



HAL
open science

Structure et fonctionnement de la niche germinale chez un Lophotrochozoaire, l'huître creuse *Crassostrea gigas*.

Maëva Cherif-Feildel

► **To cite this version:**

Maëva Cherif-Feildel. Structure et fonctionnement de la niche germinale chez un Lophotrochozoaire, l'huître creuse *Crassostrea gigas*.. Biologie animale. Normandie Université, 2018. Français. NNT : 2018NORMC282 . tel-02163506

HAL Id: tel-02163506

<https://theses.hal.science/tel-02163506>

Submitted on 24 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

**Spécialité **PHYSIOLOGIE ET BIOLOGIE DES ORGANISMES - POPULATIONS -
INTERACTIONS****

Préparée au sein de l'Université de Caen Normandie

**Structure et fonctionnement de la niche germinale chez un
Lophotrochozoaire, l'huître creuse *Crassostrea gigas*.**

**Présentée et soutenue par
Maeva CHERIF-FEILDEL**

**Thèse soutenue publiquement le 20/12/2018
devant le jury composé de**

M. PASCAL FROMENT	Chargé de recherche HDR, INRA Centre Val de Loire	Rapporteur du jury
M. JEAN-JACQUES LAREYRE	Directeur de recherche, INRA-ENSA de Rennes	Rapporteur du jury
M. VIANNEY PICHEREAU	Professeur des universités, Université de Bretagne Occidentale - UBO	Président du jury
Mme KARINE VAN DONINCK	Professeur, Université de Namur	Membre du jury
Mme CLOTHILDE HEUDE-BERTHELIN	Maître de conférences HDR, Université Caen Normandie	Directeur de thèse
Mme KRISTELL KELLNER	Maître de conférences HDR, Université Caen Normandie	Co-directeur de thèse

**Thèse dirigée par CLOTHILDE HEUDE-BERTHELIN et KRISTELL KELLNER, Biologie
des organismes et écosystèmes aquatiques (Caen)**



UNIVERSITÉ
CAEN
NORMANDIE



Normandie de Biologie Intégrative,
Santé, Environnement



Avant-propos

Ces travaux de recherche ont été réalisés au sein de l'Ecole doctorale Normandie de Biologie Intégrative, Santé et Environnement (EdNBISE) dans l'Unité Mixte de Recherche « Biologie des Organismes et Ecosystèmes Aquatiques » (UMR BOREA) de l'Université de Caen Normandie sous la direction des Drs Kristell KELLNER et Clothilde HEUDE BERTHELIN. Cette thèse a été financée par une allocation de recherche du Ministère de L'Enseignement supérieur et de la Recherche.



Remerciements

Je tiens à exprimer mes remerciements à toutes les personnes qui ont participé, à titre professionnel ou personnel, à la réalisation de cette thèse.

Dans un premier temps, je tiens à exprimer toute ma gratitude et ma sympathie à mes directrices de thèse, le Dr Kristell Kellner et le Dr Clothilde Heude Berthelin qui m'ont encadrée tout au long de ces trois années. Merci pour votre soutien, votre disponibilité, votre bienveillance et la confiance que vous m'avez accordée en me choisissant pour ce sujet. Réaliser une thèse de doctorat n'est pas uniquement conditionné par le travail personnel et la motivation mais c'est également le résultat d'un partenariat avec des personnes très impliquées ainsi qu'un encadrement de qualité.

J'adresse mes sincères remerciements à tous les membres du jury, Au Dr Jean-Jacques Lareyre, pour avoir pris le temps d'évaluer mes travaux et me permettre de bénéficier de votre expertise en tant que rapporteur de cette thèse. Au Dr Pascal Froment, que j'ai rencontré lors de mon stage de Master 2 à Tours, pour me faire l'honneur d'avoir accepté d'évaluer ce travail en tant que rapporteur. Au Pr Karine Van Doninck et au Pr Vianney Pichereau pour avoir accepté de participer à ce jury en tant qu'examineurs et d'y apporter vos évaluations constructives.

Je remercie également tous les membres du laboratoire BOREA de Caen, Tout d'abord, je tiens à remercier le Pr Pascal Sourdain pour m'avoir accueillie au sein de son laboratoire ainsi que pour ses remarques constructives.

Merci également au Pr Pascal Favrel pour mon intégration au sein de son équipe et ses conseils avisés.

Je souhaite remercier le Dr Christophe Lelong pour son aide précieuse lors des expérimentations de biologie moléculaire, notamment l'hybridation in situ, une technique pouvant être capricieuse.

J'adresse également tous mes remerciements au Dr Guillaume Rivière pour ses réponses à mes nombreuses questions en biologie moléculaire, ses conseils et sa bonne humeur communicative.

J'exprime toute ma reconnaissance au Dr Aude Gautier pour son aide lors de ma recherche d'espèces hermaphrodites alternatifs.

Je remercie particulièrement Béatrice Adeline pour avoir partagé son savoir en histologie, pour tout l'excellent travail qu'elle a réalisé pour moi, incluant un nombre de coupes incalculable, et pour tous ces bons moments devant le microtome.

Merci également à Marie-Pierre Dubos, ingénieure, pour tes conseils en biologie moléculaire et en culture cellulaire.

Je tiens à remercier Myriam Tayou pour son aide et son implication dans l'ensemble des démarches administratives.

J'exprime mes sincères remerciements à tous les membres extérieurs du laboratoire qui ont facilité la réalisation de ce travail,

Au Dr Didier Goux et au Dr Nicolas Elie, du Centre de Microscopie Appliquée à la Biologie (CMABIO), pour le temps et l'aide précieuse qu'ils m'ont apportée ainsi que pour leur expertise.

A Marilyne Guillamin, technicienne sur le plateau de cytométrie en flux du centre François Baclesse, pour ses nombreux conseils techniques.

Un grand merci à tous les doctorants et post-doctorants rencontrés pour tous les bons moments. Un remerciement particulier à Lorane pour toutes les discussions et les soirées partagées qui ne manqueront par d'être renouvelées. Merci également à tous les autres, Louis, Charles, Aurore R., Quentin, Emilie, Julie, Aurore S., Laura, Floriane, Jérôme et Coralie et à ceux avec qui j'ai partagé le bureau, Valérie, Baptiste, Alexandre, Maxime. Un mot pour Chloé qui a partagé le bureau peu de temps mais qui est maintenant une amie.

Je n'oublie pas Benoît, Julien, Brice et Bertrand pour tous les brainstorming autour d'un café à la plateforme.

Merci à toutes les personnes de Tours avec lesquels j'ai eu l'honneur et le plaisir de travailler pendant mes années de licence et de master et pour la qualité de leur enseignement. Merci au Pr Emmanuel Moyse pour m'avoir transmis la passion des cellules souches. Merci également au Dr Marina Govoroun pour m'avoir encadré en Master 2 et m'avoir soutenu et aidé lors de ma recherche de doctorat. Je tiens également à remercier le Dr Pascal Froment, aujourd'hui rapporteur de cette thèse, et Sabine Alves pour vos encouragements qui m'ont aidé dans mes choix et m'ont permis d'accéder au poste de doctorante à Caen.

Je terminerai ces remerciements par une pensée toute particulière pour mes amis et ma famille qui m'ont accompagnée tout au long de ces années. Un infini merci à ma maman Sylvie, sans qui je ne serais pas arrivée jusque là. Enfin, merci à Aurélien qui m'a soutenue (et supportée) pendant toute la rédaction de ce manuscrit et à qui je promets plus de disponibilité lorsque cette thèse sera achevée.

Sommaire

Avant-propos

Répertoire des figures et des tableaux

Introduction générale	1
1. L'origine des Cellules Germinales Souches (GSC)	1
1.1. Spécification de la lignée germinale	1
1.2. Les Cellules Germinales Primordiales (PGC)	5
1.3. Origine atypique des GSC	7
1.4. Cas des hermaphrodites alternatifs.....	13
2. Caractéristiques des Cellules Germinales Souches (GSC)	15
2.1. Propriétés générales des GSC	15
2.2. Marqueurs moléculaires de l'état souche	16
2.2.1. Le facteur Sox2.....	17
2.2.2. Le facteur Oct4/POU5F1	19
2.2.3. Le facteur KLF4.....	21
2.2.4. Le facteur c-myc.....	22
2.2.5. Le facteur Nanog.....	24
2.3. Marqueurs moléculaires spécifiques des GSC.....	25
2.3.1. Vasa	25
2.3.2. Piwi.....	27
2.3.3. Nanos	28
2.4. Caractéristiques morphologiques des GSC.....	32
3. Le microenvironnement des GSC	34
3.1. Niche stromale ou épithéliale.....	35
3.2. Structure de la niche germinale chez différentes espèces.....	36
3.3. Voies de signalisation impliquées dans le fonctionnement de la niche germinale.....	39
3.3.1. Les Wnt/ β -caténine.....	40
3.3.2. La voie des TGF- β	41
3.3.3. La voie Notch.....	43
4. Les rôles clés du système insuline dans la reproduction	45
4.1. Le système insuline	45

4.1.1.	Les ligands de la superfamille des insulines.....	45
4.1.2.	Les récepteurs d'IRPs	48
4.1.3.	Les voies de signalisation	50
4.2.	Implication des IRPs dans le contrôle de la reproduction	51
4.2.1.	Les IRPs impliqués dans la maturation ovocytaire	51
4.2.2.	Rôle androgène	52
4.2.3.	IRPs et prolifération des cellules gonadiques	52
4.2.4.	IRPs et régulation de la niche germinale	54
4.2.5.	Intégration des signaux nutritionnels	57
5.	Le modèle biologique <i>Crassostrea gigas</i>	58
5.1.	L'intérêt économique	58
5.2.	La position phylogénétique	58
5.3.	Les Cellules Germinales Primordiales (PGC) chez l'huître	59
5.4.	La reproduction de l'huître creuse	60
5.4.1.	L'aire gonadique adulte et la gamétogenèse.....	60
5.4.2.	Structure de la gonade.....	65
6.	Objectifs de thèse	66
6.1.	Identification des GSC potentielles et de leur niche	66
6.2.	Etude évolutive et fonctionnelle du système insuline	66
6.3.	Effet de l'alimentation sur les cellules germinales précoces, implication du signal insuline	67
Chapitre I :	Identification des cellules germinales souches potentielles et de leur niche.....	69
1.	Introduction	71
1.1.	Marqueurs de pluripotence et cellules souches chez les mollusques	71
1.1.1.	Chez les céphalopodes.....	71
1.1.2.	Chez les gastéropodes.....	72
1.1.3.	Chez les bivalves.....	73
1.1.4.	Données acquises chez <i>Crassostrea gigas</i>	74
1.2.	Données acquises sur les GSC chez les mollusques.....	76
1.2.1.	Caractéristiques morphologiques des cellules germinales précoces	76
1.2.2.	Etude de l'activité mitotique pour l'identification de GSC	77
1.2.3.	Données moléculaires spécifiques des GSC.....	79
1.3.	Stratégies complémentaires pour identifier les GSC.....	84

1.4. Acteurs potentiels de la niche germinale chez les mollusques	84
1.4.1. Les cellules somatiques intragonadiques	85
1.4.2. Le système nerveux.....	86
1.4.3. Les voies de signalisation	87
1.5. Objectifs	88
2. Partie A – Recherche de cellules souches dans la gonade par des approches immunologiques et étude des proliférations goniales	90
2.1. Matériels et méthodes	90
2.1.1. Analyse des bases de données.....	90
2.1.2. Animaux	90
2.1.3. Préparations histologiques	90
2.1.4. Anticorps	91
2.1.5. Immunomarquages	92
2.1.6. Western blot	92
2.1.7. Expérimentation BrdU	93
2.2. Résultats.....	94
2.2.1. Marqueurs de pluripotence	94
2.2.2. Piwi1.....	96
2.2.3. BrdU	97
2.2.1. FMRFamide	98
2.3. Discussion	99
3. Partie B – Identification de GSC et progéniteurs potentiels dans leur microenvironnement	101
3.1. Matériels et Méthodes	101
3.2. Principaux résultats	102
3.2.1. Identification d’une population de cellules germinales précoces chez l’huître adulte	102
3.2.2. Vasa, un marqueur de cellules germinales précoces chez l’huître.....	102
3.2.3. Identification de deux types de cellules germinales précoces et caractérisation d’éléments de la niche germinale.....	104
3.3. Discussion	106

Publication : Cherif-Feildel M., Kellner K., Goux D., Elie N., Adeline B., Lelong C., Heude Berthelin C. 2018. Morphological and molecular criteria allow the identification of putative germ stem cells in a lophotrochozoan, the Pacific oyster *Crassostrea gigas*.....108

Bilan du chapitre I	128
Annexe – Cas d'un hermaphrodite simultané	130
Chapitre II : Etude évolutive et fonctionnelle du système insuline.....	131
1. Introduction	133
1.1. Caractérisation des ligands IRPs	133
1.1.1. Propriétés structurales.....	134
1.1.2. Les différentes structures des IRPs chez les mollusques	134
1.2. Identification des effecteurs du signal insuline	138
1.2.1. Le récepteur des IRPs chez les mollusques.....	138
1.2.2. Les voies de signalisation	139
1.3. Données fonctionnelles	140
1.4. Objectifs	141
2. Partie A – Caractérisation des IRPs et expressions tissulaires	142
2.1. Matériels et Méthodes	142
2.2. Principaux résultats	142
2.2.1. Identification des IRPs.....	142
2.2.2. Suivi des expressions.....	144
2.2.3. Expressions tissulaires de Cg-milp7	145
2.3. Discussion	146
Publication : Cherif--Feildel M., Heude Berthelin C., Adeline B., Rivière G., Favrel P., Kellner K. <i>in press</i> . Molecular evolution and functional characterisation of Insulin Related Peptides in Mollusc: contributions of <i>Crassostrea gigas</i> genomic and transcriptomic-wide screening.....	148
Data in Brief : Cherif--Feildel M., Heude Berthelin C., Rivière G., Favrel P., Kellner K. <i>in press</i> . Data for evolutive analysis of insulin related peptides in bilaterian species.....	192
3. Partie B – Identification d'effecteurs du signal insuline	184
3.1. Matériels et méthodes	184
3.1.1. Analyse des bases de données.....	184
3.1.2. Animaux et préparations histologiques	184
3.1.3. Extraction d'ARN et reverse transcription	184
3.1.4. PCR quantitative.....	185
3.1.5. Analyses statistiques	185
3.2. Résultats.....	186

3.2.1. Le récepteur CIR.....	186
3.2.2. Les voies de signalisation	190
3.3. Discussion	193
Bilan du chapitre II	196
Chapitre III : Effet de l'alimentation sur les cellules germinales précoces, implication du signal insuline.....	199
1. Introduction	201
1.1. Alimentation et impact sur la reproduction des mollusques	201
1.2. Les IRPs, lien entre statut nutritionnel et reproduction chez les mollusques.....	203
1.3. Objectifs	203
2. Matériels et Méthodes	204
2.1.1. Animaux	204
2.1.2. Conditionnement alimentaire.....	204
2.1.3. Extraction d'ARN et reverse transcription	204
2.1.4. PCR quantitative.....	205
2.1.5. Préparations histologiques	205
2.1.6. Immunomarquages	205
2.1.7. Histologie quantitative.....	206
2.1.8. Analyses statistiques	206
3. Résultats	207
3.1. Influence de l'alimentation sur le stock de GSC potentielles	207
3.1.1. Identification d'une population de cellules germinales précoces	207
3.1.2. Validation du conditionnement trophique : suivi des marqueurs Oyvlg et PCNA ..	210
3.1.3. Effet de l'alimentation sur le compartiment souche potentiel	210
3.2. Influence de l'alimentation sur les IRPs et CIR	212
4. Discussion	213
Bilan du chapitre III	216
Conclusion générale et perspectives	219
1. Identification des acteurs de la niche germinale	221
1.1. Localisation des cellules souches dans la gonade	222
1.1.1. GSC et niche germinale potentielles.....	222
1.1.2. Cellules souches pluripotentes	222

1.1.3. Cellules en prolifération.....	223
1.2. Hypothèse de fonctionnement des cellules souches gonadiques	223
1.2.1. Régulations des GSC potentielles.....	224
2. Etude de la signalisation des IRPs	224
2.1. Identification des IRPs et analyse phylogénétique.....	224
2.2. Effecteurs du signal.....	225
2.2.1. Fonctions potentielles des IRPs	225
2.3. Hypothèse de fonctionnement des IRPs dans la gamétogenèse	226
3. Impact de l'alimentation sur le compartiment souche	227
3.1. Hypothèse de fonctionnement de la niche germinale selon le statut nutritionnel	227
Bibliographie	228

Répertoire des figures et des tableaux

Introduction générale

Figure 1 : Schémas récapitulatifs de la formation des PGC (Cellules Germinales Primordiales) chez l'embryon.....	5
Figure 2 : Schéma récapitulatif des types de cellules souches identifiés chez <i>Botryllus schlosseri</i>	9
Figure 3 : Schéma des potentiels somatique et germinale des néoblastes chez les planaires telles que <i>Dugesia japonica</i>	10
Figure 4 : Hypothèse des Cellules Souches Primordiales (PriSC).....	12
Figure 5 : Hypothèse du fonctionnement des GSC chez les poissons hermaphrodites successifs.....	14
Figure 6 : Différents modes de division des GSC (Cellules Germinales Souches).....	16
Figure 7 : Représentation schématique de la protéine Sox2.....	17
Figure 8 : Représentation schématique de la protéine Oct4.....	19
Figure 9 : Représentation schématique de la protéine KLF4.....	21
Figure 10 : Représentation schématique de la protéine c-MYC.....	22
Figure 11 : Représentation schématique de la protéine Nanog.....	24
Figure 12 : Représentation schématique de la protéine Vasa.....	25
Figure 13 : Représentation schématique de la protéine Piwi.....	27
Figure 14 : Représentation schématique de la protéine Nanos.....	28
Figure 15 : Arbre phylogénétique simplifié des Métazoaires résumant la conservation des marqueurs de pluripotence et des marqueurs de GSC.....	30
Figure 16 : Ultrastructure de cellules souches germinales.....	33
Figure 17 : Schémas des deux types décrits de niche.....	35
Figure 18 : Schémas de niches germinales chez des espèces modèles.....	37
Figure 19 : Voie de signalisation simplifiée des Wnt/ β -caténine.....	40
Figure 20 : Voie de signalisation simplifiée des TGF- β	41
Figure 21 : Voie de signalisation Notch simplifiée.....	43
Figure 22 : Schéma de la structure conservée des ligands IRPs.....	45
Figure 23 : Représentation schématique du précurseur et de la forme mature du récepteur tyrosine kinase aux insulines.....	48
Figure 24 : Voies de signalisation PI3K/mTOR et MAPK simplifiées.....	50
Figure 25 : Représentations schématiques de l'implication du signal insuline dans le fonctionnement de différentes niches germinales.....	55
Figure 26 : Représentation schématique de la localisation des transcrits <i>oyvlg</i> lors du développement de <i>C. gigas</i>	59
Figure 27 : Organisation générale de l'aire gonadique chez l'huître creuse adulte.....	60
Figure 28 : Coupes histologiques d'aire gonadique au stade 0 de gamétogenèse chez <i>C. gigas</i>	61

Figure 29 : Coupes histologiques d'aire gonadique au stade 1 de gamétogenèse chez <i>C. gigas</i> .	62
Figure 30 : Coupes histologiques d'aire gonadique au stade 2 de gamétogenèse chez <i>C. gigas</i> .	63
Figure 31 : Coupes histologiques d'aire gonadique au stade 3 de gamétogenèse chez <i>C. gigas</i> .	64
Figure 32 : Ultrastructure d'un tubule gonadique chez <i>C. gigas</i> .	65
Tableau 1 : Représentations schématiques des principales formes matures des IRPs.	47

Chapitre I

Figure I-1 : Arbre phylogénétique simplifié des mollusques résumant l'identification et les fonctions potentielles des marqueurs de pluripotence chez certaines espèces de céphalopodes, gastéropodes et bivalves.	75
Figure I-2 : Ultrastructure de deux spermatogonies de type I chez <i>C. gigas</i> .	77
Figure I-3 : Immunofluorescence MCM6 et BrdU dans les pores évacuateurs en première gamétogenèse chez <i>C. gigas</i> .	78
Figure I-4 : Localisation du transcrit <i>oyvlg</i> dans les tubules gonadiques chez <i>C. gigas</i> à maturité sexuelle.	81
Figure I-5 : Localisation du transcrit <i>nanos</i> dans les tubules gonadiques chez <i>C. gigas</i> à maturité sexuelle.	83
Figure I-6 : Ultrastructure des cellules somatiques intragonadiques (ISC) chez les mâles <i>C. gigas</i> .	85
Figure I-7 : Localisation du transcrit <i>Og-TGFβ-like</i> dans les tubules gonadiques de <i>C. gigas</i> .	87
Figure I-8 : Localisation du transcrit et de la protéine Cg-mGDF dans les tubules gonadiques de l'huître creuse mâle à maturité sexuelle.	88
Figure I-9 : Séquences des facteurs de Yamanaka identifiés chez <i>C. gigas</i> .	94
Figure I-10 : Immunomarquages avec certains facteurs de pluripotence sur l'aire gonadique de <i>C. gigas</i> .	95
Figure I-11 : Alignement des isoformes Piwi1 identifiés chez <i>C. gigas</i> .	96
Figure I-12 : Test d'injection au rouge ponceau.	97
Figure I-13 : Immunomarquages BrdU dans l'aire gonadique de <i>C. gigas</i> .	98
Figure I-14 : Immunomarquages FMRFamide dans l'aire gonadique de <i>C. gigas</i> .	99
Figure I-15 : Localisation du transcrit <i>oyvlg</i> à tous les stades de gamétogenèse.	103
Figure I-16 : Détection de BMP2/4 (anticorps hétérologue) dans le cytoplasme des cellules somatiques intragonadiques.	105
Figure I-17 : Caractéristiques des cellules germinales précoces chez un hermaphrodite simultané.	130
Tableau I-1 : Tableau récapitulatif des anticorps utilisés et de leurs dilutions.	91

Chapitre II

Figure II-1 : Représentation schématique de la structure des pré-propeptides MIP chez <i>L. stagnalis</i>	135
Figure II-2 : Représentation schématique du pré-propeptide insuline chez <i>A. californica</i> . ..	136
Figure II-3 : Représentation schématique de la structure des pré-propeptides chez <i>L. gigantea</i>	137
Figure II-4 : Représentation schématique de la structure des récepteurs aux insulines identifiés chez les mollusques.	139
Figure II-5 : Représentations schématiques des IRPs identifiés chez <i>C. gigas</i>	142
Figure II-6 : Expression de Cg-milp7 mesurée par RT-qPCR.	145
Figure II-7 : Séquence complète du récepteur CIR.	187
Figure II-8 : Expression prédite <i>in silico</i> au cours du développement de <i>C. gigas</i>	188
Figure II-9 : Expressions du récepteur CIR.	189
Figure II-10 : Voies de signalisation de l'insuline et effecteurs conservés chez <i>C. gigas</i>	191

Tableau II-1 : Comparaison relative des mesures d'expression des transcrits insuline au cours du développement, dans les ganglions viscéraux, la glande digestive et aux stades de gamétogenèses mâle et femelle.

Tableau II-2 : Tableau récapitulatif des effecteurs des voies des IRPs.

Chapitre III

Figure III-1 : Observation des caractéristiques des cellules germinales précoces lors de l'initiation de la première gamétogenèse.	207
Figure III-2 : Identification d'une population de cellules germinales précoces chez les jeunes huîtres à l'initiation de leur première gamétogenèse.	208
Figure III-3 : Suivis des marqueurs Oyvlg et PCNA sur les huîtres juvéniles en fonction du conditionnement trophique.	210
Figure III-4 : Représentation graphique de l'approche d'histologie quantitative chez les huîtres en première gamétogenèse.	211
Figure III-5 : Expressions de certains effecteurs de l'insuline mesurées par qPCR sur les huîtres juvéniles.	213

Tableau III-1 : Tableau récapitulatif des couples d'amorces utilisés.

Tableau III-2 : Comparaison de la population de cellules germinales précoces chez des individus en première gamétogenèse et des individus en 2^{ème} et 3^{ème} gamétogenèse.

Tableau III-3 : Nombre de cellules germinales mesurées selon les conditions d'alimentation.

Liste des abréviations

ADN	Acide Désoxyribonucléique	eGC	Cellules Germinales précoces
AGO	Protéines Argonautes	ELH	Egg-Laying Hormone
APC	Adrenomatous Prolypoid coli	FGSC	Cellules Souches Germinales Femelles
ARN	Acide Ribonucléique	FMRFa	Phe-Met-Arg-Phe-NH ₂
siARN	Petits ARN interférants	FSH	Hormone Folliculo-Stimulante
miARN	microARN	Fzd	Frizzled
Akt	Protéine kinase B	Gbb	Glass bottom boat
ATP	Adénosine Triphosphate	GFP	Green Fluorescent Protein
Aund	Spermatogonies souches de type A	GSC	Cellules Germinales Souches
BCN	Bag Cells Neurons	GSK3β	Glycogen Synthase Kinase 3 β
bHLH-Zip	basic Helix-Loop-Helix Leucine Zipper	Hox	Gènes à homébox
BMP	Bone Morphogenic protein	HMG-CoA	hydroxy-méthyl-glutaryl coenzyme A
BrdU	BromodeoxyUridine	HMG	High-Mobility Group
CIR	<i>C. gigas</i> Insulin Receptor	IAG	Insulin-like Androgenic Gland peptide
CK1	Casein Kinase 1	IGF	Insulin Growth Factor
cvh	Chicken Vasa Homolog	IGF-R	Insulin Growth Factor Receptor
CySC	Cellules Souches somatiques du Cyste	ILPs	Insulin-Like Peptides
DAF	Dauer Formation family member	IPC	Insulin-Producing Cells
DDX4	DEAD-box RNA hélicase 4 Vasa protein	iPSC	induced Pluripotent Stem Cells
DILPs	Drosophila Insulin-Like Peptides	IRPs	Insulin Related peptides
DIR	Drosophila Insulin Receptor	ISC	Cellules somatiques intragonadiques
Djvlg	<i>D. japonica</i> vasa-like gene	KLF	Krüppel-Like Factor
Dpp	Decapentaplegic	LGC	Light green cells
DSL	Delta/Serrate/LAG-2	M6P	Mannose-6-Phosphate
DTC	Cellules de la pointe distale	MAPK	Mitogen Activated Protein Kinase

MCM6	Minichromosome maintenance protein 6	POU	Pit-1-Oct-Unc
MIPs	Molluscan Insulin-related Peptides	PriSC	Cellules Souches Primordiales
MYC	Myelocytomatosis oncogene	PI3K	Phosphatidylinositol 3 phosphate Kinase
Myvlg	<i>M. galloprovincialis</i> vasa-like gene	Rbpj	Recombination signal binding protein for immunoglobulin kappa j region
NECD	Notch Extracellular Domain	RCPG	Récepteurs couplés aux protéines G
NICD	Notch Intracellular Domain	RXFP	Récepteurs des relaxines
NIM	NOT1-interacting motif	Smad	Mother against decapentaplegic
Oct	Protéines Octamer	Sox	SRY-related homeobox
OSC	Cellules souches ovariennes	SSC	Cellules Souches Spermatogoniales
oyvlg	Oyster Vasa-like gene	TAZ	Tafazzin
PAZ	Piwi, Argonaute et Zwillie	TGF-β	Transforming Growth Factors β
PCNA	Proliferating Cell Nuclear Antigen	β-TrCP	β -Transducing Repeat Conducting E3 ubiquitin Protein ligase
PGC	Cellules Germinales Primordiales	Vvl	ventral veinless
pPGC	Cellules Germinales Primordiales potentielles	Wnt	Wingless integration site protein
piRNA	Piwi-interacting RNA	YAP	Yes-Associated Protein
PIWI	P-element Induced Wlmpy testis		
PKB	Protéine Kinase B		

Introduction générale

Introduction générale

La reproduction permet la transmission de l'information génétique d'une génération à l'autre. Il existe deux grands modes de reproduction, la reproduction sexuée et la reproduction asexuée retrouvée par exemple chez les cnidaires et les tuniciers. La reproduction sexuée implique le croisement des patrimoines génétiques parentaux lors du processus de fécondation (Raven *et al.*, 2007). Ce mode de reproduction est possible grâce à un cycle alternant la production de gamètes matures haploïdes et le processus de fécondation permettant le brassage génétique ainsi que la restauration de la ploïdie de l'espèce (Le Moigne et Foucrier, 2009). Afin de produire les gamètes, les Cellules Germinales Souches (GSC) sont recrutées et vont engendrer des gonies différenciées qui entreront dans une phase de mitoses successives suivi du déroulement de la méiose afin de se différencier pour former l'ensemble de la lignée germinale.

1. L'origine des Cellules Germinales Souches (GSC)

1.1. Spécification de la lignée germinale

Chez la plupart des Métazoaires à reproduction sexuée, la spécification de la lignée germinale a lieu lors du développement embryonnaire. Ainsi, les cellules de la lignée germinale dérivent d'un groupe particulier de cellules souches embryonnaires (Saffman et Lasko, 1999; Wylie, 1999). Les Cellules Germinales Primordiales (PGC) correspondent aux premières cellules spécifiées en cellules de la lignée germinale. Deux modes de spécification des PGC ont été décrits, le modèle préformistique et le modèle épigénétique.

Dans le cas du modèle préformistique ou préformation, la ségrégation des cellules germinales se produit très précocement lors du développement embryonnaire et repose sur l'héritabilité de certains déterminants maternels (Extavour et Akam, 2003; Wylie, 1999). Ces déterminants constituent le plasma germinal, région particulière localisée dans le cytoplasme du pôle végétatif de l'ovocyte (Ikenishi, 1998; Wylie, 1999). Ces déterminants sont présents de façon continue au cours des générations puisqu'ils sont transmis dans les cellules embryonnaires qui formeront, à l'âge adulte, la lignée germinale. De plus, ces déterminants sont nécessaires et suffisants à la spécification des PGC et, par conséquent, à la formation de la lignée germinale. Chez la drosophile, au premier stade de développement, l'embryon est constitué d'un syncytium de noyaux, les cellules n'étant pas encore délimitées

par une membrane. Le plasma germinal est localisé dans le cytoplasme à l'extrémité postérieure de l'embryon. Lorsque ce plasma germinal est prélevé de l'embryon ou détruit par irradiation, aucune cellule de la lignée germinale n'est formée et les individus sont stériles. Des expérimentations de transplantations ont été réalisées afin de mettre en évidence la capacité du plasma germinal à former des cellules de la lignée germinale (Illmensee et Mahowald, 1974). Ainsi, du cytoplasme est prélevé au pôle postérieur d'un embryon donneur puis transplanté au niveau du pôle antérieur d'un embryon receveur. La région antérieure a été choisie car elle est exclusivement somatique. Après transplantation, des cellules morphologiquement identiques à celles normalement localisées au pôle postérieur, sont formées dans la région somatique receveuse. De plus, certains individus receveurs ont pu former une lignée germinale dite « mosaïque », c'est-à-dire contenant le patrimoine génétique de l'individu receveur ainsi que celui de l'individu donneur. Ainsi, les cellules possédant les déterminants du plasma germinal peuvent engendrer la lignée germinale indépendamment de leur localisation dans l'embryon. Ce modèle préformistique présente l'avantage de proposer un mécanisme autonome, où la présence du plasma germinal est suffisante pour la ségrégation de la lignée germinale (Extavour, 2007; Illmensee et Mahowald, 1974). Les composants du plasma germinal ont également été étudiés chez la drosophile grâce à l'utilisation de mutants. Ainsi, dix gènes ont pu être identifiés (cappuccino, mago nashi, nanos, oskar, pumilio, spire, staufen, tudor, valois et vasa) dont les transcrits sont localisés au pôle postérieur de l'ovocyte. Pour chacun de ces gènes, les femelles mutantes engendrent une descendance stérile (Boswell et Mahowald, 1985; Golumbeski *et al.*, 1991; Lehmann, 1992). Parmi eux, seulement trois sont indispensables à la formation ectopique de cellules germinales à savoir, oskar, nanos et vasa. Oskar permet le recrutement du transcrit *nanos* et de la protéine Vasa. Les autres ARN et protéines maternels sont ensuite accumulés afin de former le plasma germinal (Ephrussi et Lehmann, 1992).

Dans le cas du modèle épigénétique ou induction, l'appartenance de certaines cellules à la lignée germinale est déterminée plus tardivement dans le développement. En l'absence du plasma germinal, cette spécification résulte de l'intégration d'un signal inducteur provenant des structures environnantes qui constituent une gonade primordiale (Extavour et Akam, 2003; Nikolic *et al.*, 2016). Ce signal inducteur permet l'expression de gènes spécifiques de la lignée germinale et réprime l'expression de gènes somatiques. De plus,

chez l'embryon de souris, si des cellules sont prélevées dans la région de la gonade primordiale puis transplantées dans une autre région exclusivement somatique, elles se développeront en cellules somatiques. A l'inverse, si des cellules sont prélevées dans une région strictement somatique puis transplantées dans la région de la gonade primordiale, ces dernières pourront être spécifiées en cellules de la lignée germinale (Extavour et Akam, 2003; Tam et Zhou, 1996). Chez la souris, l'un des facteurs impliqué dans le signal inducteur du modèle épigénétique est BMP4 (Bone Morphogenetic Protein 4). Afin de démontrer la nécessité de BMP4 dans la spécification de la lignée germinale, des mutants pour cet acteur ont été utilisés et les PGC chez l'embryon de souris ont été identifiées grâce à leur activité phosphatase alcaline (Lawson *et al.*, 1999). Ainsi, les embryons mutants pour *bmp4* ne présentent aucune cellule appartenant à la lignée germinale à l'inverse des individus non mutés. De la même façon, il a été montré que BMP8b était impliqué dans le nombre de cellules spécifiées en cellules de la lignée germinale. Cependant, ce facteur ne semble pas essentiel au signal inducteur du modèle épigénétique contrairement à BMP4 (Ying *et al.*, 2000). Lors du développement embryonnaire, le facteur BMP4 permettrait donc d'inhiber l'expression de gènes de la voie somatique et d'activer celle de gènes de la voie germinale, permettant l'induction de certaines cellules embryonnaires pluripotentes en cellules de la lignée germinale (Dudley *et al.*, 2007). Afin de mieux comprendre le fonctionnement du modèle épigénétique, une étude du transcriptome des PGC en comparaison avec celui des cellules environnantes chez l'embryon de souris a été entreprise. Pour cela, le tri des PGC et des cellules somatiques a été réalisé grâce à des marqueurs incluant BMP4. Les cellules exprimant *bmp4* ainsi que *hoxb1*, un marqueur somatique, ont été désignées en tant que cellules somatiques appartenant à l'environnement des futures cellules de la lignée germinale. A l'inverse, les cellules n'exprimant aucun de ces deux marqueurs ont été désignées en tant que PGC (Saitou *et al.*, 2002). L'analyse des expressions de ces deux groupes cellulaires a permis de mettre en évidence deux gènes fortement exprimés dans les PGC à savoir, *fragilis* et *stella*. Au moment de la spécification, l'expression de *stella* apparaît seulement dans les cellules exprimant déjà fortement *fragilis* et étant négatives pour *hoxb1*. De plus, aucun signal pour *fragilis* n'a pu être détecté chez des embryons de souris déficients pour *bmp4*, ce qui suggère l'influence de BMP4 sur *fragilis* qui permettrait l'induction de cellules embryonnaires en cellules de la lignée germinale (Saitou *et al.*, 2002; Saitou et Yamaji, 2012).

Les modèles préformistique et épigénétique ne sont pas incompatibles et peuvent agir conjointement comme cela a été mis en évidence dans différents groupes phylogénétiques, notamment chez des chondrichthyens, des arthropodes ou encore des mollusques (Extavour, 2007). Une analyse phylogénétique des processus de spécification de la lignée germinale a été réalisée chez 28 phylums. Elle a permis de mettre en évidence que le modèle épigénétique est le processus retrouvé chez la plupart des groupes alors que le modèle préformistique est moins fréquemment observé. Chez les Deutérostomiens non cordés tels que les Cnidaires ou les Porifères, à ce jour, aucune preuve de préformation n'a été montrée. Chez les Echinodermes, une ségrégation tardive des cellules germinales suggère une induction de la lignée germinale. Chez les cordés, le modèle préformistique semble être retrouvé chez les téléostéens alors que le modèle épigénétique semble être le processus existant chez les mammifères (Extavour et Akam, 2003). Chez les Protostomiens, le mode de spécification retrouvé est le modèle épigénétique pour la majorité des espèces. D'après cette analyse, tenant compte de l'ensemble des Métazoaires, la spécification de la lignée germinale est possible à n'importe quel moment du développement embryonnaire. Cependant, dans 23 phylums sur les 28 étudiés, l'observation de la première population de cellules appartenant à la lignée germinale est faite lors des étapes tardives de développement. Ces données suggèrent que le modèle épigénétique serait plus fréquent dans le règne animal et pourrait constituer le caractère ancestral (Extavour et Akam, 2003). Ce dernier pourrait être le mode de spécification utilisé par l'ancêtre commun des Deutérostomiens et des Protostomiens (Extavour, 2007).

1.2. Les Cellules Germinales Primordiales (PGC)

Les Cellules Germinales Primordiales (PGC) constituent la première population de cellules embryonnaires exclusivement germinales (Figure 1A). Ces PGC sont retrouvées durant la vie embryonnaire chez la majorité des vertébrés et invertébrés (Saffman et Lasko, 1999). Une fois spécifiées, les PGC migrent afin d'atteindre les territoires somatiques qui formeront la gonade chez l'individu adulte (Moore *et al.*, 1998). Après leur migration, les PGC se différencient en Cellules Germinales Souches (GSC) qui formeront la lignée germinale adulte (Extavour, 2007; Juliano *et al.*, 2010). Une hypothèse a été avancée concernant les cellules ayant la capacité d'être spécifiées en PGC. Ces cellules ont été nommées les PGC potentielles (pPGC) et constitueraient un type cellulaire morphologiquement identique aux autres cellules embryonnaires (Figure 1B). Cependant, ces dernières possèderaient la capacité d'engendrer à la fois des cellules somatiques et des PGC chez l'embryon (Extavour et Akam, 2003).

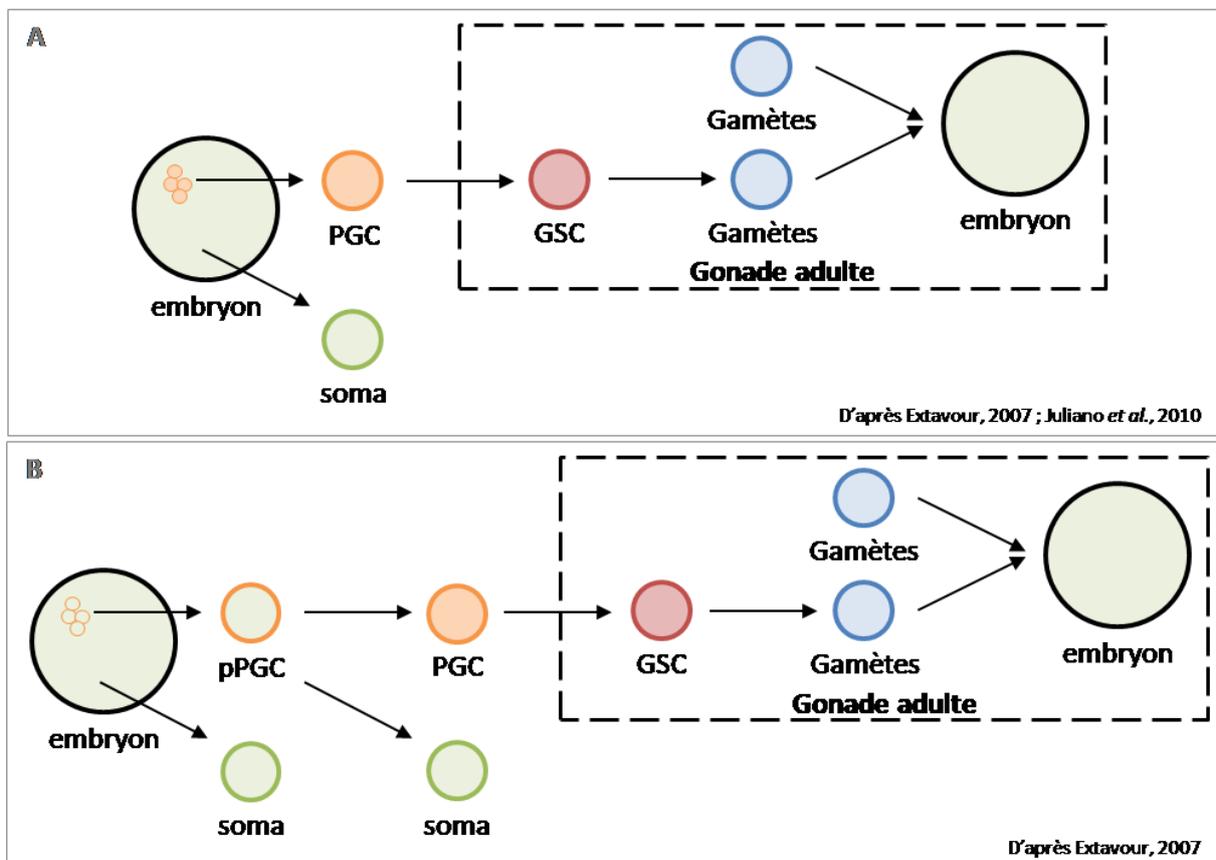


Figure 1 : Schémas récapitulatifs de la formation des PGC (Cellules Germinales Primordiales) chez l'embryon. A – Hypothèse courante concernant la formation des PGC d'après Extavour, 2007 et Juliano *et al.*, 2010. Les cellules à potentiel germinale sont spécifiées lors du développement embryonnaire et formeront les GSC (Cellules Souches Germinales) après migration et établissement dans la gonade adulte. B – Hypothèse des pPGC (PGC potentielles) qui présentent la capacité de se différencier en une cellule somatique et une PGC d'après Extavour, 2007. PGC : Cellules Germinales Primordiales ; GSC : Cellules Souches Germinales ; soma : Cellules somatiques ; pPGC : Cellules Germinales Primordiales potentielles.

Une fois spécifiées chez l'embryon, les PGC se distinguent des autres cellules embryonnaires par la présence de structures cytoplasmiques caractéristiques ressemblant à la région du plasma germinal. Ces structures sont retrouvées dans toutes les PGC, indépendamment du mode de spécification de la lignée germinale (Saffman et Lasko, 1999). Elles sont appelées granules germinaux, corps chromatoides ou encore nuage selon les auteurs et sont visibles en microscopie électronique à transmission où elles apparaissent comme des structures denses, régulièrement associées à des mitochondries. De plus, ces structures sont riches en ARN et en protéines nécessaires au développement de la lignée germinale tels que Vasa, Piwi ou encore Nanos (Gleason *et al.*, 2018; Voronina *et al.*, 2011). Les ARN de ces marqueurs ont été mis en évidence par hybridation *in situ* dans le plasma germinal de l'embryon ainsi que dans les granules germinaux des PGC chez plusieurs espèces (Harris et Macdonald, 2001; Hay *et al.*, 1988; Mosquera *et al.*, 1993).

La migration des PGC, afin d'atteindre la gonade adulte, est essentielle à leur survie et à leur prolifération et semble être contrôlée par une combinaison de mécanismes conservés, attractifs et répulsifs, provenant des tissus somatiques environnants (Marlow, 2015; Richardson et Lehmann, 2010; Wylie, 1999). Chez la drosophile, les premières étapes de migration des PGC sont dues aux remaniements morphologiques lors du développement qui amènent ces cellules au niveau de l'épithélium digestif. Suite à cela, les PGC changent de forme et la présence de pseudopodes est visible en microscopie électronique à transmission. De plus, des jonctions cellulaires peuvent être observées entre les PGC, formant alors un cluster, et l'épithélium digestif de l'embryon de drosophile. Les PGC passent ensuite au travers de cet épithélium digestif et migrent jusqu'au mésoderme gonadique qui formera la partie somatique de la gonade (Deshpete *et al.*, 2016; Jaglarz et Howard, 1995). Certains facteurs attractifs sont produits par les cellules de cette gonade primordiale. Chez la drosophile, le gène principalement impliqué est *hmgcr* codant pour l'hydroxy-méthyl-glutaryl-coenzyme A réductase (HMG-CoA). Cette enzyme est très exprimée dans les cellules somatiques de la gonade primordiale et la mutation du gène *hmgcr* entraîne de fortes perturbations dans la migration des PGC, observables par immunomarquages Vasa, les empêchant de rejoindre le mésoderme gonadique. De plus, lorsque l'expression ectopique de *hmgcr* est induite dans un autre tissu, les PGC migrent vers ce dernier (Van Doren *et al.*, 1998). La migration des PGC est également contrôlée par des facteurs répulsifs et, chez la drosophile, le gène impliqué est *wunen* qui est exprimé au niveau de l'épithélium digestif

chez l'embryon. En effet, chez les embryons déficients pour *wunen*, les PGC restent au niveau de l'épithélium digestif. De plus, quand l'expression de *wunen* est induite au niveau du mésoderme gonadique, très peu de PGC migrent jusqu'à ce tissu (Zhang *et al.*, 1997).

Chez *Caenorhabditis elegans*, les PGC ne subissent pas de réelle période de migration contrairement à la majorité des espèces modèles où ce processus a été étudié (Richardson et Lehmann, 2010). En effet, pendant le développement de *C. elegans*, les PGC se retrouvent localisées à l'endroit de la future gonade par les mouvements morphogénétiques de l'endoderme qui les poussent jusqu'au centre de l'embryon. Chez ce modèle, la gonade primordiale contient donc initialement les PGC et ce sont les précurseurs somatiques de cette gonade qui devront migrer. La migration des précurseurs somatiques a pu être suivie par microscopie confocale 4D, il a ainsi été montré qu'ils étaient initialement localisés dans la région antérieure de l'embryon. Ils migrent ensuite jusqu'à la région postérieure de l'embryon afin de rejoindre les PGC et de les envelopper par des extensions cytoplasmiques (Rohrschneider et Nance, 2013).

1.3. Origine atypique des GSC

Chez les espèces gonochoriques à reproduction sexuée, la spermatogenèse et l'ovogenèse se déroulent chez des individus différents qui conservent le même sexe tout au long de leur vie (Le Moigne et Foucrier, 2009). Dans ce cas, l'origine des GSC est embryonnaire. En effet, comme décrit précédemment, un groupe particulier de cellules embryonnaires est spécifié en cellules de la lignée germinale, nommées les PGC, selon les modèles préformistique et/ou épigénétique. Ces PGC migrent ensuite jusqu'à la future gonade et, une fois établies, elles perdent leur capacité de migration et se différencient en GSC qui engendreront les cellules germinales adultes (Figure 1) (Lin, 1997). Plusieurs termes sont utilisés pour définir la potentialité des GSC. En effet, elles peuvent être considérées pluripotentes, comme la plupart des cellules souches, car elles expriment les facteurs dits de pluripotence tels que *Nanog*, *Sox2* ou encore *Oct4* (Reik et Surani, 2015; Wylie, 1999; Yang *et al.*, 2018). Cependant, les GSC étant limitées aux types cellulaires de la lignée germinale qu'elles engendrent en se différenciant, elles peuvent également être considérées comme des cellules souches multipotentes. Certains les définissent même comme des cellules souches unipotentes car, chez les espèces gonochoriques, les GSC ne produisent qu'un seul

type cellulaire après différenciation, les gonies différenciées mâles ou femelles, selon le sexe de l'individu (Dadoune, 2013; Yuan et Yamashita, 2010).

Cette caractérisation des propriétés des GSC peut dépendre des espèces étudiées. En effet, chez les vertébrés supérieurs, les GSC permettent uniquement de produire soit les spermatozoïdes, soit les ovocytes. En revanche, certaines espèces possèdent des cellules souches adultes particulières ayant un potentiel somatique ainsi qu'un potentiel germinale (Rosner *et al.*, 2013; Shibata *et al.*, 1999; Solana, 2013). Les espèces possédant ces cellules souches bipotentielles sont capables de reproduction asexuée et de régénération. La reproduction asexuée permet de produire une descendance génétiquement identique à un parent seulement. Cette forme de reproduction n'implique pas de cellules germinales, donc pas de fécondation entre deux gamètes. Les processus regroupés dans le mode de reproduction asexuée présentent plusieurs types de fonctionnement.

Chez les tuniciers, groupe phylogénétiquement proche des vertébrés, les ascidies alternent entre reproduction sexuée et reproduction asexuée. Chez ces Métazoaires, le mode de reproduction asexuée utilisé s'appelle le bourgeonnement. Dans ce cas, le nouvel individu se forme à partir d'un bourgeon de cellules somatiques qui se développe sur une partie du corps du parent (Gasparini *et al.*, 2015; Raven *et al.*, 2007). Cet individu nouvellement formé doit grandir pour atteindre sa taille adulte. Après trois cycles de reproduction asexuée, l'ascidie mature peut se reproduire de façon sexuée (Gasparini *et al.*, 2015; Weissman, 2000). Chez une espèce d'ascidie, *Botryllus schlosseri*, la reproduction et les capacités de régénération sont assurées par un groupe de cellules souches totipotentes (TSC) qui peuvent se différencier en cellules somatiques souches ou en GSC (Figure 2). Ces TSC peuvent migrer via le système circulatoire de l'animal. Une étude de microsatellites par PCR a été réalisée sur différents tissus (bourgeon de cellules somatiques, lobes testiculaires, sang) afin d'identifier des marqueurs somatiques et germinaux. Le bourgeon et les lobes testiculaires contenaient chacun un marqueur spécifique et ces deux marqueurs étaient retrouvés dans les extraits sanguins (Stoner *et al.*, 1999). Ceci a permis de montrer que les TSC du système circulatoire pouvaient se différencier en lignée somatique ou en lignée germinale selon l'expression de marqueurs génétiques. Les futures GSC, nommées PGC-like chez *B. schlosseri*, sont positives au marquage germinale Vasa ce qui a permis de mettre en évidence leur migration (Figure 2). En effet, les PGC-like migrent à travers le système circulatoire jusqu'aux gonades en formation. Une fois établies, ces cellules envahissent les

gonades, localisées entre l'épiderme et l'épithélium péribranchial, et conduisent à la formation de la lignée germinale (Rosner *et al.*, 2013).

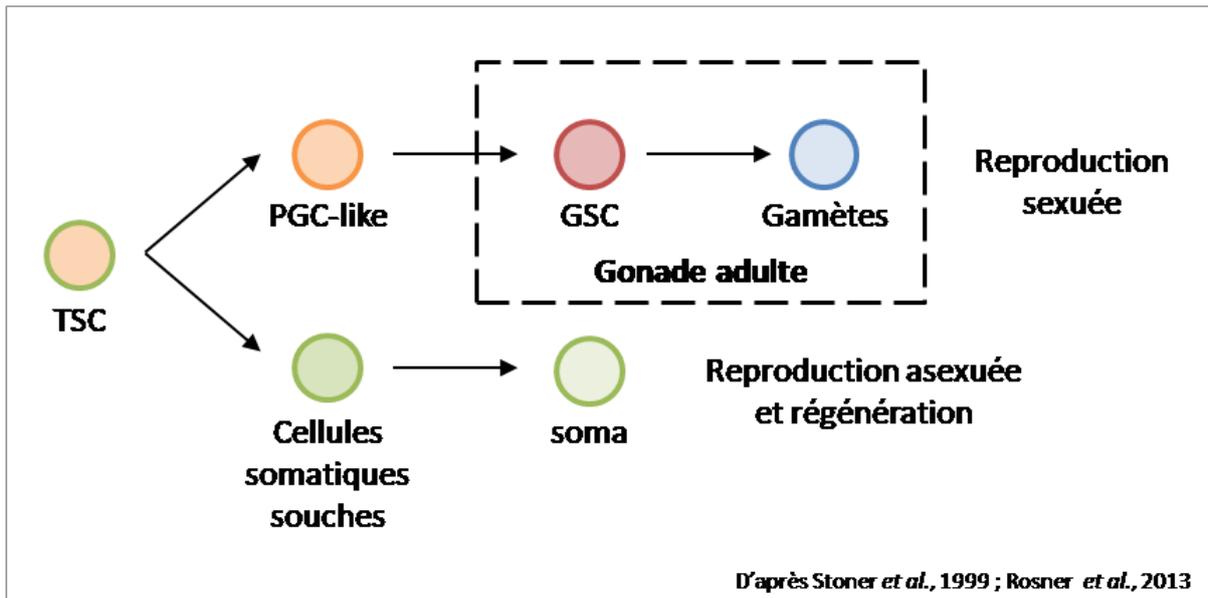


Figure 2 : Schéma récapitulatif des types de cellules souches identifiés chez *Botryllus schlosseri*. D'après Stoner *et al.*, 1999 et Rosner *et al.*, 2013. TSC : Cellules Souches Totipotentes ; PGC : Cellules Germinales Primordiales ; GSC : Cellules Germinales Souches ; soma : Cellules somatiques.

Chez les planaires, la présence de cellules souches bipotentielles a également été mise en évidence. Ces cellules souches, les néoblastes ont un potentiel somatique adapté au mode de reproduction asexuée ainsi qu'aux capacités de régénération qui caractérisent ces Lophotrochozoaires (Figure 3) (Baguña *et al.*, 1989; Shibata *et al.*, 1999).

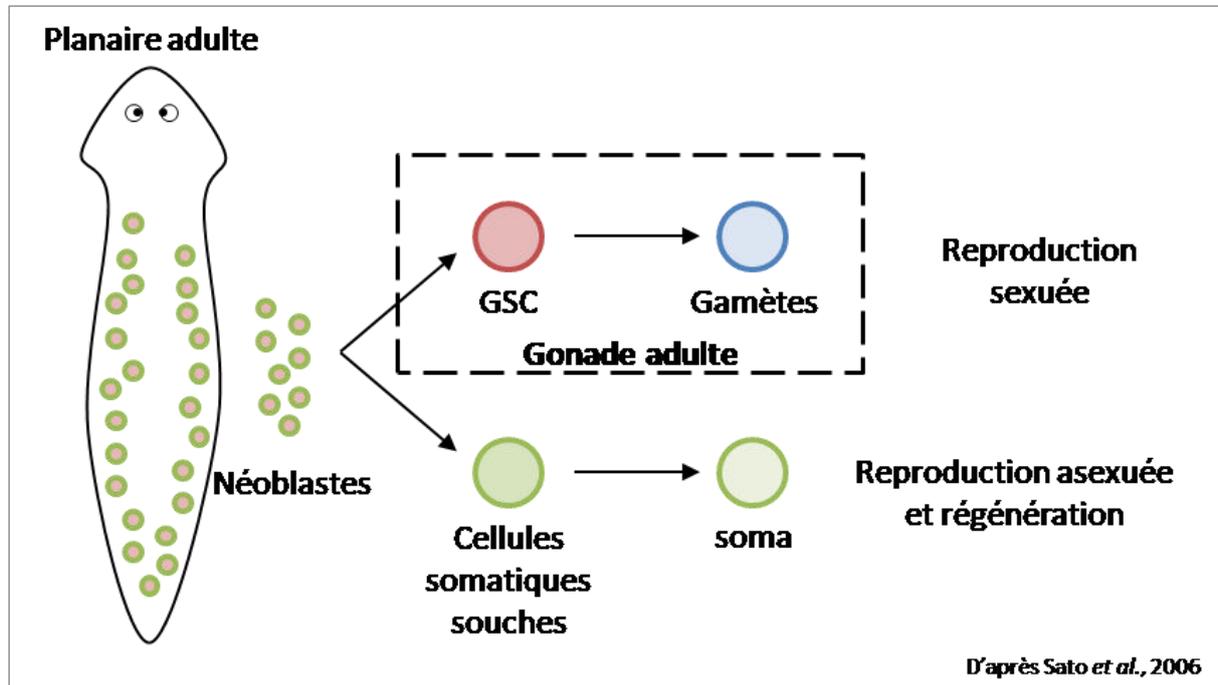


Figure 3 : Schéma des potentiels somatique et germinale des néoblastes chez les planaires telles que *Dugesia japonica*. D'après Sato *et al.*, 2006. GSC : Cellules Souches Germinales ; soma : Cellules somatiques.

Le mode de reproduction asexuée retrouvé chez certaines espèces de planaires est la scissiparité, ou scission, qui consiste en la division d'un organisme, chaque partie donnant un nouvel individu par une succession de mitoses (Raven *et al.*, 2007; Sato *et al.*, 2006). Une planaire peut être coupée en deux fragments et chacun de ces fragments peut régénérer la partie manquante en formant deux individus complets à partir des néoblastes, y compris la gonade contenant les GSC et la lignée germinale (Morita *et al.*, 1969). Après irradiation des planaires, les néoblastes ne sont plus capables de régénération et les individus meurent. Cependant, en injectant une fraction cellulaire enrichie en néoblastes dans un individu irradié, sa capacité de régénération se trouve restaurée et l'individu survit. Les néoblastes sont qualifiés de cellules totipotentes, puisqu'ils peuvent engendrer toutes les cellules d'un individu (Baguña *et al.*, 1989). Les néoblastes sont répartis tout au long du corps des planaires à l'exception de la région antérieure aux yeux. L'incorporation de BrdU (BromodeoxyUridine) a permis de mettre en évidence leur répartition ainsi que leur capacité à migrer, y compris dans la région antérieure de l'animale, afin d'assurer la régénération des

cellules somatiques et de la lignée germinale (Newmark et Alvarado, 2000). Cette capacité de migration rend difficile l'identification du microenvironnement spécifique de ces cellules souches.

Dans le groupe des planaires, le mode de reproduction sexuée est également représenté, certaines espèces ayant une reproduction sexuée saisonnée. Ces espèces se reproduisent de façon sexuée en automne/hiver puis de façon asexuée après dégénération des organes reproducteurs au printemps/été (Chong *et al.*, 2011; Kobayashi *et al.*, 1999). De plus, il a été montré que des espèces de planaires à reproduction strictement asexuée pouvaient être converties en individus à reproduction sexuée par une procédure particulière de nourrissage (Kobayashi *et al.*, 1999). En effet, la reproduction sexuée peut être induite chez des individus exclusivement asexués de l'espèce *Dugesia ryukyuensis* en les nourrissant avec des individus à maturité sexuelle de l'espèce *Bdellocephala brunnea* présentant un mode de reproduction sexuée. Après ce nourrissage, les individus *D. ryukyuensis* développent des organes sexuels différenciés (Kobayashi *et al.*, 1999). Ce processus de sexualisation, étudié sur des coupes histologiques, est possible grâce à la différenciation des néoblastes en cellules de la lignée germinale. Ceci souligne la totipotence des néoblastes, capables de former des cellules sexuelles même chez un individu n'en possédant pas initialement. Chez les planaires à reproduction sexuée, les individus sont hermaphrodites simultanés. L'hermaphroditisme simultané ou synchrone est à l'opposé du gonochorisme car, dans ce cas, les deux sexes sont retrouvés en même temps chez un même individu (Le Moigne et Foucrier, 2009). Cette situation est retrouvée par exemple chez la souche sexuée de la planaire *Schmidtea mediterranea*, où les lobes testiculaires sont distribués dorsalement le long de l'animal et la paire d'ovaires est localisée dans la région postérieure des ganglions céphaliques (Chong *et al.*, 2011; Newmark *et al.*, 2008). Bien que les gonades des deux sexes soient présentes chez une même planaire, aucun cas d'autofécondation n'a été observé. Chez ces planaires à reproduction sexuée, le potentiel germinal des néoblastes permet d'engendrer des GSC et de régénérer les lignées germinales mâle et femelle (Figure 3). L'étude du marqueur Nanos, connu pour être nécessaire aux GSC pour le développement de la lignée germinale chez plusieurs espèces, a révélé la présence de plusieurs sous-populations de néoblastes. En effet, l'hybridation *in situ* a montré que le transcrit *nanos* n'est exprimé que dans certains néoblastes localisés au niveau des gonades chez les planaires à reproduction sexuée. De plus, ce transcrit est également retrouvé dans certains

néoblastes chez des planaires asexuées lors du processus de sexualisation. Au début de ce processus de sexualisation, certains néoblastes répartis dans la planaire asexuée expriment *nanos*. Ils vont ensuite migrer pour, finalement, être restreints aux gonades nouvellement formées. Les cellules exprimant ensuite *nanos* sont identiques à celles observées dans la gonade des planaires sexuées (Sato *et al.*, 2006). Ces observations suggèrent fortement que les néoblastes peuvent engendrer deux types de cellules souches chez l'individu adulte à savoir, les cellules souches somatiques et les cellules souches germinales. Des études sont en cours concernant les différentes populations de néoblastes et leurs capacités de différenciation (Reddien, 2013; Sato *et al.*, 2006).

Ces cellules souches bipotentielles peuvent s'apparenter aux Cellules Souches Primordiales (PriSC) proposées par Solana (2013). Ce concept de PriSC a été énoncé afin de proposer une explication à l'origine des GSC chez des espèces telles que les planaires. Ces PriSC exprimeraient les marqueurs spécifiques des cellules germinales (Vasa, Piwi et Nanos) dès le développement embryonnaire où elles pourraient être spécifiées en PGC, cellules exclusivement germinales, mais également engendrer des cellules somatiques (Figure 4). De plus, ces cellules seraient conservées chez l'adulte où elles pourraient engendrer des GSC et également régénérer des tissus somatiques. Ceci permettrait les processus de reproduction sexuée ainsi que de reproduction asexuée et de régénération grâce aux potentiels germinale et somatique des PriSC, à l'instar des néoblastes (Solana, 2013).

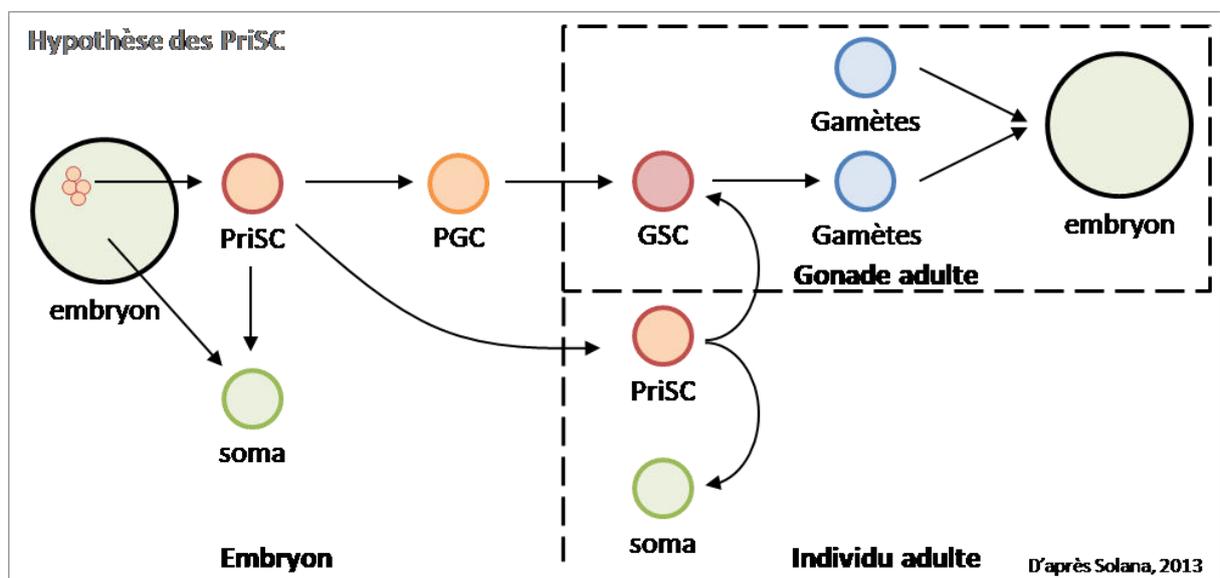


Figure 4 : Hypothèse des Cellules Souches Primordiales (PriSC). D'après Solana (2013). PriSC : Cellules Souches Primordiales ; PGC : Cellules Germinales Primordiales ; GSC : Cellules Germinales Souches ; soma : Cellules somatiques.

Trois types de PriSC ont pu être proposés. Les PriSC illimitées seraient actives tout au long de la vie de l'animal chez les espèces, telles que les planaires ou les ascidies, capables de reproduction sexuée, asexuée ainsi que de régénération. Les PriSC restreintes constituent le deuxième type. Leurs capacités seraient réduites à une région de l'animal comme c'est le cas chez certains annélides qui peuvent régénérer certaines parties de leur corps incluant la gonade. Enfin, le dernier type est appelé PriSC rudimentaires. Ces PriSC pourraient engendrer des cellules souches germinales et, éventuellement, des cellules souches somatiques. Ce type de PriSC pourrait, d'après Solana (2013), être retrouvé chez des espèces hermaphrodites alternatif à reproduction sexuée mais sans capacité de régénération décrite comme c'est le cas de l'huître *C. gigas* (Solana, 2013).

1.4. Cas des hermaphrodites alternatifs

L'étude de modèles avec différents modes de reproduction montre donc que les GSC peuvent être d'origine embryonnaire ou adulte. Un mode particulier de reproduction sexuée rencontré chez les Métazoaires est l'hermaphrodisme alternatif, aussi appelé successif ou séquentiel (Raven *et al.*, 2007; Saint-Dizier et Chastant-Maillard, 2014). Ce mode de reproduction implique que les individus peuvent changer de sexe au cours de leur vie. Plusieurs types d'inversions sexuelles sont rencontrés notamment chez certains poissons qui peuvent changer de sexe une fois au cours de leur vie (Saint-Dizier et Chastant-Maillard, 2014). Ainsi, chez les espèces hermaphrodites successives protérogynes, l'inversion sexuelle se fait du sexe femelle vers le sexe mâle. A l'inverse, les espèces hermaphrodites successives protandres passent du sexe mâle au sexe femelle.

L'anguille de rizière *Monopterus albus*, est une espèce protérogyne, elle passe donc du sexe femelle au sexe mâle. Ce changement de sexe implique un remaniement de la gonade ainsi que des lignées cellulaires somatiques et germinales. Chez cette espèce, pendant l'embryogenèse, les PGC colonisent la gonade primordiale où elles s'établissent et sont entourées pas des cellules somatiques accompagnatrices. La gonade primordiale se différencie d'abord en ovaire et les PGC engendrent les GSC qui se différencient en cellules de la lignée germinale femelle. Deux cavités sont formées dans l'ovaire. Les coupes histologiques ont permis d'observer que l'une de ces deux cavités est localisée entre les cellules germinales et les tissus somatiques. Il a été suggéré que cette cavité pourrait servir à la prolifération et à la réorganisation des cellules somatiques accompagnatrices au moment

du changement de sexe chez l'adulte. Chez cette espèce, après la ponte, les tissus ovariens régressent et se transforment en testicule (Cai *et al.*, 2017; He *et al.*, 2014; Zhu *et al.*, 2016). Chez le mérrou marbré, *Epinephelus microdon*, le mode de reproduction est également hermaphrodite protérogyne. Des cellules germinales possédant les caractéristiques morphologiques des GSC, telles que la présence d'un nuage associé à un groupe de mitochondries, ont pu être observées en microscopie électronique à transmission (Bruslé-Sicard *et al.*, 1992). Ces cellules ont été retrouvées dans la gonade indifférenciée puis, en moins grand nombre, dans les ovaires et les testicules. La présence de ces cellules de façon continue dans les gonades, y compris lors de l'inversion sexuelle, suggère la présence d'un stock permanent de GSC chez l'adulte capables de générer des ovogonies et des spermatogonies différenciées à partir d'un même type cellulaire (Figure 5A).

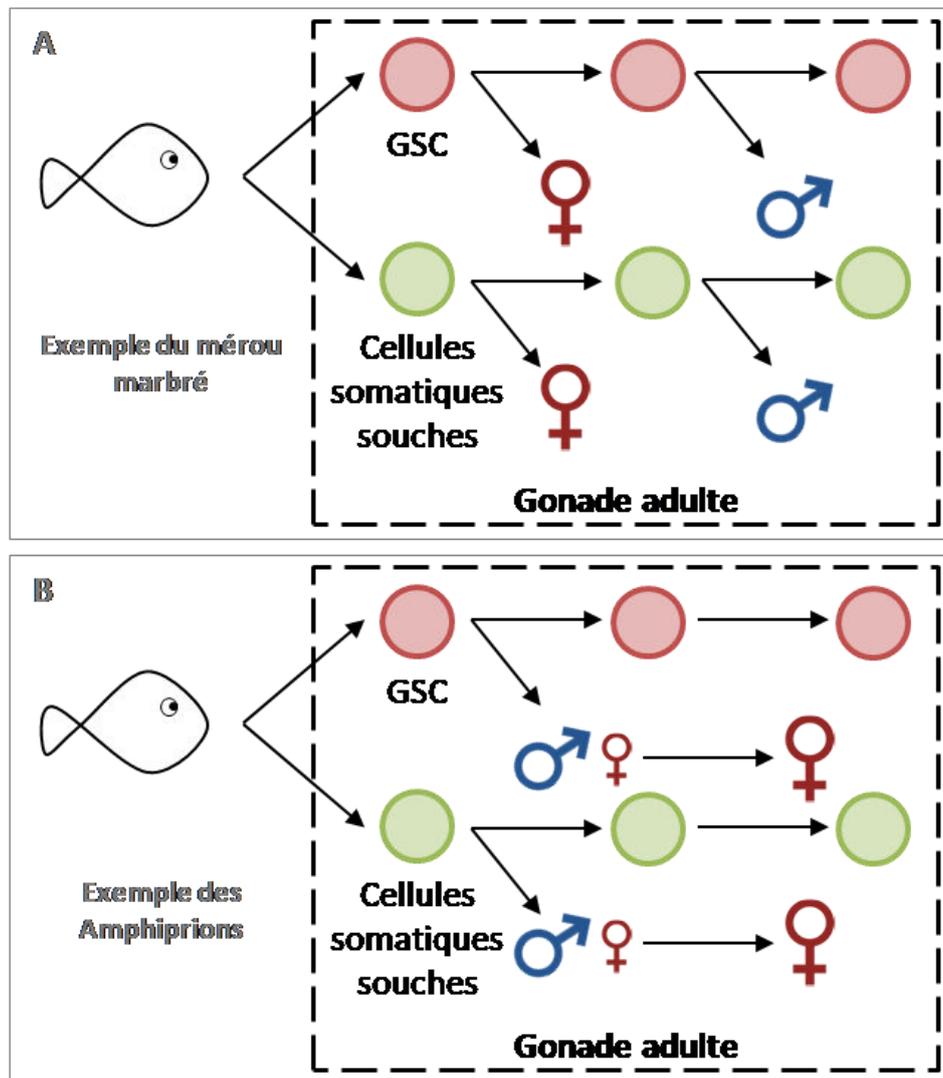


Figure 5 : Hypothèse du fonctionnement des GSC chez les poissons hermaphrodites successifs. A – Exemple du mérrou marbré, hermaphrodite successif protérogyne. B – Exemple des espèces Amphiprion, hermaphrodite successif protandre. GSC : Cellules Germinales Souches.

Chez les Amphiprions, espèces protandres qui passent donc du sexe mâle au sexe femelle, comme chez les espèces protérogynes, le tissu gonadique est remanié lors du changement de sexe. Chez ces poissons, le changement de sexe peut être induit, au moins en partie, par des facteurs sociaux. Ils vivent en groupe avec une femelle dominante, un mâle dominant et plusieurs juvéniles mâles immatures. Lorsque la femelle meurt, le mâle dominant change de sexe afin de prendre la place de la femelle et le plus grand des juvéniles devient un mâle mature dominant (Casas *et al.*, 2016). Chez ces poissons, la gonade mâle est en fait un ovotestis, constitué de tissus testiculaires et ovariens. Les tissus ovariens restent immatures jusqu'au changement de sexe, où les tissus testiculaires dégénèrent pour ne laisser que les tissus ovariens matures (Casadevall *et al.*, 2009). L'origine cellulaire des lignées germinales mâle et femelle n'est pas encore pleinement comprise chez cette espèce. Lorsqu'un inhibiteur d'aromatase est ingéré par les poissons femelles matures, un changement de sexe peut de nouveau avoir lieu, les tissus testiculaires néoformés ont été observés en histologie (Nakamura *et al.*, 2015). Ce résultat suggère que des cellules souches germinales à bipotentialité mâle et femelle existeraient chez les Amphiprion adultes (Figure 5B).

Enfin, quelques rares espèces animales présentent la capacité de changer plusieurs fois de sexe au cours de leur vie. Elles sont dites hermaphrodites successives à inversions multiples, c'est le cas du poisson épervier nain *Cirrhitichthys falco* ou encore de l'huître creuse *Crassostrea gigas*. Lors des inversions sexuelles, la gonade est remaniée et les tissus ovariens ou testiculaires peuvent dégénérer afin de donner une gonade de l'autre sexe (De Mitcheson et Liu, 2008; Saint-Dizier et Chastant-Maillard, 2014). Par conséquent, la question se pose quant à l'origine et aux mécanismes de différenciation des GSC chez ces espèces présentant une reproduction pouvant, selon les cycles de reproduction, produire des lignées germinales et somatiques mâles ou femelles.

2. Caractéristiques des Cellules Germinales Souches (GSC)

2.1. Propriétés générales des GSC

Quelque soit leur origine, les GSC adultes sont définies par trois caractéristiques à savoir : la prolifération illimitée, l'auto-renouvellement et la capacité à se différencier en gonies différenciées pour générer tous les types cellulaires de la lignée germinale afin de garantir la reproduction des individus (Lin, 1997; Yuan et Yamashita, 2010).

Le critère de prolifération illimitée se traduit par le fait que les GSC constituent un stock pérenne pour la production de gamètes tout au long de la période de reproduction de l'individu adulte (Figure 6) (Lin, 1997). Ce critère va de pair avec la capacité d'auto-renouvellement. En effet, les GSC peuvent renouveler leur stock par division asymétrique, c'est-à-dire qu'une cellule mère donnera deux cellules filles différentes entre elles. L'une de ces cellules filles sera identique à la cellule mère, donc restera GSC, tandis que l'autre constituera une gonie différenciée (Figure 6A). Les GSC peuvent également se diviser de façon symétrique. En effet, la cellule mère engendrera deux cellules filles identiques entre elles (Morrison et Kimble, 2006; Sada et Tumber, 2013). Ces deux cellules filles peuvent être identiques à la cellule mère ou être des gonies différenciées (Figure 6B). Ces dernières, grâce au processus de gamétogenèse, se multiplieront par mitoses successives et entreront en méiose afin de produire les différents types cellulaires de la lignée germinale.

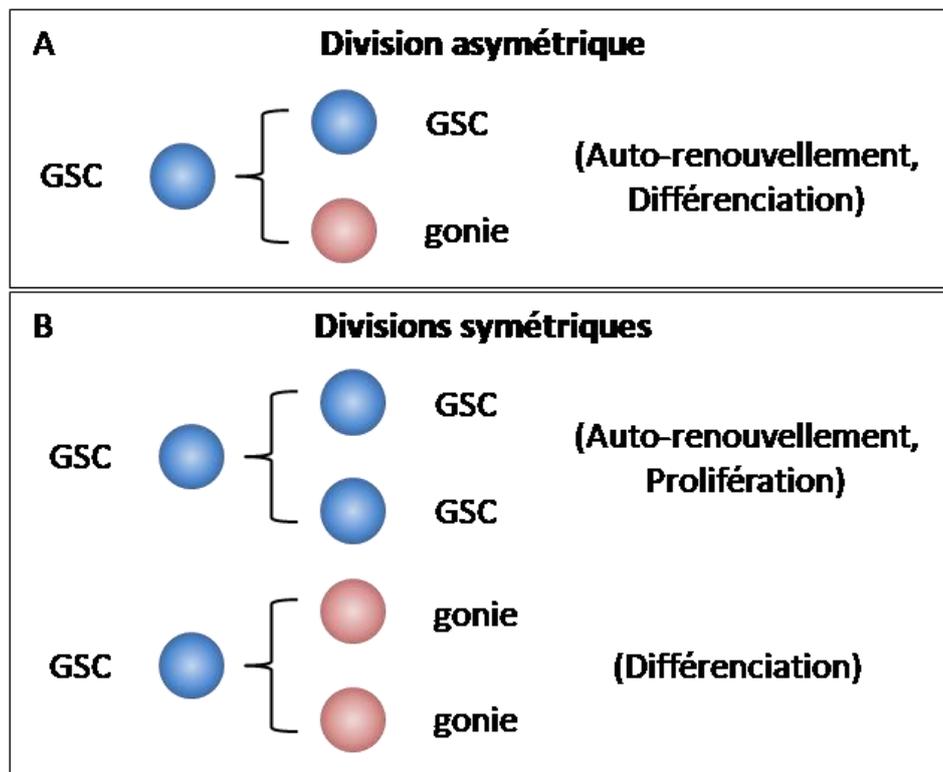


Figure 6 : Différents modes de division des GSC (Cellules Germinales Souches). A – Les GSC peuvent se diviser de façon asymétrique permettant leur auto-renouvellement et leur différenciation en gonies après une division cellulaire. B – La division symétrique des GSC leur permet soit de s'auto-renouveler et de proliférer soit de se différencier. D'après Sada et Tumber, 2013. GSC : Cellules Souches Germinales.

2.2. Marqueurs moléculaires de l'état souche

Le fonctionnement des GSC et, plus largement des cellules souches, est régulé par différents acteurs moléculaires. Chez les Métazoaires, les acteurs moléculaires bien connus

pour leur rôle dans la régulation de la pluripotence des cellules souches ont été en premier lieu identifiés chez les vertébrés et sont conservés au sein du règne animal. Ces acteurs sont Sox2, Oct4, KLF4 et c-myc, aussi appelés facteurs de Yamanaka (Liu *et al.*, 2008; Takahashi et Yamanaka, 2006; Takahashi *et al.*, 2007). Ces facteurs ont été sélectionnés parmi 24 gènes candidats exprimés par les cellules souches embryonnaires. La capacité de pluripotence dans des cellules somatiques différenciées peut être provoquée par l'induction de l'expression des facteurs de Yamanaka. Ces cellules différenciées génétiquement transformées peuvent ainsi posséder le potentiel souche et sont nommées les induced Pluripotent Stem Cells (iPSC). Nanog, bien qu'étant un facteur de pluripotence exprimé par les cellules embryonnaires chez la souris, ne semble pas indispensable à la formation d'iPSC (Takahashi et Yamanaka, 2006; Takahashi *et al.*, 2007).

2.2.1. Le facteur Sox2

La famille des Sox (SRY-related High-Mobility Group box) regroupe des acteurs impliqués dans le développement embryonnaire ainsi que la régulation des destinées cellulaires (Kamachi *et al.*, 2000). Les protéines codées par ce groupe de gènes possèdent un domaine HMG (High-Mobility Group) bien conservé entre les espèces et leur permettant de se lier à l'ADN ainsi qu'un domaine d'activation proche de la partie C-terminal (Figure 7) (Kamachi *et al.*, 2000).

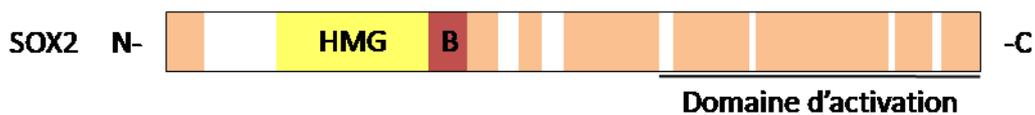


Figure 7 : Représentation schématique de la protéine Sox2. En jaune, le domaine HMG (High-Mobility Group) permettant la liaison à l'ADN. En rouge, le domaine commun au groupe B. En orange, les séquences conservées entre les protéines Sox1, Sox2 et Sox3. Le domaine d'activation est localisé dans la partie C-terminale. D'après Kamachi *et al.*, 2000.

Plus de 20 facteurs de transcription Sox ont été identifiés chez les vertébrés. Ces protéines ont été réparties dans 7 groupes nommés A à G. Le groupe B est divisé en 2 sous-groupes, celui des « activateurs » nommé B1 et celui des « répresseurs » nommé B2. Les protéines Sox1, Sox2 et Sox3 appartiennent au sous-groupe B1 (Kamachi *et al.*, 2000) et leur rôle dans la différenciation du tissu germinatif embryonnaire a été montré (Zhang et Klymkowsky, 2007). Concernant Sox1, son expression mesurée par PCR quantitative suggère qu'il aurait un rôle dans la différenciation des cellules souches neurales ainsi que des cellules cancéreuses (Ahmad *et al.*, 2017). Bien que les voies moléculaires précises liées à la fonction de Sox2 ne soient pas encore pleinement décryptées, il est certain que ce facteur de

transcription agit en étroite collaboration avec Oct4 et Nanog afin de réguler la pluripotence cellulaire (Akberdin *et al.*, 2018; Avilion *et al.*, 2003; Western *et al.*, 2005). En effet, chez l'embryon de souris, Sox2 est un marqueur de lignée cellulaire multipotente (Avilion *et al.*, 2003). Chez plusieurs mammifères, l'expression de Sox2 est détectée dans l'embryon par hybridation *in situ* à partir du stade morula et perdure jusque dans les cellules germinales (Avilion *et al.*, 2003; Western *et al.*, 2005). Sox3 apparaît impliqué dans la spermatogenèse chez la souris (Raverot *et al.*, 2005). Son expression est retrouvée dans les cellules souches spermatogoniales et sa perte chez des mutants pour Sox3 entraîne un dysfonctionnement lors de la spermatogenèse se traduisant par une stérilité de l'animal (Raverot *et al.*, 2005).

Les facteurs de transcription Sox ont également pu être identifiés chez certaines espèces d'invertébrés. En effet, une analyse phylogénétique a permis d'identifier des facteurs Sox appartenant au groupe B chez de nombreux Bilatériens incluant les Ecdysozoaires et les Lophotrochozoaires (Zhong *et al.*, 2011). D'après ces données, il semble que l'existence du groupe SoxB soit spécifique des Métazoaires. De plus, d'après l'étude du domaine HMG, la sous-division en groupes B1 et B2 est également retrouvée chez de nombreux Bilatériens. Cette sous-division s'appuie sur deux acides aminés du domaine HMG qui permettent la distinction entre les SoxB1 et les SoxB2. Dans le groupe B1, une arginine et une thréonine sont retrouvées en position 2 et 78 de l'homéodomaine respectivement alors que, dans le groupe B2, une histidine et une proline sont retrouvées aux mêmes positions (Zhong *et al.*, 2011).

Chez la drosophile, un seul gène a été identifié pour le groupe B1. Ce gène code pour la protéine Dichaete dont la séquence présente 88% d'identité avec la séquence Sox2 de souris. Le phénotype mutant engendré par la délétion du gène codant pour Dichaete peut être restauré par l'expression du Sox2 de souris chez l'embryon de drosophile (Soriano et Russell, 1998). Le rôle de Dichaete a été mis en évidence dans la neurogenèse lors du développement embryonnaire de la drosophile. De même que chez la souris, l'interaction entre les protéines Sox et Oct est conservée chez *D. melanogaster* (Soriano et Russell, 1998).

Chez le Plathelminthe *Echinococcus multilocularis*, sept gènes appartenant à la famille des Sox ont été retrouvés (Cheng *et al.*, 2017). Parmi eux, l'un code pour un homologue de la protéine Sox2 dont le domaine HMG apparaît fortement conservé par rapport aux séquences de mammifères. Chez cette espèce, se reproduisant de façon asexuée, des cellules souches similaires aux néoblastes de planaires ont été caractérisées et sont appelées

cellules germinatives. Cheng *et al.* (2017) ont utilisé le système des iPSC en substituant le facteur Sox2 de souris par celui de *E. multilocularis*, nommé EmSox2. EmSox2 est alors exprimé dans les cellules germinatives en prolifération. Ainsi, il semblerait que ce facteur soit impliqué dans la régulation de ces cellules germinatives. Les auteurs ont également montré que le facteur EmSox2 pouvait permettre l'induction de cellules somatiques en iPSC. Ces données montrent que la structure et la fonction de Sox2 sont conservées chez cette espèce (Cheng *et al.*, 2017).

Certains facteurs Sox appartenant au groupe B ont également été identifiés chez des mollusques (Yu *et al.*, 2017). Les expressions de ces facteurs ont été localisées lors du développement, dans les structures neurales ainsi que dans la gonade (Focareta et Cole, 2016; Le Gouar *et al.*, 2004; O'Brien et Degnan, 2000). A ce jour, le rôle de ces facteurs chez les mollusques n'est pas connu.

2.2.2. Le facteur Oct4/POU5F1

Les protéines Octamer (Oct) forment un groupe conservé de 11 facteurs de transcription impliqués dans de multiples processus biologiques. Ces protéines font partie d'une famille plus large, la famille des POU (Pit-1-Oct-Unc), caractérisée par la présence d'un domaine spécifique POU_S et d'un homéodomaine POU_H impliqués dans la liaison à l'ADN et les interactions entre protéines (Figure 8) (Zhao, 2013). Le facteur de transcription Oct4, aussi connu sous le nom de POU5F1, forme un complexe protéique avec Sox2 afin de réguler l'expression de leurs gènes cibles. Il agit en se liant à la région régulatrice de Sox2 afin de maintenir la pluripotence des cellules indifférenciées notamment des cellules souches embryonnaires (Tomioka *et al.*, 2002). Chez l'humain et la souris, ce marqueur est exprimé dans les cellules souches embryonnaires ainsi que dans les cellules germinales précoces et son expression diminue lorsque ces cellules se différencient (Zhao, 2013). De plus, Oct4 est impliqué dans la pluripotence ainsi que dans l'auto-renouvellement des cellules souches spermatogoniales chez plusieurs espèces de poissons (Lacerda *et al.*, 2014).

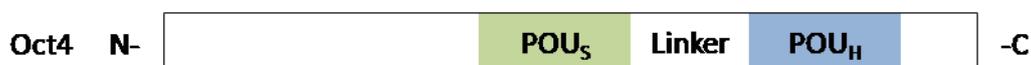


Figure 8 : Représentation schématique de la protéine Oct4. En vert, le domaine POU_S caractéristique de la famille POU. En bleu, le domaine POU_H permettant la liaison à l'ADN et les interactions entre protéines. Les deux domaines sont reliés par un domaine nommé « Linker » impliqué dans la spécificité de la liaison à l'ADN. D'après Zhao, 2013.

Les facteurs de transcription Oct apparaissent conservés chez la drosophile. En effet, plusieurs gènes codants pour des protéines contenant les domaines conservés de la famille

des POU ont été identifiés. A l'instar des vertébrés, ces acteurs sont exprimés par les cellules embryonnaires. L'utilisation de mutants a permis de mettre en évidence que les gènes codant pour des protéines POU étaient indispensables au développement embryonnaire (de Celis *et al.*, 1995; Lloyd et Sakonju, 1991). En effet, l'observation des embryons mutants pour ventral veinless (vvl) a mis en évidence plusieurs défauts dans le développement notamment au niveau des connexions du système nerveux central.

Chez la planaire *Schmidtea mediterranea*, un homologue potentiel de Oct4, nommé Smed-POU-P1, a pu être identifié par séquençage d'une population de néoblastes en prolifération et de cellules différenciées (Önal *et al.*, 2012). Ainsi la protéine Oct4 a pu être prédite et les domaines conservés POU_S et POU_H ainsi qu'une région apparentée au linker ont été identifiés. Les expressions des marqueurs de cellules souches identifiés chez les vertébrés, incluant Oct4, ont pu être observées. Pour cela, deux populations cellulaires ont été triées par cytométrie en flux selon leur sensibilité aux irradiations. En effet, la population de néoblastes étant très sensible aux rayons X, elle disparaît presque complètement après irradiation. Ainsi, en comparant les populations cellulaires chez des planaires irradiées avec celles chez des individus contrôles, les néoblastes peuvent être identifiés (Hayashi *et al.*, 2006). Il est apparu que Smed-POU-P1 était surexprimé dans la population de néoblastes en comparaison avec celle contenant les cellules différenciées (Önal *et al.*, 2012). Bien que d'autres expérimentations soient nécessaires afin de décrire le rôle de Smed-POU-P1 chez *S. mediterranea*, ces résultats suggèrent une conservation de son implication dans le maintien des cellules souches. Chez les mollusques, plusieurs facteurs de transcription appartenant à la famille des POU ont également été identifiés sur la base des domaines conservés. A l'instar de Sox2, leurs expressions a été analysées durant le développement ainsi que dans les tissus adultes, notamment dans les ganglions nerveux et la gonade (Gao *et al.*, 2016; O'Brien et Degnan, 2000, 2002; Wollesen *et al.*, 2014). Il n'existe pas d'étude fonctionnelle de ces facteurs chez les mollusques.

D'après ces données, il semblerait que les membres de la famille POU soient exprimés dans les cellules embryonnaires chez la plupart des groupes phylogénétiques étudiés. Cependant, des études complémentaires seront nécessaires afin de compléter l'identification et d'étudier le rôle de ces acteurs chez les Lophotrochozoaires. De plus, dans ce phylum, l'implication des protéines Oct dans la régulation du fonctionnement des cellules souches germinales reste encore méconnue.

2.2.3. Le facteur KLF4

Les facteurs Krüppel-like (KLF) appartiennent à une famille de facteurs de transcription contenant plus de 20 membres répartis en sept groupes, nommés de A à G, chez les vertébrés (Netan et Yang, 2009; Pei et Grishin, 2015). Ces facteurs présentent généralement trois domaines en doigt de zinc leur permettant de se lier aux régions régulatrices de leurs gènes cibles. Ils possèdent également différents domaines d'activation ou de répression de la transcription permettant les interactions protéines-protéines et la liaison à l'ADN (Figure 9) (Ghaleb et Yang, 2017; Netan et Yang, 2009).



Figure 9 : Représentation schématique de la protéine KLF4. En bleu et en rouge les domaines d'activation et de répression respectivement permettant les interactions avec d'autres protéines. En violet, les trois domaines doigt de zinc caractéristiques permettant la liaison à leur gène cible.

KLF4 est exprimé dans plusieurs tissus somatiques adultes mais son expression a également été retrouvée au niveau du testicule chez les vertébrés. Ce facteur de transcription est connu pour être impliqué dans plusieurs processus physiologiques tels que la régulation de la transcription, la différenciation ou encore la détermination de la destinée cellulaire (Netan et Yang, 2009). Le facteur KLF4 est l'un des quatre facteurs de Yamanaka essentiels à la production d'iPSC. Il est exprimé dans les cellules souches embryonnaires et spermatogoniales chez la souris avec une expression plus élevée dans ces dernières (Corbineau *et al.*, 2016). Chez cette espèce, il semble que la délétion de KLF4 au niveau testiculaire affecte la différenciation des cellules de Sertoli bien que ces cellules soient toujours capables de supporter la spermatogenèse (Godmann *et al.*, 2008). De plus, les souris mutantes présentent un taux réduit en testostérone ce qui semble indiquer que la communication paracrine entre les cellules de Sertoli et les cellules de Leydig pourrait être affectée (Godmann *et al.*, 2008).

Une analyse phylogénétique a mis en évidence la conservation des facteurs KLF chez les Bilatériens (Pei et Grishin, 2015). Les sept groupes présents chez les vertébrés sont représentés chez les Ecdysozoaires ainsi que les Lophotrochozoaires. Chez la drosophile, le gène Krüppel identifié code pour une protéine possédant trois domaines successifs en doigt de zinc. Cette protéine présente une forte homologie avec les facteurs KLF4 et KLF5 (Schuh *et al.*, 1986; Sur, 2009). La protéine Krüppel intervient lors du développement embryonnaire et serait impliquée dans la formation du thorax et de l'abdomen (Wieschaus *et al.*, 1984). En

effet, lorsque ce facteur est absent, ces deux structures présentent des malformations qui apparaissent dès la segmentation. De plus, Krüppel pourrait également intervenir dans le développement du système nerveux en permettant la mise en place correcte de l'innervation des muscles ainsi qu'en influençant la destinée des cellules gliales et neuronales chez la drosophile (Hartmann *et al.*, 1997). Krüppel présente donc un rôle de facteur de transcription essentiel lors du développement de la drosophile. Cette protéine agirait sur ses gènes cibles de façon concentration dépendante. En effet, lorsque la concentration est faible, Krüppel agirait en tant qu'activateur comme KLF4 chez les vertébrés. A l'inverse, lorsque la concentration est élevée, ce facteur agirait en tant qu'inhibiteur, comme KLF5 (Ghaleb *et al.*, 2005; Hartmann *et al.*, 1997).

Un homologue de KLF4 a été identifié chez un mollusque bivalve, *Chlamys farreri*. Son expression a été mesurée par PCR quantitative et apparaît plus élevée dans la gonade mâle. De plus, l'hybridation *in situ* a montré la localisation du transcrit dans les cellules de la lignée germinale mâle à l'exception des spermatozoïdes. Enfin, l'utilisation d'ARN interférent sur ce facteur KLF4 a provoqué une diminution du nombre de spermatozoïdes et une augmentation de l'apoptose des spermatocytes. Ainsi, l'implication de KLF4 dans le fonctionnement de la spermatogenèse est potentiellement conservée chez les mollusques (Yang *et al.*, 2017).

2.2.4. Le facteur c-myc

La famille MYC regroupe les facteurs de transcription c-MYC, N-MYC et L-MYC (Poole et van Riggelen, 2017). La protéine c-MYC (myelocytomatosis oncogene) est caractérisée par cinq domaines conservés nommés MYC box (MB) et numérotés de 0 à IV (Figure 10). Ces domaines s'étendent de la partie N-terminale jusqu'à la région centrale de la protéine (Poole et van Riggelen, 2017; Tu *et al.*, 2015). Ces domaines sont essentiels aux interactions protéine-protéine.



Figure 10 : Représentation schématique de la protéine c-MYC. En vert, les cinq domaines conservés MYC box. En orange, le domaine bHLH-Zip permettant la régulation des gènes cibles. D'après Poole et van Riggelen, 2017.

Le domaine MB0 présente une conservation élevée avec les autres protéines MYC de même que le domaine caractéristique bHLH-Zip (basic Helix-Loop-Helix Leucine Zipper) localisé au niveau C-terminal (Poole et van Riggelen, 2017). Ce dernier se lie généralement à

la séquence canonique CACGTG des boîtes E de certains complexes protéiques permettant leur recrutement afin de réguler les gènes cibles (Tu *et al.*, 2015). La protéine c-MYC est l'un des quatre facteurs de Yamanaka essentiels à la reprogrammation de cellules différenciées en iPSC. De plus, son implication dans la prolifération cellulaire a également été montrée. En effet, une augmentation de l'expression de c-myc engendre une augmentation du taux de prolifération (Singh *et al.*, 2015). De la même façon que KLF4, l'expression de c-myc est retrouvée dans les cellules souches spermatogoniales à un niveau plus élevé que dans les cellules souches embryonnaires. Les cellules souches spermatogoniales chez la souris étant réfractaires à la reprogrammation induite, il se pourrait que les expressions élevées de KLF4 et c-myc soient dues à une reprogrammation spontanée de ces cellules (Corbineau *et al.*, 2016).

Chez la drosophile, un seul gène *myc* a été identifié à ce jour. L'expression de ce facteur apparaît élevée dans les cellules souches en comparaison avec les cellules filles plus différenciées. La surexpression de *myc* bloque la différenciation des cellules souches suggérant que ce facteur serait impliqué dans la prolifération de ces cellules. En effet, lors de la délétion de *myc* chez la drosophile, les proliférations sont stoppées (Bellosta et Gallant, 2010). Les résultats obtenus confirment donc une conservation du rôle de Myc chez cet Ecdysozoaire.

Chez le Lophotrochozoaire *Platynereis dumerilii*, la protéine Myc a été identifiée par similarité de séquence (Altincicek et Vilcinskas, 2007). Il a été montré que ce facteur était phylogénétiquement proche de la protéine C-MYC humaine. De plus, les résultats obtenus ont permis de montrer une augmentation de l'expression de *myc* lors de l'induction de la réponse immunitaire. Ainsi, les auteurs ont suggéré une possible conservation du rôle de C-MYC dans la régulation de la prolifération et de la division des cellules immunitaires (Altincicek et Vilcinskas, 2007). Un homologue de MYC a également pu être identifié chez le mollusque bivalve *C. gigas* (David *et al.*, 2005).

2.2.5. Le facteur Nanog

Le facteur de transcription Nanog est une protéine contenant un homéodomaine permettant la liaison à l'ADN ainsi que, au sein du domaine d'activation, une région riche en tryptophane nécessaire à l'homodimérisation (Figure 11) (Chang *et al.*, 2009; Kim *et al.*, 2005; Zhang *et al.*, 2016).



Figure 11 : Représentation schématique de la protéine Nanog. En jaune, l'homéodomaine permettant la liaison à l'ADN. En violet, les deux sous-domaines de la partie C-terminale formant le domaine d'activation de la protéine. En bleu, la région riche en tryptophane (WRD) impliqué dans la dimérisation. D'après Zhang *et al.*, 2016.

Initialement, Nanog a été considéré comme spécifique des mammifères mais il apparaît que ce facteur de transcription est conservé dans différents groupes phylogénétiques (Scerbo *et al.*, 2014). Cependant, la similarité des séquences Nanog apparaît faible entre les espèces. En effet, chez les vertébrés, la protéine Nanog retrouvée chez le poulet ainsi que le poisson zèbre ne possède pas de région riche en tryptophane contrairement à celui de la souris. Dans le cas du poisson zèbre, il a été montré que malgré la région manquante, Nanog est capable de se dimériser et nécessite pour cela l'homéodomaine ainsi qu'un domaine N-terminal (Saunders *et al.*, 2013; Schuff *et al.*, 2012). Nanog est l'un des facteurs de pluripotence pouvant réguler la balance entre la prolifération et différenciation des cellules souches, notamment des cellules embryonnaires souches et germinales chez la souris (Kuroda *et al.*, 2005; Yamaguchi *et al.*, 2005; Zhang *et al.*, 2016). En effet, lorsque les cellules embryonnaires de souris sont mises en culture dans un milieu permettant leur différenciation, l'expression induite de *nanog* permet de les conserver dans un état indifférencié. De plus, dans ces conditions, ces cellules continuent d'exprimer des facteurs de pluripotence tel que *oct4* (Mitsui *et al.*, 2003). Les facteurs Nanog, Oct4 et Sox2 pourraient également former des complexes de régulation et contrôler le maintien de l'état souche (Kashyap *et al.*, 2009). Concernant la conservation phylogénétique du facteur Nanog, aucun homologue n'a pu être identifié chez les invertébrés (Ma *et al.*, 2014). Il semblerait que ce marqueur de cellules souches, qui n'est pas indispensable à la formation des iPSC, ne soit présent que chez les Deutérostomiens et plus précisément dans le groupe des vertébrés (Scerbo *et al.*, 2014).

2.3. Marqueurs moléculaires spécifiques des GSC

Parmi les différents marqueurs moléculaires des cellules souches, certains sont classiquement utilisés pour identifier les cellules germinales souches. Trois marqueurs phylogénétiquement conservés sont particulièrement étudiés : Vasa, Piwi et Nanos. Ces facteurs sont présents dans le nuage des GSC et régulent leur fonctionnement en contrôlant l'entrée en différenciation ou encore l'auto-renouvellement (Juliano *et al.*, 2010).

2.3.1. *Vasa*

Les DEAD-box constituent la plus grande des six superfamilles d'hélicases de l'ARN. Cette superfamille est caractérisée par douze motifs conservés dont quatre permettent l'utilisation de l'ATP, incluant le motif spécifique « Asp-Glu-Ala-Asp » (DEAD), et six permettent les interactions avec l'ARN (Figure 12). Les deux motifs restants sont impliqués dans les interactions moléculaires (pour revue voir Linder et Fuller-Pace, 2013). Les hélicases de l'ARN sont présentes chez la majorité des organismes. Ce sont des enzymes impliquées dans le déroulement des structures secondaires de l'ARN de façon ATP-dépendante. La protéine DEAD-box RNA hélicase Vasa, aussi nommée DDX4, contient tous les motifs des DEAD-box RNA hélicases mais également une région riche en glycine et en arginine du côté N-terminal composé de motifs « Arg-Gly-Gly » (RGG) (Figure 12) ainsi qu'un domaine acide au niveau de l'extrémité de la partie C-terminale (Lasko, 2013). Certaines protéines DDX4 peuvent contenir une ou plusieurs régions en doigt de zinc dans leur région N-terminal.

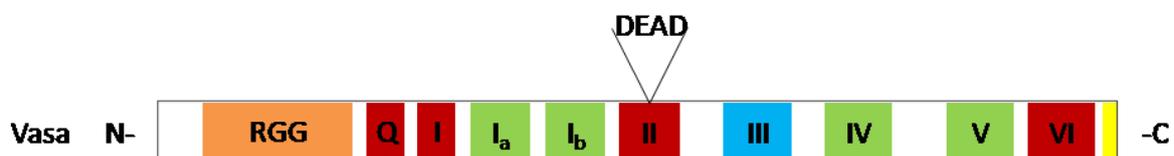


Figure 12 : Représentation schématique de la protéine Vasa. En orange, le domaine riche en motifs RGG spécifique des RNA hélicases. En rouge, vert et bleu, les domaines spécifiques des protéines DEAD-box impliqués dans les interactions avec l'ATP, l'ARN et les interactions intramoléculaires respectivement. Le motif « D-E-A-D » est localisé dans le domaine II. En jaune, le domaine acide à l'extrémité C-terminale. D'après Lasko, 2013 et Linder et Fuller-Pace, 2013.

Dans les gonades mâle et femelles des mammifères adultes, la protéine Vasa a été localisée dans le cytoplasme des cellules germinales précoces, avant leur entrée en méiose (Noce *et al.*, 2001; Toyooka *et al.*, 2000). Chez la souris, la délétion de Vasa entraîne une augmentation de l'apoptose des cellules germinales ainsi qu'un arrêt de la spermatogenèse indiquant un rôle essentiel de ce facteur pour la survie germinale et la progression méiotique. Bien que Vasa soit également retrouvée dans les cellules germinales femelles,

cette protéine ne semble pas essentielle au bon déroulement de l'ovogenèse chez la souris (Tanaka *et al.*, 2000). Chez le poulet, l'homologue de Vasa nommé cvh (Chicken Vasa Homolog) a été identifié et a permis d'observer un marquage spécifique de la lignée germinale. En effet, la protéine Cvh a été retrouvée dans le cytoplasme des PGC chez l'embryon ainsi que dans les cellules germinales adultes (Tsunekawa *et al.*, 2000). Chez le poisson zèbre adulte, *vasa* est exprimé dans les cellules germinales mâles et femelles. Lors de la délétion de *vasa*, les individus mâles deviennent stériles à cause de l'absence de cellules germinales. Ceci confirme la nécessité de Vasa pour la survie des cellules germinales (Hartung *et al.*, 2014).

Vasa a été identifié initialement chez *D. melanogaster* et ses homologues ont ensuite été caractérisés chez de nombreuses espèces de vertébrés et d'invertébrés (Raz, 2000). Leur lien avec une multitude de processus physiologiques tels que la spécification germinale, la gamétogenèse ainsi que le fonctionnement des cellules souches a été démontré (Lasko, 2013). Chez certains Ecdysozoaires, la protéine Vasa est également localisée dans le cytoplasme des cellules germinales adultes à l'exception des gamètes matures mâles. Comme chez les vertébrés, Vasa est un marqueur germinal nécessaire au bon déroulement de la gamétogenèse et à la différenciation des cellules germinales incluant les GSC chez l'adulte (Lasko, 2013; Liu *et al.*, 2009; Salinas *et al.*, 2007, 2012). Vasa a également été retrouvé chez un Métazoaire Cténophore hermaphrodite et capable de régénération, *Pleurobrachia pileus*. Chez cette espèce, l'homologue de *vasa* nommé Ppivasa, est exprimé dans les cellules des lignées germinales femelle et mâle à l'exception des spermatozoïdes matures ainsi que dans des cellules souches somatiques (Alié *et al.*, 2011).

La conservation de Vasa a permis l'identification d'homologues chez des Plathelminthes. Chez la planaire *Macrostomum lignano*, l'expression de *macvasa* a été démontrée dans les gonades ainsi que dans les cellules souches bipotentielles (somatique et germinal), que sont les néoblastes (Pfister *et al.*, 2008). Cependant la présence de cette protéine ne semble pas essentielle au maintien et à la formation de la gonade chez cette espèce. Ceci peut être dû à un mécanisme de redondance permettant le bon fonctionnement de la gonade malgré l'absence de *Macvasa* (Pfister *et al.*, 2008). Chez *Dugesia japonica*, une autre planaire, deux homologues de *vasa* ont été identifiés et leurs expressions ont été caractérisées dans les cellules germinales (Shibata *et al.*, 1999). En comparant des individus à reproduction sexuée et d'autres à reproduction asexuée, ces auteurs ont suggéré que l'un des homologues de

vasa (DjvlgA) pourrait être plus spécifique des néoblastes alors que le deuxième (DjvlgB) serait plus spécifique des cellules germinales précoces de l'ovaire et des lobes testiculaires (Shibata *et al.*, 1999). Chez les mollusques, un homologue de vasa a été identifié chez plusieurs espèces de bivalves. Son expression est retrouvée dans les cellules germinales précoces mâles et femelles (Fabioux *et al.*, 2004a, 2004b, Milani *et al.*, 2015, 2017; Obata *et al.*, 2010).

