteins (Capt, Passamonti, & Breton, 2015; Gissi, Iannelli, & Pesole, 2008), novel sex-associated mtDNA-encoded proteins have been found in bivalves with DUI; *F-orf* in the F mtDNA and *M-orf* in the M mtDNA, both of which are relatively conserved across species within a family (Breton et al., 2009; Breton et al., 2011; Milani et al., 2013; Mitchell et al., 2016). These genes were hypothesized to be key elements of a sex-determination system involving mitochondria, i.e. they may play a role in the maintenance of a gonochoric reproductive system with separate male and female sexes (Breton et al., 2011a). This hypothesis came from the observation that gonochorism in freshwater unionoid mussels is absolutely correlated with the presence of DUI and of these novel sex-specific proteins, whereas closely related hermaphroditic species lack the M mtDNA (i.e. possess strict maternal inheritance of mtDNA) and have macromutations in the *F-orf* gene in their F mtDNA (Breton et al., 2011a). If true, this would make DUI the first animal sex-determination system involving the mitochondrial genome, and would explain its long-term persistence in bivalves. However, this hypothesis still remains to be rigorously tested.

Ghiselli et al. (2012) published the first whole transcriptome analysis by RNA-Seq to better understand the mechanisms underlying DUI and sex determination in the Manila clam *R. philippinarum*. Because previous studies demonstrated that elimination of paternal mitochondria

in mammal species with strict maternal inheritance of mtDNA is dependent on a mechanism involving the ubiquitin-proteasome system, the principal system for protein degradation (Sutovsky et al., 1999), it was proposed that a modification of the ubiquitination process in male bivalves with DUI would allow sperm mitochondria and their M genomes to escape this mechanism and invade germinal cells (Ghiselli et al., 2012; Kenchington et al., 2002). Furthermore, because ubiquitination is also involved in the regulation of gene expression and plays a role in sex determination of several animals (e.g. *C. elegans* and *Drosophila*; see Ghiselli et al. 2012 for a review of the literature), Ghiselli et al. (2012) proposed a relationship between ubiquitination, sex bias, and mitochondrial inheritance, and identified candidate ubiquitination genes, such as the ubiquitin activating enzyme 1 (*uba-1*) and proteasome subunit alpha 6 (*psa-6*), for further investigation. To date, this study remains the only transcriptomic analysis performed to identify genes involved in bivalve sex determination and DUI. Although an in-depth study of sex-specific genes has recently been conducted on male and female gonads of the gonochoric, DUI-possessing freshwater mussel *Hyriopsis schlegelii* (Shi et al., 2015), the link between DUI and sex determination has not been investigated. To sum up, a total of 45,422 genes differentially expressed in male and female gonads in *H. schlegelii* and key genes reported to govern sex-determination pathways in mammals were identified, including *Sry*, *Dmrt1* and *Sox9*, upregulated in males, and *Foxl2* and β -catenin, upregulated in females (Shi et al., 2015).

With the aim of unraveling the mechanisms underlying sex determination and DUI in freshwater mussels, we thus performed a comparative gonad transcriptome analysis for the two DUI species *Venustaconcha ellipsiformis* (subfamily Ambleminae) and *Utterbackia peninsularis* (subfamily Anodontinae). Furthermore, we compared our results with published sex-specific transcriptomics data from the DUI freshwater mussel *H. schlegelii* and marine clams *R. philippinarum* and *Ruditapes decussatus* to identify conserved genes involved in the regulation of sex-specific aspects of DUI systems. Because of the many similarities found among the distantly-related DUI species (e.g. sex-ratio bias, mitochondrial behavior in early embryos, rates of evolution of the two genomes; see Zouros 2013 for a review), it was indeed expected that sets of genes associated with sex determination and DUI will be similar in distantly-related DUI species. We hypothesized that if a modification of the ubiquitination mechanism is responsible for the retention of the paternal mtDNA in male bivalves, then “transcriptionnal signatures” of this modification

should be discernable in all DUI species.

Materials and Methods

Sample collection and RNA isolation

Adult specimens of *V. ellipsiformis* and *U. peninsularis* were collected in July 2014 from Straight River (Minnesota, USA; Lat 44.006509, Long -93.290899) and Suwannee River (Florida, USA; Lat 29.58684, Long -82.94095), respectively. Mussels were shipped alive to the Université de Montréal, where they were opened and sexed by microscopic examination of gonad smears. Mature gonads of all unambiguously sexed specimens were frozen at -80°C until further analyses.

Total RNA was extracted from frozen gonad tissues with RNeasy Plus Universal Mini Kit (QIAGEN Inc., Valencia, CA) and treated with Turbo DNase (AMBION, Austin, TX) following the provided protocols. A total of 25 RNA extractions were performed, 14 for *U. peninsularis* (8 females and 6 males) and 11 for *V. ellipsiformis* (7 females and 4 males). The quality and quantity of extracted RNA were examined via electrophoresis on 1% agarose gel and BioDrop µLITE spectrophotometer. Sixteen samples (4 males and 4 females for each species) were sent to the Institute for Research in Immunology and Cancer Institute (IRIC, Université de Montréal, Montréal, Canada) for Illumina library preparation and paired-end (2 X 100 bp) sequencing (Illumina HiSeq2000; San Diego, CA).

De novo transcriptome assemblies

Prior to assembly, raw paired-end reads were trimmed with Trimmomatic (v0.32; Bolger et al., 2014), which is available as part of the Trinity transcriptome assembler (trinityrnaseq_r20140717; Grabherr et al., 2011), to remove low-quality reads (i.e. with Phred quality scores less than 30, or if the read contained more than 5 ambiguous nucleotides “N”). As part of the trimming process, reads similar to known polymerase chain reaction primers and Illumina adapter sequences were also removed by specifying the appropriate “ILLUMINACLIP” parameters. Reads quality was visualized with FastQC (v0.11.2; Andrews, 2010). For each species, high quality reads were assembled *de novo* to provide species-specific reference transcriptomes using Trinity (Haas et al.,

2013). Parameters were kept as default, except for the --min_contig_length, which was set at 300 to remove contigs shorter than 300 bp. This value was chosen for further annotation purposes, i.e. according to the parameters fixed by TransDecoder that identifies ORFs with a minimum length of 100 amino acids to avoid false positives (<https://transdecoder.github.io/>).

Briefly, Trinity combines overlapping reads to form longer fragments called contigs, which are subsequently clustered based on sequence similarity (i.e. grouping of related contigs that correspond to portions of alternatively-spliced transcripts or otherwise unique portions of paralogous genes). Then these data were processed to extract full-length alternatively spliced isoforms and tease apart transcripts derived from paralogous genes (Grabherr et al., 2011) to acquire non-redundant unigenes.

The completeness of transcriptome assemblies was tested by Benchmarking Universal Single-Copy Orthologs (BUSCO) software (busco_v3, using the mode –tran), by comparing known core eukaryotic and metazoan genes and the contigs assembled by Trinity. BUSCO provides quantitative measures for the assessment of transcriptome completeness, based on evolutionarily-informed expectations of gene content from near-universal single-copy orthologs selected from different downloadable datasets (Simão et al., 2015).

Annotation

Assembled contigs were searched against two databases, i.e. Swiss-Prot and TrEMBL (Uniref90 clusters), of the Universal Protein Resource (UniProt) protein database (The UniProt Consortium 2008; 2017) with BLASTp (E-value set at e-05) (Altschul et al., 1990). To do this, ORFs and associated protein-coding sequences were first predicted from the contigs using TransDecoder (r20140704; Haas et al., 2013) and the longest ORF candidate was considered. BLASTx searches of assembled contigs against the SwissProt database were also performed to increase the proportion of annotated genes. Finally, BLASTp searches for both species were performed against the PANMDB database. Specifically, we used the reference dataset constructed by Patnaik et al. (2016), which combines protein sequence data of Arthropoda, Nematoda, and Mollusca downloaded from the Taxonomy browser of the NCBI nr database and stored in the PANMDB

(freely downloadable from the amino acid database BLAST web-interface of the Malacological Society of Korea; <http://malacol.or.kr/blast/aminoacid.html>; Kang et al., 2015).

Read mapping, differential expression and functional enrichment analyses

The *de novo* transcriptomes assembled for each species served as reference for read mapping. Specifically, raw reads were mapped using Bowtie2 (v2.2.4; Langmead & Salzberg, 2012), keeping the default parameters. Reads were sorted by name, and indexed with Samtools (v1.1; H. Li et al., 2009). Samtools idxstat was used to extract the numbers of reads mapped per contig and the results were then normalised according to Robinson & Oshlack (2010).

Two approaches were used to perform differential expression analyses between male and female gonads for each species, i.e. one with the R package edgeR (version edgeR_3.12.0; Robinson, McCarthy, & Smyth, 2010) and one with DESeq2 (version DESeq2_1.10.1; Love, Huber, & Anders, 2014). We used both approaches for comparative purposes, because multiple methods have been used to identify differentially expressed genes between male and female gonads in bivalves [e.g. edgeR by Ghiselli et al. (2012) for the marine clam *R. philippinarum*; DESeq2 by Teaniniuraitemoana et al. (2014) for the oyster *P. margaritifera*; and Shi et al. (2015) used a statistical method developed by Audic & Claverie (1997) for the freshwater mussel *H. schlegelii*]. DESeq2 is more sensitive to outliers and generally gives lower true positive rates, whereas edgeR performs better to uncover true positives (Soneson & Delorenzi, 2013; Z. H. Zhang et al., 2014). For a complete comparison of both software programs, see Z.H. Zhang et al. (2014).

Fragments per kilobase of exon per million fragment sequenced (FPKM) were obtained using the DESeq2 Bioconductor package (Love et al., 2014), specifying “robust=F” to use sum of raw counts, allowing analyses of differentially expressed unigenes (hereafter defined as differentially expressed genes or DEG). With edgeR, DEG were selected based on FPKM values > 1 in one or the other sex, and log₂ fold change ≥ 1 and ≤ -1 . Since DESeq2 removes low expression data when calculating FDR values, no filter has been performed on FPKM values, and DEG were also selected based on log₂ fold change ≥ 1 and ≤ -1 . Only unigenes with p-values < 0.05 after adjustment by False-Discovery Rate (FDR) (Benjamini & Hochberg, 1995) were considered as differentially

expressed. The package edgeR (plotSmear) was used to visualize differential expression between gonad samples for each species. The proportion of unigenes that were differentially expressed according to both analytical methods was also measured.

To further understand these biased expressed unigenes, gene ontology (GO) enrichment analyses were performed for each species to test whether specific GO terms (biological process, molecular function, and cellular component levels) were over-represented in DEG using GOseq (v1.28.0; Young et al., 2010). An FDR corrected *p*-value of 0.05 was used for significance.

Blast reciprocal analysis

A whole transcriptome reciprocal-best-BLAST-hits (RBH) analysis was applied to identify orthologous genes between *V. ellipsiformis* and *U. peninsularis* (e.g. to filter possible contamination- and species-specific transcripts). Specifically, each contig from *V. ellipsiformis* was searched against all contigs from *U. peninsularis* using the parameter *p=blastn+* (default parameters were kept according to the proteinortho5.pl script, with an E-value set at e-05; Lechner et al., 2011), and conversely each sequence of *U. peninsularis* was searched against all sequences of *V. ellipsiformis*. This step allowed for identification of orthologous genes between the species (e.g. Zhao et al., 2014). From these orthologous gene sets, the orthologs found to be differentially expressed between males and females both in *U. peninsularis* and *V. ellipsiformis* using edgeR and DESeq2 were studied more thoroughly. Transcription level and annotation were individually verified for each DEG in each species to confirm the reliability of our approach, to identify orthologous sex-biased genes that may be involved in sex determination, molecular controls of DUI, or both.

Results and discussion

Sequencing and assemblies of the transcriptomes

For *U. peninsularis*, approximately 43 million paired end reads per individual were produced by the Illumina HiSeq2000 sequencing, generating a total of 348 652 616 raw reads of 2 X 100 bp. After trimming, 288 320 738 (83%) high-quality reads were retained for transcriptome assembly.

The *de novo* assembled transcriptome using Trinity resulted in 200 961 contigs with a N50 of 1,270 bp and average contig length of 709 bp (Table 1.1). These contigs were assembled into a total of 165 788 unigenes.

Table 1.1. *De novo* assembly quality statistics of *Uutterbackia peninsulae* and *Venustaconcha ellipsiformis* transcriptomes.

	<i>U. peninsulae</i>	<i>V. ellipsiformis</i>
Total raws	348,652,616	333,708,540
Average raws	43,581,577	41,713,567
Total trimmed raws	288,320,738	320,230,962
Average trimmed raws	36,040,092	40,028,870
Total trinity "unigenes"	165,788	221,362
Total trinity contig	200,961	285,260
Percent GC	33.36	37.58
Contig N10	5,581	6,494
Contig N20	3,895	4,650
Contig N30	2,916	3,560
Contig N40	2,224	2,731
Contig N50	1,655	2,083
Median contig length	558	595
Average contig	1,026	1,171
Total assembled bases	206,277,984	334,013,631
Unigene N10	4,882	5,792
Unigene N20	3,298	3,896
Unigene N30	2,409	2,789
Unigene N40	1,750	2,019
Unigene N50	1,251	1,433
Median unigene length	510	516
Average unigene	884	944
Total assembled bases	146,628,566	208,707,657

For *V. ellipsiformis*, 41 million reads per individual were obtained on average, giving a total of 333 708 540 raw reads generated by Illumina sequencing. After trimming, 320 230 962 (96%) clean reads were obtained and assembled into 285 260 contigs (N50 of 2,083 bp, average contig length of 1 171 bp), of which 221 362 were considered as unigenes and kept for further analyses (Table 1.1).

Assembly quality (assembled contigs) assessed by the representation of 978 core eukaryotic

and metazoan genes using BUSCO resulted in 99.9% complete for *U. peninsularis* (with 24.6% duplicated, 0.3% fragmented, and 0.7% missing), and 99.8% complete for *V. ellipsiformis* (with 35.2% duplicated, 0% fragmented, and 0.2% missing) (Supplementary Figure S1.1). In theory, the relatively high number of duplicates may represent alternatively spliced forms, gene duplication, or allelic variation (heterozygosity) in the samples used to construct the assemblies. Therefore, we ran the same BUSCO analysis using the assembled unigenes (i.e. filtered for alternatively spliced forms and paralogs) to better understand the origin of our duplicates (Supplementary Figure S1.1). This analysis still showed a similar number of completely recovered genes (98.8% for *U. peninsularis* and 99.7% for *V. ellipsiformis*), but a markedly lower number of duplicates (14.2% and 24.8%, respectively), suggesting that gene duplication and/or alternative splicing may contribute to the relatively high number of duplicates. Nevertheless, the high number of complete genes that was recovered provides an important validation of the completeness of both assemblies.

For both species, the raw sequence data in FASTQ format have been submitted to the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database and are accessible under accession numbers SRR6279376-83 (*U. peninsularis*) and SRR6279374-75 and SRR6279384-89 (*V. ellipsiformis*).

Sequence annotation

BLASTp searches against SwissProt and Uniref90, BLASTx against SwissProt, and BLASTx against PANMDB, respectively, allowed annotating 26 469 (13.1%), 32 515 (16.2%), 28 495 (14.2%) and 32 503 (16.1%) contigs for *U. peninsularis*. All databases together gave a total annotation success of 16.8% contigs. With SwissProt, most of the matches came from the extensively studied taxon Mammalia (Supplementary Figure S1.2). With Uniref90, the largest proportion of hits (46.4%) came from the oyster species *C. gigas* (Figure 1.1), for which whole-genome sequencing and annotation were recently completed (G. Zhang et al., 2012). The second largest proportion of hits (23.6%) came from the molluscan class Gastropoda (Figure 1.1). These results are comparable with those reported in other recent *de novo* transcriptome sequencing studies for freshwater mussels or other bivalve species (i.e. 15-22% annotated contigs/unigenes; Shi et al.,

2015; Patnaik et al., 2016; Yarra et al., 2016; Wang, Liu, & Wu, 2017) and indicate that it can be challenging to provide good annotations for non-model organisms for which few genomic resources are available. Also, very few contigs matched bacterial sequences (or other possible contaminating species such as trematodes), suggesting negligible contamination for *U. peninsularis*, and indicating that the assembly for this species comprised mainly gonadal coding RNAs.

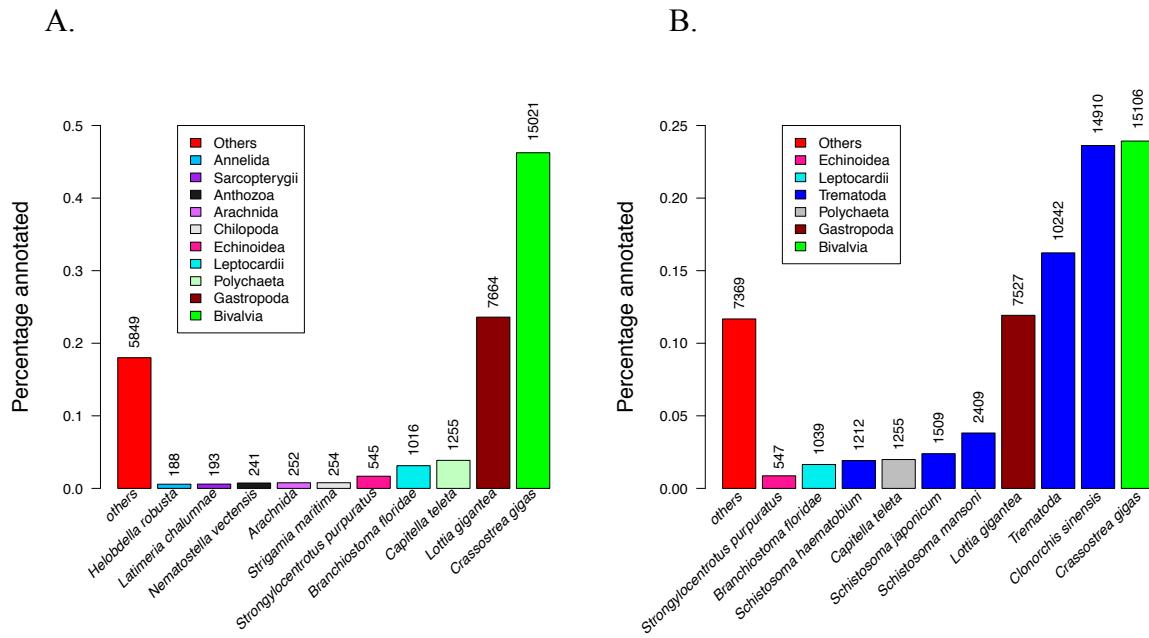


Figure 1.1. Taxonomic distribution of top BLASTp hits in Uniref90 protein database for *U. peninsularis* contigs (A) and *V. ellipsiformis* contigs (B). Numbers of unique hits in each group are shown.

These results obtained for *U. peninsularis*, however, are in sharp contrast with those obtained for *V. ellipsiformis*, for which 63 126 (22.1%) contigs were annotated with Uniref90 but with almost half of them matching trematode species (Figure 1.1). The fractions of contigs annotated with SwissProt and PANMDB were also slightly larger for *V. ellipsiformis*, with 49 904 (17.5%) and 50 444 (17.7%) contigs annotated with BLASTp and BLASTx searches against SwissProt (see Supplementary Figure S1.2), and 57 215 (20%) contigs annotated with PANMDB. Again, several matches came from *C. gigas* with Uniref90 (Figure 1.1), but the discovery that ~40% of the annotated contigs had a significant match with trematode species, which is not the case for *U. peninsularis*, strongly suggest a trematode infection in *V. ellipsiformis*, a phenomenon that is often observed in freshwater mussels, which are intermediate hosts of several trematode species

(Gangloff, Lenertz, & Feminella, 2008). This contamination most likely explains the higher number of duplicates obtained with BUSCO as well as the higher percentage of annotated contigs for *V. ellipsiformis*. We performed additional analyses to verify the extent of this contamination in our samples, and found that three of four female samples, but no male, produced several hits when blasted against the three most expressed trematode transcripts (data not shown). This is consistent with recent report that mature female freshwater mussels are more susceptible to trematode infection (Müller et al., 2015).

We discarded contaminant trematode sequences from the transcriptome of *V. ellipsiformis* by relying on annotation to remove Platyhelminthes transcripts (e.g. Sayadi et al., 2016). Although this approach may result in the significant loss of mussel-specific data, we considered it to be a valid approach for retaining the maximum number of sequences related to our focal species (*V. ellipsiformis*). Alternatively, this approach can lead to poor decontamination, i.e. with many non-annotated contigs belonging to contaminant. However, because the ultimate goal of our study was to identify similar sets of genes associated with sex determination and DUI in two different freshwater mussel species, we tried to minimize the contamination problem by using reciprocal BLAST best matching between both species to filter out contamination-specific (and species-specific) contigs (see below). A total of 25 179 (40%) annotated contigs were considered as probable contaminants and discarded. In total, we retained 37 947 genes with significant blast hits in our downstream analyses.

Differentially expressed genes (DEG) between male and female gonads in each species

The differential gene expression analysis between sexes for *U. peninsularis* revealed 7 281 DEG with edgeR, and 23 107 DEG with DESeq2. The DEG identified using DESeq2 contained all the DEG identified with edgeR (Figure 1.2). A total of 4 315 unigenes were shown to be up-regulated in males vs. 2 966 in females with edgeR, whereas 14 216 genes were considered up-regulated in males vs. 8 891 in females with DESeq2. This is a smaller number than what was found in the DUI-possessing freshwater mussel *H. schlegelii* using a statistical method (45 422 DEG, with 19 511 unigenes up-regulated in male gonads vs. 25 911 in female gonads; Shi et al., 2015), but a substantial number considering what has been revealed for example in the oyster species *Pinctada*

margaritifera using DESeq2 (1 993 DEG, with 1 419 contigs up-regulated in male gonads vs. 574 in female gonads; Teaniniuraitemoana et al., 2014) and in the DUI-possessing marine clam *R. philippinarum* using edgeR (1 575 DEG, with most of the genes being male biased; Ghiselli et al., 2012). Although different DEG analytical methods and experimental setups may explain these dissimilarities, it was suggested that a low number of DEG should be expected in bivalves considering their general lack of secondary sexual characters and sexual dimorphism, and also because sex-specific function of reproductive genes in these organisms is limited to gonad development, gametogenesis, and fertilization (Ghiselli et al., 2012). However, contrary to most bivalve species, unionoid freshwater mussels have a unique reproductive mode that may cause pronounced sexual dimorphism: (i) their larvae mature in certain areas unique to female gills called marsupia (which can also alter shell morphology) and (ii) because most species require a host fish for their larvae, many gravid females, including those of *V. ellipsiformis* (Allen et al., 2007), display adaptations such as mantle-derived lures to actively attract a prospective fish for their parasitic larvae (Haag & Warren, 1999; Barnhart, Haag, & Roston, 2008; Haag, 2012; Zieritz & Aldridge, 2011). Interestingly, sexual dimorphism in burrowing behavior has also been reported in the freshwater mussel DUI species *Elliptio complanata* (Flynn et al., 2013). To our knowledge, however, no sexual dimorphism has been reported in *U. peninsularis*. Most of the differentially expressed genes found in this species are male-biased, which is consistent with what has been found in most bivalve species, as well as in many other animals (Ghiselli et al., 2012; Teaniniuraitemoana et al., 2014; Peng et al., 2015; Y. Li et al., 2016). This phenomenon may be explained by a higher proportion of essential functions for female-biased genes (those shared by the two sexes) than for male-biased genes (see Ghiselli et al., 2012). This is also consistent with the previous assumption that male development is associated with an activation of several testis-specific genes and/or a repression of several genes vital for ovarian development (see Peng et al., 2015).

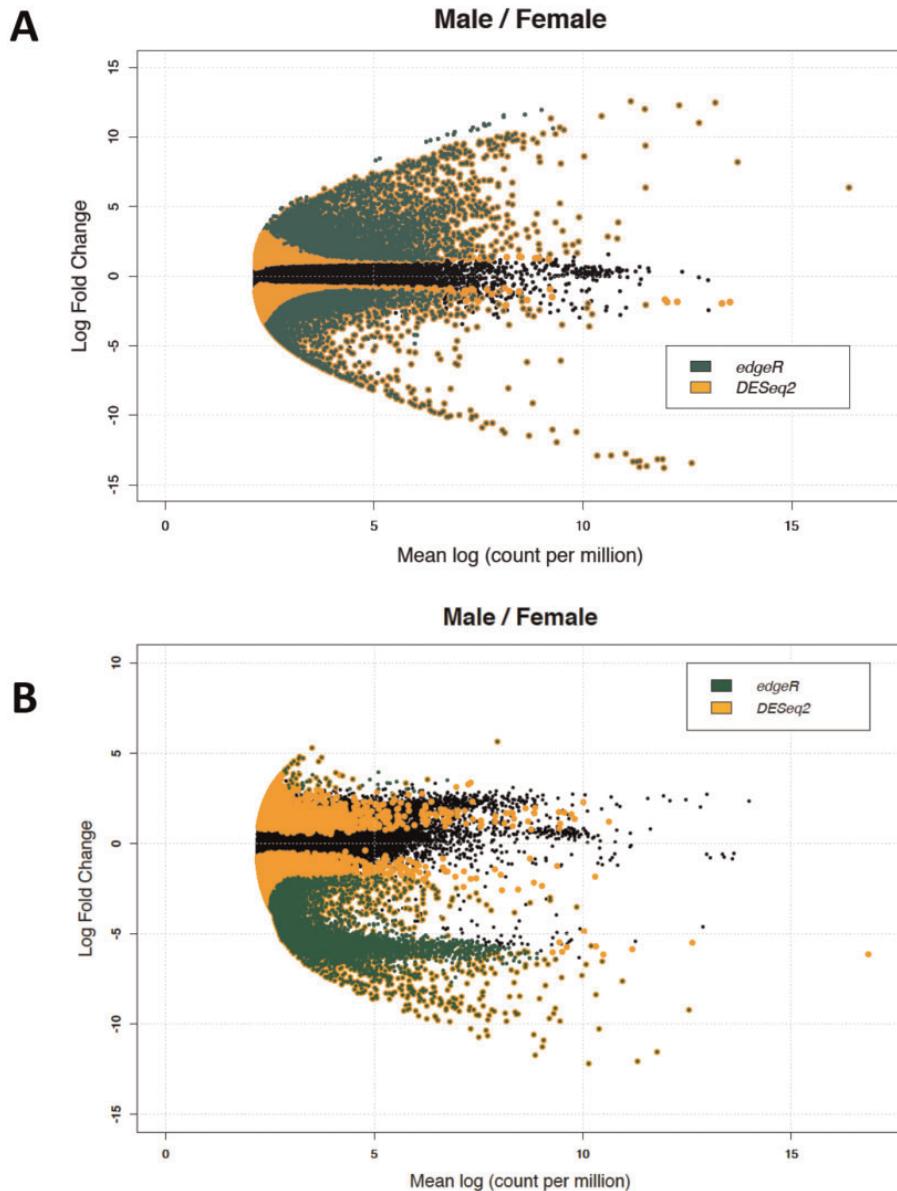


Figure 1.2. Smear plots of DEG between male and female gonads in *U. peninsularis* (A) and *V. ellipsiformis* (B). Black dots represent genes normally expressed in both tissues. Green dots, DEG revealed by edgeR software; orange dot, DEG revealed by DESeq2. Green dots surrounded by orange represent genes considered as DEG by both software packages. Positive fold-changes represent genes upregulated in males, while negative ones represent genes upregulated in females.

For *V. ellipsiformis*, 11 408 DEG were revealed with edgeR and 52 257 DEG with DESeq2 (Figure 1.2). Only 60 unigenes were up-regulated in males vs. 11 348 in females with edgeR, whereas DESeq2 revealed 3 134 unigenes up-regulated in males and 49 123 in females. This important female-biased gene expression in *V. ellipsiformis* is most likely explained by trematode contamination found in three out of four female gonad samples (see the annotation section above). After having filtered out 25 179 Platyhelminthes contigs from our data, the number of differentially

expressed genes did not vary much with 9 769 DEG revealed by edgeR (765 up-regulated in males and 9 004 in females) and 52 652 revealed by DESeq2 (3 317 up-regulated in males and 49 335 in females).

These numbers, although not surprising since the percentage of annotation (< 22%) was very low, either suggest (i) that the contamination could not be completely removed from our data using the approach described above, or (ii) that changes in female host gene expression occurred in response to parasite infection (e.g. Wijayawardena, Minchella, & DeWoody, 2016). Fortunately, because our *U. peninsularis* samples were not infected by trematodes, our BLAST reciprocal analysis discussed below allowed us to mitigate this problem and to accomplish our main objective, i.e. identify similar sets of genes associated with sex determination and DUI in two different freshwater mussel species.

Blast reciprocal analysis and enriched gene ontology terms and pathways in male and female gonads

To identify the main genes possibly involved in sex determination and/or DUI in *U. peninsularis* and *V. ellipsiformis*, we performed a whole transcriptome BLAST analysis. From the reciprocal BLAST results, we selected the genes found to be differentially expressed in *U. peninsularis* in our previous analyses using edgeR and DESeq2, and verified their sex-bias in *V. ellipsiformis* DEG filtered for “flatworm contigs” (i.e. we verified that every male-biased or female-biased gene in one species was also male-biased or female-biased in the other).

A total of ~12,000 orthologous relationships was found between both species, 16 740 when isoforms were included. From these orthologous genes, 2 342 were found to be differentially expressed in *U. peninsularis* by edgeR and 2 583 (including the 2 342 from edgeR) by DESeq2. Specific attention was given to these 2 583 genes (1 568 male-biased and 1 015 female-biased) to increase the probability of identifying candidate genes that could have a role in sex determination and/or DUI. Of the 1 568 male-biased orthologous genes, 1 432 (91.3%) had a significant BLASTp (or BLASTx) match in the SwissProt and/or UniRef90 protein databases (E-value set at e-05), whereas of the 1 015 female-biased orthologous genes, 903 (89%) had a significant BLASTp (or BLASTx) match in the SwissProt and/or UniRef90 protein databases (Supplementary Table S1.1).

The same genes were also found to be differentially expressed in *V. ellipsiformis*.

Using all annotated and differentially expressed orthologous genes, GOSeq analyses revealed 57 ontology categories significantly enriched in both species (Supplementary Table S1.2). Top GO terms enriched in female and male gonads of both species are shown in Figure 1.3 (for complete GOSeq results for *U. peninsularis* and *V. ellipsiformis* after trimming for trematode contigs, see Supplementary Table S1.3 and S1.4, respectively). In transcriptomes of both species, GO terms show higher counts for male-biased genes. At the biological process level, single-organism cellular process, cellular component organization or biogenesis, and organelle organization had the top counts. Intracellular component and organelle component represented the majority of terms at the cellular component level, whereas at the molecular function level, protein binding was the most represented term (Figure 1.3). Similar results were also obtained for other bivalve species, including the oyster *Crassostrea hongkongensis* (Tong et al., 2015), the scallop *Patinopecten yessoensis* (Li et al., 2016), as well as the DUI-containing marine clam *R. philippinarum* (Ghiselli et al., 2012) and freshwater mussel *Cristaria plicata* (Patnaik et al., 2016).

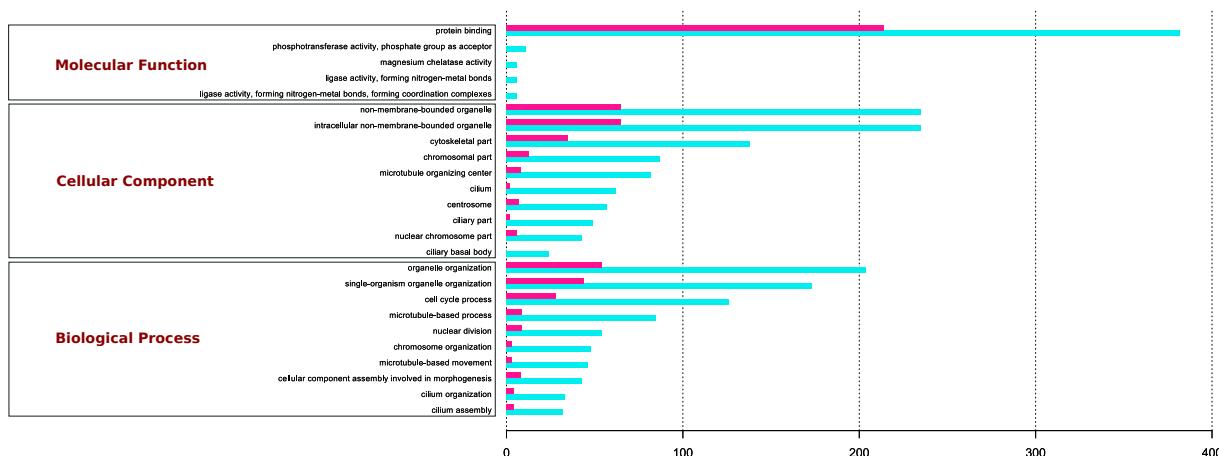


Figure 1.3. Significantly enriched-GO terms involved in molecular function (top), cellular component (middle), and biological process (bottom) in ovary (pink) and testis (blue) of both species. Analyses were performed on DEGs from blast reciprocal analyses between *V. ellipsiformis* and *U. peninsularis*. X-axis shows the number of DEGs contained in each GO category.

Unexpectedly, “reproduction” and “ubiquitination”, the two categories on which Ghiselli et al. (2012) previously focused because of their direct relationship with sex determination and DUI, were not significantly overrepresented in our results. We thus concentrated our effort on our reciprocal BLAST/DEG results.

Identification of candidate genes involved in sex determination

To identify candidate genes that contribute to sex determination/differentiation, or maintenance in mature gonads, we first searched in our reciprocal BLAST/DEG results for sex-specific expression patterns (i.e. transcribed only in one sex) – as expected for sex-determining genes (Table 1.2). In addition, we identified the top ten up-regulated (and annotated) genes in male and female gonads of both freshwater mussel species (Table 1.3).

Table 1.2. Sex-specifically expressed genes in male and female gonads of *Uterbackia peninsularis* and *Venustaconcha ellipsiformis*.

Unigene ID	Gene symbol	Description	Sex (fold change)
c43183_g1	-	Uncharacterized protein	M (10)
c63770_g1	-	Uncharacterized protein	M (10)
c106662_g8	TSSK3_MOUSE	Testis-specific serine/threonine-protein kinase 3	M (11)
c69491_g1	RDS2_CHICK	Photoreceptor outer segment membrane glycoprotein 2	M (12)
c106233_g1	EDR1_ARATH	Serine/threonine-protein kinase EDR1	M (12)
c55198_g1	CE049_MOUSE	Uncharacterized protein C5orf49 homolog	M (12)
c85464_g1	CALL3_MOUSE	Calmodulin-like protein 3	F (9)
c66486_g1	SMCO3_MOUSE	Single-pass membrane and coiled-coil domain-containing protein 3	F (13)
c237103_g1	-	Uncharacterized protein	F (13)

NOTE – Unigene ID indicates the names of the unigene given by Trinity to the assembly files. Gene symbol indicates the annotation format provided by the Uniprot database followed by a quick short description. Sex indicates in which sex the unigene is specifically expressed, fold change values (obtained from DESeq2 software) are based on the results obtained for the non-contaminated *U. peninsularis*.

Table 1.3. Top ten genes showing the greatest difference in expression in male and female gonads of *Utterbackia peninsularis* and *Venustaconcha ellipsiformis*.

Unigene ID	Gene symbol	Description	Sex (fold change)
c1056_g1	K1S6Q8_CRAGI	Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B	M (12)
c235886_g1	KLH10_HUMAN	Kelch-like protein 10	M (11)
c104796_g5	K1RU43_CRAGI	WD repeat-containing protein on Y chromosome	M (11)
c47982_g1	SAM15_MACFA	Sterile alpha motif domain-containing protein 15	M (11)
c847_g1	K1Q7E2_CRAGI	Sperm motility kinase X	M (11)
c46585_g1	K1RFU6_CRAGI	Proteasome activator complex subunit 3	M (11)
c173681_g1	K1RFD4_CRAGI	Probable 4-coumarate--CoA ligase 3	M (11)
c44943_g1	PYG_DROME	Glycogen phosphorylase	M (11)
c36459_g1	CC151_BOVIN	Coiled-coil domain-containing protein 151	M (11)
c110789_g1	K1R8H1_CRAGI	Testis-specific serine/threonine-protein kinase 5	M (11)
c69487_g1	LECG_THANI	Galactose-specific lectin nattectin;	F (14)
c92390_g1	LECG_THANI	Galactose-specific lectin nattectin;	F (14)
c82290_g1	LECM2_ERYPO	C-type lectin lectoxin-Lio2;	F (14)
c83693_g1	PSM_MYTCA	Shell matrix protein;	F (14)
c79356_g1	LECM1_PHIOL	C-type lectin lectoxin-Phi1;	F (14)
c99224_g1	PLC_HALLA	Perlucin;	F (13)
c173670_g1	LPSBP_PERAM	Hemolymph lipopolysaccharide-binding protein;	F (13)
c96153_g1	LECG_THANI	Galactose-specific lectin nattectin;	F (13)
c105339_g1	LECM2_ERYPO	C-type lectin lectoxin-Lio2;	F (13)
c55454_g1	CL17A_HUMAN	C-type lectin domain family 17, member A;	F (13)

NOTE – Unigene ID indicates the names of the unigene given by Trinity to the assembly files. Gene symbol indicates the annotation format provided by the SwissProt or Uniprot database followed by a short description. Sex indicates in which sex the unigene is up-regulated, fold change values (obtained from DESeq2 software) are based on the results obtained for the non-contaminated *U. peninsularis*.

Among these, six genes were exclusively expressed in male gonads and three in female gonads. Not surprisingly, these sex-specifically expressed and up-regulated genes appear to play roles in spermatogenesis and oogenesis, which are the ultimate biological processes in the male and female gonads, and in reproduction in general. In males, for example *Kelch-like protein 10* and *Proteasome activator complex* are members of the E3 ubiquitin-protein ligase complex process, which mediates the ubiquitination and subsequent proteasomal degradation of target proteins during spermatogenesis (Arama et al., 2007). In addition, *sperm motility kinase 2B* and *WD repeat-containing protein on Y chromosome* are involved in sperm fertility and motility, whereas *testis-specific serine/threonine-protein kinases 3* and *5* are involved in germ cell survival during meiosis

(e.g. Peng et al., 2015). Among other male-biased sequences, we also identified genes homologous to *spermatogenesis-associated protein* (*SPATA* 4, 17, 22 and 24), *sperm flagellar protein 1/2*, *sperm motility kinase*, *sperm-associated antigen and surface protein*, and *spermine oxidase* potentially expressed in the male reproductive tissues (Supplementary Table S1.4).

Specifically expressed and top up-regulated genes in females were mostly lectin-type genes, which are often highly represented in ovary transcriptomes in animal species, including bivalves. These genes have a number of biological functions such as immune defense, block to polyspermy at fertilization, and sperm-egg recognition (Luckenbach et al., 2008; Moy et al., 2008; Patnaik et al., 2016). Among other female-biased genes, we identified *vitellogenin*, *putative vitellogenin receptor*, and *Hsp90 organizing protein*, which are involved in vitellogenesis, a central process in oogenesis. We also found *cyclin-dependent kinase 2-associated protein 2*, which is involved in another vital process in oogenesis, i.e. oocyte maturation, and *superoxide dismutase* [Cu-Zn], which could neutralize reactive oxygen species and ensure oocyte quality (Peng et al., 2015 and above).

To better identify genes related to sex-determining pathways, we also searched in our reciprocal BLAST/DEG results for candidate genes previously identified in the literature in (i) the model organisms *Caenorhabditis elegans*, *Drosophila melanogaster*, *Mus musculus* (see Zhang et al., 2014 for the list), (ii) four bivalve species without DUI, i.e. the oysters *C. gigas* and *C. honkongensis*, the scallop *P. yessoensis* and the clam *Ruditapes decussatus* (Zhang et al., 2014; Tong et al., 2015; Li et al., 2016; Ghiselli et al. unpublished), and (iii) two bivalve species with DUI, i.e. the marine clams *R. philippinarum* and the freshwater mussel *H. schlegelii* (Ghiselli et al., 2012; Shi et al., 2015). Of the twenty-six key sex-determining pathway genes examined, putative homologs were found for eight genes or gene families in our two freshwater mussel species (Table 1.4; Supplementary Table S1.5). Of these, four genes (*Sry*, *Fem-1*, *MAB-3* and *FoxL2*) were differentially expressed but did not show sex-specific expression in our samples, although we cannot rule out the possibility that they may have sex-specific expression at an earlier developmental stage. For example, *Sry/Sox30* was found to be expressed only in male oysters and clams (Table 1.4).

Table 1.4. Key sex-determining pathway genes in model organisms and bivalve species.

