

THESIS / THÈSE

MASTER EN SCIENCES BIOLOGIQUES DES ORGANISMES ET ÉCOLOGIE

**Nothobranchius furzeri : élevage, méthylation de l'ADN et analyse bioinformatique.
Étude préliminaire à une approche épigénétique**

Ody, Jessica

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.

Nothobranchius furzeri : élevage,
méthylation de l'ADN et analyse
bioinformatique. Etude préliminaire à
une approche épigénétique.



Année académique 2017-2018

Mémoire défendu par **Jessica Ody** dans le cadre de l'obtention du
grade de master en biologie des organismes et écologie

PROMOTEUR : FRÉDÉRIC SILVESTRE
ENCADRANT : ANTOINE WITTORSKI

19/08/2019



Table des matières

Table des matières.....	1
1. Introduction.....	3
1.1 Evolution et environnement.....	3
1.2 Epigénétique.....	6
1.2.1 Définition.....	6
1.2.2 Mécanismes épigénétiques.....	7
1.2.3 Le mécanisme de méthylation.....	8
1.2.4 La reprogrammation.....	10
1.3 <i>Nothobranchius furzeri</i>	12
1.3.1 Description morphologique.....	12
1.3.2 Phylogénie.....	13
1.3.3 Habitat et mode de vie.....	14
1.3.4 Particularité du cycle de vie : Les trois diapauses.....	17
1.3.5 Méthode d'élevage en laboratoire.....	18
1.3.6 Génome de <i>N. furzeri</i>	19
1.3.7 Intérêt de <i>N. furzeri</i> en laboratoire.....	20
1.4 Objectifs.....	20
2. Zootechnie.....	20
2.1 Système d'élevage.....	21
2.2 Milieu et « stand alone ».....	22
2.3 Gestion des souches de <i>N. furzeri</i>	24
2.4 Gestion des reproductions et des œufs.....	26
2.5 Ecllosion et élevage des jeunes poissons.....	27
2.6 Embryologie.....	28
3. Méthylation de l'ADN de <i>Nothobranchius furzeri</i>	32
3.1 Matériel et méthode.....	32
3.1.1 Echantillonnage.....	32
3.1.2 Extraction d'ADN.....	34
3.1.3 LUMA.....	35
3.2 Résultats.....	37
3.3 Discussion.....	44
3.4 Perspectives.....	47
3.5 Conclusion.....	48
4. Bioinformatique.....	49
4.1 Méthode.....	49

4.1.1	Analyses des séquences protéiques DNMT chez <i>N. furzeri</i>	49
4.1.2	Phylogénie des protéines DNMT	50
4.1.3	Analyse des îlots CpG	50
4.2	Résultats.....	51
4.2.1	Analyses des séquences protéiques DNMT chez <i>N. furzeri</i>	51
4.2.2	Phylogénie des protéines DNMT	56
4.2.3	Analyse des îlots CpG	58
4.3	Perspectives	62
4.4	Conclusion.....	62
5.	Conclusion générale.....	62
	Remerciement.....	64
	Bibliographie.....	65
	Annexes.....	77

1. Introduction

1.1 Evolution et environnement

La question de l'origine des espèces est une question fondamentale et de nombreuses réponses furent apportées depuis l'Antiquité. Durant le 4ème siècle avant J.C., Aristote et Platon furent les premiers à mettre en place une théorie grâce à l'essentialisme qui proclame que chaque chose possède une essence inaltérable, de même que les espèces vivantes présentent des caractères fixes. À partir du 2ème siècle de notre ère, la religion chrétienne est passée dans les mœurs instaurant l'idée de la création des espèces par Dieu. Ce qui eut pour effet de faire disparaître l'idée d'une évolution des espèces car tout changement signifierait que les espèces n'étaient pas parfaites comme elles furent créées (Douglas, 2013). Toujours dans cette pensée, les espèces apparaîtraient à partir de la matière inerte, toujours de la même manière et avec le même aspect. Cette origine de la vie fut appelée la génération spontanée (Brack, 1998). Cette théorie fut acceptée pendant plusieurs siècles avant que d'autres théories fassent leur apparition. Au XVIIIème siècle, Carl von Linné et Georges-Marie Leclerc mettent en place des idées opposées sur l'organisation du vivant. Alors que Linné continue sur la théorie de Platon et Aristote en classant les espèces animales et végétales dans des compartiments, Leclerc pense à une organisation plus linéaire des êtres vivants, les espèces se succédant par la reproduction en opposition à la génération spontanée. Leclerc fut ainsi clairement l'un de ceux qui influença les théories évolutionnistes (Cachel, 1986). De la même manière, Pierre Maupertuis initia l'idée que des modifications chez les parents puissent être transmises à la descendance à travers la reproduction (Bowler, 2003). De nombreuses expériences furent menées et mirent peu à peu à mal la génération spontanée qui fut définitivement abandonnée après les travaux de Pasteur au XIXème siècle (Brack, 1998).

Avec ces nouvelles pensées et les systèmes de classification, des théories évolutionnistes ont vu le jour. Apparue au début du XIXème siècle, la théorie de Jean-Baptiste de Lamarck est la première dont on va parler. Il mit en place une idée transformiste à la base de la diversification des espèces, ce qui est vu par beaucoup comme le réel début des idées évolutionnistes (Corsi, 2001). A la base de son raisonnement, Lamarck était opposé aux idées de Linné, pour lui « la nature ne connaît ni classes ni genres » dans la même continuité que Leclerc. Seuls les individus sont reconnus par la nature même s'il est possible de les regrouper en espèces via la reproduction car celle-ci permet la constance des traits entre parents et progéniture. Lamarck développe sa théorie transformiste dans *Discours d'ouverture*

et *Recherche du l'organisation des corps vivants* de 1802. Il y défend le fait que le développement des organismes est dépendant de deux facteurs : le temps et les circonstances, qu'il définit comme les influences du milieu, des relations avec le reste du biote ainsi que leurs variations. Ces circonstances permettent aux facultés ou caractéristiques des organismes de se développer, se propager et se diversifier. Ainsi Lamarck déclare que le milieu et les habitudes de vie des organismes en modifient le corps et les capacités et que ces modifications se transmettent de génération en génération.

En 1860, Charles Robert Darwin émit la théorie qui est la base de celles utilisées actuellement, la sélection naturelle. La théorie de Darwin est basée sur plusieurs principes, le premier étant la présence de descendance continue avec modifications et le suivant est la présence d'un ancêtre commun à toute espèce vivante. Les changements se font de manière graduelle et s'appliquent aux populations. Enfin, ces changements au sein des populations sont sélectionnés par la pression de l'environnement, c'est la sélection naturelle proprement dite. La difficulté de cette théorie fut d'expliquer comment ces changements se transmettent à la descendance et permettent l'adaptation des populations. Darwin résolut ce problème en avançant l'existence des gemmules qui sont des particules présentes dans les organismes parents contenant toutes leurs caractéristiques. Lors de la reproduction, les gemmules des deux parents fusionnent et créent un nouvel organisme contenant tous les changements de l'ascendance (Darwin, 1859 ; Douglas, 2013). Grâce aux travaux réalisés sur la génétique, les théories de Darwin trouvent un réel support dans la transmission des caractères via les gènes et les allèles. La fusion de la connaissance de ces supports génétiques aux théories de Darwin est communément appelée le néodarwinisme (Douglas, 2013).

D'après la théorie de Darwin, l'environnement joue un rôle important dans la sélection des phénotypes les mieux adaptés à leur environnement immédiat dans le temps et l'espace. Ce rôle de l'environnement, lorsqu'il se déroule à un niveau populationnel, peut s'appliquer à l'évolution des espèces. En effet, un des exemples les plus connus pour illustrer ce phénomène sont les pinsons de Darwin, où pour lesquels différentes populations insulaires présentent des morphologies de becs adaptées à leur alimentation. Même d'un point de vue génétique, on peut observer que la sélection du génotype tend à rendre l'individu adapté à son environnement, avec un certain niveau de polymorphisme dû aux migrations entre populations (Lenormand, 2002). Un exemple concret de la participation du génotype est celui de l'épinoche (*Gasterosteus aculeatus*) qui suivant l'environnement, marin ou eau douce, possède en plus grande majorité un allèle au locus du gène de l'ectodysplasine codant pour des plaques latérales soit réduites soit complètes (Barret *et al.*, 2008).

Mais la sélection naturelle n'est probablement pas la seule influence de l'environnement sur l'évolution des espèces. En effet, le néodarwinisme ne semble pas suffisant pour expliquer certains phénomènes observés comme des changements hérités au travers des générations. Ces changements ne semblant pas pouvoir être associés à des modifications dans la séquence ADN, il devient donc intéressant d'étudier les transmissions non génétiques (Bonduriansky et Day, 2009). Le nombre de cas de nouveaux mécanismes de transmissions sans l'implication des gènes et des allèles augmentent constamment et semblent assez répandus chez les animaux. La plupart de ces mécanismes ont été regroupés sous l'appellation d'épigénétique, ce qui signifie, étymologiquement « au-dessus de la génétique ».

Depuis l'avancée importante dans le domaine de l'épigénétique, celle-ci semble fournir un support intéressant pour une seconde influence de l'environnement sur l'évolution des espèces mais à un niveau plus individuel. Les mécanismes épigénétiques sont considérés comme le soutien physique des théories de Lamarck qui, à l'instar des théories de Darwin, ont été revisitées avec les connaissances actuelles pour devenir le néolamarckisme. Désormais, les deux théories semblent nécessaires pour expliquer la plupart des phénomènes de l'évolution (Figure 1) (Skinner, 2011 ; Skinner, 2015). Cet intérêt croissant de l'épigénétique dans l'histoire évolutive est l'une des raisons qui a motivé ce travail.

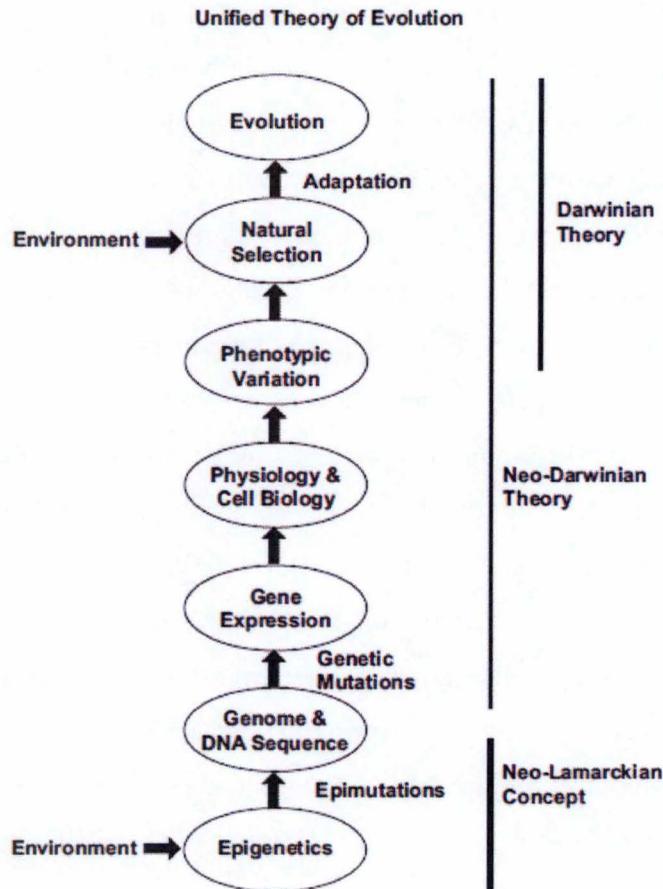


Figure 1: La théorie unifiée du néodarwinisme et du néolamarckisme (Skinner, 2015).

1.2 Epigénétique

1.2.1 Définition

L'épigénétique consiste en toutes modifications de l'expression du génome sans impacter la nature de la séquence ADN, par des changements de régulation de la transcription ou de la traduction (Holliday, 2006). A l'inverse du génome, l'épigénome, c'est-à-dire, l'état épigénétique, peut être différent pour chaque cellule selon le tissu et le niveau de développement ou l'âge de l'organisme. L'épigénétique prend une part importante lors du développement de la différenciation cellulaire étant donné que ce sont les mécanismes épigénétiques qui permettent à des cellules contenant pourtant le même génome et provenant de la division du même zygote, de se différencier en des types cellulaires différents tels que des hépatocytes et des neurones (Bird, 2007 ; Reik, 2007).

1.2.2 Mécanismes épigénétiques

Il existe de nombreux mécanismes considérés comme épigénétiques (Tableau 1). La plupart de ces mécanismes se transmettent au sein de la lignée cellulaire comme l'illustre l'inactivation d'un des chromosomes X par l'ARN xist qui se déroule lors des premières étapes de développement et reste fixé durant la vie entière de l'organisme (Lee et Jaenisch, 1997). La méthylation de l'ADN est l'un des mécanismes les plus connus qui sera détaillé par après. Ensuite, d'autres modifications de l'ADN existent comme la phosphorylation, l'acétylation qui influencent le niveau de condensation de la chromatine. Les histones (H2A, H2B, H3 et H4) peuvent également être modifiées et influencer l'état de la chromatine (Holliday, 2006 ; Keverne et Curley, 2008). Les nucléosomes (ADN et octamères d'histones) possèdent des queues N-terminal qui peuvent être modifiées à chaque acide aminé. Chaque modification mène à une réaction particulière de l'état de liaison du nucléosome, ce code extrêmement complexe est encore mal connu (Jenuwein et Allis, 2001). La paramutation est également considérée comme l'un des mécanismes d'épigénétique. Il s'agit d'un état particulier des allèles qui semble être transmissible d'un allèle paramutagène à un autre paramutable. La paramutation semble liée à la méthylation des cytosines (Brink, 1973 ; Stam et Scheid, 2005). Ces nombreux mécanismes sont souvent associés à des ARN non codant tel que les *piwi-interacting RNA* (piRNA), qui forment un complexe empêchant la transcription en association avec des protéines PIWI (Iwasaki *et al.*, 2015), et les *micro-RNA* (miRNA), de court ARN se fixant au ARNm afin d'en empêcher la traduction (Cai *et al.*, 2009).

Tableau 1: Mécanismes épigénétiques ainsi que leurs cibles et leur fonction (liste non-exhaustive).

Mécanisme épigénétique	Cible	Fonction	Littérature
Méthylation	ADN/histone	Influence la condensation de la chromatine	Holliday, 2006 ; Keverne et Curley, 2008
Phosphorylation	ADN/histone	Influence la condensation de la chromatine	Holliday, 2006 ; Keverne et Curley, 2008
Acétylation	ADN/histone	Influence la condensation de la chromatine	Holliday, 2006 ; Keverne et Curley, 2008
Paramutation	Allèle	Etat particulier transmissible entre allèles	Brink, 1973 ; Stam et Scheid, 2005.
Piwi-interacting RNA	Complexe de Transcription	Empêche la transcription	Iwasaki <i>et al.</i> , 2015
Micro-RNA	ARNm	Empêche la traduction	Cai <i>et al.</i> , 2009

La plupart de ces mécanismes semblent se transmettre au sein d'une même lignée cellulaire chez les organismes multicellulaires et permettent la plasticité développementale,

processus qui permet, à partir d'un même génotype, de modifier son phénotype en fonction de conditions environnementales (Moczek *et al.*, 2011). Mais pour relier ces mécanismes à l'évolution, il faudrait que ceux-ci soient transmissibles à travers les générations (Bird, 2007 ; Jablonka et Raz, 2009). Pour certains, il semblerait que ce soit le cas, par exemple chez *Drosophila* où une modification de l'expression d'une protéine liée à la conformation de la chromatine, est maintenue pendant les 11 générations testées (Dorn *et al.*, 1993). Une autre expérience plus récente a été réalisée chez *C. elegans*, où une inhibition de la transcription d'un gène par une injection de piRNA en F1 se poursuit jusqu'à la F20 (Ashe, 2012).

1.2.3 Le mécanisme de méthylation

La méthylation de l'ADN consiste en l'ajout d'un groupe méthyl sur les bases cytosines au niveau du cinquième carbone, soit sur les sites CpG (grande concentration en cytosines et guanines) soit sur d'autres cytosines. La méthylation des îlots CpG semble être liée à la régulation de l'expression de gènes (Holliday et Pugh, 1975 ; Riggs, 1975), en inhibant leur transcription, et pourrait même agir sur le comportement (Day et Sweatt, 2010). La concentration en CpG chez les vertébrés est inférieure à celle attendue et certaines régions présentent une plus grande concentration en CpG que d'autre sur le génome (Gardiner-Garden et Frommer, 1987). Ces régions sont appelées îlots CpG.

Le phénomène de méthylation est assez répandu chez les eucaryotes. Les champignons ainsi que les plantes et les métazoaires possèdent des régions méthylées (Suzuki et Bird, 2008). Les patterns de méthylation sont nombreux et variés, les plantes possédant parfois jusqu'à 50% de cytosines méthylées (5mC) afin de contrôler les éléments transposables. Les animaux peuvent avoir un niveau de méthylation faible comme chez *C. elegans* ou de manière plus importante pour d'autres invertébrés. La méthylation peut également être différente selon la disposition des sites méthylés, soit ceux-ci sont groupés et la méthylation est appelée globale, soit ils sont dispersés au travers du génome et la méthylation est mosaïque (Bird, 2002 ; Suzuki et Bird, 2008). La méthylation chez les vertébrés apparaît presque exclusivement sur les sites CpG et les îlots CpG avec une estimation de sites méthylés à 80% (Maegawa *et al.*, 2010). La méthylation globale chez les mammifères tend à diminuer avec l'âge tandis que la méthylation plus locale tend à augmenter sur certains gènes (Cruickshanks *et al.*, 2013).

La méthylation de l'ADN est assurée par des enzymes particulières, les *DNA methyltransferases* (DNMT). Il existe deux types de méthylation (Figure 2), soit une DNMT réalise une méthylation *de novo*, la DNMT3A et DNMT3B, soit une autre DNMT entretient la

méthylation après la division de la cellule, la DNMT1, car l'ADN fille est seulement hémiméthylé (Day et Sweatt, 2010 ; Jones, 2012 ; Suzuki et Bird, 2008), ce qui assure la transmission du pattern de méthylation entre les cellules d'une même souche (Holliday, 2006). Le taux de maintenance pour la méthylation de la cellule mère aux cellules filles des sites CpG pour des lymphocytes sur des allèles hyperméthylés, a été estimé à 96% (Laird, 2004). Cela signifie que lors de la réplication de l'ADN, 96 cytosines sur 100 qui étaient méthylées chez la cellule mère le sont toujours dans les cellules filles. Des expériences menées sur un site CpG particulier ont mis en évidence une différence de 4% entre chaque génération de cellule (Riggs, *et al.*, 1998). Il semble donc tout à fait possible que la méthylation soit une modification transmissible entre les générations de cellules (Bird 2007 ; Holliday, 2006). Cependant, lors de la prolifération cellulaire, le profil de méthylation peut fortement changer entre les cellules issues d'une même lignée, dû aux erreurs de maintenance mais surtout à la méthylation de *novo* (Bird, 2002). De ce fait, les cellules possèdent une méthylation différente pour chaque CpG même au sein du même tissu, avec certaines régions privilégiées pour la méthylation, probablement importantes pour la régulation de la transcription (Bird, 2002 ; Cedar et Bergman, 2012 ; Laird, 2004 ; Pfeifer *et al.*, 1990). La différence de méthylation peut se voir sur l'ensemble d'un organisme comme chez *Apis mellifera*. Dû à une différence d'alimentation lors du développement, les ouvrières ont un profil de méthylation global et gène spécifique différents des reines provenant pourtant du même génome (Foret *et al.*, 2012).

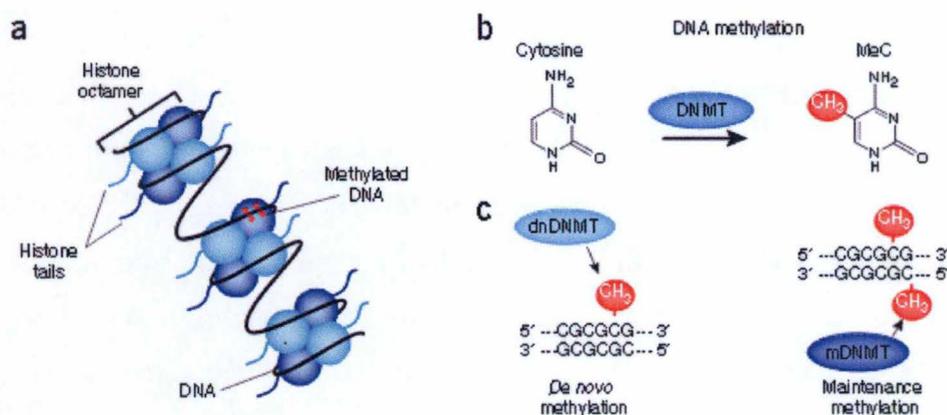


Figure 2: Méthylation d'ADN. (A) les nucléosomes sont présents au sein des cellules. (B) La méthylation est catalysée par les DNMT. (C) Deux types de méthylation existent, la méthylation de novo et celle de maintenance (Day et Sweatt, 2010).

Différents mécanismes semblent être liés pour agir sur la condensation de l'ADN (Li, 2002). Ainsi, la méthylation *de novo* semble nécessiter la présence de nucléosome, qui condense une première fois l'ADN et empêche la transcription, rendant le gène silencieux,

avant que celui-ci soit silencé par la méthylation (Jones, 2012). D'autres éléments semblent nécessaires pour cette méthylation, comme certaines protéines ou séquences ADN. La DNMT1 méthyle plus probablement les séquences particulières comme celles générées par les transposons, afin de protéger la cellule de dégâts pouvant être provoqués par ceux-ci (Bird, 2002). La présence de petits ARNs interférents corrèle avec des régions méthylées de l'ADN en plus d'une interaction avec *RNA polymerase V* chez les plantes (Matzke et Mosher, 2014). La méthylation *de novo* se déroule principalement durant les phases de développement mais peut aussi avoir lieu chez les adultes, cependant il s'agit d'un phénomène plus long car la prolifération cellulaire est moins active (Bird, 2002).

La déméthylation se réalise soit de manière passive entre chaque division en absence de DNMT (Teperek-Tkacz, 2011) soit de manière active par l'enzyme déméthylase, l'ADN glycosylase ou par oxydation des 5mC par les enzymes TET (Gu *et al.*, 2011 ; Jones, 2012 ; Kohli et Zhang, 2013 ; Morgan *et al.*, 2004 ; Rai *et al.*, 2008 ; Wu et Zhang, 2017). La déméthylation passive existe chez tous les organismes possédant des 5mC et la déméthylation active existe chez de nombreux eucaryotes, comme les plantes, les poissons et les mammifères (Gu *et al.*, 2011 ; Hajkova *et al.*, 2002 ; Jones, 2012 ; Rai *et al.*, 2008). Une grande diminution de la méthylation se réalise lors des stades précoces de développement et dans les cellules germinales (Bird, 2002 ; Jones, 1999). Par après, la méthylation *de novo* rétablit un certain niveau de méthylation. Cette période est déterminante pour le profil de méthylation de l'organisme, on parle de reprogrammation (Bird, 2002).

1.2.4 La reprogrammation

La reprogrammation contient deux phases. La première consiste en une suppression des marques épigénétiques entre deux générations afin de récupérer la totipotence du zygote et permettre la prolifération correcte des cellules durant le développement. La seconde permet la différenciation des cellules et la diminution de leur potentiel de division via une nouvelle méthylation (Cantone et Fisher, 2013 ; Huang et Fan, 2010 ; Youngson et Whitelaw, 2008).

Il existe deux moments durant lesquels se déroule la reprogrammation : lors de la fusion des gamètes chez le zygote et plus tard, lorsque les cellules germinales primordiales migrent et formeront les gamètes suivants (Cantone et Fisher, 2013).

Avant la fécondation, le niveau de méthylation est différent entre les oocytes et les spermatozoïdes. Chez *Danio rerio* par exemple, les oocytes sont méthylés entre 75 et 80% alors que les spermatozoïdes sont méthylés à 91-95% (Hackett et Surani, 2015 ; Jiang *et al.*, 2013). Durant la fusion des gamètes, il existe des différences de déméthylation des génomes

suivant leur origine. Chez *Danio rerio*, la reprogrammation est plus profonde pour le génome provenant de la femelle que pour le génome provenant du mâle. Du coup, les marqueurs épigénétiques sont principalement hérités par la lignée paternelle (Jiang *et al.*, 2013 ; O'Neill, 2013). En plus, le processus de déméthylation est différent selon l'origine du génome avec une déméthylation passive du génome maternel et une déméthylation active du génome paternel (Blewitt *et al.*, 2006 ; Gu *et al.*, 2011 ; Hackett et Surani, 2015 ; Mayer *et al.*, 2000). Ces différences selon l'origine induisent probablement un bon développement de l'embryon et la pluripotence des cellules (Cantone et Fisher, 2013). Après quelques divisions, ces différences s'effacent car le génome est essentiellement déméthylé pour permettre une grande activité transcriptionnelle nécessaire au développement de l'embryon (Cantone et Fisher, 2013 ; Huang et Fan, 2010 ; Smith *et al.*, 2012). Il est possible de visualiser par immunofluorescence l'évolution de l'état de la méthylation des cytosines qui diminue après le stade quatre cellules (Rougier *et al.*, 1998).

La deuxième phase de déméthylation se déroule chez les cellules germinales primordiales, qui formeront plus tard les gamètes de la prochaine génération. Il existe chez ces cellules des régions imprimées qui résistent à la déméthylation et permettraient une transmission de pattern de méthylation à la génération suivante (Hajkova *et al.*, 2002 ; Seisenberger *et al.*, 2012).

La reprogrammation lors du développement a principalement été étudiée chez les mammifères (Hajkova *et al.*, 2002 ; Santos et Dean, 2004), mais d'autres vertébrés ont eu leur profil de reprogrammation étudié tel que le *Kryptolebias marmoratus* où la méthylation des îlots CpG diminue vers les stades blastula et gastrula pour ensuite augmenter vers un niveau stable peu après la formation du foie (Fellous *et al.*, 2018). La reprogrammation a également été étudiée chez *Danio rerio* et *Oryzias latipes* où le niveau de méthylation était plus faible après la fécondation pour remonter lors du stade de gastrula (Jiang *et al.*, 2013 ; Maegawa *et al.*, 2010 ; Potok *et al.*, 2013 ; Wang et Bhandari, 2019). *O. latipes* et *K. marmoratus* présente des pourcentages de méthylation faible du zygote par rapport à *D. rerio* (22.91% et 27.8% contre 83% respectivement) témoignant d'une reprogrammation plus profonde. Il existe une reprogrammation induite *in vitro* (Jaenisch et Yobung, 2008). Cette technique a été mise au point afin de recréer des cellules pluripotentes à partir de cellules somatiques et ainsi mieux comprendre la reprogrammation.

Certaines influences de l'environnement, agissant sur les parents, peuvent impacter les générations futures (Benayoun *et al.*, 2015 ; Dias et Ressler, 2014 ; Sultan, 1996 ; Painter *et al.*, 2008 ; Tobi *et al.*, 2014). Chez les souris, un contact avec une molécule olfactive

l'acétophenone, qui traduit la présence de prédateurs, associé à un conditionnement chez les parents peut induire une reconnaissance chez la progéniture sans contact préalable (Dias et Ressler, 2014). Chez l'organisme modèle *C. elegans*, des marqueurs épigénétiques âge-dépendent semblent être transmis entre les générations, comme la longévité, la gestion du stress oxydatif et des maladies neurodégénératives (Benayoun *et al.*, 2015). Ces modifications accumulées au cours de la vie des parents ne peuvent être transmises que par des mécanismes d'origine épigénétique car les influences de l'environnement agissent lorsque les organismes sont matures et ces influences ne peuvent alors agir sur le génome. Si la transmission existe, alors des régions de l'épigénome doivent résister aux reprogrammations (Youngson et Whitelaw, 2008). Il existe beaucoup de régions résistantes à la reprogrammation comme les centromères (Rougier *et al.*, 1998), des transposons et certaines régions imprimées tel que le chromosome X (Reik, 2007).

1.3 *Nothobranchius furzeri*

1.3.1 Description morphologique

Le killifish turquoise, *Nothobranchius furzeri*, est un poisson mesurant en moyenne 5 cm et présentant un dimorphisme sexuel comme les autres représentants du genre (Figure 3). Les mâles sont en général plus grands que les femelles, avec une moyenne de 4,9 cm pour les premiers et 4,3 cm pour les secondes (Polacik *et al.*, 2011), et plus colorés. D'ailleurs, La différenciation des espèces se fait principalement par la différence entre les mâles. Il existe deux colorations différentes chez les mâles selon la souche. Les mâles GRZ du Zimbabwe (Figure 3B) possèdent des écailles gris bleuté avec des bords rouge ainsi qu'une bande de couleur jaune et une noire sur la queue. Les mâles de l'autre souche MZM du Mozambique (Figure 3C) ont une queue de couleur rouge (Dorn *et al.*, 2011).