2.3.2. Piwi

Les gènes piwi forment une sous-famille appartenant au groupe codant pour les protéines Argonautes (AGO). Ces protéines AGO, caractérisées par le motif conservé « Asp-Asp-His » (DDH) chez les vertébrés, sont impliquées dans l'inactivation de gènes cibles en se liant à de petits ARN non-codants tels que les siARN (petits ARN interférants) et les miARN (microARN). Les protéines Piwi, caractérisées par le motif conservé « Asp-Asp-Leu » (DDL), s'associent également aux piRNA (Piwi-interacting RNA) qui sont de petits ARN non codants d'une longueur comprise entre 24 et 31 nucléotides (Iwasaki *et al.*, 2015). Tous les membres des protéines AGO, y compris les protéines Piwi, contiennent trois domaines conservés au sein de leur séquence : le domaine PAZ (Piwi, Argonaute et Zwiller), le domaine MID (middle) et le domaine PIWI (P-element Induced Wimpy testis) (Bamezai *et al.*, 2012; Samji, 2009). Le domaine PAZ se lie à l'extrémité 3' des ARN non-codants et le domaine MID se lie à leur extrémité 5' (Figure 13). Le domaine PIWI possède une Ribonucléase H-like impliquée dans plusieurs processus d'interaction avec l'ARN (Samji, 2009).



Figure 13 : Représentation schématique de la protéine Piwi. En rouge et en bleu, les domaines PAZ et MID qui se lient aux extrémités 3' et 5' des ARN non-codants respectivement. En vert, le domaine PIWI impliqué dans les interactions avec l'ARN. D'après Bamezai *et al.*, 2012.

Les protéines Piwi sont conservées chez les Métazoaires. En effet, plusieurs homologues ont été identifiés chez différentes espèces et il apparaît qu'ils présentent des fonctions conservées. Les Piwi ont un rôle important dans le maintien et la division asymétrique des GSC permettant leur auto-renouvellement (Samji, 2009; Thomson et Lin, 2009; van Wolfswinkel, 2014; Yakushev *et al.*, 2013) ainsi que dans l'inactivation de transposons (Bamezai *et al.*, 2012; Juliano *et al.*, 2011). Chez les planaires, piwi est exprimé dans les néoblastes et apparaît indispensable à la régénération des individus (Juliano *et al.*, 2011; van Wolfswinkel, 2014). L'irradiation des néoblastes ou l'ARN interférence de piwi chez *S.*

mediterranea provoque une incapacité de régénération pour l'animal (Reddien *et al.*, 2005). Chez d'autres espèces capables de régénération, piwi est retrouvé dans les cellules souches somatiques ainsi que dans les GSC et son expression apparaît régulièrement co-localisée avec celle de vasa dans ces cellules (van Wolfswinkel, 2014).

2.3.3. Nanos

Les protéines Nanos, caractérisées chez les vertébrés, possèdent un motif en doigt de zinc spécifique. En plus d'être un domaine conservé, il est indispensable pour la liaison à l'ARN et l'interaction avec Pumilio. En effet, les protéines Nanos se lient au domaine C-terminal de Pumilio afin de former un complexe de régulation post-transcriptionnelle capable de se lier à l'ARN. Les protéines Nanos chez les vertébrés et certaines chez les invertébrés contiennent un domaine NIM (NOT1-interacting motif) dans la région N-terminale (Figure 14). Ce domaine est impliqué dans la liaison avec ses partenaires protéiques afin de réguler la transcription des gènes cibles. Jusqu'à quatre protéines Nanos sont retrouvées chez les espèces où des homologues ont été identifiés (De Keuckelaere *et al.*, 2018).



Figure 14 : Représentation schématique de la protéine Nanos. En rouge, le domaine NIM (NOT1-interacting motif) impliqué dans les interactions avec des partenaires protéiques. En vert, le domaine en doigt de zinc conservé chez les protéines Nanos et impliqué dans les liaisons à l'ARN et l'interaction avec le partenaire Pumilio. D'après De Keuckelaere *et al.*, 2018.

Les protéines Nanos sont connues pour leur rôle dans le développement de la lignée germinale et particulièrement dans la pluripotence et la survie de la lignée germinale. De la même façon que Vasa et Piwi, les protéines Nanos sont retrouvées chez diverses espèces animales et leur fonction semble conservée (De Keuckelaere *et al.*, 2018). Chez les vertébrés, les protéines Nanos et Pumilio sont retrouvées dans les cellules germinales précoces et apparaissent essentielles à leur survie et leur prolifération. Ces acteurs ont également été identifiés dans les cellules germinales indifférenciées chez différentes espèces de poissons telles que la truite et le poisson zèbre (Beer et Draper, 2013; Bellaiche *et al.*, 2014). De plus, chez la petite roussette *Scyliorhinus canicula*, deux facteurs Nanos ont été identifiés et l'un d'eux, Nanos1B, apparaît présent dans le noyau de cellules germinales précoces aussi bien mâles que femelles (Gribouval *et al.*, 2018).

Au sein du groupe des Ecdysozoaires, des homologues de Nanos ont été caractérisés notamment chez *D. melanogaster* et *C. elegans* (De Keuckelaere *et al.*, 2018; Parisi et Lin, 2000). En effet, chez la drosophile adulte, les mutants *nanos* possèdent peu de GSC et ces dernières présentent des taux de prolifération et de survie réduits (Forbes et Lehmann, 1998; Wang et Lin, 2004).

Comme Vasa et Piwi, Nanos a également été identifié chez les planaires et localisé dans les GSC (Nakagawa *et al.*, 2012; Sato *et al.*, 2006). L'étude de *nanos* chez *Dugesia ryukyuensis* a montré qu'il est exprimé dans les cellules germinales précoces mâles et femelles. Après traitement par ARN interférence pour *nanos*, les animaux ne peuvent pas acquérir la capacité de reproduction sexuée et ne produisent pas de cellules germinales, le processus de sexualisation est bloqué. De façon intéressante, les cellules somatiques des organes sexuels se développent normalement ce qui suggère fortement que *nanos* est présent dans les GSC mais pas dans les néoblastes engendrant les cellules somatiques (Nakagawa *et al.*, 2012). De même, chez *Dugesia japonica*, *nanos* est exprimé seulement dans une sous-population de néoblastes pouvant former les cellules germinales. Il semblerait donc que deux groupes de néoblastes soient présent chez les planaires, l'un permettant la production de cellules somatiques et l'autre, qui exprime *nanos*, permettant la production de cellules germinales souches après sexualisation (Sato *et al.*, 2006). Des homologues de Nanos ont également été identifiés chez quelques mollusques (Kranz *et al.*, 2010; Matsumoto *et al.*, 2013; Rabinowitz *et al.*, 2008) sans pour autant préciser leur rôle dans ce groupe phylogénétique.

En résumé, tous les facteurs de pluripotence ainsi que les marqueurs de GSC ont été caractérisés chez les modèles vertébrés et plus particulièrement chez les mammifères mais également certaines espèces invertébrées (Figure 15).

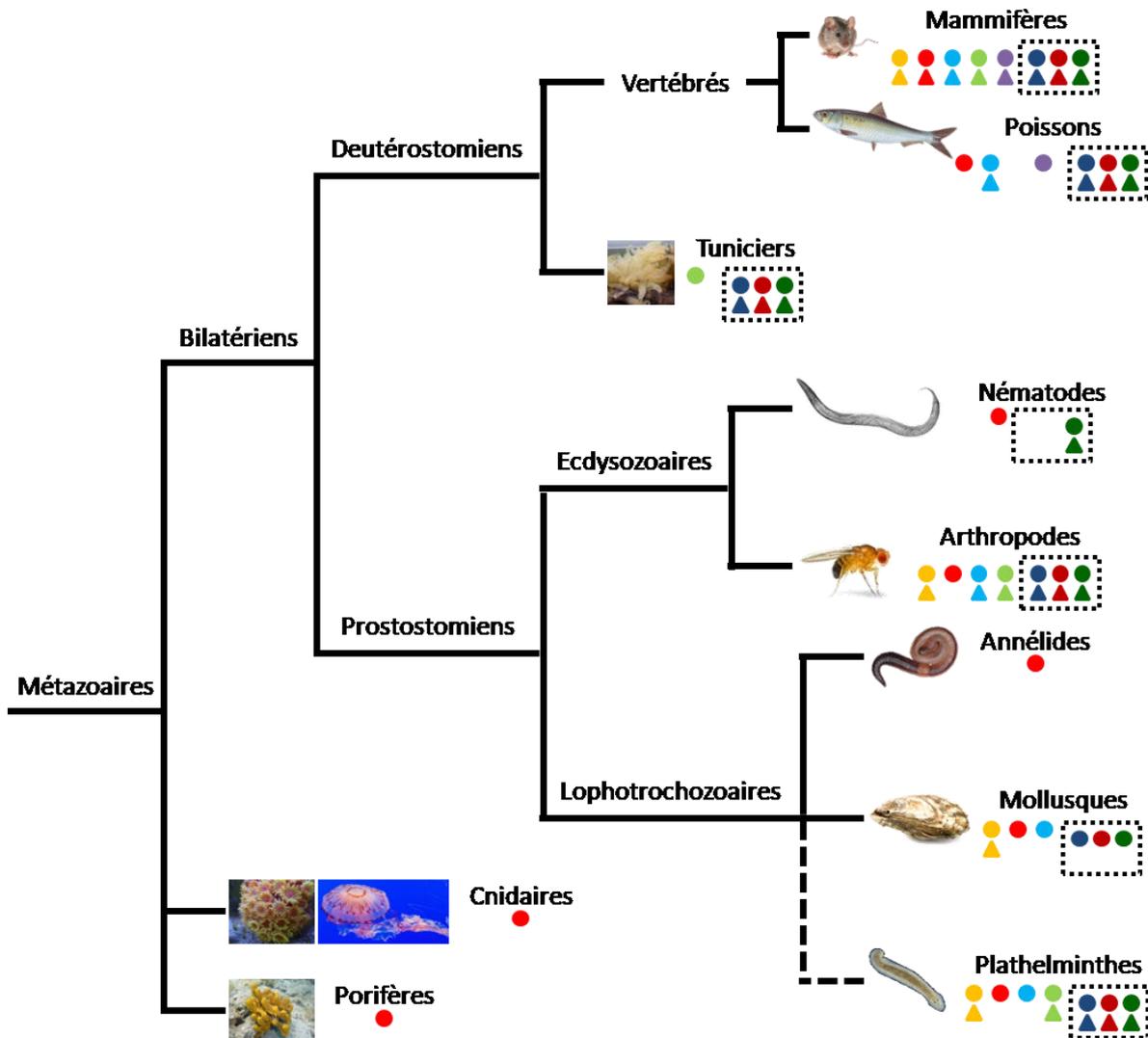


Figure 15 : Arbre phylogénétique simplifié des Métazoaires résumant la conservation des marqueurs de pluripotence et des marqueurs de GSC. Sox2 : Jaune, KLF4 : Rouge, Oct4 : Bleu, c-myc : Vert, Nanog : Violet, Vasa : Bleu foncé, Piwi : Rouge foncé, Nanos : Vert foncé. Les marqueurs spécifiques des GSC sont entourés par des pointillés. Pour chaque marqueur, la conservation de la séquence est schématisée par un rond et la conservation de la fonction est schématisée par un triangle.

La conservation des séquences a pu être mise en évidence grâce aux domaines protéiques spécifiques de chaque facteur. Bien que les séquences apparaissent relativement bien conservées au sein des Métazoaires, leurs fonctions restent encore à caractériser chez plusieurs espèces. Pour chacun des trois groupes constituant les Bilatériens, les facteurs de pluripotence indispensables à la formation d'iPSC, à savoir Sox2, Oct4, KLF4 et c-MYC, ont été identifiés. Les séquences homologues de KLF4 apparaissent les plus représentées et semblent donc les mieux conservées phylogénétiquement (Pei et Grishin, 2015). A l'inverse,

Nanog n'a pu être identifié que chez les vertébrés et présente une faible conservation de séquences entre les espèces (Ma *et al.*, 2014; Scerbo *et al.*, 2014). Cette observation est sûrement en lien avec le fait que ce facteur ne soit pas indispensable à la reprogrammation en iPSC. Concernant les marqueurs spécifiques des GSC, Vasa et ses homologues ont été étudiés chez de nombreuses espèces phylogénétiquement éloignées des vertébrés. Il en ressort une conservation structurale de la séquence protéique et en particulier des domaines des protéines RNA hélicases de la famille des DEAD-box. De plus, Vasa reste un marqueur spécifique des lignées germinales chez l'adulte ce qui tend vers une fonction conservée de cette protéine chez de nombreux organismes. Il est intéressant de noter que le transcrit et la protéine sont retrouvés dans les GSC mais également chez certaines espèces, comme les planaires, dans des cellules souches bipotentiels somatiques et germinales (Pfister *et al.*, 2008; Shibata *et al.*, 1999). A l'instar de Vasa, les deux autres marqueurs de GSC, Piwi et Nanos, sont également retrouvés chez plusieurs groupes phylogénétiques. D'après les données disponibles, ces facteurs sont tous impliqués dans le fonctionnement de la lignée germinale chez les espèces où la fonction a été décrite. Les caractérisations de ces marqueurs de cellules souches et de GSC chez les mollusques seront décrites de manière approfondie dans l'introduction du chapitre I.

2.4. Caractéristiques morphologiques des GSC

Les GSC, comme toutes les cellules souches, possèdent des caractéristiques morphologiques particulières. En effet, elles possèdent un ratio nucléo-cytoplasmique élevé ainsi qu'un nucléole particulièrement développé et visible en microscopie photonique et électronique à transmission (Figure 16) (Mercer et Cross, 1995). Les GSC les mieux décrites sont les Cellules Spermatogoniales Souches (SSC) chez les vertébrés mâles qui présentent une reproduction sexuée gonochorique. Ces SSC reposent sur la lame basale des tubes séminifères (Chiarini-Garcia et Russell, 2001, 2002; Mercer et Cross, 1995). Cette organisation est typique du microenvironnement de type épithélial des cellules souches (Scadden, 2006). Chez d'autres espèces, comme chez la drosophile par exemple, les GSC ne sont pas en contact direct avec la lame basale mais restent proches, cette organisation correspond au microenvironnement de type stromal des cellules souches. D'autres caractéristiques, communes à toutes les GSC, sont observables en étudiant leur ultrastructure telle que la faible activité de synthèse protéique des cellules quiescentes entraînant peu de structures cytoplasmiques (Chiarini-Garcia et Russell, 2001, 2002).

La morphologie des GSC femelles (FGSC), aussi appelées cellules souches ovariennes (OSC), est beaucoup moins caractérisée chez les vertébrés en raison de leur présence controversée par certains auteurs. En effet, du fait de la ménopause, la présence ou non d'une niche germinale femelle adulte chez les vertébrés est en débat. Cependant, des études montrent une conservation de critères morphologiques et moléculaires dans les OSC en comparaison avec leurs équivalents mâles, les SSC (Bhartiya, 2015; Decotto et Spradling, 2005; Gottardo *et al.*, 2017; Xie *et al.*, 2014). Les OSC sont de larges cellules, localisées proche de la membrane basale et présentant un ratio nucléo-cytoplasmique élevé. Chez la femelle du poisson *Xiphophorus nigrensis*, les OSC présentent des caractéristiques morphologiques de cellules germinales souches (Figure 16D) (Flores et Burns, 1993). Comme les SSC, les OSC possèdent un large noyau ainsi qu'un nucléole proéminent. Comme toutes les cellules souches, la chromatine présente un aspect décondensé et homogène visible en microscopie électronique à transmission (Flores et Burns, 1993). De plus, elles expriment des marqueurs moléculaires communs aux SSC, et aux GSC en général, tel que Vasa ou encore Oct4 (Virant-Klun, 2015; Ye *et al.*, 2017). Chez la drosophile où les GSC des deux sexes ont été identifiées, ces dernières partagent également des caractéristiques morphologiques spécifiques des cellules souches (Szakmary *et al.*, 2009).

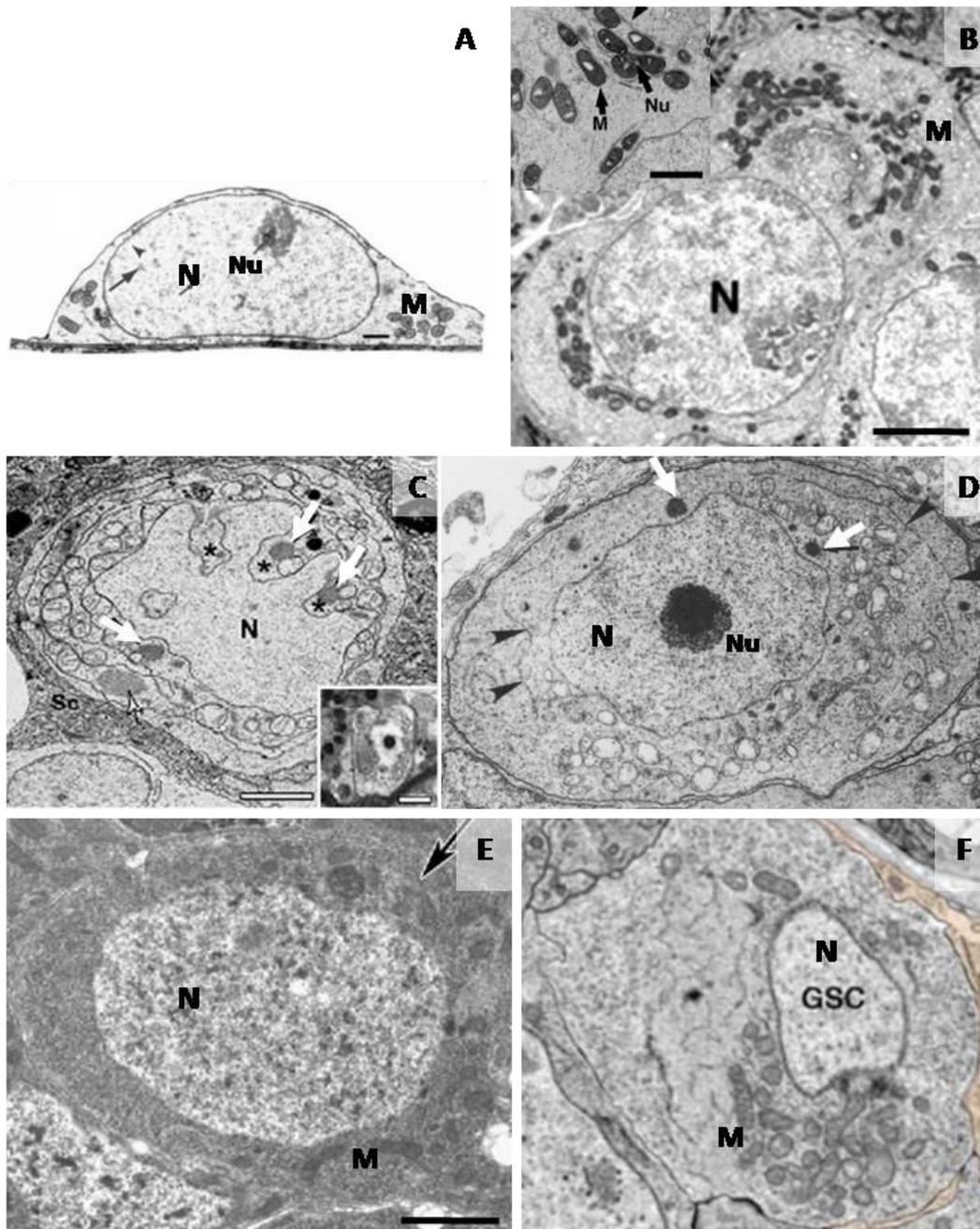


Figure 16 : Ultrastructure de cellules souches germinales. A – Spermatogonie souche de type A chez la souris (*Mus musculus*) mâle d'après Chiarini-Garcia *et al.*, 2002. B – Ovocyte néonatal chez la souris (*Mus musculus*) femelle d'après Pepling *et al.*, 2007. C – Spermatogonie souche de type A chez le tilapia (*Oreochromis niloticus*) mâle d'après Schulz *et al.*, 2010. D – Ovogonie souche chez la femelle *Xiphophorus nigrensis* d'après Flores et Burns, 1993. E – Cellule souche germinale mâle chez la drosophile (*Drosophila melanogaster*) d'après Gottardo *et al.*, 2017. F – Cellule souche germinale femelle chez la drosophile (*Drosophila melanogaster*) d'après Decotto et Spradling, 2005. N : Noyau ; Nu : Nucléole ; M : Mitochondries ; flèches blanches : nuage. Echelles : 0,5 μ m (B insert), 1 μ m (A et E), 1,5 μ m (C), 2 μ m (D), 5 μ m (B).

Le nuage qui est également nommé granules germinaux ou corps chromatoides, a été initialement décrit comme une structure cytoplasmique spécifique de la lignée germinale, dépourvue de membrane, dense aux électrons et composée d'ARN et de protéines. Dans les cellules souches, il apparaît que le nuage est régulièrement entouré de mitochondries regroupées (Figure 16) (Cannon *et al.*, 2011; Ramalho-Santos *et al.*, 2009). Chez plusieurs espèces, les composants du nuage incluent généralement les trois marqueurs spécifiques

des GSC : Vasa, Piwi et Nanos (Gao et Arkov, 2013). Cet élément cytoplasmique particulier a été initialement considéré comme une caractéristique morphologique spécifique des GSC (Figure 16). Chez la souris femelle, une étude réalisée sur les ovocytes en période néonatale a mis en évidence la présence d'un nuage dans ces cellules (Pepling *et al.*, 2007). Cependant, le nuage est aussi retrouvé dans certaines cellules souches somatiques chez des espèces capables de régénération. Ainsi, ce type de structure cytoplasmique est observé dans les néoblastes chez les planaires (Shibata *et al.*, 2010). De plus, certaines espèces présentent également un nuage dans leurs cellules neurales, incluant les cellules souches neurales (Gao et Arkov, 2013). Cette structure cytoplasmique contenant les principaux marqueurs des GSC apparaît ainsi plus largement essentielle à l'identité des cellules souches (Gao et Arkov, 2013).

L'aspect du noyau des GSC a été décrit chez plusieurs espèces. Il apparaît décondensé avec de rares régions d'hétérochromatine localisées le long de l'enveloppe nucléaire. Le noyau, ovale chez la souris mâle, peut contenir un ou deux nucléoles et sa taille diminue au cours de la spermatogenèse (Figure 16, A et B) (Chiarini-Garcia et Russell, 2001, 2002). Chez la drosophile, le noyau des GSC est également de forme ronde aussi bien chez les mâles que chez les femelles (Figure 16, E et F). Chez les poissons, la forme du noyau des SSC peut être irrégulière comme chez le tilapia par exemple. Chez cette espèce, le noyau des SSC est nettement irrégulier (Figure 16C) et s'arrondit au cours de l'avancée en spermatogenèse (Lacerda *et al.*, 2014; Schulz *et al.*, 2010). Ces caractéristiques sont également décrites dans les néoblastes des planaires (Sato *et al.*, 2006). Deux types de néoblastes ont été décrits sur la base de leur aspect nucléaire ainsi que par la structure des corps chromatoïdes localisés dans le cytoplasme, ce qui pourrait refléter deux états de différenciation différents (Higuchi *et al.*, 2007; Shibata *et al.*, 2010).

3. Le microenvironnement des GSC

La notion de microenvironnement spécifique des cellules souches et plus particulièrement le terme de « niche » a été défini pour la première fois par Schofield en 1978. À l'émergence de cette notion, la niche était définie comme une région physique de maintien permettant aux cellules, qui y sont localisées, de conserver leur état souche. Dans ce concept de niche, ces cellules entrent en différenciation lorsqu'elles quittent ce microenvironnement particulier.

3.1. Niche stromale ou épithéliale

L'organisation de la niche est relativement bien conservée dans le règne animal. L'amélioration des connaissances de différentes niches chez les espèces modèles a permis d'établir les bases de connaissances essentielles pour la compréhension de l'architecture d'une niche de cellules souches chez d'autres organismes. Plusieurs éléments sont retrouvés à savoir, la matrice extracellulaire sur laquelle reposent les cellules souches, un groupe de cellules somatiques accompagnatrices, des cellules circulantes ainsi qu'un réseau nerveux (Ferraro *et al.*, 2010; Scadden, 2006). Deux types de niche ont pu être décrits, la niche stromale et la niche épithéliale (Figure 17). La distinction entre ces deux types est liée au positionnement des cellules somatiques accompagnatrices par rapport aux cellules souches. Ainsi, les cellules souches de la niche stromale reposent sur un groupe de cellules somatiques accompagnatrices qui, elles, sont accolées à la membrane basale (Figure 17A). C'est le cas des niches germinales mâles et femelles chez la drosophile. A l'inverse, les cellules souches de la niche épithéliale reposent directement sur la membrane basale et sont entourées de cellules somatiques comme par exemple la niche germinale mâle chez la souris (Figure 17B).

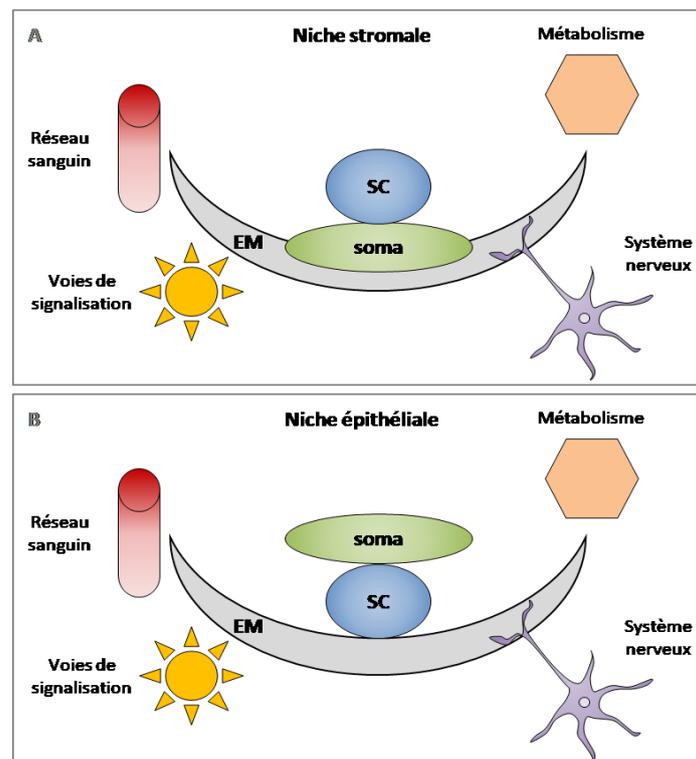


Figure 17 : Schémas des deux types décrits de niche. A – Schéma d'une niche stromale. B – Schéma d'une niche épithéliale. Les deux niches possèdent les acteurs acellulaires conservés et impliqués dans leur fonctionnement (Matrice extracellulaire, Réseau sanguin, Voies de signalisation, Système nerveux, Métabolisme) ainsi que les acteurs cellulaires principaux (Cellules souches et cellules somatiques accompagnatrices) d'après Scadden, 2006. SC : Cellules Souches ; soma : Cellules somatiques accompagnatrices ; EM : Matrice Extracellulaire.

3.2. Structure de la niche germinale chez différentes espèces

Les premières descriptions de niche germinale ont été réalisées dans la gonade de deux modèles Ecdysozoaires, *D. melanogaster* et *C. elegans* (Kimble et White, 1981; Xie et Spradling, 1998 respectivement).

Chez la drosophile femelle, la niche germinale est constituée des cellules somatiques du filament terminal, de la coiffe, des cellules somatiques accompagnatrices et des GSC (Figure 18A). Le bon fonctionnement des GSC nécessite le contact direct de ces dernières avec les cellules de la coiffe via des jonctions serrées (Hsu et Drummond-Barbosa, 2009). Chez la drosophile mâle, la niche germinale est localisée à la pointe apicale du testicule. Elle contient un groupe de cellules stromales, nommées cellules du hub, qui sont impliquées dans la régulation des cellules souches de la niche (Figure 18B). Entre six et neuf GSC, chacune entourée de deux cellules souches somatiques, sont en contact direct avec les cellules du hub. D'après cette organisation, la niche formée par les cellules du hub pourrait être commune aux GSC et aux cellules souches somatiques intragonadique. Ces cellules souches somatiques produisent les cellules somatiques du cyste (CySC) (Shim *et al.*, 2013; Xie *et al.*, 2005). Après divisions asymétriques des GSC, environ seize spermatogonies interconnectées sont formées et sont entourées de CySC (Shim *et al.*, 2013; Xie *et al.*, 2005).

Le nématode *C. elegans* est une espèce modèle qui se reproduit de façon sexuée et comprend des individus majoritairement hermaphrodites et une minorité d'individus mâles. Les GSC sont maintenues dans leur niche constituée d'une ou deux cellules somatiques, les cellules de la pointe distale (DTC), qui sont essentielles aux divisions mitotiques des cellules germinales (Figure 18C). Chez les individus hermaphrodites, une seule DTC est située à l'extrémité de chaque gonade tandis que chez les individus mâles, deux DTC sont localisées à l'extrémité de la gonade (Hubbard, 2007). Malgré qu'aucune jonction entre la DTC et les GSC ne soit actuellement caractérisée, ces cellules communiquent étroitement et sont localisées dans la zone des cellules souches (Stem Cell Zone) (Cinquin, 2009; Kimble et Crittenden, 2007). Chez cette espèce, les cellules germinales forment un syncytium malgré une communication restreinte et malgré le fait que l'avancée en gamétogenèse de ces cellules ne soit pas synchrone. Les cellules en mitose sont contenues dans la zone proliférative et les cellules ayant débuté la méiose sont situées dans la zone de transition (Hubbard, 2007). Plus les cellules se différencient, plus elles s'éloignent de la région distale de l'animal (Cinquin, 2009).

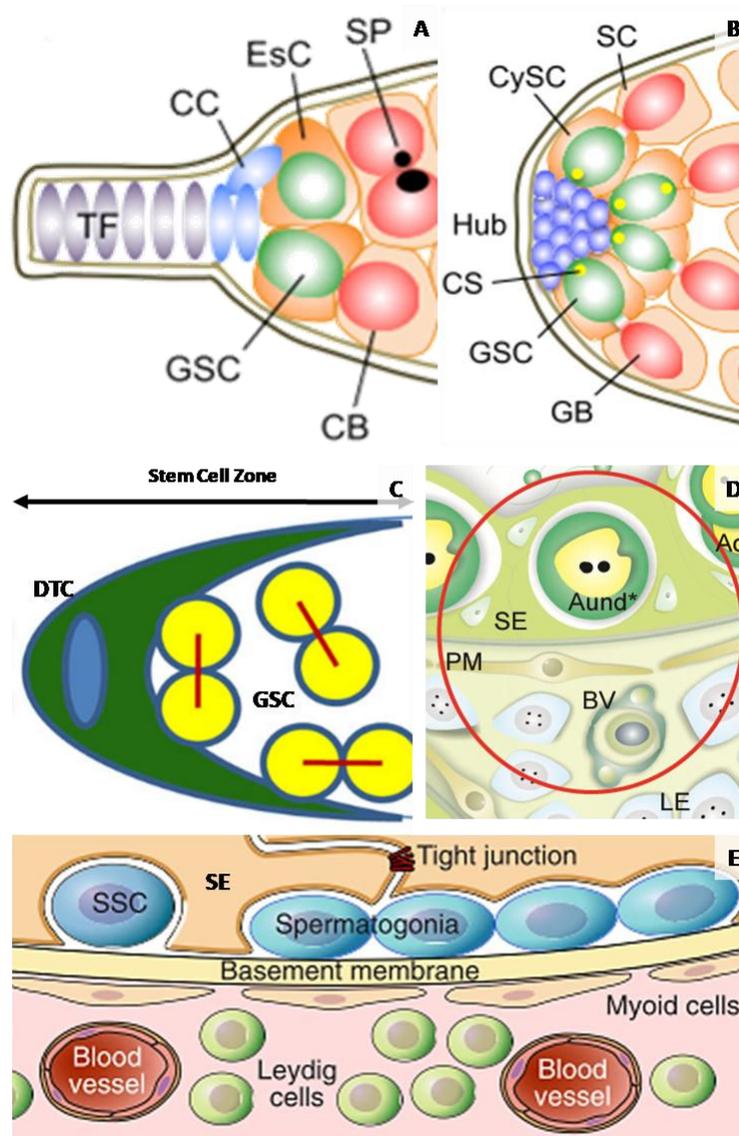


Figure 18 : Schémas de niches germinales chez des espèces modèles. A – Niche germinale femelle chez la drosophile (*Drosophila melanogaster*) d'après Shim *et al.*, 2013. B – Niche germinale mâle chez la drosophile (*Drosophila melanogaster*) d'après Shim *et al.*, 2013. C – Niche germinale chez l'hermaphrodite *Caenorhabditis elegans* d'après Joshi *et al.*, 2010. D – Niche germinale mâle chez le poisson zèbre (*Danio rerio*) d'après Nobrega *et al.*, 2010. E – Niche germinale mâle chez la souris (*Mus musculus*) d'après Oatley et Brinster, 2012. SSC : Cellule Souche Spermatogoniale ; Aund : Spermatogonie souche de type A ; SE : Cellules de Sertoli ; BM : Membrane Basale ; LE : Cellules de Leydig ; BV : Vaisseaux Sanguins ; DTC : Cellule Somatique de la pointe distale ; GSC : Cellule Germinale Souche ; TF : Cellules du Filament Terminal ; CC : Cellules de la coiffe ; EsC : Cellules Somatiques accompagnatrices ; CB : Cystoblaste ; Hub : Cellules somatiques du hub ; CySC : Cellules Souches du Cyste ; GB : Gonialblaste.

Chez les vertébrés, la structure de la niche germinale est décrite de manière plus approfondie chez les mâles. Chez certains poissons, le testicule est formé de cystes, comme par exemple chez le poisson zèbre *Danio rerio* (Leal *et al.*, 2009) ou encore chez la roussette *Scyliorhinus canicula* (Bosseboeuf *et al.*, 2014). Dans un cyste, les cellules de Sertoli entourent les SSC qui se diviseront et se différencieront de manière synchrone. Ainsi, un même cyste contient un seul type de cellules germinales. Contrairement aux mammifères, les cellules de Sertoli prolifèrent au cours de la spermatogenèse afin de s'adapter à l'augmentation du nombre de cellules germinales mâles au sein du cyste (Schulz *et al.*,

2010). La niche germinale chez le poisson zèbre mâle est composée des spermatogonies A indifférenciées accompagnées de cellules de Sertoli. Ces spermatogonies sont localisées au bord de la membrane basale (Figure 18D). À l'extérieur des cystes, sont retrouvées les cellules de Leydig, les cellules myoïdes ainsi que le réseau sanguin (Nóbrega *et al.*, 2010). Chez le poisson zèbre femelle, la niche germinale est encore peu connue. Cependant, l'existence de cellules souches germinales femelles (FGSC) a été montrée. Ces cellules maintiennent une production d'ovocytes chez l'adulte (Kawasaki *et al.*, 2012).

Chez la souris, comme chez la majorité des mammifères, les spermatogonies souches sont localisées au bord des tubes séminifères et se différencient en migrant vers la lumière des tubes (Figure 18E). Des jonctions serrées entre les cellules de Sertoli, qui soutiennent le développement germinale, séparent le compartiment contenant les cellules germinales avancées en différenciation (spermatocytes, spermatides et spermatozoïdes) du compartiment souche. Les cellules myoïdes qui forment le tube séminifère sont nécessaires à la structure de cette niche. Le tissu conjonctif, situé entre les tubes séminifères, contient les cellules de Leydig qui sécrètent la testostérone mais également des réseaux vasculaires et nerveux, ces éléments pouvant faire partie d'une vision plus large de la niche germinale (Oatley et Brinster, 2012). Les acteurs de la niche permettent de réguler le devenir des SSC en influençant leur auto-renouvellement, leur prolifération et leur différenciation. La possibilité d'une niche germinale femelle et de cellules souches ovogoniales (OSC) chez les mammifères a émergé récemment au vu des études réalisées chez les invertébrés comme la drosophile (Bukovsky *et al.*, 2004; Horan et Williams, 2017). Ce concept est encore très novateur et peu de données sont obtenues sur la caractérisation et le fonctionnement de cette potentielle niche. Cependant, certaines études indiquent que le pool de follicules primordiaux chez l'adulte serait une population dynamique. En effet, la souris adulte posséderait des cellules germinales femelles capables de prolifération et ce phénomène serait régulé par un microenvironnement particulier (Bukovsky *et al.*, 2004; Kerr *et al.*, 2013). Ces cellules germinales possèderaient donc des caractéristiques de GSC mais avec un fonctionnement un peu différent puisqu'elles ne présenterait pas la capacité d'auto-renouvellement mais seulement celle de différenciation (Kerr *et al.*, 2013). Certains suggèrent tout de même la possibilité pour ces OSC de former de nouveaux ovocytes à l'âge adulte. Dans ce cas, la limite de la reproduction femelle avec l'âge chez les vertébrés

supérieurs, pourrait être due à la dégradation des cellules somatiques de la niche et non pas à l'épuisement du stock de FGSC (Horan et Williams, 2017; Kerr *et al.*, 2013).

3.3. Voies de signalisation impliquées dans le fonctionnement de la niche germinale

Comme énoncé précédemment, les GSC possèdent la capacité de produire des cellules progénitrices, spermatogonies et/ou ovogonies. Ces gonies pourront ensuite se différencier afin de produire les gamètes matures. L'intégration de différents signaux moléculaires par les composants de la niche germinale confère à cette dernière son rôle fonctionnel sur le devenir des GSC. En effet, ce microenvironnement particulier permet le maintien du stock de cellules souches en régulant leurs capacités de prolifération, d'auto-renouvellement et de différenciation (Morrison et Spradling, 2008; Scadden, 2006). Tous ces mécanismes sont contrôlés par diverses voies de signalisation incluant celles des Wnt/ β -caténine, des Transforming Growth Factors β (TGF- β), la voie Notch ainsi que les voies de l'insuline (Ferraro *et al.*, 2010).

3.3.1. Les Wnt/ β -caténine

La voie de signalisation Wnt/ β -caténine est phylogénétiquement conservée à travers le règne animal. Les protéines Wnt font partie d'une famille de glycoprotéines et présentent plusieurs caractéristiques structurales essentielles à leur fonction telles que la conservation des cystéines et sérines impliquées dans l'interaction de ces protéines avec les lipides membranaires (Mills *et al.*, 2017). Une fois liés au récepteur Fzd (Frizzled), le ligand Wnt entraîne l'activation de la cascade de signalisation des Wnt/ β -caténine. Lorsque la voie de signalisation est activée, la protéine Dvl (Dishevelled) est recrutée afin d'inhiber le complexe de dégradation (β -TrCP couplé aux GSK3 β , CK1, Axine, APC, YAP et TAZ) ce qui permet aux β -caténines d'être stables dans le cytoplasme et, par la suite, de se lier à un co-facteur afin d'être transloquées dans le noyau et d'agir sur l'expression des gènes cibles (Figure 19). Dans le cas où la voie n'est pas activée, les β -caténines sont phosphorylées au niveau de leur domaine N-terminal puis dégradées (Lien et Fuchs, 2014).

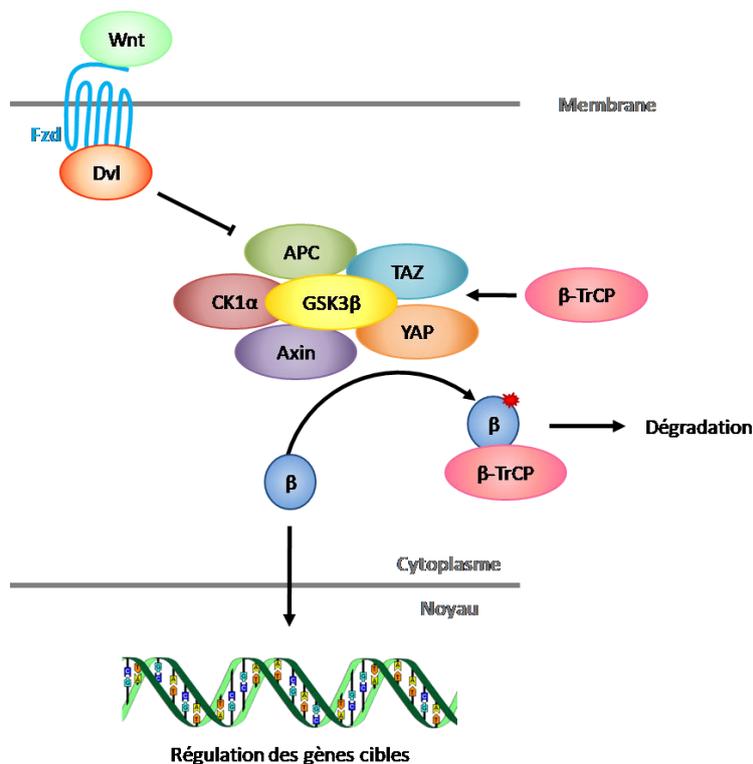


Figure 19 : Voie de signalisation simplifiée des Wnt/ β -caténine. La voie de signalisation comprend les acteurs décrits. La phosphorylation de la β -caténine (β) se fait sur sa partie N-terminale et est symbolisée par une étoile rouge.

Cette voie de signalisation est impliquée dans les mécanismes de communication cellulaire à travers les liaisons cellule à cellule et permet également le maintien de l'état souche des cellules au sein de la niche en régulant leur auto-renouvellement ainsi que leur

différenciation (Clevers *et al.*, 2014; MacDonald *et al.*, 2009; Mills *et al.*, 2017). La β -caténine intervient dans le remodelage de la chromatine et régule l'expression des facteurs de pluripotence, Oct4, Nanog et Sox2. Lorsque les Wnt sont absents, l'expression de ces facteurs est régulée négativement et, à l'inverse, lorsque la voie est activée l'expression de ces facteurs est régulée positivement ce qui stimule la prolifération des cellules souches (Lien et Fuchs, 2014; Mosimann *et al.*, 2009). La niche est une structure essentielle pour le bon fonctionnement de ce système. En effet, les spermatogonies indifférenciées chez la souris sont des cellules cibles de la voie Wnt/ β -caténine et ce signal semble être initié par Wnt6 qui est exprimé dans les cellules de Sertoli, l'un des composants cellulaires de la niche germinale (Takase et Nusse, 2016). De même, chez la drosophile, le signal Wnt est transmis aux GSC via les cellules somatiques accompagnatrices de la niche germinale (Upadhyay *et al.*, 2018).

3.3.2. La voie des TGF- β

Les ligands Transforming Growth Factors β (TGF- β) forment une famille de cytokines divisée en deux sous-familles, celle des Activines et celle des BMP (Bone Morphogenetic Proteins) définies selon les similarités de séquences entre ces cytokines et les cascades de signalisation qu'elles activent (Shi et Massagué, 2003). La voie de signalisation des TGF- β implique la liaison des ligands à deux types de récepteurs transmembranaires sérine-thréonine kinase, les types I et II (Figure 20).

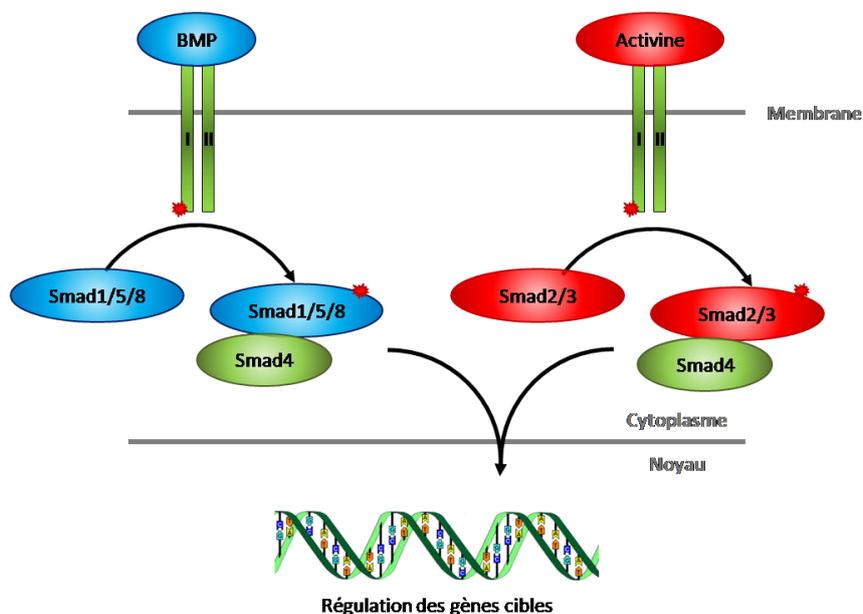


Figure 20 : Voie de signalisation simplifiée des TGF- β . La voie de signalisation est activée soit par les protéines BMP (gauche) soit par les protéines Activine (droite). Le récepteur et les protéines sont activés par phosphorylation (étoiles rouges) pour se lier à la smad4 et réguler leurs gènes cibles.

Les récepteurs de cette voie sont constitués d'une partie N-terminal extracellulaire servant à la liaison du ligand, d'une partie transmembranaire et d'une partie C-terminal intracellulaire contenant le domaine sérine-thréonine kinase (Shi et Massagué, 2003). Après liaison du ligand, le récepteur de type II phosphoryle le domaine kinase du récepteur de type I ce qui induira l'activation des protéines smads. Une fois phosphorylées, ces smads pourront se lier à la smad 4, ou co-smad, pour réguler la transcription des gènes cibles (Shi et Massagué, 2003). Les smads régulatrices activées ne sont pas les mêmes selon le type de ligand ce qui entraîne la transmission d'un signal différent. En effet, la liaison de l'Activine entraîne la phosphorylation des smads régulatrices nommées smads 2 et 3 alors que la liaison de BMP entraîne l'activation des smads 1, 5 et 8 (Rejon *et al.*, 2013; Shi et Massagué, 2003).

Les TGF- β contrôlent un large éventail de processus cellulaires tels que la prolifération, la différenciation ou encore les mécanismes d'apoptose (Shi et Massagué, 2003). D'une manière générale, il apparaît que les BMP, via leur action sur la niche, interviennent dans la régulation des propriétés inhérentes aux cellules souches telles que la prolifération ou la différenciation (Zhang et Li, 2005). Concernant la lignée germinale, les BMPs et plus précisément BMP4, jouent un rôle dans la formation des PGC chez l'embryon de souris (Dudley *et al.*, 2007; Wagner, 2007). Chez l'adulte, certains acteurs des TGF- β , comme BMP4 et l'Activine A, interviennent également dans la différenciation des SSC qui expriment leurs récepteurs. BMP4 participe à la régulation des cellules souches spermatogoniales chez la souris adulte et l'Activine A est nécessaire au maintien de cette population de spermatogonies différenciées (Oatley et Brinster, 2012). Les cellules somatiques de la niche germinale réguleraient la sécrétion de ces facteurs et par conséquent la balance entre auto-renouvellement et différenciation des GSC (Oatley et Brinster, 2012). Chez la drosophile, ce rôle des BMP est retrouvé avec Decapentaplegic (dpp), l'homologue du BMP2/4 humain. Dpp est essentiel pour le maintien des cellules souches, leurs division et différenciation dans l'ovaire de drosophile adulte (Xie et Spradling, 1998). Ce facteur est exprimé par les cellules somatiques du filament terminal et de la coiffe dans la niche germinale femelle et également dans les cellules somatiques du hub dans la niche germinale mâle (Jagut et Huynh, 2007). Chez la drosophile mâle, Dpp et Gbb (Glass bottom boat), deux facteurs BMP, agissent conjointement pour contrôler le maintien des GSC. Dpp et Gbb sont exprimés dans les cellules somatiques du hub et, lors de l'étude de mutants pour ces facteurs, peu ou pas de

GSC sont identifiables dans la niche. Aucune augmentation d'apoptose n'a été observée ce qui suggère que quand Gbb et Dpp sont absents les GSC entrent en différenciation (Kawase *et al.*, 2004).

3.3.3. La voie Notch

La voie de signalisation Notch est également conservée. Ses constituants, ligands et récepteurs Notch, permettent le transfert du signal vers le noyau des cellules. Les récepteurs Notch sont des protéines transmembranaires composées d'un domaine extracellulaire, nommé NECD, un domaine transmembranaire et un domaine intracellulaire nommé NICD (Bi et Kuang, 2015; Chillakuri *et al.*, 2012). Les ligands sont également des protéines transmembranaires appartenant la famille des DSL (Delta/Serrate/LAG-2). Ils se lient au domaine NECD du récepteur Notch d'une cellule voisine induisant ainsi la libération du domaine NICD qui migre ensuite jusqu'au noyau (Figure 21). Une fois transloqué dans le noyau, le domaine NICD se lie avec Rbpj (recombination signal binding protein for immunoglobulin kappa j region) et recrute un complexe de transcription afin de réguler l'expression des gènes cibles (Bi et Kuang, 2015).

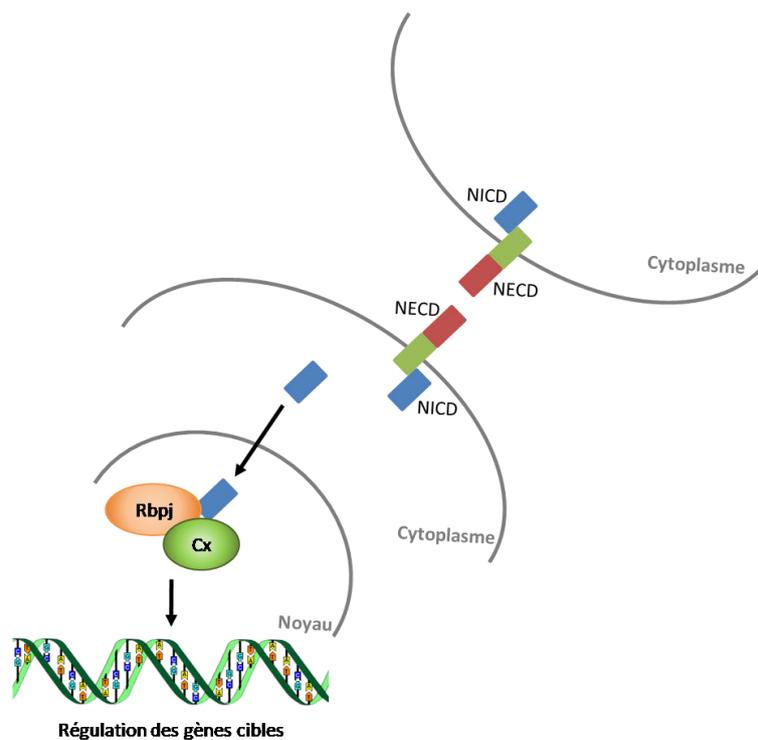


Figure 21 : Voie de signalisation Notch simplifiée. Après activation, le domaine intracellulaire de Notch (NICD) se détache de la partie extracellulaire (NECD). Il est ensuite transloqué dans le noyau pour se lier avec Rbpj qui recrute un complexe de transcription (Cx) qui régule les gènes cibles. D'après Bi et Kuang, 2015.

Notch est essentiel pour la communication cellulaire et est un acteur clé du métabolisme (Bi et Kuang, 2015). En effet, il est impliqué dans la survie cellulaire, la prolifération et la différenciation. Au niveau des cellules souches, cette voie de signalisation est impliquée dans le maintien de la quiescence cellulaire (Bi et Kuang, 2015). Chez la souris mâle, l'expression de Notch a été décrite dans les cellules somatiques et germinales, de même que son implication dans la différenciation des cellules germinales mâles (Huang *et al.*, 2013). En ce qui concerne les cellules souches spermatogoniales, le fait qu'elles expriment faiblement Notch suggère une activation de la voie Notch dans les cellules voisines ce qui conduit à une inhibition de la différenciation de ces SSC (Dirami *et al.*, 2001). Chez *C. elegans*, la voie de signalisation Notch est la plus décrite dans la régulation des GSC. Le ligand LAG-2 de la voie Notch est exprimé dans la cellule somatique de la pointe distale et le récepteur Notch dans les GSC. Cette voie de signalisation permet la prolifération et l'auto-renouvellement des GSC. Lorsque Notch est absent, les cellules germinales entrent en méiose et poursuivent la gamétogenèse (Byrd et Kimble, 2010; Kimble et Crittenden, 2007). La signalisation Notch peut fonctionner en lien étroit avec les voies BMP, Wnt ainsi que celles de l'insuline (Andersson *et al.*, 2011). Dans la niche germinale de drosophile, l'activation du récepteur à l'insuline entraîne la répression de Notch et cette balance de régulations permet de moduler la prolifération et la différenciation des cellules souches (Hsu et Drummond-Barbosa, 2011). Chez la drosophile femelle, les peptides insuline sont nécessaires au maintien des cellules somatiques de la coiffe, via la modulation de la signalisation Notch (Hsu et Drummond-Barbosa, 2011). Les cellules somatiques de la niche germinale produisent les ligands de la voie Notch afin de réguler la stabilité des GSC via les cellules de la coiffe (Hsu et Drummond-Barbosa, 2011). De plus, une perte des fonctions Notch chez la drosophile adulte femelle entraîne une diminution du nombre de cellules de la coiffe au sein de la niche (Song *et al.*, 2007). Chez la drosophile mâle, il a été montré que la voie de signalisation Notch intervient dans la régulation de la balance entre la différenciation et l'auto-renouvellement des cellules souches spermatogoniales (Okuda *et al.*, 2015). Ces éléments indiquent que la voie de signalisation de l'insuline est également un acteur clé de la régulation du fonctionnement des GSC.

4. Les rôles clés du système insuline dans la reproduction

4.1. Le système insuline

4.1.1. Les ligands de la superfamille des insulines

La superfamille des insulines regroupe un large éventail de molécules pour chaque espèce et apparaît conservée au sein des Métazoaires. Ces nombreux ligands appelés IRPs (Insulin Related Peptides) possèdent une structure caractéristique et conservée. Ils se composent d'un peptide signal, une chaîne B reliée à une chaîne A par des ponts disulfures et d'un peptide C qui peut être clivé ou non dans la forme active (Figure 22) (Claeys *et al.*, 2002).

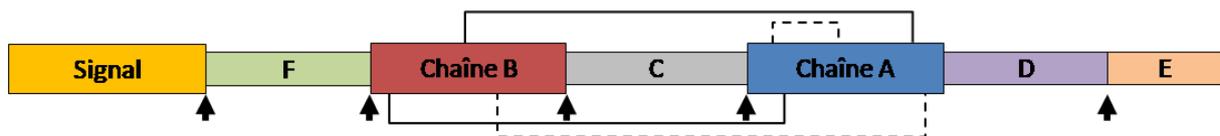


Figure 22 : Schéma de la structure conservée des ligands IRPs. Le peptide signal, les chaînes B et A ainsi que les deux ponts disulfures représentés par un trait noir sont retrouvés dans les ligands insuline. Les domaines F, D et E, le peptide C ainsi que les ponts disulfures représentés en pointillés sont facultatifs et dépendent des espèces étudiées. Les flèches représentent les sites de clivage possibles.

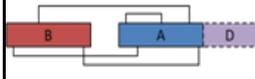
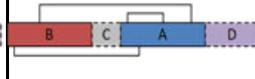
Les IRPs possèdent toujours deux ponts disulfures entre les chaînes B et A. Selon les espèces et les ligands étudiés, un pont disulfure supplémentaire peut être observé, soit entre les chaînes B et A, soit entre deux résidus cystéines de la chaîne A (Matsunaga, 2017). Des domaines supplémentaires nommés F, D et E sont facultatifs et dépendent également des espèces étudiées. A l'exception du domaine D, tous les autres domaines facultatifs sont clivés dans la forme mature des IRPs.

Des ligands de la superfamille des insulines ont été identifiés chez un grand nombre de vertébrés et d'invertébrés (Tableau 1). Chez certaines espèces, le nombre d'IRPs identifiés est très élevé. C'est le cas notamment pour *C. elegans* qui dénombre 40 ligands insuline (Li *et al.*, 2003; Matsunaga, 2017; Pierce *et al.*, 2001) et *Bombyx mori* qui en possède 32 (Kondo *et al.*, 1996; Maruyama *et al.*, 1988; Mizoguchi et Okamoto, 2013; Yoshida *et al.*, 1997, 1998). L'analyse structurale des différents IRPs chez *C. elegans* a permis de distinguer trois types sur la base de la structure de leurs formes matures. Ces trois types sont nommés α , β , γ et, à ce jour, *C. elegans* est la seule espèce où la coexistence des trois a été décrite (Tableau 1) (Matsunaga, 2017). Le type α est caractérisé par trois ponts disulfures entre les chaînes B et A, ainsi que la présence possible d'un domaine D non clivé après la chaîne A. Le type β possède les deux ponts disulfures canoniques entre les chaînes B et A ainsi qu'un pont supplémentaire à l'intérieur de la chaîne A. Enfin, le type γ est caractérisé par la présence des deux ponts disulfures canoniques ainsi qu'un pont à l'intérieur de la chaîne A. Dans ce dernier groupe, le peptide C n'est pas toujours clivé dans la forme mature et la présence d'un domaine D peut également être observée pour certains peptides insuline (Matsunaga, 2017). Différents processus de clivage du peptide C ont été décrits. Généralement, les pro-hormones convertases, PC1 et PC2 ou leurs homologues, sont impliqués dans le processus de protéolyse du peptide C induisant la maturation des ligands insuline (Malide *et al.*, 1995; Matsunaga, 2017). Pour bon nombre d'espèces dont les génomes ne sont pas disponibles, une partie seulement de ligands est identifiée.

Chez les invertébrés, les IRPs sont majoritairement synthétisés par des cellules neurosécrétrices du système nerveux central (Claeys *et al.*, 2002; Smit *et al.*, 1998). Par exemple, chez la drosophile, les DILPs 1, 2, 3 et 5 sont détectés dans des îlots de cellules neurosécrétrices de la *pars intercerebralis* (Liu *et al.*, 2016; Nässel *et al.*, 2015). Chez la lymnée *Lymnaea stagnalis*, les MIPs (Molluscan Insulin-related Peptides) sont détectés dans les LGC (light green cells) des ganglions cérébraux (Smit *et al.*, 1988, 1991). Cependant, des expressions localisées au niveau de tissus ou d'organes cibles sont également possibles. Ainsi, certains IRPs sont exprimés dans le tractus digestif comme par exemple les DILPs 2, 3 et 5 de drosophile (Nässel *et al.*, 2015; Veenstra *et al.*, 2008) ou les insulin-like de bivalves chez *Anodonta cygnea* et *Unio pictorum* (Plisetskaya *et al.*, 1978) ou *Mytilus edulis* (Fritsch *et al.*, 1976). Enfin, les IRPs sont parfois exprimés dans la gonade, comme DILP5 dans les

ovaires de drosophiles (Nässel *et al.*, 2015) où ils interviennent dans le contrôle de la gamétogenèse.

Tableau 1 : Représentations schématiques des principales formes matures des IRPs. Schémas selon la classification de Matsunaga *et al.*, 2017. Tous les IRPs ont deux ponts disulfures canoniques entre les chaînes B et A. Les types α et β ont un pont disulfure additionnel entre les chaînes B et A. Les types β et γ possèdent un pont disulfure interne à la chaîne A. Les chaînes B et A peuvent être reliées par un peptide C. Un domaine D supplémentaire peut être présent dans certains IRPs. AG : Glande Androgénique ; IAG : Insulin-like Androgenic Gland peptide ; ILP : Insulin-Like Peptide ; MIP : Molluscan Insulin-related Peptide ; IRPs : Insulin Related Peptides ; RLP : Relaxine ; R : Revue ; U : Données NCBI non publiées.

Formes matures			Type α	Type β	Type γ	Bibliographie
			2 ponts disulfures entre les chaînes A et B			
			Non	1 pont disulfure interne à la chaîne A		
			1 pont disulfure additionnel entre les chaînes A et B			
Phylogénie						
Deutérostomiens	Vertébrés	<i>H. sapiens</i>			Insulin, IGFs, INSL, Relaxin	Shabanpoor <i>et al.</i> , 2009 R
	Tuniciens	<i>C. intestinalis</i>			Ci-INS-L1, Ci-INS-L2, Ci-INS-L3	Olinski <i>et al.</i> , 2006
Ecdysozoaires	Nématodes	<i>C. elegans</i>	INS-20 to 31, INS 33 to 36, INS 39	INS 1 to 10, Daf 28	INS 11 to 19, INS 32, INS 37, INS 38	Pierce <i>et al.</i> , 2001; Li <i>et al.</i> , 2003; Matsunaga <i>et al.</i> , 2017 R
		Hexapodes	<i>B. mori</i>			Bombyxins family A to G
	<i>S. gregaria</i>				Scg-IRP	Badisco <i>et al.</i> , 2013
	<i>L. migratoria</i>				ILP	Lagueux <i>et al.</i> , 1990
	<i>A. aegyptii</i>				AaILP 1 to 8	Riehle <i>et al.</i> , 2006 R
	<i>D. melanogaster</i>				DILP 1 to 8	Broggiolo <i>et al.</i> , 2001; Grönke <i>et al.</i> , 2010 R
	Crustacés	<i>S. verreauxi</i>			Sv-IAG, Sv-ILP1, Sv-ILP2	Chandler <i>et al.</i> , 2017
		<i>A. vulgare</i>			Arv AGH (avec un pont disulfure additionnel intra-B)	Ventura <i>et al.</i> 2011
<i>M. rosenbergii</i>			Mr IAG	Mr ILP, Mr RLP	Phoungpetchara <i>et al.</i> , 2011; Ventura <i>et al.</i> , 2009	
Lophotrochozoaires	Mollusques	<i>L. stagnalis</i>		MIP 1, 2, 3, 4, 5, 7		Smit <i>et al.</i> , 1988, 1991, 1992, 1993, 1996, 1998; Li <i>et al.</i> , 1992
		<i>A. californica</i>		AI		Floyd <i>et al.</i> , 1999
		<i>C. geographus</i>		Con-Ins G2	Con-Ins G1	Safavi-Hemami <i>et al.</i> , 2015
		<i>S. officinalis</i>		So IRP		Zatylny-Gaudin <i>et al.</i> , 2015
		<i>C. gigas</i>		oIRP		Hamano <i>et al.</i> , 2005
		<i>L. gigantea</i>		ins 1, ins 2, ins 4	ins 3	Veenstra <i>et al.</i> , 2010
	Annélides	<i>P. dumerilii</i>		IRP2, IRP3, IRP4, IRP5	IRP1	Conzelmann <i>et al.</i> , 2013
	Plathelminthes	<i>S. mediterranea</i>			ilp-1	Miller and Newmark, 2012
		<i>T. solium</i>			ILP-1, ILP-2 (avec un pont disulfure additionnel intra-A)	Wang <i>et al.</i> , 2013
		<i>S. mansoni</i>			ILP (avec un pont disulfure additionnel intra-A)	Wang <i>et al.</i> , 2013
Porifères	<i>G. cydonium</i>			sponge insulin	Robitzki <i>et al.</i> , 1989	
Cnidaires	<i>N. vectensis</i>			NV 199266; NV 207484	Ancill, 2009	
	<i>H. magnipapillata</i>			Hv-ILP1; Hv-ILP2; Hv-ILP3	Steele <i>et al.</i> , 2009- U	

4.1.2. Les récepteurs d'IRPs

Les IRPs agissent via différents types de récepteurs. Les principaux récepteurs aux insulines sont des dimères de protéines transmembranaires à domaine tyrosine kinase. Ils sont constitués de deux chaînes α extracellulaires et deux chaînes β transmembranaires liées par des ponts disulfures (Claeys *et al.*, 2002). Les sous-unités α permettent la liaison du ligand au récepteur ce qui entraîne l'autophosphorylation des résidus tyrosine de la région intracellulaire (Figure 23) (Hubbard, 1999).

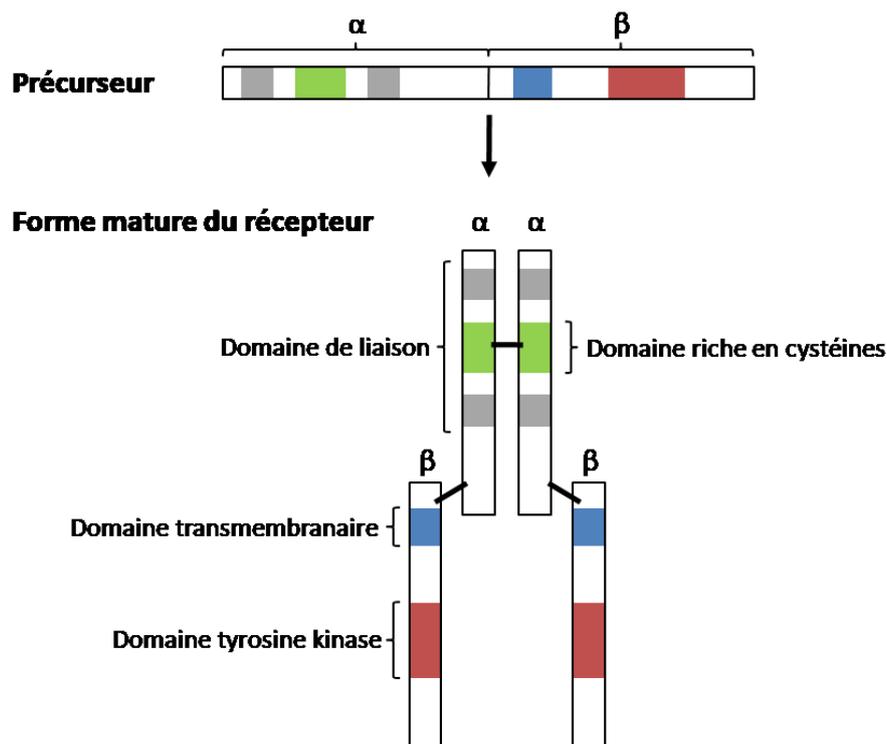


Figure 23 : Représentation schématique du précurseur et de la forme mature du récepteur tyrosine kinase aux insulines. Le précurseur unique contient les régions codantes pour les sous-unités α et β . Les sous-unités α contiennent le domaine de liaison aux IRPs constitué d'un domaine riche en cystéines (vert) et de deux séquences l'encadrant (gris) qui diffèrent selon le type de récepteur. Les sous-unités β contiennent le domaine transmembranaire (bleu) ainsi que le domaine tyrosine kinase (rouge).

Chez les vertébrés, plusieurs récepteurs permettent de lier les différents IRPs. Les IGFs (Insulin Growth Factors) peuvent se lier aux récepteurs des insulines ainsi qu'à des récepteurs spécifiques des IGFs, mais avec des affinités différentes (Shabanpoor *et al.*, 2009). Les récepteurs aux IGFs sont de deux types : l'IGF-1R et l'IGF-2R/M6P (Mannose-6-Phosphate). L'IGF-1R fait partie des récepteurs à domaine tyrosine kinase et présente une forte homologie avec le récepteur aux insulines (Foulstone *et al.*, 2005). Cependant, il diffère du récepteur aux insulines par son domaine de liaison du ligand, notamment les séquences encadrant la région riche en cystéines. Le récepteur IGF-2R/M6P est constitué d'une partie

N-terminale extracellulaire contenant 15 segments répétés d'environ 147 acides aminés chacun, d'un domaine transmembranaire unique et d'une région C-terminale intracytoplasmique (Hawkes et Kar, 2004). Ce récepteur lie les ligands contenant le M6P ainsi que les IGF-2 sur deux sites différents. Ainsi, IGF-2 se lie sur le segment 11 du domaine extracellulaire. Enfin, les relaxines, incluant les INSL, activent les récepteurs aux peptides de la famille des relaxines (RXFP) qui sont des récepteurs couplés aux protéines G (RCPG) (Bathgate *et al.*, 2013). Les RCPG sont constitués d'une partie N-terminale extracellulaire, de 7 hélices transmembranaires et d'une partie C-terminale intracellulaire.

Contrairement aux vertébrés qui possèdent plusieurs récepteurs susceptibles de lier les différents IRPs, chez les invertébrés un seul récepteur aux insulines a été décrit, liant potentiellement de multiples IRPs (De Meyts, 2004). Chez la drosophile, le récepteur DIR (*Drosophila* Insulin Receptor) présente, comme le récepteur humain d'insuline, deux chaînes α et deux chaînes β . Cependant, sa séquence présente un domaine supplémentaire à l'extrémité C-terminale des sous-unités β . Selon les types cellulaires dans lesquels le récepteur est exprimé, cette extension peut être clivée dans la forme active du récepteur DIR. Ainsi, deux isoformes de DIR peuvent être identifiées chez *D. melanogaster*, l'une très proche du récepteur aux insulines trouvé chez les vertébrés et l'autre présentant l'extension C-terminale (Claeys *et al.*, 2002). Chez *C. elegans*, le récepteur DAF-2 présente les domaines conservés des récepteurs aux insulines et aux IGFs de vertébrés (Kimura *et al.*, 1997). De même que le récepteur DIR, la séquence de DAF-2 possède une extension dans la partie C-terminale. Chez les Lophotrochozoaires, certains récepteurs aux insulines ont également été identifiés, notamment chez les mollusques. Les récepteurs aux insulines chez les mollusques seront détaillés dans l'introduction du chapitre II.

4.1.3. Les voies de signalisation

L'activation des voies de signalisation du récepteur aux IRPs est engendrée par la liaison du ligand sur le récepteur qui déclenche son autophosphorylation. Les principales voies sont la voie PI3K (Phosphatidylinositol 3 phosphate Kinase) et la voie des MAPK (Mitogen Activated Protein Kinase) (Figure 24). Ces voies de signalisation interagissent entre elles et régulent leurs effecteurs cibles via des cascades de phosphorylation (Templeman et Murphy, 2018). L'activation de la voie PI3K peut induire de multiples effets, essentiellement métaboliques mais qui peuvent également cibler l'expression génique. L'activation de cette voie est impliquée dans le contrôle de la synthèse de glycogène ou le transport du glucose, elle agit aussi sur les synthèses protéiques et la survie cellulaire. L'activation de la voie MAPK régule principalement des expressions de gènes impliqués dans les proliférations et la différenciation cellulaires.

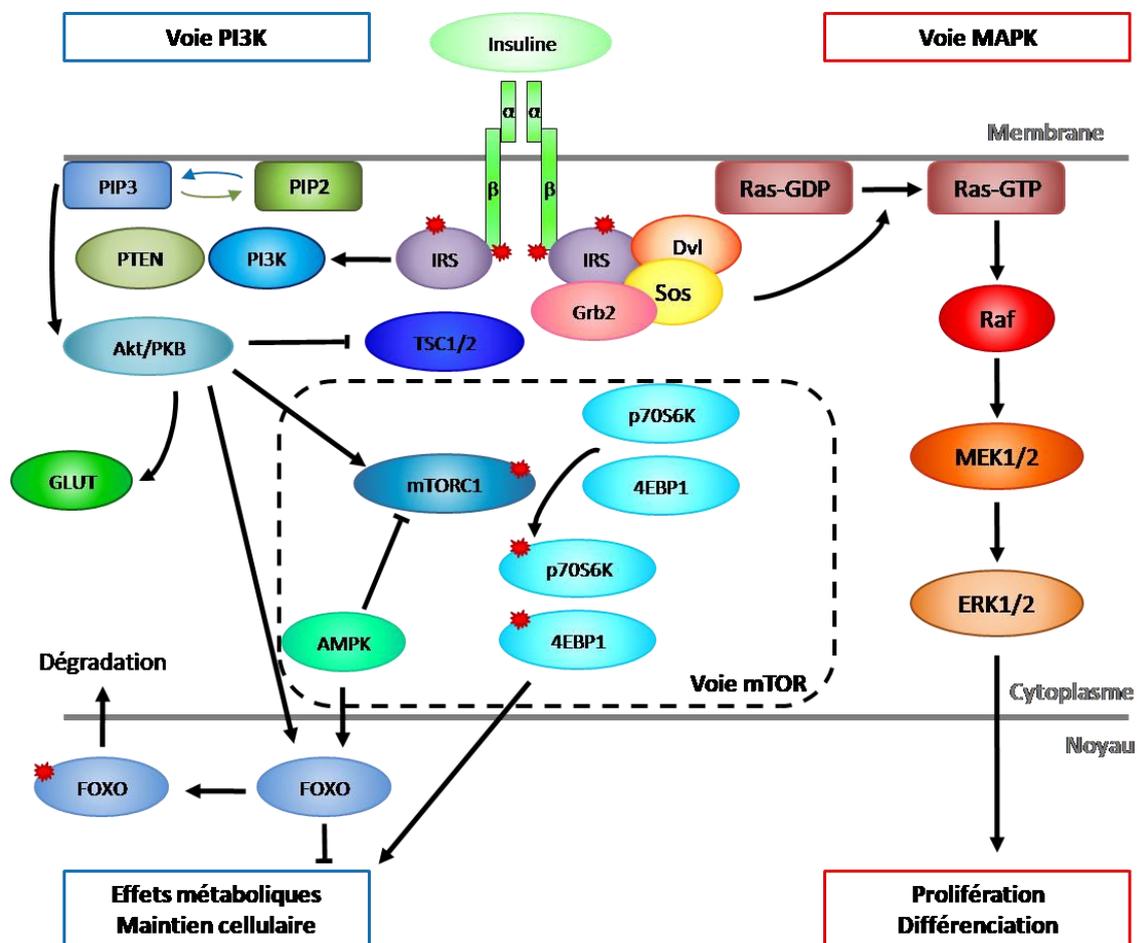


Figure 24 : Voies de signalisation PI3K/mTOR et MAPK simplifiées. Les voies de signalisation sont activées par liaison d'un peptide insuline sur le récepteur insuline à domaines tyrosine kinase. Elles comprennent les acteurs décrits. A gauche, la voie PI3K régulant les effets métaboliques et les réponses cellulaires. A droite, la voie MAPK régulant la prolifération et la différenciation cellulaire. Entourée par des pointillées, la voie mTOR phosphorylé par Akt/PKB pour l'activer et inhiber par AMPK. Elle régule, avec la voie PI3K, les effets métaboliques et le maintien cellulaire.

4.2. Implication des IRPs dans le contrôle de la reproduction

Depuis la découverte de l'insuline, la compréhension de ses mécanismes d'action a suscité un grand intérêt du fait de son implication dans la régulation de multiples processus biologiques comme le métabolisme glucidique, le métabolisme lipidique, la croissance, le développement, la durée de vie, la réponse immunitaire et la reproduction (pour revue voir Wilcox, 2005). Ces différents rôles sont retrouvés pour la plupart des clades phylogénétiques où cette signalisation est étudiée (Agrogiannis *et al.*, 2014; Dimitriadis *et al.*, 2011; Saltiel et Kahn, 2001; Sliwowska *et al.*, 2014). En ce qui concerne plus spécifiquement le processus de reproduction, l'implication des IRPs a été décrite dans le contrôle de plusieurs étapes clés incluant le fonctionnement de la niche germinale.

4.2.1. Les IRPs impliqués dans la maturation ovocytaire

Chez la plupart des Métazoaires, les voies de l'insuline sont impliquées, chez les individus femelles, dans le déroulement de l'ovogenèse et la maturation des ovocytes (Das et Arur, 2017; Dupont et Scaramuzzi, 2016). Le récepteur d'insuline a été localisé par immunohistochimie au niveau de la membrane des ovocytes humains, des cellules de la granulosa et de la thèque. L'avancée en ovogenèse entraîne une augmentation du marquage dans ces cellules indiquant l'implication de l'insuline dans la maturation ovocytaire (Samoto *et al.*, 1993). Chez le poisson zèbre, le récepteur de l'insuline est également localisé à la surface des ovocytes, dans les cellules de la thèque et de la granulosa (Das et Arur, 2017; Das *et al.*, 2016). Chez cette espèce, l'ajout d'IGF-1 ou d'IGF-2 sur des ovocytes immatures maintenus *in vitro* entraîne une augmentation du pourcentage de rupture de la vésicule germinale, indiquant une maturation des ovocytes (Nelson et Van Der Kraak, 2010). Chez *D. melanogaster*, les DILPs régulent différents processus de reproduction incluant la vitellogenèse et le développement ovocytaire. La délétion des peptides DILPs dans les cellules neuroendocrines entraîne une diminution du nombre d'ovocytes et, par conséquent, de la fécondité (Ikeya *et al.*, 2002; Wu et Brown, 2006). Chez *C. elegans*, il a été montré que les voies de l'insuline sont essentielles à la vitellogenèse. En effet, des individus hermaphrodites mutants pour le récepteur de l'insuline DAF-2, présentent une diminution des niveaux de transcrits de vit-2 et vit-5, deux gènes impliqués dans la vitellogenèse (DePina *et al.*, 2011).

4.2.2. Rôle androgène

Chez certains crustacés, des effecteurs appartenant à la superfamille des insulines ont été identifiés par homologie de séquences et sont impliqués dans la différenciation sexuelle. Chez ces espèces, la différenciation sexuelle est contrôlée par une glande androgène, seulement présente chez les mâles. L'effecteur identifié pour chaque espèce a été nommé IAG (Insulin-like Androgenic Gland peptide) puisqu'il possède les caractéristiques structurales des IRPs et que son expression est localisée par hybridation *in situ* dans la glande androgène (Chetler *et al.*, 2015, 2017; Manor *et al.*, 2007; Ventura *et al.*, 2011). En inhibant l'expression du gène d'IAG par ARN interférence, la spermatogenèse est bloquée ainsi que la régénération des caractères sexuels mâles ce qui montre l'implication de l'IAG dans le maintien de ces caractères (Ventura *et al.*, 2009). Après ablation de la glande androgène chez les mâles au stade larvaire, ces individus perdent leurs caractéristiques mâles et développent un système reproducteur femelle où se forment des ovocytes. Chez des individus femelles, l'injection d'IAG induit une masculinisation du système reproducteur (Ventura *et al.*, 2012).

4.2.3. IRPs et prolifération des cellules gonadiques

L'implication du système insuline sur les proliférations des cellules gonadiques, somatiques et germinales, a été mise en évidence chez plusieurs espèces, principalement chez les mâles. Chez la souris, les transcrits des récepteurs de l'insuline et IGF-1R ont été détectés dans les spermatogonies ainsi que dans les cellules de Sertoli. En période néonatale, les souris mâles mutantes pour ces récepteurs présentent une réduction de 75% de la taille du testicule et une diminution du taux de prolifération des cellules de Sertoli immatures (Escott *et al.*, 2014; Lucas *et al.*, 2014; Pitetti *et al.*, 2013). La signalisation insuline serait une des principales voies impliquées dans la mise en place du stock de cellules de Sertoli (Griffeth *et al.*, 2014). Chez le poisson zèbre, la protéine IGF-3 a été détectée dans les cellules de Sertoli ainsi que dans les spermatogonies de type A. L'ajout de FSH (Hormone Folliculo-Stimulante) dans des cultures *ex-vivo* de tissus testiculaires entraîne une augmentation des transcrits IGF-3, aussi bien dans les spermatogonies A que dans les cellules de Sertoli, et s'accompagne d'une hausse du taux de prolifération des spermatogonies A. L'utilisation d'inhibiteurs de MAPK ou du récepteur IGF-3R inverse ce processus et entraîne une diminution du taux de prolifération des spermatogonies A. Ces

résultats suggèrent que la FSH stimule la production d'IGF-3 notamment dans les cellules de Sertoli et que, l'action de cet IRP via le récepteur IGF-1R et la voie MAPK induit une augmentation de la prolifération des spermatogonies A (Cannarella *et al.*, 2018; Nóbrega *et al.*, 2015). Chez *C. elegans*, les individus mutants pour *daf-2*, le récepteur aux insulines, présentent une forte diminution du nombre de progéniteurs de la lignée germinale. Ce phénotype peut être inversé en diminuant l'activité de *daf-18* (PTEN) ou de *daf-16* (FOXO), démontrant que la voie PI3K est impliquée dans les proliférations germinales chez cette espèce (Cannarella *et al.*, 2018; Hubbard, 2011).

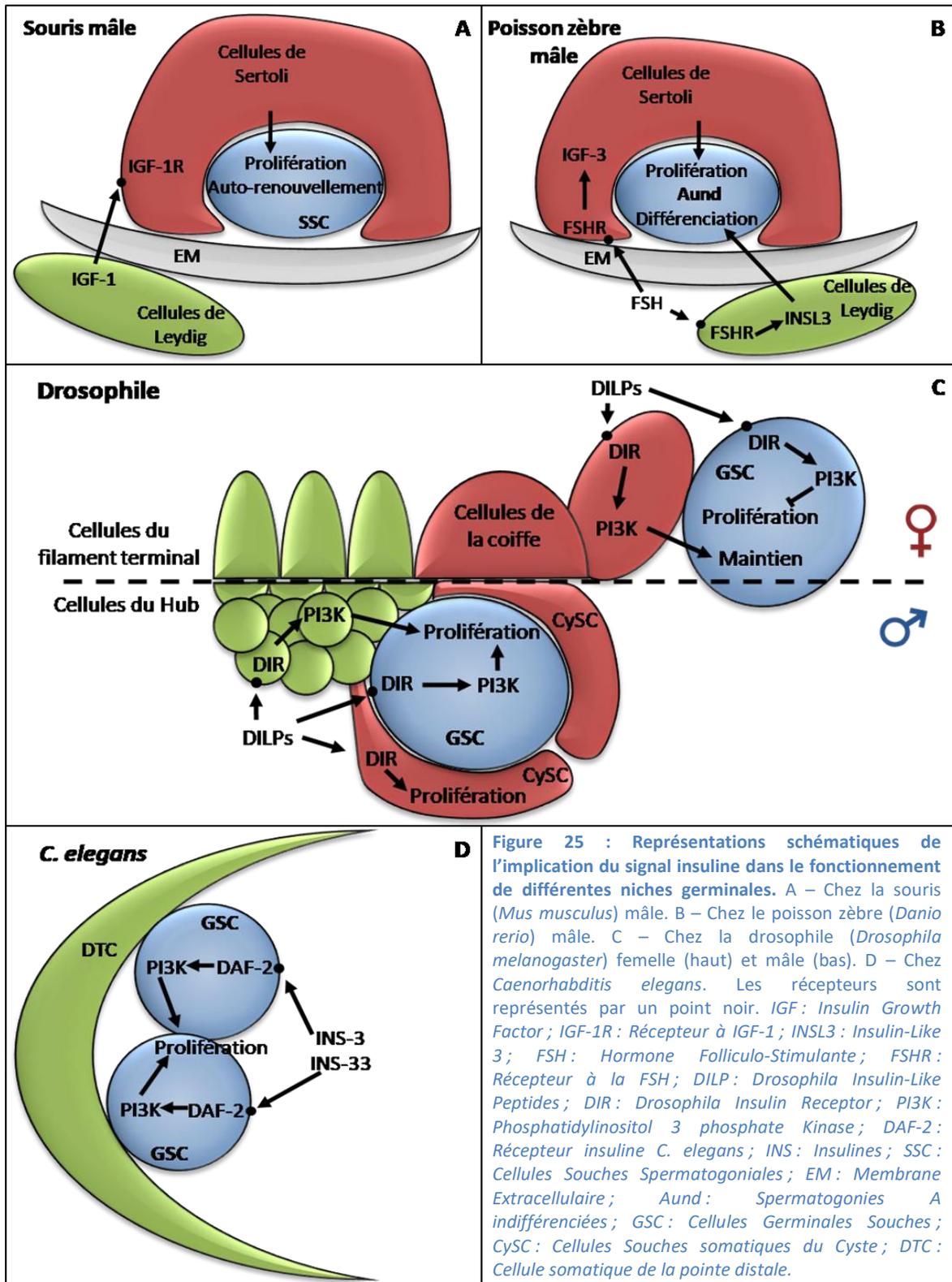
4.2.4. IRPs et régulation de la niche germinale

Les voies des IRPs font partie des voies de signalisation impliquées dans la régulation des GSC. Leur action sur le fonctionnement des GSC permet de réguler leurs capacités de prolifération, d'auto-renouvellement et de différenciation chez les vertébrés et les invertébrés. Ainsi, la régulation du taux de prolifération des GSC semble être une fonction conservée des IRPs chez les Bilatériens. Selon les espèces, l'action des IRPs sur les GSC peut être directe ou indirecte, par l'intermédiaire des cellules somatiques localisées dans la niche germinale (Figure 25).

Ainsi, chez la souris mâle, Kubota *et al.* (2004) ont montré *in vitro* que l'ajout d'IGF-1 sur des SSC en culture entraîne une augmentation de leur taux de prolifération. De plus, Huang *et al.* (2009) ont également montré que l'IGF-1, produit par les cellules de Leydig, est impliqué dans le maintien de la pluripotence des SSC. En effet, le blocage de la phosphorylation du récepteur IGF-1R ou de la voie de signalisation PI3K par des inhibiteurs réduit notablement l'expression des marqueurs de l'état souche Oct-4, Blimp1 (PRDM1) et Nanog. Les cellules de Sertoli exprimant le récepteur IGF-1R serviraient d'intermédiaire pour transmettre le signal IGF aux SSC (Figure 25A) (Oatley et Brinster, 2012; Pitetti *et al.*, 2013).

Comme énoncé précédemment, chez le poisson zèbre mâle, la production d'IGF-3 est stimulé par la FSH. L'augmentation d'IGF-3 dans les cellules de Sertoli induit la prolifération et la différenciation des SSC qui expriment le récepteur IGF-3R (Nóbrega *et al.*, 2010, 2015). De plus, sous contrôle de la FSH, les cellules de Leydig expriment l'INSL3 (Insulin-like 3) qui induirait la différenciation des SSC (Figure 25B) (Assis *et al.*, 2016; Crespo *et al.*, 2016).

Chez la drosophile, les DILPs sont produits dans les IPC (Insulin-Producing Cells) et transportés dans les tissus périphériques jusqu'à leur récepteur (Ikeya *et al.*, 2002). Chez les mâles, la mutation du récepteur aux DILPs ou l'ablation des cellules neuroendocrines produisant les DILPs entraîne une réduction rapide du taux de division des GSC dont le nombre diminue. Ceci induit la stérilité des animaux (McLeod *et al.*, 2010; Ueishi *et al.*, 2009). De même, le nombre de CySC (Cellules Souches somatiques du Cyste) diminue (Shim *et al.*, 2013). La régulation des GSC par l'insuline peut être directe ou indirecte par l'intermédiaire des cellules du hub (Figure 25C) (de Cuevas et Matunis, 2011; Shim *et al.*, 2013).



Chez les drosophiles femelles, les voies des IRPs sont aussi impliquées dans le maintien des GSC. Les mutants pour le récepteur aux DILPs présentent une diminution de la taille et un retard de développement de la gonade (LaFever et Drummond-Barbosa, 2005). Ces animaux contiennent peu de GSC à l'éclosion et leur nombre diminue plus rapidement chez les adultes par comparaison aux individus contrôles (Hsu et Drummond-Barbosa, 2009). Cependant, ces auteurs n'observent pas d'augmentation de la mort cellulaire des GSC chez les mutants, ce qui suggère que la perte de GSC serait due à une augmentation du taux de différenciation. Le signal insuline contrôlerait le maintien des GSC chez les drosophiles femelles (Figure 25C). Les ligands DILP2, DILP3 et DILP5 présentent des effets compensatoires. En effet, la délétion de DILP2 entraîne une augmentation de la transcription de DILP3 et DILP5 qui pallient au manque de DILP2 (Nässel *et al.*, 2015). Hsu et Drummond-Barbosa (2011) ont montré que la voie PI3K du signal IRP agit avec celle de Notch pour réguler la niche germinale. En effet, les DILPs stimulent les cellules de la coiffe qui possèdent des récepteurs aux DILPs. Celles-ci deviennent alors réceptives au ligand Notch impliqué dans la régulation des GSC. Les DILPs peuvent également agir directement sur les GSC qui possèdent des récepteurs de ces ligands (Hsu et Drummond-Barbosa, 2011). En résumé, les voies des IRPs agissent sur la niche germinale chez la drosophile, régulant à la fois le maintien des GSC et leur taux de prolifération.

Chez *C. elegans*, les IRPs sont exprimés dans des neurones de la région antérieure. Les mutants pour *Ins-3* et *Ins-33* présentent un phénotype similaire à celui des mutants du récepteur de l'insuline nommé DAF2. La zone contenant les GSC est réduite et le taux des cellules germinales est altéré (Hubbard, 2011; Michaelson *et al.*, 2010). Les insulines *INS-3* et *INS-33* peuvent se lier à leur récepteur (DAF-2), localisé dans les GSC, et déclencher les cascades de phosphorylations de la voie PI3K comme décrit précédemment. Akt/PKB pourra donc phosphoryler FOXO (DAF-16) qui sera dégradé dans le cytoplasme et ne pourra pas inhiber la prolifération des GSC (Figure 25D) (Hubbard, 2011).

Chez la planaire, *S. mediterranea*, un peptide de la famille des insulines (ILP-1) a été identifié et son expression est observée au niveau des lobes testiculaires (Collins *et al.*, 2010; Miller et Newmark, 2012). Des approches par RNA interférence portant soit sur ILP-1 soit sur son récepteur montrent une diminution de taille des lobes testiculaires associée à un arrêt de la spermatogenèse. Parallèlement, les auteurs ont observé une diminution du taux de prolifération des néoblastes qui pourrait refléter l'effet de l'ILP-1 sur les GSC.

4.2.5. Intégration des signaux nutritionnels

Chez la plupart des Métazoaires, le métabolisme énergétique influence les fonctions de reproduction. Chez les mammifères, par exemple, l'insuline qui est un régulateur clé du métabolisme énergétique agit également sur la fertilité mâle et femelle (Dupont *et al.*, 2014). De même, chez les insectes, les apports nutritifs via les voies des IRPs ont un impact sur le développement ovocytaire (Badisco *et al.*, 2013). Chez la planaire *S. mediterranea*, l'alimentation et les effecteurs des voies des IRPs influencent la dynamique des lignées germinales mâle et femelle (Miller et Newmark, 2012). Les voies de signalisation de l'insuline permettent généralement de faire le lien entre le statut nutritif de l'individu et le développement de la lignée germinale. Cet effet peut cibler en particulier le fonctionnement de la niche germinale.

Chez la drosophile femelle, un régime alimentaire appauvri peut engendrer une perte de GSC, à l'instar des individus mutants pour le récepteur DIR. La surexpression de DILP2 permet de contrebalancer cet effet (Hsu et Drummond-Barbosa, 2009). De plus, DILP5 est exprimé dans le tractus digestif et l'ovaire de drosophile. Sa transcription diminue lorsque les animaux sont mis à jeun suggérant son implication dans le lien entre nutrition et reproduction (Ikeya *et al.*, 2002; Nässel *et al.*, 2015). Dans le cas d'un apport nutritif favorable, les DILPs peuvent stimuler la prolifération des GSC de façon directe via les récepteurs insuline exprimés par ces cellules ou de façon indirecte en agissant sur les cellules de la coiffe qui possèdent également les récepteurs DIR (Figure 25C). Les GSC mâles chez la drosophile sont également impactées par le statut nutritif. De même que pour la femelle, un régime appauvrit chez les mâles entraîne une forte réduction du nombre de GSC mais également des CySC au sein de la niche. Ce même effet est observé lors de la mutation du récepteur DIR. L'activation de la signalisation insuline permet d'inverser cet effet (McLeod *et al.*, 2010; Shim *et al.*, 2013).

Les connaissances acquises chez les mollusques concernant l'influence de l'alimentation sur la reproduction et l'implication des insulines dans ce processus seront détaillées dans l'introduction du chapitre III.

5. Le modèle biologique Crassostrea gigas

5.1. L'intérêt économique

L'huître creuse *Crassostrea gigas* représente un atout économique mondial avec 5,2 millions de tonnes produites par an (FAO, 2014). La France est le quatrième pays producteur d'huître creuse derrière la Chine, Le Japon et la Corée. La majorité de la production est réalisée entre Novembre et Janvier. La Normandie est le troisième site de production après le Poitou-Charentes et la Bretagne. Deux méthodes principales d'obtention de juvéniles d'huître coexistent en France. Les ostréiculteurs peuvent s'approvisionner en juvéniles en installant des collecteurs en mer, à proximité des parcs de production des adultes en période de reproduction, afin de réaliser du captage naturel. Les larves d'huître vont en effet venir se fixer à ces collecteurs au moment de la métamorphose, après la période de vie larvaire en pleine eau. Cette méthode permet l'acquisition d'individus dans le milieu naturel sans aucune forme de sélection. Les ostréiculteurs peuvent par ailleurs acquérir des juvéniles auprès des écloseries. Pour produire ces juvéniles, les écloseries sélectionnent des individus et les conditionnent pour qu'ils atteignent la maturité sexuelle et se reproduisent. Une meilleure connaissance des mécanismes de reproduction de l'huître, en particulier des étapes précoces de gamétogenèse qui ont un impact notamment sur l'effort de reproduction des individus (Jouaux *et al.*, 2013), est favorable à la maîtrise du cycle de production de cette espèce.

5.2. La position phylogénétique

C. gigas est un Métazoaire Bilatérien qui présente une symétrie bilatérale (Raven *et al.*, 2007). Elle appartient à l'embranchement des Lophotrochozoaires caractérisé par un développement de type spiralé, un lophophore puis une couronne ciliée autour d'une larve trochophore (Giribet, 2008). Les données acquises chez les espèces Lophotrochozoaires sont intéressantes d'un point de vue phylogénétique en comparaison avec les données disponibles chez les Ecdysozoaires et les Deutérostomiens, qui possèdent la majorité des espèces modèles. *C. gigas* appartient au phylum des Mollusques qui comportent plus de 100 000 espèces répertoriées et classifiées en sept classes (Polyplacophora, Chaetodermomorpha, Neomeniomorpha, Cephalopoda, Scaphopoda, Gastropoda et Bivalvia) (Kocot *et al.*, 2011). L'huître creuse fait partie des bivalves qui se caractérisent par

la présence d'une coquille à deux valves. Un des atouts de *C. gigas* tient au fait que son génome ait été séquencé en 2012 et que plusieurs transcriptomes soient disponibles chez cette espèce (Rivière *et al.*, 2015; Zhang *et al.*, 2012; Zhao *et al.*, 2012). Parmi les mollusques, peu d'espèces possèdent de séquences disponibles aussi complètes, ce qui facilite la recherche et la caractérisation des acteurs moléculaires.

5.3. Les Cellules Germinales Primordiales (PGC) chez l'huître

Au cours de l'embryogenèse, l'origine des PGC a été étudiée grâce au marqueur *oyvlg* (oyster-vasa-like gene) identifié chez *C. gigas* (Fabioux *et al.*, 2004a). Ces auteurs ont montré que *oyvlg* était présent au pôle végétatif de l'ovocyte puis transmis dans l'embryon. Ces transcrits de *oyvlg* restent groupés dans un seul blastomère, le plus grand macromère, dès le stade deux cellules et perdure jusqu'au stade blastula (Figure 26). A partir du stade gastrula, les transcrits *oyvlg* sont observés dans deux groupes de cellules qui pourraient correspondre à des PGC. Ces PGC sont localisés de façon symétrique autour de la ligne médiane de l'embryon. Ces cellules vont ensuite migrer le long de la masse viscérale pour se placer, au stade véligère, de part et d'autre de la glande digestive, lieu de la future gonade. Ainsi, d'après Fabioux *et al.* (2004a), ces PGC pourront se différencier en GSC après la métamorphose afin d'assurer la gamétogenèse.

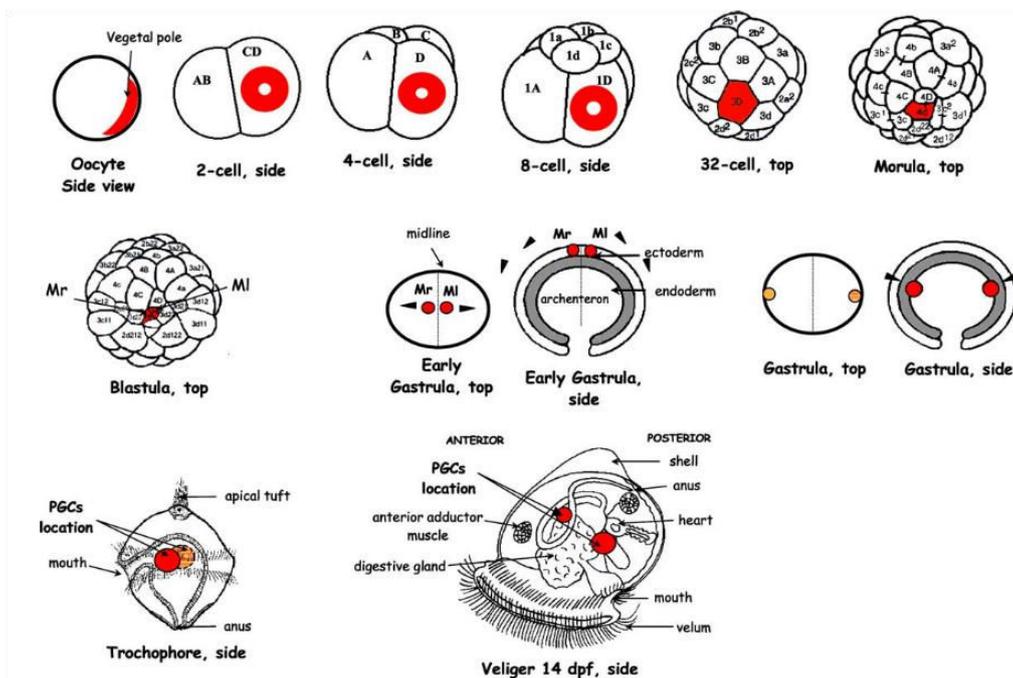


Figure 26 : Représentation schématique de la localisation des transcrits *oyvlg* lors du développement de *C. gigas*. En rouge/orange, la localisation de l'ARN *oyvlg* au cours de l'embryogenèse. Le sens de migration est indiqué par les pointes de flèche. Fabioux *et al.*, 2004a.

5.4. La reproduction de l'huître creuse

L'huître creuse est un animal hermaphrodite alternatif irrégulier présentant un cycle de reproduction annuel et saisonné. L'huître a la capacité de changer de sexe d'une année sur l'autre et quelques cas d'hermaphrodisme simultané peuvent également être observés. Park *et al.* (2012) ont montré que chez une population d'huîtres *C. gigas* collectées en Corée, 40,2% des animaux changeaient de sexe dont 66,1% passaient de femelle à mâle et 21,1% de mâle à femelle. La capacité de l'huître a reformé les lignées gonadiques, somatiques et germinales, chaque année suite ou non à un changement de sexe, mène à la question de l'identification et du fonctionnement de ses GSC.

5.4.1. L'aire gonadique adulte et la gamétogenèse

La gonade chez l'huître est située dans l'aire gonadique qui comprend également du tissu conjonctif de réserve (Figure 27) (Franco, 2009; Jouaux, 2010). Cette aire gonadique est bordée extérieurement par des pores évacuateurs qui possèdent un épithélium cilié en face de l'épithélium germinal. Ils permettent l'émission des gamètes matures lors de la ponte des animaux.

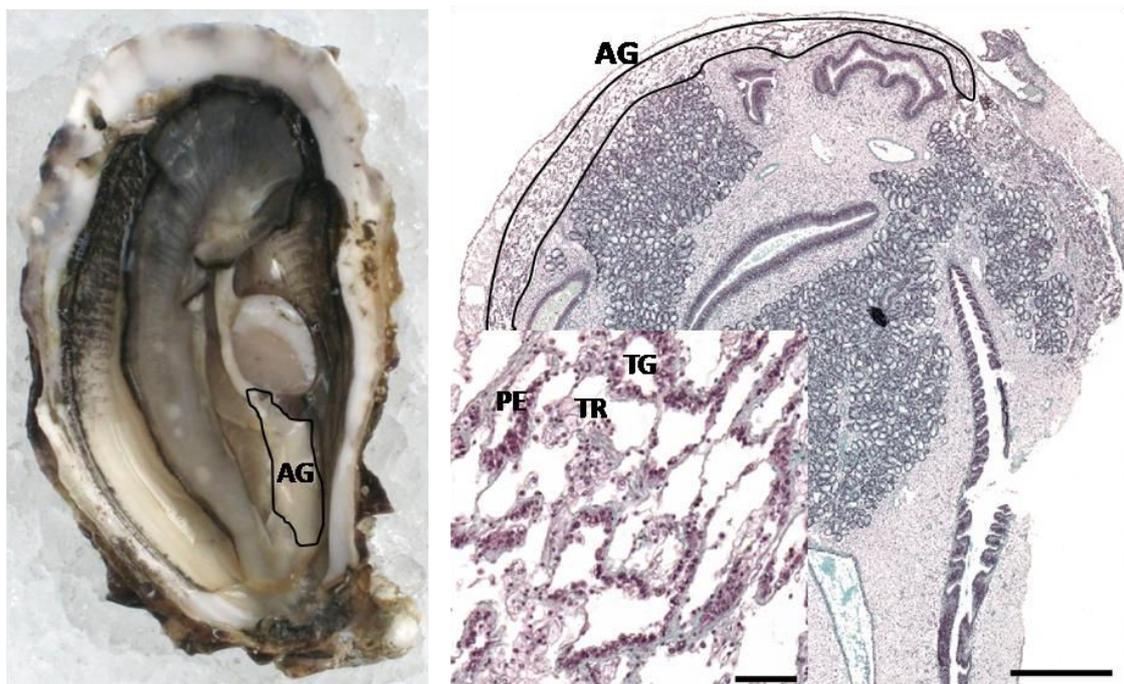


Figure 27 : Organisation générale de l'aire gonadique chez l'huître creuse adulte. Localisation de l'aire gonadique (AG) chez l'huître *Crassostrea gigas* à gauche et lame virtuelle d'une coupe histologique transversale complète d'huître à droite colorée au Trichrome de Prenant-Gabe. L'aire gonadique est délimitée par un trait noir. Insert – L'aire gonadique comprend la gonade formée de tubules gonadiques (TG) et le tissu conjonctif de réserve (TR). Les pores évacuateurs (PE) sont nécessaires à l'émission des gamètes lors de la ponte. A/P : axe antéro-postérieur ; V/D : axe dorso-ventral ; AG : Aire gonadique ; TG : Tubules Gonadiques ; TR : Tissu conjonctif de réserve. Echelles : 1mm (droite), 50µm (insert).

Le développement du tissu de réserve dans l'aire gonadique est inversement corrélé à l'avancée en gamétogenèse (Berthelin *et al.*, 2000a, 2000b; Heude Berthelin *et al.*, 2001). La progression des tubules gonadiques en lien avec la régression du tissu de réserve montre une relation étroite entre les réserves énergétiques et la reproduction de l'huître creuse. En effet, le compartiment de réserve accumule l'énergie sous forme de glycogène pour l'allouer par la suite à la reproduction. La régulation de cette balance énergétique a fait l'objet de recherches et des acteurs moléculaires ont ainsi été identifiés chez *C. gigas*. Afin de mieux comprendre les variations des tissus de l'aire gonadique au cours du cycle de gamétogenèse chez *C. gigas*, une étude a été réalisée en utilisant le marqueur de prolifération PCNA (Proliferating Cell Nuclear Antigen). L'expression de PCNA a été mesurée par PCR quantitative après microdissection laser permettant de séparer tissu de réserve et tubules gonadiques. Elle s'est révélée plus importante dans les tubules gonadiques, aussi bien mâles que femelles, en comparaison avec le tissu de réserve (Franco *et al.*, 2010).

L'avancée en gamétogenèse chez l'huître creuse peut être divisée en quatre stades. Le premier stade de gamétogenèse, le stade 0, correspond à la phase quiescente qui a lieu en Automne/Hiver sur les côtes françaises (Figure 28). Les tubules gonadiques sont très peu développés et ne contiennent que quelques cellules germinales précoces et cellules somatiques intragonadiques (Figure 28A). A cette période, le tissu de réserve est en restructuration. Les individus ayant eu une ponte partielle l'année précédente peuvent présenter quelques gamètes résiduels qui seront dégradés. A ce stade de repos sexuel, aucun immunomarquage PCNA n'a été détecté dans les tubules gonadiques (Figure 28B) (Franco *et al.*, 2010).

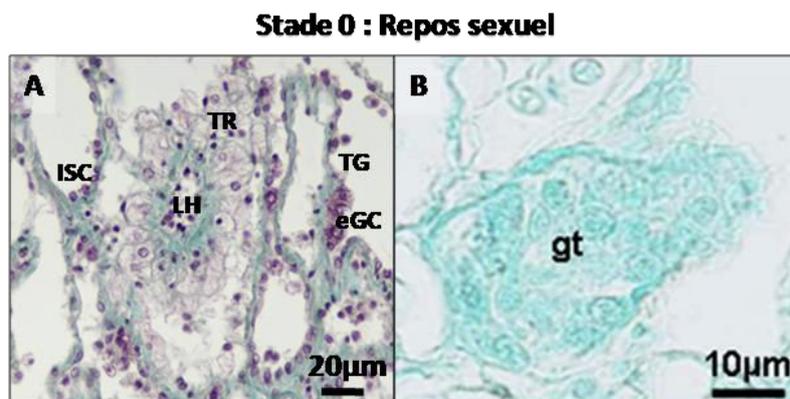


Figure 28 : Coupes histologiques d'aire gonadique au stade 0 de gamétogenèse chez *C. gigas*. A – Coupe histologique colorée au Trichrome Prenant-Gabe. B – Coupe histologique après immunomarquage PCNA d'après Franco *et al.*, 2010. LH : lacune hémolympatique ; eGC : cellule germinale précoce ; ISC : cellule somatique intragonadique ; TG/gt : tubule gonadique ; TR : tissu de réserve.

Le stade suivant, stade 1, est la phase de proliférations goniales à partir de laquelle le sexe peut être, ou non, identifié selon l'avancement des proliférations (Figure 29). Les tubules se développent et les spermatogonies ou les ovogonies prolifèrent afin de former un pool de cellules germinales qui pourra se différencier au stade suivant. Le tissu de réserve est alors reconstitué et forme un réseau conjonctif dense qui occupe la majorité de l'aire gonadique (Figure 29A, B). Au début des stades 1 de gamétogenèse mâle et femelle, un immunomarquage de PCNA a été observé dans les cellules myoïdes (Figure 29C, D). Ce marquage est probablement lié à la capacité mitotique des cellules myoïdes pour permettre l'expansion des tubules gonadiques dans l'aire gonadique. Le marquage PCNA est observé dans les ovogonies ainsi que dans les spermatogonies à la fin de ce stade de proliférations goniales. De plus, lors des stades précoces de gamétogenèse, l'expansion du tissu de réserve a pu être illustrée par la présence du marquage PCNA dans certaines cellules de réserve. Ces données ont permis une meilleure compréhension du remodelage des tissus de l'aire gonadique, notamment lors des étapes précoces de gamétogenèse, mais d'autres recherches sont nécessaires à l'identification de cellules souches à l'origine de ces tissus.