Gene	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>M. musculus</i>	<i>C. gigas</i>	<i>C. hongkongensis</i>	<i>P. yessoensis</i>	<i>R. decussatus</i>	<i>R. philippinarum</i>	<i>H. schlegelii</i>	<i>U. peninsularis / V. ellipsiformis</i>
<i>Amh (Mmu)</i>			ms							
<i>Amhr2 (Mmu), Bar-1 (Cel), Arm (Dme), Ctnnb1 (Mmu)</i>										
<i>Cbx2 (Mmu)</i>			fs						fb	
<i>Doa (Dme)</i>		fs								
<i>Dsx (Dme), Dmrt1 (Mmu), Mab-3 (Cel)</i>	ms	ms	ms	ms	ms	mb		mb	mb	
<i>Fog-1,-2,-3 (Cel, Dme, Mmu)</i>	ms		ms							
<i>Fem-1 (Cel, Dme), Fem1b (Mmu)</i>	ms						mb		mb	mb
<i>FoxL2 (Mmu)</i>			fs	fs	fs	fb			fb	fb
<i>Fru (Dme)</i>		ms								
<i>Zglp1 (Mmu)</i>										
<i>Her-1 (Cel), Her (Dme)</i>	fs									
<i>Nr0b1 (Mmu)</i>			ms							
<i>Rspo1 (Mmu)</i>			fs						fb	
<i>Rnt-1 (Cel), Run (Dme)</i>		fs								
<i>Sdc (Cel, Dme, Mmu)</i>	fs									
<i>Sxl (Dme)</i>		fs								
<i>Sry/Sox30/SoxH (Mmu)</i>			ms	ms	ms	mb	ms	ms	mb	mb
<i>Sis-a (Dme)</i>		fs								
<i>Sf1 (Mmu)</i>			ms							
<i>Sox9 (Mmu), Sox100B (Dme)</i>			ms						mb	
<i>Gata4 (Mmu)</i>							mb		mb	
<i>Tra-1 -2 -3 (Cel), Tra2 (Dme)</i>	fs	fs							fb	
<i>Wt1 (Mmu)</i>			ms						mb	
<i>Wnt4 (Dme, Mmu)</i>			fs						fb	
<i>Xol-1 (Cel)</i>	ms									

NOTE – Names in parentheses indicate the species in which each gene has been identified; shaded box indicates presence of the gene in the genome/transcriptome; fs indicates female-specific expression (i.e. transcribed only in females); ms indicates male-specific expression; fb indicates female-biased expression; mb indicates male-biased expression. See supplementary Table S5 for a short description of each gene.

As previously mentioned, both *Sry* (sex determining region-y)-box 30 (*Sox 30*) and *Dmrt1* (*MAB-3* and *dsx* related transcription factor 1) are known to be involved in male sex determination in nematodes, fruit flies, and vertebrates (Kopp, 2012; Wallis et al., 2008), whereas *FoxL2* is known

for its role in ovarian determination in vertebrates (Uhlenhaut & Treier, 2006). *Fem-1* encodes an ankyrin repeat-containing protein orthologous to human *FEM1A* and is involved in male sex-determination in *C. elegans*, and also in the protein ubiquitination pathway (Doniach & Hodgkin, 1984). Two other genes that seem well represented in bivalves, and more generally in mollusks (Kim et al., 2017), are *Bar-1/Arm/Ctnnb1* (*Beta-catenin*) and *Tra-1* (Table 1.4). *Tra-1* controls female somatic sex differentiation in Drosophila, whereas Drosophila *Beta-catenin* homolog is named *Armadillo* and is known to be involved (among other processes) in oogenesis and ovarian follicle cell development. The *Rspo1/Wnt/Beta-catenin* pathway is also involved in bird sex determination (Beukeboom & Perrin, 2014).

Collectively, the results presented in Table 1.4 suggest shared sex-determining pathway genes in distantly-related, as well as in SMI and DUI bivalves. Zhang et al. (2014) suggested a deeply conserved role in sex determination for *Dmrt1/Dsx/MAB-3* in both invertebrates and vertebrates, but also for *Sry* and *FoxL2*, even if they were thought to be new recruits to sex-determining pathways in vertebrates or placental mammals (Gamble & Zarkower, 2012; Matson & Zarkower, 2012). Indeed, some steps of the sex determination pathway may be deeply conserved in metazoans, despite rapid evolution of the regulatory pathways that in bivalves, for example, may involve both genetic and environmental factors (Zhang et al., 2014; Breton et al., 2018).

Zhang et al. (2014) proposed a working model for sex determination in the oyster *C. gigas*, in which *CgSoxH*, a *Sry*-like gene that is strictly expressed in testis, would play a leading role in the sex-determining pathway by directly or indirectly activating *CgDsx*, a DM domain gene like those (e.g. *Dmrt1*) that have been identified as master switches for testis development in animal species studied so far. Both *CgSoxH* and *CgDsx* would interact with, or inhibit, *CgFoxL2*, which is usually specifically expressed in ovaries (this scenario is consistent with the reported interaction among *Sry*, *Dmrt1*, and *FoxL2* in mammals). Even if the number of species studied is still limited, we suggest that this general model probably applies to the majority of bivalve species. We also propose that major interspecific differences in bivalves will be found in cis/trans genetic elements and/or environmental factors that may control these three “sex-determining key genes”. For example, a mito-nuclear sex determination system in which mitochondrially-encoded elements distorting sex ratios has been hypothesized to occur in species with DUI (e.g. Breton et al., 2011a;

Milani et al., 2016; Pozzi et al., 2017). In both plants and animals, cases of mitochondrial genomes acting on germ line development and/or with sex ratio distortion properties have been documented (e.g. Chase, 2007; Perlman et al., 2015). In bivalves with DUI, the existence of two sex-specific mitochondrial genomes on which selection can operate represents a possible source of novel sex determination mechanisms.

Identification of candidate genes involved in DUI

As mentioned earlier, a modification of the ubiquitin-proteasome system has been hypothesized as responsible for the retention of sperm mitochondria in male embryos of species with DUI (Kenchington et al., 2002; Ghiselli et al., 2012). To identify putative factors that could be involved in the degradation of sperm mitochondria through ubiquitination in DUI bivalves, Milani et al. (2013) analyzed structure and localization of three transcripts – i.e. baculoviral IAP repeat-containing 4 (*birc*), proteasome subunit alpha 6 (*psa*), and AN1 zinc finger ubiquitin-like domain (*anubl1*) – previously shown to present sex-specific and family biases (i.e. family producing predominantly females or males) in the Manila clam *R. philippinarum*. *In situ* hybridization confirmed the localization of these transcripts in gametogenic cells (Milani et al., 2013). Also, homologs of these genes were shown to be involved in reproduction and ubiquitination in other animals: *PSA* is a subunit of the proteasome, a complex known to be involved in male sexual differentiation, whereas *ANUBL1* possesses an ubiquitin domain, and *BIRC* acts as an inhibitor of apoptosis (Milani et al., 2013). For these reasons, the authors hypothesized that these genes could have a role in sex determination and could also be responsible for the maintenance/degradation of spermatozoon mitochondria during embryo development of the DUI species *R. philippinarum*. It is worth mentioning that *PSA* has also been reported in male-biased families in the DUI species *Mytilus edulis* (Diz et al., 2013), supporting the hypothesis of an involvement of the ubiquitination system (and of the proteasome subunit alpha) in sex determination in bivalves.

To identify genes related to the DUI mechanism, we searched our reciprocal BLAST results for genes previously suggested to be involved in this process in the DUI species *R. philippinarum* and *M. edulis*, i.e. for genes involved in ubiquitination that show male-biased transcription, indicating their potential role in spermatogenesis (Ghiselli et al., 2012; Diz et al., 2013; Milani et

al., 2013). Annotation, function, and expression ratio of male-biased (and female-biased) ubiquitination genes are reported in Supplementary Table S1.1. Compared to females in which 18 ubiquitination genes were differentially expressed, 53 ubiquitination genes were male-biased. Among them, *psa6* and *anubl1*, two male-biased ubiquitination genes were absent from female-biased genes, as were *birc5* and *birc2*. These results, consistent with the hypothesis of Milani et al. (2013), suggest that *psa6* might be important in sex determination and/or differentiation in DUI species, that *anubl1* could be a factor tagging sperm mitochondria that differentiate them from egg mitochondria, and that *birc* could be an additional mitochondrial tag that would protect the midpiece mitochondria from degradation. Our results are also consistent with our hypothesis that sets of genes associated with sex determination and DUI are similar in distantly-related DUI species, as exemplified here with freshwater mussels vs. a marine clam and a marine mussel. However, additional RNA- and protein-based studies such as those conducted by Milani et al. (2013) and Diz et al. (2013) will be needed to confirm our results.

Lastly, although speculative, one of the factors that could also influence inheritance of mitochondrial genes is DNA methylation. For example, in the green alga *Chlamydomonas*, maternal inheritance of chloroplast DNA is thought to be influenced by DNA methylation (Sager & Grabowy, 1983; Umen & Goodenough, 2001). A growing body of literature has also shown that epigenetic modifications, such as DNA methylation, can also regulate sex determination and differentiation (see Liu et al., 2015). Strikingly, in the leafhopper *Zyginidia pullula*, the symbiont *Wolbachia pipiensis* disrupts male imprinting by modifying methylation patterns in gonads; as a result, *Wolbachia* feminizes genetic males of *Z. pullula* (Negri et al., 2009). In our results, transcripts associated to DNA methyltransferases, histone deacetylases and histone acetyltransferases were identified, and most showed male-biased expression (Supplementary Table S1.1). Among these male-biased transcripts, *DNMT1*, a DNA methyltransferase associated to mitochondria, was detected. However, because the epigenetic mechanisms underlying organellar DNA inheritance and/or sex determination/differentiation are still poorly understood, we cannot infer any detailed functions of these genes from their expression patterns. Future studies, currently underway in our laboratory, are clearly needed to reveal the complex interplay between methylation, sex bias, and mitochondrial inheritance in bivalves with DUI.

Conclusion

In this study, we present the first comparative gonadal transcriptomic analysis in two freshwater mussel species aimed to assess the link between doubly uniparental inheritance of mtDNA and sex determination in Unionoida. Using a BLAST reciprocal analysis to identify orthologs between *V. ellipsiformis* and *U. peninsularis* and comparing our results with previously published sex-specific transcriptomics data from distantly-related DUI species, we were able to identify elements possibly involved in the regulation of sex-specific aspects of DUI systems. Our search for candidate genes involved in sex determination suggests that some steps of the sex determination pathway may be deeply conserved in metazoans, whereas our search for candidate genes involved in DUI support previous hypotheses that a modification of the ubiquitination mechanism could be responsible for the retention of the paternal mtDNA in male bivalves. Our results also suggest that DNA methylation could be involved in the maintenance of DUI in bivalves.

Acknowledgments

This work was supported by Natural Sciences and Engineering Research Council Discovery Grants awarded to S Breton [grant number RGPIN/435656-2013], and DT Stewart [grant number RGPIN/217175-2013]. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supplementary Material

Supplementary tables

Table S1.1. Blast reciprocal results.

En fichier complémentaire, nommé « Tables Suppl Chapitre 1 », accessible au premier feuillet.

Table S1.2. Ontology categories significantly enriched in both species.

En fichier complémentaire, nommé « Tables Suppl Chapitre 1 », accessible au deuxième feuillet.

Table S1.3. GOSeq results for *U. peninsularis*.

En fichier complémentaire, nommé « Tables Suppl Chapitre 1 », accessible au troisième feuillet.

Table S1.4. GOSeq results for *V. ellipsiformis*.

En fichier complémentaire, nommé « Tables Suppl Chapitre 1 », accessible au quatrième feuillet.

Table S1.5. Short description of key sex determining genes in animals.

En fichier complémentaire, nommé « Tables Suppl Chapitre 1 », accessible au cinquième feuillet.

Supplementary figures

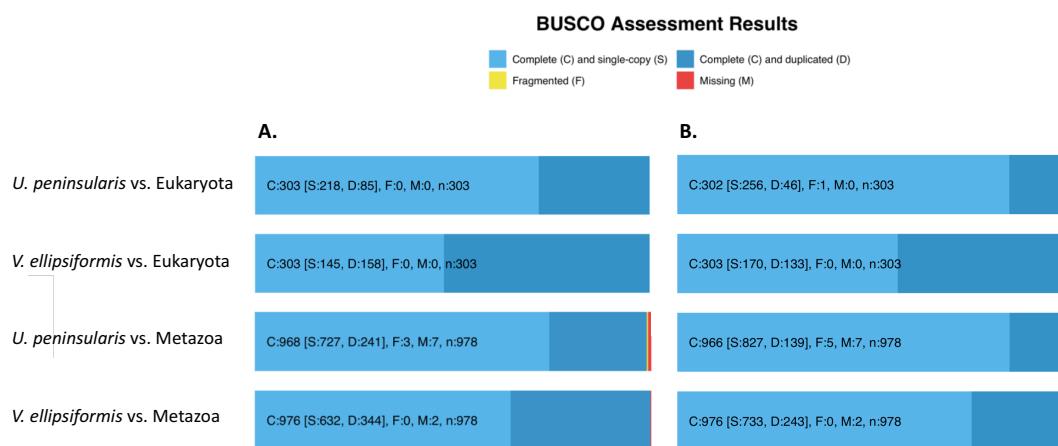


Figure S1.1. The results of BUSCO analysis of *U. peninsularis* and *V. ellipsiformis* transcriptomes. (A) Analysis of contigs. (B) Analysis of unigenes. Light-blue: complete and single-copy genes; dark-blue: complete and duplicated genes; yellow: fragmented genes; red: missing genes.

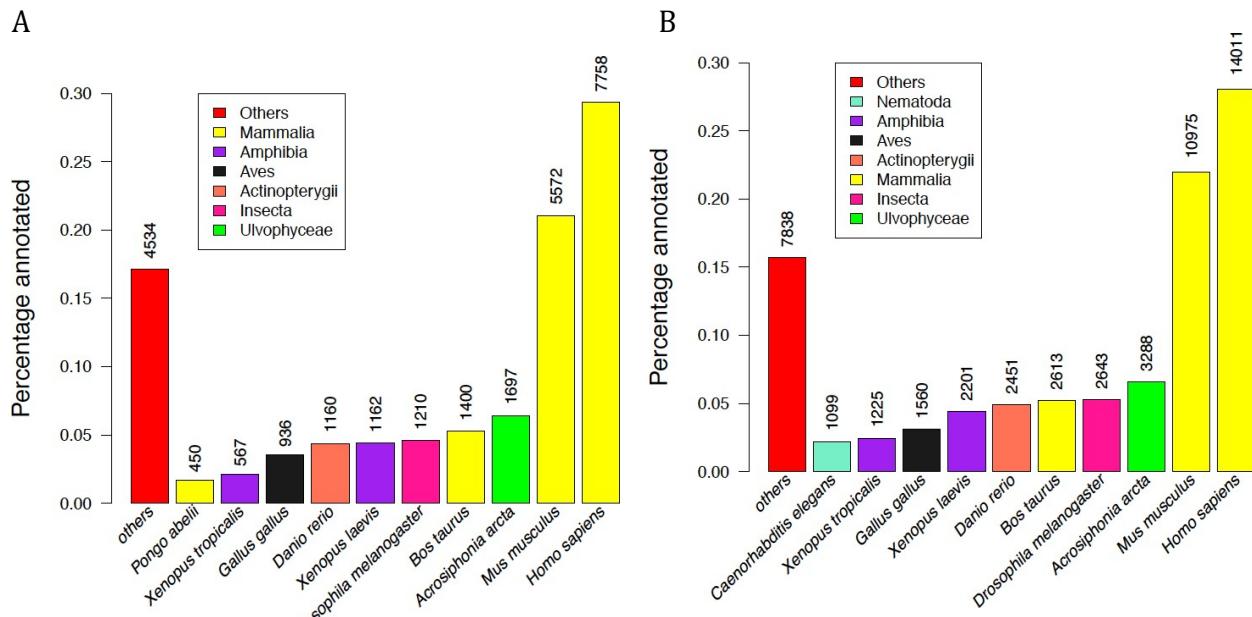


Figure S1.2. Taxonomic distribution of top BLASTp hits in SwissProt protein database for *U. peninsularis* contigs (A) and *V. ellipsiformis* contigs (B). Numbers of unique hits in each group are shown.

Chapitre 2.

***Détermination mitochondriale du sexe
présumée chez Bivalvia : Informations tirées
d'un assemblage hybride de transcriptomes
d'espèces de moules d'eau douce***

Putative mitochondrial sex determination in the Bivalvia: insights from a hybrid transcriptome assembly in freshwater mussels

Charlotte Capt^{1*}, Sébastien Renaut^{1,2}, Donald T. Stewart³, Nathan A. Johnson⁴, and Sophie Breton^{1*}

¹ Department of Biological Sciences, Université de Montréal, Montréal, QC, Canada

² Centre de la Science de la Biodiversité du Québec, Université de Montréal, Montréal, QC, Canada

³ Department of Biology, Acadia University, Wolfville, NS B4P 2R6 Canada

⁴ Wetland and Aquatic Research Center, U.S. Geological Survey, Gainesville, FL, USA

Published in Frontiers in Genetics 10:840.

doi:10.3389/fgene.2019.00840.

Minor edits to the published text have been made.

Author Contributions

CC, SR, and SB designed experiments. NJ conducted sampling. CC performed laboratory experiments. CC and SR analyzed data. CC, SR, NJ, DTS and SB wrote the manuscript. SB conceived and managed the project. All authors read and approved the final version of the manuscript.

Abstract

Bivalves exhibit an astonishing diversity of sexual systems, with genetic and environmental determinants of sex, and possibly the only example of mitochondrial genes influencing sex determination pathways in animals. In contrast to all other animal species in which strict maternal inheritance (SMI) of mitochondria is the rule, bivalves possess a system known as doubly uniparental inheritance (DUI) of mitochondria in which maternal and paternal mitochondria (and their corresponding female-transmitted or F mtDNA and male-transmitted or M mtDNA genomes) are transmitted within a species. Species with DUI also possess sex-associated mtDNA-encoded proteins (in addition to the typical set of 13), which have been hypothesized to play a role in sex determination. In this study, we analyzed the sex-biased transcriptome in gonads of two closely-related freshwater mussel species with different reproductive and mitochondrial transmission modes: the gonochoric, DUI species, *Utterbackia peninsularis*, and the hermaphroditic, SMI species, *Utterbackia imbecillis*. Through comparative analysis with other DUI and non-DUI bivalve transcriptomes already available, we identify common male and female-specific genes, as well as SMI and DUI-related genes, that are probably involved in sex determination and mitochondrial inheritance in this animal group. Our results contribute to the understanding of what could be the first animal sex determination system involving the mitochondrial genome.

Introduction

Only one putative sex determination system involving mitochondria and their genome (mtDNA) has been reported in animals (Breton et al., 2011a; Breton et al., 2018). It implicates the unorthodox system of doubly uniparental inheritance (DUI) of mitochondria in bivalve molluscs, which is also the only known exception to strict maternal inheritance (SMI) of mitochondria in animals (Birky, 2001; Breton & Stewart, 2015). Over 100 DUI species have been reported to date (Gusman et al., 2016). In these taxa, both maternal and paternal mitochondria (and their corresponding female-transmitted or F mtDNA and male-transmitted or M mtDNA) are transmitted uniparentally through eggs and sperm, respectively (see Breton et al., 2007; Passamonti & Ghiselli, 2009; Zouros, 2013 for reviews). In all DUI species, the F mtDNA is found in female gonads and in the somatic tissues of both sexes, whereas the M mtDNA is found in sperm and is sometimes detected in small quantities in male soma (see Breton et al., 2017 for a review). The M and F mtDNAs can be extremely divergent, with up to 40% intraspecific DNA divergence (Doucet-Beaupré et al., 2010; Guerra et al., 2017). The mechanism(s) by which such molecular divergence of mtDNAs is tolerated within a species are of considerable interest considering that heteroplasmy, *i.e.* the presence of two or more different mtDNAs within an organism, can cause severe disease or death in humans (Picard, Wallace, & Burelle, 2016; Wallace, 2016).

Species with DUI also possess sex-associated mtDNA-encoded proteins (in addition to the typical set of 13) that are not apparently homologous to any other known genes (*i.e.* they are mitochondrial ORFans) (Breton, Ghiselli, et al., 2011; Breton, Stewart, & Blier, 2009; Milani et al., 2013). Specifically, the F mtDNA contains the *F-orf* gene and the M mtDNA contains the *M-orf* gene, both of which are relatively conserved across species within a family (Breton et al., 2009, 2011; Milani et al., 2013; Mitchell et al., 2016). In freshwater mussels (Unionida), gonochorism is perfectly correlated with the presence of DUI and these novel sex-specific proteins, whereas hermaphroditic species lack the M mtDNA (*i.e.* they possess SMI of mtDNA) and have macromutations in the region of their mtDNA genome that is homologous to the *F-orf* gene of the F mtDNA in their gonochoric relatives (Breton et al., 2011; Breton et al., 2014). In keeping with the transition from gonochorism to hermaphroditism, the mitochondrial genome present in these species is referred to as an H mtDNA and the gene of interest is known as an *H-orf* (Breton et al.,

2011). These mtORFans have been hypothesized to be part of the sex determination system (Breton et al., 2011), and if true, this would make DUI the first animal sex determination system involving the mtDNA (heteromorphic sex chromosomes are absent in bivalves). However, we are still far from fully understanding the underlying link between DUI and sex determination.

A first transcriptomic study of mature male and female gonads of the DUI clam *Ruditapes philippinarum* proposed a relationship between sex, mitochondrial inheritance, and ubiquitination (Ghiselli et al., 2012). Ubiquitination is well known for being involved in the elimination of paternal mitochondria in mammals (Sutovsky et al., 1999). It is also known to play a role in sex determination in *Caenorhabditis elegans* and *Drosophila* (Ghiselli et al., 2012). Ghiselli et al. (2012) found a relationship between sex and differential expression of several ubiquitination genes in *R. philippinarum*, making them likely candidate genes for the regulation of sex-specific aspects of the DUI system. Following this work, Milani et al. (2013) focused on three of these genes that showed the most significant transcriptional bias and the highest levels of transcription (baculoviral IAP repeat-containing 4 [*birc*]; proteasome subunit alpha 6 [*psa*]; AN1 zinc finger ubiquitin-like domain [*anubl1*]). They hypothesized that these genes could play a role in sex determination and be responsible for the maintenance or degradation of sperm mitochondria during embryo development in DUI species. In addition, Capt et al. (2018) hypothesized that if a modification of the ubiquitination mechanism is responsible for the retention of the paternal mtDNA in male bivalves, then molecular signatures of this modification should be discernable in all DUI species. The same three differentially expressed candidate genes were indeed identified in their comparative analysis of male and female gonad transcriptomes for the two DUI freshwater mussel species *Venustaconcha ellipsiformis* and *Utterbackia peninsularis* (Capt et al., 2018). This study also revealed that DNA methylation could be involved in the regulation of DUI (Capt et al., 2018) and/or in sex determination and differentiation, as seen in other species (Capt et al., 2018; Liu et al., 2015).

Ubiquitination is not the only process involved in SMI of mitochondria in animals. The mechanisms underlying SMI are widely variable and suggest recurrent loss and restoration and/or several independent origins of this nearly ubiquitous phenomenon (Birky, 2001; Breton & Stewart, 2015). For example, paternal mtDNA can be eliminated by a nuclease-mediated mechanism either

during spermatogenesis or after fertilization, or paternal mitochondria can be selectively degraded by autophagy after fertilization (Sato & Sato, 2017). To better understand the molecular basis of DUI, Punzi et al. (2018) selected classes of proteins known to be involved in the maintenance of SMI and compared their features in two closely-related gonochoric clam species differing in their mitochondrial inheritance mechanism, that is, the SMI species *Ruditapes decussatus* and the DUI species *R. philippinarum*. Their analysis of gonad transcriptomes focused on nucleases and polymerases, ubiquitination, and proteins involved in autophagy/mitophagy. Their results provided no evidence for a role of nucleases/polymerases or autophagic machinery in the enforcement of SMI in *R. decussatus*, but transcripts of ubiquitinating enzymes were retrieved, providing additional support to previous transcriptomic studies (Punzi et al., 2018).

In the present study, we take advantage of the existence of closely-related gonochoric-DUI and hermaphroditic-SMI freshwater mussel species to better understand the genetic mechanisms that govern DUI and sex determinations in bivalves. Transcriptomic analyses were performed on three types of gonads, which will also be referred to as “sexes”, i.e. hermaphroditic gonad samples of *Utterbackia imbecillis* and available male and female gonadal transcriptome data from the species *U. peninsularis* (Capt et al., 2018). We produced a hybrid assembly that included all libraries from both species, which was used as a reference to map the sequencing reads of each sex. Male-female (MF), hermaphrodite-female (HF) and hermaphrodite-male (HM) comparisons were performed to identify candidate genes involved in sex determination and/or mitochondrial inheritance. In light of information from previous studies (Ghiselli et al., 2012; Milani et al., 2013; Capt et al., 2018; Punzi et al., 2018), particular attention was given to the following gene ontology (GO) categories to identify candidate genes putatively involved in sex determination or mitochondrial inheritance: reproduction, ubiquitination, nucleases, autophagy/mitophagy and methylation.

Methods

Sample Collection and RNA Extraction

Adult *U. imbecillis* mussels used for this study were collected in the Suwannee River (Florida, USA; Lat 29.58684, Long -82.94095) in July 2014. All mussels were kept alive during their shipment to the Université de Montréal (QC, Canada). Mussels were dissected and gonads inspected under the microscope. Only mature gonads were kept in the current study and frozen at -80°C until further analyses.

Total RNA extraction from frozen gonad tissues was performed as described in Capt et al. (2018) on eight *U. imbecillis* samples. Quality and quantity of extracted RNA was assessed via electrophoresis on a 1% agarose gel and using a BioDrop µLITE spectrophotometer; one sample with an absorbance reading A260/A280 below two was discarded (i.e. seven “high quality” RNA samples remained). Because *U. imbecillis* specimen can be confused visually with *U. peninsularis* (both species can be found in this sampling site), to validate that these samples were true hermaphrodites, the presence of the mitochondrial hermaphrodite open reading frame (*h-orf*) specific to *U. imbecillis* (Breton et al., 2011) was verified by PCR (primer sequences: UttimbHorf_F 5'-ATCTTTCTAATTAAGTGTATAACG-3' and UttimbHorf_R 5'-TTGGGCTAGGTTGTTACGG-3') and sequencing. Three PCR-confirmed *U. imbecillis* samples were sent to the Institute for Research in Immunology and Cancer Institute (IRIC, Université de Montréal, Montréal, Canada) for Illumina library preparation (mRNA-Seq) and paired-end (2 X 100 bp) sequencing (Illumina HiSeq2000; San Diego, CA). Paired-end sequencing was chosen for these three samples (no replicate) because it increases the quality of *de novo* assembly in non-model species and allows for more accurate isoform analyses (Conesa et al., 2016). The same procedure was used in Capt et al. (2018) for *U. peninsularis* (four female and four male gonad samples without replicates).

Reads Quality Assessment, De Novo Assemblies and Completeness Appraisal

Paired-end reads were trimmed with Trimmomatic (v0.32; Bolger et al., 2014) with the options ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:6:10 MINLEN:36. Illumina paired-end adaptors were removed and trailing and leading sequence ends were trimmed if quality (Phred score) dropped below 30. If mean quality dropped below ten in a six base wide sliding window, nucleotides were replaced by N, and finally,

if a trimmed read dropped below a length of 36bp, both paired-end reads were removed. The quality of reads was controlled and visually inspected with FASTQC (v0.11.2; Andrews, 2010), as described in Capt et al. (2018).

Two distinct *de novo* assemblies were performed with the trimmed reads for each species, using Trinity (Haas et al., 2013). The analytical pipeline used to obtain the transcriptome of *U. imbecillis* and the main reference transcriptome, which is referred to as the “hybrid” transcriptome for the rest of the study, is the same as the one previously described for *U. peninsularis* (Capt et al., 2018), i.e. with a minimum length fixed at 300pb for each contig, and other parameters left as default. This was done to allow TransDecoder to recognize Open Reading Frames (ORFs) and reduce the rate of false positives (<https://transdecoder.github.io/>). The hybrid transcriptome was obtained with the reads of eleven individuals (i.e. our three *U. imbecillis* hermaphrodites and four male and four female *U. peninsularis*), and allowed for differential gene expression analysis among the three sexes to avoid bias that would show up from comparing two different transcriptomes. A similar approach was used to study head transcriptomes of two fruit fly species of the genus *Anastrepha* (Rezende et al., 2016). Most analyses were thus performed on the hybrid transcriptome whereas the gonochoric (*U. peninsularis*) and the hermaphrodite (*U. imbecillis*) transcriptomes were used to assess the validity of the hybrid transcriptome, or to check for the presence of species-specific genes. The hybrid transcriptome underwent a clustering of contigs that shared >90% similarity, using CD-HIT-EST (v4.6; Li & Godzik, 2006). This step was added to avoid having multiple orthologous copies of genes from *U. imbecillis* and *U. peninsularis* in the hybrid assembly, and consequently false-positives in differential gene expression analyses. The cutoff choice was based on the divergence between mitochondrial female and hermaphrodite nucleotide sequences.

The degree of completeness of the transcriptomes was assessed with Benchmarking Universal Single-Copy Ortholog (BUSCO) software (busco_v3, mode -tran). Two core gene databases, metazoa (978 genes) and eukaryota (303 genes), were used and compared with the Trinity-based *de novo* assemblies. This comparison measures the presence of near-universal highly conserved single-copy orthologs, and provides information about the completeness of the transcriptomes based on what is expected, depending on the database (Simão et al., 2015).

Reads Alignment and Differential Gene Expression

Raw reads from each individual were aligned to the reference hybrid transcriptome using Bowtie2 (v2.2.4; Langmead & Salzberg, 2012) with default parameters. Next, reads were sorted, indexed, and the corresponding numbers of mapped reads by contig were extracted with Samtools software (v1.1; H. Li et al., 2009), following the pipeline in Capt et al. (2018). Results were normalized based on Robinson & Oshlack (2010). Given that we previously observed that both DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2010) provide similar gene expression patterns (Capt et al., 2018), only edgeR (version edgeR_3.12.0) was used in the current study to simplify analyses and interpretation. EdgeR selects for a more restricted number of differentially expressed genes (DEGs) and increases identification of true positives (Soneson & Delorenzi, 2013; Z. H. Zhang et al., 2014). DESeq2 was used only to acquire fragments per kilobase of exon per million fragments sequenced (FPKM) values, using “robust = F” to sum the raw counts. DEGs were obtained with edgeR by selecting all contigs with a FPKM value greater than 1 and a \log_2 fold-change ≥ 1 or ≤ -1 in one of the three sexes, for each paired-comparison. Gene expression comparisons were performed two per two between hermaphrodites and males (HM), hermaphrodites and females (HF), and males and females (MF). To do so, contrast was fixed at 1 for one sex for each comparison, -1 for the other, and 0 for the sex absent in the comparison, and no intercept was given for the design because there was no ‘control’ value in the current design.

A False Discovery Rate (FDR) adjustment (Benjamini & Hochberg, 1995) was performed to account for multiple comparisons; only values with a corrected p-value <0.05 were kept as DEG. Unigenes that were expressed only in one sex, i.e. with a FPKM value > 1 in one sex and a FPKM value of 0 for the two others, were considered as sex-specific. When genes were hermaphrodite-specific, the associated annotation was checked in the *U. imbecillis* transcriptome to confirm its validity and specific role in hermaphrodites. A Principal Component Analysis (PCA) was also performed on regularized log transform (rlog) values of the thousand first most expressed genes, based on their FPKM, to verify if individuals were clustering according to their sex.

Annotations and Gene Ontology Enrichment

Assembled contigs from each transcriptome were searched against protein and nucleotide sequences databases. First, potential ORFs, and their linked protein-coding sequences, were predicted using TransDecoder (r20140704; Haas & Papanicolaou, 2017). If several ORFs were found in one contig, the longest was selected. This new TransDecoder list of ORFs allowed a BLASTp search (*e*-value <1e-5) against protein databases (Swiss-prot and TrEMBL [uniref90 cluster]) available in the Universal Protein Resource (UniProt, April 2018). Swiss-prot is a high quality, manually annotated and non-redundant protein sequence database that focuses mainly on model species to ensure quality annotation. Uniref90 provides computationally generated annotations unfound in Swiss-prot and allows a higher annotation success, but the database is not manually curated. To maximize the hybrid transcriptome annotation success, an additional BLAT search (BLAST-like alignment tool) was performed against NCBI non-redundant (NR) protein database (Kent, 2002). A BLASTx search (*e*-value <1e-5) on unigene sequences against Swiss-prot nucleotide database was also performed to complete the annotation (this additional step was also done for *U. peninsularis* [Capt et al., 2018]).

Gene ontology enrichment analyses were performed by combining annotations and differential gene expression information using GOseq (v1.28.0; Young et al., 2010). After FDR correction on *p*-values, GO terms significantly over-represented were identified with a FDR corrected *p*-value < 0.05.

Results and Discussion

Transcriptome Assemblies and Completeness

A total of 188 337 818 paired-end reads was produced for the three *U. imbecillis* samples and 348 652 616 for male and female *U. peninsularis* samples (Capt et al., 2018), giving a total of 536 990 434 paired-end reads for both species. The FASTQ raw sequences are accessible under accession numbers SRR8782945-7 (*U. imbecillis*) and SRR6279376-83 (*U. peninsularis*; Capt et al., 2018)

in the National Centre for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database.

Table 2.1. *De novo* assembly quality and annotation statistics of *Utterbackia imbecillis*, *Utterbackia peninsulae* and the hybrid transcriptomes.

	Hybrid	<i>U. imbecillis</i>	<i>U. peninsulae</i> (Capt et al. 2018)
Assembly statistics			
Total raws	536,990,434	188,337,818	348,652,616
Average raws	44,749,203	47,084,454	43,581,577
Total trimmed raws	439,968,848	151,648,110	288,320,738
Average trimmed raws	36,664,071	37,912,028	36,040,092
Total trinity 'genes'	197,379	111,958	165,788
Total trinity transcripts	215,908	130,160	200,961
Percent GC	34.71	35.23	33.36
Contig N10	4,222	5,326	5,581
Contig N20	2,883	3,724	3,895
Contig N30	2,123	2,804	2,916
Contig N40	1,551	2,144	2,224
Contig N50	1,128	1,623	1,655
Median contig length	497	557	558
Average contig length	834.10	1,015.56	1,026.46
Total assembled bases	180,089,785	132,185,742	206,277,984
Contig N10	4,124	4,765	4,882
Contig N20	2,794	3,278	3,298
Contig N30	2,014	2,430	2,409
Contig N40	1,444	1,808	1,750
Contig N50	1,034	1,333	1,251
Median contig length	482	516	510
Average contig	798.47	904.79	884
Total assembled bases	157,600,779	101,298,523	146,628,566
Annotations statistics			
TransDecoder	30,964	31,679	40,360
Swiss-prot - BLASTp	26,648/215,908 (12.3%)	20,683/130,160 (15.9%)	26,469/200,961 (13.1%)
Uniref90 - BLASTp	32,430/215,908 (15.0%)	24,982/130,160 (19.2%)	32,515/200,961 (16.2%)
Nr - BLATp	19,274/215,908 (8.9 %)	-	-
Swiss-prot - BLASTx	21,338/215,908 (9.9 %)	-	28,495/200,961 (14.2%)

After the trimming process, 439 968 848 high-quality raw reads (81.9%) were retained to build a hybrid assembly of 215 908 transcripts (after the 90% clustering with CD-HIT; before this

step, the hybrid assembly provided 245 596 transcripts). The final hybrid transcriptome had an average contig length of 798 bp and a N50 of 1 034. The number of transcripts and median contig length were similar to the values obtained for *U. peninsularis* (Capt et al., 2018). Complete statistics for all transcriptomes are available in Table 2.1.

Supplementary Figure S2.1 shows the results of the completeness assessment performed using BUSCO. This analysis resulted in 97.8% completely recovered genes for *U. imbecillis* (20.6% duplicated, 1.6% fragmented, 0.6% missing), and 98.4% completely recovered genes for the hybrid transcriptome (32.2% duplicated, 0.6% fragmented, 1% missing). Similar results were also obtained for *U. peninsularis* (Capt et al., 2018). These results validated the completeness of our assemblies and allowed further and reliable analyses of gene expression from the hybrid transcriptome.

Annotation

Transdecoder recognized 30 964 ORFs from the hybrid transcriptome, which were used to perform BLASTp searches against Swiss-Prot and Uniref90, and BLAT searches against NCBI NR. A total success annotation of 12.3 %, 15% and 9% were respectively obtained with these three approaches (Table 2.1). BLASTx against Swiss-prot annotated 9% of the unigenes. Most hits came from mollusk species (data not shown), particularly from the oyster *Crassostrea gigas* for which whole-genome sequencing and annotation were recently completed (G. Zhang et al., 2012). When compared with the recently reported genomes for the freshwater mussels *Venustaconcha ellipsiformis* (Bivalvia Unionida: Renaut et al., 2018) and *Limnoperna fortunei* (Bivalvia: Mytilidae; Uliano-Silva et al., 2018), both annotated and assessed for their completeness using the same approach that we did (i.e., respectively with Benchmarking Universal Single-Copy Ortholog [BUSCO] software and BLASTing against uniprot and NCBI NR), our results with BUSCO were comparable and even a bit better (81.9%-85% of complete BUSCOs for Renaut et al. (2018) and Uliano-Silva et al. (2018) and 97.7% complete BUSCOs for the present study). Similarly, Uliano-Silva et al. annotated 26 198 out of 60 717 predicted protein-coding genes (43%) and Renaut et al. (2018) annotated around 39% of predicted protein genes (>300bp) using BLAST against Uniprot and, for example, we annotated 24 982 of our *U. imbecillis* transcripts (out of 130,160 = 19.2%).

Our percentage of annotation is rather low because we considered all our transcripts (as in other transcriptomic studies). However, if we consider only ORFs identified by Transdecoder (i.e., 31 679 ORFs >300bp from our transcriptome), our percentage of annotation would have been comparable to Uliano-Silva et al. (2018) and Renaut et al. (2018). Also, our results were similar to what was obtained for the gonadal transcriptome of *U. peninsularis* (Capt et al., 2018). Success annotation in *U. imbecillis* was also similar to what was obtained for *U. peninsularis* and the hybrid transcriptome (Table 2.1). Very few hits came from bacteria or trematodes species, indicating that our gonad samples were not contaminated. Overall, our results are similar to those reported in other transcriptomics studies of bivalve species (Patnaik et al., 2016; J. Shi et al., 2015; Wang et al., 2017; Yarra et al., 2016), and confirm the difficulty to annotate non-model species, given that few well-characterized model organisms are closely related to bivalves.

Screening of Differentially Expressed and Sex-Specific Genes

Differential gene expression analyses showed that the highest number of differentially expressed genes (DEGs) was observed between hermaphrodites and females (16 573 for the HF comparison), with 9 818 genes up-regulated in hermaphrodites and 6 755 in females (Figure 2.1). We found 12 352 DEGs between males and females (MF comparison), with 7 710 genes up-regulated in males and 4 642 in females. Finally, 9 762 genes were differentially expressed between hermaphrodites and males (HM comparison), with 5 626 genes up-regulated in hermaphrodites and 4,136 in males. A higher number of DEGs in hermaphrodites is in agreement with previous studies (e.g. Y. Shi, Liu, & He, 2018), and can be explained by the production of both gametes in hermaphrodite gonads. Figure 2.1 shows the overlap in DEGs for each comparison. For example, for the HF comparison 16 573 DEGs were found, 6 370 of which were also DEGs in the MF comparison, 1 682 were also DEGs in the HM comparison, and 1 423 were DEGs in all comparisons.