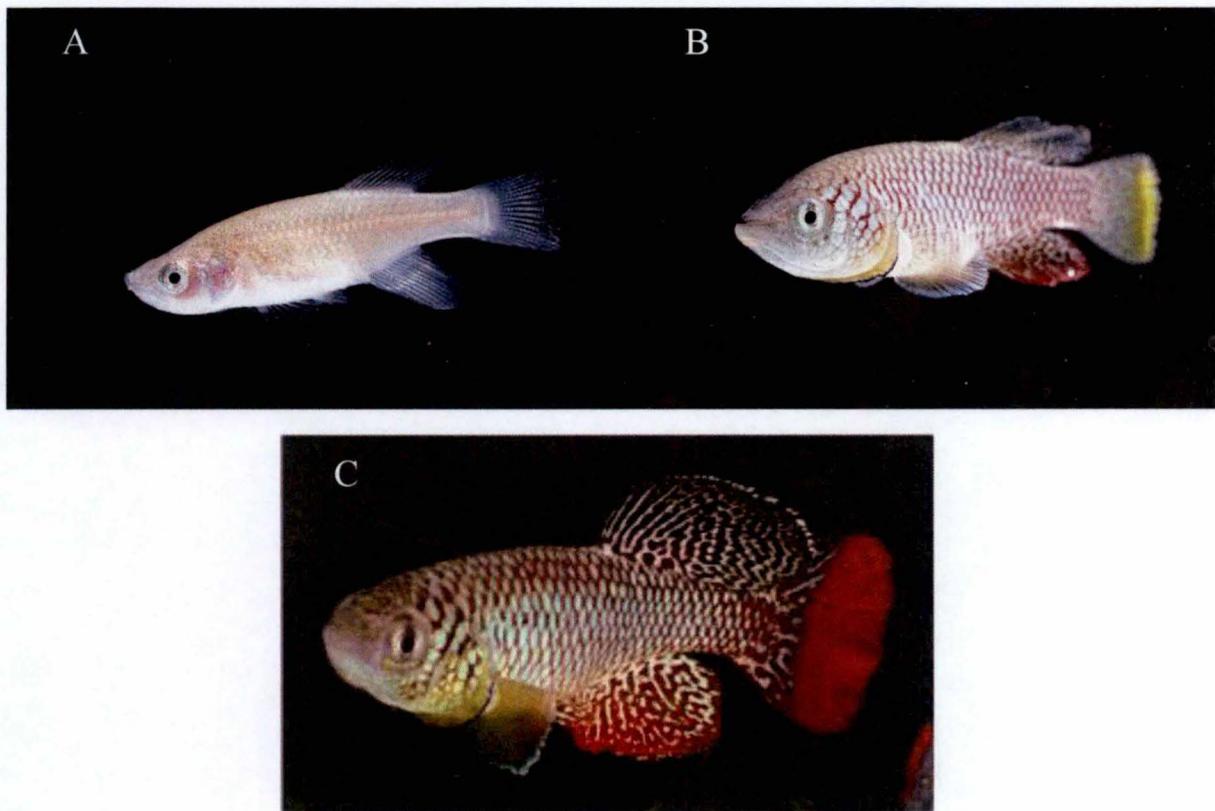


Figure 3 : Femelle *N. furzeri* (A) et mâle *N. furzeri* (B) aux environs de 4 semaines après l'éclosion (crédit photo : F. Silvestre). Mâle MZM (C) (Polacik et al., 2016).

1.3.2 Phylogénie

Nothobranchius furzeri est un actinoptérygien de l'ordre des Cyprinodontiformes (Genade et al., 2005 ; Costa, 2018), dans la famille Nothobranchiidae (Furness et al., 2015a), au sein de laquelle 289 espèces sont validées selon le catalogue des espèces de la *California academy of science* (Fricke et al., 2018), et sont souvent désignés comme poissons annuels. Il existe de nombreuses espèces au sein du genre *Nothobranchius* constamment revisités (Figure 4) (Costa, 2018 ; Dorn et al., 2011) avec 75 espèces identifiées (Froese et al., 2018). Les genres les plus proches semblent être *Pronothobranchius* et *Fundulosoma*, qui sont très proches physiquement et possèdent le même type de cycle de vie (Costa, 2018 ; Genade et al., 2005). Ces espèces présentant le même style de cycle de vie court que *N. furzeri*, semblent proches dans la séquence de gène apparemment lié à la longévité et au développement, suggérant une pression de sélection sur ces gènes (Sahm et al., 2017).

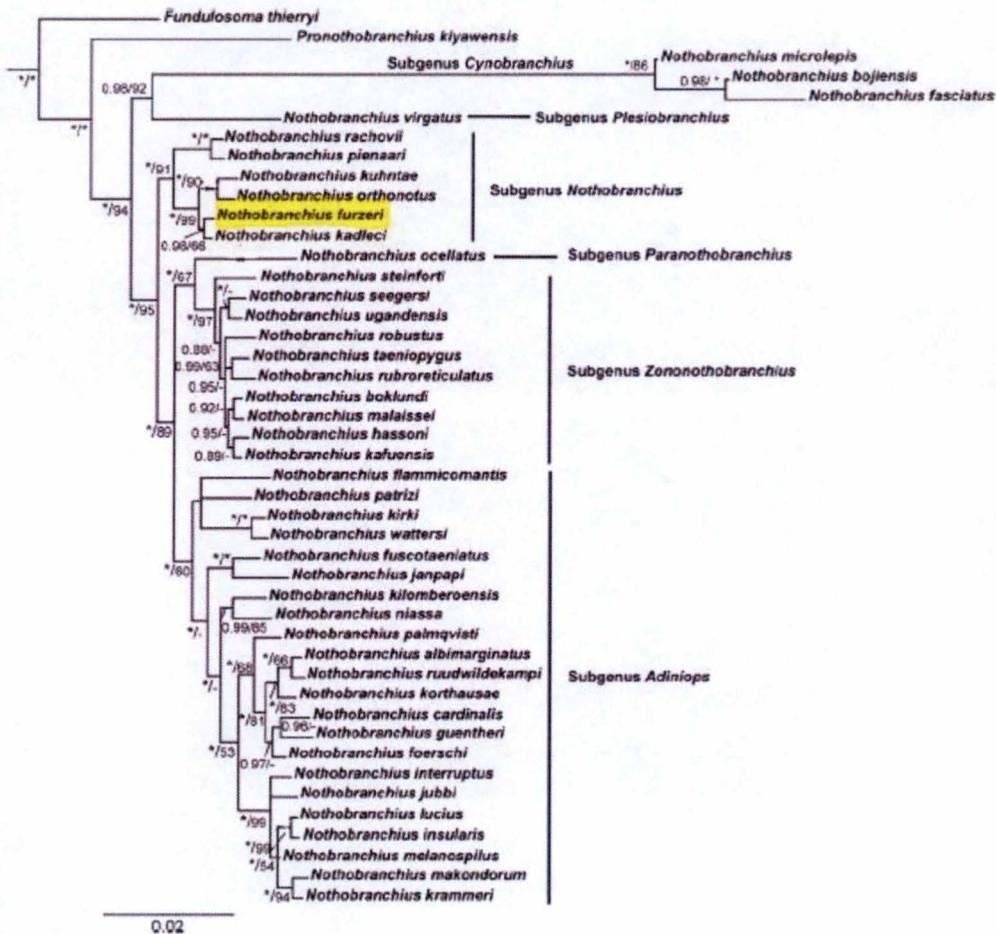


Figure 4: Arbre phylogénétique de 45 espèces de Nothobranchiini selon une analyse Bayésienne sur base de 138 caractères morphologique et un total de 3327 pb de gènes nucléaires. Les nombres indiquent la valeur de probabilité postérieure et les résultats du bootstrap supérieurs à 50%. Les symboles * et – correspondent respectivement au support values Maximum et aux valeurs de bootstrap inférieures à 50% (d'après Costa, 2018).

La souche élevée en laboratoire provient de spécimens collectés au Zimbabwe dans le parc de Gona Re Zhou, d'où le nom de la souche GRZ, en 1968 (Jubb, 1971) et depuis maintenu en élevage sans presque aucune hybridation (Genade *et al.*, 2005), ce qui résulte en une population captive presque complètement homozygote (Reichwald *et al.*, 2015 ; Valenzano *et al.*, 2015). Récemment, de nouveaux échantillons ont été réalisés entre 2004 et 2012 au Mozambique où d'autres populations ont été recensées telle que celle présentant la queue rouge (Reichard *et al.*, 2009 ; Cellerino *et al.*, 2016).

1.3.3 Habitat et mode de vie

Nothobranchius furzeri vit dans des bassins d'eau temporaires, dans le sud-est de l'Afrique entre le Soudan et l'Afrique du Sud (Genade *et al.*, 2005). L'espèce est décrite dans le bassin du Sazale dans la réserve du Gona Re Zhou au Zimbabwe, qui est drainé par quelques rivières avant de couler dans le bassin de la rivière Chefu qui finit en terrain

La présence et la disposition du réseau hydraulique est dépendante de la saison des pluies et des précipitations, qui peut varier entre 200 et 600 mm/an (Cellerino *et al.*, 2016). Dû à son habitat relativement temporaire, *Nothobranchius furzeri* possède une espérance de vie aux alentours de six mois (Tozzini *et al.*, 2013), mais certaines souches en captivité vivent entre 9 et 12 semaines après l'éclosion (Genade *et al.*, 2005 ; Valdesalici et Cellerino, 2003). Récemment de nouveaux spécimens échantillonnés au Mozambique possèdent une espérance de vie de 13 semaines, supportant la théorie qu'une si courte espérance en captivité est sûrement dû à la faible hétérogénéité de la population prélevée en 1968 (Terzibasi *et al.*, 2007). Cependant, la souche GRZ qui est l'une des souches ayant une espérance de vie de 9 à 12 semaines, est également une souche vivant dans les plus hautes altitudes, des régions plus arides où les précipitations sont plus hasardeuses (Cellerino *et al.*, 2016). La courte espérance de vie serait liée à ces conditions plus difficiles. Du fait de cette faible longévité, dans les régions proches de l'équateur, deux générations se succèdent par an puisqu'en quelques mois, les poissons d'une première génération ont éclos, ont pu se reproduire et une deuxième génération succède à la première (Costa, 2018).

Les espèces de *Nothobranchius* sont généralistes, près de 11 catégories d'invertébrés ont été retrouvés dans leur alimentation, avec une grande abondance de crustacés. Il ne semble pas que les espèces aient une proie de prédilection cependant, au sein d'un même bassin, différentes espèces présentent des niches d'alimentation différentes, se répartissant ainsi les différentes proies. *N. furzeri* semble cependant préférer les petits crustacés tels que les copépodes (Polacik et Reichard, 2010).

En milieu naturel, il n'est pas rare que les mâles soient moins nombreux que les femelles avec en moyenne un mâle pour 3 femelles, pouvant aller jusqu'à 10 femelles pour un mâle sur certains sites (Reichard *et al.*, 2009). Le mâle chez *N. furzeri* est très agressif envers les autres mâles et les femelles qui refusent de pondre. Lorsque le mâle rencontre des femelles, il réagit souvent en voulant se frotter à celles-ci afin de les stimuler à pondre. Si la femelle est réceptive, le mâle va alors l'entourer de ses nageoires dorsale et anale pendant quelques secondes et c'est après cela que la femelle pond un unique œuf (Cellerino *et al.*, 2016). L'agressivité des dominants peut être réduite dans les grands groupes (Genade *et al.*, 2005). Les femelles peuvent pondre dans des substrats boueux dès 18 jours après l'éclosion et jusqu'à 50 œufs par jour vers six semaines après l'éclosion, le taux de fécondité augmentant aussi avec l'âge avec plus d'un œuf sur deux fécondé après 6 semaines (Blažek *et al.*, 2013 ; Cellerino *et al.*, 2016 ; Reichard *et al.*, 2009). Si ces poissons peuvent se reproduire si rapidement, c'est grâce à un développement accéléré et non pas à cause d'une maturation

prématurée des gonades chez les larves (Blažek *et al.*, 2013 ; Genade *et al.*, 2005). De nombreux biomarqueurs dénotent le vieillissement rapide de *N. furzeri*, telles que la capacité de mouvement, l'expression de β -galactosidase lié à la vieillesse, la présence de lipofuscine (association de biomolécules fluorescent à 360-470nm) et la neurodégénérescence (Terzibasi *et al.*, 2007).

Les œufs ont la possibilité de se mettre en diapause ce qui leur permet de survivre à la saison sèche contrairement aux adultes (Reichard *et al.*, 2009 ; Wourms, 1972b). Il existe des différences de moment d'éclosions des œufs entre et au sein de populations (Polacik *et al.*, 2011). Ces différences au sein d'une population peuvent être due à une asynchronisation du développement des différents œufs. L'éclosion semble se dérouler au milieu de la saison des pluies plutôt qu'au début, probablement pour éviter une possible dessiccation après les premières pluies.

1.3.4 Particularité du cycle de vie : Les trois diapauses

De nombreux autres killis possèdent cette capacité de diapause afin de survivre à la dessiccation de leur environnement comme *Austrofundulus limnaeus* (Furness *et al.*, 2015a, Naumann et Englert, 2018). Durant son développement, l'œuf de *N. furzeri* passe par trois diapauses afin de survivre à l'assèchement de son environnement et perpétuer l'espèce (Genade *et al.*, 2005).

La diapause est constituée de trois phases. Après une phase de dispersion, les cellules embryonnaires se regroupent et l'embryon entre dans la première diapause qui se déroule lors du stade de gastrulation. L'œuf entre dans cette première phase lorsqu'il est présent dans les boues du bassin encore rempli d'eau et elle est causée par l'hypoxie du milieu. La deuxième diapause est celle qui dure le plus longtemps et pendant laquelle l'œuf est le plus résistant. Elle se déroule après la formation des somites et permet à l'œuf de résister à la sécheresse. Durant la diapause II, le métabolisme est fortement diminué, dont la synthèse protéique qui diminue de plus de 93% (Podrabsky et Hand, 2000). La consommation d'oxygène est également très faible, diminuant de 90% (Cellerino *et al.*, 2016 ; Furness *et al.*, 2015a). Lors de la troisième diapause, l'organogenèse est terminée et l'embryon est prêt à éclore. On appelle ce stade *golden eyes* dû à la couleur dorée que prennent les yeux de l'embryon. L'œuf éclore pendant la saison des pluies lorsque le niveau d'eau sera suffisant (Furness *et al.*, 2015a ; Genade *et al.*, 2005 ; Wourms, 1972b).

Il a été observé que des variations dans des paramètres tels que la photopériode ou la température durant le développement des œufs pouvaient influencer le temps de

développement des embryons (Furness *et al.*, 2015b). Ces mêmes paramètres influencent également le passage par la diapause II ou un développement direct. Par exemple, si la température est de 20°C, tous les œufs entrent en diapause II. Cette température correspond à celle de la saison des pluies et est perçue par l'embryon comme signal pour entrer en diapause II. Cela va permettre aux embryons d'attendre la saison des pluies prochaine afin de finir correctement leur développement. Cela laisse supposer que les œufs sont capables de percevoir le changement de saison et d'adapter leur développement afin de mieux répondre à leur environnement.

1.3.5 Méthode d'élevage en laboratoire

Il existe trois systèmes d'élevage des *N. furzeri* mise en place en laboratoire décrit dans la littérature. La première consiste en un grand bac commun avec environ dix poissons pour 20L avec deux ou trois mâles avec sept ou huit femelles (Genate *et al.*, 2005 ; Genade, 2016 ; Valdesalici et Cellerino, 2003 ; Graf *et al.*, 2009). Ce système est celui qui correspond le plus aux conditions naturelles (Cellerino *et al.*, 2016). Les deux autres systèmes d'élevages sont sensiblement similaires entre eux et ne diffèrent que dans la manière de gérer les reproductions. Dans les deux cas, les aquariums sont du type « stand- alone ». Le circuit est fermé, ce qui permet une excellente homogénéisation et maîtrise des conditions physico-chimiques du milieu. Chaque bac est séparé des autres avec un système de renouvellement de l'eau des aquariums en passant par trois filtres successifs avant d'être redistribuée, ce qui diminue la quantité de déchets au sein des aquariums et facilite l'entretien. Tous les paramètres sont maîtrisables et modifiables à tout moment. Ensuite la reproduction peut être gérée de deux manières différentes. La première peut être comparée à du « speed-dating ». Tous les matins, un mâle et une femelle sont prélevés et placés dans un aquarium isolé avec un fond sableux pour une période de deux heures (Blazek *et al.*, 2013 ; Polačik *et al.*, 2016). Cette période va permettre au mâle de stimuler la femelle à pondre et de féconder les œufs pondus. Au bout des deux heures, les poissons sont remplacés dans leur bac respectif et les œufs récoltés en filtrant le sable. Cette méthode est très efficace pour un suivi accru de la reproduction. Pour chaque poisson, il est en effet possible de faire un suivi exact du nombre d'œufs pondus, fécondés et qui arrivent jusqu'à l'éclosion. La deuxième méthode de reproduction consiste en la création de « harems » (Grégoir *et al.*, 2017 ; Hartmann et Englert, 2012). Des aquariums sont consacrés à la reproduction avec un mâle pour deux à trois femelles constamment présent ensemble dans l'aquarium. La production des œufs fécondés se fait alors de manière constante dans un bac de sable présent dans l'aquarium. Les œufs sont

récoltés de manière journalière. Les harems restent ainsi formés jusqu'à la mort naturelle des poissons.

1.3.6 Génome de *N. furzeri*

Dû au développement grandissant de l'utilisation de *N. furzeri*, l'intérêt de caractériser et de séquencer son génome c'est renforcé. Le génome de *N. furzeri* est constitué de 19 paires de chromosomes ($2n=38$). Un premier séquençage du génome de *N. furzeri* a donné une longueur approximative de 1,5Gb (avec une taille maximale entre 1,6 et 1,9Gb), ce qui est plus grand que les génomes du médaka, du poisson zèbre, du rivulus des mangroves et de l'épinoche (Reichwald *et al.*, 2009). Les séquences du génome découlant du séquençage peuvent être trouvées sur le site <http://nfingb.leibniz-fli.de/>. En se basant sur l'estimation de 1,5Gb de longueur, un nouvel assemblage du génome s'est déroulé en 2015 et la séquence totale assemblée a été de 1,24Gb (Reichwald *et al.*, 2015). Une autre étude a mené un assemblage du génome de la même souche GRZ et a réussi à assembler un génome d'une longueur de 1,02 Gb, avec une taille maximale comprise entre 1,3 et 2,2 Gb (Valenzano *et al.*, 2015). De cet assemblage, 28 494 gènes codant pour des protéines et 5859 séquences de gènes correspondant à des ARN non codants ont pu être identifiés. Le génome de *N. furzeri* contient également énormément de régions répétées avec 45% du génome constitués de séquences répétées (Reichwald *et al.*, 2009). Pour la partie séquencée, le pourcentage de GC est de 44,9% avec des séquences particulièrement riches en GC, avec des pourcentages de plus de 60%. Un pourcentage de GC élevé est souvent lié à un génome plus petit (Costantini *et al.*, 2007). Le génome de *N. furzeri* est l'un des plus grands par rapport à celui du tétraodon (220Mb), du médaka (725Mb), du poisson zèbre (1,5Gb), du rivulus des mangroves (915Mb) et de l'épinoche (352Mb). Pourtant le pourcentage en GC chez *N. furzeri* est l'un des plus élevés, proche de celui de tétraodon qui possède un petit génome (Costantini *et al.*, 2007 ; Kelley *et al.*, 2016). L'explication la plus plausible est l'importance des répétitions qui sont riches en GC et assez fréquentes dans le génome de *N. furzeri* (Reichwald *et al.*, 2009).

Certains mécanismes épigénétiques ont été étudiés chez *N. furzeri*. Principalement l'acétylation et la méthylation des histones ont été étudiées comme la triméthylation de H3K27 (Cencioni *et al.*, 2019 ; Platzer et Englert, 2016 ; Zupkovitz *et al.*, 2018). Des outils de manipulation génétique ont été adaptés pour *N. furzeri* tel que la technologie CRISPR-cas9 (Platzer et Englert, 2016). Peu d'études ont été réalisées sur les gènes codants pour les protéines DNMT excepté pour leur expression (Api *et al.*, 2017) et autres gènes impliqués dans la méthylation de l'ADN et la reprogrammation.

1.3.7 Intérêt de *N. furzeri* en laboratoire

La première raison de l'utilisation de *Nothobranchius furzeri* en laboratoire vient de sa courte espérance de vie de seulement trois mois pour certaines souches, la plus courte découverte chez les vertébrés. Il est devenu alors un modèle de choix dans l'étude du vieillissement chez les vertébrés. Cette espèce présente d'autres intérêts, comme l'étude des tumeurs ou des impacts sur le transgénérationnel (Api *et al.*, 2017 ; Platzer et Englert, 2016). La publication de génomes de référence ainsi que d'outils de manipulation génétique renforcent l'intérêt de l'utilisation de *N. furzeri* en laboratoire. Les intérêts d'étudier l'espèce ne se limitent pas à des perspectives médicales. En effet, *N. furzeri* présente également un intérêt d'un point de vue écologique avec un développement rapide et la présence de diapauses permettant une survie des œufs lors des périodes de sécheresse (Blazek *et al.*, 2013). Ces particularités peuvent impliquer des régulations épigénétiques durant le développement de *N. furzeri*.

1.4 Objectifs

L'objectif général de ce mémoire a pour but l'étude de la reprogrammation épigénétique du méthylome chez *Nothobranchius furzeri*, dont les mécanismes sont encore inconnus, afin de mettre en évidence une potentielle influence du profil de reprogrammation sur le développement et le vieillissement du *N. furzeri* ainsi que ces implications dans l'évolution de *N. furzeri*.

Le premier objectif spécifique du travail consiste en la mise au point d'un élevage de *N. furzeri* de la souche GRZ.

Le deuxième objectif est de caractériser les différents stades embryonnaires de *N. furzeri*. Cette caractérisation permettra de définir les stades auxquels le méthylome sera étudié.

Le troisième objectif spécifique consistera en l'évaluation du niveau de méthylation global chez les adultes ainsi que celui des stades embryonnaires, précédemment définis, afin de définir le profil de reprogrammation durant le développement rapide du *N. furzeri*.

Le quatrième objectif sera de caractériser les acteurs DNMT de la méthylation chez *N. furzeri* sur base des données disponibles dans les bases de données informatiques.

2. Zootechnie

Ce chapitre présente le travail zootechnique effectué lors de ce mémoire. Bien qu'hors des sentiers battus dans le cadre d'un mémoire, cette étape s'avère être cruciale et déterminante pour la suite du projet. Ayant nécessité une grande rigueur, elle mérite quelques

lignes descriptives. En effet, le projet d'étude sur *N. furzeri* au sein du laboratoire LEAP de Frédéric Silvestre est très récent, dès lors au début du mémoire l'installation zootechnique pour *N. furzeri* n'en était qu'au stade embryonnaire. Il a donc été nécessaire de mettre en place toutes les modalités techniques essentielles à l'élevage.

2.1 Système d'élevage

Pour la mise en place de notre élevage, nous avons testé les trois systèmes d'élevage des poissons. Le premier consistant en un grand bac commun (Blazek *et al.*, 2013 ; Genate *et al.*, 2005 ; Genade, 2016 ; Valdesalici et Cellerino, 2003 ; Graf *et al.*, 2009). Cette méthode s'avère intéressante lorsque, comme dans notre cas, le facteur limitant principal est l'espace. Néanmoins, l'inconvénient majeur est l'impossibilité de faire un suivi précis des poissons. Nous n'avons utilisé cette méthode que par nécessité lorsque les éclosions étaient fructueuses. Dans notre cas, en suivant cette méthode, nous avons remarqué que le nombre d'œufs pondus par jour était très faible, voire nul sans pour autant en connaître la raison. Sans doute ce genre de système nécessite un décor riche que nous ne possédions pas. De plus, dans notre cas, ce système ne possédait pas de système de filtration et les déchets s'accumulaient très rapidement.

Les deux autres systèmes d'élevages testés sont les aquariums du type « stand-alone » avec un circuit est fermé. Dans notre cas, le principal bémol de cette installation est la place nécessaire pour la mettre en œuvre. En effet, chaque poisson est élevé dans un bac d'environ 3 litres sans contact visuel avec leurs congénères. Les deux techniques de gestion de reproduction ont été testé. Pour la première, nous avons prélevés tous les matins un mâle et une femelle que nous plaçons dans un berlin de 3 litres sur fond sableux pour une période de 2 heures (Blazek *et al.*, 2013 ; Polačik *et al.*, 2016). Au bout des 2 heures, nous remplaçons les poissons dans leur bac respectif et récoltons les œufs. Cette méthode est très efficace pour un suivi accru de la reproduction. Pour chaque poisson, il est en effet possible de faire un suivi exact du nombre d'œufs pondus, fécondés et qui arrivent jusqu'à l'éclosion. Cependant, cette méthode reste stressante pour le poisson, cela représente beaucoup de manipulations et ne peut donc être efficace que si nous possédons un nombre suffisant de mâles et de femelles que pour pouvoir faire une rotation. Notre système étant composé de seulement 12 bacs, pour obtenir suffisamment d'œufs nous devons procéder à deux reproductions par jour par poisson ce qui représentait beaucoup trop de manipulations. La deuxième méthode en « harems » (Grégoir *et al.*, 2017 ; Hartmann et Englert, 2012) a été testée en dédiant des aquariums à la reproduction. Dans notre cas ces bacs étaient de 6L dans lesquels nous placions un mâle pour

deux à trois femelles de la même génération. Les harems restaient ainsi formés jusqu'à la mort naturelle des poissons. Nous avons déposés dans ces bacs une boîte de sable d'environ 10x10cm pour la ponte et un tube de PVC pour permettre aux femelles d'échapper au mâle parfois très agressif (Figure 6). Cette méthode est celle que nous avons finalement adoptée car elle présente l'avantage principal de ne quasiment pas manipuler les poissons, seuls les boîtes de sables étant récupérées quotidiennement pour la récolte des œufs. Avec cette méthode de reproduction, nous avons observé une meilleure production d'œufs.

2.2 Milieu et « stand alone »

Le système d'aquarium dans lequel sont élevés les poissons est un stand-alone (Figure 6). La température de l'eau est maintenue à $27 \pm 1^\circ\text{C}$ dans un cycle photopériodique de 12/12. Nous utilisons une eau reconstituée en ajoutant du sel (InstantOcean®) à de l'eau milliQ jusqu'à obtention d'une dureté de $600\mu\text{S}$. Une fois dans le circuit, l'eau est filtrée selon 3 méthodes :

- Mécanique : Deux systèmes de filets sont mis en place. Un manchon « chaussette » directement relié à l'arrivée d'eau dans le bac filtre permettant de retenir les débris volumineux. Une plaque de mousses micro perforée d'environ 4cm permettant une filtration plus efficace des résidus alimentaires plus petits et fèces des poissons.
- Organique : Après les filtrations mécaniques, nous avons installé un compartiment rempli de céramique, support poreux idéal pour le développement des bactéries dénitrifiantes.
- Physique : Enfin, avant d'être ré-envoyé dans le système, l'eau passe par un filtre UV permettant d'éliminer les éventuelles bactéries résiduelles.

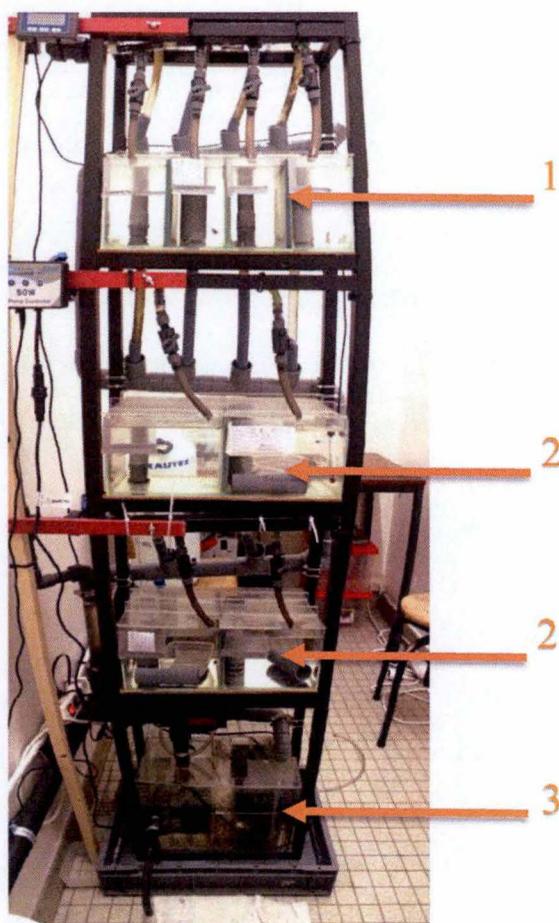


Figure 6: Système stand-alone avec les bacs individuels de 3L (1), les bacs de reproduction contenant un bac de sable et un tuyau de PVC pour l'enrichissement (2) et le bac des trois filtres et de la pompe (3).

Deux fois par jour en semaine et une fois le week-end, les poissons sont nourris *ad libitum*. Dans sa publication de 2016, Polacik recommande l'utilisation de vers de vase (*Chironomus* spp.). Les vers de vases, en plus d'être riches en lipides, ont l'avantage d'être rouge vif ce qui est attractif pour le poisson. Nous avons utilisé cette espèce mais très rapidement nous avons constaté ses limites dans notre système. Ces larves vivent enfouies dans la vase, elles sont donc incapables de nager en pleines eaux. Dans nos bacs, elles coulent dans le fond, dès lors en voulant l'attraper, le poisson se trouve dirigé tête en bas et perçoit son reflet dans le fond du bac ce qui l'amène à adopter une position d'attaque. C'est un évènement stressant qui pourrait avoir comme impact une diminution de la prise de nourriture. Dans les bacs de reproduction, les larves s'enfouissent dans le sable de reproduction, empêchant ainsi les poissons de bien les attraper. De plus, les vers de vases vivent en eau salée, s'ils ne sont pas mangés rapidement, ils meurent et polluent le bac. Dès lors, nous avons essayé une autre espèce, les larves de moustiques blanches (*Chaoborus* sp.). Celles-ci ont l'inconvénient d'être blanches et donc moins visibles mais ce sont des larves nageuses et vivent dans la colonne

d'eau, ce qui est très stimulant pour le poisson qui peut donc les chasser. Elles sont également très riches en lipides et vivent en eaux douces ce qui leur donne l'avantage de survivre dans l'eau de l'aquarium (environ 12 heures). Celles-ci créent donc moins de déchet, sont tout aussi nutritives et bien plus stimulantes.

