



THESIS / THÈSE

MASTER EN SCIENCES BIOLOGIQUES DES ORGANISMES ET ÉCOLOGIE

Variabilité épigénétique durant le reprogramming chez le vertébré à autofécondation, *Kryptolebias marmoratus*

Goujon, Virgilio

Award date:
2019

Awarding institution:
Universite de Namur

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Variabilité épigénétique durant le reprogramming chez le vertébré à autofécondation, *Kryptolebias marmoratus*

Laboratory of Evolutionary and Adaptive Physiology

Goujon Virgilio

Mémoire présenté en vue de l'obtention du diplôme
de Master en Biologie des Organismes et Écologie



Promoteur : Prof. Frédéric Silvestre

Encadrant : Dr. Victoria Suarez-Ulloa

Année académique 2018-2019



Remerciements

Arrivé au terme de ce premier travail scientifique qu'est le mémoire, je tiens à remercier les nombreuses personnes ayant contribué à son aboutissement.

Bien sûr, mes premiers remerciements sont adressés au Professeur Frédéric Silvestre, promoteur de ce travail, qui a toujours fait preuve d'écoute et de bons conseils malgré ses nombreuses contraintes professionnelles, et qui a su me réorienter lorsqu'il le fallait.

Je remercie également mon encadrante officielle, Victoria Suarez-Ulloa, pour son aide précieuse dans le début de mémoire et pour sa bonne humeur.

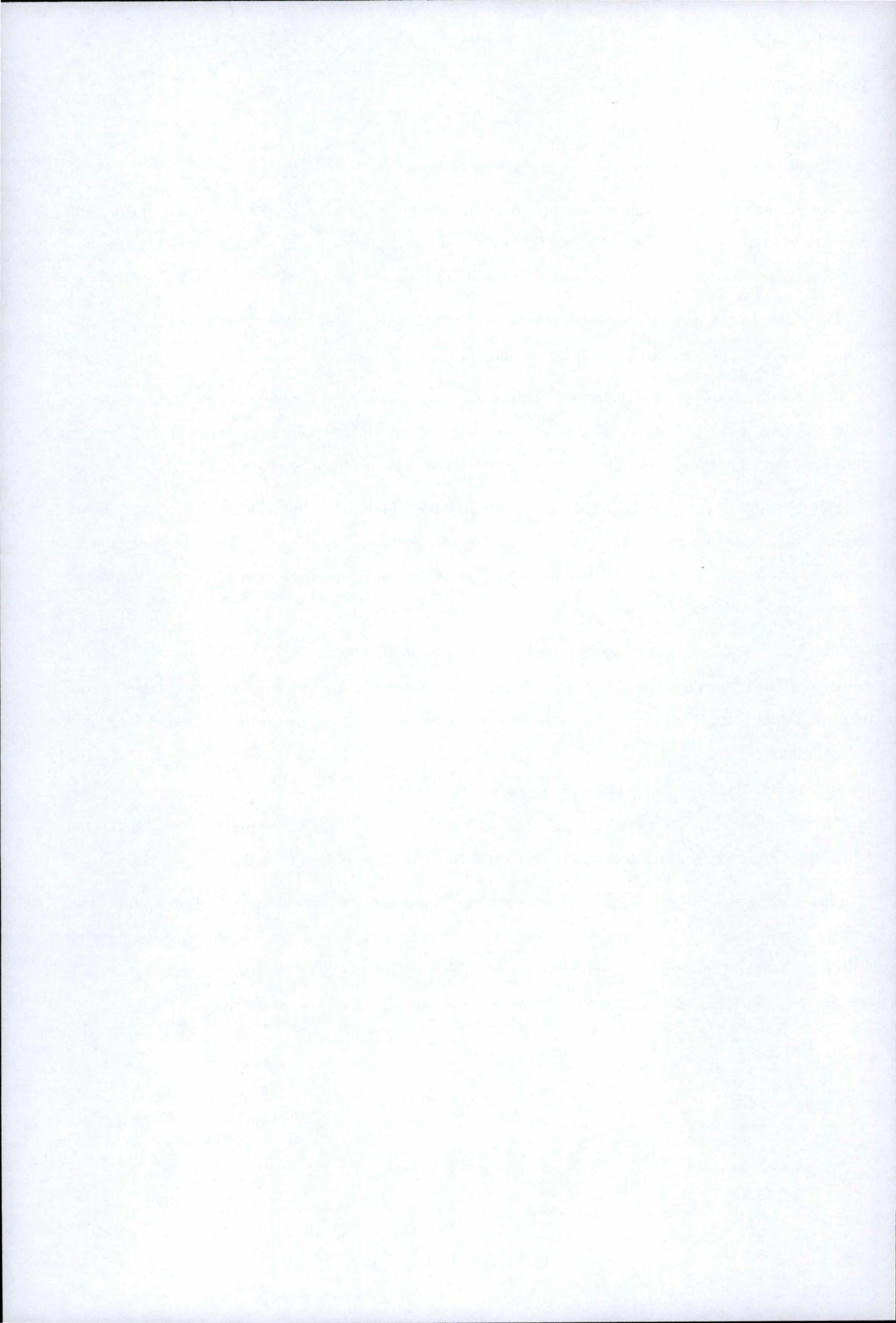
Que Alessandra Carion et Ivan Blanco trouvent en ces quelques lignes l'expression de toute ma reconnaissance pour le temps qu'ils m'ont consacré sans compter, pour leurs conseils judicieux et pour leur constante amabilité qui m'auraient fait défaut s'ils n'avaient pas été là.

Toute l'expression de ma gratitude à tous les membres de l'URBE pour l'accueil que vous m'avez réservé. Entre autres, merci à Antoine Wittorski, Valentine Chapelle et à tous ces membres de passages que sont les autres mémorants. Merci pour ce sentiment de camaraderie que vous avez entretenu et qui m'aura fait me sentir intégré dans le laboratoire.

Merci à mes parents de m'avoir permis, pour le meilleur et pour le pire, de suivre des études universitaires et de m'avoir soutenu dans toutes mes entreprises. De même, je ne peux passer à côté de mes grands parents, qui m'ont, depuis toujours, encouragés et apportés une précieuse *assistance orthographique*.

Aucun remerciement ne sera à la hauteur de ton aide et de ton soutien tout au long de ces cinq dernières années, Guddu. Au-delà de ton sourire, de tes doutes, de tes synthèses et de tes encouragements à demi-dissimulés, c'est pour ton indéfectible présence que je veux te dire merci.

Finalement, je ne saurais oublier les professeurs et assistants rencontrés durant ces cinq années au sein du département de Biologie de l'Université de Namur et de Louvain-la-Neuve pour les connaissances, et surtout, pour l'esprit critique qu'ils m'auront permis de développer, qualité ingrate à acquérir et, pourtant, tellement essentielle pour faire la part des choses de nos jours.



Abstract

The subject of this thesis is the study of post-fertilization *reprogramming* phenomenon in mangrove killifish. For that, three embryonic stages (stage Gast, OL and 27) characterizing the gradual recovery of epigenetic marks to their basal level were studied using two techniques : the study of a previous RRBS analysis allowed to highlight the methylation of the majority of CpG islands of *K. marmoratus* and pyrosequencing allowed the RRBS results reproduction of selected fragments for future transgenerational analyzes.

Analysis of the RRBS data revealed that the CpGi methylome of *K. marmoratus* was similar to that of *D. rerio* and differed from that of mammals in having higher methylation levels than the latter and that, during *reprogramming*, the distribution of hypomethylated and hypermethylated CpGs within the different genomics regions did not vary but that promoter regions appeared to exclude hypermethylated CpGs as is the case in humans or mice. The majority of RRBS fragments do not see their level of methylation vary during *reprogramming*. This leads us to question the sensitivity and ability of CpG islands to maintain an epimutation in a transgenerational way as well as the potential role of the epigenetic markings of male gametes on the methylome of the embryo and its consequences on the absence of consanguinity depression in our rivulus. The various filters used to the RRBS fragments highlighted two sets : the constant fragments and the non-constant fragments during the *reprogramming*. 24 fragments from each set were selected for pyrosequencing testing. Viable primers could only be created for five fragments and pyrosequencing analysis revealed results very similar to RRBS data. This work allowed to set up the main steps necessary to classify, isolate and specifically study RRBS data by pyrosequencing. Enrichment analyzes were performed on both sets of fragments and in both cases, the vast majority of the highlighted GO terms related regulation of biological processes, cellular communication and development. These analyzes made it possible to emphasize the specificity of non constant fragments in embryonic developmental processes.

This thesis thus constitutes a first approach to the study of the reprogramming phenomenon using recent and relevant sequencing technologies concerning the study of methylomes of a vertebrate in which the epigenetic processes would have a particularly important place.

Résumé

L'objet de ce mémoire est l'étude du phénomène de reprogrammation post-fécondation chez le killifish des mangroves. Pour ce faire, trois stades embryonnaires (stade Gast, OL et 27) caractérisant le rétablissement progressif des marques épigénétiques jusqu'à leur niveau basal furent étudiés à l'aide de deux techniques : l'étude d'une précédente analyse RRBS a permis de mettre en évidence la méthylation de la majorité des ilots CpG de *K. marmoratus* et le pyroséquençage a permis de reproduire les résultats RRBS pour certains fragments choisis dans l'optique de futures analyses transgénérationnelles.

L'analyse des données RRBS a révélé que le méthylome des CpGi de *K. marmoratus* était semblable à celui de *D. rerio* et différait de celui des mammifères en ayant des niveaux de méthylation plus élevés que ces derniers et qu'au cours du *reprogramming*, la répartition des CpG hypométhylés et hyperméthylés au sein des différentes régions génomiques ne variait pas mais que les régions promotrices semblaient exclure les CpG hyperméthylés comme c'est le cas chez l'homme ou la souris. La majorité des fragments RRBS ne voient pas leur niveau de méthylation varier durant le *reprogramming*. Ce constat nous pousse à nous questionner sur la sensibilité et la capacité des ilots CpG à conserver une épimutation de manière transgénérationnelle ainsi que sur le potentiel rôle des marques épigénétiques des gamètes mâles sur le méthylome de l'embryon et de ses conséquences sur l'absence de dépression de consanguinité évidente chez notre rivulus.

Les différents filtres apposés aux fragments RRBS ont permis de mettre en évidence deux sets : celui des fragments constants et celui des fragments non constants durant le *reprogramming*. 24 fragments de chaque set ont été sélectionnés de manière à être testés par pyroséquençage. Des amorces viables n'ont pu être créées que pour cinq fragments et l'analyse par pyroséquençage a révélé des résultats très similaires aux données RRBS. Ce travail a ainsi permis de mettre en place les principales étapes nécessaires pour classer, isoler et étudier de manière spécifique des données RRBS par pyroséquençage.

Des analyses d'enrichissement ont été réalisées sur les deux sets de fragments et dans les deux cas, la grande majorité des termes GO mis en évidence concernait la régulation des processus biologiques, la communication cellulaire et le développement. Ces analyses ont permis de souligner la spécificité des fragments non constants dans les processus développementaux embryonnaires.

Ce mémoire constitue ainsi une première approche quant à l'étude du phénomène de reprogrammation à l'aide de technologies de séquençage récentes et pertinentes concernant l'étude des méthylomes d'un vertébré chez qui les processus épigénétiques auraient une place particulièrement importante.

Table des matières

1	Introduction générale	1
1.1	Épigénétique	1
1.1.1	Définition	1
1.1.2	Principaux mécanismes	2
1.1.3	Mécanismes interconnectés	7
1.1.4	Hérédité épigénétique	7
1.1.5	Variabilité épigénétique	9
1.1.6	Implications évolutives liées à l'épigénétique	11
1.2	Le modèle : le Killifish des mangroves (<i>Kryptolebias marmoratus</i>)	14
1.2.1	Choix du modèle	15
1.2.2	Phylogénie	15
1.2.3	Distribution géographique, habitat et écologie	16
1.2.4	Cycle de vie	17
1.2.5	Reprogrammation chez <i>K. marmoratus</i>	19
1.3	Objectif général et hypothèses	20
2	Matériels et méthodes	21
2.1	Manipulations préliminaires des données RRBS	21
2.2	Méthylation gène-spécifique par pyroséquençage bisulfite	23
2.2.1	Identification des régions stables et variables différenciellement méthylées	23
2.2.2	Entretien général des rivulus	24
2.2.3	Échantillonnage/Récolte des œufs	24
2.2.4	Extraction d'ADN	25
2.2.5	Traitement bisulfite	26
2.2.6	Amorces et amplification	26
2.2.7	Pyroséquençage	27
2.3	Analyse d'enrichissement	27
2.3.1	Identification des régions stables et variables différenciellement méthylées	28
3	Résultats	30
3.1	Variabilité et méthylation intra et inter-stades	30
3.2	Méthylation gène spécifique	32
3.2.1	Identification des régions stables et variables différenciellement méthylées	32
3.2.2	Extraction d'ADN	33
3.2.3	Amorces	33
3.2.4	Traitement bisulfite et amplification de l'ADN bisulfité	34
3.2.5	Pyroséquençage	34
3.3	Analyse d'enrichissement	35
4	Discussion	38
4.1	Variabilité et méthylation intra et inter-stades	38
4.2	Analyse méthylation gène spécifique	41
4.3	Analyse d'enrichissement	43
5	Conclusion	45

Annexes	46
A	46
B	47
C	48
D	49
E	49
F	50
G	51

1 Introduction générale

1.1 Épigénétique

Le code génétique, plus petit dénominateur commun de l'information génétique, ne permet pas à lui seul de comprendre le vivant. Quel déterminisme conduit une cellule souche embryonnaire à devenir potentiellement tout type cellulaire ? Comment expliquer que certains caractères phénotypiques *acquis* par un individu peuvent être hérités à la façon de Lamarck ? Ces questions trouvent des réponses depuis la découverte des mécanismes épigénétiques, et démontrent que l'ADN n'est pas l'unique déterminant de notre phénotype.

1.1.1 Définition

Bien que l'auteur ait ignoré la nature exacte des gènes et de leurs relations avec l'ARN et les protéines, on attribue la paternité des concepts d'épigénétique à l'embryologiste Conrad Hal Waddington. A l'origine, ce terme désignait les interactions complexes entre environnement et génome impliquées dans le développement et la différenciation des organismes supérieurs [Bird, 2007]. Historiquement, cette définition fut utilisée pour décrire les mécanismes développementaux susceptibles de donner lieu à des phénotypes alternatifs et venait combler les lacunes du modèle génétique qui ne pouvait expliquer tous les phénomènes liés à la différenciation cellulaire. Elle est illustrée par le concept de *paysage épigénétique* du même auteur [Waddington, 1942] (Figure 1).

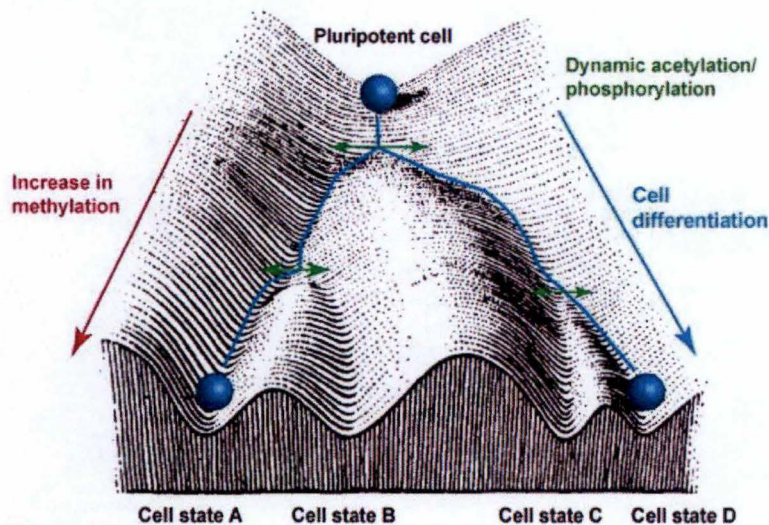


FIGURE 1 – *Illustration du paysage épigénétique de Waddington. La bille représente une cellule pluripotente et la colline, les nombreuses bifurcations et trajectoires épigénétiques possibles tout au long de son parcours, conduisant à différentes cellules spécialisées [Srinageshwar et al., 2016].*

Au cours des années, l'épigénétique de Waddington s'est considérablement modifiée. Aussi, de nouvelles définitions furent proposées et de nos jours, il est communément admis que ce terme

désigne l'étude des changements héréditaires de l'expression génique ne pouvant être expliqués par des modifications dans les séquences nucléotidiques [Jablonka et Lamb, 2002]. Dans ce contexte, le mot hérédité est défini comme impliquant à la fois l'héritage méiotique et mitotique. Ainsi, les mécanismes épigénétiques ne doivent pas se limiter aux processus transmis d'une génération à l'autre, mais concernent aussi ceux hérités à travers les divisions cellulaires au sein de l'individu [Metzger et Schulte, 2016].

1.1.2 Principaux mécanismes

Les mécanismes constituant un signal ou un processus épigénétique peuvent être classés en trois catégories ayant comme point commun la régulation des interactions ADN-protéines : la modification covalente de l'ADN par méthylation des cytosines ; la modification covalente des histones via l'acétylation, la phosphorylation, la méthylation, etc et la régulation de l'expression génique au niveau transcriptionnel et post-transcriptionnel par des ARN non codants [Srinageshwar *et al.*, 2016]. Associés, ces trois mécanismes (Figure 2) peuvent définir des zones euchromatiques favorables à l'activation de la transcription, et des zones hétérochromatiques favorables à la répression de la transcription. Cette vision est toutefois simpliste, et plusieurs travaux démontrent que la dynamique associée à l'état chromatidien est en réalité plus complexe [Hubert *et al.*, 2010].

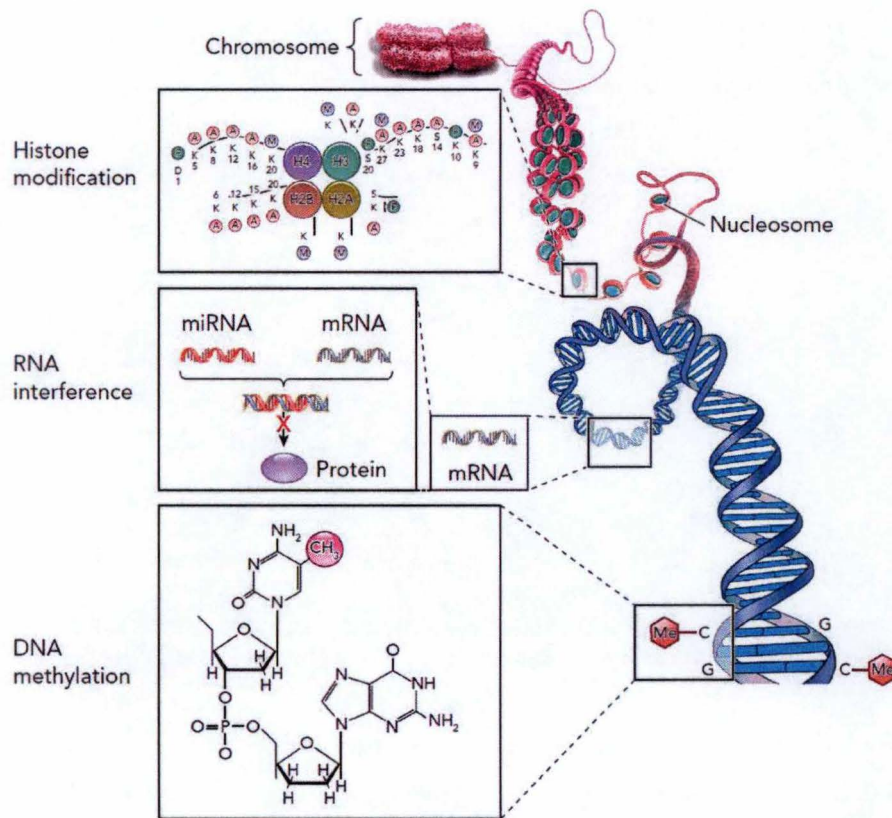


FIGURE 2 – Principaux mécanismes épigénétiques. La chromatine rajoute une information à celle de l'ADN sous la forme de méthylation de l'ADN, de modifications post-traductionnelles des histones et des ARN non codants [Puumala et Hoyme, 2015].

Modifications des histones

Les histones sont des protéines extrêmement conservées, on les retrouve depuis les archées jusqu'aux eucaryotes supérieurs. Elles sont organisées en nucléosomes, structure octamérique autour de laquelle s'enroule le matériel génétique (**Figure 3**). La succession des nucléosomes le long de la molécule d'ADN lui confère une structure en *collier de perles*, premier niveau de compaction de la chromatine.

Un octamère d'histone est un complexe protéique composé de doublets de molécules d'histone, dites *de cœurs*, les H2A, H2B, H3 et H4. Un autre type d'histones, les H1, sont présentes hors du *cœur* du nucléosome. Elles sont dites *de liaison* car leur rôle est de condenser les gènes entre chaque nucléosome et de *sceller* cette structure [Fan et Roberts, 2006].

Les histones peuvent subir des modifications post-traductionnelles sur certains acides aminés de leurs queues N et C terminales. Ces modifications sont de type acétylation, méthylation, phosphorylation, sumoylation, ou encore ubiquitylation [Berger, 2007]. Elles sont réversibles et agissent en synergie pour interagir avec le degré de compaction de la chromatine [Strahl et Allis, 2000].

La combinaison des variants d'histones et les modifications post-traductionnelles qu'elles portent constituent une information combinatoire complexe résumée sous le terme de *code des histones*. Ce concept fut apporté pour essayer de donner un sens à la complexité des modifications des histones, mais aujourd'hui, il tend à être considéré comme un abus de langage, aucun code n'ayant été découvert malgré des années de recherche [Henikoff et Shilatifard, 2011].

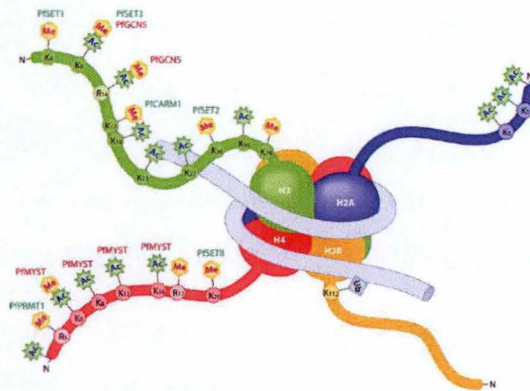


FIGURE 3 – *Représentation d'un nucléosome et des histones associés. L'histone H1 n'est pas représenté* (<http://www.amsbio.com>).

ARN non codants

L'ARN non codant (RNAnc pour *non-coding RNA*) fait référence à un ARN issu de la transcription n'étant pas traduit en protéine par les ribosomes et ayant une fonction complètement différente de l'ARN messager, de l'ARN de transfert et de l'ARN ribosomique. Il en existe au moins deux grandes classes : les petits de 20 à 30 nucléotides de longueur (ou *short ncRNA*), et les grands d'envi-

ron 200 nucléotides (ou *long ncRNA*). Actuellement, ces ARN s'imposent comme une nouvelle classe de gènes de régulation impliqués dans divers processus cellulaires tels que le développement, la différenciation et le métabolisme cellulaire [Hubert *et al.*, 2010], tout en jouant un rôle important dans la régulation dynamique des chromosomes et les modifications de la chromatine [Greiner, 2012].

Les petits ARNnc agissent en séquestrant ou en dégradant les ARN messagers via un appariement bases-séquences spécifiques ciblées, permettant une régulation de l'expression des gènes au niveau post-transcriptionnel [Siomi et Siomi, 2009].

Les premiers *short ncRNA* découverts furent les petits ARN interférents (siRNA pour *small interfering RNA*), de nouvelles familles, se distinguant par leur taille, par l'enzyme qui les clive et par la protéine qui les prend en charge, sont régulièrement mises à jour. Ainsi, en plus des siRNA, les micro ARN (ou miRNA), les ARN interagissant avec Piwi (piRNA pour *Piwi-interacting RNA*), ou encore les petits ARN interférents endogènes (endo-siRNA pour *endogenous siRNA*) sont impliqués dans différentes voies de régulation [Frias-Lasserre et Villagra, 2017].

Les *long ncRNA* régulent l'expression génique par divers mécanismes encore mal définis. Ces transcrits ont une longueur d'environ 200 pb et sont impliqués dans la régulation post-traductionnelle des gènes en participant à des processus tels que la synthèse protéique ou la maturation des ARN. Ils formeraient aussi des complexes ribonucléoprotéiques interagissant avec la chromatine, afin de réguler les modifications des histones et l'architecture de la chromatine [Wang et Chang, 2011].

Méthylation de l'ADN

La recherche concernant la méthylation de l'ADN est devenue extrêmement florissante depuis la découverte de son implication dans une grande variété de fonctions cellulaires incluant l'expression génique spécifique au tissu, la différenciation cellulaire, le développement, la reprogrammation, l'empreinte génomique, l'inactivation du chromosome X, la régulation du degré de condensation de la chromatine et son rôle dans certaines pathologies [Varriale, 2014].