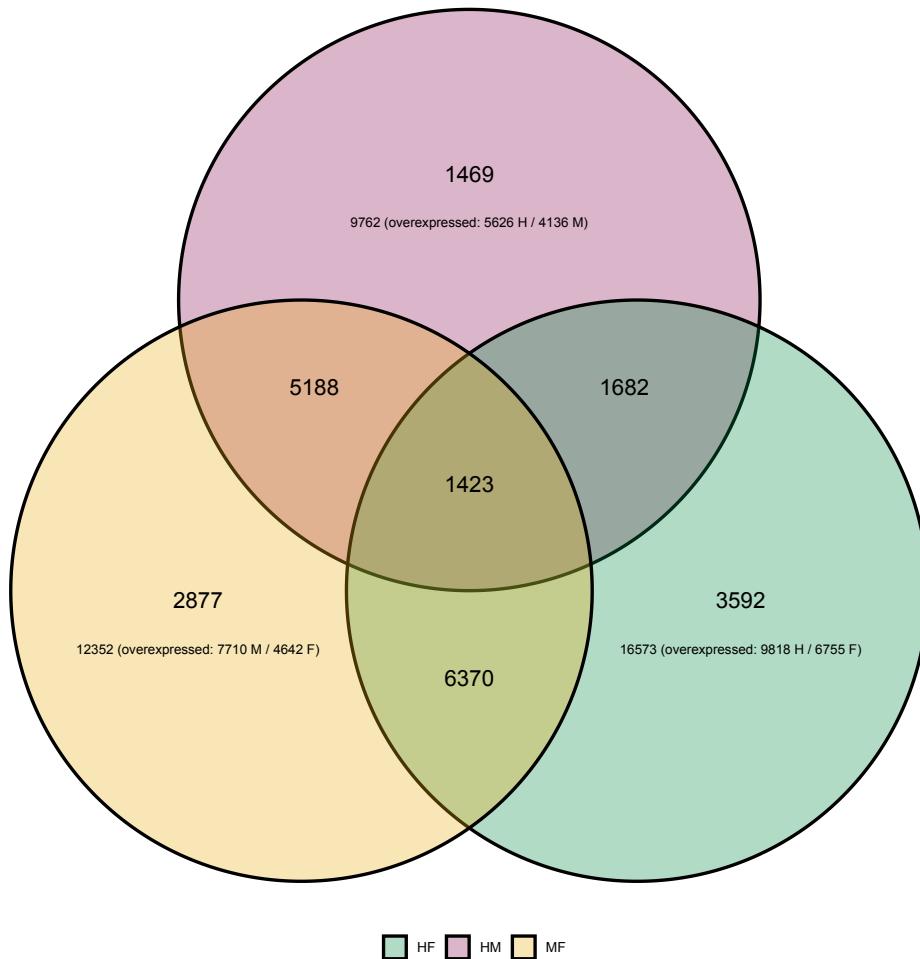


Figure 2.1. Differential expression analysis of *Utterbackia* spp unigenes. Venn diagram shows differentially expressed genes (DEGs) between sexes: hermaphrodite-female (HF in green), hermaphrodite-male (HM in pink) and male-female (MF in orange) comparisons. The figure shows the number of DEGs unique for one comparison (e.g. 2 877 for MF), the number of DEGs shared between two comparisons (e.g. 5 188 between MF and HM comparisons) and the number of DEGs shared by all comparisons (1 423).

The Principal Component Analysis (PCA) performed on regularized log transform (rlog) values of the most expressed genes confirmed that individuals clustered according to their sex (Figure 2.2), reinforcing the reliability of our results and suggesting low biological variability among samples of the same sex.

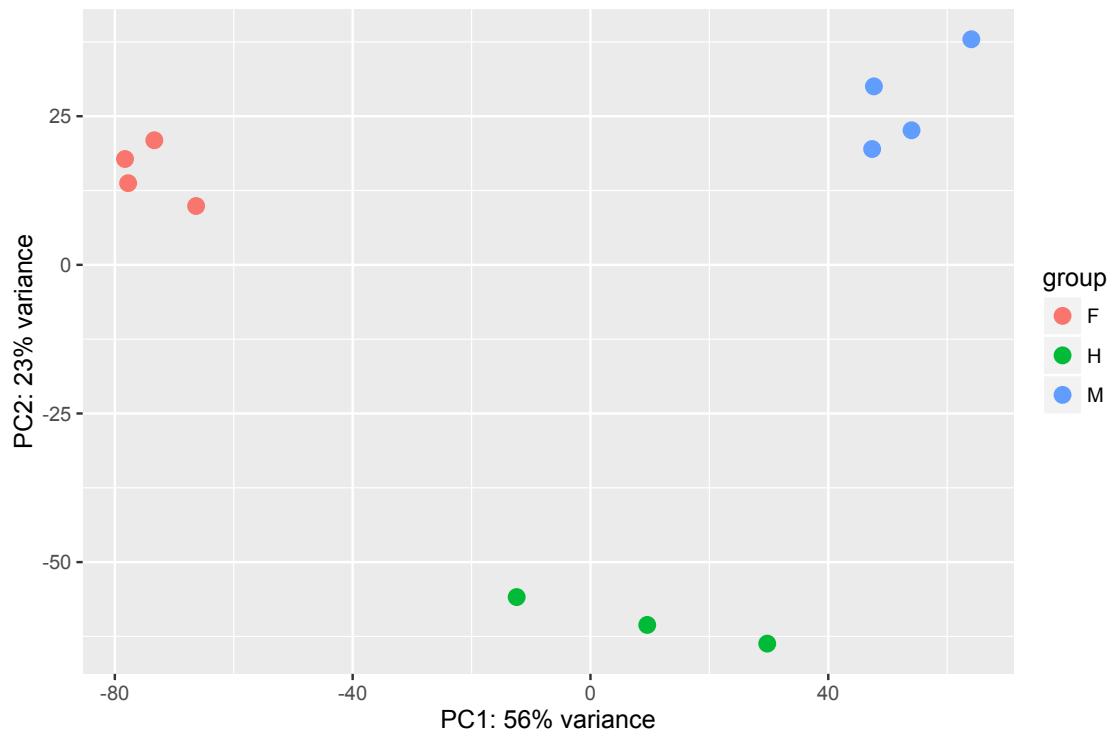


Figure 2.2. Principal Component Analysis (PCA) performed on regularized log transform (rlog) values of the first thousand most expressed genes per pairwise comparison between the three sexes (males, females and hermaphrodites). The four males are shown in blue, the four females in red and the three hermaphrodites in green.

To unravel potential candidate genes involved in sex determination and/or DUI in *Utterbackia* spp., we first focused on sex-specific genes (i.e. those expressed only in one sex and not in the two others). Table 2.2 shows sex-specific genes based on FPKM values. Our analyses revealed 567 genes specific to hermaphrodites, only eleven of which were annotated. DUI individuals showed fewer sex-specific genes with 59 for males (7 annotated) and 100 for females (10 annotated) (Table 2.2). The detection of the *h-orf* in hermaphrodite gonads, a mitochondrial gene specific to hermaphroditic freshwater mussel species (Breton et al., 2011), supports the

hypothesis that this gene could have an essential role in the development of hermaphrodites in freshwater mussels (Breton et al., 2011; Mitchell et al., 2016). Similarly, the detection of genes involved in ubiquitination in hermaphrodite gonads reinforces the hypothesis that this process could be involved in mitochondrial inheritance and sex determination in bivalves (Ghiselli et al., 2012; Capt et al., 2018; Punzi et al., 2018). However, the few annotated genes in Table 2.2 (i.e. only 4% of the sex-specific genes were annotated and many of these were uncharacterized proteins) clearly point to the need for further studies of these sex-specific factors to elucidate the molecular mechanisms of sex determination and mitochondrial inheritance in freshwater mussels.

As a second step to unravel potential candidate genes involved in sex determination and/or DUI in *Utterbackia* spp, we searched for the 20 most up-regulated differentially expressed genes, based on their log₂ Fold-Change (log₂FC) values, in each pairwise-comparison; Male-Female (MF), Hermaphrodite-Male (HM) and Hermaphrodite-Female (HF) (Supplementary Table S2.1). For the MF comparison, our results were highly similar to those obtained in our previous study comparing male and female *U. peninsularis* gonadal transcriptomes (Capt et al., 2018), with up-regulated genes in males including factors involved in spermatogenesis, sperm motility and ubiquitination processes, and up-regulated genes in females including C-type lectin domains, which are known to be involved in gamete recognition and polyspermy blocking during fertilization. One particular new female-biased gene found in the present study was *Temptin* (c175440_g1_i1; up-regulated mostly in females, and also to a lesser extent in hermaphrodites compared to males [MF; log₂FC = -14.7, HM; log₂FC = +11.3, HF; log₂FC = -3.4]). Temptin is a water-borne protein pheromone that provokes stimulation of mating behavior in some mollusk species (e.g. the sea slug *Aplysia* sp. [Cummins et al., 2004] and the gastropod *Biomphalaria glabrata* [Pila, Peck, & Hanington, 2017]). Because pheromones are thought to play a critical role in triggering spawning in bivalves (e.g. Rice et al., 2002), it is possible that such a pheromone produced by eggs could stimulate male mussels to spawn and/or attract sperm in marsupia gill chambers where eggs are fertilized in freshwater mussels (Haag, 2012).

Table 2.2. Sex-specific genes in hermaphrodites, males and females.

Unigene	Annotation (sprot)	e-value	Description	FPKM H (Mean)
c404212_g1_i1	F4ZG36_9BIVA	2,00E-60	H-orf protein	8565.0
c180137_g1_i1	PAR14_HUMAN	2,00E-12	Poly [ADP-ribose] polymerase 14	1295.3
c1018548_g1_i1	DCXR_BOVIN	1,00E-44	L-xylulose reductase	503.3
c168047_g2_i2	PAR14_MOUSE	4,00E-15	Poly [ADP-ribose] polymerase 14	129.0
c213214_g2_i1	POL2_DROME	1,00E-11	Retrovirus-related Pol polyprotein from transposon 297	57.7
c304119_g1_i1	UBIQ_WHEAT	1,00E-08	Ubiquitin	43.7
c703049_g1_i1	BIRC1_HUMAN	9,00E-10	Baculoviral IAP repeat-containing protein 1	39.7
c231908_g1_i1	RN213_MOUSE	4,00E-09	E3 ubiquitin-protein ligase RNF213	25.0
c226075_g1_i1	PIF3_TRYB2	3,00E-07	ATP-dependent DNA helicase PIF3	18.0
c763331_g1_i1	NTF2_ARATH	5,00E-40	Nuclear transport factor 2	16.0
c131247_g1_i1	PLXA3_MOUSE	2,00E-07	Plexin-A3	11.0
Unigene	Annotation (sprot)	e-value	Description	FPKM M (Mean)
c168636_g1_i1	W4Y5B2_STRPU	1,00E-19	Uncharacterized protein	111.8
c230214_g5_i1	G3TY18_LOXAF	1,00E-52	Uncharacterized protein	57.8
c148972_g1_i1	D7GXV5_TRICA	3,00E-06	Putative uncharacterized protein	55.0
c159235_g1_i1	V3ZGW1_LOTGI	4,00E-07	Uncharacterized protein	19.5
c148972_g2_i1	W4Z2T5_STRPU	3,00E-10	Uncharacterized protein	18.3
c150368_g1_i2	K1PNI5_CRAGI	1,00E-06	Putative ankyrin repeat protein L93	24.3
c219930_g7_i1	W4YBE1_STRPU	7,00E-10	Uncharacterized protein	60.5
Unigene	Annotation (sprot)	e-value	Description	FPKM F (Mean)
c209524_g1_i6	ATS12_MOUSE	2,00E-11	A disintegrin and metalloproteinase with thrombospondin motifs 12	120.0
c227946_g1_i1	KY_MOUSE	2,00E-28	Kyphoscoliosis peptidase	92.75
c179701_g1_i2	UPSA_UNIPI	2,00E-73	Upsalin	90.25
c178426_g1_i2	PLC_HALLA	7,00E-12	Perlucin	85.5
c136248_g1_i1	ZNF80_PANTR	2,00E-19	Zinc finger protein 80	72.75
c155883_g1_i1	ISL2_CHICK	3,00E-06	Insulin gene enhancer protein ISL-2	52.75
c137229_g1_i1	A4_RAT	8,00E-10	Amyloid beta A4 protein	40.75
c222051_g4_i1	PCF11_HUMAN	3,00E-17	Pre-mRNA cleavage complex 2 protein Pcf11	30.5
c119723_g1_i2	SNAII_XENLA	2,00E-25	Protein snail homolog Sna	24.75
c398528_g1_i1	LHX6_MOUSE	2,00E-11	LIM/homeobox protein Lhx6	11.0

Several top DEGs identified in the MF comparison were also present in hermaphrodites (i.e. many up-regulated genes in the hermaphrodites compared to females were “male-type” genes such as testis-specific kinases, whereas many up-regulated genes in the hermaphrodites compared

to males were “female-type” genes such as C-type lectin domains and *Temptin*, Supplementary Table S1). These results suggest similar factors involved in gonadal differentiation between the gonochoric *U. peninsularis* and hermaphrodite *U. imbecillis*.

To decipher potential candidate genes involved in mitochondrial inheritance in *Utterbackia* spp, we clustered all the genes found up-regulated in hermaphrodites compared to males and females grouped together (i.e. SMI vs. DUI). Only annotated genes were kept, and if a similar annotation was also found in male- or female-specific DEGs, it was discarded from the list. This step allowed us to focus on genes with a unique role in hermaphrodites, thus potentially revealing genes specific to SMI of mitochondria (Supplementary Table S2.2). We found 167 genes specifically over-expressed in hermaphrodite gonads. Of these, one unigene was involved in autophagy, seven in nuclease activity and six were involved in polymerase activity, representing three gene ontology categories previously described as being linked to mitochondrial inheritance strategies (Punzi et al., 2018). A similar pipeline was applied for males to find potential candidates directly linked to DUI and retention of sperm mitochondria in male gonads (Supplementary Table S2.2). We found 153 genes specifically over-expressed in males. One of them has a role in methylation events and another in ubiquitination activity. All of these results are discussed in the section “candidate genes involved in mitochondrial inheritance”.

Gene Ontology (GO) and Enrichment Analyses

Figure 2.3 shows the top ten enriched GO (biological processes and molecular function) categories for each MF, HM and HF pairwise comparison. For the complete GOseq results, see supplementary Table S2.1. Overall, GOseq analyses between males and females revealed 110 gene ontology categories significantly enriched out of 621 listed. HF and HM comparisons showed reduced numbers of enriched categories, with 41 (out of 526) and 25 (out of 456), respectively (Table S2.1). The top ten enriched categories were similar for MF and HF comparisons, potentially revealing male gonad specific processes and functions, and included microtubule dynamics (GO:0007017, GO:0007018, GO:0005815, GO:0005874, GO:0044450) and mitotic events (GO:190304, GO:0007067) (Figure 2.3). Microtubule dynamics is important in spermatogenesis and male germ cell development by supervising the division of sperm cells and by modulating sperm shape and

motility, leading to correct male fertility (O'Donnell & O'Bryan, 2014; O'Donnell et al., 2012). The up-regulation of microtubule dynamic process in males and hermaphrodites compared to females is thus not surprising. Similar results were also found in previous studies comparing male and female gonad transcriptomes in gonochoric bivalves (Capt et al., 2018; Ghiselli et al., 2012; Y. Li et al., 2016; Shi et al., 2015; Tong et al., 2015; N. Zhang et al., 2014). In species with DUI, a role of microtubules in sperm mitochondria aggregation in male gonads has also been suggested (Ghiselli, Milani, & Passamonti, 2011; Milani et al., 2016; Obata & Komaru, 2005), but this still remains to be confirmed. In the HM comparison, over-represented GO categories mostly belonged to response to biotic and external stimuli (GO:0009605, GO:0009607, GO:0043207, GO:0051707, GO:0009617) and ribosome cellular component and activity (GO:0005840, GO:0030529, GO:0044391, GO:0022625, GO:0015934, GO:0003735). Up-regulation of these GO categories in hermaphrodites could be related to some epigenetic mechanisms with key roles in the differentiation of gametes in hermaphroditic individuals (e.g. Gomot & Griffond, 1993).

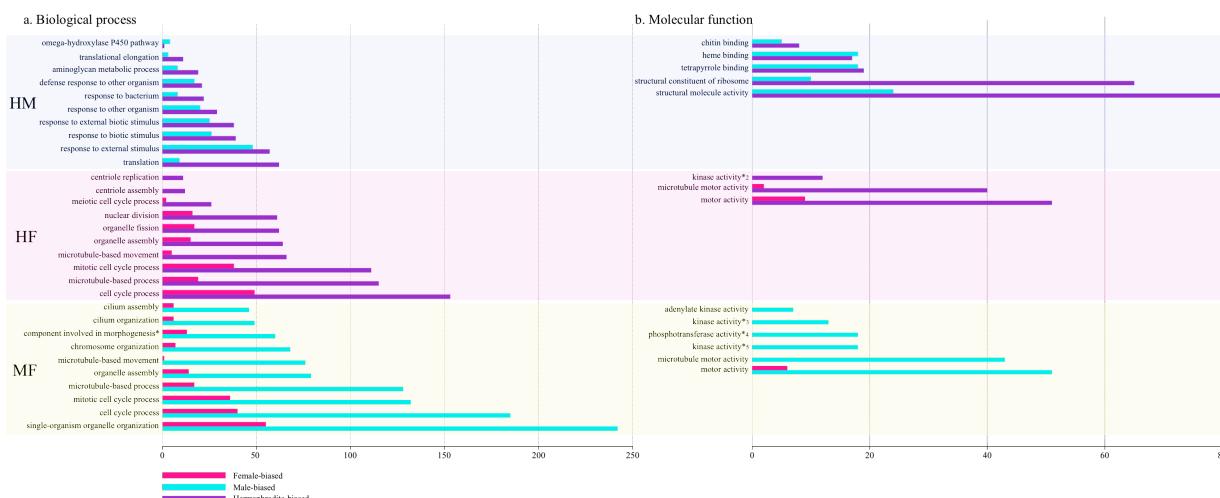


Figure 2.3. Top ten of most significantly enriched Gene Ontology (GO) categories in each comparison; hermaphrodite-male (HM), hermaphrodite-female (HF), and male-female (MF), involved in both biological processes (a) and molecular functions (b). Differentially expressed genes (DEGs) numbers involved in each GO categories are shown by x-axis; female-biased genes are illustrated in pink, male-biased in blue and hermaphrodite-biased in purple. Complete GO terms are written here, *; cellular component assembly involved in morphogenesis, *²; nucleobase-containing compound kinase activity, *³; nucleoside diphosphate kinase activity, *⁴; phosphotransferase activity, phosphate group as acceptor, *⁵; nucleobase-containing compound kinase activity.

Candidate Genes Involved in Sex Determination

Our approach to identify candidate genes involved in sex determination consisted of searches in the GO categories, terms that contained ‘reproduction’ and ‘sex’. In addition, candidate genes previously identified as sex-determining genes in the literature in bivalves and in model organisms *C. elegans*, *Drosophila melanogaster* and *Mus musculus* were searched in differentially expressed genes (Table 2.3; see also Capt et al., 2018). Of the 26 key sex-determining pathway genes examined, putative homologs were found for 15 genes or gene families in freshwater mussel species (Table 2.3). Based on log2FC values, three main factors were up-regulated in males and hermaphrodites compared to females, i.e. *Dsx/Dmrt1/mab3*, *Sry* and *Fem1c*, all of which are all known to play an important role in male sex determination (Kopp, 2012; Wallis et al., 2008). In contrast, *Foxl2*, which is a well-known gene implied in oogenesis differentiation in a large variety of species (Uhlenhaut & Treier, 2006), was up-regulated in both females and hermaphrodites when compared to males. Overall, our results are quite similar to what has been observed in previous transcriptomic studies of freshwater mussels or other bivalve species (Capt et al., 2018). One novelty in the present study is the discovery of a highly up-regulated gene in females and hermaphrodites compared to males; *Zglp1*, which codes for a GATA-4 protein and has a role in early-gonadal differentiation in mammals (Viger et al., 1998).

Table 2.3. Candidate genes involved in sex determination pathway in freshwater mussels and other bivalves.

Gene symbol	Name	Gonochoric FWM species	Hermaphroditic FWM species	Other gonochoric bivalve species
<i>Amh, Amrh2 (Mmu)</i>	Anti-Mullerian hormone and receptor type 2			
<i>Bar-1 (Cel), Arm (Dme), Ctnnb1 (Mmu)</i>	Beta catenin			
<i>Cbx2 (Mmu)</i>	Chromobox 2			
<i>Doa (Dme)</i>	Darkener of apricot			
<i>Dsx (Dme), Dmrt1 (Mmu), Mab-3 (Cel)</i>	Doublesex- and mab-3-related transcription factor 1			
<i>Fog-1,-2,-3 (Cel, Dme, Mmu)</i>	Feminization of germline			
<i>Fem-1 (Cel, Dme), Fem1b (Mmu)</i>	Feminization of XX and XO animal			
<i>FoxL2 (Mmu)</i>	Forkhead box L2			
<i>Fru (Dme)</i>	Fruitless			
<i>Zglp1 (Mmu)</i>	GATA-type zinc finger protein 1			
<i>Her-1 (Cel), Her (Dme)</i>	Hermaphrodization of XO animals			
<i>Nr0b1 (Mmu)</i>	Nuclear receptor subfamily 0, group B, member 1			
<i>Rspo1 (Mmu)</i>	R-spondin 1			
<i>Rnt-1 (Cel), Run (Dme)</i>	Segmentation protein Runt			
<i>Sdc (Cel, Dme, Mmu)</i>	Sex determination and Dosage Compensation defect			
<i>Sxl, fl(2)d (Dme)</i>	Sex lethal, female-lethal(2)d			
<i>Sry/Sox30/SoxH (Mmu)</i>	Sex-determining region of chromosome Y, SRY (sex determining region Y)-box 30			
<i>Sis-a (Dme)</i>	Sisterless			
<i>Sf1 (Mmu)</i>	Splicing factor 1			
<i>Sox9 (Mmu), Sox100B (Dme)</i>	SRY (sex determining region Y)-box 9, SRY-box 100B			
<i>Gata4 (Mmu)</i>	Transcription factor GATA-4			
<i>Tra-1 -2 -3 (Cel), Tra2 (Dme)</i>	Transformer			
<i>Wt1 (Mmu)</i>	Wilms tumor 1 homolog			
<i>Wnt4 (Dme, Mmu)</i>	Wingless-type MMTV integration site family, member 4			
<i>Xol-1 (Cel)</i>	Xol-1 XO Lethal			

NOTE – Names in parentheses indicate the species in which each gene has been identified (Cel, *C. elegans*; Dme, *Drosophila melanogaster*; Mmu, *Mus musculus*) (the list was first provided by Zhang et al., 2014); Gonochoric FWM (freshwater mussel) species include *Venustaconcha ellipsiformis* and *Utterbackia peninsularis* (Capt et al., 2018; present study) and *Hyriopsis schlegelii* (Shi et al., 2015); Hermaphroditic species include *Utterbackia peninsularis* (present study); Other gonochoric bivalve species include the oysters *Crassostrea gigas* and *C. honkongensis*, the scallop *Patinopecten yessoensis* and the marine clams *Ruditapes philippinarum* and *R. decussatus* (see Capt et al., 2018); pink box indicates female-specific or female-biased expression; pale pink box indicates hermaphrodite-biased expression compare to males; blue box indicates male-specific or male-biased expression; pale blue box indicates hermaphrodite-biased expression compare to female; grey box indicates presence in the transcriptome but not differentially expressed.

Candidate Genes Involved in Mitochondrial Inheritance

To identify candidate genes involved in mitochondrial inheritance, we used the approach proposed by Punzi et al. (2018), i.e. to compare our SMI (*U. imbecillis*) vs. our DUI (*U. peninsularis*) species, while considering presence and transcription patterns of proteins belonging to pathways known to be involved in mitochondrial inheritance in animal species. We thus analyzed transcripts belonging to the GO categories “nucleases/polymerases”, “autophagy and mitophagy”, and “ubiquitination pathway” and compared our results with those obtained for the marine clam SMI species *Ruditapes decussatus* and the DUI species *R. philippinarum* (Ghiselli et al., 2012; Milani et al., 2013b; Punzi et al., 2018). As mentioned above, we also searched the hermaphrodite-specific and male-specific DEGs for genes related to these categories. Furthermore, because autophagy and mitophagy pathways rely on evolutionarily conserved core components, we used the compiled lists of orthologs related to these processes generated previously (Lee & Han, 2017; Punzi et al., 2018) and used successfully by Punzi et al. (2018) for the bivalve species *Ruditapes philippinarum* and *R. decussatus*, to search for their presence in *Utterbackia* spp based on our annotations. Finally, we discuss transcripts belonging to “methylation GO categories” because our previous study identified this process as potentially important in influencing mtDNA inheritance and/or regulating sex determination and differentiation in bivalves (Capt et al., 2018).

Ubiquitination-related factors: The proposed mode of action for ubiquitination in SMI is to tag paternal mitochondria for degradation during spermatogenesis and/or after fertilization (Punzi et al., 2018; Sato & Sato, 2017), and a modification of this mechanism has been proposed to be responsible for DUI of mitochondria in bivalves (Ghiselli et al., 2012; Milani et al., 2013b). For example, three main male-biased ubiquitination-related factors (*birc* - baculoviral IAP repeat-containing 4 [inhibitor of apoptosis], *psa* - proteasome subunit alpha 6, and *anUBL* - AN1 zinc finger ubiquitin-like domain) have previously been hypothesized to be responsible for the maintenance/degradation of spermatozoon mitochondria during embryo development of the DUI species *R. philippinarum* (Ghiselli et al., 2012; Milani et al., 2013b). These three factors were also reported as male-biased in the freshwater mussel *Utterbackia peninsularis* (Capt et al., 2018). In the present study, a transcriptional bias (up-regulation) for ubiquitination-related genes is thus expected in males and hermaphrodites compared to females. Our Goseq results revealed seven

ubiquitination-related categories enriched (but not significantly) in DUI males (Table S2.3). The only DEG corresponding to our expectation was annotated as protein MAD2A and is involved in mitotic events (Tables S2.2 and S2.3). In our complete list of DEGs, several baculoviral IAP repeat-containing protein-coding genes were found, including male-biases *birc* genes such as *birc-5*, which was up-regulated in males but also in hermaphrodites when compared to females (Table 2.4). *Anubl1* was also found up-regulated in males and hermaphrodites compared to females (Table 2.4). According to previous studies (Capt et al., 2018; Milani et al., 2013b), *birc* could be an additional mitochondrial tag in DUI species that would protect the M midpiece mitochondria from degradation and *anubl1* could be a factor tagging sperm mitochondria that differentiate them from egg mitochondria in DUI embryos. Our results indeed suggest that these genes are mainly male-related but their link to paternal mitochondrial inheritance remains to be clearly demonstrated since they were also found in hermaphrodite and female individuals, both of which lack the M mtDNA. One other ubiquitination-related factor up-regulated in males and hermaphrodites compared to females was the F-box only protein 39 (Table 2.4). F-box proteins are important components of the protein degradation pathway involving ubiquitination (Ho, Tsai, & Chien, 2006). Hermaphrodite-specific ubiquitination-related genes were also found (Table S2.2), reinforcing the hypothesis that this process might be involved in mitochondrial inheritance in freshwater mussels. Finally, nine unigenes were annotated with *psa*, but only one of them was overexpressed in males and hermaphrodites compared to females (Table 2.4), supporting the hypothesis that subunits alpha of the proteasome may have a role in mitochondrial inheritance and/or male sex determination or differentiation in bivalve species (Milani et al., 2013b).

Table 2.4. Putative candidate genes involved in mitochondrial inheritance in freshwater mussels.

Contig ID	Gene name	HM	HF	MF	GO category
c211580_g1_i1	Birc5	-	(+1.7)	(+2.0)	Ubiquitination
c223534_g1_i1	Anubl1	-	(+5.0)	(+6.2)	Ubiquitination
c211818_g1_i1	FBX039	-	(+10.2)	(+11.0)	Ubiquitination
c224112_g6_i2	Psa	-	(+1.5)	(+2.0)	Ubiquitination
c154451_g1	Nuclease EXOG, mitochondrial	(+5.8)	(+7.6)	-	Nuclease, polymerase
c226166_g1_i2	TANK-binding kinase 1	(+1.8)	(+1.4)	-	Autophagy, mitophagy
c235196_g1_i1	Ras-like GTP-binding protein YPT1	(+7.6)	(+4.4)	-	Autophagy, mitophagy
c231455_g1_i6	DNMT1	-	(+2.9)	(+3.8)	Methylation

NOTE –The annotation was provided by the Uniprot database. HM, HF and MF respectively indicate hermaphrodites vs. males, hermaphrodites vs. females and males vs. females comparisons, log2 fold change values (obtained from

edgeR software) are in parentheses, a positive value means that the expression is biased towards the first sex in the comparison, a negative value indicates that the expression is biased towards the second sex in the comparison.

Nucleases and polymerases: The proposed mode of action of nucleases and polymerases is to degrade mtDNA during spermatogenesis to prevent transmission of paternal mtDNA (Punzi et al., 2018; Sato & Sato, 2017). For example, endonuclease G (*endoG* gene) has been shown to eliminate paternal mitochondrial nucleoids during spermatozoid production in *D. melanogaster*, thus resulting in SMI of mitochondria (Chan & Schon, 2012; Deluca & O'Farrell, 2011). *Tamas*, a gene that codes for one subunit of the mtDNA polymerase (Pol γ - α), has also been identified as a key factor in this process (Yu et al., 2017). Here, a transcriptional bias for such genes (an up-regulation) is expected in SMI hermaphrodites compared to DUI males and females. Our GOseq results and hermaphrodite-specific DEGs revealed 13 genes that corresponded to this situation (supplementary Table S2.3), including ten genes coding for Gag or Pol proteins. These genes represent signatures of retrovirus and transposon activity, suggesting a key role for these proteins in hermaphrodite gonads. One promising gene overexpressed in hermaphrodites compared to DUI males and females was *exoG* (Table 2.4), which is a paralog of *endoG* described in higher eukaryotes (Cymerman et al., 2008). It must be noted that *endoG* was also found but was up-regulated in males and females compared to hermaphrodites (data not shown). We propose that *exoG* could represent a potential candidate involved in SMI of mitochondria in hermaphroditic freshwater mussels. Like in *Drosophila*, paternal mtDNA could be mostly degraded by a nuclease-mediated mechanism during the process of spermatogenesis, and paternal mitochondrial structure could be degraded by ubiquitination-related mechanisms after fertilization (Sato & Sato, 2017).

Autophagy and mitophagy: The proposed mode of action of autophagy and mitophagy is to degrade mitochondria during spermatogenesis and/or after fertilization to prevent transmission of paternal mtDNA (Punzi et al., 2018; Sato & Sato, 2017). If these processes are involved in mitochondria degradation during spermatogenesis to maintain SMI, it is expected that they will be up-regulated in SMI hermaphrodites compared to DUI males and females. In addition with GO terms searches, and because autophagy and mitophagy are well-conserved processes, specific autophagy-orthologs listed in Lee & Han (2017) and Punzi et al. (2018) were also searched in our DEGs. Only two genes involved in mitophagy/autophagy processes were indeed up-regulated in the SMI species (Table 4 and supplementary Table S2.3): (1) TANK-binding kinase 1, a protein involved in mitophagy

regulation in the PINK1-Parkin pathway (Heo et al., 2015), and (2) Ras-like GTP-binding protein YPT1, which is linked to organelle inheritance in yeast (Lewandowska, Macfarlane, & Shaw, 2013). Combined with *exoG*, these proteins could belong to the pre-fertilization mechanisms involved in the strict maternal inheritance of mitochondria in hermaphroditic freshwater mussels.

On the other hand, if autophagy and mitophagy are involved in sperm mitochondria degradation after fertilization, it is expected that they will be over-expressed in DUI females and SMI hermaphrodites compared to DUI males. None of our data fulfilled these criteria, a result similar to what was found in *Ruditapes* spp. (Punzi et al., 2018).

Methylation: As previously suggested, methylation could play an important role in mitochondrial inheritance and/or sex determination in animal species, including freshwater mussels (Capt et al., 2018; Liu et al., 2015.). In *Chlamydomonas reinhardtii*, there are two mating types (mt+ and mt-), and uniparental inheritance of chloroplast DNA (cpDNA) is achieved by active elimination of the mt- cpDNA, with a possible role of DNA methylation in guaranteeing strict uniparental inheritance by enhancing the replication of the mt+ cpDNA (while unmethylated cpDNAs is being degraded) (Nishimura, 2010). If methylation is involved in DUI of mitochondria in bivalve species, with methylated paternal mtDNA being protected from degradation, we could expect an up-regulation of genes related to this process in DUI males compared to SMI hermaphrodites and DUI females. Our GOseq results and male-specific DEGs revealed two annotated genes fulfilling these criteria, i.e. two methyltransferases, *protein arginine N methyltransferase 1/6* that methylates proteins and *putative methyltransferase C9orf114 homolog isoform XI* that methylates RNA (supplementary Table S2.3). Interestingly, the methyltransferase DNMT1, for which a mitochondrial isoform has been described (Maresca et al., 2015), was found to be overexpressed in males and hermaphrodites compared to females. If DNMT1 is differentially expressed in males and hermaphrodites' male-parts gonads, it is not related with the assumption that paternal methylated mtDNA would be protected from nucleases, except if another mechanism in SMI hermaphrodites can overcome this event, for example through the endonuclease activity of *exoG*. Further studies will be necessary to unravel this possibility.

Conclusion

In this study, we reported the transcriptome of male, female, and hermaphrodite gonads of two freshwater mussel species, i.e. the gonochoric, DUI species *Utterbackia peninsularis* and the hermaphrodite, SMI species *U. imbecillis*. The results were used to generate a “hybrid transcriptome”, which allowed analyzing differential gene expression among the three sexes. Our results were also compared to published transcriptomes of bivalve species showing different mitochondrial transmission modes (DUI and SMI). This approach permitted us to highlight sex-specific and sex-biased expression patterns, as well as expression patterns associated to SMI or DUI within bivalves having alternative reproductive and mitochondrial transmission modes. Our results revealed similar combinations of genes with conserved sex-linked roles in gonochoric and simultaneous hermaphrodite freshwater mussel species and suggested that steps of the sex-determination pathway are conserved in distantly-related bivalve species. Our results also revealed new candidate genes showing potential roles for SMI or DUI of mitochondria in bivalves, including nucleases and factors involved in autophagy/mitophagy and methylation. Further investigation of these genes will help to unravel the mechanisms underlying mitochondrial inheritance and sex determination in bivalves.

Author Contributions

CC, SR, and SB designed experiments. NJ conducted sampling. CC performed laboratory experiments. CC and SR analyzed data. CC, SR, NJ, DTS and SB wrote the manuscript. SB conceived and managed the project. All authors read and approved the final version of the manuscript.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We would like to thank Stefano Bettinazzi for his very helpful support and advice and Chase Smith for comments on a previous version of this manuscript. We also thank two anonymous reviewers for suggestions. Financial support for this study has been provided by Natural Sciences and Engineering Research Council Discovery Grant awarded to S Breton [grant number RGPIN/435656-2013]. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S.

Supplementary Material

Supplementary tables

Table S2.1. Most up-regulated genes based on their logFold-Change (logFC) values, in each pairwise-comparison: Male-Female (MF), Hermaphrodite-Female (HF) and Hermaphrodite-Male (HM).

En fichier complémentaire, nommé « Tables Suppl Chapitre 2 », accessible au premier feuillet.

Table S2.2. Hermaphrodite- and male-specific DEGs.

En fichier complémentaire, nommé « Tables Suppl Chapitre 2 », accessible aux deuxième (hermaphrodite) et troisième (mâle) feuillets.

Table S2.3. Male-specific and hermaphrodite-specific DEGs and enriched GO categories related to "ubiquitination pathway", "nucleases/polymerases", "autophagy and mitophagy", and "methylation".

En fichier complémentaire, nommé « Tables Suppl Chapitre 2 », accessible au quatrième feuillet.

Supplementary figures

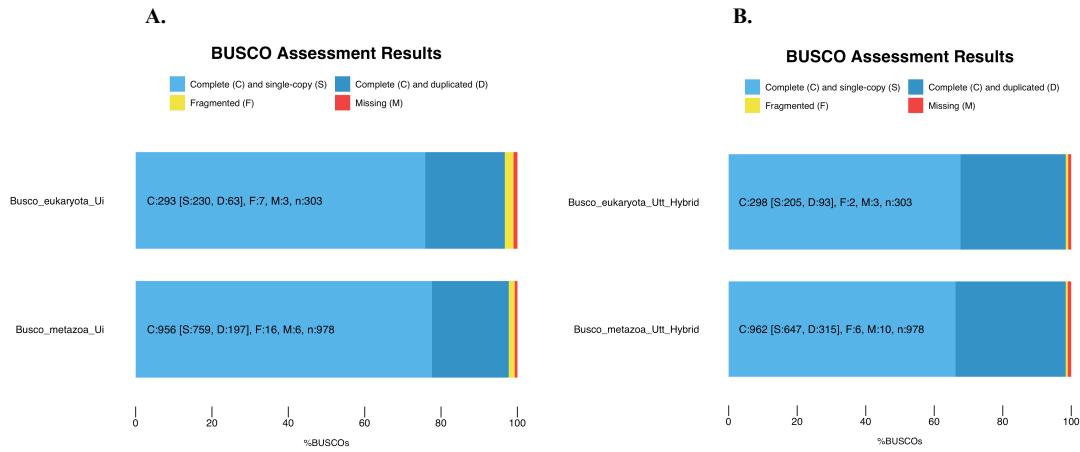


Figure S2.1. Quality evaluation of unigenes from *U. imbecillis* and the hybrid transcriptome based on two lineage-specific datasets Benchmarking Universal Single-Copy Orthologs (BUSCO). A. Quality assessment of *U. imbecillis* transcriptome based on Eukaryota and Metazoa datasets, respectively. B. Quality assessment of hybrid transcriptome based on Eukaryota and Metazoa datasets, respectively.

Chapitre 3.

L'analyse mitogénomique comparative de deux bivalves vénéridés utilisant la transmission mitochondriale doublement uniparentale révèle une insertion record dans le gène cox2 !

Comparative mitogenomics reveals unorthodox features in two venerid bivalves with doubly uniparental inheritance of mitochondria

This submission is intended as a Research Article

Charlotte Capt^{1,a}, Karim Bouvet¹, Davide Guerra¹, Brent M. Robicheau², Donald T. Stewart³, Eric Pante^{4,*} Sophie Breton^{1,*}

¹ Department of Biological Sciences, Université de Montréal, Montréal, QC, Canada

² Department of Biology, Dalhousie University, Halifax, NS, Canada

³ Department of Biology, Acadia University, Wolfville, NS B4P 2R6 Canada

⁴ Littoral, Environnement et Sociétés (LIENSs), UMR 7266 CNRS–La Rochelle Université, 2 rue Olympe de Gouges, 17000 La Rochelle, France

Accepted in *Scientific Reports*

Author Contributions

CC, EP and SB conceived and designed the experiments. CC analyzed data and drafted the manuscript. KB, DG and BMR participated in the experimental work. All authors contributed feedback to the writing process and approved the manuscript.

Abstract

In the animal kingdom, strictly maternal inheritance of mitochondria is typically the rule, but one exception (termed doubly uniparental inheritance or DUI) has been reported in several orders of bivalve molluscs and is marked by the maintenance and transmission of two separate sex-linked mitochondrial genomes. Associated with DUI is a frequent modification of the *cox2* gene within said genomes, as well as additional sex-specific mitochondrial genes that are not associated with oxidative phosphorylation. With the exception of freshwater mussels (for 3 families of the order Unionida), these two DUI-associated features have only been shown in few species [within Mytilidae (order Mytilida) and within Veneridae (order Venerida)]. This is because of the relatively few complete male- and female-transmitted mitochondrial genomes that are published for these orders. In this study, we present the complete sex-specific mitochondrial genomes of two recently-discovered DUI species in two additional families of the order Venerida, *Scrobicularia plana* (Semelidae) and *Limecola balthica* (Tellinidae). Our data show that these species display the largest differences in genome size between the sex-specific mitotypes within a single species with DUI (>10kb), as well as the highest sex-specific mtDNA divergences (sometimes reaching >50%). The large differences in genome size are due in part because of an important in-frame insertion (>3.5kb long) in the male *cox2* gene of both species. The *S. plana* *cox2* gene is the largest reported so far in the Kingdom Animalia. Our data also indicate that the newly sequenced mitogenomes may be carrying sex-specific genes, and support previous data indicating that general mitochondrial features are shared among DUI species.

Introduction

Animal mitochondrial DNA (mtDNA) is typically depicted as a strictly maternally inherited (SMI) circular DNA molecule that is relatively small (~16kb) and genomically streamlined with almost invariant gene content (13 protein-coding genes and 24 structural RNAs) (Boore, 1999; Gissi et al. 2008). However, important deviations do occur in the mtDNAs of bivalve molluscs, which not only display dramatic variation in size (<14.7kb to >67 kb; Park & Ahn, 2015; McCartney et al. 2019) and gene arrangement (e.g., Plazzi et al., 2016), but also the presence of additional protein-coding genes not associated with oxidative phosphorylation (Breton et al., 2009; Breton et al., 2011; Milani et al., 2013; Breton et al., 2014). An even more extreme departure from the norm in bivalve mitochondrial genomes is their mode of doubly uniparental inheritance (DUI) — both egg and sperm mitochondria are transmitted from generation to generation in several bivalve species, but only male offspring retain paternally-transmitted mitochondria (with male or M mtDNA) in their gametes (see Breton et al., 2007; Passamonti & Ghiselli, 2009; Zouros, 2013 for reviews). Adult females of DUI-exhibiting species usually possess only the female-transmitted mtDNA (F mtDNA) in their soma and gametes whereas males possess F mtDNA in their soma and M mtDNA in their gametes (Breton et al., 2007; Passamonti & Ghiselli, 2009; Zouros, 2013; Breton et al., 2017). The DNA divergence between F and M mtDNAs usually vary from about 8% to 40% depending on the species (Doucet-Beaupré et al., 2010; Zouros, 2013). Genetic analyses suggested that both F and M mtDNAs in DUI bivalves evolve at a faster rate than typical metazoan mtDNA, and that M mtDNA evolves faster than F mtDNA (Hoeh et al., 1996; Stewart et al., 1996; Passamonti & Scali, 2001). One factor explaining this observation may be that the M genome is subject to weaker selective pressures than the F genome due to an unequal “division of labor” in the DUI system (Stewart et al., 1996). Typical animal mtDNA functions in gonads and somatic tissues of both sexes whereas under DUI, F mtDNA functions in female gonads and the soma of both sexes, while M mtDNAs functions primarily in spermatozoa of male gonads and only partially in the male soma (Stewart et al., 1996; Dalziel & Stewart, 2002; Breton et al., 2017).