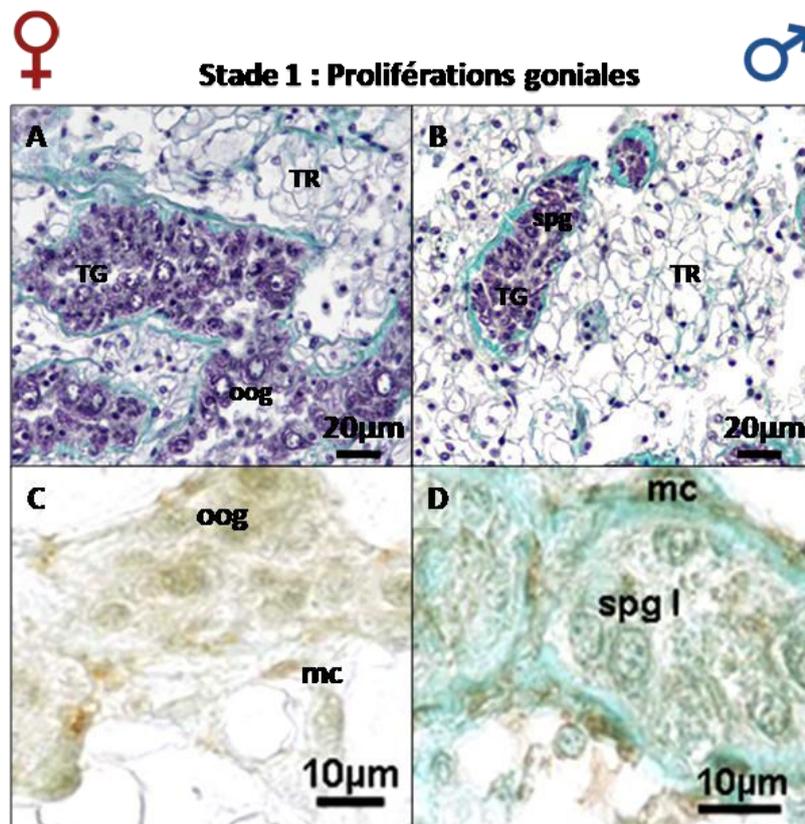


Figure 29 : Coupes histologiques d'aire gonadique au stade 1 de gamétogenèse chez *C. gigas*. A et B – Coupes histologiques colorées au Trichrome Prenant-Gabe aux stades 1 femelle et mâle respectivement. C et D – Coupes histologiques après immunomarquage PCNA aux stades 1 femelle et mâle respectivement d'après Franco *et al.*, 2010. TR : Tissu de réserve ; TG : Tubule gonadique ; oog : ovogonies ; spg : spermatogonies ; mc : cellules myoïdes.

Le stade 2 de gamétogenèse correspond au stade de différenciation germinale. En effet, les cellules germinales entrent en méiose et le tissu de réserve régresse en corrélation avec l'avancée en gamétogenèse. Chez la femelle, le stade 2 correspond à une étape de vitellogenèse (Figure 30A). Chez le mâle, les cellules germinales visibles sont les spermatogonies, spermatocytes, spermatides et quelques spermatozoïdes (Figure 30B).

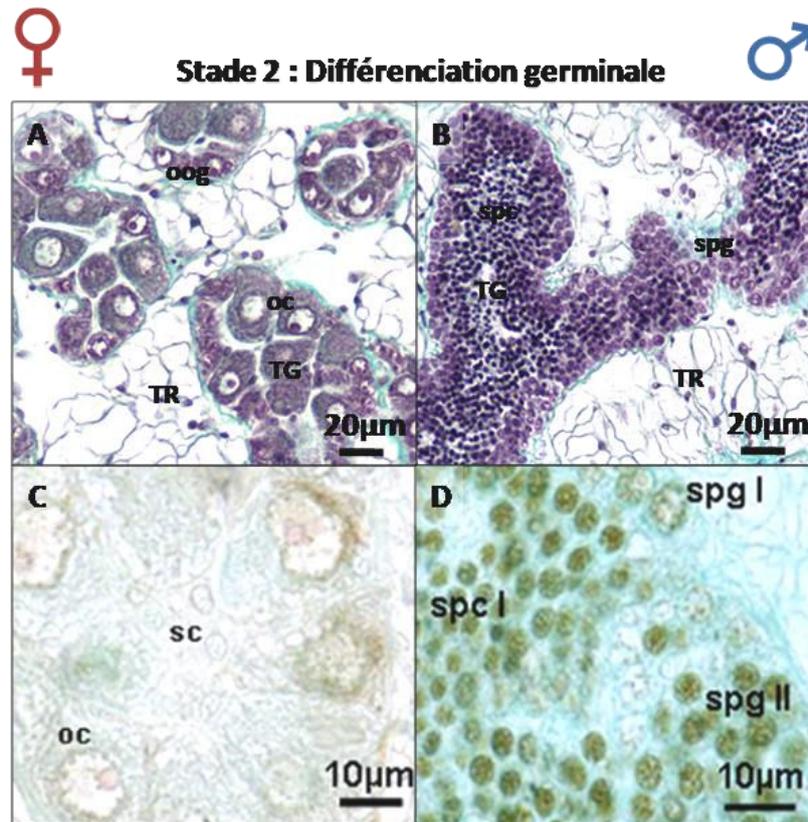


Figure 30 : Coupes histologiques d'aire gonadique au stade 2 de gamétogenèse chez *C. gigas*. A et B – Coupes histologiques colorées au Trichrome Prenant-Gabe aux stades 2 femelle et mâle respectivement. C et D – Coupes histologiques après immunomarquage PCNA aux stades 2 femelle et mâle respectivement d'après Franco *et al.*, 2010. TR : Tissu de réserve ; TG : Tubule gonadique ; oog : ovogonies ; spg : spermatogonies ; spc : spermatocytes ; oc : ovocytes ; sc : Cellules somatiques.

Chez les femelles, le marquage PCNA est plus faible que chez les mâles et seulement associé aux cellules germinales (Figure 30C). Le marquage observé dans le cytoplasme des ovocytes en développement pourrait être dû à un transfert de la protéine PCNA du noyau vers le cytoplasme (Franco *et al.*, 2010). Chez les mâles, deux types de spermatogonies ont été identifiées (Franco *et al.*, 2008) et il est apparu que le marquage PCNA différait selon ces types. En effet, au stade 1 de gamétogenèse, PCNA est présent dans les deux types alors qu'aux stades 2 et 3 seuls les spermatogonies II sont positives pour PCNA révélant une différence temporelle dans leur activité mitotique (Figure 30D). De façon intéressante, certaines cellules somatiques intragonadiques sont apparues positives pour PCNA suggérant

une prolifération de ces cellules en lien avec un rôle de support de la gamétogenèse mâle, à l'image des cellules de Sertoli.

Le dernier stade, le stade 3, a lieu en été et correspond à la maturité sexuelle des individus. Les tubules gonadiques occupent la majorité de l'aire gonadique et les gamètes matures sont prêts à être émis. Quelques cellules précoces restent encore observables au bord des tubules gonadiques. Durant ce stade de gamétogenèse, le tissu de réserve n'est finalement présent que sous forme de quelques îlots de cellules de réserve (Figure 31A, B). Le marquage PCNA au stade 3 femelle est plus hétérogène et peut être détecté dans les noyaux et le cytoplasme des ovocytes (Figure 31C). Chez les mâles matures, le marquage PCNA est restreint aux spermatogonies de type II et aux spermatocytes I. Le marquage PCNA des spermatogonies II reflète une prolifération continue de ces cellules au cours du cycle (Figure 31D) (Franco *et al.*, 2010).

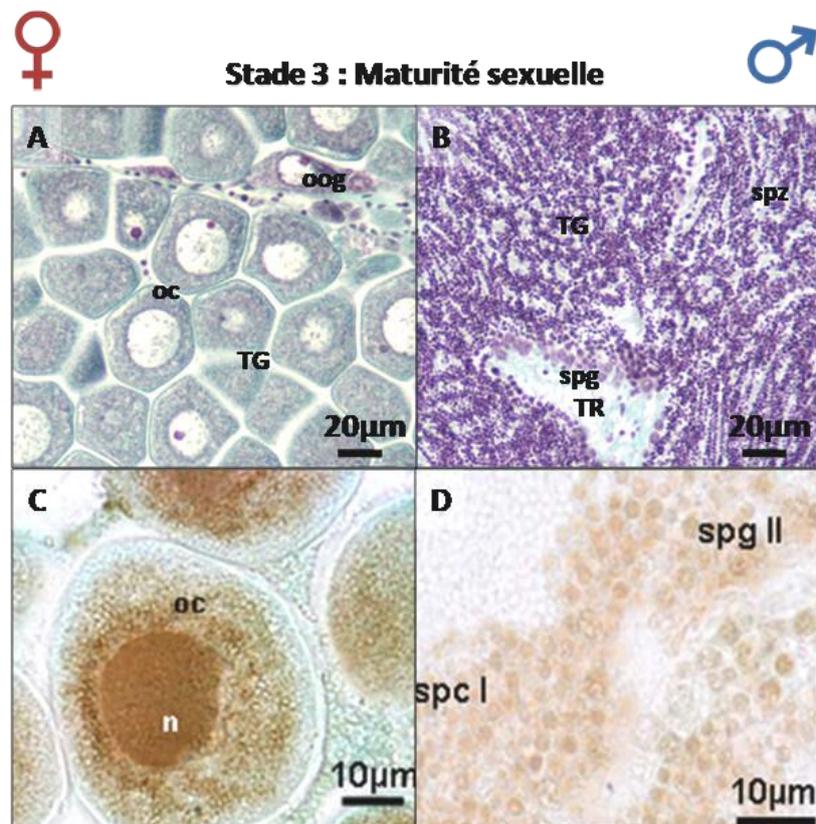


Figure 31 : Coupes histologiques d'aire gonadique au stade 3 de gamétogenèse chez *C. gigas*. A et B – Coupes histologiques colorées au Trichrome Prenant-Gabe aux stades 3 femelle et mâle respectivement. C et D – Coupes histologiques après immunomarquage PCNA aux stades 3 femelle et mâle respectivement d'après Franco *et al.*, 2010. TR : Tissu de réserve ; TG : Tubule gonadique ; oog : ovogonies ; spg : spermatogonies ; oc : ovocytes ; spc : spermatocytes ; spz : spermatozoïdes ; n : noyau.

5.4.2. Structure de la gonade

La gonade chez l'huître est formée de l'ensemble des tubules gonadiques. Ces derniers sont délimités par la lame basale, une matrice acellulaire, elle-même entourée de cellules myoïdes (Figure 32). Les cellules germinales et les cellules somatiques intragonadiques sont observées à l'intérieur des tubules gonadiques. Les cellules somatiques intragonadiques peuvent être identifiées grâce à la forme digitée de leurs noyaux ainsi qu'à leurs prolongements cytoplasmiques, où des phagolysosomes peuvent être observés. Des jonctions cellulaires ont été observées entre les cellules somatiques (Figure 32, insert) mais pas avec les cellules germinales. Les cellules germinales précoces sont localisées au bord des tubules gonadiques. Lors de la gamétogenèse, les cellules germinales se différencient de façon centripète pour former les gamètes qui sont localisées au centre des tubules. La spermatogenèse a été caractérisée chez *C. gigas*, ce qui a permis d'identifier deux types de spermatogonies, les spermatocytes primaires et secondaires, les spermatides ainsi que les spermatozoïdes (Franco *et al.*, 2008).

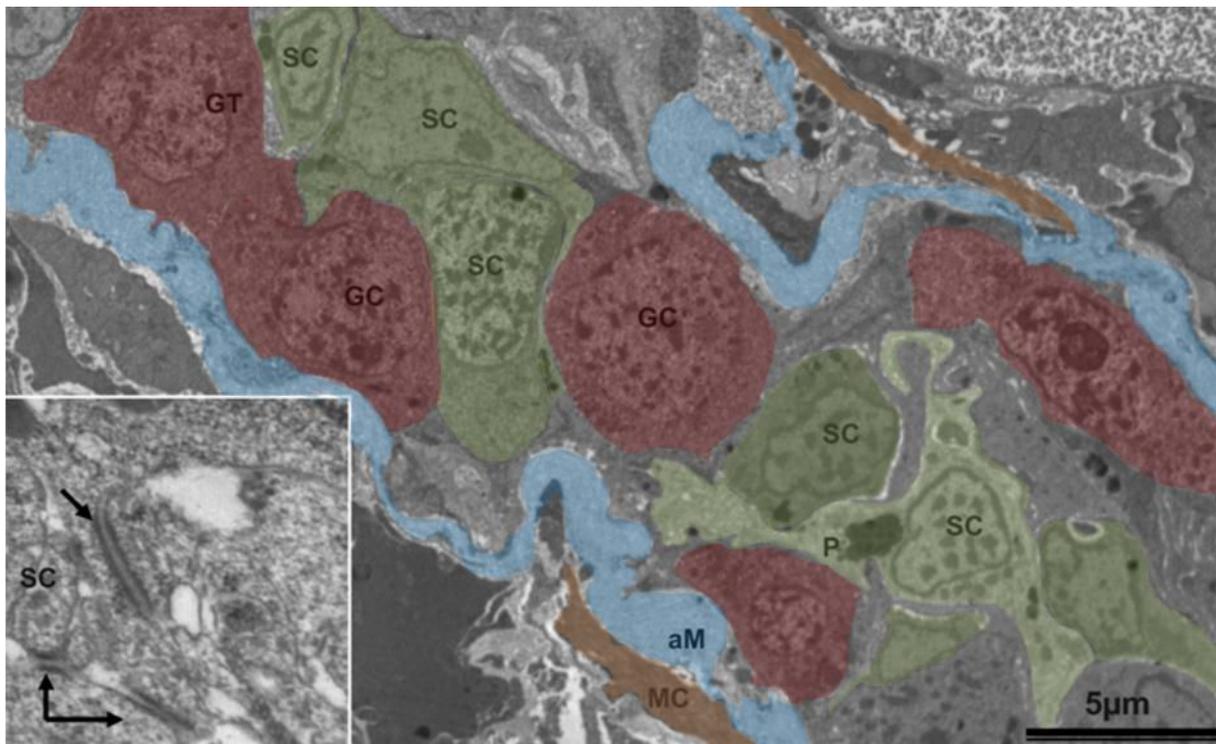


Figure 32 : Ultrastructure d'un tubule gonadique chez *C. gigas*. D'après Franco *et al.*, 2008. TG : Tubule gonadique ; GC, rouge : Cellules Germinales ; aM, bleu : Lame basale ; SC, vert : Cellules somatiques intragonadiques ; P : Phagolysosomes ; flèches : jonctions entre les cellules somatiques intragonadiques.

6. Objectifs de thèse

Le cycle de reproduction et le lien de ce dernier avec les réserves énergétiques a été bien étudié chez *C. gigas*. Cependant, la compréhension des étapes précoces de gamétogenèse reste encore incomplète. La capacité de *C. gigas* à renouveler chaque année sa lignée germinale qui peut être soit mâle soit femelle, amène le questionnement de la structure et du fonctionnement de la niche germinale chez cette espèce. Par ailleurs, plusieurs études ont également montré la conservation, chez l'huître, d'acteurs et de voies de signalisation des IRPs identifiés chez les groupes phylogénétiques majoritairement étudiés, vertébrés ou invertébrés, et l'implication des IRPs dans le contrôle des étapes précoces de gamétogenèse. L'acquisition des données génomiques et transcriptomiques chez *C. gigas* est un atout pour confirmer la conservation des acteurs du signal IRP et leur implication dans le contrôle de la gamétogenèse chez les Lophotrochozoaires.

6.1. Identification des GSC potentielles et de leur niche

Le premier objectif de cette thèse sera d'identifier et de localiser les cellules souches pouvant être à l'origine des lignées germinales. Au cours de ce travail, la caractérisation des GSC sera basée sur des critères morphologiques et moléculaires. Les acteurs moléculaires incluent le marqueur germinale Vasa mais également plusieurs marqueurs de cellules souches qui ont été analysés notamment grâce à des techniques d'immunohistochimie. De plus, le microenvironnement des cellules souches permettant la régulation de leurs capacités de prolifération, d'auto-renouvellement et de différenciation sera caractérisé chez *C. gigas* dans le but d'identifier des types cellulaires pouvant appartenir à la niche germinale et pouvant avoir un rôle dans son fonctionnement.

6.2. Etude évolutive et fonctionnelle du système insuline

La disponibilité des banques de données génomiques permet de rechercher de façon exhaustive l'ensemble des effecteurs appartenant à la famille des IRPs et d'appréhender leur histoire évolutive. L'implication fonctionnelle des différents ligands dans le contrôle de la reproduction ou de processus associés peut être abordée en suivant leur expression au cours du cycle de gamétogenèse. La caractérisation moléculaire des éléments des voies de signalisation, incluant le récepteur et les multiples effecteurs du signal permettra d'évaluer la conservation des voies de signalisations de l'insuline chez les mollusques et les données

d'expression pourront renseigner sur les rôles potentiels des IRPs dans le contrôle de la reproduction.

6.3. Effet de l'alimentation sur les cellules germinales précoces, implication du signal insuline

D'après les études antérieures, il est apparu que les IRPs sont impliqués dans le contrôle de la balance entre réserves et reproduction chez l'huître. Le dernier objectif de ce travail sera donc de déterminer si le niveau d'alimentation a un impact sur le fonctionnement des cellules germinales précoces et même plus particulièrement sur le devenir des GSC et leur niche et si les IRPs sont impliqués dans ce contrôle.

Chapitre I :
Identification des cellules
germinales souches potentielles
et de leur niche

(La deuxième partie de ce chapitre a fait l'objet d'une publication dans *Histochemistry et Cell Biology*.)

Chapitre I :

Identification de cellules germinales souches potentielles et de leur niche

1. Introduction

Peu de données existent chez les mollusques en ce qui concerne les GSC et leur niche. Ainsi, pour rechercher et identifier ce type particulier de cellules souches, il convient de prendre en compte l'ensemble des données acquises chez les mollusques sur des cellules souches, notamment concernant les marqueurs de pluripotence. Les caractéristiques morphologiques des cellules germinales précoces établies chez les mollusques, leurs propriétés mitotiques ainsi que les marqueurs spécifiques des GSC conservés sont également d'intérêt dans cette recherche. Des données plus générales sur des acteurs potentiellement impliqués dans la niche germinale doivent également être prises en compte.

1.1. Marqueurs de pluripotence et cellules souches chez les mollusques

Afin de caractériser les cellules souches, des marqueurs de pluripotence ont été recherchés chez les mollusques. Les séquences des facteurs de Yamanaka, Sox2, Oct4, KLF4 et c-myc sont ainsi présentes dans le génome de plusieurs espèces de mollusques. Le dernier marqueur de pluripotence, Nanog, n'a pas été identifié chez ces espèces ce qui va dans le sens de la non conservation de ce facteur chez les espèces invertébrés.

1.1.1. Chez les céphalopodes

Dans la classe des Céphalopodes, un facteur SoxB1 (Sof-SoxB1) a été identifié chez la seiche *Sepia officinalis* pendant le développement embryonnaire (Focareta et Cole, 2016). L'expression de Sof-SoxB1 est retrouvée dans les lobes à l'origine du système nerveux central ce qui pourrait indiquer son implication, comme chez d'autres espèces, dans la neurogenèse. Chez un calmar, *Idiosepius notoides*, quatre facteurs POU ont été identifiés par blast dans le transcriptome (Wollesen *et al.*, 2014). Après prédiction de leurs séquences protéiques, les domaines conservés spécifiques des POU ont pu être caractérisés et une analyse phylogénétique a été réalisée. Ainsi, ces facteurs ont pu être catégorisés dans les

différentes classes de la famille des protéines POU. Ils ont été nommés Ino-POU2, Ino-POU3, Ino-POU4 et Ino-POU6. Les transcrits correspondants ont été localisés par hybridation *in situ* au cours du développement de *I. notoides*. Ces facteurs apparaissent tous exprimés lors du développement embryonnaire et plus précisément dans les régions qui formeront le système nerveux central de *I. notoides*. Ces observations suggèrent l'implication, comme pour le SoxB1 de seiche, de ces facteurs POU dans le processus de neurogenèse chez les céphalopodes (Wollesen *et al.*, 2014).

1.1.2. Chez les gastéropodes

Certains effecteurs du groupe des SoxB1, incluant Sox2, ont été identifiés chez les gastéropodes. Chez l'ormeau, *Haliotis asinina*, deux transcrits ont été identifiés et leurs séquences protéiques prédites ont montré que l'un d'eux, nommé HasSox-B, appartient au groupe B, comme le montre la conservation de son domaine HMG (O'Brien et Degnan, 2000). L'expression de HasSox-B a été retrouvée dans plusieurs tissus adultes incluant les ganglions nerveux et la gonade. Les ganglions nerveux peuvent avoir un rôle dans le contrôle de la croissance et du développement, par conséquent, une conservation fonctionnelle de HasSox-B dans la régulation de ces processus peut être envisagée. L'expression de ce facteur dans la gonade suggère son implication dans les processus de reproduction (O'Brien et Degnan, 2000). Chez cette espèce, deux membres de la famille POU ont été également identifiés à savoir, HasPOU-III et HasPOU-IV, et sont classés dans les groupes III et IV des POU d'après l'analyse de leurs domaines conservés (O'Brien et Degnan, 2000, 2002). Leurs expressions sont retrouvées au cours du développement embryonnaire et plus particulièrement à partir du stade larvaire trochophore. Ce profil d'expression pourrait être en lien avec la capacité d'initier la morphogenèse lors de ce stade chez *H. asinina*. La localisation des transcrits dans les cellules embryonnaires qui formeront les cellules sensorielles et nerveuses suggère encore un rôle de ces facteurs dans la neurogenèse (O'Brien et Degnan, 2002). De plus, chez l'adulte HasPOU-III et HasPOU-IV sont exprimés dans les ganglions nerveux. Ces deux facteurs présentent des profils d'expression similaires ce qui pourrait signifier une possible interaction entre eux afin de réguler le développement comme c'est le cas chez la drosophile (O'Brien et Degnan, 2000, 2002). Chez un autre gastéropode, la patelle *Patella vulgata*, un facteur Sox a également été identifié chez l'embryon (Le Gouar *et al.*, 2004). D'après une analyse phylogénétique, il est proche des

SoxB et a ainsi été nommé PvuSoxB. Son expression au cours du développement a été suivie par hybridation *in situ* et apparaît restreinte aux cellules embryonnaires qui engendreront les structures neurales. Ces résultats suggèrent une conservation du rôle des SoxB dans les processus de neurogenèse chez les mollusques (Le Gouar *et al.*, 2004).

1.1.3. Chez les bivalves

Chez les bivalves, différents marqueurs de pluripotence sont également retrouvés. Chez le pétoncle *Patinopecten yessoensis*, sept gènes Sox ont été identifiés dont l'un codant pour une protéine SoxB1 (Yu *et al.*, 2017). Le transcrite de SoxB1 est retrouvé à un niveau plus élevé dans la gonade femelle en comparaison avec la gonade mâle, suggérant un rôle dans le développement ovarien chez cette espèce (Yu *et al.*, 2017). Chez l'huître perlière, *Pinctada fucata*, deux facteurs de transcription appartenant à la famille des POU ont pu être identifiés et nommés Pf-POU2F1 et Pf-POU3F4 (Gao *et al.*, 2016). Ces facteurs présentent les domaines conservés POUS et POUH. L'analyse de leurs expressions a été réalisée lors du développement ainsi que dans différents tissus adultes. Il a pu être montré que Pf-POU2F1 était plus exprimé lors du développement, au stade de la larve trochophore, suggérant une implication de ce facteur au stade larvaire précoce. Concernant Pf-POU3F4, son expression est plus élevée au stade juvénile, après la métamorphose, ainsi que dans le bord de manteau chez l'adulte. Le bord de manteau étant le tissu permettant la production de la coquille après métamorphose, ces observations suggèrent l'implication de ce facteur dans la formation de la coquille dès le stade juvénile de l'huître perlière *P. fucata* (Gao *et al.*, 2016). Chez une autre espèce de bivalve, le pétoncle *Chlamys farreri*, une séquence KLF4 a été retrouvée et l'analyse de son expression montre un dimorphisme sexuel en faveur des mâles. Chez ces derniers, l'expression de KLF4 augmente au cours de la spermatogenèse pour atteindre son maximum à maturité sexuelle alors que, chez les femelles, l'expression de KLF4 reste constante au cours de l'ovogenèse. Chez les mâles, lorsque l'expression de KLF4 est bloquée par ARN interférant, une diminution du nombre total de cellules germinales est observée malgré une proportion plus importante de spermatogonies. Ces résultats suggèrent que KLF4 pourrait être impliqué dans la différenciation de la lignée germinale mâle ainsi que dans le maintien des cellules germinales précoces chez *C. farreri* (Yang *et al.*, 2017).

1.1.4. Données acquises chez *Crassostrea gigas*

Chez l'huître creuse *Crassostrea gigas*, Jemaà et ses collaborateurs (2014) ont identifié un homologue de Sox2 dont la protéine a été localisée à l'aide d'un anticorps hétérologue dans les branchies ainsi que dans le liquide circulant chez les bivalves, l'hémolymphe. Les cellules positives pourraient correspondre à des cellules souches impliquées dans l'hématopoïèse. De plus, certaines d'entre elles ont également été identifiées dans les tubules gonadiques d'huîtres âgées de 8 mois, lors de leur première gamétogenèse (Cavelier *et al.*, 2017). Ce marquage a été observé dans l'épithélium germinale mais également dans certaines cellules de l'épithélium cilié des pores évacuateurs. La détection de cette protéine suggère fortement la présence de cellules souches dans ces tissus, cependant, le caractère somatique ou germinale de ces cellules souches n'est pas clairement établi. Par ailleurs, une séquence conservée de c-myc est retrouvée chez *C. gigas* (David *et al.*, 2005). L'étude du transcrite de c-myc a montré une surexpression de ce facteur en réponse à l'hypoxie. En effet, son expression augmente de façon significative après 17 jours chez les huîtres en condition d'apport réduit en oxygène. Les auteurs suggèrent un rôle de c-myc dans la régulation de la transcription des gènes en réponse à une condition de stress telle que l'hypoxie, comme cela peut être observé chez d'autres espèces (David *et al.*, 2005).

En résumé, les séquences de plusieurs marqueurs de pluripotence apparaissent conservées chez certains mollusques, incluant *C. gigas* (Figure I-1). Des facteurs appartenant au groupe des SoxB ont été identifiés dans les trois principales classes des mollusques (gastéropodes, céphalopodes, bivalves). Chez certaines espèces, ces facteurs sont proches des Sox2 identifiés chez les mammifères. De même, des facteurs de transcription appartenant à la famille des POU ont été caractérisés dans ces trois classes. D'autres expérimentations sont nécessaires chez les mollusques afin de pouvoir valider la fonction des effecteurs Sox et POU identifiés. Cependant, les résultats obtenus chez les espèces étudiées suggèrent une implication de ces facteurs dans les processus de neurogenèse, de développement embryonnaire et de reproduction. Les facteurs KLF4 et c-myc ont été identifiés chez les bivalves et, d'après les connaissances actuelles, pourraient être impliqués dans certains processus de reproduction pour KLF4 et dans la réponse au stress pour c-myc.

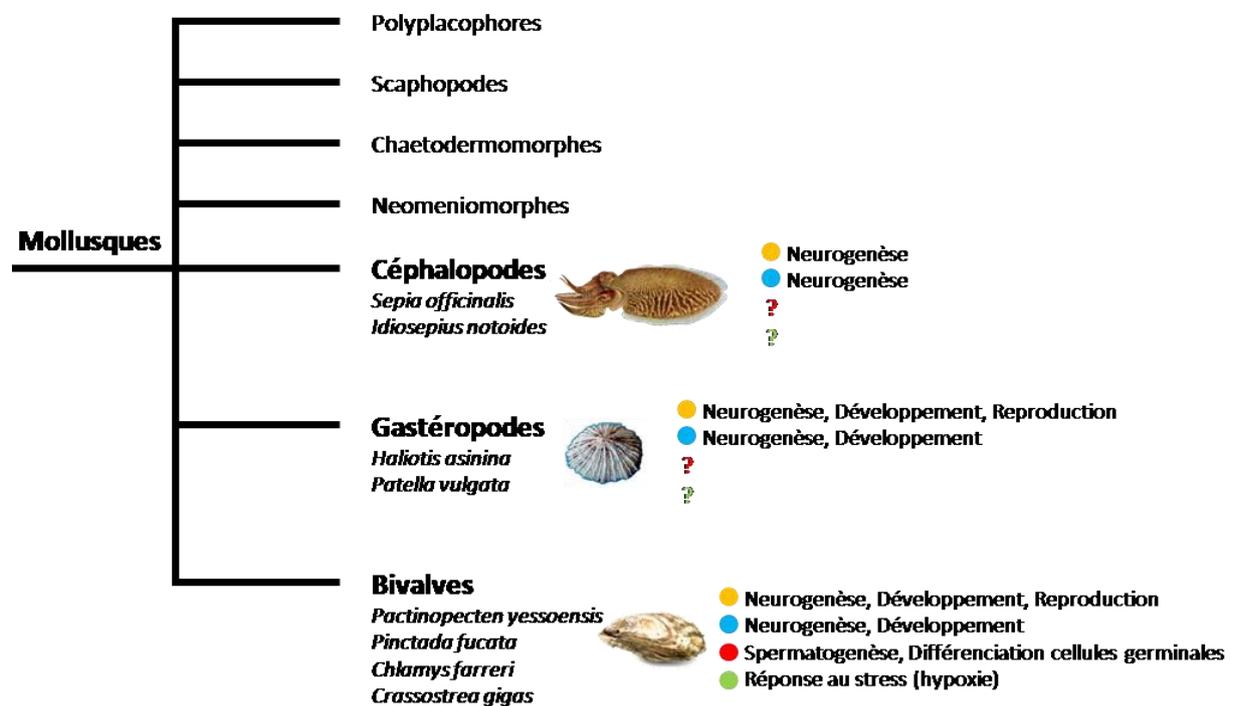


Figure I-1 : Arbre phylogénétique simplifié des mollusques résumant l'identification et les fonctions potentielles des marqueurs de pluripotence chez certaines espèces de céphalopodes, gastéropodes et bivalves. SoxB/Sox2 : jaune, POU/Oct4 : bleu, KLF4 : rouge, c-myc : vert. L'identification de séquences conservées est représentée par un rond. Les séquences qui n'ont pas été identifiées à notre connaissance sont représentées par un point d'interrogation. Les fonctions potentielles sont notifiées pour chaque marqueur.

1.2. Données acquises sur les GSC chez les mollusques

Comme pour les espèces modèles, la caractérisation des GSC chez les mollusques peut s'appuyer sur des critères morphologiques et moléculaires. Peu d'études sont disponibles concernant la caractérisation morphologique des GSC chez les mollusques. Cependant, les travaux sur les lignées germinales mâle et femelle ont permis de caractériser, chez certaines espèces, des cellules germinales précoces.

1.2.1. Caractéristiques morphologiques des cellules germinales précoces

Ainsi, chez les gastéropodes, des cellules germinales précoces ont été décrites sur la base de leurs caractéristiques morphologiques. Chez l'escargot *Helix aspersa*, des cellules nommées « PGC », bien que localisées dans la gonade établie, ont été caractérisées en période néonatale. Ces cellules ont un diamètre de 7 μ m, présentent un noyau ovoïde peu condensé et possèdent plusieurs mitochondries groupées à un pôle du cytoplasme. Ces cellules sont toujours localisées en bordure de gonade, appuyées contre la lame basale (Griffond et Bride, 1981). Chez un autre escargot, *Biomphalaria glabrata*, espèce hermaphrodite simultané, des cellules germinales précoces ont été identifiées dans l'ovotestis. Les cellules précoces femelles proviennent de l'épithélium germinatif et sont proches de la membrane basale. Elles possèdent un noyau irrégulier et deux nucléoles dont l'un disparaîtra lors de l'avancée en gamétogenèse (Jong-brink *et al.*, 1976). Dans la partie mâle de l'ovotestis, des spermatogonies localisées le long de l'épithélium germinatif, ont été caractérisées et possèdent un large noyau contenant également deux nucléoles ainsi qu'un corps chromatoïde parfois proche de mitochondries regroupées (Jong-Brink *et al.*, 1977).

Chez les bivalves, bien que les GSC ne soient pas décrites, différentes études concernent les cellules germinales incluant les cellules germinales précoces, ovogonies et spermatogonies. Ainsi, chez différentes espèces, les jeunes ovogonies possèdent un large noyau ovoïde contenant de la chromatine décondensée et de nombreuses mitochondries (Chung, 2008; Chung *et al.*, 2005, 2007; Kim, 2016; Pipe, 1987). Le diamètre de ces ovogonies est de 10 μ m à l'exception de *Pecten maximus* et *Crassostrea virginica* où il est compris entre 5 et 7 μ m (Dorange et Le Pennec, 1989; Eckelbarger et Davis, 1996a). Chez les mâles, les spermatogonies présentent également un noyau rond ainsi que de nombreuses mitochondries. Chez *Phacosoma japonicus*, leur diamètre est compris entre 6,3 et 7,2 μ m

tetis que chez *Crassostrea virginica*, leur diamètre est compris entre 4 et 5µm (Eckelbarger et Davis, 1996b; Kim *et al.*, 2011).

Chez *C. gigas*, les stades précoces de spermatogenèse ont été étudiés et deux types de spermatogonies, les spermatogonies I et II ont été décrits (Franco *et al.*, 2008). Les spermatogonies I sont celles qui présentent les caractéristiques les plus précoces (Figure I-2). Elles sont appuyées contre la lame basale du tubule gonadique, possèdent un noyau à aspect décondensé et un cytoplasme contenant de nombreuses mitochondries proches d'un nuage et regroupées à un pôle de la cellule. Le diamètre des spermatogonies I est compris entre 5 et 6µm (Franco *et al.*, 2008). Ce diamètre correspond à ceux des spermatogonies mesurées chez d'autres espèces de bivalves, *P. japonicus* et *C. virginica*. Il n'existe pas de description fine des ovogonies précoces chez *C. gigas*, cependant, chez une espèce d'huître proche, *C. virginica*, l'ultrastructure des ovogonies est décrite (Eckelbarger et Davis, 1996a).

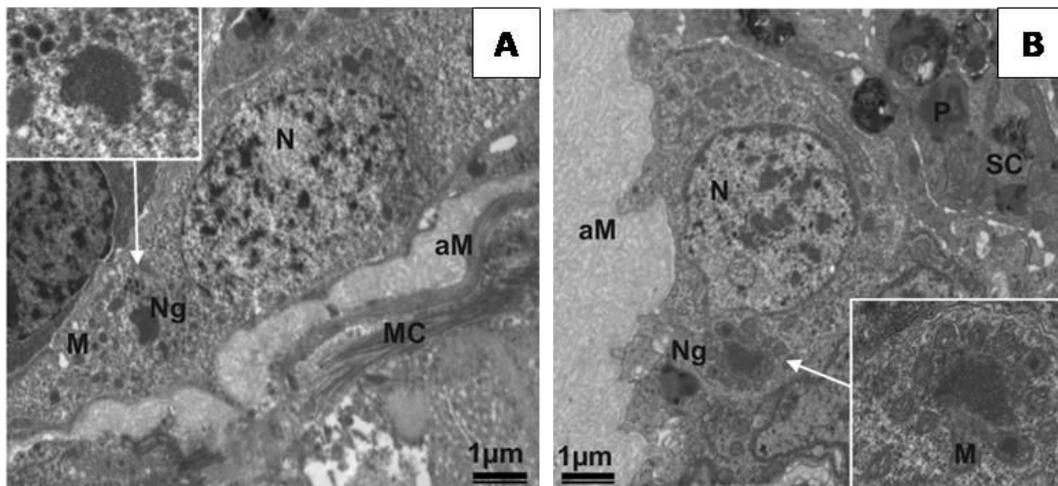


Figure I-2 : Ultrastructure de deux spermatogonies de type I chez *C. gigas*. D'après Franco *et al.*, 2008. M : Mitochondries ; Ng : Nuage ; N : Noyau ; aM : Matrice acellulaire ; MC : Cellule Myoïde ; SC : Cellule Somatique Intragonadique.

1.2.2. Etude de l'activité mitotique pour l'identification de GSC

En complément des critères morphologiques, l'étude des proliférations peut aider à la localisation des GSC dans la gonade. En effet, les GSC se divisent afin d'assurer leur auto-renouvellement et leur différenciation en progéniteurs, les gonies. Ces progéniteurs se divisent de façon plus intense afin de produire un pool de gonies qui pourront entrer en méiose (Morrison et Kimble, 2006; Spradling *et al.*, 2011). L'étude de l'activité mitotique des cellules gonadiques peut donc permettre d'identifier les GSC ainsi que leurs progéniteurs. Chez *C. gigas*, un anticorps hétérologue a été utilisé afin de cibler la protéine MCM6 (Minichromosome maintenance protein 6) qui est présente dans le noyau au moment de la

mitose, lors de la réplication de l'ADN (Cavelier *et al.*, 2017). L'immunomarquage observé chez des huîtres en première gamétogenèse, indique que certaines cellules des tubules gonadiques de l'huître mais également de l'épithélium cilié des pores évacuateurs sont engagées dans le cycle cellulaire aux stades précoces de gamétogenèse. L'incorporation de BrdU permet également de suivre les cellules en prolifération. Elle se fait lors de la phase de réplication de l'ADN du cycle cellulaire et le BrdU peut ensuite être transmis aux cellules filles permettant ainsi de suivre les processus de différenciation. De plus, le mode semi-conservatif de la réplication de l'ADN dilue le BrdU au fur et à mesure des divisions cellulaires ce qui peut permettre de discriminer les cellules quiescentes des cellules à prolifération intense. Ainsi, les GSC conserveront le marquage BrdU plus longtemps que les cellules progénitrices. Chez l'huître, l'incorporation de BrdU a été réalisée afin d'observer les processus de prolifération et de différenciation lors de l'initiation de la gamétogenèse (Cavelier *et al.*, 2017). Comme pour le marqueur MCM6, 24 heures après injection de BrdU dans la gonade, des cellules germinales ainsi que certaines cellules de l'épithélium cilié des pores évacuateurs sont positives au marquage BrdU (Figure I-3). Les marqueurs MCM6 et le BrdU ont donc permis l'observation de cellules germinales capables de proliférer lors de la première gamétogenèse. Cependant, des expérimentations supplémentaires sont nécessaires pour discriminer les potentielles GSC des gonies plus ou moins différenciées dans le tubule gonadique.

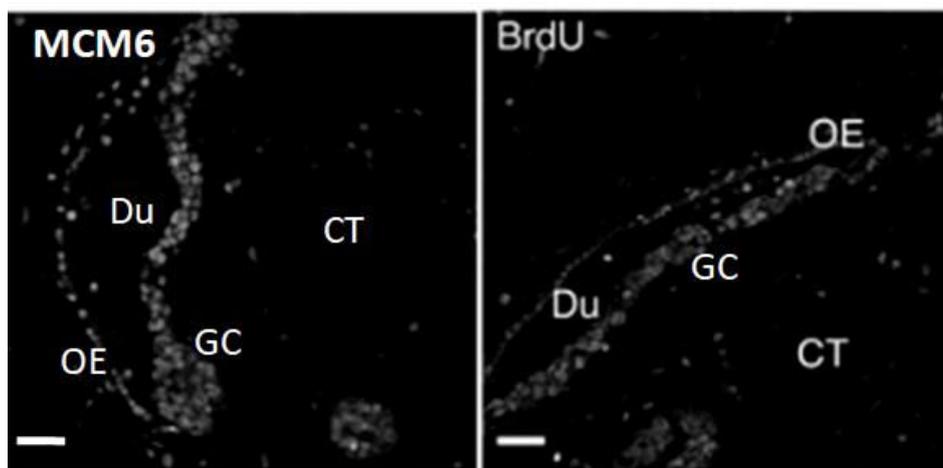


Figure I-3 : Immunofluorescence MCM6 et BrdU dans les pores évacuateurs en première gamétogenèse chez *C. gigas*. Le marquage nucléaire obtenu avec les anticorps dirigés contre MCM6 et BrdU est observé dans les pores évacuateurs (Du) au niveau de l'épithélium cilié (OE) et des cellules germinales (GC). La plupart des cellules du tissu conjonctif de réserve (CT) sont négatives. D'après Cavelier *et al.*, 2017. Du : Pores évacuateurs ; OE : épithélium cilié ; GC : Cellules Germinales ; CT : Tissu conjonctif de réserve. Echelles : 20 μ m.

L'étude menée par Franco *et al.* (2010), chez des huîtres adultes âgées de trois ans, en utilisant le marqueur de prolifération PCNA a montré que les spermatogonies I se divisent peu tout au long du cycle de reproduction, confortant leur caractère précoce dans la lignée germinale, alors que les spermatogonies II se divisent de manière importante en début de cycle les classant comme des spermatogonies plus différenciées.

1.2.3. Données moléculaires spécifiques des GSC

La recherche de marqueurs spécifiques de GSC peut aider à leur localisation dans la gonade adulte. Des données moléculaires sur des marqueurs de GSC sont disponibles chez les mollusques en plus des critères morphologiques décrits précédemment. Ainsi, plusieurs homologues de Vasa, marqueur de GSC et de cellules souches particulières comme les néoblastes de planaire, ont été retrouvés chez différentes espèces de mollusques. Chez le gastéropode *Haliotis asinina*, un transcrite *vasa* a été identifié et son expression a été caractérisée par hybridation *in situ* au cours du développement, afin d'étudier les PGC potentielles. Le transcrite *vasa* est tout d'abord localisé dans l'ovocyte non fécondé puis, lors du stade de développement 8 cellules, il est restreint aux 4 micromères formés. Par la suite, lors de la progression du développement embryonnaire, ce transcrite n'est retrouvé que dans deux clusters de cellules pouvant appartenir à la lignée germinale (Kranz *et al.*, 2010). L'étude des transcrits *vasa* chez quelques bivalves a permis de localiser l'expression de ce marqueur dans les cellules de la lignée germinale. Ainsi, chez la moule méditerranéenne, *Mytilus galloprovincialis*, le transcrite de *myvlg* (*M. galloprovincialis* vasa-like gene) est retrouvé dans les spermatogonies, spermatocytes, ovogonies et ovocytes lors de la période de reproduction ainsi que dans les cellules germinales précoces qui prolifèrent en tout début de gamétogenèse (Obata *et al.*, 2010). Un homologue de Vasa a été récemment identifié chez la palourde *Ruditapes philippinarum* (Milani *et al.*, 2015, 2017). Un anticorps homologue a été produit contre deux peptides de la région C-terminale de sa séquence. Les auteurs ont également réalisé des immunomarquages à l'aide de cet anticorps chez trois autres espèces de mollusques bivalves : *Mya arenaria*, *Anadara kagoshimensis* et *Crassostrea gigas* (Milani *et al.*, 2017). Ils ont observé un marquage positif dans l'épithélium digestif et la gonade chez les différentes espèces de bivalves étudiées (Milani *et al.*, 2017, 2018).

Plusieurs homologues de Piwi ont également été identifiés chez différentes espèces de mollusques et semblent impliqués dans le fonctionnement des cellules germinales. Chez *Chlamys farreri*, l'homologue de Piwi1 est localisé dans les spermatogonies, spermatocytes, ovogonies et les ovocytes (Ma *et al.*, 2017). L'inactivation de ce facteur, par ARN interférence, entraîne une augmentation de l'apoptose dans les cellules germinales. Cependant, une accumulation des cellules germinales précoces a été observée avant dégénération de ces dernières. Ces résultats suggèrent l'implication de Piwi1 dans l'auto-renouvellement et la différenciation des cellules germinales chez *C. farreri*. Plusieurs facteurs impliqués dans le fonctionnement de la lignée germinale ont également pu être identifiés dans le transcriptome d'un autre bivalve, *Crassostrea hongkongensis*, incluant Piwi1, nommé ChPiwi1 (Tong *et al.*, 2015). D'après une analyse phylogénétique, la protéine ChPiwi1 apparaît proche des Piwi identifiés chez les vertébrés. L'étude de ce facteur a permis de montrer une expression plus élevée dans la gonade que dans les tissus somatiques. De plus, son expression est plus forte chez les femelles et plus particulièrement dans les cellules germinales précoces. Ces résultats suggèrent un rôle de ChPiwi1 dans le fonctionnement des cellules germinales précoces et de l'ovogenèse chez *C. hongkongensis* (Tong *et al.*, 2015).

Un homologue de Nanos a pu être identifié lors du développement embryonnaire chez le gastéropode *Ilyanassa obsoleta*. Lors des premiers stades de développement, ce facteur est présent dans l'ensemble des cellules embryonnaires puis apparaît restreint à deux cellules, potentiellement à l'origine de la lignée germinale, entre les stades 24 et 36 cellules. De plus, lorsque l'expression de nanos est bloquée, les animaux présentent un développement anormal suggérant un rôle essentiel de ce facteur chez l'embryon (Kranz *et al.*, 2010; Rabinowitz *et al.*, 2008). De même, chez un autre mollusque gastéropode, *Haliotis asinina*, l'expression spatiale de nanos est presque identique à celle de vasa pendant l'embryogenèse (Kranz *et al.*, 2010). En effet, les cellules exprimant *nanos* semblent appartenir à une sous-population de cellules exprimant *vasa*. Ainsi, les cellules possédant les transcrits *nanos* et *vasa* pourraient constituer les PGC chez cette espèce. Chez le mollusque bivalve *Pinctada fucata*, le séquençage du génome a permis d'identifier trois gènes codant potentiellement pour des protéines Nanos dans les banques d'ADNc obtenues à partir des stades embryonnaires 8 à 16 cellules (Matsumoto *et al.*, 2013).

Chez *C. gigas*, ces trois marqueurs moléculaires (Vasa, Piwi1, Nanos) ont été identifiés. L'expression du gène *oyvlg* (Oyster Vasa-like gene) a été analysée au cours du

développement mais également chez l'adulte par hybridation *in situ* (Fabioux *et al.*, 2004a, 2004b). Le transcrit a été localisé dans les tubules gonadiques mâles et femelles, tout au long de la gamétogenèse, spécifiquement dans les cellules germinales. *Oyvg* est majoritairement exprimé au bord de ces tubules lors de la maturité sexuelle des individus (Figure I-4). Cette localisation correspond aux cellules germinales précoces, incluant les GSC. Cependant, les cellules exprimant *oyvg* n'ont pas été précisément caractérisées.

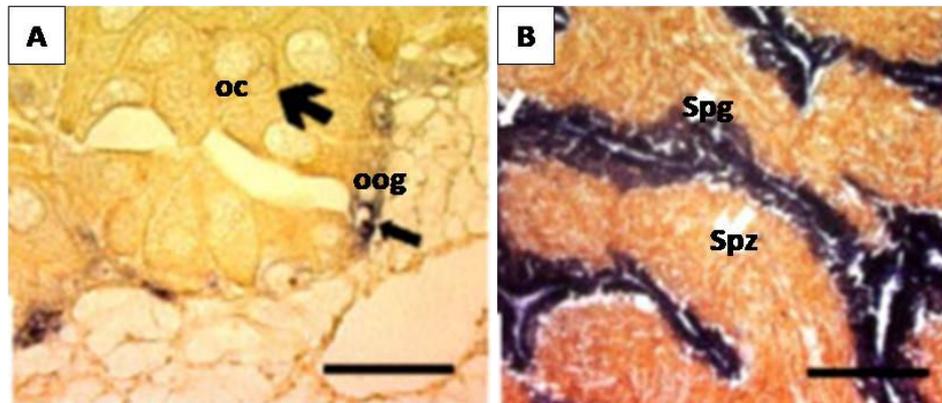


Figure I-4 : Localisation du transcrit *oyvg* dans les tubules gonadiques chez *C. gigas* à maturité sexuelle. A – Localisation du transcrit *oyvg* dans les ovogonies (oog) chez les femelles matures au stade 3 de gamétogenèse. B – Localisation du transcrit *oyvg* dans les spermatogonies (Spg) chez les mâles matures au stade 3 de gamétogenèse. D'après Fabioux *et al.*, 2004b. oog : ovogonies ; oc : ovocytes ; Spg : spermatogonies ; Spz : spermatozoïdes. Echelles : 50µm.

L'expression protéique de vasa chez *C. gigas* a été étudiée par Cavelier et ses collaborateurs (2017). Pour cette étude un anticorps hétérologue initialement produit contre le domaine central de la séquence Vasa de poisson zèbre a été utilisé. Les résultats obtenus montrent un marquage Vasa restreint aux tubules gonadiques et aux pores évacuateurs chez l'adulte. Aucun marquage Vasa n'a été observé dans le tissu digestif. Ces résultats diffèrent de ceux obtenus par Milani *et al.* (2017) chez plusieurs bivalves, incluant *C. gigas*, où un marquage positif est observé dans les tissus digestifs et gonadiques. D'après ces observations, Milani et ses collaborateurs émettent l'hypothèse d'un pool de cellules souches au niveau du tube digestif qui serait capable de migrer dans la gonade et de redonner des GSC chez l'adulte à chaque cycle de reproduction. Ces cellules pourraient être comme celles de type PriSC décrites par Solana (2013). Cavelier *et al.* (2017) utilisent un anticorps anti-vasa hétérologue différent de celui utilisé dans l'étude de Milani et collaborateurs. Cette différence de résultat pourrait être due à l'affinité différente des anticorps utilisés pour le vasa d'huître dans ces deux études (Cavelier *et al.*, 2017). Par ailleurs, un anticorps homologue dirigé spécifiquement contre la protéine Vasa de *C. gigas* a

été développé par Fabioux *et al.* (2009). Ces auteurs ont validé l'expression spécifique de la protéine Oyvlg dans les tissus gonadiques mâle et femelle par western blot mais n'ont pas caractérisé les cellules marquées dans la gonade. De plus, le blocage de *oyvlg*, par ARN interférence, lors de l'initiation de la gamétogenèse a montré une réduction du nombre de cellules germinales dans les tubules gonadiques ainsi que l'arrêt de leur développement en début de méiose aussi bien chez les mâles que chez les femelles. Oyvlg a donc un rôle dans la régulation des proliférations des cellules germinales précoces et joue probablement également un rôle dans la régulation de la différenciation des cellules germinales chez l'huître creuse (Fabioux *et al.*, 2009).

Chez *C. gigas* la séquence de Piwi1 a été identifiée et son transcrit a été étudié (Khider, 2010). Il est apparu que piwi1 est plus exprimé dans les tubules gonadiques au stade quiescent de gamétogenèse, stade qui présente une majorité de cellules germinales précoces au sein de sa population gonadique. De plus, le transcrit est localisé dans les cellules germinales au bord des tubules gonadiques à différents stades de gamétogenèse aussi bien mâles que femelles. A ce jour, la protéine n'a pas été étudiée et la production d'un anticorps à façon spécifique de cette dernière chez *C. gigas* pourrait permettre de compléter les résultats obtenus avec le marqueur Vasa.

Récemment, le marqueur Nanos a également été identifié chez l'huître creuse et nommé Cg-Nanos-like (Xu *et al.*, 2018). L'étude de son expression lors du développement a montré, qu'à partir du stade gastrula, les transcrits Cg-Nanos-like sont répartis dans deux groupes de cellules pouvant correspondre aux PGC potentielles initialement identifiées avec le marqueur vasa chez *C. gigas* (Fabioux *et al.*, 2004a). L'expression de Cg-Nanos-like chez l'huître adulte est localisée dans toutes les cellules germinales femelles et seulement dans les spermatogonies chez les mâles au stade de maturité sexuelle, suggérant un rôle de Nanos dans le développement des cellules germinales (Figure I-5). Le marquage obtenu chez le mâle ne permet pas d'associer spécifiquement l'expression de ce marqueur aux cellules germinales très précoces et/ou aux spermatogonies plus différenciées.

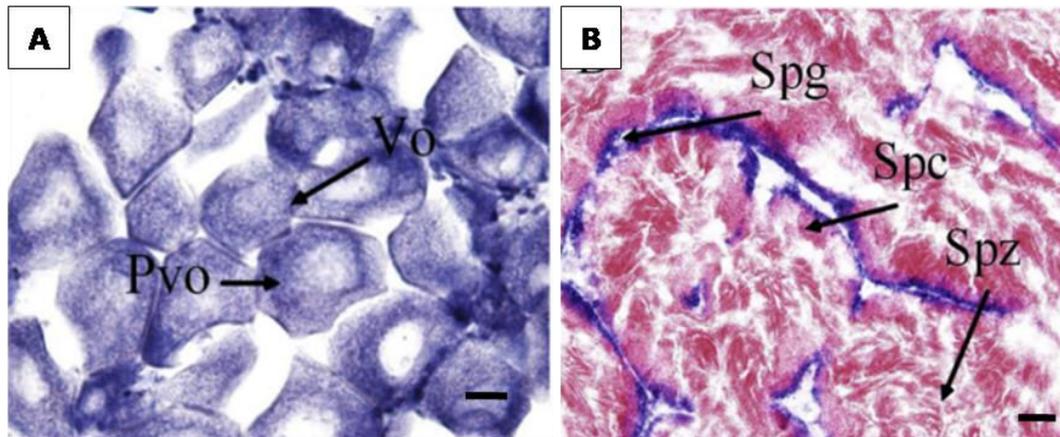


Figure I-5 : Localisation du transcrit *nanos* dans les tubules gonadiques chez *C. gigas* à maturité sexuelle. A – Localisation du transcrit *Cg-Nanos-like* dans les cellules germinales femelles au stade 3 de gamétogenèse. B – Localisation du transcrit *Cg-Nanos-like* dans les spermatogonies (Spg) chez les mâles au stade 3 de gamétogenèse. D'après Xu *et al.*, 2018. Vo : ovocyte en vitellogenèse; Pvo: ovocyte en pré-vitellogenèse ; Spg: Spermatogonies; Spc: spermatocytes; Spz: spermatozoïdes. Echelles : 20µm.

Pour identifier les GSC chez l'huître, la recherche de cellules exprimant des marqueurs moléculaires spécifiques peut donc être un des éléments à prendre en compte. L'expression de transcrits peut ainsi être étudiée en hybridation *in situ* et ces études peuvent être complétées par des approches en immunohistochimie, en utilisant des anticorps hétérologues ou homologues. Chez les mollusques, la conservation de séquences de ces marqueurs moléculaires, identifiés chez des espèces éloignées phylogénétiquement, doit tout d'abord être confirmée avant de pouvoir les utiliser pour l'identification des cellules d'intérêt. L'utilisation d'anticorps hétérologues nécessite de connaître les séquences protéiques afin de pouvoir sélectionner l'anticorps reconnaissant un épitope conservé de la séquence d'intérêt chez l'huître. Les séquences antigéniques ciblées par ces anticorps hétérologues sont principalement issues d'espèces vertébrées et ne sont donc pas toujours bien conservées chez des espèces éloignées comme le sont les mollusques. La production d'un anticorps homologue à partir de la séquence protéique identifiée chez une espèce étudiée peut être une alternative à l'utilisation d'anticorps hétérologues. Cependant, cette solution reste coûteuse et l'anticorps produit peut également ne pas fonctionner. Des méthodes complémentaires d'identification des cellules d'intérêt peuvent donc s'ajouter aux approches en hybridation *in situ* et en immunohistochimie.

1.3. Stratégies complémentaires pour identifier les GSC

D'autres outils peuvent être appliqués à l'identification des GSC en complément des marqueurs moléculaires, des caractéristiques morphologiques et de l'étude de l'activité mitotique. En effet, l'analyse de caractéristiques cellulaires telles que leur nombre, leur volume (cellulaire ou nucléaire) ainsi que leur distribution spatiale peut être réalisée par des approches stéréologiques. La stéréologie est une méthode non biaisée permettant d'estimer des caractéristiques 3D à partir de l'étude de sections en 2D. Ces approches sont classiquement utilisées pour l'identification de cellules anormales en pathologie clinique et cancérologie. Ainsi, dans l'étude de certaines pathologies cérébrales, l'analyse stéréologique de coupes histologiques permet de déterminer des régions cérébrales pathologiques de façon non biaisée (Schmitz et Hof, 2005). Cette méthode est appliquée également aux cellules souches cancéreuses dans le cas de tumeurs du sein. Elle permet de mesurer les cellules cancéreuses afin d'obtenir leur nombre et leur taille et donc de pouvoir déterminer la morphologie des cellules malignes. Les résultats obtenus reposent sur de très nombreuses mesures et permettent une analyse statistique robuste (Ladekarl, 1998). Ces méthodes sont également utilisées en biologie de la reproduction. Ainsi, chez l'ocelot *Leopardus pardalis*, les approches stéréologiques ont permis une meilleure connaissance de la spermatogenèse (Silva *et al.*, 2010). Pour cela, les cellules germinales et somatiques des tubes séminifères ont été mesurées et comptées sur coupes histologiques. Les résultats obtenus ont permis d'estimer les volumes de ces structures et de discriminer les différents types de cellules germinales mâles au cours de la spermatogenèse. La stéréologie présente l'avantage de pouvoir mesurer, de façon objective, un grand nombre de cellules ce qui assure la reproductibilité et la robustesse de cette approche (Brown, 2017). De plus, cette méthode peut être réalisée à partir de techniques histologiques de routine ce qui la rend facilement adaptable à des modèles atypiques tels que *C. gigas*.

1.4. Acteurs potentiels de la niche germinale chez les mollusques

Chez les mollusques, la niche germinale n'a pas été caractérisée à ce jour. Cependant, les acteurs qui peuvent participer à cette niche sont régulièrement retrouvés chez différentes espèces étudiées. Certains acteurs des voies de signalisation connues pour être

impliquées dans le fonctionnement de la niche germinale ont également pu être identifiés, notamment chez *C. gigas*.

1.4.1. Les cellules somatiques intragonadiques

Concernant les cellules somatiques intragonadiques, plusieurs types ont été décrits chez les mollusques. Ces cellules accompagnatrices sont généralement nommées cellules folliculaires chez les femelles et cellules de Sertoli chez les mâles sans que leur fonction précise ne soit connue chez ces espèces (Chung, 2008; Chung *et al.*, 2005, 2007; Jong-brink *et al.*, 1976; de Jong-Brink *et al.*, 1977; Pipe, 1987). Chez *C. virginica*, plusieurs types cellulaires somatiques différents semblent présents dans la gonade sans qu'il soit établi une origine commune ou différente chez les mâles et les femelles (Eckelbarger et Davis, 1996a, 1996b).

Chez *C. gigas*, des cellules somatiques intragonadiques (ISC) ont été décrites au cours de la spermatogenèse. Ces cellules présentent un noyau digité ainsi qu'un cytoplasme étendu contenant des phagolysosomes et des gouttelettes lipidiques (Figure I-6).

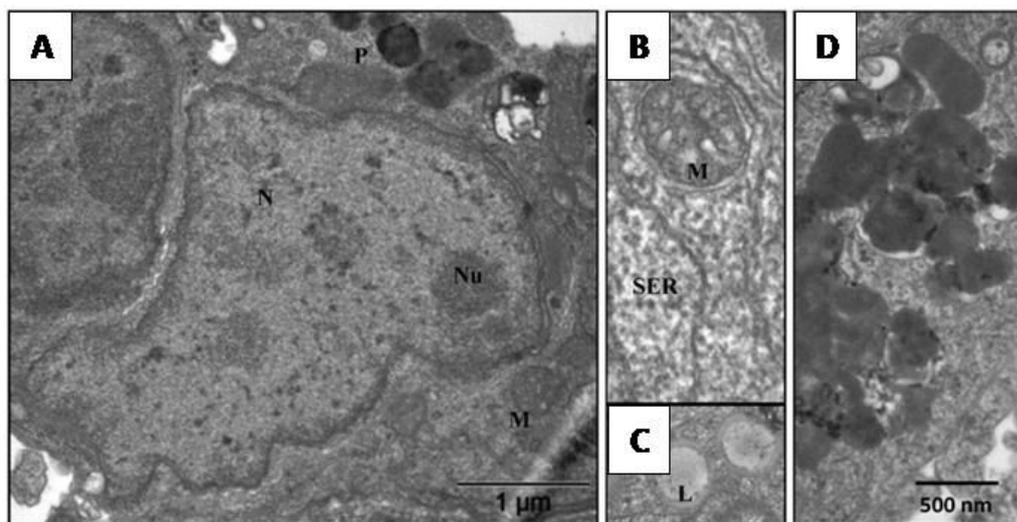


Figure I-6 : Ultrastructure des cellules somatiques intragonadiques (ISC) chez les mâles *C. gigas*. A – Aspect du noyau (N) digité et du nucléole (Nu). B – Aspect des mitochondries (M) C – Aspect des gouttelettes lipidiques (L) présentes dans le cytoplasme. D – Aspect des phagolysosomes (P). D'après Franco *et al.*, 2011. N : Noyau ; Nu : Nucléole ; P : Phagolysosomes ; M : Mitochondries ; L : Gouttelettes lipidiques.

Lors de la reprise de gamétogenèse, au stade de proliférations goniales, les ISC sont présentes au bord de la matrice acellulaire des tubules gonadiques. Ainsi, il a pu être montré que, au cours de l'avancée en spermatogenèse, les ISC forment des colonnes à partir du bord du tubule vers la lumière de ce dernier. Les phagolysosomes sont plus nombreux dans le cytoplasme des ISC au stade quiescent de gamétogenèse, ce qui est probablement dû à

l'élimination des gamètes résiduels de la ponte précédente. Les gouttelettes lipidiques sont, quant à elles, plus nombreuses dans le cytoplasme des ISC au stade de maturité sexuelle. Cette évolution permet probablement d'augmenter la disponibilité des réserves en fonction de l'avancée en gamétogenèse. Des jonctions adhérentes ont été observées entre les ISC en microscopie électronique à transmission. Cependant, aucune jonction cellulaire n'a été observée entre les cellules germinales et les ISC (Franco *et al.*, 2011). Chez la femelle, les cellules somatiques accompagnatrices n'ont pas été décrites chez *C. gigas*.

1.4.2. Le système nerveux

Le système nerveux peut également être impliqué dans le fonctionnement de la niche germinale comme énoncé en introduction. Chez les mollusques, les peptides FMRFamide (Phe-Met-Arg-Phe) ont été identifiés et semblent impliqués dans plusieurs processus physiologiques incluant la reproduction (López-Vera *et al.*, 2008). Ces peptides ont tout d'abord été détectés dans les ganglions nerveux du bivalve *Mercenaria mercenaria* et leur rôle contractile a été mis en évidence sur des cœurs isolés chez cette espèce en mesurant l'amplitude des battements cardiaques (Frontali *et al.*, 1967). Leur appartenance aux FMRFamides a ensuite été établie par l'identification d'une substance permettant les contractions cardiaques dans les ganglions nerveux de *Macrocallista nimbosa*. La caractérisation de cette substance a permis l'identification de la séquence Phe-Met-Arg-Phe-NH₂ (FMRFamide) (Price et Greenberg, 1977). Chez le céphalopode *Octopus vulgaris*, des fibres nerveuses ont été localisées par immunohistochimie avec un anticorps dirigé contre le FMRFamide. Ces fibres s'étendent des ganglions nerveux jusqu'aux organes reproducteurs aussi bien chez les mâles que les femelles. Ces observations suggèrent l'implication du FMRFamide dans la reproduction chez cette espèce et notamment l'émission des gamètes (Di Cristo *et al.*, 2002). L'implication des FMRFamides dans l'étape d'émission des gamètes a également été montrée chez la seiche *Sepia officinalis* par des tests d'activité myotropique. Ainsi, les FMRFamides font partie des peptides permettant la contraction de l'oviducte pour le transport des ovocytes chez cette espèce (Henry *et al.*, 1999).

1.4.3. Les voies de signalisation

Différentes voies de signalisation sont connues pour être impliquées dans la régulation des GSC. Parmi elles, la voie des TGF- β a été étudiée chez *C. gigas* et plusieurs membres de cette famille ont été identifiés. Parmi eux, Og-TGF β -like a été étudié et son expression caractérisée dans les tubules gonadiques au cours du cycle de reproduction. L'expression de cet acteur augmente au cours des différents stades de gamétogenèse, atteignant son maximum à la maturité sexuelle aussi bien mâle que femelle. Le transcrite de ce facteur a été localisé dans les cellules intragonadiques suggérant un rôle dans le déroulement de la gamétogenèse chez l'huître creuse (Figure I-7) (Fleury *et al.*, 2008). L'inhibition de la protéine, par injection d'anticorps, entraîne une diminution de la surface occupée par les tubules gonadiques suggérant son rôle dans le développement de la gonade et de la lignée germinale (Corporeau *et al.*, 2011).

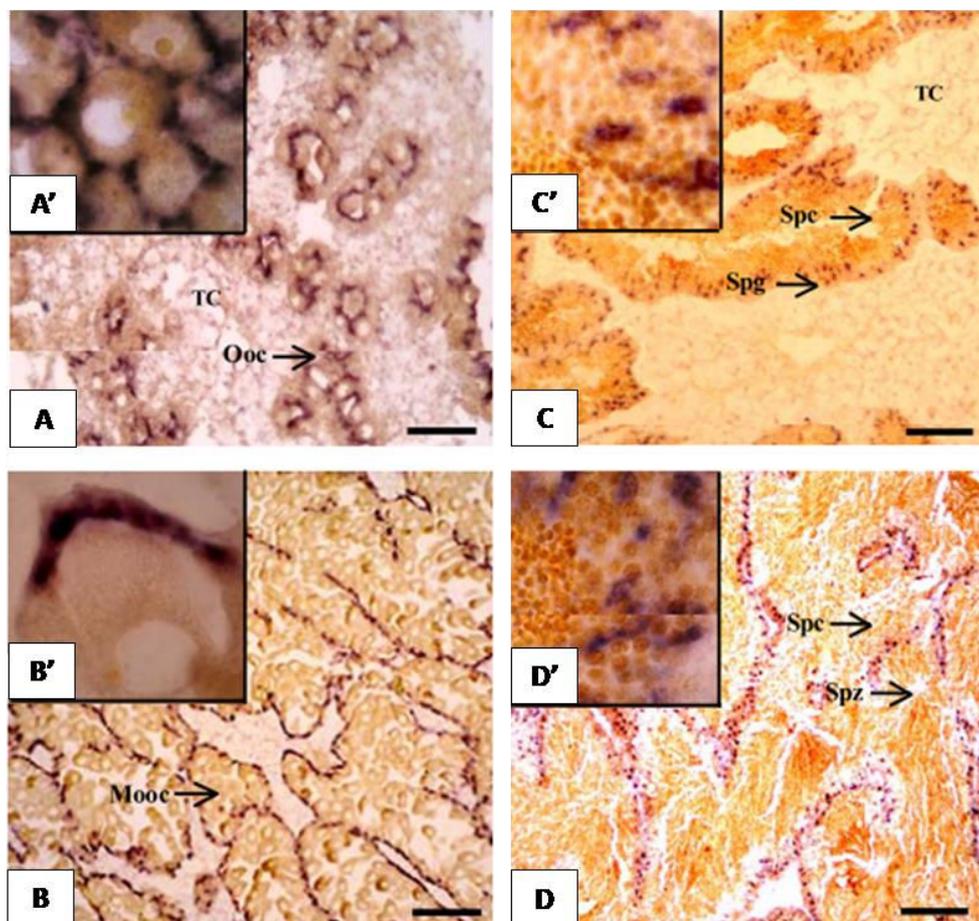


Figure I-7 : Localisation du transcrite *Og-TGF β -like* dans les tubules gonadiques de *C. gigas*. A et A' – Stade 2 femelle. B et B' – Stade 3 femelle. C et C' – Stade 2 mâle. D et D' – Stade 3 mâle. Les transcrits *Og-TGF β -like* sont localisés dans les cellules somatiques intragonadiques chez les deux sexes. TC : Tissu conjonctif ; Ooc : ovocytes ; Mooc : Ovocytes matures ; Spg : Spermatogonies ; Spc : Spermatocytes ; Spz : Spermatozoïdes. Echelles : 15 μ m. D'après Fleury *et al.*, 2008.

Un autre membre de la famille des TGF- β présents chez l'huître est nommé Cg-mGDF (Lelong *et al.*, 2000). Ce facteur présente une forte homologie avec le BMP2 humain aussi nommé Dpp chez la drosophile. Il est impliqué dans les mécanismes de régulation des GSC. L'hybridation *in situ* a permis de localiser le transcrite dans les ISC au cours de la spermatogenèse de l'huître (Figure I-8A) (Franco, 2009). La localisation de la protéine a également été suivie avec un anticorps hétérologue anti-BMP2/4, capable de lier ce facteur mais également deux autres facteurs de la famille des TGF- β , Cg-Gbb et Og-TGF β -like (Figure I-8B) (Franco, 2009). L'ensemble des ISC intragonadiques présente un marquage avec cet anticorps de la base du tubule jusqu'à la lumière, ce qui est un peu différent des résultats obtenus pour le transcrite. Ceci suggère que les facteurs TGF- β chez *C. gigas* pourraient être impliqués dans la régulation de la gamétogenèse via les cellules somatiques présentes dans la gonade.

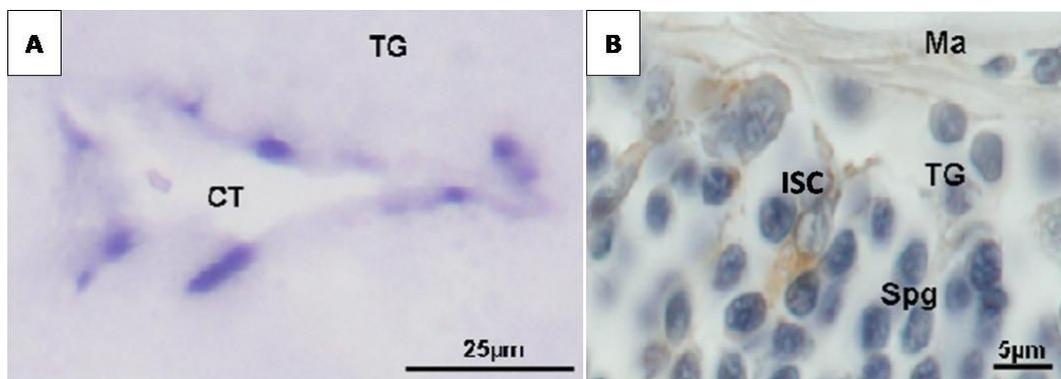


Figure I-8 : Localisation du transcrite et de la protéine Cg-mGDF dans les tubules gonadiques de l'huître creuse mâle à maturité sexuelle. A – Marquage en hybridation *in situ* observé dans des cellules en bord de tubule gonadique (TG). B – Immunomarquages observés dans des cellules somatiques intragonadiques (ISC) en bord du tubule gonadique (TG), proche de la matrice acellulaire (Ma). CT : Tissu conjonctif ; Spg : Spermatogonies. D'après la thèse de Franco, 2009.

Certains acteurs des voies des IRPs ont également été identifiés chez *C. gigas* et ont montré un rôle probable dans les processus de reproduction chez cette espèce. L'étude de cette voie de signalisation fera l'objet du chapitre II.

1.5. Objectifs

Afin d'étudier le fonctionnement du compartiment souche dans la gonade de l'huître creuse *C. gigas*, il convient d'acquérir une connaissance plus fine des GSC et de leur niche. Cette connaissance est un préalable à des études fonctionnelles du dialogue entre cellules souches et acteurs cellulaires et acellulaires au sein de la niche.

Le premier objectif de ce chapitre est donc l'identification et la localisation dans l'ensemble de la gonade de GSC potentielles chez l'huître. Pour cela, les cellules exprimant différents marqueurs de pluripotence et de GSC spécifiquement seront recherchées dans la gonade. Cette approche sera combinée avec l'observation de critères morphologiques des GSC. Pour compléter cette identification, l'étude de la cinétique des mitoses des cellules germinales précoces dans le tubule sera entreprise. Ceci dans le but de discriminer les GSC potentielles, qui se divisent peu, des gonies plus différenciées qui présentent des vagues de prolifération plus importantes. Le deuxième objectif de ce chapitre sera de caractériser les premiers éléments de la niche germinale chez l'huître par des approches en microscopie et immunohistochimie.

Ce premier chapitre est divisé en deux parties. La première partie illustre la recherche de différents marqueurs pouvant permettre d'identifier des cellules souches dans la gonade de l'huître creuse. Pour cela, des anticorps hétérologues et homologues ont été testés par immunohistochimie dans la gonade. En complément de cette recherche de cellules souches, une étude expérimentale a également été menée pour étudier les divisions des cellules germinales, ayant incorporées le BrdU, à différents temps de cinétique. La deuxième partie de ce chapitre concerne l'identification d'une population de GSC potentielles, commune aux mâles et aux femelles. Cette identification d'appuie sur un marquage par l'anticorps homologue dirigé contre Oyvlg, couplé aux critères morphologiques et aux outils stéréologiques. Les caractéristiques du microenvironnement de ces GSC potentielles ont également été précisées.

2. Partie A – Recherche de cellules souches dans la gonade par des approches immunologiques et étude des proliférations goniales

Cette première partie comporte les différentes études immunohistochimiques réalisées dans le but d'observer des cellules souches ainsi que des acteurs de leur microenvironnement dans la gonade de *C. gigas*.

2.1. Matériels et méthodes

2.1.1. Analyse des bases de données

La recherche des séquences protéiques Piwi1, Sox2, Oct4, c-myc et Nanog a été réalisée par homologie de séquences en utilisant pour chaque protéine, une séquence issue d'une espèce modèle, *M. musculus* et/ou *D. melanogaster*. Pour chacune de ces séquences, des blast ont été réalisés sur le protéome prédit (NCBI) de *C. gigas*. Les alignements ont été faits avec Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

2.1.2. Animaux

Les huîtres creuses diploïdes proviennent de la ferme ostréicole GAEC La Tatihou à Saint-Vaast La Hougue (Manche, France). Trente huîtres en deuxième ou troisième gamétogenèse ont été échantillonnées tous les deux mois et leur ploïdie a été individuellement vérifiée selon le protocole de Jouaux *et al.* (2010).

2.1.3. Préparations histologiques

Pour chaque individu, une section transversale d'huître a été fixée dans du Davidson (10% de glycérol, 20% de formaldéhyde, 30% d'éthanol 95°, 30% d'eau de mer stérile et 10% d'acide acétique glacial) pendant 24 à 72 heures. Les échantillons sont ensuite déshydratés dans des concentrations croissantes d'éthanol puis transférés dans du butanol (Carlo Erba, France) avant d'être inclus en paraffine (Roth, France). Des coupes de 3µm d'épaisseur ont été réalisées au microtome puis colorées selon le protocole du trichrome Prenant-Gabe (Gabe, 1968) permettant la détermination du sexe et du stade de gamétogenèse selon Heude Berthelin *et al.* (2001).

2.1.4. Anticorps

L'anticorps à façon cobaye polyclonal anti-Piwi1 a été produit par Eurogentec selon leur programme speedy. Cet anticorps est dirigé contre une séquence peptidique (YRQEDADWSRDMRGK) spécifique de la protéine Piwi1 de *C. gigas*. L'anticorps secondaire utilisé pour l'immunohistologie enzymatique DAB est un lapin anti-cobaye HRP (P0141, DAKO Agilent technologies) dilué au 1/100^{ème} et celui utilisé pour l'immunofluorescence est un anti-cobaye FITC dilué au 1/600^{ème} (gracieusement fourni par Eric Maubert, Normandie Université, UNICAEN, INSERM, UMR-S 1237, Cyceron).

Des anticorps primaires hétérologues ont été utilisés pour Sox2 (lapin polyclonal, ab97959, Abcam), Oct4 (lapin polyclonal, ab53028, Abcam) et KLF4 (chèvre polyclonal, sc-1905, Santa Cruz Biotechnology) aux dilutions 1/200^{ème}, 1/250^{ème} et 1/100^{ème} respectivement. L'anticorps secondaire porc anti-lapin HRP (P0399, DAKO Agilent technologies), dilué au 1/100^{ème}, est utilisé pour l'immunohistologie enzymatique des anticorps primaires anti-Sox2 et anti-Oct4. L'anticorps secondaire utilisé pour l'immunofluorescence de Sox2 et Oct4 est un anti-lapin Alexa Fluor® 488 produit chez l'âne (A-21206, Thermo Scientific) et dilué au 1/1000^{ème}. Les anticorps secondaires utilisés pour l'anti-KLF4 sont soit un anti-chèvre HRP produit chez l'âne (A-16005, Thermo Scientific) et dilué au 1/1000^{ème}, soit un anti-chèvre Alexa Fluor® 488 produit chez l'âne (A-11055, Thermo Scientific) et dilué au 1/2000^{ème}. L'anticorps dirigé contre FMRFamide (FMRFa, T-4322, Peninsula Laboratories) est un anticorps homologue de FMRFa de mollusque dilué au 1/1000^{ème}. Les anticorps secondaires sont les mêmes que pour Sox2 et Oct4. Les anticorps utilisés sont résumés dans le Tableau I-1.

Tableau I-1 : Tableau récapitulatif des anticorps utilisés et de leurs dilutions.

Protéine cible	Anticorps primaire	Anticorps secondaires
Piwi1 (Eurogentec)	Cobaye polyclonal Anticorps à façon	DAB : lapin α -cobaye HRP dilué au 1/100 ^{ème} ----- Fluo : α -cobaye FITC dilué au 1/600 ^{ème}
Sox2 (ab97959, Abcam)	Lapin polyclonal dilué au 1/200 ^{ème}	DAB : porc α -lapin HRP dilué au 1/100 ^{ème} ----- Fluo : Âne α -lapin Alexa Fluor® 488 dilué au 1/1000 ^{ème}
Oct4 (ab53028, Abcam)	Lapin polyclonal dilué au 1/250 ^{ème}	DAB : porc α -lapin HRP dilué au 1/100 ^{ème} ----- Fluo : Âne α -lapin Alexa Fluor® 488 dilué au 1/1000 ^{ème}
KLF4 (sc-1905, Santa Cruz Biotechnology)	Chèvre polyclonal dilué au 1/100 ^{ème}	DAB : Âne α -chèvre HRP dilué au 1/1000 ^{ème} ----- Fluo : Âne α -chèvre Alexa Fluor® 488 dilué au 1/2000 ^{ème}
FMRFa (T-4322, Peninsula Laboratories)	Lapin polyclonal dilué au 1/1000 ^{ème}	DAB : Porc α -lapin HRP dilué au 1/100 ^{ème} ----- Fluo : Âne α -lapin Alexa Fluor® 488 dilué au 1/1000 ^{ème}

2.1.5. Immunomarquages

Des coupes paraffines de 3µm et 10µm ont été réalisés pour les immunomarquages enzymatique DAB et à fluorescence. Les coupes sont déparaffinées et l'immunohistologie enzymatique est réalisée selon Franco *et al.* (2010). Pour l'immunofluorescence, après réhydratation, les coupes sont incubées dans 3% H₂O₂/PBS salin (500mM NaCl, 10mM Na₂HPO₄, 3,1mM KH₂PO₄, pH 7,4) toute la nuit à 4°C. Les coupes sont ensuite rincées au PBS salin avant incubation dans la solution de saturation (10% sérum normal de chèvre, 0,25% BSA, 1% Triton X-100) pendant 1 heure à 4°C. Les anticorps primaires sont dilués dans la solution de saturation avant d'être incubés toute la nuit à 4°C. Les lames sont ensuite rincées et incubées pendant 2 heures à température ambiante avec l'anticorps secondaire adéquat, dilué dans du PBS salin. Les coupes sont finalement rincées au PBS et montées entre lame et lamelle avec du milieu de montage Vectashield® contenant du DAPI.

Un démasquage est réalisé seulement pour l'anticorps primaire anti-Oct4. Pour cela, du tampon citrate 10mM est chauffé au bain-marie jusqu'à ébullition puis les coupes sont immergées dans cette solution pendant 45 minutes, jusqu'à refroidissement. Les contrôles sont réalisés en omettant soit l'anticorps primaire soit l'anticorps secondaire.

Les photographies en lumière blanche sont obtenues en utilisant un microscope Nikon eclipse 80i et une caméra Nikon DS-Ri2 et celle pour la fluorescence sont obtenues grâce au microscope confocal Olympus FLUOVIEW FV1000.

2.1.6. Western blot

Le protocole de Western blot a été réalisé d'après Franco *et al.* (2010). Des échantillons d'aire gonadique issus de 12 animaux en stade quiescent de gamétogenèse ont été groupés pour l'extraction des protéines totales. Deux quantités protéiques, 100µg et 150µg, ont été testées et déposées sur gel polyacrilamide 10% après dénaturation des protéines. L'anticorps primaire anti-Piwi1 a été testé au 1/250^{ème} et incubé 3 heures sur la membrane PVDF (difluorure de polyvinylidène). L'anticorps secondaire utilisé est l'anti-cobaye HRP de lapin dilué au 1/1000^{ème} et incubé pendant 2 heures. Le témoin positif utilisé est l'anticorps anti-Oyvlg d'après les conditions de Fabioux *et al.* (2009).

2.1.7. Expérimentation BrdU

Des injections de BrdU ont été réalisées *in vivo* sur des huîtres lors de l'initiation de leur troisième gamétogenèse. Afin de pouvoir les injecter, les huîtres ont été préalablement anesthésiées avec une solution d'eau de mer contenant 50g/L de MgCl₂ (Sigma Aldrich). Des lots de 5 huîtres ont été répartis pour chaque condition. Les doses de 100mg/kg et 500mg/kg de BrdU ont été testées avec, pour chacune, une injection de 200µL soit dans la gonade soit dans le sinus du muscle adducteur. Les animaux contrôles ont été injectés de la même façon avec de l'eau de mer filtrée. Pour chacune des conditions préalablement énoncées, 7 temps de cinétiques ont été testés (T0, 4 heures, 8 heures, 24 heures, 48 heures, 72 heures et 96 heures). Après chaque injection, les huîtres sont laissées hors de l'eau pendant 15 minutes puis remises dans l'eau de mer pour le reste de leur temps de cinétique. Un prélèvement est ensuite réalisé dans les branchies pour vérifier la ploïdie puis une coupe transversale d'huître est faite pour l'histologie. Les préparations histologiques sont réalisées comme énoncé précédemment.

Le protocole d'immunomarquages BrdU est adapté de celui de Jemaà *et al.* (2014). Les coupes sont déparaffinées puis réhydratées avant une incubation pendant 5 minutes dans du TBST (Tris-HCl 0,05M, NaCl 0,15M, Tween 20 0,05%, pH 8). Les coupes sont ensuite immergées 30 minutes à 37°C dans une solution d'HCl 2N puis 5 minutes dans du tampon borate 0,1M à pH 9. Après rinçage dans du TBST, les coupes sont incubées toute la nuit à température ambiante avec l'anticorps primaire souris anti-BrdU (B-8434, Sigma Aldrich) dilué au 1/1000^{ème} dans du TBST. Les lames sont ensuite rincées et incubées 2 heures à température ambiante avec l'anticorps secondaire anti-souris HRP de chèvre dilué au 1/100^{ème} (P044701, DAKO Agilent technologies) avant d'être rincées de nouveau dans du TBST. Le DAB est déposé sur les coupes pendant 15 minutes à température ambiante avant rinçage. Les coupes sont ensuite déshydratées et montées entre lame et lamelle. Les observations sont faites avec le microscope Nikon eclipse 80i et une caméra Nikon DS-Ri2.

2.2. Résultats

2.2.1. Marqueurs de pluripotence

Les quatre facteurs de Yamanaka ont été recherchés dans le protéome prédit de *C. gigas* (Figure I-9). Une séquence Sox2 a pu être identifiée avec les deux principaux domaines conservés à savoir l'homéodomaine permettant la liaison à l'ADN ainsi que le domaine spécifique des Sox appartenant au groupe B (Figure I-9A). La recherche d'une séquence homologue de Oct4, aussi nommée POU5F1 chez les vertébrés, a permis d'identifier une séquence de la famille des Oct chez *C. gigas*. Cette séquence est apparue plus proche de POU3F4 que de POU5F1. Les domaines spécifiques POU_S et POU_H ont pu être identifiés de même que le linker situé entre ces deux domaines (Figure I-9B).

<p>A – Sox2</p> <p>MEYQDLPEQDLPKPGQTQLPPSSQGMHVPVN PQHAGMHGSHSHGQPQQHSPNPQTNLNGSAGSPNSG ASKDADR VKRPMNAFMVWSRGQRKMAQENPKMHNSEISKRLGAEWKLLSEAEKRPFIDEAKRLRAIHM KEHPDYKYRPRRKTTLMKDKYALPGMPSGAPLQPGREGQMYPMGGYMPNGYPMMDANAYMNQ MGLAGQYGYMPSQQMPGNQMTSGSYMNGSSSYTMSMAGMGYPSPMSPSQMTGGVKREPTPGSGPQG PGRPTNPGDLREMISMYLPTGDGSDPHSAQQRLMHQAHYQPPPPSSEGVNNTVPLTHM</p>
<p>B – Oct4/POU3F4</p> <p>MMTMATSAGLYMPGT PPSLNSPLGDPVHGMQCSVGGQPQYIHRDIDQELKYMPPQVQNGHFNGHWGIPT SADSWNTGMPQLHDMRDLRDNFVRDTRDIKPGQGNE MGHQVTLHHRGPQQIHAPHGWNTPIPGTHLNG GGQLQHGGYVPLNGMLSPHQLPHAMRDPDHDLDQHTPSDDDAPSSDD LEAFKQFKQRRIKLGFQADV GLALGTLYGNVFSQTTICRFEALQLSFKNMCKLPLLQKWLLEADSTTGSTTSIDKIAAQGRKRKRTSIEVTVK GALENHFMKQPKPAAQEISQLAEQLLEKEVVRVWFCNRRQKEKRMTPPGQIGAN GEIIMPDESDGSSRAD SPLQSRPNGAMNGSEHTGLNMAVSSSQTLMSPNYHSVGQLLNAHNPQSGGHPHYNTQNHSPTH</p>
<p>C – KLF4/KLF3</p> <p>MINVDSLLESGVGYSGQGINSPINENIDMQSGWHDIESLMMYQDNTPYENGMLQALDQMRAANMSGKHEI SDVAGNRSPCEQRETYDELDDYDFILANTVESQSLYEDNCLKIKQEQQSNMIGGDALPDFHSTFLDIPDIKFDN DVNNSLPNEMETMKMDFNKYIVPKQEFGHATVSSSCASYMGNKSPVMTMSRLTCNATPVSPQHQQMHYPV NIPSNMSPSSPEMQNDLPKYRAPLQMLQQQQSHFHRMAPQQIHPLQNQQYLMNTNPYMPKMQQQH MLQVPHQMITPPSSPQLVLLLLPPPTVDSTAAQPKKRGRRTWGRKRQTSHTCTHPGCNKTYTKSSHLKAHLRT HTGEKPYACTWKGCWKFARSD ELTRHYRKHTGDRPFQCHLCERAFSRSDHLSLHMKRHI</p>
<p>D – C-myc</p> <p>MLIGRKNE KTAIMDYEMFQPCFYENEPETISPSLPLEDMWKKFELLPTPPLSPRRDDDYDMNSLDLEDFNLI FDEFFEKESAIPKIMTLSETPPNLNSLIQDCMWSGDSGRYPTGDKNAACPDVNSADCVDPAAVFPYPIHSL NSIHSETKLHSLGTETPSDSEEEIDVVTVEKKQLQVVPVAPT KKT VFTVQSVKTERLDGHDEKFGAAKTTVTVK VKVDCTDEHPYSLPVSSLKRVRSLPNSPMSPLPSKRCKRELSDRDMKKVCHKLRASGSSASSSRNSSDSED YPEGKRTQHNVLEKRRNDLKFSSFLRDSVPELVKQERAPKVQILKKASEYIRKLTADERRLESEREMLQEKH EQLRRTLEKLHAREF</p>

Figure I-9 : Séquences des facteurs de Yamanaka identifiés chez *C. gigas*. A – Séquence Sox2 avec en jaune l'homéodomaine et en rouge le domaine spécifique du groupe B. B – Séquence Oct4 identifiée en tant que POU3F4 chez *C. gigas* avec en vert le domaine POU_S, en noir le linker et en bleu le domaine POU_H. C – Séquence KLF4 identifiée en tant que KLF3 chez *C. gigas* avec en violet le domaine de doigt de zinc. D – Séquence C-myc avec en vert les différents domaines spécifiques de Myc et en orange le domaine HLH.

La recherche d'une séquence homologue de KLF4 chez *C. gigas* a permis l'identification d'une séquence proche de KLF3. Seul le domaine en doigt de zinc a pu être observé dans cette séquence (Figure I-9C). La présence de C-myc a également été confirmée chez *C. gigas* (Figure I-9D). La région contenant les cinq domaines conservés spécifique de Myc a pu être observée, de même que le domaine HLH permettant la régulation des gènes cibles. Nanog, un autre facteur de pluripotence, n'a pas pu être identifié dans les bases de données chez *C. gigas*.

Trois anticorps hétérologues dirigés contre les parties C-terminales de Sox2, Oct4 et KLF4 ont été testés afin d'observer les cellules marquées par ces facteurs de pluripotence. Seuls les anticorps anti-KLF4 et anti-Oct4 ont permis d'observer un marquage dans l'aire gonadique (Figure I-10). Aucun anticorps hétérologue satisfaisant n'a pu être sélectionné pour le facteur de transcription C-myc. L'anticorps anti-KLF4 a permis d'observer un marquage cytoplasmique des cellules germinales précoces dans l'épithélium germinatif des pores évacuateurs au stade quiescent de gamétogenèse, suggérant la pluripotence de ces dernières (Figure I-10A). Ces cellules apparaissent localisées le long de la membrane basale. Concernant l'anticorps anti-Oct4, un marquage a pu être observé au stade de proliférations goniales chez les mâles seulement. Ce marquage est localisé dans le cytoplasme de certaines cellules se trouvant à proximité de l'épithélium germinal des pores évacuateurs (Figure I-10B).

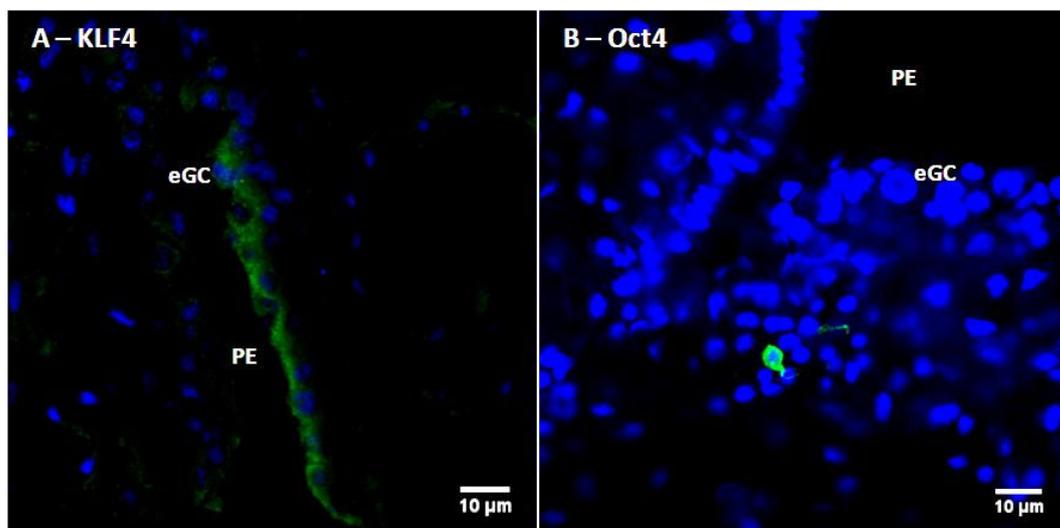


Figure I-10 : Immunomarquages avec certains facteurs de pluripotence sur l'aire gonadique de *C. gigas*. A – Marquage avec l'anticorps anti-KLF4 au stade 0 de gamétogenèse. B – Marquage avec l'anticorps anti-Oct4 au stade 1 mâle de gamétogenèse. TG : Tubules gonadiques ; PE : Pores évacuateurs ; eGC : Cellules germinales précoces.

2.2.2. *Piwi1*

Deux isoformes de la protéine Piwi1 ont été identifiés chez l’huître et présentent 98% d’identité (Figure I-11). Une séquence signal identique aux deux isoformes a pu être prédite, de même que le domaine conservé PAZ skpris entre les acides aminés 286 et 402. Le domaine MID n’a pu être observé, cependant le domaine conservé Piwi a pu être identifié entre les acides aminés 411 et 856. Les acides aminés à l’origine des 2% de différence entre les isoformes sont localisés dans ce domaine Piwi. Concernant la production d’un anticorps à façon, une seule séquence peptidique a pu être validée. La séquence retenue pour l’antigène est accessible dans la molécule pour une reconnaissance par un anticorps mais également spécifique des deux isoformes de Piwi1.

Piwi1_isoforme1	MSGRGRARARGRARGASEDQARRPGEQPAPQQQPPQSSAPAPAPAGGPPAASGRASYRG	60
Piwi1_isoforme2	MSGRGRARARGRARGASEDQARRPGEQPAPQQQPPQSSAPAPAPAGGPPAASGRASYRG	60

Piwi1_isoforme1	GAKEPRPGIASGSGDVPADALSKMSIGTKDAGERDRDLFYSDPECKPAWLSDKRGTSGRA	120
Piwi1_isoforme2	GAKEPRPGIASGSGDVPADALSKMSIGTKDAGERDRDLFYSDPECKPAWLSDKRGTSGRA	120

Piwi1_isoforme1	LPVVTNYFKLEMTPDWHLYQYDVKNFPPIDSRKMRMALLMTHENLLGRTKAFDGMILYLP	180
Piwi1_isoforme2	LPVVTNYFKLEMTPDWHLYQYDVKNFPPIDSRKMRMALLMTHENLLGRTKAFDGMILYLP	180

Piwi1_isoforme1	HRLQEQTVEVFSVRKTDDVQIRITITLTNELPSSPQVMQVYNIIFRRVLAMIEMKQIGR	240
Piwi1_isoforme2	HRLQEQTVEVFSVRKTDDVQIRITITLTNELPSSPQVMQVYNIIFRRVLAMIEMKQIGR	240

Piwi1_isoforme1	NYFNPALSVDIPQHKLIVMPGFVTAIARYETDTLLCADISHKIIRSDTLLEDLMEYLYQQA	300
Piwi1_isoforme2	NYFNPALSVDIPQHKLIVMPGFVTAIARYETDTLLCADISHKIIRSDTLLEDLMEYLYQQA	300

Piwi1_isoforme1	RGDSFYDDCVRKFFVGSIVLTRYNNKTYRVDDFDWDKRPDHSFKLRNDTTITIAEYKKS	360
Piwi1_isoforme2	RGDSFYDDCVRKFFVGSIVLTRYNNKTYRVDDFDWDKRPDHSFKLRNDTTITIAEYKKS	360

Piwi1_isoforme1	NIEVKDMNQPLVVSRRPKKDIRMGRTEPIFLLPELCTVTGLSDEARADFGVMKDVGAHTR	420
Piwi1_isoforme2	NIEVKDMNQPLVVSRRPKKDIRMGRTEPIFLLPELCTVTGLSDEARADFGVMKDVGAHTR	420

Piwi1_isoforme1	VPPEGRNRTLQGFINQINQNEKVKGAEMQWGLAFSQTLMTLNARVAPQENIYQKNAQLS	480
Piwi1_isoforme2	VPPEGRNRTLQGFINQINQNEKVKGAEMQWGLAFSQTLMTLNARVAPQENIYQKNAQLS	480

Piwi1_isoforme1	YRQEDADWSRDMRGKDLITFVNLENWVIVFTRRNSAQAQDLVQTLRSVGPMPGMRINSPT	540
Piwi1_isoforme2	YRQEDADWSRDMRGKDLITFVNLENWVIVFTRRNSAQAQDLVQTLRSVGPMPGMRINSPT	540

Piwi1_isoforme1	ICELQDDRNDYSITALKQYVIPTQLALTLILPTNRKDRYDAIKKFCVDHVPVPSQCVVQR	600
Piwi1_isoforme2	ICELQDDRNDYSITALKQYVIPTQLALTLILPTNRKDRYDAIKKFCVDHVPVPSQCVVQR	600

Piwi1_isoforme1	TLSKKQMLMSVATKIAIQLNCKLGGVWCLDIPKKNMVGIDTYHDSAKKGRSVGGVIC	660
Piwi1_isoforme2	TLSKKQMLMSVATKIAIQLNCKLGGVWCLDIPKKNMVGIDTYHDSAKKGRSVGGVIC	660

Piwi1_isoforme1	STNNLLTRYYSRTTFQHTGEELVNGIVISMRGALEKYHEVNGVLPERIVFRDGVGDGQL	720
Piwi1_isoforme2	STNNLLTRYYSRTTFQHTGEELVNGIVISMRGALEKYHEVNGVLPERIVFRDGVGDGQL	720

Piwi1_isoforme1	RAVFEHEVPQLNECFKRSGGQDYNPKVAIVVVKGRINTRFFARSGPSSLNPPCGTIVDIT	780
Piwi1_isoforme2	RAVFEHEVPQLNECFKRSGGQDYNPKVAIVVVKGRINTRFFARSGPSSLNPPCGTIVDIT	780

Piwi1_isoforme1	VTRPIWYDFFLVQSQSVRQGTIVTPTHYNIWIDITGLKPDHMQRLAYKMCCHLYYWPFGTIRV	840
Piwi1_isoforme2	VTRPIWYDFFLVQSQSVRQGTIVTPTHYNIWIDITGLKPDHMQRLAYKMCCHLYYWPFGTIRV	840

Piwi1_isoforme1	PAPCQYAHKLAFLVGQSIHKDPHMALSDRLYFL 873	
Piwi1_isoforme2	PAPCQYAHKLAFLVGQSIHKDPHMALSDRLYFL 873	

Figure I-11 : Alignement des isoformes Piwi1 identifiés chez *C. gigas*. En rouge : le domaine PAZ. En bleu : le domaine Piwi. Encadré noir : la séquence peptidique utilisée pour la production de l’anticorps à façon.

Le Western blot réalisé avec l'anticorps à façon n'a pas permis de visualiser de bande à 98kDa, taille attendue pour la protéine Piwi1. Cependant, Oyvlg a été utilisé comme témoin positif et une bande à la taille attendue, 79kDa, a bien été observée. Des immunomarquages enzymatique DAB et à fluorescence ont également été réalisés afin de tester l'anticorps à façon anti-Piwi1. Différentes concentrations ont été testées (1/50^{ème}, 1/100^{ème}, 1/250^{ème} et 1/500^{ème}) sur des coupes histologiques d'individus en stade quiescent, cependant aucune n'a permis d'observer de marquage spécifique.

2.2.3. BrdU

L'expérimentation *in vivo* a été réalisée par des injections de BrdU soit directement dans l'aire gonadique soit dans le sinus du muscle adducteur. Afin de vérifier la propagation de la solution injectée dans l'animal, du rouge ponceau a été dissous dans de l'eau de mer filtrée puis injecté comme décrit préalablement. Les individus testés ont été observés 48 heures après l'injection (Figure I-12). La coloration légèrement rouge des parties molles des huîtres a permis de confirmer que la solution injectée se propageait dans l'animal quelle que soit la zone d'injection.

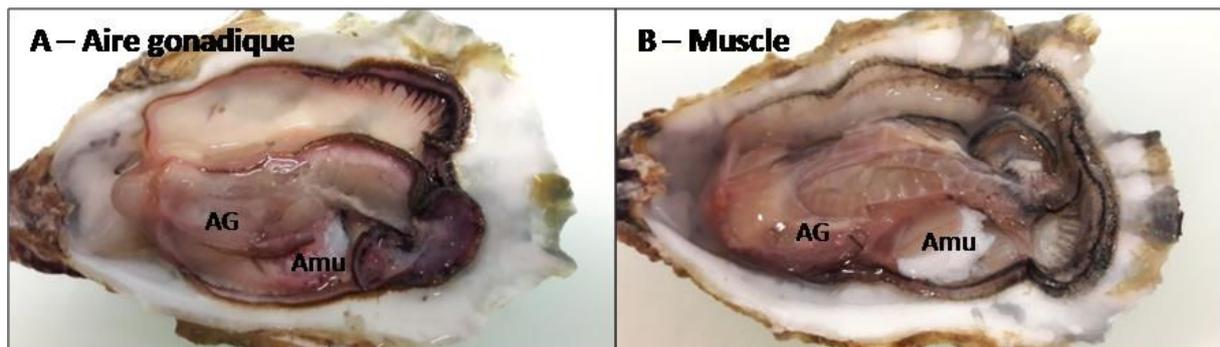


Figure I-12 : Test d'injection au rouge ponceau. A – Huître injectée dans l'aire gonadique. B – Huître injectée dans le sinus du muscle adducteur. AG : Aire Gonadique ; Amu : Muscle adducteur.

Les tests préliminaires d'immunomarquages avec l'anticorps anti-BrdU ont été réalisés sur les individus incubés 24 heures après injection du BrdU dans l'aire gonadique (Figure I-13). Ces premiers marquages ont permis de déterminer la meilleure condition pour l'observation des cellules ayant intégré le BrdU à savoir l'injection directement dans l'aire gonadique avec 500mg/kg de BrdU. Les cellules marquées ont été principalement observées à l'extrémité des pores évacuateurs, certaines du côté de l'épithélium cilié et d'autres du côté de l'épithélium germinal.

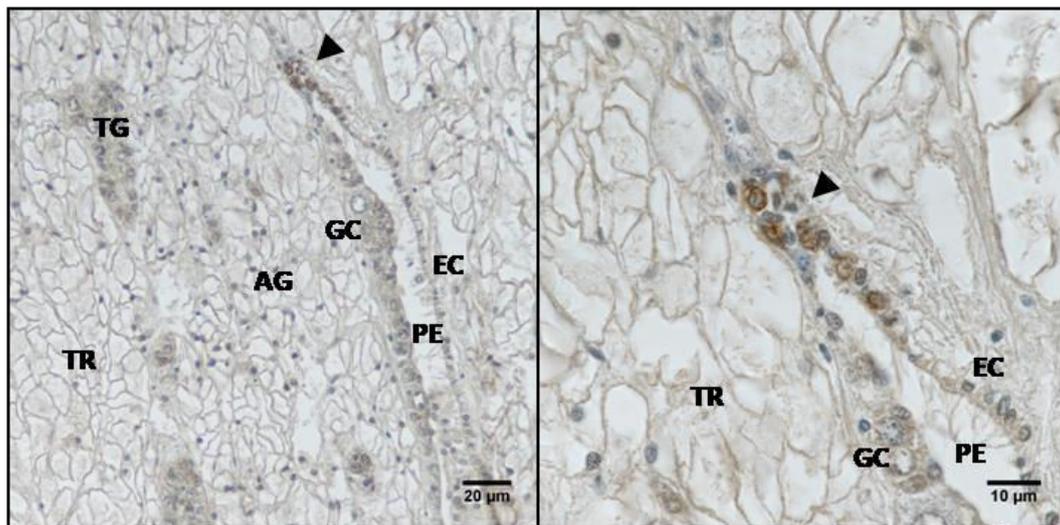


Figure I-13 : Immunomarquages BrdU dans l'aire gonadique de *C. gigas*. Les échantillons ont été prélevés 24 heures après injection de 500mg/kg de BrdU dans l'aire gonadique. Le marquage (flèche) est observé à l'objectif 40 (gauche) et à l'objectif 100 (droite) dans les pores évacuateurs (PE). AG : Aire Gonadique ; TG : Tubules gonadiques ; TR : Tissu conjonctif de réserve ; GC : Cellules Germinales ; EC : Epithélium cilié ; PE : Pore évacuateur.

2.2.1. FMRFamide

L'anticorps dirigé contre le peptide FMRFamide a été testé dans le but d'observer du tissu nerveux proche des tubules gonadiques. Le marquage fluorescent obtenu (vert) montre la présence de peptides FMRFamide au niveau de l'aire gonadique (Figure I-14). Ce marquage est majoritairement observé dans des fibres nerveuses du côté de l'épithélium cilié des pores évacuateurs. Quelques fibres nerveuses positives pour FMRF sont tout de même observées à côté de tubules gonadiques mais toujours dans une région proche des pores évacuateurs.

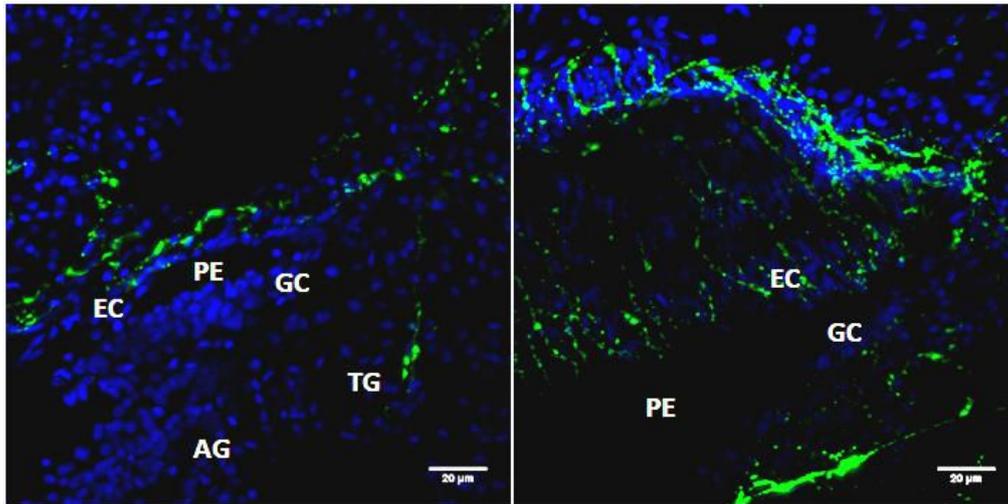


Figure I-14 : Immunomarquages FMRFamide dans l'aire gonadique de *C. gigas*. Le marquage FMRF observé est localisé au niveau des pores évacuateurs (PE), majoritairement du côté de l'épithélium cilié (EC). Quelques marquages FMRF sont parfois observés dans l'aire gonadique (AG), proche de tubules gonadiques (TG). PE : Pores évacuateurs ; TG : Tubule Gonadique ; EC : Épithélium cilié ; GC : Cellules Germinales ; AG : Aire Gonadique ; FMRF : vert ; DAPI : bleu.