Chez les eucaryotes, l'addition covalente d'un groupement méthyle provenant d'une S-Adenosyl Methionine (SAM) à une base cytosine, définissant la réaction de méthylation de l'ADN, est catalysée par une famille d'enzymes appelée DNMT (pour *DNA methyltransferase*). Le carbone en position 5' des dinucléotides CpG (une cytosine directement suivie d'une guanosine) est la principale cible de ces enzymes. Le résultat de sa méthylation est une 5-méthylcytosine (5-mC) [Bestor, 2000] bien que des cytosines méthylées puissent également apparaître sur des sites non CpG tels que les motifs palindromiques CpHpG et asymétriques CpHpH (où H = A, C ou T) ; ces derniers motifs, rares chez les animaux, sont prédominants dans le paysage épigénétique des plantes [Law et Jacobsen, 2010]. A son tour, une 5-mC peut être déméthylée. Bien que les mécanismes exacts de déméthylation soient encore incertains, ils font intervenir des dioxygénases, des enzymes de la famille des TET (pour *Ten-*

eleven Translocation Enzyme) et des TDG (pour *G/T mismatch-specific thymine DNA glycosylase*). La voie la mieux caractérisée fait intervenir un processus en plusieurs étapes passant tout d'abord par l'hydroxylation des 5-mC en 5-hydroxyméthylcytosine (5-hmC) via les protéines TET. Le rôle des 5-hmC reste encore flou [Cadet et Wagner, 2014] mais elles ont été principalement identifiées dans les cellules souches embryonnaires et les neurones adultes, où elles représentent 15 à 40% des 5-mC; elles sont aussi présentes dans tous les types cellulaires mais n'y constituent que 1 à 5% des 5-mC [Tahiliani *et al.*, 2009]. Leur spécificité tissulaire suggère qu'elles ne peuvent pas être exclusivement un état intermédiaire d'un processus de déméthylation actif mais, au contraire, qu'elles ont probablement une fonction unique encore à élucider [Munzel *et al.*, 2011]. Une 5-hmC peut encore être oxydée en 5-formylcytosine (5-fC) et ensuite en 5-caC (5-carboxylcytosine). Enfin, ces derniers peuvent être ramenés à la cytosine originelle par le biais du système de *base excision-repair* et des enzymes de la famille TDG. En outre, une déméthylation passive de l'ADN peut se produire lorsque les DNMT1 ne parviennent pas à reproduire les motifs de méthylation au cours de la réplication. Un autre mécanisme possible comprend l'action de l'enzyme AID/APOBEC (pour *Activation-induced deaminase and Apolipoprotein B mRNA editing enzyme*) capable soit de convertir une 5-mC en thymine, soit de convertir une 5-hmC en 5-hydroxyméthyluracile (5-hmU). La thymine et la 5-hmU peuvent également être clivées par les enzymes TDG et remplacées par une cytosine non méthylée [Moore *et al.*, 2013] (Figure 4).

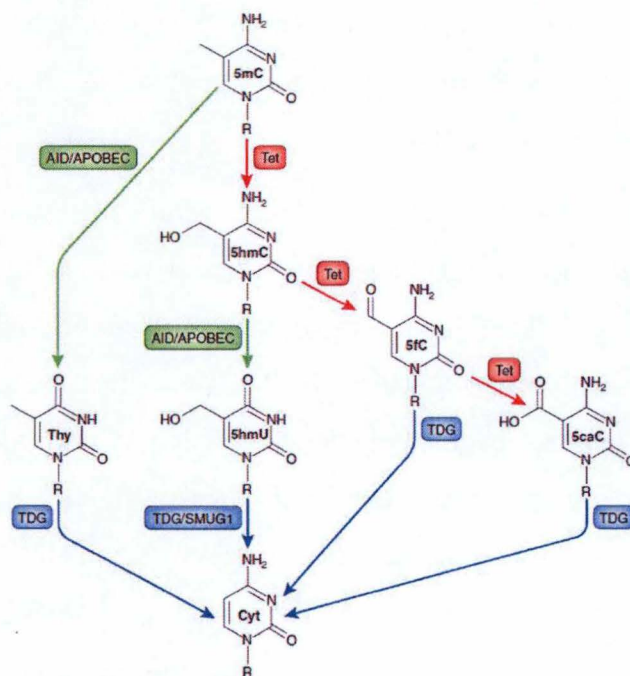


FIGURE 4 – *Voies actives de déméthylation de l'ADN.* Les enzymes AID/APOBEC et/ou les enzymes TET peuvent convertir les cytosines méthylées en thymine, 5-hmU, 5-fC ou 5-caC. Finalement, ces produits sont reconnus et remplacés par une cytosine non méthylée grâce au système de base excision-repair et à des enzymes TDG. [Moore *et al.*, 2013].

Alors que les processus de méthylation et de déméthylation sont de mieux en mieux compris, les mécanismes par lesquels ces enzymes ciblent des locus génomiques spécifiques pour être méthylés ou déméthylés restent inconnus [Liu *et al.*, 2016].

La cytosine méthylée est une modification dynamique non seulement parce qu'elle est ajoutée et éliminée par des processus enzymatiques, mais aussi parce qu'elle peut spontanément muter en se désaminant (perte du groupement amine NH₂) en thymine [He *et al.*, 2015] (en revanche, la désamination spontanée de la cytosine non méthylée est facilement reconnue par la machinerie de réparation de l'ADN et corrigée [Stier et Kiss, 2013]).

Chez les vertébrés, l'établissement et la maintenance de ce type de modification épigénétique sont sous la responsabilité de trois familles de méthyltransférases de l'ADN (DNMT) : les DNMT1, les DNMT2, et les DNMT3 comprenant les DNMT3a, 3b, 3c et 3l. Leur activité est régulée transcriptionnellement et post-transcriptionnellement par des modifications (phosphorylation et ubiquitination) ainsi que par des cofacteurs protéiques, notamment via le domaine N-terminal des méthyltransférases [Trelle et Jensen, 2007]. Les DNMT1 sont impliquées dans le maintien de la méthylation après chaque division cellulaire, au moment où les molécules d'ADN sont des substrats hémiméthylés [Reik *et al.*, 2001], alors que les DNMT3a et 3b sont responsables des méthylation *de novo* et jouent un rôle crucial dans les cellules germinales et au cours du développement précoce. Les DNMT3l sont exprimées spécifiquement dans les cellules germinales et sont, notamment, responsables de l'empreinte génomique. Les DNMT2 sont considérées comme un type de DNMT en raison de leurs fortes similarités structurales mais leur activité catalytique est faible, de mécanismes flous, et ne semble toucher que les ARNt [Kiani *et al.*, 2013].

Historiquement, la méthylation de l'ADN est liée à la répression transcriptionnelle [Mcghee et Ginder, 1979]. Cependant, l'analyse comparative de méthylomes appartenant à différents taxons a révélé des schémas de méthylation très divergents selon les espèces [Zemach *et al.*, 2010; Lee *et al.*, 2010], étayant la conclusion selon laquelle la méthylation ne peut être un mécanisme universel de répression de l'expression génique. En effet, la méthylation de l'ADN est fréquente chez les eucaryotes, allant des champignons aux vertébrés, mais sa fonction varie considérablement chez ces derniers. Cependant, deux caractéristiques du méthylome semblent être hautement conservées : l'hypométhylation des régions riches en CpG, appelées îlots CpG (CpGi), et l'hyperméthylation des éléments transposables (ET) [Zemach *et al.*, 2010]. En moyenne, la teneur en CpG dans le génome est relativement faible avec 40% pour *Xenopus*, 42% chez la souris/homme et 37% chez le poisson-zèbre [Potok *et al.*, 2013]. Le pourcentage de méthylation globale des CpG dans le génome des vertébrés est compris entre 60 et 90% [Head, 2014]. Si la majorité (85%) de ces CpG méthylés se situent dans les séquences répétées, on les retrouve dans toutes les régions du génome [Hendrich et Tweedie, 2003].

Chez l'humain, 75% de tous les promoteurs se trouvent dans les îlots CpG et sont largement non méthylés, tandis que le reste des promoteurs porte des densités en CpG très faibles [Edwards *et al.*, 2017]. La méthylation est généralement exempte des régions codantes, permettant un état permissif sur le plan transcriptionnel [Illingworth et Bird, 2009].

Chez le poisson-zèbre, seulement 15% des CpG sont trouvés dans les CpGi des régions promotrices de quelques 70% des gènes codants [Sandelin *et al.*, 2007], suggérant que de nombreux gènes peuvent ne pas être régulés par la méthylation de l'ADN au niveau des promoteurs. En fait, l'expression génique a également été fortement corrélée à la méthylation des amplificateurs, des régions régulatrices ou du corps des gènes dans des contextes différents [Yang *et al.*, 2014, Elliott *et al.*, 2015]. Par ailleurs, des gènes réduits au silence ont des promoteurs entièrement non méthylés et, inversement, de nombreux gènes hautement exprimés présentent des niveaux élevés de méthylation des promoteurs [Bestor *et al.*, 2015]. Au sein même des vertébrés, si la compréhension de l'impact de la méthylation de l'ADN sur l'expression des gènes évolue constamment, le méthylome, en tant que régulateur de la transcription, devient aussi de plus en plus complexe.

Les poissons semblent également avoir moins d'îlots CpG et des pourcentages de méthylation plus élevés que les autres vertébrés tels que les mammifères et les oiseaux. Cette différence est attribuée à un taux de désamination plus élevé des 5-mC (et donc à un taux de 5-mC moins élevé) chez les animaux endothermiques. Cette hypothèse est soutenue par l'observation des niveaux de 5-mC qui sont inversement proportionnels à la température corporelle chez les espèces polaires, tempérées et tropicales [Head, 2014].

1.1.3 Mécanismes interconnectés

Les différents mécanismes épigénétiques sont fortement connectés. Il est maintenant établi que ces processus ne se produisent pas indépendamment les uns des autres mais sont en relation pour moduler l'expression génique. Certaines évidences tendent à démontrer que méthylation et histones travaillent de concert [Nan *et al.*, 1998], ainsi les DNMT1 peuvent agir en synergie avec les DNMT3a et 3b, certaines histones méthyltransférases et déacétylases [Groth *et al.*, 2007]. De même les siRNA sont capables d'activer la méthylation de l'ADN grâce au recrutement de DNMT par le complexe RISC (pour *RNA-induced silencing complex*) [Vandegheuchte et Janssen, 2011]. L'existence de ces multiples boucles de régulation rend difficile l'établissement de liens de cause à effet entre un mécanisme épigénétique et un phénotype particulier.

1.1.4 Hérité épigénétique

Pour persister transgénérationnellement, les marques épigénétiques doivent être méiotiquement et mitotiquement stables afin d'échapper aux deux cycles de déméthylation et de reméthylation qui reprogramment la lignée germinale et le développement embryonnaire précoce pour produire un

zygote totipotent, ou bien, occasionner le transfert de l'information essentielle au rétablissement des marques épigénétiques à la génération suivante [Yohn et Blendy, 2016].

Dynamique de reprogrammation

Le terme *reprogrammation* désigne les étapes de suppression et rétablissement des marques épigénétiques se produisant lors de la gamétogenèse et lors de l'initiation du développement embryonnaire [Reik *et al.*, 2001].

Chez les mammifères, la première phase de déméthylation survient au cours de la migration et de la prolifération des cellules germinales primordiales. Elle est suivie d'une reméthylation post-migratoire. Une deuxième phase de déméthylation survient peu de temps après la fécondation, pendant les étapes de clivage jusqu'au stade blastocyste. Finalement, une méthylation *de novo* rétablit les profils de méthylation à partir du stade morula [Edwards *et al.*, 2017] (Figure 5).

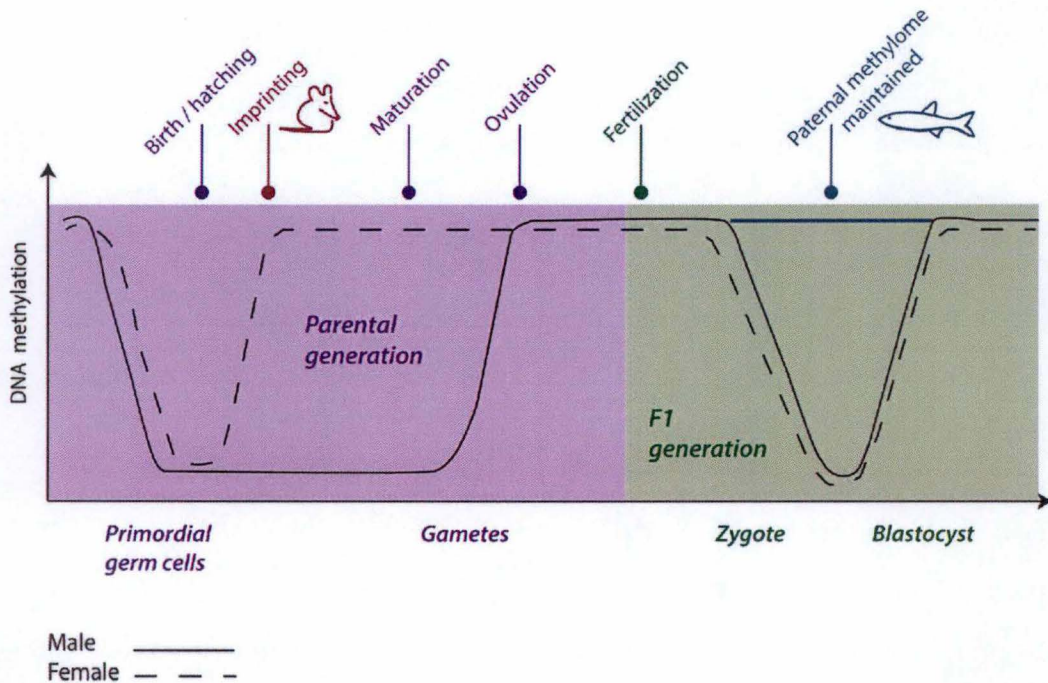


FIGURE 5 – *Dynamique de la reprogrammation du méthylome depuis les gamètes jusqu'à l'embryogenèse chez les mammifères et les poissons.* Au moment de la genèse de la génération F0 (parentale), le méthylome est complètement éliminé dans les cellules germinales primordiales. Les gamètes subissent alors une méthylation *de novo* suivie du phénomène d'empreinte chez les mammifères que l'on ne retrouve pas chez les poissons, probablement en raison du manque d'homologue à la DNMT-3l). Un deuxième cycle de déméthylation/reprogrammation se produit peu après la fécondation en génération F1 chez les mammifères. Chez les poissons, le méthylome maternel subit ce cycle, alors que l'équivalent paternel persiste. [Brander *et al.*, 2017].

Ces événements permettent aux cellules germinales de se différencier correctement et sont cruciaux au bon développement de l'embryon [Seisenberger *et al.*, 2012]. Bien que le génome ait initialement été supposé uniformément affecté par la reprogrammation, il semble aujourd'hui que les différentes régions génomiques, ainsi que les génomes paternel et maternel, diffèrent dans le moment

et l'ampleur de leur reprogrammation. Ainsi durant cette période, l'individu subit une réinitialisation complète du méthylome, à l'exception des gènes à empreintes (pour les mammifères), des régions centromériques/péricentromériques, et des rétrotransposons [Edwards *et al.*, 2017].

Malgré une dynamique de reprogrammation semblable entre la souris et le poisson zèbre, des différences de reprogrammation distinguent ces deux organismes. Une des dissemblances majeures réside dans le fait que le zygote de *D. rerio* atteint la totipotence lorsque son génome est activé en conservant les degrés de méthylation du génome paternel tout en reprogrammant le génome maternel (ovocyte) pour qu'il coïncide avec l'état de méthylation des spermatozoïdes contrairement à la souris, chez qui ce serait le génome maternel qui servirait de modèle.

Le phénomène de reprogrammation pourrait également jouer un rôle important durant cette fenêtre critique (ou période sensible) qu'est le développement embryonnaire au cours duquel le phénotype d'un organisme peut réagir à des facteurs intrinsèques ou extrinsèques (environnementaux). En effet, le méthylome pourrait être particulièrement sensible à des signaux environnementaux, et les altérations épigénétiques qui en résulteraient pourraient influencer le développement, le comportement, la structure et/ou la fonction des tissus et des organes voire les propriétés épigénomiques futures de l'individu en question [Nagy et Turecki, 2012, Champagne, 2013].

1.1.5 Variabilité épigénétique

On distingue trois types de variation épigénétique en fonction de leur dépendance vis-à-vis de leur contexte génomique : obligatoire, facilitée et pure.

Dans la mesure où l'état épigénétique est strictement déterminé par le génotype, la variation épigénétique est décrite comme *obligatoire*. Il existe alors une correspondance entre l'épigénotype et la variation génétique agissant en *cis* ou en *trans*¹). Un exemple de ce type de variation est la floraison hâtive en absence de vernalisation et d'expression de l'allèle récessif pour le gène FRI chez *Arabidopsis* : la transition entre le stade végétatif et le stade reproductif, définissant la floraison, est contrôlée par des processus endogènes et des signaux environnementaux dont la vernalisation (période de froid subie par une plante nécessaire pour enclencher sa floraison). Plusieurs gènes sont impliqués dans ce processus dont FRIGIDA (FRI) et le locus de floraison C (FLC) [Liu *et al.*, 2004]. Le gène FLC est un répresseur de la floraison et est négativement régulé par la vernalisation tandis que le gène FRI est lié à la régulation de FLC [Sheldon *et al.*, 2002]. Une floraison hâtive n'est possible que si, pour au moins un locus, l'allèle récessif est exprimé. La souche *Landsberg erecta* possède un allèle nul pour le gène FRI (c'est-à-dire, qu'il fournit un produit génique non fonctionnel) et même en absence de

1. Une séquence d'ADN ayant un effet régulateur sur la transcription d'un gène est appelée élément *cis*-régulateur (agit en *cis*). En revanche, lorsque le taux de transcription est modifié par des protéines se fixant sur ces séquences régulatrices, ces protéines sont des facteurs *trans*-régulateurs (agit en *trans*). La régulation en *cis* fait intervenir des séquences spécifiques alors que la régulation en *trans* fait intervenir des protéines se fixant sur les séquences *cis* [Li *et al.*, 2012].

vernalisation, la floraison est hâtive. Ceci est rendu possible par la présence d'un élément transposable dans le premier intron du gène FLC. Son insertion en cis engendre une plus faible expression du gène FLC par l'intermédiaire de siRNA. En présence de cet élément transposable, la floraison est toujours hâtive chez cette souche [Liu *et al.*, 2004]. L'état épigénétique est donc strictement déterminé par le génotype puisque l'expression du gène FLC est obligatoirement régulée par les siARN produits en présence du transposon.

Lorsque la variation épigénétique est encouragée à apparaître dans un contexte génomique spécifique, on la décrit comme *facilitée*. Le génotype dirige ou potentialise l'épigénotype d'une manière probabiliste mais pas strictement déterministe. Un exemple de variation épigénétique facilitée est celui du gène Agouti chez la souris. La variation génétique, sous la forme de l'insertion d'un rétrotransposon, plus ou moins méthylé, en amont de la séquence codante Agouti, conduit à une sous/surexpression de ce gène et affecte la couleur du pelage. Une mosaïque de phénotypes allant du brun (normal, rétrotransposon méthylé) à jaune (mutant, rétrotransposon non méthylé) est dès lors possible [Morgan *et al.*, 1999]. En absence de l'élément rétroviral, le phénotype sauvage est exprimé tandis que la régulation épigénétique de celui-ci facilite l'apparition de toute une gamme de phénotypes.

Enfin, si la variation épigénétique n'est d'aucune manière sous l'influence du génotype, elle est décrite comme *pure* et est le résultat d'événements stochastiques indépendants de la variation génétique [Richards, 2006]. Par exemple, le degré de méthylation d'un gène peut être altéré en réponse à l'environnement et/ou par des processus aléatoires. Les épimutations, définies comme des modifications de l'état épigénétique initial de séquences d'ADN qui en résultent peuvent avoir un effet sur le phénotype des individus, et le résultat peut être soit avantageux (par exemple, en augmentant la plasticité phénotypique dans un contexte de stress environnemental [Angers *et al.*, 2010]), soit neutre, soit désavantageux (par exemple, la perte de l'empreinte génomique sous l'influence d'agents tératogènes [Kappil *et al.*, 2015]).

Tout comme les mutations, les épimutations apparaissent et s'établissent de façon aléatoire au cours du développement et/ou pendant la vie adulte des organismes. Mais, alors que les mutations somatiques sont rares et irréversibles, leurs homologues épigénétiques seraient 100 fois plus fréquentes et réversibles (bien que les mécanismes exacts d'inversion soient encore incertains). Les épimutations représentent ainsi un potentiel important de variations [Baker *et al.*, 2003]. Il n'est cependant pas évident de prévoir l'effet ponctuel d'une épimutation sur le phénotype. En effet, la mesure dans laquelle l'épigénotype contribue à un changement du phénotype est quelque peu controversée et est difficile à démêler des contributions génétiques et environnementales. En général, les preuves sont indirectes, c'est-à-dire que l'effet épigénétique est considéré comme une source non négligeable de variabilité phénotypique après que des effets génétiques et environnementaux aient été exclus

[Peaston et Whitelaw, 2006]. Cependant, la variation de l'état de méthylation d'un seul dinucléotide CpG, ainsi que la densité de méthylation au niveau du promoteur d'un gène, peuvent avoir une influence capitale sur son expression comme le montre la tolérance au stress provoquée par les comportements maternels chez *Rattus norvegicus* [Weaver *et al.*, 2004].

Il est possible de se faire une idée de l'impact, à l'échelle du génome, des effets transcriptionnels épigénétiques par l'étude de souches animales particulières comme l'hybride *Chrosomus eos-neogaeus* à reproduction clonale, ou le killifish *Kryptolebias marmoratus* à autofécondation dans des conditions environnementales définies. Ces derniers sont un moyen prometteur d'isoler la variabilité non génétique.

Aussi [Massicotte *et al.*, 2011] a inspecté l'étendue de la variabilité épigénétique, et plus particulièrement du polymorphisme de méthylation de l'ADN non lié à la variabilité génétique dans des populations asexuées de *C. eos-neogaeus* en milieu naturel. Cette étude suggère une importante part de hasard dans l'établissement des profils de méthylation entre les individus, et ce, à deux niveaux. Premièrement, un nombre élevé de dinucléotides CpG a été observé, pour lesquels l'état de méthylation est variable entre les cellules d'un même tissu, occasionnant une mosaïque de patrons de méthylations possibles (chaque cellule d'un même tissu pouvant présenter un profil de méthylation différent). Ensuite, les données montrèrent une association entre le profil épigénétique et une condition physico-chimique de l'environnement local (le pH des lacs où furent prélevés les poissons). Ainsi, les variations naturelles des conditions environnementales pourrait mener à du polymorphisme de méthylation. Cette dernière constatation fut depuis lors réalisée à plusieurs reprises chez d'autres organismes montrant des profils de méthylation spécifique à un environnement [Schrey et Richards, 2012, Skinner *et al.*, 2014, Smith *et al.*, 2016]. La variabilité épigénétique se produisant à la suite d'un changement environnemental semble ainsi établir de manière aléatoire des mutations épigénétiques.

1.1.6 Implications évolutives liées à l'épigénétique

« Natural selection may explain the survival of the fittest, but it cannot explain the arrival of the fittest. »

Hugo De Vries

La théorie synthétique de l'évolution ou néodarwinienne argue que seule la diversité génétique est source de variation héréditaire : des mutations aléatoires créent une variabilité génétique qui promeut l'évolution par le biais de la sélection naturelle [Mayr et Provine, 1998]. Cependant, il est aujourd'hui évident que la variation génétique n'est pas la seule source de diversité phénotypique héréditaire. Les mécanismes épigénétiques sont de potentiels générateurs de variabilité phénotypique, et auraient donc une éventuelle contribution dans l'évolution des espèces. A tel point que certains ont proposé une théorie unifiée de l'évolution expliquant comment l'environnement peut d'une part agir

pour influencer directement la variation phénotypique, et d'autre part, faciliter la sélection naturelle [Skinner, 2015] (Figure 6).

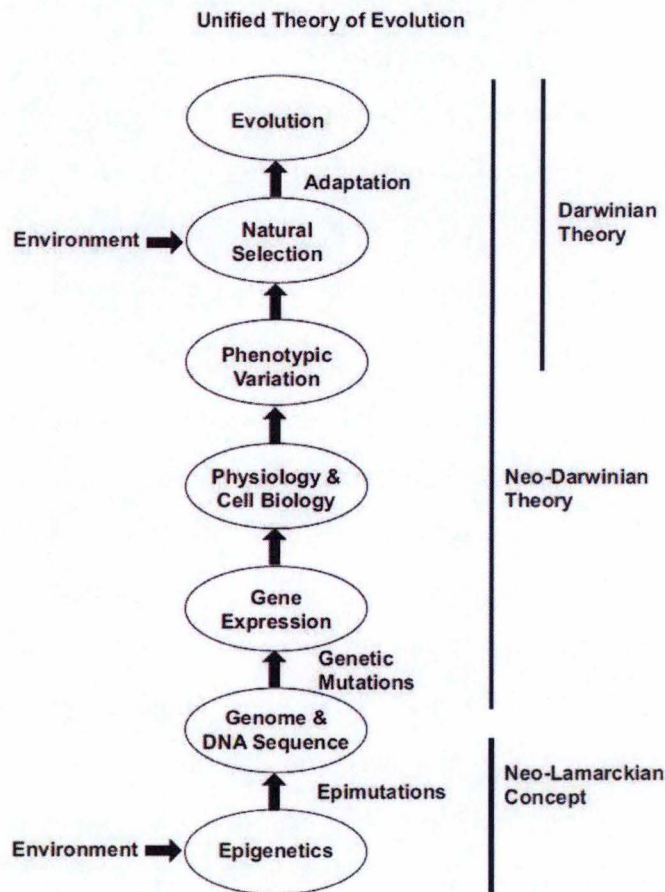


FIGURE 6 – *Schéma de la théorie unifiée de l'évolution.* Toutes les composantes (symbolisées par des cercles) de cette théorie sont représentées comme ayant une importance égale. Aucune dominance n'est suggérée comme, par exemple, entre épimutations et génétique [Skinner, 2015].

L'un des principes de base du concept d'évolution est que la sélection naturelle façonne les populations et les espèces au cours du temps en agissant sur les phénotypes améliorés (ou diminués) par l'accumulation de mutations. Contrairement à une mutation qui se manifeste aléatoirement à basse fréquence, qui peut être facilement perdue par dérive génétique et qui n'affecte qu'un individu à la fois pour n'être que lentement amplifiée par la sélection naturelle sur des centaines voire des milliers de générations si elle est avantageuse [Rando et Verstrepen, 2007], les modifications épigénétiques peuvent apparaître beaucoup plus rapidement en réponse à un stress environnemental et toucher simultanément plusieurs générations, sinon la totalité d'une population d'individus [Klironomos *et al.*, 2013]. Théoriquement, un phénotype avantageux hérité d'origine épigénétique est aussi rapidement réversible selon que la plasticité phénotypique est coûteuse à maintenir et que la perturbation environnementale se dissipe (ou resurgisse) [Burggren, 2014].

L'avantage de l'hérédité épigénétique dans un environnement dynamique, où les stress environnementaux peuvent ne durer que quelques générations, est illustré dans la **Figure 7** à travers plusieurs

scénarios.