2.3 Gestion des souches de *N. furzeri*

Après avoir validé le bon fonctionnement de notre système, nous avons reçu 50 œufs de la souche GRZ_AD (Gonarezhou National Park, Zimbabwe) du laboratoire d'écologie, évolution and Conservation de la biodiversité de l'Université catholique de Louvain (KUL). 25 de ces œufs ont été mis à l'éclosion et 25 ont été stockés à 17°C sur de la fibre de noix de coco à un taux d'hygrométrie d'environ 80%. Dans ces conditions, les embryons entrent en diapause et constitue un stock de réserve. Sur les 25 éclos, seuls six poissons ont atteint l'âge adulte : trois femelles et trois mâles. Seul un mâle était en bonne santé et ne présentait pas de malformations (Figure 8). Malgré tous nos efforts, très peu d'œufs ont été pondus et quasiment aucun n'a atteint le dernier stade embryonnaire pré-éclosion (stade *golden eyes*). Nous avons donc recommencé une 2^{ème} éclosion avec notre réserve (25 oeufs au stade de diapause I) et seul un mâle a atteint l'âge adulte. Ce premier essai avec une survie de la génération F0 à 12% et un taux de reproduction atteignant les 0% était bien plus faible que ceux décrit par d'autres élevages de *N.furzeri* (Genate *et al.*, 2005, Polačik *et al.*, 2016). Après discussion avec nos collaborateurs de la KUL, il s'avère qu'ils rencontrent exactement les mêmes problèmes avec la même génération, laissant penser à un problème de consanguinité, plutôt qu'un problème de gestion d'aquarium et de technique d'éclosion. Lorsque l'on travaille sur un petit système, une attention toute particulière doit être apportée à la gestion des reproductions. En effet, la première faiblesse du système de reproduction en « speed-dating » ou en « harem » est la consanguinité très prononcée. Au fil des générations, un *bottleneck* se fait ressentir, se traduisant par l'apparition de maladies, d'un taux de fécondation faible, d'œufs en mauvais état etc. (Figure 7).



Figure 7 : Comparaison d'œufs non fécondés et infectés (F1) de la souche GRZ KUL (F0) (Gauche) avec un œuf non fécondé (Milieu) et d'un œuf fécondé (Droite) de la même génération. On peut observer dans tous les œufs infectés, la présence d'un champignon au sein du chorion (flèches rouges).

Nous avons ensuite reçu 78 œufs au stade *golden eyes* de la même souche GRZ_AD provenant du laboratoire du Docteur Valenziano au *Max planck institute* (MPI). L'avantage au MPI c'est qu'ils possèdent une très grande installation, il leur est donc plus simple de contrer les problèmes de consanguinité. Après être passé par toutes les étapes nécessaires à l'éclosion et la gestion des juvéniles, dont nous discuterons plus loin, hors de 78 œufs, nous avons obtenu 31 *N. furzeri* adultes, 12 femelles et 19 mâles en parfaite santé qui nous ont permis de démarrer notre élevage. Ce deuxième essai avec un taux de survie de 39,7% fut bien plus concluant dans la mise en place de l'élevage des *N. furzeri*. Ce nombre de femelles nous a permis de mettre en place quatre aquariums de reproduction, contenant chacun trois femelles et un mâle. Dans chaque aquarium, la production journalière d'œufs était entre 20 et 50 œufs pondus, ce qui est moins qu'attendu selon la littérature (Blazek *et al.*, 2013 ; Cellerino *et al.*, 2016).



Figure 8: Mâle adulte (gauche) et femelle (droite) GRZ-AD de la KUL en bonne santé.

2.4 Gestion des reproductions et des œufs

La maturité sexuelle chez *N. furzeri* est observable aux environs de quatre semaines post éclosion par une coloration qui apparaît sur les nageoires des mâles. Ce n'est qu'à ce moment précis que les harems sont formés, un mâle pour deux à trois femelles. *N. furzeri* possède une parade nuptiale qui consiste à coincer la femelle sous sa nageoire dorsale, basculer sur le flanc et à secouer vigoureusement sa nageoire caudale (Blazek *et al.*, 2013). Cette petite danse stimule la femelle pour qu'elle dépose ses œufs sur un substrat meuble (*e.g.* : du sable) que le mâle viendra féconder ensuite. La période de reproduction dure toute la vie adulte du poisson mais présente un pic d'activité et un taux d'œufs fécondés et viables supérieurs entre la 3^{ème} et la 6^{ème} semaine de vie (Blazek *et al.*, 2013 ; Cellerino *et al.*, 2016).

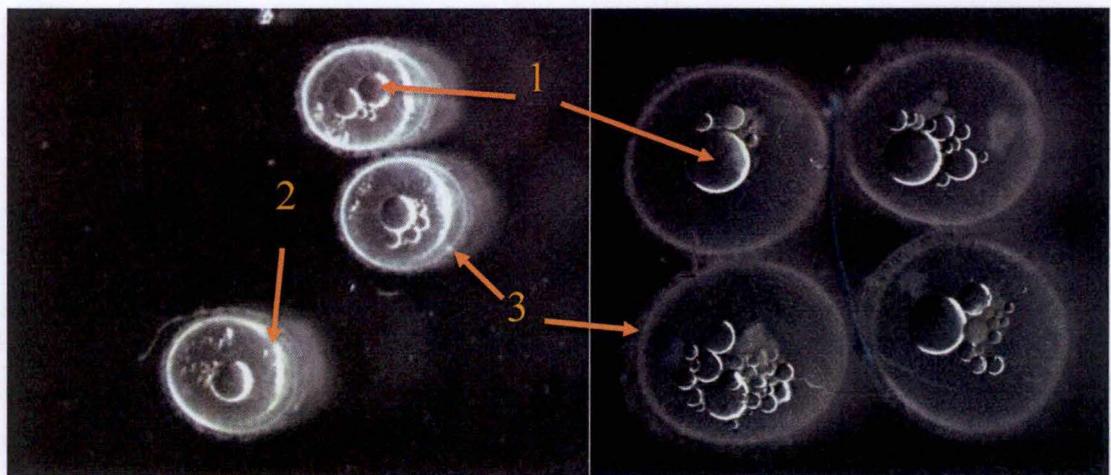


Figure 9: Œufs fécondés (à gauche) et non fécondés (à droite). Dans les œufs, on peut observer les gouttelettes lipidiques (1), le sac vitellin (2) le tout entouré du chorion (3).

Les œufs sont récoltés tous les jours. Le bac de sable est tamisé avant d'être remis

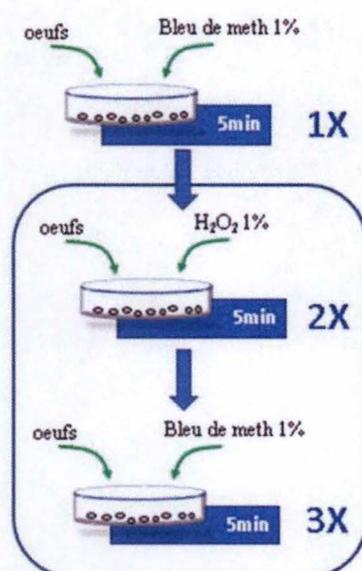


Figure 10 : Protocole de bleaching utilisé. Le premier bain de bleu de méthylène (*meth.*) correspond au tri des œufs fécondés et non-fécondés.

dans les bacs de reproduction et les œufs sont ensuite triés dans une solution de Bleu de Méthylène. Lors du tri, les œufs non fécondés sont reconnaissables car ils ne possèdent pas de double membrane, la membrane du chorion et la membrane du sac vitellin en plus de prendre une teinte bleutée par rapport aux œufs fécondés (Figure 9). Les œufs morts apparaissent complètement opaques voir jaunâtres. Après ce premier tri, il est nécessaire de les nettoyer par bleaching (Figure 10). Le bleaching consiste en deux bains successifs dans une solution d'H₂O₂ à 1% pendant 5 minutes et trois bains dans une solution de bleu de méthylène à 100µL/L. Le bleaching permet

de limiter la propagation de moisissures sur les œufs et d'éliminer les éventuels parasites et pathogènes.

Une fois le bleaching terminé, les œufs sont stockés dans un bain de bleu de méthylène à 27°C pendant plus ou moins une semaine. Quotidiennement, la solution de bleu de méthylène est renouvelée et les œufs morts sont éliminés. Au bout d'environ une semaine, les embryons ont atteint le stade dit « black eyes » facilement reconnaissable par les deux grosses taches noires que sont les yeux sur la tête de l'embryon (Figure 12J). Les embryons à ce stade sont transférés sur de la fibre de noix de coco humide dans une boîte de Pétri à raison d'une cinquantaine d'œufs par boîte. Ils seront également placés à 27°C jusqu'à la fin de leur développement embryonnaire. Ici aussi les boîtes sont contrôlées quotidiennement afin d'éliminer les œufs morts. Au bout d'environ trois semaines post-fécondation, nous pouvons apercevoir à travers le chorion, un poisson juvénile bien formé enroulé sur lui-même. Ce juvénile possède deux gros yeux brillants, c'est le stade « golden eyes » (Figure 12K). Le stade golden eyes est le stade pré-éclosion. A ce stade le poisson juvénile est capable de rentrer en diapause, c'est la diapause III. Il ralentit son rythme métabolique et peut attendre quelques semaines que les bonnes conditions soient réunies pour éclore. Lorsqu'un bon nombre d'embryons est au stade golden eyes, les œufs sont récoltés pour passer à l'étape éclosion. Les embryons qui ne sont pas prêts sont simplement redéposés à 27°C

2.5 Eclosion et élevage des jeunes poissons

Le protocole d'éclosion est principalement basé sur la publication de Polacik *et al.* (2016) et adapté après une formation auprès des laboratoires de la KUL et du MPI ainsi qu'auprès de Charlotte Philippe. L'éclosion s'effectue dans une solution d'acide humique 0,33g/L à 4°C. Une solution mère d'acide humique d'une concentration 1g/L est préparée et conservée au frigo. Cette solution est diluée 3x avec de l'eau du système pour constituer le milieu d'éclosion qui lui aussi sera gardé au frigo. Les œufs au stade *golden eyes* sont déposés dans une cuve d'environ 10x20 cm, un volume du milieu d'éclosion est versé par-dessus pour atteindre une hauteur de maximum 1cm. Cette faible hauteur est très importante car dès l'éclosion les juvéniles doivent être en mesure de rejoindre la surface pour remplir leur vessie natatoire. Si la distance à parcourir pour atteindre la surface est trop longue, ils risquent soit de mourir d'épuisement soit ne pas pouvoir remplir leur vessie natatoire et donc ensuite d'être incapable de flotter entre deux eaux, ces poissons sont nommés « belly-swimmer » et sont incapables de se reproduire. Utiliser un milieu à 4°C permet de créer un choc thermique qui

pourrait être équivalent à celui qu'engendrerait la pluie dans le milieu naturel, stimulant l'éclosion.

Le bac est ensuite stocké dans l'incubateur à 27°C avec un flux d'air conséquent afin de fournir une agitation du milieu. Au bout de 24h les poissons ont éclos, il est dès lors important de diminuer le flux d'air, *N. furzeri* n'étant pas un grand nageur il préférera une eau plus calme (Genade *et al.*, 2005). Pendant les cinq jours suivant l'éclosion, le milieu sera quotidiennement siphonné des déchets et dilué 2x avec de l'eau du système (600 µS, 27°C). Les juvéniles sont nourris 2x par jour *ad libitum* avec des nauplies d'artémias fraîchement écloses (Figure 11) (*Artemia salina*).

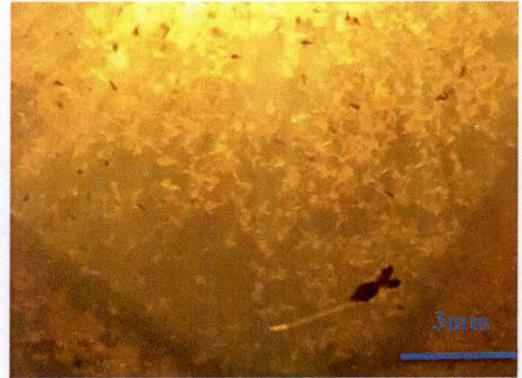


Figure 11: juvénile *N. furzeri* avec des nauplies d'artémias juste après l'éclosion des poissons.

Durant les premiers jours, le nombre d'artémias n'est pas très élevé, une dizaine par poisson et nourrissage, mais l'extrême rapidité de croissance du *N. furzeri* impose d'augmenter rapidement la quantité administrée. Il est possible d'observer *de visu* la satiété des juvéniles car le ventre des jeunes poissons attrape une couleur orange visible lorsqu'ils ont correctement mangé. Après une semaine, les jeunes poissons sont suffisamment grands pour être installés dans le système *stand-alone* avec un maximum de cinq poissons dans un volume de 6 litres. A ce stade, il est encore impossible de déterminer le sexe. Les poissons sont toujours nourris deux fois par jours avec des artémias en plus de quelques larves de moustiques afin de les habituer le plus tôt possible à un nouveau régime alimentaire. Au bout de deux semaines post-éclosion, seules les larves de moustiques blanches leur seront administrées, toujours à raison de deux fois par jour. Lorsque les premières couleurs apparaissent sur les nageoires des mâles, il est nécessaire de répartir les poissons. Soit un mâle pour plusieurs femelles dans un bac de 6L, soit un poisson mâle par bac de 3L afin d'éviter les agressions parfois très violentes entre deux congénères, pouvant causer la mort d'un des individus. Dès ce moment, les poissons seront gardés en bacs individuels ou seront placés dans des bacs de reproduction et ce, jusqu'à leur mort naturelle.

2.6 Embryologie

La partie suivante est consacrée à la description des stades embryologiques observés chez *N. furzeri*. Le développement du *N. furzeri* passe par des étapes similaires au développement chez le poisson comme *K. marmoratus* (Mourabit *et al.*, 2011) et *Danio rerio*

(Kimmel *et al.*, 1995). Les principaux stades définis sont le zygote, les premières divisions, la blastula suivie de la gastrula, de la segmentation/neurulation, le stade *black eyes* et le stade *golden eyes* (Api *et al.*, 2017 ; Hartmann et Englert, 2012). La différence de *N. furzeri* avec les autres espèces de poissons modèles est la présence de certains stades comme les diapauses et le stade de dispersion lors de l'épibolie 100% (Naumann et Englert, 2018). Les nombreuses photographies illustrent les stades les plus importants du développement de manière non-exhaustive (Figure 12). Elles ont été réalisées via le programme NIS-D par un appareil photographique Nikon USB3 1/2.5 15 IM/SEC sur un microscope Nikon SMZ1270.

Le zygote apparaît après la fusion des deux gamètes, après la pénétration du spermatozoïde dans l'ovule (Figure 12A). L'œuf présente deux couches, la couche extérieure d'apparence plus épaisse est le chorion et la couche interne est la membrane du sac vitellin, séparées par l'espace périvitellin. Le point de pénétration du spermatozoïde est visible ainsi que de nombreuses gouttelettes lipidiques. Le stade une cellule, appelée blastodisque (Figure 12B) est visible environ deux heures après la fécondation sur la membrane vitelline dans l'espace périvitellin. Le stade deux cellules apparaît après la première division qui se déroule à 4 heures post-fécondation (hpf), 2h après l'apparition du blastodisque, et les deux cellules conjointes sont visibles dans l'espace périvitellin (Figure 12C). Le stades quatre cellules apparaît après une division perpendiculaire (Figure 12D). Le stade huit cellules présente des cellules de taille égale, le stade 16 cellules laisse apparaître deux couches de cellules de tailles différentes.

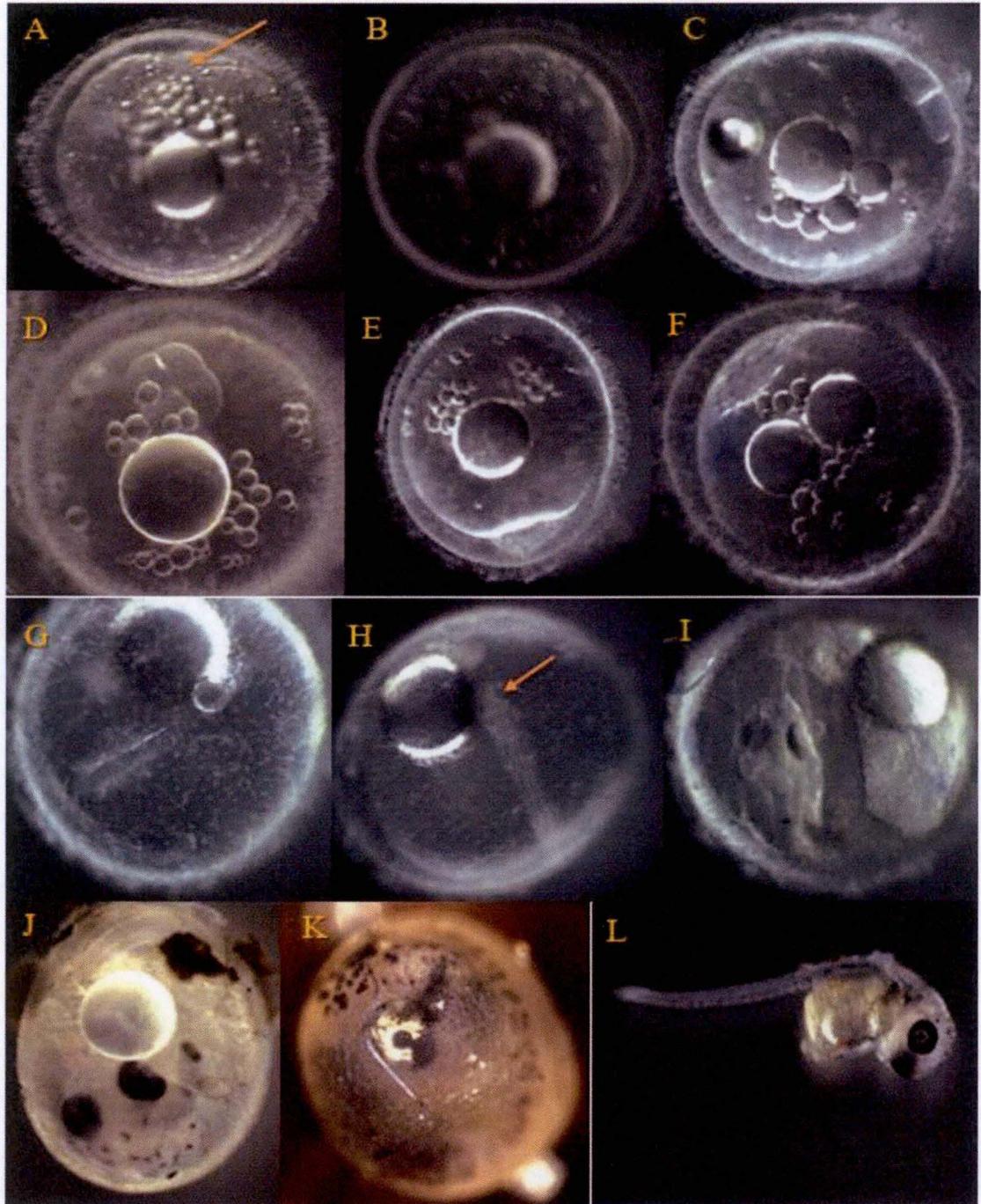


Figure 12: Grandes étapes du développement embryonnaire observés chez *N. furzeri*. Pénétration du spermatozoïde avec la visualisation de la zone d'ancrage du spermatozoïde (A), Blastodisque (B), stade 2 cellules (C), stade 4 cellules (D), blastula (E), gastrulation et début d'épibolie (F), neurulation et diapause II (G), vésicules optiques et somites (H), début d'organogenèse avec apparition du battement cardiaque et de la pigmentation des yeux et de la peau (I), organogenèse presque complète qui correspond au stade black eyes (J), stade pré-éclosion avec l'organogenèse complète et apparition de la pigmentation claire des yeux qui correspond au stade golden eyes et à la diapause III (K), alevin à l'éclosion (L).

La cinquième division monte le nombre de cellules à 32 (stades qui n'ont pas été observés chez les œufs issus de notre élevage), le nombre de cellules continue à augmenter au

fur et à mesure des divisions mais leur nombre n'est plus discernable. Le stade blastula est atteint après à 24hpf (Figure 12E) et est reconnaissable par l'aplatissement du disque cellulaire. Après ce stade, les cellules continuent de se diviser et commencent l'épibolie, une migration cellulaire des blastomères profonds conduisant d'une part au premier stade de diapause et ensuite au commencement du stade de gastrulation (Figure 12F) (Naumann et Englert, 2018). A ce stade d'épibolie, il est possible d'observer une des différences entre les poissons non-annuels tels que *K. marmoratus* et *D. rerio* avec *N. furzeri* (Blazek *et al.*, 2013 ; Cellerino *et al.*, 2019). Chez *N. furzeri*, lorsque l'embryon atteint les 100% d'épibolie, deux à trois jours après la fécondation (Api *et al.*, 2017 ; Hartmann et Englert, 2012), les cellules sont dispersées sur toute la surface du sac vitellin sans organisation apparente avec aucun axe embryonnaire visible, les cellules n'entrant que rarement en contact les unes avec les autres (Cellerino *et al.*, 2019 ; Hartmann et Englert, 2012 ; Kimmel *et al.*, 1995 ; Mourabit *et al.*, 2011). Contrairement aux espèces non-annuelles où l'embryon est organisé autour d'un axe embryonnaire et les cellules rassemblées. La deuxième différence de développement avec les poissons non-annuels peut être observée à ce stade correspondant au moment de la première diapause de l'embryon (Naumann et Englert, 2018 ; Wourms, 1972). Dans notre cas, le stade de dispersion a pu être observé mais sans passer par le stade de diapause car l'embryon n'arrête son développement qu'en cas d'hypoxie du milieu, lorsque celui-ci est enfui dans le substrat de ponte (*i.e.* : sable), et nous mettions nos œufs dans les bonnes conditions afin qu'ils continuent leur développement. Ainsi, l'embryon esquive cette diapause et commence sa réaggrégation le jour d'après, formant alors l'axe embryonnaire à partir duquel la gastrulation prend place (Api *et al.*, 2017 ; Cellerino *et al.*, 2019 ; Hartmann et Englert, 2012). Lorsque la gastrulation est finie, on retrouve les étapes clés du développement embryonnaire du poisson que sont la neurulation et la segmentation (Figure 12G) avec les régions associées à la tête et la queue qui deviennent visibles. La diapause II se déroule à ce stade somitique précoce (Cellerino *et al.*, 2016 ; Furness *et al.*, 2015a ; Naumann et Englert, 2018 ; Podrabsky et Hand, 2000 ; Wourms, 1972b). En laboratoire, les conditions y sont établies afin d'éviter cette diapause pour permettre aux embryons d'esquiver cette diapause et de continuer leur développement. En même temps que la formation des somites, on peut observer l'apparition des vésicules optiques (Figure 12H) de part et d'autre du cerveau. Les yeux ainsi que la peau se pigmentent et les premiers battements cardiaques apparaissent, c'est le début de l'organogenèse (Figure 12I). Lorsque les yeux sont entièrement noirs et visibles à l'œil nu, il s'agit du stade de *black eyes* durant lequel l'organogenèse touche à sa fin (Figure 12J). Le stade suivant est le stade de *golden eyes* (Figure 12K), l'organogenèse ainsi que la

pigmentation larvaire sont complètes. C'est durant ce dernier stade avant l'éclosion « hatching-stage » (Figure 12L) que se déroule la diapause III. Cette diapause est caractérisée par la diminution du nombre de battements cardiaques et des mouvements de l'embryon, ainsi qu'une utilisation faible voire nulle des réserves vitellines (Wourms, 1972b).

Si la durée intrinsèque des différents stades de développement ne varie pas, le temps post-fécondation peut être extrêmement variable entre les embryons. Cela est dû aux différents stades de diapauses possibles ainsi que la durée qu'ils passent dans ces états. Les facteurs expliquant l'entrée ou la sortie des différentes diapauses ne sont pas encore complètement maîtrisés et la notion de facteur « nécessaire » ou « suffisants » est encore floue. Certains facteurs tels que les conditions environnementales comme la température ou la lumière ainsi que l'âge des parents peuvent ainsi influencer l'entrée en diapause mais également le temps de la diapause (Api *et al.*, 2017 ; Furness *et al.*, 2015b ; Wourms, 1972b). Les techniques de laboratoires nous permettent d'obtenir un taux d'éclosions sans de passages par les diapauses assez élevé. Cependant, les œufs restent parfois en diapauses même si les conditions semblent réunies pour les éviter et inversement dû à la stratégie du « *bet-hedging* » (Furness *et al.*, 2015b).

La particularité de développement qu'est la phase dispersée des cellules chez *N. furzeri* peut être observé chez d'autres espèces comme *Austrofundulus limnaeus* (Cleaver *et al.*, 2017), un autre killi vivant en Amérique du Sud qui possède un cycle de vie annuel ainsi que la possibilité d'entrer en diapause. Comme cette particularité se retrouve chez ces deux poissons annuels mais pas chez *D. rerio* et *K. marmoratus*, il est possible que cette absence d'organisation soit liée au cycle annuel de vie.

3. Méthylation de l'ADN de *Nothobranchius furzeri*

L'un des objectifs de ce travail est d'étudier le niveau de méthylation global de *N. furzeri* ainsi que de permettre une visualisation de la reprogrammation lors du développement de *N. furzeri*. Afin de permettre de remplir ces objectifs, la technique du LUMA : LUMinometric Methylation Assay (Head *et al.*, 2014) sera utilisée.

3.1 Matériel et méthode

3.1.1 Echantillonnage

Afin d'étudier la reprogrammation de la méthylation de l'ADN chez *N. furzeri*, des œufs à différents stades de développement ont été récoltés. Au vu des difficultés d'élevage,

seulement quatre stades ont été échantillonnés : le stade blastodisque, qui correspond au stade une cellule (B), le stade blastula (Ba), le stade *black eyes* (BE) et le *golden eyes* (GE).

Afin de récolter les stades blastodisque et blastula, les œufs étaient récoltés deux fois sur une même journée, au matin juste avant le nourrissage et début d'après-midi, ainsi les œufs qui se trouvaient dans le substrat de reproduction n'étaient âgés que de maximum quelques heures. Pour le stade blastodisque, les œufs de quelques heures étaient triés, ceux possédant une cellule étaient sélectionnés, ceux sans cellule étaient laissés au développement jusqu'à ce qu'une cellule soit visible et ceux plus développés étaient stockés à 28°C pour atteindre le stade *black eyes*. Pour le stade blastula, les œufs étaient aussi récoltés l'après-midi et laissés 24h à l'incubateur, temps correspondant à l'apparition du stade blastula après la fécondation.

Juste après la sélection, les œufs ont été congelés par immersion dans de l'azote liquide à une température de -196°C, puis conservés au congélateur à -80°C jusqu'à obtenir un nombre suffisant d'œufs pour les analyses (Tableau 2). Les œufs supplémentaires ont été utilisés pour les premiers tests et ajustement. Le nombre d'œufs utilisés par réplica est différent entre les stades. Celui-ci a été ajusté par rapport au développement de l'embryon, plus celui-ci est développé, plus le nombre de cellules est important et la concentration en ADN est élevée. De cette manière, nous obtenons plus d'ADN pour les stades plus précoces tout en évitant une trop forte concentration dans les stades plus tardifs.

Tableau 2: Œufs récoltés et utilisés par stade dans les analyses LUMA

Stade	Blastodisque	Blastula	Black eyes	Golden eyes
Nombre maximum	224	200	80	57
Nombre œufs par réplica	32	32	9	9
Nombre de réplica	5	5	5	5
Nombre total utilisé	160	160	45	45

Des poissons adultes de *N. furzeri* ont été utilisés afin de réaliser les analyses sur des poissons matures. Dix poissons femelles de cinq et dix poissons femelles de neuf semaines ainsi que dix poissons mâles de cinq et dix poissons mâles neuf semaines ont été élevés sans contact visuel entre les différents sexes. Les mâles étaient séparés dans des aquariums individuels et les femelles étaient élevées par groupe de quatre. Tous ont été nourris exclusivement d'artémias *ad libitum* deux fois par jour. Ces poissons ont été euthanasiés une fois l'âge désiré atteint et disséqués endéans les cinq minutes pour récupérer les gonades

directement congelées à l'azote liquide et stockée à -80°C . Ces organes ainsi que ces poissons ont été pesés et sont repris en annexe (Annexe 1). Trois mâles de 124 jours ont également été élevé sans contact visuel et nourris avec *Chaoborus* sp. deux fois par jour *ad libitum*. Les poissons ont été euthanasiés et disséqués endéans les cinq minutes. Le foie, le cerveau ainsi que les gonades ont été prélevés, congelés à l'azote liquide et stocké à -80°C . Ces trois mâles ont été utilisé pour évaluer la méthylation chez les poissons très âgés.