In addition to a different mode of mitochondrial transmission and evolution rate of mtDNA, two other remarkable differences have been reported between the F and M mtDNAs in DUI bivalves. First, the COX2 protein encoded by M mtDNA (*Mcox2* gene) is longer than the

FCOX2 protein, which is about the same size as other animal COX2 proteins, although this pattern is not shared by all DUI species for which sex-specific mtDNAs have been completely sequenced (reviewed in Bettinazzi et al., 2016). For example, male freshwater mussels (order Unionida) have an approximately 550 bp 3'-coding extension to the *cox2* gene (*Mcox2e*), that is absent from other animal mtDNAs (Curole & Kocher, 2002; Chapman et al., 2008; Guerra et al., 2017). This extension proved to be translated and localized in both inner and outer mitochondrial membranes (Chakrabarti et al., 2006; 2007), leading to the hypothesis that it could act as a mitochondrial tag implicated in paternal mitochondria survival in male embryos (Chapman et al., 2008). Such a 3'-coding extension of the *Mcox2* gene has also been found in the mytilid mussel *Musculista senhousia*, but in a duplicated version of the *cox2* gene (Passamonti et al., 2011). The extension is apparently absent from M mtDNAs of other mytilid DUI species (e.g., *Mytilus* spp.) (Breton et al., 2006; Zouros, 2013). In the family Veneridae, the M mtDNA of *Meretrix lamarckii* presents an insertion of 100 codons within the *cox2* gene (Bettinazzi et al., 2016), whereas a duplicated version of the *cox2* gene, similar to that of *M. senhousia* (i.e., longer at 3'), has been found in *Ruditapes philippinarum*, but located in the F genome (unpublished GenBank annotation mentioned in Passamonti et al. [2011]). As suggested by Bettinazzi et al. (2016), non-canonical features of the *cox2* gene are often coupled with DUI, but it is difficult to propose a general function because each major lineage of bivalves that possesses the DUI system exhibits some novel features. Clearly further analyses involving additional species are required to better understand the relationship between these structural variations in the *cox2* gene and DUI, as well as other general features of DUI-exhibiting mitochondrial genomes.

To date, DUI has been found in over one hundred bivalve species representing four taxonomic orders and twelve bivalve families (reviewed in Gusman et al., 2016). In these species, complete F and M mtDNAs have been sequenced for ~25 DUI spp. of the families Mytilidae (order Mytilida), Veneridae (order Venerida), Unionidae, Margaritiferidae and Hyriidae (order Unionida) (see Breton et al., 2009; Breton et al., 2011; Soroka & Burzyński, 2015; Soroka & Burzyński, 2016; Guerra et al., 2017; Śmiertanka & Burzyński, 2017; Chase et al., 2018; Soroka & Burzyński, 2018; Gomes-dos-Santos et al., 2019; Plazzi & Passamonti, 2019). To our knowledge, all these species share one DUI-specific feature: they contain additional sex-specific mitochondrial genes without recognizable homologies to other known genes (hereafter called mitochondrial ORFans or

mtORFans). These F- and M-specific mtORFans (which have been shown to be expressed) have been respectively called F-*orf* and M-*orf* (Breton et al., 2009; Breton et al., 2011; Milani et al., 2013; Milani et al., 2014; Minoiu et al., 2016; Ouimet, 2018). This discovery is particularly interesting because these mtORFans could be responsible for the different mode of transmission of the mtDNAs and/or the functioning of DUI in bivalves. In unionid freshwater mussels, their discovery established a strong link between DUI and sex determination because of a perfect correlation between gonochorism and DUI, and hermaphroditism and SMI with a highly modified F-*orf* gene (so called H-*orf*) being present in the mitochondrial genome of hermaphrodites, and also, in part, because bivalves lack heteromorphic sex chromosomes (Breton et al., 2011; Guerra et al., 2017). Otherwise, their predicted functions support their direct involvement in the DUI mechanism: F-*ORF* proteins are suggested to interact with nucleic acids, adhere to membranes, and have roles in signalling, and M-*ORFs* are suggested to interact with the cytoskeleton and take part in ubiquitination and apoptosis (Milani et al., 2013; Mitchell et al., 2016; Guerra et al., 2017). However, the precise nature of the link between DUI and sex determination and the functions of the F and M mtORFans remains currently unknown. Again, further analyses involving additional species are needed to shed light on this.

Recently, DUI has been discovered in a species of the family Semelidae, *Scrobicularia plana* (Gusman et al., 2016), as well as in a species of the family Tellinidae, *Limecola balthica* (Pante et al., 2017), and except for the complete F mtDNA of *L. balthica* (Saunier et al., 2014), their M mtDNAs (and F mtDNA for *S. plana*) have not previously been sequenced or reported on until now. In this study, we present the complete M and F mitochondrial genomes of these two DUI species. Our main objective was to highlight both unique features and characteristics shared among DUI species of different bivalve families. Obtaining the complete sex-specific mitogenome sequences of the species *S. plana* was particularly interesting to better understand the hypothesized link between DUI and sex determination since an “intersex” condition, i.e., the appearance of oocytes in male gonads following endocrine disruption, has been reported in this species and was associated with down-regulation of male mitochondrial transcripts in males exhibiting intersex compared to “normal” males (see Gusman et al., 2016 for details). Overall, besides indicating that the newly sequenced mitogenomes may be carrying sex-specific genes like in other DUI species, our data reveal that the *cox2* gene in the M mitogenome of *S. plana* is the largest reported so far in

the Kingdom Animalia. It remains to be demonstrated if such unorthodox features play key roles in DUI and sex determination in bivalves.

Materials and methods

Specimen collection and sample preparation

Adult specimens of *Limecola balthica* were collected at low tide on the Aytré sandy mudflat (France; 46.1259N, 1.1284W) in April 2016 and May 2017 (these are periods during which gonads are known to be near sexual maturation (Pante et al., 2017). Individuals were carefully opened and gonad tissue nicked with a micro-scalpel. Gametes were then sampled with a micropipette and a small amount was used to sex the animal with a light microscope; the remainder of the sample was flash-frozen in liquid nitrogen and stored at -80°C. Adult specimens of *Scrobicularia plana* were collected in May 2013 from Concarneau (France; 47.8728° N, 3.9207° W). Individuals were dissected and gonads inspected under the microscope. Female and male mature gonads were conserved in 95% Ethanol and sent to the Université de Montréal.

For *S. plana*, total genomic DNA was extracted separately from male and female gonad tissues with a Qiagen DNeasy Blood & Tissue Kit (QIAGEN Inc., Valencia, CA) using the kit provided animal tissue protocol. The quality and quantity of DNA were assessed by electrophoresis on a 1% agarose gel and also with a BioDrop µLITE spectrophotometer. One female and one male gonad sample with the highest purity and concentration were chosen for sequencing. For *L. balthica*, total genomic DNA from male gonadal tissues (majority of gametic cells, and some somatic carry-over) was extracted using the phenol-chloroform protocol described in Rosel & Block (1996), resuspended in molecular grade water, and quantified using a Nanodrop 2000 and a Qubit fluorometer (dsDNA HS assay kit, Molecular Probes). DNA integrity was also checked by electrophoresis on a 1% agarose gel. One sample characterised by the highest purity and concentration was chosen for sequencing.

Mitochondrial genome sequencing

Total DNA extractions from *S. plana* male and female gonads were respectively used to prepare two DNA libraries that were paired-end sequenced (2×100 bp) using an Illumina HiSeq platform (Génome Québec Innovation Centre, Montréal, QC, Canada). Paired reads were trimmed with Trimmomatic 0.32 (Bolger et al., 2014) and merged with Pear 0.9.6 (J. Zhang et al., 2014) on the Galaxy online platform (Afgan et al., 2018), checking the quality of reads at every step with FastQC 0.11.5 (Andrews, 2015). MITObim 1.8 (Hahn, Bachmann, & Chevreux, 2013) was used to assemble the complete genome sequence, starting from a *S. plana* partial *cox1* sequence (GenBank accession number KX447421 for the male mtDNA and KX447423 for the female mtDNA) as an initial seed.

For *L. balthica*, the library preparation was done following the SQK-LSK108 Nanopore protocol. A DNA shearing step (using Covaris gTubes) was included to linearize circular mtDNA molecules (1min at 11k RPM in a 5415R Eppendorf centrifuge). The resulting library was sequenced on the Minion MIN-101B sequencer using a R9 flowcell. Basecalling was performed using Albacore 2.1.3 (Pomerantz et al., 2018), followed by quality control using Minion_QC 1.0 (Lanfear et al., 2019). Nanopore adapters were removed and chimeric reads were discarded using Porechop 0.2.3 (Wick R. Porechop, available at: <https://github.com/rrwick/Porechop>). Reads with quality scores <9 were removed using Nanofilt 2.0.0 (De Coster et al., 2018). Cleaned reads were assembled using Canu 1.6 (Koren et al., 2017) considering an approximate genome size of 2Gb and disregarding all reads smaller than 500 bp. Resulting unitigs were blasted locally using blastn (Altschul et al., 1997), using male and female *rrnL* sequences (Genbank accession numbers KX831969 and KX831970, respectively; Pante et al., 2017). Based on these results, we identified tig00000009 (length of 26,120 bp, assembly based on 61 nanopore reads) as the male mitogenome. We did not retrieve the female mitochondrial genome as one contig but as two (tig00000888: 16,790 bp, 34 reads; tig00000885: 6,319 bp, 1 read). Basecalling errors were corrected using signal-level data with Nanopolish 0.9.0 (Loman, Quick, & Simpson, 2015). We used these two cleaned contigs as reference genomes to map Illumina 2×150 bp HiSeq sequences (data quality control and pre-processing step as in Viricel et al., 2018) produced from a second individual using Pilon 1.21 (Walker et al., 2014). HiSeq sequencing of cDNA was performed based on mRNAs

purified from sperm and foot tissue (for mapping onto the male and female mitogenomes, respectively) using the Macherey-Nagel Nucleospin RNA kit for NucleoZOL (RNA quality control, cDNA synthesis, library preparation, sequencing, data quality control and pre-processing as in (Vagner et al., 2019). By mapping high-quality Illumina reads on our draft nanopore contigs, we were able to further correct sequencing errors. The two F-type contigs (tig00000885 and tig00000888) were then merged. The quality of the mitogenome assembly was checked at each step of the pipeline (initial Canu assembly, Nanopolish polishing, Pilon polishing) by calculating the percent identity between our F-type unitigs and a published F-type mitogenome from the same sampling locality (KM373201; Saunier et al., 2014) using MUMmer 3.23 (Kurtz et al., 2004).

Mitochondrial genome sequences were deposited in GenBank (accession codes KXXXXXX and KXXXXXY for the F and M mtDNAs of *S. plana*, respectively, and accession codes KXXXXXX and KXXXXXY for the F and M mtDNAs of *L. balthica*, respectively).

Characterization of mtDNAs and sequence analyses

The male and female mitogenome sequences obtained were annotated with MITOS2 (Bernt et al., 2013). Protein-coding genes (PCG) were also predicted using NCBI's Open Reading Frame Finder (ORF Finder), choosing the invertebrate mitochondrial genetic code and alternative start codons. Candidate transfer RNA (tRNA) genes were also identified using tRNAScan-SE (Lowe & Eddy, 1997). The ends of 16S and 12S rRNA genes were assumed to extend to the boundaries of their flanking genes. The complete mitochondrial genomes, with their annotated PCG, tRNAs and rRNAs, were illustrated with GenomeVx online tool (Conant & Wolfe, 2008).

The degree of genetic conservation among the M-versus-F genes in each species was assessed using two approaches: (i) pairwise *p*-distances (or average nucleotide divergence), and (ii) the rates of synonymous and non-synonymous substitutions, i.e. the number of synonymous substitutions per synonymous sites [K_s] and number of non-synonymous substitutions per non-synonymous sites [K_a], which were calculated in MEGA v7 (Kumar, Stecher, & Tamura, 2016). Since there were various indels present between the individual M- and F-type genes within both species, pairwise deletion was used to account for gaps/missing data-points and codon-alignments

were used to facilitate accurate calculations of genetic distances, as well as Ka and Ks values. Amino acid sequences were determined using EMBOSS Transeq and by assuming an invertebrate mitochondrial code (Madeira et al., 2019). Codon-alignments were generated by aligning M and F amino acid sequences for each gene using MUSCLE (Edgar, 2004) in the online portal found at <http://www.ebi.ac.uk/Tools/msa/muscle/> (Madeira et al., 2019). Using these protein alignments as references, nucleotide sequences for each respective gene were then codon-aligned via PAL2NAL v14 (Suyama, Torrents, & Bork, 2006). The exceptionally large indel within the *Mcox2* gene was manually excluded from the gene alignments. MEGA was also used to conduct a Z-test of Selection [via the Nei-Gojobori method] for each gene (Nei & Gojobori, 1986; Tamura et al., 2013). All plots relating to gene conservation were generated in R version 3.4.1 (R Core Team 2017) using the package ggplot2 v3.2.0 (Wickham, 2009).

For comparative purpose, p -distances of individual genes between F and M mtDNAs were also calculated for three species known for having the greatest F versus M DNA divergences in the families Mytilidae, Veneridae and Unionidae: i.e. the mytilid *Modiolus modiolus* (GenBank accession: F, KX821782; M, KX821783; Robicheau et al., 2017), the venerid *Ruditapes philippinarum* (GenBank accession: F, AB065375; M, AB065374; Passamonti & Scali, 2001), and the unionid *Quadrula quadrula* (GenBank accession: F, FJ809750; M, FJ809751; Doucet-Beaupré et al., 2010). These comparisons allowed a better idea of the selection pressures acting on the mitochondrial-encoded genes in these species, since it is already known that some mitochondrial protein-coding genes are more conserved than others in animal species, including bivalves (Bettinazzi et al., 2016; Doucet-Beaupré et al., 2010; Plazzi & Passamonti, 2019; Robicheau et al., 2017). Finally, the presence of supernumerary mitochondrial protein-coding genes (and mtORFans) in unassigned regions was examined using ORF Finder (NCBI). We retained only ORFs with a minimal length of 150bp and localized on the coding strand, as mitochondrial genes are only encoded on one strand in our two species. Transmembrane domains were predicted using the TMHMM server v2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). Alignments between predicted ORFs were done with MUSCLE (Edgar, 2004) and the percentage of identity was visualized with MView (Brown, Leroy, & Sander, 1998), via the EMBL-EBI online web services (Madeira et al., 2019) keeping default parameters.

Results and Discussion

Main genomic features

The complete F mtDNAs of *Scrobicularia plana* and *Limecola balthica* are 16 170 bp and 17 492 bp in length, respectively, whereas the complete M mtDNAs are respectively 26 270 bp and 24 792bp long (Fig. 3.1).

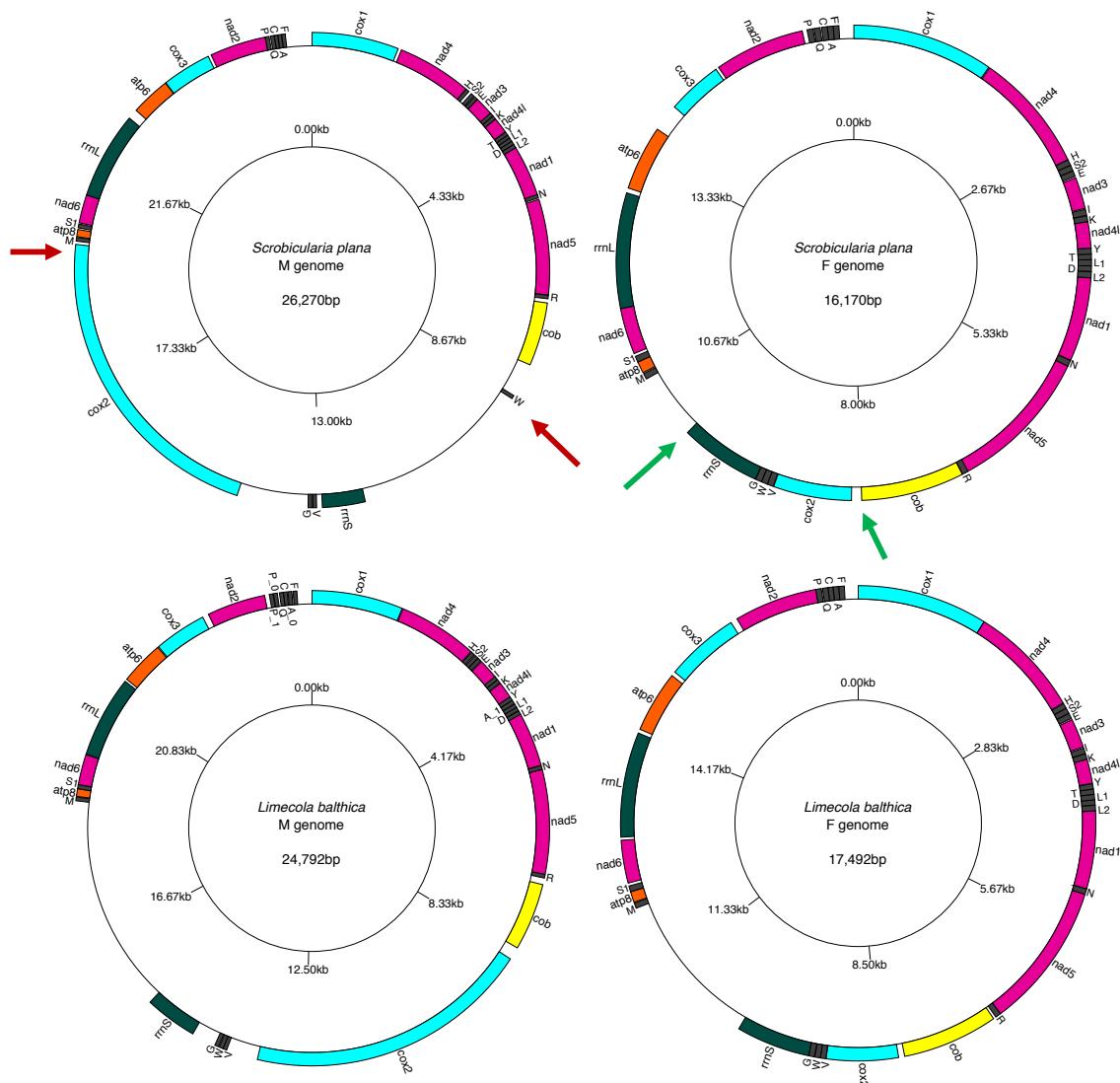


Figure 3.1. Gene maps of M and F genomes of *Scrobicularia plana* and *Limecola balthica*. All genes are encoded on the heavy strand, total genome lengths are reported inside their corresponding genome. Gene colors correspond to functional groups (OXPHOS gene families, tRNAs and rRNAs). Arrows show mitochondrial genes rearrangement between paternal (red arrows; *trnM-trnW*) and maternal (green arrows; *rrnS-cox2*) mitochondrial genomes of *S. plana*.

Table 3.1. Mitochondrial genome sizes in DUI species exhibiting extreme length differences between M and F types.

Species	Mitotype	Genome Size (bp)	Genome size difference (bp)	GenBank Accession Number	Reference
Freshwater mussels (Unionida, Unionidae)					
<i>Anodonta anatina</i>	F	15,637		KF030965	Soroka & Burzynski 2015
<i>Anodonta anatina</i>	M	16,906	1,269	KF030963	Soroka & Burzynski 2016
Freshwater mussels (Unionida, Hyriidae)					
<i>Hyridella menziesi</i>	F	16,031		KU873121	Guerra et al. 2017
<i>Hyridella menziesi</i>	M	18,140	2,109	KU873122	Guerra et al. 2017
Freshwater mussels (Unionida, Margaritiferidae)					
<i>Cumberlandia monodonta</i>	F	16,099		KU873123	Guerra et al. 2017
<i>Cumberlandia monodonta</i>	M	17,575	1,476	KU873124	Guerra et al. 2017
Marine mussels (Mytilida, Mytilidae)					
<i>Musculista senhousia</i>	F	21,557		GU001953	Passamonti et al. 2011
<i>Musculista senhousia</i>	M	20,621	936	GU001954	Passamonti et al. 2011
Marine clams (Venerida, Veneridae)					
<i>Meretrix lamarckii</i>	F	20,025		KP244451	Bettinazzi et al. 2016
<i>Meretrix lamarckii</i>	M	19,688	337	KP244452	Bettinazzi et al. 2016
Marine clams (Venerida, Tellinidae)					
<i>Limecola balthica</i>	F	17,492		X	Present study
<i>Limecola balthica</i>	M	24,792	7,300	X	Present study
Marine clams (Venerida, Semelidae)					
<i>Scrobicularia plana</i>	F	16,170		X	Present study
<i>Scrobicularia plana</i>	M	26,270	10,100	X	Present study

To our knowledge, these values represent the biggest differences in genome size between the F and M mtDNAs in species with DUI (Table 3.1). These differences in length are partly explained by the presence of an insertion in the PCG *cox2* in the M mtDNA of both species ($F_{cox2} = 855$ bp and $M_{cox2} = 4\ 815$ bp in *L. balthica*, whereas $F_{cox2} = 861$ bp and $M_{cox2} = 5\ 679$ bp in *S. plana*) (Table 3.2), which is discussed in more details below. Otherwise, the two sex-specific mtDNAs in each species possess the 13 typical mitochondrial PCGs of metazoans (Boore, 1999), all present on the same strand, as is the case in most bivalve species (Bettinazzi et al., 2016; Doucet-Beaupré

et al., 2010; Jameson et al., 2003). Gene organisation is similar among the four genomes, except for the region between *cytb* and *trnM* in the M mtDNA of *S. plana*. Reorganisation events, such as inversion or transposition, as well as tandem duplication followed by random loss events, are common in animals mtDNAs (Boore, 1999, 2000).

The initiation and termination codons for the typical 13 PCGs encoded by the four mitogenomes are presented in Table 3.2. There are differences between sex-specific mtDNAs within species and also between species. For example, in *S. plana* a similar start codon between the F and M mtDNA for a particular gene is observed in 4 cases out of 13 (only two in *L. balthica*) whereas 6 PCGs out of 13 have a similar stop codon (4 in *L. balthica*). Overall, most of the PCGs use the ATD start codon found in metazoan mtDNAs (Wolstenholme, 1992): ATA occurs the most (15/52), followed by ATG (14/52) and ATT (11/52). TTG is also found in seven cases, and GTG in five, an observation that has been previously reported in molluscan species (Grande et al., 2002; Yuan et al., 2012). Most of the PCGs are terminated with the TAA (27/52) and TAG (20/52) codons. Some incomplete termination codons are also found, which are assumed to be completed by polyadenylation of their mRNAs (Ojala, Montoya, & Attardi, 1981).

Table 3.2. General features of *Scrobicularia plana* and *Limecola balthica* M and F mitochondrial genomes.

<i>Scrobicularia plana</i>			<i>Scrobicularia plana</i>			<i>Limecola balthica</i>			<i>Limecola balthica</i>		
F mt genome		M mt genome				F mt genome		M mt genome			
Gene/ Region	Length	Start/ Stop	Gene/ Region	Length	Start/ Stop	Gene/ Region	Length	Start/ Stop	Gene/ Region	Length	Start/ Stop
<i>cox1</i>	1554	ATA/TAA	<i>cox1</i>	1551	ATT/TAA	<i>cox1</i>	1554	ATA/TAA	<i>cox1</i>	1536	ATT/TAA
<i>nad4</i>	1323	TTG/TAG	<i>nad4</i>	1335	GTG/TAA	<i>nad4</i>	1341	ATA/TAG	<i>nad4</i>	1341	TTG/TAA
<i>trnH</i>	65		<i>trnH</i>	63		<i>trnH</i>	64		<i>trnH</i>	64	
<i>trnS2</i>	64		<i>trnS2</i>	65		<i>trnS2</i>	65		<i>trnS2</i>	64	
<i>trnE</i>	66		<i>trnE</i>	64		<i>trnE</i>	67		<i>trnE</i>	66	
<i>nad3</i>	348	ATA/TAA	<i>nad3</i>	348	ATA/TAA	<i>nad3</i>	348	ATA/TAG	<i>nad3</i>	348	ATG/TAA
<i>trnI</i>	68		<i>trnI</i>	68		<i>trnI</i>	69		<i>trnL</i>	68	
<i>trnK</i>	64		<i>trnK</i>	66		<i>trnK</i>	65		<i>trnK</i>	65	
<i>nad4L</i>	291	ATT/TAA	<i>nad4L</i>	285	ATT/TAG	<i>nad4L</i>	291	GTG/TAA	<i>nad4L</i>	276	TTG/TAG
<i>trnY</i>	64		<i>trnY</i>	64		<i>trnY</i>	63		<i>trnY</i>	67	
<i>trnT</i>	63		<i>trnT</i>	64		<i>trnT</i>	64		<i>trnA_1</i>	71	
<i>trnL1</i>	66		<i>trnL1</i>	65		<i>trnL1</i>	66		<i>trnL1</i>	65	
<i>trnD</i>	63		<i>trnD</i>	66		<i>trnD</i>	63		<i>trnD</i>	65	

<i>trnL2</i>	66	<i>trnL2</i>	66	<i>trnL2</i>	66	<i>trnL2</i>	65
<i>nad1</i>	924	GTG/TAG	<i>nad1</i>	927	ATT/TAG	<i>nad1</i>	924
<i>trnN</i>	66		<i>trnN</i>	65		<i>trnN</i>	65
<i>nad5</i>	1635	TTG/TAA	<i>nad5</i>	1704	ATG/TAG	<i>nad5</i>	1743
<i>trnR</i>	67		<i>trnR</i>	65		<i>trnR</i>	65
<i>cob</i>	1135	ATA/T**	<i>cob</i>	1135	ATG/T**	<i>cob</i>	1140
<i>cox2</i>	861	ATG/TAA	<i>trnW</i>	64		<i>cox2</i>	855
<i>trnV</i>	69		<i>rrnS</i>	765		<i>trnV</i>	66
<i>trnW</i>	66		<i>trnV</i>	66		<i>trnW</i>	66
<i>trnG</i>	68		<i>trnG</i>	68		<i>trnG</i>	66
<i>rrnS</i>	896		<i>cox2</i>	5679	ATG/TAG	<i>rrnS</i>	880
<i>trnM</i>	66		<i>trnM</i>	67		<i>trnM</i>	65
<i>atp8</i>	129	ATT/TAG	<i>atp8</i>	129	ATG/TAA	<i>atp8</i>	129
<i>trnS1</i>	71		<i>trnS1</i>	69		<i>trnS1</i>	69
<i>nad6</i>	510	ATA/TAA	<i>nad6</i>	510	ATT/TAG	<i>nad6</i>	510
<i>rrnL</i>	1269		<i>rrnL</i>	1556		<i>rrnL</i>	1249
<i>atp6</i>	714	ATG/TAG	<i>atp6</i>	705	ATT/TAA	<i>atp6</i>	735
<i>cox3</i>	882	ATA/TAG	<i>cox3</i>	882	ATA/TAG	<i>cox3</i>	882
<i>nad2</i>	996	ATG/TAA	<i>nad2</i>	1002	ATT/TAA	<i>nad2</i>	990
<i>trnP</i>	66		<i>trnP</i>	67		<i>trnP</i>	66
<i>trnQ</i>	67		<i>trnQ</i>	67		<i>trnQ</i>	66
<i>trnC</i>	63		<i>trnC</i>	61		<i>trnC</i>	61
<i>trnA</i>	65		<i>trnA</i>	65		<i>trnA</i>	67
<i>trnF</i>	63		<i>trnF</i>	65		<i>trnF</i>	66
						<i>trnF</i>	65

NOTE – T** stands for incomplete stop codon; # The exact size of the *Mcox2* gene in *Limecola balthica* is uncertain.

The usual 22 tRNAs are found in the four mitochondrial genomes, ranging from 61 to 71 bp, and most of them can be folded into the typical secondary structures (not shown). There is an exception in the M mtDNA of *L. balthica*, which is lacking *trnT*, but instead possess a second copy of *trnA*, which has been identified by MITOS in the same area, between *trnY* and *trnL1* (Fig. S3.1). This second copy of *trnA* (*trnA_1*) does not possess the typical secondary structure and it is not clear if it could be functional. This same mitochondrial genome contains also another copy of *trnP* (*trnP_1*) located next to the first *trnP* (*trnP_0*), and both possess usual secondary structures (Fig. S3.1). Additional tRNA copies are common in bivalve mitochondrial genomes (Bettinazzi et al., 2016; Milbury & Gaffney, 2005; Sun et al., 2015; Xu, Kanno et al., 2011). Concerning the missing *trnT*, it is possible that the *trnA_1* could be a mutated *trnT* (with a mutation in its anticodon that

changed it from *trnT* to a “secondary” *trnA*). It is also possible that an unusual secondary structure makes it unidentifiable by the tRNA search programs, and this will require further analyses.

Length variations among rRNA genes range from 765 bp to 896 bp and from 1249 bp to 1556 bp for 12S (*rrnS*) and 16S (*rrnL*), respectively (Table 3.2). Their locations are almost identical in the four genomes; *rrnL* is found between *nad6* and *atp6* whereas *rrnS* is found between *trnG* and *trnM*, except for the M genome of *S. plana*, in which it is located between *trnV* and *trnW* (Fig. 3.1; Tables 3.2 and 3.3).

Table 3.3. Number of nucleotides at gene boundaries in F and M mitochondrial genomes of *Scrobicularia plana* and *Limecola balthica*.

Boundary	<i>S. plana</i> F mtDNA	Boundary	<i>S. plana</i> M mtDNA	Boundary	<i>L. balthica</i> F mtDNA	Boundary	<i>L. balthica</i> M mtDNA
<i>cox1</i> <i>nd4</i>	10	<i>cox1</i> <i>nd4</i>	60	<i>cox1</i> <i>nd4</i>	0	<i>cox1</i> <i>nd4</i>	15
<i>nd4</i> <i>trnH</i>	2	<i>nd4</i> <i>trnH</i>	4	<i>nd4</i> <i>trnH</i>	6	<i>nd4</i> <i>trnH</i>	13
<i>trnH</i> <i>trnS2</i>	0	<i>trnH</i> <i>trnS2</i>	55	<i>trnH</i> <i>trnS2</i>	1	<i>trnH</i> <i>trnS2</i>	16
<i>trnS2</i> <i>trnE</i>	0	<i>trnS2</i> <i>trnE</i>	15	<i>trnS2</i> <i>trnE</i>	1	<i>trnS2</i> <i>trnE</i>	2
<i>trnE</i> <i>nd3</i>	15	<i>trnE</i> <i>nd3</i>	6	<i>trnE</i> <i>nd3</i>	15	<i>trnE</i> <i>nd3</i>	18
<i>nd3</i> <i>trnI</i>	10	<i>nd3</i> <i>trnI</i>	-17	<i>nd3</i> <i>trnI</i>	13	<i>nd3</i> <i>trnI</i>	-17
<i>trnI</i> <i>trnK</i>	3	<i>trnI</i> <i>trnK</i>	1	<i>trnI</i> <i>trnK</i>	2	<i>trnI</i> <i>trnK</i>	2
<i>trnK</i> <i>nd4L</i>	0	<i>trnK</i> <i>nd4L</i>	2	<i>trnK</i> <i>nd4L</i>	0	<i>trnK</i> <i>nd4L</i>	10
<i>nd4L</i> <i>trnY</i>	1	<i>nd4L</i> <i>trnY</i>	7	<i>nd4L</i> <i>trnY</i>	4	<i>nd4L</i> <i>trnY</i>	1
<i>trnY</i> <i>trnT</i>	-3	<i>trnY</i> <i>trnT</i>	-2	<i>trnY</i> <i>trnT</i>	-2	<i>trnY</i> <i>trnA_1</i>	-2
<i>trnT</i> <i>trnL1</i>	0	<i>trnT</i> <i>trnL1</i>	1	<i>trnT</i> <i>trnL1</i>	0	<i>trnA_1</i> <i>trnL1</i>	3
<i>trnL1</i> <i>trnD</i>	-1	<i>trnL1</i> <i>trnD</i>	-1	<i>trnL1</i> <i>trnD</i>	-1	<i>trnL1</i> <i>trnD</i>	-1
<i>trnD</i> <i>trnL2</i>	2	<i>trnD</i> <i>trnL2</i>	2	<i>trnD</i> <i>trnL2</i>	1	<i>trnD</i> <i>trnL2</i>	6
<i>trnL2</i> <i>nd1</i>	1	<i>trnL2</i> <i>nd1</i>	1	<i>trnL2</i> <i>nd1</i>	1	<i>trnL2</i> <i>nd1</i>	-14
<i>nd1</i> <i>trnN</i>	3	<i>nd1</i> <i>trnN</i>	-19	<i>nd1</i> <i>trnN</i>	1	<i>nd1</i> <i>trnN</i>	5
<i>trnN</i> <i>nd5</i>	0	<i>trnN</i> <i>nd5</i>	25	<i>trnN</i> <i>nd5</i>	0	<i>trnN</i> <i>nd5</i>	0
<i>nd5</i> <i>trnR</i>	-40	<i>nd5</i> <i>trnR</i>	4	<i>nd5</i> <i>trnR</i>	1	<i>nd5</i> <i>trnR</i>	-10
<i>trnR</i> <i>cob</i>	3	<i>trnR</i> <i>cob</i>	72	<i>trnR</i> <i>cob</i>	31	<i>trnR</i> <i>cob</i>	111
<i>cob</i> <i>cox2</i>	110	<i>cob</i> <i>trnW</i>	610	<i>cob</i> <i>cox2</i>	71	<i>cob</i> <i>cox2</i>	216
<i>cox2</i> <i>trnV</i>	5	<i>trnW</i> <i>rrnS</i>	3231	<i>cox2</i> <i>trnV</i>	3	<i>cox2</i> <i>trnV</i>	533
<i>trnV</i> <i>trnW</i>	0	<i>rrnS</i> <i>trnV</i>	101	<i>trnV</i> <i>trnW</i>	-1	<i>trnV</i> <i>trnW</i>	0
<i>trnW</i> <i>trnG</i>	2	<i>trnV</i> <i>trnG</i>	18	<i>trnW</i> <i>trnG</i>	5	<i>trnW</i> <i>trnG</i>	15
<i>trnG</i> <i>rrnS</i>	4	<i>trnG</i> <i>cox2</i>	1272	<i>trnG</i> <i>rrnS</i>	3	<i>trnG</i> <i>rrnS</i>	410
<i>rrnS</i> <i>trnM</i>	732	<i>cox2</i> <i>trnM</i>	57	<i>rrnS</i> <i>trnM</i>	1890	<i>rrnS</i> <i>trnM</i>	3697

<i>trnM</i>	<i>atp8</i>	11	<i>trnM</i>	<i>atp8</i>	13	<i>trnM</i>	<i>atp8</i>	9	<i>trnM</i>	<i>atp8</i>	11
<i>atp8</i>	<i>trnS1</i>	-2	<i>atp8</i>	<i>trnS1</i>	28	<i>atp8</i>	<i>trnS1</i>	3	<i>atp8</i>	<i>trnS1</i>	0
<i>trnS1</i>	<i>nd6</i>	32	<i>trnS1</i>	<i>nd6</i>	20	<i>trnS1</i>	<i>nd6</i>	34	<i>trnS1</i>	<i>nd6</i>	-2
<i>nd6</i>	<i>rrnL</i>	0	<i>nd6</i>	<i>rrnL</i>	0	<i>nd6</i>	<i>rrnL</i>	35	<i>nd6</i>	<i>rrnL</i>	12
<i>rrnL</i>	<i>atp6</i>	54	<i>rrnL</i>	<i>atp6</i>	180	<i>rrnL</i>	<i>atp6</i>	47	<i>rrnL</i>	<i>atp6</i>	40
<i>atp6</i>	<i>cox3</i>	30	<i>atp6</i>	<i>cox3</i>	8	<i>atp6</i>	<i>cox3</i>	52	<i>atp6</i>	<i>cox3</i>	17
<i>cox3</i>	<i>nd2</i>	34	<i>cox3</i>	<i>nd2</i>	56	<i>cox3</i>	<i>nd2</i>	81	<i>cox3</i>	<i>nd2</i>	105
<i>nd2</i>	<i>trnP</i>	52	<i>nd2</i>	<i>trnP</i>	3	<i>nd2</i>	<i>trnP</i>	0	<i>nd2</i>	<i>trnP_1</i>	78
<i>trnP</i>	<i>trnQ</i>	11	<i>trnP</i>	<i>trnQ</i>	25	<i>trnP</i>	<i>trnQ</i>	13	<i>trnP_1</i>	<i>trnP_0</i>	7
<i>trnQ</i>	<i>trnC</i>	15	<i>trnQ</i>	<i>trnC</i>	0				<i>trnP_0</i>	<i>trnQ</i>	35
<i>trnC</i>	<i>trnA</i>	0	<i>trnC</i>	<i>trnA</i>	4	<i>trnQ</i>	<i>trnC</i>	-1	<i>trnQ</i>	<i>trnC</i>	0
<i>trnA</i>	<i>trnF</i>	-1	<i>trnA</i>	<i>trnF</i>	3	<i>trnC</i>	<i>trnA</i>	1	<i>trnC</i>	<i>trnA_0</i>	13
<i>trnF</i>	<i>cox1</i>	161	<i>trnF</i>	<i>cox1</i>	372	<i>trnA</i>	<i>trnF</i>	-1	<i>trnA_0</i>	<i>trnF</i>	20
						<i>trnF</i>	<i>cox1</i>	163	<i>trnF</i>	<i>cox1</i>	250

NOTE – Negative values indicate that genes overlap.

Intraspecific divergences

Genetic distances of individual genes between F and M mtDNAs were analyzed in each species. The results show that the level of conservation is higher for both ribosomal RNA genes (*rrnS* and *rrnL*), and also for the PCG *cox1*, whereas NADH dehydrogenase subunit genes (*nad* series) are generally less conserved (Fig. 3.2a, c). These results are consistent with the general findings in DUI bivalves (Plazzi & Passamonti, 2019). For *S. plana* the average nucleotide divergence (measured as *p*-distance) of combined protein coding and ribosomal RNA genes is 0.43 ± 0.08 SD (based on values in Fig. 3.2a). For *L. balthica* the average *p*-distance of combined protein coding and ribosomal RNA genes is 0.41 ± 0.10 SD (based on values in Fig. 3.2c). The number of nonsynonymous substitutions per nonsynonymous sites (*Ka*) relative to the number of synonymous substitutions per synonymous sites (*Ks*) were also calculated (shown in Fig. 3.2b, d). This analysis provides an estimate of the degree of selection (either neutral, positive, or purifying) on each PCG. For both species and for all PCGs, all data points fall below the line representing neutrality (expressed as *Ka* = *Ks*; see dashed line in Fig. 3.2b,d), suggesting that the PCGs in *S. plana* and *L. balthica* accumulate more synonymous substitutions over evolutionary time relative to nonsynonymous substitutions. Therefore, all *Ka* and *Ks* rates supported an initial hypothesis of purifying selection for mitochondrial genes. To statistically test if the genes were indeed under

purifying selection, we conducted a Z-test of Selection (Kumar et al., 2016; Tamura et al., 2013). For all PCGs in both species, with the exception of *atp8* in *L. balthica*, there is strong statistical support (at $\alpha = 0.05$) for rejecting a null hypothesis of neutrality ($H_n: Ka = Ks$) in support of an alternate hypothesis for purifying selection ($H_a: Ka < Ks$). This is because all pairwise Z-test p-values were <0.01 for both hypotheses (Tamura et al., 2013). For *atp8* in *L. balthica* the Z-test of Selection did not support rejecting neutrality (p-value = 0.277 for $H_n: Ka = Ks$), a result that might be affected by the small size of this gene (129 bp), which in turn altered the statistical power of the test. With regards to the large insertion within the male *cox2* gene for both species, these data are of even greater interest as they suggest that this insertion likely does not render the gene functionless. Rather, *cox2* (the alignable parts) actually remains among one of the more relatively conserved mt genes (Fig. 3.2). Overall, our results are in line with what has been observed in other bivalves with DUI, i.e. despite the considerable divergence of DNA and amino acid sequences of PCGs there is evidence of strong purifying selection acting on these mitochondrial genes (Plazzi & Passamonti, 2019).