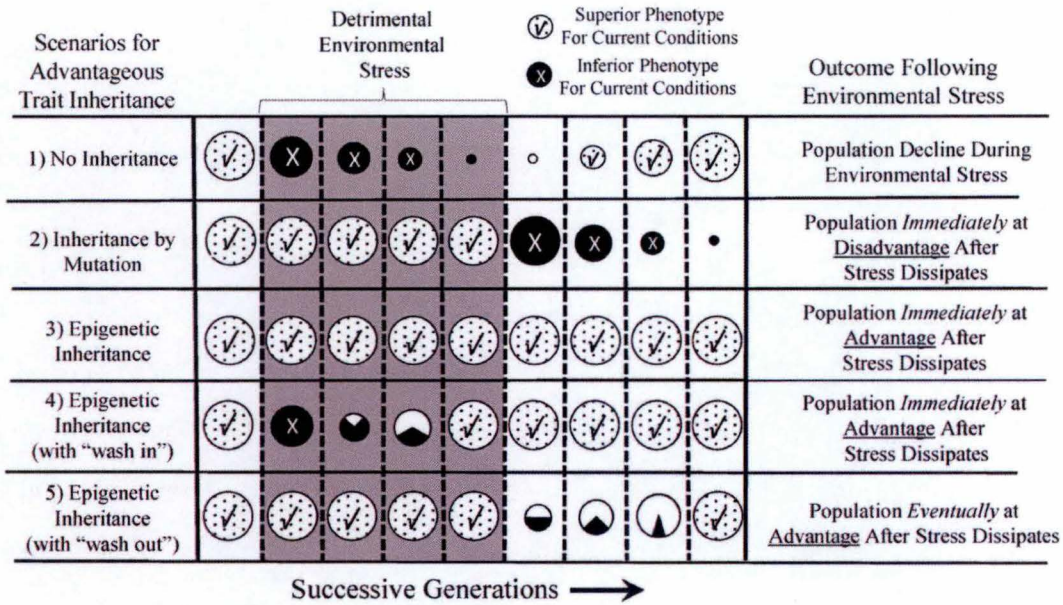


FIGURE 7 – Représentation de cinq scénarios envisageables où l'avantage de l'hérédité épigénétique pourrait se révéler dans une population vivant dans un environnement stressant. Chaque ligne représente une population en transition sur plusieurs générations, un cercle représentant une génération dont le diamètre reflète la taille de la population. Toutes les générations F0 (la première colonne de cercles à gauche) sont supposées avoir un phénotype adapté à l'environnement précédant la perturbation. Dès l'apparition du stress environnemental, les générations successives de chaque population déclinent, ou bien, prospèrent avec un phénotype acquis par mutation ou par hérédité épigénétique. De même, à mesure que le stress s'estompe jusqu'à disparaître, les générations suivantes ont soit un phénotype avantageux ou désavantageux par rapport à l'environnement original (exempt de stress) [Burggren, 2016].

En général, on considère que la persistance d'un phénotype avantageux lié aux phénomènes épigénétiques résulte de la transmission héréditaire de marqueurs épigénétiques sur de nombreuses générations. Pourtant, une question-clé reste largement sans réponse : dans quelle mesure les phénotypes épigénétiquement hérités se fixent-ils dans le génome, plutôt que d'avoir une influence indirecte sur celui-ci ? Depuis peu, la recherche porte une attention considérable sur les CpGi méthylés, considérés comme des *hot spots* mutationnels en raison de la conversion spontanée des cytosines méthylées en thymines. Ces mutations, non nécessairement reconnues comme ADN lésé, ne sont pas sujettes à l'excision ou à la correction [Bateson, 2012]. Toutefois, ce type d'épimutation n'a encore été lié à la genèse d'aucun phénotype contribuant à l'évolution ou à la spéciation d'une espèce [Burggren, 2016].

A ce point, il est intéressant de préciser qu'il n'est pas nécessaire que le phénotype associé aux phénomènes épigénétiques soit héréditaire pour influencer l'évolution d'une espèce. Dans la mesure où les épimutations environnementales occasionnent l'apparition de phénotypes alternatifs favorables, les individus portant ce(s) génotype(s) sensible(s) aux changements environnementaux auront une plus grande chance de survie et de reproduction. En ce sens, il apparaît moins crucial que les épimutations environnementales soient héréditaires puisque le(s) génotype(s) sensible(s) augmentera(ont) en

fréquence permettant ainsi de faire face aux conditions de l'environnement à la génération suivante [Massicotte *et al.*, 2011].

L'environnement induit aussi de multiples effets épigénétiques que l'on pourrait qualifier d'*indirects* jouant sur la stabilité des gènes et les taux mutationnels; indirects parce que l'effet des modifications épigénétiques générées par un stress environnemental n'est pas nécessairement lié au phénotype réel créé par ces marqueurs. Par exemple, l'élévation à des niveaux de cytosines plus méthylées augmentent les fréquences mutationnelles en raison des taux élevés de transition cytosine-thymine [He *et al.*, 2015]. De même, la méthylation de l'ADN peut à la fois interférer avec les gènes responsables des mécanismes de réparation de l'ADN [Meng *et al.*, 2015]. En tant que tels, ces phénomènes peuvent contribuer à des changements évolutifs au travers de leurs effets indirects sur la stabilité des gènes. Un second type d'effet indirect est lié à la création de phénotypes conduisant à une sélection naturelle atypique des génotypes dont l'exemple le plus connu est l'effet Baldwin, dans lequel la capacité d'un animal à apprendre de nouveaux comportements (dont nous savons qu'ils peuvent être hérités de manière épigénétique [Sznajder *et al.*, 2012]) peut affecter son succès reproducteur. Il y a aussi l'étude du cas classique de la ratte qui, en procurant des soins de type maternel à ses ratons, transmet ce trait à sa progéniture en déclenchant des modifications épigénétiques [Champagne, 2008].

Ainsi, se pose une autre question : qu'en est-il de l'héritage d'un *potentiel* épigénétique? En d'autres termes, certaines populations devraient porter codée dans leur génome une capacité plus ou moins importante d'hériter d'un phénotype par transfert épigénétique transgénérationnel, ou de créer de la variabilité épigénétique par épimutations [Kilvitis *et al.*, 2014].

De nombreuses études ont montré des différences dans le degré global de méthylation entre les populations de la même espèce microbienne [Schrey et Richards, 2012], les plantes [Alonso *et al.*, 2014] et les animaux (y compris les humains) [Oey *et al.*, 2015, Vogt, 2015]. Bien qu'il y ait indubitablement une composante stochastique à cette variation des marqueurs épigénétiques entre les populations, cette variation peut être à la fois une preuve et un mécanisme de la sélection naturelle agissant sur les phénotypes épigénétiques [Burggren, 2016].

1.2 Le modèle : le Killifish des mangroves (*Kryptolebias marmoratus*)

En 1961, Robert Harrington publia un rapport sur un poisson cryptique à longueur maximale de 60mm, aplati dorso-ventralement, de physionomie élancée, et observé pondant des œufs isolés. Dans celui-ci, il décrit *Rivulus marmoratus* (Poey 1880) : un petit cyprinodontiforme hermaphrodite ovipare à autofécondation interne. Les adultes contenaient des zygotes, et il observa qu'ils étaient oviposés à divers stades de développement. Quelques mois plus tard, les petits étaient eux-mêmes

féconds [Harrington, 1968]. Désormais connu sous le nom de *Kryptolebias marmoratus*, le killifish des mangroves est un organisme surprenant vivant là où on ne l'attend pas, dans des conditions difficiles et privilégiant l'autofécondation même s'il est capable de fécondation croisée.

1.2.1 Choix du modèle

Son mode de reproduction, *unique* chez les vertébrés, fait de *K. marmoratus* un modèle puissant pour les études évolutives et biomédicales, car de longues périodes d'autofécondation aboutissent à des génotypes naturellement homozygotes et des lignées isogéniques sans limitations significatives associées à la dépression de consanguinité. En d'autres termes, elles engendrent des lignées clonales neutralisant la majeure partie de la variabilité génétique, et mettant en évidence un autre type de variabilité dite *non génétique*. En effet, plus de 400 lignées isogéniques actuellement détenues dans les laboratoires à travers le monde et pourtant maintenues dans des conditions similaires de vie, montrent des variations physiologiques, comportementales et d'histoires de vie considérables.

La température influence le développement des mâles primaires ainsi que le changement de sexe entre les hermaphrodites et les mâles secondaires, ce qui rend cet organisme idéal pour l'étude de la détermination du sexe et de la plasticité sexuelle.

K. marmoratus présente également des adaptations remarquables à la vie dans des environnements extrêmes, et le système promet de faire la lumière sur l'évolution de la locomotion terrestre, la respiration amphibie et les larges tolérances à l'hypoxie, la salinité, la température et les polluants environnementaux.

Aussi son génome fut séquencé et permet l'étude de gènes et des familles géniques associés aux particularités décrites ci-dessus [Kelley *et al.*, 2016].

1.2.2 Phylogénie

Kryptolebias marmoratus fait partie d'un petit clade de killifish cyprinodontiformes de la famille des Rivulidés et de sous-ordre Aplocheiloidei [Costa, 2006] (Figure 8).

L'histoire taxonomique du genre *Kryptolebias* est très confuse, et sa cladistique n'est pas encore clairement établie, en cause, les erreurs d'identification, l'emploi de synonymes abusifs [Costa, 1994] et le manque d'études phylogénétiques des espèces proches (telles que *Rivulus heyei* Nicols 1914, *Rivulus bonairensis* Hoedeman 1958 ou *Rivulus garciai* de la Cruz Dubitsky 1976 qui sont suspectées de faire en réalité partie des *Kryptolebias* [Costa, 2006]). De plus, les révisions fréquentes de ce genre conduisent souvent à des réarrangements taxonomiques radicaux [Costa, 2011]. C'est pour cette raison que ce mémoire suit la classification taxonomique publiée par la version la plus récente du *Catalog of Fishes* [Eschmeyer *et al.*, 2018] où *K.* (= *Rivulus*) *marmoratus*, et *Kryptolebias caudomarginatus* (Seegers 1984) sont considérés comme des synonymes de *K. marmoratus*.

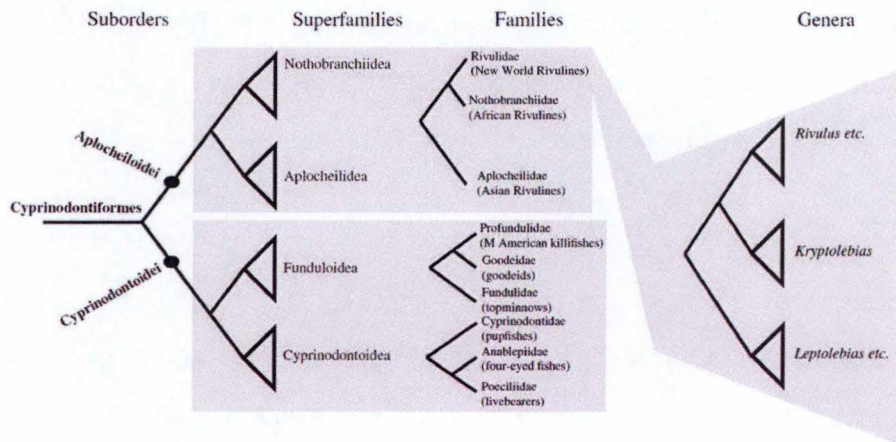


FIGURE 8 – Macro-phylogénie des *Cyprinodontiformes* représentant la disposition générale du genre *Kryptolebias*. *K. marmoratus* appartient à la famille Rivulidae, également connue sous le nom de New World Rivulines. [Avisé et Tatarenkov, 2015].

1.2.3 Distribution géographique, habitat et écologie

Leur aptitude à former une population à partir d'un seul individu, ainsi que leurs grandes capacités d'adaptation [Wright, 2012] procurent à ce poisson une distribution géographique étendue : *K. marmoratus* est retrouvé dans les mangroves et zones côtières d'eau saumâtre [Davis *et al.*, 1990] du bassin tropical et subtropical de l'Atlantique Est, depuis la Floride jusqu'au sud-est du Brésil en occupant tout le pourtour des Iles Caraïbes [Avisé et Tatarenkov, 2015].

Les populations de *Kryptolebias marmoratus* sont caractérisées par un grand nombre d'occurrences (sous-populations) et d'emplacements (tels que définis par l'UICN) dans leur habitat de prédilection [Turner *et al.*, 1992]).

Comme son nom l'indique, le killifish des mangroves est étroitement associé aux mangroves (et plus précisément, aux mangroves rouges, *Rhizophora mangle*) dans lesquelles on le retrouve dans toute une série de micro-habitats incongrus tels que les mares d'eau stagnante peu profondes et/ou temporaires, la litière de la mangrove, les rigoles inondées, à l'intérieur/sous les buches, dans des cannettes, dans des noix de coco, etc [Taylor, 2000]. Les terriers de crabes terrestres, en particulier ceux de *Cardisoma guanhumi* et *Ucides cordatus*, sont très appréciés par ces poissons [Costa, 2006]. Ces lieux parfois insolites ont pour points communs d'être des milieux de vie ichtyologique difficiles présentant des conditions physico-chimiques variables voire extrêmes (rythme d'immersion/émersion, températures élevées, variation importante de salinité et d'oxygène dissous, taux élevé de sulfate d'hydrogène et d'ammoniac).

Il a été démontré que ces poissons étaient capables de résister à des changements extrêmes de salinité, allant de 0 à 114 ppt (trois fois la salinité moyenne de l'eau de mer [King *et al.*, 1989]). Les eaux qu'ils affectionnent en font une exception chez les rivulidés, dont les autres espèces sont confinées aux habitats d'eau douce [Costa, 2006]). Il est à noter que *K. marmoratus* a également été signalé

dans les eaux douces, mais il y est considéré comme rare [Loftus et Kushlan, 1987]. De plus, bien que les rivulus vivent de préférence à 25°C, ils tolèrent des variations de 7 à 38°C dans l'eau [Taylor *et al.*, 1995]. Leurs tolérances et comportements face aux conditions environnementales difficiles et variables ont conduit certains scientifiques à classer *K. marmoratus* comme extrémophile [Wright, 2012].

Leur régime alimentaire est également vaste et est constitué d'insectes, de crustacés, de poissons juvéniles et de gastéropodes [Taylor, 1992].

Le killifish des mangroves a de remarquable qu'il est capable d'alterner sa respiration branchiale avec une respiration cutanée grâce à un réseau étendu de capillaires sous la peau et de nageoires permettant l'absorption d'oxygène hors de l'eau [Wright, 2012]. Capacité qu'il exploite lorsqu'il opère son comportement amphibie. Deux modèles d'émersion ont été répertoriés : lorsque les poissons étaient sous des débris, à l'intérieur de *galeries* ou dans la litière, et une émersion sur les côtés ou près de l'embouchure des terriers de crabes [Taylor, 2012].

Le comportement résultant du premier cas est probablement une réponse généralisée au stress environnemental, et se déroulerait lorsque les bassins temporaires ont séché, ou que les conditions de qualité de l'eau se sont fortement dégradées. Même si les stimulus amenant à ces comportements sont pour l'instant mal compris, la durée d'émersion a été corrélée aux concentrations en H₂S du milieu [Taylor, 1990] et une combinaison de fortes concentrations d'H₂S, de basses températures et d'hypoxie légère amènerait également à ce résultat [Regan *et al.*, 2011]. Dans ce cas, l'émersion est à plus long terme et est synonyme de stratégie d'évitement. Il a d'ailleurs été mis en évidence que *K. marmoratus* peut survivre jusqu'à 66 jours hors de l'eau [Taylor, 1990].

L'émersion dans le deuxième cas est à plus court terme, le poisson transitant dans et hors de l'eau plusieurs fois en une heure. Ce comportement reste incompris pour l'instant [Taylor, 2000].

1.2.4 Cycle de vie

K. marmoratus possède un temps de génération court, atteignant la maturité environ trois à quatre mois après l'éclosion et pouvant produire plus d'une dizaine d'œufs par jour pour les individus les plus fertiles [Moore *et al.*, 2012]. Il est l'un des deux seuls vertébrés androdioïques connus capables de s'autoféconder, l'autre étant une espèce étroitement apparentée, nommée *Kryptolebias hermaphroditus* (anciennement *Kryptolebias ocellatus* [Costa, 2011]). Autrement dit, leurs populations ne comprennent *que* des individus mâles et des individus hermaphrodites [Tatarenkov *et al.*, 2007]. Trois phénotypes sexuels sont observés dans ce système androdioïque : les mâles primaires, les hermaphrodites et les mâles secondaires.

Les mâles primaires, c'est-à-dire les individus unisexués aux gonades mâles, sont rares dans la plupart

des populations sauvages [Turner *et al.*, 2006]. Les mâles secondaires peuvent provenir d'hermaphrodites perdant leur fonction reproductrice femelle à un âge avancé de leur cycle de vie (après 3 à 4 ans), ou d'individus exposés à des températures plus élevées et à une photopériode raccourcie [Lee *et al.*, 2008].

Ces différents phénotypes sexuels s'accompagnent d'un dimorphisme de même type qui peut être vu comme une différence de coloration. Les mâles primaires sont d'une couleur orange, comme dans une moindre mesure pour les secondaires, et les hermaphrodites sont brunâtres avec une ocelle caudale (Figure 9). Cette différence n'est toutefois pas toujours nette et, pour distinguer avec précision les mâles des hermaphrodites, il est préférable de considérer conjointement la coloration et la présence/absence d'une ocelle caudale [Grageda *et al.*, 2004]. Aussi, des nuances sont observées selon le substrat sur lequel ils évoluent, lui procurant un excellent camouflage dans son écosystème de prédilection [Taylor, 2000].

La production de mâles est probablement influencée par des facteurs génétiques et environnementaux, mais les mécanismes régissant leur différenciation sexuelle sont encore à élucider [Turner *et al.*, 2006]. Tout comme ce qui concerne la production des œufs, même si des données récentes suggèrent qu'elle peut varier de façon saisonnière et est probablement influencée par la composition génétique des autres hermaphrodites à proximité [Lomax *et al.*, 2017].

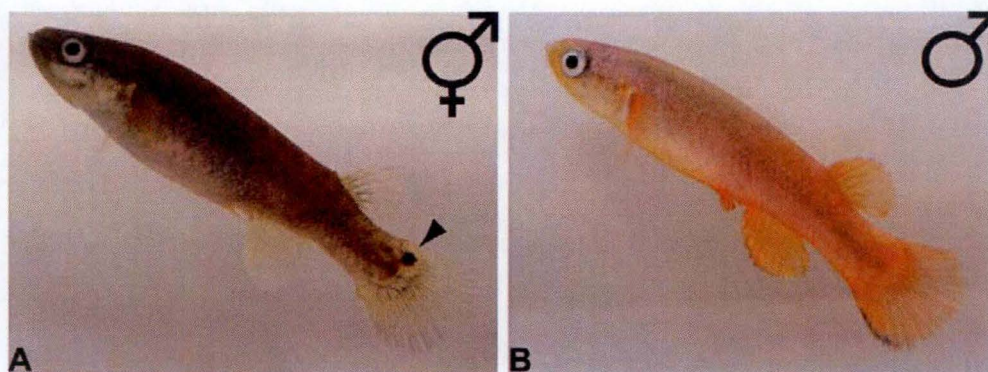


FIGURE 9 – *K. marmoratus* hermaphrodite (A) et mâle primaire (B). En A, la pointe de la flèche indique l'ocelle caudale [Mourabit, 2012].

Des études réalisées à l'aide de microsatellites ont révélé qu'il existe différents degrés d'auto-fécondation parmi les populations de *K. marmoratus*, mais aussi qu'on observe, globalement, un faible taux de fécondation croisée et, par conséquent, une faible hétérozygotie en milieu naturel [Lins *et al.*, 2018]. Toutefois, des croisements occasionnels entre mâles et hermaphrodites sont possibles et semblent se dérouler préférentiellement entre individus de lignées éloignées [Ellison *et al.*, 2018]. Bien que les mâles soient rares ($\pm 5\%$ des individus d'une population), il est des cas où le sexe-ratio peut être plus équilibré comme à Twin Cays au Belize ($>24\%$ de mâles; cas exceptionnel). Ces variations de proportion hermaphrodites/mâles engendrent des taux de croisement (et donc d'hété-

rozygotie) très variables entre les populations de *K. marmoratus* : certains sites affichent un taux de *seulement* 10% de fécondation croisée (Floride) mais il peut grimper jusque 60% dans d'autres, comme à Twin Cays au Belize [Mackiewicz *et al.*, 2006].

En absence de fécondation croisée, comme c'est le cas en laboratoire, l'autofécondation engendre des lignées homozygotes et isogéniques. En d'autres termes, tous les individus d'une même lignée partagent le même patrimoine génétique si l'on exclut les mutations *de novo* [Avisé et Tatarenkov, 2015].

1.2.5 Reprogrammation chez *K. marmoratus*

Le phénomène de reprogrammation du méthylome de *Kryptolebias marmoratus* a été caractérisé dans notre laboratoire en utilisant la technique LUMA. Les ovotestis des hermaphrodites semblent hypométhylés en comparaison aux testicules des individus mâles (respectivement, 79.6 et 87.2%) et rejoignent ce qui peut être observé chez le poisson zèbre et la souris, soit une hypométhylation de l'ovaire [Laing *et al.*, 2016] (Figure 10).

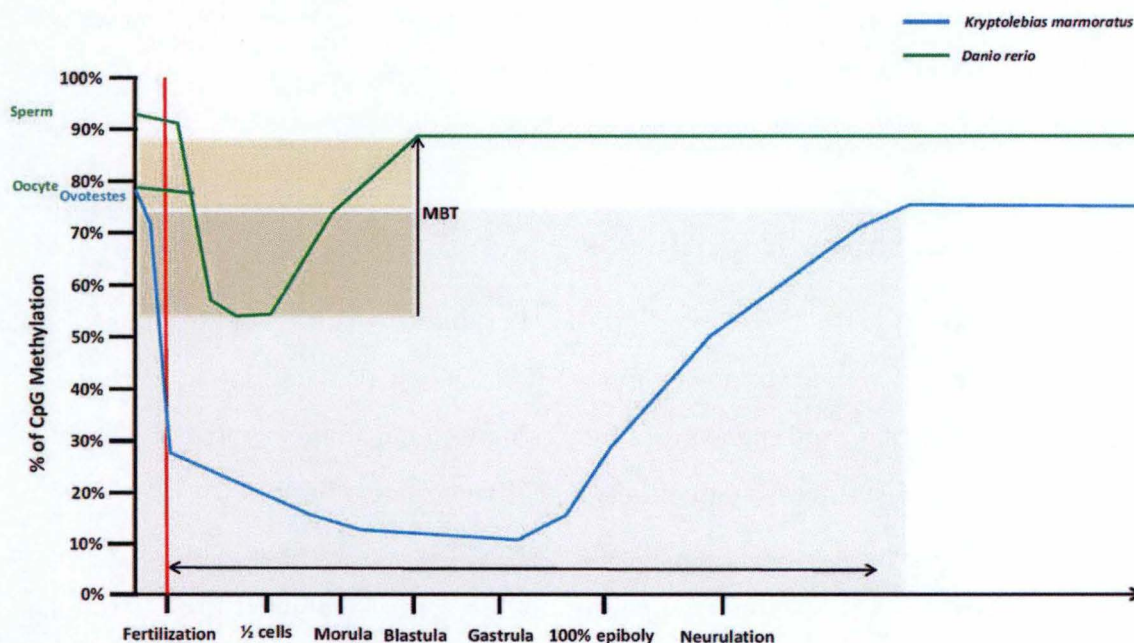


FIGURE 10 – Comparaison des dynamiques de méthylation de l'ADN au cours du développement embryonnaire de *Danio rerio* et de *Kryptolebias marmoratus*. (MBT, Mid Blastula Transition); adapté de [Fellous *et al.*, 2018].

L'étude de pools d'œufs fécondés à différents stades embryonnaires révèle une diminution immédiate de la méthylation pour chuter à 15,8% au stade gastrula, suivie d'un rétablissement progressif du niveau de méthylation jusqu'à atteindre un plateau stable de 70% à partir du stade 26 (correspondant à la formation du foie et précédant le stade 27 caractérisé par l'augmentation de la pigmentation de l'embryon et de ses premiers mouvements rapides). La période de reprogrammation semble aussi plus longue, plus considérable et plus tardive chez *K. marmoratus* que ce qui a été rapporté chez le poisson zèbre [Fellous *et al.*, 2018] (Figure 11).

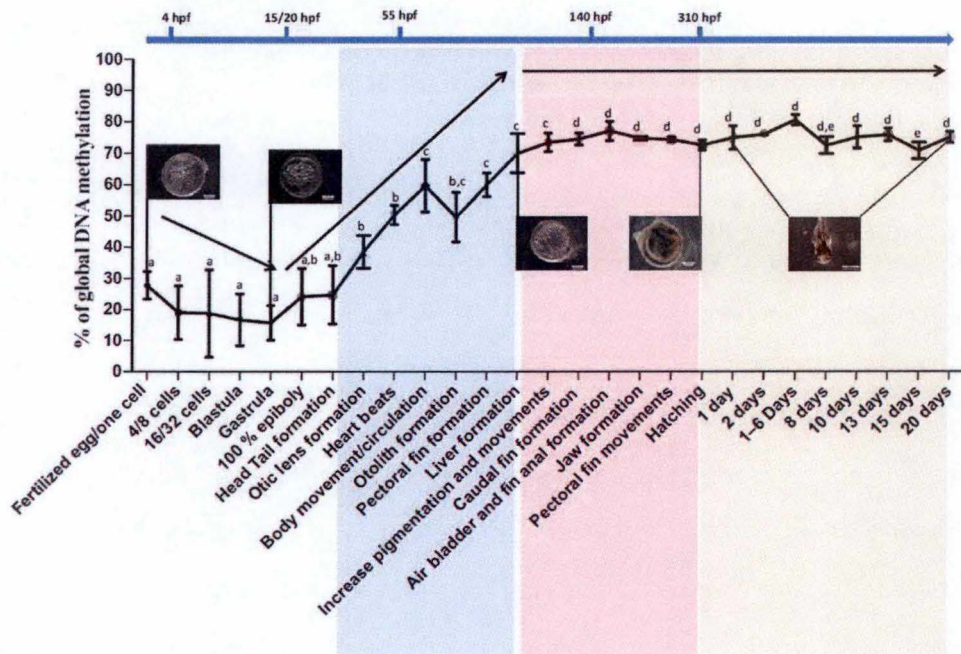


FIGURE 11 – Représentation des niveaux de méthylation de l'ADN au cours du développement embryonnaire de *Kryptolebias marmoratus*. Les stades développementaux sont indiqués sous l'axe des x. Les moyennes étiquetées avec la même lettre ne présentent pas de différences statistiquement significatives [Fellous et al., 2018].