2.3. Discussion

Plusieurs marqueurs spécifiques sont utilisés classiquement pour identifier les cellules souches et notamment les GSC. Certains acteurs sont exprimés par les cellules souches tels que les facteurs de pluripotence Sox2, KLF4, Oct4, C-myc et Nanog. La séquence Nanog n'a pas pu être identifiée chez l'huître, ce facteur n'est pas non plus présent chez d'autres espèces invertébrées étudiées et ceci pourrait être lié aux différences développementales entre les vertébrés et les invertébrés (Ma *et al.*, 2014). Les résultats obtenus lors de ce travail concernant la protéine Sox2 n'ont pas permis de confirmer les observations faites par de précédentes études chez *C. gigas* (Cavelier *et al.*, 2017). Etant donné que l'anticorps est le même dans les deux études, des améliorations du protocole d'immunohistochimie utilisé sont à envisager. De plus, chez des individus en première gamétogenèse (Cavelier *et al.*, 2017), le nombre de cellules souches est peut-être plus important que chez les adultes étudiés lors de ce travail de thèse. Ces cellules souches pourraient aussi présenter une localisation un peu différente, la gonade étant restreinte aux pores évacuateurs et à quelques tubules gonadiques chez les individus en première gamétogenèse. En ce qui concerne KLF4 et Oct4, certaines cellules de la gonade présentent un marquage positif lors des stades précoces de gamétogenèse. Ces cellules sont localisées dans les tubules gonadiques ainsi que dans l'épithélium germinatif des pores évacuateurs. Aucun marquage au niveau de l'épithélium cilié des pores évacuateurs n'a été observé pour ces deux facteurs. Ceci pourrait signifier que ces deux derniers pourraient permettre de marquer les cellules

souches dans la gonade de l'huître. Le marquage strictement cytoplasmique obtenu pour le facteur Oct4 est surprenant puisque la localisation attendue de ce facteur de transcription est nucléaire. Cependant, une étude réalisée chez l'humain a montré que deux isoformes de Oct4, nommées Oct4A et Oct4B, peuvent être issues d'épissage alternatif et présentent une localisation cellulaire différente (Bhartiya *et al.*, 2010). En effet, le marquage Oct4A observé dans le testicule humain est présent dans le noyau de petites cellules alors que celui de Oct4B est situé dans le cytoplasme de cellules plus grandes. Ces cellules positives pour Oct4B sont localisées au bord du tube séminifère et sont des spermatogonies souches et ce marquage disparaît lorsque ces cellules se différencient (Bhartiya *et al.*, 2010). Ainsi, l'anticorps utilisé lors de notre étude pourrait cibler une protéine similaire à Oct4B chez l'huître ce qui expliquerait l'observation d'un marquage cytoplasmique seulement en début de gamétogenèse mâle. Ce marquage de possibles spermatogonies précoces chez l'huître peut constituer un bon outil pour cibler ces cellules dans des études ultérieures.

L'incorporation de BrdU dans cette étude a permis de confirmer la présence de cellules à potentiel souche qui prolifèrent dans l'épithélium germinatif des pores évacuateurs de l'huître (Cavelier *et al.*, 2017). Ces cellules capables de prolifération intègrent le BrdU et le conserve pendant au moins 24 heures. Ceci tend à suggérer une division cellulaire lente de ces cellules puisque Jemaà *et al.* (2014) ont montré que seulement deux heures étaient nécessaires à l'incorporation du BrdU dans des cellules progénitrices. De plus, la localisation de ces cellules semble similaire à celle observée pour les cellules marquées par KLF4 et Oct4 et pour celles marquées par Sox2 dans l'étude de Cavelier *et al.* (2017). Ces résultats indiqueraient donc la présence de cellules à potentiel souche dans l'épithélium germinatif, côté externe de la gonade, contre le manteau. Cependant, pour affirmer la nature germinale ou somatique de ces cellules potentiellement souches, des deux côtés des pores évacuateurs, un marqueur spécifique des GSC ainsi que des critères morphologiques sont nécessaires.

En ce qui concerne les marqueurs spécifiques des GSC, les résultats obtenus en immunohistochimie pour Vasa feront l'objet d'un volet de la deuxième partie de ce chapitre. La production d'un anticorps homologue dirigé contre Piwi1 d'huître ne nous a pas permis d'obtenir de cellule marquée dans la gonade. D'autres protocoles d'immunomarquages seront envisagés notamment en pratiquant un démasquage pour augmenter l'accès de l'anticorps à la protéine ciblée. En complément des marqueurs Vasa et Piwi1 étudiés dans ce

travail, un immunomarquage Nanos pourra être réalisé chez l'huître afin de compléter les résultats présentés dans cette thèse.

Dans l'étude de l'environnement proche des cellules germinales précoces incluant les GSC, le marquage FMRFamide obtenu lors de ce travail a permis de montrer la présence de fibres nerveuses localisées au niveau de l'aire gonadique. Le marquage est observé du côté cilié des pores évacuateurs contre le manteau ce qui suggère une implication des peptides FMRFamide dans l'émission des gamètes matures en agissant probablement sur les contractions du tractus reproducteur, comme cela a été proposé chez les céphalopodes *Octopus vulgaris* (Di Cristo *et al.*, 2002) et *Sepia officinalis* (Henry *et al.*, 1999). Très peu de fibres nerveuses ont pu être observées au niveau des tubules gonadiques après immunohistochimie avec l'anticorps anti-FMRFamide. Si des structures nerveuses sont proches des niches germinales, elles sont sans doute très petites et difficiles à localiser. Les observations faites en microscopie électronique, qui seront présentées dans la deuxième partie de ce chapitre, n'ont pas permis par ailleurs de repérer de fibres nerveuses aux abords des tubules gonadiques.

3. Partie B – Identification de GSC et progéniteurs potentiels dans leur microenvironnement

Cette deuxième partie du chapitre I a fait l'objet d'une publication dans *Histochemistry* et *Cell Biology*. Elle comporte la détection immunologique dans l'aire gonadique de l'homologue de Vasa chez *C. gigas* couplée à une approche stéréologique d'histologie quantitative développée lors de ce travail de thèse. L'identification de GSC potentielles a permis de les localiser à l'échelle de la gonade et de caractériser des acteurs de la niche germinale chez l'huître.

3.1. Matériels et Méthodes

La partie « matériels et méthodes » est détaillée dans la publication.

3.2. Principaux résultats

3.2.1. Identification d'une population de cellules germinales précoces chez l'huître adulte

Une approche d'histologie quantitative a été développée afin d'identifier une population commune aux différents stades de gamétogenèse chez le mâle et la femelle et d'estimer leurs volumes nucléaires. Les cellules mesurées ont été sélectionnées sur la base de certains critères de GSC à savoir : leur position le long du bord des tubules gonadiques, l'aspect décondensé de la chromatine ainsi que le marquage positif de ces cellules par l'anticorps homologue anti-Oyvlg. Par cette approche, une population de cellules germinales précoces commune aux mâles et femelles a pu être mise en évidence.

3.2.2. Vasa, un marqueur de cellules germinales précoces chez l'huître

Les marquages réalisés avec l'anticorps à façon anti-Oyvlg ont permis de visualiser les cellules positives dans l'aire gonadique. Il est apparu que le marquage cytoplasmique est spécifique des cellules germinales et plus précisément des cellules germinales précoces à tous les stades de gamétogenèses mâle et femelle. L'étude du transcrit *oyvlg* par qPCR a révélé que son expression dans l'aire gonadique augmente en corrélation avec l'avancée en gamétogenèse. Afin de comparer la localisation de la protéine avec celle du transcrit, des hybridations *in situ* ont été réalisées à tous les stades de gamétogenèse (Figure I-15). De la même façon que pour le marquage protéique, le transcrit *oyvlg* est localisé dans les cellules au bord des tubules gonadiques et concerne les cellules germinales précoces incluant les GSC, les spermatogonies ainsi que les ovogonies. Ces résultats obtenus en hybridation *in situ* sont en accord avec ceux de Fabioux *et al.* (2004b).

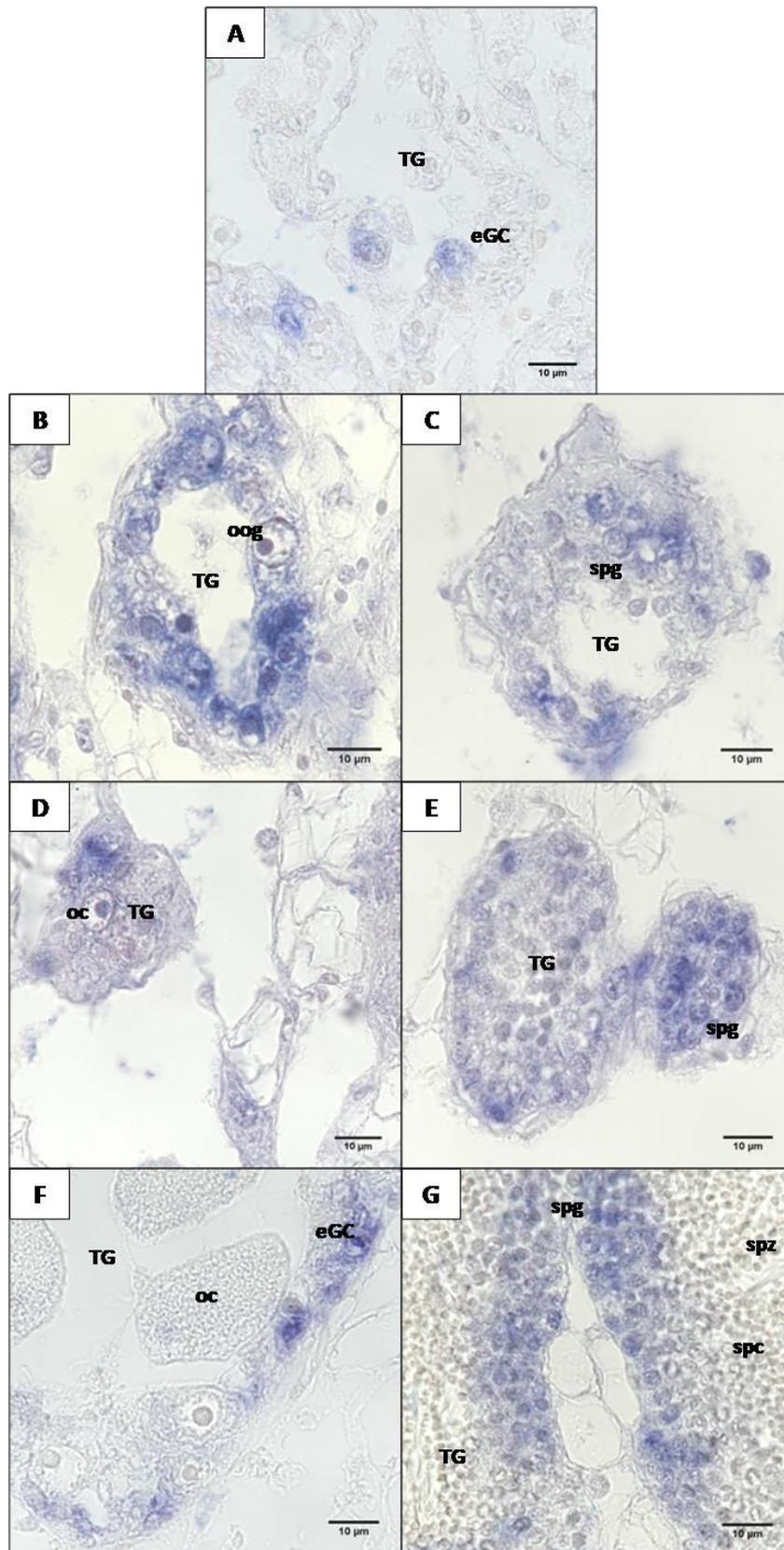


Figure I-15 : Localisation du transcrit *oyvlg* à tous les stades de gamétogenèse. A – Stade 0, repos sexuel. B et C – Stades 1 femelle et mâle respectivement, proliférations goniales. D et E – Stades 2 femelle et mâle respectivement, différenciation germinale. F et G – Stades 3 femelle et mâle respectivement, maturité sexuelle. TG : Tubules gonadiques ; eGC : Cellule germinale précoce ; oog : ovogonie ; oc : ovocyte ; spg : spermatogonie ; spc : spermatocyte ; spz : spermatozoïde.

3.2.3. Identification de deux types de cellules germinales précoces et caractérisation d'éléments de la niche germinale

L'aspect nucléaire des cellules constituant la population de cellules germinales précoces a été particulièrement étudié en stéréologie. En effet, en divisant le plus petit diamètre par le plus grand diamètre nucléaire, les deux mesurés pour chaque cellule de cette population, le ratio d'élongation a été obtenu et utilisé comme indicateur de la forme du noyau. Grâce à ce ratio, deux types cellulaires ont été discriminés au sein de la population de cellules germinales précoces. L'un des deux types cellulaires présente un noyau de forme ronde alors que l'autre type possède un noyau de forme irrégulière. La différence de forme nucléaire est bien visible en microscopie électronique à transmission ce qui a également permis d'identifier différents acteurs pouvant constituer leur microenvironnement respectif. Les deux types de cellules germinales précoces identifiés présentent un marquage Oyvlg, retrouvé en immunogold, proche du nuage et des mitochondries. Autour de ces cellules germinales précoces appuyées à la lame basale, une cellule somatique accompagnatrice est toujours associée à la cellule germinale. L'utilisation d'un anticorps hétérologue anti-BMP2/4 a permis d'obtenir un marquage positif de ces cellules somatiques associées aux cellules germinales précoces des deux types (Figure I-16). Les deux types de cellules germinales précoces (noyaux irréguliers et ronds) sont proches d'une cellule myoïde qui forme la structure du tubule gonadique. Dans un deuxième temps, ces deux types cellulaires ont pu être localisés dans l'aire gonadique totale en travaillant sur une coupe transversale entière numérisée. Cette approche a permis de montrer que les deux types cellulaires sont répartis dans toute la gonade, sans localisation privilégiée. Les cellules à noyau irrégulier représentent 27% de la population totale de cellules précoces identifiée.

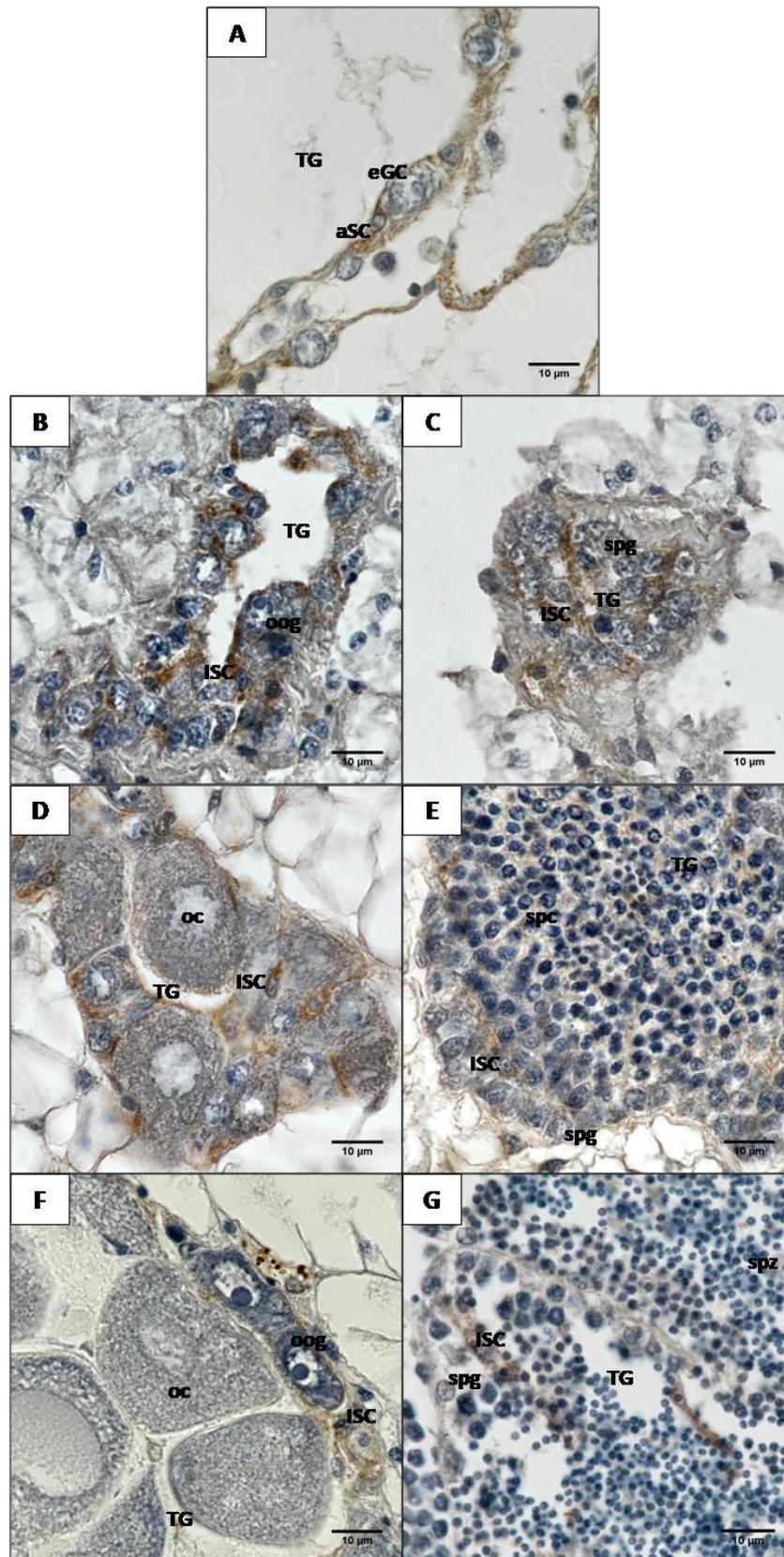


Figure I-16 : Détection de BMP2/4 (anticorps hétérologue) dans le cytoplasme des cellules somatiques intragonadiques. A – Stade 0, repos sexuel. B et C – Stades 1 femelle et mâle respectivement, proliférations goniales. D et E – Stades 2 femelle et mâle respectivement, différenciation germinale. F et G – Stades 3 femelle et mâle respectivement, maturité sexuelle. TG : Tubules gonadiques ; eGC : Cellule germinale précoce ; aSC : Cellule somatique associée ; ISC : Cellule somatique intragonadique ; oog : ovogonie ; oc : ovocyte ; spg : spermatogonie ; spc : spermatocyte ; spz : spermatozoïde.

3.3. Discussion

Les GSC forment une population cellulaire pérenne chez chaque individu permettant de redonner une lignée germinale soit mâle soit femelle à chaque cycle de reproduction. Ce travail a permis d'identifier chez un organisme hermaphrodite successif, l'huître *C. gigas*, une population de cellules germinales précoces sur la base de critères spécifiques (localisation au bord des tubules gonadiques, aspect décondensé de la chromatine dans le noyau, détection du facteur vasa). Le marquage obtenu avec l'anticorps anti-Oyvlg confirme la présence de cellules germinales précoces tout au long du cycle de gamétogenèse aussi bien chez les mâles que chez les femelles. De plus, ce marquage apparaît restreint à la gonade et plus précisément à la lignée germinale ce qui est en accord avec les travaux de Cavelier *et al.* (2017) et ceux de Fabioux *et al.* (2009). Cependant, ces résultats diffèrent de ceux décrits par Milani *et al.* (2017) probablement à cause de la différence d'anticorps. L'étude du transcrite *oyvlg* au cours du cycle de gamétogenèse a révélé sa présence dans les cellules germinales précoces au sein des tubules gonadiques, tout au long des gamétogenèses mâle et femelle, comme cela a été observé pour la protéine Oyvlg.

La population de cellules germinales précoces identifiée a pu être divisée en deux types cellulaires sur la base de l'aspect du noyau, irrégulier ou rond. Ces deux types cellulaires sont répartis de manière identique dans l'ensemble des tubules gonadiques constituant la gonade, les cellules présentant un noyau irrégulier constituent 27% de la population de cellules germinales précoces. Ces résultats confrontés à des données bibliographiques nous ont conduits à définir les cellules germinales précoces à noyau irrégulier comme les GSC potentielles chez l'huître. Les cellules germinales précoces présentant un noyau rond représentent 73% de la population totale identifiée, ont les mêmes caractéristiques chez les mâles et chez les femelles et correspondent probablement aux progéniteurs. De plus, d'après les observations en microscopie électronique à transmission, ces cellules forment souvent des groupes et ne semblent pas directement en contact avec la lame basale. Chez le mâle, les progéniteurs décrits dans notre travail correspondent probablement aux spermatogonies I décrites par Franco *et al.* (2008), avec un noyau rond de diamètre équivalent et les mêmes caractéristiques morphologiques. Il serait intéressant de caractériser la lignée germinale femelle chez *C. gigas* et particulièrement les ovogonies afin de déterminer s'il en existe également plusieurs types. En effet, chez une huître proche, *C.*

virginica, les spermatogonies et ovogonies précoces ont un même diamètre nucléaire qui correspond aussi à celui des progéniteurs décrits dans notre travail (Eckelbarger et Davis, 1996a, 1996b). Ces gonies précoces mâles et femelles pourraient donc correspondre au type cellulaire de progéniteurs décrit dans notre travail chez *C. gigas*.

L'observation des GSC et progéniteurs potentiels en microscopie électronique à transmission a permis d'identifier un microenvironnement similaire incluant une cellule somatique très proche de la cellule germinale. Les cellules somatiques ainsi associées aux GSC ou progéniteurs potentiels présentent un marquage positif pour BMP2/4. Cependant, ces résultats étant obtenus avec un anticorps hétérologue, l'identité de la protéine marquée n'est pas certaine. En effet, une analyse de séquence avait montré que cet anticorps pouvait également cibler deux autres membres des TGF- β : Cg-Gbb et Og-TGF β -like (Franco, 2009). Ces résultats suggèrent l'implication des voies de signalisation TGF- β dans les étapes précoces de gamétogenèse sans pouvoir affirmer précisément le ou les facteur(s) engagés. Il semblerait que ce ou ces facteur(s) agissent via les cellules somatiques appartenant à la niche germinale chez l'huître.

L'étude d'un cas d'hermaphrodisme simultané (données en annexe de ce chapitre) a permis d'observer les deux types, GSC et progéniteurs potentiels identifiés chez le mâle et la femelle. Ceci montrerait qu'une population pérenne de GSC potentielles est commune à tous les individus et leur permettrait de redonner une lignée germinale chaque année.

A l'issue de ce premier chapitre, des questions restent en suspens quant à la nature des cellules souches à l'origine du renouvellement des cellules germinales et somatiques dans la gonade d'huître. Notre travail a permis d'identifier une GSC potentielle dans le tubule gonadique. Cette GSC pourrait dériver d'une cellule souche plus amont, située dans l'épithélium germinatif des pores évacuateurs et qui pourrait être celle présentant un marquage anti-KLF4, dans ce travail, et Sox2, dans le travail de Cavalier *et al.* (2017). Pour poursuivre ce travail, il sera intéressant d'étudier les lignées somatiques mâles et femelles afin de déterminer si leur origine est commune ou différente de celle de la lignée germinale et si elles comportent un ou plusieurs types cellulaires au sein d'un même sexe ou selon le sexe.

*Morphological and molecular criteria allow the identification of putative germ stem cells in a lophotrochozoan, the Pacific oyster *Crassostrea gigas**

Maëva Cherif-Feidel, Kristell Kellner, Didier Goux, Nicolas Elie, Béatrice Adeline, Christophe Lelong & Clothilde Heude Berthelin

Histochemistry and Cell Biology

ISSN 0948-6143

Histochem Cell Biol

DOI 10.1007/s00418-018-1740-3



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Germany, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



Morphological and molecular criteria allow the identification of putative germ stem cells in a lophotrochozoan, the Pacific oyster *Crassostrea gigas*

Maëva Cherif-Feildel^{1,2} · Kristell Kellner^{1,2} · Didier Goux^{1,3} · Nicolas Elie^{1,3} · Béatrice Adeline^{1,2} · Christophe Lelong^{1,2} · Clothilde Heude Berthelin^{1,2}

Accepted: 6 October 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

While our knowledge of bivalve gametogenesis recently progressed, data on early stages of gametogenesis remain to be developed, especially when dealing with germinal stem cells (GSC) and their niche in these organisms. Here, we wish to develop a strategy to identify putative GSC in Pacific oyster *Crassostrea gigas* based on morphological criteria combined with vasa marker expression. A histological quantitative approach, based on stereology, allowed us to identify two types of early germ cells in the germinal epithelium, one presenting round nuclei and the other irregular ones. Both early germ cell types present slightly condensed chromatin in nucleus, are vasa-positive and the *Oyvlg* (oyster vasa-like gene) expression in these cells is recorded throughout the whole gametogenesis process. The microenvironment of an early germ cell in oyster includes an associated somatic cell presenting an immunolabeling for BMP2/4 and a close myoid cell. In agreement with the GSC characteristics in other species, we postulate that putative germ stem cells in *C. gigas* correspond to the early germ cell type with irregular nucleus shape; those early germ cells with a round nucleus may consist in progenitors.

Keywords Germinal stem cell · Germinal niche · Lophotrochozoans · *Crassostrea gigas* · Quantitative histology · Vasa

Introduction

Reproduction ensures the sustainability of a species, it presents much diversified processes but the earliest stages of gametogenesis are well conserved among species. It is of the utmost importance to know the characteristics and functioning of GSC in an organism to understand the regulation of these early stages of gametogenesis. There is only few data

available on GSC and their microenvironment in lophotrochozoans. In terms of development, organization and physiology, lophotrochozoans differ largely from the two other major phylogenetic clades of bilaterian animals (Adoutte et al. 2000). However, their relatively simple organization and their ancestral position in phylogenetic trees make them an interesting case to study in terms of evolutive and fundamental biology.

Among lophotrochozoans, in phylum of Platyhelminths, planarian stem cells were widely studied and some remarkable adult stem cells named neoblasts could be described in these flatworms. These cells are pluripotent cells by nature and are able to differentiate between all zygotic cell types included somatic and germ cells (Rossi et al. 2008; Baguñà 2012; Rink 2013). These organisms present asexual and sexual reproduction modes, depending on the species. To enhance our understanding of GSC nature in lophotrochozoans, studying a species exempt of asexual reproduction mode is particularly interesting. The oyster *Crassostrea gigas* reproduces only by sexual reproduction mode and presents an annual reproductive cycle. In this alternative hermaphrodite, germ and somatic lines present in the gonad may be

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00418-018-1740-3>) contains supplementary material, which is available to authorized users.

✉ Clothilde Heude Berthelin
clothilde.heude@unicaen.fr

¹ Normandy University, Caen, France

² Université de Caen Normandie, Unité mixte de recherche Biologie des Organismes et Ecosystèmes Aquatiques MNHN, Sorbonne Université, UCN, CNRS-7208, IRD-207, UA, Caen, France

³ Université de Caen Normandie, UNICAEN, SF 4206 ICORE, CMABIO3, 14000 Caen, France

male or female lines depending on the adult reproductive cycle. Gonadal tubules embedded in a dense conjunctive tissue specialized in glycogen storage form the gonad located between the visceral mass and the mantle (Berthelin et al. 2000a, b). The course of gametogenesis as well as the gonad structure have been described (Steele and Mulcahy 1999; Heude Berthelin et al. 2001; Franco et al. 2008).

Each year, the storage tissue redevelops at the time of the first gonial mitosis, and then development of germline leads to a progressive regression of the storage tissue. At the end of each gametogenetic cycle, after the emission of gametes, the storage tissue is broadly depleted and gonadal tubules appear almost empty (Berthelin et al. 2000a). Different stages of gametogenesis are described in this species (Heude Berthelin et al. 2001; Franco et al. 2008). At the quiescent stage (stage 0) the early germ cells present in the tubules are not sufficiently differentiated to allow sex identification and no mitotic activity is noted. The early development of germinal epithelium (stage 1) is characterised by the beginning of the gonial proliferation. During the stage 2 of gametogenesis, all the germ cell types are observed in the germinal epithelium with a minor proportion of mature gametes. The last stage of gametogenesis (stage 3) corresponds to the mature reproductive stage where mature gametes constitute the majority of the germ cells present in the tubules.

Recent sequencing of the genome of *C. gigas* (Zhang et al. 2015) and the availability of several transcriptomes (Dheilly et al. 2012; Zhang et al. 2012; Zhao et al. 2012) give access to stem cell putative markers which are also helpful to localize and characterize GSC in this species. Markers of stemness are usually used in histochemistry and immunochemistry approaches to identify the GSC. Recent publication of Cavelier et al. (2017) reports alkaline phosphatase activity in germ cells localized in outer genital *C. gigas* gonad ducts. The dynamics of stem cells was also classically investigated using BrdU or phosphorylated-histone H3 labelling to detect quiescence or reduced rate of proliferation (Newmark and Alvarado 2000; Jemaà et al. 2014). The expression measurement of specific and conserved markers of germ stemness, among bilaterians (Vasa, Nanos, Piwi for example) was also commonly used to localize GSC (Juliano et al. 2010; Brown and Swalla 2012; Maceren-Pates et al. 2015). In particular, Vasa, an evolutionary conserved DEAD-box RNA helicase that plays different functions in germ cell formation and germline maintenance is one of evolutionarily conserved germ cell marker (reviewed in Lasko 2013) therefore, GSC. In *C. gigas*, it has recently been shown that the nanos transcript was detected in early germinal cells and the authors suggested its implication in the differentiation of cells with stem cell-like properties (Xu et al. 2018). In the oyster and more broadly in bivalves, Vasa is for the moment the only one marker of germ cells at the protein level. Fabioux et al. (2004b) showed that, in the particular

case of oyster *C. gigas*, the vasa-like gene (*Oyvlg*) is specifically expressed in gonad tissue as well as the protein (Fabioux et al. 2009) but the cellular location of this protein was not investigated. More recently, Cavelier et al. (2017) identified Vasa-positive germ cells in gonad ducts in the earliest stages of gametogenesis. Likewise, Obata et al. (2010) detected Vasa-positive cells in the *Mytilus galloprovincialis* mussel and Milani et al. (2015, 2017) confirmed by immunodetection that Vasa is a germ cell marker in several bivalve species. The morphological criteria of stem cells are also commonly used for GSC identification as elevated nucleocytoplasmic ratio, slightly condensed chromatin, the presence of a nuage in the cytoplasm, associated with numerous and round mitochondria (Chiarini-Garcia and Russell 2001, 2002). In *C. gigas*, Franco et al. (2008) described an early spermatogonia type; however, for the time being, the germ stem cell at the origin of both male and female germlines in this alternative hermaphrodite are neither identified nor described. Stereological approach is applied to characterize cell types or structural changes in cells fields such neurology (Schmitz and Hof 2005; Golub et al. 2015), cancer cell biology (Ladekarl 1998; Meijer et al. 1998) and reproductive biology (Johnson et al. 2008; Silva et al. 2010). The stereological approach allows investigating a large number of samples from a routine technique like histology. Stereology is based on the principles of geometry and statistics and provides an unbiased quantitative tool allowing access to estimation of 3D parameters based on the structures of 2D sections (Garcia et al. 2007; Brown 2017). This method could be of interest to investigate GSC identity in *C. gigas*.

GSC commonly reside in a special microenvironment, termed germinal niche, provided by somatic support cells, in which they both self-renew and produce progenitors that start the differentiation process while leaving the niche (for review see Spradling et al. 2011). Some elements of the niche appear conserved in the protostomes (Lophotrochozoa and Ecdysozoa) and in deuterostomes as the extracellular matrix which supports the stem cells, some associated somatic cells and myoid cells and a nervous network (Scadden 2006; Ferraro et al. 2010).

The different actors of the niche integrate several molecular signals to ensure functional and structural roles in stem cells destiny. This particular microenvironment controls the stem cells stock by the regulation of their properties of proliferation, auto-renewal and differentiation (Scadden 2006; Morrison and Spradling 2008). These mechanisms are controlled by different signaling pathways as Wnt/ β -catenin, Notch, insulin and Transforming Growth Factors β (TGF- β) (Ferraro et al. 2010). TGF- β (activins and Bone Morphogenetic Proteins BMP), are particularly involved in the control of cell proliferation and in the maintenance of stem cells in numerous species. In *C. gigas*, previous findings demonstrated the occurrence of a diversity of TGF- β signaling

components including various ligands and two subfamilies of receptors: type I receptors, type II BMP/activin, and most of the Smad downstream transducers (Herpin et al. 2004, 2005; Le Quere et al. 2009). Oyster-gonadal-TGF- β (og-TGF- β) was classified as a derived member of the TGF- β superfamily (Fleury et al. 2008). According to the authors, its expression appeared to be gonad-specific and located in associated somatic cells surrounding the germ cells in both male and female. The *in vivo* og-TGF- β dsRNA injection into the oyster's gonad, when injected at the initiation of gametogenesis, leads to a reduction of the number of germ cells in both male and female (Huvet et al. 2012). Among identified ligands in oyster, *Crassostrea gigas*-Glass bottom boat (Cg-Gbb) was found to present a homology with Gbb of *drosophila* as well as with the human Bone Morphogenetic Protein 8 (BMP8). Cg-mGDF, a homologous of the *drosophila* Decapentaplegic (Dpp) and of the human BMP2/4, is also conserved in *C. gigas*. This ligand in *drosophila* is essential for the regulation of germ stem cells stock in the gonad (Xie and Spradling 1998) and is expressed by somatic cells in the germinal niche in male and female (Jagut and Huynh 2007).

Here, we report an original approach to locate and characterize putative GSC in a sexually reproducing lophotrochozoan organism, the Pacific oyster *Crassostrea gigas*. To identify putative GSC we used a histological quantitative approach based on morphological criteria of GSC combined with the detection of Vasa-positive germ cells in the gonad. Microenvironment of identified putative GSC was also specified using an anti-BMP2/4 antibody.

Materials and methods

Animals

Pacific oysters (*Crassostrea gigas*, 2 years old) were provided by a commercial oyster farm in Saint-Vaast la Hougue (Manche, France). Every 2 months, 30 diploid oysters were sampled. Ploidy was individually verified on gill tissue samples according to the protocol set up by Jouaux et al. (2010).

Light microscopy

Transverse sections were fixed for histology in Davidson's fixative at 4 °C (10% glycerol, 20% formaldehyde, 30% ethanol 95°, 30% sterile seawater, 10% acetic acid) for 48 h. The tissue samples were dehydrated in successive dilutions of ethanol, then transferred in butanol (Carlo Erba, France) and embedded in paraffin wax (Roth, France). Three micrometers sections were stained according to the Prenant-Gabe trichrome or Feulgen staining protocols (Gabe 1968). The

stages of the gametogenetic cycle were individually determined according to Heude Berthelin et al. (2001).

Quantitative histology

The quantitative histological approach consists of two steps (online resource 1). The first step consisted in selecting three animals at each stage of male and female gametogenesis, and each corresponding slide was treated according to the Feulgen staining protocol (Gabe 1968). Each slide was used to photograph a series of 25 fields (objective 100) dispatched on the whole gonadal area using a Nikon eclipse 80i microscope and a Nikon DS-Ri2 camera with the Nikon software. According to the literature, the fields were chosen along the gonad tubule wall to reveal the preferential location of GSC. The nucleus located along the tubule wall and exhibiting a slightly condensed chromatin (stemness criteria) were measured by line-intercept sampling according to the method described by Gundersen and Jensen (1985). Thanks to computer program designed to automate measurement of targeted cells, a total of 196,149 cells were measured for the whole set of gametogenetic stages studied in this first step. A stereological estimation of the volume-weighted mean volume was thus calculated using the Cruz-Orive and Hunziker (1986) ruler for each stage and sex. The second step in the histological quantitative study on the selected Feulgen stained slides consisted in the measurement of an ellipse surrounding each nucleus of interest for computer calculations of the longest and the shortest diameters. A total of 1410 cells were measured for the second step after Feulgen staining. The calculation of elongation ratios, consisting in dividing the shortest diameter by the longest diameter, gives an accurate idea of the nucleus shape (round nucleus > 0.8 , irregular nucleus ≤ 0.8). The elongation ratios were also calculated for Oyvlg positive cells after immunolabelling for one selected slide for each gametogenesis stage. A total of 952 cells were measured for the second step after Oyvlg immunolabelling. All measures were performed using Fiji software. The reproducibility of the measurements was evaluated by measuring all the cells of one field (10 cells), 10 times each, and no significant difference was observed in the variance (variance test, $\alpha = 0.05$). The coefficient of variation was calculated on the basis of these measurements, by dividing the standard deviation by the mean.

Whole Slide Image (WSI) was obtained using an Aperio scanner (magnification 40 \times) and nuclei of interest were spotted using QuPath software (Bankhead et al. 2017).

Oyvlg antibody

Anti-Oyvlg antibody was kindly provided by C. Fabioux (UBO, France) and C. Corporeau (Ifremer, France). This rabbit polyclonal antibody, produced by Millegen (Labege,

France), targeted two peptides (GSKNDGESSGFGGG and EEGHFARECPEPRK) specific of the Oyvlg amino acid sequence. The specificity of this antibody in oyster gonad was attested in Fabioux et al. (2009).

Immunohistochemistry

Three and ten-micrometer sections were deparaffinized for immunolabelling and immunofluorescence procedures, respectively. Immunolabelling was performed according to Franco et al.'s indications (2010). Control without primary or secondary antibodies were performed. Rabbit polyclonal primary antibodies, anti-BMP2/4 (sc-9003, Santa Cruz Biotechnology, USA) and anti-Oyvlg were diluted at 1:50 and 1:500, respectively. For anti-BMP2/4, antigen retrieval was done by immersing the sections in saponine 0.05% for 30 min. The secondary antibody used was the horseradish-peroxidase conjugated polyclonal swine anti-rabbit Immunoglobulins (P0399, Dako Agilent Technologies, USA) diluted at 1:100. For immunofluorescence labelling with Oyvlg, rehydrated sections were incubated in 3% H₂O₂/PBS (500 mM NaCl, 10 mM Na₂HPO₄, 3.1 mM KH₂PO₄, pH 7.4) overnight at 4 °C as prescribed by Chong et al. (2011). The sections were washed with PBS and incubated in the saturation solution (10% normal goat serum, 0.25% BSA, 1% Triton X-100) for 1 h at 4 °C. The Oyvlg antibody, diluted at 1:500 in the saturation solution, was dropped on the incubated sections overnight at 4 °C. The sections were washed and incubated with the donkey anti-rabbit IgG secondary antibody Alexa Fluor[®] 488 conjugate (A-21,206, Thermo Fisher Scientific, France) diluted 1:1000 in PBS for 2 h at room temperature. The sections were washed in PBS and mounted in Vectashield[®] with DAPI. Immunolabelling were repeated at least three times. The photographs in light microscopy were obtained using a Nikon eclipse 80i microscope and a Nikon DS-Ri2 camera with the Nikon software and those in fluorescence were obtained using the Olympus FLUOVIEW FV1000 confocal laser scanning microscope and the Olympus software FV10-ASW 4.2. Image processing was done using Fiji software.

RNA probe synthesis

Digoxigenin-labelled RNA probes were synthesized from the clone of an oyvlg fragment (1745 bp) by PCR using a cDNA library normalized to all tissues and developmental stages of *C. gigas* and specific primers (Forward: 5'-ACGTGCGTA AAGCCAGTAT-3' and Reverse: 5'-CAATCCGCCGAA CATTACTT-3'). RNA probes were labelled using the DIG RNA labelling kit (SP6/T7) (Roche) as recommended by Fabioux et al. (2004b).

In situ hybridization (ISH)

Three-micrometer sections mounted on Superfrost Ultra Plus slides were heated overnight at 37 °C, 2 min at 56 °C and 30 min at room temperature before dewaxing and rehydration in serial ethanol dilutions. The sections were treated with Proteinase K (5 µg/mL in Tris-EDTA Buffer pH 8) for 5 min and rinsed in PBS. Post-fixation in PBS/PFA 4% (7 min) was done before rinsing in PBS (3 min) and in 2X SSC (2×5 min). The sections were pre-hybridized for 60 min at 50 °C in hybridization buffer (4X SSC, 10% dextran sulfate, 1X Denhardt's solution, 2 mM EDTA, 50% deionized formamide, 500 µg/mL salmon genomic DNA) before the hybridization step with 200 ng/mL of RNA probes in hybridization buffer overnight at 50 °C. Sections were rinsed in 2X SSC (5 min) and treated with 60% formamide in 0.2X SSC (3×5 min), at 55 °C before rinses in 2X SSC (2×5 min) and in 100 mM Tris-HCl pH 7.5 with 150 mM NaCl (5 min) at room temperature. The slides were then treated for 30 min with blocking buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% blocking reagent) before incubation with the alkaline phosphatase-conjugated anti-DIG antibody (polyclonal, Fab fragments) diluted 1:2000 in blocking buffer. Rinses were performed with 100 mM Tris-HCl, 150 mM NaCl (2×5 min) and with detecting buffer (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM MgCl₂, 10 min). An overnight incubation with a BCIP/NBT STOCK solution diluted at 1:50 and a rinsing process in Tris-EDTA buffer (5 min) were done. Dehydration was made in 100% ethanol (2×1 min) and in Roti-Histol (5 min) before mounting. Negative controls were performed with sense probes. ISH was performed in triplicates.

RNA extraction and cDNA synthesis

For each of the gametogenesis stages, four pools of three gonadal area samples were done. Tissues were ground in 1 mL of TriReagent at room temperature (Sigma Aldrich, France). One hundred microliters of 1-Bromo-3-chloropropane (Sigma Aldrich, France) were added and samples were vortexed before a 15 min centrifugation at 4 °C. The aqueous phase was recovered and the total RNA extraction was performed using Nucleospin RNA II kit (Macherey-Nagel). The RNA samples were reverse transcribed as recommended by Dheilly et al. (2012).

Real-time polymerase chain reaction

The primers sequences for Oyvlg (Forward: 5'-CACAAG TCAGACCGAGTGCA-3' and Reverse: 5'-GCGATG GTGTGATGATTGGC-3') and EF1α (Forward: 5'-ACG ACGATCGCATTCTCTT-3' and Reverse: 5'-ACCACC CTGGTGAGATCAAG-3') were designed using Primer3

software and folding analysis (<http://unafold.rna.albany.edu/?q=mfold>). Quantitative PCR were done using 1X GoTaq SYBR Green Mix (Promega, France), 8.85 ng cDNA and 900 nM of each primer in a final volume of 15 μ L. Each run consisted of one cycle of 5 min at 95 °C followed by 45 cycles of 15 s at 95 °C and 45 s at 60 °C. Specific amplification of the target sequence was estimated by constructing melting curves (80 cycles of 10 s, increasing the temperature by 0.5 °C step from 55 °C to 95 °C). The reference gene (EF1 α) was selected among the three tested housekeeping genes (GAPDH, Arf1, EF1 α) after the checking of constant expression and in agreement with previous studies using the same tissues and species (Fabioux et al. 2009; Dheilly et al. 2011; Dubos et al. 2016).

Electron microscopy

Small pieces of gonadal area tissue were sampled and immersed for 16 h at 4 °C in a fixative solution (2.5% glutaraldehyde, 0.31 M cacodylate, 0.25 M saccharose, pH 7.4). The samples were rinsed in 0.38 M cacodylate buffer with 0.28 M saccharose at 4 °C and post-fixed for 2 h with 1% osmium tetroxide in 0.2 M cacodylate buffer containing 0.365 M saccharose. Tissues were dehydrated in ascending acetone concentrations and embedded in Epon. Ultrathin sections were mounted on coated copper grids and contrasted with uranyl acetate followed by lead citrate. Ultrathin sections were examined with a JEOL1011 transmission electron microscope. Image processing was done using Fiji software.

Statistical analysis

Statistical analyses were performed using a variance test and a non-parametric one-way analysis with permutation test on R 3.4.3 and RStudio software ($\alpha=0.05$).

Results

Quantitative histology

The first step of the histological quantitative approach allowed us to obtain the volume-weighted mean volume of germ cells nuclei which we wish to focus on here (early germ cells) at each stage of the gametogenesis (Fig. 1a). It appeared that most of the stages of the gametogenesis shared some common nucleus volumes (range between 76.91 and 97.67 μ m³). This common population of early germ cells was found in four of the seven stages of the gametogenesis (no significant difference): in the quiescence stage (St0), the male gonial mitosis initiation stage (St1M), the female germline development stage (St2F) and the female sexual

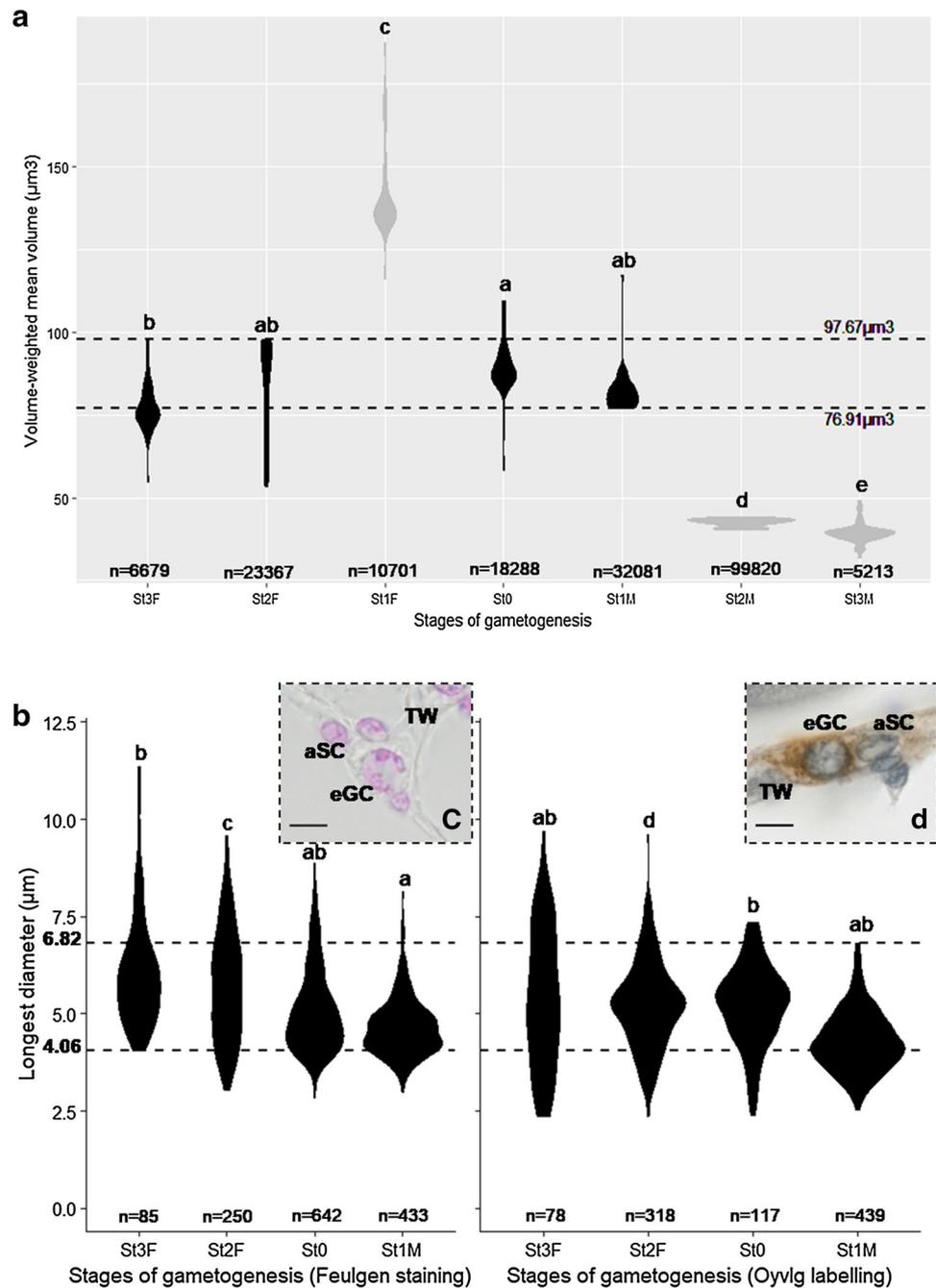
maturity stage (St3F). To characterize sub-populations of early germ cells, the second measurement step was conducted by studying only the four stages of interest mentioned above (Fig. 1b). For these four stages, the longest diameters of the nuclei of interest were measured first on the Feulgen stained slides and then on Oyvlg immunolabelled slides and further compared (Fig. 1b). The statistical analysis revealed that, for each stage, the nucleus measurements did not display any difference between Feulgen staining and Oyvlg immunolabelling, except in the gonial mitosis initiation stage in male. Moreover, no significant difference was found between the quiescence stage, the development and sexual maturity stages in female. The nuclei diameters shared in the four stages ranged from 4.06 to 6.82 μ m. Early germ cells presenting determined diameter were located along the tubule wall, presented a nucleus with slightly condensed chromatin (Fig. 1c) and an Oyvlg positive labelling (Fig. 1d), they were closely associated to a differentiated somatic cell (aSC) (Fig. 1c, d).

Nucleus elongation ratios (shortest diameter/longest diameter) of the early germ cell identified population were calculated for cells of interest stained with Feulgen or immunolabelled with Oyvlg. These elongation ratios ranged from 0.42 to 1 (Fig. 2). This result reflected that early germ cells present a round or irregular nucleus. To discriminate sub-populations in early germ cells, a coefficient of variation of 0.09 was calculated according to the evaluation of measurements reproducibility. Using this coefficient, we distinguished two sub-populations of early germ cells depending on their nucleus elongation ratios. The first sub-population presented round nuclei with an elongation ratio that is strictly superior to 0.8 and corresponded to 73% of the total early germ cells population. The second sub-population of early germ cells presented irregular nuclei with an elongation ratio equal or below to 0.8 and represented 27% of the total early germ cells population.

Oyvlg immunolabelling

Immunofluorescence detection using the Oyvlg antibody showed a specific signal in germ cells (Fig. 3). Cytoplasmic signal in germ cells was observed in the gonad through the reproductive cycle. During the quiescence stage, only large cells presenting an aspect of early germ cells were labelled (Fig. 3b). These cells are located along the wall of the tubule and present nuclei with slightly condensed chromatin underlined by DAPI staining (Fig. 3b inset). In the following stages of gonial mitosis initiation (Fig. 3c, male and Fig. 3d, female), germline development (Fig. 3e, male and Fig. 3f, female) and sexual maturity (Fig. 3g, male and Fig. 3h, female), Oyster Vasa positive labelling is observed only in early germ cells. It could be noted that the slight

Fig. 1 Graphical representation of the two-step histological quantitative approach. **a** Step 1 of the quantitative approach led us to estimate the volume-weighted mean volume of each field (75 fields per gametogenesis stage). The common population at the most stages of gametogenesis (four stages, colored in black) is represented by the dotted lines. **b** Step 2 of the quantitative histology led us to measure the longest diameter of each nucleus of interest on the four selected stages of gametogenesis (*St0* quiescence stage, *St1M* male gonial mitosis initiation stage, *St2F* female development stage, *St3F* female sexual maturity stage). The longest diameters of nucleus measured after Feulgen staining are represented on the left-end side of the figure and those measured after Oyvlg immunolabelling are represented on the right-end side of the figure. The common nucleus population in both Feulgen staining and Oyvlg immunodetection is represented by the dotted lines. The letters indicate the significant differences; the number of cells measured is annotated (n). Insets **c** (Feulgen staining) and **d** (Oyvlg immunodetection): Aspect of early germ cells belonging to the common population. Tubule wall (TW), early germ cell (eGC) and associated somatic cell (aSC). Scale bars: 5 μm (c, d)



staining observed in DAB immunolabelling around GSC was not noticed in immunofluorescence detection of Oyvlg.

Oyvlg expression

qPCR data showed that oyvlg transcript appears to be expressed throughout the gametogenesis for both male and female with significant increase correlated to the gametogenesis progress and with an expression that is significantly higher in the sexual maturity stage in female (Fig. 4a). To compare protein and transcript locations, in situ

hybridization was performed during the earliest stages of gametogenesis (Fig. 4b–d). Early germ cells were stained with the same pattern as Oyvlg protein expression labelling.

Germinal niche

The microenvironment or germinal niche of both types of early germ cells identified by quantitative histology was studied. At photonic scale, early germ cells appear to be closely surrounded by at least one associated somatic cell and are closed to a myoid cell (Fig. 5a). The associated

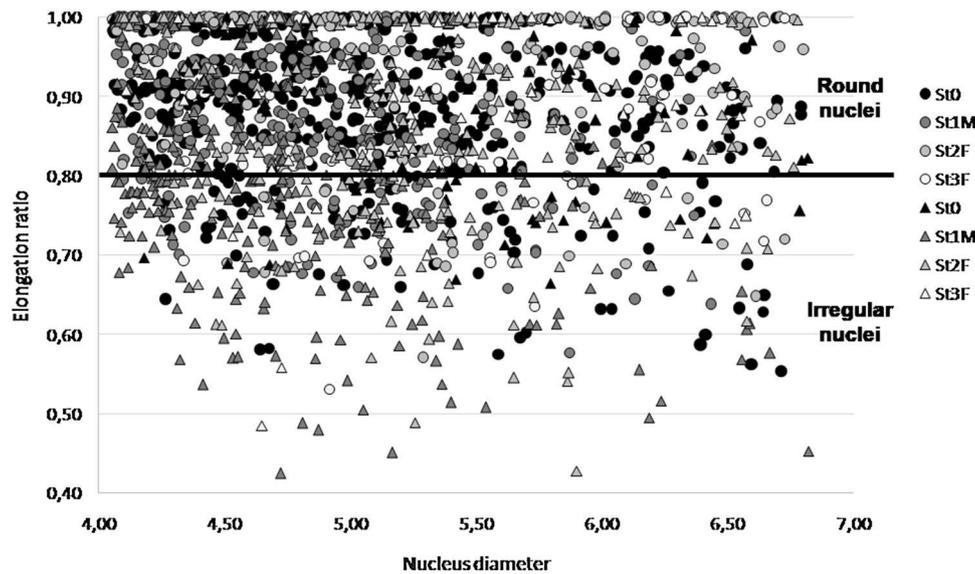


Fig. 2 Graphical representation of the elongation ratios for each nucleus diameter between 4.06 μm and 6.82 μm corresponding to the cell population of interest for the four gametogenesis stages selected (*St0* quiescence stage, *St1M* male gonial mitosis initiation stage, *St2F* female development stage, *St3F* female sexual maturity stage). The

elongation ratios obtained after Feulgen staining measurements are represented by circles and those obtained after Oyvlg immunolabelling measurements are represented by triangles. The elongation ratios that are strictly superior to 0.8 correspond to round nuclei while the elongation ratios that are below 0.8 correspond to irregular nuclei

somatic cells surrounding early germ cells are immunolabelled by a BMP2/4 antibody (Fig. 5b). Electron microscopy allowed to precise the structure of the germinal niche in oyster (Fig. 6). A continuum of the two types of early germ cells (with irregular or round nucleus) was observed several times in the gonadal tubule (Fig. 6a). Microenvironment of both early germ cell types (Fig. 6b, c) shared common elements: the location of early germ cell against the *basal lamina* near a myoid cell (Fig. 6d, e) and the presence of at least one closed associate somatic cell (Fig. 6f, g). Somatic cell associated with early germ cell sometimes contained some phagolysosomes in its cytoplasm (Fig. 6e). Both early germ cells types presented a nucleus with slightly condensed chromatin and a distinctive nuage in cytoplasm associated with numerous and round mitochondria (Fig. 6h, i). It has been noted that the early germ cells with round nucleus were quite often combined in two's or four's.

Finally, the distribution of the both types of early germ cells with round (Fig. 7a, b, red circles) and irregular (Fig. 7a, b, blue circles) nuclei was studied on a whole gonad section (whole slide image, WSI). Both early germ cells types are uniformly distributed in the oysters' gonad tubules, from the outer (against mantle) to the inner (against digestive area) parts of the gonad. The specificity of the vasa detection is evaluated by a negative control (Fig. 7c, d).

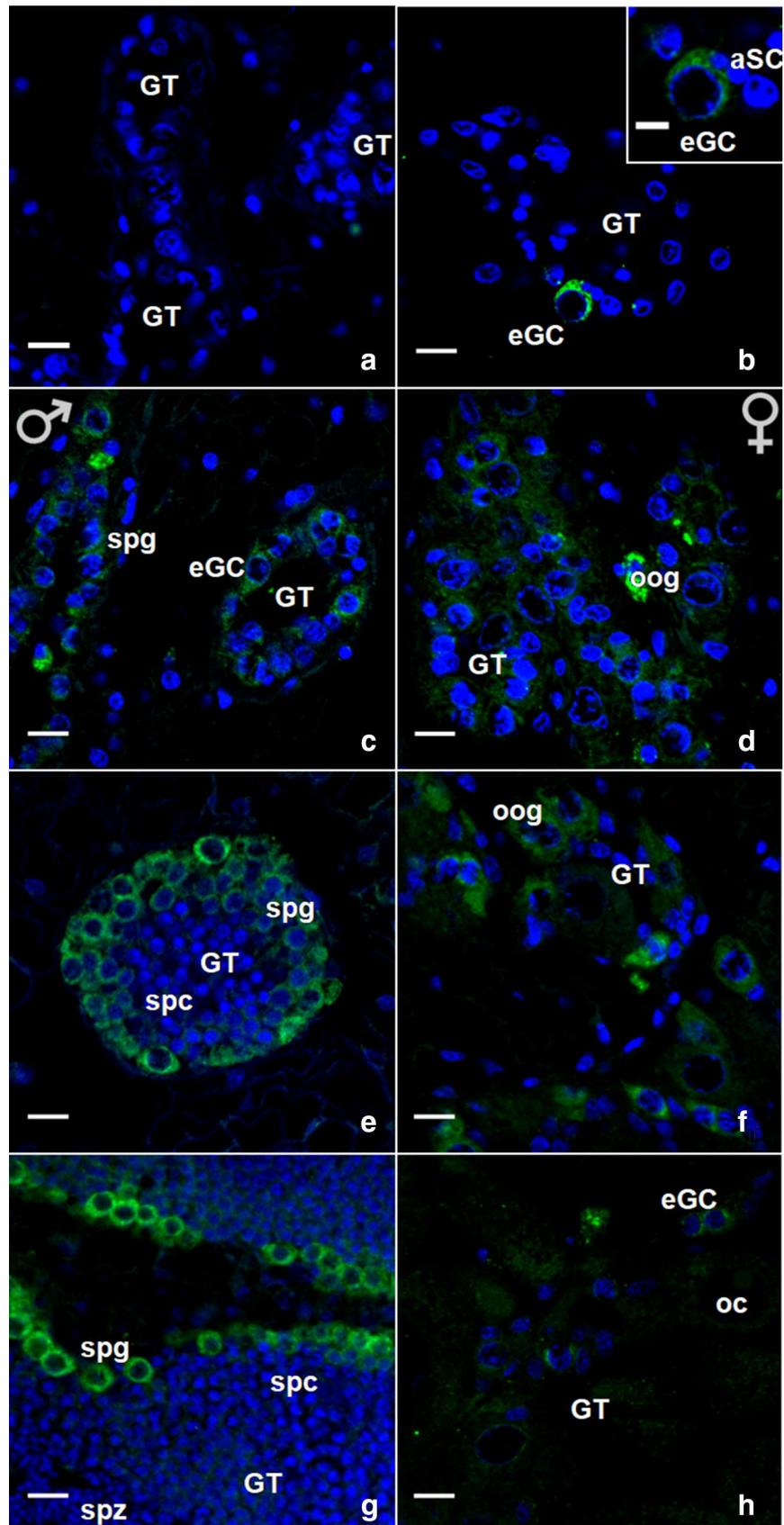
Discussion

The oyster *C. gigas* belongs to the lophotrochozoans, which presents very heterogeneous reproduction modes. In this alternative hermaphrodite, gonad cells (somatic and germ cells) are able to differentiate in female or male lines depending on the seasonal reproductive cycle. The germline is renewed each year and this study presents a quantitative and qualitative approach to better characterize the GSC in their niche in adult oyster. Previous studies carried out on the spermatogenesis of two molluscan species (*C. gigas* and *Mytilus galloprovincialis*) described two types of spermatogonia in the germinal epithelium (Franco et al. 2008; Prisco et al. 2017).

First, a population of early GC is identified in the oyster adult gonad

Germinal stem cells form a sustainable and scarce population that is common to all sexes and stages of *C. gigas*' reproductive cycle. The first result concerning the mean volume of the germ cells nuclei of interest revealed a common population of germ cells apparent in four stages of the gametogenesis: the quiescence stage, the gonial mitosis initiation stage in male, the developmental and

Fig. 3 Immunofluorescence detection of Oyvlg in male and female gonad tubules (GT) of *C. gigas* throughout the gametogenesis. **a** Negative control **b** Quiescence stage, few early germ cells (eGC) are positive to Oyvlg antibody and are found close to an associated somatic cell (aSC) (**b** inset). In male (**c, e, g**) and female (**d, f, h**) gametogenesis stages, Oyvlg immunodetection was restricted to germ cells. For gonial mitosis initiation (**c** male and **d** female), germline development (**e** male and **f** female) and sexual maturity stages (**g** male and **h** female), Oyvlg immunolabelling is observed from gonia to early oocytes and spermatocytes. Spermatogonia (spg), spermatocytes (spc) spermatozoa (spz), oogonia (oog) and oocyte (oc). Scale bars: 10 μ m (**a-h**) and 5 μ m (**b** inset)



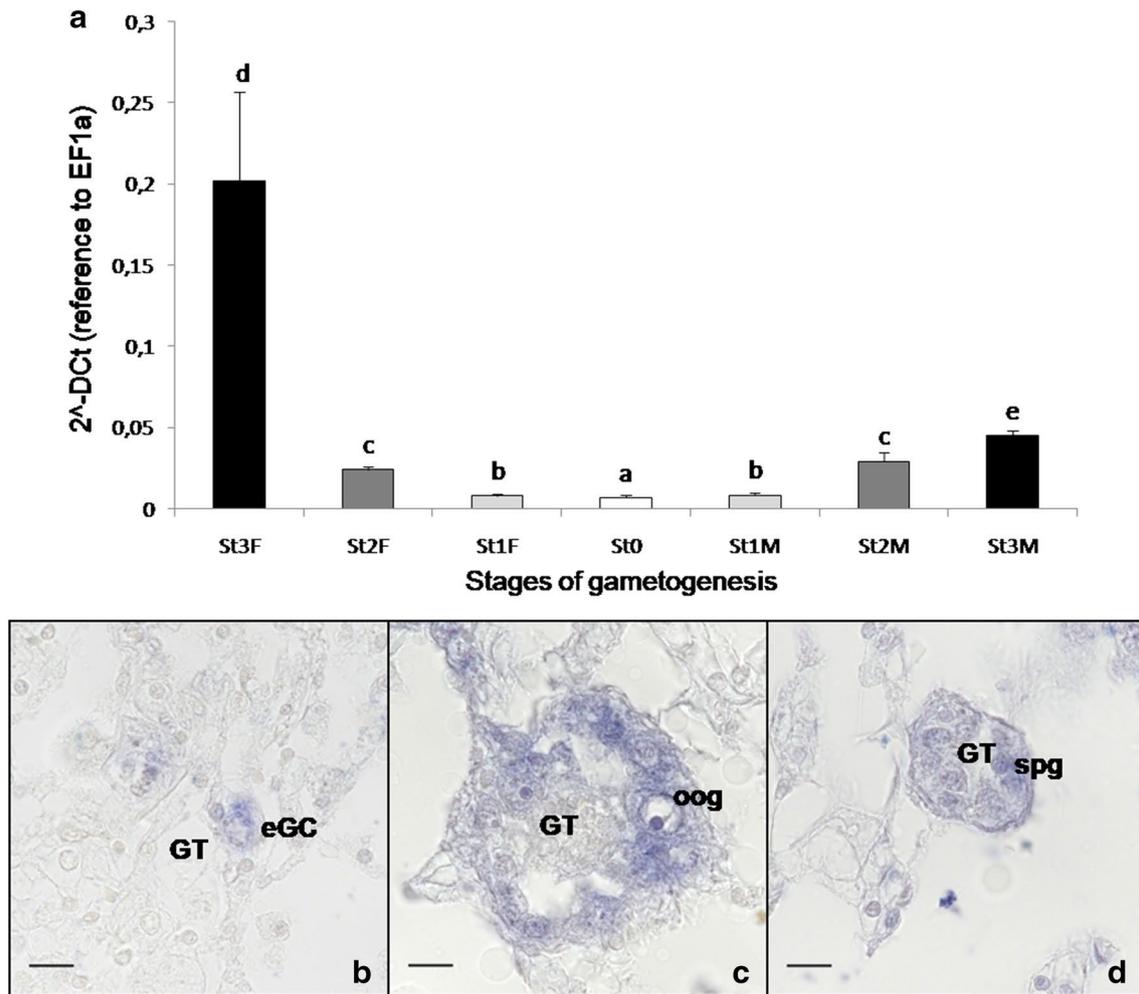


Fig. 4 Expression and location of *oylvg* transcript in *C. gigas*. **a** Quantification of *oylvg* transcript by qPCR at each stage of gametogenesis. Location of *oylvg* transcript in the *C. gigas* gonad tubules by in situ hybridization at the earliest stages of gametogenesis, the quiescent stage of gametogenesis (**b**), the gonial mitosis initiation stages in female (**c**) and male (**d**). *St0* quiescence stage, *St1F* and

M female and male gonial mitosis initiation stages, *St2F* and *2M* female and male development stages, *St3F* and *3M* female and male sexual maturity stages. Letters indicate significant differences. Gonad tubules (GT), early germ cells (eGC), oogonia (oog), spermatogonia (spg). Scale bars: 10 μm

Fig. 5 Characteristics of the early Germ Cells (eGC) and their microenvironment in *C. gigas*. **a** Feulgen staining of an early germ cell (eGC) closely related to an associated somatic cell (aSC) and a myoid cell (MC). **b** BMP2/4 immunodetection in the aSC cytoplasm in gonad tubules and negative control (inset). *eGC* early germ cell, *aSC* associated somatic cell, *MC* myoid cell, *TW* tubule wall, *GT* gonad tubule. Scale bars: 10 μm

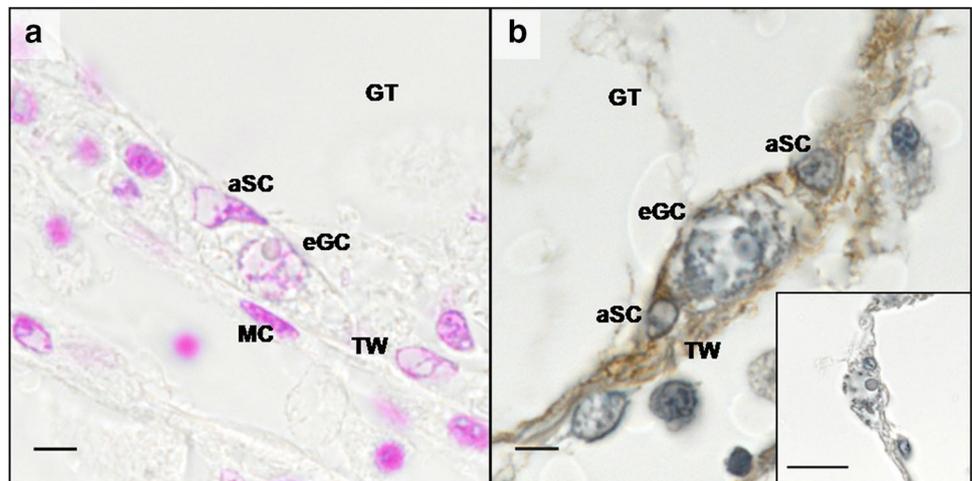


Fig. 6 Ultrastructural characteristics of the germinal niche in *C. gigas*. **a** Continuum of the two early germ cell (eGC) types in their microenvironment. **b** Ultrastructure of the niche of early germ cell with irregular nucleus (eGCir). **c** Ultrastructure of the microenvironment of early germ cell with round nucleus (eGCr). **d, e** Details of the basal lamina (BL) and the myoid cell (MC) of the microenvironment of both early germ cell types (irregular and round nucleus, respectively). **f, g** Characteristics of each associated somatic cell (aSC) observed. **h, i** Both early germ cells with irregular nucleus (eGCir) and early germ cells with round nucleus (eGCr) present a nuage (n) associated with numerous mitochondria (m). *eGCir* early germ cell with irregular nucleus, *eGCr* early germ cell with round nucleus, *aSC* associated somatic cell, *MC* myoid cell, *BL* basal lamina, *m* mitochondria, *n* nuage. Scale bars: 1 μm (**a–c**); 0.5 μm (**d–g**) and 0.2 μm (**h, i**)

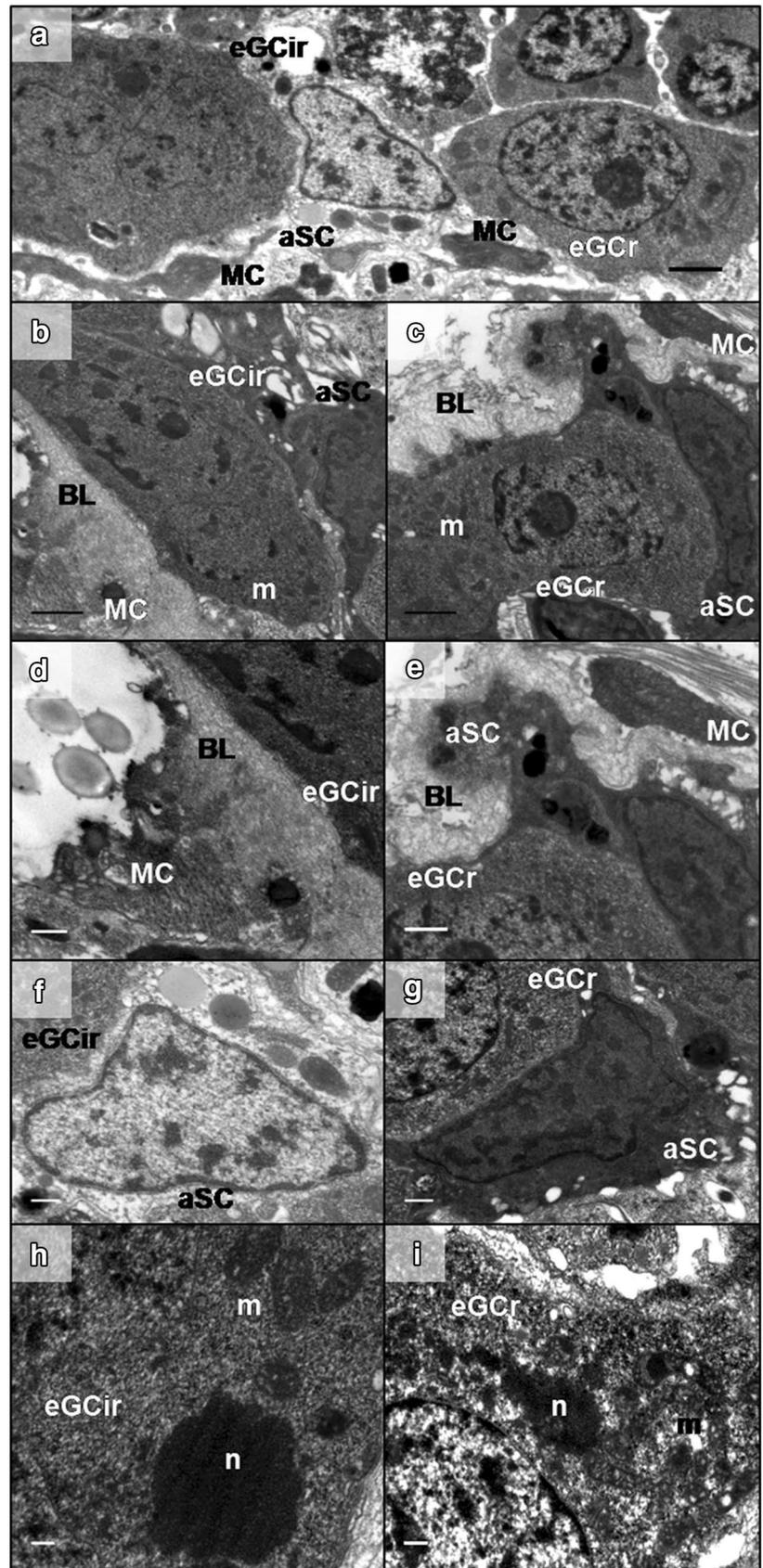
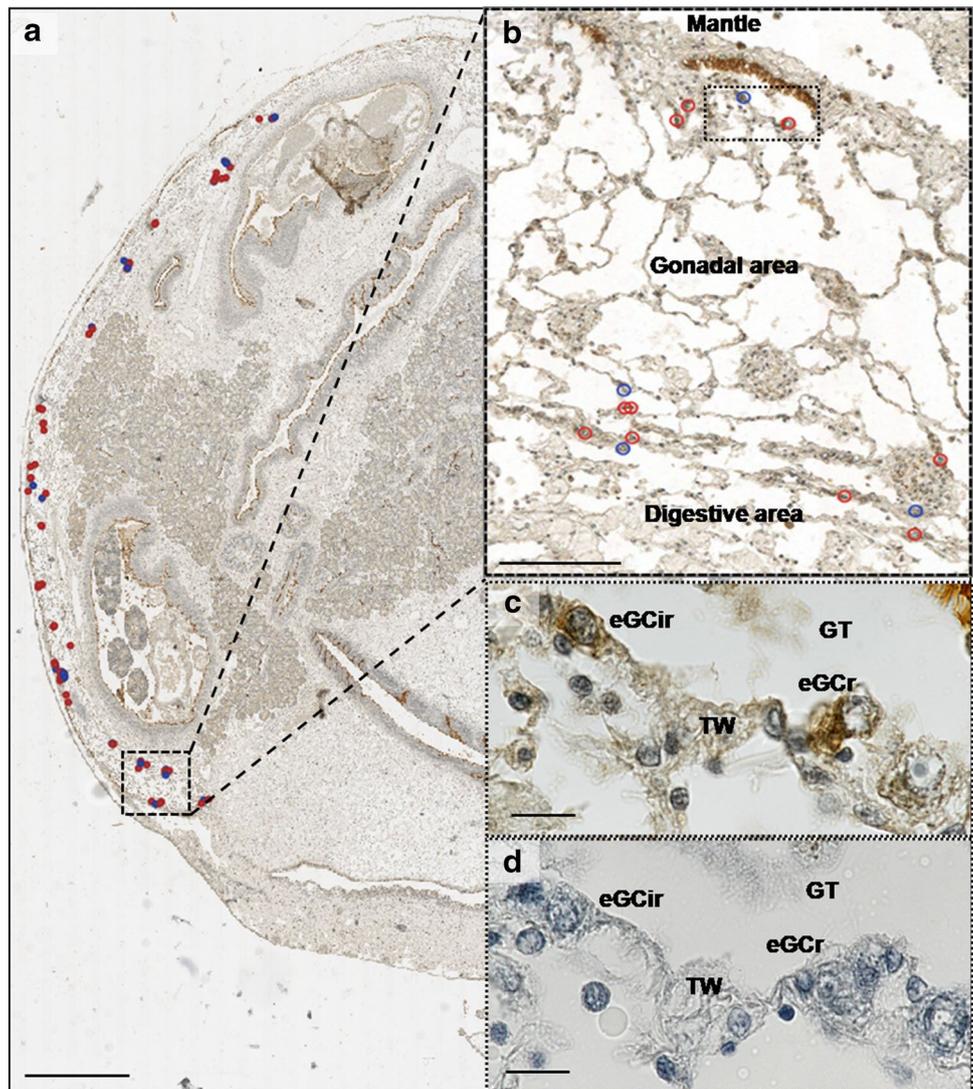


Fig. 7 Location of putative germinal niches in the gonad of *C. gigas*. **a** Whole slide image of Oyvlg labelling at the quiescence stage of gametogenesis (Stage 0). **b** Inset of the whole slide image. **c, d** Details and negative control of the dashed box area in the inset (magnification 100). Each putative germ stem cell belonging to the population of interest at this stage (nucleus diameter between 4.06 μm and 6.82 μm) are circled in the gonadal area. The red circles correspond to the round nuclei with an elongation ratio which is strictly superior to 0.8 and the blue circles represent the irregular nuclei with an elongation ratio below 0.8. *GT* gonad tubules, *TW* tubule wall, *pGSC* putative germ stem cell, *eGC* early germ cell. Scale bars: 1 mm (**a**), 0.1 mm (**b**) and 10 μm (**c, d**)



mature stages in female. The mean volumes obtained for this population of early germ cells nuclei are range from 76.91 to 97.67 μm^3 and might be partially consistent with a putative GSC type. The fact that we were unable to find any common cell nucleus diameters among all stages of gametogenesis is probably due to a technical bias. Indeed, we based our measures on a selection of cells presenting characteristics of early germ cells in each studied field. In the female, we probably selected young oogonia among the cells of interest, due to their position and low chromatin density. However, oogonia have larger nuclei leading to overestimation of mean volume. For male development and maturity stages, spermatogonia may be present among the considered cells of interest. These spermatogonia have a reduced nuclei (Franco et al. 2008) compared with early germ cells. This is why the quiescence stage, the male gonial mitosis initiation stage and the female development and maturity stages seem to be more relevant for the

further quantitative identification of the early germ cells in *C. gigas*.

Oyster vasa is expressed in the population of early germ cells

We performed Oyvlg immunodetection through the reproductive cycle to detect germinal cells in gonadal tubules. First detection of Oyvlg with DAB coloration showed a highly staining in early germ cells and a slight one just around them that could be due to a staining of associated somatic cells. However, the second immunodetection using fluorescence allowed us to conclude that Oyster Vasa positive labelling is exclusively observed in early germ cells, therefore, in the GSC. Recent data of Cavalier et al. (2017) report similar immunolabelling of Vasa in the oyster's gonad. In contrast, unlike Milani et al. (2017), we did not find any Vasa labelling in other tissues than in the gonad.

These differences between our study and the results previously obtained by Milani et al. might be due to differences in sensitivity of antibodies used. Indeed, the antibody used in our study is a *C. gigas* homologous Oyvlg antibody which links with different antigenic sequences than in Milani and collaborators' work. In complement, we noted that oylvlg transcript expression in germ cells was in accordance with the results obtained by Fabioux et al. (2004b). In our experiments, we noticed that at the quiescence stage, oylvlg transcript expression tends to appear in few germ cells. Oyvlg transcript expression increases during the gametogenesis course, probably because of the increasing amount of germ cells. Moreover, qPCR showed a peak of oylvlg expression at the female sexual maturity stage probably due to the oylvlg maternal mRNA accumulation. We noticed that vasa labelling is present in adult oyster gonad at all stages of gametogenesis that seek with a permanent gonad, germ line and GSC pool in adult in this species.

A common population of putative GSC is identified among early germ cells in male and female

For the four stages of gametogenesis (quiescence, male gonial mitosis initiation and female development and maturity stages) in which an early germ cells population was identified, we combined criteria of chromatin condensation, highlighted by Feulgen staining, and oyster Vasa immunodetection to identify putative GSC. The results confirmed the presence of a population of early germ cells that could be identified through some stemness criteria. The study of elongation ratios of nuclei in this population of early germ cells allowed us to identify two cell types, one presenting round nuclei (ratios strictly superior to 0.8) and the other with irregular nuclei (ratios equal or below to 0.8). An ultrastructural study showed that both early germ cells types present specific morphological criteria of GSC as a nucleus with slightly condensed chromatin and numerous grouped mitochondria close to a nuage structure in the cytoplasm. Both types are located against the *basal lamina*, are immunogold labelled by oyster Vasa (online resource 2) and their micro-environment consists in at least a closely associated somatic cell and a surrounding myoid cell. The only morphological difference found between the two types of early germ cells identified by the quantitative and molecular approaches is the shape of their nucleus (round or irregular). In the male, the two types of spermatogonia previously described in oyster present clearly round nuclei (Franco et al. 2008). Moreover, in tilapia, it has been shown that differentiated spermatogonia present a round nucleus while the undifferentiated spermatogonia of type A possesses an irregular nucleus well observed under electron microscopy but difficult to visualize under light microscopy (Schulz et al. 2010). In addition, in planarians, a group that is phylogenetically closely related to

the molluscan, the neoblasts, the stem cells able to differentiate in germ and somatic lines, present an irregular nucleus with slightly condensed chromatin becoming particularly apparent under transmission electron microscopy (Sato et al. 2006). All these elements enabled us to say that the identified early germ cells presenting an irregular nucleus correspond to the putative GSC in oyster and those presenting a round nucleus to the putative progenitors. This idea is supported by the more frequent observation of groups of putative progenitors (with round nucleus) in the tubule and by the fact that putative GSC (with irregular nucleus) represent only 27% of the total early germ cells population.

The germinal niche and its location are specified in the whole gonad

By the use of a heterologous antibody against BMP2/4, we observed the characteristics of a labelling specifically associated somatic cells closely related to the early germ cells identified. This result suggests the eventual involvement of TGF- β signaling pathway in the functioning of germinal niche in *C. gigas*. In *drosophila*, the TGF- β signal transduction pathway has been determined to impact the number of germline stem cells and the size of the stem cell niche (Schulz et al. 2004). Moreover Dpp, homologous of the human BMP2, well known for its implication in the germinal niche functioning in *D. melanogaster* (Xie and Spradling 1998), presents a strong homology with Cg-mGDF in *C. gigas* (Lelong et al. 2000; Fleury et al. 2008). Further studies are required to describe the proliferation and its regulation of identified putative GSC and progenitors in oyster. In addition, the somatic cells associated with the germline in oyster should be further studied to determine their functional features.

The study of the distribution of putative GSC on a Whole Slide Image (WSI) allowed us to conclude that the germinal niches in adult oyster are uniformly distributed on a gonad scale. The distribution of germinal niches in oyster from the first gametogenesis and the following ones could expand from the outermost part of the gonad (against the mantle) (Santerre et al. 2013; Cavelier et al. 2017) to the innermost part (against the digestive area).

Conclusions

Concerning the origin of GSC in *C. gigas* during development, the mRNA from the oyster vasa gene is found in vegetal pole of oocytes and becomes progressively restricted to the 4d blastomere. These two daughter cells persist without further proliferation during the subsequent larval stages and, are therefore, believed to be the PGC of *C. gigas* (Fabioux et al. 2004a). Concerning Cg-Nanos-like, transcripts

were also localized in the vegetal pole of oocytes but are uniformly expressed at early cleavage stages and appear restricted to two cells clusters at gastrulation until umbo-larvae which could be the putative PGC (Xu et al. 2018). Classically, after their migration in the pre-adult gonad, PGC with niche cellular components become GSC which become differentiated germ cells during gametogenesis (Saffman and Lasko 1999).

However, the issue of the nature of stem cells and the renewal of the somatic and germlines at each reproductive cycle remains open in *C. gigas*. Two scenarios should be taken into account. The first one consists on the persistence of a single stem cell able to give rise of both somatic and germinal lineages in the adult gonad, which could be in accordance with the definition of primordial stem cells (PriSC) of Solana (2013). In this case, our study identified a putative GSC that could derive from this single stem cell still unidentified in our work. The second possibility considers that the PGC in oyster are strictly germinal and become GSC at the first gametogenesis in adult. In this case, we characterised in our study the putative GSC in adult oyster and the somatic stem cell at the origin of the intragonadal somatic line is yet to be identified. The use of complementary conserved markers of stemness and GSC in *C. gigas* (for example KLF, Piwi, Nanos) could allow identifying cells to validate one or other of these possibilities.

Acknowledgements PhD of M. Cherif-Feildel was financially supported by the French Research Minister (PhD grant). The authors are grateful to C. Fabioux (Université de Bretagne occidentale, Institut universitaire européen de la mer, LEMAR) and C. Corporeau (Ifremer, LPI) for the oyster vasa antibody supplying. The authors thank J.P. Robin for essential advice on all the statistical aspects of this study. Proofreading of English text has been realised by C. Quint of the translation centre of the Carré international UNICAEN.

Compliance with ethical standards

Conflict of interest The authors declare no competing financial interests.

References

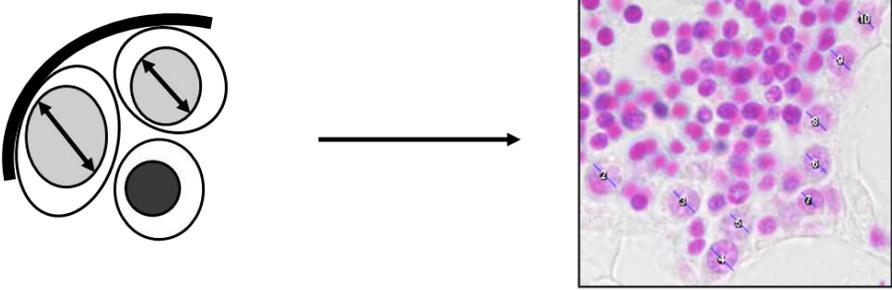
- Adoutte A, Balavoine G, Lartillot N, Lespinet O, Prud'homme B, de Rosa R (2000) The new animal phylogeny: reliability and implications. *Proc Natl Acad Sci USA* 97:4453–4456. <https://doi.org/10.1073/pnas.97.9.4453>
- Baguñà J (2012) The planarian neoblast: the rambling history of its origin and some current black boxes. *Int J Dev Biol* 56:19–37. <https://doi.org/10.1387/ijdb.113463jb>
- Bankhead P, Loughrey MB, Fernández JA, Dombrowski Y, McArt DG, Dunne PD, McQuaid S, Gray RT, Murray LJ, Coleman HG, James JA, Salto-Tellez M, Hamilton PW (2017) QuPath: open source software for digital pathology image analysis. *Sci Rep* 7:1–7. <https://doi.org/10.1038/s41598-017-17204-5>
- Berthelin C, Kellner K, Mathieu M (2000a) Histological characterization and glucose incorporation into glycogen of the Pacific oyster *Crassostrea gigas* storage cells. *Mar Biotechnol* 2:136–145. <https://doi.org/10.1007/s101269900017>
- Berthelin C, Kellner K, Mathieu M (2000b) Storage metabolism in the Pacific oyster (*Crassostrea gigas*) in relation to summer mortalities and reproductive cycle (West Coast of France). *Comp Biochem Physiol B Biochem Mol Biol* 125:359–369. [https://doi.org/10.1016/S0305-0491\(99\)00187-X](https://doi.org/10.1016/S0305-0491(99)00187-X)
- Brown DL (2017) Bias in image analysis and its solution: unbiased stereology. *J Toxicol Pathol* 30:183–191. <https://doi.org/10.1293/tox.2017-0013>
- Brown FD, Swalla BJ (2012) Evolution and development of budding by stem cells: Ascidian coloniality as a case study. *Dev Biol* 369:151–162. <https://doi.org/10.1016/j.ydbio.2012.05.038>
- Cavelier P, Cau J, Morin N, Delsert C (2017) Early gametogenesis in the Pacific oyster: new insights using stem cell and mitotic markers. *J Exp Biol*. <https://doi.org/10.1242/jeb.167734>
- Chiari-Garcia H, Russell LD (2001) High-resolution light microscopic characterization of mouse spermatogonia. *Biol Reprod* 65:1170–1178
- Chiari-Garcia H, Russell LD (2002) Ultrastructure cellulaire et tissulaire. *Approch Reprod* 123:567–577. <https://doi.org/10.1530/rep.0.1230567>
- Chong T, Stary JM, Wang Y, Newmark P (2011) Molecular markers to characterize the hermaphroditic reproductive system of the planarian Schmidtea mediterranea. *BMC Dev Biol* 11:69. <https://doi.org/10.1186/1471-213X-11-69>
- Cruz-Orive L-M, Hunziker EB (1986) Stereology for anisotropic cells: application to growth cartilage. *J Microsc* 143:47–80. <https://doi.org/10.1111/j.1365-2818.1986.tb02765.x>
- Dheilly NM, Lelong C, Huvet A, Favrel P (2011) Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns. *BMC Genom* 12:468. <https://doi.org/10.1186/1471-2164-12-468>
- Dheilly NM, Lelong C, Huvet A, Kellner K, Dubos M-P, Riviere G, Boudry P, Favrel P (2012) Gametogenesis in the Pacific oyster *Crassostrea gigas*: a microarrays-based analysis identifies sex and stage specific genes. *PLoS One* 7:e36353. <https://doi.org/10.1371/journal.pone.0036353>
- Dubos MP, Bernay B, Favrel P (2016) Molecular characterization of an adipokinetic hormone-related neuropeptide (AKH) from a mollusk. *Gen Comp Endocrinol* 243:15–21. <https://doi.org/10.1016/j.ygcen.2016.11.002>
- Fabioux C, Huvet A, Lelong C, Robert R, Pouvreau S, Daniel J, Minguant C, Le Pennec M (2004a) Oyster vasa-like gene as a marker of the germline cell development in *Crassostrea gigas*. *Biochem Biophys Res Commun* 320:592–598. <https://doi.org/10.1016/j.bbrc.2004.06.009>
- Fabioux C, Pouvreau S, Le Roux F, Huvet A (2004b) The oyster vasa-like gene: a specific marker of the germline in *Crassostrea gigas*. *Biochem Biophys Res Commun* 315:897–904. <https://doi.org/10.1016/j.bbrc.2004.01.145>
- 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>
- Ferraro F, Lo Celso C, Scadden D (2010) Adult stem cells and their niches. *Adv Exp Med Biol* 695:155–168. <https://doi.org/10.1007/978-1-4419-7037-4>
- Fleury E, Fabioux C, Lelong C, Favrel P, Huvet A (2008) Characterization of a gonad-specific transforming growth factor- β superfamily member differentially expressed during the reproductive cycle of the oyster *Crassostrea gigas*. *Gene* 410:187–196

- Franco A, Heude Berthelin C, Goux D, Sourdain P, Mathieu M (2008) Fine structure of the early stages of spermatogenesis in the Pacific oyster, *Crassostrea gigas* (Mollusca, Bivalvia). *Tissue Cell* 40:251–260. <https://doi.org/10.1016/j.tice.2007.12.006>
- Franco A, Jouaux A, Mathieu M, Sourdain P, Lelong C, Kellner K, Heude Berthelin C (2010) Proliferating cell nuclear antigen in gonad and associated storage tissue of the Pacific oyster *Crassostrea gigas*: seasonal immunodetection and expression in laser microdissected tissues. *Cell Tissue Res* 340:201–210. <https://doi.org/10.1007/s00441-009-0923-6>
- Gabe M (1968) *Techniques histologiques*, 6th edn. Masson, Paris
- Garcia Y, Breen A, Burugapalli K, Dockery P, Pandit A (2007) Stereological methods to assess tissue response for tissue-engineered scaffolds. *Biomaterials* 28:175–186. <https://doi.org/10.1016/j.biomaterials.2006.08.037>
- Golub VM, Brewer J, Wu X, Kuruba R, Short J, Manchi M, Swonke M, Younus I, Reddy DS (2015) Neurostereology protocol for unbiased quantification of neuronal injury and neurodegeneration. *Front Aging Neurosci* 7:1–14. <https://doi.org/10.3389/fnagi.2015.00196>
- Gundersen H, Jensen E (1985) Stereological estimation of the volume-weighted mean volume of arbitrary particles observed on random sections. *J Microsc* 138:127–142. <https://doi.org/10.1111/j.1365-2818.1985.tb02607.x>
- Herpin A, Lelong C, Favrel P (2004) Transforming growth factor- β -related proteins: an ancestral and widespread superfamily of cytokines in metazoans. *Dev Comp Immunol* 28:461–485. <https://doi.org/10.1016/j.dci.2003.09.007>
- Herpin A, Lelong C, Becker T, Rosa F, Favrel P, Cunningham C (2005) Structural and functional evidence for a singular repertoire of BMP receptor signal transducing proteins in the lophotrochozoan *Crassostrea gigas* suggests a shared ancestral BMP/activin pathway. *FEBS J* 272:3424–3440. <https://doi.org/10.1111/j.1742-4658.2005.04761.x>
- Heude Berthelin C, Laisney J, Espinosa J, Martin O, Hernandez G, Mathieu M, Kellner K (2001) Storage and reproductive strategy in *Crassostrea gigas* from two different growing areas (Normandy and the Atlantic coast, France). *Invertebr Reprod Dev* 40:79–86. <https://doi.org/10.1080/07924259.2001.9652500>
- Huvet A, Fleury E, Corporeau C, Quillien V, Daniel JY, Riviere G, Boudry P, Fabioux C (2012) In vivo RNA interference of a gonad-specific transforming growth factor- β in the Pacific oyster *Crassostrea gigas*. *Mar Biotechnol* 14:402–410. <https://doi.org/10.1007/s10126-011-9421-4>
- Jagut M, Huynh JR (2007) Régulation des cellules souches de la lignée germinale. *Med Sci*. <https://doi.org/10.1051/medsci/20072367611>
- Jemaà M, Morin N, Cavelier P, Cau J, Strub JM, Delsert C (2014) Adult somatic progenitor cells and hematopoiesis in oysters. *J Exp Biol* 217:3067–3077. <https://doi.org/10.1242/jeb.106575>
- Johnson L, Thompson DL, Varner DD (2008) Role of Sertoli cell number and function on regulation of spermatogenesis. *Anim Reprod Sci* 105:23–51. <https://doi.org/10.1016/j.anireprosci.2007.11.029>
- Jouaux A, Heude Berthelin C, Sourdain P, Mathieu M, Kellner K (2010) Gametogenic stages in triploid oysters *Crassostrea gigas*: Irregular locking of gonial proliferation and subsequent reproductive effort. *J Exp Mar Bio Ecol* 395:162–170. <https://doi.org/10.1016/j.jembe.2010.08.030>
- Juliano CE, Swartz SZ, Wessel GM (2010) A conserved germline multipotency program. *Development* 137:4113–4126. <https://doi.org/10.1242/dev.047969>
- Ladekarl M (1998) Objective malignancy grading: a review emphasizing unbiased stereology applied to breast tumors. *APMIS Suppl* 79:1–34
- Lasko P (2013) The DEAD-box helicase vasa: evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochim Biophys Acta Gene Regul Mech* 1829:810–816. <https://doi.org/10.1016/j.bbagr.2013.04.005>
- Le Quere H, Herpin A, Huvet A, Lelong C, Favrel P (2009) Structural and functional characterizations of an Activin type II receptor orthologue from the Pacific oyster *Crassostrea gigas*. *Gene* 436:101–107. <https://doi.org/10.1016/j.gene.2009.01.010>
- Lelong C, Mathieu M, Favrel P (2000) Structure and expression of mGDF, a new member of the transforming growth factor-beta superfamily in the bivalve mollusc *Crassostrea gigas*. *Eur J Biochem* 267:3986–3993. <https://doi.org/10.1046/j.1432-1327.2000.01432.x>
- Maceren-Pates M, Kurita Y, Pates G, Yoshikuni M (2015) A model for germ cell development in a fully segmented worm. *Zool Lett* 1:34. <https://doi.org/10.1186/s40851-015-0035-y>
- Meijer GA, Baak JPA, Talbot IC, Atkin WS, Meuwissen SGM (1998) Predicting the risk of metachronous colorectal cancer in patients with rectosigmoid adenomas using quantitative pathological features. A case-control study. *J Pathol* 184:63–70
- Milani L, Maurizii MG, Pecci A, Ghiselli F, Passamonti M (2015) Localization and dynamics of primordial germ cells in the bivalve species *Ruditapes philippinarum*. *Mol Reprod Dev* 82:406–407. <https://doi.org/10.1002/mrd.22502>
- Milani L, Pecci A, Ghiselli F, Passamonti M, Bettini S, Franceschini V, Maurizii MG (2017) VASA expression suggests shared germ line dynamics in bivalve molluscs. *Histochem Cell Biol* 148:157–171. <https://doi.org/10.1007/s00418-017-1560-x>
- Morrison SJ, Spradling AC (2008) Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* 132:598–611. <https://doi.org/10.1016/j.cell.2008.01.038>
- Newmark PA, Alvarado AS (2000) Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev Biol* 220:142–153. <https://doi.org/10.1006/dbio.2000.9645>
- Obata M, Sano N, Kimata S, Nagasawa K, Yoshizaki G, Komaru A (2010) The proliferation and migration of immature germ cells in the mussel, *Mytilus galloprovincialis*: observation of the expression pattern in the *M. galloprovincialis* vasa-like gene (Myvlg) by in situ hybridization. *Dev Genes Evol* 220:139–149. <https://doi.org/10.1007/s00427-010-0335-3>
- Prisco M, Agnese M, De Marino A, Andreuccetti P, Rosati L (2017) Spermatogenic cycle and steroidogenic control of spermatogenesis in *Mytilus galloprovincialis* collected in the bay of Naples. *Anat Rec* 300:1881–1894. <https://doi.org/10.1002/ar.23626>
- Rink JC (2013) Stem cell systems and regeneration in planaria. *Dev Genes Evol* 223:67–84. <https://doi.org/10.1007/s00427-012-0426-4>
- Rossi L, Salvetti A, Batistoni R, Deri P, Gremigni V (2008) Planarians, a tale of stem cells. *Cell Mol Life Sci* 65:16–23. <https://doi.org/10.1007/s00018-007-7426-y>
- Saffman EEF, Lasko P (1999) Germline development in vertebrates and invertebrates. *Cell Mol Life Sci* 55:1141–1163. <https://doi.org/10.1007/s000180050363>
- Santerre C, Sourdain P, Marc N, Mingant C, Robert R, Martinez A (2013) Oyster sex determination is influenced by temperature—First clues in spat during first gonadic differentiation and gametogenesis. *Comp Biochem Physiol Part A* 165:61–69. <https://doi.org/10.1016/j.cbpa.2013.02.007>
- Sato K, Shibata N, Orii H, Amikura R, Sakurai T, Agata K, Kobayashi S, Watanabe K (2006) Identification and origin of the germline stem cells as revealed by the expression of nanos-related gene in planarians. *Dev Growth Differ* 48:615–628. <https://doi.org/10.1111/j.1440-169X.2006.00897.x>
- Scadden DT (2006) The stem-cell niche as an entity of action. *Nature* 441:1075–1079. <https://doi.org/10.1038/nature04957>
- Schmitz C, Hof PR (2005) Design-based stereology in neuroscience. *Neuroscience* 130:813–831. <https://doi.org/10.1016/j.neurosci.2004.08.050>

- Schulz C, Kiger AA, Tazuke SI, Yamashita YM, Pantalena-Filho LC, Jones DL, Wood CG, Fuller MT (2004) A misexpression screen reveals effects of bag-of-marbles and TGF β class signaling on the Drosophila male germ-line stem cell lineage. *Genetics* 167:707–723. <https://doi.org/10.1534/genetics.103.023184>
- Schulz RW, de França LR, Lareyre JJ, LeGac F, Chiarini-Garcia H, Nobrega RH, Miura T (2010) Spermatogenesis in fish. *Gen Comp Endocrinol* 165:390–411. <https://doi.org/10.1016/j.ygcen.2009.02.013>
- Silva RC, Costa GMJ, Andrade LM, França LR (2010) Testis stereology, seminiferous epithelium cycle length, and daily sperm production in the ocelot (*Leopardus pardalis*). *Theriogenology* 73:157–167. <https://doi.org/10.1016/j.theriogenology.2009.08.009>
- Solana J (2013) Closing the circle of germline and stem cells: the primordial stem cell hypothesis. *Evodevo* 4:2. <https://doi.org/10.1186/2041-9139-4-2>
- Spradling A, Fuller MT, Braun RE, Yoshida S (2011) Germline stem cells. *Cold Spring Harb Perspect Biol* 3:1–20. <https://doi.org/10.1101/cshperspect.a002642>
- Steele S, Mulcahy M (1999) Gametogenesis of the oyster *Crassostrea gigas* in southern Ireland. *J Mar Biol Ass UK* 79:673–686. <https://doi.org/10.1017/S0025315498000836>
- Xie T, Spradling AC (1998) Decapentaplegic is essential for the maintenance and division of germline stem cells in the Drosophila ovary. *Cell* 94:251–260. [https://doi.org/10.1016/S0092-8674\(00\)81424-5](https://doi.org/10.1016/S0092-8674(00)81424-5)
- Xu R, Li Q, Yu H, Kong L (2018) Oocyte maturation and origin of the germline as revealed by the expression of Nanos-like in the Pacific oyster *Crassostrea gigas*. 663:41–50. <https://doi.org/10.1016/j.gene.2018.04.021>
- Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Wang J, Zhang L, Wang X, Qi H, Xiong Z, Que H, Xie Y, Holland PWH, Paps J, Zhu Y, Wu F, Chen Y, Wang J, Peng C, Meng J, Yang L, Liu J, Wen B, Zhang N, Huang Z, Zhu Q, Feng Y, Mount A, Hedgecock D, Xu Z, Liu Y, Domazet-Lošo T, Du Y, Sun X, Zhang S, Liu B, Cheng P, Jiang X, Li J, Fan D, Wang W, Fu W, Wang T, Wang B, Zhang J, Peng Z, Li Y, Li N, Wang J, Chen M, He Y, Tan F, Song X, Zheng Q, Huang R, Yang H, Du X, Chen L, Yang M, Gaffney PM, Wang S, Luo L, She Z, Ming Y, Huang W, Zhang S, Huang B, Zhang Y, Qu T, Ni P, Miao G, Wang J, Wang Q, Steinberg CEW, Wang H, Li N, Qian L, Zhang G, Li Y, Yang H, Liu X, Yin Y, Wang J (2012) The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490:49–54. <https://doi.org/10.1038/nature11413>
- Zhang L, Li L, Guo X, Litman GW, Dishaw LJ, Zhang G (2015) Massive expansion and functional divergence of innate immune genes in a protostome. *Sci Rep* 5:1–11. <https://doi.org/10.1038/srep08693>
- Zhao X, Yu H, Kong L, Li Q (2012) Transcriptomic responses to salinity stress in the Pacific oyster *Crassostrea gigas*. *PLoS One*. <https://doi.org/10.1371/journal.pone.0046244>

Quantitative histology, step 1 (Feulgen staining)

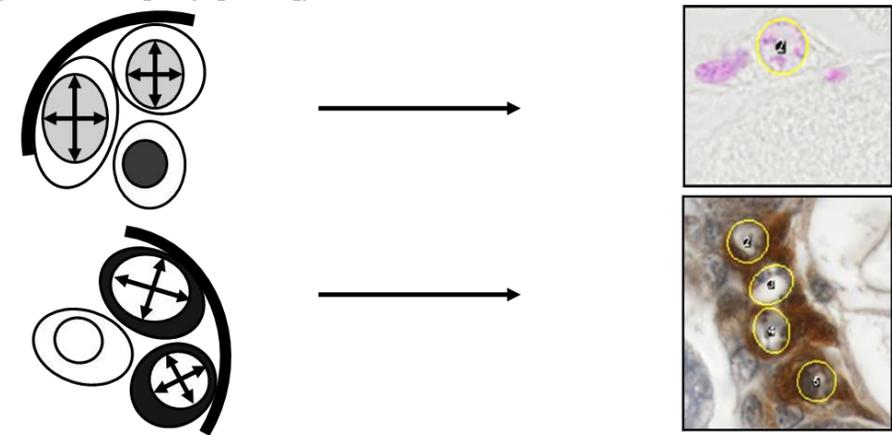
1- Measurement by line-intercept sampling (random diameters) of each nucleus of interest into cells along the tubule wall with a slightly condensed chromatin at each stage of the gametogenesis.



2- Calculation of the volume-weighted mean volume for each field.

Quantitative histology, step 2 (Feulgen staining and Oyvlg immunolabelling)

1- Ellipse track for computer calculations of the longest and shortest diameters of each nucleus of cells of interest (along the tubule wall presenting a slightly condensed chromatin after Feulgen staining or Oyvlg positive according to Oyvlg labelling).



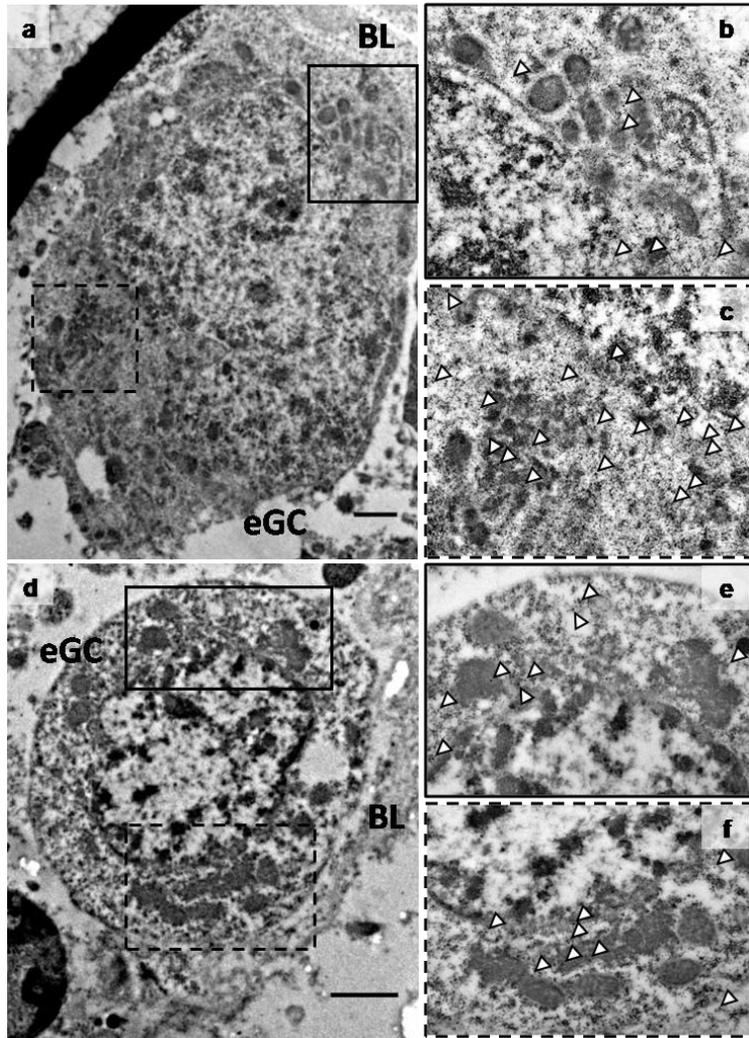
2- Calculation of the elongation ratio by dividing the shortest diameter by the longest for each nucleus.