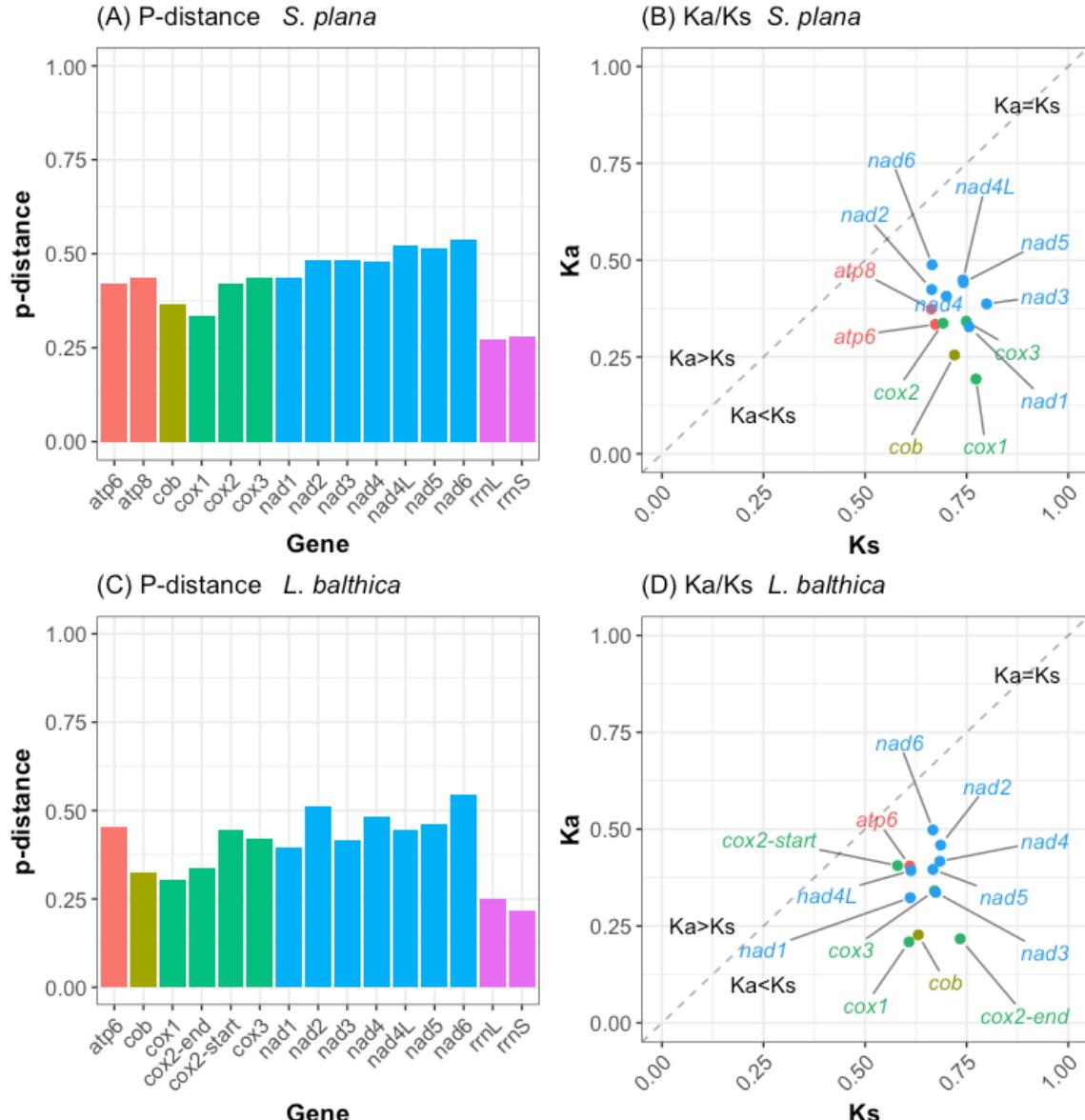


Figure 3.2. Individual M-versus-F gene conservation expressed as p-distance and the rates of synonymous and nonsynonymous substitutions within these same mt genes for *Scrobicularia plana* (**A & B, respectively**) and *Limecola balthica* (**C & D, respectively**). Genes are color-coded by gene family.

For comparative purpose, p -distances of individual genes (PCGs only) between F and M mtDNAs were also calculated for three species known for having the greatest F versus M DNA divergences in the families Mytilidae, Veneridae and Unionidae: i.e., the mytilid *Modiolus modiolus* (Robicheau et al., 2017), the venerid *Ruditapes philippinarum* (Passamonti & Scali, 2001), and the unionid *Quadrula quadrula* (Doucet-Beaupré et al., 2010) (Table 3.4). Our results show an average p -distance (for all 13 PCGs) of 0.45 for nucleotide (nt) and 0.53 for amino acid (aa) in *S. plana* and 0.44 (nt) and 0.53 (aa) in *L. balthica* (Table 3.4). To our knowledge, the

nucleotide and amino acid divergences reported in *S. plana* and *L. balthica* are the greatest reported among DUI species, surpassing those in freshwater mussels (Table 3.4), which were previously thought to exhibit the greatest divergences between their sex-specific mtDNAs (Doucet-Beaupré et al., 2010). On the other hand, the average uncorrected nucleotide divergence observed between the F and M PCGs of the marine mussel *Mytilus edulis* is about 0.23 (Breton et al., 2006). This low level of divergence has been proposed to be a consequence of masculinization events, which are characterized by an invasion of the male route of inheritance by an F mtDNA that becomes transmitted through sperm as a standard M mtDNA (Zouros, 2000; Zouros, 2013). These events reset the level of divergence between the F and M mitochondrial genomes to zero (Zouros, 2000; Zouros, 2013). Conversely, the high level of divergence observed in freshwater mussels has been hypothesized to be a consequence of a complete absence of masculinization events for over 200 million years in Unionida (Curiale & Kocher, 2002; Doucet-Beaupré et al., 2010). According to these studies, masculinization would be no longer possible in this taxon because of the existence of the M-specific extension of the *cox2* gene, a specialized feature of the unionid M mtDNA that would prevent recombination between the F and M mtDNAs, i.e. a step necessary for masculinization to occur (Curiale & Kocher, 2002; Doucet-Beaupré et al., 2010; Zouros, 2013). We thus propose that the high divergences observed between the F and M mtDNAs in *S. plana* and *L. balthica* are also related to an absence of masculinization events in these species because of the insertion in their *Mcox2* gene (described below).

Table 3.4. Intraspecific divergence (uncorrected *p*-distances) of the 13 mitochondrial protein-coding genes between M and F mtDNAs in DUI species. (a) nucleotides (b) amino acids.

a.

	<i>cox1</i>	<i>nad4</i>	<i>nad3</i>	<i>nad4L</i>	<i>nad1</i>	<i>nad5</i>	<i>cob</i>	<i>cox2</i>	<i>atp8</i>	<i>nad6</i>	<i>atp6</i>	<i>cox3</i>	<i>nad2</i>	Average
<i>Scrobicularia plana</i>	0.335	0.466	0.482	0.524	0.435	0.516	0.336	0.419	0.437	0.536	0.420	0.436	0.483	0.449
<i>Limecola balthica</i>	0.305	0.481	0.426	0.446	0.391	0.467	0.329	0.410	0.540	0.544	0.453	0.421	0.513	0.440
<i>Quadrula quadrula</i>	0.287	0.383	0.374	0.404	0.414	0.390	0.365	0.348	0.358	0.475	0.384	0.361	0.443	0.384
<i>Modiolus modiolus</i> *	0.280	0.375	0.373	-	-	-	-	-	-	-	0.466	0.375	-	0.374
<i>Ruditapes philippinarum</i>	0.222	0.300	0.278	0.468	0.240	0.299	0.281	0.336	-	0.313	0.298	0.274	0.347	0.305

b.

	<i>cox1</i>	<i>nad4</i>	<i>nad3</i>	<i>nad4L</i>	<i>nad1</i>	<i>nad5</i>	<i>cob</i>	<i>cox2</i>	<i>atp8</i>	<i>nad6</i>	<i>atp6</i>	<i>cox3</i>	<i>nad2</i>	Average
<i>Scrobicularia plana</i>	0.288	0.610	0.604	0.500	0.480	0.618	0.329	0.506	0.595	0.701	0.493	0.517	0.614	0.527
<i>Limecola balthica</i>	0.300	0.610	0.541	0.527	0.474	0.621	0.332	0.515	0.619	0.690	0.571	0.485	0.663	0.534
<i>Quadrula quadrula</i>	0.318	0.591	0.492	0.763	0.549	0.572	0.380	0.420	0.500	0.722	0.523	0.440	0.661	0.533
<i>Modiolus modiolus*</i>	0.233	0.445	0.449	-	-	-	-	-	-	-	0.576	0.422	-	0.425
<i>Ruditapes philippinarum</i>	0.123	0.334	0.303	0.779	0.216	0.346	0.273	0.419	-	0.560	0.343	0.314	0.444	0.371

NOTE – *The mtDNA sequence for *Modiolus modiolus* is incomplete.

Mitochondrial ORFans

Unassigned regions in the coding strand were searched for the presence of supernumerary PCGs and mtORFans with a minimal length of 150bp. Only ORFs encoding proteins with at least one predicted transmembrane domain were retained, because all sex-specific mtORFans characterized to date in DUI species (i.e. those encoding F-ORF in females and M-ORF in males), except for one case, possess at least one transmembrane domain or helix (TMH) (Breton et al., 2011; Milani et al., 2013; Mitchell et al., 2016; Guerra et al., 2017). In both *S. plana* and *L. balthica* F mtDNAs, only two unassigned regions were susceptible to contain supernumerary ORFs of >150bp, i.e. between *rrnS-trnM* and *trnF-cox1* (Table 3.3). However, the region between *trnF* and *cox1* was discarded because of the presence of a possible 5' extension of the *cox1* gene in all four genomes, an issue that will need to be assessed by looking at expression data. Otherwise, no ORFs corresponding to the expected profile were found in the F mtDNAs of *S. plana* and *L. balthica*. However, it is conceivable that smaller ORFs or ORFs encoding proteins without predicted transmembrane domain could be involved in DUI in these distantly-related species. The smallest F-ORF identified to date in a DUI species, i.e. in the unionid *Venustaconcha ellipsiformis*, is encoding an 89aa-long protein with one predicted TMH (Breton et al., 2011; Mitchell et al., 2016), whereas the smallest M-ORF is potentially encoding a 30aa-long protein without TMH in *Mytilus californianus* (Milani et al., 2013), although the functionality of this latter ORF remains to be demonstrated. Additional F mt sequences and expression data from *S. plana* and *L. balthica* will be necessary to clearly demonstrate the presence (or absence) of the F-*orf* gene in these species.

In both male mitochondrial DNAs, four unassigned regions contained supernumerary ORFs of >150bp, i.e. between *cob-trnW*, *trnW-rrnS*, *trnG-cox2* and *rrnL-atp6* for *S. plana* and between *cob-cox2*, *cox2-trnW*, *trnG-rrnS* and *rrnS-trnM* for *L. balthica* (Table 3.3). Five ORFs corresponding to the expected profile were found in *S. plana*, two between *trnW-rrnS* and three between *trnG-cox2*, whereas four ORFs with the expected profile were found in *L. balthica*, one between *trnG-rrnS* and three between *rrnS-trnM* (Fig. S3.2). Sequence similarity searches using PSI-BLAST (Altschul et al., 1997) against non-redundant protein sequences and SWISSPROT databases failed to detect significant sequence similarity with known proteins for all these ORFs except one (i.e. mtORFan1 in the unassigned region *trnG-rrnS* of *L. balthica* M mtDNA; Fig. S3.2). For this sequence, our results revealed moderately significant hits (E-values 2e-07) with microtubule-associated proteins. This result is interesting since previous *in silico* analyses of M-ORF sequences in other DUI species also indicated connections with cytoskeleton proteins involved in microtubule-binding and actin-binding (e.g. ankyrin) (Milani et al., 2013; Mitchell et al., 2016). These observations led to the hypothesis that M-ORF, with its predicted transmembrane domains, may target sites outside sperm mitochondria and be responsible for their cellular positioning in developing embryos (Milani et al., 2013). Indeed, studies have shown that, after fertilization, only in male DUI embryos sperm mitochondria remain grouped together, and are eventually sequestered in the germ line, whereas they are dispersed and/or destroyed in female embryos (reviewed in Zouros, 2013). M-ORFs are thus considered as ideal candidates for M mtDNA-derived masculinizing factors in DUI species (Milani et al., 2013; Mitchell et al., 2016).

At this moment, however, we cannot confirm nor disprove the presence of a M-*orf* gene in the M mtDNAs of *S. plana* and *L. balthica*. Pairwise alignments between the ORFs found in both species were performed but it was not possible to obtain satisfactory alignments that could provide support for identifying a conserved ORF (data not shown). Again, additional M mt sequences and expression data from *S. plana* and *L. balthica* will be necessary to test for the presence of a M-*orf* gene in these species.

Insertion in the male cox2 gene

Annotations with MITOS revealed an insertion of >4.5kb and >3.5kb in the *Mcox2* gene of *S. plana* and *L. balthica*, respectively, which is absent in the *Fcox2* genes of both species (Fig. 3.1 and 3.3). In *S. plana*, this in-frame insertion, if translated, means that the *cox2* gene would be 5679bp-long and, to our knowledge, would therefore encode for the longest *COX2* protein in the animal kingdom (i.e. 1893 amino acids). The sequencing of additional *S. plana* male individuals revealed that this insertion is conserved among different individuals from different populations (Tassé et al., unpublished), indicating that it is most probably functional. A multiple sequence alignment of *Homo sapiens COX2* and *S. plana FCOX2* and *MCOX2* amino acids indicates that the insertion is situated between the “heme-patch” region, containing an important residue Trp that functions as the point of electron entry from Cytochrome C, and the first Cu_a-binding center (Fig. 3.3 and Supplementary Fig. S3.3), and *in silico* analyses suggest that it does not contain any additional transmembrane domain (Tassé et al., unpublished). In-frame insertions resulting in enlarged *cox2* genes have also been reported in hymenoptera, ciliates, brown algae and microflagellates (Szafranski, 2017). As mentioned above, modifications to the *cox2* gene is a particular feature often found in DUI species. In unionid mussels, the longest *cox2* gene is found in the M mtDNA of *Hyridella menziesii* (1,380bp; Guerra et al., 2017) and is the result of a 3' extension, as in other DUI unionids (Chapman et al., 2008). Previous studies indicated an extra-mitochondrial localization of *MCOX2* as well as a possible involvement of the protein in reproduction in freshwater mussels (Chakrabarti et al., 2006; 2007), further supporting the link between gender and mtDNA transmission patterns. In the order Venerida, i.e. the order to which *S. plana* and *L. balthica* belong, an in-frame insertion of 300 nt has been reported in the species *Meretrix lamarckii* (family Veneridae) (Bettinazzi et al., 2016). However, such modifications of the *cox2* gene have not been reported in marine mytilid mussels, i.e. *Mytilus* spp., and in the venerid clam *Ruditapes philippinrum*, a duplicated copy of the *cox2* gene with a 3' extension has been reported, but in the F mtDNA (Bettinazzi et al., 2016; Zouros, 2013).

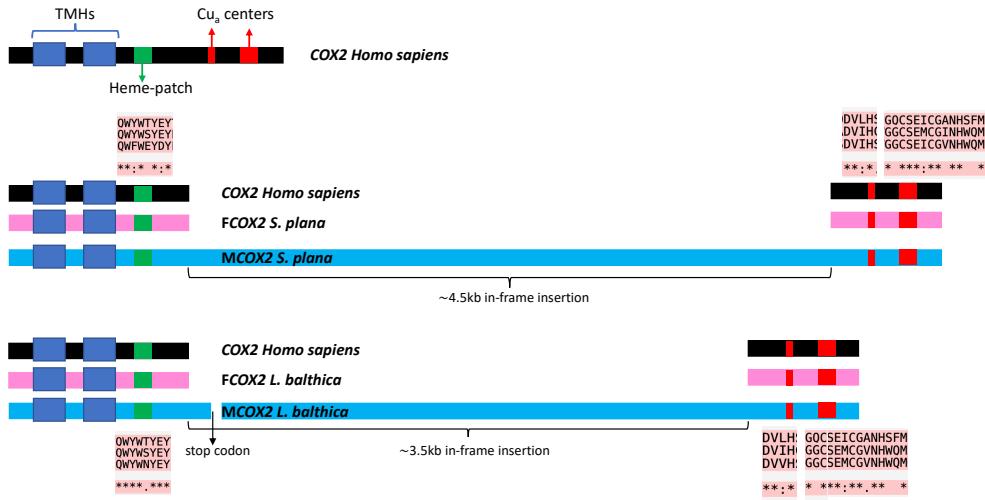


Figure 3.3. *Fcox2* and *Mcox2* structural features in *Scrobicularia plana* and *Limecola balthica*. For the sequence alignments of the “heme-patch” regions and Cu_a centers, identical amino acids are indicated by an *.

The situation in *L. balthica* is quite different than in *S. plana* regarding the *Mcox2* gene, which is split in two by an insertion. Specifically, this insertion divides *Mcox2* into *Mcox2a*, encoding the two transmembrane helices and the “heme-patch” region followed by a complete stop codon (TAA), and *Mcox2b*, encoding an enlarged intermembrane space and the Cu_a centers (Fig. 3.3 and Supplementary Fig. S3.4). This situation is confirmed by transcriptomic data (Illumina RNAseq reads from sperm cells, used to polish our nanopore reference mitogenomes), which indicate that both regions are transcribed (i.e. with discrete, non-overlapping *Mcox2a* and *Mcox2b* transcripts). Although it remains to be determined whether *Mcox2a* and *Mcox2b* are translated, it is worth noting that similar cases have been described in the freshwater mussel *Anodonta cygnea* (Chase et al., 2018) and the Hymenoptera *Campsomeris* spp. (Szafranski, 2017). For example, a translocation of a portion of the *nad5* gene has been reported in *A. cygnea*, and since the translocated portion and the rest of the gene both possessed their own start codon, the authors proposed that they could be transcribed and translated separately (Chase et al., 2018). In the genus *Campsomeris*, *cox2* is also split into two genes and this likely occurred through intragenic insertion of a cluster of several ORFs, one of which encodes a putative endonuclease that might have been directly involved in the process of *cox2* fission (Szafranski, 2017). However, no ORFs with similarity to an endonuclease were found in the *Mcox2* insertion of *L. balthica* (nor in *S. plana*).

According to Szafranski (2017), COXIIA and COXIIB polypeptides in *Campsomeris* spp. apparently assemble into a functional COXII heterodimer. Further studies will be needed to determine if this is also the case in *L. balthica*, or to clearly verify if a process of mitochondrial RNA splicing, because a group II intron has already been reported in the *cox1* gene of Annelida (Vallès, Halanych, & Boore, 2008), could be involved (i.e. *Mcox2* RNA splicing or *Mcox2a* and *Mcox2b* RNA trans-splicing into a single mRNA).

It also remains to be determined if *L. balthica* possesses an enlarged *MCOX2* or a *MCOX2* protein of a typical length. For example, the stop codon in *Mcox2a* could be read as a sense codon by a suppressor tRNA, i.e. a mutated (usually) tRNA that would insert an amino acid instead of initiating termination, as seen in several eukaryotes (Beier & Grimm, 2001; Ling, O'Donoghue, & Söll, 2015). The presence of duplicated *trn* genes in the M mtDNA of *L. balthica* might be a clue for such scenario. Modification or duplication of tRNAs have already been proven to allow codon reassignment by changing codons from stop-to-sense or sense-to-sense in yeast, green algae and metazoan mitochondria (Ling et al., 2015; Noutahi et al., 2019). Alternatively, termination can be avoided by ribosomal frame-shifting, a process that allows for protein merging from two or more overlapped ORFs (Rampersad & Tennant, 2018). This has been reported for the mitochondrial *cox1* gene of the diatom *Phaeodactylum tricornutum*: to avoid a stop codon, translational frameshift skips a nucleotide (called +1) (Secq et al., 2011). Similar strategies (+1 or -1 frameshifts) have also been recorded in metazoan mitochondrial genomes including ants, turtles and humans (Beckenbach, Robson, & Crozier, 2005; Russell & Beckenbach, 2008; Temperley et al., 2010). Further transcriptomic and proteomic analyses, and a closest look in mitochondrial tRNAs, will be required to validate any machinery involved in *Mcox2* transcription and translation in both *L. balthica* and *S. plana*. The results presented in this study clearly indicate that the relationship between *cox2* variations and DUI deserves much greater attention.

Conclusion

In summary, the description and comprehensive analysis of complete F and M mtDNAs of two newly discovered bivalve species with DUI from two additional families (Semelidae and Tellinidae), have led to new insights into DUI mitogenomics. Our results revealed uncorrected

amino acid *p*-distance of ~53% for PCGs between the M and F genomes of both species: this is the highest divergences reported in DUI species thus far. Hence the way (or ways) in which males of bivalves with DUI can tolerate heteroplasmy characterized by such high variability remains an important unanswered question (e.g. see Bettinazzi et al., 2019). Our results also highlighted an extremely unusual feature of the M genomes of *S. plana* and *L. balthica* compared to their female-transmitted counterparts. This feature is the presence of an important in-frame insertion (>3.5 to 4.5 kb) in the *Mcox2* gene. This is the longest insertion reported in the Kingdom Animalia, which remains to be functionally characterized. The reported data further indicate that the newly-sequenced M mitogenomes may be carrying lineage-specific genes (mtORFans) possibly involved in the DUI process. Analyses of complete mtDNAs from additional bivalve species and further protein-based studies are needed to elucidate the number, taxonomic distribution, evolution, and function of mtORFans and atypical *cox2* genes in this group, as well as the molecular mechanisms underlying DUI.

Acknowledgements

EP thanks the Molecular Core Facility at the University of La Rochelle. This work was funded by the French Agence Nationale de la Recherche (HySea project ANR-12-BSV7-0011; DRIVE project ANR-18-CE02-0004-01), La Rochelle Université, and by funding from the Natural Sciences and Engineering Research Council of Canada (grant no. RGPIN/435656-2013 to SB).

Supplementary Material

Supplementary figures

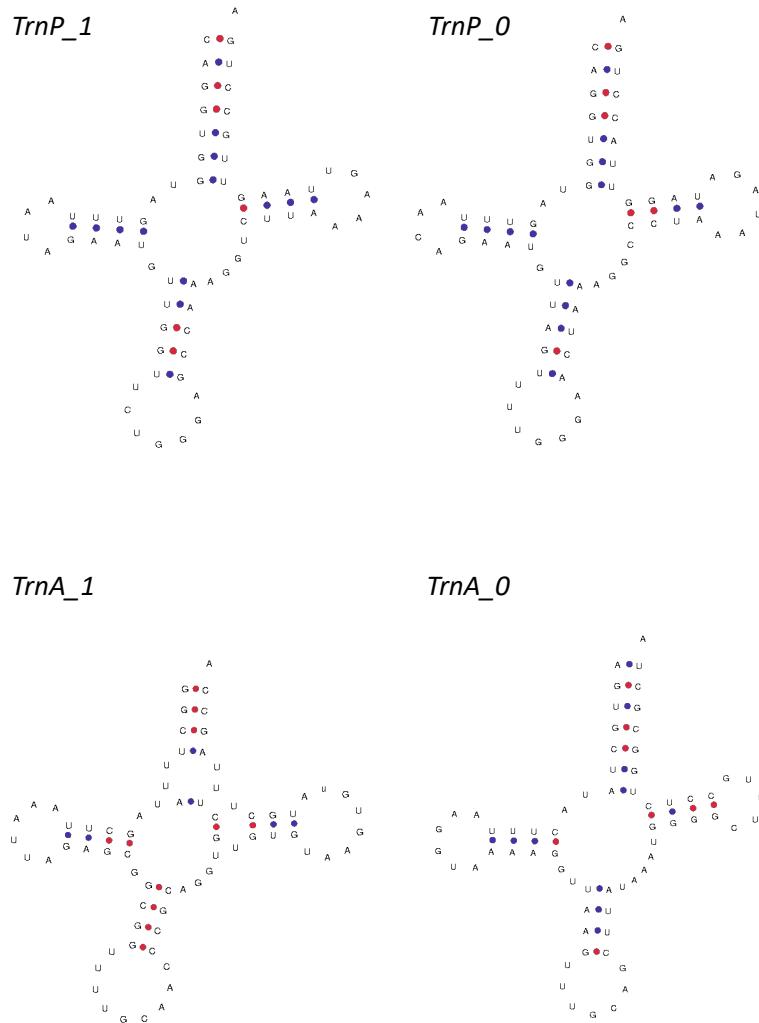


Figure S3.1. Secondary structures of *trnP* and *trnA* duplications in the M mtDNA of *L. balthica*.

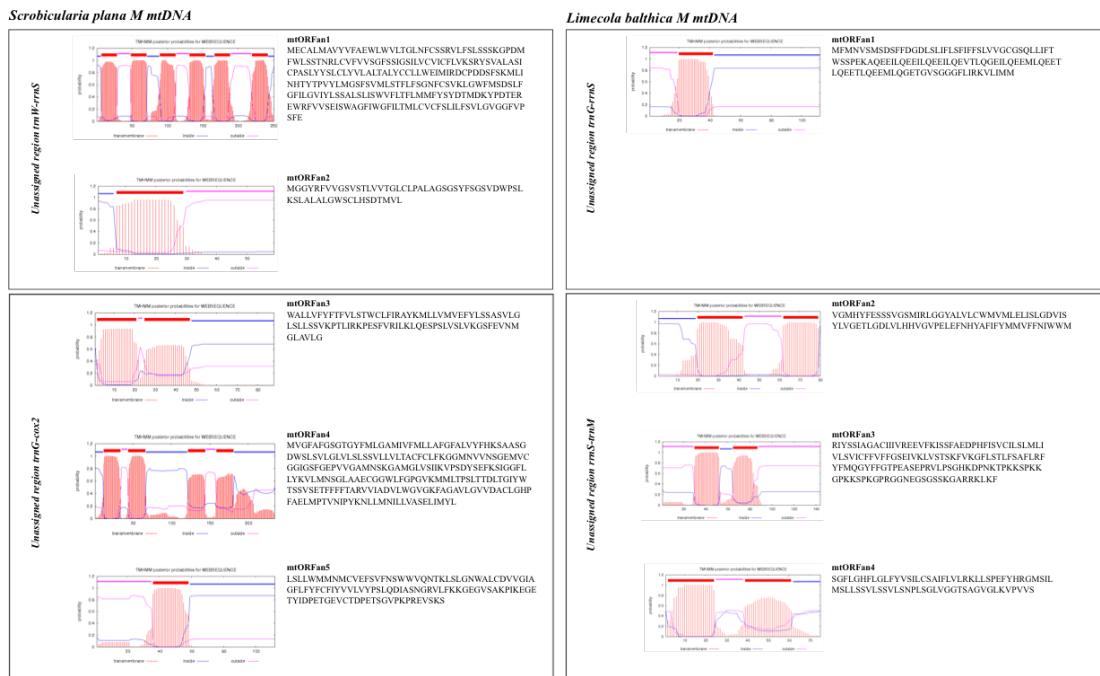


Figure S3.2. Sequence and secondary structure of predicted ORFs found in all unassigned regions longer than 150nt per sex per mitochondrial genome. Only ORFs with at least one predicted TMH were retained.

Alignment of COX2: *H. sapiens* vs *S. plana* FCOX2 and MCOX2

MSA

The multiple sequence alignment result as produced by T-coffee.

T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)

Cedric Notredame

SCORE=921

*

BAD AVG GOOD

*

Hsapie ns	:	91
Fcox2	:	90
Mcox2	:	84
cons	:	92

Hsapie ns	MAHAAQVGLQDATSP-----IMEELITFHDHALMIFLICFLVLYALFLT-----TKLTNT
Fcox2	-----MLGLIFLHDYLMCMCFMIMS VVGSLFLVVPKVSYFCGGIHFR
Mcox2	MNEVTSQGVNPRLRFVVDAVSPDWANSEGLDLILLHDYLMFCFFIMSAMLGVYVNF GKKGLM VIRGMYR
cons	

Hsapie ns	NISDAOEMETVWTLPAIIILVLIALPSLRILYMTDEVNDP-SLTIKSIGHOWYWTYEYTD-YGGL-----
Fcox2	NVYRNNSLELWWTIPIFLIIIMGYPFSVQLYAMGMNDKPKFITVKITGHQWYWSYEYLVHFPSLVSLAEEL
Mcox2	AVRNKQMLEFSAAIFAILLVIMMGYPFSVKLYSMGMNDTEKSMSLKA VGHQWFWEYDYFVD MENLELCDKY
cons	

Hsapie ns	GVFKEYNO-----
Fcox2	KVFPHWEGYGKKS LGDEGYALFDDSGPKTETVKDIVGDEEPEILASVFDLVL FDDSGPKTETVKDIVGDEEP
Mcox2	
cons	

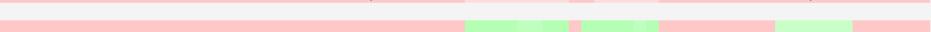
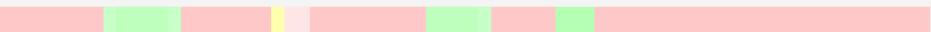
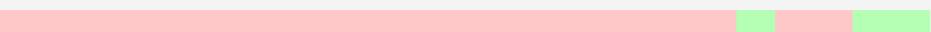
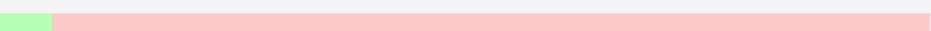
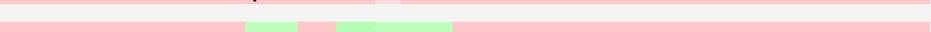
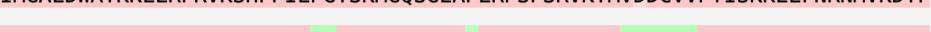
Hsapie ns	EILASFFGECALEAEMVKGGVGFSLNTYMSPPICDGKSSAPGVKGQPGSGAEGAPTPADDPRPKGSLPK
Fcox2	-----
Mcox2	-----
cons	

Hsapie ns	-----
Fcox2	VMYDPDVYRNLP PCGP KKR NRPV VETDIKNLPWSARRLLFKEVKEQSCREMLGTPEGDSSQQKG VPRLS
Mcox2	
cons	

Hsapie ns	-----
Fcox2	VLAARALPGDVVG DGD CPP QNPGNNVTPPVQGDAGVGLNP AVQDPVAMVLNNPAPAPA DAAENLAAEAPAG
Mcox2	
cons	

Hsapie ns	-----
Fcox2	GEGNAPVQEPAQAVLPSTPPRSGMEEMPSILLKKDPQLCPKF RGGPYKPRCAKQLDFNNPPHPPVPPFIPH
Mcox2	
cons	

Hsapie ns	-----
Fcox2	YLQNNLQGYIYPWP VPVGNNLGGGQAGVQGQVGGQAPAGGQVQVGGQAPVGGQVQVGGQAPAGG
Mcox2	
cons	

Hsapiens	
Fcox2	GFTKGSGSS
Mcox2	QGQGAGVDNGSGVDEDMSQCTRGMSEGVDGSGPTMLVDKSLPSGLYLYKKCEGGYNLGKGVDLLKVLKKVH
cons	
Hsapiens	
Fcox2	PEFGSGPRGEPSNDGPQPQGMGPLEFLYIFWCEGCIWYSKLNLAFFVQDLWAVNRSGRHAVNPELFSSA
Mcox2	
cons	
Hsapiens	
Fcox2	DKSFMCLWRMPLVSKVIMPLWHNLGLASQSVSEYHNDPSVLRVVTVLSPGPDNSVYMVKVIYNQGVNLESNF
Mcox2	
cons	
Hsapiens	
Fcox2	SKSMMFQGLSVGLNSDIKVALIWLVKGGPEDINGPNRPSRPFKGPSVGGTSKSDPFLAGLTNRPMDPSP
Mcox2	
cons	
Hsapiens	
Fcox2	EESRPVQDEPSNGGSGAQGGAEGSTQESEWSEGSYDGGTGRADRCLLDPSPKRSGGEAGPKEPQPADEAGEQEPS
Mcox2	
cons	
Hsapiens	
Fcox2	LPEGSNTSVSDLSCFPGLNLVFSGASNLLDDKPLGAAYDGNNMVIDENNLIELSDEEDMEGVE
Mcox2	
cons	
Hsapiens	
Fcox2	VNNPSSQLGSGFALSFIFEMMDPYDPSEHEDQVVEATIVEMPDPKTGSQLVPEEELPAPEPVQEIEPLD
Mcox2	
cons	
Hsapiens	
Fcox2	ESRSEDIPRPSAQCISSFFDMKVFCFFMSSPFLGISCPFGDFNPWDIDYALLGAGSVMGGYDSLYEDGF
Mcox2	
cons	
Hsapiens	
Fcox2	KLFMDTGPPKKNSSPGGPEQGSSGESRALLPKNSSVKKYGSLGNTPFPHTPDHSKVTEPLPGTVVPKSY
Mcox2	
cons	
Hsapiens	
Fcox2	IMGALDWAYKKLLRPRVKSHPPILPGTSKMGQSGEAPLRPSRSVKTMDDCVVPTISKLLFNRNMVKDYP
Mcox2	
cons	

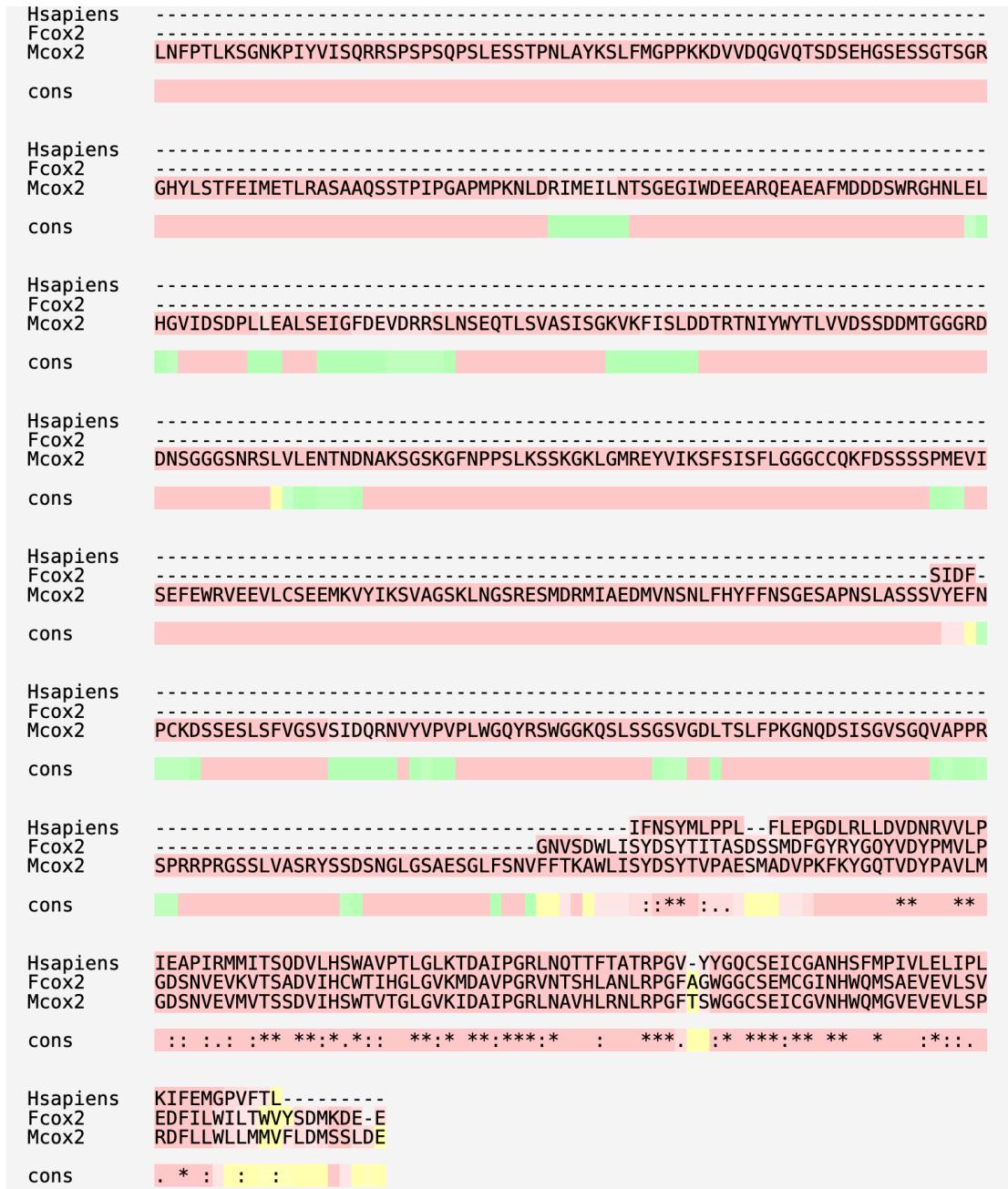


Figure S3.3. *Scrobicularia plana* F and M COX2 proteins alignment. Identical amino acids are indicated by an *.

Alignment of COX2: *H. sapiens* vs. *L. balthica* FCOX2 and MCOX2a

MSA

The multiple sequence alignment result as produced by T-coffee.

T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507)

Cedric Notredame

SCORE=848

*

BAD AVG GOOD

*

Hsapie... : 88
Fcox2Lbal : 89
Mcox2aLbal : 86
cons : 84

Hsapie... MAHAAQVGLODATSPIMEELITFHDHALMIIFLICFLVLYAL-----	FLTLTTKLNTNTNISDAOEMET
Fcox2Lbal MFG---FSISSLR-----	PCLCVFYNHISCSKCAFFLVPKVSYFCGGIHFRNVYRNNSLEL
Mcox2aLbal MKE---LCVENWGPKGKEMVLLHDYFVLFGFVIMAIIFTVLVVYLRKGGLMTRGMYRNVRVNQGLEF	

cons * [color bar] . . . [color bar] : . : . [color bar] : * : : : *
--

Hsapie... VWTILPAIIILVLIALPSLRILYMTDEVNDPS-LTIKSIGHOWYWTYEYTDYG-GLI-----	
Fcox2Lbal WWTIIPILLIIMGYPSEVQLYAMGMNDKPKFITVKVTGHQWYWSYEYLVHLPSLISLADELGVFN	NYN
Mcox2aLbal SWSVLPVMVLLVCMGYPSCSLYDGLSPKSMGLKVTGHQWYWNYEYFVNIEELVDQCLTKALESAK	

cons * : ; * : ; : . * : ** . : . * : * : * : * : * : : * : [color bar]

Hsapie... -----FNSYMLPPL--FLEPGDLRLLDVDRNRVVLPIEAPIRMMITSQDVL	
Fcox2Lbal NOGASKGELEANVOFEGVSDWLISYDSYTVNNVDGVL	ESGFRYQQFVDYPMVLPGDSNVEVKVTSADV
Mcox2aLbal GPGVLANSGPKKETLEGVIEG-----AEAEAAPMSSAK	

cons [color bar] : . * ..

Hsapie... HSWAVPTLGLKTDAIPGRLNOTTFTATRPGVY-YGQCSEICGANHSFMPIVLEIPLKIFEMGPVFTL	
Fcox2Lbal HCWTIOGLGVKMDAVPGRVNSAHFSDLRPGFSAWGGCSEMCGVNHWQMSAEVEVLTVSDFILWILT	WVY
Mcox2aLbal GPGVLANSGPKKETLEGVIEGLSQSLR-----	

cons . : * * : ; * : : * : [color bar] [color bar] [color bar] [color bar] [color bar] [color bar] : [color bar] [color bar] [color bar]
--

Hsapie... -----	
Fcox2Lbal SDIKLD	
Mcox2aLbal -----	

cons [color bar] [color bar] [color bar]
--

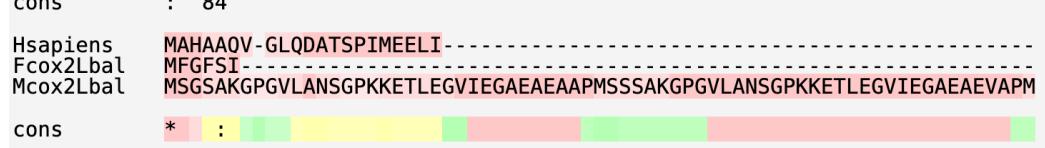
Alignment of COX2: *H. sapiens* vs. *L. balthica* FCOX2 and MCOX2b

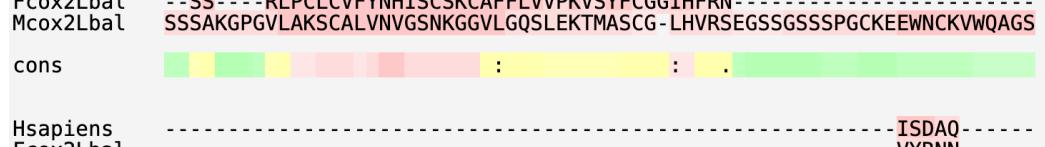
MSA

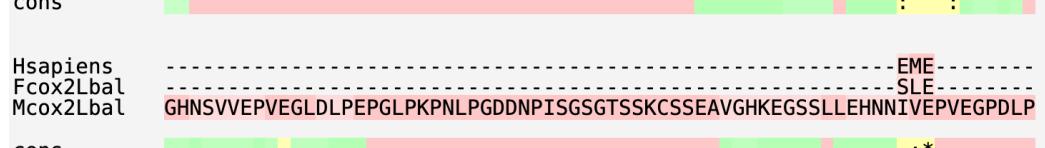
The multiple sequence alignment result as produced by T-coffee.