1.3 Objectif général et hypothèses

L'objectif principal de cette thèse est d'étudier l'origine de la variabilité épigénétique chez le Killifish des mangroves (*Kryptolebias marmoratus*). Ce poisson présente la capacité unique chez les vertébrés de s'autoféconder, ce qui mène à des lignées hautement homozygotes et isogéniques, tandis que les fécondations croisées (rares) augmentent leur diversité génétique.

En milieu naturel, des souches isogéniques distinctes coexistent au même endroit. Cette caractéristique permet d'utiliser *K. marmoratus* comme un modèle précieux pour étudier le rôle de l'épigénétique dans l'émergence de la plasticité phénotypique, c'est-à-dire la capacité d'un génotype à produire différents phénotypes en réponse à l'environnement. Les mécanismes épigénétiques sont des systèmes puissants pour amener de la variabilité et donc, potentiellement, de la plasticité phénotypique. Les mécanismes épigénétiques pourraient ainsi permettre aux organismes de réagir plus rapidement aux stress environnementaux, et de les mener plus facilement à s'adapter et à évoluer.

Chez cette espèce, le rôle de l'épigénétique pourrait même être crucial, car il existe une diversité génétique presque nulle au sein d'une même lignée établie dans un environnement stressant pour la plupart des espèces aquatiques. De cette observation, se pose la question de savoir si le génome de ces poissons correspond au modèle du *general-purpose genotype*, pour lequel un génotype choisi est hautement capable de s'adapter à un large éventail de nouvelles situations environnementales. Chez cet organisme, les mécanismes épigénétiques sont les premiers candidats à expliquer comment un

génotype générique peut produire différents phénotypes.

Le fait que *K. marmoratus* présente un événement de reprogrammation au cours de l'embryogenèse plus tardif, plus long et plus profond que ce qui a été rapporté chez le poisson zèbre, pourrait avoir été choisi pour prolonger la durée de cette fenêtre de développement critique, au cours de laquelle des signaux environnementaux pourraient être intégrés dans le méthylome. Ainsi, notre hypothèse est que *K. marmoratus* ait sélectionné un mécanisme favorisant l'intégration des signaux environnementaux durant l'embryogenèse, pour produire des phénotypes adaptés aux conditions environnementales, plutôt que de miser sur la faculté d'hériter du méthylome d'une génération à l'autre. Afin d'explorer cette hypothèse, nous cherchons à étudier la source de la variabilité épigénétique dans une seule lignée homozygote et isogénique (DC4) en milieu contrôlé. Les origines supposées de la variabilité épigénétique sont : l'héritage des parents, l'induction par des facteurs environnementaux d'une variabilité épigénétique ciblée ou aléatoire et la stockasticité par des épimutations.

Les objectifs spécifiques du mémoire peuvent être définis comme suit :

- Déterminer les régions différentiellement méthylées potentiellement hérissables de génération en génération et celles reprogrammées. Pour cela, étudier la partie stable du méthylome (qui n'est pas effacée) et la partie variable (qui est effacée) à partir d'analyses RRBS comparant les méthylomes de différents stades développementaux de *K. marmoratus* entre eux. Identifier, classifier et étudier ces fragments RRBS. Tenter de mettre en évidence un profil fonctionnel à ces différents sets. Sélectionner quelques candidats CpG stables, et donc potentiellement transmissibles. Vérifier leur niveau de méthylation par pyroséquençage.
- Déterminer quels sont les effets de stress environnementaux sur les CpG candidats chez l'adulte et l'embryon. Pour cela, exposer des embryons pendant et/ou après la reprogrammation à différents stressseurs (PbTx-2, BMAA, EE2, métal, température, etc) et étudier les adultes utilisés dans de précédentes expériences directement après exposition.

2 Matériels et méthodes

2.1 Manipulations préliminaires des données RRBS

Le matériel de départ de ce mémoire est le résultat d'une analyse RRBS (pour *Reduced representation bisulfite sequencing*) réalisée antérieurement à ce travail au sein de notre laboratoire. Cette technique utilise une ou plusieurs enzymes de restriction pour fragmenter de manière spécifique l'ADN génomique. Traditionnellement, et, comme c'est le cas ici, MspI fut utilisé et a clivé les liaisons phosphodiester en amont des dinucléotides CpG. L'ADN fragmenté a ensuite été traité/converti au bisulfite, séquencé par Illumina et aligné bioinformatiquement. Il s'agit de la méthode de choix pour étudier les régions riches en CpG d'un génome [Gu *et al.*, 2010].

Cette technique fut employée afin de mettre en évidence le méthylome de quatre stades du développement embryonnaire de *K. marmoratus* : le stade 4 (pour quatre cellules), correspondant à un des premiers stades échantillonnables post-fécondation (niveau de méthylation moyen : $27.8\% \pm 4.4$) ; le stade Gast (pour gastrula), correspondant à un recouvrement progressif du vitellus par le blastoderme (le taux de méthylation moyen est à son plus faible : $15.8\% \pm 5.6$) ; le stade OL (pour *Otic Lens formation*), correspondant à la formation du système nerveux, otique et optique, et, durant lequel la méthylation *de novo* est en cours (niveau de méthylation moyen du génome : $38.5\% \pm 5.0$) ; le stade 27 (pour 27 cellules), correspondant à une augmentation de la pigmentation et à des mouvements plus rapides de l'embryon ; il possède le niveau stable de méthylation du génome après fertilisation soit $70.0\% \pm 6.0$.

Après un premier traitement passant par l'utilisation du programme *diffmeth* du DMAP package [Stockwell *et al.*, 2014] permettant de rendre exploitable ces données RRBS, le résultat de l'analyse se présente comme une suite de 48474 lignes représentant chacune un fragment de longueur comprise entre 40 et 220 paires de bases (médiane à 78) et contenant entre 1 et 29 CpG (médiane à 6). Pour chacun de ceux-ci, nous avons leur position précise dans le génome ; le nombre de CpG qu'il contient ; les valeurs moyenne de méthylation des trois réplicats des quatre stades cités précédemment ; le résultat d'une statistique F utilisée dans une ANOVA dont l'hypothèse nulle est que le niveau de méthylation est constant au sein de ces différents stades embryonnaires ainsi que la p-valeur associée.

Étudier la partie stable du méthylome (qui n'est pas effacée) et la partie variable (qui est effacée) lors du phénomène de reprogrammation a nécessité l'ajout de certaines données comme le niveau de méthylation moyen des réplicats des stades d'intérêts ou l'écart-type réalisé sur les moyennes de ces mêmes réplicats pour chaque fragment.

Tous les fragments analysés ont été annotés en fonction du gène codant le plus proche et classés selon leur localisation génomique par rapport au site d'initiation de la transcription (ou TSS pour *Transcriptional Start Site*) à l'aide du programme *identgeneloc* du DMAP package. Le qualificatif d'*internal* fut attribué aux fragments situés dans un gène ; *promoter* s'ils se trouvaient dans la région promotrice d'un gène (de 0 à 5000 pdb en amont du TSS) ; *upstream* s'ils se trouvaient entre 5000 à 10000 pdb en amont du TSS et *intergenic* s'ils se trouvaient au dessus de 10000 pdb du TSS [Stockwell *et al.*, 2014].

Le stade 4 ne contenant que très peu d'ADN et malgré les moyens mis en place pour le concentrer et obtenir un niveau d'ADN utilisable durant l'RRBS, très peu de fragments ont pu obtenir des valeurs de méthylation pour ce stade (seulement 44 possédaient deux réplicats valides). Approfondir

l'analyse des méthylomes des quatre stades en ne tenant compte que de ces 44 fragments n'a que peu d'intérêt. En conséquence, les ANOVA présentes dans le résultat RRBS ont été refaites en excluant le stade 4. De même, seuls les fragments possédant un minimum de deux réplicats pour les stades Gast, OL et 27 ont été pris en compte.

2.2 Méthylation gène-spécifique par pyroséquençage bisulfite

Dans le but d'étudier certains fragments particulièrement stables (cst) ou particulièrement variables (ncst) mis en évidence par le traitement des données RRBS, il était intéressant de valider leur comportement lors du reprogramming à l'aide d'une autre technique de séquençage permettant également de déterminer quantitativement la méthylation de ces fragments, le pyroséquençage bisulfite. Afin d'obtenir le profil de méthylation de tous les doublets CG d'une région donnée, l'ADN est tout d'abord traité au bisulfite, de manière à convertir les cytosines des CpG non méthylés en uracyle. Ensuite, la région d'intérêt est dans un premier temps amplifiée par PCR, puis passée au pyroséquenceur pour visualiser et quantifier la méthylation de chacune des cytosines de la région en question.

2.2.1 Identification des régions stables et variables différentiellement méthylées

La sélection des fragments considérés comme ncst est réalisée en imposant une série de filtres à l'ensemble des données RRBS. Successivement, sont réalisés dans ce but :

- l'ajustement des p-valeurs par la procédure Benjamini-Hochbergun en vue de réduire l'élévation du taux de faux positifs associée aux ANOVA réalisées sur de petits échantillons ;
- l'élimination de tous les fragments ne possédant pas trois réplicats pour les stades Gast, OL et 27 ;
- l'élimination de tous les fragments dont la p-valeur ajustée > 0.05 ;
- l'élimination de tous les fragments n'ayant pas une différence de méthylation d'au moins 15% entre le stade Gast et 27 ;
- le classement des fragments selon l'écart-type des moyennes de méthylation des trois stades.

La sélection des fragments considérés comme cst est réalisée en imposant une série de filtres à l'ensemble des données RRBS. Successivement, sont réalisés dans ce but :

- l'élimination de tous les fragments ne possédant pas trois réplicats pour les stades Gast, OL et 27 ;
- l'élimination de tous les fragments dont la p-valeur < 0.05 ;
- l'élimination de tous les fragments dont le niveau de méthylation moyen au cours du reprogramming est inférieur à 90% ;
- le classement des fragments selon l'écart-type des moyennes de méthylation des trois stades.

Parallèlement aux ANOVA, la normalité des résidus de chaque modèle a été testée et un test

non paramétrique de Kruskal-Wallis utilisé au cas où cette condition d'application n'aurait pas été respectée. Finalement, cette approche fut écartée car elle rendait la création de sous groupe délicate et leurs comparaisons difficiles.

Le test exact de Fisher's a été adopté pour vérifier si une différence significative de répartition génomique des fragments hypo/hyperméthylés existe entre les différents stades de développement durant le reprogramming. Le test non paramétrique de Wilcoxon-Mann-Whitney a, quant à lui, été mis en œuvre pour tester l'hypothèse selon laquelle les médianes du niveau de méthylation de chacun des stades sont proches.

2.2.2 Entretien général des rivulus

Trois différentes lignées isogéniques de *Kryptolebias marmoratus* ont été (et sont toujours) maintenues dans une salle possédant des conditions contrôlées : une photopériode de 12/12h et une température moyenne de la pièce de $27 \pm 1^\circ\text{C}$.

Dès l'éclosion, les rivulus sont élevés individuellement dans de petits bacs en plastique de 500ml remplis à moitié d'eau à 25 ± 1 ppt de salinité. Les poissons sont nourris quotidiennement et exclusivement *ad libitum* avec des nauplies d'*Artémia salina* vivants. L'eau des bacs est laissée stagnante pour simuler au mieux le milieu naturel des killifishes. Chaque récipient est également pourvu d'un îlot d'ouate situé à l'interface air/eau, d'une part pour enrichir l'environnement, d'autre part pour faciliter la ponte et la récolte des œufs.

2.2.3 Échantillonnage/Récolte des œufs

La récolte d'œufs des différentes lignées de *K. marmoratus* s'est déroulée durant un mois et demi à raison de deux à trois fois par semaine. Les œufs récoltés furent grossièrement classés dans une plaque de six puits contenant de l'eau ($25 \text{ ppt}/27^\circ\text{C}$) suivant la visibilité et l'état de l'embryon. Ensuite, les œufs ont été isolés à l'aide d'un binoculaire et séparés en fonction des stades recherchés (**Figure 12**). Ceux ne correspondant pas aux stades d'intérêt étaient stockés dans un puits approprié et re-analysés le lendemain ou immergés dans l'azote liquide et jeté.

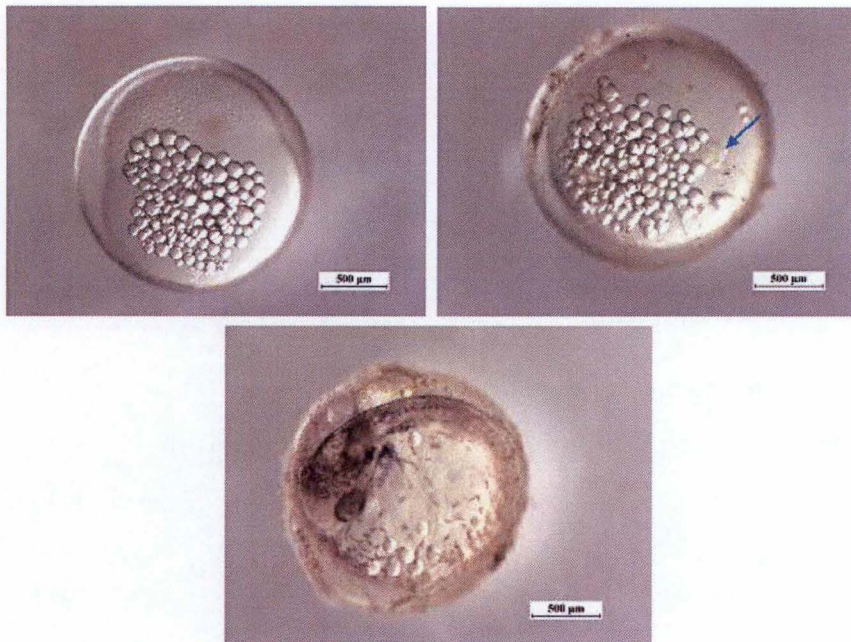


FIGURE 12 – *Vue du binoculaire des trois stades embryonnaires d'intérêt. À gauche, le stade gastrula; à droite, le stade OL; en bas, le stade 27. La flèche bleue représente l'encéphale de l'embryon avec la formation du système otique et optique.* ©Hugo Petry

2.2.4 Extraction d'ADN

L'extraction d'ADN est réalisée sur des pools de différentes quantités d'œufs selon le stade étudié. Des essais préliminaires ont permis de déterminer que des pools de 4 à 7 œufs de stade 27; 8 à 14 œufs de stade OL; et minimum 20 œufs de stades Gast permettaient d'obtenir les 35ng d'ADN par ml requis pour être visible sur un gel d'agarose 1%.

Après avoir lavé à l'eau distillée et séché au papier sans peluche par capillarité les pools d'œufs, ces derniers sont broyés grâce à des pistons adaptés (à usage unique). L'ADN est extrait à l'aide du kit Nucleospin Tissue XS (Macherey-Nagel). L'extraction commence par la lyse des échantillons en utilisant 160µl de tampon de lyse et 8µl de protéinase K. L'ensemble est incubé à 56°C durant quatre heures. L'ARN résiduel est enlevé via une ARNase et un second tampon de lyse. Une centrifugation permet de récupérer le surnageant contenant l'ADN; de l'éthanol y est ajouté, permettant à l'ADN d'adhérer à la colonne. L'échantillon est ensuite purifié par une série de rinçages/centrifugations et élué via un tampon d'éluion.

La concentration en acide nucléique, et donc, en ADN présente dans les échantillons est estimée à l'aide d'un spectrophotomètre NanoDrop. Il informe également quant à une possible contamination aux protéines via le rapport 260/280 et une possible contamination aux solvants via le rapport 260/230. L'absorbance des protéines étant de 280nm; de l'ADN, 260nm et des solvants, 230nm.

La qualité d'ADN et sa possible dégradation sont ensuite évaluées en faisant migrer ce dernier sur un gel d'agarose 1% à 100mV durant 30 minutes.

2.2.5 Traitement bisulfite

L'ADN extrait est ensuite traité au bisulfite via le kit DNA methylation-Gold kit (Zymo Research). Ce traitement convertit les cytosines non-méthylées en uracyle tandis que les cytosines méthylées restent inchangées en raison du groupement méthyle qui les protège de la réaction. Les uraciles seront ensuite reconnues en tant que thymine lors de la PCR (**Table 2.2.5**). 400ng d'ADN mis à volume par de l'eau distillée à 20µl sont utilisés. La conversion des cytosines non-méthylées se fait via l'ajout de 130µl de CT Reagent dans les tubes qui sont ensuite placés 10 minutes à 98°C puis 2h30 à 64°C. L'ADN traité au bisulfite est ensuite purifié via plusieurs rinçages/centrifugations et récolté via l'utilisation d'un tampon d'élution couplé à une dernière centrifugation.

Tableau 1 – *Exemple de séquences résultant du traitement/conversion au bisulfite et de l'amplification PCR bisulfite.*

	séquence originale	séquence bisulfitée	séquence amplifiée
ADN non méthylé	A-C-G-T-C-G-T-C-A	A-U-G-T-U-G-T-U-A	A-T-G-T-T-G-T-T-A
ADN méthylé	A-C-G-T-C-G-T-C-A	A-C-G-T-C-G-T-U-A	A-C-G-T-C-G-T-T-A

2.2.6 Amorces et amplification

Les fragments RRBS d'intérêt ont été rallongés de 100 pdb à leur extrémité 3' et 5' de manière à ajouter des portions de séquence sur lesquelles les amorces nécessaires à la PCR pourront se fixer sans qu'il y ait de perte d'information de la séquence d'intérêt.

Les amorces spécifiques à l'ADN traité au bisulfite ont été designées via le site MethPrimer [Li et Dahiya, 2002] pour les séquences le permettant (parmi les paramètres pris en compte, on retrouve : le choix de la séquence à amplifier dans le fragment allongé ; la taille de l'amplification désirée ; les *températures de Melting* des deux amorces ; la différence maximum admissible de différence entre ces deux températures). Leurs amplifications ont été testées virtuellement à l'aide du logiciel PyroMark Assay Design 2.0 (Qiagen). Celles considérées comme probantes furent commandées et testées une première fois par PCR, les amorces induisant une amplification admissible ont été commandées de nouveau avec cette fois une biotine (sur la reverse) et une troisième amorce dite *de séquençage* créée via PyroMark Assay Design 2.0 par paire d'amorces. Les primers finaux ont été ensuite contrôlés à l'aide d'une PCR adaptée à la nature de l'ADN traité au bisulfite (kit PyroMark PCR de Qiagen). Trois mix PCR furent testés : un contenant 1µl de MgCl₂ 25mM ; un autre contenant 1µl de Q-solution (un produit modifiant les températures de fusion moléculaire de l'ADN) ; et un dernier combinant 1µl de MgCl₂ 25mM et 1µl de Q-solution. Une étape d'activation de la Hot Star Taq polymérase à 95°C est nécessaire avant de commencer le premier cycle d'amplification. La température de dénaturation est de 94°C, d'annealing de 52.5°C et l'élongation se fait à 72°C. Un gel d'agarose 1%, 100 mV avec un temps de migration de 30 minutes permet de vérifier quel mix et

quelles amorces amplifient le mieux la séquence d'ADN bisulfite.

2.2.7 Pyroséquençage

L'analyse de la méthylation gène-spécifique à proprement parler se fait via l'utilisation du pyroséquenceur (Qiagen). La cartouche 11 (méthode 15) de l'appareil permettant d'injecter les nucléotides, ainsi que les mix *Substrat* et *Enzyme*, est remplie selon les volumes indiqués sur la feuille *Pre run information* fournie par le logiciel PyroMark Advanced Q-24 3.0 (Qiagen). Les produits PCR appariés à l'amorce biotinylée subissent différents traitements afin d'isoler le fragment d'ADN à séquencer. Ces derniers sont premièrement mis en présence de billes couvertes de streptavidine permettant à la biotine de s'y fixer. A l'aide de la *workstation*, les billes associées aux fragments d'ADN d'intérêt sont ensuite trempées dans de l'éthanol 70% et un tampon de dénaturation afin de séparer les deux brins d'ADN. Celui n'étant pas apparié avec l'amorce biotinylée est aspiré tandis que l'autre reste fixé aux billes. Ces dernières sont ensuite plongées dans un bain de purification et déposées dans la plaque contenant les amorces de séquençage. Cette plaque est chauffée cinq minutes à 80°C afin de permettre la liaison des amorces de séquençage aux produits PCR puis introduite dans le pyroséquenceur. Celui-ci incorpore les nucléotides en suivant le fichier de *Run* réalisé à l'aide du programme PyroMark Advanced Q-24. Finalement, les résultats sont analysés à l'aide de ce même programme et se présentent sous la forme de pyrogrammes.

2.3 Analyse d'enrichissement

Dans le but d'étudier les ensembles des fragments annotés particulièrement stables (cst) et variables (ncst) mis en évidence par le traitement des données RRBS, il était intéressant de récupérer un profil fonctionnel des ensembles de gènes associés à ces fragments, afin de mieux comprendre les processus biologiques sous-jacents à leur dynamique de (dé)méthylation lors du *reprogramming*. L'analyse d'enrichissement détermine quels termes GO (pour *Gene Ontology*) sont surreprésentés par rapport à l'ensemble des termes GO associés au génome de *K. marmoratus*. A cette fin, nous utilisons le package `g :Profiler` [Raudvere et al., 2019] installé sur R.

Les termes GO sont traditionnellement associés à trois bases de données ou *sub-ontology* : *Molecular Function*, *Cellular Component* et *Biological Process*. L'analyse fournie par `g :Profiler` permet aussi d'identifier un enrichissement dans la base de données KEGG (pour *Kyoto Encyclopedia of Genes and Genomes*) fournissant ainsi un autre type d'information pertinent à mettre en relation avec le métabolisme [Kanehisa et al., 2017].

2.3.1 Identification des régions stables et variables différentiellement méthylées

La sélection des régions considérées comme ncst et cst est réalisée en imposant deux séries de filtres à l'ensemble des fragments RRBS. Les critères de filtration vont différer de ceux utilisés pour l'identification des régions à destination du pyroéquencage, d'une part, car ces fragments ne sont pas destinés à être pyroséquencés, et que donc, une variabilité des niveaux de méthylation au sein d'un stade pour un fragment peut être plus facilement tolérée, nous avons donc limité le nombre de réplicats valides à minimum deux par stade), et, d'autre part, parce que trop peu de fragments appartenant au stade Gast possèdent au minimum ces deux réplicats (moins de 12000 sur les 48484 contre plus de 45000 pour les stades OL et 27). Prendre en considération le stade Gast dans l'analyse d'ontologie n'amène pas à l'identification de terme GO significativement enrichi. Pour contrer cela, seuls les stades OL et 27 ont été étudiés. De même, ne sont pris en compte que les fragments présents dans les régions promotrices ou dans les gènes. En effet, l'annotation des fragments upstream (5-10 kb) et intergéniques (> 10 kb) est fournie à titre indicatif et doit être considérée avec prudence car la méthylation de ces régions peut ne pas avoir d'influence sur les gènes associés.

Successivement, sont réalisés sur le set de départ en vue de mettre en évidence les fragments ncst :

- l'élimination de tous les fragments ne possédant pas au moins deux réplicats valides pour les stades OL et 27 ;
- la réalisation d'un test de Student dont l'hypothèse nulle est qu'il n'y a pas de différence entre les moyennes de méthylation des deux stades testés ;
- l'élimination de tous les fragments dont la p-valeur associée au t-test > 0.05 ;
- l'élimination de tous les fragments n'ayant pas une différence de méthylation d'au moins 25% entre les deux stades d'intérêt ;
- l'ajustement des p-valeurs associées au t-test par la procédure g :SCS ;

Successivement, sont réalisés sur le set de départ en vue de mettre en évidence les fragments cst :

- l'élimination de tous les fragments ne possédant au moins deux réplicats valides pour les stades OL et 27 ;
- la réalisation d'un test de Student dont l'hypothèse nulle est qu'il n'y a pas de différence entre les moyennes de méthylation des deux stades testés ;
- l'élimination de tous les fragments dont la p-valeur associée au t-test < 0.05 ;
- l'élimination de tous les fragments dont l'écart-type des moyennes de méthylation des trois stades est supérieur à 0.15 ;
- l'ajustement des p-valeurs associées au t-test par la procédure g :SCS ;

L'ajustement des p-valeurs est réalisé par la procédure g :SCS (pour *Set Counts and Sizes*) plutôt que par les procédures Benjamini-Hochbergun ou Bonferroni traditionnellement utilisées car ces dernières sont conçues pour plusieurs tests indépendants les uns des autres, or ce n'est pas le

cas dans une analyse GO dans laquelle les termes d'ontologie sont liés de manière hiérarchique entre eux. En d'autres termes, l'algorithme utilisé par g :SCS prend en compte la structure inégalement distribuée des ensembles de gènes annotés sur le plan fonctionnel, et donc leur *chevauchement*, quand les autres procédures ne le font pas [Reimand *et al.*, 2007].

3 Résultats

3.1 Variabilité et méthylation intra et inter-stades

Les 48474 fragments RRBS analysés correspondent à un total d'approximativement quatre millions de paires de base, 343717 CpGs et 10540 séquences de référence RefSeq (plusieurs fragments étant souvent associés à la même annotation). Le génome de *K. marmoratus* étant estimé à 915Mpdb [Kelley *et al.*, 2016], 0.43% de son génome est donc mis en évidence dans cette analyse RRBS (et ce, au moins pour les stades le permettant).

Parmi tous les fragments RRBS, la majorité est située dans le corps de gènes (46.0%), suivis des régions intergéniques (27.0%), des régions promotrices (19.7%) et des régions *upstream* (7.3%). Les niveaux de méthylation des CpG sont plus bas dans les fragments situés dans les régions promotrices (moyenne = 34.8 ± 42.2 , médiane = 6.03) par rapport aux régions *upstream* (moyenne = 62.0 ± 43.0 médiane = 93.2), intergéniques (moyenne = 70.0 ± 39.1 , médiane = 94.0) et *internal* (moyenne = 77.48 ± 35.4 , médiane = 95.6).