Online resource 1: Schematic representation of a two-step histological quantitative approach.

Supplementary material

Immunogold material and methods

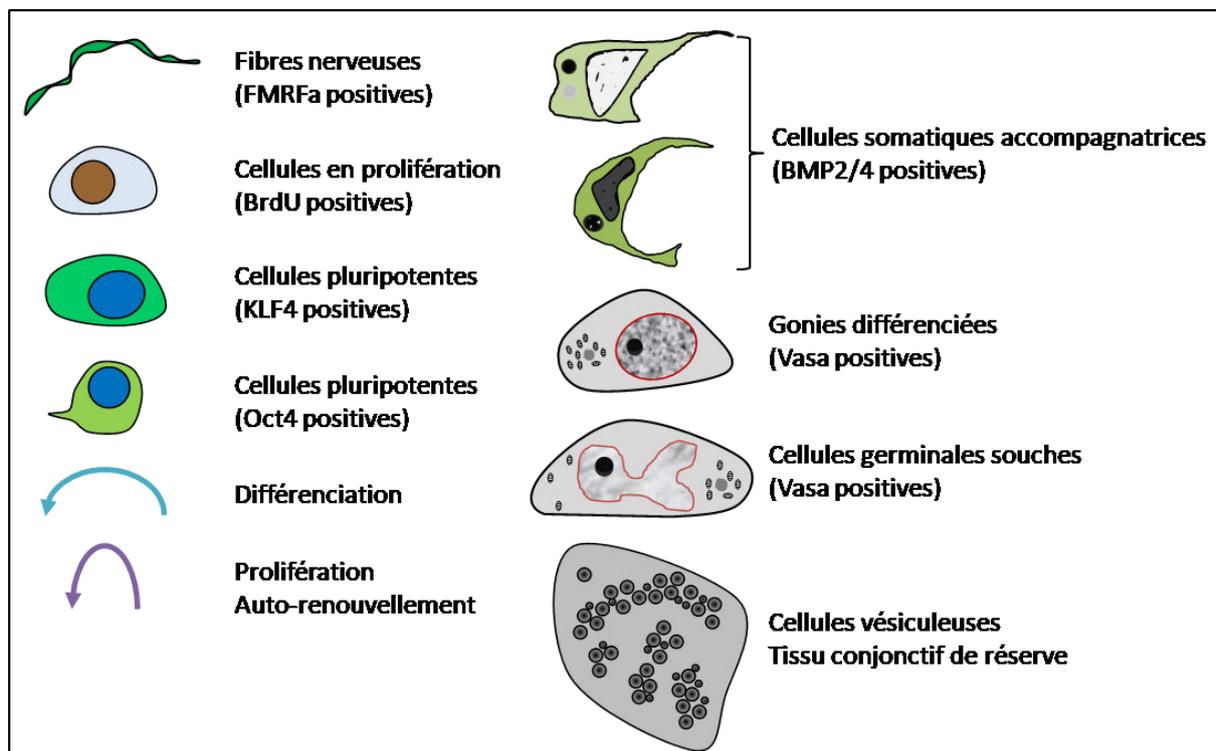
Small gonadal area tissue samples were collected and immersed in a fixative solution (4% paraformaldehyde, 0.1% glutaraldehyde, 0.2 M phosphate buffer, 0.25 M saccharose) for 3 hours. Samples were rinsed in 0.2 M phosphate buffer and 0.25 M saccharose (pH 7.4) before dehydration in serial ethanol's dilutions. Impregnation was done by 1 hour incubations in ascending concentrations of LRWhite Resin. Tissue samples were put in BEEM capsules with LRWhite HARD for 90 min before hot polymerization at 50°C for 24 hours. Dehydration and inclusion were done with cryosubstitution apparatus FS8500 RMC. Ultrathin sections of 3 individuals were treated for immunogold labeling at room temperature. A wetting step was done, 15 min in 0.05 M glycine/PBS, followed by 30 min in blocking solution (Aurion, The Netherlands). Sections were rinsed in 0.1% BSA-c (Aurion, The Netherlands)/PBS and incubated with the Oyvlg rabbit polyclonal primary antibody diluted at 1:50 in 0.1% BSA-c/PBS for 1 hour. After being rinsed as before, sections were incubated with secondary antibody goat anti-rabbit IgG gold particle size 10 nm (Aurion, The Netherlands, 810-011) diluted at 1:40 in 0.1% BSA-c/PBS for 90 min. Rinses in 0.1% BSA-c/PBS and then in PBS were performed. Post fixation was realized using a solution of PBS with 1% of glutaraldehyde for 5 min before rinses in PBS and then in distilled water. After contrasting the sections with 5% uranyl acetate for 30 min and rinses in distilled water, ultrathin sections were examined with a JEOL1011 transmission electron microscope.



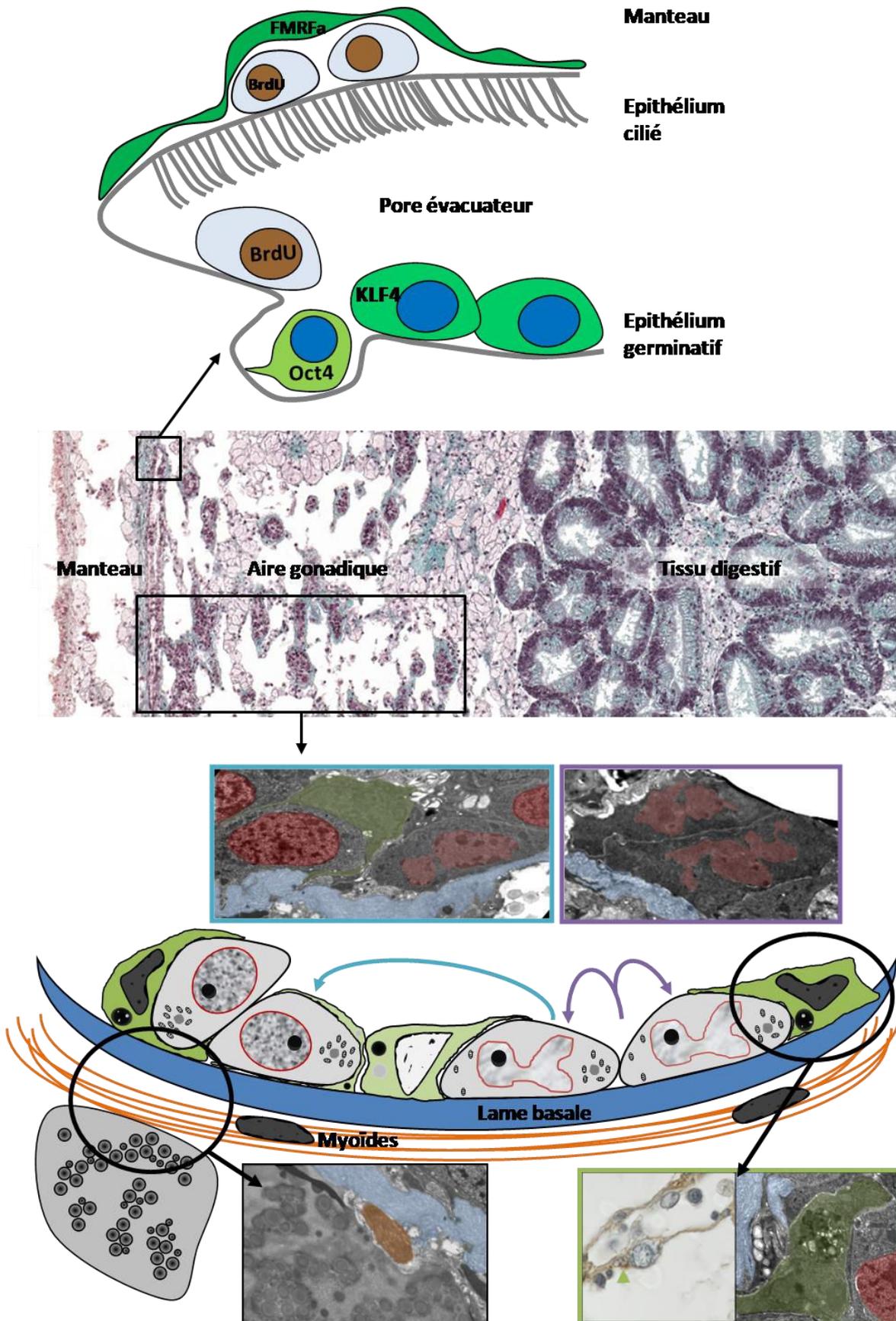
Online resource 2: Detection of Oylg protein by Immunogold labelling in early GC (eGC) with irregular nucleus (a) and with round one (d). Both are localized against the *basal lamina* (BL). Labeling, highlighted by white triangles, appears to be present near to the mitochondria (insets b, e, f) and the nuage (inset c). Scale bars: 1 μ m

Bilan du chapitre I

- ✓ Observation et localisation de cellules à potentiel souche positives pour KLF4 et Oct4 en immunohistochimie dans l'aire gonadique.
- ✓ Identification d'îlots de cellules en prolifération dans l'épithélium germinatif au niveau des pores évacuateurs par une approche BrdU.
- ✓ Identification d'une population pérenne de cellules germinales précoces chez le mâle et la femelle composée de GSC et progéniteurs potentiels.
- ✓ Répartition des GSC potentielles *a priori* aléatoire dans l'ensemble de la gonade.
- ✓ Descriptions des premiers éléments de la niche germinale à savoir, une cellule somatique accompagnatrice, la lame basale et une cellule myoïde.



Structure de l'épithélium germinatif chez *C. gigas*



Annexe – Cas d'un hermaphrodite simultané

De rares cas d'hermaphroditisme simultané peuvent être observés chez *C. gigas*. Dans ces cas, lors de la maturité sexuelle, les ovocytes et spermatozoïdes sont visibles au sein d'un même tubule gonadique ainsi que des cellules germinales précoces (eGC) proche de la membrane basale (Figure I-17A). La coloration Feulgen réalisée sur cet hermaphrodite simultané révèle la présence de cellules possédant une chromatine décondensée, proche du bord du tubule gonadique (Figure I-17B). Les cellules avec une chromatine très condensée correspondent aux spermatozoïdes au centre du tubule entourés des ovocytes. Les immunomarquages Oyvlg ont permis de marquer des cellules germinales précoces présentant les mêmes caractéristiques que les cellules de la population commune aux mâles et aux femelles (Figure I-17C). L'immunohistochimie avec l'anticorps hétérologue anti-BMP2/4 a permis d'observer des cellules somatiques positives dans le tubule gonadique, proche de cellules germinales précoces (Figure I-17D).

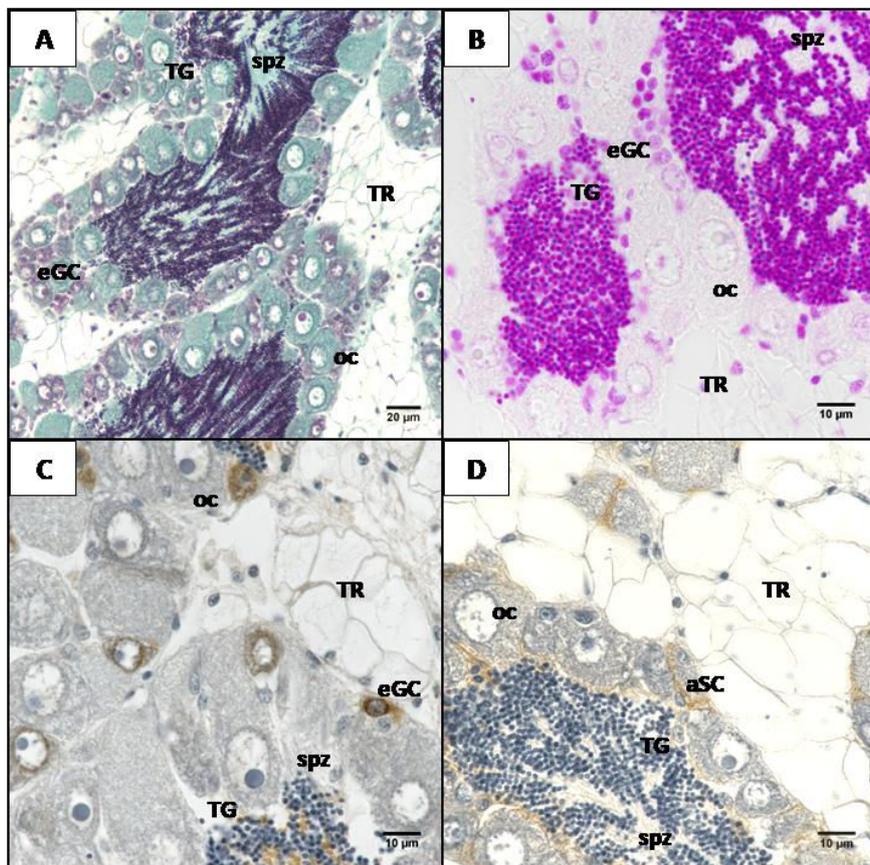


Figure I-17 : Caractéristiques des cellules germinales précoces chez un hermaphrodite simultané. A et B – Colorations Trichrome Prenant-Gabe (objectif 40) et Feulgen (objectif 100) d'une coupe histologique d'aire gonadique d'un hermaphrodite simultané respectivement. C et D – Localisation de la protéine Oyvlg (anticorps homologue) et de BMP2/4 (anticorps hétérologue) respectivement. TG : Tubules gonadiques ; TR : Tissu conjonctif de réserve ; eGC : Cellules Germinales précoces ; spz : spermatozoïdes ; oc : ovocytes ; aSC : Cellules Somatiques accompagnatrices.

Chapitre II :
Etude évolutive et fonctionnelle
du système insuline

(La première partie de ce chapitre a fait l'objet d'une publication et d'un article de type « Data in brief » accepté dans General et Comparative Endocrinology.)

Chapitre II : Etude évolutive et fonctionnelle du système insuline

1. Introduction

Comme énoncé en introduction générale, l'implication des molécules de la superfamille des insulines dans le contrôle de la reproduction et plus particulièrement dans le fonctionnement de la niche germinale a été décrite chez différentes espèces (Hsu et Drummond-Barbosa, 2011; Hubbard, 2011; Nóbrega *et al.*, 2015; Oatley et Brinster, 2012). De plus, un très grand nombre d'acteurs des voies des IRPs ont été décrits chez les Lophotrochozoaires, incluant les mollusques. Chez *Crassostrea gigas*, des travaux antérieurs de l'équipe ont montré l'implication des voies des IRPs dans les étapes précoces de gamétogenèse (Gricourt *et al.*, 2006; Jouaux *et al.*, 2012), ce qui en fait un bon candidat pour les voies de signalisation impliquées dans le fonctionnement de la niche chez cette espèce. Afin de pouvoir étudier l'implication des voies des IRPs dans la régulation des GSC chez *C. gigas*, il importait en premier lieu d'identifier les différents effecteurs conservés. La disponibilité du génome ainsi que des transcriptomes a permis la recherche *in silico* de ces effecteurs dans les bases de données (Rivière *et al.*, 2015; Zhang *et al.*, 2012; Zhao *et al.*, 2012). Les acteurs des voies des IRPs peuvent ainsi être identifiés sur la base de la conservation phylogénétique et leurs séquences peuvent être caractérisées. Ce criblage des bases de données d'expression complété par suivis directs en qPCR ou en hybridation *in situ* constitue la première étape indispensable à la compréhension des rôles fonctionnels des IRPs.

1.1. Caractérisation des ligands IRPs

Tous les ligands de la famille insuline sont regroupés sous le terme Insulin Related Peptides (IRPs). Ce terme comprend les insulines et les Insulin-like Growth Factors (IGFS), identifiés en premier lieu chez les vertébrés, mais également les relaxines ainsi qu'un grand nombre de molécules structurellement apparentées à l'insuline. Les membres de la superfamille ont été identifiés, aussi bien chez les vertébrés que chez les invertébrés, grâce à leur conservation structurale (Lu *et al.*, 2005).

1.1.1. Propriétés structurales

Du fait d'une conservation relativement faible des séquences, l'identification des IRPs est généralement basée sur la caractérisation de domaines spécifiques. La conformation des IRPs est en partie déterminée par le nombre et l'emplacement des ponts disulfures. Comme énoncé en introduction générale, les cystéines responsables des ponts disulfures sont réparties dans les chaînes B et A des insulines et sont essentielles à la configuration de la forme mature. Les séquences des insulines contiennent six à huit cystéines conservées, permettant la formation de trois ou quatre ponts disulfures entre les chaînes B et A ou à l'intérieur de la chaîne A. En plus de ces chaînes, plusieurs domaines facultatifs ont été identifiés dans certains IRPs, notamment le peptide C, lui aussi important pour la bonne conformation des insulines. Cependant, ce domaine facultatif est potentiellement clivé dans la forme mature des IRPs. Pour toutes les molécules de la famille insuline, la structure B-C-A est conservée, permettant d'identifier les ligands insuline à partir de caractéristiques structurales. D'autres domaines facultatifs ont été initialement identifiés chez les vertébrés. Ces domaines sont nommés D ou E et sont observés à l'extrémité C-terminale du peptide avant maturation. La différence majeure entre ces deux domaines est que le domaine D est conservé dans la forme mature alors que le domaine E est clivé (Claeys *et al.*, 2002).

1.1.2. Les différentes structures des IRPs chez les mollusques

Chez les mollusques, les premiers IRPs ont été identifiés chez le gastéropode *Lymnaea stagnalis* à partir d'ADNc spécifiques des cellules du système nerveux central. Ces IRPs ont été nommés MIPs (Molluscan Insulin-related Peptides). Les séquences nucléotidiques obtenues ont conduit à la prédiction des séquences en acides aminés. Leur analyse a permis l'identification des domaines caractéristiques des IRPs à savoir, un peptide signal, les chaînes B et A ainsi que le peptide C. L'alignement des MIPs avec d'autres IRPs, non issus des mollusques, a révélé la présence de six cystéines conservées pour tous et de deux supplémentaires, caractéristiques des MIPs de *L. stagnalis*. Ces deux cystéines supplémentaires permettent la formation d'un pont disulfure additionnel entre les chaînes B et A. Cette caractéristique permet de les classer dans le type β . Une étude par hybridation *in situ* montre que leur expression est exclusivement observée dans les cellules nommées Light-Green Cells (LGC) localisées dans les ganglions nerveux (Smit *et al.*, 1988, 1991, 1992, 1993, 1996, 1998). A ce jour, cinq gènes codants pour des MIPs et deux pseudogènes ont été

identifiés chez *L. stagnalis* (Smit *et al.*, 1988, 1991, 1992, 1993, 1996, 1998). Parmi les cinq MIPs caractérisés, seul MIP I possède un peptide C sans site de clivage interne. Les autres MIPs possèdent un peptide C contenant un site de clivage dibasique (KR) interne (Figure II-1). L'hypothèse de l'implication des MIPs dans la spermatogenèse et la maturation ovocytaire a été proposée mais doit être confirmée (Geraerts, 1976; Geraerts *et al.*, 1992; Van Minnen *et al.*, 1989).

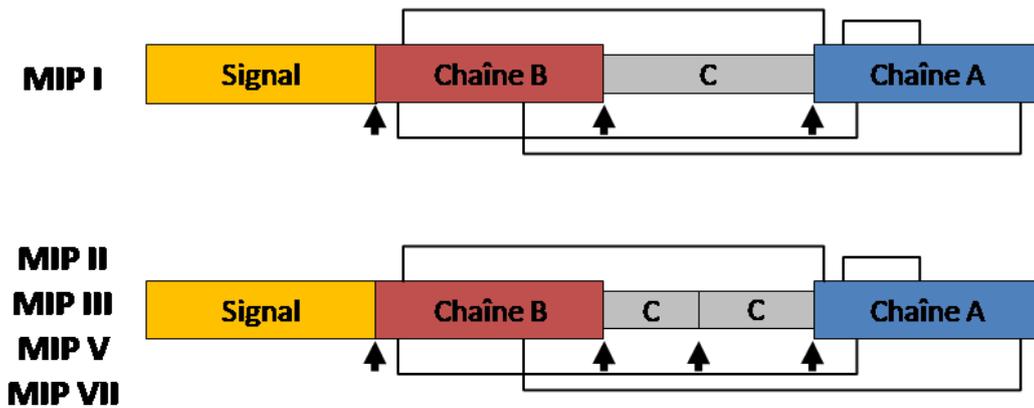


Figure II-1 : Représentation schématique de la structure des pré-peptides MIP chez *L. stagnalis*. Le peptide signal, les chaînes B et A ainsi que les ponts disulfures représentés par un trait noir sont retrouvés dans tous les MIPs. Les peptides C sont représentés en gris, un seul est identifié pour MIP I alors que deux sont caractérisés pour MIP II, MIP III, MIP V et MIP VII. Les flèches représentent les sites de clivage possibles. D'après Smit *et al.*, 1988, 1991, 1992, 1993, 1996, 1998.

Chez un autre gastéropode, *Aplysia californica*, un IRP a également été caractérisé à partir d'une banque d'ADNc puis en spectrométrie de masse (Floyd et Li, 1999). A l'instar des MIPs, l'IRP de *A. californica* est également présent dans les ganglions nerveux. La séquence en acides aminés possède tous les domaines caractéristiques des insulines et, tout comme les MIPs, huit cystéines ont été identifiées ainsi qu'un peptide C clivé en deux fragments. De façon intéressante, la prédiction de la séquence peptidique a montré la présence d'un domaine D à l'extrémité C-terminale, ce qui a ensuite été confirmé par séquençage (Figure II-2) (Floyd et Li, 1999). Ce domaine D facultatif n'avait pas encore été identifié dans une séquence IRP chez les mollusques. Chez *A. californica*, la ponte est régulée par l'hormone ELH (Egg-Laying Hormone) qui est produite dans les BCN (Bag Cells Neurons). L'utilisation d'un anticorps hétérologue dirigé contre le récepteur de l'insuline de mammifères, a permis d'observer un marquage positif important dans les BCN. D'après cela, Jonas et ses collaborateurs (1996) proposent que l'IRP d'aplysie pourrait influencer l'ELH par son action sur les BCN et par conséquent, intervenir dans le contrôle de la ponte des animaux. Ces conclusions sont cependant préliminaires. Récemment, un autre IRP présentant la même structure que l'insuline d'Aplysie, avec un domaine D, a également été identifié chez le céphalopode *Sepia officinalis* par spectrométrie de masse et séquençage (Zatylny-Gaudin *et al.*, 2016).

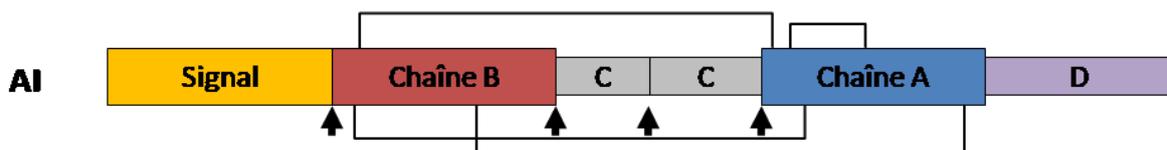


Figure II-2 : Représentation schématique du pré-propeptide insuline chez *A. californica*. Le peptide signal, les chaînes B et A ainsi que les ponts disulfures représentés par un trait noir sont identifiés dans l'insuline d'*A. californica* (AI). Les peptides C sont représentés en gris et le domaine D en violet. Les flèches représentent les sites de clivage possibles. D'après Floyd *et al.*, 1990.

Chez ces trois espèces, tous les IRPs identifiés appartiennent au type β puisqu'ils possèdent quatre ponts disulfures, trois entre les chaînes B et A et un interne à la chaîne A. Cependant, chez *Lottia gigantea*, la disponibilité récente du génome a permis l'identification de plusieurs types d'IRPs (Veenstra, 2010). Ainsi, l'analyse des séquences génomiques a conduit à l'identification de quatre séquences qui codent pour quatre IRPs (Figure II-3). Parmi eux, deux ont une structure identique à celle des MIPs identifiés chez *L. stagnalis*, ils sont nommés Ins1 et Ins2. Le troisième, Ins3, ne possède que six cystéines conservées donc seulement trois ponts disulfures, ce qui le classe dans le type γ . C'est la première fois chez

les mollusques qu'un IRP de ce type a pu être identifié, permettant ainsi de décrire la coexistence des types β et γ chez ce mollusque. Bien qu'appartenant au type β , le quatrième IRP, Ins4, diffère de Ins1 et Ins2 par la position de ses deux cystéines supplémentaires et sa forte homologie avec DILP7 (*Drosophila* Insulin-Like Peptide 7). Ces 2 derniers IRPs ont donc été décrits pour la première fois chez un mollusque, en lien avec la disponibilité de son génome et questionnent quant à l'existence de multiples formes chez les mollusques et à la phylogénie de ces molécules.

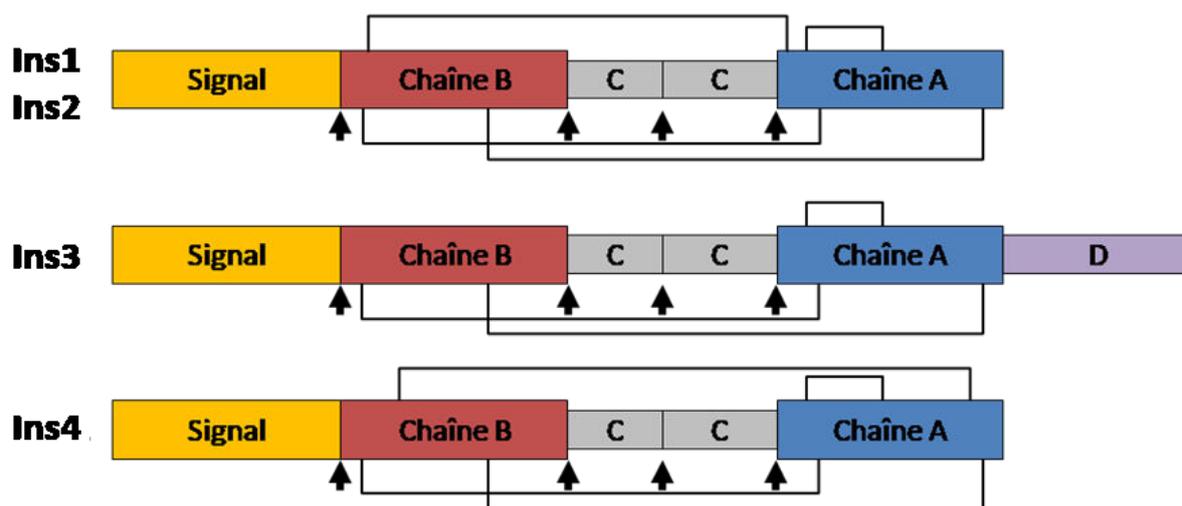


Figure II-3 : Représentation schématique de la structure des pré-propeptides chez *L. gigantea*. Le peptide signal, les chaînes B et A ainsi que les ponts disulfures représentés par un trait noir sont notifiés pour chaque IRP. Les peptides C sont représentés en gris et le domaine D en violet. Les flèches représentent les sites de clivage possibles. D'après Veenstra, 2010.

Chez l'huître creuse *C. gigas*, les travaux d'Hamano *et al.* (2005) ont permis d'identifier trois transcrits IRPs dans les ganglions viscéraux. La localisation de ces transcrits par hybridation *in situ* montre qu'ils sont exprimés dans certaines cellules neuronales. Ces transcrits sont issus du même gène et codent pour un même pré-propeptide IRP de type β contenant huit cystéines permettant la formation de quatre ponts disulfures comme pour les MIPs de *L. stagnalis*. La disponibilité récente du génome et de transcriptomes chez *C. gigas* (Zhang *et al.*, 2012; Zhao *et al.*, 2012) est un atout majeur pour rechercher d'autres IRPs, qui pourraient appartenir à différents types, et les comparer à ceux identifiés chez *L. gigantea*.

1.2. Identification des effecteurs du signal insuline

Si la conservation des IRPs est relative, les séquences des effecteurs des voies de l'insuline apparaissent mieux conservées entre les espèces et il est intéressant de les étudier pour préciser les aspects fonctionnels. Chez les vertébrés, les IRPs se lient à plusieurs types de récepteurs. A l'inverse, chez les invertébrés, un seul récepteur a pu être identifié, il semble pouvoir lier les différents ligands IRPs.

1.2.1. Le récepteur des IRPs chez les mollusques

La recherche de récepteurs d'IRPs chez les mollusques a été réalisée en se basant sur la conservation des domaines tyrosine kinase des récepteurs aux IRPs d'*H. sapiens* et *D. melanogaster*. Ainsi, un fragment codant pour la région tyrosine kinase d'un récepteur aux MIPs a pu être isolé et amplifié à partir du système nerveux central de *L. stagnalis* (Roovers *et al.*, 1995). Sa séquence protéique a ensuite été prédite et son analyse a montré que le domaine tyrosine kinase possède une forte similarité, comprise entre 73% et 75%, avec les récepteurs aux insulines des mammifères. Bien que la séquence identifiée chez *L. stagnalis* soit complète, l'analyse de MIPR n'a pas permis de déterminer si ce récepteur est plus proche des récepteurs aux insulines ou des récepteurs aux IGF-1 issus des vertébrés. A ce jour, aucun autre MIPR n'a été identifié chez *L. stagnalis* suggérant que les différents MIPs pourraient se lier à ce récepteur unique, peut-être avec différentes affinités. Chez *A. californica*, un ADNc codant pour un récepteur à domaine tyrosine kinase a été isolé à partir du système nerveux, de la même façon que chez *L. stagnalis*. La séquence protéique du récepteur identifié contient tous les domaines caractéristiques des récepteurs aux insulines. L'expression du gène a été montrée dans un groupe de cellules neuronales impliquées dans le contrôle des comportements reproducteurs (Jonas *et al.*, 1996). La protéine a bien été retrouvée dans ce même groupe de cellules par une approche immunocytoologique, basée sur l'utilisation d'un anticorps hétérologue dirigé contre la sous-unité α du récepteur d'insuline humain. Toujours sur la base de la forte conservation du domaine tyrosine kinase, un récepteur aux IRPs a été identifié dans la lignée de cellules embryonnaires de *Biomphalaria glabrata* (Lardans *et al.*, 2001) et chez l'huître perlière *Pinctada fucata* (Shi *et al.*, 2013).

Toutes les séquences de récepteurs aux insulines identifiées chez ces mollusques contiennent les domaines spécifiques de ces récepteurs, incluant une région riche en cystéines, un site de clivage (2 à 4 acides aminés basiques entre les chaînes α et β), une région transmembranaire et un domaine tyrosine kinase (Figure II-4). Les ponts disulfures dus aux multiples cystéines regroupées dans la partie N-terminale du récepteur, sont déterminants pour la structure tertiaire de ce dernier et notamment pour la liaison des IRPs. De plus, ce domaine riche en cystéines de la chaîne α ainsi que les deux régions l'encadrant peuvent renseigner sur l'identité du récepteur (Andersen *et al.*, 1992; Roovers *et al.*, 1995; Schumacher *et al.*, 1993).



Figure II-4 : Représentation schématique de la structure des récepteurs aux insulines identifiés chez les mollusques. La partie codant pour la sous-unité α contient un peptide signal (SP) ainsi que la région riche en cystéines. La région codant pour la sous-unité β contient le domaine transmembranaire et le domaine tyrosine kinase. Le site de clivage est indiqué par une flèche.

En 2003, une séquence incomplète d'un récepteur insuline unique a également été identifié chez *C. gigas* et nommé CIR (*C. gigas* insulin receptor-related) (Gricourt *et al.*, 2003). Son domaine tyrosine kinase apparaît bien conservé par rapport aux différentes espèces où le récepteur a été identifié, présentant des pourcentages d'identité compris entre 42% (*vs* DAF-2, *C. elegans*) et 66% (*vs* MIPR, *L. stagnalis*) (Gricourt *et al.*, 2003). Cette séquence de CIR incomplète puisque le domaine riche en cystéines des sous-unités α n'est pas identifié, ne renseigne pas sur le type de récepteur présent chez *C. gigas*.

1.2.2. Les voies de signalisation

Des effecteurs intracellulaires des voies des IRPs ont déjà été identifiés chez des mollusques. Ainsi, chez le gastéropode *A. californica*, une séquence Ras a pu être identifiée et localisée dans les ganglions cérébraux ainsi que dans l'ovotestis, plus particulièrement au niveau des ovocytes (Swanson *et al.*, 1986). Chez les bivalves *Mytilus edulis* et *Mytilus trossulus*, des séquences codantes pour l'effecteur Ras ont également été caractérisées à partir d'échantillons de glande digestive et d'hémolymphe respectivement. Ces séquences Ras présentent une forte conservation des régions fonctionnelles par comparaison avec la protéine Ras de plusieurs vertébrés (Ciocan et Rotchell, 2005; Ciocan *et al.*, 2006).

Plus récemment, Jouaux *et al.* (2012) ont identifiés certains effecteurs des voies des IRPs chez *C. gigas*, à savoir CgRas potentiellement impliqué dans la voie MAPK mais également CgPTEN et Cgp70S6K, acteurs de la voie PI3K. L'effecteur CgRas contient le domaine effecteur et le domaine G2 spécifiques des protéines de la famille Ras. CgPTEN, quant à lui, contient quatre cystéines conservées dont deux sont impliquées dans son activité catalytique. Le domaine catalytique a également été identifié pour Cgp70S6K (Jouaux *et al.*, 2012). Les expressions de ces effecteurs ont été mesurées par qPCR. CgRas est principalement exprimé dans le bord de manteau et les ganglions viscéraux et Cgp70S6K est majoritairement exprimé dans les branchies et, dans une moindre mesure, au niveau du bord de manteau, des palpes labiaux et des ganglions viscéraux. Enfin, le transcrite de CgPTEN est essentiellement retrouvé dans les branchies même si son expression est détectée dans les autres tissus (Jouaux *et al.*, 2012). La disponibilité des bases de données génomiques et transcriptomiques permet d'envisager la recherche de l'ensemble des effecteurs de ces voies de signalisation. Les informations concernant leur expression seront un premier élément pour identifier les cibles du signal des IRPs et discuter des effets potentiels dans le contrôle de la reproduction.

1.3. Données fonctionnelles

Certains rôles fonctionnels du système insuline en lien avec la reproduction ont pu être proposés chez *C. gigas*. L'étude de l'expression du ligand oIRP (oyster Insulin Related Peptide) dans les ganglions viscéraux au cours du cycle de reproduction suggère qu'il pourrait être impliqué dans le contrôle central de la reproduction et du statut énergétique (Hamano *et al.*, 2005).

Le suivi de l'expression du récepteur CIR a révélé que le transcrite est présent dans la gonade, en période de gamétogenèse. CIR apparaît fortement exprimé lors des phases de proliférations goniales et de maturation des ovocytes, suggérant l'implication des IRPs dans ces processus chez *C. gigas* (Gricourt *et al.*, 2006).

Les deux effecteurs du signal présentant la plus forte expression, CIR et CgRas, ont été suivis dans les deux compartiments constituant l'aire gonadique (Jouaux *et al.*, 2012). La technique de microdissection laser a permis de séparer les tubules gonadiques et le tissu de réserve. L'expression des deux gènes a été suivie en qPCR au cours du cycle annuel dans ces deux compartiments. Ces expressions apparaissent plus élevées dans les tubules gonadiques

que dans le tissu de réserve. Chez les mâles, le stade 1 de gamétogenèse, correspondant aux mitoses goniales, est celui où l'expression de CIR est la plus élevée et l'expression de CgRas augmente entre les stades quiescent et de proliférations goniales. Chez les femelles, c'est au stade de maturité sexuelle que CIR et CgRas sont les plus exprimés (Jouaux *et al.*, 2012). Ces données suggèrent l'implication de ce système, notamment via la voie Ras, dans les proliférations des spermatogonies ainsi que dans le processus de vitellogenèse des ovocytes (Jouaux *et al.*, 2012).

1.4. Objectifs

Dans ce contexte, les objectifs de ce travail porteront sur les différents niveaux du signal insuline et s'appuieront essentiellement sur des analyses *in silico* des banques de données génomiques et transcriptomiques et sur des données d'expression.

La disponibilité de l'ensemble des ligands permettra d'approfondir l'histoire évolutive des IRPs et de renseigner sur leur forme ancestrale. D'autre part, les profils d'expression renseigneront sur les IRPs les plus probablement impliquées dans la reproduction et même plus précisément dans le contrôle des étapes précoces de gamétogenèse.

Concernant le récepteur, l'analyse de la séquence du récepteur CIR et notamment de la partie N-terminale manquante sera complétée. Son expression sera recherchée non seulement dans la gonade mais aussi dans les ganglions nerveux qui sont impliqués dans le contrôle de la reproduction chez d'autres mollusques.

Un troisième objectif sera d'identifier les acteurs de toutes les voies potentiellement impliquées dans les multiples fonctions physiologiques des IRPs. Le contrôle du fonctionnement des GSC décrit dans la littérature fait plutôt intervenir la voie PI3K comme chez *D. melanogaster* (Hsu et Drummond-Barbosa, 2009, 2011) ou encore *C. elegans* (Hubbard, 2011).

2. Partie A – Caractérisation des IRPs et expressions tissulaires

Cette partie a fait l'objet d'une publication acceptée et sous presse ainsi que d'un article de type « Data in brief » accepté et sous presse dans la revue *General and Comparative Endocrinology*. Ce travail de thèse comprend l'identification *in silico* et la caractérisation des IRPs identifiés chez *C. gigas* ainsi que leurs expressions mesurées dans différents tissus adultes. De plus, afin d'identifier les IRPs impliqués dans la reproduction, les mesures d'expressions ont été réalisées au cours d'un cycle complet de gamétogenèse dans les ganglions viscéraux et l'aire gonadique.

2.1. Matériels et Méthodes

La partie « matériels et méthodes » est détaillée dans la publication.

2.2. Principaux résultats

2.2.1. Identification des IRPs

La recherche des IRPs a été faite par homologie de séquences dans les bases de données génomiques et transcriptomiques disponibles (Rivière *et al.*, 2015; Zhang *et al.*, 2012; Zhao *et al.*, 2012) ainsi que dans le protéome prédit. Au total, six séquences peptidiques présentant la structure typique des IRPs ont ainsi été identifiées. Ces six IRPs possèdent une chaîne B, une chaîne A un peptide C divisé en plusieurs fragments (Figure II-5).

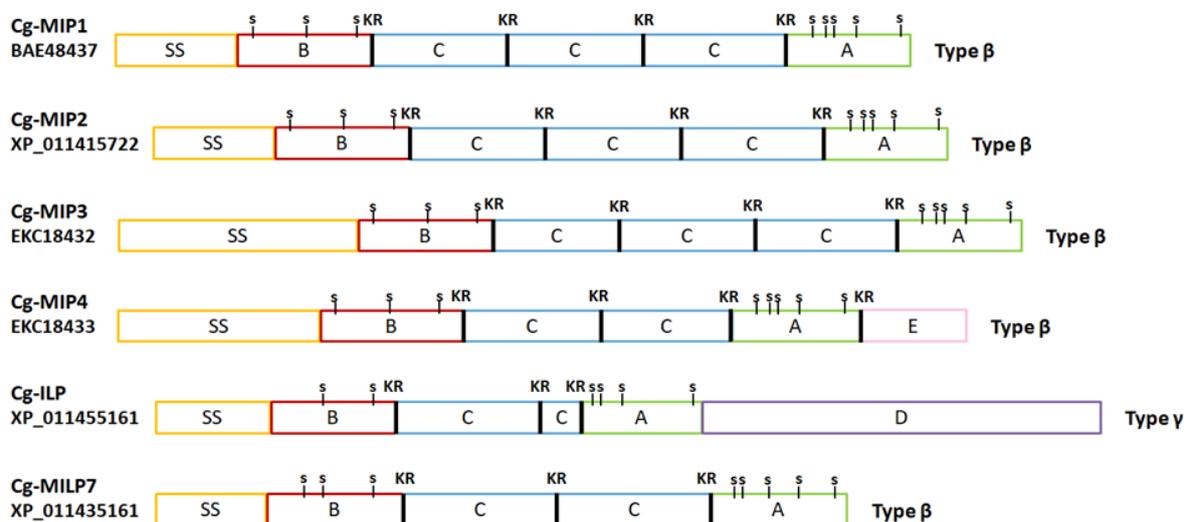


Figure II-5 : Représentations schématiques des IRPs identifiés chez *C. gigas*. Pour chaque IRP, le nom, le numéro d'accèsion NCBI et le type sont notifiés à côté du schéma des domaines de la séquence. SS : Séquence Signal ; B : Chaîne B ; C : Peptide C ; A : Chaîne A ; E : Domaine E ; D : Domaine D ; s : cystéines impliquées dans les ponts disulfures ; KR : sites de clivage dibasiques.

Quatre de ces IRPs, nommés Cg-MIP1, Cg-MIP2, Cg-MIP3 et Cg-MIP4, possèdent huit cystéines et peuvent donc former quatre ponts disulfures comme les MIPs décrits chez *L. stagnalis*. Ces caractéristiques structurales sont reflétées par le motif signature des MIPs, C(X)CC(3X)C(8X)C, contenu dans leurs chaînes A. Cg-MIP4 présente la particularité de posséder un domaine E à son extrémité C-terminale, domaine qui sera clivé dans la forme mature. Ces IRPs sont classiquement identifiés chez les espèces mollusques.

Le cinquième IRP identifié possède également 8 cystéines mais son motif, CC4XC8XC, est différent des séquences Cg-MIPs. De plus, cette séquence présente une forte homologie avec DILP7. Puisqu'il associe la présence de quatre ponts disulfure et une forte homologie avec DILP7, cet IRP a été nommé Cg-MILP7. Quelques IRPs de ce type ont été retrouvés, notamment chez *L. gigantea* (Veenstra, 2010), mais également en recherchant dans les banques d'expression de *L. stagnalis* et *A. californica*. Ces IRPs sont particulièrement intéressants car ils possèdent le même motif que DILP7 mais pas le même nombre de cystéines.

Enfin, le sixième IRP a été nommé Cg-ILP en raison de la présence de six cystéines conservées seulement et du motif signature des IRPs vertébrés, CC3XC8XC. De plus, Cg-ILP possède un domaine D à son extrémité C-terminale. A ce jour, un seul autre IRP de ce type a été retrouvé chez les mollusques et plus précisément chez *L. gigantea*.

La caractérisation de ces IRPs a permis de confirmer la coexistence des types β et γ chez un deuxième mollusque. En effet, les Cg-MIP et Cg-MILP7 appartiennent au type β puisqu'ils possèdent quatre ponts disulfures alors que Cg-ILP appartient au type γ avec ses trois ponts disulfures.

Une phylogénie a été réalisée avec les IRPs identifiés chez différentes espèces de gastéropodes, bivalves et céphalopodes, les trois principaux groupes de mollusques étudiés. Les séquences de l'insuline humaine et de DILP7 de drosophile ont également été incluses dans l'arbre phylogénétique. Cette analyse a mis en évidence que la ségrégation phylogénétique se faisait principalement selon les motifs de la chaîne A. De façon intéressante, DILP7 est également regroupé avec les IRPs de mollusques partageant son motif (CC4XC8XC). De plus, Cg-ILP et l'Ins3 de *L. gigantea* sont groupés avec leur orthologue humain.

2.2.2. Suivi des expressions

L'analyse des bases de données *in silico* a permis d'identifier trois gènes et trois transcrits au total. Toutes les séquences génomiques présentent la structure classique des gènes de l'insuline à savoir, au minimum deux exons, l'un codant pour la chaîne B et un autre codant pour la chaîne A. Tous les Cg-MIPs partagent la même séquence génomique. Cependant, Cg-MIP1, Cg-MIP2 et Cg-MIP3 dérivent de la même séquence transcriptomique, nommée Cg-mip123, alors que Cg-MIP4 provient d'un autre transcrit nommé Cg-mip4. Cg-ILP a été associé à une séquence génomique et une séquence transcriptomique, nommée Cg-ilp. Une séquence génomique a pu être associée à Cg-MILP7 mais aucun transcrit correspondant à cet IRP n'a pu être identifié dans les transcriptomes.

L'expression de Cg-mip123, Cg-mip4 et Cg-ilp a été mesurée en PCR quantitative dans plusieurs tissus adultes ainsi que dans l'aire gonadique à tous les stades de gamétogenèses mâle et femelle (Tableau II-1). Les trois transcrits sont exprimés au niveau des ganglions viscéraux suggérant une régulation centrale faite par les IRPs chez *C. gigas*, notamment Cg-mip123 dont l'expression est la plus élevée. La localisation en hybridation *in situ* a révélé que les cellules neurosécrétrices exprimant Cg-mip123 sont localisées dans un lobe des ganglions viscéraux. L'expression de ce transcrit augmente dans l'aire gonadique en fonction de l'avancée en spermatogenèse atteignant son maximum au stade de maturité sexuelle. Le transcrit est localisé dans des cellules au bord des tubules gonadiques, zone où se trouvent les GSC et des cellules somatiques accompagnatrices. Le transcrit Cg-mip4 est fortement exprimé lors du développement et n'apparaît que peu exprimé chez l'adulte à l'exception du stade 3 de gamétogenèse chez les femelles. Le seul transcrit exprimé dans la glande digestive est Cg-ilp et ce dernier apparaît également exprimé dans l'aire gonadique et son expression augmente en corrélation avec l'avancée en gamétogenèses mâle et femelle.

Tableau II-1 : Comparaison relative des mesures d'expression des transcrits insuline au cours du développement, dans les ganglions viscéraux, la glande digestive et aux stades de gamétogenèses mâle et femelle. - - : Très faiblement exprimé ; - : Faiblement exprimé ; + : Exprimé ; ++ : Fortement exprimé ; +++ : Très fortement exprimé ; F : Chez les femelles ; M : Chez les mâles.

	Développement	Ganglions viscéraux	Glande digestive	Aire gonadique			
				Stade 0	Stade 1	Stade 2	Stade 3
Cg-mip123	--	+++	-	+	F : + M : +	F : +- M : ++	F : - M : +++
Cg-mip4	+++	++	--	--	F : - M : -	F : - M : -	F : +- M : -
Cg-ilp	-	+	+++	+	F : + M : +	F : ++ M : ++	F : +++ M : +++

2.2.3. Expressions tissulaires de Cg-milp7

L'analyse des bases de données pour le peptide Cg-MIP7 n'a pas permis de trouver de séquence ARN mais une séquence génomique codant pour ce peptide a été identifiée, permettant de construire des amorces (Sens : 5'-TCGACGTGTTGATCGAGTTC-3' ; Anti-sens : 5'-GGAGGAATTTGCCGAATTTT-3') dans les régions correspondant aux exons. Après amplification de la banque PAL17, un transcrite nommé Cg-milp7 a pu être obtenu. L'expression de cette séquence a été mesurée en PCR quantitative dans plusieurs tissus d'huîtres adultes ainsi que dans les ganglions viscéraux et l'aire gonadique à tous les stades de gamétogenèse de la même façon que pour les autres transcrits.

Cg-milp7 est également exprimé dans les ganglions viscéraux mais son expression ne montre pas de variation significative selon les stades de gamétogenèse (Figure II-6A). La mesure de son expression dans l'aire gonadique a révélé que le transcrite Cg-milp7 est en moyenne 10 fois moins exprimé dans ce tissu que dans les ganglions viscéraux. Le profil d'expression dans l'aire gonadique est très proche de celui mesuré pour Cg-mip4 (Figure II-6B). Cg-milp7 n'est que faiblement exprimé au stade quiescent ainsi que dans les stades de gamétogenèse mâle mais son expression augmente en corrélation avec l'avancée de la gamétogenèse femelle. Dans les tissus adultes, Cg-milp7 apparaît très faiblement exprimé (Figure II-6C).

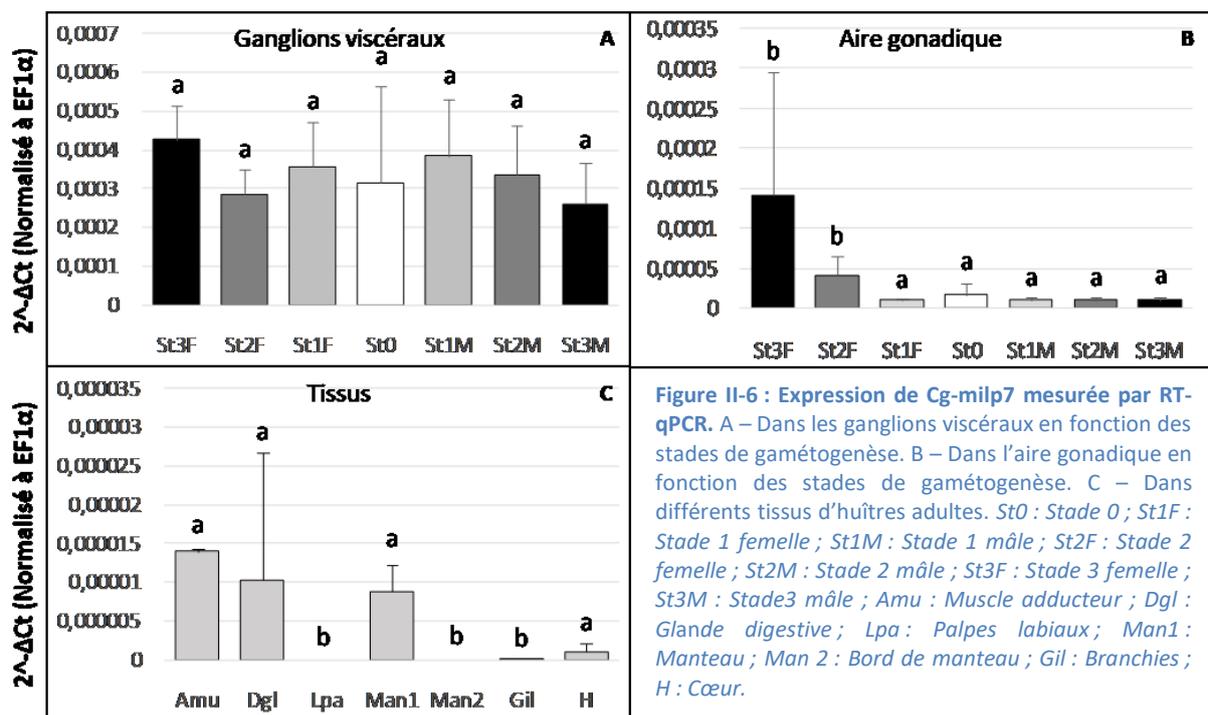


Figure II-6 : Expression de Cg-milp7 mesurée par RT-qPCR. A – Dans les ganglions viscéraux en fonction des stades de gamétogenèse. B – Dans l'aire gonadique en fonction des stades de gamétogenèse. C – Dans différents tissus d'huîtres adultes. St0 : Stade 0 ; St1F : Stade 1 femelle ; St1M : Stade 1 mâle ; St2F : Stade 2 femelle ; St2M : Stade 2 mâle ; St3F : Stade 3 femelle ; St3M : Stade 3 mâle ; Amu : Muscle adducteur ; Dgl : Glande digestive ; Lpa : Palpes labiaux ; Man1 : Manteau ; Man 2 : Bord de manteau ; Gil : Branchies ; H : Cœur.

2.3. Discussion

La recherche de ligands des voies des IRPs chez l'huître creuse *C. gigas* a permis de mettre en évidence la présence de plusieurs IRPs dont les structures peptidiques et génomiques sont caractéristiques des membres de la famille des insulines chez différentes espèces (Steiner *et al.*, 1985). Les IRPs de *C. gigas* regroupent trois types de motifs spécifiques de la chaîne A. A ce jour, un seul autre mollusque, *Lottia gigantea*, possède également des IRPs avec les trois motifs représentés (Veenstra, 2010). Chez les autres mollusques, les IRPs identifiés sont majoritairement de type MIP. Le fait que les IRPs de types ILP et MILP7 n'aient été retrouvés que chez *L. gigantea* et *C. gigas* doit être rapproché de la stratégie d'identification par criblage exhaustif des bases de données. De plus, l'identification d'IRPs a souvent été réalisée en ciblant les ganglions nerveux comme chez *L. stagnalis* par exemple, ce qui favorise l'identification de peptides de type MIP (Smit *et al.*, 1988, 1991, 1992, 1993, 1996, 1998) qui, chez *C. gigas*, apparaissent majoritairement exprimés dans ce tissu. Chez *S. officinalis*, l'identification d'IRPs a été faite par recherche de motif (Zatylny-Gaudin *et al.*, 2016). Les MIPs étant jusqu'alors considérés comme la forme spécifique des mollusques, seul le motif typique de la chaîne A des MIPs a été recherché, conduisant à l'identification exclusive de ce type de séquences.

L'arbre phylogénétique regroupant les IRPs de plusieurs mollusques ainsi que l'insuline humaine et le DILP7 de drosophile a permis de mettre en évidence que la ségrégation de ces IRPs se fait principalement en fonction du motif de la chaîne A. Les trois groupes d'IRPs identifiés sont présents chez plusieurs mollusques, suggérant une ségrégation phylogénétique antérieure à ces espèces. De plus, les ILPs étant retrouvés chez les vertébrés, les Ecdysozoaires et certains Lophotrochozoaires comme *L. gigantea* (Veenstra, 2010), *C. gigas* et *Platynereis dumerilii* (Conzelmann *et al.*, 2013), ils pourraient constituer la forme ancestrale des ligands des voies de l'insuline. L'arbre phylogénétique montre que les MILP7 des mollusques sont des orthologues de DILP7, ce qui suggère que le motif CC4XC8XC serait apparu avant la séparation des Ecdysozoaires et des Lophotrochozoaires.

L'étude fonctionnelle a été initiée en précisant les profils d'expression des IRPs qui reflètent leurs rôles potentiels. Tous les transcrits sont exprimés dans le système nerveux mais c'est Cg-mip123 qui présente la plus forte expression dans les ganglions viscéraux en accord avec les résultats publiés par Hamano *et al.* (2005). Ces MIPs doivent donc être

impliqués dans le contrôle central de multiples processus physiologiques. L'étude du transcrit Cg-mip123 dans l'aire gonadique montre qu'il est majoritairement exprimé au stade de maturité sexuelle mâle. Il est également détecté dans l'aire gonadique de certains individus mâles au stade de maturité sexuelle et ce profil suggère que Cg-mip123 pourrait avoir un rôle androgène dans le changement de sexe qui affecte certains mâles avant l'année de reproduction suivante. En effet, le rôle androgène des IRPs a été décrit chez certaines espèces de crustacés, où l'IRP nommé IAG (Insulin-like Androgenic Gland peptide) permet le maintien des caractéristiques mâles (Chetler *et al.*, 2015, 2017).

Cg-MIP4 et Cg-MILP7 présentent des profils d'expression proches, ils sont tous les deux très peu exprimés dans les tissus adultes mais, bien que faible, leur expression augmente lors du stade de maturité sexuelle femelle. Ce profil suggère qu'ils pourraient être impliqués dans la maturation ovocytaire, rôle connu des IRPs chez le poisson zèbre (Nelson et Van Der Kraak, 2010) ou chez la drosophile (Ikeya *et al.*, 2002; Wu et Brown, 2006).

Cg-ilp est le seul transcrit exprimé à la fois dans la glande digestive et dans l'aire gonadique. Or, chez *Crassostrea gigas*, l'alimentation joue un rôle clé dans le contrôle de la gamétogenèse, et les conditionnements de géniteurs en écloserie s'appuient justement sur des périodes de forte alimentation (Muranaka et Lannan, 1984). Les travaux précédents ont montré que les réserves préalablement établies conditionnent l'avancée en gamétogenèse (Heude Berthelin *et al.*, 2001) et l'effort de reproduction chez cette espèce (Jouaux *et al.*, 2013). Le profil d'expression très spécifique de Cg-ilp en fait donc un bon candidat comme régulateur de la balance entre alimentation et reproduction chez l'huître creuse.

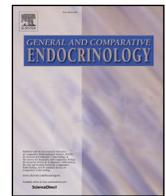
Ainsi, ces données d'expression ont permis d'identifier les ligands et de préciser leurs implications potentielles dans le contrôle de la gamétogenèse.



ELSEVIER

Contents lists available at ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen

Molecular evolution and functional characterisation of insulin related peptides in molluscs: Contributions of *Crassostrea gigas* genomic and transcriptomic-wide screening

Maëva Cherif-Feildel^{a,b}, Clothilde Heude Berthelin^{a,b}, Beatrice Adeline^{a,b}, Guillaume Rivière^{a,b}, Pascal Favrel^{a,b}, Kristell Kellner^{a,b,*}

^a Normandy University, Caen, France

^b University of Caen Normandie, Unity Biology of Organisms and Aquatic Ecosystems (BOREA), MNHN, Sorbonne University, UCN, CNRS, IRD, Esplanade de la Paix, 14032 Caen, France

ARTICLE INFO

Keywords:

Insulin related peptides
Crassostrea gigas
 Molluscs
 Phylogeny
 Gametogenesis

ABSTRACT

Insulin Related Peptides (IRPs) belong to the insulin superfamily and possess a typical structure with two chains, B and A, linked by disulphide bonds. As the sequence conservation is usually low between members, IRPs are classified according to the number and position of their disulphide bonds. In molluscan species, the first IRPs identified, named Molluscan Insulin-related Peptides (MIPs), exhibit four disulphide bonds. The genomic and transcriptomic data screening in the Pacific oyster *Crassostrea gigas* (Mollusc, Bivalvia) allowed us to identify six IRP sequences belonging to three structural groups. Cg-MIP1 to 4 have the typical structure of MIPs with four disulphide bonds. Cg-ILP has three disulphide bonds like vertebrate Insulin-Like Peptides (ILPs). The last one, Cg-MILP7 has a significant homology with *Drosophila* ILP7 (DILP7) associated with two additional cysteines allowing the formation of a fourth disulphide bond. The phylogenetic analysis points out that ILPs may be the most ancestral form. Moreover, it appears that ILP7 orthologs are probably anterior to lophotrochozoa and ecdysozoa segregation. In order to investigate the diversity of physiological functions of the oyster IRPs, we combine *in silico* expression data, qPCR measurements and *in situ* hybridization. The Cg-*ilp* transcript, mainly detected in the digestive gland and in the gonadal area, is potentially involved in the control of digestion and gametogenesis. The expression of Cg-*mip4* is mainly associated with the larval development. The Cg-*mip* transcript shared by the Cg-MIP1, 2 and 3, is mainly expressed in visceral ganglia but its expression was also observed in the gonads of mature males. This pattern suggested the key roles of IRPs in the control of sexual reproduction in molluscan species.

1. Introduction

Since their discovering in mammals, the molecules belonging to the insulin family and their potential physiological roles have given rise to a huge interest. The abundant literature dealing with this subject occurs for two main reasons. Firstly, insulin family members are involved in the control of a large variety of physiological processes including reproduction, development, growth, carbohydrate and lipid metabolism (reviewed in Agrogiannis et al., 2014; Dimitriadis et al., 2011; Saliel and Kahn, 2001; Sliwowska et al., 2014). The involvement of insulin in the control of life span has also been demonstrated (Avogaro et al., 2010). Secondly, insulin belongs to a superfamily including a large number of members. Insulin and Insulin-like Growth Factors (IGFs)

were first described in mammals and other members were later identified as relaxins and a variety of related peptides (Bathgate et al., 2013; Shabanpoor et al., 2009; Wilkinson et al., 2005). Insulin related peptides (IRPs) were also found in almost all animal phyla, so this family of molecules is of particular interest for phylogenetic analysis and study of the evolutionary history. From its large phylogenetic repartition, Jekely (2013) concluded that IRP may be ancestral to all metazoan species.

Invertebrate IRPs sequences have firstly been identified based on biochemical purification or cDNA identification thus producing a large number of IRPs of various taxonomic clades. Another approach relies on genomic and transcriptomic-wide screening. This enabled to complete the panel of IRPs in some species or to identify sequences in new species. These different approaches helped identifying and

* Corresponding author.

E-mail address: kristell.kellner@unicaen.fr (K. Kellner).

<https://doi.org/10.1016/j.ygcen.2018.10.019>

Received 9 July 2018; Received in revised form 25 October 2018; Accepted 26 October 2018

0016-6480/© 2018 Elsevier Inc. All rights reserved.

Table 1

Schematic structure of the main putative mature IRPs sequences, according to the classification of Matsunaga (2017). All IRPs groups have the 2 canonical disulfide bridges connecting the A and B chains. α and β types have an additional inter A and B chains disulfide bridge, only the β type include a fourth intra A chain disulfide bridge. A and B chains may be linked or not depending on the presence and cleavage of a C peptide between them. A supplementary D domain is also present in several members. AG: androgenic gland; IAG: Insulin-like Androgenic gland peptide, ILP: insulin like peptide, IRP: insulin related peptides, RLP: relaxin like peptide. R: Review article, U: Unpublished NCBI data.

			α type		β type		γ type		Bibliography
			2 canonic disulfide bridges between A and B chains						
			No			1 disulfide intra-A chain bridge			
			1 additional disulfide bridge between A and B chains			No			
Various mature forms									
Species and phylogeny									
Chordata	Vertebrata	<i>H. sapiens</i>					Insulin, IGFs, INSL, Relaxin	Shabanpoor et al. (2009) R	
	Tunicata	<i>C. intestinalis</i>					Ci-INS-L1, Ci-INS-L2, Ci-INS-L3	Olinski et al. (2006)	
Ecdysozoa	Nematoda	<i>C. elegans</i>	INS-20 to 31, INS 33 to 36, INS 39	INS 1 to 10, Daf28			INS 11 to 19, INS 32, INS 37, INS 38	Pierce et al. (2001), Li et al. (2003), Matsunaga et al. (2017) R	
	Hexapods	<i>B. mori</i>					Bombyxins family A to G	Maruyama et al. (1988) Kondo et al. (1996) Yoshida et al. (1997, 1998), Bayazit (2009) R	
		<i>S. gregaria</i>					Scg-IRP	Badisco et al. (2013)	
		<i>L. migratoria</i>					ILP	Lagueux et al. (1990)	
		<i>A. aegyptii</i>					AaILP 1 to 8	Riehle et al. (2006) R	
		<i>D. melanogaster</i>					DILP 1 to 8	Brogiolo et al. (2001), Grönke et al. (2010) R	
	Crustacea	<i>S. verreauxi</i>					Sv-IAG, Sv-ILP1, Sv-ILP2	Chandler et al. (2017)	
		<i>A. vulgare</i>					Arv AGH (with additional intra B disulfide bridge)	Ventura et al. (2011)	
		<i>M. rosenbergii</i>			Mr IAG		Mr ILP, Mr RLP	Phoungpetchara et al. (2011), Ventura et al. (2009)	
Lophotrochozoa	Mollusca	<i>L. stagnalis</i>		MIP 1, 2, 3, 4, 5, 7				Smit et al. (1988, 1991, 1992, 1993, 1996, 1998), Li et al. (1992)	
		<i>A. californica</i>		AI				Floyd et al., 1999	
		<i>C. geographus</i>		Con-Ins G2		Con-Ins G1		Safavi-Hemami et al. (2015)	
		<i>S. officinalis</i>		So IRP				Zatylny-Gaudin et al. (2016)	
		<i>C. gigas</i>		olRP				Hamano et al. (2005)	
		<i>L. gigantea</i>		ins 1, ins 2, ins 4		ins 3		Veenstra (2010)	
	Annelida	<i>P. dumerilii</i>		IRP2, IRP3, IRP4, IRP5		IRP1		Conzelmann et al. (2013)	
	Platyhelminths	<i>S. mediterranea</i>					ilp-1	Miller and Newmark (2012)	
		<i>T. solium</i>					ILP-1, ILP-2 (with an additional intra A disulfide bridge)	Wang et al. (2013)	
		<i>S. mansoni</i>					ILP (with an additional intra A disulfide bridge)	Wang et al. (2013)	
Porifera	<i>G. cydonium</i>					sponge insulin	Robitzki et al. (1989)		
Cnidaria	<i>N. vectensis</i>					NV 199266; NV 207484	Ancill (2009)		
	<i>H. magnipapillata</i>					Hv-ILP1; Hv-ILP2; Hv-ILP3	Steele et al. (2009)- U		

(See above-mentioned references for further information.)

characterizing IRPs in a large number of species belonging to all clades of the phylogenetic tree. The most studied members can be found in Table 1.

A comparative analysis of IRP sequences revealed that the percentage of identity among species as well as between insulin related members within one species is relatively low. However, the IRPs display common features. First, they are composed of two chains named B and A which are linked by at least two canonical disulfide bonds (De

Meyts, 2004). Chains B and A originated in the post-translational proteolytic cleavage from a unique pre-propeptide that contained a signal peptide and some facultative additional domains named C peptide, D and E domains. The C peptide is located between the B and A sequences and may be cleaved or not in the mature form. The D and E domains are located at the C-terminal extremity, the latter one being post-translationally cleaved through dibasic residues. One of the most diversified series of IRPs was described in the nematode *Caenorhabditis elegans* and

led to the description of 40 insulins (INS) named INS-1 to INS-39 and DAF28 (Matsunaga, 2017; Pierce et al., 2001). *C. elegans* IRPs can be assigned into three groups named α , β and γ according to their structural properties (Duret et al., 1998; Matsunaga, 2017). The IRPs belonging to α and β types have an additional disulphide bond between the A and B chains. In the β type, a fourth disulphide bond occurs between two residues of the A chain (intra-A disulphide bond). In the γ type, only the two canonical disulphide bonds between A and B are present but the intra-A chain bond also exists (Matsunaga, 2017).

So far, the co-existence of the three types of IRPs in the same species has only been reported in *C. elegans*. Interestingly, the situation regarding IRP types is very different depending on the species considered (Table 1). In deuterostomes, IRPs are only assigned to the γ type, as well as in most of ecdysozoan species, with the exception of the crustacea *Macrobrachium rosenbergii* which has a β type IRP. However, the presence of β type IRPs appears more common in lophotrochozoan species. On the contrary, the γ IRPs were scarcely described in lophotrochozoan and only for species with a sequenced genome. Interestingly, this co-existence of both β and γ types might constitute an evolutionary signature of IRPs in protostomian species. Moreover, distribution of IRPs across the phylogenetic tree led to hypothesise that γ type members constitute the ancestral form.

Investigation of physiological roles of IRPs may improve our understanding of their evolutionary history. Indeed, the multiple biological functions of different members of the insulin family observed in vertebrate species are generally conserved in invertebrates. Indeed, protostomian IRPs assigned to different types present similar functions compared to deuterostome IRPs which are all grouped in the γ type. This functional conservation was clearly demonstrated for *Drosophila melanogaster* and *C. elegans*, since these species benefit from functional approaches mainly based on targeted mutations of insulin-like pathway genes (ligands, receptor or effectors of signal) or RNAi experiments. In both species, IRPs played key roles in the control of growth and development (Böhni et al., 1999; Brogiolo et al., 2001; Garelli et al., 2012; Grönke et al., 2010; Li et al., 2003; Shingleton et al., 2005), energy storage (Ashrafi et al., 2003; Broughton et al., 2005), stress resistance (Liu et al., 2016; Söderberg et al., 2011), response to diet restriction (Broughton et al., 2010; Hibshman et al., 2016), lifespan (Apfeld and Kenyon, 1998; Clancy et al., 2001; Grönke et al., 2010; Partridge et al., 2011) and fecundity (Hsu and Drummond-Barbosa, 2009; Luo et al., 2010; Michaelson et al., 2010; Narbonne et al., 2015). Unlike their vertebrate counterparts, the primary source of IRPs in invertebrates is the nervous system. These assertions rely mainly on immunocytochemical and *in situ* hybridization labelling of the IRPs-producing cells located in the nervous ganglia. However, local expressions in other tissues also exist. Some of them are linked to the role of IRPs in the control of the sexual reproduction. Evidence of this can be seen in crustacean species in which IRPs are expressed in the androgenic gland, an organ involved in the control of sexual differentiation (Manor et al., 2007; Ventura et al., 2011). In the planaria *Schmidtea mediterranea*, *ilp-1* expression has also clearly been demonstrated in testis lobes and RNAi inhibition of Insulin-like signal resulted in the reduction of the testis lobe size and in a decrease in the spermatid number in mature animals (Miller and Newmark, 2012).

The pacific oyster *Crassostrea gigas* (e.g. *Magallana gigas*) is a mollusc bivalve belonging to lophotrochozoan species. Previous research in this oyster resulted in the identification of three IRPs transcripts issued from the same precursor (Hamano et al., 2005) and signalling pathway related members (Gricourt et al., 2003; Jouaux et al., 2012). The IRP transcripts are expressed in the visceral ganglia during the annual cycle according to the reproductive and energetic status (Hamano et al., 2005). Moreover, the phylogenetic position of this species and its particular reproduction as an alternative hermaphrodite made *C. gigas* an interesting model for the study of IRPs at molecular and functional levels. In addition, our group has already identified the tyrosine kinase receptor for *C. gigas* IRPs (CIR) as well as three effectors of the

signalling pathway and confirmed the involvement of IRPs in the control of growth, reproduction and nutritional status (Gricourt et al., 2003; Jouaux et al., 2012).

Here, we took advantage of the availability of the genome (Zhang et al., 2012) as well as various transcriptomic data in *C. gigas* (Riviere et al., 2015; Zhang et al., 2012, 2014; Zhao et al., 2012) to better understand the evolutionary history of IRP sequences. For this purpose, we first performed *in silico* analyses to explore the range of IRP sequences and precise their structural characteristics. Next, the phylogenetic analysis of molluscan IRPs was performed. Finally, we investigated the expression level and location of the corresponding transcripts using quantitative PCR and *in situ* hybridization in order to gain insights into their physiological roles. Altogether, our results shed light on structural and functional evolution of the insulin superfamily.

2. Material and method

2.1. Analysis of molecular databases

We used a strategy of sequence similarity to identify sequences encoding protein homologous to model vertebrate species (*Mus musculus* INS and IGFs), ecdysozoan species (*Drosophila melanogaster* DILPs) and lophotrochozoan species (*Lymnaea stagnalis* MIP) insulin superfamily members in the predicted proteome (NCBI), in the transcriptome (GIGATON <http://ngspipelines-Sigenae.toulouse.inra.fr>) and genome (NCBI) of *C. gigas* (Thunberg, 1793, taxid: 29159) using reciprocal BLASTP or TBLASTN. We used the BLOSUM62 matrix for NCBI analysis of the databases SWISSPROT for proteomic search, and Oyster_v9 for genomic search. The sequences used for similarity are listed in the Supplementary data Table S1. Predicted expressions were found using Gigaton databases which allow to retrieve annotations, expression levels in adult tissues and developmental stages (obtained by Zhang et al. (2012)) as well as polymorphism data like described in Riviere et al. (2015).

2.2. Multiple sequence alignment and phylogenetic analysis

The protein sequences of the IRPs family identified in *C. gigas* were aligned by the use of Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) using default parameters (Sievers et al., 2011). The phylogenetic analysis was performed with IRP sequences from four other molluscan species (*Lymnaea stagnalis*, *Lottia gigantea*, *Sepia officinalis* and *Aplysia californica*), one vertebrate species (*Homo sapiens*) and one ecdysozoan species (*Drosophila melanogaster*). The evolutionary analyses were conducted in MEGA X (Kumar et al., 2018) using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992) with the Poisson substitution model. The bootstrap consensus tree inferred from 500 replicates is taken to represent the evolutionary history of the taxa analysed (Felsenstein, 1985).

2.3. Analysis of protein sequences

The signal peptide was predicted using PREDISI (<http://www.predisi.de/>). KR potential cleavage sites were identified and compared with those of *L. Stagnalis* (Smit et al., 1988, 1991, 1992, 1993, 1996, 1998). A and B chains as well as C peptide were deduced using the same strategy. Molecular weights were calculated using scansite (http://scansite.mit.edu/calc_mw_pi.html) and verified using ProteinProspector (<http://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct>).

2.4. Genomic organization

The WEBSICIO server was used for the prediction of the genomic organization of scaffolds related to IRP sequences in *C. gigas* (Hatje et al., 2011). The RNA organization was deduced from blastn analysis

linked to genomic organization.

2.5. Animals

Diploid Pacific oysters (*C. gigas*) were provided by a commercial oyster farm in Saint-Vaast la Hougue (Manche, France). Experiments were carried out on animals during their second or third cycle of gametogenesis. Thirty oysters were sampled every two months. The ploidy was individually tested using gill samples as recommended by Jouaux et al. (2010).

2.6. Light microscopy

Visceral ganglia and transverse sections of the gonadal area were fixed for histology in Davidson's fixative at 4 °C (10% glycerol, 20% formaldehyde, 30% ethanol 95°, 30% sterile sea water, 10% acetic acid) for 24 to 72 h. The tissue samples were dehydrated in serial ethanol dilutions, transferred in butanol (Carlo Erba) and embedded in paraffin wax (Roth). Three micrometers sections were stained according to the Prenant-Gabe trichrome protocol (Gabe, 1968). The stages of the gametogenetic cycle were determined according to Heude Berthelin et al. (2001).

2.7. RNA extraction and cDNA synthesis

Four pools of three samples were collected for each tissue sample (adductor muscle, digestive gland, labial palps, inner mantle, edge mantle, gills and heart). Similarly, samples were done for visceral ganglia and gonadal area at each stage of gametogenesis. Each pool was crushed and incubated for 5 min at room temperature in TriReagent® (Sigma Aldrich). One hundred microliters of 1-Bromo-3-chloropropane were added, samples were vortexed and centrifuged (15 min at 4 °C). The aqueous phase was recovered and total RNA extraction was performed using the Nucleospin RNA II kit (Macherey-Nagel). RNA samples were reverse transcribed according to Dheilly et al. (2012).

2.8. Real-time quantitative polymerase chain reaction

The sequences of forward and reverse primers were designed using the Primer3 software and folding was analysed (<http://unafold.rna.albany.edu/?q=mfold>). Quantitative PCR was done using 1X GoTaq SYBR Green Mix (Promega), 8.85 ng cDNA and 900 nM of each primer (Table S2) in a final volume of 15 µL. Each run consisted in one cycle of 5 min at 95 °C followed by 45 cycles of 15 sec at 95 °C and 45 sec at 60 °C. The specific amplification of the target sequences was verified by constructing melting curves (80 cycles of 10 sec, temperature increased by 0.5 °C step from 55 °C to 95 °C). Three reference genes (GAPDH, Arf1, EF1α) were tested for qPCR for each tissue and each condition. They were chosen among the best reference genes as identified by Dheilly et al. (2011). EF1α was found as a reliable normalization gene as no significant difference ($p < 0.05$) of Ct values was observed between the different samples compared. Indeed, EF1α was previously used as reference gene in various tissues of *C. gigas* (Dubos et al., 2018, 2016; Fabioux et al., 2004). Thus, the relative level of each gene expression was calculated for one copy of the EF1α reference gene by using the following formula: $N = 2^{-\Delta Ct}$ (reference to EF1α).

2.9. Normalized cDNA library

Total RNA sample was used for double stranded cDNA synthesis using SMART approach (Zhu et al., 2001). SMART prepared amplified cDNA was then normalized using Duplex Specific Nuclease (DSN) normalization method (Zhulidov et al., 2004). Normalization included cDNA denaturation/reassociation treatment by DNS (Shagin et al., 2002) and amplification of normalized fraction by PCR. Resulted normalized cDNA was used for the “all developmental stages and adult

central nervous system” library preparation of *C. gigas* (Fleury et al., 2009).

2.10. RNA probe preparation

The sequences of interest were amplified by PCR using specific primers (Table S2), with a “all developmental stages and adult central nervous system” directional and normalized cDNA library inserted into the Pal 17.3 plasmid (Evrogen) (Fleury et al., 2009). The samples contained 1X PCR Buffer, 0.2 mM dNTP, primers at 0.2 µM each, 4 mM MgCl₂, 10 ng cDNA, 1U Taq Polymerase. The cycling conditions were as follows: 2 min at 96 °C, 40 cycles at 96 °C, 58 °C and 72 °C during 30 sec each followed by a final step at 72 °C. The PCR products were purified from 1% agarose gel using the Nucleospin® Gel and PCR clean-up kit (Macherey-Nagel). Ligation and transformation were realized according to the protocol of the pGEM®-T Easy Vector Systems kit in JM109 chemocompetent bacteria (Promega). The integration of recombinant plasmids in the transformed bacteria was checked by PCR with T7 and SP6 universal primers. The integrated fragment was then validated by sequencing (GENEWIZ Genomics). For each validated colony, a mini-prep was performed using the Nucleospin® Plasmid kit (Macherey-Nagel). Isolated plasmids were linearized with restriction enzyme either Pst I or Sac II during 2 h at 37 °C and 5 min at 65 °C. After clean up purification (Nucleospin® Gel and PCR Clean-up protocol), the RNA probes were labelled using the protocol of the DIG RNA labelling kit (SP6/T7) (Roche).

2.11. In situ hybridization (ISH)

ISH was performed with 3 µm sections mounted on Superfrost Ultra Plus slides. The sections were heated overnight at 37 °C, 2 min at 56 °C and 30 min at room temperature before dewaxing and rehydration in increasing ethanol dilutions. The sections were treated with Proteinase K (Macherey-Nagel) at 5 µg/mL in Tris-EDTA (Ethylenediaminetetraacetic acid, Sigma-Aldrich) Buffer pH 8 for 10 min between two rinses of 5 min in 1X PBS (Phosphate Buffered Saline, P5493 Sigma-Aldrich). Slides were post fixed with a solution of 1X PBS/PFA 4% (Paraformaldehyde, Fisher Scientific 115–867-11) for 7 min and rinsed for 3 min in PBS and 5 min twice in 2X SSC (Saline Sodium Citrate, 93,017 Sigma-Aldrich). The sections were pre-hybridized for 60 min at 50 °C in hybridization buffer consisting of 4X SSC, 10% dextran sulfate, 1X Denhardt's solution, 2 mM EDTA, 50% deionized formamide (F9037 Sigma-Aldrich) and 500 µg/mL salmon genomic DNA (D1626 Sigma-Aldrich) before the hybridization step with 400 ng/mL of RNA probes overnight at 50 °C. The sections were treated once with 2X SSC and three times with 60% formamide in 0.2X SSC each 5 min at 55 °C before being rinsed twice in 2X SSC and once in 100 mM Tris-HCl pH 7.5 with 150 mM NaCl, 5 min each at room temperature. The slides were treated for 30 min with blocking buffer consisting of 100 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% blocking reagent before incubation with alkaline phosphatase-conjugated anti-DIG antibody (polyclonal, Fab fragments) diluted 1:2000 in blocking buffer. Two rinses of 5 min were made with 100 mM Tris-HCl, 150 mM NaCl and one of 10 min with detecting buffer containing 100 mM Tris-HCl pH 9.5, 100 mM NaCl and 50 mM MgCl₂. The sections were incubated overnight with a BCIP (5-bromo-4-chloro-3'-indolylphosphate)/NBT (nitro-blue tetrazolium) solution (11 681 451 001 Roche) diluted at 1:50 and rinsed 5 min with Tris-EDTA buffer. The dehydration process was then carried out with two baths in 100% ethanol and 5 min in Roti-Histol before mounting.

2.12. Statistical analysis

Statistical analyses were performed using R 3.4.3 and RStudio software. A Wilcoxon test and a non-parametric one-way analysis with permutation test were done to determinate significant differences

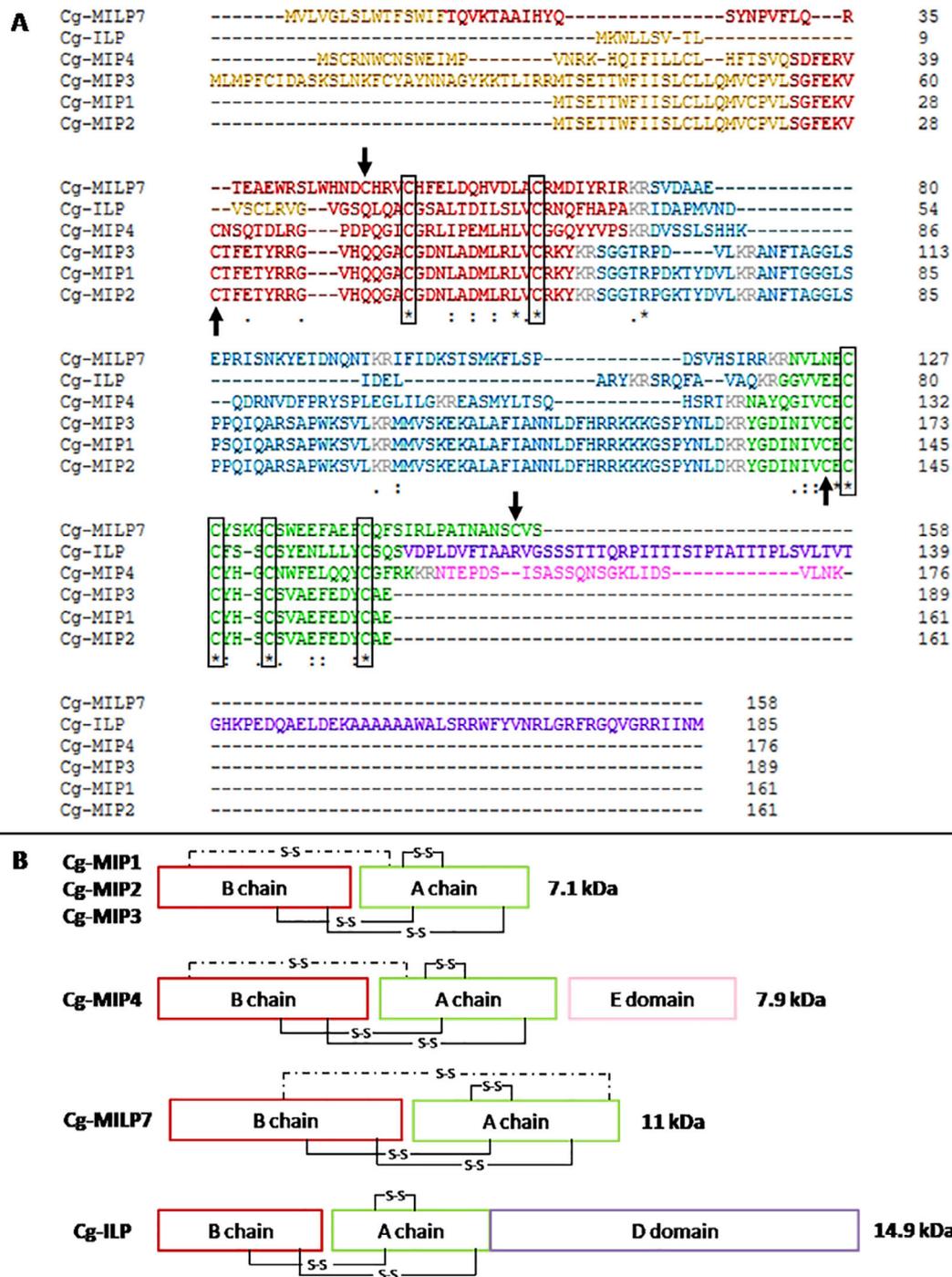


Fig 1. Amino acid alignments (A) and schematic representation of the predicted mature forms (B) for *C. gigas* IRPs. Signal sequences (yellow), B chains (red), C peptides (blue), A chain (green), putative D (purple) and E (pink) domains are specified for each sequence. The putative cleavage sites are either in grey (A) or represented by spaces between chains and domains (B). In the amino acid alignment (A), the six conserved cysteine residues are framed in black and the two additional ones are highlighted by arrows. For the predicted mature forms (B), the three conserved disulphide bonds are annotated by a full line while the putative additional bonds are noted by a dotted line. The estimation of the molecular weight is annotated at the right of each corresponding *C. gigas* IRP.

($\alpha = 0.05$).

3. Results

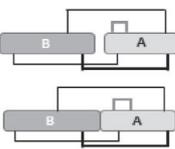
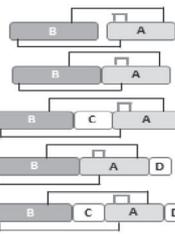
3.1. Identification and classification of IRPs in *C. gigas*

We used sequence similarity and the conserved cysteine residues to identify six sequences of pre-propeptides in the predicted proteome of

C. gigas (NCBI IDs: XP_011455161; BAE48437; XP_011415722; EKC18432; EKC18433; XP_011435161). The functional analysis with INTERPROSCAN (Quevillon et al., 2005) confirmed that they belong to the insulin superfamily. With these sequences as queries, three mRNA sequences were found in *C. gigas* transcriptome GIGATON (Zhang et al., 2012; Zhao et al., 2012; Riviere et al., 2015) (GIGATON IDs: CHOY-P_ILP.1.1, CHOYP_MPI5.1.1. and CHOYP_INS.1.1). In the genome of *C. gigas*, three scaffolds corresponding to these transcripts were detected

Table 2

NCBI accession number, type of IRP with the number of cysteine residues in the B and A chains, signature motif and name for each *C. gigas* IRP.

NCBI IDs	Type	Motif	Name
BAE48437	 β (8 Cys)	C(1X)CC(3X)C(8X)C	Cg-MIP1
XP_011415722			Cg-MIP2
EKC18432			Cg-MIP3
EKC18433			Cg-MIP4
XP_011435161	 γ (6 Cys)	CC(4X)C(8X)C	Cg-MILP7
XP_011455161			Cg-ILP
		CC(3X)C(8X)C	

(NCBI IDs: NW_011937261.1; NW_011937734.1; NW_011935790.1).

The protein sequences were first aligned to identify conserved domains and structural characteristics of the insulin superfamily (Fig. 1A). The sequences exhibited different signal peptides whose length varies from 15 to 55 amino acids. The B chains, range from 26 to 56 amino acids, always end with a pair of basic amino acids (KR) that allows the cleavage of C peptides which present one or two potential internal cleavage sites. The A chains, ranging from 24 to 37 amino acids, display different signature motifs formed according to the cysteine residues number and position. The identified sequences display 6 or 8 cysteine residues distributed in the B and A chains, which allows the formation of 2 or 3 disulphide bonds between both chains and an additional one intra A chain. Therefore, oyster IRPs belong to both β and γ types (Table 2). Among the six identified sequences, only one (NCBI ID: XP_011455161) has a total of 6 cysteine residues that allows the formation of 3 disulphide bonds. So it belongs to the γ type. Its A chain displays the signature motif of vertebrate Insulin-Like Peptides (ILPs): CC(3X)C(8X)C, this is why it is called Cg-ILP. Moreover, the Cg-ILP sequence presents an uncleaved D domain (88 aa). The other five sequences have 8 cysteine residues allowing the formation of an additional disulphide bonds between A and B chains. For this reason, they belong to the β type. Four of them (NCBI IDs: BAE48437; XP_011415722; EKC18432 and EKC18433) have the signature motif of Molluscan Insulin-related peptides (MIPs) of *Lymnea stagnalis*: C(1X)CC(3X)C(8X)C. Hence, these sequences were called Cg-MIP1 to Cg-MIP4. Cg-MIP1, Cg-MIP2 and Cg-MIP3 exhibit strong identity (> 98%, Table S3). They differ widely from Cg-MIP4 (32.54% identity in average). This last sequence holds an E domain (24 aa) potentially cleaved using the dibasic site in the active form. The last β type sequence (NCBI ID: XP_011435161) presents a different A chain motif CC(4X)C(8X)C and a homology with DILP7 of *Drosophila* stronger than with the other Cg-IRPs (Table S3). The possible existence of four disulphide bonds associated to DILP7 homology led us to name this sequence Cg-MILP7 (Molluscan Insulin-Like Peptide 7). Whereas the identity between sequences is low, the six key cysteine residues are conserved in accordance with insulin superfamily characteristics.

Mature forms were predicted according to the sequence similarity and the identification of cleavage sites (Fig. 1B). The molecular weight of each of these predicted mature forms was calculated. As expected and due to their strong homology, Cg-MIP1, Cg-MIP2 and Cg-MIP3

presented the same molecular weight (7.1 kDa). The Cg-MIP4 differs from the other Cg-MIPs by 16 aa and 24 aa in the A and B chains respectively. The molecular weight of the Cg-MIP4 mature form was estimated at 7.9 kDa after cleavage of the E domain. Cg-MILP7 and Cg-ILP molecular weights were estimated at 11 kDa and 14.9 kDa respectively.

3.2. Genomic organisation of *C. gigas* IRPs

The schematic representation of IRPs genomic location (Fig. S1) showed that IRPs were dispatched on three scaffolds in relation to their structural characteristics. Indeed, all the Cg-MIPs were found on the scaffold 43798. It appeared that Cg-MIP1, Cg-MIP2 and Cg-MIP3 share the same location on this scaffold. Cg-MIP4 was found at a different location, but still on scaffold 43798. Cg-MILP7 and Cg-ILP were localized on scaffold 43,272 and scaffold 737 respectively.

A detailed analysis of the organization of genes, transcripts and peptides confirmed the similarity between Cg-MIP1, Cg-MIP2 and Cg-MIP3 (Fig. 2). Besides sharing the same scaffold and the same transcript, named Cg-mip123 (GIGATON ID: CHOYP_MPI5.1.1), the pre-propeptides structures are close together as well. They all contain B and A chains with three and five cysteine residues respectively and four cleavage sites (KR) in the C peptide and at its extremities. Cg-MIP3 differed from Cg-MIP1 and Cg-MIP2 in that the signal sequence is longer and issued from two different exons (exons 1 and 2) whereas Cg-MIP1 and Cg-MIP2 signal peptides originated from exon 2 only. Both genomic sequences encoding Cg-MIP1 and Cg-MIP2 shared the same predicted organization based on two exons (exons 2 and 3) separated by the intron 2. Interestingly, genomic sequence encoding Cg-MIP4 was located in a different region of scaffold 43798. It displayed three exons and two introns but, while using BLASTN, it appeared to be associated with a different RNA sequence, namely Cg-mip4 (GIGATON ID: CHOYP_INS.1.1). The pre-propeptide sequence Cg-MIP4 appeared to be slightly different from other Cg-MIPs. Its C peptide may be cleaved in two parts instead of three and a potentially cleaved E domain was identified at the C-terminal extremity of the sequence.

The Cg-MILP7 was related to three exons and two introns of scaffold 43272. The associated transcript Cg-milp7 was not present in the databases. Cg-MILP7 presented an amino acid sequence with B and A chains containing three and five cysteine residues respectively and a C peptide potentially cleaved in two parts.

The Cg-ILP was related to its own DNA sequence (scaffold 737) and its own transcript named Cg-ilp (GIGATON ID: CHOYP_ILP.1.1). The genomic sequence possesses three exons and two introns. The amino acid sequence differs due to its six cysteine residues, two of which found on the B chain and four on the A chain. Cg-ILP presents a C peptide, which is potentially cleaved in two parts. A D domain is also present at the C-terminal extremity.

Whatever the IRPs of *C. gigas*, the A and B chains were respectively encoded by two successive exons from the genomic sequences.

3.3. Phylogenetic analysis of molluscan IRPs

The phylogenetic analysis of molluscan IRPs is shown in Fig. 3. For this purpose, we selected IRPs from five molluscan species, *C. gigas* (Cg-MIP1 to 4, Cg-ILP and Cg-MILP7), *S. officinalis* (Insulin), *L. stagnalis* (MIP I to VII and Ls-MILP7), *A. californica* (ILP and Ac-MILP7), *L. gigantea* (Insulin 1 to 4), the vertebrate representative insulin of *Homo sapiens* (Insulin) and the *Drosophila* ILP7 (DILP7). The predicted MILP7 of *A. californica* and *L. stagnalis* were previously described by Veenstra (2010) (ESTs NCBI IDs: EB230021 and ES573590 for *A. californica* and *L. stagnalis* respectively) and the prediction of peptides related to these sequences using EXPASY revealed that they both belong to the MILP7 group. The phylogenetic classification revealed that the Cg-ILP of *C. gigas* and the Insulin 3 of *L. gigantea* are close to each other and grouped with vertebrate insulin. In the same way, the Cg-MIPs are phylogenetically close to the Insulins 1 and 2 of *L. gigantea*, the *S. officinalis* and

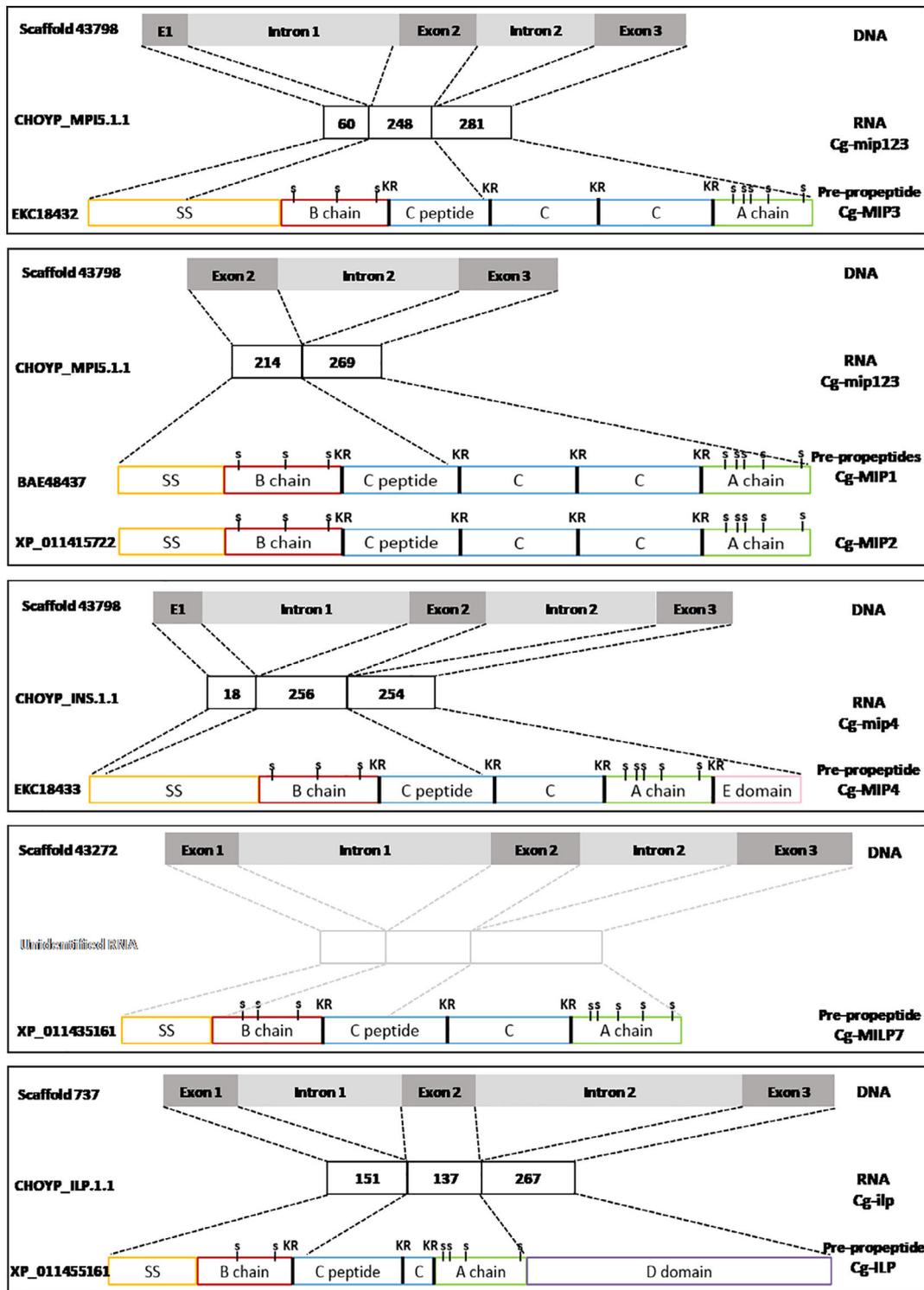


Fig 2. Schematic representation of *C. gigas* scaffolds correlated with transcripts and insulin-like pre-propeptide sequences using WEBSCIPIO. Cysteine residues (s), cleavage sites (KR) and signal sequences (SS) are annotated. Transcript not found in databases were annotated “Unidentified RNA” and represented in light grey.

A. californica IRPs and all the MIPs of *L. stagnalis*. Cg-MIP1, Cg-MIP2 and Cg-MIP3 were very close with each other but phylogenetically more distant from Cg-MIP4. Likewise, *L. stagnalis* MIP I, MIP II, MIP III, MIP V and MIP VII were grouped together. Cg-MILP7 was grouped with the Insulin 4 of *L. gigantea*, the Ls-MILP7 of *L. stagnalis*, the Ac-MILP7 of *A. californica* and they are phylogenetically close to DILP7.

3.4. Quantification and location of *C. gigas* IRPs expression

The *in silico* expression of transcripts Cg-mip4, Cg-ilp and Cg-mip123 were recorded using GIGATON expression data for the developmental stages of *C. gigas* (Fig. 4) obtained by Zhang et al. (2012a,b). Cg-mip4 and Cg-ilp were found to be expressed in all developmental stages. Both transcripts presented the same patterns but Cg-mip4 expression level is clearly higher. The Cg-mip4 transcript (and to a lesser extent Cg-ilp) increased from the beginning of the trochophore larvae

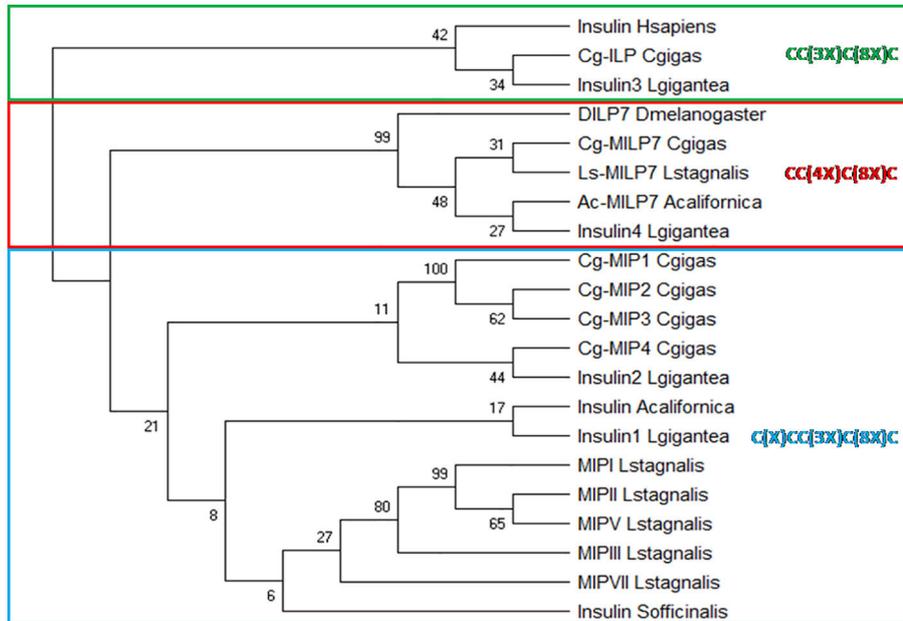


Fig 3. Phylogenetic analysis of some vertebrate, ecdysozoan and molluscan IRPs. *Homo sapiens* (Hsapiens), *Drosophila melanogaster* (Dmelanogaster), *Crassostrea gigas* (Cgigas), *Lymnaea stagnalis* (Lstagnalis), *Sepia officinalis* (Sofficinalis), *Aplysia californica* (Acalifornica) and *Lottia gigantea* (Lgigantea). The phylogenetic tree was constructed using MEGA X software. Bootstraps (500 replicates) are notified next to the branches and signature motifs are annotated for each box.

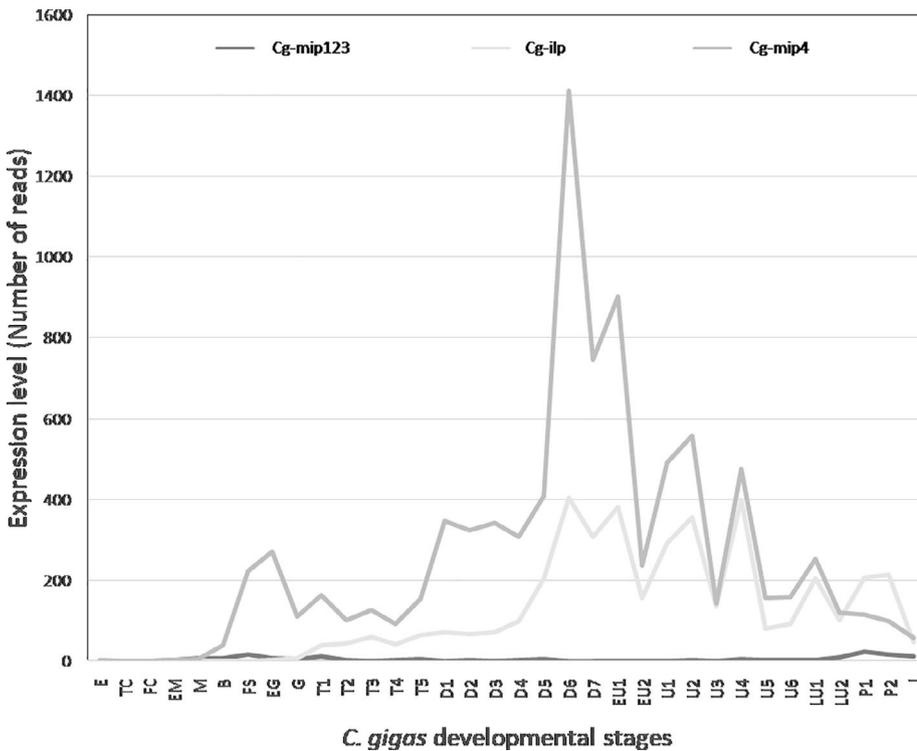


Fig 4. Predicted expressions of Cg-mip123, Cg-ilp and Cg-mip4 transcripts found *in silico* using GIGATON during the developmental stages of *C. gigas*. Unfertilized oocytes (E), two cell embryo (TC), four cell embryo (FC), early morula (EM), Morula (M), Blastula (B), Swimming blastula (FS), early gastrula (EG), gastrula (G), trochophore larvae from 9.5h to 14.5h post-fertilization (T1 to T5), D-shaped larvae from 17.58h to 3.77 days post-fertilization (D1 to D7), early umbo larvae at 5 and 7 days post-fertilization (EU1 and EU2 respectively), umbo larvae from 8 to 14 days post-fertilization (U1 to U6), veliger larvae at 15 and 16 days post-fertilization (LU1 and LU2 respectively), pediveliger larvae at 18.03 and 18.19 days post-fertilization (P1 and P2 respectively), juvenile at 215 days post-fertilization (J) according to Zhang et al. (2012).

stages (T1 to T5) until the juvenile stage (J). The highest expressions of this transcript were observed at the late stages of D-shaped larvae (D6 and D7) and at two umbo larvae stages (U2 and U4). By comparison with the two other transcripts, Cg-mip123 was only weakly expressed during development.

Expressions of Cg-mip123, Cg-ilp and Cg-mip4 transcripts in adult tissues were analysed on the basis of GIGATON transcriptomic data associated to quantitative PCR measurements (Fig. 5). *In silico* expressions showed that all transcripts seemed to be mostly expressed in the digestive gland and in the male mature gonad (Fig. 5A1 to C1). Real time qPCR tends to confirm the overexpression of Cg-mip123 in the digestive gland and the male mature gonadal area, as predicted by *in silico* data. Real time qPCR also revealed that this transcript is expressed

in the adductor muscle and in the visceral ganglia (Fig. 5A1 to A4). The Cg-mip4 is the least transcript expressed in the adult tissues, ten times less than the two others (Fig. 5B2). Concerning Cg-ilp, it appeared that this transcript was significantly overexpressed in the digestive gland as predicted by *in silico* data (Fig. 5C2).