3.1.2 Extraction d'ADN

L'extraction d'ADN a été réalisée avec le kit NucléoSpin tissue XS de chez Macherey-Nagel. Possédant un chorion résistant, les œufs doivent être préalablement écrasés avant de réaliser la lyse de cellules car le tampon de lyse B1 et la protéinase K sont incapables de dégrader le chorion. La lyse des cellules c'est fait à 56°C pendant 2 heures 30, avec 5 secondes de vortex toutes les 30 minutes. Cette période de temps est la durée minimum pour que plus aucun tissu ne soit visible dans les échantillons d'organes. Les volumes des solutions ont été doublé à $160\mu\text{L}$ pour les tampons B1 et B3 ainsi que la quantité d'éthanol pure. Chaque échantillon est élué deux fois, chacun dans $20\mu\text{L}$ du tampon d'élution afin de récupérer un maximum d'ADN de la colonne de silice. Vu le nombre élevé d'échantillons, plusieurs extractions ont été nécessaire, toutes basées selon le protocole décrit précédemment.

L'extraction de chaque organe s'est réalisée séparément. Le nombre d'œufs par réplica par stade de développement est repris dans le Tableau 2. Afin de tester une technique d'extraction pouvant récolter une plus grande quantité d'ADN, l'extraction des répliques 3, 4 et 5 des stades blastodisques et blastulas ont été divisés en quatre échantillons plus petits de huit œufs. Les éluions de chaque sous-échantillon ont été rassemblés en échantillons plus grands de $80\mu\text{L}$ et concentrés par évaporation au Speedvac (Thermo Fisher), avec un volume final d'environ $20\mu\text{L}$. Le Speedvac consiste en une centrifugation des échantillons afin de rassembler les molécules au fond du tube puis la solution s'évapore sous vide sans emporter d'ADN.

Juste après l'extraction, une analyse des échantillons par NanodropTM (Thermo Fisher Scientific NanoDrop 2000c Spectrophotometer) a été appliquée afin d'évaluer la quantité en acide nucléique et l'ampleur des contaminations. L'état de dégradation de l'ADN a été évalué par migration sur gel d'agarose (1%, 100V pendant 30 minutes, avec SYBR Safe comme agent intercalant). Après cela, les échantillons ont été conservés au congélateur -20°C .

3.1.3 LUMA

Pour évaluer la méthylation globale lors du développement de *N. furzeri*, nous avons utilisé la technique LUMA : LUminometric Methylation Assay (Head *et al.*, 2014), basée sur la méthode définie par Karimi *et al.* (2006).

La première étape de la technique du LUMA consiste en une restriction de l'ADN par HpaII ou MspI. Chaque échantillon nécessite deux restrictions, une par HpaII et une par MspI. Chaque réaction contient 20µl avec 450µg d'ADN et 3µl de mix réactionnel contenant 2µl de Tango Buffer, 0,5µl de MunI 2,5U et soit 0,5µl de HpaII 5U, soit 0,5µl de MspI 5U. Chaque réaction est laissée à l'incubation pendant 4 heures à 37°C sans agitation. Durant ces 4 heures, les enzymes coupent le même site de restriction, C|CGG mais HpaII étant sélective, elle ne scinde pas si la cytosine intérieure est méthylée. MunI permet de normaliser les résultats obtenus par HpaII et MspI, et ainsi rend les résultats indépendants de la quantité d'ADN utilisée.

Après la réaction de restriction, 15µl d'Annealing buffer sont ajoutés aux réactifs et mélangé par pipetage afin de stopper la réaction de restriction. 30µl de chaque solution est prélevé et déposé dans la plaque de pyroséquençage. Une fois la plaque chargée, celle-ci est placée dans le pyroséquenceur Pyromark Q24 (Qiagen). La cartouche de pyroséquençage est chargée selon les indications fournies par le programme informatique Pyromark Q24 (Qiagen). Lors de la réaction réalisée dans le pyroséquenceur, les fragments de restriction sont complétés par insertion de nucléotides, dans l'ordre défini au préalable et présenté dans le Tableau 3. Lors de l'insertion de ces nucléotides, deux pyrophosphates (PPi) sont libérés et deux enzymes, la sulfurylase et la luciférase, transforment ces Ppi en ATP puis en lumière mesurée par le pyroséquenceur (Karimi *et al.*, 2006).

Tableau 3 : Séquence d'injection des nucléotides du LUMA et leur utilité.

1	G	Mesure le bruit de fond et les fragments d'ADN non générés par restriction
2	T	
3	G	
4	T	
5	C	Complète les G des restriction HpaII et MspI et les T des restriction MunI, ainsi que les fragments non générés par restriction
6	A	
7	C	
8	A	
9	T	T complète les séquences générées par MunI et G complète les séquences générées HpaII et MspI
10	G	
11	T	Mesure le bruit de fond et que toutes les incorporations se sont déroulées correctement
12	G	
13	T	
14	G	

La séquence d'injection de nucléotide dans les puits est importante (Tableau 3), on utilise les valeurs des insertions 9 et 10 pour mesurer le taux de méthylation car les insertions 5 et 6 peuvent comprendre des insertions non spécifiques. La formule ci-dessous (Equation 1) est utilisée afin de calculer le pourcentage de méthylation.

$$1 - \left[\frac{\left(\frac{HpaII(G)}{MunI(T)} \right)}{\left(\frac{MspI(G)}{MunI(T)} \right)} \right] \times 100$$

Equation 1 : Calcul du pourcentage de méthylation.

Lorsque des répliques techniques ont pu être fait, une moyenne est faite entre deux rapports HpaII/MunI ou MspI/MunI de chaque réplique puis la moyenne est utilisée dans le calcul de méthylation. Il est indispensable de réaliser un contrôle qualité de la restriction et de vérifier la significativité du signal enregistré par le pyroséquenceur, car il faut que celui-ci soit différenciable du bruit de fond. Ce contrôle est réalisé en comparant la hauteur du pic G10 avec la moyenne des pics G1 et G3 (Equation 2). Idéalement, la hauteur des pics G1 et G3 doivent être proche de 0 et donc, le rapport de leur moyenne avec la hauteur de G10 doit être élevé. En pratique, il existe toujours du signal correspondant au bruit de fond. La valeur minimum de ce rapport est de 6,3, assurant ainsi une différenciation suffisante du signal avec le bruit de fond.

$$\frac{HpaII(G10)}{HpaII(\text{moy}(G1, G3))}$$

Equation 2 : Calcul du contrôle qualité G10/Gbf (bruit de fond).

Les différences des niveaux de méthylation entre les différents échantillons sera évaluer par analyse statistique ANOVA comparaisons deux à deux.

3.2 Résultats

Les valeurs obtenues lors de l'analyse des extractions au Nanodrop™ sont présentés dans le tableau 4.

Tableau 4 : Concentrations en ADN des échantillons et les rapports 260/280 (contamination en protéines) et 260/230 (contamination en solvant). Blastodisque (B), Blastula (Ba), Black eyes (BE) golden eyes (GE) et des échantillons de gonades. L'ajout d'un prime après le nom de l'échantillon signifie qu'il s'agit de l'élution 2 du même échantillon. Les valeurs recommandées des rapports 260/230 et 260/280 sont situées entre 1,8 et 2 (Desjardins et Conklin, 2010).

Échantillon	Concentration µg/µl	$\frac{260}{280}$	$\frac{260}{230}$	Échantillon	Concentration µg/µl	$\frac{260}{280}$	$\frac{260}{230}$
B1	43,3	1,6	0,28	BE5	237,5	1,84	1,9
B2	27,9	1,58	0,45	BE5'	96,7	1,86	1,94
B4	43,7	1,46	0,37	BE3	165	1,84	1,75
B4'	10,6	1,62	0,48	BE3'	61,3	1,84	1,67
B5	52,6	1,46	0,46	GE1	234,7	1,82	1,85
B5'	60,5	1,43	0,61	GE2	376,7	1,83	1,99
B3	27,9	1,65	0,23	GE4	172,1	1,83	1,98
B3'	76,7	1,17	0,47	GE4'	105,4	1,86	1,96
Ba1	71,9	1,6	0,26	GE5	231,7	1,8	1,84
Ba2	65,1	1,64	0,24	GE5'	274,1	1,78	1,76
Ba4	29,6	1,56	0,41	GE3	433,2	1,8	3,2
Ba4'	3,9	2,13	0,46	GE3'	455,4	1,82	1,98
Ba5	31,7	1,54	0,36	Cerveau1	69,7	1,85	2,33
Ba5'	5,4	1,75	0,45	Cerveau2	34,6	1,84	1,82
Ba3	33,7	1,66	0,28	Cerveau3	47,8	1,88	2,31
Ba3'	11,7	1,75	0,36	Cerveau3'	43,5	1,93	2,09
BE1	360,3	1,84	1,79	Gonade1	776,5	1,83	2,31
BE2	236,4	1,81	1,8	Gonade2	282,1	1,82	2,19
BE4	121,3	1,83	1,63	Gonade3	464,5	1,82	2,29
BE4'	92,8	1,86	1,84	Foie1	1009,3	1,76	1,19
				Foie2	595,2	1,78	1,59
				Foie3	332,7	1,77	1,15
15M	30,5	1,71	1,19	15M'	36,6	1,71	1,19
35M	56,8	1,72	1,17	35M'		1,72	1,17
45M	19	1,55	0,66	45M'	25,4	1,55	0,66
55M	19,8	1,79	1,19	55M'	10,1	1,79	1,19

65M	12,1	1,60	1,08	65M'	27,7	1,60	1,08
75M	43,2	1,39	1,59	75M'	12,2	1,39	1,59
85M	40	1,65	1,32	85M'	8,1	1,65	1,32
95M	28,3	1,60	1,19	95M'	20,8	1,60	1,19
105M	53,1	1,68	1,08	105M'	85,7	1,68	1,08
75F	308,3	1,61	0,62	75F'	696,5	1,61	0,62
15F	54,3	1,70	1,15	15F'	51,7	1,70	1,15
25F	59,7	1,78	0,79	25F'	96,2	1,78	0,79
35F	86,3	1,75	1,25	35F'	66,2	1,75	1,25
45F	69,7	1,81	1,38	45F'	82,5	1,81	1,38
55F	267,7	1,81	1,27	55F'	70,3	1,81	1,27
65F	66,9	1,78	1,45	65F'	35,2	1,78	1,45
85F	147,2	1,78	1,32	85F'	188	1,78	1,32
95F	123,6	1,80	1,49	95F'	50,1	1,80	1,49
105F	103,3	1,80	1,40	105F'	39	1,80	1,40
115F	173,9	1,82	1,70	115F'	57,3	1,82	1,70
19M	103,8	1,81	1,84	19M'	49,9	1,81	1,84
29M	74	1,77	1,65	29M'	37,8	1,77	1,65
39M	78,3	1,48	0,89	39M'	26	1,48	0,89
49M	54,8	1,53	0,81	49M'	9,5	1,53	0,81
59M	1880	1,45	0,57	59M'	41,6	1,45	0,57
69M	39,9	1,53	1,57	69M'	103,5	1,53	1,57
79M	39,9	1,45	0,86	79M'	98,8	1,45	0,86
89M	147,4	1,72	2,28	89M'	67,3	1,72	2,28
99M	110,8	1,77	2,27	99M'	23,6	1,77	2,27
109M	262,8	1,82	1,92	109M'	25,2	1,82	1,92
19F	240,6	1,80	1,49	19F'	12,5	1,80	1,49
29F	432,3	1,81	1,35	29F'	95,3	1,81	1,35
39F	209,4	1,80	1,30	39F'	337,6	1,80	1,30
49F	448,7	1,79	1,18	49F'	140,5	1,79	1,18
59F	26,1	1,67	1,18	59F'	324,1	1,67	1,18
69F	127,3	1,69	1,19	69F'	323,9	1,69	1,19
79F	147,1	1,78	1,09	79F'	34,3	1,78	1,09
89F	578,9	1,74	0,89	89F'	43	1,74	0,89
99F	482,9	1,87	1,49	99F'	259,5	1,87	1,49
109F	216,5	1,77	1,16	109F'	508,7	1,77	1,16
1bis9F	129,8	1,73	1,07	1bis9F'	82,6	1,73	1,07



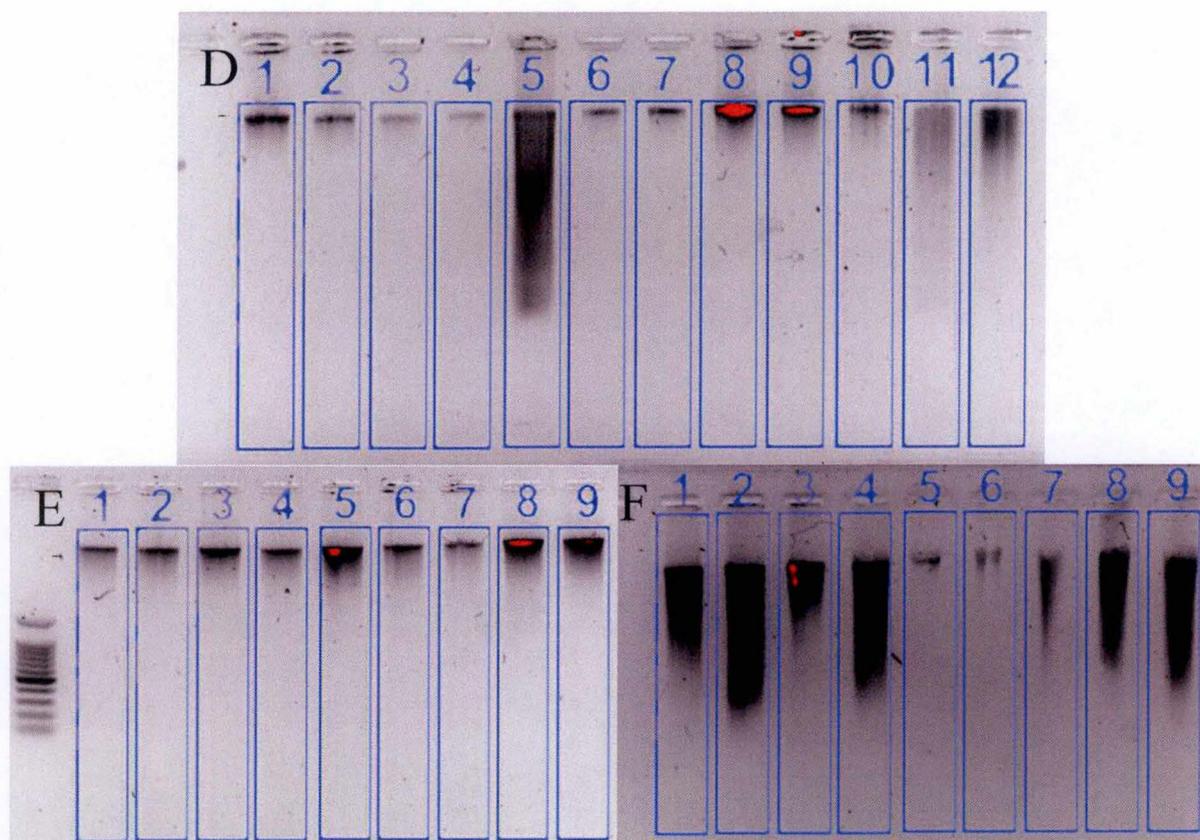


Figure 13 : (A) Extraction ADN des échantillons des œufs *black eyes* et des œufs *golden eyes* et des cerveaux, gonades et foies. Bandes : 1= étalon ; 2 = blanc ; 4-5 = blastodisque ; 6-7= Blastula ; 8-9= *black eyes* ; 10-11= *golden eyes* ; 13-14-15=cerveau. 16-17-18= gonades ; 19-20-21= foie. (B) Extraction ADN des échantillons des stades *black eyes* et *golden eyes* 1= étalon ; 2 = blanc ; 4-5-6 = *black eyes* ; 8-9-10= *golden eyes*. (C)Extraction des échantillons gonades. 1-9= 15M 35M 45M 55M 65M 75M 85M 95M ; 11-12 = 109F et 119F (D) Extraction des échantillons gonades. 1-10 = 19M 29M 39M 49M 59M 69M 79M 89M 99M 109M ; 11-12 = 109F, 119F (E) Extraction des échantillons gonades. 1-9 = 15F, 25F, 35F, 45F, 55F, 65F, 75F, 85F5 et 95F (F) Extraction des échantillons gonades. 1-9 = 19F, 29F, 39F, 49F, 59F, 69F, 79F, 89F, 99F.

Comme le montre la figure 13, la plupart des échantillons montrent une importante dégradation de l'ADN visualisable sur les gels d'agarose avec la présence de *smears* pour les échantillons de *black eyes*, *golden eyes*, foie, cerveau et les gonades de certains poissons adultes (59M, 29F, 49F, 79F, 89F, 99F, 109F).

Enfin, les résultats de l'analyse LUMA sont présentés dans le tableau 5.

Tableau 5: Résultats du LUMA sur les échantillons d'ADN extrait. Les valeurs en rouge traduisent d'une absence de signal, ou insuffisant pour être séparé du bruit de fond. Les ratios G10/Gbf inférieurs à 5 sont également indiqués en rouge.

Échantillon	Q DNA LUMA	Hauteur de G HPAll (pic 10)	Hauteur du bruit de fond (bf) G Hpall (moyenne pic 1 et 3)	ratio $\frac{G10}{Gbf}$
B3	450	0,16	0,13	1,23
B4	450	0,65	0,25	2,60
B5	450	0,53	0,27	1,96
BE1	450	19,40	7,91	2,45
BE2	450	13,79	5,49	2,51
BE3	450	21,84	8,76	2,49
BE4	450	19,95	7,33	2,72
BE5	450	16,26	5,65	2,88
GE1	450	15,37	5,60	2,74
GE2	450	1,16	0,43	2,67
GE5	450	6,76	2,10	3,21
Cerv1	450	4,07	0,18	22,61
Cerv3	450	3,91	0,19	20,58
Gonade1	450	9,56	1,18	8,10
Gonade2	450	5,94	0,67	8,83
Gonade3	450	7,90	0,83	9,55
15M	450	1,21	0,34	3,56
35M	450	4,24	1,48	2,86
95M	450	0,11	0,19	0,59
105M	450	2,10	0,11	19,09
75F	450	1,06	0,17	6,24
15F	450	0,04	0,16	0,26
25F	450	0,08	0,05	1,60
35F	450	0,13	0,06	2,17
45F	450	5,15	1,02	5,07
55F	450	8,21	2,34	3,51
65F	450	13,58	4,36	3,12
85F	450	7,92	1,90	4,18
95F	450	8,26	2,08	3,98
105F	450	8,70	2,09	4,17
115F	450	10,08	2,29	4,41
19M	450	5,29	1,58	3,36
29M	450	2,46	0,31	7,94
39M	450	0,59	0,14	4,21
49M	450	1,67	0,68	2,46
69M	450	2,67	0,42	6,43
79M	450	1,67	0,31	5,39
89M	450	0,42	0,05	9,33
109M	450	7,74	0,46	17,01

99M	450	5,20	1,90	2,74
19F	450	6,64	2,06	3,22
39F	450	6,41	2,12	3,03
59F	450	0,04	0,22	0,19
69F	450	7,31	2,11	3,46

Afin de réaliser une analyse LUMA, il est nécessaire d'avoir des échantillons avec un minimum de 900ng et un minimum de dégradation. Seuls les échantillons présentant une concentration et une qualité d'extraction suffisante ont été analysés.

La majorité des échantillons, dont les échantillons B, Ba, BE, GE possèdent un ratio G10/Gbf inférieur à la valeur seuil de 5 (Head *et al.*, 2014). Certains des échantillons ne possèdent pas une valeur de G au pic 10 suffisante pour être discriminé du bruit de fond, car celles-ci sont proches de 0. Les échantillons de cerveaux 1 et 3 possèdent une hauteur de pic 10 suffisante pour être discriminés du bruit de fond avec un ratio G10/Gbf supérieur à 5. Mais lorsqu'on observe les hauteurs de pic des pyrogrammes de ces échantillons avec la restriction MspI, on observe que les valeurs de ces pics sont très faibles et que le signal ne peut être discriminé du bruit de fond. Pour les échantillons gonades, on peut observer une hauteur de pic G10 beaucoup plus élevée. Cependant, on peut observer que le ratio G(10)/Gbf n'est pas toujours supérieur à 5. Lorsque le ratio G(10)/Gbf est supérieur à 3 avec une différence entre le bruit de fond et le pic G10, le pourcentage de méthylation fut calculé.

Les pourcentages de méthylation calculés selon les résultats LUMA sont présentés dans le tableau 6.

Tableau 6 : Pourcentage de méthylation des échantillons Gonades. Gonade1, Gonade2 et Gonade3 sont les gonades des mâles âgés de 124 jours. 105M et 75F sont des mâles âgés de 5 semaines. 15F, 55F, 65F, 85F, 95F, 105F et 115F sont des femelles âgées de 5 semaines. 19M, 29M, 39M, 49M, 59M, 79M, 89M et 99M sont des mâles âgés de 9 semaines et 19F, 29F et 49F sont des femelles âgées de 9 semaines. Les pourcentages de méthylation où le rapport de G(10)/Gbf est supérieur à 3 mais inférieur à 5 sont indiqué en orange.

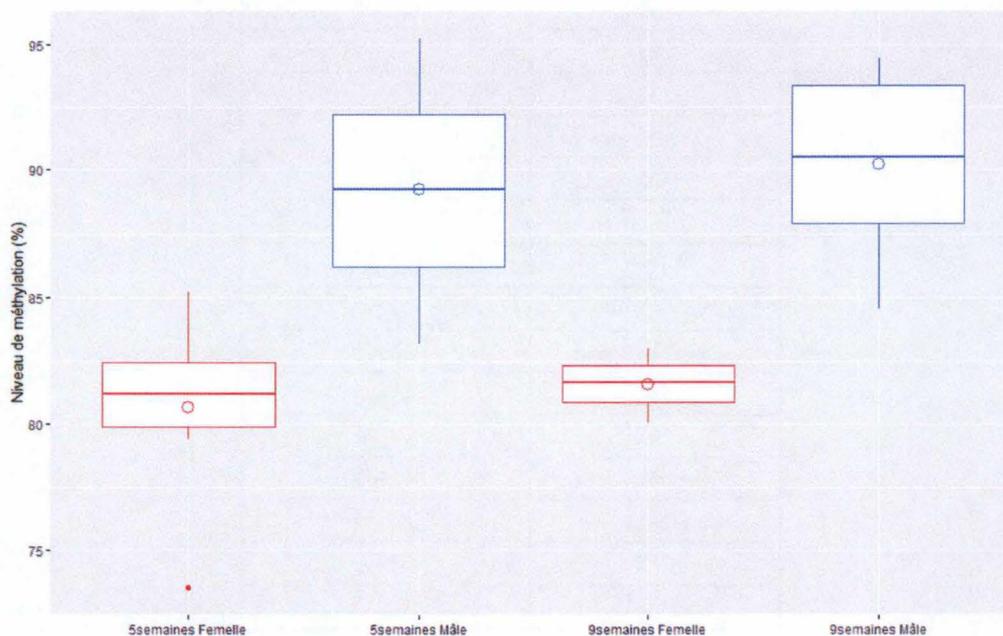
Échantillon	% méthylation
Gonade1	93,9
Gonade2	93,45
Gonade3	93,89
105M	95,23
75F	83,21
45F	82,06
55F	79,37
65F	73,56
85F	81,22
95F	80,36
105F	85,22
115F	82,83
19M	84,56
29M	91,45
39M	93,38
49M	86,47
59M	87,92
79M	90,51
89M	88,67
99M	94,45
19F	80,05
29F	81,66
49F	82,98

Des résultats de l'analyse LUMA, le pourcentage de méthylation a pu être calculé pour les trois échantillons de gonades des mâles âgés de 124 jours. La moyenne du pourcentage de méthylation d'ADN des gonades mâles a été calculée avec les valeurs de la deuxième analyse, donnant un pourcentage de 93,74% \pm 0,26 avec un coefficient de variabilité (CV) de 0,27%.

Les résultats de l'analyse LUMA sur les gonades des adultes *N. furzeri* sont présentés dans le graphique ci-dessous (Figure 14). Les femelles de cinq semaines et de neuf semaines ont une méthylation moyenne de 80,66% \pm 3,65 et 81,56% \pm 1,47 respectivement (CV= 4,5% et 1,80%) et les mâles une moyenne de 89,22% \pm 8,50 à cinq semaines et 90,21% \pm 3,56 à neuf

semaines (CV= 9,53% et 3,94%). Par analyse ANOVA, on observe une différence très significative (p valeur=0.00061) entre les deux sexes quelques soit l'âge des poissons. Lorsqu'on réalise des comparaisons statistiques deux à deux, aucune différence significative ne différencie les mêmes sexes aux différents âges (p valeur entre femelles= 0,74 et p valeur entre mâles = 0,75). La comparaison la plus significative est entre les femelles de cinq semaines et les mâles de neuf semaines (p valeur = 0.00015). Ensuite vient la comparaison entre les femelles et mâles de neuf semaines (p valeur=0.00396), les mâles et les femelles de cinq semaines (p valeur=0.01389) et enfin, les femelles de neuf semaines et les mâles de cinq semaines (p valeur=0.04584).

Figure 14 : Pourcentage de méthylation des échantillons gonades des adultes *N. furzeri* mâles (bleu) et femelles (rouge) âgés de cinq et neuf semaines.



3.3 Discussion

L'objectif de ce travail était de caractériser la reprogrammation de la méthylation de l'ADN chez *N. furzeri* en regardant le pourcentage de méthylation à différents stades de développement, ainsi que la caractérisation de la méthylation chez l'adulte. Dans le cadre de cet objectif, des pré-tests ont été réalisés.

De nombreuses difficultés ont été observées lors de la réalisation de ces pré-tests. Pour les différents stades embryonnaires il n'a malheureusement pas été possible de mesurer le taux de méthylation pour plusieurs raisons. Pour commencer, les rapports 260/280 et 260/230 du Nanodrop™ laissent supposer une contamination de nombreux échantillons dont B et Ba (Desjardins et Conklin, 2010). Ensuite l'extraction d'ADN de tous les échantillons d'œufs en

stade blastodisque et blastula, ne fut pas assez efficace pour fournir les 900ng d'ADN nécessaire pour l'analyse du LUMA. Aucun pic G10 n'a pu être différencié du bruit de fond sur les pyrogrammes lors des essais avec les échantillons B3, 4 et 5. Ces trois échantillons semblaient pourtant contenir suffisamment d'ADN mais la concentration a pu être induite par la présence de la contamination. Tous ces facteurs ont empêché le calcul des pourcentages de méthylation des stades blastodisque et blastula.

Aucune analyse n'a pu être menées pour les stades plus développés de *black eyes* et *golden eyes*, l'ADN des échantillons BE et GE étant très fortement dégradés. La dégradation de l'ADN empêche une mesure correcte de la restriction des sites CCGG par les enzymes HpaII et MspI dû à l'apparition de restrictions non-spécifiques. Selon J. Head, les échantillons peuvent rester 24h à température ambiante (21°C) et on peut obtenir des résultats de LUMA sans trop de dégradation. Nos œufs ont directement été congelés après la récolte par immersion à l'azote liquide puis conservé au congélateur -80°C, décongelés uniquement lors de l'extraction d'ADN. Les échantillons d'ADN élué dans un tampon ont été placés au congélateur -20°C jusqu'à la réalisation de l'analyse LUMA. Cette dégradation ne vient donc probablement pas d'un problème après l'extraction car l'ADN est très stable à ces températures, mais éventuellement d'un problème lors de l'extraction. Lors de l'extraction, les œufs écrasés sont incubés dans une solution de lyse pour dégrader les protéines à 56°C pendant 2h30. À cette température inférieure à 100°C, l'ADN ne se dégrade que très peu (Karni *et al.*, 2013).

La concentration insuffisante en ADN des œufs en stade blastodisque et blastula ainsi que la dégradation de l'ADN des œufs en stade *black eyes* et *golden eyes* laissent penser que le protocole ne soit pas adapté à l'extraction d'ADN sur des œufs. Cependant, ce protocole a été mis en place pour des échantillons contenant peu d'ADN et des précédentes expériences avec ce protocoles sur des œufs de *Kryptolebias marmoratus*, un taxon proche de *N. furzeri*, ont déjà été réalisées (Fellous *et al.*, 2018). D'une part, il se pourrait que le manque de concentration en ADN soit dû à l'utilisation d'un broyage mécanique des œufs avec un pilon. Ce broyage est nécessaire à cause de l'épaisseur du chorion mais causerait une perte d'ADN dû à du tissu resté collé sur le pilon. Ce problème pourrait probablement être résolu en augmentant le nombre d'œufs des stades blastodisque et blastula par répliques tout en subdivisant en de plus nombreux échantillons lors de l'extraction.

Pour évaluer le niveau de méthylation de l'ADN chez l'adulte, nous disposions d'échantillons de trois organes : le foie, le cerveau et les gonades, chez des mâles âgés de 124 jours ainsi que les gonades de femelles et de mâles âgés de cinq et neuf semaines.