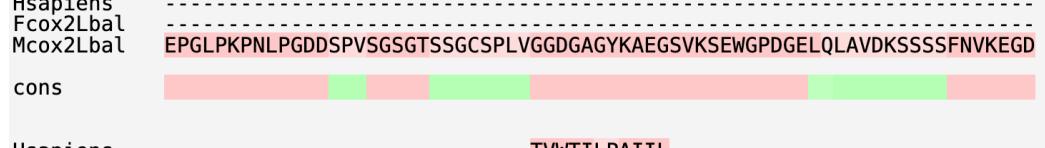
T-COFFEE, Version_11.00.d625267 (2016-01-11 15:25:41 - Revision d625267 - Build 507
 Cedric Notredame
 SCORE=845
 *
 BAD AVG GOOD

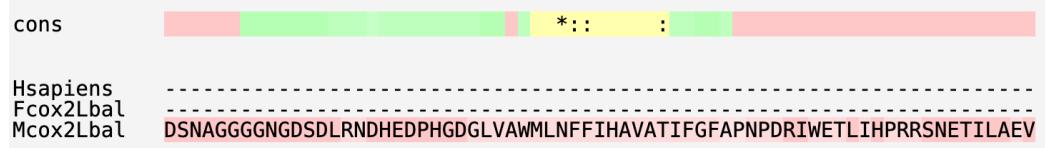
*
 Hsapiens : 79
 Fcox2Lbal : 75
 Mcox2Lbal : 44
 cons : 84

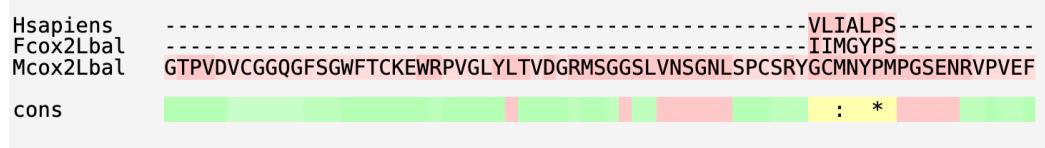
Hsapiens MAHAAQV -GLQDATSPIMEELI-----
 Fcox2Lbal MFGFSI-----
 Mcox2Lbal MSGSAKGPVLANSGPKKETLEGVIEGAEAEAAPMSSAKGPVLANSGPKKETLEGVIEGAEAEVAPM
 cons * : : -----


Hsapiens -----TFHDHALMIIFLICFLVLYALFLTTLTTKLNTN-----
 Fcox2Lbal --SS--RLPCLCVFYNHISCSKCAFFLVPKVSYFCGGIHFRN
 Mcox2Lbal SSSAKGPVLAKSCALVNVGNSNKGGVQLGQSLEKTMASCG-LHVRSEGSSSSPGCKEENNCKVWQAGS
 cons : : : -----


Hsapiens -----ISDAQ-----
 Fcox2Lbal -----VYRNN-----
 Mcox2Lbal RAGSAEQPAKAWTWSGSSWIVKAEGDKSRVFTCFESVDGKTRYKANYPVDENNNNFMGVLNNECFSFL
 cons : : : -----


Hsapiens -----EME-----
 Fcox2Lbal -----SLE-----
 Mcox2Lbal GHNSVVEPVGLDLPEPGLPKPNLPGDDNPISGSGTSSKCSSEAVGHKEGSSLLEHNНИVEPVEGPDLF
 cons : : : -----


Hsapiens -----
 Fcox2Lbal EPGLPKPNLPGDDSPVSGGTSSGCPLVGGDGAGYKAEGSVKSEWGPDGEQLAVDKSSSFNVKEGD
 Mcox2Lbal
 cons : : : -----


Hsapiens -----TVWTILPAIIL-----
 Fcox2Lbal -----LWWTIIPILLI-----
 Mcox2Lbal DMSGSRCCGPWVKAGNFSPYSCTFFGGSWWSFWIHCMDNGGDANPVGGTGGGLDTPDQQDNAENE
 cons : : : -----


Hsapiens -----
 Fcox2Lbal DSNAGGGGNGDSLRNDHEDPHGDGLVAWMLNFFIHAVATIFGFAPNPDRIWETLIHPRRSNETILAEV
 Mcox2Lbal
 cons : : : -----


Hsapiens -----VIALPS-----
 Fcox2Lbal -----IIMGYP-----
 Mcox2Lbal GTPVDVCGGQGFSGWFTCKEWRPVGLYLTVDGRMSGGLVNSGNLSPCSRGCNYPMPGSENRPVVEF
 cons : : : -----




Figure S3.4. *Limecola balthica* F and M COX2 proteins alignment. Identical amino acids are indicated by an *.

Discussion

L'importance des bivalves comme indicateurs écologiques de la pollution des eaux (Campos et al., 2012) et leur impact économique lié à leur valeur nutritive (Naylor et al., 2000), rendent les connaissances de leurs stratégies reproductives essentielles pour favoriser le maintien de leurs populations naturelles, ainsi qu'un élevage efficace et responsable de leurs populations en culture (Shumway, 2011). L'avancée des technologies de ces dernières années, notamment le séquençage nouvelle génération, a donné naissance à une optimisation des investigations des facteurs responsables de la détermination du sexe chez les espèces animales. Pourtant, les études réalisées sur les bivalves, qui représente le deuxième groupe le plus diversifié des mollusques, demeurent peu nombreuses (Breton et al., 2018). L'écologie et la biologie des bivalves sont suffisamment diverses pour constituer une riche source d'informations permettant de mieux comprendre l'évolution du sexe et les mécanismes de la détermination du sexe en général. Aussi, avec plus de 25,000 espèces vivantes (Breton et al., 2018), ce groupe taxonomique est suffisamment diversifié pour fournir des exemples uniques de déterminisme sexuel, comme par exemple le seul système de détermination du sexe qui impliquerait directement les mitochondries. Tel que suggéré par Breton et al. (2011) le système atypique de double transmission uniparentale de l'ADNmt (DUI) retrouvé chez plus d'une centaine d'espèces de bivalves (Gusman et al., 2016), serait fort probablement impliqué dans la détermination du sexe, mais cela n'a pas encore été clairement démontré. L'objectif général de cette thèse de doctorat consistait à mieux comprendre les mécanismes sous-jacents à la DUI et la détermination du sexe chez les bivalves. Deux groupes taxonomiques ont été ciblés : (1) les moules d'eau douce, étant donné que des espèces proches parentes hermaphrodites (sans DUI) et gonochoriques (avec DUI) ont été rapportées dans ce groupe seulement et (2) les palourdes marines, avec deux nouvelles espèces à DUI découvertes très récemment dans deux nouvelles familles, et qui représentent une nouvelle piste à explorer afin d'identifier des éléments génétiques mitochondriaux partagés entre des espèces éloignées qui pourraient être responsables de la régulation de la DUI.

Aspects nucléaires sous-jacents à la détermination du sexe chez les bivalves

Un des objectifs des deux premiers chapitres de cette thèse était de définir quels facteurs nucléaires pourraient être responsables de la détermination du sexe chez les moules d'eau douce, qu'elles soient gonochoriques à DUI ou hermaphrodites à SMI. Pour ce faire, deux études transcriptomiques ont été réalisées ; la première a permis d'identifier un set de gènes orthologues aux espèces gonochoriques à DUI *Venustaconcha ellipsiformis* et *Utterbackia peninsularis*, responsables de la différentiation des gonades mâles et femelles. L'étude des gènes différentiellement exprimés entre les sexes pour chacune des espèces, ainsi que l'ajout d'une analyse blast réciproques entre espèces, a permis d'amoindrir l'influence de facteurs environnementaux ou espèce-spécifiques non-désirés pour cette question. Globalement, nous avons pu démontrer l'influence de *Sry*, *Fem1* et *dmrt1* dans le développement des mâles, comme chez les nématodes, les mouches à fruits et les vertébrés (Doniach & Hodgkin, 1984; Kopp, 2012; Wallis et al., 2008), et de *FoxL2*, *Beta-catenin* et *Tra-1* dans celui des femelles, comme chez les mollusques, les nématodes, les oiseaux et les drosophiles (Beukeboom & Perrin, 2014; Kim et al., 2017; Uhlenhaut & Treier, 2006). Ces résultats viennent s'ajouter à ceux de d'autres études qui ont démontré que les gènes du déterminisme sexuel qui étaient considérés longtemps comme uniquement associés aux mammifères ou aux oiseaux, sont en fait également présent chez d'autres espèces dont le système sexuel n'implique pas nécessairement des chromosomes sexuels XY ou ZW. Il a donc été déduit que certaines parties du déterminisme sexuel pouvaient être bien conservées entre taxons très lointains, et que la divergence à l'origine de la grande diversité des systèmes sexuels porteraient surtout sur des mécanismes régulatoires (Capt et al., 2018). Pour un meilleur récapitulatif des gènes impliqués chez nos espèces, voir le Tableau 1.4 (Chapitre 1) qui présente également les résultats d'études similaires chez d'autres espèces de bivalves, tels que *R. philippinarum* et *H. schlegelli* (Ghiselli et al., 2012 ; Shi et al., 2015).

L'intérêt de la deuxième étude transcriptomique, est la réalisation d'un transcriptome hybride comportant les données de deux espèces afin de pouvoir réaliser des analyses d'expression différentes de gènes entre celles-ci. Le but étant d'identifier les gènes responsables de la détermination et différentiation sexuelles entre mâles et femelles d'espèces gonochoriques et ceux de la différentiation gonadique des hermaphrodites, le choix s'est porté sur deux espèces proches

parentes ; l'espèce gonochorique *Utterbackia peninsularis*, également utilisée dans le chapitre 1, et l'espèce hermaphrodite *Utterbackia imbecillis*. L'avantage d'utiliser des espèces proches-parentes était de limiter le temps de divergence entre les séquences afin d'avoir une meilleure idée des facteurs initiaux responsables de ces processus. Les résultats ont démontré des mécanismes similaires au sein des deux espèces puisque la plupart des gènes les plus exprimés chez les mâles et les femelles d'*U. peninsularis* se retrouvent également surexprimés chez *U. imbecillis*, mais à moindre mesure. Ceci s'explique par le fait que les tissus gonadiques des hermaphrodites nécessitent l'action combinée des gènes responsables de la différentiation des tissus testiculaires et ovariens, contrairement aux tissus gonochoriques qui concentrent leurs efforts sur un seul sexe. Un gène responsable de la différentiation des mâles était donc surexprimé chez les mâles par rapport aux hermaphrodites et aux femelles, et surexprimé chez les hermaphrodites comparés aux femelles. Il est intéressant de constater que les gènes impliqués dans la régulation sexuelle des espèces hermaphrodites sont les mêmes que ceux régissant les espèces gonochoriques (voir Tableau 2.3 du Chapitre 2).

Malgré une contribution importante de ces études, il est important de souligner qu'elles ont été réalisées sur des gonades d'individus matures et que les facteurs clés responsables de la détermination du sexe s'enclenchent évidemment plus tôt, durant le développement embryonnaire. À ma connaissance, il n'existe qu'une seule étude qui a su établir une fenêtre durant laquelle le sexe est déterminé chez les bivalves, et qui s'est intéressée aux facteurs génétiques et environnementaux qui agissent à ce moment-là. Cette étude, réalisée chez l'huître *Crassostrea gigas*, a pu montrer que l'expression d'un orthologue au gène *sox9*, Cg-SoxE, caractérisé dans une étude précédente (Naimi, Martinez, Specq, Mrac, et al., 2009; Naimi, Martinez, Specq, Diss, et al., 2009; Santerre, Sourdaine, & Martinez, 2012), était à son maximum dans les larves lorsque le sexe n'était pas encore discernable (Santerre et al., 2013). Cela fait de lui, pour le moment, un candidat clé agissant potentiellement comme activateur de la détermination du sexe. Toutefois, les facteurs génétiques ne sont pas les seuls acteurs impliqués, car la température aussi contribuerait au destin sexuel de ces individus, des résultats soutenus par une étude sur l'huître *Pinctada margaritifera* (Teaniniuraitemoana et al., 2016). Un projet de recherche plus récent s'est également intéressé au niveau d'expression de plusieurs des gènes liés à la différentiation du sexe chez les bivalves (*foxl2*, *wnt4*, β -*catenin*, *fst*, *dmrt*, *dax1*, *soxe*, and *soxh*; Li et al., 2016; Ni et al., 2012; Patnaik et al., 2016;

Teaniniuraitemoana, 2014; Tong et al., 2015; N. Zhang et al., 2014), chez des juvéniles du pétoncle *Patinopecten yessoensis* (R. Li et al., 2018). Bien qu'ils aient pu définir le rôle essentiel de *foxl2* et *dmrt1l* dans les différentiations des gonades femelles et mâles, respectivement, aucun résultat n'a été décrit comme déclencheur de cette cascade moléculaire puisqu'aucun de ces gènes ne présentait une expression différente lors des stades de gonades indifférenciées (Li et al., 2018). Ceci démontre l'importance de réaliser des études centrées sur des embryons, afin d'identifier plus précisément les véritables acteurs de la détermination du sexe chez les bivalves, dont les espèces à DUI.

Aspects nucléaires sous-jacents à la DUI

Un deuxième objectif de cette thèse était d'identifier des gènes candidats potentiellement liés à la double transmission uniparentale des mitochondries. Ainsi, nous avons focalisé nos analyses sur des catégories fonctionnelles et des gènes connus comme étant associés à des processus responsables de l'élimination des mitochondries paternelles chez les espèces animales (Punzi et al., 2018 ; Lee & Han, 2017). Nous avons ensuite comparé nos résultats issus d'espèces à DUI vs. espèces à SMI, afin de mettre en lumière les mécanismes qui permettraient la rétention des mitochondries paternelles dans les embryons mâles à DUI. Un de ces mécanismes, identifié lors de l'analyse transcriptomique du chapitre 1 comparant deux espèces de moules d'eau douce à DUI, est une modification des processus du système d'ubiquitination-protéasome (Capt et al., 2018). Chez les espèces animales à SMI, le mode d'action proposé pour l'ubiquitination consiste à marquer les mitochondries paternelles en vue de leur dégradation pendant la spermatogenèse et / ou après la fécondation (Sato & Sato, 2017 ; Punzi et al., 2018). Les résultats du chapitre 1 supportent l'hypothèse proposée par Kenchington et al. (2002) qu'un processus d'ubiquitination altéré serait impliqué dans la rétention des mitochondries paternelles dans les embryons mâles à DUI. Trois facteurs particuliers, initialement présentés dans une étude similaire à la nôtre chez la palourde marine *R. philippinarum* (Ghiselli et al., 2012), ont été trouvés surexprimés dans nos transcriptomes mâles. Ces facteurs sont associés à l'ubiquitination ou la reproduction chez d'autres espèces animales (Milani et al., 2013). Le premier, *psa6* (proteasome subunit alpha 6 ou sous-unité alpha 6 du protéasome 20S), serait important dans la différentiation et/ou la détermination du sexe, notamment dans la différentiation des gonades mâles. Le deuxième, *anubl1* (AN1 zinc finger ubiquitin-like domain ou domaine AN1 en doigt de zinc de type ubiquitine), aurait une fonction de

marquage des mitochondries des spermatozoïdes, via son domaine ubiquitine, pour les différencier de celles de la lignée maternelle issues des œufs, et agirait de concert avec *birc* (baculoviral IAP repeat-containing 4 [inhibitor of apoptosis]), qui grâce à sa fonction anti-apoptique, protègerait ainsi les mitochondries paternelles de la dégradation (Capt et al., 2018; Ghiselli et al., 2012; Milani et al., 2013).

Dans le deuxième chapitre, nous avons comparé les transcriptomes de deux espèces proches parentes, une à DUI et l'autre à SMI, permettant ainsi de différencier les acteurs responsables de la DUI vs. ceux de la SMI. Les résultats liés aux mécanismes d'ubiquitination ont été réévalués, mais comme ce processus n'est pas le seul qui est responsable du maintien de la SMI chez les espèces animales, nous avons aussi pris en compte ces autres processus dans nos résultats, telles que les nucléases, les polymérases, et les facteurs impliqués dans l'auophagie/mitophagie (Lee & Han, 2017; Punzi et al., 2018). Étant donné que ces processus ont tous des implications à des moments différents du développement (durant la spermatogenèse, ou après la fertilisation), cela définit chez quel parent il faut investiguer chacun des acteurs responsables de l'élimination des mitochondries. Les hypothèses attendues, quant aux gènes différemment exprimés pour chacun de ces mécanismes, sont présentées dans le Tableau 4.1.

Tableau 4.1. Tableau récapitulatif des gènes différemment exprimés attendus pour chaque paire de comparaison (MF, HF, HM), selon chaque catégorie fonctionnelle liée à la transmission mitochondriale.

	MF	HF	HM
Ubiquitination	+	+	- / 0
Nucléases / Endonucléases	- / 0	+	+
Autophagie / Mitophagie - Avant fertilisation	- / 0	+	+
Autophagie / Mitophagie - Après fertilisation	-	- / 0	+
Méthylation	+	- / 0	-

NOTE – Les processus d'autophagie/mitophagie pouvant avoir lieu autant durant la spermatogenèse qu'après la fertilisation, deux hypothèses sont proposées. MF : comparaison mâle-femelle, HF : hermaphrodite-femelle, HM : hermaphrodite-mâle. + : le gène est surexprimé dans le premier sexe de la comparaison, - : le gène est surexprimé dans le deuxième sexe de la comparaison, 0 : signifie que la différence d'expression du gène dans cette comparaison est nulle. Par exemple, HM + signifie que le gène est supposé être surexprimé chez les hermaphrodites comparé aux mâles.

Par exemple, une surexpression des gènes liés à l'ubiquitination est attendue chez les mâles à DUI et les hermaphrodites à SMI par rapport aux femelles. Nos résultats suggèrent que les gènes *psa* et le facteur d'ubiquitination *F-box39*, pourraient être des éléments clé de ce processus, puisque différemment exprimés dans les tissus testiculaires de nos mâles et hermaphrodites, et connus pour être impliqués dans la voie de dégradation « ubiquitine-protéasome » des protéines (Ho et al., 2006). Ces résultats sont en accord avec ceux obtenus pour la palourde marine *R. philippinarum* (Punzi et al., 2018). En revanche, nous n'avons pas identifié de facteurs directement liés à la transmission paternelle des mitochondries chez l'espèce à DUI dans cette catégorie « ubiquitination ».

Concernant la catégorie nucléases, il est connu que chez l'espèce modèle *Drosophila melanogaster*, c'est l'action des nucléases qui se charge d'assurer une transmission strictement maternelle de l'ADNmt. Par exemple, l'endonucléase G (gène *endoG*) dégrade les nucléoides (le contenu génétique des mitochondries) pendant la spermatogenèse, avant même la fertilisation (Chan & Schon, 2012; DeLuca & O'Farrell, 2012; Zhou et al., 2016). Une surexpression de tels gènes était donc attendue chez les hermaphrodites SMI par rapport aux mâles et femelles à DUI. Correspondant à cette situation, nous avons identifié *exoG*, qui est un paralogue d'*endoG* (Cymerman et al., 2008). *ExoG* pourrait donc participer à la SMI chez les individus hermaphrodites, en s'attaquant au matériel génétique des mitochondries, lors de la spermatogenèse, et les structures seraient ensuite dégradées dans l'embryon par des facteurs d'ubiquitination codés par les femelles, comme décrit chez la mouche à fruit (Sato & Sato, 2017).

Un autre processus lié à la dégradation des mitochondries chez les espèces à SMI, est le phénomène d'autophagie, ou plus spécifiquement la mitophagie, dont les facteurs impliqués ont été listés dans une revue de littérature (Lee & Han, 2017). La mitophagie agit normalement durant la spermatogénèse ou après la fertilisation pour prévenir la transmission des mitochondries paternelles (Lee & Han, 2017 ; Sato & Stao, 2017). On s'attendait donc à une surexpression de tels facteurs chez les hermaphrodites SMI par rapport aux mâles et femelles à DUI. Nous avons ainsi identifié les protéines TANK-binding kinase 1 et Ras-like GTP-binding YPT1, qui sont toutes deux impliquées dans le processus de mitophagie chez d'autres espèces (Heo et al., 2015; Lewandowska et al., 2013). Ces facteurs pourraient donc également être responsables du maintien de la SMI chez

les moules d'eau douce hermaphrodites et pourraient agir pendant la spermatogenèse au côté du gène *exoG* ou encore après la fertilisation.

Une dernière catégorie d'intérêt était la méthylation, au vu des résultats obtenus dans le chapitre 1 (Capt et al., 2018), mais également parce qu'un tel processus est responsable de la transmission de l'ADN chloroplastique chez l'algue *Chlamydomonas* (Nishimura, 2010). Notre hypothèse était que si la méthylation est impliquée dans la DUI des mitochondries chez les bivalves, en offrant une protection contre la dégradation à l'ADNmt paternel méthylique (Nishimura, 2010), on pourrait s'attendre à une surexpression des gènes liés à ce processus chez les mâles DUI par rapport aux hermaphrodites SMI et aux femelles DUI. Nos résultats ont permis de mettre en évidence DNMT1, une méthyl-transférase surexprimée chez les mâles, résultat d'autant plus surprenant que sa fonction est associée aux mitochondries (Maresca et al., 2015). Cependant, ce gène est également surexprimé chez nos hermaphrodites, ce qui semble contredire un possible rôle dans la protection de la dégradation, à moins que la fonction de ce facteur soit annulée par un autre facteur spécifique aux hermaphrodites, fonction qui pourrait être tenue par *exoG*.

Dans l'ensemble, la plupart des résultats issus du chapitre 2 ont permis de révéler plusieurs facteurs qui pourraient être impliqués spécifiquement dans la SMI, mais aucun n'a pu clairement être associé au phénomène de la DUI, à l'exception peut-être de la protéine F-box39. Cependant, l'activité de celle-ci nécessiterait l'action combinée d'un facteur additionnel qui reste à être identifié, c'est-à-dire un facteur qui « masquerait » le tag ubiquitine pour éviter la dégradation des mitochondries paternelles, comme proposé par Milani et collègues (2013).

Aspects mitochondriaux sous-jacents à la DUI

La caractérisation des génomes mitochondriaux F et M complets de deux nouvelles espèces à DUI, *Scrobicularia plana* et *Limecola balthica*, de l'ordre Venerida, un taxon beaucoup moins étudié chez les bivalves à DUI en général, a permis de mettre en lumière des éléments génétiques mitochondriaux largement partagés chez les espèces à DUI. Ces éléments sont décrits dans les sections suivantes.

Une insertion dans le gène cox2

Nos résultats du chapitre 3 ont révélé, chez les deux espèces *S. plana* et *L. balthica*, une importante différence de taille ($>10\text{kb}$) entre les génomes mitochondriaux mâle et femelle. Cette différence en nombre de nucléotides est due en majeure partie à une insertion dans le gène *cox2* des ADNmt M. Spécifiquement, l'insertion est de 4 818 pb pour *S. plana* et 3 960 pb pour *L. balthica*. Une étude en préparation (Tassé et al., non publié), ainsi que des données transcriptomiques (Pante et al., non publié), suggèrent respectivement que ces insertions sont conservées chez plusieurs individus et populations et qu'elles sont transcris, donc que la protéine qui en résulte est fonctionnelle. Chez *S. plana*, cette insertion fait du gène *Mcox2* le plus long gène mitochondrial *cox2* de tout le règne animal. Chez *L. balthica*, l'insertion présente une originalité supplémentaire avec un codon stop au milieu de la séquence du gène, qui coderait donc pour deux transcrits : *Mcox2a* et *Mcox2b* (Capt et al., soumis). On ne sait pas, par contre, si ces transcrits sont traduits en deux protéines séparées ou si, malgré le codon stop, la totalité de la séquence est traduite en une seule fois, comme par exemple grâce à des ARNt suppresseurs mutés qui incorporeraient un acide aminé à la place du codon stop (Capt et al., soumis). Des analyses transcriptomiques, protéomiques et immunohistochimiques supplémentaires seront nécessaires pour déterminer l'expression tissulaire et subcellulaire de *MCOX2*, mais également temporelle, liée aux stades de développement de l'individu, afin d'établir des hypothèses plus précises sur son rôle exact dans le système DUI. Des analyses protéomiques, de type western blot, sont également essentielles pour vérifier la traduction de l'insertion.

La situation observée chez les individus mâles *S. plana* et *L. balthica* n'est pas unique chez les bivalves à DUI. Une extension codante en 3' d'approximativement 555 pb (185 acides aminés) du gène *cox2* a été découverte dans l'ADNmt M chez les moules d'eau douce (Unionida) (Chapman et al., 2008). Cette extension représente une augmentation de 80% de la longueur du gène en comparaison à celui du génome F (Curole & Kocher, 2002; 2005). À l'aide de la microscopie électronique, il a été démontré que la protéine *MCOX2e* est localisée dans les membranes internes et externes des mitochondries contenue dans le sperme mature (Chakrabarti et al., 2007). Considérant que *MCOX2e* est en partie localisée dans la membrane mitochondriale externe (les protéines codées par l'ADNmt se retrouvent normalement que dans la membrane interne), il a été

proposé qu'elle pourrait jouer un rôle de ‘marqueur’ pour les mitochondries mâles DUI (Chakrabarti et al., 2007). Ce ‘marqueur’ permettrait à une machinerie cellulaire, qui reste à identifier, de localiser les mitochondries paternelles dans les œufs fécondés des espèces à DUI et ainsi les séquestrer dans la lignée germinale mâle (Cao, Kenchington, & Zouros, 2004; Chakrabarti et al., 2007; Cogswell, Kenchington, & Zouros, 2006). Outre chez les moules d'eau douce, une insertion a également été découverte dans le gène *Mcox2* chez la palourde marine *Meretrix lamarckii* (famille Veneridae, ordre Venerida) (Bettinazzi et al., 2016). L'extension est d'environ 300 pb (100 acides aminés) et est située au centre du gène (Bettinazzi et al., 2016), tout comme celle de *S. plana* et *L. balthica* qui sont du même ordre. Il est à noter que ces extensions ne sont pas seulement uniques aux génomes paternels des espèces à DUI mais elles sont aussi uniques chez les métazoaires.

MCOX2 chez *S. plana*, *L. balthica* et *M. Lamarckii* se situe dans l'espace intermembranaire de la mitochondrie, mais ne possède pas d'hélices transmembranaires dans la membrane mitochondriale externe (Tassé et al., en préparation). Ces particularités ne permettent donc pas de savoir si la fonction attribuée à MCOX2 est identique chez toutes les espèces et souligne l'importance d'établir les profils mitochondriaux d'espèces à DUI supplémentaires. Néanmoins, les résultats de nos transcriptomes (Capt et al., 2018, 2019) ont montré que *Mcox2* est largement surexprimé chez les mâles comparés aux individus hermaphrodites, suggérant que ce gène pourrait jouer un rôle important chez les mâles des espèces DUI.

Nouveaux ORFs

Parmi tous les ORFs potentiels analysés dans les génomes F et M des deux espèces, un seul ORF, présent dans le génome M de *L. balthica* (mtORFan1), présente des similitudes avec une protéine associée aux microtubules, fonction également attribuée aux M-ORF des espèces de moules d'eau douce et de *R. philippinarum* où ces gènes ont été particulièrement bien étudiés (Milani et al., 2013, Mitchell et al., 2016). Tel que proposé par Milani et al. (2013), la protéine M-ORF pourrait à la fois être impliquée dans le processus de séquestration des mitochondries paternelles lors du développement des embryons mâles (c'est-à-dire les regrouper et les rediriger ensuite vers la cellule primordiale qui donnera la lignée germinale), et être un facteur de masculinisation, dont les

différents niveaux d'accumulation du produit de ce gène dans les spermatozoïdes des espèces de bivalves à DUI, détermineraient le sexe de l'embryon.

Malgré une localisation subcellulaire de M-ORF très peu étudiée chez les espèces à DUI, des analyses *in silico* suggèrent qu'elle pourrait être ciblée à la surface des mitochondries des spermatozoïdes et être responsable de leur distribution et de leur positionnement dans les embryons en développement (Milani et al., 2013; Mitchell et al., 2016). Il a été suggéré que la dynamique mitochondriale, y compris la motilité, la fusion, la fission et l'autophagie, des processus qui nécessitent des interactions avec les microtubules, devrait être contrôlée, au moins en partie, par une « signalisation » provenant de la mitochondrie elle-même (voir Mitchell et al., 2016)

Nous proposons l'hypothèse que cet ORF (mtORFan1) identifié dans le génome M de *L. balthica*, qui possède des similitudes aux protéines associées aux microtubules, correspond au M-ORF décrit chez les autres espèces à DUI. Bien que F-*orf* et M-*orf* soient les gènes mitochondriaux qui évoluent le plus rapidement chez les espèces à DUI (Breton et al., 2011 ; Mitchell et al., 2016), avec des divergences importantes observées au niveau des séquences en acides aminés entre les espèces, les protéines M-ORFs et F-ORFs semblent être conservées au niveau de la structure, ce qui suggère une conservation inter-espèces de leurs fonctions également. Par exemple, excepté pour une seule espèce de moule d'eau douce (plus de 50 espèces ont été étudiées : Milani et al., 2013; Mitchell et al., 2016; Guerra et al., 2017), les protéines M-ORFs ont toutes une hélice transmembranaire prédite dans leur portion N- terminale, ce qui est également le cas pour la séquence en acides aminés de mtORFan1 chez *L. balthica* (voir Figure S3.2 du Chapitre 3). De plus, les analyses *in silico* ont associé la fonction de M-ORF non seulement à la dynamique du cytosquelette, mais aussi à l'interaction avec les acides nucléiques, à la signalisation, à la différenciation cellulaire et développementale, à l'ubiquitination, à l'apoptose et à la réponse immunitaire (Milani et al., 2013; Mitchell et al., 2016). Une analyse rapide avec le serveur I-TASSER (server for protein structure and function prediction ; Zhang, 2008), prédit des fonctions similaires (liaison aux acides nucléiques et apoptose) pour la protéine (possiblement) codée par mtORFan1 dans l'ADNmt M chez *L. balthica* (données non montrées), supportant notre hypothèse.

À contrario, il a été suggéré que la protéine F-ORF permettrait de féminiser les embryons destinés à devenir des femelles en inhibant la formation des tissus testiculaires (Mitchell et al., 2016). Nous n'avons pas identifié d'ORF chez *S. plana* et *L. balthica* qui pourrait correspondre au gène F-*orf* identifié chez les autres espèces à DUI. Cependant, nos analyses ont considéré les ORFs >150bp (50aa) et il est possible qu'un cadre de lecture plus petit puisse correspondre à F-*orf*. Le plus petit F-ORF identifié à ce jour chez une espèce DUI, c'est-à-dire chez la moule d'eau douce *Venustaconcha ellipsiformis*, code pour une protéine d'une longueur de 89aa (Breton et al., 2011; Mitchell et al., 2016), tandis que le plus petit M-ORF code potentiellement pour une protéine de 30 aa (chez la moule marine *Mytilus californianus*; Milani et al., 2013). Des séquences ADNmt F et des analyses d'expression supplémentaires provenant de *S. plana* et de *L. balthica* seront nécessaires pour démontrer clairement la présence (ou l'absence) du gène F-*orf* chez ces espèces. Aussi, des études biochimiques approfondies chez *S. plana* et *L. balthica*, mais aussi chez les autres espèces à DUI, seront nécessaires pour élucider les fonctions de M-ORF et F-ORF.

Nos données transcriptomiques des chapitres 1 et 2 nous renseignent également un peu sur ces nouveaux ORFs mitochondriaux chez les bivalves. Par exemple, nos analyses ont montré que M-*orf* ainsi que H-*orf* sont exprimés uniquement dans leur gonade respective, tandis que F-*orf* semble agir dans les gonades mâles et femelles (Capt et al., 2018; 2019). Notamment, le gène H-*orf* était un des gènes les plus différentiellement exprimé chez les hermaphrodites (Capt et al., 2019). Pour rappel, la protéine H-ORF provient du génome mitochondrial des espèces de moules d'eau douce hermaphrodites à SMI (Breton et al., 2011 ; Mitchell et al., 2016). F-ORF et M-ORF sont soupçonnées d'être responsables du maintien de la séparation des sexes, tandis que H-ORF (une version hautement modifiée de F-ORF) aurait perdu sa fonction inhibitrice de formation des testicules, permettant ainsi l'hermaphrodisme simultané (Mitchell et al., 2016). Évidemment, d'autres analyses transcriptomiques et biochimiques, notamment sur les embryons mâles, femelles et hermaphrodites, permettront de mieux comprendre les fonctions de ces nouveaux gènes.

Nos résultats amènent beaucoup de nouvelles questions, notamment s'il y a des parallèles à faire entre la DUI chez les bivalves et le seul autre système de détermination du sexe connu à ce jour qui implique les mitochondries, soit la stérilité cytoplasmique mâle ou CMS pour *cytoplasmic male sterility* chez les plantes. La prochaine section retrace brièvement ce qu'on connaît de ce

système CMS et les similitudes que l'on peut faire avec le système de la DUI et les résultats de cette thèse.

CMS – Stérilité cytoplasmique mâle

La CMS a été découverte chez plus de 200 espèces de plantes, initialement hermaphrodites, et bien que le fonctionnement diffère entre les espèces, un aspect commun est l'induction de la stérilité mâle par un gène mitochondrial, tandis qu'un gène d'origine nucléaire (communément appelé *Rf* pour «*fertility restorer* ») est responsable de rétablir la fertilité (Budar, Touzet, & De Paepe, 2003; Chase, 2007; Hu et al., 2014 pour des revues de littérature). Spécifiquement, les gènes mitochondriaux de la CMS empêchent la production de pollen fonctionnel par différents mécanismes (on se retrouve donc avec des plants femelles par défaut) alors que les gènes *Rf*, quant à eux, rétablissent la fertilité mâle en agissant sur les gènes de la CMS et ses produits, de façon directe ou indirecte (et rétablit donc l'hermaphrodisme). Au niveau mitochondrial, la CMS coïncide avec la présence de nouveaux ORFs, une caractéristique largement exposée dans les génomes mitochondriaux de nos espèces DUI (voir section précédente). Chez les plantes, ces nouveaux ORFs proviendraient d'un réarrangement du génome mitochondrial (Hanson & Bentolila, 2004; Tuteja et al., 2013), une autre caractéristique retrouvée chez les espèces à DUI. En effet, deux études ont appuyé l'hypothèse que l'origine des gènes M-*orf* et F-*orf*, chez les moules d'eau douce, pourrait respectivement provenir d'une duplication suivit d'une néo-fonctionnalisation (gain d'une nouvelle fonction des copies dupliquées) des gènes *atp8* et *nad2* (Mitchell et al., 2016; Guerra et al., 2017).

Certains ORFs mitochondriaux responsables de la CMS chez les plantes, sont des gènes chimériques, qui contiennent des bouts de séquences d'un gène typique, comme *cox2* ou *atp9*, associés à une séquence de provenance inconnue (Hanson & Bentolila, 2004). Certains de ces gènes chimériques ont été associés à des processus apoptotiques qui induiraient une mort précoce des cellules à l'origine du développement des anthères (organe reproducteur mâle des plantes) (Balk & Leaver, 2001; Ji et al., 2013; Wan et al., 2007). Les protéines encodées par les gènes de la CMS peuvent aussi entraîner une baisse de production d'ATP au niveau de la mitochondrie, soit en altérant la fonction de la chaîne de transport d'électrons / ATP synthase, ou encore par la formation

d'un pore dans la membrane interne de la mitochondrie (Chen & Liu, 2014). Plusieurs protéines CMS ont donc été identifiées avec des domaines transmembranaires (Li et al., 2018), une caractéristique également partagée par toutes les protéines F-ORF et M-ORF des mitogénomes d'espèces à DUI (Mitchell et al., 2016). Il est également tentant de penser aux modifications importantes de la protéine COX2, codée par l'ADNmt M des espèces à DUI, qui pourrait altérer la fonction mitochondriale. Cependant, il existe plusieurs formes de CMS, et donc de *Rf* associés. Les mécanismes vont donc diverger selon le type de CMS qui est apparu, une hypothèse que l'on pourrait également proposer pour expliquer les différences recensées entre les taxons à DUI.

Liste des facteurs Rf

Les *Rf* d'origine nucléaires, quant-à-eux, font majoritairement partie des protéines pentatricopeptidiques (PPR) qui permettent des liaisons protéines-protéines ou protéines-ARN (O'Toole et al., 2008). On retrouve de nombreuses copies des gènes codant ces protéines dans les génomes végétaux. Par exemple, *Arabidopsis thaliana* possède environ 450 protéines PPR dont la majorité ont été prédictes comme étant localisées au niveau des mitochondries ou des plastes (Carlsson et al., 2008; Chase, 2007). Les gènes *Rf* peuvent agir de nombreuses façons afin de rétablir la fertilité, que ce soit via le génome, la transcription, la traduction, ainsi qu'au niveau du métabolisme. Par exemple, certains gènes *Rf* vont cliver et dégrader des transcrits de gènes de la CMS, alors que d'autres vont s'associer aux ARNm CMS afin de bloquer leur traduction (Hu et al., 2004).

Un fait intéressant, des annotations de protéines pentatricopeptidiques (PPR) ont été trouvées dans nos analyses transcriptomiques des chapitres 1 et 2. En particulier, nous avons identifié un contig PPR uniquement associé aux gonades mâles et annoté comme « leucine rich PPR motif containing protein mitochondrial » (Tableau 4, unigène c220030_g1_i3 du chapitre 2), une protéine jouant un rôle dans le métabolisme des ARN chez l'humain. Chez les plantes, le rôle des protéines *Rf* est de restaurer la fertilité mâle et donc de restaurer l'hermaphrodisme au niveau du plant. Il est possible que ce type de protéines pourraient avoir un rôle à jouer dans la fertilité mâle et le maintien du gonochorisme chez les espèces à DUI, mais ceci demeure à ce stade, purement spéculatif.

Un scénario évolutif qui expliquerait l'apparition de la DUI et du gonochorisme chez les bivalves avait été suggéré par Milani et collègues (2016), en soulignant l'aspect « égoïste » des gènes mitochondriaux. L'ajout des résultats de mon projet de doctorat, et les similitudes avec le système CMS, viennent renforcer la plausibilité de cette hypothèse, ainsi qu'une origine multiple de la DUI, qui aurait débuté avec l'apparition d'une nouvelle séquence dans le génome mitochondrial d'une population de bivalves hermaphrodites. Cette insertion, d'origine virale (Milani et al. 2014; 2016) ou mitochondriale (Mitchell et al. 2016), lui aurait conféré une capacité à se protéger de toute dégradation, ainsi que de se transmettre dans les générations suivantes, autant bien dans les lignées germinales femelles que mâles. Puis, une production accrue des gamètes mâles et de leurs divisions cellulaires par rapports aux gamètes femelles, aurait permis une conquête plus rapide et plus efficace dans la lignée mâle (Katzouraki et al., 2007 ; Milani et al., 2016). Le nombre de spermatozoïdes portant ces mitochondries infectées aurait augmenté, grâce à leur nouvelle capacité à s'orienter vers le blastomère responsable de produire la lignée germinale mâle, un phénomène déjà recensé chez des espèces à DUI (Milani et al., 2012 ; Zouros, 2013). Ces nouvelles mitochondries seraient donc soumises à une pression de sélection, qui favoriserait ainsi l'expansion de celles qui portent des mutations augmentant la valeur sélective des mâles. Le jumelage avec une mutation qui rendrait les fonctions femelles stériles, provoquerait une transition d'hermaphrodisme à androadioécie, la production de mâles étant favorisée sur celle des hermaphrodites puisqu'elle permet de maximiser la diversité génétique en évitant l'auto-fertilisation, et que la production de gamètes mâles est moins coûteuse énergétiquement (Charlesworth & Charlesworth, 1978; Weeks, 2012). Une deuxième transition aurait alors pu se produire d'androadioécie à gonochorisme, une fois que la production de spermatozoïdes serait devenue trop importante, la pression sélective favorisant alors la production d'œufs chez les hermaphrodites en sélectionnant les mutations rendant les organes mâles stériles. Une telle distorsion du sex-ratio devrait générer l'évolution des gènes nucléaires qui contrecarrent leur effet, afin de rétablir un équilibre dans le sex-ratio (Hurt, 1992). Ce genre de gènes compensatoires peuvent créer des générations exclusivement mâles ou femelles (Werren & Beukeboom, 1998), un phénomène retrouvé chez certaines espèces à DUI, où le biais du sex-ratio chez une portée dépend uniquement de la femelle, peu importe le mâle avec lequel elle se reproduit (Kenchington et al., 2002; Ghiselli et al., 2012; Machordom et al., 2015). Chez

Ruditapes philippinarum, grâce à des études transcriptionnelles, traductionnelles et fonctionnelles, il a été proposé que le nouveau cadre de lecture ouvert propre au génome mitochondrial mâle, Rphm21, pourrait être le gène qui permette aux mitochondries paternelles d'éviter leur dégradation, tandis que Rphf22, le cadre de lecture ouvert des mitochondries maternelles, jouerait un rôle dans le rétablissement des sex-ratio (Milani et al., 2016). Cependant, il est généralement convenu qu'un tel rôle est d'ordre nucléaire, ce qui laisse supposer que Rphf22 n'est pas le seul à prendre part à un tel mécanisme (Milani et al., 2016).