The expression of the three transcripts was measured in the visceral ganglia (Fig. 5A3 to C3) and in the gonadal area (Fig. 5A4 to C4) at each stage of the gametogenic cycle for both male and female. The results showed that, of the three transcripts, the Cg-mip123 is the one that is mostly expressed in both the visceral ganglia and the gonadal area (Fig. 5A3 and A4). Moreover, Cg-mip123 appeared to be a hundred time more expressed in the visceral ganglia than in the gonadal area. Its expression in the visceral ganglia reaches its highest level at the

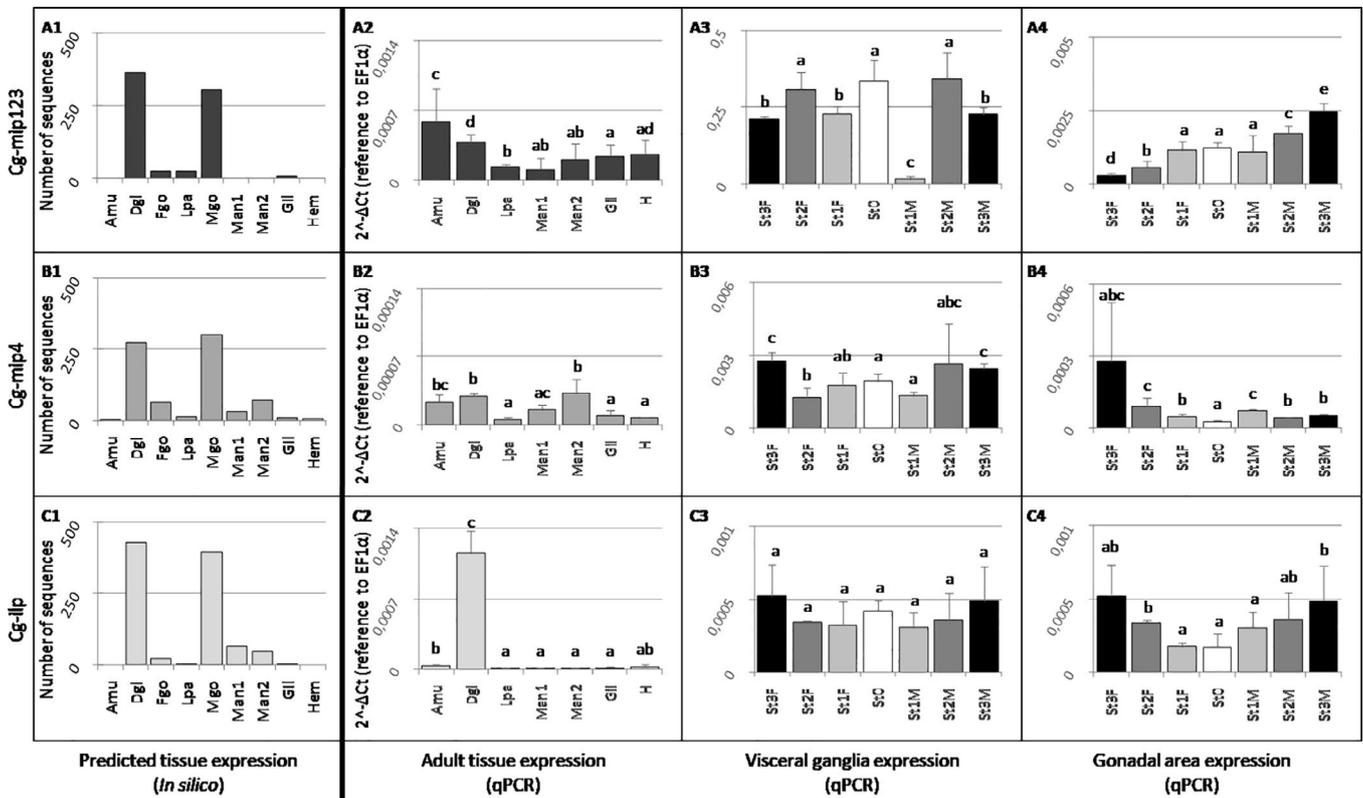


Fig 5. Expressions analyses of Cg-mip123, Cg-mip4 and Cg-ilp transcripts. Predicted expressions found *in silico* using GIGATON (Cg-mip123: A1, Cg-mip4: B1, Cg-ilp: C1) and expressions analyzed by real-time qPCR (Cg-mip123: A2, Cg-mip4: B2, Cg-ilp: C2) in adult tissues of *C. gigas*. Adductor muscle (Amu), Digestive gland (Dgl), female and male gonadal area (Fgo and Mgo respectively), labial palps (Lpa), inner mantle (Man1), mantle edge (Man2), Gill (Gil), hemocytes (Hem), gonadal area (Ago) and heart (H). Expressions measured by real-time qPCR in the visceral ganglia (Cg-mip123: A3, Cg-mip4: B3, Cg-ilp: C3) and in the gonadal area (Cg-mip123: A4, Cg-mip4: B4, Cg-ilp: C4) of animals selected at each reproductive stages. Quiescent stage (St0), male and female gonial mitosis initiating stages (St1M and St1F respectively), male and female germline development stages (St2M and St2F respectively) and male and female sexual maturity stages (St3M and St3F respectively). Each bar is the mean of 4 pools of 3 samples each tested in triplicate. The letters indicate significant differences between bars based on a non-parametric permutation test (p -value < 0.05)

quiescent stage (stage 0) and the male and female germline development stages (stages 2F and 2M). It reaches its lowest level at the initiation of male gonial mitosis (stage 1M). Interestingly, Cg-mip123 expression levels were constant in the gonadal area over the early gametogenetic stages (stage 0 corresponding to the quiescent stage, stages 1M and 1F corresponding to male and female gonial mitosis stages respectively). A significant decrease of Cg-mip123 expression was then observed during the female gametogenesis progression when on the contrary, a significant increase was noted during the male gametogenesis progression. Stage 3M (male at sexual maturity) presented the Cg-mip123 highest expression level in the gonadal area. The Cg-mip4 transcript showed constant levels of expression in the visceral ganglia at the early gametogenetic stages and a tendency to increase at male and female sexual maturity (stages 3F and 3M). In the gonadal area, the Cg-mip4 transcript levels were lower than in the visceral ganglia but tended to increase at the female sexual maturity stage (Fig. 5B3 and B4). Concerning Cg-ilp, no significant difference was shown between the gametogenetic stages in the visceral ganglia. Moreover, the transcript levels were in the same order of magnitude in both the visceral ganglia and the gonadal area (Fig. 5C3 and C4). The Cg-ilp transcript expression in the gonadal area tended to increase as the gametogenesis progressed, in both male and female oysters.

In order to better understand the involvement of Cg-IRPs in reproduction, *in situ* hybridization labelling was only investigated for the transcripts putatively involved in reproduction processes whom expression significantly increase during adult gametogenesis. Subsequently, Cg-mip123 and Cg-ilp expression were explored in the gonadal area and in visceral ganglia. Only the Cg-mip123 transcript

could be detected in the both tissues regardless of the stage of gametogenesis for both males and females. In the visceral ganglia, Cg-mip123 labelled cells appeared to be grouped together in a localised area whatever the stage of gametogenesis (Fig. 6). At the cellular level, Cg-mip123 labelling was localised mainly in the cytoplasm of neurosecretory cells. Variations of labelling intensity may result from the position of the cutting plan in the ganglia. A cartography of visceral ganglia has been established based upon the location of the labelled (Fig. 7). The Cg-mip123 producing cells are located in a lobe of the visceral ganglia, near the epidermal surface. In this area, the positive cells do not constitute a cohesive group but appear to be relatively dispersed. The Cg-mip123 transcript was also clearly detected into the gonadal tubules of mature males at stage 3 (Fig. 8). The labelled cells were located at the periphery of the gonadal tubules in the area containing early germinal cells (spermatogonia) and intra-gonadal somatic cells.

4. Discussion

The present study reports the identification of six IRP sequences in *C. gigas* and the detection of their conserved motif in databases. Moreover, the expression patterns of these IRPs were specified in order to gain insight into their putative biological functions and more precisely their involvement in the adult reproduction.

4.1. Comparison of *C. gigas* IRPs and phylogenetic analysis

Despite the low sequence homology of the members, IRPs

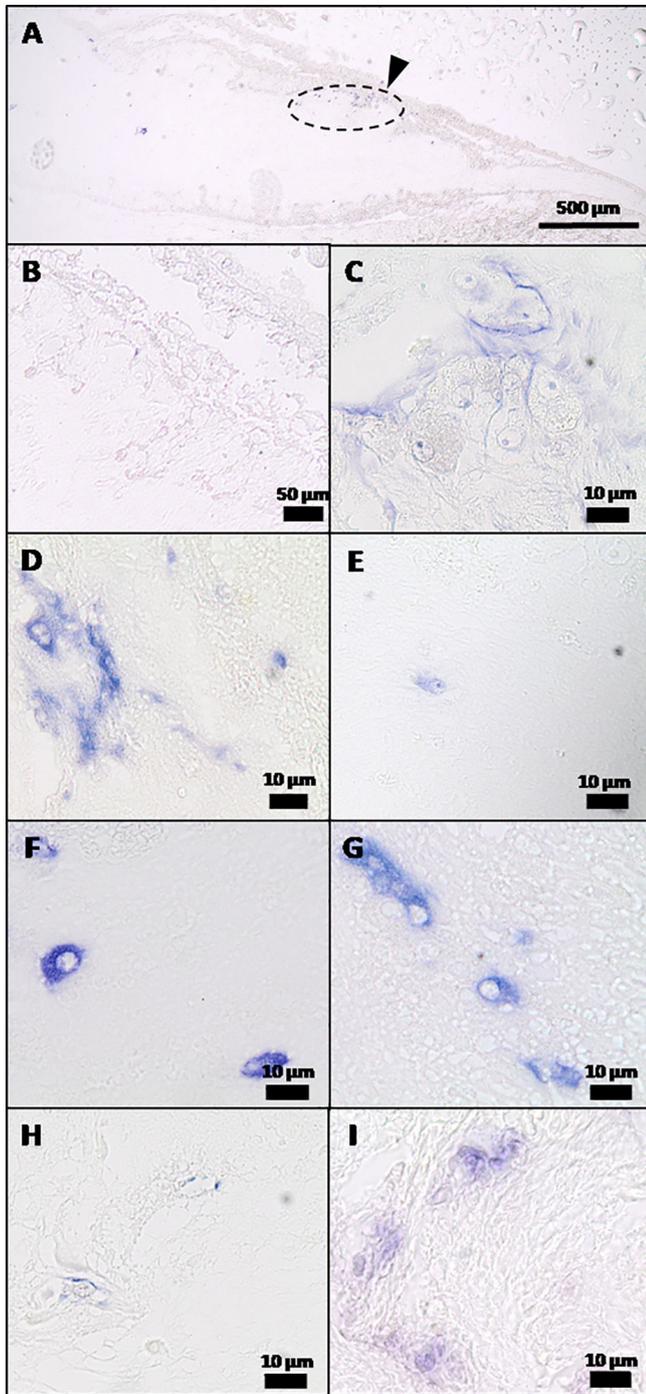


Fig 6. Transcript location of Cg-mip123 in the visceral ganglia by ISH at each stages of gametogenesis. General view of ISH labeling with antisense probe (A) with the labelled area designated by an arrowhead and ISH control with sense probe (B). Quiescent stage (Stage 0, C), gonial mitosis stages female and male (Stage 1, D and E respectively), female and male germline development stages (Stages 2, F and G respectively) and female and male sexual maturity stages (Stages 3, H and I respectively). Observation of neurosecretory cells labelled by Cg-mip123 transcript at magnification 100 for each gametogenic stage (C to I).

configuration led to the identification of conserved signature motifs containing cysteine residues involved in the canonical disulphide bonds which are CC(3X)C(8X)C for Cg-ILP, C(1X)CC(3X)C(8X)C for Cg-MIPs and CC(4X)C(8X)C for Cg-MILP7. The cysteines are responsible for the formation of three or four disulphide bonds depending on the IRPs. They include two canonical bonds between the A and B chains, one intra-A chain and a facultative additional bond between the A and B

chains. Interestingly, the principal segregation in the phylogenetic tree depends first on the IRP signature motif, then on the species. This observation and the fact that all the motifs are found in *L. gigantea* and *C. gigas*, demonstrate that these A chain motifs were present before the segregation of molluscan classes. ILPs were not identified in *L. stagnalis* nor *A. californica*. However, the existence of ILPs cannot be excluded in these species, seen the absence of a whole sequenced genome. Moreover, IRPs identification was targeted in the nervous system. In other gastropod species belonging to the group of the cone snail molluscs, some sequences with ILP motifs were also identified (Table S1). Thus, in *Conus geographus*, sequences with MIP or ILP motifs were identified (Safavi-Hemami et al., 2016). However, these IRPs were produced as venom for predation according to prey preference. They mimic IRPs of their prey, thus interfering with the insulin signal or leading to hypoglycemic shock. However, the authors postulate that the genes of the venom insulins differ from canonical insulin genes and should be issued from a duplication event early in the history of conus. For this reason, we deliberately discarded these molluscan IRPs from our molluscan phylogenetic tree because of their adaptive function.

One of the IRP sequences in *C. gigas*, named Cg-ILP, presented a structural homology with the vertebrates' insulin and a non-cleavable D domain at the C-terminal end of the sequence observed in vertebrates' IGF (Shabanpoor et al., 2009). According to its characteristics, we assigned Cg-ILP in the γ type. The genomic sequence of Cg-ILP presented similarities with the insulin gene structure described in vertebrates. Moreover, as currently observed in other species (Chan and Steiner, 2000), two among the three exons code for the A and B chain sequences and the last one appeared to code for the uncleavable D domain. The identification of Cg-ILP in a lophotrochozoan species demonstrates the ancestral conservation of this IRP gene organization coding for γ type, prior to vertebrate segregation. Indeed, other ILPs with the same structure than the vertebrates' insulins were also found in invertebrate species as *L. gigantea* (Veenstra, 2010), *P. dumerilii* (Conzelmann et al., 2013) and *Nematostella vectensis* (Ancil, 2009).

In this paper, we identified Cg-MIPs belonging to the β type named Cg-MIP 1 to 4. Before *C. gigas* genome availability, Hamano et al. (2005) already identified a cDNA of MIPs, named oIRP, in the visceral ganglia of the oyster. We confirmed here the protein encoded by oIRP corresponds to the Cg-MIP1 sequence in the predicted proteome of *C. gigas*. All Cg-MIPs possess the classical structure of the IRPs of *L. stagnalis* (Smit et al., 1988, 1991, 1992, 1993, 1996, 1998) with an additional disulphide bond between A and B chains. Their designation as MIPs for Molluscan Insulin-related Peptides is due to their first identification in molluscan species but recent data showed that they are rather extended to the lophotrochozoa (Conzelmann et al., 2013). The four Cg-MIP sequences display some differences. Cg-MIP1 and Cg-MIP2 differ only by three amino acids located in the C peptide. These substitutions result from Single Nucleotide Polymorphism (SNPs) as predicted in GIGATON databases (Rivière et al., 2015; Zhang et al., 2012). Indeed, *C. gigas* possesses one of the highest level of genomic polymorphism (Qi et al., 2017; Zhang et al., 2012) and in the Cg-mip123 transcript, SNPs are regularly distributed along the whole sequence. However, amino acid substitutions only occur in the C peptide, cleaved in the mature protein. This situation of higher rate of mutations acceptance in the C peptide was already described by Steiner et al. (1985) for vertebrate IRPs. Cg-MIP3 appeared to be issued from a splicing variant of its transcript. The presence of a cleavable E domain in the Cg-MIP4 structure was also found in *P. dumerilii* IRP-2, IRP-3 and IRP-5 (Conzelmann et al., 2013). The genomic organization presents the classical organization of the IRP family with one exon coding for the B chain and another coding for the A chain (Steiner et al., 1985). Moreover, the Cg-MIPs C peptides are issued from two different exons, as described for MIP II and MIP V genes in *L. stagnalis* (Smit et al., 1998).

The last Cg-IRP sequence identified in our study was called Cg-MILP7. It exhibited three potential A-B disulphide bonds like the Molluscan Insulin-like Peptides but also the A-chain motif of DILP7 of

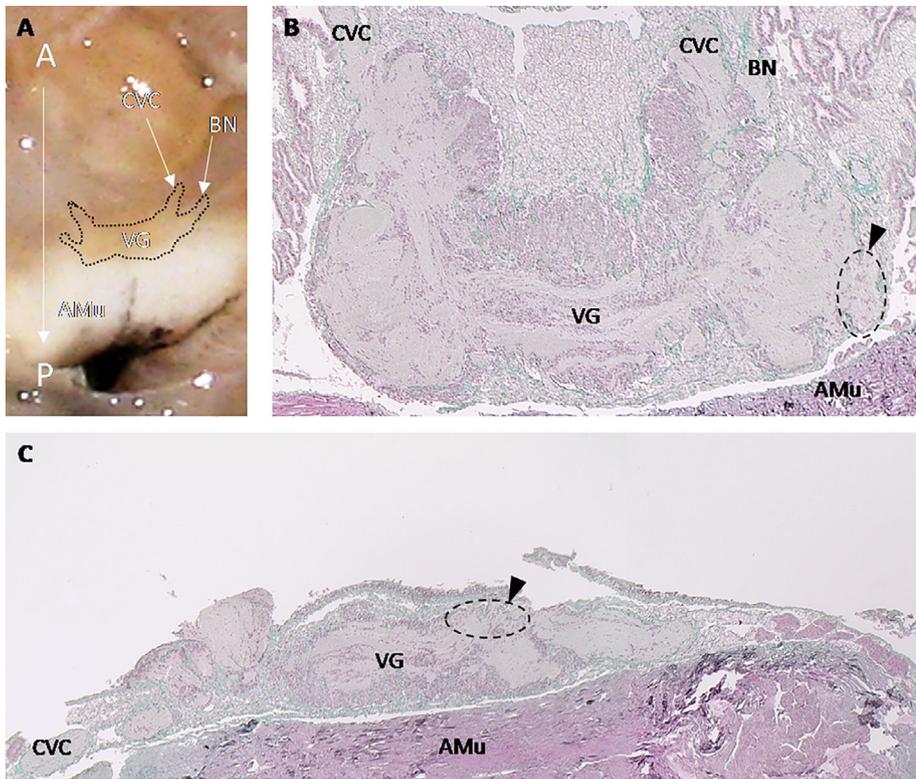


Fig 7. Photography (A) and trichrome Prenant-Gabe coloration of transverse (B) and longitudinal (C) sections of *C. gigas* visceral ganglia (VG). The antero-posterior (A/P) axis, the adductor muscle (AMu), cerebro-visceral connective (CVC) and branchial nerve (BN) are annotated. The arrowheads indicate the location where the labelling is visible.

drosophila [CC(4X)C(8X)C]. This ILP7 ortholog was also identified in the genome of *L. gigantea* (Insulin 4), in the ESTs databases of *L. stagnalis* (ES573590 and ES573220) and *A. californica* (EB230021) (Veenstra, 2010). As expected, all these molluscan MILP7 sequences appeared grouped together as shown in the phylogenetic tree. Moreover, when the phylogenetic analysis includes the DILP7 of *D. melanogaster*, it appears to be grouped with its homologous molluscan sequences, leading to suppose that the motif should be conserved upstream the ecdysozoan specification. The *Cg-milp7* transcript could not be found whatever the databases suggesting a very low or targeted expression. Further experiments would consist in the characterization

of its expression pattern of this atypical IRP.

Until recently, most scientific research on IRPs in molluscan species referred to MIP molecules. The existence of ILP and MILP7 was firstly reported in *L. gigantea*, which was the first mollusc to have its genome sequenced. Our work on the oyster *C. gigas* confirmed the existence of three types of molecules belonging to IRPs in molluscs. The three types are present in several molluscan species and considering the phylogenetic segregation, these types would probably be ancestral to molluscs providing new questions concerning the evolution of the insulin family in metazoan. This superfamily of molecules is known to be involved in multiple physiological regulations. Subsequently, the various IRPs we

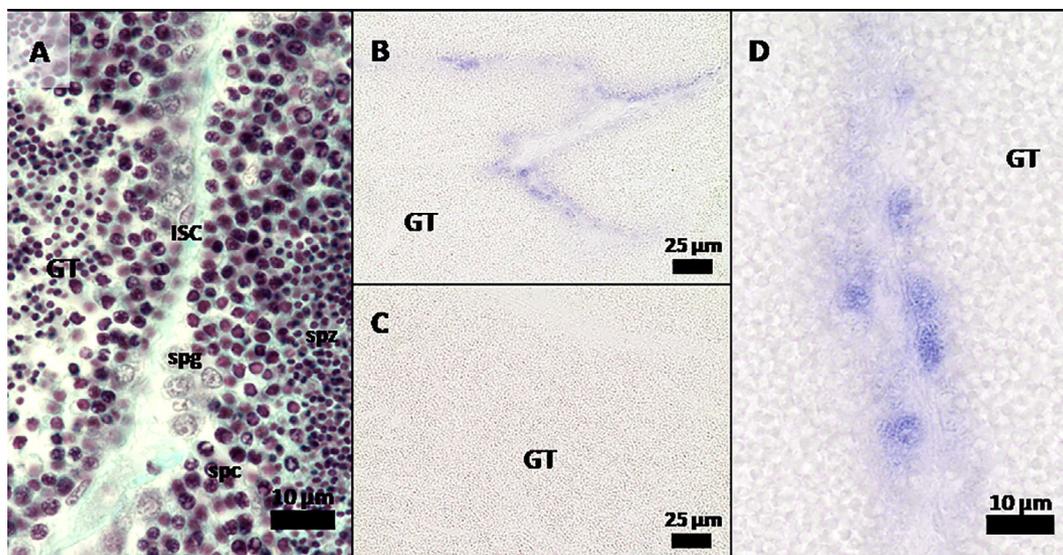


Fig 8. Trichrome Prenant-Gabe coloration of the gonadal tubules (GT) at male sexual maturity (Stage III, A). Spermatogonia (spg), spermatocytes (spc), spermatozoa (spz) and intra-gonadal somatic cell (ISC) are annotated. General view of the transcript location of *Cg-mip123* in the gonadal area by ISH at the male sexual maturity stage with antisense probe (B) and control with sense probe (C). *Cg-mip123* transcript location of the gonadal tubule at male sexual maturity at magnification 100 (D).

had identified in *C. gigas* probably support specific and conserved physiological roles related to their structural properties, including development, reproduction, sex change and energetic metabolism.

4.2. Evidence of multiple roles of IRPs in *C. gigas*

We also characterized Cg-IRPs expression during development and in various adult tissues in order to investigate their physiological involvement. Predicted expressions were investigated using GIGATON databases and completed by qPCR experiments for adult tissues, as GIGATON databases did not inform about expression in the visceral ganglia and in the gonadal area along the gametogenesis. Indeed, in invertebrate species, IRPs expression was mainly observed in the nervous system. The expression of *C. gigas* IRPs was also mainly detected in the nervous system, with Cg-MIPs being the most expressed IRPs in the visceral ganglia throughout the year. Cg-MIPs should be involved in the central control of multiple physiological processes, in agreement with the results published by Hamano et al. (2005) on the same species. In complement, we also demonstrated local expressions of the other Cg-IRPs in the digestive and gonadal tissues as well as during larval development.

4.2.1. A putative role of Cg-MIP4 during development

According to the databases, Cg-mip4 transcript is mostly expressed during the larval development. The expression of this transcript increased mainly at the end of the D-larvae stage, when larvae start to feed themselves, with minor peaks during umbo stages. The involvement of IRPs in the control of development is highly conserved: in vertebrate species, Insulin-like Growth Factors (IGFs) provide essential signals for the control of embryonic and postnatal development (Dupont and Holzenberger, 2003). In insects and nematodes, the insulin signalling pathway plays essential roles in the control of development and organ growth in relation with the nutritional status (Lin and Smaghe, 2018; Schindler et al., 2014). In the oyster *C. gigas*, Cg-MIP4 might be a key actor in developmental control.

4.2.2. The putative role of Cg-ILP in digestive processes

As expected by *in silico* expression data, Cg-ilp was mainly expressed in the digestive gland and the gonadal area. These results fit with the digestive expression of some IRPs described in various molluscan species as for example *Mytilus edulis* (Fritsch et al., 1976), *Anodonta cygnea*, *Unio pictorum* (Plisetskaya et al., 1978) and *Helisoma duryi* (Abdraba and Saleuddin, 2006). Giard et al. (1998) observed that vertebrate IGF-1 increase amino acid incorporation into proteins in dissociated digestive cells of *Pecten maximus*. More recently, Jouaux et al. (2012) observed the expression of the insulin receptor in the digestive gland of *C. gigas* and suggested a paracrine function of the IRPs produced during digestion. This digestive role of IRPs could also be observed in other invertebrate species. Among the eight DILPs from *D. melanogaster*, DILP3 is expressed in the brain but also in the midgut (Veenstra et al., 2008) and plays a critical role in promoting sugar homeostasis (Kim and Neufeld, 2015). This role in the regulation of nutritional state and carbohydrate metabolism is also crucial and conserved in vertebrate species.

4.2.3. Potential involvement of Cg-ILP and Cg-MIPs in the control of reproduction

In *C. gigas*, Cg-ilp and Cg-mip123 are both expressed in the gonadal area with differential expression related to gametogenetic cycle. Cg-ilp expression significantly increase according to the course of gametogenesis in females as well as in males, allowing to postulate that gonadal Cg-ILP might play a part in the control of gametogenesis and/or associated metabolic pathways. The relatively low Cg-ilp expression did not allow the identification of cells producing Cg-ilp by *in situ* hybridization. However, the involvement of IRPs in the control of gametogenesis is well known in numerous invertebrate species. In the

lophotrochozoan planaria *S. mediterranea*, ilp-1 is expressed in germ cells of the testis lobe and ILP disruption impaired testis growth and spermatogenesis (Miller and Newmark, 2012). In *D. melanogaster* DILP2, DILP3 and DILP5 (presenting the ILPs signature motif [CC(3X)C(8X)C]) are involved in the control of germ stem cell proliferation defining the final reproductive effort (Grönke et al., 2010; Hsu and Drummond-Barbosa, 2009). In the nematode *C. elegans*, Michaelson et al. (2010) indicate that, under replete reproductive conditions, the larval germline responds to insulin signalling (INS-3 and INS-33) to ensure robust germline proliferation and amplify the germline stem cell population. The activation of insulin signalling pathway acts in parallel with Notch signalling, thus affecting the number of cells in the proliferative zone. Moreover, for all these invertebrate species, germline proliferation is closely linked to nutritional status via insulin signalling. If we consider its expression pattern in *C. gigas*, Cg-ILP may be the best candidate to assume the dual control of nutritional status and reproductive effort. Further measurements of Cg-ilp expression based on the nutritional status and germline cells proliferation should be performed to confirm this hypothesis.

The expression of Cg-mip123 was also detected in the gonadal area even though much lower than in visceral ganglia. The Cg-mip123 expression is mainly related to male sexual maturity, both for qPCR and ISH labelling approaches. Interestingly, ISH positive cells were localized near the tubule wall in the location of early germinal cells and associated intragonadal somatic cells (Franco et al., 2008). The small size and low number of labelled cells did not allow us to identify with certainty the nature of the cells but a co-labelling with the germ cells marker Oyvlg (Fabioux et al., 2004) could clarify their belonging to the germinal or somatic lineage. In crustacean species, IRPs were found to be involved in the maintenance of male characteristics (Chandler et al., 2015, 2017; Manor et al., 2007; Sharabi et al., 2016; Ventura et al., 2009). Therefore, the specific detection of Cg-mip123 in the male gonad questions the androgenic role of Cg-mip123 in the oyster, which is an annual alternative hermaphrodite.

5. Conclusion

Regarding the phylogenetic analysis, the most commonly distributed motif is that of ILPs, leading to postulate its ancestry and conservation in metazoan. ILP7 orthologs were found in a diversity of protostomian species, from insects (DILP7) to molluscs (MILP7). ILP7 orthologs may be derived from ILPs resulting from duplication and evolutionary pressure in protostomes. Molluscan MIPs constitutes another group of ILPs, with a different evolutionary history. These MIPs were also found in Annelida (Conzelmann et al., 2013) but were not identified in Platyhelminthes (Wang et al., 2013) at the moment. Our work would lead us gain insight into the evolutionary history of MIPs in lophotrochozoa. So far, our data provided some clues concerning the multiple and conserved roles of IRPs in molluscs. Our expression data suggested the involvement of Cg-ILP and Cg-MIPs in the control of the adult gametogenesis. The roles of IRPs in the control of these physiological processes were rarely mentioned in molluscan species and have now to be explored.

Acknowledgements

PhD of M. Cherif-Feildel was financially supported by the French research ministry (PhD grant). The authors are grateful to M.P. Dubos for providing some visceral ganglia slides used for ISH. Proofreading of the English text has been realised by C. Quint of the centre of translation of UNICAEN.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ygcen.2018.10.019>.

References

- Abdraba, A.M., Saleuddin, S., 2006. Insulin-like material in the digestive tract of the freshwater snail *Helisoma duryi* (Mollusca: Pulmonata). *Pure Sci.* 33, 85–92.
- Agrogiannis, G.D., Sifakis, S., Patsouris, E.S., Konstantinidou, A.E., 2014. Insulin-like growth factors in embryonic and fetal growth and skeletal development (Review). *Mol. Med. Rep.* 10, 579–584. <https://doi.org/10.3892/mmr.2014.2258>.
- Ancil, M., 2009. Chemical transmission in the sea anemone *Nematostella vectensis*: a genomic perspective. *Comp. Biochem. Physiol. - Part D Genomics Proteomics* 4, 268–289. <https://doi.org/10.1016/j.cbd.2009.07.001>.
- Apfeld, J., Kenyon, C., 1998. Cell nonautonomy of *C. elegans* *daf-2* function in the regulation of diapause and life span. *Cell* 95, 199–210. [https://doi.org/10.1016/S0092-8674\(00\)81751-1](https://doi.org/10.1016/S0092-8674(00)81751-1).
- Ashrafi, K., Chang, F.Y., Watts, J.L., Fraser, A.G., Kamath, R.S., Ahringer, J., Ruvkun, G., 2003. Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature* 421, 268–272. <https://doi.org/10.1038/nature01279>.
- Avogaro, A., De Kreutzenberg, S.V., Fadini, G.P., 2010. Insulin signaling and life span. *PLoS Arch. Eur. J. Physiol.* 459, 301–314. <https://doi.org/10.1007/s00424-009-0721-8>.
- Bathgate, R.A.D., Halls, M.L., van der Westhuizen, E.T., Callander, G.E., Kocan, M., Summers, R.J., 2013. Relaxin family peptides and their receptors. *Physiol. Rev.* 93, 405–480. <https://doi.org/10.1152/physrev.00001.2012>.
- Bayazit, V., 2009. Evaluation of the Bombyxin Gene and Bombyxin Insulin-like Peptide in Silkworm (*Bombyx Mori*). *Aust. J. of Basic Appl. Sci.* 3, 1032–1042.
- Böhni, R., Riesgo-Escovar, J., Oldham, S., Brogiolo, W., Stocker, H., Andrus, B.F., Beckingham, K., Hafen, E., 1999. Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1–4. *Cell* 97, 865–875. [https://doi.org/10.1016/S0092-8674\(00\)80799-0](https://doi.org/10.1016/S0092-8674(00)80799-0).
- Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., Hafen, E., 2001. An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* 11, 213–221. [https://doi.org/10.1016/S0960-9822\(01\)00068-9](https://doi.org/10.1016/S0960-9822(01)00068-9).
- Broughton, S.J., Piper, M.D.W., Ikeya, T., Bass, T.M., Jacobson, J., Driege, Y., Martinez, P., Hafen, E., Withers, D.J., Leevers, S.J., Partridge, L., 2005. Longer lifespan, altered metabolism, and stress resistance in *Drosophila* from ablation of cells making insulin-like ligands. *Proc. Natl. Acad. Sci.* 102, 3105–3110. <https://doi.org/10.1073/pnas.0405775102>.
- Broughton, S.J., Slack, C., Alic, N., Metaxakos, A., Bass, T.M., Driege, Y., Partridge, L., 2010. DILP-producing median neurosecretory cells in the *Drosophila* brain mediate the response of lifespan to nutrition. *Aging Cell* 9, 336–346. <https://doi.org/10.1111/j.1474-9726.2010.00558.x>.
- Chan, S.J., Steiner, D.F., 2000. Insulin through the ages: phylogeny of a growth promoting and metabolic regulatory hormone. *Am. Zool.* 40, 213–222. [https://doi.org/10.1668/0003-1569\(2000\)040\[0213:ITTAPO\]2.0.CO;2](https://doi.org/10.1668/0003-1569(2000)040[0213:ITTAPO]2.0.CO;2).
- Chandler, J.C., Aizen, J., Elizur, A., Hollander-Cohen, L., Battaglene, S.C., Ventura, T., 2015. Discovery of a novel insulin-like peptide and insulin binding proteins in the Eastern rock lobster *Sagmariasus verreauxi*. *Gen. Comp. Endocrinol.* 215, 76–87. <https://doi.org/10.1016/j.ygcen.2014.08.018>.
- Chandler, J.C., Gandhi, N.S., Mancera, R.L., Smith, G., Elizur, A., Ventura, T., 2017. Understanding insulin endocrinology in decapod crustacea: molecular modelling characterization of an insulin-binding protein and insulin-like peptides in the eastern spiny lobster, *Sagmariasus verreauxi*. *Int. J. Mol. Sci.* 18, 1–19. <https://doi.org/10.3390/ijms18091832>.
- Clancy, D.J., Gems, D., Harshman, L.G., Oldham, S., Stocker, H., Hafen, E., Leevers, S.J., Partridge, L., 2001. Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. *Science* 292, 104–106. <https://doi.org/10.1126/science.1057991>.
- Conzelmann, M., Williams, E.A., Krug, K., Franz-Wachtel, M., Macek, B., Jékely, G., 2013. The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* 14, 906. <https://doi.org/10.1186/1471-2164-14-906>.
- De Meys, P., 2004. Insulin and its receptor: structure, function and evolution. *BioEssays* 26, 1351–1362. <https://doi.org/10.1002/bies.20151>.
- Dheilly, N.M., Lelong, C., Huvet, A., Favrel, P., 2011. Development of a Pacific oyster (*Crassostrea gigas*) 31,918-feature microarray: identification of reference genes and tissue-enriched expression patterns. *BMC Genomics* 12, 468. <https://doi.org/10.1186/1471-2164-12-468>.
- Dheilly, N.M., Lelong, C., Huvet, A., Kellner, K., Dubos, M.-P., Riviere, G., Boudry, P., Favrel, P., 2012. Gametogenesis in the Pacific oyster *Crassostrea gigas*: a microarrays-based analysis identifies sex and stage specific genes. *PLoS One* 7, e36353. <https://doi.org/10.1371/journal.pone.0036353>.
- Dimitriadis, G., Mitrou, P., Lambadiari, V., Maratou, E., Raptis, S.A., 2011. Insulin effects in muscle and adipose tissue. *Diabetes Res. Clin. Pract.* 93, S52–S59. [https://doi.org/10.1016/S0168-8227\(11\)70014-6](https://doi.org/10.1016/S0168-8227(11)70014-6).
- Dubos, M.P., Bernay, B., Favrel, P., 2016. Molecular characterization of an adipokinetic hormone-related neuropeptide (AKH) from a mollusk. *Gen. Comp. Endocrinol.* 243, 15–21. <https://doi.org/10.1016/j.ygcen.2016.11.002>.
- Dubos, M.P., Zels, S., Schwartz, J., Pasquier, J., Schoofs, L., Favrel, P., 2018. Characterization of a tachykinin signalling system in the bivalve mollusk *Crassostrea gigas*. *Gen. Comp. Endocrinol.* 266, 110–118. <https://doi.org/10.1016/j.ygcen.2018.05.003>.
- Dupont, J., Holzenberger, M., 2003. Biology of insulin-like growth factors in development. *Birth Defects Res. Part C – Embryo Today Rev.* 69, 257–271. <https://doi.org/10.1002/bdrc.10022>.
- Duret, L., Guex, N., Peitsch, M.C., Bairoch, A., 1998. New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling. *Genome Res.* 8, 348–353. <https://doi.org/10.1101/gr.8.4.348>.
- Fabioux, C., Pouvreau, S., Le Roux, F., Huvet, A., 2004. Oyster vasa-like gene: a specific marker of the germline in *Crassostrea gigas*. *Biochem. Biophys. Res. Commun.* 315, 897–904. <https://doi.org/10.1016/j.bbrc.2004.01.145>.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution (N. Y.)* 39, 783. <https://doi.org/10.2307/2408678>.
- Fleury, E., Huvet, A., Lelong, C., de Lorange, J., Boulo, V., Gueguen, Y., Bachère, E., Tanguy, A., Moraga, D., Fabioux, C., Lindeque, P., Shaw, J., Reinhardt, R., Prunet, P., Davey, G., Lapègue, S., Sauvage, C., Corporeau, C., Moal, J., Gavory, F., Wincker, P., Moreews, F., Klopp, C., Mathieu, M., Boudry, P., Favrel, P., 2009. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics* 10, 341. <https://doi.org/10.1186/1471-2164-10-341>.
- Floyd, P., Li, L., 1999. Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. *J. Neurosci.* 19, 7732–7741.
- Franco, A., Heude Berthelin, C., Goux, D., Sourdaire, P., Mathieu, M., 2008. Fine structure of the early stages of spermatogenesis in the Pacific oyster, *Crassostrea gigas* (Mollusca, Bivalvia). *Tissue Cell* 40, 251–260. <https://doi.org/10.1016/j.tice.2007.12.006>.
- Fritsch, H.A.R., Van Noorden, S., Pearce, A.G.E., 1976. Cytochemical and immunofluorescence investigations on insulin-like producing cells in the intestine of *Mytilus edulis* L. (Bivalvia). *Cell Tissue Res.* 165, 365–369. <https://doi.org/10.1007/BF00222439>.
- Gabe, M., 1968. *Techniques histologiques*, 6th ed. Paris.
- Garelli, A., Gontijo, A.M., Miguela, V., Caparros, E., Dominguez, M., 2012. Imaginal discs secrete insulin-like peptide 8 to mediate plasticity of growth and maturation. *Science* 336 (80), 579–582. <https://doi.org/10.1126/science.1216735>.
- Giard, W., Lebel, J.M., Boucaud-Camou, E., Favrel, P., 1998. Effects of vertebrate growth factors on digestive gland cells from the mollusk *Pecten maximus* L.: an in vitro study. *J. Comp. Physiol. – B Biochem. Syst. Environ. Physiol.* 168, 81–86. <https://doi.org/10.1007/s003600050123>.
- Gricourt, L., Bonnet, G., Boujard, D., Mathieu, M., Kellner, K., 2003. Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrlGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor. *Gen. Comp. Endocrinol.* 134, 44–56. [https://doi.org/10.1016/S0016-6480\(03\)00217-X](https://doi.org/10.1016/S0016-6480(03)00217-X).
- Grönke, S., Clarke, D.-F., Broughton, S., Andrews, T.D., Partridge, L., 2010. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet.* 6, e1000857. <https://doi.org/10.1371/journal.pgen.1000857>.
- Hamano, K., Awaji, M., Usuki, H., 2005. cDNA structure of an insulin-related peptide in the Pacific oyster and seasonal changes in the gene expression. *J. Endocrinol.* 187, 55–67. <https://doi.org/10.1677/joe.1.06284>.
- Hatje, K., Keller, O., Hammesfahr, B., Pillmann, H., Waack, S., Kollmar, M., 2011. Cross-species protein sequence and gene structure prediction with fine-tuned Webscipro 2.0 and Scipio. *BMC Res. Notes* 4, 265. <https://doi.org/10.1186/1756-0500-4-265>.
- Heude Berthelin, C., Laisney, J., Espinosa, J., Martin, O., Hernandez, G., Mathieu, M., Kellner, K., 2001. Storage and reproductive strategy in *Crassostrea gigas* from two different growing areas (Normandy and the Atlantic coast, France). *Invertebr. Reprod. Dev.* 40, 79–86. <https://doi.org/10.1080/07924259.2001.9652500>.
- Hibshman, J.D., Hung, A., Baugh, L.R., 2016. Maternal diet and insulin-like signaling control intergenerational plasticity of progeny size and starvation resistance. *PLOS Genet.* 12, e1006396. <https://doi.org/10.1371/journal.pgen.1006396>.
- Hsu, H.-J., Drummond-Barbosa, D., 2009. Insulin levels control female germline stem cell maintenance via the niche in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1117–1121. <https://doi.org/10.1073/pnas.0809144106>.
- Jékely, G., 2013. Global view of the evolution and diversity of metazoan neuropeptide signaling. *Proc. Natl. Acad. Sci.* 110, 8702–8707. <https://doi.org/10.1073/pnas.1221833110>.
- Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics* 8, 275–282. <https://doi.org/10.1093/bioinformatics/8.3.275>.
- Jouaux, A., Heude Berthelin, C., Sourdaire, P., Mathieu, M., Kellner, K., 2010. Gametogenic stages in triploid oysters *Crassostrea gigas*: irregular locking of gonial proliferation and subsequent reproductive effort. *J. Exp. Mar. Biol. Ecol.* 395, 162–170. <https://doi.org/10.1016/j.jembe.2010.08.030>.
- Jouaux, A., Franco, A., Heude Berthelin, C., Sourdaire, P., Bin, J.L., Mathieu, M., Kellner, K., 2012. Identification of Ras, Pten and p70S6K homologs in the Pacific oyster *Crassostrea gigas* and diet control of insulin pathway. *Gen. Comp. Endocrinol.* 176, 28–38. <https://doi.org/10.1016/j.ygcen.2011.12.008>.
- Kim, J., Neufeld, T.P., 2015. Dietary sugar promotes systemic TOR activation in *Drosophila* through AKH-dependent selective secretion of Dilp3. *Nat. Commun.* 6, 6846. <https://doi.org/10.1038/ncomms7846>.
- Kumar, S., Stecher, G., Li, M., Niyaz, C., Tamura, K., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. <https://doi.org/10.1093/molbev/msy096>.
- Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., Iwami, M., 1996. Multiple gene copies for Bombyxin, an insulin-related peptide of the silkworm *Bombyx mori*: structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of Bombyxin. *J. Mol. Biol.* 259, 926–937. <https://doi.org/10.1006/jmbi.1996.0370>.
- Lagueux, M., Lwoff, L., Meister, M., Goltzené, F., Hoffmann, J.A., 1990. cDNAs from neurosecretory cells of brains of *Locusta migratoria* (Insecta, Orthoptera) encoding a novel member of the superfamily of insulins. *Eur. J. Biochem.* 187, 249–254. <https://doi.org/10.1111/j.1432-1033.1990.tb15302.x>.
- Li, K.W., Geraerts, W.P.M., Joosse, J., 1992. Purification and sequencing of molluscan

- insulin-related peptide II from the neuroendocrine light green cells in *Lymnaea stagnalis*. *Endocrinology* 130, 3427–3432. <https://doi.org/10.1210/en.130.6.3427>.
- Li, W., Kennedy, S.G., Ruvkun, G., 2003. *daf-28* encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. *Genes Dev.* 17, 844–858. <https://doi.org/10.1101/gad.1066503>.
- Lin, X., Smaghe, G., 2018. Roles of the insulin signaling pathway in insect development and organ growth. *Peptides* 0–1. <https://doi.org/10.1016/j.peptides.2018.02.001>.
- Liu, Y., Liao, S., Veenstra, J.A., Nüssel, D.R., 2016. *Drosophila* insulin-like peptide 1 (DILP1) is transiently expressed during non-feeding stages and reproductive dormancy. *Sci. Rep.* 6, 26620. <https://doi.org/10.1038/srep26620>.
- Luo, S., Kleemann, G.A., Ashraf, J.M., Shaw, W.M., Murphy, C.T., 2010. TGF- β and insulin signaling regulate reproductive aging via oocyte and germline quality maintenance. *Cell* 143, 299–312. <https://doi.org/10.1016/j.cell.2010.09.013>.
- Manor, R., Weil, S., Oren, S., Glazer, L., Aflalo, E.D., Ventura, T., Chalifa-Caspi, V., Lapidot, M., Sagi, A., 2007. Insulin and gender: an insulin-like gene expressed exclusively in the androgenic gland of the male crayfish. *Gen. Comp. Endocrinol.* 150, 326–336. <https://doi.org/10.1016/j.ygcen.2006.09.006>.
- Matsunaga, Y., 2017. Diverse insulin-like peptides in *Caenorhabditis elegans*. *Int. Biol. Rev.* 1, 1–15. <https://doi.org/10.18103/ibr.v1i1.1276>.
- Maruyama, K., Hietter, H., Nagasawa, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., 1988. Isolation and primary structure of bombyxin-iv, a novel molecular species of bombyxin from the silkworm, *bombyx mori*. *Agric. Biol. Chem.* 52, 3035–3041. <https://doi.org/10.1080/00021369.1988.10869178>.
- Michaelson, D., Korta, D.Z., Capua, Y., Hubbard, E.J.A., 2010. Insulin signaling promotes germline proliferation in *C. elegans*. *Development* 137, 671–680. <https://doi.org/10.1242/dev.042523>.
- Miller, C.M., Newmark, P.A., 2012. An insulin-like peptide regulates size and adult stem cells in planarians. *Int. J. Dev. Biol.* 56, 75–82. <https://doi.org/10.1387/ijdb.113443can>.
- Narbonne, P., Maddox, P.S., Labbe, J.-C., 2015. *daf-18/PTEN* locally antagonizes insulin signalling to couple germline stem cell proliferation to oocyte needs in *C. elegans*. *Development* 4230–4241. <https://doi.org/10.1242/dev.130252>.
- Olinski, R.P., Dahlberg, C., Thorndyke, M., Hallböök, F., 2006. Three insulin-relaxin-like genes in *Ciona intestinalis*. *Peptides* 27, 2535–2546. <https://doi.org/10.1016/j.peptides.2006.06.008>.
- Partridge, L., Alic, N., Bjedov, I., Piper, M.D.W., 2011. Ageing in *Drosophila*: the role of the insulin/Igf and TOR signalling network. *Exp. Gerontol.* 46, 376–381. <https://doi.org/10.1016/j.exger.2010.09.003>.
- Phoungpetchara, I., Tinikul, Y., Poljaroen, J., Chotwiwatthanakun, C., Vanichviriyakit, R., Sroyraya, M., Hanna, P.J., Sobhon, P., 2011. Cells producing insulin-like androgenic gland hormone of the giant freshwater prawn, *Macrobrachium rosenbergii*, proliferate following bilateral eyestalk-ablation. *Tissue Cell* 43, 165–177. <https://doi.org/10.1016/j.tice.2011.02.001>.
- Pierce, S.B., Costa, M., Wisotzky, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquini, A.A., Liu, L.X., Doberstein, S.K., Ruvkun, G., 2001. Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* 672–686. <https://doi.org/10.1101/gad.867301.4>.
- Plisetskaya, E., Kazakov, V.K., Soltitskaya, L., Leibson, L.G., 1978. Insulin-producing cells in the gut of freshwater bivalve molluscs *Anodonta cygnea* and *Unio pictorum* and the role of insulin in the regulation of their carbohydrate metabolism. *Gen. Comp. Endocrinol.* 35, 133–145.
- Qi, H., Song, K., Li, C., Wang, W., Li, B., Li, L., Zhang, G., 2017. Construction and evaluation of a high-density SNP array for the Pacific oyster (*Crassostrea gigas*). *PLoS One* 12, 1–16. <https://doi.org/10.1371/journal.pone.0174007>.
- Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. *Nucl. Acids Res.* 33, 116–120. <https://doi.org/10.1093/nar/gki442>.
- Riehle, M.A., Fan, Y., Cao, C., Brown, M.R., 2006. Molecular characterization of insulin-like peptides in the yellow fever mosquito, *Aedes aegypti*: expression, cellular localization, and phylogeny. *Peptides* 27, 2547–2560. <https://doi.org/10.1016/j.peptides.2006.07.016>.
- Rivière, G., Klopp, C., Ibouniyamine, N., Huvet, A., Boudry, P., Favrel, P., 2015. GigaTON: an extensive publicly searchable database providing a new reference transcriptome in the pacific oyster *Crassostrea gigas*. *BMC Bioinf.* 16, 401. <https://doi.org/10.1186/s12859-015-0833-4>.
- Robitzki, A., Schröder, H.C., Ugarkovic, D., Pfeifer, K., Uhlenbruck, G., Müller, W.E., 1989. Demonstration of an endocrine signaling circuit for insulin in the sponge *Geodia cydonium*. *EMBO J.* 8, 2905–2909.
- Safavi-Hemami, H., Lu, A., Li, Q., Fedosov, A.E., Biggs, J., Showers Corneli, P., Seger, J., Yandell, M., Olivera, B.M., 2016. Venom insulins of cone snails diversify rapidly and track prey taxa. *Mol. Biol. Evol.* 33, 2924–2934. <https://doi.org/10.1093/molbev/msw174>.
- Saltiel, A.R., Kahn, C.R., 2001. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 414, 799–806. <https://doi.org/10.1038/414799a>.
- Schindler, A.J., Baugh, L.R., Sherwood, D.R., 2014. Identification of late larval stage developmental checkpoints in *Caenorhabditis elegans* regulated by insulin/IGF and steroid hormone signaling pathways. *PLoS Genet.* 10, 13–16. <https://doi.org/10.1371/journal.pgen.1004426>.
- Shabanpoor, F., Separovic, F., Wade, J.D., 2009. In: Chapter 1 The Human Insulin Superfamily of Polypeptide Hormones, 1st ed, Vitamins and Hormones. Elsevier Inc. [https://doi.org/10.1016/S0083-6729\(08\)00601-8](https://doi.org/10.1016/S0083-6729(08)00601-8).
- Shagin, D.A., Rebrikov, D.V., Kozhemyako, V.B., Altschuler, I.M., Shcheglov, A.S., Zhulidov, P.A., Bogdanova, E.A., Staroverov, D.B., Rasskazov, V.A., Lukanov, S., 2002. A novel method for SNP detection using a new duplex-specific nuclease from crab hepatopancreas. *Genome Res.* 12, 1935–1942. <https://doi.org/10.1101/gr.547002>.
- Sharabi, O., Manor, R., Weil, S., Aflalo, E.D., Lezer, Y., Levy, T., Aizen, J., Ventura, T., Mather, P.B., Khalaila, I., Sagi, A., 2016. Identification and characterization of an insulin-like receptor involved in crustacean reproduction. *Endocrinology* 157, 928–941. <https://doi.org/10.1210/en.2015-1391>.
- Shingleton, A.W., Das, J., Vinicius, L., Stern, D.L., 2005. The temporal requirements for insulin signaling during development in *Drosophila*. *PLoS Biol.* 3, 1607–1617. <https://doi.org/10.1371/journal.pbio.0030289>.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. <https://doi.org/10.1038/msb.2011.75>.
- Sliwowska, J.H., Fergani, C., Gawalek, M., Skowronska, B., Fichna, P., Lehman, M.N., 2014. Insulin: Its role in the central control of reproduction. *Physiol. Behav.* 133, 197–206. <https://doi.org/10.1016/j.physbeh.2014.05.021>.
- Smit, A.B., Vreugdenhil, E., Ebberink, R.H., Geraerts, W.P., Klootwijk, J., Jooisse, J., 1988. Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* 331, 535–538. <https://doi.org/10.1038/331535a0>.
- Smit, A.B., Geraerts, P.M., Meester, I., van Heerikhuizen, H., Jooisse, J., 1991. Characterization of a cDNA clone encoding molluscan insulin-related peptide II of *Lymnaea stagnalis*. *Eur. J. Biochem.* 199, 699–703.
- Smit, A.B., Thijsen, S.F., Geraerts, W.P., Meester, I., van Heerikhuizen, H., Jooisse, J., 1992. Characterization of a cDNA clone encoding molluscan insulin-related peptide V of *Lymnaea stagnalis*. *Brain Res. Mol. Brain Res.* 14, 7–12.
- Smit, A.B., van Marle, A., van Elk, R., Bogerd, J., van Heerikhuizen, H., Geraerts, W.P., 1993. Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*. *J. Mol. Endocrinol.* 11, 103–113.
- Smit, A.B., Spijker, S., Van Minnen, J., Burke, J.F., De Winter, F., Van Elk, R., Geraerts, W.P., 1996. Expression and characterization of molluscan insulin-related peptide VII from the mollusc *Lymnaea stagnalis*. *Neuroscience* 70, 589–596. [https://doi.org/10.1016/0306-4522\(95\)00378-9](https://doi.org/10.1016/0306-4522(95)00378-9).
- Smit, A.B., van Kesteren, R.E., Li, K.W., Van Minnen, J., Spijker, S., Van Heerikhuizen, H., Geraerts, W.P., 1998. Towards understanding the role of insulin in the brain: lessons from insulin-related signaling systems in the invertebrate brain. *Prog. Neurobiol.* 54, 35–54. [https://doi.org/10.1016/S0301-0082\(97\)00063-4](https://doi.org/10.1016/S0301-0082(97)00063-4).
- Söderberg, J.A.E., Birse, R.T., Nüssel, D.R., 2011. Insulin production and signaling in renal tubules of *drosophila* is under control of tachykinin-related peptide and regulates stress resistance. *PLoS One* 6. <https://doi.org/10.1371/journal.pone.0019866>.
- Steiner, D.F., Chan, S.J., Welsh, J.M., Kwok, S.C.M., 1985. Structure and evolution of the insulin gene. *Annu. Rev. Genet.* 19, 463–484. <https://doi.org/10.1146/annurev.ge.19.120185.002335>.
- Veenstra, J.A., 2010. Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen. Comp. Endocrinol.* 167, 86–103. <https://doi.org/10.1016/j.ygcen.2010.02.010>.
- Veenstra, J.A., Agrícola, H.-J., Sellami, A., 2008. Regulatory peptides in fruit fly midgut. *Cell Tissue Res.* 334, 499–516. <https://doi.org/10.1007/s00441-008-0708-3>.
- Ventura, T., Manor, R., Aflalo, E.D., Weil, S., Raviv, S., Glazer, L., Sagi, A., 2009. Temporal silencing of an androgenic gland-specific insulin-like gene affecting phenotypic gender differences and spermatogenesis. *Endocrinology* 150, 1278–1286. <https://doi.org/10.1210/en.2008-0906>.
- Ventura, T., Rosen, O., Sagi, A., 2011. From the discovery of the crustacean androgenic gland to the insulin-like hormone in six decades. *Gen. Comp. Endocrinol.* 173, 381–388. <https://doi.org/10.1016/j.ygcen.2011.05.018>.
- Wang, S., Luo, X., Zhang, S., Yin, C., Dou, Y., Cai, X., 2013. Identification of putative insulin-like peptides and components of insulin signaling pathways in parasitic platyhelminths by the use of genome-wide screening. *FEBS J.* 281, 877–893. <https://doi.org/10.1111/febs.12655>.
- Wilkinson, T.N., Speed, T.P., Tregear, G.W., Bathgate, R.A., Sherwood, O., Bathgate, R., Samuel, C., Burazin, T., Layfield, S., Claasz, A., Reyntomas, I., Dawson, N., Zhao, C., Bond, C., Summers, R., Parry, L., Wade, J., Tregear, G., Bullesbach, E., Schwabe, C., Bullesbach, E., Schwabe, C., Hansell, D., Bryant-Greenwood, G., Greenwood, F., Winslow, J., Shih, A., Bourell, J., Weiss, G., Reed, B., Stults, J., Goldsmith, L., Sherwood, O., Zimmermann, S., Steding, G., Emmen, J., Brinkmann, A., Nayernia, K., Holstein, A., Engel, W., Adham, I., Nef, S., Parada, L., Bieche, I., Laurent, A., Laurendeau, I., Duret, L., Giovannardi, Y., Frenedo, J., Olivieri, M., Fausser, J., Evain-Brion, D., Vidau, M., Conklin, D., Lofton-Day, C., Haldeman, B., Ching, A., Whitmore, T., Lok, S., Jaspers, S., Lok, S., Johnston, D., Conklin, D., Lofton-Day, C., Adams, R., Jelmberg, A., Whitmore, T., Schrader, S., Griswold, M., Jaspers, S., Hsu, S., Nakabayashi, K., Nishi, S., Kumagi, J., Kudo, M., Sherwood, O., Hsueh, A., Sudo, S., Kumagai, J., Nishi, S., Layfield, S., Ferraro, T., Bathgate, R., Hsueh, A., Kumagai, J., Hsu, S., Matsumi, H., Roh, J., Fu, P., Wade, J., Bathgate, R., Hsueh, A., Bogatcheva, N., Truong, A., Feng, S., Engel, W., Adham, I., Agoulinc, A., Liu, C., Eriste, E., Sutton, S., Chen, J., Roland, B., Kuei, C., Farmer, N., Jornvall, H., Sillard, R., Lovenberg, T., Liu, C., Chen, J., Sutton, S., Roland, B., Kuei, C., Farmer, N., Sillard, R., Lovenberg, T., Hsu, S., Kudo, M., Chen, T., Nakabayashi, K., Bhalla, A., van der Spek, P., van Duijn, M., Hsueh, A., Schwabe, C., Bullesbach, E., Heyn, H., Yoshioka, M., Schwabe, C., LeRoith, D., Thompson, R., Shiloach, J., Roth, J., Georges, D., Viguier-Martinez, M., Poirier, J., Georges, D., Schwabe, C., Roche, P., Crawford, R., Tregear, G., Klönisch, T., Froehlich, C., Tetens, F., Fischer, B., Hombach-Klonisch, S., Hsu, S., Bathgate, R., Scott, D., Chung, S., Ellyard, D., Garreffa, A., Tregear, G., Bathgate, R., Siebel, A., Tovote, P., Claasz, A., Macris, M., Tregear, G., Parry, L., de Rienzo, G., Aniello, F., Branno, M., Minucci, S., Bielawski, J., Yang, Z., Schwabe, C., Bullesbach, E., Schwabe, C., Gowan, L., Reinig, J., Schwabe, C., Warr, G., Dores, R., Rubin, D., Quinn, T., Nagasawa, H., Kataoka, H., Isogai, A., Tamura, S., Suzuki, A., Mizoguchi, A., Fujiwara, Y., Suzuki, A., Takahashi, S., Ishizaki, H., Brogiolo, W.,

- Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., Hafen, E., Pierce, S., Costa, M., Wisotzky, R., Devadhar, S., Homburger, S., Buchman, A., Ferguson, K., Heller, J., Platt, D., Pasquinelli, A., Liu, L., Doberstein, S., Ruvkan, G., Hombach-Klonisch, S., Abd-Elnaeim, M., Skidmore, J., Leiser, R., Fischer, B., Klonisch, T., Bravo, P., Stewart, D., Lasley, B., Fowler, M., Abd-Elnaeim, M., Saber, A., Hassan, A., Abou-Elmagd, A., Klisch, K., Jones, C., Leiser, R., Brackett, K., Fields, P., Dubois, W., Chang, S., Peer de, Y. Van, Taylor, J., Meyer, A., Bullesbach, E., Schwabe, C., Callard, I., Steinetz, B., Schwabe, C., Callard, I., Goldsmith, L., Bullesbach, E., Schwabe, C., Lacy, E., Reinig, J., Daniel, L., Schwabe, C., Gowen, L., Steinetz, B., O'Byrne, E., Suzuki, Y., Nei, M., Suzuki, Y., Nei, M., Zhang, J., Wyckoff, G., Wang, W., Wu, C., Ting, C., Tsaur, S., Wu, M., Wu, C., Swanson, W., Clark, A., Waldrup-Dail, H., Wolfner, M., Aquadro, C., Rooney, A., Zhang, J., Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., Lipman, D., Thompson, J., Higgins, D., Gibson, T., Felsenstein, J., Schmidt, H., Strimmer, K., Vingron, M., Haeseler, A., Page, R. von, Page, R., Hedges, S., Yang, Z., Nielsen, R., Yang, Z., 2005. Evolution of the relaxin-like peptide family. *BMC Evol. Biol.* 5, 14. <https://doi.org/10.1186/1471-2148-5-14>.
- Yoshida, I., Tsuzuki, S., Abdel Salam, S.E., Ino, M., Korayem, A.M., Sakurai, S., Iwami, M., 1997. Bombyxin F1 Gene: structure and Expression of a New Bombyxin Family Gene That Forms a Pair with Bombyxin B10 Gene. *Zool. Sci.*
- Yoshida, I., Moto, K., Sakurai, S., Iwami, M., 1998. A novel member of the bombyxin gene family: structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkworm *Bombyx mori*. *Dev. Genes Evol.* 208, 407–410. <https://doi.org/10.1007/s004270050197>.
- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Wang, J., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P.W.H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Lošo, T., Du, Y., Sun, X., Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Y., Li, N., Wang, J., Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P.M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, J., Wang, Q., Steinberg, C.E.W., Wang, H., Li, N., Qian, L., Zhang, G., Li, Y., Yang, H., Liu, X., Yin, Y., Wang, J., 2012a. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49–54. <https://doi.org/10.1038/nature11413>.
- Zhang, L., Hou, R., Su, H., Hu, X., Wang, S., Bao, Z., 2012b. Network analysis of oyster transcriptome revealed a cascade of cellular responses during recovery after heat shock. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0035484>.
- 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 Genes|Genomes|Genetics* 4, 2207–2217. <https://doi.org/10.1534/g3.114.013904>.
- Zhao, X., Yu, H., Kong, L., Li, Q., 2012. Transcriptomic responses to salinity stress in the pacific oyster *Crassostrea gigas*. *PLoS One* 7. <https://doi.org/10.1371/journal.pone.0046244>.
- Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., Siebert, P.D., 2001. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques* 30, 892–897. <https://doi.org/10.1126/science.aam8999>.
- Zhulidov, P.A., Bogdanova, E.A., Shcheglov, A.S., Vagner, L.L., Khaspekov, G.L., Kozhemyako, V.B., Matz, M.V., Meleshkevitch, E., Moroz, L.L., Lukyanov, S.A., Shagin, D.A., 2004. Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucl. Acids Res.* 32. <https://doi.org/10.1093/nar/gnh031>. 37e–37.
- Zatylny-Gaudin, C., Cornet, V., Leduc, A., Zanuttini, B., Corre, E., Le Corguillé, G., Bernay, B., Garderes, J., Kraut, A., Couté, Y., Henry, J., 2016. Neuropeptidome of the cephalopod *sepia officinalis*: identification, tissue mapping, and expression pattern of neuropeptides and neurohormones during egg laying. *J. Proteome Res.* 15, 48–67. <https://doi.org/10.1021/acs.jproteome.5b00463>.

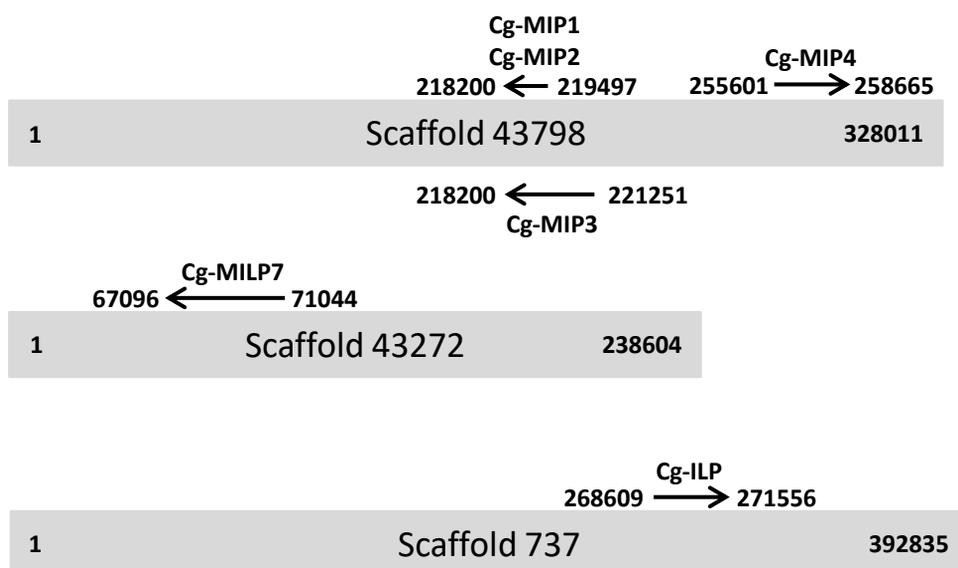


Figure S1: Schematic representation of the *C. gigas* IRPs location on scaffolds obtained using WEBSCIPIO.

A - Sequence similarity strategy		
Name	Species	Amino acid sequence
Insulin 1	<i>Mus musculus</i>	MALLVHFLPLLALLALWEPKPTQAFVKQHLGPHLVEALYLVCGERGFYTPKSRREVEDPQVQLELGGSPGDLQTLALEVARQK RGIVDQCCTISCSLYQLENYCN
Insulin 2	<i>Mus musculus</i>	MALWMRFLPLLALLFLWESHPTQAFVKQHLGSHLVEALYLVCGERGFYTPMSRREVEDPQVAQLELGGSPGAGDLQTLALEV AQQRKRVGDQCCTISCSLYQLENYCN
IGF I	<i>Mus musculus</i>	MGKISSLPTLQFKICDFLKIKHIMSSSHLFYALCLLTFSTTAGPETLCAEALVDALQFVCGPRGFYNKPTGYGSSIRRAPQTGI VDECCFRSCLRLREMYCAPLKPTKAARSIRAQRHTDMPKTKQSPSLSTNKTKLQRRRKGSTFEHK
IGF II	<i>Mus musculus</i>	MGGSVAGFQVPMGIPVGKMLVLLISLAFALCIAAYGPGETLCGGELVDLQFVCSDRGFYFSPSSRANRRSRGIVEECFRSCD LALLEYCATPAKSERDVSQAVLPDFFRYVGVKFFQYDTRWQSAAGLRRLRLLARRRRMLAKELKEFREAKRHRPLVLP KDPAHGGASSEMSNHQ
DILP1	<i>Drosophila melanogaster</i>	MFSQHNGAAVHGLRLQSLIAAMLTAAMAMVPTGSGHQLLPPGNHKLCPALSDAMDVVCPHGFNTLPKRRESLLGNSSDDE DTEQEVQDSSMWQLTDGAGYSFPLTLNYGSEVLKMRHRRLTGGVYDECCVKTCSYLELAIYCLPK
DILP2	<i>Drosophila melanogaster</i>	MSKPLSFISMVAVILLASSTVKAQGLTCEKLNELVSMVCEEYNPVPHKRAMPGADSDLDALNPLQFVQEEEDNSISEPLRSAL FPGSYLGGVLSNAEVRRTQRQGIIVERCKKSCDMKALREYCSVVRN
DILP3	<i>Drosophila melanogaster</i>	MGIEMRCQDRILLPSLLLLMIGGVQATMKLGRKLPETLSKLCVGFNAMTKRTLDPVNFNQDGFEDRSLLERLSSDVSQML KTRRLRDGVFDECCLSCTMDEVLRVYCAAKPRT
DILP4	<i>Drosophila melanogaster</i>	MSLIRLGLALLLTLVSLQQLPQVGRKRMCGEALIQALDVICVNGFTRRRVRSASAKDARVRLIRKLPQDPEDEIQTETGRKQK HTDADTEKGVPPAVGSGRKLRRHRIAHECCKEGCTYDDILDYCA
DILP5	<i>Drosophila melanogaster</i>	MFRSVIPLFLIPLLSQAANSRLACGPALMDLVRVACPNFNGNSMFAKRGTLGLFDYEDHLADDSSESHMNSLSSIRDFRG VVDSCRCSCSFSTRAYCDS
DILP6	<i>Drosophila melanogaster</i>	MVLKVPSTKVVLLVATLFAVAAMISSWMPQVAASPLAPTEYEQRRMCMCTGSDVGIQIKVCGVVALGDVFPNFGKRRKRLQ VTDLCKSGGCTYRELLQYCKG
DILP7	<i>Drosophila melanogaster</i>	MTRMIQNSGSWTLGAVLFLPLIPTPEALQHTEEGLEMLFRERSQSDWENVWHQETHSRCDKLRVQLYWACEKIYLRTR NKKRTGNDDEAWIKTTTEPDGSTWLHVNYANMFLRSRRSDGNTPSISNECTKAGCTWEEYAEYCPNKRNNHY
DILP8	<i>Drosophila melanogaster</i>	MSSKLMCRWMLLVIGVCLMGSSGSFCLERMKFKAMEACEHLFQADEGARRRRSIEFAHHHLNRLSGSKGTHNKHHYISRS SYPMGGYLVKTRHFNRLSELDIFPRYKPIKPHHEKHRFRKDRHSSRSYNNIPYCCNLQCEEEFFC
MIP I	<i>Lymnaea stagnalis</i>	MAGVRLVFTKAFMVTVLTLLNIGVKAPEGQFSACNINDRPHRRGVCSALADLVDFACSSNQPMVKRNETDLDPLRNLIK LSSESALTYLTKRQGTINIVCECMKPTLSLRQYCP
MIP II	<i>Lymnaea stagnalis</i>	MVGVRVLTNAFVTVLTLTLLDVVVKPAEQSSCSLSSRPHRPGICGSNLAGFRACFCSNQNSPMVKRDAETGWLLPETMVKR NAETDLDLDRNLIKLSSESALTYLTKRQRTNLVCECCFNCTPDVVRKYCY
MIP III	<i>Lymnaea stagnalis</i>	MASVHLTLTKAFMVTVLTLLNIVSTRGTTQHTCSILSRPHRGLCGSTLANMVQWLCSTYTTSSKVKRQAEPEDEDDAMSKIMIS KKRALSYLTKRESRPSIVCECCFNCTVQELLAYC
MIP V	<i>Lymnaea stagnalis</i>	MAGVRLVFTKAFMVTVLTLLNIGVKAPEGQFSACFSSRPHRPGICGSADLADRFACSRNQPAMVKRDAETGWLLPETMVKR NAQTDLDLDRNLIKLSSESALTYLTKRQRTNLVCECCYNCTVDVFEYCY
MIP VII	<i>Lymnaea stagnalis</i>	MNASVESCLTFTFVLVALCVGLTIGQVNTCTMFSRQHPRLCGNRLARAHANLCLFRNTYDIFPRKRSVDNTEFKVYSIPLSVL AELDLSDDWVAGYVSKKIDIPYRSETNGLSGANFESSAFDKLELPAMKSTTSQFRILKRLGRSLKREVMAPESLVDCDCCYNECSVR KLATYC

B - Phylogenetic tree		
Name	Species	Amino acid sequence
Insulin	<i>Homo sapiens</i>	MALWMRLLPLLALLALWGPDAAFVFNQHLGSHLVEALYLVCGERGFYTPKTRREAEDLQVQVLELGGSPGAGSLQPLALEG SLQQRKRVGEQCTISCSLYQLENYCN
DILP7	<i>Drosophila melanogaster</i>	MTRMIQNSGSWTLGAVLFLPLIPTPEALQHTEEGLEMLFRERSQSDWENVWHQETHSRCDKLRVQLYWACEKIYLRTR NKKRTGNDDEAWIKTTTEPDGSTWLHVNYANMFLRSRRSDGNTPSISNECTKAGCTWEEYAEYCPNKRNNHY
Cg-ILP	<i>Crassostrea gigas</i>	MKWLVSVLVCLRVGVGSQQAQGSALTDILSLVCRNQFHAPAKRIDAPMVDNDLDELARYKRSRQFAVAQKRGVVEECCFSSCS YENLLYCSQSDPLDVFTAARVGSSTTQRPIITTTSTPTATTTPLSVLTVGKHPEDQAELEKAAAAAVALSRWFVYVNLGR FRGQVGRRIIM
Cg-MIP1	<i>Crassostrea gigas</i>	MTSETTWFIISLCLLQMVCPVLSGFKEVCTFETRYRRGVHQQACGDNLDMLRLVCRKYSRGGTRPKDYVLRANFTAGGLS PSQIQARSAPWKSVLKRMVMSEKALAFIANNLDLFRHRKKGSPYNLDKRYGDIINVCCEYHSCVAEFEDYCAE
Cg-MIP2	<i>Crassostrea gigas</i>	MTSETTWFIISLCLLQMVCPVLSGFKEVCTFETRYRRGVHQQACGDNLDMLRLVCRKYSRGGTRPKDYVLRANFTAGGLS PPQIQARSAPWKSVLKRMVMSEKALAFIANNLDLFRHRKKGSPYNLDKRYGDIINVCCEYHSCVAEFEDYCAE
Cg-MIP3	<i>Crassostrea gigas</i>	MLMPFCIDASKSNKFCYANNAGYKKTURMTSETTWFIISLCLLQMVCPVLSGFKEVCTFETRYRRGVHQQACGDNLDMLRL VCRKYSRGGTRPKDYVLRANFTAGGLSPPQIQARSAPWKSVLKRMVMSEKALAFIANNLDLFRHRKKGSPYNLDKRYGDIINVC ECCYHSCVAEFEDYCAE
Cg-MIP4	<i>Crassostrea gigas</i>	MSCRNWCNSWEIMVFNKRNKHQIFILLCHFTSVQSDFERVCNSQDRLRGPDPQIGCRILPEMLHLVCGGQYVPSKRDVSSLSHH KQDRNVDFPRYSPLGELIUKREASMYLTSQHSRTKRNAYQIVCECCYHGCNWFELQYCGFRKRNTEPDISASSQNSGKLI SVLNLK
Cg-MILP7	<i>Crassostrea gigas</i>	MVLVGLSLWTFSWIFTQVKTAAIHQYQSNPVLQRTAEAWRSLWHNDCHRVCHFELDQHDVCLACRMDIYRKRSDVAEEPRIS NKYETDNQNTKRIFIDKSTSMKFLSPDVSHSIRKRNVLNECCYKSGCSWEEFAEFCQJRLPATNANSCVS
MIP I	<i>Lymnaea stagnalis</i>	MAGVRLVFTKAFMVTVLTLLNIGVKAPEGQFSACNINDRPHRRGVCSALADLVDFACSSNQPMVKRNETDLDPLRNLIK LSSESALTYLTKRQGTINIVCECMKPTLSLRQYCP
MIP II	<i>Lymnaea stagnalis</i>	MVGVRVLTNAFVTVLTLTLLDVVVKPAEQSSCSLSSRPHRPGICGSNLAGFRACFCSNQNSPMVKRDAETGWLLPETMVKR NAETDLDLDRNLIKLSSESALTYLTKRQRTNLVCECCFNCTPDVVRKYCY
MIP III	<i>Lymnaea stagnalis</i>	MASVHLTLTKAFMVTVLTLLNIVSTRGTTQHTCSILSRPHRGLCGSTLANMVQWLCSTYTTSSKVKRQAEPEDEDDAMSKIMIS KKRALSYLTKRESRPSIVCECCFNCTVQELLAYC
MIP V	<i>Lymnaea stagnalis</i>	MAGVRLVFTKAFMVTVLTLLNIGVKAPEGQFSACFSSRPHRPGICGSADLADRFACSRNQPAMVKRDAETGWLLPETMVKR NAQTDLDLDRNLIKLSSESALTYLTKRQRTNLVCECCYNCTVDVFEYCY
MIP VII	<i>Lymnaea stagnalis</i>	MNASVESCLTFTFVLVALCVGLTIGQVNTCTMFSRQHPRLCGNRLARAHANLCLFRNTYDIFPRKRSVDNTEFKVYSIPLSVL AELDLSDDWVAGYVSKKIDIPYRSETNGLSGANFESSAFDKLELPAMKSTTSQFRILKRLGRSLKREVMAPESLVDCDCCYNECSVR KLATYC
Ls-MILP7	<i>Lymnaea stagnalis</i>	MAGVRLVFTKAFMVTVLTLLNIGVKAPEGQFSACFSSRPHRPGICGSADLADRFACSRNQPAMVKRDAETGWLLPETMVKR NAQTDLDLDRNLIKLSSESALTYLTKRQRTNLVCECCYNCTVDVFEYCY
Insulin	<i>Aplysia californica</i>	MSKFLLQSHSANACLLTLLASNLDISLANFEHSCNGYMRPHRPLCGLEDLHVHISNLCSLGGNRRFLAKYMKVDRDTEVNDKL RGILLNKEAFSFKREASGISTCECCFNCRIFELAQYCRPLPDHFRSIRSTRGNSGHAQLLEDNFS
Ac-MILP7	<i>Aplysia californica</i>	MATSRFQTSARPELIEIWHDCRRRCRWQLIRHVEIACQFDPVYQSKRRKSPMKLAPSREHEQVKVIAVLRNMTIPNDRDPQT TSPQSDTRPPIPQNFPLPKRATGFLTSHGGRPGGGRKRLRRKRGIMTECECFKPAWEEFAEYCHDNRRRLSERDNQCT
Insulin1	<i>Lottia gigantea</i>	MAESVWMLKLLSLCIVGTPVAFYERSCNLERQQEDPRGLCGSQLESILHNVCFYLNKRASPKVGPVERVKRMDQSKLDFKFSALS FLVLRKGDGFGNTGITCECCNLRNLISSYCYL
Insulin2	<i>Lottia gigantea</i>	MEVTKCPLVLGVFLNFGTVLTHLEWTCTLETKRSPRGVCGORLPEVLSMCKRYGGYRDTWFRKRNREGTNSRNLNIGLGR DAFSYLKGRGQSYGEGQITCECCYHSCFRRELRYCRNSQQRISIKK
Insulin3	<i>Lottia gigantea</i>	MKTYQLVALAIFLSGTTIEQAEPERLGRNLANALDLCVGRGFHVEVSKRSDINHPAKRRNRRMIDVFDKRNHYHRRGVVEE CCYIGCTYETLESYCAEPGDEALVRPGRNFRLLMGRITTRTPVAATPRPPVVTTPGVHFSAVFNSHPRPGRSKRYFYVQLGTRP TRSPLI
Insulin4	<i>Lottia gigantea</i>	MERNGITFKLQINVLSTLIIHLTKLVQTSNGNDLLVQAFDRFTQASPEDLLALWHTDCHRRCSRLIEHVNIACIFDQYKRYK RSVDNQNSNKRNLRIDKSRATFLATLADTRGRKRGHSIKKRGVMECCYHSCFRRELRYCRNSQQRISIKK
Insulin	<i>Sepia officinalis</i>	MKSTVCMGTFLATLSIVNWQVNVAGLEHTCNEETIRQGAQGAHCGVEIPNLQLLCPAGYNERMSDRQRNLPVTSRAIG RRGNGLRDIISKQAKSYLTKRDRNWTGIVCECCYNKILELDDYCKDPSYFSKQLRS

Table S1: Sequences used for the identification of *C. gigas* IRPs (A) and for the phylogenetic tree (B). For each sequence, the name, species and amino acid sequence are provided.

Transcript name	Forward (5' - 3')	Reverse (5' - 3')
qPCR		
Cg-mip123	CAAGATCCGCTCCCTGGAAG	TCCTTTCTTCTTGCGGCGAT
Cg-ilp	TCGGGAGGCGCATAATCAAT	GCCGTGTCTGCAATACAAGC
Cg-mip4	CATAACGACTGTCACCGGGT	TTCCTCCGCTGCATCAACTG
EF1 α	ACGACGATCGCATTCTCTT	ACCACCCTGGTGAGATCAAG
ISH		
Cg-mip123	ATTCAAAGGACTTTGCTCCCG	TTGCCTGTCCCATGTCTTCC
Cg-ilp	AGTGGCTCAGAAAAGAGGGG	ATGATCGCGTGAGACACAAAAC
Cg-mip4	CATAACGACTGTCACCGGGT	TTCCTCCGCTGCATCAACTG

Table S2: List of the forward and reverse primers of each sequence used for qPCR and *in situ* hybridization experiments.

A	Insulin	DILP7	Cg-MIP1	Cg-MIP2	Cg-MIP3	Cg-MIP4	Cg-ILP	Cg-MILP7
Insulin	100%	18,68%	26,42%	26,42%	27,45%	26,60%	29,55%	17,20%
DILP7		100%	18,18%	18,18%	17,24%	23,57%	15,56%	25%
Cg-MIP1			100%	98,14%	98,73%	32,84%	27,66%	18,55%
Cg-MIP2				100%	99,36%	32,84%	26,60%	18,55%
Cg-MIP3					100%	31,94%	28,89%	17,52%
Cg-MIP4						100%	21,93%	18,75%
Cg-ILP							100%	20,41%
Cg-MILP7								100%

B	DILP1	DILP2	DILP3	DILP4	DILP5	DILP6	DILP7	DILP8
Cg-MIP1	18,25%	17,69%	21,82%	19,23%	22,33%	20%	19,53%	16,79%
Cg-MIP2	18,25%	17,69%	21,82%	19,23%	23,30%	20%	19,53%	16,79%
Cg-MIP3	17,69%	18,25%	21,82%	19,05%	22,77%	19,23%	19,08%	17,32%
Cg-MIP4	18,70%	17,70%	20,21%	19,47%	15,05%	18,95%	20,80%	17,59%
Cg-ILP	25,77%	25,77%	29,35%	27,37%	31,18%	21,74%	16,16%	14,67%
Cg-MILP7	16,38%	13,21%	18,37%	19,81%	16,09%	17,71%	25%	13,59%

Table S3: Matrix of the identity percentages between insulin pre-propeptide sequences. A: IRPs of *C. gigas* (Cg-MIP1 to Cg-MIP4, Cg-ILP and Cg-MILP7) are compared with IRP of *H. sapiens* (Insulin) and DILP7 of *D. melanogaster*. B: IRPs of *C. gigas* are compared with the DILPs of *D. melanogaster* (DILP1 to DILP8).

Data Article

Title: *Data for evolutive analysis of insulin related peptides in bilaterian species*

Authors: Maëva Cherif--Feildel, Clothilde Heude Berthelin, Guillaume Rivière, Pascal Favrel and Kristell Kellner

Affiliations: Normandy University, Caen, France. University of Caen Normandie, Unity Biology of organisms and aquatic ecosystems (BOREA), MNHN, Sorbonne University, UCN, CNRS, IRD, Esplanade de la Paix, 14032 Caen, France

Contact email: kristell.kellner@unicaen.fr

Abstract

In bilaterian species, the amino acid sequence conservation between Insulin related peptides is relatively low except for the cysteine residues involved in the disulphide bonds. In the A chain, the conserved cystein residues are included in a signature motif. Investigating the variations in this motif would give insight into the phylogenetic history of the family. The table presented in this paper contains a large set of insulin-related peptides in bilateral phylogenetic groups (deuterostomian, ecdysozoan, lophotrochozoan). NCBI databases *in silico* wide screening combined with bibliographic researches provided a framework for identifying and categorising the structural characteristics of these insulin related peptides. The dataset includes NCBI IDs of each sequence with hyperlinks to FASTA format. Moreover, the structural type (α , β or γ), the A chain motif, the total number of cysteins, the C peptide cleavage mode and the potential additional domains (D or E) are specified for each sequence. The data are associated with the research article “Molecular evolution and functional characterisation of insulin-related peptides in molluscs: contributions of *Crassostrea gigas* genomic and transcriptomic-wide screening” [1]. The table presented here can be found at <http://dx.doi.org/10.17632/w4gr8zcpk5.4#file-21c0f6a5-a3e3-4a15-86e0-e5a696458866>.

Specifications Table

Subject area	<i>Biology</i>
More specific subject area	<i>Insulin</i>
Type of data	<i>Table (online data)</i>
How data was acquired	<i>Screening in NCBI's databases</i>
Data format	<i>Analysed data</i>
Experimental factors	<i>In silico wide screening and bibliographic researches</i>
Experimental features	<i>Research of insulin related peptides was done in NCBI's databases and identification of each domain was determined according to insulin structural characteristics and bibliography.</i>
Data source location	<i>Data were collected on NCBI and in cited bibliography.</i>
Data accessibility	Data is available via the following link: http://dx.doi.org/10.17632/w4gr8zcpk5.4#file-21c0f6a5-a3e3-4a15-86e0-e5a696458866
Related research article	Cherif--Feidel, M., Heude Berthelin, C., Adeline, B., Rivière, G., Favrel, P., Kellner, K., 2018. Molecular evolution and functional characterisation of Insulin Related Peptides in molluscs: contributions of Crassostrea gigas genomic and transcriptomic-wide screening. Gen. Comp. Endocrinol. 0–1. https://doi.org/10.1016/j.ygcen.2018.10.019 [1]

Data Value

- These data provide a database for the structural characteristics of a large set of insulin related peptides in deuterostomian, ecdysozoan and lophotrochozoan species.
- The data may be used to investigate the conservation of structural motifs in bilaterian insulin related peptides.
- The data could give a basis for further phylogenetic analyses.

Data

The dataset provides a list of hyperlinks to insulin-related peptides NCBI ID originating from various metazoan species (referenced as [1] to [49]). For each sequence, the key motif of the A chain associated to the number and position of disulphide bonds was specified. The structural type, the total number of cysteins, the cleavage of C peptide and the possible additional domains were also indicated. The data is available at <http://dx.doi.org/10.17632/w4gr8zcpk5.4#file-21c0f6a5-a3e3-4a15-86e0-e5a696458866>.

Experimental Design, Materials, and Methods

All sequences were identified either using wide screening of NCBI sequence databases or on the basis of bibliographic analysis (for NCBI unreferenced sequences). Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) was used to help identify domains, cleavage sites

and conserved cysteins. The structural type α , β or γ was determined according to Matsunaga et al. [2] for *Caenorhabditis elegans* insulin-related peptides. The C peptide cleavage mode and the presence of additional domains D (uncleaved) or E (cleaved) resulted from this information. The variations in the A chain signature motif characterising the family [50] were specified for each sequence.

Acknowledgments

PhD of M. Cherif-Feidel was financially supported by the French research ministry (PhD grant). Proofreading of the English text has been realised by C. Quint of the centre of translation of UNICAEN.

References

- [1] Cherif-Feidel, M., Heude Berthelin, C., Adeline, B., Rivière, G., Favrel, P., Kellner, K., 2018. Molecular evolution and functional characterisation of Insulin Related Peptides in molluscs: contributions of *Crassostrea gigas* genomic and transcriptomic-wide screening. *Gen. Comp. Endocrinol.* 0–1. <https://doi.org/10.1016/j.ygcen.2018.10.019>
- [2] Shabanpoor, F., Separovic, F., Wade, J.D., 2009. Chapter 1 The Human Insulin Superfamily of Polypeptide Hormones, 1st ed, *Vitamins and Hormones*. Elsevier Inc. [https://doi.org/10.1016/S0083-6729\(08\)00601-8](https://doi.org/10.1016/S0083-6729(08)00601-8)
- [3] Lecroisey, C., Le Pétilon, Y., Escriva, H., Lammert, E., Laudet, V., 2015. Identification, evolution and expression of an insulin-like peptide in the cephalochordate branchiostoma lanceolatum. *PLoS One* 10, 1–15. <https://doi.org/10.1371/journal.pone.0119461>
- [4] Leroith, D., Roberts, C.T., 1993. Phylogeny of the Insulin-Like Growth Factors (IGFs) and Receptors : A Molecular Approach 338, 332–338.
- [5] Wilkinson, T.N., Speed, T.P., Tregear, G.W., Bathgate, R.A., Sherwood, O., Bathgate, R., Samuel, C., Burazin, T., Layfield, S., Claasz, A., et al., 2005. Evolution of the relaxin-like peptide family. *BMC Evol. Biol.* 5, 14. <https://doi.org/10.1186/1471-2148-5-14>

- [6] Hahn, V., Winkler, J., Rapoport, T.A., Liebscher, D.H., Coutelle, C., Rosenthal, S., 1983. Carp preproinsulin cDNA sequence and evolution of insulin genes. *Nucleic Acids Res.* 11, 4541–52.
- [7] Papasani, M.R., Robison, B.D., Hardy, R.W., Hill, R.A., 2006. Early developmental expression of two insulins in zebrafish (*Danio rerio*). *Physiol. Genomics* 27, 79–85. <https://doi.org/10.1152/physiolgenomics.00012.2006>
- [8] Wang, D.S., Jiao, B., Hu, C., Huang, X., Liu, Z., Cheng, C.H.K., 2008. Discovery of a gonad-specific IGF subtype in teleost. *Biochem. Biophys. Res. Commun.* 367, 336–341. <https://doi.org/10.1016/j.bbrc.2007.12.136>
- [9] Hrytsenko, O., Pohajdak, B., Wright, J.R., 2016. Ancestral genomic duplication of the insulin gene in tilapia: An analysis of possible implications for clinical islet xenotransplantation using donor islets from transgenic tilapia expressing a humanized insulin gene. *Islets* 8, 97–103. <https://doi.org/10.1080/19382014.2016.1187352>
- [10] Chen, M.H.C., Lin, G.H., Gong, H.Y., Weng, C.F., Chang, C.Y., Wu, J.L., 2001. The characterization of prepro-insulin-like growth factor-1 Ea-2 expression and insulin-like growth factor-1 genes (devoid 81 bp) in the zebrafish (*Danio rerio*). *Gene* 268, 67–75. [https://doi.org/10.1016/S0378-1119\(01\)00433-4](https://doi.org/10.1016/S0378-1119(01)00433-4)
- [11] Yuan, Y., Hong, Y., 2017. Medaka insulin-like growth factor-2 supports self-renewal of the embryonic stem cell line and blastomeres in vitro. *Sci. Rep.* 7, 78. <https://doi.org/10.1038/s41598-017-00094-y>
- [12] Plisetskaya, E., Pollock, H.G., Rouse, J.B., Hamilton, J.W., Kimmel, J.R., Gorbman, A., 1985. Characterization of coho salmon (*Oncorhynchus kisutch*) insulin. *Regul. Pept.* 11, 105–116. [https://doi.org/10.1016/0167-0115\(85\)90071-0](https://doi.org/10.1016/0167-0115(85)90071-0)

- [13] Mommsen, T.P., Silverstein, J.T., Plisetskaya, E.M., Whittaker, L.J., Whittaker, J., Conlon, J.M., 2001. Two insulins from channel catfish: Purification, structures, receptor-binding and cDNA sequences. *Fish Physiol. Biochem.* 25, 61–70. <https://doi.org/10.1023/A:1019708017696>
- [14] Lance, V.A., Hamilton, J.W., Rouse, J.B., Kimmel, J.R., Pollack, H.G., 1984. Isolation and characterization of reptilian insulin, glucagonistic, and pancreatic polypeptide: complete amino acid sequence of alligator (*Alligator mississippiensis*) insulin and pancreatic polypeptides. *Can. J. Zool* 61, 1744–1751.
- [15] Peterson, J.D., Steiner, D.F., Emdin, S.O., Falkmer, S., 1975. The amino acid sequence of the insulin from a primitive vertebrate, the Atlantic hagfish (*Myxine glutinosa*). *J. Biol. Chem.* 250, 5183–5191.
- [16] Olinski, R.P., Dahlberg, C., Thorndyke, M., Hallböök, F., 2006. Three insulin-relaxin-like genes in *Ciona intestinalis*. *Peptides* 27, 2535–2546. <https://doi.org/10.1016/j.peptides.2006.06.008>
- [17] Li, W., Kennedy, S.G., Ruvkun, G., 2003. *daf-28* encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. *Genes Dev.* 17, 844–858. <https://doi.org/10.1101/gad.1066503>
- [18] Matsunaga, Y., Iwasaki, T., Kawano, T., 2017. Diverse insulin-like peptides in *Caenorhabditis elegans* 1, 1–15.
- [19] Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., Liu, L.X., Doberstein, S.K., Ruvkun, G., 2001. Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* 672–686. <https://doi.org/10.1101/gad.867301.4>

- [20] Riehle, M.A., Fan, Y., Cao, C., Brown, M.R., 2006. Molecular characterization of insulin-like peptides in the yellow fever mosquito, *Aedes aegypti*: Expression, cellular localization, and phylogeny. *Peptides* 27, 2547–2560. <https://doi.org/10.1016/j.peptides.2006.07.016>
- [21] Brogiolo, W., Stocker, H., Ikeya, T., Rintelen, F., Fernandez, R., Hafen, E., 2001. An evolutionarily conserved function of the drosophila insulin receptor and insulin-like peptides in growth control. *Curr. Biol.* 11, 213–221. [https://doi.org/10.1016/S0960-9822\(01\)00068-9](https://doi.org/10.1016/S0960-9822(01)00068-9)
- [22] Nasonkin, I.O., Alikasifoglu, A., Barrette, T., Cheng, M.M., Thomas, P.M., Nikitin, A.G., 2002. Cloning, characterization, and embryonic expression analysis of the *Drosophila melanogaster* gene encoding insulin/relaxin-like peptide. *Biochem. Biophys. Res. Commun.* 295, 312–318. [https://doi.org/10.1016/S0006-291X\(02\)00653-8](https://doi.org/10.1016/S0006-291X(02)00653-8)
- [23] Sajid, W., Kulahin, N., Schluckebier, G., Ribel, U., Henderson, H.R., Tatar, M., Hansen, B.F., Svendsen, A.M., Kiselyov, V. V., Nørgaard, P., Wahlund, P.O., Brandt, J., Kohanski, R.A., Andersen, A.S., De Meyts, P., 2010. Structural and biological properties of the drosophila insulin-like peptide 5 show evolutionary conservation. *J. Biol. Chem.* 286, 661–673. <https://doi.org/10.1074/jbc.M110.156018>
- [24] Garelli, A., Gontijo, A.M., Miguela, V., Caparros, E., Dominguez, M., 2012. Imaginal Discs Secrete Insulin-Like Peptide 8 to Mediate Plasticity of Growth and Maturation. *Science* (80-). 336, 579–582. <https://doi.org/10.1126/science.1216735>
- [25] Mizoguchi, A., Okamoto, N., 2013. Insulin-like and IGF-like peptides in the silkworm *Bombyx mori*: Discovery, structure, secretion, and function. *Front. Physiol.* 4 AUG, 1–11. <https://doi.org/10.3389/fphys.2013.00217>
- [26] Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., Iwami, M., 1996. Multiple gene copies for Bombyxin, an insulin-related peptide of the silkworm *Bombyx mori*: Structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of Bombyxin. *J. Mol. Biol.* 259, 926–937. <https://doi.org/10.1006/jmbi.1996.0370>

- [27] Maruyama, K., Hietter, H., Nagasawa, H., Isogai, A., Tamura, S., Suzuki, A., Ishizaki, H., 1988. Isolation and primary structure of bombyxin-iv, a novel molecular species of bombyxin from the silkworm, *bombyx mori*. *Agric. Biol. Chem.* 52, 3035–3041. <https://doi.org/10.1080/00021369.1988.10869178>
- [28] Yoshida, I., Tsuzuki, S., Abdel Salam, S.E., Ino, M., Korayem, A.M., Sakurai, S., Iwami, M., 1997. Bombyxin F1 Gene: Structure and Expression of a New Bombyxin Family Gene That Forms a Pair with Bombyxin B10 Gene. *Zoolog. Sci.*
- [29] Yoshida, I., Moto, K., Sakurai, S., Iwami, M., 1998. A novel member of the bombyxin gene family: Structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkworm *Bombyx mori*. *Dev. Genes Evol.* 208, 407–410. <https://doi.org/10.1007/s004270050197>
- [30] Lagueux, M., Lwoff, L., Meister, M., Goltzené, F., Hoffmann, J.A., 1990. cDNAs from neurosecretory cells of brains of *Locusta migratoria* (Insecta, Orthoptera) encoding a novel member of the superfamily of insulins. *Eur. J. Biochem.* 187, 249–254. <https://doi.org/10.1111/j.1432-1033.1990.tb15302.x>
- [31] Ventura, T., Manor, R., Aflalo, E.D., Weil, S., Raviv, S., Glazer, L., Sagi, A., 2009. Temporal silencing of an androgenic gland-specific insulin-like gene affecting phenotypical gender differences and spermatogenesis. *Endocrinology* 150, 1278–1286. <https://doi.org/10.1210/en.2008-0906>
- [32] Ventura, T., Rosen, O., Sagi, A., 2011. From the discovery of the crustacean androgenic gland to the insulin-like hormone in six decades. *Gen. Comp. Endocrinol.* 173, 381–8. <https://doi.org/10.1016/j.ygcn.2011.05.018>

- [33] Chandler, J.C., Gandhi, N.S., Mancera, R.L., Smith, G., Elizur, A., Ventura, T., 2017. Understanding insulin endocrinology in decapod crustacea: Molecular modelling characterization of an insulin-binding protein and insulin-like peptides in the eastern spiny lobster, *Sagmariasus verreauxi*. *Int. J. Mol. Sci.* 18, 1–19. <https://doi.org/10.3390/ijms18091832>
- [34] Hamano, K., Awaji, M., Usuki, H., 2005. cDNA structure of an insulin-related peptide in the Pacific oyster and seasonal changes in the gene expression. *J. Endocrinol.* 187, 55–67. <https://doi.org/10.1677/joe.1.06284>
- [35] Smit, A.B., Vreugdenhil, E., Ebberink, R.H., Geraerts, W.P., Klootwijk, J., Joosse, J., 1988. Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* 331, 535–8. <https://doi.org/10.1038/331535a0>
- [36] Smit, A.B., Geraerts, P.M., Meester, I., van Heerikhuizen, H., Joosse, J., 1991. Characterization of a cDNA clone encoding molluscan insulin-related peptide II of *Lymnaea stagnalis*. *Eur. J. Biochem.* 199, 699–703.
- [37] Smit, A.B., van Marle, A., van Elk, R., Bogerd, J., van Heerikhuizen, H., Geraerts, W.P., 1993. Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*. *J. Mol. Endocrinol.* 11, 103–113.
- [38] Smit, A.B., Spijker, S., Van Minnen, J., Burke, J.F., De Winter, F., Van Elk, R., Geraerts, W.P.M., 1996. Expression and characterization of molluscan insulin-related peptide VII from the mollusc *Lymnaea stagnalis*. *Neuroscience* 70, 589–596. [https://doi.org/10.1016/0306-4522\(95\)00378-9](https://doi.org/10.1016/0306-4522(95)00378-9)
- [39] Smit, A.B., Thijsen, S.F., Geraerts, W.P., Meester, I., van Heerikhuizen, H., Joosse, J., 1992. Characterization of a cDNA clone encoding molluscan insulin-related peptide V of *Lymnaea stagnalis*. *Brain Res. Mol. Brain Res.* 14, 7–12.

- [40] Veenstra, J.A., 2010. Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen. Comp. Endocrinol.* 167, 86–103. <https://doi.org/10.1016/j.ygcen.2010.02.010>
- [41] Floyd, P., Li, L., 1999. Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. *J. Neurosci.* 19, 7732–7741.
- [42] Zatylny-Gaudin, C., Cornet, V., Leduc, A., Zanuttini, B., Corre, E., Le Corguillé, G., Bernay, B., Garderes, J., Kraut, A., Couté, Y., Henry, J., 2016. Neuropeptidome of the Cephalopod *Sepia officinalis*: Identification, Tissue Mapping, and Expression Pattern of Neuropeptides and Neurohormones during Egg Laying. *J. Proteome Res.* 15, 48–67. <https://doi.org/10.1021/acs.jproteome.5b00463>
- [43] Safavi-Hemami, H., Gajewiak, J., Karanth, S., Robinson, S.D., Ueberheide, B., Douglass, A.D., Schlegel, A., Imperial, J.S., Watkins, M., Bandyopadhyay, P.K., Yandell, M., Li, Q., Purcell, A.W., Norton, R.S., Ellgaard, L., Olivera, B.M., 2015. Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc. Natl. Acad. Sci.* 112, 1743–1748. <https://doi.org/10.1073/pnas.1423857112>
- [44] Safavi-Hemami, H., Lu, A., Li, Q., Fedosov, A.E., Biggs, J., Showers Corneli, P., Seger, J., Yandell, M., Olivera, B.M., 2016. Venom Insulins of Cone Snails Diversify Rapidly and Track Prey Taxa. *Mol. Biol. Evol.* 33, 2924–2934. <https://doi.org/10.1093/molbev/msw174>
- [45] Conzelmann, M., Williams, E.A., Krug, K., Franz-Wachtel, M., Macek, B., Jékely, G., 2013. The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* 14, 906. <https://doi.org/10.1186/1471-2164-14-906>
- [46] Miller, C.M., Newmark, P.A., 2012. An insulin-like peptide regulates size and adult stem cells in planarians. *Int. J. Dev. Biol.* 56, 75–82. <https://doi.org/10.1387/ijdb.113443cm>
- [47] Wang, S., Luo, X., Zhang, S., Yin, C., Dou, Y., Cai, X., 2013. Identification of putative insulin-like peptides and components of insulin signaling pathways in parasitic platyhelminths by the use of genome-wide screening. *FEBS J.* 281, 877–893.