Il a été impossible d'obtenir des résultats LUMA pour les échantillons de foies et de cerveaux des poissons vieux de 124 jours. Les échantillons de foie étaient trop dégradés pour pouvoir réaliser une analyse LUMA et les échantillons de cerveaux trop peu concentrés. Le taux de méthylation des gonades des mâles de 124 jours a pu être calculé à 93,74% avec un coefficient de variabilité très faible (0,27%). L'utilisation de poissons vieux de 124 jours peut amener une réflexion supplémentaire vu que l'âge de ces poissons est supérieur à l'espérance de vie moyenne de la souche utilisée (Genade *et al.*, 2005 ; Valdesalici et Cellerino, 2003). L'effet du vieillissement peut alors impacter nos résultats. Les effets du vieillissement ont été énormément étudiés chez *N. furzeri* (Api *et al.*, 2017 ; Baumgart *et al.*, 2014 ; Di Cicco *et al.*, 2011 ; Genade *et al.*, 2005 ; Reichwald *et al.*, 2015 ; Terzibasi *et al.*, 2007) mais peu sur le méthylome de *N. furzeri*. Les études sur l'évolution du niveau de méthylation en fonction de l'âge suggérant une hypométhylation globale chez les mammifères (Cruickshanks *et al.*, 2013) et chez les gonades mâles (Gunes *et al.*, 2016), on peut hypothétiser une diminution du niveau global de méthylation des gonades de poissons âgés de 124 jours par rapport aux poissons âgés de cinq et neuf semaines. Cependant, la méthylation globale des gonades des poissons âgés de 124 jours n'est pas significativement différente de la méthylation globale des gonades des poissons plus jeunes. Cette observation pourrait s'expliquer par le fait d'une transmission des marqueurs épigénétiques de méthylation par la lignée paternel et donc un maintien du haut taux de méthylation (Jiang *et al.*, 2013 ; O'Neill, 2013).

Les valeurs de méthylation pour les *N. furzeri* femelles et mâles de cinq et neuf semaines présentent des différences significatives en fonction de l'âge. Les gonades des femelles de cinq semaines et de neuf semaines ont une méthylation moyenne de 80,66 et 81,56% respectivement et les gonades des mâles une moyenne de 89,22 et 90,21%. Il faut cependant analysés ces résultats avec précaution vu que le ratio G10/Gbf n'était pas toujours au-delà de 5. Les valeurs étant cohérentes entre-elles, et le ratio évalué par corrélation dans la publication de J. Head (2014), les résultats de méthylation globale seront comparés à ceux présent dans la littérature chez d'autres organismes. Ces valeurs sont un peu plus élevé que ce qui a pu être observé chez *Kryptolebias marmoratus* (Fellous *et al.*, 2018), à raison de 87,22% de méthylation pour les gonades mâles et 79,55% pour les gonades des hermaphrodites. Ces niveaux de méthylation restent relativement équivalant en comparant les femelles *N. furzeri* et les hermaphrodites *K. marmoratus* et les mâles des deux espèces. De la même manière, la méthylation globale est similaire chez *Danio rerio*, avec une méthylation des gonades mâles de 87,76%, ce qui est plus élevé que la méthylation globale des gonades femelles, 82,78% (Laing *et al.*, 2018). La différence de méthylation observée entre les organismes pourrait

provenir de la concentration plus importante en GC dans le génome de *N. furzeri* par rapport aux autres organismes (Reichwald *et al.*, 2009). Ces différences entre les deux sexes peuvent s'expliquer par le fait que les spermatozoïdes sont plus méthylés que les ovules chez les poissons (Hackett et Surani, 2015). Les spermatozoïdes étant fort méthylés, les gonades mâles possèdent alors un niveau de méthylation global plus élevé et implique que les marqueurs épigénétiques de la méthylation de l'ADN pourraient être principalement hérités via la lignée paternelle (Jiang *et al.*, 2013 ; O'Neill, 2013). Ce qui peut être soutenu par le fait que le génome paternel ne subit pas une déméthylation aussi profonde que le génome maternel chez *D. rerio* mais également chez *Oryzias latipes* lors de la reprogrammation du zygote (Hackett et Surani, 2015 ; Wang et Bhandari, 2019).

L'absence de mesure du niveau de méthylation lors du développement de l'embryon de *N. furzeri* empêche la visualisation du profil de reprogrammation de celui-ci. Cependant, on peut observer des concordances du niveau de méthylation des gonades mâles et femelles avec les espèces *K. marmoratus* (Fellous *et al.*, 2018) et *O. latipes* (Wang et Bhandari, 2019). Les espèces étant également des taxons proches (Genade *et al.*, 2005 ; Terzibasi *et al.*, 2007), on peut estimer que le profil de reprogrammation de l'embryon en développement de *N. furzeri* sera similaire. Hypothétiquement, le niveau de méthylation globale devrait atteindre les 20-25% de méthylation chez le zygote juste après la fusion et remonter progressivement à un niveau de méthylation similaire à celui des spermatozoïdes aux environs de 90% de méthylation à l'étape de gastrulation. Bien sûr, le développement de *N. furzeri* comporte quelques singularités qui pourrait modifier ce profil de reprogrammation théorique. Le fait que *N. furzeri* soit un poisson au cycle de vie annuel avec la présence de diapauses dans son développement pourrait influencer le profil de reprogrammation. La diapause a des effets sur le métabolisme de *N. furzeri* qui pourrait être impacté par la méthylation de l'ADN (Podrabsky et Hand, 2000 ; Cellerino *et al.*, 2016 ; Forest *et al.*, 2012 ; Furness *et al.*, 2015a) et il a été découvert un impact de la diapause chez *Austrofundulus limnaeus* sur la méthylation des histones (Reynolds, 2017).

3.4 Perspectives

Outre les problèmes d'adaptations de l'extraction d'ADN pour les œufs de *N. furzeri*, il serait surtout intéressant d'augmenter le nombre de stades du début de développement embryonnaire afin de correctement visualiser la reprogrammation de la méthylation de l'ADN qui, si l'on se base sur le profil de reprogrammation de *K. marmoratus* se situerait entre la

fécondation et le stade de dispersion, ou d'épibolie 100%. Il serait également intéressant d'effectuer l'analyse sur les gamètes oocytes et spermatozoïdes.

Il serait également intéressant de mesurer le pourcentage de méthylation lors des trois diapauses afin d'observer s'il existe une influence de ce stade particulier sur ces mécanismes épigénétiques comme observé chez *Austrofundulus limnaeus*.

Enfin, il faudrait évaluer le pourcentage de méthylation de gènes spécifiques chez *N. furzeri* tels que les gènes DNMT, gènes responsables de la méthylation de l'ADN.

3.5 Conclusion

Le pourcentage de méthylation des gonades de spécimens adultes de *N. furzeri* est plus important que chez les autres organismes et pourrait être dû à la concentration en GC plus importante chez *N. furzeri*. La différence observée entre les deux sexes est cohérente avec ce qui est observé chez les autres organismes tels que *K. marmoratus* et *D. rerio*. Il faut cependant éviter d'accorder trop de confiance à ces résultats dû à la présence de contrôle qualité faible.

4. Bioinformatique

Les protéines DNA méthyltransférase (DNMT) sont des protéines bien documentées chez de nombreuses espèces, tels que *Danio rerio* ou *Mus musculus* (Kamstra *et al.*, 2015 ; Mhanni *et al.*, 2001 ; Law et Jacobsen 2010). Cependant leur analyse chez *N. furzeri* est relativement peu documentée. Seul l'expression de leur gène a été étudiée (Baumgart *et al.*, 2014). Ces protéines sont extrêmement importantes dans la méthylation de l'ADN (Day et Sweatt, 2010 ; Jones, 2012 ; Suzuki et Bird, 2008), soit par méthylation *de novo* assuré par la DNMT3A ou DNMT3B soit en l'entretenant entre les divisions par la DNMT1, c'est pour cela qu'il est important de réaliser une analyse de ces protéines et de leur gène chez *N. furzeri*.

Dans un premier temps, les séquences protéiques des DNMT seront analysées d'un point de vue fonctionnel et seront comparée entre-elles via un programme d'alignement multiple. Ensuite les séquences protéiques seront analysées selon leur phylogénie carseules quelques analyses phylogénétiques ont été réalisée sur le génome de *N. furzeri* (Costa, 2018 ; Dorn *et al.*, 2011 ; Furness *et al.*, 2015a ; Terzibasi *et al.*, 2007), et il peut être intéressant d'en réaliser une sur les séquences protéiques de DNMT définies plus tôt afin d'évaluer l'histoire évolutive des différentes DNMT. Enfin une analyse des ilots CpG *in silico* sera réalisée. L'analyse des ilots CpG est utile lors de l'étude de la méthylation de l'ADN car elle permet de visualiser la position préférentielle de la méthylation, considérant que celle-ci se déroule sur les régions riches en CpG (Holliday et Pugh, 1975 ; Riggs, 1975).

4.1 Méthode

4.1.1 Analyses des séquences protéiques DNMT chez *N. furzeri*

Les séquences protéiques des DNMT de *Nothobranchius furzeri* ont été obtenues par recherche BLASTX dans la base de données NCBI (*National Center for Biotechnology Information*) *non-redondant protein sequences*, depuis les séquences nucléotidiques ARN de *N. furzeri* annotées et identifiées par Reichwald *et al.*, (2015) dans le projet de séquençage PRJNA314891. Les séquences protéiques ont été confirmées par un *reciprocal best blast hit* et comparée aux séquences présentent dans la base de données Uniprot. Les séquences ont ensuite été analysées via le HMMER web server (Potter *et al.*, 2018), afin de rechercher les sites homologues sur les séquences protéiques et de visualiser les domaines trouvés dans la base de données de protéomes de références de HMMER. Une E-value est également fournie afin de supporter l'alignement réalisé dans la base de données. L'E-value caractérise la

probabilité de réaliser l'alignement par hasard, plus celui-ci est proche de 0, plus le résultat de l'alignement est significatif.

Ensuite, pour évaluer la ressemblance entre les différentes DNMT, les séquences sont comparées deux par deux par Dialign Genomatix qui permet un alignement multiple de séquences protéiques et nucléiques sur l'ensemble des séquences détectant ainsi même de petites homologies. Les séquences protéiques déterminées plus tôt ainsi que les séquences ARN identifiées par Reichwald *et al.*, (2015) et les séquences ADN correspondantes (Annexe 2, 3 et 4) ont été implémentées dans le logiciel (<http://www.genomatix.de/cgi-bin/dialign/dialign.pl>). Les coordonnées des gènes protéiques correspondants sont présentées dans le tableau 7.

Tableau 7 : Gènes, ID des gènes dans la base de données NCBI ainsi que les coordonnées de ces gènes.

Gene	Gene ID	Sequence	Start	End
DNMT3B-like	107372626	scaffold01559	317	30700
DNMT3A-like	107377776	sgr04	56040084	56086670
DNMT1	107378352	sgr05	3305854	3321094
DNMT3B	107390569	sgr15	568335	605813
LOC107391197	107391197	sgr15	30851095	30869300
DNMT3A	107392003	sgr16	19809244	19895742

4.1.2 Phylogénie des protéines DNMT

L'analyse phylogénique a été menée sur les séquences protéiques de DNMT obtenues sur de nombreux organismes via une recherche BLASTP dans la base de données NCBI. Certaines séquences de certains organismes ne sont pas reprises dans l'analyse car l'homologie n'était pas suffisante pour être détectée par la recherche BLASTP. Ces séquences ont été alignées avec Muscle (Edgar, 2004) en utilisant le programme MEGA7 (Kumar, Stecher, & Tamura, 2016). L'analyse phylogénétique a été réalisée selon la méthode du Maximum Likelihood (ML) basée sur le modèle de matrice JTT (Jones *et al.*, 1992) avec un bootstrap de 3000 répliques (Felsenstein, 1985).

4.1.3 Analyse des îlots CpG

Les analyses ont été réalisées par Emboss CpGplot sur https://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/ (Rice *et al.*, 2000) sur les séquences ADN des gènes DNMT1, DNMT3a et DNMT3b de *Nothobranchius furzeri*. Les différents paramètres ont été définis aux valeurs par défaut, établis selon Gardiner-Garden et Frommer

(1987). Le ratio Observé/attendu a été fixé à 0.60, seuls les îlots contenant plus de 50% GC ainsi qu'une longueur supérieure à 200 sont représentés.

4.2 Résultats

4.2.1 Analyses des séquences protéiques DNMT chez *N. furzeri*

Les protéines DNMT trouvées par la recherche BLASTX et confirmées par *reciprocal best blast hit* ainsi que les résultats de l'alignement par HMMER sont présentés dans la figure 15 et le tableau 8. Les séquences protéiques sont précisées en Annexe 2.

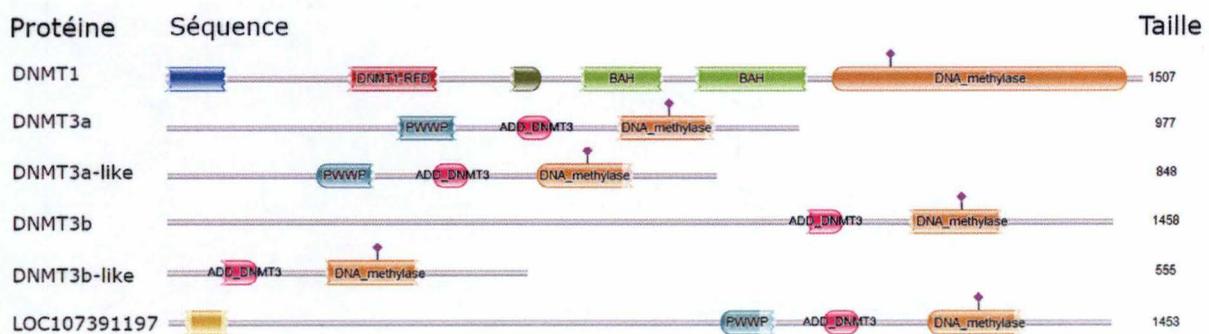


Figure 15 : Nom des séquences DNMT présentes chez *N. furzeri* ainsi que leur taille en paire de base, avec les domaines identifiés. Domaine ADN méthyltransférase (orange) avec le site actif (losange rose), domaine BAH (Vert clair), domaine C-5 cytosine-specific DNA methylase (Bleu foncé), domaine Cytosine specific DNA methyltransferase replication foci (DNMT1-RFC)(rouge foncé), domaine DMAP1-binding (vert foncé), Domaine ADD-DNMT3 (rose) et domaine PWWP (Bleu clair), domaine CAMSAP CH.

Tableau 8 : Description des domaines trouvés par alignement ainsi que les E-values associées.

Protéine	Description des domaines	E-value
DNMT1	C-5 cytosine-specific DNA methylase	4.7e-50
	Cytosine specific DNA methyltransferase replication foci domain	4.3e-45
	BAH domain	1.1e-16
	BAH domain	2.3e-15
	DMAP1-binding Domain	1.2e-16
	CXXC zinc finger domain	9.3e-14
DNMT3a	Cysteine rich ADD domain in DNMT3	4.6e-24
	C-5 cytosine-specific DNA methylase	4.1e-12
	PWWP domain	1.6e-12
DNMT3a-like	Cysteine rich ADD domain in DNMT3	2.9e-26
	C-5 cytosine-specific DNA methylase	7.8e-12
	PWWP domain	5.0e-11
DNMT3b	Cysteine rich ADD domain in DNMT3	5.4e-25
	C-5 cytosine-specific DNA methylase	2.9e-16
DNMT3b-like	Cysteine rich ADD domain in DNMT3	1.6e-25
	C-5 cytosine-specific DNA methylase	6.7e-17
LOC107391197	Cysteine rich ADD domain in DNMT3	8.3e-24
	C-5 cytosine-specific DNA methylase	6.2e-13
	PWWP domain	6.5e-08
	CAMSAP CH domain	0.00015

Les protéines DNMT présentent de nombreux domaines, dont le domaine *C-5 cytosine-spécifique DNA methylase* que toutes les DNMT ont en commun. Ce domaine est celui permettant de catalyser la réaction de transfert du groupement méthyl sur la cytosine (Holliday et Pugh, 1975 ; Riggs, 1975). La DNMT1 présente des domaines spécifiques tels que le domaine *Cytosine specific DNA methyltransferase replication foci*, le domaine BAH, le domaine DMAP1-binding et le domaine *CXXC zinc finger*. Respectivement associés à la reconnaissance correcte des résidus (Rountree *et al.*, 2000), à l'association entre protéines pour la régulation de la transcription (Callebaut *et al.*, 1999), à la liaison d'un corépresseur (Rountree *et al.*, 2000) et à la liaison à des CpG non-méthylé (Cross *et al.*, 1997). Entre les séquences de DNMT3a, DNMT3a-like, la séquence LOC107391197, la séquence DNMT3b et DNMT3b-like, le domaine PWWP, lié à des interactions protéines-protéines (Stec *et al.*, 2000) est commun. Le domaine *Cysteine rich ADD* est quant à lui commun entre les protéines DNMT3a, DNMT3a-like et LOC107391197. Ce domaine a une structure proche d'un *PHD finger* et est lié à la méthylation de l'ADN (Taverna *et al.*, 2007). La plupart de ces domaines possèdent des E-values élevées, caractérisant un alignement très significatif. Sauf le domaine CAMSAP CH, aligné avec la séquence LOC107391197 avec une E-value de 0,00015 l'alignement de ce domaine pourrait être dû au hasard. De plus, ce domaine est lié à la

structure cellulaire (Marchler-Bauer *et al.* 2017), ce qui corrobore l'hypothèse d'un alignement dû au hasard. Contrairement aux autres domaines qui ont des fonctions associées à la fonction des protéines DNMT.

Les différences de domaines entre la séquence de DNMT1 et celles de DNMT3 peuvent s'expliquer par la différence de fonction, la protéine DNMT1 assurant l'entretien de la méthylation et les DNMT3 catalysant la réaction de méthylation *de novo* (Day et Sweatt, 2010 ; Jones, 2012 ; Suzuki et Bird, 2008). La séquence LOC107391197 a été trouvée dans la base de données NCBI via *le reciprocal best blast hit* contre la séquence de la protéine DNMT3b. La séquence LOC107391197 contient le domaine *Cysteine rich ADD* comme les séquences DNMT3a et DNMT3a-like. Cette séquence est structurellement plus proche de DNMT3b mais est fonctionnellement plus proche d'une DNMT3a.

Les domaines alignés sur la séquence DNMT1 de *N. furzeri* sont cohérents avec ceux trouvés pour la protéine DNMT1 de *D. rerio* et *K. marmoratus* (Kim *et al.*, 2015 ; Mhanni *et al.*, 2001) ainsi que ceux des protéines DNMT3 (Goll et Halpern, 2011 ; Kim *et al.*, 2015), ce qui renforce notre analyse sur les protéines DNMT de *N. furzeri*. Toutefois, il faut faire attention au fait que les séquences trouvées dans la base de données NCBI sont des séquences modèles, inférées par homologie sur base du séquençage du génome, ce qui peut induire possiblement des erreurs de séquence dû à l'apparition de différences dans la séquence génomique entre des espèces différentes à la suite des phénomènes évolutifs (Day et Tuite, 1998 ; Gallego et Virshup, 2007).

Les résultats d'identité et de similarité des alignements des séquences protéiques sont présentés dans le tableau 9. Les résultats d'identité et de similarité des alignements des séquences ARN sont présentés dans le tableau 10. Les résultats d'identité et de similarité des alignements des séquences sont présentés dans le tableau 11.

Tableau 9: Comparaison deux à deux des séquences protéiques des DNMT de *N. furzeri* par Dialign Genomatix. Pour chaque alignement, la similarité (relative à la similarité maximale) et l'identité (en %) sont données. Les valeurs maximales sont soulignées.

	DNMT3b (1458 aa)	DNMT3a (977 aa)	DNMT3a-like (848 aa)	DNMT3b-like (555 aa)	LOC107391197 (1453 aa)
DNMT1 (1507 aa)	0.007 4 %	0.009 7 %	0.009 7 %	0.006 11 %	0.005 3 %
DNMT3b (1458 aa)		0.412 31 %	0.427 37 %	<u>1.000</u> <u>100 %</u>	0.582 32 %
DNMT3a (977 aa)			0.980 70 %	0.395 47 %	0.553 37 %
DNMT3a-like (848 aa)				0.421 48 %	0.522 41 %
DNMT3b-like (555 aa)					0.496 55 %

Après alignement local des séquences protéiques DNMT trouvées chez *N. furzeri*, on peut observer la présence d'un alignement de 100% entre la séquence DNMT3b et DNMT3b-like. Les alignements suivant les plus proches sont ceux entre les séquences DNMT3a et DNMT3a-like avec 70% d'identité et 0,98% de similarité. La séquence de DNMT1 quant à elle, ne présente que très peu de similarité ainsi qu'un pourcentage faible d'identité avec les autres séquences protéiques identifiées. La séquence la plus proche de la séquence non-identifiée LOC107391197 est celle de DNMT3b-like.

Tableau 10 : Comparaison deux à deux des séquences ARN des DNMT de *N. furzeri* par Dialign Genomatix. Pour chaque alignement, la similarité (relative à la similarité maximale) et l'identité (en %) sont données. Les valeurs maximales sont soulignées.

	DNMT3b (4593 bp)	DNMT3a (5954 bp)	DNMT3a-like (6670 bp)	DNMT3b-like (1830 bp)	LOC107391197 (5019 bp)
DNMT1 (4953 bp)	0.007 5 %	0.002 5 %	0.003 5 %	0.004 9 %	0.002 4 %
DNMT3b (4593 bp)		0.087 14 %	0.082 15 %	<u>1.000</u> <u>97 %</u>	0.139 22 %
DNMT3a (5954 bp)			0.309 25 %	0.093 39 %	0.107 17 %
DNMT3a-like (6670 bp)				0.100 35 %	0.095 17 %
DNMT3b-like (1830 bp)					0.133 40 %

On peut observer que les séquences ARN les plus proches sont les séquences de DNMT3b et DNMT3b-like, rejoignant ainsi l'alignement des séquences protéiques. De même, les comparaisons de la séquence de DNMT1 avec les autres séquences montrent également de très faibles d'identité et de similitude (<10%). La séquence la plus proche de la séquence non-identifiée LOC107391197 semble toujours être celle de DNMT3b-like. Cependant, on peut observer que les valeurs d'identité et de similarité sont plus faible pour l'ensemble des comparaisons par rapport aux comparaisons des séquences protéiques, probablement dû à la redondance du code nucléotidique. La séquence ARN DNMT3a-like est

plus proche de la séquence DNMT3b-like, contrairement à l'alignement des séquences protéiques.

Tableau 11 : Comparaison deux à deux des séquences ADN génomique des DNMT de *N. furzeri* par Dialign Genomatix. Pour chaque alignement, la similarité (relative à la similarité maximale) et l'identité (en %) sont données. Les valeurs maximales sont soulignées.

	DNMT 3a (2787 bp)	DNMT 3a-like (2547 bp)	DNMT 3b (2322 bp)	DNMT 3b-like (1124 bp)	LOC107391197 (4023 bp)
DNMT 1 (4524 bp)	0.008 8 %	0.012 5 %	0.016 4 %	0.010 16 %	0.008 4 %
DNMT 3a (2787 bp)		<u>1.000</u> <u>56 %</u>	0.014 8 %	0.270 45 %	0.328 27 %
DNMT 3a-like (2547 bp)			0.017 8 %	0.260 43 %	0.289 30 %
DNMT 3b (2322 bp)				0.009 10 %	0.058 13 %
DNMT 3b-like (1124 bp)					0.295 45 %

L'alignement des séquences ADN montre que les séquences DNMT3b et DNMT3b-like, ne sont plus les séquences les plus proches avec un pourcentage d'identité de seulement 10%. La séquence DNMT3b-like est plus proches des autres séquences avec la plus grande similarité (0,295) et identité (45%) avec la séquence LOC107391197. Les séquences avec les plus grands scores de similarité et d'identité sont les séquences DNMT3a et DNMT3a-like avec un identité de 56% et une similarité de 1,0. Les pourcentages sont à nouveau plus faibles que ceux de l'alignement des séquences protéiques dû à la redondance du code nucléique. Les alignements de la séquence de DNMT1 avec les autres séquences présentent une identité et une similarité faibles, similaires aux pourcentages des alignements protéiques et ARN.

Les comparaisons sont différentes suivant le type de séquences. Les séquences protéiques montrent clairement des résultats de similarité et d'identité supérieures par rapport aux séquences nucléotidiques. Ce qui semble logique, considérant la redondance du code nucléotidique. Avec les résultats d'alignement des séquences protéiques, on peut estimer que la séquence DNMT3b-like est un morceau de DNMT3b et serait un pseudogène. La séquence LOC107391197 présentant une similarité et identité plus importante envers la séquence DNMT3b-like que les autres, il est probable que cette séquence soit également un pseudogène de la séquence DNMT3b. Le tout étant en corrélation avec les résultats trouvés lors de l'analyse par HMMER. La séquence DNMT1 est plus éloignée des autres DNMT, suggérant une différenciation du gène plus précoce dans l'évolution.

4.2.2 Phylogénie des protéines DNMT

Les résultats de l'analyse phylogénétique sont visibles à la figure 16 et 17.

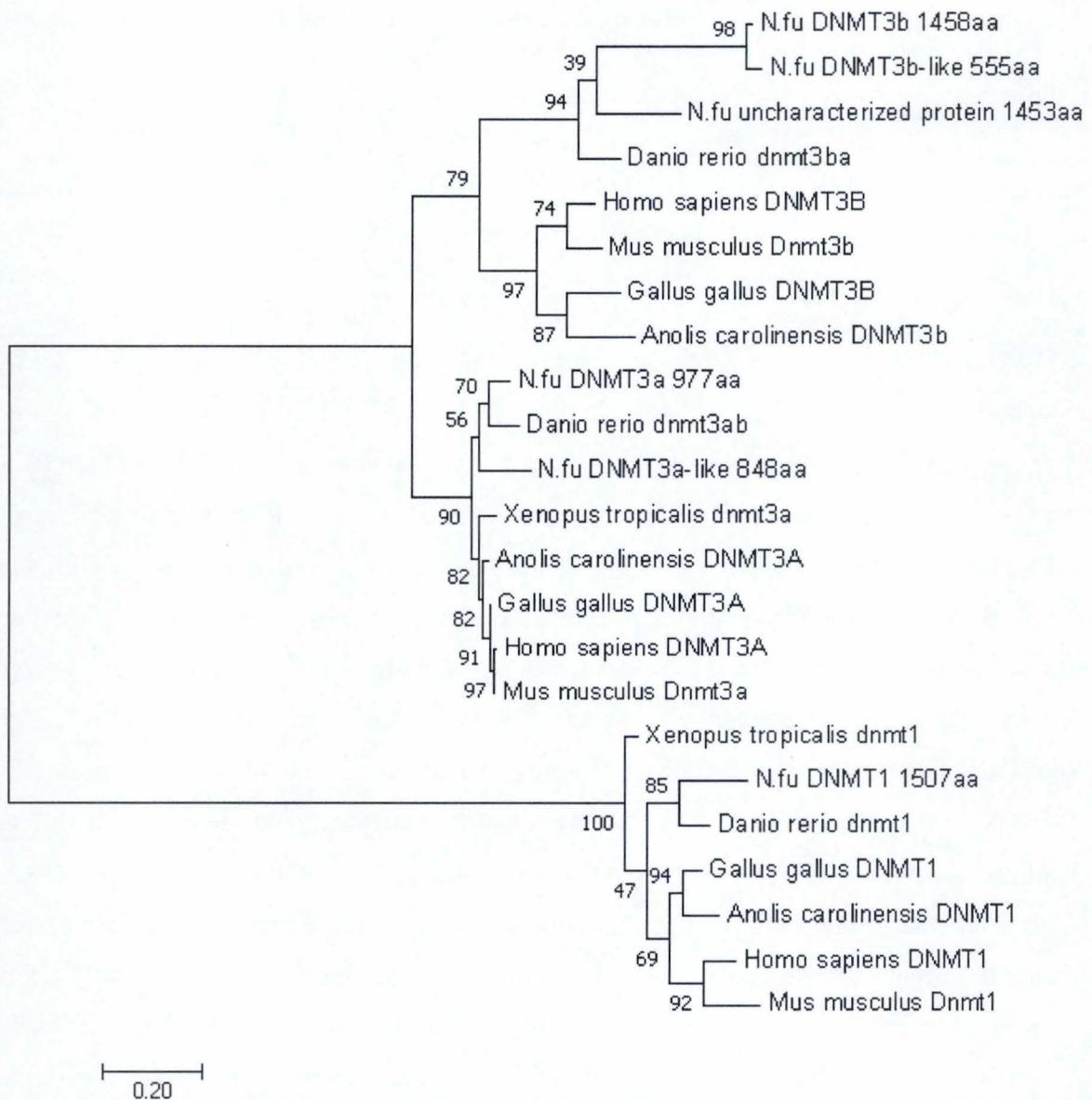


Figure 16 : Analyse phylogénétique de 23 séquences protéiques par la méthode Maximum Likelihood basée sur le modèle JTT (Jones et al., 1992), avec une valeur de bootstrap de 3000 replicas (Felsenstein, 1985). L'arbre montrant la valeur log likelihood la plus élevée (-7587.66) est montré. La valeur présente à côté des noeuds de l'arbre représente le pourcentage d'arbre contenant ce taxa. L'arbre de départ de cette analyse a été obtenu par Neighbor-Joining avec la valeur log likelihood la plus élevée. La longueur des branches mesure le nombre de substitutions par site. N.fu uncharacterized protein 1453aa correspond à la séquence LOC107391197. Les organismes représentés sont *Xenopus tropicalis*, *Nothobranchius furzeri*, *Danio rerio*, *Gallus gallus*, *Anolis carolinensis*, *Mus musculus* et *Homo sapiens*.