Les stades Gast, OL et 27 ne présentent pas de différence significative ($p > 0.05$, test de wilcoxon) au niveau de leurs moyennes de méthylation qui sont respectivement de $66.6\% \pm 44$ (médiane à 96%), $66.8\% \pm 43$ (médiane à 95%), $64.0\% \pm 41$ (médiane à 91%).

Sur les 48484 fragments, 9015 possèdent minimum deux réplicats valides pour les stades Gast, OL et 27, 189 sont considérés comme ncst et 7165 comme cst. Ces deux sets présentent une répartition analogue de leurs fragments à celle de la totalité des fragments RRBS (internal > intergenic > promoter > upstream). On retrouve significativement ($p < 0.001$, Fisher's exact test) plus de fragments méthylés dans les régions intergéniques des fragments ncst que dans les régions intergéniques des fragments cst. Différence contrebalancée par une diminution significative ($p < 0.001$, FET) des fragments méthylés dans les régions promotrices des fragments ncst comparées aux régions promotrices des fragments cst (Figure 13).

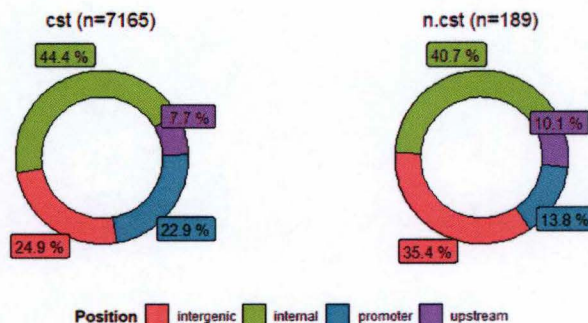


FIGURE 13 – Proportion des fragments ncst et cst selon leur position génomique. ncst = fragments dont minimum deux réplicats/stade valides ; p -valeur ajustée < 0.05 ; \bar{X} methyl.Gast - \bar{X} methyl.27 $> |0.15|$ et cst = fragments dont minimum deux réplicats/stades valides ; p -valeur > 0.05 ; $\sigma < 0.15$.

La densité des fragments cst selon leur niveau de méthylation représente une distribution bimodale, *identique* entre les trois stades d'intérêt. Si les fragments ncst sont isolés, une répartition bimodale est toujours observée avec une prédominance d'hyperméthylés. Profil qui s'atténue au stade OL et s'estompe complètement au stade 27 (Figure 14).

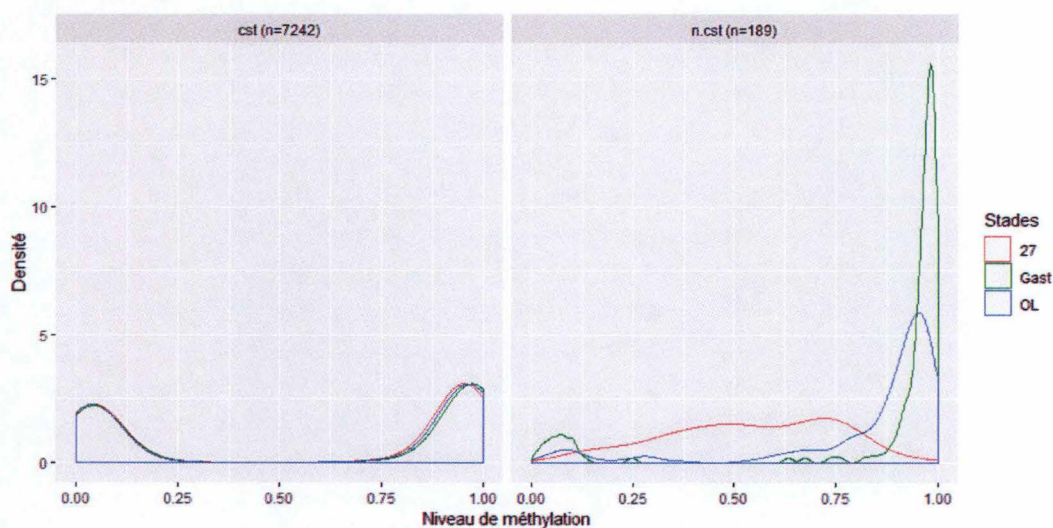


FIGURE 14 – Densité des fragments ncst et cst (à minimum deux répliquats/stade valides) selon le niveau de méthylation moyen des répliquats des stades *Gastrula*, OL et 27. Pour le graphique de gauche, tous les fragments considérés comme cst sont représentés (cst = p-valeur > 0.05).

La majorité des fragments hyperméthylés se trouve dans les gènes contrairement aux fragments hypométhylés qui se trouvent majoritairement dans les régions promotrices ($p < 0.0001$, FET). Aucune différence significative dans la répartition des fragments n'a été mise en évidence entre les stades Gast, OL et 27 au sein des conditions hypo et hyperméthylées (p-valeur > 0.05, test de Wilcoxon) (Figure 15).

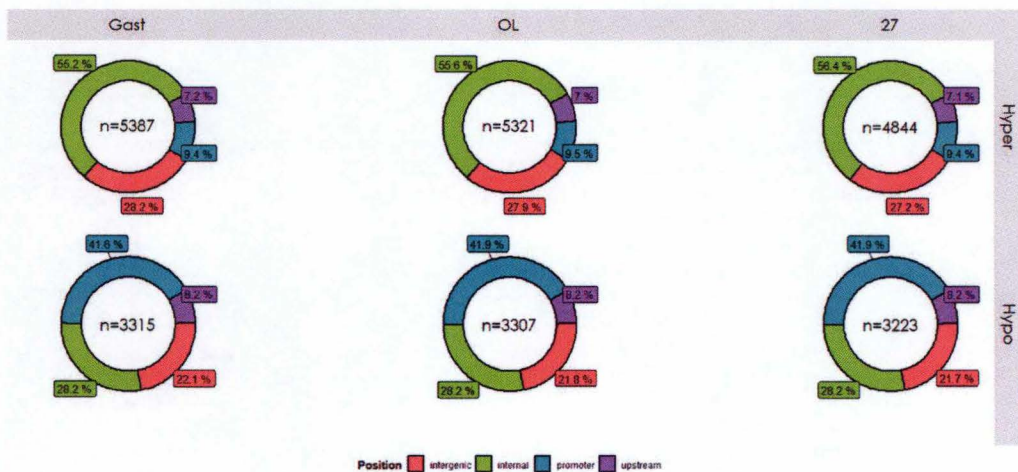


FIGURE 15 – Proportion des fragments hyper et hypométhylés des différents stades selon leur position génomique. hyper = fragments dont minimum deux répliquats/stade valides ; \bar{X} methylation > 0.85 et hypo = fragments dont minimum deux répliquats/stade valides ; \bar{X} methylation < 0.15.

3.2 Méthylation gène spécifique

3.2.1 Identification des régions stables et variables différentiellement méthylées

Les différents critères de sélection apposés aux 48484 fragments ont mis en évidence 400 fragments cst (Figure 16) et 24 fragments ncost (Table 2; Figure 17) potentiellement intéressants pour leur pyroséquençage.

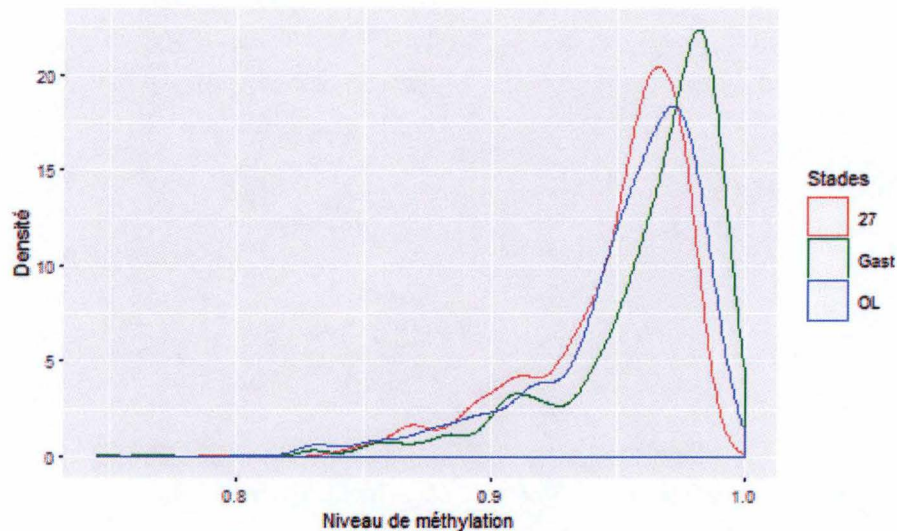


FIGURE 16 – Densité des 400 fragments cst (à minimum trois réplicats/stades valides) selon le niveau de méthylation moyen des réplicats des stades Gastrula, OL et 27.

Tableau 2 – Fragments considérés comme ncost entre les stades Gast, OL et 27. "Description" correspond à la protéine codée par le gène le plus proche du fragment en question.

GeneID	Lg.	CpG	Description	Position	Gast	OL	27	p.adj
XP_017292837.1	49	5	keratin, type I cytoskeletal 13-like	promoter	0.96	0.93	0.48	0.01
XP_017267115.1	57	9	enoyl-CoA hydratase domain-containing protein 3, mitochondrial	promoter	0.04	0.02	0.31	0.01
XP_017295260.1	61	4	polycomb group RING finger protein 2	promoter	0.97	0.97	0.70	0.02
XP_017259587.1	69	4	formin-like protein 2 isoform X2	promoter	0.99	0.95	0.69	0.01
XP_017269904.1	54	4	collagen alpha-1(XXV) chain isoform X9	internal	0.99	0.98	0.80	0.03
XP_017267236.1	78	10	matrilin-3-like isoform X1	internal	0.03	0.16	0.56	0.002
XP_017291541.1	73	6	protein MB21D2-like	internal	0.99	0.97	0.76	0.03
XP_017265633.1	56	4	SLAIN motif-containing protein 1 isoform X2	internal	0.99	0.99	0.83	0.01
XP_017275286.1	63	5	glypican-4-like	internal	0.97	0.77	0.20	0.01
XP_017277105.1	69	7	pleckstrin homology domain-containing family H member 1 isoform X1	upstream	0.13	0.53	0.91	0.02
XP_017272167.1	58	4	zinc finger CCCH domain-containing protein 7B isoform X2	upstream	0.95	0.87	0.39	0.01
XP_017284678.1	66	6	BRCA2-interacting transcriptional repressor EMSY isoform X2	upstream	0.98	0.63	0.23	0.01
XP_017275635.1	79	9	endothelin B receptor-like	upstream	0.89	0.66	0.07	0.01
XP_017290220.1	53	4	centrosomal protein of 55 kDa	upstream	0.94	0.95	0.66	0.02
XP_017269904.1	52	8	collagen alpha-1(XXV) chain isoform X9	intergenic	0.08	0.06	0.54	0.01
XP_017274649.1	98	4	olfactomedin-like protein 3B	intergenic	0.98	0.95	0.75	0.02
XP_017270117.1	62	4	methylcytosine dioxygenase TET3	intergenic	0.92	0.79	0.12	0.01
XP_017286230.1	62	5	cytochrome b-c1 complex subunit 6, mitochondrial-like	intergenic	0.11	0.30	0.54	0.02
XP_017287526.1	98	5	suppressor protein SRP40-like	intergenic	0.92	0.90	0.76	0.01
XP_017264678.1	99	4	cholesterol 24-hydroxylase-like	intergenic	0.95	0.96	0.79	0.02
XP_017261708.1	52	7	DNA damage-binding protein 2	intergenic	0.98	0.86	0.37	0.005
XP_017285013.1	49	6	disco-interacting protein 2 homolog C isoform X2	intergenic	0.96	0.92	0.51	0.04
XP_017272217.1	73	7	phosphoinositide 3-kinase regulatory subunit 5-like isoform X2	intergenic	0.94	0.95	0.68	0.04

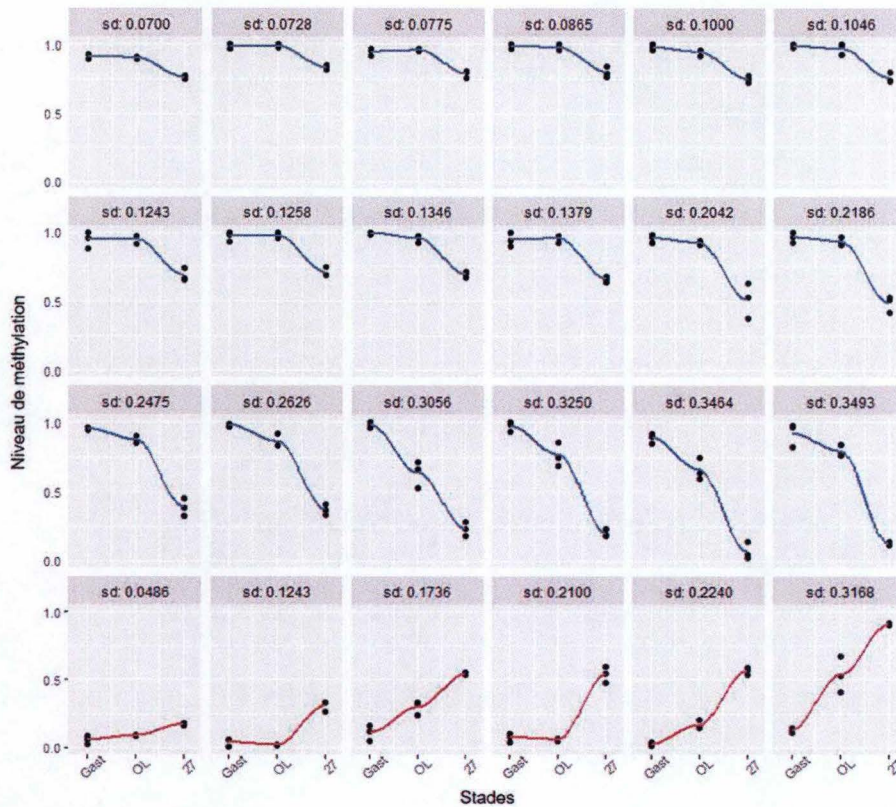


FIGURE 17 – *Fragments considérés comme ncst classés selon un écart-type croissant des moyennes de méthylation de chaque fragment des stades Gast, OL et 27. Les fragments dont les courbes sont en bleu présentent une diminution de leur niveau de méthylation au cours du reprogramming. En rouge, le contraire. Pour chaque fragment, les trois réplicats de chaque stade sont représentés.*

3.2.2 Extraction d'ADN

Les quantités d'ADN estimées par Nanodrop et extraites à partir des pools d'œufs varient entre 31.6ng/µl (stade Gast) et 106.1ng/µl (stade 27). Le gel d'agarose 1% sur lequel l'ADN a migré révèle que la qualité d'ADN est acceptable, il n'est pas dégradé et une seule bande est visible pour les stades d'intérêts. Le blanc ne *ressort* pas, indiquant qu'il n'y a pas eu de contamination externe importante (Annexe A).

3.2.3 Amorces

Sur les 24 fragments ncst bioinformatiquement allongés, huit d'entre eux se sont vus élaborer des primers par MethPrimer et seuls quatre de ces fragments ont vu leurs différentes paires amorces passer les critères de sélection du logiciel PyroMark Assay Design 2.0.

Sur les 400 fragments cst, les 24 plus stables au cours du reprogramming ont été étudiés (c'est-à-dire, ceux présentant les écart-types des moyennes de méthylation des stades Gast, OL et 27 les plus faibles), cinq ont franchi MethPrimer, et un seul de ces fragments a vu ses différentes paires amorces passer les critères de sélection du second logiciel (Table 3).

Trois paires d'amorces furent commandées pour chacun de ces cinq fragments. Elles furent testées

Tableau 3 – *Fragments sélectionnés en vue d'être analysés par pyroséquençage. La colonne "nom" correspond à leur identifiant, il est à mettre en relation avec leur catégorie (cst ou ncst).*

nom	GeneID	Lg.	CpG	Description	Position	Gast	OL	27
cst1	XP_017286220.1	76	12	growth arrest and DNA damage-inducible protein GADD45 alpha	intergenic	0.93	0.92	0.93
ncst1	XP_017261708.1	52	7	DNA damage-binding protein 2	intergenic	0.98	0.86	0.37
ncst2	XP_017272167.1	58	4	zinc finger CCCH domain-containing protein 7B isoform X2	upstream	0.95	0.87	0.39
ncst3	XP_017277105.1	69	7	pleckstrin homology domain-containing family H member 1 isoform X1	upstream	0.13	0.53	0.91
ncst4	XP_017284678.1	66	6	BRCA2-interacting transcriptional repressor EMSY isoform X2	upstream	0.98	0.63	0.23

par PCR simple (**Annexe B, Annexe C**). La meilleure paire de chaque fut commandée avec une biotine sur le brin *reverse* et son primer de séquençage (**Annexe E**).

3.2.4 Traitement bisulfite et amplification de l'ADN bisulfité

L'amplification de l'ADN traité au bisulfite a nécessité plusieurs mises au point. Le mix contenant 1µl de MgCl₂ a permis l'amplification optimale des fragments d'ADN traité au bisulfite pour quatre des cinq fragments testés. Le seul dont l'amplification était meilleure sans l'ajout de MgCl₂ (ncst2) est tout de même suffisamment amplifié dans cette dernière condition pour être utilisable (**Annexe C**). De manière, à ce que les cinq fragments soient amplifiés dans les mêmes conditions, le mix avec ajout de MgCl₂ pour le fragment ncst2 ainsi que les amorces associées ont été sélectionnées (**Annexe D**).

La PCR réalisée sur l'ADN traité au bisulfite montre une amplification pour tous les échantillons ainsi qu'un manque de spécificité de la paire d'amorces ncst2 et, dans une moindre mesure, de la paire d'amorces ncst4 (**Annexe F**).

3.2.5 Pyroséquençage

Les pyrogrammes fournis par le pyroséquenceur indiquent le pourcentage de méthylation de chaque cytosine méthylée au sein du fragment RRBS. Tous les pyrogrammes montrent un mauvais début de séquençage qui s'améliore considérablement lorsqu'est analysée la séquence donnée par le RRBS (**Figure 18**).

Les dynamiques de reprogrammation obtenus par pyroséquençage pour les cinq fragments (sur bases des trois stades d'intérêt) présentent des profils similaires à ceux obtenus par RRBS, excepté pour le fragment ncst4 (**Figure 19**). Les valeurs de méthylation de chaque CpG d'intérêt pour les différents réplicats et les différentes paires d'amorces se trouvent en **Annexe G**.

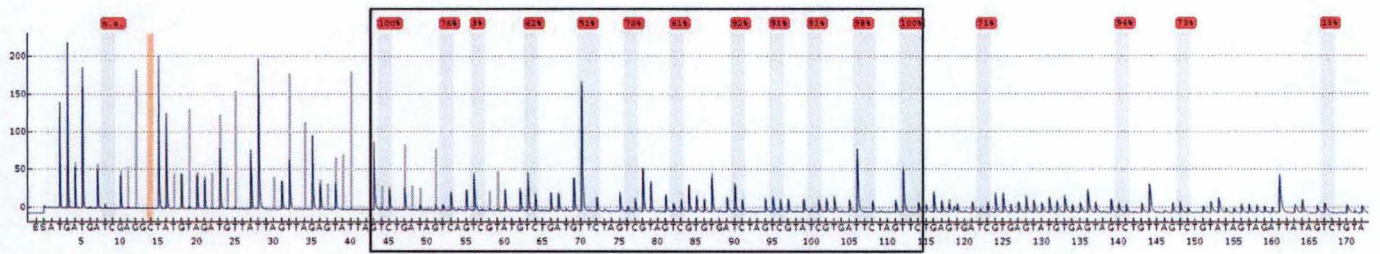


FIGURE 18 – *Pyrogramme de la séquence bisulfitee et amplifiée correspondant au fragments cst1. L'ADN en question provient d'un des trois répliquats des pools d'œufs de stade Gast. L'histogramme en gris et en arrière plan correspond aux pics attendus sur bases de la séquence fournie. Les pics en bleu foncé représentent l'intensité de lumière générée lorsque un ou plusieurs dNTP spécifiques (dATP, dTTP, dGTP ou dCTP) ont été incorporés. Si plusieurs mêmes dNTP ont été incorporés, la taille du pic sera "proportionnelle" aux nombres de nucléotides incorporés. Le pic du substrat (S) procure une échelle aux autres pics. Le contrôle bisulfite (en rose) permet de vérifier que le traitement bisulfite s'est bien déroulé. Les nucléotides pour lequel aucun pic n'est attendu sont des contrôles négatifs. Les bandes verticales bleues claires représentent la position de chaque cytosine méthylée, les pourcentages indiquent leur niveau de méthylation et leur couleur, la qualité du run (bleu : OK, jaune : moyen, rouge : KO). Le cadre noir situe la séquence du fragments RRBS.*

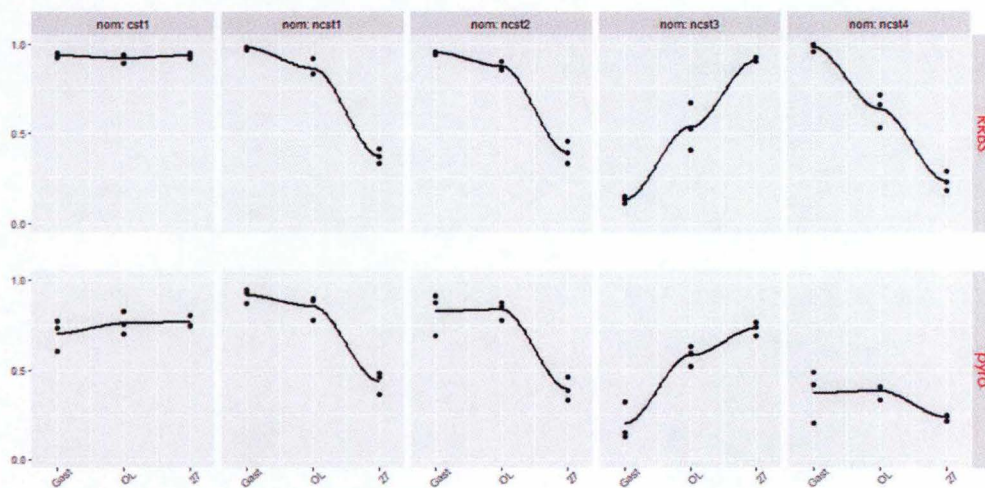


FIGURE 19 – *Comparaison des profils de reprogrammation basés sur l'analyse des stades Gast, OL et 27, des fragments RRBS et de ces mêmes fragments étudiés par pyroséquençage. Pour chaque graphique, les trois répliquats de chaque stades sont représentés.*

3.3 Analyse d'enrichissement

Une analyse d'enrichissement a été réalisée avec le package `g:Profiler` depuis R en utilisant les sets de gènes associés aux régions promotrices et géniques des fragments considérés comme `ncst` et `cst`. Pour le set des `ncst`, 481 fragments ont été mis en évidence et 26 termes GO ont été significativement enrichis (Table 4) (Figure 20). Pour le set des `cst`, 322 fragments ont été mis en évidence et 12 termes GO ont été significativement enrichis (Table 5) (Figure 21).

Dans les deux analyses, la grande majorité des termes GO mis en évidence concernent les processus biologiques et sont liés à la réponse aux stimuli externes, à la communication cellulaire et

au développement. Trois fonctions moléculaires et une information provenant de la base de données KEGG ressortent également.

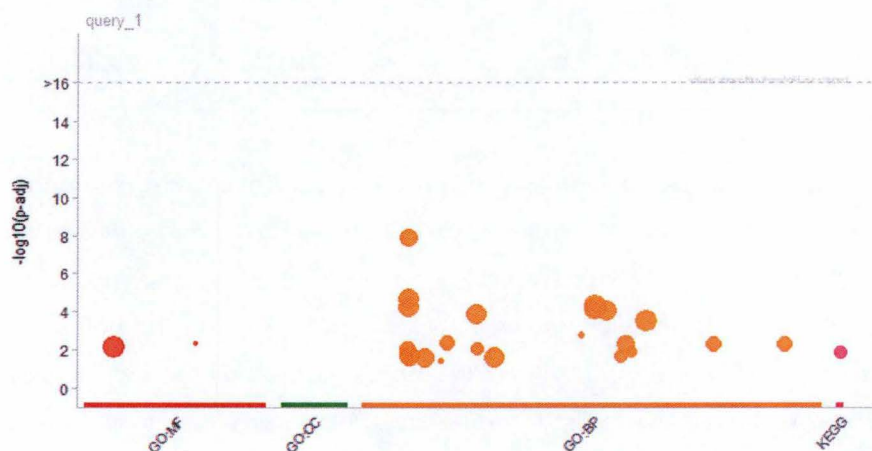


FIGURE 20 – *Graphique Manhattan résultant de l'analyse d'ontologie des fragments ncst.* L'axe des X montre les 26 termes fonctionnels regroupés par source de données et la p-valeur (en $-\log_{10}$) associée à leur enrichissement sur l'axe des Y. Chaque source de données possède sa couleur (rouge pour "Molecular Function", vert pour "Cellular Component" et jaune pour "Biological Process"). Chaque cercle représente un terme GO. La taille d'un cercle dépend du nombre de gènes annotés pour ce terme.