L'ajout constant d'études portant sur un des aspects de la DUI permet d'affiner les hypothèses quant à son scénario évolutif, et démontrent ainsi l'importance d'analyses comparatives.

Conclusion et perspectives

Le sujet de cette thèse regroupe trois grandes aires de recherche : la transmission mitochondriale, la détermination du sexe et les conflits génomiques. Si des systèmes de déterminisme sexuel tels que la CMS comportent des mécanismes aussi complexes et diversifiés avec une transmission mitochondriale strictement maternelle, il est légitime de penser que la situation pourrait se complexifier avec un système qui comprend une double transmission uniparentale de l'ADNmt. Cette thèse a permis de mettre en avant plusieurs processus et caractéristiques qui pourraient être impliqués dans cette mystérieuse mécanique liant la DUI à la détermination du sexe chez les bivalves, d'autant plus que tous les chapitres comportaient une partie comparative afin de tenter de regrouper les éléments clés partagés par tous les taxons à DUI. Parmi ces éléments clés, il semble évident que les processus d'ubiquitination-mitophagie, ainsi que la méthylation, doivent être étudiés plus en profondeur, notamment au niveau des embryons lors de leurs tout premiers stades de développement, lorsque les mitochondries paternelles sont éliminées chez les embryons femelles et conservés chez les mâles. Il serait également pertinent d'étudier les protéines mitochondrielles MCOX2, F-ORF et M-ORF. Par exemple, définir leur distribution tissulaire par immunobuvardage durant différents stades du développement, leur localisation tissulaire et subcellulaire par immunohistochimie, et leurs partenaires d'interaction par immunoprecipitation, autant chez les mâles que chez les femelles, permettrait éventuellement d'élucider leurs fonctions. D'autres expériences pertinentes en ce sens, pourraient être de les exprimer chez la levure, ou encore chez *E. coli*, afin de vérifier si elles ont des effets particuliers, des expériences qui ont permis notamment de démontrer les effets cytotoxiques des gènes CMS (Hu et al., 2014). Aussi, l'inhibition de l'expression de ces gènes serait une façon de mieux cerner si leur rôle est de stimuler le développement d'un sexe, et/ou d'inhiber celui du sexe opposé. Évidemment, il faudra continuer à investiguer la présence de la DUI, ainsi que de caractériser de nouveaux mitogénomes et transcriptomes chez des espèces gonochoriques à DUI, mais aussi chez des espèces proches-parentes hermaphrodites (ou gonochoriques) à SMI, afin de renforcer le tableau dressé par les présentes études, et les précédentes.

Références

- Afgan, E., Baker, D., Batut, B., Van Den Beek, M., Bouvier, D., Ech, M., ...
- Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, 46(W1), W537–W544.
<https://doi.org/10.1093/nar/gky379>
- Ahmed, M., and Sparks, A.K. (1967). A Preliminary Study of Chromosomes of Two Species of Oysters (*Ostrea lurida* and *Crassostrea gigas*). *J. Fish. Res. Bd. Can.* 24, 2155–2159.
- Allen, D. C., Sietman, B. E., Kelner, D. E., Hove M.C., Kurth, J. E., Davis, J. M., Weiss, J. L., Hornbach, D. J. (2007). Early Life-history and Conservation Status of *Venustaconcha Ellipsiformis* (Bivalvia, Unionidae) in Minnesota. *The American Midland Naturalist*, 157(1), 74–91. [https://doi.org/10.1674/0003-0031\(2007\)157\[74:elacso\]2.0.co;2](https://doi.org/10.1674/0003-0031(2007)157[74:elacso]2.0.co;2)
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool.pdf. *Journal of Molecular Biology*. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Altschul, S. F., Madden, T. L., Schaffer, A. a, Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Blast and Psi-Blast: Protein Database Search Programs. *Nucleic Acid Research*, 25(17), 2289–4402.
- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology* 11, R106.
- Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., et al. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290, 457–465.
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinform*. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Andrews, S. (2015). FASTQC A Quality Control tool for High Throughput Sequence Data. Babraham Institute.

Arama, E., Bader, M., Rieckhof, G. E., & Steller, H. (2007). A Ubiquitin Ligase Complex Regulates Caspase Activation During Sperm Differentiation in Drosophila. *PLOS Biology*, 5(10), e251. <https://doi.org/10.1371/journal.pbio.0050251>

Ashman, T.-L., Bachtrog, D., Blackmon, H., Goldberg, E.E., Hahn, M.W., Kirkpatrick, M., Kitano, J., Mank, J.E., Mayrose, I., Ming, R., et al. (2014). Tree of Sex: A database of sexual systems. *Scientific Data* 1.

Audic, S., & Claverie, J.-M. (1997). The significance of digital gene expression profiles. *Genome Research*, 7(10), 986–995.

Bachtrog, D., Mank, J.E., Peichel, C.L., Kirkpatrick, M., Otto, S.P., Ashman, T.-L., Hahn, M.W., Kitano, J., Mayrose, I., Ming, R., et al. (2014). Sex Determination: Why So Many Ways of Doing It? *PLoS Biol* 12, e1001899.

Balk, J., & Leaver, C. J. (2001). The PET1-CMS Mitochondrial Mutation in Sunflower Is Associated with Premature Programmed Cell Death and Cytochrome c Release. *The Plant Cell*, 13(8), 1803–1818. <https://doi.org/10.1105/TPC.010116>

Bar-Yaacov, D., Blumberg, A., and Mishmar, D. (2012). Mitochondrial-nuclear co-evolution and its effects on OXPHOS activity and regulation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1819, 1107–1111.

Barnhart, M. C., Haag, W. R., & Roston, W. N. (2008). Adaptations to host infection and larval parasitism in Unionoida. *Am. Benthol. Soc*, 27(2), 370–394. <https://doi.org/10.1899/07-093.1>

Batista, F.M., Lallias, D., Taris, N., Guedes-Pinto, H., and Beaumont, A.R. (2011). Relative quantification of the M and F mitochondrial DNA types in the blue mussel *Mytilus edulis* by real-time PCR. *J. Mollus. Stud.* 77, 24–29.

Bayrer, J.R., Zhang, W., and Weiss, M.A. (2005). Dimerization of Doublesex Is Mediated

by a Cryptic Ubiquitin-associated Domain Fold IMPLICATIONS FOR SEX-SPECIFIC GENE REGULATION. *J. Biol. Chem.* 280, 32989–32996.

Beckenbach, A. T., Robson, S. K. A., & Crozier, R. H. (2005). Single nucleotide +1 frameshifts in an apparently functional mitochondrial cytochrome b gene in ants of the genus Polyrhachis. *Journal of Molecular Evolution*, 60(2), 141–152. <https://doi.org/10.1007/s00239-004-0178-5>

Beier, H., & Grimm, M. (2001). Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. *Nucleic Acids Research*, 29(23), 4767–4782. <https://doi.org/10.1093/nar/29.23.4767>

Bell, G. (1982). *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* (CUP Archive).

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., ... Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*. <https://doi.org/10.1016/j.ympev.2012.08.023>

Bettinazzi, S., Plazzi, F., & Passamonti, M. (2016). The complete female- and male-transmitted mitochondrial genome of Meretrix lamarckii. *PLoS ONE*, 11(4). <https://doi.org/10.1371/journal.pone.0153631>

Bettinazzi, S., Rodríguez, E., Milani, L., Blier, P. U., & Breton, S. (2019). Metabolic remodelling associated with mtDNA: Insights into the adaptive value of doubly uniparental inheritance of mitochondria. *Proceedings of the Royal Society B: Biological Sciences*, 286(1896). <https://doi.org/10.1098/rspb.2018.2708>

Beukeboom, L.W., and Perrin, N. (2014). *The Evolution of Sex Determination* (Oxford University Press).

Beukeboom, L.W., and Zande, L.V.D. (2010). Genetics of sex determination in the haplodiploid wasp *Nasonia vitripennis* (Hymenoptera: Chalcidoidea). *J Genet* 89, 333–339.

Beukeboom, L.W., Ellers, J., and Van Alphen, J.J.M. (2000). Absence of single-locus complementary sex determination in the braconid wasps *Asobara tabida* and *Alysia manducator* - Beukeboom - 2001 - Heredity - Wiley Online Library.

Beye, M., Hasselmann, M., Fondrk, M.K., Page, R.E., and Omholt, S.W. (2003). The gene csd is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* 114, 419–429.

Birky, C. W. (2001). The Inheritance of Genes in Mitochondria and Chloroplasts: Laws, Mechanisms, and Models. *Annual Review of Genetics*, 35(1), 125–148.
<https://doi.org/10.1146/annurev.genet.35.102401.090231>

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120.
<https://doi.org/10.1093/bioinformatics/btu170>

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8), 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>

Boore, J. L. (2000). The Duplication/Random Loss Model for Gene Rearrangement Exemplified by Mitochondrial Genomes of Deuterostome Animals (pp. 133–147).
https://doi.org/10.1007/978-94-011-4309-7_13

Borsig, P., and Thiriot-Quiévreux, C. (1990). Karyological and allozymic characterization of *Ruditapes philippinarum*, *R. aureus* and *R. decussatus* (Bivalvia, Veneridae). *Aquaculture* 90, 209–227.

Breton, S. (2017). Sex Determining Mechanisms in Bivalves, (July), 0–23.
<https://doi.org/10.20944/preprints201706.0127.v1>

Breton, S., & Stewart, D. T. (2015). Atypical mitochondrial inheritance patterns in eukaryotes. *Genome*, 58(10), 423–431. <https://doi.org/10.1139/gen-2015-0090>

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. <https://doi.org/10.1016/j.tig.2007.05.011>

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. <https://doi.org/10.1534/genetics.109.110700>

Breton, S., Beaupré, H.D., Stewart, D.T., Hoeh, W.R., and Blier, P.U. (2007). The unusual system of doubly uniparental inheritance of mtDNA: isn't one enough? *Trends Genet.* 23, 465–474.

Breton, S., Beaupré, H.D., Stewart, D.T., Piontkivska, H., Karmakar, M., Bogan, A.E., Blier, P.U., and Hoeh, W.R. (2009a). 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, 1575–1589.

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). <https://doi.org/10.1371/journal.pone.0183529>

Breton, S., Burger, G., Stewart, D. T., & Blier, P. U. (2006). Comparative analysis of gender-associated complete mitochondrial genomes in marine mussels (*Mytilus* spp.). *Genetics*, 172(2), 1107–1119. <https://doi.org/10.1534/genetics.105.047159>

Breton, S., Capt, C., Guerra, D., & Stewart, D. (2018). Sex-Determining Mechanisms in Bivalves. In *Transitions Between Sexual Systems* (pp. 165–192). Springer International Publishing. https://doi.org/10.1007/978-3-319-94139-4_6

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. <https://doi.org/10.1371/journal.pone.0019365>

Breton, S., Stewart, D. T., & Blier, P. U. (2009). Role-reversal of gender-associated mitochondrial DNA affects mitochondrial function in *Mytilus edulis* (Bivalvia: Mytilidae). *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 312B(2), 108–117. <https://doi.org/10.1002/jez.b.20251>

Breton, S., Stewart, D. T., Shepardson, S., Trdan, R. J., Bogan, A. E., Chapman, E. G., ... Hoeh, W. R. (2011b). Novel protein genes in animal mtDNA: A new sex determination system in freshwater mussels (Bivalvia: Unionoida)? *Molecular Biology and Evolution*, 28(5), 1645–1659. <https://doi.org/10.1093/molbev/msq345>

Breton, S., Stewart, D.T., and Blier, P.U. (2009c). Role-reversal of gender-associated mitochondrial DNA affects mitochondrial function in *Mytilus edulis* (Bivalvia: Mytilidae). *J. Exp. Zool.* 312B, 108–117.

Breton, S., Stewart, D.T., Shepardson, S., Trdan, R.J., Bogan, A.E., Chapman, E.G., Ruminas, A.J., Piontkivska, H., and Hoeh, W.R. (2011). Novel Protein Genes in Animal mtDNA: A New Sex Determination System in Freshwater Mussels (Bivalvia: Unionoida)? *Mol Biol Evol* 28, 1645–1659.

Brookes, P.S., Levonen, A.-L., Shiva, S., Sarti, P., and Darley-Usmar, V.M. (2002). Mitochondria: regulators of signal transduction by reactive oxygen and nitrogen species 1, 2. *Free Radical Biology and Medicine* 33, 755–764.

Brown, N. P., Leroy, C., & Sander, C. (1998). MView: A web-compatible database search or multiple alignment viewer. *Bioinformatics*, 14(4), 380–381. <https://doi.org/10.1093/bioinformatics/14.4.380>

Brush, S.G. (1978). Nettie M. Stevens and the Discovery of Sex Determination by Chromosomes. *Isis* 69, 163–172.

Budar, F., Touzet, P., & De Paepe, R. (2003). The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. *Genetica*, 117(1), 3–16. <https://doi.org/10.1023/A:1022381016145>

Bull, J. J. (1983). Evolution of sex determining mechanisms. 316 pp.

Burnell, G.M. (1995) Age-related protandry in the scallop *Chlamys varia* (L.) on the west coast of Ireland. ICES. *Mar. Sci Symp.* 199: 26–30

Burt, A. (2000). Perspective: Sex, Recombination, and the Efficacy of Selection—Was Weismann Right? *Evolution* 54, 337–351.

Bustin, S.A., Beaulieu, J.-F., Huggett, J., Jaggi, R., Kibenge, F.S., Olsvik, P.A., Penning, L.C., and Toegel, S. (2010). MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Molecular Biology* 11, 74.

Cahais, V., Gayral, P., Tsagkogeorga, G., Melo-Ferreira, J., Ballenghien, M., Weinert, L., Chiari, Y., Belkhir, K., Ranwez, V., and Galtier, N. (2012). Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. *Molecular Ecology Resources* 12, 834–845.

Campos, A., Tedesco, S., Vasconcelos, V., & Cristobal, S. (2012). Proteomic research in bivalves: towards the identification of molecular markers of aquatic pollution. *Journal of Proteomics*, 75(14), 4346–4359. <https://doi.org/10.1016/j.jprot.2012.04.027>

Cao, L., Kenchington, E., & Zouros, E. (2004). Differential Segregation Patterns of Sperm Mitochondria in Embryos of the Blue Mussel (*Mytilus edulis*). *Genetics*, 166(2), 883–894. <https://doi.org/10.1534/genetics.166.2.883>

Capt, C., Passamonti, M., & Breton, S. (2016). The human mitochondrial genome may code for more than 13 proteins. *Mitochondrial DNA*, 27(5), 1–4. <https://doi.org/10.3109/19401736.2014.1003924>

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. <https://doi.org/10.1093/gbe/evy019>

Capt, C., Renaut, S., Stewart, D. T., Johnson, N. A., & Breton, S. (2019). Putative mitochondrial sex determination in the Bivalvia: insights from a hybrid transcriptome assembly in freshwater mussels. *Frontiers in Genetics*, 10, 840.
<https://doi.org/10.3389/FGENE.2019.00840>

Carlsson, J., Leino, M., Sohlberg, J., Sundström, J. F., & Glimelius, K. (2008). Mitochondrial regulation of flower development. *Mitochondrion*, 8(1), 74–86.
<https://doi.org/10.1016/j.mito.2007.09.006>

Chabi, B., Ljubicic, V., Menzies, K.J., Huang, J.H., Saleem, A., and Hood, D.A. (2008). Mitochondrial function and apoptotic susceptibility in aging skeletal muscle. *Aging Cell* 7, 2–12.

Chakrabarti, R., Walker, J. M., Chapman, E. G., Shepardson, S. P., Trdan, R. J., Curole, J. P., Watters G.T., Stewart, D.T., Vijayaraghavan, S., and Hoeh, W. R. (2007). Reproductive function for a C-terminus extended, male-transmitted cytochrome c oxidase subunit II protein expressed in both spermatozoa and eggs. *FEBS Letters*, 581(27), 5213–5219.
<https://doi.org/10.1016/j.febslet.2007.10.006>

Chakrabarti, R., Walker, J. M., Stewart, D. T., Trdan, R. J., Vijayaraghavan, S., Curole, J. P., & Hoeh, W. R. (2006). Presence of a unique male-specific extension of C-terminus to the cytochrome c oxidase subunit II protein coded by the male-transmitted mitochondrial genome of *Venustaconcha ellipsiformis* (Bivalvia: Unionoidea). *FEBS Letters*, 580(3), 862–866.
<https://doi.org/10.1016/j.febslet.2005.12.104>

Chan, D. C., & Schon, E. A. (2012). Eliminating Mitochondrial DNA from Sperm. *Developmental Cell*. <https://doi.org/10.1016/j.devcel.2012.02.008>

Chapman, E. G., Piontkivska, H., Walker, J. M., Stewart, D. T., Curole, J. P., & Hoeh, W. R. (2008). Extreme primary and secondary protein structure variability in the chimeric male-transmitted cytochrome c oxidase subunit II protein in freshwater mussels: evidence for an elevated amino acid substitution rate in the face of domain-specific purifying selection. *BMC Evolutionary Biology*, 8, 165. <https://doi.org/10.1186/1471-2148-8-165>

Charnier, M. (1966). [Action of temperature on the sex ratio in the *Agama agama*

(Agamidae, Lacertilia) embryo]. C. R. Seances Soc. Biol. Fil. 160, 620–622.

Charnov, E.L. (1982). The theory of sex allocation. Princeton University Press, Princeton.

Chase, C. D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends in Genetics, 23(2), 81–90.
<https://doi.org/10.1016/j.tig.2006.12.004>

Chase, E. E., Robicheau, B. M., Veinot, S., Breton, S., & Stewart, D. T. (2018). The complete mitochondrial genome of the hermaphroditic freshwater mussel *Anodonta cygnea* (Bivalvia: Unionidae): In silico analyses of sex-specific ORFs across order Unionoida. BMC Genomics, 19(1). <https://doi.org/10.1186/s12864-018-4583-3>

Chávez-Villalba, J., Soyez, C., Huvet, A., Gueguen, Y., Lo, C., & Moullac, G. Le. (2011). Determination of Gender in the Pearl Oyster *Pinctada margaritifera*. Journal of Shellfish Research, 30(2), 231–240. <https://doi.org/10.2983/035.030.0206>

Chen, L., & Liu, Y.-G. (2014). Male Sterility and Fertility Restoration in Crops. Annual Review of Plant Biology, 65(1), 579–606. <https://doi.org/10.1146/annurev-arplant-050213-040119>

Ciocan, C.M., Cubero-Leon, E., Peck, M.R., Langston, W.J., Pope, N., Minier, C., Rotchell, J.M. (2012) Intersex in *Scrobicularia plana*: transcriptomic analysis reveals novel genes involved in endocrinedisruption. *Environ Sci Technol* 46:12936–12942

Clayton, D.A. (1991). Replication and Transcription of Vertebrate Mitochondrial DNA. Annual Review of Cell Biology, 7, 453–478.

Cline, T.W., & Meyer, B.J. (1996). Vive la différence: Males vs Females in Flies vs Worms. Annual Review of Genetics, 30, 637–702.

Coe, W. R. (1943). Sexual Differentiation in Mollusks. I. Pelecypods. The Quarterly Review of Biology, 18(2), 154–164. <https://doi.org/10.1086/394673>

Collin R (2013) Phylogenetic patterns and phenotypic plasticity of molluscan sexual

systems. *IntegrComp Biol* 53:723–735

Cogswell, A. T., Kenchington, E. L. ., & Zouros, E. (2006). Segregation of sperm mitochondria in two- and four-cell embryos of the blue mussel *Mytilus edulis*: implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA. *Genome*, 49(7), 799–807. <https://doi.org/10.1139/g06-036>

Cogswell, A.T., Kenchington, E.L., and Zouros, E. (2006). Segregation of sperm mitochondria in two- and four-cell embryos of the blue mussel *Mytilus edulis*: implications for the mechanism of doubly uniparental inheritance of mitochondrial DNA. *Genome* 49, 799–807.

Collin, R. (1995). Sex, Size, and Position: A Test of Models Predicting Size at Sex Change in the Protandrous Gastropod *Crepidula fornicata*. *The American Naturalist* 146, 815–831.

Collin, R., Chaparro, O.R., Winkler, F., and Véliz, D. (2007). Molecular Phylogenetic and Embryological Evidence That Feeding Larvae Have Been Reacquired in a Marine Gastropod. *Biol Bull* 212, 83–92.

Conant, G. C., & Wolfe, K. H. (2008). GenomeVx: simple web-based creation of editable circular chromosome maps. *Bioinformatics* (Oxford, England), 24(6), 861–862. <https://doi.org/10.1093/bioinformatics/btm598>

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., Mcpherson, A., ... Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. <https://doi.org/10.1186/s13059-016-0881-8>

Cummins, S. F., Nichols, A. E., Amare, A., Hummon, A. B., Sweedler, J. V., & Nagle, G. T. (2004). Characterization of *Aplysia* enticin and temptin, two novel water-borne protein pheromones that act in concert with attractin to stimulate mate attraction. *Journal of Biological Chemistry*, 279(24), 25614–25622. <https://doi.org/10.1074/jbc.M313585200>

Curole, J. P., & Kocher, T. D. (2002). Ancient sex-specific extension of the cytochrome c oxidase II gene in bivalves and the fidelity of doubly-uniparental inheritance. *Molecular Biology and Evolution*, 19(8), 1323–1328. <https://doi.org/10.1093/oxfordjournals.molbev.a004193>

Curole, J. P., & Kocher, T. D. (2005). Evolution of a unique mitotype-specific protein-coding extension of the cytochrome c oxidase II gene in freshwater mussels (Bivalvia: Unionoida). *Journal of Molecular Evolution*, 61(3), 381–389. <https://doi.org/10.1007/s00239-004-0192-7>

Curole, J.P. (2004). Universal primers for the specific amplification of the male mitotype of the Unionoidea (Bivalvia). *Conservation Genetics* 5, 733–735.

Cymerman, I. A., Chung, I., Beckmann, B. M., Bujnicki, J. M., & Meiss, G. (2008). EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Research*, 36(4), 1369–1379. <https://doi.org/10.1093/nar/gkm1169>

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. <https://doi.org/10.1139/g01-159>

De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, 34(15), 2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>

De Wit, P., Pespeni, M.H., and Palumbi, S.R. (2015). SNP genotyping and population genomics from expressed sequences –current advances and future possibilities. *Mol Ecol*, 24(10), 2310-2323. <https://doi.org/10.1111/mec.13165>

De Wit, P., Pespeni, M.H., Ladner, J.T., Barshis, D.J., Seneca, F., Jaris, H., Therkildsen, N.O., Morikawa, M., and Palumbi, S.R. (2012). The simple fool's guide to population genomics via RNA-Seq: an introduction to high-throughput sequencing data analysis. *Mol Ecol Resour.*, 12, 1058–1067.

DeLuca, S. Z., & O'Farrell, P. H. (2012). Barriers to Male Transmission of Mitochondrial DNA in Sperm Development. *Developmental Cell*, 22(3), 660–668. <https://doi.org/10.1016/j.devcel.2011.12.021>

Desvages, G., Girondot, M., and Pieau, C. (1993). Sensitive Stages for the Effects of Temperature on Gonadal Aromatase Activity in Embryos of the Marine Turtle Dermochelys

coriacea. General and Comparative Endocrinology 92, 54–61.

Diz, A. P., Dudley, E., Cogswell, A., MacDonald, B. W., Kenchington, E. L. R. R., Zouros, E., & Skibinski, D. O. F. F. (2013). Proteomic Analysis of Eggs from *Mytilus edulis* Females Differing in Mitochondrial DNA Transmission Mode. *Molecular & Cellular Proteomics*, 12(11), 3068–3080. <https://doi.org/10.1074/mcp.M113.031401>

Doniach, T., & Hodgkin, J. (1984). A sex-determining gene, *fem-1*, required for both male and hermaphrodite development in *Caenorhabditis elegans*. *Developmental Biology*, 106(1), 223–235. [https://doi.org/10.1016/0012-1606\(84\)90077-0](https://doi.org/10.1016/0012-1606(84)90077-0)

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), 1–19. <https://doi.org/10.1186/1471-2148-10-50>

Dournon, C., Houillon, C., and Pieau, C. (1990). Temperature sex-reversal in amphibians and reptiles. *Int. J. Dev. Biol.* 34, 81–92.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>

Eppley, S.M., and Jesson, L.K. (2008). Moving to mate: the evolution of separate and combined sexes in multicellular organisms. *Journal of Evolutionary Biology* 21, 727–736.

Ezaz, T., Stiglec, R., Veyrunes, F., and Marshall Graves, J.A. (2006). Relationships between Vertebrate ZW and XY Sex Chromosome Systems. *Current Biology* 16, R736–R743.

Fabioux, C., Huvet, A., Le Souchu, P., Le Pennec, M., and Pouvreau, S. (2005). Temperature and photoperiod drive *Crassostrea gigas* reproductive internal clock. *Aquaculture* 250, 458–470.

Fabioux, C., Huvet, A., Lelong, C., Robert, R., Pouvreau, S., Daniel, J. Y., ... Le Pennec, M. (2004). Oyster vasa-like gene as a marker of the germline cell development in *Crassostrea gigas*. *Biochemical and Biophysical Research Communications*, 320(2), 592–598.

<https://doi.org/10.1016/j.bbrc.2004.06.009>

Fabioux C, Corporeau C, Quillien V, Favrel P, Huvet A. (2009). In vivo RNA interference in oyster-vasa silencing inhibits germ cell development. *FEBS J* 276:2566–2573. <https://doi.org/10.1111/j.1742-4658.2009.06982.x>

Flynn, K., Wedin, M. B., Bonventre, J. A., Dillon-White, M., Hines, J., Weeks, B. S., ... Gagné, F. (2013). Burrowing in the Freshwater Mussel *Elliptio complanata* is Sexually Dimorphic and Feminized by Low Levels of Atrazine. *Journal of Toxicology and Environmental Health, Part A*, 76(20), 1168–1181. <https://doi.org/10.1080/15287394.2013.845865>

Gamble, T., & Zarkower, D. (2012). Sex determination. *Current Biology*, 22(8), R257–R262. <https://doi.org/10.1016/j.cub.2012.02.054>

Gangloff, M. M., Lenertz, K. K., & Feminella, J. W. (2008). Parasitic mite and trematode abundance are associated with reduced reproductive output and physiological condition of freshwater mussels. *Hydrobiologia*, 610(1), 25. <https://doi.org/10.1007/s10750-008-9419-8>

Garrido-Ramos, M.A., Stewart, D.T., Sutherland, B.W., and 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, 818–824.

Ghiselin, M.T. (1969). The evolution of hermaphroditism among animals. *Q Rev Biol.* 44:189–208

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. <https://doi.org/10.1093/molbev/msq271>

Ghiselli, F., Milani, L., and 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, 949–961.

Ghiselli, F., Milani, L., Chang, P. L., Hedgecock, D., Davis, J. P., Nuzhdin, S. V., & Passamonti, M. (2012). De Novo Assembly of the Manila Clam *Ruditapes philippinarum* Transcriptome Provides New Insights into Expression Bias, Mitochondrial Doubly Uniparental Inheritance and Sex Determination. *Molecular Biology and Evolution*, 29(2), 771–786.
<https://doi.org/10.1093/molbev/msr248>

Giese, A. (2012). Reproduction of Marine Invertebrates V5: Molluscs : Pelecypods and Lesser Classes (Elsevier).

Gilbert, S.F. (2000). Environmental Sex Determination. in: *Developmental Biology*. 6th edition. Sunderland (MA).

Gilbert S.F. (2005). Mechanisms for the environmental regulation of gene expression: Ecological aspects of animal development; *J. Biosci.*(30) 65–74.

Gissi, C., Iannelli, F., & Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity*, 101(4), 301–320.
<https://doi.org/10.1038/hdy.2008.62>

Gomes-dos-Santos, A., Froufe, E., Amaro, R., Ondina, P., Breton, S., Guerra, D., ... Lopes-Lima, M. (2019). The male and female complete mitochondrial genomes of the threatened freshwater pearl mussel *Margaritifera margaritifera* (Linnaeus, 1758) (Bivalvia: Margaritiferidae). *Mitochondrial DNA Part B: Resources*, 4(1), 1417–1420.
<https://doi.org/10.1080/23802359.2019.1598794>

Gomot, L., & Griffond, B. (1993). Action of epigenetic factors on the expression of hermaphroditism in the snail *Helix aspersa*. *Comparative Biochemistry and Physiology -- Part A: Physiology*. [https://doi.org/10.1016/0300-9629\(93\)90303-L](https://doi.org/10.1016/0300-9629(93)90303-L)

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644–652. <https://doi.org/10.1038/nbt.1883>

Grande, C., Templado, J., Cervera, J. L., & Zardoya, R. (2002). The complete mitochondrial genome of the nudibranch *Roboastra europaea* (Mollusca: Gastropoda) supports

the monophyly of opisthobranchs. *Molecular Biology and Evolution*, 19(10), 1672–1685.

<https://doi.org/10.1093/oxfordjournals.molbev.a003990>

Guerra, D., Plazzi, F., Stewart, D. T., Bogan, A. E., Hoeh, W. R., & Breton, S. (2017). Evolution of sex-dependent mtDNA transmission in freshwater mussels (Bivalvia: Unionida). *Scientific Reports*, 7(1), 1–13. <https://doi.org/10.1038/s41598-017-01708-1>

Gusman, A., Lecomte, S., Stewart, D. T., Passamonti, M., & Breton, S. (2016). Pursuing the quest for better understanding the taxonomic distribution of the system of doubly uniparental inheritance of mtDNA. *PeerJ*, 4(12), 1–27. <https://doi.org/10.7717/peerj.2760>

Haag, W. R. (2012). North American Freshwater Mussels: Natural History, Ecology, and Conservation. Cambridge University Press.

Haag, W. R., & Warren, M. L. (1999). Mantle displays of freshwater mussels elicit attacks from fish. *Freshwater Biology*, 42(1), 35–40. <https://doi.org/10.1046/j.1365-2427.1999.00454.x>

Haas, B. J., and Papanicolaou, A. (2017). TransDecoder. Retrieved from <http://transdecoder.github.io>.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... Regev, A. (2013). De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nature Protocols*, 8(8). <https://doi.org/10.1038/nprot.2013.084>

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. <https://doi.org/10.1093/nar/gkt371>

Hanson, M. R., & Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell*. <https://doi.org/10.1105/tpc.015966>

Hayashi, J.-I., Yonekawa, H., Gotoh, O., Watanabe, J., and Tagashira, Y. (1978). Strictly maternal inheritance of rat mitochondrial DNA. *Biochemical and Biophysical Research*

Communications 83, 1032–1038.

Hayes, T.B. (1998). Sex determination and primary sex differentiation in amphibians: Genetic and developmental mechanisms. *Journal of Experimental Zoology*.

Heller, J. (1993). Hermaphroditism in molluscs. *Biological Journal of the Linnean Society* 48, 19–42.

Heo, J. M., Ordureau, A., Paulo, J. A., Rinehart, J., & Harper, J. W. (2015). The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Molecular Cell*, 60(1), 7–20.
<https://doi.org/10.1016/j.molcel.2015.08.016>

Ho, M. S., Tsai, P. I., & Chien, C. T. (2006). F-box proteins: The key to protein degradation. *Journal of Biomedical Science*. <https://doi.org/10.1007/s11373-005-9058-2>

Hoeh, W. R., Stewart, D. T., Sutherland, B. W., & Zouros, E. (1996). Cytochrome c oxidase sequence comparisons suggest an unusually high rate of mitochondrial DNA evolution in *Mytilus* (Mollusca: Bivalvia). *Molecular Biology and Evolution*, 13(2), 418–421.
<https://doi.org/10.1093/oxfordjournals.molbev.a025600>

Hu, J., Huang, W., Huang, Q., Qin, X., Yu, C., Wang, L., ... Zhu, Y. (2014). Mitochondria and cytoplasmic male sterility in plants. *Mitochondrion*, 19(PB), 282–288.
<https://doi.org/10.1016/j.mito.2014.02.008>

Huang, X.-C., Rong, J., Liu, Y., Zhang, M.-H., Wan, Y., Ouyang, S., Zhou, C.-H., and Wu, X.-P. (2013). The Complete Maternally and Paternally Inherited Mitochondrial Genomes of the Endangered Freshwater Mussel *Solenia carinatus* (Bivalvia: Unionidae) and Implications for Unionidae Taxonomy. *PLoS ONE* 8, e84352.

Jakob, S., and Lovell-Badge, R. (2011). Sex determination and the control of Sox9 expression in mammals. *FEBS Journal* 278, 1002–1009.

Jameson, D., Gibson, A. P., Hudelot, C., & Higgs, P. G. (2003). OGRe: a relational database for comparative analysis of mitochondrial genomes. *Nucleic Acids Research*, 31(1),

202–206. <https://doi.org/10.1093/nar/gkg077>

Ji, J., Huang, W., Yin, C., & Gong, Z. (2013). Mitochondrial cytochrome c oxidase and F1Fo-ATPase dysfunction in peppers (*Capsicum annuum L.*) with cytoplasmic male sterility and its association with *orf507* and *Ψatp6-2* genes. *International Journal of Molecular Sciences*, 14(1), 1050–1068. <https://doi.org/10.3390/ijms14011050>

Juchault, P., & Legrand, J.J. (1976). Modification de la sex ratio dans les croisements entre différentes populations du crustacé oniscoïde *Armadillidium vulgare*. Latr. Notion de déterminisme polygénique et épigénétique du sexe. *Arch. Zool. Exp. Gen.* 117:81–93.

Kang, S. W., Park, S. Y., Patnaik, B. B., Hwang, H. J., Kim, C., Kim, S., ... Lee, Y. S. (2015). Construction of PANM Database (Protostome DB) for rapid annotation of NGS data in Mollusks. *The Korean Journal of Malacology*, 31(3), 243–247.
<https://doi.org/10.9710/kjm.2015.31.3.243>

Kayal, E., Bentlage, B., Collins, A.G., Kayal, M., Pirro, S., and Lavrov, D.V. (2012). Evolution of Linear Mitochondrial Genomes in Medusozoan Cnidarians. *Genome Biol Evol* 4, 1–12.

Kenchington, E., MacDonald, B., Cao, L., Tsagkarakis, D., & Zouros, E. (2002). Genetics of mother-dependent sex ratio in blue mussels (*Mytilus spp.*) and Implications for doubly uniparental inheritance of mitochondrial DNA. *Genetics*, 161(4), 1579–1588.

Kenchington, E.L., Hamilton, L., Cogswell, A., and Zouros, E. (2009). Paternal mtDNA and Maleness Are Co-Inherited but Not Causally Linked in Mytilid Mussels. *PLoS ONE* 4, e6976.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. *Genome Research*, 12(4), 656-664. <https://doi.org/10.1101/gr.229202>

Kim, M. A., Rhee, J.-S., Kim, T. H., Lee, J. S., Choi, A.-Y., Choi, B.-S., ... Sohn, Y. C. (2017). Alternative Splicing Profile and Sex-Preferential Gene Expression in the Female and Male Pacific Abalone *Haliotis discus hannai*. *Genes*, 8(3). <https://doi.org/10.3390/genes8030099>

Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P., and Lovell-Badge, R. Transgenic for Sry.

Kopp, A. (2012). Dmrt genes in the development and evolution of sexual dimorphism. Trends in Genetics: TIG, 28(4), 175–184. <https://doi.org/10.1016/j.tig.2012.02.002>

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ -mer weighting and repeat separation. Genome Research, 27(5), 722–736. <https://doi.org/10.1101/gr.215087.116>

Kroon, F.J., Munday, P.L., Westcott, D.A., Hobbs, J.-P.A., and Liley, N.R. (2005). Aromatase pathway mediates sex change in each direction. Proceedings of the Royal Society of London B: Biological Sciences 272, 1399–1405.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Molecular Biology and Evolution, 33(7), 1870–1874. <https://doi.org/10.1093/molbev/msw054>

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. Genome Biology, 5(2), 12.

Kyriakou, E., Zouros, E., and Rodakis, G.C. (2010). The atypical presence of the paternal mitochondrial DNA in somatic tissues of male and female individuals of the blue mussel species *Mytilus galloprovincialis*. BMC Research Notes 3, 222.

Lalli, E., Ohe, K., Latorre, E., Bianchi, M.E., and Sassone-Corsi, P. (2003). Sexy splicing: regulatory interplays governing sex determination from *Drosophila* to mammals. Journal of Cell Science 116, 441–445.

Lanfear, R., Schalamun, M., Kainer, D., Wang, W., & Schwessinger, B. (2019). MinIONQC: fast and simple quality control for MinION sequencing data. Bioinformatics (Oxford, England), 35(3), 523–525. <https://doi.org/10.1093/bioinformatics/bty654>

Lang, J.W., and Andrews, H.V. (1994). Temperature-dependent sex determination in

crocodilians. *Journal of Experimental Zoology* 270, 28–44.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*, 12, 124. <https://doi.org/10.1186/1471-2105-12-124>

Lee, C., Zeng, J., Drew, B.G., Sallam, T., Martin-Montalvo, A., Wan, J., Kim, S.-J., Mehta, H., Hevener, A.L., de Cabo, R., et al. (2015). The Mitochondrial-Derived Peptide MOTS-c Promotes Metabolic Homeostasis and Reduces Obesity and Insulin Resistance. *Cell Metabolism* 21, 443–454.

Lee, S. R., & Han, J. (2017). Mitochondrial Nucleoid: Shield and Switch of the Mitochondrial Genome. *Oxidative Medicine and Cellular Longevity*, 2017, 8060949. <https://doi.org/10.1155/2017/8060949>

Legrand, J.J., & Juchault, P. (1969). Le déterminisme de l'intersexualité chez les crustacés isopodes terrestres; corrélation entre intersexualité et monogénie. *CR Acad. Sci. Paris*. 268:1647–1649

Legrand, J.J., Juchault, P. (1984). Nouvelles données sur le déterminisme génétique et épigénétique de la monogénie chez le crustace isopode terrestre *Armadillidium vulgare*. *Latr. Genet. Sel. Evol.* 1:16:57–84. doi: 10.1186/1297-9686-16-1-57.

Lehtonen, J., Jennions, M.D., and Kokko, H. (2012). The many costs of sex. *Trends in Ecology & Evolution* 27, 172–178.

Lewandowska, A., Macfarlane, J., & Shaw, J. M. (2013). Mitochondrial association, protein phosphorylation, and degradation regulate the availability of the active Rab GTPase Ypt11 for mitochondrial inheritance. <https://doi.org/10.1091/mbc.E12-12-0848>

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-Bio]*.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Subgroup, 1000 Genome Project Data Processing. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (Oxford, England), 25(16), 2078–2079.
<https://doi.org/10.1093/bioinformatics/btp352>

Li, R., Zhang, L., Li, W., Zhang, Y., Li, Y., Zhang, M., ... Bao, Z. (2018). FOXL2 and DMRT1L are yin and yang genes for determining timing of sex differentiation in the bivalve mollusk *Patinopecten yessoensis*. *Frontiers in Physiology*, 9.

<https://doi.org/10.3389/fphys.2018.01166>

Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659.
<https://doi.org/10.1093/bioinformatics/btl158>

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659.

Li, Y., Zhang, L., Sun, Y., Ma, X., Wang, J., Li, R., ... Bao, Z. (2016). Transcriptome Sequencing and Comparative Analysis of Ovary and Testis Identifies Potential Key Sex-Related Genes and Pathways in Scallop *Patinopecten yessoensis*. *Marine Biotechnology*, 18(4), 453–465.
<https://doi.org/10.1007/s10126-016-9706-8>

Ling, J., O'Donoghue, P., & Söll, D. (2015). Genetic code flexibility in microorganisms: novel mechanisms and impact on physiology. *Nature Reviews. Microbiology*, 13(11), 707–721.
<https://doi.org/10.1038/nrmicro3568>

Liu, H., Lamm, M. S., Rutherford, K., Black, M. A., Godwin, J. R., & Gemmell, N. J. (2015). Large-scale transcriptome sequencing reveals novel expression patterns for key sex-related genes in a sex-changing fish. *Biology of Sex Differences*, 6(1).
<https://doi.org/10.1186/s13293-015-0044-8>

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of Relative Gene Expression Data

Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25, 402-408.

Leclercq, S., Theze, J., Chebbi, M.A., Giraud, I., Moumen, B., Ernenwein, L., Greve, P., Gilbert, C., Cordaux, R. (2016). Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome. *Proc. Natl. Acad. Sci. USA.* 113:15036–15041. doi: 10.1073/pnas.1608979113.