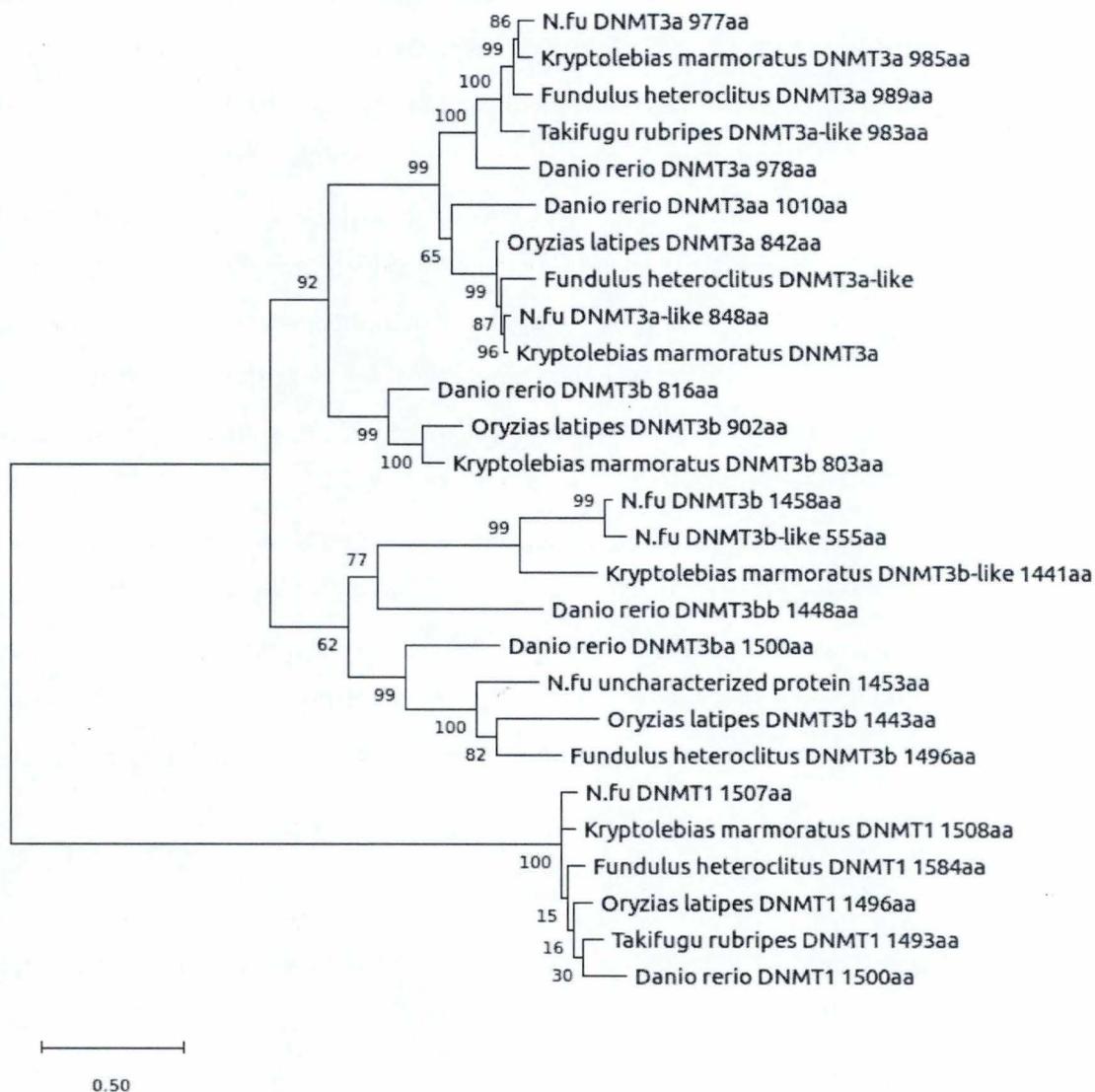


Figure 17 : Analyse phylogénétique de 27 séquences protéiques par la méthode Maximum Likelihood basée sur le modèle JTT (Jones et al., 1992), avec une valeur de bootstrap de 3000 replicas (Felsenstein, 1985). L'arbre montrant la valeur log likelihood la plus élevée (-42430.07) est montré. La valeur présente à côté des noeuds de l'arbre représente le pourcentage d'arbre contenant ce taxa. L'arbre de départ de cette analyse a été obtenu par Neighbor-Joining avec la valeur log likelihood la plus élevée. La longueur des branches mesure le nombre de substitutions par site. N.fu uncharacterized protein 1453aa correspond à la séquence LOC107391197. Les organismes représentés sont *Nothobranchius furzeri*, *Fundulus heteroclitus*, *Kryptolebias marmoratus*, *Oryzias latipes*, *Takifugu rubripes* et *Danio rerio*.

Les arbres phylogénétiques ont permis de mettre en évidence trois groupes de protéines pour les deux analyses. Le 1^{er} consiste en l'ensemble des séquences de DNMT1, le 2^{ème} aux séquences DNMT3a et DNMT3a-like et le 3^{ème} aux séquences de DNMT3b et DNMT3b-like. On peut observer de nombreux orthologues présent dans chacun de ces groupes entre les différentes espèces. La séquence LOC107391197 est associée avec les séquences de DNMT3b et DNMT3b-like. L'organisation des séquences orthologues pour la 1^{ère} analyse correspond à d'autres analyse phylogénétiques menées sur d'autres séquences (Valenzano *et al.*, 2015). Cette organisation des séquences dans l'arbre phylogénétique selon

d'abord l'identité de la protéine puis selon l'espèce permet de supporter l'idée d'une divergence des différentes DNMTs avant la divergence des espèces sélectionnées (Ponger et Li, 2005). La deuxième analyse concerne plus spécifiquement le groupe des Actinoptérygiens. *N. furzeri* appartient à l'ordre des Cyprinodontiformes comme *Fundulus heteroclitus* (Terzibasi *et al.*, 2007) et *K. marmoratus* (Vermeulen et Hrbek, 2005). *Takifugu rubripes* appartient à l'ordre des Tetraodontiformes. Ces deux groupes sont sœurs du groupe des Beloniformes auquel appartient *Orizyas latipes*. Tous appartiennent au superordre Achantopterygii. *Danio rerio* est un représentant du superordre Ostariophysii (Steinke *et al.*, 2006). Cependant, les différents orthologues dans la 2^{ème} analyse phylogénétique correspondent moins à cette phylogénie.

Les résultats de l'analyse phylogénétique confirment ce qui a été établis précédemment par l'alignements des séquences, la divergence des différentes DNMTs s'est déroulé avant la divergence des vertébrés. Le gène DNMT1 ayant divergé plus tôt que les gènes de DNMT3, divergence liée à la différenciation de fonction dans la méthylation avec DNMT1 impliqué dans le maintien de la méthylation et les DNMT3 dans la méthylation *de novo*.

4.2.3 Analyse des îlots CpG

Les résultats de l'analyse *in silico* des îlots CpG sont présentés dans les figures 18, 19, 20. Les coordonnées des îlots sont présentées dans le tableau 12.

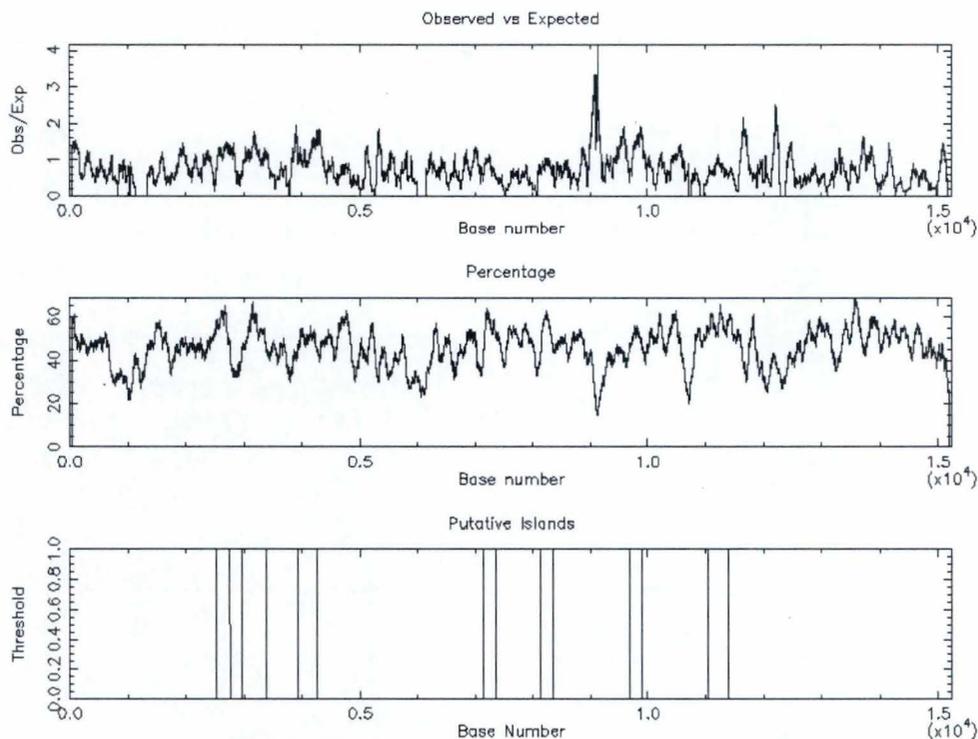


Figure 18 : CpG plot de la séquence du gène DNMT1 obtenue par Emboss CpGplot (Rice et al., 2000). Les différents paramètres ont été réglés aux valeurs par défaut (Window Size : 100, Min length : 200 (taille minimale de l'îlot CpG pour être reporté). Min obs/exp: 0.6 (moyenne minimale du ratio Obs/Exp). Min percentage: 50 (minimum de moyenne de pourcentage de GC sur 10 fenêtres).

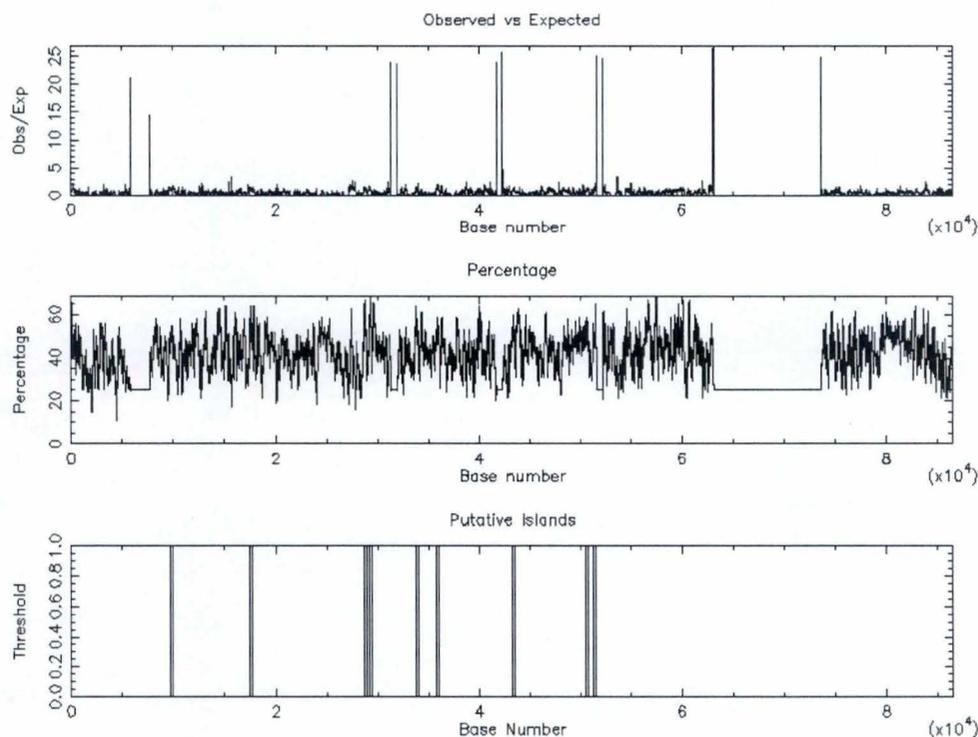


Figure 19: CpG plot de la séquence du gène DNMT3a obtenu par Emboss CpGplot (Rice et al., 2000). Les différents paramètres ont été réglés aux valeurs par défaut (Window Size : 100, Min length : 200 (taille minimale de l'îlot CpG pour être reporté). Min obs/exp: 0.6 (moyenne minimale du ratio Obs/Exp). Min percentage: 50 (minimum de moyenne de pourcentage de GC sur 10 fenêtres).

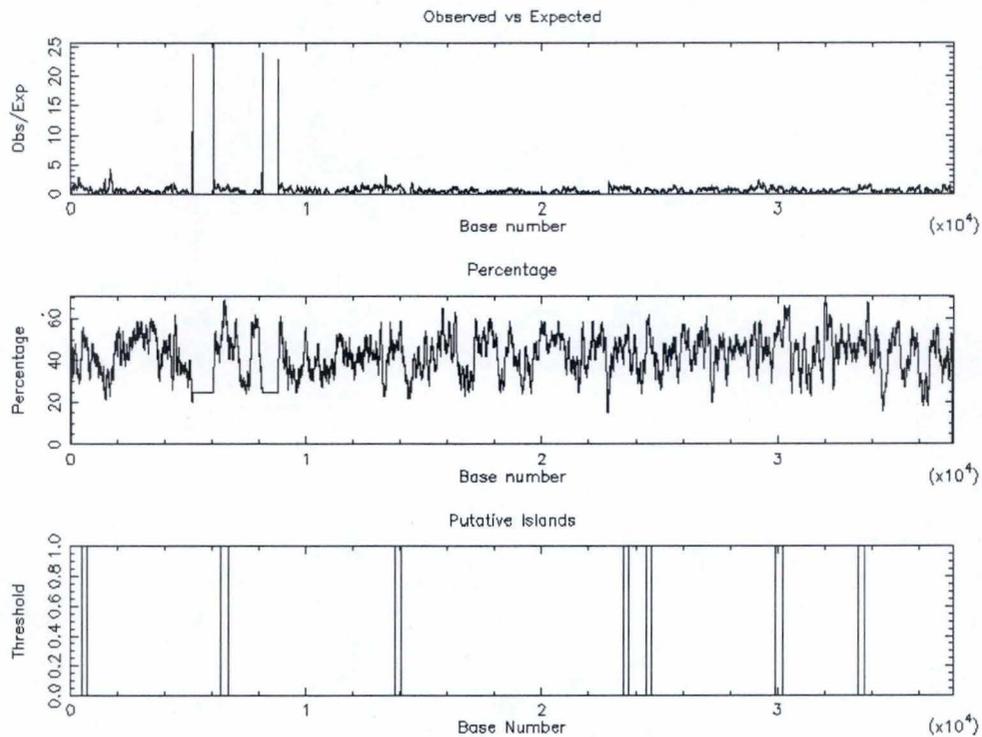


Figure 20 : CpG plot de la séquence du gène DNMT3b obtenu par Emboss CpGplot (Rice et al., 2000). Les différents paramètres ont été réglés aux valeurs par défaut (Window Size : 100, Min length : 200 (taille minimale de l'îlot CpG pour être reporté). Min obs/exp: 0.6 (moyenne minimale du ratio Obs/Exp). Min percentage: 50 (minimum de moyenne de pourcentage de GC sur 10 fenêtres).

Tableau 12: longueur et position des îlots CpG identifiés par emboss CpGplot (Rice et al., 2000) sur les séquences ADN de DNMT3a, DNMT3b et DNMT1.

DNMT3a (86 499 pb)		DNMT3b (37 479 pb)		DNMT1 (15 241 pb)	
Longueur (pb)	Position	Longueur (pb)	Position	Longueur (pb)	Position
230	9723 à 9952	253	465 à 717	236	2520 à 2755
273	17468 à 17740	309	6377 à 6685	412	2967 à 3378
260	28609 à 28868	279	13776 à 14054	313	3939 à 4251
220	29130 à 29349	214	23504 à 23717	212	7144 à 7355
279	33708 à 33986	207	24443 à 24649	209	8134 à 8342
313	35754 à 36066	307	29931 à 30237	229	9679 à 9907
204	43272 à 43475	280	33411 à 33690	346	11046 à 11391
254	50503 à 50756				
343	51174 à 51516				

La présence de pic sur graphe Obs/Exp ainsi que les plateaux de pourcentage calculé s'explique par la présence de larges régions de séquences non définies (NNN).

La localisation des îlots CpG sur les gènes DNMT ne sont pas identiques entre les séquences de même que leurs nombres. Le nombre d'îlot est un peu plus grand pour DNMT3a

ce qui est probablement lié à la taille de la séquence qui est plus grande également. Cependant, les îlots identifiés sur la séquence de DNMT3a semblent se concentrer sur la première moitié de la séquence, contrairement aux deux autres séquences où les îlots semblent plus dispersés. Il se peut que cette observation soit dû à la présence d'une large séquence non définie (NNN). Les îlots de DNMT1 (279,6 pb) semblent avoir une taille moyenne plus importante que les DNMT3a (264,0 pb) et b (264,1 pb).

Les îlots CpG sont cités dans la littérature comme souvent placés à la fin des gènes (Bird 1986 ; Cross et Bird 1995 ; Maegawa *et al.*, 2010), mais les îlots des gènes DNMT chez *N. furzeri* semblent plus dispersés sans être liés aux promoteurs, souvent situés en amont des gènes. Les promoteurs des gènes peuvent contenir des îlots CpG et de ce fait être silencés lors de la méthylation de ces îlots. Chez les mammifères, environ 50% des gènes sont associés à des îlots CpG (Delgado *et al.*, 1998) et ceux-ci sont rarement méthylés (Toyota et Issa 2005 ; Jones et Baylin 2007). En effet, la méthylation des îlots aux promoteurs est souvent associée à la régulation des cellules cancéreuses et au silençage de gènes transposables.

La méthylation gène spécifique peut impacter le développement d'un organisme (Foret *et al.*, 2012), et peut également être un marqueur de vieillissement (Madrigano *et al.*, 2012). L'existence d'îlots CpG sur les DNMTs permettent d'hypothétiser une implication de la méthylation de l'ADN sur les gènes qui en sont responsables.

4.3 Perspectives

La méthylation possède de nombreux intervenants autres que les protéines DNMT comme TET3 ou Aid/apobec. Étudier ces protéines chez *N. furzeri* peut être intéressant dans l'étude de la reprogrammation de la méthylation.

Les HDAC (histones déacétylase) sont impliqués dans les phénomènes de méthylation d'histones qui sont des régulations épigénétiques et peuvent également être des intervenants intéressants à étudier dans le cadre d'une étude des mécanismes épigénétiques agissant chez *N. furzeri*.

4.4 Conclusion

Les gènes DNMT sont des gènes proches entre eux car ceux-ci sont conservés depuis l'apparition de la méthylation chez les eucaryotes. On peut cependant observer des différences fonctionnelles et structurelles entre les familles principales. Chez *N. furzeri* seulement trois familles sont présentes, DNMT1, DNMT3A et DNMT3B. DNMT3A et DNMT3B sont les plus proches que ce soit au niveau fonctionnel ou séquentiel, car les deux sont responsables de la méthylation *de novo* alors que DNMT1 est responsable du maintien de la méthylation entre les divisions cellulaires. Ces trois groupes sont relativement anciens d'après la phylogénie des vertébrés mais peuvent présenter des différences entre des groupes animaux phylogénétiquement proches. Le nombre et la taille des îlots CpG ne semblent pas être liés à la taille ou la fonction des DNMTs. La présence d'îlots CpG sur les gènes DNMT peut laisser hypothétiser une implication de la régulation par méthylation de l'ADN des gènes qui en sont responsables.

5. Conclusion générale

Dans cette étude préliminaire, les objectifs étaient de mettre au point un élevage efficace et l'analyse ainsi que la reconnaissance de stades de développement embryonnaire, afin de pouvoir évaluer la méthylation globale de *Nothobranchius furzeri*, un poisson au cycle de vie court. Le système d'élevage est actuellement fonctionnel et en voie d'être agrandi. Les stades de développement embryonnaire ont été observés et il est désormais possible de récupérer les différents stades de développement de *N. furzeri*.

Un autre objectif était la visualisation de la reprogrammation de l'ADN lors du développement de l'embryon. Cependant, devant l'incapacité à obtenir des échantillons d'ADN de suffisamment bonne qualité pour permettre l'analyse, il a été impossible de

mesurer le pourcentage de méthylation. La méthylation globale a pu être évaluée pour les gonades des organismes adultes et ces résultats montrent une corrélation avec d'autres espèces dont la méthylation globale des organes sexuels a été mesurée. La méthylation plus importante des gonades mâles est cohérente avec les éléments de la littérature démontrant une hyperméthylation des spermatozoïdes envers les ovules. Ce qui renforce la théorie d'une hérédité paternelle du méthylome.

Enfin l'analyse *in silico* du génome de *N. furzeri* présent dans les bases de données informatiques ont permis de réaliser une analyse préliminaire sur les gènes d'intérêt que sont les DNMT dans l'étude des modifications épigénétiques et de la transmission du méthylome. Trois familles de DNMT ont pu être mises en évidence : la DNMT1, responsable du maintien de la méthylation au cours des divisions et les DNMT3A et DNMT3B, responsables de la méthylation *de novo*. Les deux DNMT3 montrant une plus forte ressemblance entre elles qu'avec DNMT1. La mise en évidence des îlots CpG pouvant se trouver sur les gènes des DNMT peut être une pré-analyse de l'évaluation de la méthylation des gènes DNMT spécifiques.

Remerciement

Merci à Antoine Wittorski et Frédéric Silvestre pour leur encadrement et leurs nombreux conseils. Merci à Victoria Suarez-Ulloa pour son aide pour les analyses bioinformatiques. Merci au laboratoire d'écologie, évolution and Conservation de la biodiversité de la KUL et plus particulièrement au docteur Eli Thore et à Charlotte Philippe pour leurs nombreux conseils sur l'entretien de *N. furzeri* ainsi qu'au docteur Sophie Vanhunsel et le laboratoire de neurobiologie et physiologie animale. Merci au laboratoire du Docteur Valenziano D.R. du Max Planck Institute, et à Luca Dolfi pour ces nombreux conseils sur l'entretien de *N. furzeri*.

Bibliographie

- Api, M., Biondi, P., Olivotto, I., Terzibasi, E., Cellerino, A., & Carnevali, O. (2017). Effects of Parental Aging During Embryo Development and Adult Life: The Case of *Nothobranchius furzeri*. *Zebrafish*, *00*(0), 112–123. <https://doi.org/10.1089/zeb.2017.1494>
- Ashe, A., Sapetschnig, A., Weick, E. M., Mitchell, J., Bagijn, M. P., Cording, A. C., ... Miska, E. A. (2012). PiRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell*, *150*(1), 88–99. <https://doi.org/10.1016/j.cell.2012.06.018>
- Barrett, R. D. H., Rogers, S. M., & Schluter, D. (2008). Natural Selection on a Major Armor Gene in Threespine Stickleback. *Science*, *322*(5899), 255 LP-257. Retrieved from <http://science.sciencemag.org/content/322/5899/255.abstract>
- Baumgart, M., Groth, M., Priebe, S., Testa, G., Dix, A., Ripa, R., ... Goethe-universit, K. D. J. W. (2014). RNA-seq of the aging brain in the short-lived fish *N. furzeri* - conserved pathways and novel genes associated with neurogenesis. *Aging Cell*, *13*, 965–974. <https://doi.org/10.1111/acel.12257>
- Benayoun, B. A., Pollina, E. A., & Brunet, A. (2015). Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nature Reviews. Molecular Cell Biology*, *16*(10), 593–610. <https://doi.org/10.1038/nrm4048>
- Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature*, *321*(6067), 209–213. <https://doi.org/10.1038/321209a0>
- Bird, A. (2002). DNA methylation patterns and epigenetic memory, *Genes and development*, *16*, 6–21. <https://doi.org/10.1101/gad.947102.6>
- Bird, A. (2007). Perceptions of epigenetics. *Nature*, *447*(7143), 396–398. <https://doi.org/10.1038/nature05913>
- Blažek, R., Polačik, M., & Reichard, M. (2013). Rapid growth, early maturation and short generation time in African annual fishes. *EvoDevo*, *4*(1), 1–7. <https://doi.org/10.1186/2041-9139-4-24>
- Blewitt, M. E., Vickaryous, N. K., Paldi, A., Koseki, H., & Whitelaw, E. (2006). Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. *PLoS Genetics*, *2*(4), 399–405. <https://doi.org/10.1371/journal.pgen.0020049>
- Bonduriansky, R., & Day, T. (2009). Nongenetic Inheritance and Its Evolutionary Implications. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 103–125. <https://doi.org/10.1146/annurev.ecolsys.39.110707.173441>
- Bowler, Peter J. (2003). *Evolution: the history of an idea*. Berkeley, CA. p73–75
- Brack, A. (1998). *The Molecular Origins of Life*, Cambridge University Press. p1-2

- Brink R. A. (1973). Paramutation. *Annual Review of Genetics*, 7(1), 129-152
- Cachel, S. (1986). The Growth of Biological Thought revisited. *American Anthropologist*, 88. <https://doi.org/10.1525/aa.1986.88.2.02a00160>
- Cai, Y., Yu, X., Hu, S., & Yu, J. (2009). A Brief Review on the Mechanisms of miRNA Regulation. *Genomics, Proteomics and Bioinformatics*, 7(4), 147–154. [https://doi.org/10.1016/S1672-0229\(08\)60044-3](https://doi.org/10.1016/S1672-0229(08)60044-3)
- Callebaut I., Courvalin J.-C. and Mornon J.-P. (1999), The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation, *FEBS Letters*, 446, doi: 10.1016/S0014-5793(99)00132-5
- Cantone, I., & Fisher, A. G. (2013). Epigenetic programming and reprogramming during development. *Nature Structural and Molecular Biology*, 20(3), 282–289. <https://doi.org/10.1038/nsmb.2489>
- Cellerino, A., Valenzano, D. R., & Reichard, M. (2016). From the bush to the bench: The annual *Nothobranchius* fishes as a new model system in biology. *Biological Reviews*, 91(2), 511–533. <https://doi.org/10.1111/brv.12183>
- Cellerino, A., Dolfi, L., Ripa, R., Antebi, A., & Valenzano, D. R. (2019). Cell cycle dynamics during diapause entry and exit in an annual killifish revealed by FUCCI technology. *BioRxiv*, 522417. <https://doi.org/10.1101/522417>
- Cedar, H., & Bergman, Y. (2012). Programming of DNA Methylation Patterns. *Annu. Rev. Biochem.*, 81, 97–117. <https://doi.org/10.1146/annurev-biochem-052610>
- Chilundo, M., Kelderman, P., & Ókeeffe, J. H. (2008). Design of a water quality monitoring network for the Limpopo River Basin in Mozambique. *Physics and Chemistry of the Earth*, 33(8–13), 655–665. <https://doi.org/10.1016/j.pce.2008.06.055>
- Cleaver, T. G., Riggs, C. L., Romney, A. L., Woll, S. C., Wagner, J. T., & Culpepper, K. M. (2017). Embryonic Development of the Annual Killifish *Austrofundulus limnaeus* : An Emerging Model for Ecological and Evolutionary Developmental Biology Research and Instruction, 779–801. <https://doi.org/10.1002/DVDY.24513>
- Corsi, P. (2001) Lamarck, genèse et enjeux du transformisme, 1770-1830. CNRS ÉDITIONS, Paris.
- Costa, F. F. (2008). Non-coding RNAs, epigenetics and complexity. *Gene*, 410(1), 9–17. <https://doi.org/https://doi.org/10.1016/j.gene.2007.12.008>
- Costa, W. J. E. M. (2018). Comparative morphology, phylogeny and classification of African seasonal killifishes of the tribe Nothobranchiini (Cyprinodontiformes : Aplocheilidae). *Zoological Journal of the Linnean Society*, (February), 1–21. <https://doi.org/10.1093/zoolinnean/zlx102/4857373>

- Costantini, M., Auletta, F., & Bernardi, G. (2007). Isochore patterns and gene distributions in fish genomes. *Genomics*, *90*(3), 364–371. <https://doi.org/10.1016/j.ygeno.2007.05.006>
- Cross, S. H., & Bird, A. P. (1995). CpG islands and genes. *Current Opinion in Genetics & Development*, *5*(3), 309–314. [https://doi.org/https://doi.org/10.1016/0959-437X\(95\)80044-1](https://doi.org/https://doi.org/10.1016/0959-437X(95)80044-1)
- Cross SH, Meehan RR, Nan X, Bird A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. *Nat Genet.* *16*(3) 256-259. doi:10.1038/ng0797-256. PMID: 9207790.
- Cruickshanks, H. A., McBryan, T., Nelson, D. M., VanderKraats, N. D., Shah, P. P., van Tuyn, J., ... Adams, P. D. (2013). Senescent cells harbour features of the cancer epigenome. *Nature Cell Biology*, *15*, 1495-1506. <https://doi.org/10.1038/ncb2879>
- Cubas P, Vincent C, Coen E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, *401*, 157–61
- Darwin, C. (1859). *On the origin of species by means of natural selection*. London: Murray.
- Das OP, Messing J. (1994). Variegated phenotype and developmental methylation changes of a maize allele originating from epimutation. *Genetics*, *136*, 1121–41
- Day, D. A., & Tuite, M. F. (1998). Post-transcriptional gene regulatory mechanisms in eukaryotes : an overview. *Journal of Endocrinology*, *157*, 361–371.
- Day, J. J., & Sweatt, J. D. (2010). DNA methylation and memory formation. *Nature Neuroscience*, *13*(11), 1319–1323. <https://doi.org/10.1038/nn.2666>
- Delgado, S., Bird, A., & Antequera, F. (1998). Initiation of DNA replication at CpG islands in mammalian chromosomes. *The EMBO Journal*, *17*(8), 2426–2435.
- Desjardins, P., Conklin, D. (2010). NanoDrop Microvolume Quantitation of Nucleic Acids. *J. Vis. Exp.* (45), e2565, DOI : 10.3791/2565
- Di Cicco, E., Tozzini, E. T., Rossi, G., & Cellerino, A. (2011). The short-lived annual fish *Nothobranchius furzeri* shows a typical teleost aging process reinforced by high incidence of age-dependent neoplasias. *Experimental Gerontology*. <https://doi.org/10.1016/j.exger.2010.10.011>
- Dias, B. G., & Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations : *Nature Neuroscience* : Nature Publishing Group, *17*(1), 89–96. Retrieved from <http://www.nature.com/neuro/journal/vaop/ncurrent/full/nn.3594.html>
- Dorn, R., Krauss, V., Reuter, G., & Saumweber, H. (1993). The enhancer of position-effect variegation of *Drosophila*, E(var)3-93D, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *Proceedings of the*

National Academy of Sciences of the United States of America, 90(23), 11376–11380.
<https://doi.org/10.1073/pnas.90.23.11376>

Dorn, A., Ng'oma, E., Janko, K., Reichwald, K., Polačik, M., Platzer, M., ... Reichard, M. (2011). Phylogeny, genetic variability and colour polymorphism of an emerging animal model: The short-lived annual *Nothobranchius* fishes from southern Mozambique. *Molecular Phylogenetics and Evolution*, 61(3), 739–749.
<https://doi.org/10.1016/j.ympev.2011.06.010>

Douglas J. Futuyma (2013). *Evolution*. Troisième édition. Massachusetts USA : Sinauer Associates, Inc., 656.