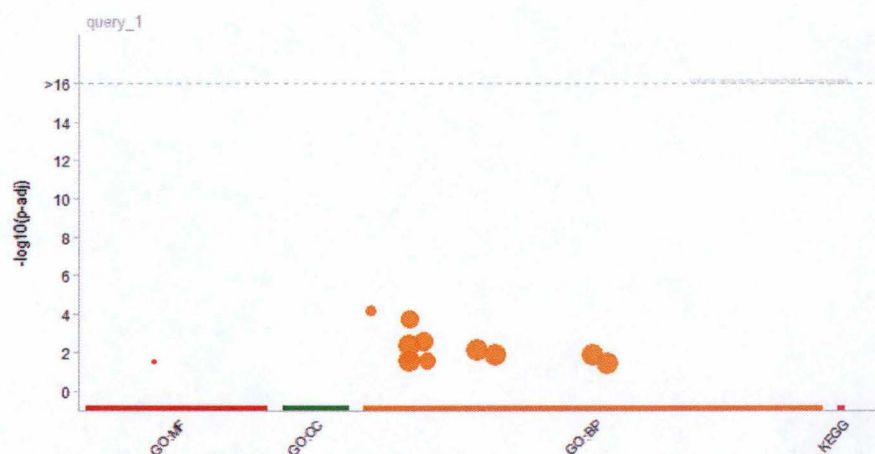


FIGURE 21 – *Graphique Manhattan résultant de l'analyse d'ontologie des fragments cst.* L'axe des X montre les 12 termes fonctionnels regroupés par source de données et la p-valeur (en $-\log_{10}$) associée à leur enrichissement sur l'axe des Y. Chaque source de données possède sa couleur (rouge pour "Molecular Function", vert pour "Cellular Component" et jaune pour "Biological Process"). Chaque cercle représente un terme GO. La taille d'un cercle dépend du nombre de gènes annotés pour ce terme.

Tableau 4 – *Description, p-valeur ajustée et type d'enrichissement des termes GO mis en évidence par l'analyse g :Profiler des fragments ncst.*

Description	p.adj.	type
cell surface receptor signaling pathway	1.28x10 ⁻⁸	GO :BP
signal transduction	0.00002	GO :BP
regulation of biological process	0.00005	GO :BP
cell communication	0.0001	GO :BP
response to stimulus	0.0001	GO :BP
regulation of cellular process	0.0001	GO :BP
cellular response to stimulus	0.0001	GO :BP
signaling	0.0001	GO :BP
biological regulation	0.0003	GO :BP
ephrin receptor signaling pathway	0.002	GO :BP
Wnt signaling pathway	0.004	GO :BP
cell surface receptor signaling pathway involved in cell-cell signaling	0.005	GO :BP
cell-cell signaling by wnt	0.005	GO :BP
epithelium development	0.005	GO :BP
regulation of Wnt signaling pathway	0.009	GO :BP
enzyme linked receptor protein signaling pathway	0.010	GO :BP
regulation of canonical Wnt signaling pathway	0.014	GO :BP
canonical Wnt signaling pathway	0.020	GO :BP
multicellular organism development	0.020	GO :BP
multicellular organismal process	0.026	GO :BP
tissue development	0.026	GO :BP
establishment or maintenance of cell polarity	0.035	GO :BP
response to activity	0.037	GO :BP
ephrin receptor binding	0.005	GO :MF
protein binding	0.008	GO :MF
Hedgehog signaling pathway	0.013	KEGG

Tableau 5 – *Description, p-valeur ajustée et type d'enrichissement des termes GO mis en évidence par l'analyse g :Profiler des fragments cst.*

Description	p.adj	type
ameboidal-type cell migration	0.0001	GO :BP
cell surface receptor signaling pathway	0.0002	GO :BP
anatomical structure morphogenesis	0.003	GO :BP
signal transduction	0.004	GO :BP
signaling	0.008	GO :BP
developmental process	0.013	GO :BP
anatomical structure development	0.014	GO :BP
enzyme linked receptor protein signaling pathway	0.018	GO :BP
regulation of signal transduction	0.026	GO :BP
cell communication	0.026	GO :BP
cellular response to stimulus	0.037	GO :BP
transmembrane receptor protein kinase activity	0.034	GO :MF

4 Discussion

Le *paysage* du méthylome est structuré de manière dynamique au cours du développement et l'on sait que des variations dans les profils de méthylation distinguent les cellules saines des cellules malades. Cependant, on ignore si les profils de méthylation sont conservés d'une espèce à l'autre ni comment ces profils de méthylation s'établissent durant l'embryogenèse. Ce dernier point pourrait être critique pour *K. marmoratus* chez qui les mécanismes épigénétiques sont les premiers candidats à expliquer comment un génotype pourrait produire différents phénotypes au sein d'une même lignée établie dans un environnement variable à partir d'une diversité génétique presque nulle.

Nous nous basons ici sur l'analyse RRBS du méthylome de différents stades embryonnaires chez le killifish des mangroves afin d'identifier et d'analyser de plusieurs façons les régions constantes, non constantes, hypométhylées et hyperméthylées lors du phénomène de reprogrammation post-fécondation afin d'en déterminer un peu plus sur cette période cruciale du développement embryonnaire.

4.1 Variabilité et méthylation intra et inter-stades

L'analyse de tous les CpG RRBS a montré une distribution bimodale de leur méthylation, avec près de 45% des CpG classés comme hyperméthylés et 33% classés comme hypométhylés comme il peut l'être observé chez *D. rerio*. Ce type de distribution est aussi retrouvé chez la souris et l'humain mais avec la tendance inverse à avoir une majorité d'hypométhylés et une minorité d'hyperméthylés [Potok *et al.*, 2013, Feng *et al.*, 2010]. Cette observation fut attribuée à une prévalence des régions intergéniques chez le zebrafish qui concorde avec un plus grand nombre d'éléments transposables (ET) observables (typiquement hyperméthylés) dans son génome (51%) comparé à celui de la souris (33%) [Zhang *et al.*, 2016]. Cependant, il doit en être autrement chez *K. marmoratus* puisque seul 27% de son génome est constitué d'ET [Rhee *et al.*, 2017].

0.43% du génome de *K. marmoratus* est présenté dans cette étude, cette valeur approximée est inférieure à celle retrouvée dans une autre analyse RRBS également réalisée sur cet organisme dans notre laboratoire (données non publiées). Cette analyse basée sur des échantillons de foie d'adultes comportait 114000 fragments de tailles similaires à notre étude, représentant ainsi 0.97% de son génome. Cette différence pourrait provenir des faibles quantités d'ADN présentes dans les œufs résultant en une profondeur et un recouvrement moindre des *reads* lors de l'*alignement avec référence* et donc, *in fine*, en une représentation plus faible de son génome.

Pour déterminer si le *paysage* des CpG RRBS différait entre les stades Gast, OL et 27, leur distribution génomique par rapport au gène codant le plus proche fut analysée et comparée. La majorité

des CpG se trouvaient dans le corps des gènes (46.0%) puis dans les régions intergéniques (27.0%) contre 19.7% dans les promoteurs et 7.3% dans les régions *upstream*. Ces pourcentages sont cohérents avec l'enrichissement typique des CpG qui caractérise les données RRBS [Lea *et al.*, 2017]. Cela reflète aussi la distribution des CpG hyperméthylés dans le génome de *K. marmoratus* (internal > intergenic > promoter > upstream), ce qui n'est pas très étonnant dans la mesure où ces derniers représentaient la majorité (58%) des CpG analysés.

Les régions promotrices semblaient exclure les CpG hyperméthylés, comme il peut l'être également observé chez les mammifères. Ce dernier point concorde avec l'étude de Zhang *et al.*, 2016 qui a comparé le rapport des fragments d'hyperméthylés et hypométhylés entre différents organes du poisson zèbre et de la souris et qui a également observé cette différence de répartition. Ces observations rejoignent le consensus selon lequel les îlots CpG ne sont généralement pas méthylés au niveau des promoteurs de manière à permettre la liaison de facteurs de transcription tout en empêchant la machinerie de méthylation de l'ADN de cibler ces régions [Messerschmidt *et al.*, 2014].

Comme attendu, les fragments *cst* exhibaient un profil de méthylation *identique* entre les différents stades étudiés, contrairement aux fragments *ncst* dont le stade Gast suivait une distribution bimodale s'atténuant au stade OL et s'estompant complètement au stade 27. Il serait intéressant d'avoir des données transcriptomiques (au moins) des stades Gast, OL et 27 de manière à déterminer si une corrélation existe entre cette baisse de méthylation chez les *ncst* et l'expression génique des gènes associés durant le *reprogramming*. Ce sujet est un peu plus approfondie dans le chapitre suivant où l'analyse d'enrichissement des fragments *ncst* est discutée.

Il peut paraître surprenant qu'une majorité des fragments RRBS semblait avoir un niveau de méthylation constant durant le *reprogramming* (7165 *cst* contre 189 *ncst*) et, de même, il peut paraître surprenant que le niveau de méthylation moyen des fragments RRBS des trois stades restait stable et équivalait au niveau de méthylation moyen des CpG RRBS du poisson à l'âge adulte : 66.0% mesuré à partir du foie et 66.1% mesuré à partir de cerveaux (données RRBS non publiées) alors que les données LUMA (Figure 22) indiquent que la majorité des gènes du point de vue de leur méthylation sont *ncst*. Cependant, il ne faut pas oublier que lorsque le RRBS cible les zones denses en CpG (CpGi) et ne tient compte que de 10 à 15% maximum de l'ensemble des CpG dans le génome humain [Harris *et al.*, 2010] ou dans celui de *D. rerio* [Zhang *et al.*, 2016], le LUMA mesure le niveau de méthylation moyen de l'ensemble d'un méthylome. Ceci explique aussi pourquoi les niveaux de méthylation génomique moyen mesurés à l'aide du RRBS sont sous-estimés par rapport aux mesures LUMA (niveau de méthylation génomique moyen de foie : 81.67% [Fellous *et al.*, 2018]). De ce fait, les résultats de ces deux techniques ne doivent pas être directement comparés. En effet, si la technique RRBS est une méthode relativement peu onéreuse comparée au *gold standard* de l'analyse du méthylome, le *Whole Genome Bisulfite Sequencing*, il faut bien garder à l'esprit que le *Reduced Representa-*

tion Bisulfite Sequencing a une couverture limitée des régions intergéniques et des régions entourant les CpGi. Régions, cependant, importantes d'un point de vue fonctionnel [Stadler *et al.*, 2011] (d'autres limitations à l'RRBS peuvent être retrouvées dans l'article de [Paun *et al.*, 2019]). Pour palier au manque de recouvrement de l'RRBS, l'ERRBS (pour *Enhanced-RRBS*) a été développé et est une solution permettant de cibler les îlots CpG comme les régions plus pauvres en CpG tout en étant pertinente d'un point de vue pécurier [Garrett *et al.*, 2015].

La prévalence des fragments est une question. Si les CpGi subissent le phénomène de reprogrammation post-fécondation, alors la remise à niveau des niveaux de méthylation de ces CpG se déroule plus tôt que celle de la grande majorité du génome et si elles ne le subissent pas, est-ce qu'une modification de ces régions serait transmissible aux générations futures ? (se référer à l'introduction en ce qui concerne les implications évolutives liés à l'épigénétique). Le second objectif de ce mémoire aurait, en partie, pu répondre à cette question en étudiant l'effet de stress sur certains CpG candidats chez les adultes et leur descendance, mais par manque de temps, cela n'a pas pu se faire et doit être considéré comme une perspective intéressante à étudier en prolongement de ce travail. Dans tous les cas, il est aujourd'hui acquis que le méthylome d'origine paternel du poisson zèbre est hérité de manière stable, sans changements importants au cours du développement précoce alors que le méthylome maternel subit un remodelage complet et se calque sur l'état épigénétique du génome paternel [Hackett et Surani, 2013] contrairement aux mammifères chez qui une déméthylation *complète* de l'ADN du sperme et de l'ovocyte est observée [Smith *et al.*, 2012]. Ainsi, une nouvelle question se pose : est-ce que l'état épigénétique des fragments est provient des gamètes mâles, des gamètes femelles ou est indépendant de ces derniers ? La réponse à cette question pourrait nous éclairer sur la capacité de *K. marmoratus* à générer des génotypes homozygotes et des lignées isogéniques sans limitations significatives associées à la dépression de consanguinité. En effet, outre les raisons génétiques conduisant aux effets délétères de la consanguinité, il a été démontré que la méthylation de l'ADN intervient dans ce phénomène [Vergeer *et al.*, 2012] et même s'il n'existe aucune explication mécanistique claire entre la variation épigénétique et la dépression de consanguinité, le fait que les phénomènes de consanguinité semblent moindre chez les poissons comparés aux mammifères [Brown *et al.*, 2012, Shinya et Sakai, 2011] pousse à investiguer du côté de l'implication de la transmission de l'épigénome paternel chez la descendance de notre killifish. Cela implique un autre challenge qui consiste à isoler les gamètes mâles et femelles de l'ovotestis de *K. marmoratus*, ce qui n'a encore jamais été réalisé de manière probante pour l'instant.

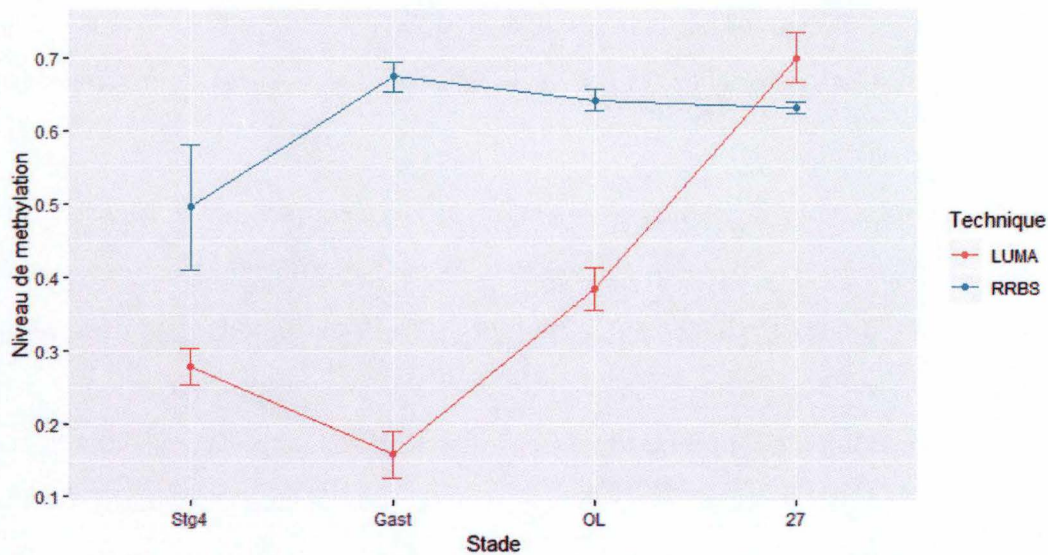


FIGURE 22 – Niveau de méthylation moyen des CpG analysés par LUMA versus ceux analysés par RRBS (à minimum un réplicat/stade valide) pour les stades *Gastrula*, *OL* et *27*. Données LUMA provenant de [Fellous et al., 2018].

4.2 Analyse méthylation gène spécifique

L'étude de la méthylation gène-spécifique de certains fragments *cst* et *ncst* visait à valider leurs résultats RRBS en vue de les analyser ultérieurement avec des échantillons provenant d'adultes et d'embryons soumis à un stress.

Les filtres apposés en vue de déterminer les fragments *cst* et *ncst* ont permis de mettre en évidence 400 fragments *cst* et 24 fragments *ncst* possédant trois réplicats valides pour les stades *Gast*, *OL* et *27* (c'est à dire, respectivement, 41.0% et 2.46% des fragments répondant à cette dernière condition). Sur les 24 fragments *ncst*, seuls quatre se sont vu attribuer des amorces fonctionnelles, de même sur les 24 fragments considérés comme les plus *cst*, seule, une paire permettait virtuellement la PCR. Et pourtant, même avec ces paramètres exigeants, deux paires d'amorces (*ncst2* et *ncst4*) sur les cinq testées ont montré un manque de spécificité visible sur les gels de migration (qui n'avait pas été décelable lors des tests antérieurs).

La comparaison des profils de méthylation des fragments étudiés à l'aide du RRBS avec ceux étudiés par pyroséquençage a révélé des profils similaires excepté pour le fragment *ncst4*. Différence probablement imputable au manque de spécificité de ces amorces.

Les résultats du pyroséquençeur montrent une bonne répétabilité au niveau des réplicats techniques alors que de nombreuses sources d'erreurs entrent en jeu dans l'analyse. La plus importante étant probablement la variabilité des stades développementaux des œufs constituant un pool (trois pools sont utilisés pour chaque stade). Le stade *Gast* demande une vingtaine d'œufs afin d'obtenir, en fin d'extraction, une concentration en ADN suffisante pour les étapes suivantes, et ce, combiné au

fait qu'il s'agisse d'un stade se confondant quelque peu avec certains stades plus tardifs, explique probablement pourquoi c'est aussi chez lui que la variabilité des mesures de méthylation est la plus élevée.

Sur base de ces résultats, la procédure permettant de sélectionner des gènes *cst/ncst* issus d'une analyse RRBS et de les tester par pyroséquençage semble être au point. Cependant, nous n'avons pas eu le luxe de choisir les fragments testés sur base de critères fonctionnels. Par exemple, le fragment associé à la *methylcytosine dioxygenase TET3* qui a été isolé parmi les 24 fragments *ncst* aurait été un candidat idéal d'un point de vue fonctionnel en raison de son implication dans le contrôle de la méthylation de l'ADN via l'hydroxylation des 5-mC, de son rôle clé dans la régulation génique et le développement embryonnaire mais aussi dans bien d'autres processus [Rasmussen et Helin, 2016]. Qui plus est, la méthylation de ce fragment RRBS diminue entre les stades Gast (92%), OL (79%) et 27 (12%) quand une augmentation progressive de l'expression de l'ARNm de la TET3v1 (une des deux isoformes de TET3 chez *K. marmoratus*) est constatée au même moment dans le cerveau des hermaphrodites. Cette donnée combinée au rôle supposément conservé des protéines TET dans la neurogénèse des rivulus, suggère une fonction spécifique de la TET3v1 dans le cerveau des hermaphrodites [Fellous *et al.*, 2018].

Dans l'optique d'augmenter les chances de tester des gènes sélectionnés aussi sur bases de critères fonctionnels, il sera nécessaire de se satisfaire de deux réplicats valides au lieu de trois par stade et d'être plus restrictif sur d'autres critères, de manière à augmenter les probabilités de définir des paires d'amorces valides pour ceux-ci. Parmi ces critères, on pourrait retrouver l'élimination des fragments dont les régions adjacentes contiendraient des régions répétées, l'élimination des fragments RRBS trop longs (> 90 pdb), qui, une fois allongés, ne permettent pas d'être analysés par le pyroséquençeur de manière optimale ou encore, l'élimination des fragments associés aux régions *upstream* et intergéniques qui sont peut-être moins pertinentes d'un point de vue fonctionnel.

La raison des mauvaises qualités de *run* affichées par les pyrogrammes (la couleur rouge des labels au dessus des CpG) proviennent des mauvais débuts de séquençages. Il est probable qu'ils soient dus à l'assemblage relativement incomplet du génome de *K. marmoratus* [Kelley *et al.*, 2016] sur lequel se base, entre autres, l'élongation des séquences RRBS et, au fait que les cinq fragments testés appartiennent à des régions *upstream* et intergéniques. Le nombre important de séquences répétées en amont des séquences d'intérêt sont peut être aussi à mettre en cause.

4.3 Analyse d'enrichissement

Sur les 481 fragments ncst et les 322 fragments cst identifiés, seuls les fragments ncst sont associés à des gènes directement liés aux marques épigénétiques d'intérêt (comme les enzymes DNMT et TET). Est-ce que ces fragments ont été mis en évidence car il existe de nombreux *gènes épigénétiques* dans le génome et qu'ils ont donc une forte probabilité d'être retrouvés dans un set de longueur similaire, ou sont-ils là parce que leur déméthylation a été déclenchée? L'analyse d'enrichissement tente de répondre à cette question. S'il y a beaucoup plus de gènes liés à l'épigénétique que par hasard, nous considérerons qu'il est enrichi et que, par conséquent, les fragments pourraient être liés de manière fonctionnelle à la période de reprogrammation.

Dans les deux analyses, la grande majorité des termes GO mis en évidence sont liés à la régulation des processus biologiques, à la communication cellulaire et au développement. Beaucoup de termes GO sont partagés entre les deux sets dont les plus significativement enrichis alors qu'ils sont constitués de fragments différents. Cela sous-entend l'importance et la complexité des mécanismes de régulation épigénétique lors du développement embryonnaire. De même, les termes de *haut niveau hiérarchique* (c'est-à-dire incluant beaucoup de gènes) tels que *cell communication* ou *cellular response to stimulus* retrouvés dans les deux analyses signifient que les CpGi sont situés dans les régions promotrices et géniques de très nombreux gènes. D'ailleurs, environ 70% des promoteurs annotés sont associés à un CpGi, ce qui en fait le type de promoteur le plus répandu dans le génome des vertébrés [Saxonov *et al.*, 2006].

Plus de termes GO et plus de termes GO très significativement enrichis se retrouvent dans le set des fragments ncst par rapport au set des cst. Les gènes associés aux fragments ncst interviennent dans plus de processus biologiques, mais aussi de manière plus spécifique que les fragments cst, puisque moins de fragments ncst sont simultanément liés à plusieurs termes GO. C'est une information difficilement interprétable dans l'état, qui semble refléter la spécificité des fragments ncst dans la régulation transcriptionnelle de processus biologiques spécifiques à l'embryogenèse. Le nombre de termes GO associés à la voie de signalisation Wnt (*Wnt signaling pathway*) va dans ce sens (un cinquième des termes GO de l'analyse ON des ncst). Il s'agit d'une voie ancienne et conservée depuis les métazoaires, impliquée dans l'induction de la différenciation cellulaire afin de favoriser la formation d'organes importants tels que les poumons, le cœur et les ovaires au cours du développement embryonnaire. Les 19 protéines, constituant la grande famille des Wnt chez l'homme, assurent en outre le développement de ces tissus grâce à une régulation de la différenciation, de la prolifération et de la migration cellulaire [Komiya et Habas, 2008].

De même, plusieurs termes GO spécifiques à l'analyse des fragments ncst font référence à la voie de signalisation des récepteurs ephrin (*ephrin receptor signaling pathway*). Cette voie est impliquée

dans la régulation de nombreux processus essentiels au développement embryonnaire [Klein, 2012]. Comme les autres voies de signalisation qui jouent un rôle crucial au cours de l'embryogenèse sont étroitement réglementées, l'expression de ces protéines (et de leurs antagonistes) est extrêmement restreinte dans le temps et dans l'espace au cours du développement [Cadigan et Liu, 2006]. Une dernière observation appuyant l'idée selon laquelle les fragments ncst participent de manière plus spécifique que les fragments cst au développement embryonnaire est le seul terme KEGG significativement enrichie dans l'analyse des ncst : la voie de signalisation de Hedgehog (*Hedgehog signaling pathway*). Il s'agit d'une voie complexe faisant intervenir de nombreux acteurs ayant un impact dans une grande variété d'événements développementaux chez l'embryon de nombreux organismes dont les vertébrés [McMahon, 2000].

De façon intéressante, la très grande majorité des gènes ncst mis en évidence voient leur niveau de méthylation diminuer durant le *reprogramming*. De même, les gènes directement associés au terme GO *changes to DNA methylation involved in embryo development* ne se retrouvent que dans le set des ncst. Parmi ces gènes, on retrouve la *methylcytosine dioxygenase TET3* présentée un peu plus haut, mais aussi la *DNA (cytosine-5)-methyltransferase 3a-like isoform X1* (une isoforme de la DNMT3a participant à la reméthylation de l'embryon pendant le *reprogramming*) dont la chute du niveau de méthylation entre le stade OL et 27 pourrait être corrélée avec le pic d'expression de l'ARNm associé à l'enzyme DNMT3av1 à ce même moment [Fellous *et al.*, 2018]. De même, plusieurs fragments associés aux *histone-lysine N-methyltransferase* présentent le même profil que les gènes précédemment cités parmi ceux-ci, on retrouve EZH2 qui méthyle les Lys-9 (formant H3K9me) et Lys-27 (formant H3K27me) de l'histone H3, entraînant une répression transcriptionnelle des gènes ciblés par ces complexes. Plus intéressant, EZH2 peut recruter certaines DNMT catalysant la méthylation de l'ADN de régions ciblées provoquant leur répression transcriptionnelle [Liu *et al.*, 2018]. Si aucun terme GO ne fait directement référence à des modifications de la méthylation de l'ADN, c'est probablement que les gènes liés n'étaient pas assez nombreux pour être significativement enrichi.

5 Conclusion

Le killifish des mangroves, de par son mode de reproduction unique chez les vertébrés permettant de générer des lignées hautement homozygotes et isogéniques, offre l'opportunité de mettre en évidence les contributions épigénétiques à la variation phénotypique et d'approfondir les relations entre environnement, épigénome et phénotype. Dans ce travail, cet organisme fut utilisé pour étudier une période cruciale du développement embryonnaire des vertébrés, le *reprogramming*, sur bases d'une analyse RRBS réalisée sur trois stades embryonnaires caractérisant le rétablissement progressif des marques épigénétiques jusqu'à leur niveau basal.

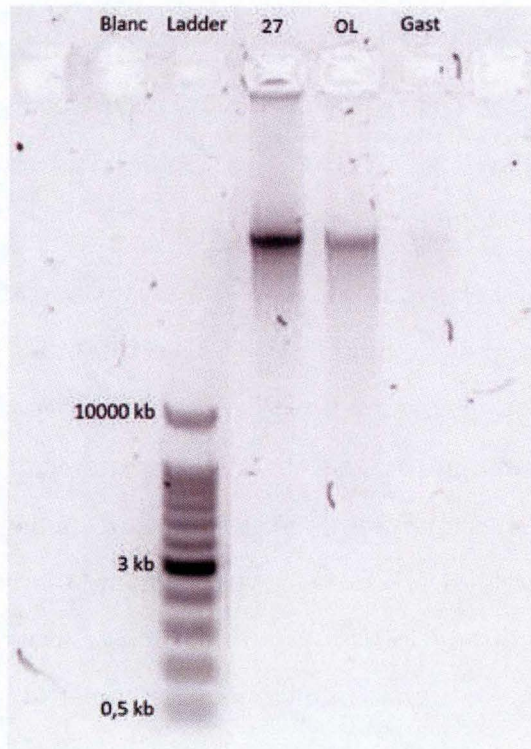
Il fut mis en évidence que le méthylome de *K. marmoratus* au niveau des CpGi était semblable à celui *D. rerio* et différait de celui des mammifères en ayant des niveaux de méthylation plus élevés que ces derniers et qu'au cours du *reprogramming*, la répartition des CpG hypo et hyperméthylés au sein des différentes régions génomiques ne variait pas mais que les régions promotrices semblaient exclure les CpG hyperméthylés. Les analyses d'enrichissement ont, quant à elles, permis de souligner la spécificité des fragments nct dans les processus développementaux embryonnaires.