<https://doi.org/10.1111/febs.12655>

- [48] Robitzki, A., Schröder, H.C., Ugarkovic, D., Pfeifer, K., Uhlenbruck, G., Müller, W.E., 1989. Demonstration of an endocrine signaling circuit for insulin in the sponge *Geodia cydonium*. *EMBO J.* 8, 2905–9.
- [49] Anctil, M., 2009. Chemical transmission in the sea anemone *Nematostella vectensis*: A genomic perspective. *Comp. Biochem. Physiol. - Part D Genomics Proteomics* 4, 268–289. <https://doi.org/10.1016/j.cbd.2009.07.001>
- [50] Lu, C., Lam, H.N., Menon, R.K., 2005. New members of the insulin family: Regulators of metabolism, growth and now...reproduction. *Pediatr. Res.* 57, 70–73. <https://doi.org/10.1203/01.PDR.0000159573.55187.CA>

Name	NCBI IDs	Type	A chain motif	Cystein residues number	C peptide(s) number	Possible C peptide cleavage ? Yes (Y) or No (N)	Possible D domain ? Yes (Y) or No (N)	Possible E domain ? Yes (Y) or No (N)	Bibliography
<i>Homo sapiens (human)</i>									
Insulin	AAA59172.1	γ	CC3XC8XC	6	1	Y	N	N	
IGF I	CAA40342.1	γ	CC3XC8XC	6	1	N	Y	Y	
IGF II	AAB34155.1	γ	CC3XC8XC	6	1	N	Y	Y	
Relaxin 1	AAH05956.1	γ	CC3XC8XC	6	1	Y	N	N	
Relaxin 2	AAI26416.1	γ	CC3XC8XC	6	1	Y	N	N	
Relaxin 3	AAL40345.1	γ	CC3XC8XC	6	1	N	N	N	
INSL3	AAI06723.1	γ	CC3XC8XC	6	2	Y (only for 1)	N	N	
INSL4	AAH26254.1	γ	CC3XC8XC	6	1	N	N	N	
INSL5	AAQ89389.1	γ	CC3XC8XC	6	2	Y (only for 1)	N	N	
INSL6	AAD39003.1	γ	CC3XC8XC	6	2	Y (only for 1)	Y	Y	
<i>Mus musculus (house mouse)</i>									
Insulin 1	EDL01683.1	γ	CC3XC8XC	6	1	Y	N	N	[2]
Insulin 2	EDL18180.1	γ	CC3XC8XC	6	1	Y	N	N	[3]
IGF I	NP_034642.2	γ	CC3XC8XC	6	1	N	Y	Y	
IGF II	NP_034644.2	γ	CC3XC8XC	6	1	N	Y	Y	[4]
Relaxin 1	EDL41676.1	γ	CC3XC9XC	7	1	N	N	N	
Relaxin 3	EDL10929.1	γ	CC3XC8XC	6	1	N	N	N	
INSL3	AAH49540.1	γ	CC3XC8XC	6	1	N	N	N	
INSL5	AAD29687.1	γ	CC3XC8XC	6	1	Y	N	N	
INSL6	EDL41675.1	γ	CC3XC8XC	6	2	Y (only for 1)	Y	N	
<i>Cyprinus carpio (common carp)</i>									
Insulin	XP_018955433.1	γ	CC3XC8XC	6	1	Y	N	N	[5]
<i>Danio rerio (zebrafish)</i>									
Insa	NP_571131.1	γ	CC3XC8XC	6	1	Y	N	N	[6]
Insb	XP_017214511	γ	CC3XC8XC	6	1	Y	N	N	[7]
IGF-1	NP_571900.1	γ	CC3XC8XC	6	1	N	Y	N	
IGF-2a	NP_571508.1	γ	CC3XC8XC	6	1	N	Y	Y	
IGF-2b	NP_001001815.1	γ	CC3XC8XC	6	1	N	Y	Y	
IGF-3	NP_001108522.1	γ	CC3XC8XC	6	1	N	Y	Y	
<i>Oreochromis niloticus (Nile tilapia)</i>									
NTins1	XP_019219195.1	γ	CC3XC8XC	6	1	N	,	N	[8]
NTins2	XP_003445871.3	γ	CC3XC8XC	6	2	Y	N	N	[7]; [9]
IGF-1	ABY88872.1	γ	CC3XC8XC	6	1	N	Y	Y	
IGF-2	ABY88873.1	γ	CC3XC8XC	6	1	N	Y	Y	
IGF-3	ABY88870.1	γ	CC4XC8XC	6	1	N	Y	Y	[7]

<i>Oryzias latipes (Japanese medaka)</i>										
IGF-1	XP_011476003.1	γ	CC3XC8XC	6	1	N	Y	Y	[7]	
IGF-2	Bibliography	γ	CC3XC8XC	6	1	N	Y	Y	[10]	
IGF-3	Bibliography	γ	CC4XC8XC	6	1	N	Y	Y	[7]	
<i>Oncorhynchus kisutch (coho salmon)</i>										
Insulin	XP_020340217.1	γ	CC3XC8XC	6	1	N	N	N	[11]	
<i>Ictalurus punctatus (channel catfish)</i>										
Insulin I	XP_017314995.1	γ	CC3XC8XC	6	1	N	N	N	[12]	
Insulin II	XP_017315312.1	γ	CC3XC8XC	6	1	N	N	N		
<i>Alligator mississippiensis (American alligator)</i>										
Insulin	XP_019332234.1	γ	CC3XC8XC	6	1	Y	N	N	[13]	
<i>Myxine glutinosa (Atlantic hagfish)</i>										
Insulin	P01342.1	γ	CC3XC8XC	6	1	Y	N	N	[14]	
<i>Ciona intestinalis (vase tunicate)</i>										
Ci-INS-L1	NP_001123204.1	γ	CC3XC8XC	6	1	N	Y	Y	[15]	
Ci-INS-L2	ABG35121.1	γ	CC3XC9XC	6	1	Y	N	N		
Ci-INS-L3	NP_001123344.2	γ	CC3XC8XC	6	1	N	N	N		
<i>Caenorhabditis elegans</i>										
ins-1	NP_501926.1	β	CC3XC8XCC	8	1	Y	N	N	[16]; [17]	
ins-2	NP_495194.1	β	CC3XC8XCC	8	0	N	N	N		
ins-3	NP_495195.1	β	CC3XC8XCC	8	0	N	N	N		
ins-4	NP_495196.1	β	CC3XC8XCC	8	0	N	N	N		
ins-5	NP_495197.4	β	CC3XC8XCC	8	0	N	N	N		
ins-6	NP_495198.1	β	CC3XC8XCC	8	0	N	N	N		
ins-7	NP_001255417.1	β	CC3XC8XCC	8	0	N	N	N		
ins-8	NP_001255416.1	β	CC3XC8XCC	8	0	N	N	N		
ins-9	NP_001024362.1	β	CC3XC8XCC	8	0	N	N	N		
ins-10	NP_001024125.1	β	CC3XC8XCC	8	0	N	N	N		
ins-11	NP_495071.1	γ	CC3XC8XC	6	0	N	Y	N	[17]	
ins-12	NP_001021963.1	γ	CC3XC8XC	6	0	N	Y	N		
ins-13	NP_001021962.1	γ	CC3XC8XC	6	0	N	Y	N		
ins-14	NP_001022153.1	γ	CC3XC8XC	6	0	N	Y	N		
ins-15	NP_001022154.1	γ	CC3XC8XC	7	0	N	Y	Y		
ins-16	NP_001022841.1	γ	CC3XC5XC	6	0	N	N	N		
ins-17	NP_497911.1	γ	CC3XC8XC	6	1	N	N	N		
ins-18	NP_492231.1	γ	CC3XC8XC	6	1	Y	N	N		[16]; [17]
ins-19	NP_001022339.1	γ	CC3XC9XC	6	0	N	Y	N		
ins-20	NP_001022529.2	α	C12XCC	6	0	N	N	N		
ins-21	NP_499222.2	α	C12XCC	6	0	N	N	N		
ins-22	NP_499223.1	α	C12XCC	6	0	N	N	N		

ins-23	NP_499224.1	α	C12XCC	6	0	N	N	N	
ins-24	NP_493443.2	α	C12XCC	6	0	N	N	N	[17]
ins-25	NP_001021849.1	α	C12XCC	6	0	N	N	N	
ins-26	NP_493445.1	α	C12XCC	6	0	N	N	N	
ins-27	NP_001021847.1	α	C14XCC	6	0	N	Y	N	
ins-28	NP_001021850.1	α	C12XCC	6	0	N	N	N	
ins-29	NP_001021846.1	α	C12XCC	6	0	N	N	N	
ins-30	NP_493444.1	α	C12XCC	6	0	N	N	N	
ins-31	NP_494454.1	α	C15XCC	18	0	N	Multiple chains (ins-31a, ins-31b and ins-31c)		[17]; [18]
ins-32	NP_494655.1	γ	CC4XC9XC	6	0	N	Y	N	
ins-33	NP_493389.1	α	C12XCC	6	0	N	N	N	
ins-34	NP_502702.2	α	C13XCC	6	0	N	Y	N	
ins-35	NP_507926.1	α	C14XCC	6	0	N	Y	N	[17]
ins-36	NP_001251580.1	α	C14XCC	6	0	N	Y	N	
ins-37	CAB04062.2	γ	CC3XC9XC	10	0	N	Y	Y	
ins-38	NP_001021964.1	γ	CC3XC9XC	6	0	N	N	N	
ins-39	NP_508247.1	α	C13XCC	6	0	N	Y	N	
DAF-28	NP_507840.1	β	CC3XC8XCC	8	0	N	N	N	[16]; [17]
<i>Aedes aegyptii</i> (yellow fever mosquito)									
AaegILP1	ABI64116.1	γ	CC3XC8XC	6	1	Y	N	N	
AaegILP2	ABI64118.2	γ	CC3XC8XC	6	1	Y	N	N	
AaegILP3	ABI64117.2	γ	CC3XC8XC	6	1	Y	N	N	
AaegILP4	ABI64119.1	γ	CC3XC9XC	6	1	Y	N	N	
AaegILP5	ABI64124.1	γ	CC4XC8XC	6	1	Y	N	N	[19]
AaegILP6a	ABI64121.1	γ	CC3XC8XC	6	1	Y	Y	N	
AaegILP6b	ABI64122.1	γ	CC3XC8XC	6	1	Y	Y	N	
AaegILP7	ABI64123.1	γ	CC3XC9XC	6	1	Y	N	N	
AaegILP8	ABI64120.1	γ	CC3XC8XC	6	1	Y	N	N	
<i>Drosophila melanogaster</i> (fruit fly)									
DILP1	NP_648359.1	γ	CC3XC8XC	6	1	Y	N	N	
DILP2	AAF50204.1	γ	CC3XC8XC	6	1	Y	N	N	[20]
DILP3	NP_648360.2	γ	CC3XC8XC	6	1	Y	N	N	
DILP4	CCF74199.1	γ	CC3XC8XC	6	1	Y	N	N	[20]; [21]
DILP5	NP_996037.2	γ	CC3XC8XC	6	1	Y	N	N	[20]; [22]
DILP6	NP_570000.1	γ	CC4XC8XC	6	1	N	N	N	[20]
DILP7	NP_570070.1	γ	CC4XC8XC	6	2	Y	Y	N	
DILP8	NP_648949.2	γ	CC3XC5XC	6	1	Y	N	N	[23]

***Bombyx mori* (domestic silkworm)**

A1	XP_004934106.1	γ	CC3XC8XC	6	1	Y	N	N
A2	AAA27825.1	γ	CC3XC8XC	6	1	N	N	N
A3	AAA27827.1	γ	CC3XC8XC	6	1	N	N	N
A4	NP_001121607.2	γ	CC3XC8XC	6	1	Y	N	N
A5	NP_001121608.2	γ	CC3XC8XC	6	1	N	N	N
A6	BAA00667.1	γ	CC3XC8XC	6	1	Y	N	N
A7	NP_001121629.2	γ	CC3XC8XC	6	1	Y	N	N
A8	NP_001121630.1	γ	CC3XC8XC	6	1	N	N	N
A9	BAA00672.1	γ	CC3XC8XC	6	1	N	N	N

[24]

Pseudogene

B1	NP_001121791.1	γ	CC3XC8XC	6	1	Y	N	N
B2	NP_001121793.1	γ	CC3XC8XC	6	1	Y	N	N
B3	BAA00674.1	γ	CC3XC8XC	6	1	Y	N	N
B4	NP_001121792.2	γ	CC3XC8XC	6	1	Y	N	N
B5	NP_001166891.1	γ	CC3XC8XC	6	1	Y	N	N
B6	NP_001121795.1	γ	CC3XC8XC	6	1	Y	N	N
B7	NP_001121789.1	γ	CC3XC8XC	6	1	Y	N	N
B8	NP_001121790.1	γ	CC3XC8XC	6	1	Y	N	N
B9	BAA00682.1	γ	CC3XC8XC	6	1	Y	N	N
B10	NP_001121788.1	γ	CC3XC8XC	6	1	Y	N	N
B11	NP_001121606.1	γ	CC3XC12XC	6	1	Y	N	N
B12	NP_001121794.2	γ	CC3XC8XC	6	1	Y	N	N
C1	NP_001119736.1	γ	CC3XC8XC	6	1	Y	N	N
C2	NP_001119735.1	γ	CC3XC8XC	6	1	Y	N	N

[25]

Pseudogenes

D1	NP_001121635.1	γ	CC3XC8XC	6	1	Y	N	N
E1	NP_001119733.1	γ	CC3XC8XC	6	1	Y	N	N
F1	NP_001119734.1	γ	CC3XC8XC	6	1	Y	N	N
G1	NP_001121634.1	γ	CC3XC8XC	6	1	Y	N	N

[26]

[27]

[28]

***Locusta migratoria* (migratory locust)**

LIRP	P15131.2	γ	CC3XC8XC	6	1	Y	N	N
------	--------------------------	---	----------	---	---	---	---	---

[29]

***Macrobrachium rosenbergii* (giant freshwater prawn)**

Mr-IAG	ACJ38227.1	β	CC7XC8XC	8	1	N	N	N
--------	----------------------------	---	----------	---	---	---	---	---

[30]

***Armadillidium vulgare* (common pillbug)**

Arv-AGH	BAA86893.1	γ	CC7XC8XC	8	1	Y	N	N
---------	----------------------------	---	----------	---	---	---	---	---

[31]

<i>Sagmariasus verreauxi</i> (green rock lobster)									
Sv-IAG	AHY99679.1	γ	CC7XC8XC	6	1	N	N	N	
Sv-ILP1	AIU40991.1	γ	CC4XC8XC	6	2	Y	N	N	[32]
Sv-ILP2	AIU40994.1	γ	CC3XC8XC	6	1	Y	N	N	
<i>Crassostrea gigas</i> (Pacific oyster)									
Cg-ILP	XP_011455161.1	γ	CC3XC8XC	6	2	Y	Y	N	[33]
Cg-MIP1	BAE48437.1	β	CXCC3XC8XC	8	3	Y	N	N	[33]; [34]
Cg-MIP2	XP_011415722.1	β	CXCC3XC8XC	8	3	Y	N	N	
Cg-MIP3	EKCI8432.1	β	CXCC3XC8XC	8	3	Y	N	N	[33]
Cg-MIP4	EKCI8433.1	β	CXCC3XC8XC	8	2	Y	N	Y	
Cg-MILP7	XP_011435161.1	β	CC4XC8XC	8	2	Y	N	N	
<i>Lymnaea stagnalis</i> (great pond snail)									
MIP I	P07223.2	β	CXCC3XC8XC	8	1	Y	N	N	[35]
MIP II	P25289.1	β	CXCC3XC8XC	8	2	Y	N	N	[36]
MIP III	P80090.2	β	CXCC3XC8XC	8	2	Y	N	N	[37]
MIP IV						<i>Pseudogene</i>			[38]
MIP V	P31241.1	β	CXCC3XC8XC	8	2	Y	N	N	[39]
MIP VI						<i>Pseudogene</i>			[38]
MIP VII	AAB46831.1	β	CXCC3XC8XC	8	2	Y	N	N	[38]
LS-MILP7	ESS73590.1 (ESTs)	β	CC4XC8XC	8	4	Y	N	N	According to [40]
<i>Aplysia californica</i> (California sea hare)									
Insulin	NP_001191615.1	β	CXCC3XC8XC	8	2	Y	Y	N	[41]
Ac-MILP7	EB230021.1 (ESTs)	β	CC4XC8XC	8	2	Y	N	N	According to [40]
<i>Lottia gigantea</i> (owl limpet)									
Insulin 1	XP_009056008.1	β	CXCC3XC8XC	8	2	Y	N	N	
Insulin 2	XP_009065889.1	β	CXCC3XC8XC	8	2	Y	N	N	
Insulin 3	XP_009060386.1	γ	CC3XC8XC	6	2	Y	Y	N	[40]
Insulin 4	XP_009065388.1	β	CC4XC8XC	8	2	Y	N	N	
<i>Sepia officinalis</i> (common cuttlefish)									
Insulin	Bibliography	β	CXCC3XC8XC	8	3	Y	N	N	[42]
<i>Conus geographus</i> (geography cone)									
Insulin 1	AJD85832.1	γ	CC3XC8XC	6	3	Y	N	N	
Insulin 1b	AJD85830.1	γ	CC3XC8XC	6	3	Y	N	N	
Insulin 1c	AJD85839.1	γ	CC3XC8XC	6	3	Y	N	N	
Insulin 2	AJD85826.1	β	CXCC3XC8XC	8	3	Y	N	N	[43]; [44]
Insulin 2b	AJD85823.1	β	CXCC3XC8XC	8	3	Y	N	N	
Insulin 3	AJD85820.1	γ	CC3XC8XC	6	2	Y	N	N	
Insulin 3b	AJD85822.1	γ	CC3XC8XC	6	2	Y	N	N	

<i>Conus floridulus</i>									
Insulin 1	AJD85829.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AJD85828.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
Insulin 3	AJD85824.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 4	AJD85837.1	β	CXCC3X8XC	8	3	Y	N	N	
<i>Conus marmoreus</i>									
Insulin 1	AJD85835.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40158.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus tulipa (tulip cone)</i>									
Insulin 1	AJD85833.1	γ	CC3XC8XC	6	2	Y	N	N	
Insulin 2	AJD85827.1	γ	CC3XC8XC	6	2	Y	N	N	[44]
Insulin 3	AJD85821.1	γ	CC3XC8XC	6	2	Y	N	N	
<i>Conus victoriae</i>									
Insulin 1	JAI08985.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus textile (cloth-of-gold cone)</i>									
Insulin 1	AJD85834.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40157.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus bandanus</i>									
Insulin 1	AOF40151.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40152.1	β	CXCC3X8XC	8	2	Y	N	N	[44]
<i>Conus tessulatus</i>									
Insulin 1	AOF40164.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40163.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus pulicarius</i>									
Insulin 1	AOF40162.1	β	CXCC3X8XC	9	3	Y	N	N	[44]
<i>Conus eburneus</i>									
Insulin 1	AOF40159.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40160.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus virgo</i>									
Insulin 1	AOF40168.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40156.1	β	CXCC3X8XC	8	2	Y	N	N	[44]
<i>Conus quercinus</i>									
Insulin 1	AJD85836.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AJD85831.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus tribblei</i>									
Insulin	AOF40165.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus varius</i>									
Insulin 1	AOF40167.1	β	CXCC3X8XC	8	3	Y	N	N	
Insulin 2	AOF40166.1	β	CXCC3X8XC	8	3	Y	N	N	[44]
<i>Conus planorbis</i>									
Insulin 1	AOF40161.1	β	CXCC3X8XC	8	3	Y	N	N	[44]

Platynereis dumerilii (Dumeril's clam worm)									
IRP-1	Bibliography	γ	CC3XC8XC	6	1	Y	Y	Y	
IRP-2	AHB62360.1	β	CXCC3XC8XC	8	2	Y	N	Y	
IRP-3	AHB62359.1	β	CXCC3XC8XC	8	2	Y	N	Y	[45]
IRP-4	Bibliography	β	CXCC3XC8XC	5	2	Y	N	Y	
IRP-5		β	CXCC3XC8XC	8	2	Y	N	Y	
Schmidtea mediterranea									
ILP	DAA33923.1	γ	CCC2XC8XC	8	1	N	Y	N	[46]
Schistosoma mansoni									
ILP	Bibliography	γ	CCC2XC8XC	8	1	Y	N	N	According to [47]
Schistosoma japonicum									
ILP	Bibliography	γ	CCC2XC8XC	8	1	Y	N	N	According to [47]
Taenia solium (pork tapeworm)									
ILP-1	Bibliography	γ	CCC2XC8XC	8	0	N	N	N	According to [47]
ILP-2		γ	CCC2XC8XC	8	1	Y	N	N	
Taenia saginata									
ILP1	Bibliography	γ	CCC2XC8XC	8	1	Y	N	N	According to [47]
ILP2		γ	CCC2XC8XC	8	0	N	N	N	
Taenia asiatica									
ILP1	Bibliography	γ	CCC2XC8XC	8	0	N	N	N	According to [47]
ILP2		γ	CCC2XC8XC	8	1	Y	N	N	
Hymenolepis microstoma									
ILP1	Bibliography	γ	CCC2XC8XC	8	0	Y	N	N	According to [47]
ILP2		γ	CCC2XC8XC	8	1	N	N	N	
Echinococcus multilocularis									
ILP1	Bibliography	γ	CCC2XC8XC	8	0	N	N	N	According to [47]
ILP2		γ	CCC2XC8XC	8	1	Y	N	N	
Echinococcus granulosus									
ILP1	Bibliography	γ	CCC2XC8XC	8	0	N	N	N	According to [47]
ILP2		γ	CCC2XC8XC	8	1	Y	N	N	
Geodia cydonium									
Sponge insulin	S09278	γ	C4XC7XC	5	1	Y	N	N	[48]
Nematostella vectensis (starlet sea anemone)									
NV199266	XP_001639325.1	γ	CC3XC8XC	6	1	Y	N	N	According to [49]
NV207484	XP_001632788.1	γ	CC3XC8XC	6	1	Y	N	N	
Hydra vulgaris									
ILP1	ADA67985.1	γ	CC3XC9XC	6	1	N	N	N	Steele et al. , 2009 (unpublished)
ILP2	ADA67986.1	γ	CC4XC8XC	6	1	N	N	N	
ILP3	ADA67987.1	γ	CC4XC8XC	6	1	N	N	N	

Abbreviations: Insulin Related Peptide (IRP); Insulin Growth Factor (IGF); INSulin Like (INSL); Insulin (Ins/ins); Insulin-Like Peptide (ILP); Insulin-Like Androgenic Gland factor (IAG); Androgenic Gland Hormone (AGH); Molluscan Insulin related Peptide (MIP); Molluscan Insulin-Like Peptide 7 (MILP7); Nile Tilapia (Nt); *Ciona intestinalis* (Ci); *Aedes aegyptii* (Aeeg); *Drosophila* Insulin-Like Peptide (DILP); Locust Insulin Related Peptide (LIRP); *Macrobrachium rosenbergii* (Mr); *Armadillidium vulgare* (Arv); *Sagmaria verreauxi* (Sv); *Crassostrea gigas* (Cg); *Lymnaea stagnalis* (Ls); *Aplysia californica* (Ac); *Nematostella vectensis* (NV).

3. Partie B – Identification d'effecteurs du signal insuline

3.1. Matériels et méthodes

3.1.1. Analyse des bases de données

L'identification des effecteurs des voies de l'insuline a été réalisée en se basant sur les effecteurs de la signalisation insuline de référence décrits dans KEGG GENES database (www.genome.jp/kegg/). Les séquences sont identifiées par blast en utilisant, pour chaque effecteur, les séquences issues de vertébrés (*H. sapiens* ; *M. musculus*), d'Ecdysozoaires (*Drosophila busckii* ; *D. melanogaster* ; *C. elegans* ; *Ascaris suum*) ou de Lophotrochozoaires (*A. californica* ; *B. glabrata* ; *M. edulis* ; *Lingula anatina* ; *Pinctada martensii* ; *Echinococcus multilocularis*). Pour chacune de ces séquences, des blast réciproques ont été réalisés sur le transcriptome (GIGATON : <http://ngspipelines-130Sigenae.toulouse.inra.fr>) et le protéome prédit (NCBI) de *C. gigas*. Toutes les séquences obtenues ont été prédites en utilisant l'outil de traduction ExPasy (<https://web.expasy.org/translate/>). Les alignements ont été faits avec Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) et la séquence signal protéique du récepteur CIR a été prédite en utilisant PREDISI (<http://www.predisi.de/>). Les voies de signalisation de *C. gigas* ont été cartographiées sur la base du modèle de référence KEGG (map04910).

3.1.2. Animaux et préparations histologiques

Les huîtres creuses diploïdes utilisées sont les mêmes que celles décrites dans le chapitre I, de même que les protocoles de fixation et de coloration au Trichrome Prenant-Gabe des préparations histologiques.

3.1.3. Extraction d'ARN et reverse transcription

Pour chaque tissu (muscle adducteur, glande digestive, palpes labiaux, manteau, bord de manteau, branchies et cœur), quatre pools de trois échantillons ont été collectés en Novembre. Pour les ganglions viscéraux et l'aire gonadique, les pools ont été réalisés en regroupant des animaux mâles ou femelles pour chaque stade de gamétogenèse. Chaque pool est broyé puis incubé 5 minutes dans du TriReagent® (Sigma Aldrich) avant l'ajout de 100µL de 1-Bromo-3-Chloropropane (Sigma Aldrich). Les échantillons sont ensuite vortexés et centrifugés à 11 000g pendant 15 minutes à 4°C. La phase aqueuse ainsi obtenue est

prélevée pour l'extraction d'ARN selon le kit Nucleospin RNA II (Macherey-Nagel). La transcription inverse est réalisée selon le protocole de Dheilly *et al.* (2012).

3.1.4. PCR quantitative

Le couple d'amorces pour CIR (Sens : 5'-AACCTGCCTGACAAAAATG-3' et Anti-sens 5'-TACACGATCACCTCCCCTG-3') est le même que celui de Jouaux *et al.* (2012). Le gène de référence EF1 α a été sélectionné parmi les trois gènes de référence testés (GAPDH, Arf1, EF1 α). Le couple d'amorces pour EF1 α (Sens : 5'-ACGACGATCGCATTTCTCTT-3' et Anti-sens : 5'-ACCACCCTGGTGAGATCAAG-3') a été préparé en utilisant Primer3. Les qPCR ont été réalisées en utilisant le mix 1X GoTaq SYBR Green (Promega), 8,85ng d'ADNc et 900nM de chaque amorce dans un volume final de 15 μ L. Chaque qPCR consiste en un cycle de 5 minutes à 95°C suivi de 45 cycles de 15 secondes à 95°C et 45 secondes à 60°C. La courbe de fusion a été obtenue par 80 cycles de 10 secondes en augmentant la température de 0,5°C à chaque cycle à partir de 55°C jusqu'à 95°C.

3.1.5. Analyses statistiques

Les analyses statistiques ont été réalisées sur les logiciels R 3.4.3 et RStudio. Un test de Shapiro a été réalisé afin de vérifier la normalité de distribution des valeurs suivi soit d'un test post-hoc de Wilcoxon soit d'un test non-paramétrique de permutations permettant de définir les différences significatives ($\alpha = 0,05$).

3.2. Résultats

3.2.1. Le récepteur CIR

La séquence protéique complète du récepteur CIR a été retrouvée dans les bases de données. Six isoformes prédites (XP_011456321.1 ; XP_011456322.1 ; XP_011456323.1 ; XP_011456324.1 ; XP_011456325.1 ; XP_011456326.1) avec peptide signal et une séquence putative (EKC21734.1) sans peptide signal sont décrites. L'alignement montre que leurs séquences sont très proches, avec moins de 2,6% de différence. L'alignement de la séquence putative de CIR (EKC21734.1) avec celle identifiée par Gricourt *et al.* (2003) (CAD59674.1) montre 96,77% d'identité dans la région commune. De plus, EKC21734.1 contient la région riche en cystéines impliquée dans la liaison des ligands (Figure II-7). La séquence prédite correspondante (XP_011456323.1) permet d'identifier le peptide signal.

Les pourcentages d'identité entre les chaînes α du récepteur CIR complet et des récepteurs IGF-1R et INSR humains sont identiques et d'une valeur de 37%. Ce pourcentage est de 33% avec le récepteur DIR de drosophile et 29% avec DAF-2 de *C. elegans*.

Peptide signal M I C E D V D I R N D V N A F A E L E N C T V I E G N L K I L L I E G A S H Q Q Y E N L Y F P D L V E I T D
→ α
 F F L L Y R V Y G L K S L R H I F P N L S I I R G H K L F F N F A L V V F E L M D L E E L G L I G L T T I E R G A V R F E K N P Q L C Y I D T I
 D W T K I A V G V Q G M S E H Y F K Q N R I A **K E C V D V C P S D C A S T P V I A L T G E V S K E V K K R C W T K S Y C Q K I L T C E**
G L N R P C Q S G V C D N G K C C H P N C I G G C T G E T D R D C K V C K N V S M I N N Q A I C Q K Q C L P G T Y T L R R R R C
L L E Y E C T D V I H Y K L V K N N Q T G L P G D C V D K L A G Y L E S A D G R T C N K C A D K C P K S C N G K K L E T A A Q S Q D
 M F G C T I I V N H L E I R I T R G S N V A Q E L E R N L G Q I E E V Q G Y I W I H S N T L L S L G F F K N L R K I H G K N S N S N N F
 S L L V H D N M N L Q E L F P E N V Q N N L E I L N G N M F F H Y N R K L C Y N K I K S F E S K V K M D N K S K N D I S D S S N G D
 Q M P C T V T T L N L T V V K I A A R I A F L R W D N F K S S D P R Q L L S Y V I S W K E A P F K N L S I Y E G R D A C D N D V W Y T K
 D I L N K M D K D E D Y A T A L L L N L K P W T Q Y A V Y V Q T Y T T S S A K Y G A I S S I V Y F T T L P Y Y P T I P D K L R V K A L A P G T
 L K I M W N P P A N P H G N V T H Y E L Y W R K R P L V H E K F N E R D Y C S E P L D T L L D Q S D D L Y D E E D K M D S S R D L G L
 S N N K T I Q G E C A C P K T F E E E E E R K E R E M Q I A F E N F L H N Y V Y L K R T D Y G P P K P L P D L P T D L P T T S M D S K R
 A E T L F K K N R R K F S N I T L P T N L P L K P N P T L **R R Q R** **A R R Y S R E A S L P Y N A S A V N G L K I N L S D S E E N D T D V V N E**
→ β
T E K D L P F N K V I V Q T L V Q K N D M S Y I I S N L G H F Q D Y R I E V I A C Q E R D P R D P A K M K L C S I R A I S L G R T L P D K N A
D N I D P K T I K I Y K M S N K T G E V I V Y W E E P K E P N G L I L S Y D I Q Y S K V M K D V K P N T L C V P H L K Y K K I G G Y K L Q K
L D P G N Y S V Q I R A I S L A F T T N Y T E P R Y F V V S A P T T T A H W P P E T I I A V T V S G V L L L I V V I G V W F I A R S R F K K
I P D R P D F T S I N P D Y E H Y N P D E W E V E R D K I Q L I R E L G Q G S F G M V Y E G V A R D L Y G K G G E I K V A V K T V N E H
A T Y R E R M E F L S E A S R M K A F S C N H V V R L L G V V S D G Q P A L L I M E L M E K G D L K N F L R M H R P V E E D I D S T
F A G I K E H I S R T P P T I K R I I Q M A G E I A D G M A Y L A D K K F V H R D L A A R N C M V A E D L T V K I A D F G M T R D I Y
E T D Y Y R K G G K A L L P V R W M A P E S L K D G I F T S L S D V W S Y G V V L W E M A T L A A Q P Y Q G L S N E E V L R Y V L
N G R V M E K P E D C P D R L F E L M Q K C W R Y R P K Q R P T F K E I L E E L V P E L D P S F A N K S Y F F S D E N K Q D S L E Y I E E
G D E D M D E S K T P F I Q N E D E C D M K G A E G G V D L P E L P D Q D I D D D I D L G E D Y Y P P G Y Q E A Q H S Y Q N G E L F
P E L R I D K V N N V L Y S R S Q A D K L M L N G T E P C D C I L T Q E P N P D S H P D Q R F S S C S N P Y S A I G S S D D S K D S S K S S
S S S Y A Q M N G V H I A N G H I A M H H L K T T K C

Figure II-7 : Séquence complète du récepteur CIR. Numéro d'accèsion NCBI : EKC21734.1. Le peptide signal (jaune), les séquences (gris) encadrant la région riche en cystéines (vert) de la partie N-terminale. La partie correspondant à l'identification de Gricourt *et al.* (2003) (NCBI ID : CAD59674) est en italique. Cette partie contient le site tétrabasique de clivage potentiel (noir et souligné) qui détermine le début de la région codante pour les sous-unités β, les domaines transmembranaire (bleu) et tyrosine kinase (rouge). Les motifs MVA et MTR (rouge et soulignés) sont caractéristiques des récepteurs tyrosine kinase IGF-1R chez l'humain et DIR chez la drosophile.

L'expression du transcrit de CIR a été prédite *in silico* au cours du développement. Les résultats (Figure II-8) montrent une expression plus forte lors des stades précoces de développement et une baisse d'expression du stade morula jusqu'au stade gastrula. L'expression augmente à nouveau à partir des stades larves D tardifs (Figure II-8, D6 et D7) jusqu'au stade juvénile (Figure II-8, J).

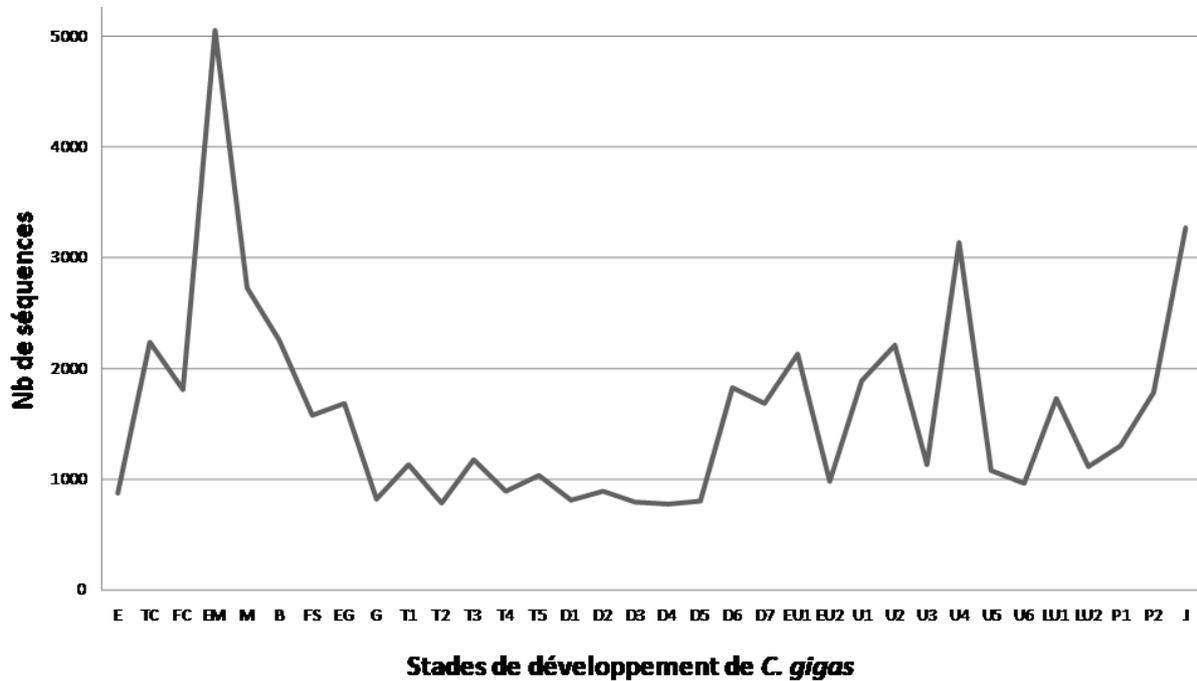


Figure II-8 : Expression prédite *in silico* au cours du développement de *C. gigas*. E : Ovocytes non fécondés ; TC : Stade deux cellules ; FC : Stade quatre cellules ; EM : Jeune morula ; M : Morula ; B : Blastula ; FS : Blastula en mouvement ; EG : Jeune gastrula ; G : Gastrula ; T1 à T5 : Larves trochophore de 9h30 à 14h30 post-fécondation ; D1 à D7 : Larves D de 17h35 à 3,77 jours post-fécondation ; EU1 et EU2 : Jeune larve umbo à 5 et 7 jours post-fécondation respectivement ; U1 à U6 : Larves umbo de 8 à 14 jours post-fécondation ; LU1 et LU2 : Larves véligères à 15 et 16 jours post-fécondation respectivement ; P1 et P2 : Larves pédivéligères à 18,03 et 18,19 jours post-fécondation ; J : Juvéniles à 215 jours post-fécondation selon Zhang *et al.*, 2012.

Le profil d'expression *in silico* d'après GIGATON a également été décrit pour différents tissus adultes et mis en relation avec les expressions en qPCR (Figure II-9A). Le récepteur d'IRPs est plus exprimé dans les palpes labiaux, l'aire gonadique mâle, le bord de manteau ainsi que les branchies. Les mesures faites en qPCR pour le même transcrite ont confirmé que CIR est plus exprimé dans le cœur et les branchies (Figure II-9B). Le récepteur CIR présente le même niveau d'expression global et le même profil dans les ganglions viscéraux et l'aire gonadique (Figure II-9, C et D) au cours du cycle annuel. L'expression du récepteur augmente du stade quiescent (Stade 0) jusqu'aux stades de différenciation germinale (Stades 2) aussi bien mâles que femelles (Figure II-9C). Chez les femelles, cette expression continue d'augmenter jusqu'au stade de maturité sexuelle (Stade 3). A l'inverse, chez les mâles, le maximum d'expression est atteint au stade de proliférations goniales (Stade 1) et de différenciation germinale (Stade 2) puis l'expression diminue significativement au stade de maturité sexuelle (Stade 3).

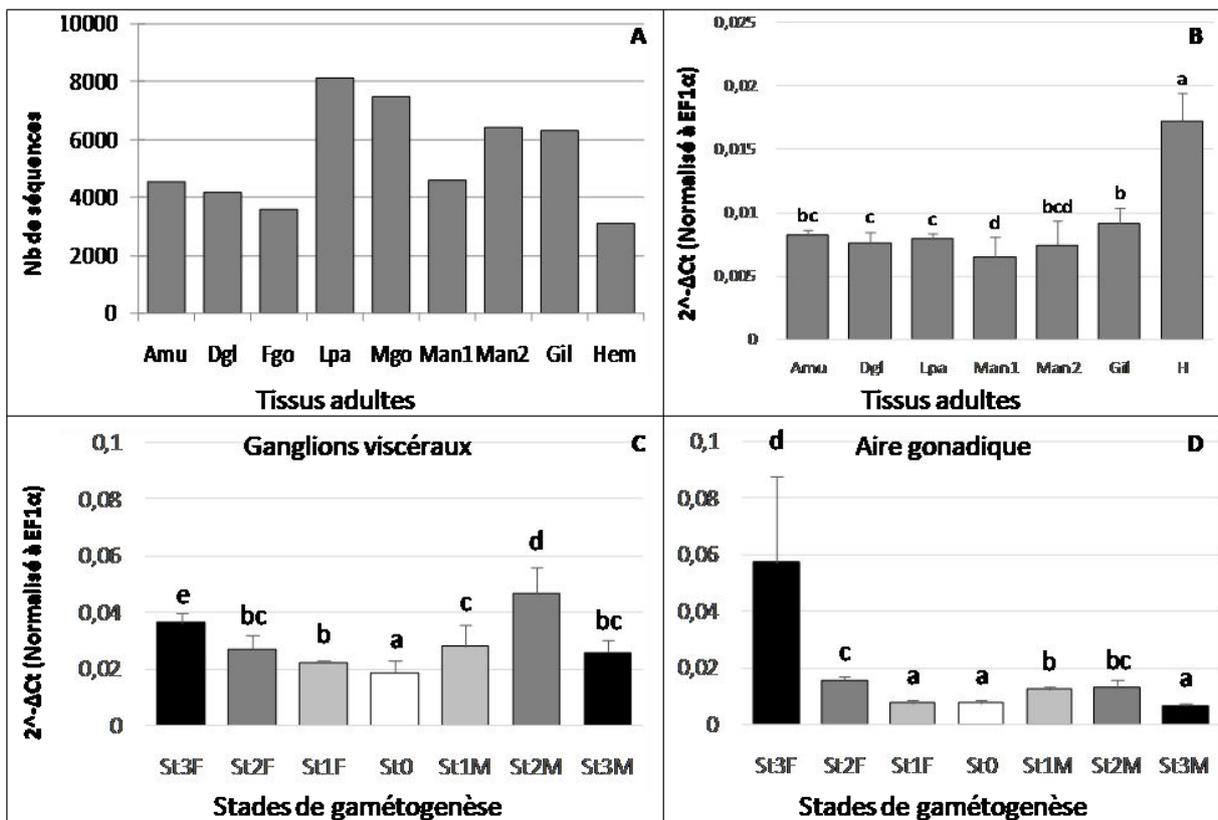


Figure II-9 : Expressions du récepteur CIR. A – Expression prédite *in silico* du transcrite CIR en utilisant GIGATON dans les tissus adultes de *C. gigas*. B – Expressions de CIR mesurées en qPCR dans les tissus adultes de *C. gigas*. Amu : Muscle adducteur ; Dgl : Glande digestive ; Fgo et Mgo : Gonade femelle et mâle respectivement ; Lpa : Palpes labiaux ; Man1 : Manteau ; Man2 : Bord de manteau ; Gil : Branchies ; Hem : Hémocytes ; H : Cœur. C – Expressions de CIR mesurées en qPCR dans l'aire gonadique à tous les stades de gamétogenèse. D – Expressions de CIR mesurées en qPCR dans les ganglions viscéraux à tous les stades de gamétogenèse. St0 : Stade 0 ; St1F et St1M : Stades 1 femelle et mâle respectivement ; St2F et St2M : Stades 2 femelle et mâle respectivement ; St3F et St3M : Stades 3 femelle et mâle respectivement.

3.2.2. Les voies de signalisation

La recherche dans GIGATON a permis d'identifier la quasi-totalité des effecteurs des différentes voies associées à la signalisation insuline (Figure II-10) ce qui confirme la conservation de cette signalisation chez les mollusques. Les détails des séquences identifiées et les fonctions prédites sont indiqués dans le Tableau II-2. L'IRS1, sa cible PI3K, Akt et ses protéines cibles (TSC2, mTORC1, FOXO) sont retrouvées chez *C. gigas* ce qui suggère la conservation des premières étapes de ces voies de signalisation. De même, les effecteurs de la voie Ras/MAPK ont été retrouvés à l'exception de SHC. Les gènes codant pour la plupart des éléments fonctionnels impliqués dans le métabolisme du glycogène sont retrouvés (cadres rouges), à l'exception du GLUT4, transporteurs du glucose inductibles par le signal insuline chez les vertébrés (Dupont et LeRoith, 2001). Toutefois, chez *C. gigas*, deux GLUT sont connus et apparentés aux GLUT1 et GLUT12 de vertébrés. Les acteurs du métabolisme lipidique (cadres bleus) sont également retrouvés dans leur ensemble, ainsi que les acteurs impliqués dans les synthèses protéiques (cadre vert) et dans le contrôle de la prolifération et de la différenciation cellulaire (cadre orange). L'effecteur BAD impliqué dans le contrôle de l'apoptose (cadre jaune) n'a pas été retrouvé.

L'expression de ces effecteurs a été recherchée *in silico* dans les banques d'expression d'aire gonadique d'animaux matures, ces données sont indiquées dans le Tableau II-2 et la Figure II-10. Ainsi, les acteurs les plus exprimés concernent majoritairement trois processus : le métabolisme du glucose (glycolyse et glycogénèse), le métabolisme des protéines ainsi que la prolifération et la différenciation cellulaires. De façon intéressante, les expressions prédites pour les acteurs GK (glycolyse) et FAS (lipogénèse) sont plus élevées dans l'aire gonadique mature des femelles. A l'inverse, les effecteurs finaux FBP (glycolyse), PYG (glycogénèse), MNK et ELK1 (prolifération/différenciation) présentent des expressions prédites plus élevées dans l'aire gonadique des mâles matures.

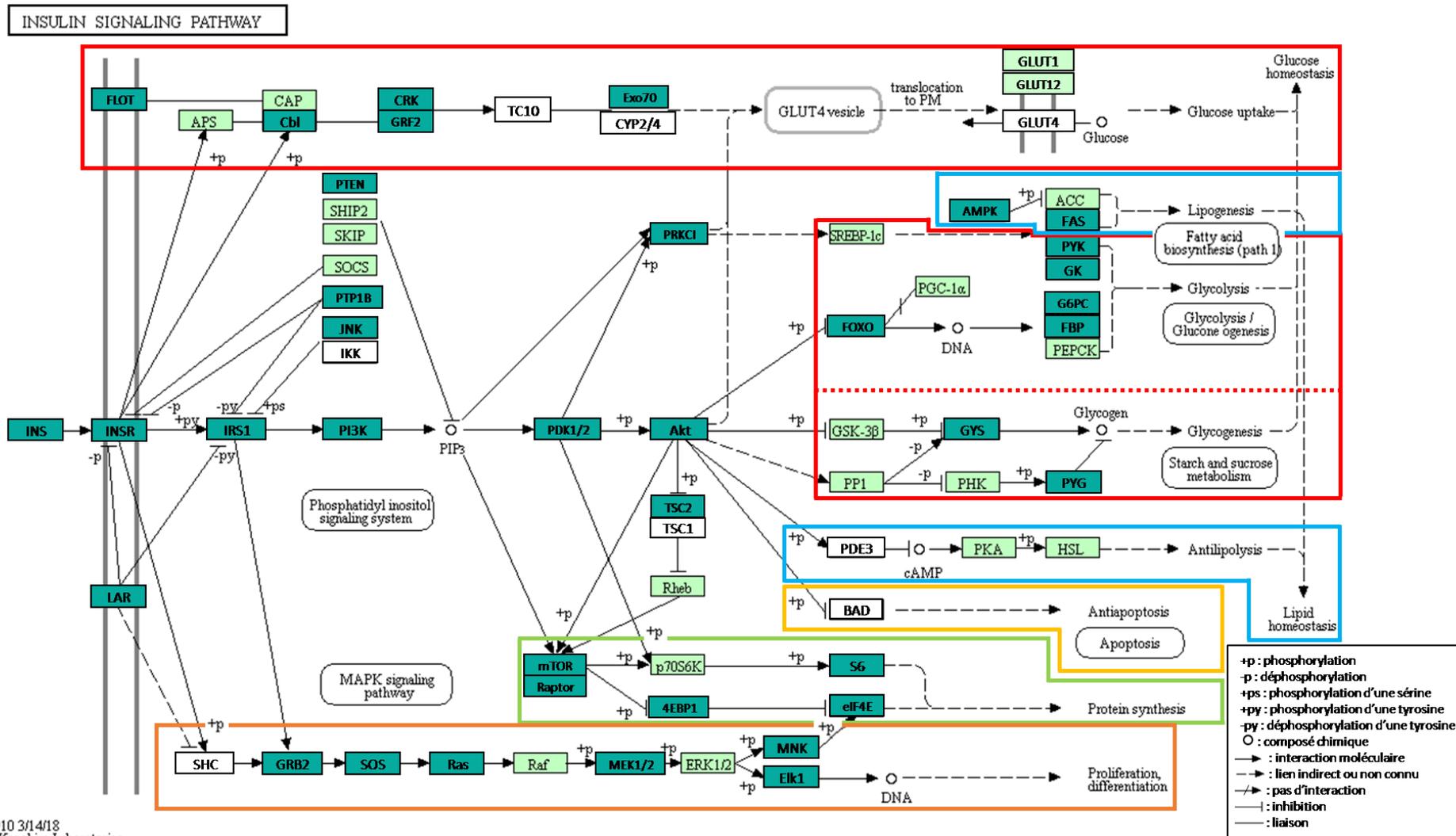


Figure II-10 : Voies de signalisation de l'insuline et effecteurs conservés chez *C. gigas*. La cartographie des effecteurs a été réalisée sur la base du modèle de référence map04910 issu de KEGG GENES database (www.genome.jp/kegg/). Les effecteurs en vert sont conservés chez *C. gigas*, ceux en vert foncé sont exprimés dans l'aire gonadique (données qPCR ou GIGATON). Cadre rouge : métabolisme du glycogène, cadres bleus : métabolisme lipidique, cadre jaune : voie de l'apoptose, cadre vert : synthèse protéique, cadre orange : contrôle de la prolifération et de la différenciation cellulaire.

Tableau II-2 : Tableau récapitulatif des effecteurs des voies des IRPs. La notation, le KEGG ID, le nom, la définition, la conservation ou non chez *C. gigas*, des informations complémentaires et les expressions qPCR (rouge) ou GIGATON (bleu) sont notifiés pour chaque séquence issue du schéma de signalisation KEGG. + : fortement exprimé ; - : faiblement exprimé ; (M) et (F) : expression prédite plus fortement dans l'aire gonadique mature mâle et femelle respectivement.

Notation (schéma KEGG)	KEGG ID	Nom	Définition (KEGG reference pathway)	<i>Crossostrea gigas</i> NCBI ID	Remarques	Expressions AG
INS	K04526	INS	insulin	XP_011435161.1 ; XP_011145516.1 ; EK18433.1 ; EK18432.1 ; XP_011415722.1 ; BAE48437	<i>in review</i>	+
INSR	K04527	INSR	insulin receptor	XP_011456323.1		+
FLOT	K07192	FLOT1, FLOT2	Flotillin 1 et 2	EKC35912.1 ; XP_011435140.1	La séquence de Flot1 est prédite	+, +
APS	K07193	SH2B2, APS	SH2B adaptor protein 2	EKC29126.1	SH2B adaptor protein 1	-
CAP	K06086	SORBS1, SH3D5, PONSIN, CAP	sorbin and SH3 domain containing 1	EKC30775.1		
Cbl	K04707	CBL	E3 ubiquitin-protein ligase CBL	EKC24300.1		+
CRK	K04438	CRK	CRK proto-oncogene, adaptor protein	XP_011430292	séquence prédite	+
GRF2	K06277	RAPGEF1, GRF2	Rap guanine nucleotide exchange factor 1	EKC31451.1		+
TC10	K07194	TC10	Ras homolog gene family, member Q	non identifié		
Exo70	K07195	EXO70, EXO70	exocyst complex component 7	EKC30356.1		+
CIP4/2	K07196	TRIP10, CIP4	thyroid hormone receptor interactor 10	non identifié		
GLUT4	K07191	SLC2A4	solute carrier family 2 member 4	non identifié	autres GLUT existent : GLUT1 et	
PTEN		PTEN	Phosphatase and TENsin homolog	XP_011431707.1	séquence prédite	+
SHIP2	K15909	SHIP2, INPPL1	inositol polyphosphate phosphatase like 1	EKC37240.1		-
Socs	K04694	Socs1, 2, 3, 4	suppressor of cytokine signaling	AKA59675.1 (SOCS2)	EKC39889.1 (SOCS5) ; EKC24066.1 (SOCS6) ; EKC24859.1 (SOCS7)	
PTP1B	K05696	PTP1B, PTPN1	protein tyrosine phosphatase, non-receptor type 1	EKC41394.1		+
JNK	K04440	JNK	mitogen-activated protein kinase 8/9/10 (c-Jun N-terminal kinase)	XP_011442365.1	séquence prédite	+
IKK	K07209	IKKB, IKK β	inhibitor of nuclear factor kappa B kinase subunit beta	non identifié		
IRS1	K16172	IRS1	insulin receptor substrate 1	EKC40064.1		+
PI3K	K02649	PI3KR1	phosphoinositide-3-kinase regulatory subunit 1	EKC33889.1		+
	K00922	PI3KCA_B_D	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha/beta/delta	EKC39772.1 ; EKC40180.1		- ; +
PDK1/2	K06276	PDK1, PDK1/2	3-phosphoinositide dependent protein kinase 1	EKC19681.1		+
aPKC	K18952	PRKCZ	protein kinase C zeta	non identifié		
	K06069	PRKCI	protein kinase C iota	XP_011436250.1	séquence prédite	+
SREB-1	K07197	SREBP1, SREBF1	sterol regulatory element-binding transcription factor 1	XP_011452106.1	séquence prédite	-
AMPK	K07198	PRKAA1	protein kinase AMP-activated catalytic subunit alpha	APP95933.1 ; EKC27728.1		+
	K07199	PRKAB1	protein kinase AMP-activated non-catalytic subunit beta	EKC41094.1		+
	K07200	PRKAG	protein kinase AMP-activated non-catalytic subunit gamma 1	AGH13204.1	séquence partielle	-
ACC	K11262	ACACA ACAC	acetyl-CoA carboxylase	XP_011441549.2		-
FAS	K00665	FASN	fatty acid synthase	EKC24795.1		+(F)
PYK	K12406	PKLR	pyruvate kinase L/R	CAJ28914.1		+
HK	K00844	HK1, 2, 3	hexokinase	CAJ28915.1 ; XP_011412815.1 ; XP_011431075.1	Les séquences d'hexokinase 1 et 2 sont prédites	+ ; + ; -
GK	K12407	GCK	glucokinase	EKC25684.1		+(F)
AKT	K04456	AKT	AKT serine/threonine kinase	XP_011430945	séquence prédite	+
FOXO	K07201	FOXO1	forkhead box O1	EKC23646.1		+
PGC-1a	K07202	PPARGC1A, PGC1A	PPARG coactivator 1 alpha	XP_011427714.1	séquence prédite	-
G6PC	K01084	G6PC1, G6PC2, G6PC3	glucose-6-phosphatase catalytic subunit 1 à 3	EKC25490.1		+
FBP	K03841	FBP1, FBP2	fructose-bisphosphatase 1 et 2	XP_011419913.1	séquence prédite	+(M)
PEPCK	K01596	PCK1, PCK2	phosphoenolpyruvate carboxykinase 1 et 2	CAJ28913.1		-
GSK3B	K03083	GSK3B	glycogen synthase kinase 3 beta	EKC35379.1		+
GYS	K00693	GYS1, GYS2	glycogen synthase 1 et 2	AA593900.1		+
PP1	K06269	PPP1CA, PPP1CB	protein phosphatase 1 catalytic subunit	EKC23784.1 ; EKC31188.1		- ; -
PHK	K00871	PHKG1, PHKG2, PHKB	phosphorylase kinase catalytic subunit gamma et beta	EKC30288.1		-
	K07190	PHKA1, PHKA2	phosphorylase kinase regulatory subunit alpha	non identifié		
	K02183	CALML1 à 6	calmodulin	EKC32562.1		-
PYG	K00688	PYGL, PYGM, PYG	glycogen phosphorylase	AA593901.1 ; EKC30247.1		+(M)
PDE3	K13296	PDE3B	cGMP-inhibited 3',5'-cyclic phosphodiesterase B	non identifié		
PKA	K04345	PRKACA, PRKACB, PRKACG	protein kinase cAMP-activated catalytic subunit alpha, beta, gamma	XP_011439327	séquence prédite	-
PRKAR	K04739	PRKARIA, PRKAR2A, PRKAR2B, PRKAR1B	protein kinase cAMP-dependent type I regulatory subunit alpha	EKC24170.1 ; EKC18724.1		- ; -
HSL	K07188	LIPE	lipase E, hormone sensitive type	XP_011451299.1	séquence prédite	-
BAD	K02158	BAD	BCL2 associated agonist of cell death	non identifié		
TSC1	K07206	TSC1	TSC complex subunit 1	non identifié		
TSC2	K07207	TSC2	TSC complex subunit 2	EKC33134.1		+
RHEB	K07208	RHEB	Ras homolog, mTORC1 binding	XP_011422874	séquence prédite	-
mTOR	K07203	mTOR	mechanistic target of rapamycin kinase	EKC35519.1		+
Raptor	K07204	RPTOR	regulatory associated protein of mTOR complex 1	EKC19393.1		+
LAR	K05695	PTPRF	protein tyrosine phosphatase, receptor type F	XP_019923274.1	séquence prédite	+(M)
P7056K	K04688	RPS6KB1, RPS6KB2	ribosomal protein S6 kinase B1	EKC41560.1		-
S6	K02991	RPS6	ribosomal protein S6	XP_011421218.1	séquence prédite	+
4EBP1	K07205	EIF4EBP1	eukaryotic translation initiation factor 4E binding protein 1	EKC38736.1		+
EIF4E	K03259	EIF4B, EIF422	eukaryotic translation initiation factor 4E family member 2	XP_011426855	séquence prédite	+
SHC	K06279	SHC1	SHC adaptor protein 1	non identifié		
	K17447	SHC2	SHC adaptor protein 2	non identifié		
	K17448	SHC3	SHC adaptor protein 3	non identifié		
	K17449	SHC4	SHC adaptor protein 4	non identifié		
GRB2	K04364	GRB2	growth factor receptor bound protein 2	EKC35648.1		+(M)
SOS	K03099	SOS1, SOS2	SOS Ras/Rac guanine nucleotide exchange factor	EKC39507		+
RAS	K02833	HRAS	HRas proto-oncogene, GTPase	XP_011447097.1	séquence prédite	+
	K07827	KRAS	KRAS proto-oncogene, GTPase	non identifié		
	K07828	NRAS	NRAS proto-oncogene, GTPase	non identifié		
RAF	K08845	ARAF	A-Raf proto-oncogene, serine/threonine kinase	non identifié		
	K04365	BRAF	B-Raf proto-oncogene, serine/threonine kinase	EKC22769.1		-
MEK1/2	K04366	RAF1	Raf-1 proto-oncogene, serine/threonine kinase	XP_011418545.2	séquence prédite	
	K04368	MAP2K1	mitogen-activated protein kinase kinase 1	XP_011443360.1	séquence prédite	+
	K04369	MAP2K2	mitogen-activated protein kinase kinase 2	non identifié		
ERK	K04371	MAPK1, MAPK3	mitogen-activated protein kinase 1 et 3	EKC28575.1 ; XP_011416487.1	La séquence de MAPK3 est prédite	+ ; +
MNK	K04372	MKNK1, MKNK2	MAP kinase interacting serine/threonine kinase 1 et 2	EKC18571.1	Une seule séquence identifiée	+(M)
ELK1	K04375	ELK1	ELK1, ETS transcription factor	EKC24101.1		+(M)

3.3. Discussion

La récente disponibilité du génome et des transcriptomes de *C. gigas* a permis d'obtenir la séquence complète du récepteur CIR ainsi que les séquences de la plupart des effecteurs décrits des voies de signalisation de l'insuline. Le fait que plusieurs isoformes du récepteur soit présentes dans les bases de données est probablement dû au fort taux de polymorphisme décrit chez *C. gigas* (Zhang *et al.*, 2012). L'identification de CIR lors de cette étude a permis d'obtenir la partie N-terminale du récepteur impliquée dans la liaison au ligand. A l'instar du récepteur d'*A. californica* (Jonas *et al.*, 1996), deux motifs contenus dans le domaine tyrosine kinase ont pu être identifiés (MVA et MTR). Ces motifs sont également retrouvés dans le récepteur IGF-1R humain ainsi que dans le récepteur DIR de drosophile. Bien qu'un seul récepteur d'IRPs ait été détecté, il ne peut être exclu que les IRPs de mollusques puissent se lier à d'autres types de récepteurs. Chez les vertébrés, les IRPs de type relaxine se lient à des récepteurs couplés aux protéines G (RCPG) (Bathgate *et al.*, 2013). Ce type de récepteurs existe chez *C. gigas* et, au moins l'un des RCPG, est exprimé dans la gonade (Rodet *et al.*, 2005). D'après ces données, il sera intéressant de déterminer si les IRPs identifiés chez *C. gigas* se lient tous au même récepteur CIR ou s'ils peuvent se lier à d'autres types de récepteurs. De plus, afin de mieux connaître leurs rôles fonctionnels, la voie de signalisation activée lors de la liaison de ces IRPs devra être identifiée.

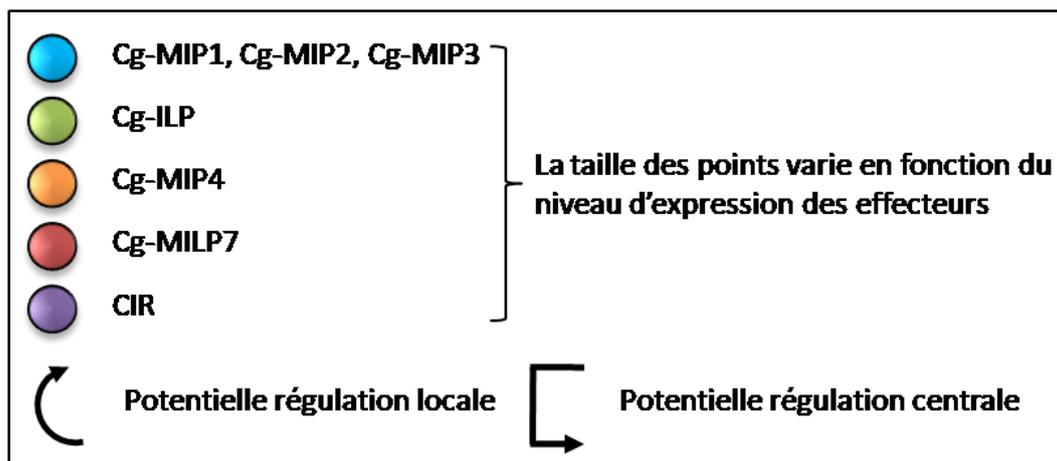
Dans les tissus adultes, les expressions de CIR obtenues en utilisant GIGATON sont en accord avec celles mesurées par qPCR lors de cette étude et des études précédentes (Jouaux *et al.*, 2012). En effet, l'expression dans l'aire gonadique confirme une implication du système insuline dans le contrôle de la reproduction comme chez la drosophile (Hsu et Drummond-Barbosa, 2009, 2011; Shim *et al.*, 2013) ou chez *C. elegans* (Hubbard, 2011; Michaelson *et al.*, 2010). Concernant les données qPCR mesurées dans les ganglions viscéraux et l'aire gonadique au cours des stades de gamétogenèse, il apparaît que les profils d'expression sont similaires. Ceci suggère la possibilité d'une régulation de la reproduction par les IRPs soit directement sur la gonade, soit via des ganglions neurosécréteurs assurant une régulation centrale de la reproduction. La diminution de CIR dans l'aire gonadique des mâles à maturité sexuelle peut s'expliquer par le fait que le tissu de réserve, qui exprime également le récepteur, soit résorbé à ce stade (Jouaux *et al.*, 2012). En revanche, chez les femelles, l'expression du récepteur augmente au stade de maturité sexuelle, ce qui

confirmerait l'implication des voies des IRPs dans la maturation des ovocytes, comme pour d'autres espèces (Das et Arur, 2017; Wu et Brown, 2006).

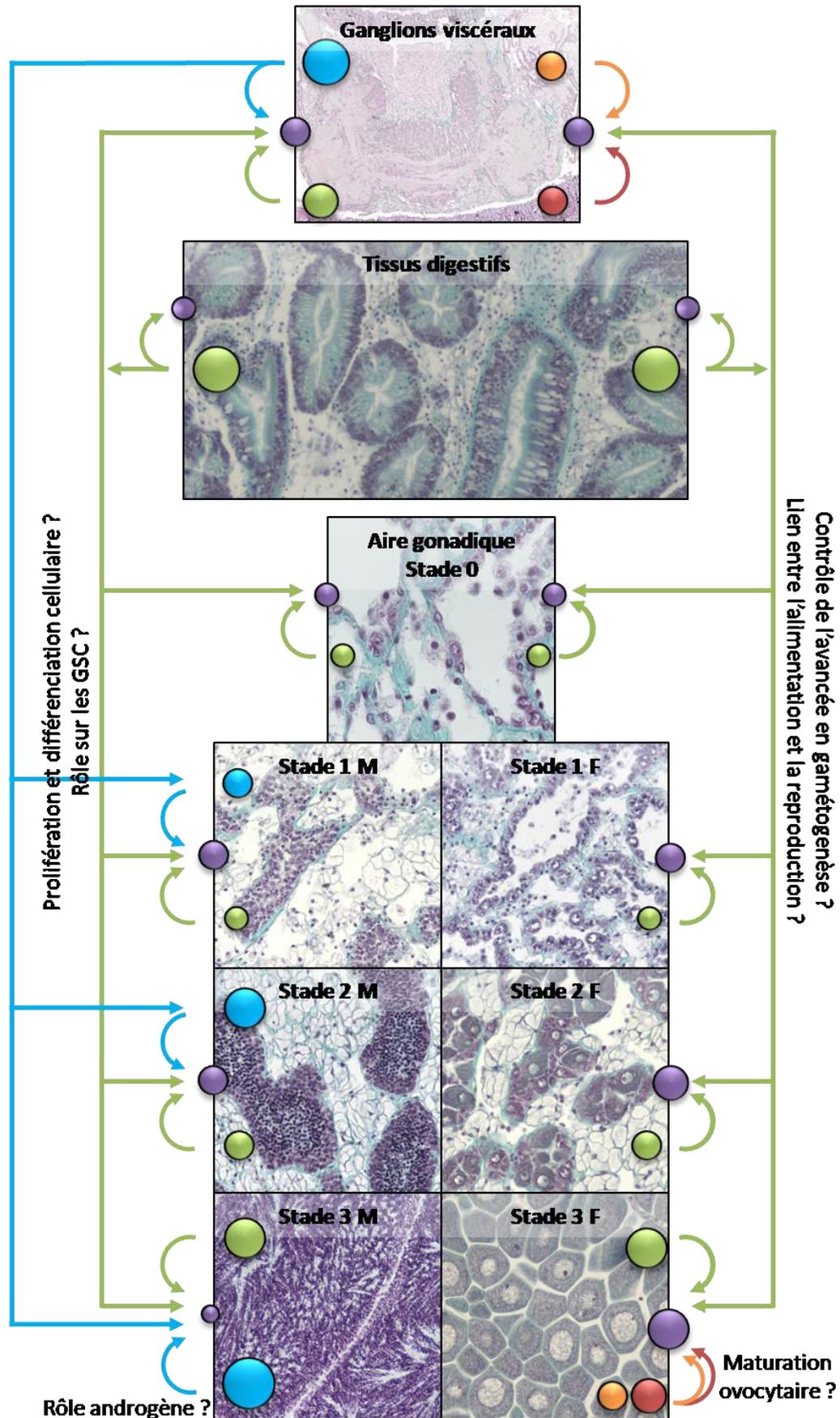
La plupart des effecteurs finaux des voies de signalisation de l'insuline sont retrouvés chez l'huître suggérant fortement une conservation des fonctions des IRPs. De plus, ces effecteurs sont potentiellement exprimés dans l'aire gonadique montrant la possible implication des voies IRPs dans les processus associés à la reproduction chez *C. gigas*. Toutefois, si les acteurs impliqués dans la glycolyse, la glycogénèse, les synthèses protéiques, la différenciation et la prolifération cellulaire ainsi que la lipogénèse sont présents dans la gonade, les acteurs des voies de l'apoptose et de l'antilipolyse n'y sont pas retrouvés. Certains effecteurs finaux de la glycolyse et la lipogénèse ont une expression prédite plus élevée dans l'aire gonadique femelle à maturité sexuelle par rapport aux mâles. C'est le cas de GK (glycolyse) et FAS (lipogénèse), ce qui confirmerait le rôle des IRPs dans le contrôle de la vitellogenèse chez *C. gigas*. Dans la gonade mâle, les effecteurs les exprimés sont associés au métabolisme du glycogène ainsi qu'aux voies de prolifération et différenciation cellulaire. Cette dernière est connue chez d'autres espèces pour être impliquée dans la régulation du compartiment souche et pourrait également être impliquée dans le fonctionnement de la niche germinale chez *C. gigas*.

Bilan du chapitre II

- ✓ Identification de six IRPs par criblage *in silico* des données génomiques et transcriptomiques.
- ✓ Trois structures identifiées selon le motif de la chaîne A, les ponts disulfures et l'homologie de séquences.
- ✓ Profils d'expressions des IRPs dans les ganglions nerveux, la glande digestive et l'aire gonadique.
- ✓ Rôles potentiels proposés pour chaque IRP : lien alimentation/reproduction, contrôle de la prolifération et différenciation cellulaire, maturation ovocytaire, rôle androgène.
- ✓ Expression de CIR dans les ganglions viscéraux et l'aire gonadique au cours du cycle de reproduction de *C. gigas* en accord avec le rôle potentiel des IRPs dans la reproduction.
- ✓ Identification des effecteurs des voies de signalisation et mise en évidence dans l'aire gonadique des effecteurs impliqués dans le métabolisme du glucose (réserve/mobilisation), les synthèses protéiques, la prolifération et la différenciation cellulaire.



Régulation de la gamétogenèse par les IRPs chez *C. gigas*



Chapitre III :

*Effet de l'alimentation sur les
cellules germinales précoces,
implication du signal insuline*

Chapitre III :

Effet de l'alimentation sur les cellules germinales précoces, implication du signal IRP

1. Introduction

Comme cela a été décrit en introduction générale, chez la plupart des Métazoaires, la gamétogenèse est étroitement dépendante du statut nutritionnel des individus, et notamment les étapes précoces puisque le fonctionnement de la niche germinale est soumis à des régulations dépendantes de l'apport nutritif. Les voies de signalisation des IRPs sont souvent impliquées dans la transmission de ce signal nutritionnel. Ce processus a été particulièrement bien décrit chez la drosophile, en combinant des conditionnements alimentaires et l'utilisation d'individus mutants pour les IRPs, le récepteur de l'insuline ou différents effecteurs des voies de signalisation (Hsu et Drummond-Barbosa, 2009, 2011; Shim *et al.*, 2013). Chez *C. elegans*, les IRPs interviennent dans le contrôle de nombreux processus physiologiques dont la reproduction. Ils agissent notamment sur le devenir des GSC (maintien et prolifération). En conditions d'alimentation favorables, le signal insuline est augmenté ce qui impacte la reproduction (pour revue voir Templeman et Murphy, 2018). Chez les vertébrés, l'insuline est un « nutrient sensor » reliant statut nutritionnel et fertilité (Dupont *et al.*, 2014).

Chez les Mollusques, les approches méthodologiques sont plus limitées. Toutefois, la possibilité de réaliser des conditionnements trophiques contrôlés, la connaissance des voies de signalisation des IRPs et la possibilité d'identifier les cellules germinales précoces permettent d'envisager l'étude du lien entre le statut nutritionnel et la reproduction pouvant impliquer les voies des IRPs.

1.1. Alimentation et impact sur la reproduction des mollusques

L'influence du régime alimentaire sur la reproduction chez les mollusques a fait l'objet de nombreuses études notamment chez les espèces d'intérêt aquacole pour lesquelles la production de géniteurs en éclosérie est un enjeu majeur. Ainsi, chez la coquille Saint-Jacques, *Pecten maximus*, la qualité du régime algal administré aux géniteurs impacte le pourcentage d'individus qui émettent leur gamètes, la qualité des ovocytes ainsi que le

développement embryonnaire après fécondation (Soudant *et al.*, 1996). En plus de la quantité et la qualité de la nourriture, la température ainsi que la saison à laquelle les coquilles sont soumises au conditionnement géniteurs sont également des facteurs influençant la reproduction (Andersen et Ringvold, 2000). Chez la moule *Mytilus edulis*, une étude ancienne a permis de mettre en évidence le lien entre l'apport nutritif et la production des gamètes (Bayne *et al.*, 1975). Pour cela, des rations nutritives contrastées ont été administrées à deux groupes d'individus pendant 14 jours. Après une mise à jeun pendant 24 heures puis une alimentation à base de microalgues marquées au ^{14}C , les deux groupes ont présenté des fécondités différentes. En condition de faible alimentation, le nombre d'ovocytes marqué au ^{14}C était plus élevé, ce qui indique qu'après une période d'apports nutritifs réduits, l'assimilation de la nourriture est principalement utilisée pour la reproduction et plus particulièrement pour le développement des gamètes chez les femelles.

Chez *C. gigas*, l'influence de l'alimentation sur la reproduction a été montrée en premier lieu pour la production de géniteurs matures. Lorsque l'alimentation est augmentée pendant la maturation, la fécondité des individus matures (ratio de la taille des tubules gonadiques sur la taille de l'aire gonadique) est augmentée de 60% (Muranaka et Lannan, 1984). Chavez Villalba (2001) a également montré que ce type de conditionnement favorise la maturation ovocytaire reflétée par une augmentation du diamètre ovocytaire moyen en condition de forte alimentation.

Concernant les étapes plus précoces de gamétogenèse, Berthelin *et al.* (2000a, 2000b) ont montré que l'avancée en gamétogenèse dépend des réserves préalablement établies dans le tissu de réserve sous forme de glycogène. De plus, Jouaux *et al.* (2013) ont réalisé un conditionnement trophique de 5 semaines en période automnale (début de gamétogenèse). Les résultats indiquaient que les huîtres nourries *ad libitum* étaient plus avancées en gamétogenèse alors que celles laissées à jeun présentaient un retard dans le déroulement de leur gamétogenèse. L'augmentation de l'expression du marqueur PCNA dans l'aire gonadique traduit un taux de proliférations goniales plus élevé chez les individus nourris et qui entraîne, en fin de gamétogenèse, un effort de reproduction plus intense. Parallèlement, un effet de l'alimentation est observé sur la restructuration du tissu de réserve de l'aire gonadique, ce qui a conduit les auteurs à formuler l'hypothèse que les réserves en glycogène préalablement établies conditionnaient les proliférations germinales (balance entre stockage et reproduction) (Jouaux *et al.*, 2013).

1.2. Les IRPs, lien entre statut nutritionnel et reproduction chez les mollusques

Les données disponibles chez les mollusques concernant le rôle potentiel des IRPs dans l'intégration des signaux liés à l'alimentation sont peu nombreuses et présentent majoritairement des hypothèses de fonctionnement.

Chez *L. stagnalis*, la mise à jeun entraîne un arrêt de la croissance, du stockage de glycogène et également de la reproduction (Geraerts, 1992; Geraerts *et al.*, 1992). Parallèlement, cette condition alimentaire provoque la disparition des transcrits de MIP II et la réduction du niveau d'expression de MIP III dans les LGC (Light Green Cells). Toutefois, ces travaux relativement anciens n'établissent pas directement le lien entre les MIPs et l'impact de l'alimentation sur la reproduction.

Chez *C. gigas*, Jouaux *et al.* (2013) ont travaillé sur l'implication du signal IRP dans l'intégration des signaux environnementaux sur la reproduction. L'implication du système insuline lors de conditionnements trophiques impactant la reproduction a été proposée après mesure, par qPCR, de l'expression du récepteur CIR dans l'aire gonadique (Jouaux *et al.*, 2013). L'expression de CIR apparaît significativement plus élevée lorsque les animaux sont mis à jeun. Ainsi, l'influence de l'alimentation sur la reproduction en lien avec le signal insuline a pu être observée chez l'huître.

Sachant que l'alimentation favorise les proliférations goniales chez l'huître, il est possible que cet effet passe par une régulation du fonctionnement du stock de GSC. Le travail présenté dans le chapitre I a permis d'identifier une population de cellules germinales précoces chez l'adulte contenant deux types cellulaires potentiels, les GSC et les progéniteurs. La caractérisation de ces types cellulaires permet maintenant d'étudier l'impact de l'alimentation sur la prolifération et la différenciation des GSC chez *C. gigas*. De plus, l'identification d'IRPs potentiellement impliqués dans la reproduction et d'effecteurs de ce signal dans le chapitre II, permet d'étudier le lien entre l'alimentation et le fonctionnement de la niche via ces effecteurs.

1.3. Objectifs

Ce chapitre aura d'abord pour objectifs de montrer l'influence de l'alimentation sur l'auto-renouvellement et la différenciation des GSC chez *C. gigas*. Pour cela, un

conditionnement alimentaire a été réalisé sur de jeunes huîtres lors de l'initiation de leur première gamétogenèse. Afin de vérifier l'augmentation des proliférations goniales, et en comparaison avec les travaux réalisés chez l'adulte, les expressions de PCNA et Oyvlg seront mesurées.

Ensuite, le rôle des IRPs dans l'intégration du signal alimentaire sera étudié en suivant l'expression de certains IRPs ainsi que du récepteur CIR, au cours du même conditionnement trophique.

2. Matériels et Méthodes

2.1.1. Animaux

Les huîtres en première gamétogenèse ont été fournies par le CFPPA de Coutances (Michel Leguillois). Ces animaux sont issus d'une ponte en éclosure, mesurent environ 2cm et présentent un poids moyen décoquillé de 70mg. Leur petite taille, donc leur moindre besoin en nourriture, facilite la réalisation des conditionnements trophiques dans nos structures expérimentales.

2.1.2. Conditionnement alimentaire

Les jeunes huîtres sont maintenues par lot de 50 individus dans 250mL d'eau de mer, filtrée à 10µm, pendant 14 jours. L'eau est changée tous les deux jours avec de l'eau de mer filtrée. Un lot de 50 individus a été nourri *ad libitum* à raison de 200 000 000 cellules/mL de microalgues *Isochrysis galbana*, distribuées en trois rations journalières. Un deuxième lot est resté à jeun pendant toute la durée du conditionnement. Des échantillonnages sont réalisés au début (T0) et à la fin (14 jours) du conditionnement, pour les individus nourris (N) et à jeun (AJ). Pour chaque échantillonnage, 12 individus sont prélevés, décoquillés et conservés à -80°C pour les suivis d'expression. Dix autres individus sont décoquillés puis immergés dans du fixateur histologique.

2.1.3. Extraction d'ARN et reverse transcription

Pour chacun des échantillonnages réalisés pour le suivi d'expression, quatre pools de trois individus ont été formés. Chaque pool est broyé dans du TriReagent® (Sigma Aldrich) puis incubé 5 minutes à température ambiante avant l'ajout de 100µL de 1-Bromo-3-Chloropropane (Sigma Aldrich). Les échantillons sont ensuite vortexés et centrifugés à

11 000g pendant 15 minutes à 4°C. La phase aqueuse est prélevée pour l'extraction d'ARN selon le kit Nucleospin RNA II (Macherey-Nagel). La transcription inverse est réalisée selon le protocole de Dheilly *et al.* (2012).

2.1.4. PCR quantitative

Les transcrits mesurés sont ceux des gènes PCNA, *oyvlg*, CIR, Cg-mip123 et Cg-ilp. Le gène de référence utilisé est EF1 α . Les couples d'amorces utilisés sont répertoriés dans le Tableau III-1). Les qPCR ont été réalisées en utilisant le mix 1X GoTaq SYBR Green (Promega), 8,85ng d'ADNc et 900nM de chaque amorce dans un volume final de 15 μ L. Chaque qPCR consiste en un cycle de 5 minutes à 95°C suivi de 45 cycles de 15 secondes à 95°C et 45 secondes à 60°C. La courbe de fusion a été obtenue par 80 cycles de 10 secondes en augmentant la température de 0,5°C à chaque cycle à partir de 55°C jusqu'à 95°C.

Tableau III-1 : Tableau récapitulatif des couples d'amorces utilisés.

Nom du transcrit	Sens (5' - 3')	Anti-sens (5' - 3')
PCNA	GCTGCTCGATTTAGACTGTGA	CTTGACTCCTTCCTTTGTACA
<i>oyvlg</i>	ACGTGCGTAAAGCCCAGTAT	CAATCCGCCGAACATTACTT
CIR	AACCCTGCCTGACAAAAATG	TACACGATCACCTCCCCTG
Cg-mip123	CAAGATCCGCTCCCTGGAAG	TCCTTTCTTCTTGCGGCGAT
Cg-ilp	TCGGGAGGCGCATAATCAAT	GCCGTGTCTGCAATACAAGC
EF1 α	ACGACGATCGCATTCTCTT	ACCACCCTGGTGAGATCAAG

2.1.5. Préparations histologiques

Chaque individu échantillonné pour l'histologie a été fixé dans du Davidson (10% de glycérol, 20% de formaldéhyde, 30% d'éthanol 95°, 30% d'eau de mer stérile et 10% d'acide acétique glacial) pendant 24 heures. Les échantillons sont ensuite déshydratés dans des concentrations croissantes d'éthanol puis transférés dans du butanol (Carlo Erba, France) avant d'être inclus en paraffine (Roth, France). Des coupes transversales sériées de 3 μ m d'épaisseur ont été réalisées au microtome dans toute la zone contenant la gonade.

2.1.6. Immunomarquages

Les coupes sont déparaffinées et l'immunohistologie enzymatique est réalisée selon Franco *et al.* (2010). L'anticorps primaire homologue polyclonal anti-Oyvlg de lapin est dilué

au 1/500^{ème}. L'anticorps secondaire utilisé est un polyclonal anti-lapin HRP produit chez le porc (P0399, DAKO Agilent Technologies), dilué au 1/100^{ème}. Les contrôles sont réalisés en omettant soit l'anticorps primaire soit l'anticorps secondaire. Les observations sont faites avec le microscope Nikon eclipse 80i et une caméra Nikon DS-Ri2.

2.1.7. Histologie quantitative

Pour trois individus par condition, une lame sur dix est colorée selon le protocole du trichrome Prenant-Gabe (Gabe, 1968) permettant de sélectionner les coupes présentant une section complète d'aire gonadique entourée de tissus digestifs. Pour chaque individu, deux coupes répondant à ces critères sont sélectionnées et colorées au Feulgen (Gabe, 1968). Ces coupes sont scannées en utilisant un scanner Aperio (objectif 40) et analysées d'après la méthode d'histologie quantitative développée dans le chapitre I. Les noyaux de toutes les cellules germinales de la gonade ont été repérés et entourés d'une ellipse pour déterminer le plus grand et le plus petit diamètre. Les ratios d'élongation sont calculés en divisant le plus petit diamètre par le plus grand ce qui permet de discriminer les cellules à noyaux irréguliers (ratio $\leq 0,8$) de celles à noyaux ronds (ratio $> 0,8$).

2.1.8. Analyses statistiques

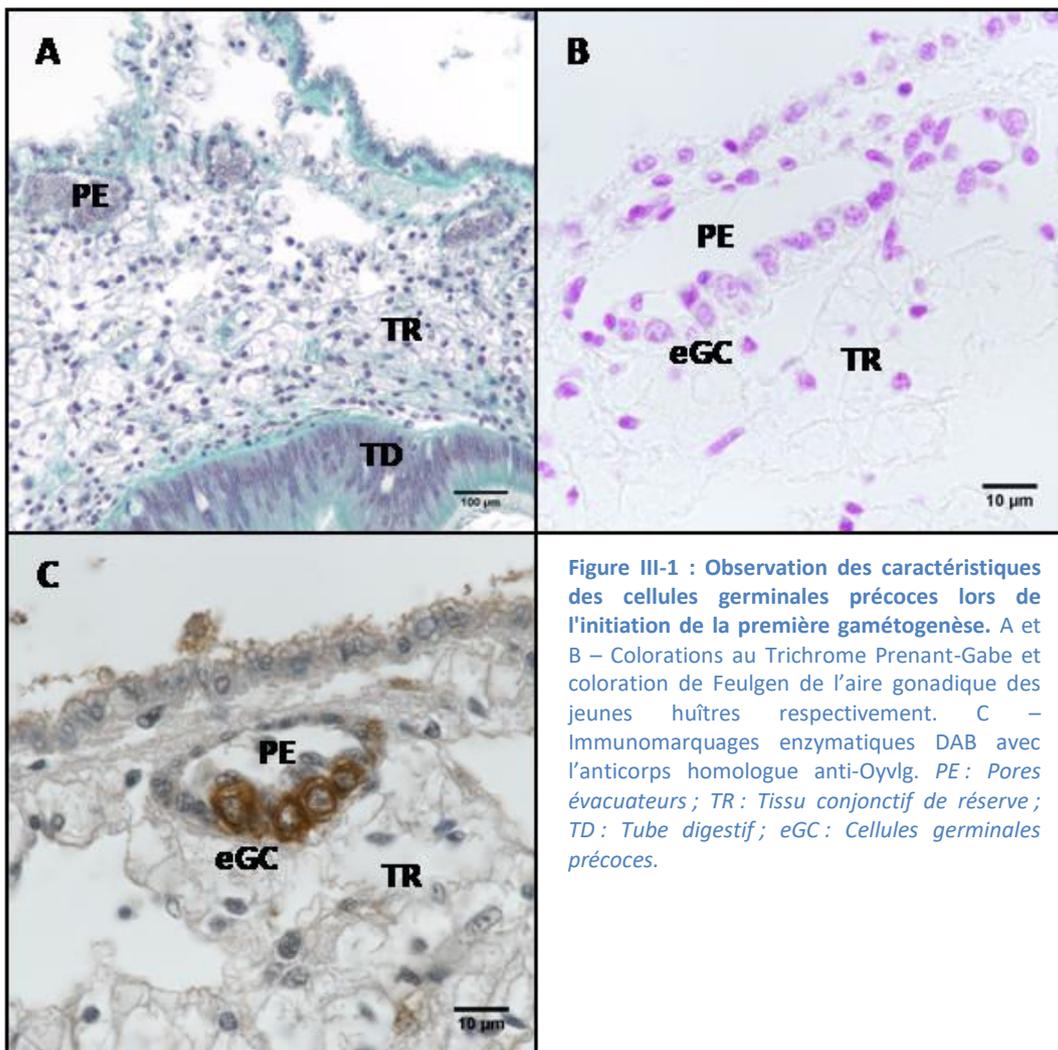
Les analyses statistiques ont été réalisées sur les logiciels R 3.4.3 et RStudio. Un test de Shapiro a été réalisé afin de vérifier la normalité de distribution des valeurs suivi soit d'un test post-hoc de Wilcoxon, pour les suivis d'expression, soit d'un test non-paramétrique de permutations, pour l'histologie quantitative, permettant de définir les différences significatives ($\alpha = 0,05$).

3. Résultats

3.1. Influence de l'alimentation sur le stock de GSC potentielles

3.1.1. Identification d'une population de cellules germinales précoces

Lors de l'initiation de la première gamétogenèse, les tubules gonadiques se développent du côté externe de l'aire gonadique, contre le manteau. Au niveau des pores évacuateurs, quelques cellules germinales sont présentes et accompagnées de cellules somatiques intragonadiques (Figure III-1A). Ces cellules germinales présentent toutes les caractéristiques des cellules germinales précoces à savoir, une chromatine à aspect décondensé (Figure III-1B) et un marquage cytoplasmique positif pour la protéine Oyvlg (Figure III-1C).



Pour étudier les populations de GSC et de progéniteurs potentiels, l'approche d'histologie quantitative développée chez l'adulte a été réalisée sur les jeunes huîtres à T0. Les mesures réalisées pour chacun des trois individus sélectionnés ne présentent aucune différence interindividuelle significative. Ainsi, tous les noyaux des cellules germinales, colorées au Feulgen, ont été mesurés et considéré comme une même population (Figure III-2). La population de cellules germinales précoces chez les jeunes huîtres est identifiée sur la base du plus grand diamètre nucléaire, compris entre 4,06 μm et 6,82 μm , comme caractérisé chez les adultes en deuxième et troisième gamétogenèses (Figure III-2A). De plus, d'après les ratios d'élongation calculés, les cellules germinales précoces présentes lors de l'initiation de la première gamétogenèse peuvent être discriminées en deux types cellulaires, comme chez l'adulte (Figure III-2B). Certaines cellules possèdent un noyau rond (ratio > 0,8) et sont de type progéniteurs, alors que les autres présentent un noyau irrégulier (ratio \leq 0,8) et sont de type GSC.

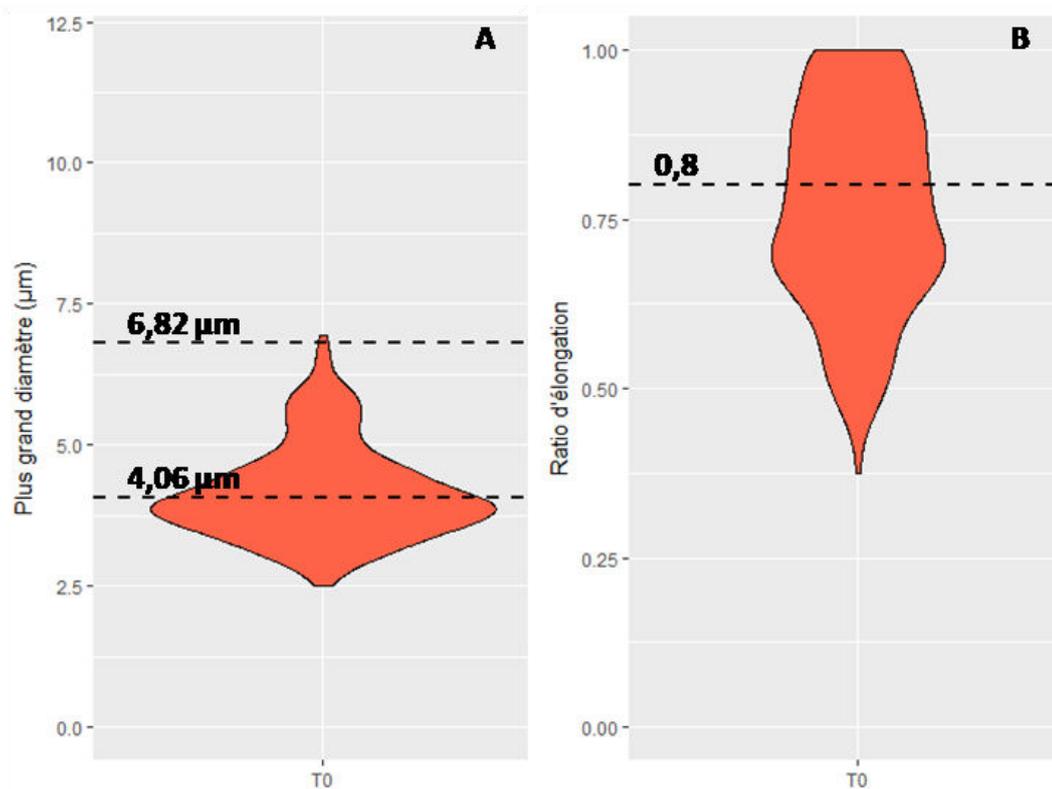


Figure III-2 : Identification d'une population de cellules germinales précoces chez les jeunes huîtres à l'initiation de leur première gamétogenèse. A – Répartition des cellules germinales mesurées à T0 (initiation de la première gamétogenèse) en fonction de leur plus grand diamètre. B – Répartition des ratios d'élongation pour les cellules germinales précoces dont le diamètre est compris entre 4,06 μm et 6,82 μm .

La population de cellules germinales précoces, identifiée aux stades quiescents des deuxième et troisième gamétogenèses (stades 0), a été comparée à celle trouvée à l'initiation de la première gamétogenèse (T0) (Tableau III-2). Pour chaque cycle de reproduction, trois individus ont été utilisés et ne présentent pas de variation interindividuelle significative. Ainsi, toutes les cellules germinales précoces (eGC) mesurées sont considérées comme appartenant à une même population. La proportion de GSC potentielles (noyaux irréguliers) diminue après la première gamétogenèse alors que celle de progéniteurs potentiels augmente (noyaux ronds).

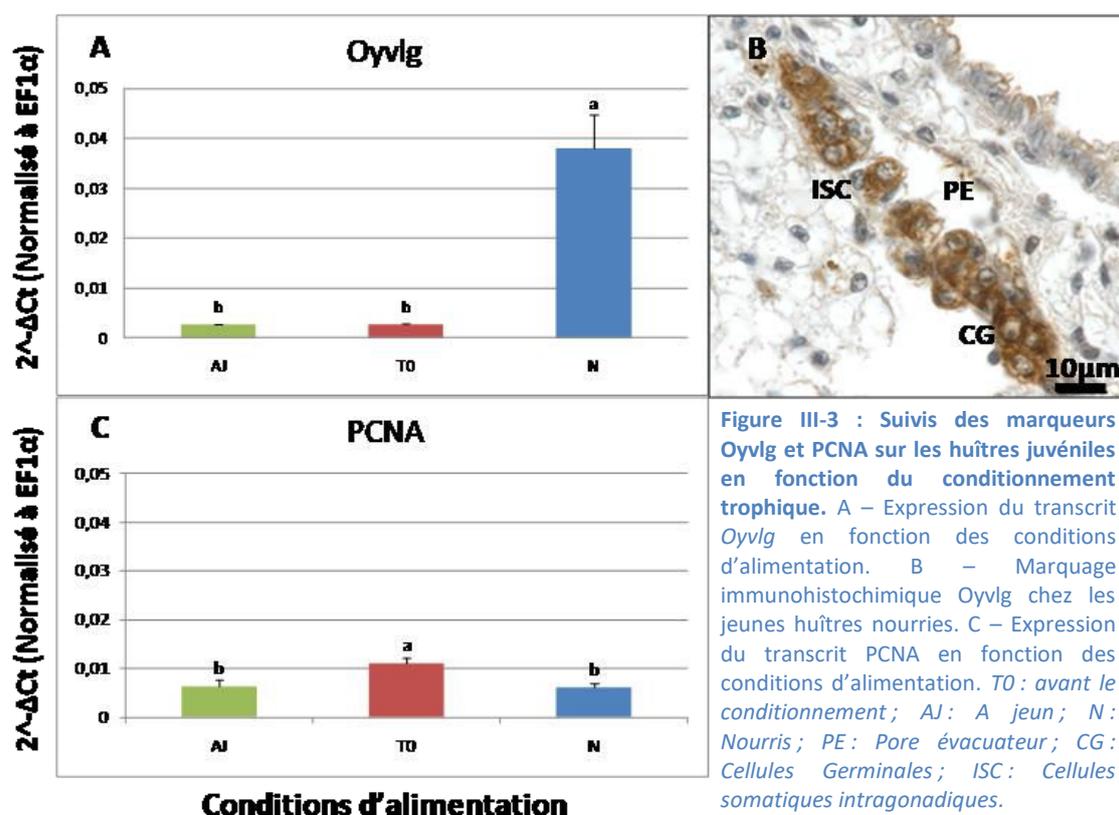
Tableau III-2 : Comparaison de la population de cellules germinales précoces chez des individus en première gamétogenèse et des individus en 2^{ème} et 3^{ème} gamétogenèse. eGC : Cellules Germinales précoces.

Cycle de reproduction	eGC	Noyaux irréguliers (GSC potentielles)	Noyaux ronds (progéniteurs potentiels)
T0 (1 ^{ère} gamétogenèse)	221	137 (62 %)	84 (38 %)
Stades 0 (2 ^{ème} et 3 ^{ème} gamétogenèse)	96	26 (27%)	70 (73%)

3.1.2. Validation du conditionnement trophique : suivi des marqueurs

Oyvlq et PCNA

Les expressions du marqueur de prolifération PCNA, du marqueur de cellules germinales VASA ont été mesurées dans l'animal entier (Figure III-3). Le marqueur germinale VASA apparaît significativement plus exprimé lorsque les jeunes huîtres sont nourries pendant 14 jours par rapport aux individus T0 ou maintenus à jeun (Figure III-3A). De même, le nombre de cellules positives pour Oyvlq est plus important chez les animaux nourris (Figure III-3B). Le marqueur de prolifération PCNA apparaît plus exprimé au début du conditionnement et aucune différence significative n'est observée entre les conditions « à jeun » et « nourris » (Figure III-3C). D'après les données transcriptomiques et immunohistochimiques, le conditionnement trophique permet d'observer une augmentation du nombre de cellules germinales dans la gonade des jeunes huîtres.



3.1.3. Effet de l'alimentation sur le compartiment souche potentiel

La discrimination des GSC et progéniteurs potentiels, après coloration au Feulgen, a été réalisée chez les jeunes huîtres en fonction des conditions d'alimentation. Ainsi, les noyaux

des cellules germinales ont été mesurés pour chaque condition d'alimentation (Figure III-4A). La population de cellules germinales précoces dont le diamètre nucléaire est compris entre $4,06\mu\text{m}$ et $6,82\mu\text{m}$ est retrouvée indépendamment des apports nutritifs. La répartition des ratios d'élongation de cette population de cellules germinales précoces présente des différences significatives selon les conditions trophiques (Figure III-4B). Les individus soumis à une forte alimentation (bleu) présente moins de noyaux irréguliers (ratio $\leq 0,8$), type GSC, et plus de noyaux ronds (ratio $> 0,8$), type progéniteurs, en comparaison avec les individus maintenus à jeun (vert).

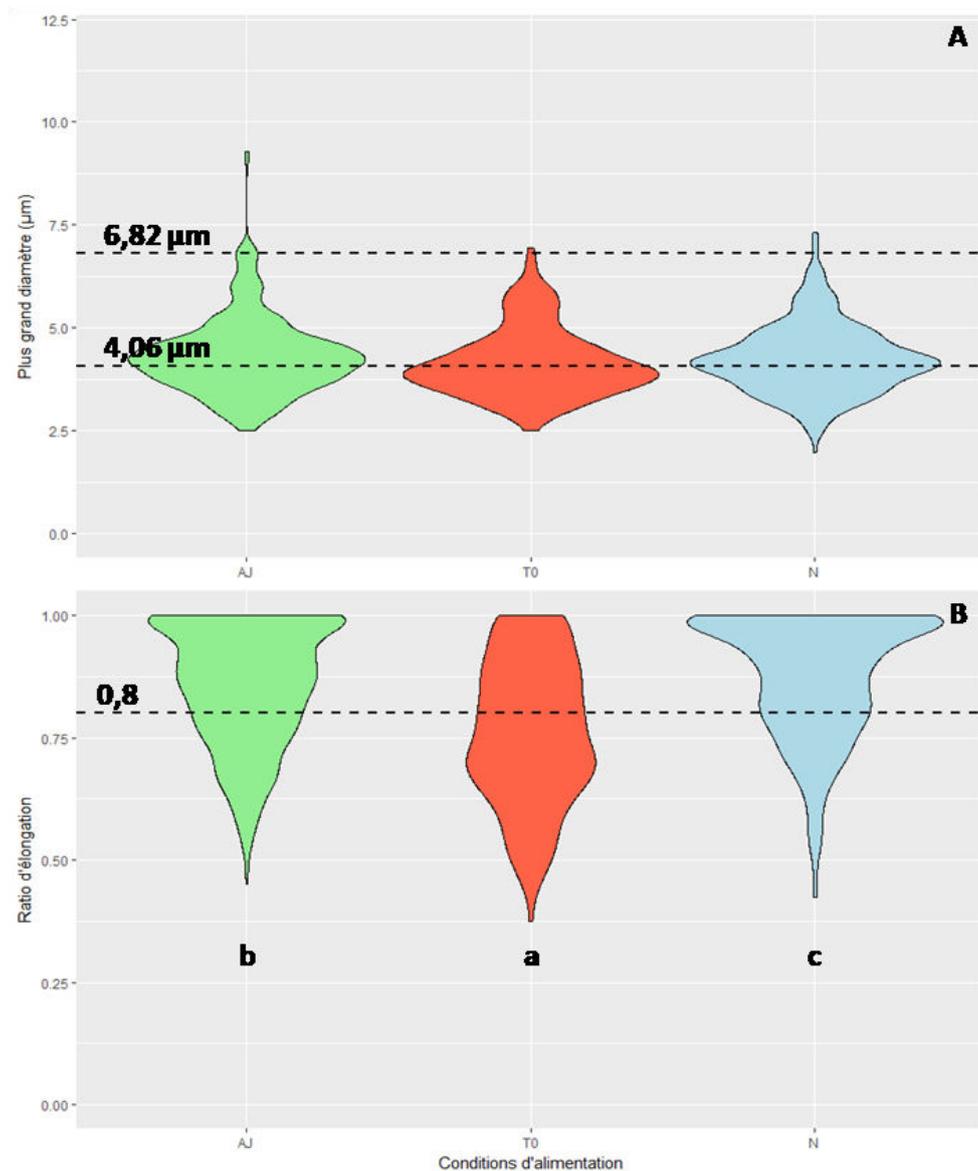


Figure III-4 : Représentation graphique de l'approche d'histologie quantitative chez les huîtres en première gamétogenèse. A – La distribution des cellules germinales est représentée selon la mesure du plus grand diamètre nucléaire en fonction du conditionnement alimentaire. B – Le ratio d'élongation est calculé pour chaque cellule germinale précoce (diamètres entre $4,06\mu\text{m}$ et $6,82\mu\text{m}$) en fonction du conditionnement alimentaire. TO : avant le conditionnement ; AJ : A jeun ; N : Nourris.

Les comptages réalisés (Tableau III-3) dénombrent les cellules germinales précoces des trois individus par condition, ces individus ne présentent par ailleurs, aucune variation interindividuelle significative. Ces comptages montrent que le nombre de cellules germinales est plus élevé après les 14 jours de conditionnement alimentaire, quelque soit la condition trophique. Cette augmentation est liée à une augmentation du nombre de cellules à noyaux ronds de type progéniteurs, alors que le nombre de cellules à noyaux irréguliers, de type GSC, reste inchangé. Les résultats de comptage montrent également que le nombre de progéniteurs est plus élevé lorsque les jeunes huîtres sont nourries par rapport aux individus laissés à jeun.

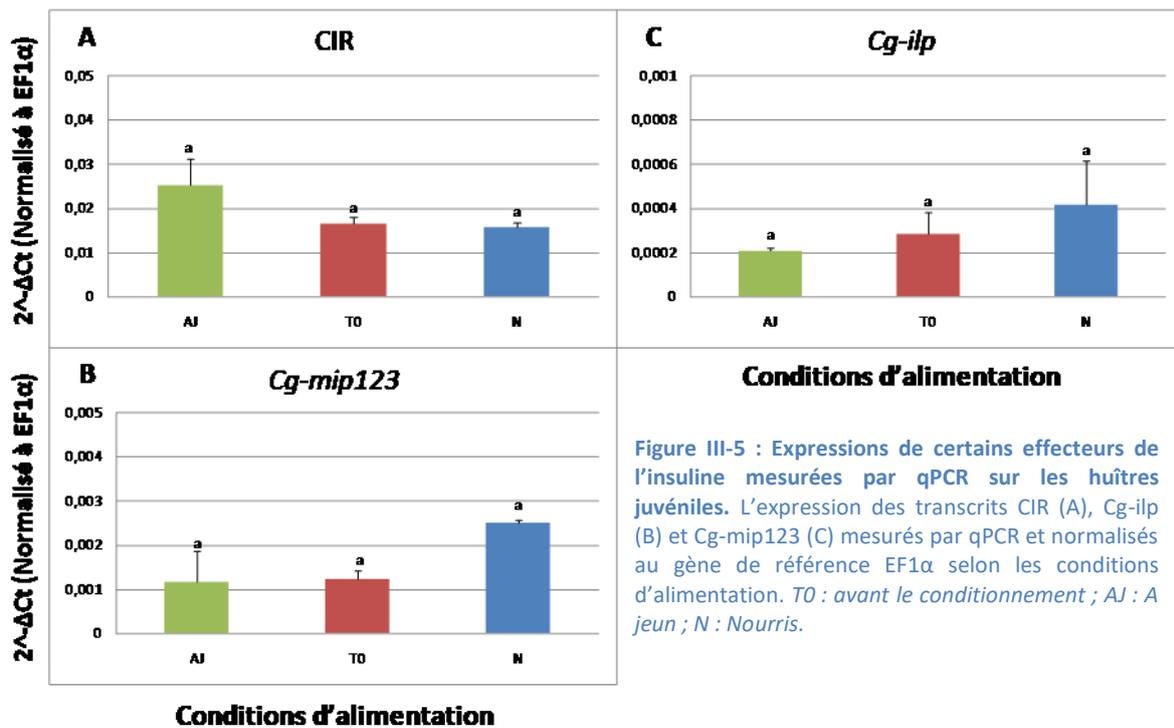
Les pourcentages de cellules à noyaux irréguliers (GSC potentielles) et à noyaux ronds (progéniteurs potentiels) ont également été comparés entre le début et la fin du conditionnement alimentaire. A T0, les cellules à noyaux irréguliers, type GSC, sont majoritaires (62%) alors qu'en fin de conditionnement ce sont les cellules à noyaux ronds, type progéniteurs, qui sont majoritaires quelque soit les conditions d'alimentation (69% et 72% pour les « à jeun » et « nourris » respectivement).

Tableau III-3 : Nombre de cellules germinales mesurées selon les conditions d'alimentation. Les nombres de cellules à noyaux irréguliers et celles à noyaux ronds ont été calculés parmi les cellules germinales totales et leur proportion est indiquée en pourcentage. T0 : avant le conditionnement ; AJ : A jeun ; N : Nourris ; eGC : Cellules Germinales précoces.

Conditions d'alimentation	eGC	Noyaux irréguliers (GSC potentielles)	Noyaux ronds (progéniteurs potentiels)
T0	221	137 (62 %)	84 (38 %)
AJ	388	119 (31 %)	269 (69%)
N	522	145 (28 %)	377 (72%)

3.2. Influence de l'alimentation sur les IRPs et CIR

Pour étudier le rôle du signal insuline lors du conditionnement alimentaire, les expressions du récepteur CIR ainsi que des transcrits *Cg-ilp* et *Cg-mip123* ont été suivies par qPCR (Figure III-5). Aucune différence significative n'a pu être observée pour les différents transcrits mesurés en fonction des conditions d'alimentation, même si des tendances générales semblent se dessiner.



4. Discussion

Ce travail a permis de caractériser les cellules germinales précoces chez l'huître dès la première gamétogenèse et de les comparer à celles identifiées dans le chapitre I. Les deux types cellulaires, GSC et progéniteurs potentiels identifiés au cours des 2^{ème} et 3^{ème} gamétogenèses, ont été retrouvées chez les jeunes huîtres, avec les mêmes caractéristiques morphologiques. Il est apparu que la proportion de GSC potentielles rapportée au nombre de cellules germinales mesurées est beaucoup plus élevée à TO (62% de type GSC) que chez les adultes (27%). Ceci démontre l'existence d'une population de GSC présente dès l'initiation de la première gamétogenèse au niveau de l'épithélium germinatif des pores évacuateurs et retrouvée lors des différents cycles de reproduction de *C. gigas* dans l'ensemble de la gonade. Ainsi, ces observations montrent que le compartiment souche potentiel est déjà en place dès la première gamétogenèse. Ce compartiment est ensuite retrouvé au cours des cycles de reproduction suivants, quels que soient le stade et le sexe.

Le conditionnement alimentaire réalisé chez de jeunes huîtres lors de l'initiation de leur première gamétogenèse a permis d'étudier l'influence de l'alimentation sur différents marqueurs. Le marqueur PCNA n'a pas montré d'augmentation des proliférations lorsque les huîtres sont nourries, à l'inverse de ce qui avait été observé chez les adultes (Jouaux *et al.*, 2013). Ceci peut être dû au fait que l'expression de PCNA est mesurée sur l'animal entier

dans notre étude, en raison de la petite taille des huîtres, ce qui peut entraîner une dilution du signal. La durée du conditionnement, réduite à deux semaines lors de ce travail, peut également être une explication à cette absence d'effet sur PCNA. Cependant, ce temps de conditionnement a tout de même permis d'observer une nette augmentation du marqueur spécifiquement germinale Vasa chez les individus nourris par rapport aux T0 et aux individus maintenus à jeun. Cette augmentation indique l'activité de proliférations goniales en lien avec l'alimentation comme cela avait également été montré chez l'adulte (Jouaux *et al.*, 2013) et valide le conditionnement.

L'influence de l'alimentation a été étudiée chez les huîtres en première gamétogenèse sur les deux types cellulaires identifiés, les GSC et les cellules progénitrices potentielles. Quelque soit le niveau d'alimentation, le nombre de GSC potentielles reste stable alors que le nombre de progéniteurs potentiels augmente au cours du conditionnement. Ainsi, la proportion de GSC diminue au cours de la première gamétogenèse, alors que la proportion de progéniteurs potentiels augmente. Ces données indiquent que le stock de GSC potentielles se maintient au cours de la gamétogenèse par auto-renouvellement. Les GSC en se divisant génèrent aussi des progéniteurs qui prolifèrent ensuite par mitoses goniales. Après 14 jours, les proportions des deux types cellulaires sont identiques à celles retrouvées chez les animaux en 2^{ème} et 3^{ème} gamétogenèse ce qui indique que le stock de GSC se maintient au fur et à mesure des gamétogenèses. Ces observations ont été faites sur une population de cellules germinales précoces issue de trois individus par condition trophique. Ainsi, il est nécessaire de confirmer ce fonctionnement avec une cohorte plus importante.

Pour la première gamétogenèse, l'apport nutritif ne semble impacter ni le nombre, ni la proportion de GSC potentielles. En revanche, on observe une augmentation du nombre de progéniteurs potentiels en conditions de forte alimentation par comparaison aux animaux à jeun. Ceci suggère que l'alimentation active la différenciation des GSC en progéniteurs et/ou la prolifération des cellules progénitrices mais que la proportion GSC/progéniteurs est maintenue constante, donc est très strictement contrôlée.

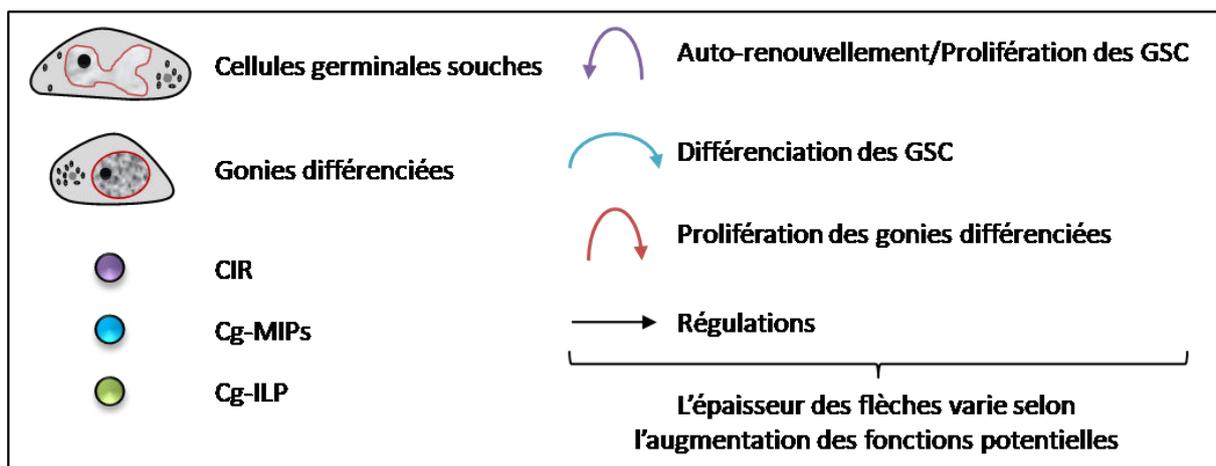
Concernant l'implication du signal insuline dans ce processus, les expressions de plusieurs effecteurs ont été mesurées sur des pools d'animaux entiers, du fait de leur petite taille. Cependant, les résultats ne sont pas significatifs et ne permettent pas d'interprétation. Ceci peut être lié à une expression très faible et localisée des gènes ciblés (notamment pour les IRPs). Or, la taille des animaux n'est pas compatible avec une dissection des tissus cibles.

Une alternative pourra consister à travailler sur des adultes ou sur des temps plus longs pour un futur conditionnement. Chez une autre huître, *Pinctada fucata*, l'expression du récepteur d'IRPs nommé Pfirr a été étudiée après l'ajout de l'inhibiteur d'IGF-1R, PQ401, afin d'évaluer son rôle fonctionnel dans le développement embryonnaire (Shi *et al.*, 2013). Ces résultats montrent qu'il est possible d'utiliser ce type d'inhibiteurs chez les huîtres. Au cours de ce travail de thèse, un essai préliminaire a été entrepris lors du conditionnement trophique des jeunes huîtres en administrant l'inhibiteur PQ401 par injection, cependant aucun effet significatif n'a pu être observé dans ces conditions expérimentales. D'autres inhibiteurs sont disponibles et spécifiques de certaines voies de signalisation (LY294002 et Wortmannin pour la voie PI3K, Rapamycine pour la voie mTOR et PD98059 pour voie MAPK) et pourront être également testés dans de futures expérimentations.