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

Fellous, A., Labeled-Veydert, T., Locrel, M., Voisin, A.-S., Earley, R.L., Silvestre, F. (2018). ORIGINAL RESEARCH DNA methylation in adults and during development of the self-fertilizing mangrove rivulus, *Kryptolebias marmoratus*. *Ecology and Evolution*, 00, 1–18. <https://doi.org/10.1002/ece3.4141>

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791. Fricke, R., Eschmeyer, W. N. & R. van der Laan (eds) (2018). CATALOG OF FISHES: GENERA, SPECIES, REFERENCES. California academy of science (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Electronic version accessed 31/12/2018.

Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S. E., Robinson, G. E., & Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proceedings of the National Academy of Sciences*, 109(13), 4968–4973. <https://doi.org/10.1073/pnas.1202392109>

Froese, Rainer, and Daniel Pauly, eds. (2018). Species of *Nothobranchius* in FishBase. World Wide Web electronic publication. www.fishbase.org, version (06/2018).

Furness, A. I., Reznick, D. N., Springer, M. S., & Meredith, R. W. (2015a). Convergent evolution of alternative developmental trajectories associated with diapause in African and South American killifish. *Proceedings of the Royal Society B: Biological Sciences*, 282(1802), 20142189–20142189. <https://doi.org/10.1098/rspb.2014.2189>

Furness, A. I., Lee, K., & Reznick, D. N. (2015b). Adaptation in a variable environment: Phenotypic plasticity and bet-hedging during egg diapause and hatching in an annual killifish. *Evolution*, 69(6), 1461–1475. <https://doi.org/10.1111/evo.12669>

Gallego, M., & Virshup, D. M. (2007). Post-translational modifications regulate the ticking of the circadian clock, 8(February), 139–148. <https://doi.org/10.1038/nrm2106>

Gardiner-Garden M, Frommer M (1987). CpG islands in vertebrate genomes. *Journal of molecular biology*, 196 (2): 261-282. [10.1016/0022-2836\(87\)90689-9](https://doi.org/10.1016/0022-2836(87)90689-9).

- Genade, T., Benedetti, M., Terzibasi, E., Roncaglia, P., Valenzano, D. R., Cattaneo, A., & Cellerino, A. (2005). Annual fishes of the genus *Nothobranchius* as a model system for aging research. *Aging Cell*, 4(5), 223–233. <https://doi.org/10.1111/j.1474-9726.2005.00165.x>
- Genade, T. (2007). LABORATORY MANUAL FOR CULTURING *N. furzeri*. *Natural History*, 16. Retrieved from www.nothobranchius.info
- Genade, T. (2016). Journal of the American Killifish Association. *Journal of the American Killifish Association*, 49(1).
- Goll, M. G., & Halpern, M. E. (2011). DNA Methylation in Zebrafish. Modifications of Nuclear DNA and Its Regulatory Proteins. *Progress in Molecular Biology and Translational Science*, 101: 193–218. doi:10.1016/b978-0-12-387685-0.00005-6
- Graf, M., Kirschner, J., Hartmann, N., Reichwald, K., Dorn, A., Terzibasi, E., ... Englert, C. (2009). Telomeres shorten while Tert expression increases during ageing of the short-lived fish *Nothobranchius furzeri*, 130, 290–296. <https://doi.org/10.1016/j.mad.2009.01.003>
- Grégoir, A. F., Philippe, C., Pinceel, T., & Reniers, J. (2017). Life stage dependent responses to desiccation risk in the annual killifish *Nothobranchius wattersi*, 1–16. <https://doi.org/10.1111/jfb.13385>
- Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., Liu, W., ... Xu, G. L. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*, 477(7366), 606–612. <https://doi.org/10.1038/nature10443>
- Gunes, S., Hekim, G. N. T., Arslan, M. A., & Asci, R. (2016). Effects of aging on the male reproductive system. *Journal of Assisted Reproduction and Genetics*, 33(4), 441–454. <https://doi.org/10.1007/s10815-016-0663-y>
- Hackett, J. A., & Surani, M. A. (2015). Europe PMC Funders Group Beyond DNA : Programming and Inheritance of Parental Methylomes, *Cell*, 153(4), 737–739. <https://doi.org/10.1016/j.cell.2013.04.044.Beyond>
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-maarri, O., & Reik, W. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development*, 117, 15–23.
- Hartmann, N., & Englert, C. (2012). A microinjection protocol for the generation of transgenic killifish (Species: *Nothobranchius furzeri*). *Developmental Dynamics*, 241(6), 1133–1141.
- Head, J. A., Mittal, K., & Basu, N. (2014). Application of the LUMinometric Methylation Assay to ecological species: Tissue quality requirements and a survey of DNA methylation levels in animals. *Molecular Ecology Resources*, 14(5), 943–952. <https://doi.org/10.1111/1755-0998.12244>
- Holliday, R. & Pugh, J. E. (1975). DNA modification mechanisms and gene activity during development. *Science* 187, 226–232.

Holliday, R. (2006). Epigenetics: A historical overview. *Epigenetics*, 1(2), 76–80. <https://doi.org/10.4161/epi.1.2.2762>

Huang, K., & Fan, G. (2010). DNA methylation in cell differentiation and reprogramming: an emerging systematic view. *Review Regen. Med.*, 5, 531–544.

Iwasaki, Y. W., Siomi, M. C., & Siomi, H. (2015). PIWI-Interacting RNA: Its Biogenesis and Functions. *Annual Review of Biochemistry*, 84(1), 405–433. <https://doi.org/10.1146/annurev-biochem-060614-034258>

Jablonka, E., & Raz, G. (2009). Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. *The Quarterly Review of Biology*, 84(2), 131–176. <https://doi.org/10.1086/598822>

Jaenisch, R., & Young, R. (2008). Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell*, 132(4), 567–582. <https://doi.org/10.1016/j.cell.2008.01.015>

Jenuwein, T., & Allis, C. D. (2001). Translating the Histone Code. *Science*, 293(5532), 1074 LP-1080. Retrieved from <http://science.sciencemag.org/content/293/5532/1074.abstract>

Jiang, L., Zhang, J., Wang, J., Wang, L., Zhang, L., Li, G., ... He, C. (2013). Sperm , but Not Oocyte , DNA Methylome Is Inherited by Zebrafish Early Embryos, 773–784. <https://doi.org/10.1016/j.cell.2013.04.041>

Jones D.T., Taylor W.R., and Thornton J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8: 275-282.

Jones, P. A. (1999). The DNA methylation paradox. *Trends in Genetics*, 15(1), 34–37. [https://doi.org/10.1016/S0168-9525\(98\)01636-9](https://doi.org/10.1016/S0168-9525(98)01636-9)

Jones, P. A., & Baylin, S. B. (2007). The Epigenomics of Cancer. *Cell*, 128(4), 683–692. <https://doi.org/https://doi.org/10.1016/j.cell.2007.01.029>

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics*, 13(7), 484–492. <https://doi.org/10.1038/nrg3230>

Jubb, R.A. (1971) A new *Nothobranchius* (Pisces, Cyprinodontidae) from Southeastern Rhodesia. *Journal of the American Killifish Association*, 8, 12–19.

Kamstra, J. H., Aleström, P., Kooter, J. M., & Legler, J. (2015). Zebrafish as a model to study the role of DNA methylation in environmental toxicology, 16262–16276. <https://doi.org/10.1007/s11356-014-3466-7>

Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grandér, D., Schalling, M., ... Ekström, T. J. (2006). LUMA (LUMinometric Methylation Assay)-A high throughput method to the analysis of genomic DNA methylation. *Experimental Cell Research*, 312(11), 1989–1995. <https://doi.org/10.1016/j.yexcr.2006.03.006>

- Karni, Moshe, Zidon, Dolev, Polak, Pazit, Zeev Zalevsky, Shefi, and O. (2013). Thermal Degradation of DNA. *DNA AND CELL BIOLOGY*, 32(6), 1–4.
<https://doi.org/10.1089/dna.2013.2056>
- Kelley, J. L., Yee, M.-C., Brown, A. P., Richardson, R. R., Tatarenkov, A., Lee, C. C., ... 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.
<https://doi.org/10.1093/gbe/evw145>
- Keverne, E. B., & Curley, J. P. (2008). Epigenetics, brain evolution and behaviour. *Frontiers in Neuroendocrinology*, 29(3), 398–412.
<https://doi.org/10.1016/j.yfrne.2008.03.001>
- Kim, B.-M., Mirbahai, L., Mally, A., Kevin Chipman, J., Rhee, J.-S., & Lee, J.-S. (2015). Correlation between the DNA methyltransferase (Dnmt) gene family and genome-wide 5-methylcytosine (5mC) in rotifer, copepod, and fish. *Genes & Genomics*, 38(1), 13–23.
doi:10.1007/s13258-015-0333-y
- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. (1995). Stages of embryonic development of the zebrafish. *Dev Dyn Off public*, 203, 253–310.
- Kohli, R. M., & Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*, 502, 472–479. <https://doi.org/10.1038/nature12750>
- Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547-1549.
- Laing, L. V, Viana, J., Dempster, E. L., Webster, T. M. U., Aerle, R. Van, & Mill, J. (2018). Part A Sex-specific transcription and DNA methylation profiles of reproductive and epigenetic associated genes in the gonads and livers of breeding zebra fish, *Comparative Biochemistry and Physiology* 222, 16–25.
<https://doi.org/10.1016/j.cbpa.2018.04.004>
- Laird, C. D., Pleasant, N. D., Clark, A. D., Sneed, J. L., Hassan, K. M. A., Manley, N. C., ... Stöger, R. (2004). Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*, 101(1), 204–209.
<https://doi.org/10.1073/pnas.2536758100>
- Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics*, 11, 204–220.
<https://doi.org/10.1038/nrg2719>
- Lee, J. T., & Jaenisch, R. (1997). The (epi)genetic control of mammalian X-chromosome inactivation. *Current Opinion in Genetics & Development*, 7(2), 274–280.
[https://doi.org/https://doi.org/10.1016/S0959-437X\(97\)80138-4](https://doi.org/https://doi.org/10.1016/S0959-437X(97)80138-4)

- Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, 17(4), 183–189. [https://doi.org/10.1016/S0169-5347\(02\)02497-7](https://doi.org/10.1016/S0169-5347(02)02497-7)
- Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Reviews Genetics*, 3(9), 662–673. <http://doi.org/10.1038/nrg887>
- Lisanti, S., Omar, W. A. W., Tomaszewski, B., Prins, S. De, Jacobs, G., Koppen, G., ... Langie, S. A. S. (2013). Comparison of Methods for Quantification of Global DNA Methylation in Human Cells and Tissues, PLoS ONE, 8 (11). <https://doi.org/10.1371/journal.pone.0079044>
- Madrigano, J., Baccarelli, A., Mittleman, M. A., Sparrow, D., Vokonas, P. S., Tarantini, L., & Schwartz, J. (2012). Aging and epigenetics: Longitudinal changes in gene-specific DNA methylation. *Epigenetics*, 7(1), 63–70. <https://doi.org/10.4161/epi.7.1.18749>
- Maegawa, S., Hinkal, G., Kim, H. S., Shen, L., Zhang, L., Zhang, J., ... Issa, J. J. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. *Genome*, 20, 332–340. <https://doi.org/10.1101/gr.096826.109.332>
- Marchler-Bauer A et al. (2017), "CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.", *Nucleic Acids Res.* 45(D1):D200-D203.
- Markofsky, J., & Mayer, M. (1979). CHANGES IN THE LIVER OF THE MALE ANNUAL CYPRINODONT FISH, NOTHOBRANCHIUS GUENTHERI. *Exp. Oecron*, 14, 11–20.
- Matzke, M. A., & Mosher, R. A. (2014). RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nature Reviews Genetics*, 15(6), 394–408. <https://doi.org/10.1038/nrg3683>
- Mayer, W., Niveleau, A., Walter, J., Fundele, R., Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* 403, 6769, 501–502.
- Maegawa, S., Hinkal, G., Kim, H. S., Shen, L., Zhang, L., Zhang, J., ... Issa, J. J. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. *Genome*, 20, 332–340. <https://doi.org/10.1101/gr.096826.109.332>
- Mhanni, A. A., Yoder, J. A., Dubesky, C., & It, D. (2001). Cloning and Sequence Analysis of a Zebrafish cDNA Encoding DNA (Cytosine-5) -Methyltransferase-1, 219, 213–219. <https://doi.org/10.1002/gene.1067>
- Mikkelsen, T. S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., ... Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature*, 454(7200), 49–55. <https://doi.org/10.1038/nature07056>
- Moczek A., Sultan S., Foster S., Ledon-Rettig C., Dworkin I., Nijhout F., Abouheif E. and Pfennig D. (2011). The role of developmental plasticity in evolutionary innovation. *Proc. R. Soc. B*, 278 : 2705–2713. <http://dx.doi.org/10.1098/rspb.2011.0971>.

- Morgan, H.D., Dean, W., Coker, H.A., Reik, W. & Petersen-Mahrt, SK. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *Journal of Biological Chemistry*, 279, 52353–52360.
- Mourabit, S., Edenbrow, M., Croft, D. P., & Kudoh, T. (2011). Embryonic Development of the Self-fertilizing Mangrove Killifish *Kryptolebias marmoratus*, (June), 1694–1704. <https://doi.org/10.1002/dvdy.22668>
- Naumann B. & Englert C., (2018). Dispersion/reaggregation in early development of annual killifishes: Phylogenetic distribution and evolutionary significance of a unique feature. *Developmental Biology*, 442 (1): 69-79. <https://doi.org/10.1016/j.ydbio.2018.07.015>
- O'Neill C. (2013). Lessons from zebrafish on reprogramming the epigenetic code after fertilisation. *Asian journal of andrology*, 15(5), 582-3.
- Painter, R., Osmond, C., Gluckman, P., Hanson, M., Phillips, D. and Roseboom, T. (2008), Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *BJOG: An International Journal of Obstetrics & Gynaecology*, 115, 1243-1249. <https://doi.org/10.1111/j.1471-0528.2008.01822.x>
- Pfeifer, G.P., Steigerwald, S.D., Hansen, R.S., Gartler, S.M., and Riggs, A.D. (1990). Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: Methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. *Proc. Nat. Acad. Sci.*, 87, 8252–8256.
- Platzer, M., & Englert, C. (2016). *Nothobranchius furzeri*: A Model for Aging Research and More. *Trends in Genetics*, 32(9), 543–552. <https://doi.org/10.1016/j.tig.2016.06.006>
- Podrabsky, J. E., & Hand, S. C. (2000). Depression of protein synthesis during diapause in embryos of the annual killifish *Austrofundulus limnaeus*. *Physiological and Biochemical Zoology : PBZ*, 73(6), 799–808. <https://doi.org/10.1086/318106>
- Polacik, M. and Reichard, M. (2010). Diet overlap among three sympatric African annual killifish species *Nothobranchius* spp. from Mozambique. *Journal of Fish Biology*, 77, 754–768. <https://doi.org/10.1111/j.1095-8649.2010.02717.x>
- Polacik, M., Donner, M. T., & Reichard, M. (2011). Age structure of annual *Nothobranchius* fishes in Mozambique : is there a hatching synchrony ?, 78, 796–809. <https://doi.org/10.1111/j.1095-8649.2010.02893.x>
- Polačik, M., Blažek, R., & Reichard, M. (2016). Laboratory breeding of the short-lived annual killifish *Nothobranchius furzeri*. *Nature Protocols*, 11(8), 1396–1413. <https://doi.org/10.1038/nprot.2016.080>
- Ponger, L., & Li, W. H. (2005). Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. *Molecular Biology and Evolution*, 22(4), 1119–1128. <https://doi.org/10.1093/molbev/msi098>

- Potok, M. E., Nix, D. A., Parnell, T. J., & Cairns, B. R. (2013). Reprogramming the Maternal Zebrafish Genome after Fertilization to Match the Paternal Methylation Pattern. *Cell*, 153(4), 759–772. <https://doi.org/10.1016/j.cell.2013.04.030>
- Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A. & Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. *Cell*, 135, 1201–1212.
- Reichwald, K., Lauber, C., Nanda, I., Kirschner, J., Hartmann, N., Schories, S., ... Platzer, M. (2009). Open Access High tandem repeat content in the genome of the short-lived annual fish *Nothobranchius furzeri*: a new vertebrate model for aging, 10(2). <https://doi.org/10.1186/gb-2009-10-2-r16>
- Reichwald, K., Petzold, A., Koch, P., Downie, B. R., Hartmann, N., Pietsch, S., ... Platzer, M. (2015). Insights into Sex Chromosome Evolution and Aging from the Genome of a Short-Lived Fish. *Cell*, 163(6), 1527–1538. <https://doi.org/10.1016/j.cell.2015.10.071>
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature*, 447(7143), 425–432. <https://doi.org/10.1038/nature05918>
- Reynolds, J. A. (2017). Epigenetic Influences on Diapause. *Insect Epigenetics* (1st ed., Vol. 53). Elsevier Ltd. <https://doi.org/10.1016/bs.aiip.2017.03.003>
- Reichard, M., Polačik, M., & Sedláček, O. (2009). Distribution, colour polymorphism and habitat use of the African killifish *Nothobranchius furzeri*, the vertebrate with the shortest life span. *Journal of Fish Biology*, 74(1), 198–212. <https://doi.org/10.1111/j.1095-8649.2008.02129.x>
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16(6), 276–277. doi:10.1016/s0168-9525(00)02024-2
- Riggs, A. D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* 14, 9–25.
- Riggs, A.D., Xiong, Z., Wang, L., and LeBon, J.M. (1998). Methylation dynamics, epigenetic fidelity and X chromosome structure. *Novartis Found. Symp.* 214, 214–225.
- Rougier, N., Bourc, D., Gomes, D. M., Niveleau, A., & Plachot, M. (1998). Chromosome methylation patterns during mammalian preimplantation? development service Chromosome methylation patterns during mammalian preimplantation development, (Bird 1986), 2108–2113. <https://doi.org/10.1101/gad.12.14.2108>
- Rountree MR, Bachman KE, Baylin SB. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet.* 25(3) 269-277. doi:10.1038/77023. PMID: 10888872.
- Sahm, A., Bens, M., Platzer, M., & Cellerino, A. (2017). Parallel evolution of genes controlling mitonuclear balance in short-lived annual fishes. *Aging Cell*, 16, 488–496. <https://doi.org/10.1111/acel.12577>

- Santos, F., & Dean, W. (2004). Epigenetic reprogramming during early development in mammals. *Reproduction*, 127(6), 643–651. <https://doi.org/10.1530/rep.1.00221>
- Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., ... Reik, W. (2012). The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. *Molecular Cell*, 48(6), 849–862. <https://doi.org/10.1016/j.molcel.2012.11.001>
- Skinner, M. K. (2011). Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics*, 6(7), 838–842. <https://doi.org/10.4161/epi.6.7.16537>
- Skinner, M. K. (2015). Environmental epigenetics and a unified theory of the molecular aspects of evolution: A neo-Lamarckian concept that facilitates neo-Darwinian evolution. *Genome Biology and Evolution*, 7(5), 1296–1302. <https://doi.org/10.1093/gbe/evv073>
- Smith, Z. D., Chan, M. M., Mikkelsen, T. S., Gu, H., Gnirke, A., Regev, A., & Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*, 484(7394), 339–344. <https://doi.org/10.1038/nature10960>
- Stam, M., & Scheid, O. M. (2005). Paramutation: An encounter leaving a lasting impression. *Trends in Plant Science*, 10(6), 283–290. <https://doi.org/10.1016/j.tplants.2005.04.009>
- Stec I, Nagl SB, van Ommen GJ, den Dunnen JT. (2000). The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett.*, 473(1) 1-5. doi:10.1016/s0014-5793(00)01449-6.
- Sultan, SE. (1996). Phenotypic plasticity for offspring traits in *Polygonum persicaria*. *Ecology*, 77,1791–807.
- Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nature Reviews Genetics*, 9(6), 465–476. <https://doi.org/10.1038/nrg2341>
- Teperek-Tkacz, M., Pasque, V., Gentsch, G., & Ferguson-Smith, A. C. (2011). Epigenetic reprogramming: Is deamination key to active DNA demethylation? *Reproduction*, 142(5), 621–632. <https://doi.org/10.1530/REP-11-0148>
- Terzibasi, E., Riccardo, D., & Cellerino, A. (2007). The short-lived fish *Nothobranchius furzeri* as a new model system for aging studies. *Experimental Gerontology* 42, 42, 81–89. <https://doi.org/10.1016/j.exger.2006.06.039>
- Tobi, E. W., Goeman, J. J., Monajemi, R., Gu, H., Putter, H., Zhang, Y., ... Heijmans, B. T. (2014). DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications*, 5, 5592. Retrieved from <https://doi.org/10.1038/ncomms6592>

Toyota, M., & Issa, J.-P. J. (2005). Epigenetic Changes in Solid and Hematopoietic Tumors. *Seminars in Oncology*, 32(5), 521–530.

<https://doi.org/https://doi.org/10.1053/j.seminoncol.2005.07.003>

Tozzini, T.E., Dorn, A., Ng'Oma, E., Polačik, M., Blažek, R., Reichwald, K., ... Cellerino, A. (2013). Parallel evolution of senescence in annual fishes in response to extrinsic mortality. *BMC Evolutionary Biology*, 13(1). <https://doi.org/10.1186/1471-2148-13-77>

Valdesalici, S., Cellerino, A. (2003). Extremely short lifespan in the annual fish *Nothobranchius furzeri*. *Proc. Roy. Soc. Lond. B*, 270, 189–191.

Valenzano, D. R., Benayoun, B. A., Singh, P. P., Zhang, E., Etter, P. D., Hu, C. K., ... Brunet, A. (2015). The African Turquoise Killifish Genome Provides Insights into Evolution and Genetic Architecture of Lifespan. *Cell*, 163(6), 1539–1554.

<https://doi.org/10.1016/j.cell.2015.11.008>

Vermeulen, F. B. M., & Hrbek, T. (2005). *Kryptolebias sepia* n. sp. (Actinopterygii: Cyprinodontiformes: Rivulidae), a new killifish from the Tapanahony River drainage in southeast Surinam, *Zootaxa* 20, 1–20.

Wang X. & Bhandari R. K. (2019): DNA methylation dynamics during epigenetic reprogramming of medaka embryo, *Epigenetics*, DOI:10.1080/15592294.2019.1605816

Wiegand M. (1996) Composition, accumulation and utilization of yolk lipids in teleost fish. *Reviews in Fish Biology and Fisheries*, 6 (3), Pages 259.

Wourms, J. P. (1972a). The Developmental Biology of Annual Fishes. II. NATURALLY OCCURRING DISPERSION AND REAGGREGATION OF BLASTOMERES DURING THE DEVELOPMENT OF ANNUAL FISH EGGS. *Experimental Zoology*, 182, 169–200.

Wourms, J. (1972b). The developmental biology of annual fishes. III. Pre-embryonic and embryonic diapause of variable duration in the eggs of annual fishes. *J. Exp. Zool.* 182, 389–414.

Wu, X., & Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nature Reviews Genetics*, 18, 517. Retrieved from

<https://doi.org/10.1038/nrg.2017.33>

Youngson, N. A., & Whitelaw, E. (2008). Transgenerational Epigenetic Effects. *Annual Review of Genomics and Human Genetics*, 9(1), 233–257.

<https://doi.org/10.1146/annurev.genom.9.081307.164445>

Zupkovitz, G., Lagger, S., Martin, D., Steiner, M., Hagelkruys, A., Seiser, C., ... Pusch, O. (2018). Histone deacetylase 1 expression is inversely correlated with age in the short-lived fish *Nothobranchius furzeri*. *Histochemistry and Cell Biology*, 150(3), 255–269.

<https://doi.org/10.1007/s00418-018-1687-4>

Annexes

Annexe 1 : *N. furzeri* adultes utilisés dans l'analyse avec le numéro d'identification, le sexe, le poids du poisson et le poids de la gonade.

N°	Semaine	Sexe	Abréviation	Poids Poisson (g)	Poids Gonade (g)	Remarques
1	5 semaines	M	15M	0,2696	0,0025	
3	5 semaines	M	35M	0,3105	0,0035	
4	5 semaines	M	45M	0,2118	0,0017	
5	5 semaines	M	55M	0,2061	0,0018	
6	5 semaines	M	65M	0,2209	0,0115	
7	5 semaines	M	75M	0,2427	0,0124	
8	5 semaines	M	85M	0,234	0,001	
9	5 semaines	M	95M	0,241	0,0025	
10	5 semaines	M	105M	0,1788	0,0116	
7	5 semaines	F	75F	ND	0,0144	Mâle à la dissection
1	5 semaines	F	15F	0,1756	0,0155	
2	5 semaines	F	25F	0,1878	0,0166	
3	5 semaines	F	35F	0,181	0,0126	
4	5 semaines	F	45F	0,138	0,0071	
5	5 semaines	F	55F	0,1961	0,0231	
6	5 semaines	F	65F	0,1652	0,0182	
8	5 semaines	F	85F	0,18	0,018	
9	5 semaines	F	95F	0,1569	0,0143	
10	5 semaines	F	105F	0,1902	0,0192	
11	5 semaines	F	115F	ND	ND	
1	9 semaines	M	19M	0,9031	0,0079	
2	9 semaines	M	29M	0,8333	0,0083	
3	9 semaines	M	39M	0,6722	0,0156	
4	9 semaines	M	49M	0,6723	0,0106	
5	9 semaines	M	59M	0,7699	0,0105	
6	9 semaines	M	69M	0,7291	0,0119	
7	9 semaines	M	79M	0,734	0,006	
8	9 semaines	M	89M	0,6956	0,0087	
9	9 semaines	M	99M	0,7622	0,0087	
10	9 semaines	M	109M	0,7244	0,0074	
1	9 semaines	F	19F	0,6723	0,08	
2	9 semaines	F	29F	0,6261	0,0705	
3	9 semaines	F	39F	0,5424	0,0719	
4	9 semaines	F	49F	0,4572	0,054	
5	9 semaines	F	59F	0,7383	0,0998	
6	9 semaines	F	69F	0,544	0,0654	
7	9 semaines	F	79F	0,655	ND	
8	9 semaines	F	89F	0,5835	0,0653	
9	9 semaines	F	99F	0,7104	0,0867	
10	9 semaines	F	109F	0,5217	0,0658	

Annexe 2 : Séquences protéiques utilisées pour l’alignement des séquences et l’identification des domaines ainsi que la comparaison deux à deux dans Dialign Genomatix.