La prévalence des fragments cst durant le *reprogramming* pousse à se questionner sur la sensibilité et la capacité des îlots CpG à conserver une épimutation de manière transgénérationnelle. De futures recherches devront investiguer cette piste en suivant sur plusieurs générations des épimutations générées soit de manière ciblée par la technologie CRISPR/Cas9 soit de manière aléatoire après en avoir identifiées. Une autre perspective à ce travail serait de déterminer si le méthylome de l'embryon de *K. marmoratus* est calqué sur celui des spermatozoïdes comme c'est le cas chez le poisson-zèbre où le méthylome du gamète femelle est effacé et remplacé par celui du mâle. Dans cette optique, ce travail a permis de mettre en place les principales étapes nécessaires pour classer, isoler et étudier de manière spécifique des données RRBS par pyroséquençage. Ce qui permettra sans nul doute d'exploiter au mieux d'autres résultats RRBS au sein de notre laboratoire.

Il est encore trop tôt pour avancer que *K. marmoratus* a sélectionné un mécanisme favorisant l'intégration des signaux environnementaux durant l'embryogenèse pour produire des phénotypes adaptés aux conditions environnementales, plutôt que de miser sur la faculté d'hériter du méthylome d'une génération à l'autre mais le phénomène de reprogrammation épigénétique soulève de nombreuses questions passionnantes et les mécanismes moléculaires impliqués commencent à peine à révéler leurs secrets. Ces nouvelles connaissances associées à l'utilisation d'organismes aussi pertinents que le killifish des mangroves pour étudier les mécanismes épigénétiques et au développement rapide des nouvelles technologies permettront probablement de répondre à ces questions dans un avenir proche.

Annexes

A



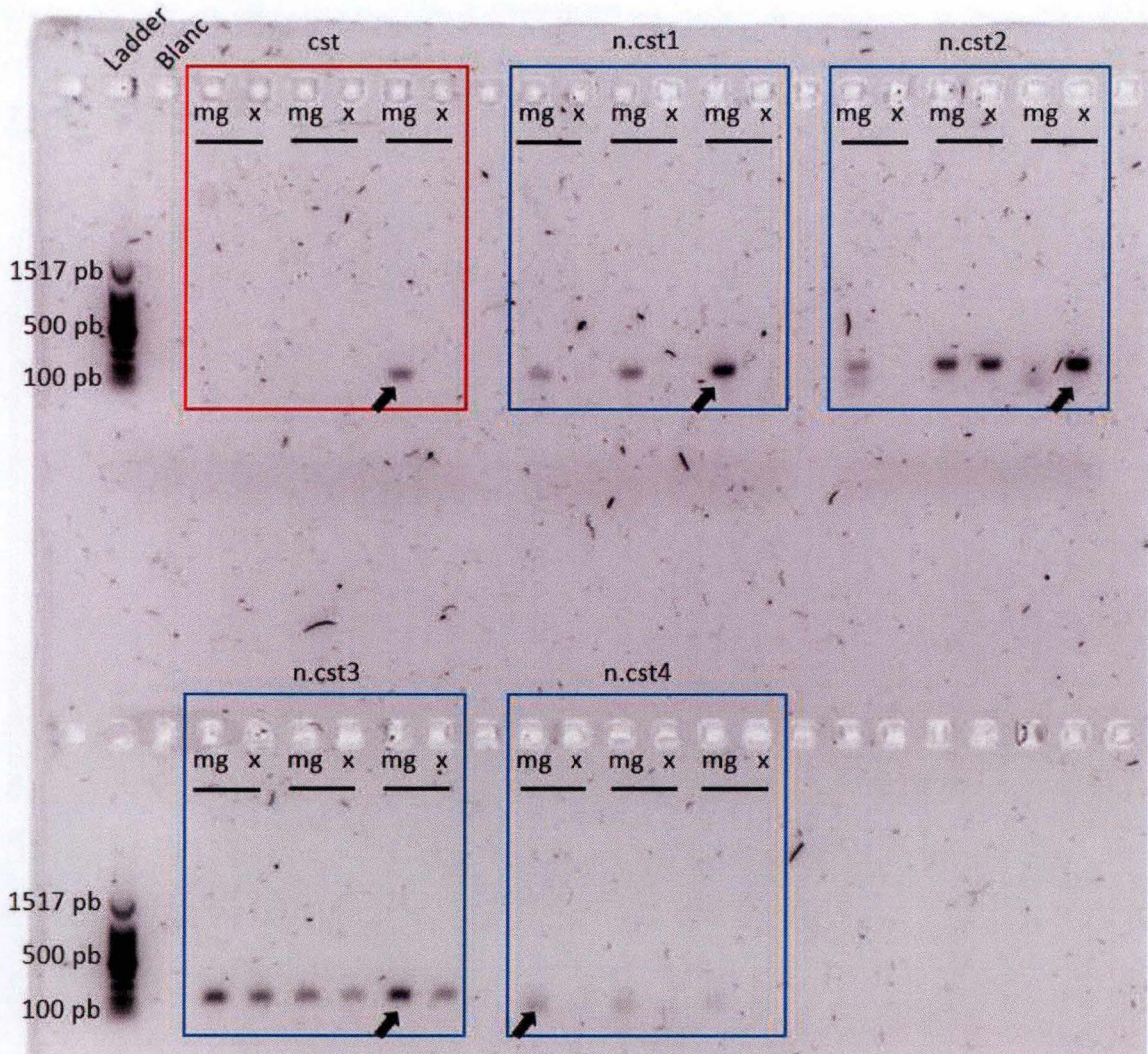
Gel d'agarose 1% sur lequel a migré 4 μ l d'ADN suite à l'extraction d'ADN de pool de sept œufs de stade 27, 13 de stade OL et 20 de stade Gast.

B

Nom	Sequence	Scale	Purification
cst1.1.Left	GTGATTTAAGAGGGGTATGTTTAG	25 nm	STD
cst1.1.Right	ACTACCACAAATAAAAAATCTACTATTAAC	25 nm	STD
cst1.2.Left	TGATTTAAGAGGGGTATGTTTAG	25 nm	STD
cst1.2.Right	AACTACCACAAATAAAAAATCTACTAT	25 nm	STD
cst1.3.Left	TGATTTAAGAGGGGTATGTTTAG	25 nm	STD
cst1.3.Right	AACTACCACAAATAAAAAATCTACTATTA	25 nm	STD
ncst1.1.Left	AGTATATTTTATGGAAAGAGGATTTTG	25 nm	STD
ncst1.1.Right	ATAACCACTTATAATTATACACTTTACC	25 nm	STD
ncst1.2.Left	AGTATATTTTATGGAAAGAGGATTTTG	25 nm	STD
ncst1.2.Right	CTATAACCACTTATAATTATACACTTTACC	25 nm	STD
ncst1.3.Left	GTATATTTTATGGAAAGAGGATTTTG	25 nm	STD
ncst1.3.Right	ACTATAACCACTTATAATTATACACTTTAC	25 nm	STD
ncst2.1.Left	GTGTGTTGAAATGGGTTTTTTTAATA	25 nm	STD
ncst2.1.Right	ACCCCACCATATTAACCTTTAAATAAAC	25 nm	STD
ncst2.2.Left	TGTGTGTTGAAATGGGTTTTTTTAATA	25 nm	STD
ncst2.2.Right	ACCCCACCATATTAACCTTTAAATAAAC	25 nm	STD
ncst2.3.Left	GAGTGTGTGTTGAAATGGGTTTTTT	25 nm	STD
ncst2.3.Right	ACCCCACCATATTAACCTTTAAATAAAC	25 nm	STD
ncst3.1.Left	TGGAATTTTGATATTTGGAATGTAAA	25 nm	STD
ncst3.1.Right	AAATCACCTCCTTACATATATAAACTT	25 nm	STD
ncst3.2.Left	TGGAATTTTGATATTTGGAATGTAAA	25 nm	STD
ncst3.2.Right	CAAATCACCTCCTTACATATATAAACTT	25 nm	STD
ncst3.3.Left	TGGAATTTTGATATTTGGAATGTAAA	25 nm	STD
ncst3.3.Right	AAATCACCTCCTTACATATATAAACTT	25 nm	STD
ncst4.1.Left	AGGAGGGAGTTTTTATTTTGTAGTT	25 nm	STD
ncst4.1.Right	CAAAAAAAAAAACTTTATTTATCCTCCTACT	25 nm	STD
ncst4.2.Left	AGGAGGGAGTTTTTATTTTGTAGTT	25 nm	STD
ncst4.2.Right	CCTCCTCCCTCCAAAAAAAAAACTTTATT	25 nm	STD
ncst4.3.Left	GTTTAGGAGGGAGTTTTTATTTTGTAGTT	25 nm	STD
ncst4.3.Right	CAAAAAAAAAAACTTTATTTATCCTCCTACT	25 nm	STD

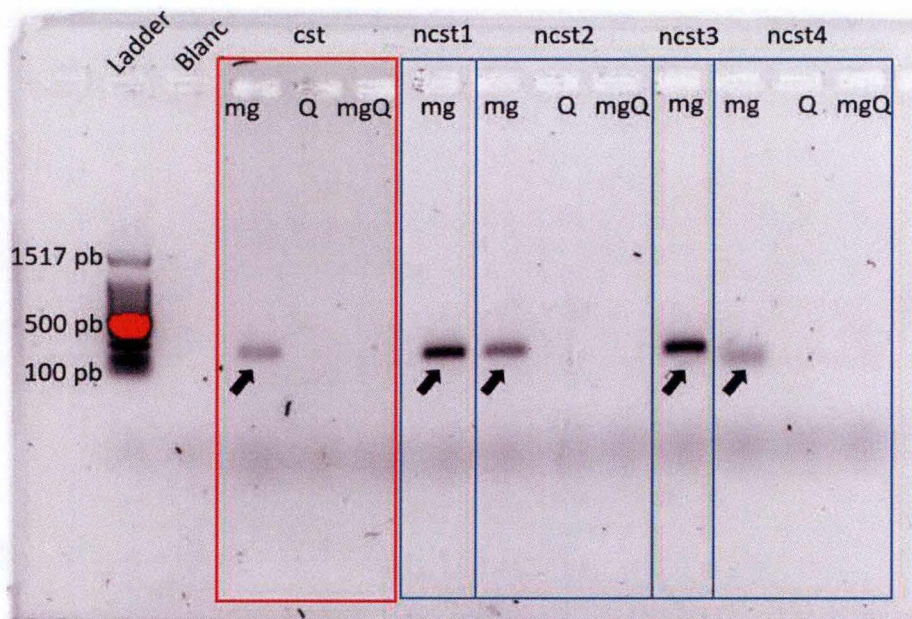
Paires d'amorces sélectionnés pour être testés par PCR simple en vue de déterminer quelle paire des trois testées par fragments sera conservées pour la PCR bisulfite. La colonne "nom" correspond à leur identifiant, il est à mettre en relation avec leur catégorie (cst ou ncst).

C



Double gel d'agarose 1% sur lequel a migré l'ADN amplifié par PCR (si les primers ont correctement fonctionné). Un cadre représente un fragment testé pour trois paires de primers différentes selon deux conditions à chaque fois : avec ou sans ajout de $MgCl_2$. Les flèches noires indiquent les paires de primers et les conditions optimales permettant l'amplification des fragments d'intérêts. L'ADN utilisé provient de pools d'œufs de stade 27.

D



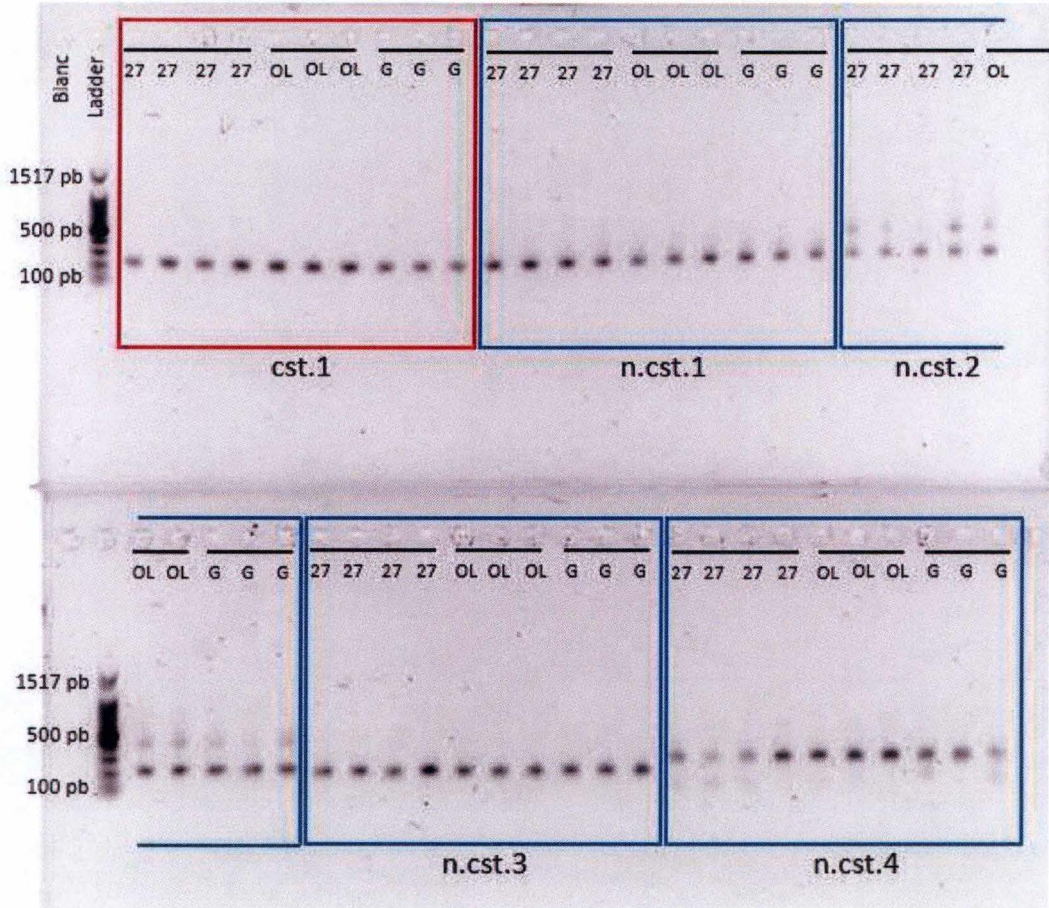
Gel d'agarose 1% sur lequel a migré les fragments amplifié par PCR (si les primers ont correctement fonctionné). Un cadre représente un fragment testé pour une paires de primers selon une ou trois conditions : avec ou sans ajout de $MgCl_2$ et de Q-solution. Les flèches noires indiquent les paires de primers et les conditions optimales permettant l'amplifications des fragments d'intérêts. L'ADN utilisé provient de pools d'œufs de stade 27.

E

Nom	Sequence	Scale	Purification
cst1.3.Left	TGATTTAAGAGGGGTATGTTT	25 nm	STD
cst1.3.Right	/5Biosg/AACTACCACAAATAAAAAATCTACTATTA	25 nm	STD
cst1.3.Sequ	GAGGGGTATGTTTAGA	25 nm	STD
ncst1.3.Left	GTATATTTTATGGAAAGAGGATTTTGG	25 nm	STD
ncst1.3.Right	/5Biosg/ACTATAACCACTTATAATTATACACTTTAC	25 nm	STD
ncst1.3.Sequ	GGAAAGAGGATTTTGGT	25 nm	STD
ncst2.2.Left	TGTGTGTTGAAATGGGTTTTTTTAATA	25 nm	STD
ncst2.2.Right	/5Biosg/ACCCCACCATATTAACCTTTAAATAAAC	25 nm	STD
ncst2.2.Sequ	AATATTAAGTGAGAGG	25 nm	STD
ncst3.3.Left	TGGAATTTTGATATTTGGAATGTAAA	25 nm	STD
ncst3.3.Right	/5Biosg/AAATCACCTCCTTACATATATAAACTT	25 nm	STD
ncst3.3.Sequ	TTTGATATTTGGAATGTAAATTTT	25 nm	STD
ncst4.1.Left	AGGAGGGAGTTTTTATTTTGTAGTT	25 nm	STD
ncst4.1.Right	/5Biosg/CAAAAAAAAAACTTTATTTATCCTCCTACT	25 nm	STD
ncst4.1.Sequ	TTTGTAGTTTTTTTTTATTTTTAG	25 nm	STD

Amorces sélectionnés pour amplifier les cinq fragments d'intérêt ainsi que leur primer de séquençage qui sera utilisé durant la méthylation gène spécifique. La colonne "nom" correspond à leur identifiant, il est à mettre en relation avec leur catégorie (cst ou ncst).

F



Double gel d'agarose 1% sur lequel a migré les ADN bisulfités amplifiés par PCR. Un cadre représente un fragment testé (et donc, une paire d'amorces) pour quatre échantillons différents d'ADN de stade 27, trois de stade OL et trois de stade Gast. L'escalier que forme les bandes d'ADN sur le gel du haut provient d'une déformation du portoir dans lequel le gel a été coulé.

G

ID	CpG												
		1	2	3	4	5	6	7	8	9	10	11	12
cst	27.1	100	87	4	79	96	80	70	97	89	77	100	89
	27.2	100	82	4	53	97	73	69	67	94	84	100	69
	27.3	100	80	4	62	98	64	53	94	85	67	95	100
	OL.1	100	86	5	74	46	73	66	93	90	77	97	100
	OL.2	100	38	24	38	100	35	43	93	91	82	100	97
	OL.3	100	89	5	83	91	79	76	91	88	88	100	99
	G.1	100	76	3	62	91	72	61	92	91	81	98	100
	G.2	100	73	3	62	100	58	50	100	85	54	100	96
	G.3	100	52	4	35	94	57	41	67	68	60	76	72
ncst1	27.1	41	33	42	27	35	37	41					
	27.2	56	45	54	66	41	53	21					
	27.3	52	41	50	37	49	47	48					
	OL.1	95	83	90	66	100	93	93					
	OL.2	88	77	89	69	47	82	93					
	OL.3	93	84	90	77	95	95	93					
	G.1	89	85	88	72	94	89	91					
	G.2	96	90	94	80	97	100	94					
	G.3	97	91	95	82	98	100	97					
ncst2	27.1	32	32	40	29								
	27.2	46	47	36	26								
	27.3	52	56	45	32								
	OL.1	92	79	90	89								
	OL.2	79	72	87	73								
	OL.3	87	79	92	84								
	G.1	90	82	12	93								
	G.2	94	84	96	91								
	G.3	89	78	94	88								
ncst3	27.1	27	85	71	87	69	91	87					
	27.2	20	87	81	89	79	92	91					
	27.3	26	78	71	77	70	84	79					
	OL.1	18	60	41	63	46	70	66					
	OL.2	21	72	57	84	59	74	75					
	OL.3	20	57	57	70	59	79	73					
	G.1	15	35	27	36	29	42	41					
	G.2	7	18	12	13	14	19	22					
	G.3	7	16	11	11	8	14	21					
ncst4	27.1	12	40	36	28	4	9						
	27.2	14	37	35	40	6	13						
	27.3	13	44	39	39	4	9						
	OL.1	8	93	54	68	7	15						
	OL.2	9	92	18	59	6	16						
	OL.3	8	93	48	64	8	19						
	G.1	43	64	76	39	7	22						
	G.2	6	100	59	80	14	32						
	G.3	10	34	25	40	4	8						

Niveau de méthylation de chaque CpG d'intérêt analysé par pyroséquençage sur les trois pools d'œufs de stade Gast, OL, et 27.

Références

- [Alonso *et al.*, 2014] ALONSO, C., PEREZ, R., BAZAGA, P., MEDRANO, M. et HERRERA, C. M. (2014). Individual variation in size and fecundity is correlated with differences in global DNA cytosine methylation in the perennial herb *Helleborus foetidus* (Ranunculaceae). *American Journal of Botany*, 101(8):1309–1313.
- [Angers *et al.*, 2010] ANGERS, B., CASTONGUAY, E. et MASSICOTTE, R. (2010). Environmentally induced phenotypes and DNA methylation : how to deal with unpredictable conditions until the next generation and after. *Molecular Ecology*, 19(7):1283–1295.
- [Avisé et Tatarenkov, 2015] AVISE, J. C. et TATARENKOV, A. (2015). Population genetics and evolution of the mangrove rivulus *Kryptolebias marmoratus*, the world's only self-fertilizing hermaphroditic vertebrate. *Journal of Fish Biology*, 87(3):519–538.
- [Baker *et al.*, 2003] BAKER, P. E., WILKOWSKI, J. et BURKE, D. T. (2003). Age-associated activation of epigenetically repressed genes in the mouse. *Genetics*, 165(4):2055–62.
- [Bateson, 2012] BATESON, P. (2012). The impact of the organism on its descendants. *Genetics Research International*, pages 1–7.
- [Berger, 2007] BERGER, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, 447:407.
- [Bestor, 2000] BESTOR, T. H. (2000). The DNA methyltransferases of mammals. *Human Molecular Genetics*, 9(16):2395–2402.
- [Bestor *et al.*, 2015] BESTOR, T. H., EDWARDS, J. R. et BOULARD, M. (2015). Reply to Wilkinson : minor role of programmed methylation and demethylation in mammalian development. *Proceedings of the National Academy of Sciences*, 112(17):E2117–E2117.
- [Bird, 2007] BIRD, A. (2007). Perceptions of epigenetics. *Nature*, 447(7143):396–398.
- [Brander *et al.*, 2017] BRANDER, S. M., BIALES, A. D. et CONNOR, R. E. (2017). The role of epigenomics in aquatic toxicology. *Environmental Toxicology and Chemistry*, 36(10):2565–2573.
- [Brown *et al.*, 2012] BROWN, A. R., BICKLEY, L. K., RYAN, T. A., PAULL, G. C., HAMILTON, P. B., OWEN, S. F., SHARPE, A. D. et TYLER, C. R. (2012). Differences in sexual development in inbred and outbred zebrafish (*Danio rerio*) and implications for chemical testing. *Aquatic Toxicology*, 112-113:27–38.
- [Burggren, 2014] BURGGREN, W. (2014). Epigenetics as a source of variation in comparative animal physiology or Lamarck is looking pretty good these days. *Journal of Experimental Biology*, 217(5):682–689.

- [Burggren, 2016] BURGGREN, W. (2016). Epigenetic inheritance and its role in evolutionary biology : re-evaluation and new perspectives. *Biology*, 5(2):24.
- [Cadet et Wagner, 2014] CADET, J. et WAGNER, J. R. (2014). TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 764-765:18–35.
- [Cadigan et Liu, 2006] CADIGAN, K. M. et LIU, Y. I. (2006). Wnt signaling : complexity at the surface. *Journal of Cell Science*, 119(3):395–402.
- [Champagne, 2008] CHAMPAGNE, F. A. (2008). Epigenetic mechanisms and the transgenerational effects of maternal care. *Frontiers in neuroendocrinology*, 29(3):386–97.
- [Champagne, 2013] CHAMPAGNE, F. A. (2013). Epigenetics and developmental plasticity across species. *Developmental psychobiology*, 55(1):33–41.
- [Costa, 2011] COSTA, W. (2011). Identity of *Rivulus ocellatus* and a new name for a hermaphroditic species of Kryptolebias from south-eastern Brazil (Cyprinodontiformes : Rivulidae). *Ichthyological Exploration of Freshwaters*, 22(2):185–192.
- [Costa, 1994] COSTA, W. J. (1994). Comments On The Proposed Conservation of the specific name of *Rivulus marmoratus* Poey, 1880 (Osteichthyes, Cyprinodontiformes). *The Bulletin of zoological nomenclature.*, 51:46–47.
- [Costa, 2006] COSTA, W. J. (2006). Redescription of *Kryptolebias ocellatus* (Hensel) and *K. caudomarginatus* (Seegers), two killifishes from mangroves of south-eastern Brazil. *Aqua, International Journal of Ichthyology*, 11:5–12.
- [Davis et al., 1990] DAVIS, W. P., TAYLOR, D. S. et TURNER, B. J. (1990). Field observations of the ecology and habits of mangrove rivulus (*Rivulus marmoratus*) in Belize and Florida. *Ecology*, I(2).
- [Edwards et al., 2017] EDWARDS, J. R., YARYCHKIVSKA, O., BOULARD, M. et BESTOR, T. H. (2017). DNA methylation and DNA methyltransferases. *Epigenetics & Chromatin*, 10(1):23.
- [Elliott et al., 2015] ELLIOTT, G., HONG, C., XING, X., ZHOU, X., LI, D., COARFA, C., BELL, R. J., MAIRE, C. L., LIGON, K. L., SIGAROUDINIA, M., GASCARD, P., TLSTY, T. D., HARRIS, R. A., SCHALKWYK, L. C., BILENKY, M., MILL, J., FARNHAM, P. J., KELLIS, M., MARRA, M. A., MILOSAVLJEVIC, A., HIRST, M., STORMO, G. D., WANG, T. et COSTELLO, J. F. (2015). Intermediate DNA methylation is a conserved signature of genome regulation. *Nature Communications*, 6(1):6363.
- [Ellison et al., 2013] ELLISON, A., JONES, J., INCHLEY, C. et CONSUEGRA, S. (2013). Choosy males could help explain androdioecy in a selfing fish. *The American Naturalist*, 181(6):855–862.