Llera-Herrera, R., García-Gasca, A., Abreu-Goodger, C., Huvet, A., & Ibarra, A. M. (2013). Identification of male gametogenesis expressed genes from the scallop *Nodipecten subnodosus* by suppressive subtraction hybridization and pyrosequencing. *PloS One*, 8(9), e73176. <https://doi.org/10.1371/journal.pone.0073176>

Loman, N. J., Quick, J., & Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nature Methods*, 12(8), 733–735. <https://doi.org/10.1038/nmeth.3444>

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>

Lowe, T. M., & Eddy, S. R. (1997). tRNAscan-SE: A Program for Improved Detection of Transfer RNA Genes in Genomic Sequence. *Nucleic Acids Research*, 25(5), 955–964. <https://doi.org/10.1093/nar/25.5.955>

Lubet, P. (1959) Recherches sur le cycle sexuel et l'émission des gamètes chez les mytilidés et les pectinidés (Mollusques bivalves). *Rev Trav Inst Pêches Marit.* 23:387–548

Lucas, A. (1965) Recherche sur la sexualité des mollusques bivalves. *Bull Biol Fr Berg.* 99:115–247

Luckenbach, J. A., Iliev, D. B., Goetz, F. W., & Swanson, P. (2008). Identification of differentially expressed ovarian genes during primary and early secondary oocyte growth in coho salmon, *Oncorhynchus kisutch*. *Reproductive Biology and Endocrinology*, 6, 2. <https://doi.org/10.1186/1477-7827-6-2>

MacDougall, C., Harbison, D., and Bownes, M. (1995). The Developmental

Consequences of Alternate Splicing in Sex Determination and Differentiation in Drosophila. Developmental Biology 172, 353–376.

Madeira, F., Park, Y. mi, Lee, J., Buso, N., Gur, T., Madhusoodanan, N., ... Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research, 47(W1), W636–W641. <https://doi.org/10.1093/nar/gkz268>

Manolakou, P., Angelopoulou, R., & Lavranos, G. (2006). Sex determinants in the genome--lessons from the animal kingdom. Collegium Antropologicum, 30(3), 649–652.

Manolakou, P., Angelopoulou, R., and Lavranos, G. (2006). Sex Determinants in the Genome—Lessons from the Animal Kingdom. Collegium Antropologicum 30, 649–652.

Maresca, A., Zaffagnini, M., Caporali, L., Carelli, V., & Zanna, C. (2015). DNA methyltransferase 1 mutations and mitochondrial pathology: Is mtDNA methylated? Frontiers in Genetics, 5(FEB), 90. <https://doi.org/10.3389/fgene.2015.00090>

Matson, C. K., & Zarkower, D. (2012). Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. Nature Reviews. Genetics, 13(3), 163–174. <https://doi.org/10.1038/nrg3161>

McCartney, M. A., Auch, B., Kono, T., Mallez, S., Zhang, Y., Obille, A., ... Gohl, D. M. (2019). The genome of the zebra mussel, *Dreissena polymorpha*: A resource for invasive species research. BioRxiv, 696732. <https://doi.org/10.1101/696732>

Milani, L., Ghiselli, F., & Passamonti, M. (2016). Mitochondrial selfish elements and the evolution of biological novelties. Current Zoology, zow044. <https://doi.org/10.1093/cz/zow044>

Milani, L., Ghiselli, F., and Passamonti, M. (2012). Sex-Linked Mitochondrial Behavior During Early Embryo Development in *Ruditapes philippinarum* (Bivalvia Veneridae) a Species with the Doubly Uniparental Inheritance (DUI) of Mitochondria. J. Exp. Zool. 318, 182–189.

Milani, L., Ghiselli, F., Guerra, D., Breton, S., & Passamonti, M. (2013a). A Comparative Analysis of Mitochondrial ORFans: New Clues on Their Origin and Role in Species with Doubly Uniparental Inheritance of Mitochondria. Genome Biology and Evolution, 5(7), evt101.

<https://doi.org/10.1093/gbe/evt101>

Milani, L., Ghiselli, F., Iannello, M., and Passamonti, M. (2014a). Evidence for somatic transcription of male-transmitted mitochondrial genome in the DUI species *Ruditapes philippinarum* (Bivalvia: Veneridae). *Curr Genet* 60, 163–173.

Milani, L., Ghiselli, F., Maurizii, M. G., & Passamonti, M. (2011). Doubly Uniparental Inheritance of Mitochondria As a Model System for Studying Germ Line Formation. *PLoS ONE*, 6(11), e28194. <https://doi.org/10.1371/journal.pone.0028194>

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. <https://doi.org/10.1093/gbe/evu021>

Milani, L., Ghiselli, F., Nuzhdin, S.V., and Passamonti, M. (2013b). Nuclear genes with sex bias in *Ruditapes philippinarum* (Bivalvia, veneridae): Mitochondrial inheritance and sex determination in DUI species. *J. Exp. Zool. (Mol. Dev. Evol.)* 320, 442–454.

Milani, L., Ghiselli, F., Pecci, A., Maurizii, M. G., & Passamonti, M. (2015). The Expression of a Novel Mitochondrially-Encoded Gene in Gonadic Precursors May Drive Paternal Inheritance of Mitochondria. *PLOS ONE*, 10(9), e0137468.
<https://doi.org/10.1371/journal.pone.0137468>

Milani, L., Ghiselli, F., Pecci, A., Maurizii, M.G., and Passamonti, M. (2015). The Expression of a Novel Mitochondrially-Encoded Gene in Gonadic Precursors May Drive Paternal Inheritance of Mitochondria. *PLOS ONE* 10, e0137468.

Milbury, C. A., & Gaffney, P. M. (2005). Complete mitochondrial DNA sequence of the eastern oyster *Crassostrea virginica*. *Marine Biotechnology*, 7(6), 697–712.
<https://doi.org/10.1007/s10126-005-0004-0>

Minoiu, I., Burzyński, A., & Breton, S. (2016). Analysis of the coding potential of the ORF in the control region of the female-transmitted *Mytilus* mtDNA. *Gene*. Elsevier.
<https://doi.org/10.1016/j.gene.2015.09.065>

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).

<https://doi.org/10.1186/s12864-016-2986-6>

Mitchell, N.J., Nelson, N.J., Cree, A., Pledger, S., Keall, S.N., and Daugherty, C.H. (2006). Support for a rare pattern of temperature-dependent sex determination in archaic reptiles: evidence from two species of tuatara (*Sphenodon*). *Frontiers in Zoology* 3.

Moy, G. W., Springer, S. A., Adams, S. L., Swanson, W. J., & Vacquier, V. D. (2008). Extraordinary intraspecific diversity in oyster sperm bindin. *Proceedings of the National Academy of Sciences of the United States of America*, 105(6), 1993–1998.

<https://doi.org/10.1073/pnas.0711862105>

Müller, T., Czarnoleski, M., Labecka, A. M., Cichy, A., Zajac, K., & Dragosz-Kluska, D. (2015). Factors affecting trematode infection rates in freshwater mussels. *Hydrobiologia*, 742(1), 59–70. <https://doi.org/10.1007/s10750-014-1965-7>

Naimi, A., Martinez, A. S., Specq, M. L., Diss, B., Mathieu, M., & Sourdaine, P. (2009). Molecular cloning and gene expression of *Cg-Foxl2* during the development and the adult gametogenetic cycle in the oyster *Crassostrea gigas*. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, 154(1), 134–142.

<https://doi.org/10.1016/j.cbpb.2009.05.011>

Naimi, A., Martinez, A.-S. S., Specq, M.-L. L., Mrac, A., Diss, B., Mathieu, M., & Sourdaine, P. (2009). Identification and expression of a factor of the DM family in the oyster *Crassostrea gigas*. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, 152(2), 189–196. <https://doi.org/10.1016/j.cbpa.2008.09.019>

Navarro-Martín, L., Viñas, J., Ribas, L., Díaz, N., Gutiérrez, A., Di Croce, L., and Piferrer, F. (2011). DNA Methylation of the Gonadal Aromatase (cyp19a) Promoter Is Involved in Temperature-Dependent Sex Ratio Shifts in the European Sea Bass. *PLoS Genet* 7, e1002447.

Negri, I., Franchini, A., Gonella, E., Daffonchio, D., Mazzoglio, P. J., Mandrioli, M., &

Alma, A. (2009). Unravelling the Wolbachia evolutionary role: the reprogramming of the host genomic imprinting. *Proceedings of the Royal Society B: Biological Sciences*, 276(1666), 2485–2491. <https://doi.org/10.1098/rspb.2009.0324>

Nei, M., & Gojobori, T. (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, 3(5), 418–426. <https://doi.org/10.1093/oxfordjournals.molbev.a040410>

Ni, J., Zeng, Z., Han, G., Huang, H., & Ke, C. (2012). Cloning and characterization of the follistatin gene from *Crassostrea angulata* and its expression during the reproductive cycle. *Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology*, 163(2), 246–253. <https://doi.org/10.1016/j.cbpb.2012.06.006>

Nishimura, Y. (2010). Uniparental inheritance of cpDNA and the genetic control of sexual differentiation in *Chlamydomonas reinhardtii*. *Journal of Plant Research*, 123(2), 149–162. <https://doi.org/10.1007/s10265-009-0292-y>

Noutahi, E., Calderon, V., Blanchette, M., El-Mabrouk, N., & Lang, B. F. (2019). Rapid Genetic Code Evolution in Green Algal Mitochondrial Genomes. *Molecular Biology and Evolution*, 36(4), 766–783. <https://doi.org/10.1093/molbev/msz016>

O'Donnell, L., & O'Bryan, M. K. (2014). Microtubules and spermatogenesis. *Seminars in Cell & Developmental Biology*, 30, 45–54. <https://doi.org/10.1016/j.semcdb.2014.01.003>

O'Donnell, L., Rhodes, D., Smith, S. J., Merriner, D. J., Clark, B. J., Borg, C., ... O'Bryan, M. K. (2012). An essential role for Katanin p80 and microtubule severing in male gamete production. *PLoS Genetics*, 8(5). <https://doi.org/10.1371/journal.pgen.1002698>

O'Toole, N., Hattori, M., Andres, C., Iida, K., Lurin, C., Schmitz-Linneweber, C., ... Small, I. (2008). On the expansion of the pentatricopeptide repeat gene family in plants. *Molecular Biology and Evolution*, 25(6), 1120–1128. <https://doi.org/10.1093/molbev/msn057>

Obata, M., & Komaru, A. (2005). Specific location of sperm mitochondria in mussel *Mytilus galloprovincialis* zygotes stained by MitoTracker. *Development Growth and Differentiation*, 47(4), 255–263. <https://doi.org/10.1111/j.1440-169X.2005.00801.x>

Obata, M., Kamiya, C., Kawamura, K., and Komaru, A. (2006). Sperm mitochondrial DNA transmission to both male and female offspring in the blue mussel *Mytilus galloprovincialis*. *Development, Growth & Differentiation* 48, 253–261.

Obata, M., Sano, N., and 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, 878–886.

Ojala, D., Montoya, J., & Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. *Nature*, 290(5806), 470–474.
<https://doi.org/10.1038/290470a0>

Oldfield E (1961) The functional morphology of *Kellia suborbicularis* (Montagu), *Montacuta ferruginosa* (Montagu) and *M. substriata* (Montagu), (Mollusca, Lamellibranchiata). *J Moll Stud*. 34: 255–295.

Ospina-Álvarez, N., and Piferrer, F. (2008). Temperature-Dependent Sex Determination in Fish Revisited: Prevalence, a Single Sex Ratio Response Pattern, and Possible Effects of Climate Change. *PLoS ONE* 3, e2837.

Pante, E., Poitrimol, C., Saunier, A., Becquet, V., & Garcia, P. (2017). Putative sex-linked heteroplasmy in the tellinid bivalve *Limecola balthica* (Linnaeus, 1758). *Journal of Molluscan Studies*, 83(2), 226–228. <https://doi.org/10.1093/mollus/eyw038>

Pardue, M.L. (1990). Thirtieth Annual Meeting of the Japan Society of Histochemistry and Cytochemistry. *Acta Histochemica et Cytochemica* 23, 1–16.

Park, H., & Ahn, D. H. (2015). Complete mitochondrial genome of the Antarctic soft-shelled clam, *Laternula elliptica* (Bivalvia; Laternulidae). *Mitochondrial DNA*, 26(4), 642–643. <https://doi.org/10.3109/19401736.2013.836515>

Passamonti, M., & Ghiselli, F. (2009). Doubly Uniparental Inheritance: Two Mitochondrial Genomes, One Precious Model for Organelle DNA Inheritance and Evolution. *DNA and Cell Biology*, 28(2), 79–89. <https://doi.org/10.1089/dna.2008.0807>

Passamonti, M., & Scali, V. (2001). Gender-associated mitochondrial DNA heteroplasmy in the venerid clam *Tapes philippinarum* (Mollusca Bivalvia). Current Genetics, 39(2), 117–124. <https://doi.org/10.1007/s002940100188>

Passamonti, M., Ricci, A., Milani, L., & Ghiselli, F. (2011). Mitochondrial genomes and Doubly Uniparental Inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (Bivalvia Mytilidae). BMC Genomics, 12(1), 442. <https://doi.org/10.1186/1471-2164-12-442>

Passamonti, M., Ricci, A., Milani, L., and Ghiselli, F. (2011). Mitochondrial genomes and Doubly Uniparental Inheritance: new insights from *Musculista senhousia* sex-linked mitochondrial DNAs (Bivalvia Mytilidae). BMC Genomics 12, 442.

Passamonti, M., Stewart, D. T., Breton, S., Milani, L., Ghiselli, F., Guerra, D., ... Passamonti, M. (2014). A resourceful genome: Updating the functional repertoire and evolutionary role of animal mitochondrial DNAs. Trends in Genetics, 30(12), 555–564. <https://doi.org/10.1016/j.tig.2014.09.002>

Patnaik, B. B., Wang, T. H., Kang, S. W., Hwang, H. J., Park, S. Y., Park, E. B., ... Lee, Y. S. (2016). Sequencing, de novo assembly, and annotation of the transcriptome of the endangered freshwater pearl bivalve, *Cristaria plicata*, provides novel insights into functional genes and marker discovery. PLoS ONE, 11(2), 1–28. <https://doi.org/10.1371/journal.pone.0148622>

Penalva, L.O.F., and Sánchez, L. (2003). RNA Binding Protein Sex-Lethal (Sxl) and Control of Drosophila Sex Determination and Dosage Compensation. Microbiol. Mol. Biol. Rev. 67, 343–359.

Peng, J., Wei, P., Zhang, B., Zhao, Y., Zeng, D., Chen, X., ... Chen, X. (2015). Gonadal transcriptomic analysis and differentially expressed genes in the testis and ovary of the Pacific white shrimp (*Litopenaeus vannamei*). BMC Genomics, 16. <https://doi.org/10.1186/s12864-015-2219-4>

Perlman, S. J., Hodson, C. N., Hamilton, P. T., Opit, G. P., & Gowen, B. E. (2015).

Maternal transmission, sex ratio distortion, and mitochondria. Proceedings of the National Academy of Sciences, 112(33), 10162–10168. <https://doi.org/10.1073/pnas.1421391112>

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucl. Acids Res.* 29, e45–e45.

Pfaffl, M.W., Tichopad, A., Prgomet, C., and Neuvians, T.P. (2004). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnology Letters* 26, 509–515.

Picard, M., Wallace, D. C., & Burelle, Y. (2016). The rise of mitochondria in medicine. *Mitochondrion*. <https://doi.org/10.1016/j.mito.2016.07.003>

Picard, M.A.-L., Cosseau, C., Mouahid, G., Duval, D., Grunau, C., Toulza, È., Allienne, J.-F., and Boissier, J. (2015). The roles of Dmrt (Double sex/Male-abnormal-3 Related Transcription factor) genes in sex determination and differentiation mechanisms: Ubiquity and diversity across the animal kingdom. *Comptes Rendus Biologies*.

Pieau, C., and Dorizzi, M. (1981). Determination of temperature sensitive stages for sexual differentiation of the gonads in embryos of the turtle, *Emys orbicularis*. *J. Morphol.* 170, 373–382.

Pieau, C., and Dorizzi, M. (2004). Oestrogens and temperature-dependent sex determination in reptiles: all is in the gonads. *J Endocrinol* 181, 367–377.

Pila, E. A., Peck, S. J., & Hanington, P. C. (2017). The protein pheromone temptin is an attractant of the gastropod *Biomphalaria glabrata*. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, 203(10), 855–866. <https://doi.org/10.1007/s00359-017-1198-0>

Plazzi, F., & Passamonti, M. (2019). Footprints of unconventional mitochondrial inheritance in bivalve phylogeny: Signatures of positive selection on clades with doubly uniparental inheritance. *Journal of Zoological Systematics and Evolutionary Research*, 57(2), 258–271. <https://doi.org/10.1111/jzs.12253>

Plazzi, F., Puccio, G., & Passamonti, M. (2016). Comparative large-scale mitogenomics evidences clade-specific evolutionary trends in mitochondrial DNAs of Bivalvia. *Genome Biology and Evolution*, 8(8). <https://doi.org/10.1093/gbe/evw187>

Pollack, M., and Leeuwenburgh, C. (2001). Apoptosis and Aging Role of the Mitochondria. *J Gerontol A Biol Sci Med Sci* 56, B475–B482.

Policansky, D. (1982). Sex change in plants and animals. *Annu Rev Ecol Syst* 13:471–495

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., ... Prost, S. (2018). Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, 7(4). <https://doi.org/10.1093/gigascience/giy033>

Pozzi, A., Plazzi, F., Milani, L., Ghiselli, F., & Passamonti, M. (2017). SmithRNAs: Could Mitochondria “Bend” Nuclear Regulation? *Molecular Biology and Evolution*, 34(8), 1960–1973. <https://doi.org/10.1093/molbev/msx140>

Punzi, E., Milani, L., Ghiselli, F., & Passamonti, M. (2018). Lose it or keep it: (how bivalves can provide) insights into mitochondrial inheritance mechanisms. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 330(1), 41–51. <https://doi.org/10.1002/jez.b.22788>

Radonić, A., Thulke, S., Mackay, I.M., Landt, O., Siegert, W., and Nitsche, A. (2004). Guideline to reference gene selection for quantitative real-time PCR. *Biochem. Biophys. Res. Commun.* 313, 856–862.

Rampersad, S., & Tennant, P. (2018). Replication and Expression Strategies of Viruses. In *Viruses: Molecular Biology, Host Interactions, and Applications to Biotechnology* (pp. 55–82). Elsevier. <https://doi.org/10.1016/B978-0-12-811257-1.00003-6>

Renaut, S., Guerra, D., Hoeh, W. R., Stewart, D. T., Bogan, A. E., Ghiselli, F., ... Breton, S. (2018). Genome survey of the freshwater mussel *Venustaconcha ellipsiformis* (Bivalvia: Unionida) using a hybrid de novo assembly approach. *Genome Biology and Evolution*. Oxford

University Press. <https://doi.org/10.1093/gbe/evy117>

Rezende, V. B., Congrains, C., Lima, A. L. A., Campanini, E. B., Nakamura, A. M., de Oliveira, J. L., ... de Brito, R. A. (2016). Head transcriptomes of two closely related species of fruit flies of the *anastrepha fraterculus* group reveals divergent genes in species with extensive gene flow. *G3: Genes, Genomes, Genetics*, 6(10), 3283–3295.
<https://doi.org/10.1534/g3.116.030486>

Rigaud, T., & Juchault, P. (1993). Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare* Latr. *Genetics*. 133:247–252

Rice, P., Ray, S. M., Painter, S. J., & Nagle, G. T. (2002). An intrinsic membrane protein in oyster sperm stimulates spawning behaviors in *Crassostrea virginica*: Implications for aquaculture. *Journal of Shellfish Research*, 21(2), 715-718.

Rico-Villa, B., Le Coz, J.R., Mingant, C., and Robert, R. (2006). Influence of phytoplankton diet mixtures on microalgae consumption, larval development and settlement of the Pacific oyster *Crassostrea gigas* (Thunberg). *Aquaculture* 256, 377–388.

Robicheau, B. M., Breton, S., & Stewart, D. T. (2017). Sequence motifs associated with paternal transmission of mitochondrial DNA in the horse mussel, *Modiolus modiolus* (Bivalvia: Mytilidae). *Gene*, 605, 32–42. <https://doi.org/10.1016/j.gene.2016.12.025>

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139–140. <https://doi.org/10.1093/bioinformatics/btp616>

Robinson, M., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, 11(3), R25. <https://doi.org/10.1186/gb-2010-11-3-r25>

Rosel, P. E., & Block, B. A. (1996). Mitochondrial control region variability and global population structure in the swordfish, *Xiphias gladius*. *Marine Biology*, 125(1), 11–22.

<https://doi.org/10.1007/BF00350756>

Russell, R. D., & Beckenbach, A. T. (2008). Recoding of translation in turtle mitochondrial genomes: programmed frameshift mutations and evidence of a modified genetic code. *Journal of Molecular Evolution*, 67(6), 682–695. <https://doi.org/10.1007/s00239-008-9179-0>

Saavedra, C., Reyero, M.-I., and Zouros, E. (1997). Male-Dependent Doubly Uniparental Inheritance of Mitochondrial DNA and Female-Dependent Sex-Ratio in the Mussel *Mytilus galloprovincialis*. *Genetics* 145, 1073–1082.

Saccone, G., Pane, A., and Polito, L.C. (2002). Sex Determination in Flies, Fruitflies and Butterflies. *Genetica* 116, 15–23.

Sager, R., & Grabowy, C. (1983). Differential methylation of chloroplast DNA regulates maternal inheritance in a methylated mutant of Chlamydomonas. *Proceedings of the National Academy of Sciences of the United States of America*, 80(10), 3025–3029.

Sanchez, L. (2008). Sex-determining mechanisms in insects. *The International Journal of Developmental Biology* 52, 837–856.

Santerre, C., Sourdaine, P., & Martinez, A. S. (2012). Expression of a natural antisense transcript of Cg-Foxl2 during the gonadic differentiation of the oyster *Crassostrea gigas*: First demonstration in the gonads of a lophotrochozoa species. *Sexual Development*, 6(4), 210–221. <https://doi.org/10.1159/000338085>

Santerre, C., Sourdaine, P., Adeline, B., & Martinez, A. S. (2014). Cg-SoxE and Cg- β -catenin, two new potential actors of the sex-determining pathway in a hermaphrodite lophotrochozoan, the Pacific oyster *Crassostrea gigas*. *Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology*, 167, 68–76. <https://doi.org/10.1016/j.cbpa.2013.09.018>

Santerre, C., Sourdaine, P., Marc, N., Mingant, C., Robert, R., & Martinez, A.-S. (2013). Oyster sex determination is influenced by temperature - first clues in spat during first gonadic differentiation and gametogenesis. *Comparative Biochemistry and Physiology. Part A, Molecular*

& Integrative Physiology, 165(1), 61–69. <https://doi.org/10.1016/j.cbpa.2013.02.007>

Sato, K., & Sato, M. (2017). Multiple ways to prevent transmission of paternal mitochondrial DNA for maternal inheritance in animals. *Journal of Biochemistry*. Oxford University Press. <https://doi.org/10.1093/jb/mvx052>

Saunier, A., Garcia, P., Becquet, V., Marsaud, N., Escudié, F., & Pante, E. (2014). Mitochondrial genomes of the Baltic clam *Macoma balthica* (Bivalvia: Tellinidae): Setting the stage for studying mito-nuclear incompatibilities. *BMC Evolutionary Biology*, 14(1). <https://doi.org/10.1186/s12862-014-0259-z>

Sayadi, A., Immonen, E., Bayram, H., & Arnqvist, G. (2016). The De Novo Transcriptome and Its Functional Annotation in the Seed Beetle *Callosobruchus maculatus*. *PLOS ONE*, 11(7), e0158565. <https://doi.org/10.1371/journal.pone.0158565>

Scheffler, I.E. (2007). Wiley: Mitochondria, 2nd Edition - Immo E. Scheffler.

Secq, M. P. O. Le, Green, B. R., Oudot-Le Secq, M.-P., & Green, B. R. (2011). Complex repeat structures and novel features in the mitochondrial genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Gene*, 476(1–2), 20–26. <https://doi.org/10.1016/j.gene.2011.02.001>

Seyednasrollah, F., Laiho, A., and Elo, L.L. (2013). Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform bbt086*.

Shi, J., Hong, Y., Sheng, J., Peng, K., & Wang, J. (2015). De novo transcriptome sequencing to identify the sex-determination genes in *Hyriopsis schlegelii*. *Bioscience, Biotechnology, and Biochemistry*, 0(0), 1–9. <https://doi.org/10.1080/09168451.2015.1025690>

Shi, Y., Liu, W., & He, M. (2018). Proteome and Transcriptome Analysis of Ovary, Intersex Gonads, and Testis Reveals Potential Key Sex Reversal/Differentiation Genes and Mechanism in Scallop *Chlamys nobilis*. *Marine Biotechnology*, 20(2), 220–245. <https://doi.org/10.1007/s10126-018-9800-1>

Shumway SE, Parsons GJ (eds) (2011) Scallops: biology, ecology and aquaculture. 40.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., ... Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature*, 346(6281), 240–244. <https://doi.org/10.1038/346240a0>

Śmietanka, B., & Burzyński, A. (2017). Disruption of doubly uniparental inheritance of mitochondrial DNA associated with hybridization area of European *Mytilus edulis* and *Mytilus trossulus* in Norway. *Marine Biology*, 164(11). <https://doi.org/10.1007/s00227-017-3235-5>

Soneson, C., & Delorenzi, M. (2013). A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics*, 14(1), 91. <https://doi.org/10.1186/1471-2105-14-91>

Soroka, M., & Burzyński, A. (2015). Complete female mitochondrial genome of *Anodonta anatina* (Mollusca: Unionidae): confirmation of a novel protein-coding gene (F ORF). *Mitochondrial DNA*, 26(2), 267–269. <https://doi.org/10.3109/19401736.2013.823176>

Soroka, M., & Burzyński, A. (2018). Doubly uniparental inheritance and highly divergent mitochondrial genomes of the freshwater mussel *Unio tumidus* (Bivalvia: Unionidae). *Hydrobiologia*, 810(1), 239–254. <https://doi.org/10.1007/s10750-017-3113-7>

Steele, S., and Mulcahy, M. F. (1999). Gametogenesis of the oyster *Crassostrea gigas* in southern Ireland. *Journal of the Marine Biological Association of the United Kingdom* 79, 673–686.

Stewart, D. T., Kenchington, E. R., Singh, R. K., & Zouros, E. (1996). Degree of selective constraint as an explanation of the different rates of evolution of gender-specific mitochondrial DNA lineages in the mussel mytilus. *Genetics*, 143(3), 1349–1357.

Sun, S., Kong, L., Yu, H., & Li, Q. (2015). The complete mitochondrial DNA of *Tegillarca granosa* and comparative mitogenomic analyses of three Arcidae species. *Gene*, 557(1), 61–70. <https://doi.org/10.1016/j.gene.2014.12.011>

Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., & Schatten, G. (1999). Ubiquitin tag for sperm mitochondria. *Nature*, 402(6760), 371–372. <https://doi.org/10.1038/46466>

Sutovsky, P., Moreno, R.D., Ramalho-Santos, J., Dominko, T., Simerly, C., and Schatten, G. (2000). Ubiquitinated Sperm Mitochondria, Selective Proteolysis, and the Regulation of Mitochondrial Inheritance in Mammalian Embryos. *Biol Reprod*, 63, 582–590.

Suyama, M., Torrents, D., & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, 34(Web Server issue), W609-12. <https://doi.org/10.1093/nar/gkl315>

Szafranski, P. (2017). Evolutionarily recent, insertional fission of mitochondrial cox2 into complementary genes in bilaterian Metazoa. *BMC Genomics*, 18(1), 269. <https://doi.org/10.1186/s12864-017-3626-5>

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., Zhou, B., ... NVIDIA. (2013). Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin Human Influenza A Viruses Single-Reaction Genomic Amplification Accelerates Sequencing and Vaccine Production for Classical and Swine Origin. Society, 28(12), 2725–2729. <https://doi.org/10.1128/JVI.01109-09>

Teaniniuraitemoana, V. (2014). Reproduction de l’huître perlière *Pinctada margaritifera* : étude des déterminants du sexe femelle chez l’adulte. <http://www.theses.fr>.

Teaniniuraitemoana, V., Huvet, A., Levy, P., Klopp, C., Lhuillier, E., Gaertner-Mazouni, N., ... Le Moullac, G. (2014). Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: Identification of potential sex differentiation and sex determining genes. *BMC Genomics*, 15(1), 1–20. <https://doi.org/10.1186/1471-2164-15-491>

Teaniniuraitemoana, V., Leprêtre, M., Levy, P., Vanaa, V., Parrad, S., Gaertner-Mazouni,

N., ... Le Moullac, G. (2016). Effect of temperature, food availability, and estradiol injection on gametogenesis and gender in the pearl oyster *Pinctada margaritifera*. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 325(1), 13–24.
<https://doi.org/10.1002/jez.1992>

Temperley, R., Richter, R., Dennerlein, S., Lightowers, R. N., M, Z., & Chrzanowska-Lightowers, Z. M. (2010). Supporting Online Material for hungry codons promote frameshifting in human mitochondrial ribosomee. *Science*, 301(January), 2010–2010.
<https://doi.org/10.1126/science.1180674>

Thellin, O., Zorzi, W., Lakaye, B., De Borman, B., Coumans, B., Hennen, G., Grisar, T., Igout, A., and Heinen, E. (1999). Housekeeping genes as internal standards: use and limits. *Journal of Biotechnology* 75, 291–295.

Thiriot-Quiévreux, C. (2003). Advances in Chromosomal Studies of Gastropod Molluscs. *J Mollus Stud*, 69, 187–202.

Tong, Y., Zhang, Y., Huang, J., Xiao, S., Zhang, Y., Li, J., ... Yu, Z. (2015). Transcriptomics analysis of crassostrea hongkongensis for the discovery of reproduction-related genes. *PLoS ONE*, 10(8). <https://doi.org/10.1371/journal.pone.0134280>

Turabelidze, A., Guo, S., and DiPietro, L.A. (2010). Importance of housekeeping gene selection for accurate reverse transcription-quantitative polymerase chain reaction in a wound healing model. *Wound Repair and Regeneration* 18, 460–466.

Tuteja, R., Saxena, R. K., Davila, J., Shah, T., Chen, W., Xiao, Y.-L., ... Varshney, R. K. (2013). Cytoplasmic male sterility-associated chimeric open reading frames identified by mitochondrial genome sequencing of four Cajanus genotypes. *DNA Research : An International Journal for Rapid Publication of Reports on Genes and Genomes*, 20(5), 485–495.
<https://doi.org/10.1093/dnares/dst025>

Uhlenhaut, N. H., & Treier, M. (2006). Foxl2 function in ovarian development. *Molecular Genetics and Metabolism*, 88(3), 225–234. <https://doi.org/10.1016/j.ymgme.2006.03.005>

Uliano-Silva, M., Dondero, F., Dan Otto, T., Costa, I., Lima, N. C. B., Americo, J. A., ...

Rebelo, M. de F. (2018). A hybrid-hierarchical genome assembly strategy to sequence the invasive golden mussel, *Limnoperna fortunei*. *GigaScience*, 7(2), 1–10.
<https://doi.org/10.1093/GIGASCIENCE/GIX128>

Umen, J. G., & Goodenough, U. W. (2001). Chloroplast DNA methylation and inheritance in Chlamydomonas. *Genes & Development*, 15(19), 2585–2597.

Vagner, M., Pante, E., Viricel, A., Lacoue-Labarthe, T., Zambonino-Infante, J.-L., Quazuguel, P., ... Imbert-Auvray, N. (2019). Ocean warming combined with lower omega-3 nutritional availability impairs the cardio-respiratory function of a marine fish. *The Journal of Experimental Biology*, 222(8). <https://doi.org/10.1242/jeb.187179>

Valenzuela, N. (2008). Sexual Development and the Evolution of Sex Determination. *Sexual Development* 2, 64–72.

Vallès, Y., Halanych, K. M., & Boore, J. L. (2008). Group II introns break new boundaries: Presence in a bilaterian's genome. *PLoS ONE*, 3(1).
<https://doi.org/10.1371/journal.pone.0001488>

Van Wilgenburg, E., Driessen, G., and Beukeboom, L.W. (2006). Single locus complementary sex determination in Hymenoptera: an “unintelligent” design. *Front Zool* 3.

Vandesompele, J., Preter, K.D., Pattyn, F., Poppe, B., Roy, N.V., Paepe, A.D., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* 3, research0034.

Venetis, C., Theologidis, I., Zouros, E., and Rodakis, G.C. (2006). No evidence for presence of maternal mitochondrial DNA in the sperm of *Mytilus galloprovincialis* males. *Proc Biol Sci* 273, 2483–2489.

Viger, R. S., Mertineit, C., Trasler, J. M., & Nemer, M. (1998). Transcription factor GATA-4 is expressed in a sexually dimorphic pattern during mouse gonadal development and is a potent activator of the Müllerian inhibiting substance promoter. *Development* (Cambridge, England), 125(14), 2665–2675.

Viricel, A., Becquet, V., Dubillot, E., & Pante, E. (2018). De novo assembly and functional annotation of the transcriptome of *Mimachlamys varia*, a bioindicator marine bivalve. *Marine Genomics*, 41, 42–45. <https://doi.org/10.1016/J.MARGEN.2018.04.002>

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., ... Earl, A. M. (2014). Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*, 9(11). <https://doi.org/10.1371/journal.pone.0112963>

Walker, S.M., Prodöhl, P.A., Fletcher, H.L., Hanna, R.E.B., Kantzoura, V., Hoey, E.M., and Trudgett, A. (2007). Evidence for multiple mitochondrial lineages of *Fasciola hepatica* (liver fluke) within infrapopulations from cattle and sheep. *Parasitology Research* 101, 117–125.

Wallace, D. C. (2016). Genetics: Mitochondrial DNA in evolution and disease. *Nature*. Nature Publishing Group. <https://doi.org/10.1038/nature18902>

Wallace, H., Badawy, G.M.I., and Wallace, B.M.N. (1999). Amphibian sex determination and sex reversal. *CMLS, Cell. Mol. Life Sci.* 55, 901–909.

Wallis, M. C., Waters, P. D., & Graves, J. a. M. (2008). Sex determination in mammals — Before and after the evolution of SRY. *Cellular and Molecular Life Sciences*, 65(20), 3182–3195. <https://doi.org/10.1007/s00018-008-8109-z>

Wan, C., Li, S., Wen, L., Kong, J., Wang, K., & Zhu, Y. (2007). Damage of oxidative stress on mitochondria during microspores development in Honglian CMS line of rice. *Plant Cell Reports*, 26(3), 373–382. <https://doi.org/10.1007/s00299-006-0234-2>

Wang, C., and Croll, R.P. (2004). Effects of sex steroids on gonadal development and gender determination in the sea scallop, *Placopecten magellanicus*. *Aquaculture* 238, 483–498.

Wang, X., Liu, Z., & Wu, W. (2017). Transcriptome analysis of the freshwater pearl mussel (*Cristaria plicata*) mantle unravels genes involved in the formation of shell and pearl. *Molecular Genetics and Genomics : MGG*, 292(2), 343–352. <https://doi.org/10.1007/s00438-016-1278-9>

Warner, D.A., Shine, R., and Schwenk, K. (2005). The adaptive significance of

temperature-dependent sex determination: experimental tests with a short-lived lizard. *Evolution* 59, 2209–2221.

Warner, R.R. (1984). Mating Behavior and Hermaphroditism in Coral Reef Fishes: The diverse forms of sexuality found among tropical marine fishes can be viewed as adaptations to their equally diverse mating systems. *American Scientist* 72, 128–136.

West, Lively, and Read (1999). A pluralist approach to sex and recombination. *Journal of Evolutionary Biology* 12, 1003–1012.

Wickham, J. (2018). Introduction. In *Unequal Europe* (pp. 1–6).
<https://doi.org/10.4324/9781315636702-1>

Wijayawardena, B. K., Minchella, D. J., & DeWoody, J. A. (2016). The influence of trematode parasite burden on gene expression in a mammalian host. *BMC Genomics*, 17.
<https://doi.org/10.1186/s12864-016-2950-5>

Wolstenholme, D. R. (1992). Animal Mitochondrial DNA: Structure and Evolution. *International Review of Cytology*, 141(C), 173–216. [https://doi.org/10.1016/S0074-7696\(08\)62066-5](https://doi.org/10.1016/S0074-7696(08)62066-5)

Xu, K., Kanno, M., Yu, H., Li, Q., & Kijima, A. (2011). Complete mitochondrial DNA sequence and phylogenetic analysis of Zhikong scallop *Chlamys farreri* (Bivalvia: Pectinidae). *Molecular Biology Reports*, 38(5), 3067–3074. <https://doi.org/10.1007/s11033-010-9974-8>

Yarra, T., Gharbi, K., Blaxter, M., Peck, L. S., & Clark, M. S. (2016). Characterization of the mantle transcriptome in bivalves: *Pecten maximus*, *Mytilus edulis* and *Crassostrea gigas*. *Marine Genomics*, 27, 9–15. <https://doi.org/10.1016/j.margen.2016.04.003>

Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Method Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol*, 11, R14.

Yu, F.-F., Wang, M.-F., Zhou, L., Gui, J.-F., & Yu, X.-Y. (2011). Molecular Cloning and Expression Characterization of *Dmrt2* in Akoya Pearl Oysters, *Pinctada martensii*. *Journal of Shellfish Research*, 30(2), 247–254. <https://doi.org/10.2983/035.030.0208>

Yu, Z., O'Farrell, P. H., Yakubovich, N., & DeLuca, S. Z. (2017). The Mitochondrial DNA Polymerase Promotes Elimination of Paternal Mitochondrial Genomes. *Current Biology*, 27(7), 1033–1039. <https://doi.org/10.1016/j.cub.2017.02.014>

Yuan, Y., Li, Q., Yu, H., & Kong, L. (2012). The complete mitochondrial genomes of six heterodont bivalves (tellinoidea and solenoidea): Variable gene arrangements and phylogenetic implications. *PLoS ONE*, 7(2). <https://doi.org/10.1371/journal.pone.0032353>

Yusa, Y., Breton, S., and Hoeh, W.R. (2013). Population Genetics of Sex Determination in *Mytilus* Mussels: Reanalyses and a Model. *J Hered* 104, 380–385.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., ... Wang, J. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*, 490(7418), 49–54. <https://doi.org/10.1038/nature11413>

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30(5), 614–620. <https://doi.org/10.1093/bioinformatics/btt593>

Zhang, N., Xu, F., & Guo, X. (2014). Genomic analysis of the Pacific oyster (*Crassostrea gigas*) reveals possible conservation of vertebrate sex determination in a mollusc. *G3 (Bethesda)*, 4(11), 2207–2217. <https://doi.org/10.1534/g3.114.013904>

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9. <https://doi.org/10.1186/1471-2105-9-40>

Zhang, Z. H., Jhaveri, D. J., Marshall, V. M., Bauer, D. C., Edson, J., Narayanan, R. K., ... Zhao, Q. Y. (2014). A comparative study of techniques for differential expression analysis on RNA-seq data. *PLoS ONE*, 9(8). <https://doi.org/10.1371/journal.pone.0103207>

Zhao, X., Yu, H., Kong, L., Liu, S., & Li, Q. (2014). Comparative transcriptome analysis of two oysters, *Crassostrea gigas* and *Crassostrea hongkongensis* provides insights into adaptation to hypo-osmotic conditions. *PLoS ONE*, 9(11). <https://doi.org/10.1371/journal.pone.0111915>

Zhou, Q., Li, H., Li, H., Nakagawa, A., Lin, J. L. J., Lee, E. S., ... Xue, D. (2016). Mitochondrial endonuclease G mediates breakdown of paternal mitochondria upon fertilization. *Science*, 353(6297), 394–399. <https://doi.org/10.1126/science.aaf4777>

Zieritz, A., & Aldridge, D. C. (2011). Sexual, habitat-constrained and parasite-induced dimorphism in the shell of a freshwater mussel (*Anodonta anatina*, Unionidae). *Journal of Morphology*, 272(11), 1365–1375. <https://doi.org/10.1002/jmor.10990>

Zimmerman, J.K. (1991). Ecological Correlates of Labile Sex Expression in the Orchid *Catasetum Viridiflavum*. *Ecology* 72, 597–608.

Zouros, E. (2000). The exceptional mitochondrial DNA system of the mussel family Mytilidae. *Genes & Genetic Systems*, 75(6), 313–318.

Zouros, E. (2013). Biparental Inheritance Through Uniparental Transmission: The Doubly Uniparental Inheritance (DUI) of Mitochondrial DNA. *Evol Biol* 40, 1–31.
<https://doi.org/10.1007/s11692-012-9195-2>

Zouros, E., Ball, A.O., Saavedra, C., and Freeman, K.R. (1994). An unusual type of mitochondrial DNA inheritance in the blue mussel *Mytilus*. *PNAS* 91, 7463–7467.