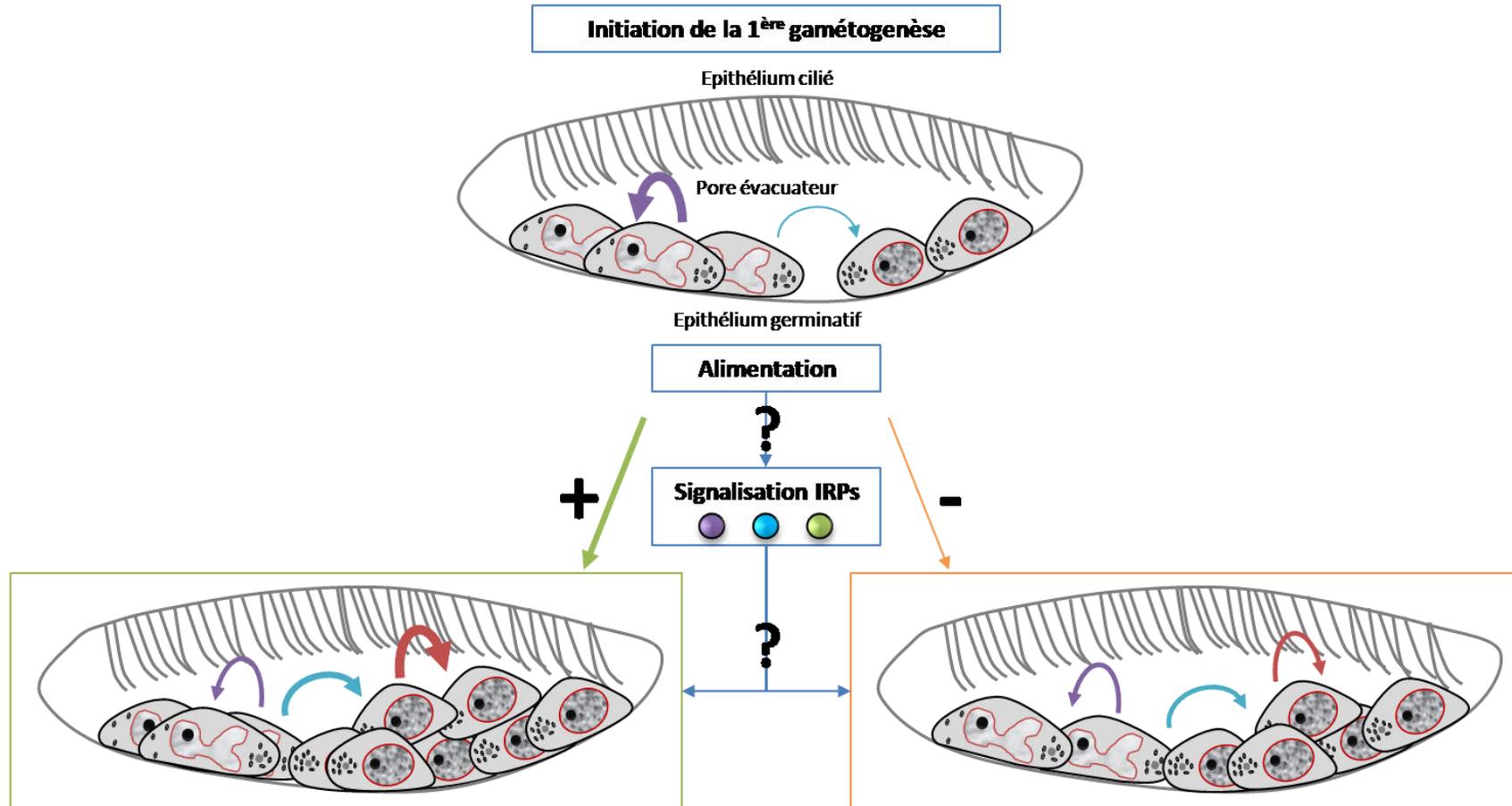
En conclusion, les résultats obtenus dans ce travail indiquent clairement que l'alimentation influence la différenciation et la prolifération des GSC et ce dès la première gamétogenèse. L'impact de l'alimentation porte directement sur la différenciation des GSC en gonies différenciées. L'implication du signal des IRPs comme intégrateur des signaux nutritionnels n'a pas pu être mise en évidence lors de ce premier essai de conditionnement mais il pourra être envisagé de reprendre ces conditionnements soit sur des temps plus longs, soit sur des animaux de plus grande taille pour travailler sur les tissus cibles.

Bilan du chapitre III

- ✓ Identification dès la première gamétogenèse des deux types de cellules germinales précoce, GSC et progéniteurs potentiels, déjà décrits chez l'adulte.
- ✓ Mise en évidence de l'augmentation du nombre de gonies différenciées en condition d'alimentation favorable.
- ✓ Nécessité de refaire des expérimentations pour étudier le rôle des voies de signalisation des IRPs.



Impact de l'alimentation sur la niche germinale en première gamétogenèse



*Conclusion générale et
perspectives*

Conclusion générale et perspectives

1. Identification des acteurs de la niche germinale

Au cours de cette thèse, la niche germinale potentielle a été identifiée chez l'huître *Crassostrea gigas* en combinant des critères morphologiques et moléculaires spécifiques des GSC. L'éloignement phylogénétique, entre l'huître et les espèces qui ont fait l'objet de recherche de marqueurs de GSC, rend difficile l'utilisation de marqueur moléculaire, c'est pourquoi une approche originale d'histologie quantitative a été développée. Cette approche est basée sur les principes de stéréologie et combine la localisation des GSC (en bord de tubule gonadique), l'aspect décondensé de la chromatine dans leur noyau et le marqueur germinale Oylg. Une population de cellules germinales précoces commune aux mâles et aux femelles a été identifiée par l'étude en histologie quantitative. Les mesures des noyaux ont permis de discriminer deux types cellulaires, les GSC potentielles et leurs cellules filles, les progéniteurs potentiels.

Des acteurs du microenvironnement des GSC ont ensuite pu être caractérisés. Ainsi, des cellules somatiques accompagnatrices, proches des GSC et progéniteurs potentiels, ont été identifiées de même que des cellules myoïdes à l'extérieur des tubules gonadiques, le long de la lame basale. Concernant les cellules somatiques accompagnatrices, l'utilisation d'un anticorps hétérologue anti-BMP2/4 a révélé un marquage cytoplasmique positif de ces cellules. Cet anticorps peut cibler différents facteurs à savoir, Cg-mGDF, Cg-Dpp et Cg-Gbb, tous appartenant à la famille des TGF- β , suggérant une implication de cette voie de signalisation dans le fonctionnement de la niche germinale chez *C. gigas*. L'organisation de la niche germinale chez *C. gigas* ressemble à celles identifiées chez les vertébrés, de type épithélial, puisqu'aucune cellule somatique n'a été observée entre les GSC et la lame basale.

Afin de confirmer l'identité des GSC caractérisées lors de ce travail, il sera nécessaire de développer des approches pour les isoler puis les maintenir et les faire proliférer en culture cellulaire. Dans un second temps, des approches de transplantations dans des individus receveurs pourront être envisagées afin de suivre les GSC potentielles et confirmer leur capacité à engendrer la lignée germinale.

1.1. Localisation des cellules souches dans la gonade

1.1.1. GSC et niche germinale potentielles

L'identification de GSC a été réalisée à partir de coupes histologiques transversales entières d'huître. L'approche d'histologie quantitative permet de localiser précisément les GSC potentielles dans la gonade sur ces coupes.

En première gamétogenèse, la gonade est restreinte aux pores évacuateurs, du côté du manteau. Ces pores évacuateurs se développent lors de la première gamétogenèse et des suivantes afin de former les tubules gonadiques qui s'étendent dans l'aire gonadique jusqu'au niveau des tissus digestifs. La population de cellules germinales précoces, incluant les GSC et progéniteurs potentiels, a été retrouvée dans les pores évacuateurs lors de cette première gamétogenèse. Cette observation confirme que les GSC sont établies dans la gonade dès que les huîtres sont capables de reproduction. Lors des gamétogenèses suivantes les GSC potentielles sont réparties dans toute la gonade, au sein des pores évacuateurs ainsi que dans les tubules gonadiques. Le fait que plusieurs niches germinales soient observées dans toute la gonade suggère que les GSC potentielles colonisent les tubules gonadiques au fur et à mesure de leur formation. De plus, elles restent présentes tout au long des cycles de reproduction ce qui leur permet de redonner une nouvelle lignée germinale chaque année.

1.1.2. Cellules souches pluripotentes

D'autres cellules à potentiel souche ont été observées dans la gonade d'huîtres, en deuxième et troisième gamétogenèses, grâce à des anticorps hétérologues ciblant des marqueurs de pluripotence, KLF4 et Oct4. Le marqueur Oct4 n'étant retrouvé qu'au stade mâle de proliférations goniales, il est possible que l'anticorps hétérologue ne cible pas les GSC potentielles. Concernant le marqueur KLF4, il est retrouvé dans l'épithélium germinatif des pores évacuateurs seulement au stade quiescent de gamétogenèse. Ceci suggère que les cellules germinales présentes dans les pores évacuateurs au premier stade de gamétogenèse seraient pluripotentes. Puisqu'elles ne sont pas retrouvées aux stades suivants, il est probable qu'elles se différencient et que leur potentialité change lors du cycle de gamétogenèse. Par conséquent, ces cellules positives pour KLF4 seraient plus amont que les

GSC identifiées. Malheureusement, les anticorps anti-KLF4 et anti-Oyvlg sont tous les deux produits chez le lapin ce qui rend la possibilité d'un co-marquage impossible à ce jour.

1.1.3. Cellules en prolifération

L'expérimentation BrdU initiée au cours de cette thèse montre des cellules capables de proliférer dans les pores évacuateurs au stade quiescent de gamétogenèse. Ces cellules sont présentes au niveau de l'épithélium germinatif mais également au niveau de l'épithélium cilié. Il est fort probable que ces cellules constituent deux types cellulaires différents. Concernant les cellules présentes dans l'épithélium germinatif, elles semblent être dans un état de quiescence plus marqué que les GSC potentielles qui n'ont pas été marquées par le BrdU pour ce temps long d'incorporation. Il conviendra d'étudier des temps plus courts d'incorporation du BrdU afin de retrouver un marquage des GSC potentielles caractérisées chez le juvénile et l'adulte. Afin de confirmer les potentiels somatique et/ou germinatif de ces cellules en prolifération, des co-marquages devront être réalisés.

1.2. Hypothèse de fonctionnement des cellules souches gonadiques

Les résultats obtenus lors de cette thèse, confrontés aux connaissances bibliographiques, permettent de proposer une hypothèse concernant le fonctionnement des cellules souches gonadiques chez l'huître.

Chez l'embryon, des PGC potentielles avaient été identifiées très tôt dans le développement par hybridation *in situ* avec le marqueur Oyvlg, suggérant un modèle préformistique pour l'origine de la lignée germinale chez l'huître (Fabioux *et al.*, 2004a). Cependant, le potentiel somatique de ces PGC n'a pas été recherché. Il est donc possible d'envisager que ces PGC potentielles soient en réalité proches de cellules souches type PriSC (Solana, 2013) qui seraient présentes tout au long de la vie de l'animal. D'après ce scénario, les PriSC sont restreintes à deux groupes de cellules en fin de développement qui pourraient former la gonade par la suite. Les cellules positives pour le marquage BrdU au niveau de l'épithélium cilié des pores évacuateurs pourraient constituer les types cellulaires engendrés par ces PriSC. Afin de confirmer cette hypothèse, des co-marquages, notamment avec l'anticorps anti-Oyvlg, devront être réalisés. De plus, l'identification de ces mêmes cellules BrdU positives en première gamétogenèse est nécessaire pour confirmer la continuité de

leur présence dans la gonade. Enfin, dans le but d'identifier les cellules engendrées par les BrdU positives, il est nécessaire de poursuivre le suivi cinétique.

En considérant que des PriSC engendrent des cellules à potentiel germinatif capables de proliférations chez l'huître, il est possible que les GSC potentielles identifiées lors de cette thèse soit issues de ces PriSC. Ceci signifierait que, dès que l'huître est capable de reproduction, les PriSC colonisent l'épithélium germinatif des pores évacuateurs et se différencient en GSC, ce qui expliquerait la forte proportion de GSC à l'initiation de la première gamétogenèse. Lors de l'avancée en première gamétogenèse, ces GSC se différencient en progéniteurs ou gonies afin d'engendrer les lignées germinales mâles et femelles. Ceci est illustré par le fait que le pourcentage de progéniteurs augmente 14 jours après l'initiation de la première gamétogenèse. En parallèle, certaines GSC s'auto-renouvellent afin de perdurer au cours des différents cycles de reproduction. Ceci explique que le pourcentage de GSC (environ 30%) soit stable au cours des gamétogenèses.

1.2.1. Régulations des GSC potentielles

Le fonctionnement du stock de GSC potentielles avec celui des progéniteurs potentiels ainsi que la pluripotence des cellules souches gonadiques est bien sûr soumis à différentes régulations. Concernant les pourcentages GSC/progéniteurs, il est probable qu'un dialogue soit établi entre ces types cellulaires afin de réguler leurs proportions qui restent stables au cours des différentes gamétogenèses. De plus, les acteurs TGF- β sont probablement impliqués dans la régulation de ces cellules germinales précoces, en accord avec le marquage BMP2/4 dans les cellules somatiques accompagnatrices. Cependant, il est fort probable que d'autres voies de signalisation soient impliquées.

2. Etude de la signalisation des IRPs

2.1. Identification des IRPs et analyse phylogénétique

Au cours de cette thèse, plusieurs IRPs ont été identifiés chez *C. gigas*. Leur caractérisation a permis de mettre en évidence la présence de ligands de types β et γ chez les mollusques. Ceci avait déjà été observé chez *L. gigantea* (Veenstra, 2010) mais aucune analyse phylogénétique n'avait été faite. L'apport phylogénétique de l'étude des IRPs de *C. gigas* a permis de proposer l'ILP de type γ comme forme ancestrale aux Bilatériens. L'un des

IRPs de type β commun à *L. gigantea* et *C. gigas* (MILP7), retrouvé chez *L. stagnalis* et *A. californica*, présente une forte homologie avec DILP7 et l'analyse phylogénétique a montré que les orthologues ILP7 étaient antérieurs à la séparation des Lophotrochozoaires et des Ecdysozoaires. Enfin, une recherche bibliographique a montré que les MIP, initialement considérés comme spécifiques des mollusques et identifiés chez *C. gigas*, sont en réalité retrouvés chez d'autres Lophotrochozoaires. L'identification des IRPs a été réalisée *in silico* par criblage des bases de données, il sera donc intéressant de vérifier la présence de ces peptides par analyse en spectrométrie de masse dans les différents tissus, notamment la gonade.

2.2. Effecteurs du signal

Comme pour les autres invertébrés, un seul récepteur des IRPs a été identifié chez *C. gigas* (Gricourt *et al.*, 2003). Au cours de cette thèse, la séquence complète a pu être retrouvée, permettant d'identifier la région de liaison des ligands. Le suivi d'expression de CIR a montré l'implication de cet effecteur dans les ganglions viscéraux, suggérant une régulation centrale, de la même façon que pour les Cg-MIPs. Le récepteur CIR est également exprimé dans l'aire gonadique lors des stades précoces de gamétogenèse mais également au stade de maturité sexuelle chez les femelles. Ces résultats suggèrent une régulation locale de différents aspects de la gamétogenèse tels que la prolifération et la différenciation cellulaire ainsi que la maturation ovocytaire. Ces différentes fonctions sont potentiellement relayées par différents ligands. Ainsi, la liaison des IRPs au récepteur CIR devra être démontrée par désorphanisation du récepteur. Certaines techniques de transfert d'énergie par résonance peuvent être utilisées et permettent l'observation du récepteur dans les cellules en microscopie confocale (Chung, 2017). Cette observation permettrait d'identifier le type cellulaire contenant le récepteur afin de savoir si les voies des IRPs sont relayées par les cellules du tissu de réserve et/ou agissent directement sur les cellules germinales. Il sera également intéressant de rechercher si d'autres types de récepteurs peuvent être activés par les IRPs identifiés, comme c'est le cas chez les vertébrés par exemple.

2.2.1. Fonctions potentielles des IRPs

Des rôles potentiels, en lien avec la reproduction, ont été proposés pour les différents IRPs identifiés chez *C. gigas*, d'après le suivi de leurs expressions. Ainsi, Cg-MIP4 et Cg-MILP7

pourrait être impliqués dans le développement embryonnaire mais également participer au processus de maturation ovocytaire. Ce dernier rôle est également suggéré après identification de la majorité des effecteurs des voies de signalisation des IRPs chez *C. gigas*. En effet, les voies de glycolyse, glycogénèse et lipogénèse sont conservées et les deux dernières semblent spécifiques de la gamétogenèse femelle. Concernant Cg-MIP1, Cg-MIP2 et Cg-MIP3, leur expression spécifique des mâles suggère un rôle androgène, passant éventuellement par une régulation locale dans l'aire gonadique des mâles à maturité sexuelle. De plus, ils présentent une expression très élevée dans les ganglions viscéraux ce qui pourrait refléter leur implication dans une régulation centrale chez l'huître. Enfin, d'après les expressions de Cg-ILP, il constitue le meilleur candidat pour faire le lien entre l'alimentation et la gamétogenèse. Puisqu'il est exprimé dès le stade quiescent de gamétogenèse, il est possible qu'il intervienne dans la prolifération et la différenciation des cellules germinales, et potentiellement des GSC. En effet, les effecteurs finaux de ces fonctions des voies des IRPs ont également été retrouvés chez *C. gigas*. Afin de vérifier quelles voies des IRPs sont précisément impliquées dans la gamétogenèse, notamment dans le fonctionnement du compartiment souche, l'utilisation d'inhibiteurs de ces voies ou encore d'ARN interférence pourra être envisagée. Il sera intéressant d'étudier, selon la voie bloquée, si un arrêt de la gamétogenèse, de la mise en réserve ou encore de la maturation ovocytaire est observé. De même, l'impact de l'inhibition des voies de signalisation pourra être observé sur les acteurs de la niche germinale chez *C. gigas* et comparé avec l'influence de l'alimentation sur cette dernière.

2.3. Hypothèse de fonctionnement des IRPs dans la gamétogenèse

D'après les résultats obtenus concernant la signalisation des IRPs, il est fort probable que ce signal soit impliqué dans une régulation centrale par l'intermédiaire des ganglions viscéraux comme le suggère l'expression élevée des Cg-MIPs dans ce tissu. L'expression de CIR retrouvée dans les ganglions et l'aire gonadique pourrait permettre de transmettre cette régulation centrale et d'influencer plusieurs processus physiologiques, incluant la reproduction. Cette hypothèse est supportée par le fait que les profils d'expression de CIR sont similaires dans les ganglions viscéraux en comparaison avec l'aire gonadique. Une

régulation locale dans l'aire gonadique est également possible d'après l'expression des effecteurs exprimés localement dans ces tissus. Cette expression est similaire chez les mâles et les femelles aux stades quiescent, de proliférations goniales et de différenciation germinale. L'accumulation de glycogène dans le tissu de réserve à ces stades peut être en lien avec les voies conservées des IRPs et impliquées dans le métabolisme du glucose. Parallèlement, les effecteurs finaux des voies de prolifération et différenciation cellulaires étant conservés chez *C. gigas*, il est possible que les IRPs influencent l'avancée en gamétogenèse en agissant directement sur les cellules intragonadiques. Aux stades de maturité sexuelle, le récepteur CIR est moins exprimé chez les mâles que chez les femelles. A ce stade, le tissu de réserve est régressé et les ovocytes sont matures. Il est probable que certains IRPs interviennent dans ce processus de maturation ovocytaire. Cg-ILP est quant à lui particulièrement exprimé dans les tissus digestifs, ce qui suggère son implication possible dans l'intégration des signaux nutritifs. Il pourrait faire le lien entre la disponibilité de l'énergie issue de l'alimentation et l'effort de reproduction, avec un effet possible sur le fonctionnement de la niche germinale.

3. Impact de l'alimentation sur le compartiment souche

L'influence de l'alimentation sur le compartiment souche a été étudiée chez des animaux en première gamétogenèse. Dans ce contexte, certains effecteurs des voies des IRPs ont également été étudiés dans le but de faire le lien entre l'alimentation, les IRPs et les GSC chez l'huître. A ce jour, l'implication des IRPs dans l'intégration de l'état nutritionnel des individus n'a pas pu être mise en évidence. Cependant, l'alimentation a un impact direct sur les GSC dès le moment où l'huître est capable de se reproduire. Le fait que l'alimentation influence le fonctionnement du stock de GSC dès l'initiation de la première gamétogenèse suggère que la mise en réserve du glycogène se fait en parallèle du fonctionnement du compartiment souche. Il est possible que l'alimentation ait un effet sur ces deux processus.

3.1. Hypothèse de fonctionnement de la niche germinale selon le statut nutritionnel

Après l'établissement des GSC potentielles dans les pores évacuateurs, leur pourcentage diminue en faveur des progéniteurs potentiels, ce qui suggère la différenciation des GSC et donc l'avancée en gamétogenèse. En conditions favorables d'alimentation, il semblerait que

les GSC soient recrutées pour se différencier en progéniteurs et que, ces derniers, prolifèrent de façon intensive. Une fois le stock de gonies formé, ces cellules pourront entrer en méiose afin d'achever la gamétogenèse. Si le taux de différenciation des GSC est plus élevé, il semblerait que leur auto-renouvellement le soit également puisque la proportion de GSC potentielles est stable entre la première gamétogenèse et les suivantes. En conditions d'alimentation défavorables, le recrutement de GSC potentielles est probablement plus faible et les progéniteurs potentiels se divisent de façon moins intense comme le suggère la réduction de leur nombre. Cependant, la proportion de GSC potentielles ne change pas, ce qui suggère que des régulations supplémentaires puissent participer au contrôle du stock de GSC. Cette hypothèse est probable considérant le mode de vie des huîtres. En effet, puisqu'elles sont fixes, elles doivent s'adapter aux changements environnementaux. Ainsi, il est cohérent que plusieurs voies de signalisation puissent intervenir dans le maintien du stock de GSC, même en conditions défavorables, et permettre une reproduction des individus ajustée en fonction des conditions.

A l'issue de ce travail, il est nécessaire de continuer à étudier l'effet des conditions trophiques sur les étapes précoces de gamétogenèse. Ainsi en repartant de conditionnements trophiques, sur des temps plus longs, des animaux nourris et à jeun pourront être comparés en terme d'expression des effecteurs du signal IRP identifiés dans ce travail, particulièrement Cg-MIPs et Cg-ILP, très probablement impliqués dans le lien alimentation/reproduction. Le fait de pouvoir discriminer le compartiment souche du compartiment prolifératif de l'épithélium germinatif ouvre des perspectives d'études plus globales de suivi d'expression d'effecteurs impliqués dans le maintien de l'état souche ou dans la différenciation (RNA seq). Ces approches permettront d'aller plus loin sur la compréhension de la régulation par l'alimentation des étapes précoces de la gamétogenèse chez l'huître *C. gigas*.

Bibliographie

A

Agrogiannis, G.D., Sifakis, S., Patsouris, E.S., and Konstantinidou, A.E. (2014). Insulin-like growth factors in embryonic and fetal growth and skeletal development (Review). *Mol. Med. Rep.* *10*, 579–584.

Ahmad, A., Strohbuecker, S., Tufarelli, C., and Sottile, V. (2017). Expression of a SOX1 overlapping transcript in neural differentiation and cancer models. *Cell. Mol. Life Sci.* *74*, 4245–4258.

Akberdin, I.R., Omelyanchuk, N.A., Fadeev, S.I., Leskova, E., Oschepkova, E.A., Kazantsev, F. V, and Matushkin, Y.G. (2018). Pluripotency gene network dynamics : System views from parametric analysis. *PLoS One* 1–24.

Alié, A., Leclère, L., Jager, M., Dayraud, C., Chang, P., Le Guyader, H., Quéinnec, E., and Manuel, M. (2011). Somatic stem cells express Piwi and Vasa genes in an adult ctenophore: Ancient association of “germline genes” with stemness. *Dev. Biol.* *350*, 183–197.

Altincicek, B., and Vilcinskis, A. (2007). Analysis of the immune-related transcriptome of a lophotrochozoan model, the marine annelid *Platynereis dumerilii*. *Front. Zool.* *4*.

Andersen, S., and Ringvold, H. (2000). Seasonal differences in effect of broodstock diet on spawning success in the great scallop. *Aquac. Int.* *8*, 259–265.

Andersen, A.S., Kjeldsen, T., Wiberg, F.C., Vissing, H., Schaffer, L., and Rasmussen, J.S. (1992). Identification of determinants that confer ligand specificity on the insulin receptor. *J. Biol. Chem.* *267*, 13681–13686.

Andersson, E.R., Sandberg, R., and Lendahl, U. (2011). Notch signaling: simplicity in design, versatility in function. *Development* *138*, 3593–3612.

Assis, L.H.C., Crespo, D., Morais, R.D.V.S., França, L.R., Bogerd, J., and Schulz, R.W. (2016). INSL3 stimulates spermatogonial differentiation in testis of adult zebrafish (*Danio rerio*). *Cell Tissue Res.* *363*, 579–588.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development on SOX2 function. *Genes Dev.* *17*, 126–140.

B

Badisco, L., Van Wielendaele, P. V., and Broeck, J. Vanden (2013). Eat to reproduce: A key role for the insulin signaling pathway in adult insects. *Front. Physiol.* *4* AUG, 1–16.

Baguña, J., Salo, E., and Auladell, C. (1989). Regeneration and pattern formation in planarians III. Evidence that neoblasts are totipotent stem cells and the source of blastema cells. *Development* *107*, 77–86.

Bamezai, S., Rawat, V.P.S., and Buske, C. (2012). Concise review: The Piwi-piRNA Axis: Pivotal beyond transposon silencing. *Stem Cells* *30*, 2603–2611.

- Bathgate, R.A.D., Halls, M.L., van der Westhuizen, E.T., Callander, G.E., Kocan, M., and Summers, R.J. (2013). Relaxin Family Peptides and Their Receptors. *Physiol. Rev.* *93*, 405–480.
- Bayne, B.L., Widdows, J., and Gabbott, P.A. (1975). Some effects of stress in the adult on the eggs and larvae of *Mytilus edulis* L. *J. Mar. Biol. Assoc. United Kingdom* *55*, 675–689.
- Beer, R.L., and Draper, B.W. (2013). Nanos3 maintains germline stem cells and expression of the conserved germline stem cell gene nanos2 in the zebrafish ovary. *Dev. Biol.* *374*, 308–318.
- Bellaiche, J., Lareyre, J.-J., Cauty, C., Yano, A., Allemand, I., and Le Gac, F. (2014). Spermatogonial Stem Cell Quest: nanos2, Marker of a Subpopulation of Undifferentiated A Spermatogonia in Trout Testis. *Biol. Reprod.* *90*, 79–79.
- Bellosta, P., and Gallant, P. (2010). Myc function in *Drosophila*. *Genes and Cancer* *1*, 542–546.
- Berthelin, C., Kellner, K., and Mathieu, M. (2000a). Storage metabolism in the Pacific oyster (*Crassostrea gigas*) in relation to summer mortalities and reproductive cycle (West Coast of France). *Comp. Biochem. Physiol. - B Biochem. Mol. Biol.* *125*, 359–369.
- Berthelin, C., Kellner, K., and Mathieu, M. (2000b). Histological Characterization and Glucose Incorporation into Glycogen of the Pacific Oyster *Crassostrea gigas* Storage Cells. *Mar. Biotechnol.* *2*, 136–145.
- Bhartiya, D. (2015). Ovarian stem cells are always accompanied by very small embryonic-like stem cells in adult mammalian ovary. *J. Ovarian Res.* *8*, 1–4.
- Bhartiya, D., Kasiviswanathan, S., Unni, S.K., Pethe, P., Dhabalia, J. V, Patwardhan, S., and Tongaonkar, H.B. (2010). Newer insights into premeiotic development of germ cells in adult human testis using Oct-4 as a stem cell marker. *J. Histochem. Cytochem.* *58*, 1093–1106.
- Bi, P., and Kuang, S. (2015). Notch signaling as a novel regulator of metabolism. *Trends Endocrinol. Metab.* *26*, 248–255.
- Bosseboeuf, A., Gautier, A., Auvray, P., Mazan, S., and Sourdaïne, P. (2014). Characterization of spermatogonial markers in the mature testis of the dogfish (*Scyliorhinus canicula* L.). *Reproduction* *147*, 125–139.
- Boswell, R.E., and Mahowald, A.P. (1985). tudor, a gene required for assembly of the germ plasm in *Drosophila melanogaster*. *Cell* *43*, 97–104.
- Brown, D.L. (2017). Bias in image analysis and its solution: unbiased stereology. *J. Toxicol. Pathol.* *30*, 183–191.
- Bruslé-Sicard, S., Debas, L., Fourcault, B., and Fuchs, J. (1992). Ultrastructural study of sex inversion in a protogynous hermaphrodite, *Epinephelus microdon* (Teleostei, Serranidae). *Reprod. Nutr. Dev.* *32*, 393–406.
- Bukovsky, A., Caudle, M.R., Svetlikova, M., and Upadhyaya, N.B. (2004). Origin of germ cells and formation of new primary follicles in adult human ovaries. *Reprod. Biol. Endocrinol.* *2*, 20.

Byrd, D.T., and Kimble, J. (2010). Scratching the Niche That Controls *C. elegans* Germline Stem Cells. *Semin Cell Dev Biol.* 20, 1–15.

C

Cai, J., Yang, W., Chen, D., Zhang, Y., He, Z., Zhang, W., and Zhang, L. (2017). Transcriptomic analysis of the differentiating ovary of the protogynous ricefield eel *Monopterus albus*. *BMC Genomics* 18, 1–16.

Cannarella, R., Condorelli, R.A., La Vignera, S., and Calogero, A.E. (2018). Effects of the insulin-like growth factor system on testicular differentiation and function: a review of the literature. *Andrology* 6, 3–9.

Cannon, M. V., Takeda, K., and Pinkert, C.A. (2011). Mitochondrial biology in reproduction. *Reprod. Med. Biol.* 10, 251–258.

Casadevall, M., Delgado, E., Colleye, O., Monserrat, S.B., and Parmentier, E. (2009). Histological Study of the Sex-Change in the Skunk Clownfish *Amphiprion akallopisos*. *Open Fish Sci. J.* 2, 55–58.

Casas, L., Saborido-Rey, F., Ryu, T., Michell, C., Ravasi, T., and Irigoien, X. (2016). Sex Change in Clownfish: Molecular Insights from Transcriptome Analysis. *Sci. Rep.* 6, 1–19.

Cavelier, P., Cau, J., Morin, N., and Delsert, C. (2017). Early gametogenesis in the Pacific oyster: new insights using stem cell and mitotic markers. *J. Exp. Biol.* jeb.167734.

de Celis, J.F., Llimargas, M., and Casanova, J. (1995). *ventral veinless*, the gene encoding the Cf1a transcription factor, links positional information and cell differentiation during embryonic and imaginal development in *Drosophila melanogaster*. *Development* 121, 3405–3416.

Chandler, J.C., Aizen, J., Elizur, A., Hollander-Cohen, L., Battaglione, S.C., and Ventura, T. (2015). Discovery of a novel insulin-like peptide and insulin binding proteins in the Eastern rock lobster *Sagmariasus verreauxi*. *Gen. Comp. Endocrinol.* 215, 76–87.

Chandler, J.C., Gandhi, N.S., Mancera, R.L., Smith, G., Elizur, A., and Ventura, T. (2017). Understanding insulin endocrinology in decapod crustacea: Molecular modelling characterization of an insulin-binding protein and insulin-like peptides in the eastern spiny lobster, *Sagmariasus verreauxi*. *Int. J. Mol. Sci.* 18, 1–19.

Chang, D.F., Tsai, S.C., Wang, X.C., Xia, P., Senadheera, D., and Lutzko, C. (2009). Molecular Characterization of the Human NANOG Protein. *Stem Cells* 27, 812–821.

Cheng, Z., Liu, F., Dai, M., Wu, J., Li, X., Guo, X., Tian, H., Heng, Z., Lu, Y., Chai, X., et al. (2017). Identification of EmSOX2, a member of the Sox family of transcription factors, as a potential regulator of *Echinococcus multilocularis* germinative cells. *Int. J. Parasitol.* 47, 625–632.

Chiarini-Garcia, H., and Russell, L.D. (2001). High-resolution light microscopic characterization of mouse spermatogonia. *Biol. Reprod.* 65, 1170–1178.

- Chiarini-Garcia, H., and Russell, L.D. (2002). Characterization of mouse spermatogonia by transmission electron microscopy. *Reproduction* *123*, 567–577.
- Chillakuri, C.R., Sheppard, D., Lea, S.M., and Handford, P.A. (2012). Notch receptor-ligand binding and activation: Insights from molecular studies. *Semin. Cell Dev. Biol.* *23*, 421–428.
- Chong, T., Stary, J.M., Wang, Y., and Newmark, P. a (2011). Molecular markers to characterize the hermaphroditic reproductive system of the planarian *Schmidtea mediterranea*. *BMC Dev. Biol.* *11*, 69.
- Chung, E.-Y. (2008). Ultrastructural Studies Of Oogenesis and Sexual Maturation In Female *Chlamys (azumapecten) farreri farreri* (Jones & Preston , 1904) (Pteriomorphia: Pectinidae) On The Western Coast Of Korea. *50*, 279–292.
- Chung, I. (2017). Optical measurement of receptor tyrosine kinase oligomerization on live cells. *Biochim. Biophys. Acta - Biomembr.* *1859*, 1436–1444.
- Chung, E., Park, Y., Lee, J., and Ryu, D. (2005). Germ Cell Differentiation and Sexual Maturation of the Hanging Cultured Female Scallop *Patinopecten yessoensis* on the East Coast of Korea. *J. Shellfish Res.* *24*, 913–921.
- Chung, E., Koh, C.H., and Park, G.M. (2007). Oogenesis, oocyte degeneration and sexual maturation in female *Cyclina sinensis* (Gmelin, 1971) (Bivalvia: Veneridae) in Korea. *Integr. Biosci.* *11*, 191–198.
- Cinquin, O. (2009). Purpose and regulation of stem cells: a systems-biology view from the *Caenorhabditis elegans* germ line. *J. Pathol.* *217*, 186–198.
- Ciocan, C.M., and Rotchell, J.M. (2005). Conservation of cancer genes in the marine invertebrate *Mytilus edulis*. *Environ. Sci. Technol.* *39*, 3029–3033.
- Ciocan, C.M., Moore, J.D., and Rotchell, J.M. (2006). The role of ras gene in the development of haemic neoplasia in *Mytilus trossulus*. *Mar. Environ. Res.* *62*, 147–150.
- Claeys, I., Simonet, G., Poels, J., Van Loy, T., Vercammen, L., De Loof, A., and Vanden Broeck, J. (2002). Insulin-related peptides and their conserved signal transduction pathway. *Peptides* *23*, 807–816.
- Clevers, H., Loh, K.M., and Nusse, R. (2014). An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science (80-.)*. *346*, 1248012–1248012.
- Collins, J.J., Hou, X., Romanova, E. V., Lambrus, B.G., Miller, C.M., Saberi, A., Sweedler, J. V., and Newmark, P.A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLoS Biol.* *8*.
- Conzelmann, M., Williams, E.A., Krug, K., Franz-Wachtel, M., Macek, B., and Jékely, G. (2013). The neuropeptide complement of the marine annelid *Platynereis dumerilii*. *BMC Genomics* *14*, 906.

Corbineau, S., Lassalle, B., Givelet, M., Souissi-Sarahoui, I., Firlej, V., Henri Romeo, P., Allemand, I., Riou, L., and Fouchet, P. (2016). Spermatogonial stem cells and progenitors are refractory to reprogramming to pluripotency by the transcription factors Oct3/4, c-Myc, Sox2 and Klf4. *Oncotarget* 8, 1–14.

Corporeau, C., Groisillier, A., Jeudy, A., Barbeyron, T., Fleury, E., Fabioux, C., Czjzek, M., and Huvet, A. (2011). A Functional Study of Transforming Growth Factor-Beta from the Gonad of Pacific Oyster *Crassostrea gigas*. *Mar. Biotechnol.* 13, 971–980.

Crespo, D., Assis, L.H.C., Furmanek, T., Bogerd, J., and Schulz, R.W. (2016). Expression profiling identifies Sertoli and Leydig cell genes as Fsh targets in adult zebrafish testis. *Mol. Cell. Endocrinol.* 437, 237–251.

D

Di Cristo, C., Paolucci, M., Iglesias, J., Sanchez, J., and Cosmo, A. Di (2002). Presence of two neuropeptides in the fusiform ganglion and reproductive ducts of *Octopus vulgaris*: FMRFamide and gonadotropin-releasing hormone (GnRH). *J. Exp. Zool.* 292, 267–276.

de Cuevas, M., and Matunis, E.L. (2011). The stem cell niche: lessons from the *Drosophila* testis. *Development* 138, 2861–2869.

Dadoune, J.-P. (2013). Cellules souches et gamètes. *Med. La Reprod.* 15, 4–16.

Das, D., and Arur, S. (2017). Conserved insulin signaling in the regulation of oocyte growth, development, and maturation. *Mol. Reprod. Dev.* 84, 444–459.

Das, D., Nath, P., Pal, S., Hajra, S., Ghosh, P., and Maitra, S. (2016). Expression of two insulin receptor subtypes, *insra* and *insrb*, in zebrafish (*Danio rerio*) ovary and involvement of insulin action in ovarian function. *Gen. Comp. Endocrinol.* 239, 21–31.

David, E., Tanguy, A., Pichavant, K., and Moraga, D. (2005). Response of the Pacific oyster *Crassostrea gigas* to hypoxia exposure under experimental conditions. *FEBS J.* 272, 5635–5652.

Decotto, E., and Spradling, A.C. (2005). The *Drosophila* ovarian and testis stem cell niches: Similar somatic stem cells and signals. *Dev. Cell* 9, 501–510.

DePina, A.S., Iser, W.B., Park, S., Maudsley, S., Wilson, M.A., and Wolkow, C.A. (2011). Regulation of *Caenorhabditis elegans* vitellogenesis by DAF-2/IIS through separable transcriptional and posttranscriptional mechanisms. *BMC Physiol.* 11, 11.

Deshpande, G., Manry, D., Jourjine, N., Mogila, V., Mozes, H., Bialistoky, T., Gerlitz, O., and Schedl, P. (2016). Role of the ABC transporter *Mdr49* in Hedgehog signaling and germ cell migration. *Development* 143, 2111–2120.

Dheilly, N.M., Lelong, C., Huvet, A., Kellner, K., Dubos, M.-P., Riviere, G., Boudry, P., and Favrel, P. (2012). Gametogenesis in the Pacific Oyster *Crassostrea gigas*: A Microarrays-Based Analysis Identifies Sex and Stage Specific Genes. *PLoS One* 7, e36353.

Dimitriadis, G., Mitrou, P., Lambadiari, V., Maratou, E., and Raptis, S.A. (2011). Insulin effects in muscle and adipose tissue. *Diabetes Res. Clin. Pract.* *93*, S52–S59.

Dirami, G., Ravindranath, N., Achi, M. V, and Dym, M. (2001). Expression of Notch pathway components in spermatogonia and Sertoli cells of neonatal mice. *J Androl* *22*, 944–952.

Dorange, G., and Le Pennec, M. (1989). Ultrastructural study of oogenesis and oocytic degeneration in *Pecten maximus* from the Bay of St. Brieuc. *Mar. Biol.* *103*, 339–348.

Van Doren, M., Broihier, H.T., Moore, L.A., and Lehmann, R. (1998). HMG-CoA reductase guides migrating primordial germ cells. *Nature* *396*, 466–469.

Dudley, B.M., Runyan, C., Takeuchi, Y., Schaible, K., and Molyneaux, K. (2007). BMP signaling regulates PGC numbers and motility in organ culture. *Mech. Dev.* *124*, 68–77.

Dupont, J., and LeRoith, D. (2001). Insulin and insulin-like growth factor I receptors: similarities and differences in signal transduction. *Horm. Res.* *55 Suppl 2*, 22–26.

Dupont, J., and Scaramuzzi, R.J. (2016). Insulin signalling and glucose transport in the ovary and ovarian function during the ovarian cycle. *Biochem. J.* *473*, 1483–1501.

Dupont, J., Reverchon, M., Bertoldo, M.J., and Froment, P. (2014). Nutritional signals and reproduction. *Mol. Cell. Endocrinol.* *382*, 527–537.

E

Eckelbarger, K., and Davis, C. (1996a). Ultrastructure of the gonad and gametogenesis in the eastern oyster, *Crassostrea virginica*. I. Ovary and oogenesis. *Mar. Biol.* *127*, 79–87.

Eckelbarger, K., and Davis, C. (1996b). Ultrastructure of the gonad and gametogenesis in the eastern oyster, *Crassostrea virginica*. II. Testis and spermatogenesis. *Mar. Biol.* *127*, 89–96.

Ephrussi, A., and Lehmann, R. (1992). Induction of germ cell formation by oskar. *Nature* *358*, 387–392.

Escott, G.M., da Rosa, L.A., and Loss, E. da S. (2014). Mechanisms of hormonal regulation of sertoli cell development and proliferation: a key process for spermatogenesis. *Curr. Mol. Pharmacol.* *7*, 96–108.

Extavour, C.G.M. (2007). Evolution of the bilaterian germ line: Lineage origin and modulation of specification mechanisms. *Integr. Comp. Biol.* *47*, 770–785.

Extavour, C.G., and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* *130*, 5869–5884.

F

Fabioux, C., Huvet, A., Lelong, C., Robert, R., Pouvreau, S., Daniel, J., Minguant, C., and Le Pennec, M. (2004a). Oyster vasa-like gene as a marker of the germline cell development in *Crassostrea gigas*. *Biochem. Biophys. Res. Commun.* *320*, 592–598.

- Fabioux, C., Pouvreau, S., Le Roux, F., and Huvet, A. (2004b). Oyster vasa-like gene: a specific marker of the germline in *Crassostrea gigas*. *Biochem. Biophys. Res. Commun.* *315*, 897–904.
- Fabioux, C., Corporeau, C., Quillien, V., Favrel, P., and Huvet, A. (2009). In vivo RNA interference in oyster -vasa silencing inhibits germ cell development. *FEBS J.* *276*, 2566–2573.
- Ferraro, F., Lo Celso, C., and Scadden, D. (2010). Adult stem cells and their niches. *Adv. Exp. Med. Biol.* *695*, 155–168.
- Fleury, E., Fabioux, C., Lelong, C., Favrel, P., and Huvet, A. (2008). Characterization of a gonad-specific transforming growth factor- β superfamily member differentially expressed during the reproductive cycle of the oyster *Crassostrea gigas*. *Gene* *410*, 187–196.
- Flores, J.A., and Burns, J.R. (1993). Ultrastructural study of embryonic and early adult germ cells, and their support cells, in both sexes of *Xiphophorus* (Teleostei:Poeciliidae). *Cell Tissue Res.* *271*, 263–270.
- Floyd, P., and Li, L. (1999). Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. *J. Neurosci.* *19*, 7732–7741.
- Focareta, L., and Cole, A.G. (2016). Analyses of Sox-B and Sox-E family genes in the cephalopod *sepia officinalis*: Revealing the conserved and the unusual. *PLoS One* *11*, 1–21.
- Forbes, A., and Lehmann, R. (1998). Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells. *Development* *125*, 679–690.
- Foulstone, E., Prince, S., Zaccheo, O., Burns, J.L., Harper, J., Jacobs, C., Church, D., and Hassan, A.B. (2005). Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *J. Pathol.* *205*, 145–153.
- Franco, A. (2009). Caractérisation cellulaire et moléculaire de la spermatogenèse chez l’huître creuse *Crassostrea gigas*. Université de Caen.
- Franco, A., Heude Berthelin, C., Goux, D., Sourdaire, P., and Mathieu, M. (2008). Fine structure of the early stages of spermatogenesis in the Pacific oyster, *Crassostrea gigas* (Mollusca, Bivalvia). *Tissue Cell* *40*, 251–260.
- Franco, A., Jouaux, A., Mathieu, M., Sourdaire, P., Lelong, C., Kellner, K., and Heude Berthelin, C. (2010). Proliferating cell nuclear antigen in gonad and associated storage tissue of the Pacific oyster *Crassostrea gigas*: seasonal immunodetection and expression in laser microdissected tissues. *Cell Tissue Res.* *340*, 201–210.
- Franco, A., Kellner, K., Goux, D., Mathieu, M., and Heude Berthelin, C. (2011). Intra-gonadal Somatic Cells (ISCs) in the male oyster *Crassostrea gigas*: Morphology and contribution in germinal epithelium structure. *Micron* *42*, 718–725.
- Fritsch, H.A.R., Van Noorden, S., and Pearse, A.G.E. (1976). Cytochemical and immunofluorescence investigations on insulin-like producing cells in the intestine of *Mytilus edulis* L. (Bivalvia). *Cell Tissue Res.* *165*, 365–369.
- Frontali, N., Williams, L., and Welsh, J.H. (1967). Heart excitatory and inhibitory substances in molluscan ganglia. *Comp. Biochem. Physiol.* *22*, 833–841.

G

- Gabe, M. (1968). Techniques histologiques (Paris).
- Gao, M., and Arkov, A.L. (2013). Next generation organelles: Structure and role of germ granules in the germline. *Mol. Reprod. Dev.* *80*, 610–623.
- Gao, J., Chen, Y., Yang, Y., Liang, J., Xie, J., Liu, J., Li, S., Zheng, G., Xie, L., and Zhang, R. (2016). The transcription factor Pf-POU3F4 regulates expression of the matrix protein genes *Aspein* and *Prismalin-14* in pearl oyster (*Pinctada fucata*). *FEBS J.* *283*, 1962–1978.
- Gasparini, F., Manni, L., Cima, F., Zaniolo, G., Burighel, P., Caicci, F., Franchi, N., Schiavon, F., Rigon, F., Campagna, D., et al. (2015). Sexual and asexual reproduction in the colonial ascidian *Botryllus schlosseri*. *Genesis* *53*, 105–120.
- Geraerts, W.P.M. (1976). The role of the Lateral Lobes in the control of growth and reproduction in the hermaphrodite freshwater snail *Lymnaea stagnalis*. *Gen. Comp. Endocrinol.* *29*, 97–108.
- Geraerts, W.P.M. (1992). Neurohormonal control of growth and carbohydrate metabolism by the light green cells in *Lymnaea stagnalis*. *Gen. Comp. Endocrinol.* *86*, 433–444.
- Geraerts, W.P.M., Smit, A.B., Li, K.W., and Hordijk, P.L. (1992). The Light Green Cells of *Lymnaea*: A neuroendocrine model system for stimulus-induced expression of multiple peptide genes in a single cell type. *Experientia* *48*, 464–473.
- Ghaleb, A.M., and Yang, V.W. (2017). Krüppel-like factor 4 (KLF4): What we currently know. *Gene* *611*, 27–37.
- Ghaleb, A.M., Nandan, M.O., Chanchevalap, S., Dalton, W.B., Hisamuddin, I.M., and Yang, V.W. (2005). Krüppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. *Cell Res.* *15*, 92–96.
- Giribet, G. (2008). Assembling the lophotrochozoan (= spiralian) tree of life. 1513–1522.
- Gleason, R.J., Anand, A., Kai, T., and Chen, X. (2018). Protecting and Diversifying the Germline. *208*, 435–471.
- Godmann, M., Katz, J.P., Guillou, F., Simoni, M., Kaestner, K.H., and Behr, R. (2008). Krüppel-like factor 4 is involved in functional differentiation of testicular Sertoli cells. *Dev. Biol.* *315*, 552–566.
- Golubeski, G.S., Bardsley, A., Tax, F., and Boswell, R.E. (1991). tudor, a posterior group gene of *Drosophila melanogaster*, encodes a novel protein and an mRNA localized during mid-oogenesis. *Genes Dev.* *5*, 2060–2070.
- Gottardo, M., Callaini, G., and Riparbelli, M.G. (2017). Correction: The *Drosophila* centriole – conversion of doublets into triplets within the stem cell niche (doi: 10.1242/jcs.172627). *J. Cell Sci.* *130*, 4225–4225.

Le Gouar, M., Guillou, A., and Vervoort, M. (2004). Expression of a SoxB and a Wnt2/13 gene during the development of the mollusc *Patella vulgata*. *Dev. Genes Evol.* *214*, 250–256.

Gribouval, L., Sourdaine, P., Lareyre, J., Bellaiche, J., Le Gac, F., Mazan, S., Guiardièrre, C., Auvray, P., and Gautier, A. (2018). The *nanos1* gene was duplicated in early Vertebrates and the two paralogs show different gonadal expression profiles in a shark. *Sci. Rep.* *8*, 6942.

Gricourt, L., Bonnac, G., Boujard, D., Mathieu, M., and Kellner, K. (2003). Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrIGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor. *Gen. Comp. Endocrinol.* *134*, 44–56.

Gricourt, L., Mathieu, M., and Kellner, K. (2006). An insulin-like system involved in the control of Pacific oyster *Crassostrea gigas* reproduction: hrIGF-1 effect on germinal cell proliferation and maturation associated with expression of an homologous insulin receptor-related receptor. *Aquaculture* *251*, 85–98.

Griffeth, R.J., Bianda, V., and Nef, S. (2014). The emerging role of insulin-like growth factors in testis development and function. *Basic Clin. Androl.* *24*, 12.

Griffond, B., and Bride, J. (1981). Etude histologique et ultrastructurale de la gonade d'*Helix aspersa* Müller à l'éclosion. *Reprod. Nutr. Dev.*

H

Hamano, K., Awaji, M., and Usuki, H. (2005). cDNA structure of an insulin-related peptide in the Pacific oyster and seasonal changes in the gene expression. *J. Endocrinol.* *187*, 55–67.

Harris, a N., and Macdonald, P.M. (2001). Aubergine encodes a *Drosophila* polar granule component required for pole cell formation and related to eIF2C. *Development* *128*, 2823–2832.

Hartmann, C., Landgraf, M., Bate, M., and Jäckle, H. (1997). Kruppel target gene knockout participates in the proper innervation of a specific set of *Drosophila* larval muscles. *EMBO J.* *16*, 5299–5309.

Hartung, O., Forbes, M.M., and Marlow, F.L. (2014). Zebrafish *vasa* is required for germ-cell differentiation and maintenance. *Mol. Reprod. Dev.* *81*, 946–961.

Hawkes, C., and Kar, S. (2004). The insulin-like growth factor-II/mannose-6-phosphate receptor: Structure, distribution and function in the central nervous system. *Brain Res. Rev.* *44*, 117–140.

Hay, B., Jan, L.Y., and Jan, Y.N. (1988). A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicases. *Cell* *55*, 577–587.

Hayashi, T., Asami, M., Higuchi, S., Shibata, N., and Agata, K. (2006). Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Dev. Growth Differ.* *48*, 371–380.

He, Z., Li, Y., Wu, Y., Shi, S., Sun, C., Deng, Q., Xie, J., Wang, T., Zhang, W., and Zhang, L. (2014). Differentiation and morphogenesis of the ovary and expression of gonadal development-related genes in the protogynous hermaphroditic ricefield eel *Monopterus albus*. *J. Fish Biol.* *85*, 1381–1394.

Henry, J., Zatylny, C., and Boucaud-Camou, E. (1999). Peptidergic control of egg-laying in the cephalopod *Sepia officinalis*: Involvement of FMRFamide and FMRFamide-related peptides. *Peptides* *20*, 1061–1070.

Heude Berthelin, C., Laisney, J., Espinosa, J., Martin, O., Hernandez, G., Mathieu, M., and Kellner, K. (2001). Storage and reproductive strategy in *Crassostrea gigas* from two different growing areas (Normandy and the Atlantic coast, France). *Invertebr. Reprod. Dev.* *40*, 79–86.

Higuchi, S., Hayashi, T., Hori, I., Shibata, N., Sakamoto, H., and Agata, K. (2007). Characterization and categorization of fluorescence activated cell sorted planarian stem cells by ultrastructural analysis. *Dev. Growth Differ.* *49*, 571–581.

Horan, C.J., and Williams, S.A. (2017). Oocyte stem cells: Fact or fantasy? *Reproduction* *154*, R23–R35.

Hsu, H.-J., and Drummond-Barbosa, D. (2009). Insulin levels control female germline stem cell maintenance via the niche in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 1117–1121.

Hsu, H.J., and Drummond-Barbosa, D. (2011). Insulin signals control the competence of the *Drosophila* female germline stem cell niche to respond to Notch ligands. *Dev. Biol.* *350*, 290–300.

Huang, Y.-H., Chin, C.-C., Ho, H.-N., Chou, C.-K., Shen, C.-N., Kuo, H.-C., Wu, T.-J., Wu, Y.-C., Hung, Y.-C., Chang, C.-C., et al. (2009). Pluripotency of mouse spermatogonial stem cells maintained by IGF-1-dependent pathway. *FASEB J.* *23*, 2076–2087.

Huang, Z., Rivas, B., and Agoulnik, A.I. (2013). NOTCH1 Gain of Function in Germ Cells Causes Failure of Spermatogenesis in Male Mice. *PLoS One* *8*.

Hubbard, E.J.A. (2007). The *C. elegans* germ line: a model for stem cell biology. *Dev Dyn* *236*, 3343–3357.

Hubbard, J.A.E. (2011). Insulin and Germline Proliferation in *Caenorhabditis elegans*. *Vitam. Horm.* *87*, 61–77.

Hubbard, S.R. (1999). Structural analysis of receptor tyrosine kinases. *Prog. Biophys. Mol. Biol.* *71*, 343–358.



Ikenishi, K. (1998). Germ plasm in *Caenorhabditis elegans*, *Drosophila* and *Xenopus*. *Dev. Growth Differ.* *40*, 1–10.

Ikeya, T., Galic, M., Belawat, P., Nairz, K., and Hafen, E. (2002). Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr. Biol.* *12*, 1293–1300.

Illmensee, K., and Mahowald, A.P. (1974). Transplantation of posterior polar plasm in *Drosophila*. Induction of germ cells at the anterior pole of the egg. *Proc. Natl. Acad. Sci. U. S. A.* *71*, 1016–1020.

Iwasaki, Y.W., Siomi, M.C., and Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu. Rev. Biochem.* *84*, 405–433.

J

Jaglarz, M.K., and Howard, K.R. (1995). The active migration of *Drosophila* primordial germ cells. *Development* *121*, 3495–3503.

Jagut, M., and Huynh, J. (2007). Régulation des cellules souches de la lignée germinale. 611–618.

Jemaà, M., Morin, N., Cavelier, P., Cau, J., Strub, J.M., and Delsert, C. (2014). Adult somatic progenitor cells and hematopoiesis in oysters. *J. Exp. Biol.* *217*, 3067–3077.

Jonas, E. a, Knox, R.J., Kaczmarek, L.K., Schwartz, J.H., and Solomon, D.H. (1996). Insulin receptor in *Aplysia* neurons: characterization, molecular cloning, and modulation of ion currents. *J. Neurosci.* *16*, 1645–1658.

Jong-brink, M. De, de Wit, A., Kraal, G., and Boer, H.H. (1976). A light and electron microscope study on oogenesis in the freshwater pulmonate snail *Biomphalaria glabrata*. *Cell Tissue Res.* *171*, 195–219.

de Jong-Brink, M., Boer, H.H., Hommes, T.G., and Kodde, A. (1977). Spermatogenesis and the role of Sertoli cells in the freshwater snail *Biomphalaria glabrata*. *Cell Tissue Res.* *181*, 37–58.

Jorge Eduardo Chavez Villalba (2001). Conditionnement expérimental de l’huître *Crassostrea gigas*. Phd Thesis 0–162.

Jouaux, A. (2010). Gestion de l’allocation d’énergie à la croissance ou à la reproduction chez l’huître creuse *Crassostrea gigas* : implication d’une régulation par les insulin-like. Université de Caen.

Jouaux, A., Heude Berthelin, C., Sourdaïne, P., Mathieu, M., and Kellner, K. (2010). Gametogenic stages in triploid oysters *Crassostrea gigas*: Irregular locking of gonial proliferation and subsequent reproductive effort. *J. Exp. Mar. Bio. Ecol.* *395*, 162–170.

Jouaux, A., Franco, A., Heude Berthelin, C., Sourdaïne, P., Blin, J.L., Mathieu, M., and Kellner, K. (2012). Identification of Ras, Pten and p70S6K homologs in the Pacific oyster *Crassostrea gigas* and diet control of insulin pathway. *Gen. Comp. Endocrinol.* *176*, 28–38.

Jouaux, A., Blin, J.L., Adeline, B., Heude Berthelin, C., Sourdaïne, P., Mathieu, M., and Kellner, K. (2013). Impact of energy storage strategies on gametogenesis and reproductive effort in diploid and triploid Pacific oysters *Crassostrea gigas* — Involvement of insulin signaling. *Aquaculture* *388–391*, 173–181.

Juliano, C., Wang, J., and Lin, H. (2011). Uniting Germline and Stem Cells: The Function of Piwi Proteins and the piRNA Pathway in Diverse Organisms. *Annu. Rev. Genet.* *45*, 447–469.

Juliano, C.E., Swartz, S.Z., and Wessel, G.M. (2010). A conserved germline multipotency program. *Development* *137*, 4113–4126.

K

Kamachi, Y., Uchikawa, M., and Kondoh, H. (2000). Pairing SOX off: With partners in the regulation of embryonic development. *Trends Genet.* *16*, 182–187.

Kashyap, V., Rezende, N.C., Scotland, K.B., Shaffer, S.M., Persson, J.L., Gudas, L.J., and Mongan, N.P. (2009). Regulation of Stem Cell Pluripotency and Differentiation Involves a Mutual Regulatory Circuit of the Nanog, OCT4, and SOX2 Pluripotency Transcription Factors With Polycomb Repressive Complexes and Stem Cell microRNAs. *Stem Cells Dev.* *18*, 1093–1108.

Kawasaki, T., Saito, K., Sakai, C., Shinya, M., and Sakai, N. (2012). Production of zebrafish offspring from cultured spermatogonial stem cells. *Genes to Cells* *17*, 316–325.

Kawase, E., Wong, M.D., Ding, B.C., and Xie, T. (2004). Gbb/Bmp signaling is essential for maintaining germline stem cells and for repressing bam transcription in the *Drosophila* testis. *Development* *131*, 1365–1375.

Kerr, J.B., Myers, M., and Anderson, R.A. (2013). The dynamics of the primordial follicle reserve. *Reproduction* *146*.

De Keuckelaere, E., Hulpiau, P., Saeys, Y., Berx, G., and van Roy, F. (2018). Nanos genes and their role in development and beyond. *Cell. Mol. Life Sci.* *75*, 1929–1946.

Khider, R. (2010). La régulation des cellules souches germinales par les protéines Piwi et Tudor chez l’huitre creuse *Crassostrea gigas* Exemple de la caractérisation fonctionnelle du gène piwi1.

Kim, S.H. (2016). Ultrastructural Studies on Oocyte Development and Vitellogenesis associated with Follicle Cells in Female *Scapharca subcrenata* (Pelecypoda : Arcidae) in Western Korea. *Dev. Reprod.* *20*, 227–235.

Kim, J.H., Chung, E., Choi, M.S., Lee, K., Lee, I., and Seo, W. (2011). Germ cell Differentiation During Spermatogenesis, and Ultrastructural Characteristics of Mature Sperm in Male *Phacosoma japonicus* (Bivalvia: Veneridae). *27*, 55–65.

Kim, J.S., Kim, J., Kim, B.S., Chung, H.Y., Lee, Y.Y., Park, C.S., Lee, Y.S., Lee, Y.H., and Chung, I.Y. (2005). Identification and functional characterization of an alternative splice variant within the fourth exon of human nanog. *Exp. Mol. Med.* *37*, 601–607.

Kimble, J., and Crittenden, S.L. (2007). Controls of Germline Stem Cells, Entry into Meiosis, and the Sperm/Oocyte Decision in *Caenorhabditis elegans*. *Annu. Rev. Cell Dev. Biol.* *23*, 405–433.

Kimble, J.E., and White, J.G. (1981). On the control of germ cell development in *Caenorhabditis elegans*. *Dev. Biol.* *81*, 208–219.

Kimura, K.D., Tissenbaum, H.A., Liu, Y., and Ruvkun, G. (1997). *daf-2*, an Insulin Receptor – Like Gene That Regulates Longevity and Diapause in *Caenorhabditis elegans*. *Science* (80-). 277, 942–946.

Kobayashi, K., Koyanagi, R., Matsumoto, M., and Cabrera, J.P. (1999). Switching from Asexual to Sexual Reproduction in the Planarian *Dugesia ryukyuensis*: Bioassay System and Basic Description of Sexualizing Process Switching from Asexual to Sexual Reproduction in the Planarian *Dugesia ryukyuensis*: Bioassay System and Basi. 16, 291–298.

Kocot, K.M., Cannon, J.T., Todt, C., Citarella, M.R., Kohn, A.B., Meyer, A., Santos, S.R., Schander, C., Moroz, L.L., Lieb, B., et al. (2011). Phylogenomics reveals deep molluscan relationships. *Nature* 477, 452–456.

Kondo, H., Ino, M., Suzuki, A., Ishizaki, H., and Iwami, M. (1996). Multiple gene copies for Bombyxin, an insulin-related peptide of the silkworm *Bombyx mori*: Structural signs for gene rearrangement and duplication responsible for generation of multiple molecular forms of Bombyxin. *J. Mol. Biol.* 259, 926–937.

Kranz, A.M., Tollenaere, A., Norris, B.J., Degnan, B.M., and Degnan, S.M. (2010). Identifying the germline in an equally cleaving mollusc: *Vasa* and *Nanos* expression during embryonic and larval development of the vetigastropod *Haliotis asinina*. *J. Exp. Zool. Part B Mol. Dev. Evol.* 314 B, 267–279.

Kubota, H., Avarbock, M.R., and Brinster, R.L. (2004). Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. *Proc. Natl. Acad. Sci.* 101, 16489–16494.

Kuroda, T., Tada, M., Kubota, H., Hatano, S., Suemori, H., Nakatsuji, N., Kimura, H., and Tada, T. (2005). Octamer and Sox Elements Are Required for Transcriptional cis Regulation of *Nanog* Gene Expression Octamer and Sox Elements Are Required for Transcriptional cis Regulation of *Nanog* Gene Expression. *Society* 25, 2475–2485.

L

Lacerda, S.M. dos S.N., Costa, G.M.J., and de França, L.R. (2014). Biology and identity of fish spermatogonial stem cell. *Gen. Comp. Endocrinol.* 207, 56–65.

Ladekarl, M. (1998). Objective malignancy grading: a review emphasizing unbiased stereology applied to breast tumors. *APMIS. Suppl.* 79, 1–34.

LaFever, L., and Drummond-Barbosa, D. (2005). Direct Control of Germline Stem Cell Division and Cyst Growth by Neural Insulin in *Drosophila*. *Science* (80-). 309, 1071–1073.

Lardans, V., Coppin, J.F., Vicogne, J., Aroca, E., Delcroix, M., and Dissous, C. (2001). Characterization of an insulin receptor-related receptor in *Biomphalaria glabrata* embryonic cells. *Biochim. Biophys. Acta - Biomembr.* 1510, 321–329.

Lasko, P. (2013). The DEAD-box helicase *Vasa*: Evidence for a multiplicity of functions in RNA processes and developmental biology. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1829, 810–816.

- Lawson, K.A., Dunn, N.R., Roelen, B.A., Zeinstra, L.M., Davis, A.M., Wright, C. V, Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* *13*, 424–436.
- Leal, M.C., Cardoso, E.R., Nóbrega, R.H., Batlouni, S.R., Bogerd, J., França, L.R., and Schulz, R.W. (2009). Histological and Stereological Evaluation of Zebrafish (*Danio rerio*) Spermatogenesis with an Emphasis on Spermatogonial Generations1. *Biol. Reprod.* *81*, 177–187.
- Lehmann, R. (1992). Germ-plasm formation and germ-cell determination in *Drosophila*. *Curr. Opin. Genet. Dev.* *2*, 543–549.
- Lelong, C., Mathieu, M., and Favrel, P. (2000). Structure and expression of mGDF, a new member of the transforming growth factor-beta superfamily in the bivalve mollusc *Crassostrea gigas*. *Eur. J. Biochem.* *267*, 3986–3993.
- Li, W., Kennedy, S.G., and Ruvkun, G. (2003). daf-28 encodes a *C. elegans* insulin superfamily member that is regulated by environmental cues and acts in the DAF-2 signaling pathway. *Genes Dev.* *17*, 844–858.
- Lien, W.H., and Fuchs, E. (2014). Wnt some lose some: Transcriptional governance of stem cells by Wnt/ β -catenin signaling. *Genes Dev.* *28*, 1517–1532.
- Lin, H. (1997). The Tao of Stem Cells in the Germline. *Annu. Rev. Genet.* *31*, 455–491.
- Linder, P., and Fuller-Pace, F. V. (2013). Looking back on the birth of DEAD-box RNA helicases. *Biochim. Biophys. Acta - Gene Regul. Mech.* *1829*, 750–755.
- Liu, N., Han, H., and Lasko, P. (2009). Vasa promotes *Drosophila* germline stem cell differentiation by activating *mei-P26* translation by directly interacting with a (U)-rich motif in its 3'UTR. *Genes Dev.* *23*, 2742–2752.
- Liu, X., Huang, J., Chen, T., Wang, Y., Xin, S., Li, J., Pei, G., and Kang, J. (2008). Yamanaka factors critically regulate the developmental signaling network in mouse embryonic stem cells. *Cell Res.* *18*, 1177–1189.
- Liu, Y., Liao, S., Veenstra, J.A., and Nässel, D.R. (2016). *Drosophila* insulin-like peptide 1 (DILP1) is transiently expressed during non-feeding stages and reproductive dormancy. *Sci. Rep.* *6*, 26620.
- Lloyd, A., and Sakonju, S. (1991). Characterization of two *Drosophila* POU domain genes, related to oct-1 and oct-2, and the regulation of their expression patterns. *Mech. Dev.* *36*, 87–102.
- López-Vera, E., Aguilar, M.B., and Heimer de la Cotera, E.P. (2008). FMRFamide and related peptides in the phylum mollusca. *Peptides* *29*, 310–317.
- Lu, C., Lam, H.N., and Menon, R.K. (2005). New members of the insulin family: Regulators of metabolism, growth and now...reproduction. *Pediatr. Res.* *57*, 70–73.

Lucas, T.F.G., Nascimento, A.R., Pisolato, R., Pimenta, M.T., Lazari, M.F.M., and Porto, C.S. (2014). Receptors and signaling pathways involved in proliferation and differentiation of Sertoli cells. 1–9.

M

Ma, X., Ji, A., Zhang, Z., Yang, D., Liang, S., Wang, Y., and Qin, Z. (2017). *Piwi1* is essential for gametogenesis in mollusk *Chlamys farreri*. PeerJ 5, e3412.

Ma, Y., Zhang, X., Ma, H., Ren, Y., Sun, Y., Wang, Q., and Shi, J. (2014). Bioinformatic analysis of the four transcription factors used to induce pluripotent stem cells. Cytotechnology 66, 967–978.

MacDonald, B.T., Tamai, K., and He, X. (2009). Wnt/ β -Catenin Signaling: Components, Mechanisms, and Diseases. Dev. Cell 17, 9–26.

Malide, D., Seidah, N.G., Chretien, M., and Bendayan, M. (1995). Electron microscopic immunocytochemical evidence for the involvement of the convertases PC1 and PC2 in the processing of proinsulin in pancreatic β - cells. J. Histochem. Cytochem. 43, 11–19.

Manor, R., Weil, S., Oren, S., Glazer, L., Aflalo, E.D., Ventura, T., Chalifa-Caspi, V., Lapidot, M., and Sagi, A. (2007). Insulin and gender: An insulin-like gene expressed exclusively in the androgenic gland of the male crayfish. Gen. Comp. Endocrinol. 150, 326–336.

Marlow, F. (2015). Primordial Germ Cell Specification and Migration. F1000Research 4, 1–14.

Maruyama, K., Hietter, H., Nagasawa, H., Isogai, A., Tamura, S., Suzuki, A., and Ishizaki, H. (1988). Isolation and primary structure of bombyxin-iv, a novel molecular species of bombyxin from the silkworm, bombyx mori. Agric. Biol. Chem. 52, 3035–3041.

Matsumoto, T., Masaoka, T., Fujiwara, A., Nakamura, Y., Satoh, N., and Awaji, M. (2013). Reproduction-Related Genes in the Pearl Oyster Genome. Zoolog. Sci. 30, 826–850.

Matsunaga, Y. (2017). Diverse insulin-like peptides in *Caenorhabditis elegans*. Int. Biol. Rev. 1, 1–15.

McLeod, C.J., Wang, L., Wong, C., and Jones, D.L. (2010). Stem cell dynamics in response to nutrient availability. Curr. Biol. 20, 2100–2105.

Mercer, K.L., and Cross, P.C. (1995). Ultrastructure cellulaire et tissulaire. Approche fonctionnelle.

De Meyts, P. (2004). Insulin and its receptor: Structure, function and evolution. BioEssays 26, 1351–1362.

Michaelson, D., Korta, D.Z., Capua, Y., and Hubbard, E.J.A. (2010). Insulin signaling promotes germline proliferation in *C. elegans*. Development 137, 671–680.

Milani, L., Maurizii, M.G., Pecci, A., Ghiselli, F., and Passamonti, M. (2015). Localization and dynamics of primordial germ cells in the bivalve species *Ruditapes philippinarum*. Mol. Reprod. Dev. 82, 406–407.

- Milani, L., Pecci, A., Ghiselli, F., Passamonti, M., Bettini, S., Franceschini, V., and Maurizii, M.G. (2017). VASA expression suggests shared germ line dynamics in bivalve molluscs. *Histochem. Cell Biol.* *148*, 157–171.
- Milani, L., Pecci, A., Ghiselli, F., Passamonti, M., Lazzari, M., Franceschini, V., and Maurizii, M.G. (2018). Germ cell line during the seasonal sexual rest of clams: finding niches of cells for gonad renewal. *Histochem. Cell Biol.* *149*, 105–110.
- Miller, C.M., and Newmark, P.A. (2012). An insulin-like peptide regulates size and adult stem cells in planarians. *Int. J. Dev. Biol.* *56*, 75–82.
- Mills, K.M., Szczerkowski, J.L.A., and Habib, S.J. (2017). Wnt ligand presentation and reception: From the stem cell niche to tissue engineering. *Open Biol.* *7*.
- Van Minnen, J., Smit, A.B., and Joosse, J. (1989). Central and Peripheral Expression of Genes Coding for Egg- Laying Inducing and Insulin-Related Peptides in a Snail. *Arch. Histol. Cytol* *52*, 241–252.
- De Mitcheson, Y.S., and Liu, M. (2008). Functional hermaphroditism in teleosts. *Fish Fish.* *9*, 1–43.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* *113*, 631–642.
- Mizoguchi, A., and Okamoto, N. (2013). Insulin-like and IGF-like peptides in the silkworm *Bombyx mori*: Discovery, structure, secretion, and function. *Front. Physiol.* *4* AUG, 1–11.
- Le Moigne, A., and Foucrier, J. (2009). *Biologie du développement*.
- Moore, L.A., Broihier, H.T., Doren, M. Van, Lunsford, L.B., Lehmann, R., Van Doren, M., Lunsford, L.B., and Lehmann, R. (1998). Identification of genes controlling germ cell migration and embryonic gonad formation in *Drosophila*. *Development* *125*, 667–678.
- Morita, M., Best, J.B., and Noel, J. (1969). Electron microscopic studies of planarian regeneration. I. Fine structure of neoblasts in *Dugesia dorotocephala*. *J. Ultrastructure Res.* *27*, 7–23.
- Morrison, S.J., and Kimble, J. (2006). Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* *441*, 1068–1074.
- Morrison, S.J., and Spradling, A.C. (2008). Stem Cells and Niches: Mechanisms That Promote Stem Cell Maintenance throughout Life. *Cell* *132*, 598–611.
- Mosimann, C., Hausmann, G., and Basler, K. (2009). β -Catenin hits chromatin: Regulation of Wnt target gene activation. *Nat. Rev. Mol. Cell Biol.* *10*, 276–286.
- Mosquera, L., Forristall, C., Zhou, Y., and King, M.L. (1993). A mRNA localized to the vegetal cortex of *Xenopus* oocytes encodes a protein with a nanos-like zinc finger domain. *Development* *117*, 377–386.

Muranaka, M.S., and Lannan, J.E. (1984). Broodstock management of *Crassostrea gigas*: Environmental influences on broodstock conditioning. *Aquaculture* 39, 217–228.

N

Nakagawa, H., Ishizu, H., Chinone, A., Kobayashi, K., and Matsumoto, M. (2012). The *Dn-nanos* gene is essential for germ cell specification in the planarian *Dugesia ryukyuensis*. *Int. J. Dev. Biol.* 56, 165–171.

Nakamura, M., Miura, S., Nozu, R., and Kobayashi, Y. (2015). Opposite-directional sex change in functional female protandrous anemonefish, *Amphiprion clarkii*: effect of aromatase inhibitor on the ovarian tissue. *Zool. Lett.* 1, 30.

Nandan, M.O., and Yang, V.W. (2009). The role of Krüppel-like factors in the reprogramming of somatic cells to induced pluripotent stem cells. *Histol Histopathol.* 24, 1343–1355.

Nässel, D.R., Liu, Y., and Luo, J. (2015). Insulin/IGF signaling and its regulation in *Drosophila*. *Gen. Comp. Endocrinol.* 221, 255–266.

Nelson, S.N., and Van Der Kraak, G. (2010). The role of the insulin-like growth factor (IGF) system in zebrafish (*Danio rerio*) ovarian development. *Gen. Comp. Endocrinol.* 168, 103–110.

Newmark, P.A., and Alvarado, A.S. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev. Biol.* 220, 142–153.

Newmark, P.A., Wang, Y., and Chong, T. (2008). Germ Cell Specification and Regeneration in Planarians. *Cold Spring Harb. Symp. Quant. Biol.* LXXIII, 573–581.

Nikolic, A., Volarevic, V., Armstrong, L., Lako, M., and Stojkovic, M. (2016). Primordial Germ Cells: Current Knowledge and Perspectives. *Stem Cells Int.* 2016, 1–8.

Nóbrega, R.H., Greebe, C.D., van de Kant, H., Bogerd, J., de França, L.R., and Schulz, R.W. (2010). Spermatogonial stem cell niche and spermatogonial stem cell transplantation in zebrafish. *PLoS One* 5, 1–16.

Nóbrega, R.H., De Souza Morais, R.D.V., Crespo, D., De Waal, P.P., De França, L.R., Schulz, R.W., and Bogerd, J. (2015). Fsh stimulates spermatogonial proliferation and differentiation in zebrafish via *Igf3*. *Endocrinology* 156, 3804–3817.

Noce, T., Okamoto-Ito, S., and Tsunekawa, N. (2001). *Vasa* homolog genes in mammalian germ cell development. *Cell Struct Funct* 26, 131–136.

O

O'Brien, E.K., and Degnan, B.M. (2000). Expression of POU, Sox, and Pax genes in the brain ganglia of the tropical abalone *Haliotis asinina*. *Mar. Biotechnol.* 2, 545–557.

O'Brien, E.K., and Degnan, B.M. (2002). Developmental expression of a class IV POU gene in the gastropod *Haliotis asinina* supports a conserved role in sensory cell development in bilaterians. *Dev. Genes Evol.* *212*, 394–398.

Oatley, J.M., and Brinster, R.L. (2012). The Germline Stem Cell Niche Unit in Mammalian Testes. *Physiol. Rev.* *92*, 577–595.

Obata, M., Sano, N., Kimata, S., Nagasawa, K., Yoshizaki, G., and Komaru, A. (2010). The proliferation and migration of immature germ cells in the mussel, *Mytilus galloprovincialis*: Observation of the expression pattern in the *M. galloprovincialis* vasa-like gene (*Myvlg*) by in situ hybridization. *Dev. Genes Evol.* *220*, 139–149.

Okuda, H., Kiuchi, H., Takao, T., Miyagawa, Y., Tsujimura, A., Nonomura, N., Miyata, H., Okabe, M., Ikawa, M., Kawakami, Y., et al. (2015). A novel transcriptional factor *Nkapl* is a germ cell-specific suppressor of Notch signaling and is indispensable for spermatogenesis. *PLoS One* *10*, 1–19.

Önal, P., Grün, D., Adamidi, C., Rybak, A., Solana, J., Mastrobuoni, G., Wang, Y., Rahn, H.-P., Chen, W., Kempa, S., et al. (2012). Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J.* *31*, 2755–2769.

P

Parisi, M., and Lin, H. (2000). Translational repression: a duet of Nanos and Pumilio. *Curr. Biol.* *10*, R81–R83.

Park, J.J., Kim, H., Kang, S.W., An, C.M., Lee, S.-H., Gye, M.C., and Lee, J.S. (2012). Sex Ratio and Sex Reversal in Two-year-old Class of Oyster, *Crassostrea gigas* (Bivalvia: Ostreidae). *Dev. Reprod.* *16*, 385–388.

Pei, J., and Grishin, N. V. (2015). C2H2 zinc finger proteins of the SP/KLF, Wilms tumor, EGR, Hucbein, and Klumpfuss families in metazoans and beyond. *Gene* *573*, 91–99.

Pepling, M.E., Wilhelm, J.E., O'Hara, A.L., Gephardt, G.W., and Spradling, A.C. (2007). Mouse oocytes within germ cell cysts and primordial follicles contain a Balbiani body. *Proc. Natl. Acad. Sci.* *104*, 187–192.

Pfister, D., De Mulder, K., Hartenstein, V., Kualess, G., Borgonie, G., Marx, F., Morris, J., and Ladurner, P. (2008). Flatworm stem cells and the germ line: Developmental and evolutionary implications of *macvasa* expression in *Macrostomum lignano*. *Dev. Biol.* *319*, 146–159.

Pierce, S.B., Costa, M., Wisotzkey, R., Devadhar, S., Homburger, S.A., Buchman, A.R., Ferguson, K.C., Heller, J., Platt, D.M., Pasquinelli, A.A., et al. (2001). Regulation of DAF-2 receptor signaling by human insulin and *ins-1*, a member of the unusually large and diverse *C. elegans* insulin gene family. *Genes Dev.* *672–686*.

Pipe, R. (1987). Oogenesis in the marine mussel *Mytilus edulis*: an ultrastructural study. *Mar. Biol.* *95*, 405–414.

Pitetti, J.-L., Calvel, P., Zimmermann, C., Conne, B., Papaioannou, M.D., Aubry, F., Cederroth, C.R., Urner, F., Fumel, B., Crausaz, M., et al. (2013). An Essential Role for Insulin and IGF1 Receptors in Regulating Sertoli Cell Proliferation, Testis Size, and FSH Action in Mice. *Mol. Endocrinol.* *27*, 814–827.

Plisetskaya, E., Kazakov, V.K., Soltitskaya, L., and Leibson, L.G. (1978). Insulin-producing cells in the gut of freshwater bivalve molluscs *Anodonta cygnea* and *Unio pictorum* and the role of insulin in the regulation of their carbohydrate metabolism. *Gen. Comp. Endocrinol.* *35*, 133–145.

Poole, C.J., and van Riggelen, J. (2017). MYC—master regulator of the cancer epigenome and transcriptome. *Genes (Basel)*. *8*.

Price, D.A., and Greenberg, M.J. (1977). Structure of a molluscan cardioexcitatory neuropeptide. *Science (80-)*. *197*, 670–671.

R

Rabinowitz, J.S., Chan, X.Y., Kingsley, E.P., Duan, Y., and Lambert, J.D. (2008). Nanos Is Required in Somatic Blast Cell Lineages in the Posterior of a Mollusk Embryo. *Curr. Biol.* *18*, 331–336.

Ramalho-Santos, J., Varum, S., Amaral, S., Mota, P.C., Sousa, A.P., and Amaral, A. (2009). Mitochondrial functionality in reproduction: From gonads and gametes to embryos and embryonic stem cells. *Hum. Reprod. Update* *15*, 553–572.

Raven, P., Johnson, G., Losos, J., and Singer, S. (2007). *Biologie*.

Raverot, G., Weiss, J., Park, S.Y., Hurley, L., and Jameson, J.L. (2005). Sox3 expression in undifferentiated spermatogonia is required for the progression of spermatogenesis. *Dev. Biol.* *283*, 215–225.

Raz, E. (2000). The function and regulation of vasa-like genes in germ-cell development. *Genome Biol.* *1*, 1–6.

Reddien, P.W. (2013). Specialized progenitors and regeneration. *Development* *140*, 951–957.

Reddien, P.W., Oviedo, N.J., Jennings, J.R., Jenkin, J.C., and Sánchez-Alvarado, A. (2005). Smed WI-2 is a PIWI-like protein that regulates planarian stem cells. *Science (80-)*. *310*, 1327–1331.

Reik, W., and Surani, M.A. (2015). Germline and Pluripotent Stem Cells. *Cold Spring Harb. Perspect. Biol.* *7*, a019422.

Rejon, C. a., Hancock, M. a., Li, Y.N., Thompson, T.B., Hébert, T.E., and Bernard, D.J. (2013). Activins bind and signal via bone morphogenetic protein receptor type II (BMPRII) in immortalized gonadotrope-like cells. *Cell. Signal.* *25*, 2717–2726.

Richardson, B.E., and Lehmann, R. (2010). Mechanisms guiding primordial germ cell migration: strategies from different organisms. *Nat. Rev. Mol. Cell Biol.* *11*, 37–49.

Rivière, G., Klopp, C., Ibouniyamine, N., Huvet, A., Boudry, P., and Favrel, P. (2015). GigaTON: an extensive publicly searchable database providing a new reference transcriptome in the pacific oyster *Crassostrea gigas*. *BMC Bioinformatics* 16, 401.

Rodet, F., Lelong, C., Dubos, M.P., Costil, K., and Favrel, P. (2005). Molecular cloning of a molluscan gonadotropin-releasing hormone receptor orthologue specifically expressed in the gonad. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1730, 187–195.

Rohrschneider, M.R., and Nance, J. (2013). The union of somatic gonad precursors and primordial germ cells during *Caenorhabditis elegans* embryogenesis. *Dev. Biol.* 379, 139–151.

Roovers, E., Vincent, M.E., van Kesteren, E., Geraerts, W.P.M., Planta, R.J., Vreugdenhil, E., and van Heerikhuizen, H. (1995). Characterization of a putative molluscan insulin-related peptide receptor. *Gene* 162, 181–188.

Rosner, A., Moiseeva, E., Rabinowitz, C., and Rinkevich, B. (2013). Germ lineage properties in the urochordate *Botryllus schlosseri* - From markers to temporal niches. *Dev. Biol.* 384, 356–374.

S

Sada, A., and Tumber, T. (2013). New Insights into Mechanisms of Stem Cell Daughter Fate Determination in Regenerative Tissues. In *Int Rev Cell Mol Biol.*, pp. 1–50.

Saffman, E.E.F., and Lasko, P. (1999). Germline development in vertebrates and invertebrates. *Cell. Mol. Life Sci.* 55, 1141–1163.

Saint-Dizier, M., and Chastant-Maillard, S. (2014). *La reproduction animale et humaine*.

Saitou, M., and Yamaji, M. (2012). Primordial Germ Cells in Mice. *Cold Spring Harb. Perspect. Biol.* 4, a008375–a008375.

Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. *Nature* 418.

Salinas, L.S., Maldonado, E., Macías-Silva, M., Blackwell, T.K., and Navarro, R.E. (2007). The DEAD box RNA helicase VBH-1 is required for germ cell function in *C. elegans*. *Genesis* 45, 533–546.

Salinas, L.S., Franco-Cea, A., Láscarez-Lagunas, L.I., Villanueva-Chimal, E., Maldonado, E., and Navarro, R.E. (2012). Germ cell survival in *C. elegans* and *C. remanei* is affected when the dead box rna helicases Vbh-1 or Cre-VBH-1 are silenced. *Genesis* 50, 801–818.

Saltiel, A.R., and Kahn, C.R. (2001). Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 414, 799–806.

Samji, T. (2009). PIWI, piRNAs, and germline stem cells: what's the link? *Yale J. Biol. Med.* 82, 121–124.

- Samoto, T., Maruo, T., Ladines-Llave, C.A., Matsuo, H., Deguchi, J., Barnea, E.R., and Mochizuki, M. (1993). Insulin receptor expression in follicular and stromal compartments of the human ovary over the course of follicular growth, regression and atresia. *Endocr. J.* *40*, 715–726.
- Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S., and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed by the expression of nanos-related gene in planarians. *Dev. Growth Differ.* *48*, 615–628.
- Saunders, A., Faiola, F., and Wang, J. (2013). Concise Review: Pursuing Self-Renewal and Pluripotency with the Stem Cell Factor Nanog. *Stem Cells* *31*, 1227–1236.
- Scadden, D.T. (2006). The stem-cell niche as an entity of action. *Nature* *441*, 1075–1079.
- Scerbo, P., Markov, G. V., Vivien, C., Kodjabachian, L., Demeneix, B., Coen, L., and Girardot, F. (2014). On the origin and evolutionary history of NANOG. *PLoS One* *9*.
- Schmitz, C., and Hof, P.R. (2005). Design-based stereology in neuroscience. *Neuroscience* *130*, 813–831.
- Schuff, M., Siegel, D., Philipp, M., Bundschu, K., Heymann, N., Donow, C., and Knöchel, W. (2012). Characterization of *Danio rerio* *Nanog* and Functional Comparison to *Xenopus* *Vents*. *Stem Cells Dev.* *21*, 1225–1238.
- Schuh, R., Aicher, W., Gaul, U., Côte, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schröder, C., Kemler, R., et al. (1986). A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Krüppel, a Drosophila segmentation gene. *Cell* *47*, 1025–1032.
- Schulz, R.W., de França, L.R., Lareyre, J.J., LeGac, F., Chiarini-Garcia, H., Nobrega, R.H., and Miura, T. (2010). Spermatogenesis in fish. *Gen. Comp. Endocrinol.* *165*, 390–411.
- Schumacher, R., Soos, M.A., Schlessinger, J., Brandenburg, D., Siddle, K., and Ullrich, A. (1993). Signaling-competent chimeras allow mapping of major insulin receptor binding domains determinants. *J. Biochem.* *268*, 1087–1094.
- Shabanpoor, F., Separovic, F., and Wade, J.D. (2009). Chapter 1 The Human Insulin Superfamily of Polypeptide Hormones (Elsevier Inc.).
- Shi, Y., and Massagué, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* *113*, 685–700.
- Shi, Y., Guan, Y., and He, M. (2013). Molecular identification of insulin-related peptide receptor and its potential role in regulating development in *Pinctada fucata*. *Aquaculture* *408–409*, 118–127.
- Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K., and Agata, K. (1999). Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. *Dev. Biol.* *206*, 73–87.
- Shibata, N., Rouhana, L., and Agata, K. (2010). Cellular and molecular dissection of pluripotent adult somatic stem cells in planarians. *Dev. Growth Differ.* *52*, 27–41.

- Shim, J., Gururaja-Rao, S., and Banerjee, U. (2013). Nutritional regulation of stem and progenitor cells in *Drosophila*. *Development* *140*, 4647–4656.
- Silva, R.C., Costa, G.M.J., Andrade, L.M., and França, L.R. (2010). Testis stereology, seminiferous epithelium cycle length, and daily sperm production in the ocelot (*Leopardus pardalis*). *Theriogenology* *73*, 157–167.
- Singh, V.K., Kumar, N., Kalsan, M., Saini, A., and Chandra, R. (2015). Mechanism of Induction: Induced Pluripotent Stem Cells (iPSCs). *J. Stem Cells* *10*, 43–62.
- Sliwowska, J.H., Fergani, C., Gawalek, M., Skowronska, B., Fichna, P., and Lehman, M.N. (2014). Insulin: Its role in the central control of reproduction. *Physiol. Behav.* *133*, 197–206.
- Smit, A.B., Vreugdenhil, E., Ebberink, R.H., Geraerts, W.P., Klootwijk, J., and Joosse, J. (1988). Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature* *331*, 535–538.
- Smit, A.B., Geraerts, P.M., Meester, I., van Heerikhuizen, H., and Joosse, J. (1991). Characterization of a cDNA clone encoding molluscan insulin-related peptide II of *Lymnaea stagnalis*. *Eur. J. Biochem.* *199*, 699–703.
- Smit, A.B., Thijsen, S.F., Geraerts, W.P., Meester, I., van Heerikhuizen, H., and Joosse, J. (1992). Characterization of a cDNA clone encoding molluscan insulin-related peptide V of *Lymnaea stagnalis*. *Brain Res. Mol. Brain Res.* *14*, 7–12.
- Smit, A.B., van Marle, A., van Elk, R., Bogerd, J., van Heerikhuizen, H., and Geraerts, W.P. (1993). Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*. *J. Mol. Endocrinol.* *11*, 103–113.
- Smit, A.B., Spijker, S., Van Minnen, J., Burke, J.F., De Winter, F., Van Elk, R., and Geraerts, W.P. (1996). Expression and characterization of molluscan insulin-related peptide VII from the mollusc *Lymnaea stagnalis*. *Neuroscience* *70*, 589–596.
- Smit, A.B., van Kesteren, R.E., Li, K.W., Van Minnen, J., Spijker, S., Van Heerikhuizen, H., and Geraerts, W.P. (1998). Towards understanding the role of insulin in the brain: lessons from insulin-related signaling systems in the invertebrate brain. *Prog. Neurobiol.* *54*, 35–54.
- Solana, J. (2013). Closing the circle of germline and stem cells: the Primordial Stem Cell hypothesis. *Evodevo* *4*, 2.
- Song, X., Call, G.B., Kirilly, D., and Xie, T. (2007). Notch signaling controls germline stem cell niche formation in the *Drosophila* ovary. *Development* *134*, 1071–1080.
- Soriano, N.S., and Russell, S. (1998). The *Drosophila* SOX-domain protein Dichaete is required for the development of the central nervous system midline. *Development* *125*, 3989–3996.
- Soudant, P., Marty, Y., Moal, J., Robert, R., Quéré, C., Le Coz, J.R., and Samain, J.F. (1996). Effect of food fatty acid and sterol quality on *Pecten maximus* gonad composition and reproduction process. *Aquaculture* *143*, 361–378.

Spradling, A., Fuller, M.T., Braun, R.E., and Yoshida, S. (2011). Germline stem cells. *Cold Spring Harb. Perspect. Biol.* **3**, 1–20.

Steiner, D.F., Chan, S.J., Welsh, J.M., and Kwok, S.C.M. (1985). Structure and Evolution of the Insulin Gene. *Annu. Rev. Genet.* **19**, 463–484.

Stoner, D.S., Rinkevich, B., and Weissman, I.L. (1999). Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. *Proc. Natl. Acad. Sci.* **96**, 9148–9153.

Sur, I. (2009). Kruppel-Like Factors 4 and 5: Unity in Diversity. *Curr. Genomics* **10**, 594–603.

Swanson, M.E., Elste, A.M., Greenberg, S.M., Schwartz, J.H., Aldrich, T.H., and Furth, M.E. (1986). Abundant expression of ras proteins in *Aplysia* neurons. *J. Cell Biol.* **103**, 485–492.

Szakmary, A., Reedy, M., Qi, H., and Lin, H. (2009). The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in *Drosophila melanogaster*. *J. Cell Biol.* **185**, 613–627.

T

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* **131**, 861–872.

Takase, H.M., and Nusse, R. (2016). Paracrine Wnt/ β -catenin signaling mediates proliferation of undifferentiated spermatogonia in the adult mouse testis. *Proc. Natl. Acad. Sci.* **113**, E1489–E1497.

Tam, P.P.L., and Zhou, S.X. (1996). The allocation of epiblast cells to ectodermal and germline lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev. Biol.* **178**, 124–132.

Tanaka, S.S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., and Noce, T. (2000). The mouse homolog of *Drosophila* Vasa is required for the development of male germ cells. *Genes Dev.* **14**, 841–853.

Templeman, N.M., and Murphy, C.T. (2018). Regulation of reproduction and longevity by nutrient-sensing pathways. *J. Cell Biol.* **217**, 93–106.

Thomson, T., and Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu. Rev. Cell Dev. Biol.* **25**, 355–376.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Cloning* **30**, 3202–3213.

Tong, Y., Zhang, Y., Huang, J., Xiao, S., Zhang, Y., Li, J., Chen, J., and Yu, Z. (2015). Transcriptomics analysis of *Crassostrea hongkongensis* for the discovery of reproduction-related genes. *PLoS One* *10*, 1–24.

Toyooka, Y., Tsunekawa, N., Takahashi, Y., and Matsui, Y. (2000). Expression and intracellular localization of mouse Vasa -homologue protein during germ cell development. *Mech. Dev.* *93*, 139–149.

Tsunekawa, N., Naito, M., Sakai, Y., Nishida, T., and Noce, T. (2000). Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells. *Development* *127*, 2741–2750.

Tu, W.B., Helander, S., Pilstål, R., Hickman, K.A., Lourenco, C., Jurisica, I., Raught, B., Wallner, B., Sunnerhagen, M., and Penn, L.Z. (2015). Myc and its interactors take shape. *Biochim. Biophys. Acta - Gene Regul. Mech.* *1849*, 469–483.

U

Ueishi, S., Shimizu, H., and Inoue, Y. (2009). Male germline stem cell division and spermatocyte growth require insulin signaling in *Drosophila*. *Cell Struct. Funct.* *34*, 61–69.

Upadhyay, M., Kuna, M., Tudor, S., Martino Cortez, Y., and Rangan, P. (2018). A switch in the mode of Wnt signaling orchestrates the formation of germline stem cell differentiation niche in *Drosophila*. *PLoS Genet.* *14*, 1–29.

V

Veenstra, J.A. (2010). Neurohormones and neuropeptides encoded by the genome of *Lottia gigantea*, with reference to other mollusks and insects. *Gen. Comp. Endocrinol.* *167*, 86–103.

Veenstra, J.A., Agricola, H.-J., and Sellami, A. (2008). Regulatory peptides in fruit fly midgut. *Cell Tissue Res.* *334*, 499–516.

Ventura, T., Manor, R., Aflalo, E.D., Weil, S., Raviv, S., Glazer, L., and Sagi, A. (2009). Temporal silencing of an androgenic gland-specific insulin-like gene affecting phenotypical gender differences and spermatogenesis. *Endocrinology* *150*, 1278–1286.

Ventura, T., Rosen, O., and Sagi, A. (2011). From the discovery of the crustacean androgenic gland to the insulin-like hormone in six decades. *Gen. Comp. Endocrinol.* *173*, 381–388.

Ventura, T., Manor, R., Aflalo, E.D., Weil, S., Rosen, O., and Sagi, A. (2012). Timing Sexual Differentiation: Full Functional Sex Reversal Achieved Through Silencing of a Single Insulin-Like Gene in the Prawn, *Macrobrachium rosenbergii*. *Biol. Reprod.* *86*, 1–6.

Virant-Klun, I. (2015). Postnatal oogenesis in humans: A review of recent findings. *Stem Cells Cloning* *8*, 49–60.

Voronina, E., Seydoux, G., Sassone-Corsi, P., and Nagamori, I. (2011). RNA granules in germ cells. *Cold Spring Harb. Perspect. Biol.* *3*.

W

- Wagner, T.U. (2007). Bone morphogenetic protein signaling in stem cells - One signal, many consequences. *FEBS J.* *274*, 2968–2976.
- Wang, Z., and Lin, H. (2004). Nanos Maintains Germline Stem Cell Self-Renewal by Preventing Differentiation. *Science (80-.)*. *303*, 2016–2019.
- Weissman, I.L. (2000). Stem cells: units of development, units of regeneration, and units in evolution. *Cell* *100*, 157–168.
- Western, P., Maldonado-Saldivia, J., van den Bergen, J., Hajkova, P., Saitou, M., Barton, S., and Surani, M.A. (2005). Analysis of *Esg1* Expression in Pluripotent Cells and the Germline Reveals Similarities with *Oct4* and *Sox2* and Differences Between Human Pluripotent Cell Lines. *Stem Cells* *23*, 1436–1442.
- Wieschaus, E., Nusslein-Volhard, C., and Kluding, H. (1984). Krüppel, a gene whose activity is required early in the zygotic genome for normal embryonic segmentation. *Dev. Biol.* *104*, 172–186.
- Wilcox, G. (2005). Insulin and insulin resistance. *Clin. Biochem. Rev.* *26*, 19–39.
- van Wolfswinkel, J.C. (2014). Piwi and potency: PIWI proteins in animal stem cells and regeneration. *Integr. Comp. Biol.* *54*, 700–713.
- Wollesen, T., McDougall, C., Degnan, B.M., and Wanninger, A. (2014). POU genes are expressed during the formation of individual ganglia of the cephalopod central nervous system. *Evodevo* *5*, 1–15.
- Wu, Q., and Brown, M.R. (2006). Signaling and Function of Insulin-Like Peptides in Insects. *Annu. Rev. Entomol.* *51*, 1–24.
- Wylie, C. (1999). Germ cells. *Curr. Opin. Genet. Dev.* *10*, 410–413.

X

- Xie, T., and Spradling, A.C. (1998). decapentaplegic is essential for the maintenance and division of germline stem cells in the *Drosophila* ovary. *Cell* *94*, 251–260.
- Xie, T., Kawase, E., Kirilly, D., and Wong, M.D. (2005). Intimate relationships with their neighbors: Tales of stem cells in *Drosophila* reproductive systems. *Dev. Dyn.* *232*, 775–790.
- Xie, W., Wang, H., and Wu, J. (2014). Similar morphological and molecular signatures shared by female and male germline stem cells. *Sci. Rep.* *4*, 1–9.
- Xu, R., Li, Q., Yu, H., and Kong, L. (2018). Oocyte maturation and origin of the germline as revealed by the expression of Nano s-like in the Pacific oyster *Crassostrea gigas*. *Gene* *663*, 41–50.

Y

- Yakushev, E.Y., Sokolova, O. a, Gvozdev, V. a, and Klenov, M.S. (2013). Multifunctionality of PIWI proteins in control of germline stem cell fate. *Biochem. Biokhimiia* 78, 585–591.
- Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2005). Nanog expression in mouse germ cell development. *Gene Expr. Patterns* 5, 639–646.
- Yang, D., Zhang, Z., Liang, S., Yang, Q., Wang, Y., and Qin, Z. (2017). A novel role of Krüppel-like factor 4 in Zhikong scallop *Chlamys farreri* during spermatogenesis. *PLoS One* 1–16.
- Yang, H., Yao, X., Tang, F., Wei, Y., Hua, J., and Peng, S. (2018). Characterization of female germline stem cells from adult mouse ovaries and the role of rapamycin on them. *Cytotechnology*.
- Ye, H., Zheng, T., Li, W., Li, X., Fu, X., Huang, Y., Hu, C., Li, J., Huang, J., Liu, Z., et al. (2017). Ovarian Stem Cell Nests in Reproduction and Ovarian Aging. *Cell. Physiol. Biochem.* 43, 1917–1925.
- Ying, Y., Liu, X.M., Marble, a, Lawson, K. a, and Zhao, G.Q. (2000). Requirement of Bmp8b for the Generation of Primordial Germ Cells in the Mouse. *Mol. Endocrinol.* 14, 1053–1063.
- Yoshida, I., Tsuzuki, S., Abdel Salam, S.E., Ino, M., Korayem, A.M., Sakurai, S., and Iwami, M. (1997). Bombyxin F1 Gene: Structure and Expression of a New Bombyxin Family Gene That Forms a Pair with Bombyxin B10 Gene. *Zoolog. Sci.*
- Yoshida, I., Moto, K., Sakurai, S., and Iwami, M. (1998). A novel member of the bombyxin gene family: Structure and expression of bombyxin G1 gene, an insulin-related peptide gene of the silkworm *Bombyx mori*. *Dev. Genes Evol.* 208, 407–410.
- Yu, J., Zhang, L., Li, Y., Li, R., Zhang, M., Li, W., Xie, X., Wang, S., Hu, X., and Bao, Z. (2017). Genome-wide identification and expression profiling of the SOX gene family in a bivalve mollusc *Patinopecten yessoensis*. *Gene* 627, 530–537.
- Yuan, H., and Yamashita, Y.M. (2010). Germline stem cells: Stems of the next generation. *Curr. Opin. Cell Biol.* 22, 730–736.

Z

- Zatylny-Gaudin, C., Cornet, V., Leduc, A., Zanuttini, B., Corre, E., Le Corguillé, G., Bernay, B., Garderes, J., Kraut, A., Couté, Y., et al. (2016). Neuropeptidome of the Cephalopod *Sepia officinalis*: Identification, Tissue Mapping, and Expression Pattern of Neuropeptides and Neurohormones during Egg Laying. *J. Proteome Res.* 15, 48–67.
- Zhang, C., and Klymkowsky, M.W. (2007). The Sox axis, Nodal signaling, and germ layer specification. *Differentiation* 75, 536–545.
- Zhang, J., and Li, L. (2005). BMP signaling and stem cell regulation. *Dev. Biol.* 284, 1–11.

- Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Wang, J., Zhang, L., Wang, X., et al. (2012). The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490, 49–54.
- Zhang, N., Zhang, J., Purcell, K.J., Cheng, Y., and Howard, K. (1997). The *Drosophila* protein Wunen repels migrating germ cells. *Nature* 385, 64–67.
- Zhang, W., Sui, Y., Ni, J., and Yang, T. (2016). Insights into the Nanog gene: A propeller for stemness in primitive stem cells. *Int. J. Biol. Sci.* 12, 1372–1381.
- Zhao, F.-Q. (2013). Octamer-binding transcription factors: genomics and functions. *Front. Biosci. (Landmark Ed.)* 18, 1051–1071.
- Zhao, X., Yu, H., Kong, L., and Li, Q. (2012). Transcriptomic Responses to Salinity Stress in the Pacific Oyster *Crassostrea gigas*. *PLoS One* 7.
- Zhong, L., Wang, D., Gan, X., Yang, T., and He, S. (2011). Parallel expansions of sox transcription factor group b predating the diversifications of the arthropods and jawed vertebrates. *PLoS One* 6.
- Zhu, Y., Wang, C., Chen, X., and Guan, G. (2016). Identification of gonadal soma-derived factor involvement in *Monopterus albus* (protogynous rice field eel) sex change. *Mol. Biol. Rep.* 43, 629–637.

Résumé

L'huître creuse *Crassostrea gigas* est un mollusque hermaphrodite alternatif dont le cycle de reproduction est annuel. Sa gamétogenèse est soutenue par des réserves énergétiques, stockées dans un tissu conjonctif de réserve entourant la gonade. Des travaux antérieurs ont montré l'implication du système insuline dans ce processus liant étroitement alimentation, énergie stockée et gamétogenèse. Le fonctionnement des étapes précoces de gamétogenèse reste encore méconnu chez l'huître. Ce travail a permis d'identifier les cellules germinales souches (GSC) et les progéniteurs potentiels en se basant sur une approche d'histologie quantitative couplée au marquage des cellules germinales par un anticorps homologue anti-Oyvlg (Oyster vasa-like gene). Certains éléments constitutionnels de la niche germinale ont également été identifiés, notamment la présence d'une cellule somatique associée à la cellule germinale souche potentielle qui présente un marquage par un anticorps hétérologue anti-BMP2/4. En amont de l'étude concernant la régulation des étapes précoces de gamétogenèse par les voies des insulines, le criblage *in silico* des bases de données a permis d'identifier des ligands, un récepteur et de nombreux effecteurs du signal insuline conservés chez *C. gigas*. Six IRPs (Insulin Related Peptides) ont été caractérisés ce qui a permis de retracer l'histoire évolutive des IRPs chez les mollusques. D'après leur profil d'expression en qPCR et HIS, Cg-mip123 et Cg-ilp sont susceptibles d'être impliqués dans le contrôle de la reproduction. Ces IRPs peuvent se lier à un récepteur CIR (*C. gigas* Insulin Receptor) dont la séquence complète a été décrite. Les acteurs des voies de signalisation de l'insuline sont également conservés chez l'huître et exprimés dans la gonade. Pour aller plus loin sur le rôle de ces IRPs dans les étapes précoces de gamétogenèse, un conditionnement alimentaire (à jeun vs alimentés en *Isochrysis galbana*) a été réalisé sur des huîtres en première gamétogenèse. Les résultats de ce conditionnement montrent que l'apport nutritif augmente la différenciation des GSC ainsi que les proliférations goniales. L'implication du signal insuline dans ce contrôle devra être précisée.

Mots clés : *Crassostrea gigas* ; Cellules Germinales Souches ; Niche germinale ; Insulin Related Peptides ; Signalisation insuline

Structure and functioning of the germinal niche in a Lophotrochozoan, the Pacific oyster *Crassostrea gigas*

The Pacific oyster *Crassostrea gigas* is an alternative hermaphrodite mollusc with an annual reproduction cycle. Its gametogenesis is supported by energy reserves, store in a conjunctive storage tissue surrounding the gonad. Previous studies have shown the insulin system involvement in this process closely connecting diet, energy reserves and gametogenesis. The functioning of early gametogenetic stages stays unknown in the oyster. This work allows the identification of putative germ stem cells (GSC) et progenitors on the basis of histological quantitative approach combined with a germ cells labelling by homologous antibody against Oyvlg (Oyster vasa-like gene). The main components of the germinal niche have also been identified including a somatic cell, associated to the putative germ stem cell, with a heterologous antibody against BMP2/4 labelling. Above the study of the early gametogenetic stages regulation by insulin signalling, the genomic and transcriptomic-wide screening allows the identification of ligands, receptor et several effectors conserved in *C. gigas*. Six IRPs (Insulin Related Peptides) have been characterized which inform about the evolutionary history of molluscan IRPs. According to the expression profiles, in qPCR and ISH, Cg-mip123 and Cg-ilp may be involved in the reproduction process. These IRPs are able to bind the CIR (*C. gigas* Insulin Receptor) receptor whose sequence has been described. The insulin signalling effectors are also conserved in *C. gigas* and expressed in the gonad. To better understand the involvement of IRPs in the early stages functioning, a food conditioning (unfed vs fed with *Isochrysis galbana*) has been done with oysters in their first gametogenesis. The results showed that nutrient intake increases GSC and gonial mitosis. The involvement of insulin signalling has to be clarified.

Key words: *Crassostrea gigas*; Germ Stem Cells; Germinal niche; Insulin Related Peptides; Insulin signalling