- [Eschmeyer *et al.*, 2018] ESCHMEYER, W., VAN DER LAAN, R. et FRICKE, R. (2018). Catalog of fishes : genera, species, references.
- [Fan et Roberts, 2006] FAN, L. et ROBERTS, V. A. (2006). Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. *Proceedings of the National Academy of Sciences*, 103(22):8384–8389.
- [Fellous *et al.*, 2018] FELLOUS, A., LABED-VEYDERT, T., LOCREL, M., VOISIN, A.-S. et SILVESTRE, F. (2018). DNA methylation in adults and during development of the self-fertilizing mangrove rivulus, *Kryptolebias marmoratus*. *Ecology and Evolution*.
- [Feng *et al.*, 2010] FENG, S., COKUS, S. J., ZHANG, X., CHEN, P.-Y., BOSTICK, M., GOLL, M. G., HETZEL, J., JAIN, J., STRAUSS, S. H., HALPERN, M. E., UKOMADU, C., SADLER, K. C., PRADHAN, S., PELLEGRINI, M. et JACOBSEN, S. E. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences*, 107(19):8689–8694.
- [Frias-Lasserre et Villagra, 2017] FRIAS-LASSERRE, D. et VILLAGRA, C. A. (2017). The importance of ncRNAs as epigenetic mechanisms in phenotypic variation and organic evolution. *Frontiers in Microbiology*, 8(DEC):1–13.
- [Garrett *et al.*, 2015] GARRETT, F. E., SHERIDAN, C. K., KACMARCZYK, T. J., ISHII, J., BETEL, D., ALONSO, A., MASON, C. E., FIGUEROA, M. E. et MELNICK, A. M. (2015). Enhanced Reduced Representation Bisulfite Sequencing for assessment of DNA methylation at base pair resolution. *Journal of Visualized Experiments*, (96):e52246.
- [Grageda *et al.*, 2004] GRAGEDA, M. V. C., SAKAKURA, Y. et HAGIWARA, A. (2004). Early development of the self-fertilizing mangrove killifish *Rivulus marmoratus* reared in the laboratory. *Ichthyological Research*, 51(4):309–315.
- [Greiner, 2012] GREINER, V. (2012). *Épigénétique et méthylation de l'ADN : étude des mécanismes d'interaction du domaine SRA de UHRF1 avec l'ADN hémi-méthylé*. Thèse de doctorat, Université de Strasbourg.
- [Groth *et al.*, 2007] GROTH, A., ROCHA, W., VERREAULT, A. et ALMOUZNI, G. (2007). Chromatin challenges during DNA Replication and repair. *Cell*, 128(4):721–733.
- [Gu *et al.*, 2010] GU, H., BOCK, C., MIKKELSEN, T. S., JÄGER, N., SMITH, Z. D., TOMAZOU, E., GNIRKE, A., LANDER, E. S. et MEISSNER, A. (2010). Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nature Methods*, 7(2):133–136.
- [Hackett et Surani, 2013] HACKETT, J. A. et SURANI, M. A. (2013). Beyond DNA : programming and inheritance of parental methylomes. *Cell*, 153(4):737–9.

- [Harrington, 1968] HARRINGTON, R. (1968). Delimitation of the thermolabile phenocritical period of sex determination and differentiation in the ontogeny of the normally hermaphroditic fish *Rivulus marmoratus*. *Physiological Zoology*, 41:447–460.
- [Harris *et al.*, 2010] HARRIS, R. A., WANG, T., COARFA, C., NAGARAJAN, R. P., HONG, C., DOWNEY, S. L., JOHNSON, B. E., FOUSE, S. D., DELANEY, A., ZHAO, Y., OLSHEN, A., BALLINGER, T., ZHOU, X., FORSBERG, K. J., GU, J., ECHIPARE, L., O'GEEN, H., LISTER, R., PELIZZOLA, M., XI, Y., EPSTEIN, C. B., BERNSTEIN, B. E., HAWKINS, R. D., REN, B., CHUNG, W.-Y., GU, H., BOCK, C., GNIRKE, A., ZHANG, M. Q., HAUSSLER, D., ECKER, J. R., LI, W., FARNHAM, P. J., WATERLAND, R. A., MEISSNER, A., MARRA, M. A., HIRST, M., MILOSAVLJEVIC, A. et COSTELLO, J. F. (2010). Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology*, 28(10):1097–1105.
- [He *et al.*, 2015] HE, X., TILLO, D., VIERSTRA, J., SYED, K.-S., DENG, C., RAY, G. J., STAMATOYANNOPOULOS, J., FITZGERALD, P. C. et VINSON, C. (2015). Methylated cytosines mutate to transcription factor binding sites that drive tetrapod evolution. *Genome Biology and Evolution*, 7(11):3155–3169.
- [Head, 2014] HEAD, J. A. (2014). Patterns of DNA methylation in animals : an ecotoxicological perspective. *Integrative and Comparative Biology*, 54(1):77–86.
- [Hendrich et Tweedie, 2003] HENDRICH, B. et TWEEDIE, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends in Genetics*, 19(5):269–277.
- [Henikoff et Shilatifard, 2011] HENIKOFF, S. et SHILATIFARD, A. (2011). Histone modification : cause or cog? *Trends in Genetics*, 27(10):389–396.
- [Hubert *et al.*, 2010] HUBERT, B., LYON, B., CLAUDE, L. U. et LYON, B. (2010). *La régulation épigénétique des éléments transposables dans les populations naturelles de Drosophila simulans*. Thèse de doctorat.
- [Illingworth et Bird, 2009] ILLINGWORTH, R. S. et BIRD, A. P. (2009). CpG islands : a rough guide. *FEBS Letters*, 583(11):1713–1720.
- [Jablonka et Lamb, 2002] JABLONKA, E. et LAMB, M. J. (2002). The changing concept of epigenetics. *Annals of the New York Academy of Sciences*, 981:82–96.
- [Kanehisa *et al.*, 2017] KANEHISA, M., FURUMICHI, M., TANABE, M., SATO, Y. et MORISHIMA, K. (2017). KEGG : new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, 45(D1):D353–D361.
- [Kappil *et al.*, 2015] KAPPIL, M., LAMBERTINI, L. et CHEN, J. (2015). Environmental influences on genomic imprinting. *Current environmental health reports*, 2(2):155–162.

- [Kelley *et al.*, 2016] KELLEY, J. L., YEE, M. C., BROWN, A. P., RICHARDSON, R. R., TATARENKOV, A., LEE, C. C., HARKINS, T. T., BUSTAMANTE, C. D. et EARLEY, R. L. (2016). The genome of the self-fertilizing mangrove rivulus fish, *Kryptolebias marmoratus* : a model for studying phenotypic plasticity and adaptations to extreme environments. *Genome biology and evolution*, 8(7):2145–2154.
- [Kiani *et al.*, 2013] KIANI, J., GRANDJEAN, V., LIEBERS, R., TUORTO, F., GHANBARIAN, H., LYKO, F., CUZIN, F. et RASSOULZADEGAN, M. (2013). RNA-mediated epigenetic heredity requires the cytosine methyltransferase Dnmt2. *PLoS Genetics*, 9(5).
- [Kilvitis *et al.*, 2014] KILVITIS, H. J., ALVAREZ, M., FOUST, C. M., SCHREY, A. W., ROBERTSON, M. et RICHARDS, C. L. (2014). Ecological Epigenetics. *In Advances in experimental medicine and biology*, volume 781, pages 191–210.
- [King *et al.*, 1989] KING, J. A. C., ABEL, D. C. et DIBONA, D. R. (1989). Effects of salinity on chloride cells in the euryhaline cyprinodontid fish *Rivulus marmoratus*. *Cell and Tissue Research*, 257(2):367–377.
- [Klein, 2012] KLEIN, R. (2012). Eph/ephrin signalling during development. *Development (Cambridge, England)*, 139(22):4105–9.
- [Klironomos *et al.*, 2013] KLIRONOMOS, F. D., BERG, J. et COLLINS, S. (2013). How epigenetic mutations can affect genetic evolution : model and mechanism. *BioEssays*, 35(6):571–578.
- [Komiya et Habas, 2008] KOMIYA, Y. et HABAS, R. (2008). Wnt signal transduction pathways. *Organogenesis*, 4(2):68–75.
- [Laing *et al.*, 2016] LAING, L. V., VIANA, J., DEMPSTER, E. L., TRZNADEL, M., TRUNKFIELD, L. A., UREN WEBSTER, T. M., van AERLE, R., PAULL, G. C., WILSON, R. J., MILL, J. et SANTOS, E. M. (2016). Bisphenol A causes reproductive toxicity, decreases dnmt1 transcription, and reduces global DNA methylation in breeding zebrafish (*Danio rerio*). *Epigenetics*, 11(7):526–38.
- [Law et Jacobsen, 2010] LAW, J. A. et JACOBSEN, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, 11(3):204–220.
- [Lea *et al.*, 2017] LEA, A. J., VILGALYS, T. P., DURST, P. A. et TUNG, J. (2017). Maximizing ecological and evolutionary insight in bisulfite sequencing data sets. *Nature Ecology and Evolution*, 1(8):1074–1083.
- [Lee *et al.*, 2008] LEE, J.-S., RAISUDDIN, S. et SCHLENK, D. (2008). *Kryptolebias marmoratus* (Poey, 1880) : a potential model species for molecular carcinogenesis and ecotoxicogenomics. *Journal of Fish Biology*, 72(8):1871–1889.

- [Lee *et al.*, 2010] LEE, T.-f., ZHAI, J. et MEYERS, B. C. (2010). Conservation and divergence in eukaryotic DNA methylation. *Proceedings of the National Academy of Sciences*, 107(20):9027–9028.
- [Li *et al.*, 2012] LI, C.-M., TZENG, J.-N. et SUNG, H.-M. (2012). Effects of cis and trans regulatory variations on the expression divergence of heat shock response genes between yeast strains. *Gene*, 506(1):93–97.
- [Li et Dahiya, 2002] LI, L. et DAHIYA, R. (2002). MethPrimer : designing primers for methylation PCRs. *Bioinformatics*, 18(11):1427–1431.
- [Lins *et al.*, 2018] LINS, L. S. F., TROJAHN, S., SOCKELL, A., YEE, M.-C., TATARENKOV, A., BUSTAMANTE, C. D., EARLEY, R. L. et KELLEY, J. L. (2018). Whole-genome sequencing reveals the extent of heterozygosity in a preferentially self-fertilizing hermaphroditic vertebrate. *Genome*, 61(December):241–247.
- [Liu *et al.*, 2004] LIU, J., HE, Y., AMASINO, R. et CHEN, X. (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. *Genes & Development*, 18(23):2873–2878.
- [Liu *et al.*, 2018] LIU, X., LI, C., ZHANG, R., XIAO, W., NIU, X., YE, X., LI, Z., GUO, Y., TAN, J. et LI, Y. (2018). The EZH2- H3K27me3-DNMT1 complex orchestrates epigenetic silencing of the *wwc1* gene, a Hippo/YAP pathway upstream effector, in breast cancer epithelial cells. *Cellular Signalling*, 51:243–256.
- [Liu *et al.*, 2016] LIU, Y., DUONG, W., KRAWCZYK, C., BRETSCHNEIDER, N., BORBÉLY, G., VARSHNEY, M., ZINSER, C., SCHÄR, P. et RÜEGG, J. (2016). Oestrogen receptor β regulates epigenetic patterns at specific genomic loci through interaction with thymine DNA glycosylase. *Epigenetics & Chromatin*, 9(1):7.
- [Loftus et Kushlan, 1987] LOFTUS, W. et KUSHLAN, J. (1987). Freshwater fishes of Southern Florida. 31:300.
- [Lomax *et al.*, 2017] LOMAX, J. L., CARLSON, R. E., WELLS, J. W., CRAWFORD, P. M. et EARLEY, R. L. (2017). Factors affecting egg production in the selfing mangrove rivulus (*Kryptolebias marmoratus*). *Zoology*, 122:38–45.
- [Mackiewicz *et al.*, 2006] MACKIEWICZ, M., TATARENKOV, A., PERRY, A., MARTIN, J. R., ELDER, J. F., BECHLER, D. L. et AVISE, J. C. (2006). Microsatellite documentation of male-mediated outcrossing between inbred laboratory strains of the self-fertilizing mangrove killifish. *Journal of Heredity*, 97(5):508–513.
- [Massicotte *et al.*, 2011] MASSICOTTE, R., WHITELAW, E. et ANGERS, B. (2011). DNA methylation : a source of random variation in natural populations. *Epigenetics*, 6(4):422–428.

- [Mayr et Provine, 1998] MAYR, E. et PROVINE, W. B. (1998). The evolutionary synthesis : perspectives on the unification of biology. page 487.
- [Mcghee et Ginder, 1979] MCGHEE, J. et GINDER, G. (1979). Specific DNA methylation sites in the vicinity of the chicken β -globin genes. *Nature*, 280(5721):419–420.
- [McMahon, 2000] MCMAHON, A. P. (2000). More surprises in the Hedgehog signaling pathway. *Cell*, 100(2):185–8.
- [Meng *et al.*, 2015] MENG, H., CAO, Y., QIN, J., SONG, X., ZHANG, Q., SHI, Y. et CAO, L. (2015). DNA Methylation, its mediators and genome integrity. *International Journal of Biological Sciences*, 11(5):604–617.
- [Messerschmidt *et al.*, 2014] MESSERSCHMIDT, D. M., KNOWLES, B. B. et SOLTER, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & development*, 28(8):812–28.
- [Metzger et Schulte, 2016] METZGER, D. C. et SCHULTE, P. M. (2016). Epigenomics in marine fishes. *Marine Genomics*, 30:43–54.
- [Moore *et al.*, 2012] MOORE, G. L., SUCAR, S., NEWSOME, J. M., ARD, M. E., BERNHARDT, L., BLAND, M. J. et RING, B. C. (2012). Establishing developmental genetics in a self-fertilizing fish (*Kryptolebias marmoratus*). *Integrative and Comparative Biology*, 52(6):781–791.
- [Moore *et al.*, 2013] MOORE, L. D., LE, T. et FAN, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, 38(1):23–38.
- [Morgan *et al.*, 1999] MORGAN, H. D., SUTHERLAND, H. G., MARTIN, D. I. et WHITELAW, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics*, 23(3):314–318.
- [Mourabit, 2012] MOURABIT, S. (2012). *Establishing the mangrove killifish, Kryptolebias marmoratus, as a model species for developmental biology*. Thèse de doctorat.
- [Munzel *et al.*, 2011] MUNZEL, M., GLOBISCH, D. et CARELL, T. (2011). 5-hydroxymethylcytosine, the sixth base of the genome. *Angewandte Chemie International Edition*, 50(29):6460–6468.
- [Nagy et Turecki, 2012] NAGY, C. et TURECKI, G. (2012). Sensitive periods in epigenetics : bringing us closer to complex behavioral phenotypes. *Epigenomics*, 4(4):445–457.
- [Nan *et al.*, 1998] NAN, X., NG, H.-H., JOHNSON, C. A., LAHERTY, C. D., TURNER, B. M., EISENMAN, R. N. et BIRD, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, 393(6683):386–389.
- [Oey *et al.*, 2015] OEY, H., ISBEL, L., HICKEY, P., EBAID, B. et WHITELAW, E. (2015). Genetic and epigenetic variation among inbred mouse littermates : identification of inter-individual differentially methylated regions. *Epigenetics & Chromatin*, 8(1):54.

- [Paun *et al.*, 2019] PAUN, O., VERHOEVEN, K. J. et RICHARDS, C. L. (2019). Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. *New Phytologist*, 221(2):738–742.
- [Peaston et Whitelaw, 2006] PEASTON, A. E. et WHITELAW, E. (2006). Epigenetics and phenotypic variation in mammals. *Mammalian Genome*, 17(5):365–374.
- [Potok *et al.*, 2013] POTOK, M. E., NIX, D. A., PARNELL, T. J. et CAIRNS, B. R. (2013). Germline epigenetics, and reprogramming in zebrafish early embryos. *Epigenetics & Chromatin*, 6(Suppl 1):O23.
- [Puumala et Hoyme, 2015] PUUMALA, S. E. et HOYME, H. E. (2015). Epigenetics in pediatrics. *Pediatrics in review*, 36(1):14–21.
- [Rando et Verstrepén, 2007] RANDO, O. J. et VERSTREPÉN, K. J. (2007). Timescales of genetic and epigenetic inheritance. *Cell*, 128(4):655–668.
- [Rasmussen et Helin, 2016] RASMUSSEN, K. D. et HELIN, K. (2016). Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development*, 30(7):733–50.
- [Raudvere *et al.*, 2019] RAUDVERE, U., KOLBERG, L., KUZMIN, I., ARAK, T., ADLER, P., PETERSON, H. et VILO, J. (2019). g :Profiler : a web server for functional enrichment analysis and conversions of gene lists. *Nucleic Acids Research*, 47(W1):W191–W198.
- [Regan *et al.*, 2011] REGAN, K. S., JONZ, M. G. et WRIGHT, P. A. (2011). Neuroepithelial cells and the hypoxia emersion response in the amphibious fish *Kryptolebias marmoratus*. *Journal of Experimental Biology*, pages 2560–2568.
- [Reik *et al.*, 2001] REIK, W., DEAN, W. et WALTER, J. (2001). Epigenetic reprogramming in mammalian development. *Science*, 293(5532):1089–1093.
- [Reimand *et al.*, 2007] REIMAND, J., KULL, M., PETERSON, H., HANSEN, J. et VILO, J. (2007). g :Profiler : a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, 35(suppl_2):W193–W200.
- [Rhee *et al.*, 2017] RHEE, J.-S., CHOI, B.-S., KIM, J., KIM, B.-M., LEE, Y.-M., KIM, I.-C., KANAMORI, A., CHOI, I.-Y., SCHARTL, M. et LEE, J.-S. (2017). Diversity, distribution, and significance of transposable elements in the genome of the only selfing hermaphroditic vertebrate *Kryptolebias marmoratus*. *Scientific Reports*, 7(1):40121.
- [Richards, 2006] RICHARDS, E. J. (2006). Inherited epigenetic variation - revisiting soft inheritance. *Nature Reviews Genetics*, 7(5):395–401.
- [Sandelin *et al.*, 2007] SANDELIN, A., CARNINCI, P., LENHARD, B., PONJAVIC, J., HAYASHIZAKI, Y. et HUME, D. A. (2007). Mammalian RNA polymerase II core promoters : insights from genome-wide studies. *Nature Reviews Genetics*, 8(6):424–436.

- [Saxonov *et al.*, 2006] SAXONOV, S., BERG, P. et BRUTLAG, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences*, 103(5):1412–1417.
- [Schrey et Richards, 2012] SCHREY, A. W. et RICHARDS, C. L. (2012). Within-genotype epigenetic variation enables broad niche width in a flower living yeast. *Molecular Ecology*, 21(11):2559–2561.
- [Seisenberger *et al.*, 2012] SEISENBERGER, S., PEAT, J. R., HORE, T. A., SANTOS, F., DEAN, W. et REIK, W. (2012). Reprogramming DNA methylation in the mammalian life cycle : building and breaking epigenetic barriers. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 368(1609):20110330–20110330.
- [Sheldon *et al.*, 2002] SHELDON, C. C., CONN, A. B., DENNIS, E. S. et PEACOCK, W. J. (2002). Different regulatory regions are required for the vernalization-induced repression of Flowering Locus C and for the epigenetic maintenance of repression. *The Plant cell*, 14(10):2527–37.
- [Shinya et Sakai, 2011] SHINYA, M. et SAKAI, N. (2011). Generation of highly homogeneous strains of zebrafish through full sib-pair mating. *G3 (Bethesda, Md.)*, 1(5):377–86.
- [Siomi et Siomi, 2009] SIOMI, H. et SIOMI, M. C. (2009). On the road to reading the RNA-interference code. *Nature*, 457(7228):396–404.
- [Skinner, 2015] SKINNER, M. K. (2015). Environmental epigenetics and a unified theory of the that facilitates neo-Darwinian evolution. 7(5):1296–1302.
- [Skinner *et al.*, 2014] SKINNER, M. K., GURRERERO-BOSAGNA, C., HAQUE, M. M., NILSSON, E. E., KOOP, J. A., KNUTIE, S. A. et CLAYTON, D. H. (2014). Epigenetics and the evolution of Darwin's finches. *Genome Biology and Evolution*, 6(8):1972–1989.
- [Smith *et al.*, 2016] SMITH, T. A., MARTIN, M. D., NGUYEN, M. et MENDELSON, T. C. (2016). Epigenetic divergence as a potential first step in darter speciation. *Molecular Ecology*, 25(8):1883–1894.
- [Smith *et al.*, 2012] SMITH, Z. D., CHAN, M. M., MIKKELSEN, T. S., GU, H., GNIRKE, A., REGEV, A. et MEISSNER, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*, 484(7394):339–44.
- [Srinageshwar *et al.*, 2016] SRINAGESHWAR, B., MAITI, P., DUNBAR, G. L. et ROSSIGNOL, J. (2016). Role of epigenetics in stem cell proliferation and differentiation : implications for treating neurodegenerative diseases. *International Journal of Molecular Sciences*, 17(2).
- [Stadler *et al.*, 2011] STADLER, M. B., MURR, R., BURGER, L., IVANEK, R., LIENERT, F., SCHÖLER, A., van NIMWEGEN, E., WIRBELAUER, C., OAKELEY, E. J., GAIDATZIS, D., TIWARI, V. K. et SCHÜBELER, D. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature*, 480(7378):490–495.

- [Stier et Kiss, 2013] STIER, I. et KISS, A. (2013). Cytosine-to-Uracil deamination by SssI DNA methyltransferase. *PLoS ONE*, 8(10):e79003.
- [Stockwell et al., 2014] STOCKWELL, P. A., CHATTERJEE, A., RODGER, E. J. et MORISON, I. M. (2014). DMAP : differential methylation analysis package for RRBS and WGBS data. *Bioinformatics*, 30(13):1814–1822.
- [Strahl et Allis, 2000] STRAHL, B. D. et ALLIS, C. D. (2000). The language of covalent histone modifications. *Nature*, 403(6765):41–45.
- [Sznajder et al., 2012] SZNAJDER, B., SABELIS, M. W. et EGAS, M. (2012). How adaptive learning affects evolution : reviewing theory on the Baldwin Effect. *Evolutionary Biology*, 39(3):301–310.
- [Tahiliani et al., 2009] TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L. et RAO, A. (2009). Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science*, 324(5929):930–935.
- [Tatarenkov et al., 2007] TATARENKOV, A., GAO, H., MACKIEWICZ, M., TAYLOR, D. S., TURNER, B. J. et AVISE, J. C. (2007). Strong population structure despite evidence of recent migration in a selfing hermaphroditic vertebrate, the mangrove killifish (*Kryptolebias marmoratus*). *Molecular Ecology*, 16(13):2701–2711.
- [Taylor, 1992] TAYLOR, D. S. (1992). Diet of the killifish *Rivulus marmoratus* collected from land crab burrows, with further ecological notes. *Environmental Biology of Fishes*, 33(4):389–393.
- [Taylor et al., 1995] TAYLOR, D. S., DAVIS, W. P. et TURNER, B. J. (1995). *Rivulus marmoratus* - ecology of distributional patterns in Florida and the Central Indian River lagoon. *Bulletin of Marine Science*, 57(1):202–207.
- [Taylor, 1990] TAYLOR, S. D. (1990). Adaptive specializations of the cyprinodont fish *Rivulus marmoratus*. *Florida Scientist*, 53:239–248.
- [Taylor, 2000] TAYLOR, S. D. (2000). Biology and ecology of *Rivulus marmoratus* : new insights and a review. *Florida Scientist*, 63:242–255.
- [Taylor, 2012] TAYLOR, S. D. (2012). Twenty-four years in the mud : what have we learned about the natural history and ecology of the mangrove rivulus, *kryptolebias marmoratus*? *Integrative and Comparative Biology*, 52(6):724–736.
- [Trelle et Jensen, 2007] TRELLE, M. B. et JENSEN, O. N. (2007). Functional proteomics in histone research and epigenetics. *Expert Review of Proteomics*, 4(4):491–503.
- [Turner et al., 2006] TURNER, B., T FISHER, M., SCOTT TAYLOR, D., P DAVIS, W. et KELLY, B. (2006). *Evolution of maleness and outcrossing in a population of the self-fertilizing killifish, Kryptolebias marmoratus.*, volume 8.

- [Turner *et al.*, 1992] TURNER, B. J., DAVIS, W. P. et TAYLOR, D. S. (1992). Abundant males in populations of a selfing hermaphrodite fish, *Rivulus marmoratus*, from some Belize cays. *Journal of Fish Biology*, 40(2):307–310.
- [Vandegehuchte et Janssen, 2011] VANDEGEHUCHTE, M. B. et JANSSEN, C. R. (2011). Epigenetics and its implications for ecotoxicology. *Ecotoxicology*, 20(3):607–624.
- [Varriale, 2014] VARRIALE, A. (2014). DNA Methylation, epigenetics, and evolution in vertebrates : facts and challenges. *International Journal of Evolutionary Biology*, 2014:1–7.
- [Vergeer *et al.*, 2012] VERGEER, P., WAGEMAKER, N. C. A. M. et OUBORG, N. J. (2012). Evidence for an epigenetic role in inbreeding depression. *Biology letters*, 8(5):798–801.
- [Vogt, 2015] VOGT, G. (2015). Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *Journal of biosciences*, 40(1):159–204.
- [Waddington, 1942] WADDINGTON, C. H. (1942). The epigenotype. *Endeavour*, pages 18–20.
- [Wang et Chang, 2011] WANG, K. C. et CHANG, H. Y. (2011). Molecular mechanisms of long non-coding RNAs. *Molecular Cell*, 43(6):904–14.
- [Weaver *et al.*, 2004] WEAVER, I. C. G., CERVONI, N., CHAMPAGNE, F. A., D’ALESSIO, A. C., SHARMA, S., SECKL, J. R., DYMOV, S., SZYF, M. et MEANEY, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, 7(8):847–854.
- [Wright, 2012] WRIGHT, P. A. (2012). Integrative and Comparative biology environmental physiology of the mangrove rivulus, *Kryptolebias marmoratus*, a cutaneously breathing fish that survives for weeks out of water. 52(6):792–800.
- [Yang *et al.*, 2014] YANG, X., HAN, H., DE CARVALHO, D. D., LAY, F. D., JONES, P. A. et LIANG, G. (2014). Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer cell*, 26(4):577–590.
- [Yohn et Blendy, 2016] YOHN, N. L. et BLENDY, J. A. (2016). Multigenerational and transgenerational inheritance of drug exposure : the effects of alcohol, opiates, cocaine, marijuana, and nicotine. 118(0):21–33.
- [Zemach *et al.*, 2010] ZEMACH, A., MCDANIEL, I. E., SILVA, P. et ZILBERMAN, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, 328(5980):916–919.
- [Zhang *et al.*, 2016] ZHANG, C., HOSHIDA, Y. et SADLER, K. C. (2016). Comparative epigenomic profiling of the DNA methylome in mouse and Zebrafish uncovers high interspecies divergence. *Frontiers in Genetics*, 7(JUN):1–15.