>N. furzeri DNMT1 1507aa

MPSRTLSLPPDVRKRLELLDDEGCSSEEEQVKKKIKLVQDFLHDDAQDQLTSLLEEKIKKSEISKEYVYISKVKALLGKELHLENGSHV
 DGVERNEKRNGLSNGSHKDEEDGDGDMTQEEDTSSSVKSPSTTKGKGGRRSKGSADTKKSPASSRVTRNSVKQPTILSMFSK
 VQKRKSDEVNGEPVNETSEPKKEEADESREEKRLKVESDEKPAEEPKPEVNRSASAAKTPPPKCPDCRQYLLDSDLKFFQGD
 NALEPEMLTDERLSLFDANEDGFESYDDLPHQKITSFSVYDKHGHLCPPFDSGLIEKNVELYFSCVVKPIYDDNPCLDGGVSAKLLG
 PINAWWITGFDGGEKALIGFTTSFADYILMQPSEEYAPIFALMQEKIYMSKIVMEFLQKNPDATYEDLLNRIETTVPAGLNFCNTE
 DTLLRHAQFVVEQVESYDEAGSDSQPIIVTPCMRDLIKLAGVTLGKSMMLYWRRAARRQAIRHPTKIEKDNKGPTKATTTTTLVYQI
 FDAFFSGQIEQNDKEGGVQRQRCGVCEVCQSPDCGKCAACKDMIKFGGSGRSKQGCRRRCPNLAVKEAEDDENIEEDVPEEKLK
 KAPNTRKRKQTSKLAWIEEPVQTVGKKQFYKKSVDNDELEVGDVSVSSDDPSFPLYLARITSLWEDNNGKMFHAWFLRGIH
 TVLGESSDPLELVVDECEMMLNLYVQKVSVTYKAPSNNWSMEGGVDVLDKVIEDDGKSFYQFVYDVFARFETPPETLSE
 CRFSFCGSCTRTKEREEQEIIPRSEPLENEDSDTKALYGLVCFKREQFRVGDGVYLPDADFHFVSPASPLKRSHRKEDVDELDY
 YRKRSSDYIKGSLNDAPEPFRIGRIKIEFCHRRSTGKSDTSDVKLRLYKFRPENTHKGVKAGYHTDINQLYWSEEEVTVSVDVGL
 RCQVEYAEADLNEVQESYSSGGPDRFYFLEAYNAKSKSFEDPPNHARSSVHKGKGKGGKGGKAAQEPQDPHTEPPKVPKYR
 TLDVFSGGCGLSEGFHQAGISETLWAIEMWEPAAQAFRLNPNPVTFTEDCNVLLKLVMSGEKTNLQKLPQKGDVEMLCGGPP
 CQGFSGMNRFNRSRTYSKFKNSLVVSYLSYCDYRPFKFLLENVRFVSKSSMVLKTLRCLVRMGYQCTFGVLQAGQYGVAT
 RRRAILAAAPGEKLPYRYPELVHFAFRACQLSVVDEKKYVSNATRNGNGGIYRITVRDMSDLPEIRNGAAALEISYNGEPQSWF
 QRQIRGTQYQPILRDHICKDMSALVEARMRHIPLAPGSDWRDLNIEVRMKDGNNTKRLRYTHHDKNGRSSTGALRGVCTCAGG
 KPCDPLDRQFNTLIPWPLHTGPNRHNHAWAGLYGRLEWDGFFSTVTNPEPMGKQGRVLAHQHRVSVRECARSQGFPDTPYRF
 GNILDKHRQVGNVPPPLSRAIGLEIKKCVIERIKQEQAEDNIKQEKMEA

>N. furzeri DNMT3b 1458aa

MDRRTPEEPARNQSLTNLSQNVSVISLISEDEEDIVCVSLEQTENAKEDCECFQKQNLTAAPGVDMENAGDHHGHCDLVGFIRRC
 CGTHVYSPDPKEIVQAVSNTPFCLYLTVYVQVDEGHSTFVLISYFDQSSGRSVVLLHTLQISTDTAYPQVSGDQDAVTRAARSDA
 CLLIRLTKLELPLSNMAVFYCEPQLGEMCVSGLQTFSPNVISLCSIPGIAARASQAALTAFTVHVDLILDIHHHYSTCTISINDSLKE
 VFANTQSYNPLCPISTQCSFVTHSVQKMAVARWQDLLDYFQSLKQTEVTNQIRMRLMDENVRLHFLCHALEPLKALQELQEAQT
 KDLAVELQLTSILLQSYAASILRSSAVSGLWRQDLQLLSSEKDLLPLAEVNIQSQTDLMWAMGLNGGERDMIGFLKRAQSFYKA
 ALQSLMESLPEALGDVTLKNIKDIVKDPKNFTVRHFPKKVLSQIGVQLGLCAVGPTENQLASSYAFQLMQLSQESNWKECNSWAQ
 VLSQMRRTSLHKLTLTLLTMPSSLHRTRVFEAYPNVRMPHREENAENSPLLKPRRSRPSVLKRRPPSLARHAGQRREQELEETS
 GDSSTAAKPLSKTTQLDDRIEDTNSSDVIDITEGSKMSFTSSQVTPQRKGSQGEIGLILTSDEEDASLDTATTPRSEELPEGFTE
 ELVWGMIVGFPPWPGHIVPNKRQGHPPGMIMVEWYVQKMCQFVSLKTLRPFVAFGLHFSANFYATFITYRESIYTSLQEAALRCNK
 KFPTELEDGEELLKHLLDWAFGGFQPSGPAFGKPAANGEMTRPVVTKLVGKVDLFTNQSDPMKNTKLSSEMNNAAAAPLNNL
 SCTSFQSKTDNRNKFDPKRSKSLERKGSRTLSTEMEKHGSLEEEKDEWAKEWEKMKPVSGRRSSRGSTRVGRSLKKTFGPFY
 GLDDDMSPDFVPSKKNFVKTYNKGSAQTSVYTPQDKLREETIHRIMNLKLDIEGFCLCCGTENIEISHPLFQGSCLCKCKDNFTE
 TYRDEEDGYQSYCTICCYGLEVILCGNDSCCRYSCECLNILVGPPTFESLKHLDPWICYLCPHLPHGALVPRVDVSVQVQELF
 ANNSAMEFEPHRVYPSIPAYLRRPVRVLSLFDGIGTGYLVKDLGFKVEKYVASEICEDSVCVSTINHDGKVS HVGDVRFITQENLE
 KWGPFDDL VGGSPCNDLSIVNPIRKGLYEGTGRLLFFEFYRLLQLLPKEDNPLPFFWLFENVVMNTHDKVNICRFLECNPVLVDAV
 KVSPAHRARYFWGNIPGMSRPIIASQNDKLSLQECLEIGREARVTKVRTITTPNSLQKQKDVSLLPVLQNGSEDNLWITELEKIFGF
 PKHYTDVRNLRQQRKLLGKAWSVVPVIRHLFAPLKEYFACEELPLLTPVSTSSVSPPLPHTLQTSTSFQQ

>N. furzeri DNMT3a 977aa

MPSNTVANTNTTSFKSSDFADSPGNRLGESTILDMDLMEEKSALSASCKAGKPKGRKRKQLPVESSSLAGSLNVSKSYSGVRRPF
 MAQVHNGEMDTHEDKMPDACPQKQERREVENGISLELSSCAEDSSQDRSLSPSNENGLSSQDEQDSEKDKDDGLMMQRKRG
 RRKLERPTKYVEQKEDGCAIKTEGGRGRQGGVWEISLRQRPMPRVTFQAGDPYIYISKRTREEMAKWKMEAEKAKVMS
 VMNMTKLEQNTKQTEEEVSVIKHPQSTPQEPQHSQTLSRPKYQQSQPCLPQQQQQPTPEASPTVATTPDPVAIGVEDK
 ASPKSADAPEYEDGRGFGIGDLVWGLRGFSWVWGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSFSNAFH
 QPTYNKQPMYKKAIEYVQLQVASSRAGKSFITCPESDETETSKSVEMLNKQMIEWAMTGFQPTGPRGLQPPEEERNSEVYPMEMWYD
 EAAAATPPACKPRKSTAEKPKVKDIIDERTRELVYEVQRKRSIEDICSCGGLNVGLGHPLFAGGMCQCGCKNCFLECAQYQDD
 DGYQSYCTICCGREVLMCGNCCRCFCVCDLLVGGQAAHAIAKEDPWNCYMCCKPVGFLLERRSDVRLQHFFANNH
 DQDFDPKLYPPVMVEKRPVRVLSLFDGIATGLLVLELGIQVDRYVASEVCEDSITVIVRHQGRIMYVGDVNRVTRKHIEQWG
 PFDLVIGGSPCNDLSIVNPIRKGLYEGTGRLLFFEFYRLLHEARPKEDDRPFFWLFENVVMWVSDKRDISRFLCNPVMIDAKEVS
 AAHRARYFWGNLPGMNRPLSAMFTDRLDLQDCLEHGRTAKFGKVRTITTRNSIKQKQKQDHFVPMNEKEDILWCTEMERVFGF
 PVHYTDVSNMSRLARQLLGRSWSVPIRHLFAPLKYDFACV

>N. furzeri DNMT3a-like 848aa

MPSNSPAAAEPPETPENDASNDVSEDRAQDSPEEGTPASPRNRRVGRPSRKSQQLPEMATSDPCANASPTPKPEPNVSPSRKK
 RRRQKPEQTEKDKEDPEMDRNCDSPREGETGRLRRRPVPRVTFQAGDPYIYISRRQKEEWLSRWKMEVGDKQAERRAYRAEAINM
 MDDLSEVDFQKEEGPASPAPPPSQHTDPASPTVAVTPPEVARGDQAMPSEIEYQDGRGFSIGTLVFGKLRGFSWVWGRIVSWWMS
 GRSRAADGTRWVMWFGDGKFSVVCVEKLMPLSSFSFAHQPTYNKQPMYRKAIFEAALQVAVRAGRVPSPCDTSEDAEAVELQT
 RQMIEWAMTGFPLNGPQSLDPPEEEQNQFKEGYVEMWPEEAAATPPPAKKPRKNSAEKAKIREVIDEGTRERLIEIKKTRNIED
 ICISCGSLNVLEHPLFIGSMCQCGKNSFLECAQYQDDDGYSYCTICCGGREVLMCGNCCRCFCVCDLLVWAGSAAAAIKE
 DPWNCYMCGRSTYGLRRRDDWPRLQHFFANNHEQEFEPAKLYPPVAAEKRPVRLVSLFDGIATGLLVKDLGIQVQDYIASE
 VCEDSITVGLVRHQGRIMYVGDVNRVTRKHIEEWGPFDLVIGGSPCNDLSIVNPIRKGLYEGTGRLLFFEFYRLLHEARPKEDDRP
 FWFLENVVMWVSDKRDISRFLCNPVMIDAKEVSAHRARYFWGNLPGMNRPLSAMFTDRLDLQDCLEHGRTAKFGKVRTITTRNSIKQKQKQDHFVPMNEKEDILWCTEMERVFGF
 SNSVKQKQKDEHFPVMDNKEDILWCTEMERVFGFVHYTDVSNMSRLARQLLGRSWSVPIRHLFAPLKEYFACN

>N. furzeri DNMT3b-like 555aa

MEKHGSLEEEKDEWAKEWEKMKPVSGRRSSRGSTRVGRSLKKTFGPFYGLDDDMSPDFVPSKKNFVKTYNKGSAQTSVYTPQD
 QKLEETIHRIMNLKLDIEGFCLCCGTENIEISHPLFQGSCLCKCKDNFETLRYDEEDGYQSYCTICCYGLEVILCGNDSCCRYSYCD
 CLNIVGPTFESLKHLDIPHLPHGALVPRVDVSVQVQELFANNASAMEFEPHRVYPSIPAYLRRPVRVLSLFDGIGTGYL
 VLKDLGFKVEKYVASEICEDSVCVSTINHDGKVS HVGDVRFITQENLEKWGPFDDL VGGSPCNDLSIVNPIRKGLYEGTGRLLFFEFY
 RILQLLKPEDNPLPFFWLFENVVMNTHDKVNICRFLECNPVLVDAVKVSPAHRARYFWGNIPGMSRPIIASQNDKLSLQECLEIG
 REARVTKVRTITTPNSLQKQKDVSLLPVLQNGSEDNLWITELEKIFGFPHYTDVSNMSRLARQLLGRSWSVPIRHLFAPLKEYFACN

>N. furzeri LOC107391197 1453aa

MAGSEVVSRRSLDKSSRLYVVAWTRNLLKTDFTDVREMGSGACHCQMMDYTFPGSVDSLKVKFDAQSEDDCRHNYSLLEAFT
 ALGVTRTIPVDELIERDLKSSFEFLKWFVYAAANVNDKDYDPVKARNGQAIPLQSSAQRSSSDSGSDVEKNGERKTHPYTHKW
 KENFKWASPSSIGEYQTYCYLNCNNLNFTFKNGVAELTRHANTMKHTKAAAVWLPCSDAAVRFINSHCYTGSTEREQAPASFARCK
 LGSQYPDDIVSICRVYVYGGVAVGEEDA VSVLTVGYFNVEISRHCIRFLDALQLVDGAEDHAAA VVTLKKFDLSDNLA
 AAVYIGSNGDASEQIYSHLRELNPVVALGGLYTIADAACHAGLKLPNWAREMMVDLHAHYSSHSTNNDNLKALFGSDITENK
 AFYLNISCLRFCQLVTKILEVWTDLISYFSSCHEGDGKAKSICSQKDPK VQATFMFLQALPLQRFHSQIQAASWDNMSLILEEAS
 SLLGTYSYFLQPQAVIDYLLKPPDEILKSKELQLSFQDLNLGGMEDFIKASEATAVLSQVEEDALSFYVALTGCIVEKLPLSDLELK
 SVGQLLNPQGLDMTAAAVEKLGTKLVCSPEEVKELTAEFHEYQLTGEGAGDKTALVSLEKHWQTLDKDRTRTTSVFRKLVTL
 LSFPCPLDSQRVFAEALKDGDSEIFSETEDTAERDSDATSYSALRNSAEVKDSQDHTPETNGLDAAVKPCEVRLTKMSYLSGGVC
 GVNGGFQNEGTSRGGFGWESSLRQKPPARTVFQAGVGSWSKPIGLDADRNSGHKSKDEASPRKSTPTKGQHKYQDQGGKGFQMG
 ELVWGWKPKGFSMWPGMVMRWKTKSSPPGMRRVEWFGDMFSEIYTDKLLPFSAFSRFCFKNSFASLPLYREALYQIIEAAERCG
 KSFAHAEGKEQELKLMVDWAFEGFLPTGPKGLLPADSCPVSESADSSLSDSQPAPKRKYVFKSKAVAVMVFNRKIIENLKEKNKI
 IRDFCLSCSSPEINVEHPLFEGGLCLRCKENFSETLYRYDEDGYQSYCTVCCAGSEVILCGNSSCCRCFKDCLEILVAEGTFEKLKDI
 DPWNCFMCNPSQCEGNLKLRSDDRWRVKVQDFFANNSAMEFEPHRVYPSIPADQRRPIRVLISFDGIATGYLVLRLDGFKVERYIASEI
 CDDSIAVGMIKHEGKIEYVNDVNRNITRKHLEWGFDFLLIGGSPCNDLSMVNPLRKLGFEGTGRLFFEFYRILNFKPKKEGDDRRPFF
 WLFENVVMSAHDKSDICRFLECNPILIDAVKVSAPHRARYFWGNLPGMNRPLSSFIDDKVNLQDCLEVGRKAMFEKVRTITTKSN
 SIRQKVGSLPVTMNGKEDYLWCTEMEQIFGPKHYTDVNNMGRQRQVLRGRSWSVPVIRHLFAPLKDIFYECESGQ

Annexe 3 : Séquences nucléiques ARNm utilisées pour l'alignement des séquences et la comparaison deux à deux dans Dialign Genomatix

>N. furzeri DNMT1

CACAACACCCGGTTTCCGCGCGAAAAGACGGCCTGCGCGCGCTCGAACTGCTTGTAGCATCCGCTTGAAGAAGCGCCACTC
 GACTGGTAGTCATCTCTGAGTATTGCTGCGGCGATTGACGGCGGCTAACGCGTGTGTTTTATATCAGTTTTTATTGCCGTGTC
 ATCTGCTTCTGCTTTAGTTAGCTGGTCAACATGCCATCCAGAACGTCCTGTCTCTGCCAGACGATGTCAGGAAAAGGTTGG
 AGCTGCTGGATGAGGATGGATGCTCAGAAGAGGAAACAAGTGAAGAAGAAGATCAAGTTGGTGCAGGACTTCTGCATGATG
 ATGCTCAGGACCAACTGACCAGTCTTGAAGAAAAATAAAAAATCAGAGATCTCAAAGGAGGTTTACATCTCAAAGGTGAA
 AGCCTTGGTGGGCAAAGAACTCATCTTGAGAAATGGCTCACATGTTGATGGCGTGGAGCGTAACGAAAAAGAGAAATGGCCTC
 TCTAATGGCTCCATAAAGACGAAGAGGATGGAGATGGTGTATGTTTCCATGGATACCCAGGAAGAAGACACGTCATCATCTG
 TCAAGTCAACCAACCCTCCAAAGGGAAGGGGGGCGCAGGAGCAAGGGGAGCGCAGACACAAAAAATCACCAGCCAGC
 AGCAGAGTCACGAGGAACAGTGTGAAGCAGCAACCATCTGTCCATGTTTTCTAAAGTCCAAAAGCGCAAGTCTGATGAGG
 TGAACGGAGAACCCTGTCAATGAAACAAAGTGAACCTAAAAAAGAAGAGGCCGTTGATGAGTCTCGGGAGGAGAAACGCTGA
 AAGTGGAGTCGGATGAAAAACCTGCTGCTGAGGAACCAAGCCTGAAGTTAACAGGTCGCGTCAGCAGCTAAGACTCCGCC
 CCCCCAATGTCAGACTGCAGACAATACCTGGATGACTCCGATCTGAAATTTCTCAAGGAGATCTGATAATGCGCTGGAA
 GAGCCTGAAATGTTGACGGATGAGCGTTTTGCTCCTGTTGACGCAACAGGAGGATGGGTTTGAAGACTACGATGATCTGCCG
 AGCAACAAGTACAAGTTTCAAGTGTGACGACAAGCAGCAACTCTCTGCCGTTTCGACTCTGGGCTGATGAGAAGAAGCTG
 GGAGCTTACTTCAAGTGTGTCGTTAAACCCATCTACGACGACAACCCGCTGCTTGGACGGTGGCGTTTCTGCTAAAAAGCTCG
 GGCCATTAACGCTGGTGGATCACTGGTTTTGATGGAGGTGAAAAGGCTTTGATGGCTTACAACTTTTGGCCGACTAC
 ATCTTGATGACGCCAGCGGAGGAGTACGCTCCCTCTTCGCGCTCATGCAGGAGAAAAATCTACATGAGCAAGATGTGATGG
 AGTTTTTGCAGAAAAATCCCGACGCTACTTATGAGGACCTTCTTACAGAATAGAGACAACCTTCCGCTGCAGGACTCAAC
 TTTAACTGCTTCACTGAAGACACGCTGCTGCGCCACGCCCAGTTTGTGGTGGAGCAGGTGGAGAGTTACGACGAGGCTGGAG
 ACTCGGACGAGCAGCCCATCTTGTGACGCCCTGCATGAGAGATCTGATCAAACCTCGCTGGAGTCACTTATAGAAAAAGCAT
 GCTGCTGATCTGGCAGCCGCGCAGAAGACAGGCCATCCGTCACCAAGATAGAGAAGGACAACAAGGGCCCGGACAAA
 AGCCACCACCAACGCTGCTACTAGATCTTTGATGCTTCTTCTCAGGTACAGTTGAGCAGAACGATAAGGAGGTTGGAG
 TCAAAAGACAGCGCTGTGGAGTCTGTGAGGTGTGCCACTCTCTGACTGTGGGAAGTGTGCCGCTCAAAGACATGATCAA
 GTTTGGAGGAAGCGCAGAAGCAAGCAGGGTTGTAGGCAGAGAAGATGTCCAAACCTTGTGTGAAGGAGCGGAGGATGA
 TGAGAATATTGAAGAAGACGTTCCAGAAGAGAAGCTCAAAAAGGCTCCCAACACTAAGAGGAAGAAGCAGACCCAGTCCAA
 ACTGGCTGGATCGAAGAGCCCTGACAGTGTGGGAAGAAGCAGTCTACAAGAAGGTTCTGTGAATGACGAGGTTCTG
 GAAGTGGGAGACTGTGTTTCAAGTCTTCTGATGATCTTCTTCTTCTTACGAGGTAAGGTAAGGTAAGGTAAGGTAAGG
 AACAAATGGGAAGATGTTCCACGCCACTGGTTCCTTCGCGGATTCACACGGTGTCTGGAGAGTCTCCGATCCACTGGAGCT
 GGTGCTGTTGGACGAGGTGAGGACATGCTGCTCAACTACGTCACGGAAGAAAGTCAAGCTACGATAAGGCCCGTCTAAC
 AACTGGTCTATGGAGGGAGGGGTGGACGTTGACCTCAAGGTGATCGAGGACGATGGAAGAGGTTTCTTCTACAGTTCTGGT
 ATGATCAACTTTGCTCGATTTGAGACTCCCGGAGACCTTCCATCAGAAGAGTCCCGGTTCAAGTTCTGTGTCAGCTGC
 ACTCGTACTAAAGAGAGGGAAGAGCAGGAGATCCCTCGTCCATCTGAACCCCTGGAGAACGAAGATAGCGACACCAAGCT
 CTTTATGGTTTGTCTGCTTCAAGAGAGAGCAGTTCAGGGTGGGAGATGGTGTCTACCTGCCTCTGACGCTTCCATTTCACT
 GTGAAGCCTGCCAGTCCGCTGAAACGCTCCACAGGAAGAGGATGTGGATGAAGATCTGTATCTGAGTACTACAGGAAAT
 CCTCAGACTACATTAAGGGGTCAAACCTGGATGCTCCAGGCCGTTCCGTATCGGCCGCATCAAGGATCTTCTGTCAACGA
 CGGAGCACCGGGAAGTCCGACACCTCGGATGTCAAACCTGAGACTCTACAAGTTCTACAGGCCTGAGAACACTCACAAAGGCG
 TCAAAGCGGGTACCACACAGACATCAATCAGCTGTACTGGAGCGAAGAGGAGGTGACCGTCAAGCTGTGCGACGCTGTCGG
 CCGTGTCAAGTGGAGTACGCTGAAGACCTGAATGAATCCGTCAGGAGTATCCAGCGGTGGGCGGATCGCTTCTATTTCC
 TCGAGGCCCTACAATGCTAAATCCAAGAGCTTTGAAGATCCTCCGAACCTATGCCCGTCTCTCTGTCATAAAGGAAAGGAAAA
 AGGCAAAGGAAAAGGGAAGGGAAGGAAAGGCTGCGCAGGAACCGCAAGACCCTCACACAGAACCTCCCAAGTGCCCA
 AATATCGCACACTGGATGTGTTCTCCGGCTGTGGAGGACTCTCTGAGGGTTTCCACCAGGCGGGCATCTCCGAGACTCTCTGG
 GCCATAGAGATGTGGAACTGCAGCAGCGCTTACAGCTCAACAACCTGGCACCACGGTGTTCACGGGAGGACTGTAACG
 TCTTACTGAAGTTGGTACTGTCCGGGAGAAGCAATTCTTGGGAGAAGCTGCCTCAGAAGGGGATGTTGAGATGCT
 GTGTGGTGGACCTCCCTGTCAAGGTTTTAGTGGGATGAACCGTTCAAATCCCGGACTTACTCCAAGTCAAGAACTCTTGT
 GGTCTCGTATCTGAGCTACTGCGACTACTACAGACCAAGTTTTTCTGCTGGAGAACGTGAGGAACCTCGTTTCAATCAAAA
 GCTCCATGGTTTCTGAAGTGTGACGTTGCGCTGTCTGTCGGAATGGGCTACCAGTGCACCTTTGGGGTCTTACGGCGGGTACG
 TACGGCTGCGCCAGCCGTCGACGGGCGATCCTGCTGCTGCCCCAGGAGAGAAGTGCACCCCTTACCCTGAACCCCT
 CCACGTGTTTCTCCAGGCTGCCAGCTGACGCTGGTGGTGGAGCAAGAAAGAAATATGTACGCAATGCCACGCGGAGCAAC
 GGTGGGATCTATAGAACCATCACCGTCCGCGACACCATGTCCGATCTGCCGAGATTCGTAACGGTCCCGCAGCTTTGGAGAT
 TTCTACACCGCGAGCCTAGTCTGGTCCAGAGGCAGATCAGAGGCACGAGTACCAGCCATCCTAAGAGACCACATC
 TGCAAGGACATGAGCGCTGTGGTTGAGGCCAGAATGCGCACATCCCCTTGGCTCCAGGCTCGGAGATGGAGAGATCTTCCAA
 ATATTAGGTTCTGGATGAAAGTGGCAACACGACTAAGAAGCTGCTTACACGCATCAGTAAAAAGAGCCGCGGACGCA
 GCACTGGTGCACCTCAGAGGTGTCTGCACCTGTGACGAGGGAAACCTTGCAGCCCTTACAGAGGCAGTTTAAACCCCTGATT
 CCTGGTGTCTGCCTCACACTGAAACCCGCCAATCACTGGGCTGGACTTATGGCAGACTGGAGTGGGATGGCTTCTCAG

CACAACAGTCACCAACCTGAACCCATGGGAAAACAGGGTCGTGTTCTGCACGCCGAGCAGCACAGAGTGGTGAGTGCAGAGAGTGTGCACGCTCTCAGGGATTCCCTGACACCTACCGATTCTTTGGAAACATCCTGGACAAGCACAGACAGGTTGGAAATGCTGTTCTCTCTCTCCAGAGCCATCGGGCTGGAGATTAAGAAATGTGTCATAGAAAAGGATAAAGCAAGAACAGGCAGAGATAACATCAAGCAAGAGAAGATGGAAGCTTGATTGATCTCATCCATCCAGCAGGAATCGACTTACAAAACCCCGATGTCGGGGAAAATTCTTTAAATATTTTGGCTTTTAAACGGTCACAAAAGTGGCGCCGAAGCTCCGTGTAGTAACCATCCATTTTGAATCTTAACTGCTTTTAGTTATGCCATTAATGTTCAATGTGGGATGAATTGTATGTAGATTATATGTAATTTCCAATAAAGGTCTTATAAAATTGTA

> N. furzeri DNMT3b

GGAGCTTCGGTCCGGGCCGGGGGGTGTGAACAGTCGGTGTTTAGTAAGTGAAGGGTCCGACACTCCGCACGGGCACAGTGAGTTCCTACTGTGGAGATGGATCGCAGAACCAGAAAGAACAGCAAGGAACCAATCACTCACTAACCTCAGCCAAAATGTATCTGTGATATCTTTGATATCAGAAGATGAAGAGGACATTGTGTGTCTCTCTGGAGCAGACGGAGATGCTAAGGAGGATGAGTGTTCAGAAAACAGCAGAATCTGACGGCAGCCAAACAGGTTGGACATGGAGAATGCTGGAGATCATATGGCCAATGTGATTAGTGGGCTTATCAGAAGATGCTGTACTACTACTACTGATGACAAAAGAGATTGTGAAAGCCTGTTTCCAACACGCCTTTCTGCCTGTACCTGTACGTGGGAGTGCAGGTGGATGAAGGGCACAGCACCAGCTTTGTGCTCATTAGTTATTTGACCAAAGCTCGGGTAGGAGTGTAGTGAAGCTGCTGCACACTCTGCAGATAAGCACCGACACCGCTTATCCACAGGTCTCCGAGACACAGCAGTCACGCGTGCAGCTCGTTCTGATGCCTGCCTTTTGATTCCGACGCTAGAAAACTGGAGCTTCCATTTCCAACATGGCTGTGTTTACTGCAGCGAGCCCAAGCTGGCGAGATGTGTGTCTGGGCTCCAGACGTTACGCCCCAATGTGATATCACTCTGAGCATCCCGGAATTGCAGCAAGAGCTCCACAGGCTGCAGGAGGGCACCAAAAGACTGGTTGACTTGATCCTGGACATCCATCATCATTACTCCACTGCCTTCCATTAACGACAGCCTGAAGGAGGTCTTTGCTAACACTCAGTCTACAACCCCTCTGTCCATCTCAACGCAGTGTTCGTTCTGTCACACACAGTGTGCAGAAGATGGTTGCCCGCTGGCAGGATCTGTGGACTATTTCAATCCCTGAAACAACCGGAGGTTACTAACAGATCAGAATGCGGCTCATGGATGAGAATGTGAGACTGCATTTCTCTTCTGACGCTACGCTCTGGAGCCTCAGGACCCCTCAAGAGCTCCACAGGCTGCAGGAGGGCACCAAAAGACTGGCTGTGGAGCTTCAGCTCACGTCATTCTGCTTACGTCTATGCAGCCAGCATTTCTCCGCTCTCTGCTGTCACTGAGTGGCTTTCTGTGGAAGCAGGATTTGCAGCTGCTCAGCAGTGAGAAGGACCTCTGCCTTTAGCCGAGGTGAACATTGGTTCTCAGACCACAGACCTATGTGGGCCATGGGTGTAACCGGAGGTGAAAGGGACATGATGGCTTCTGAAGCGTGCACAGTCTTCTATAAAAGCGGCGCTCAGAGTTGATGGAGATTACCAGAAGCACTTGGGATGTGACATAAAGAACATCAAGGACATTGTAAGACCCAAAATAATTTACGGTCAGACATTTCCCAAGAAGTCTGTCCAAAATGGTGTTCAGCTGGGTTGTGTGCTGTGGACCTACAGAGAACCAACTGGCATCCAGTTACAGCGCTTCCCTCAGATGCTGAGCCAGGAATCGAAGTGGAGGAGTGAACAGCTGGCCAGGTTGTGATGGAGACTTCCAGGTTACAGCGTTCCTCCAGATGCTGAGCCAGGAATCGAAGTGGAGGAGTGAACAGCTGGCCCCAGGTTTGTGATGAGACGGTGGACAAGTCTGCACAACCTTCTTCTGACCCTGCTGACCATGCCAGCTCTTACACACACTCGGGTGTCTCTGAGGCTACCCGAACGTCGAAGTGCCTCACCCGAGGGAGGAAAACGACAGAAAACCTCCCTGTCTGAAACCTAGGAGATCAGGACCTTCTGTGCTCAAAAGGAGACCACCTTCCCTGGCCAGGCATGCAGGCCAGAGAAGAGAACAAGAGCTTGAGGAGACTTCAGGAGACAGTGACAGCACTGCAGCAAAGCCTTTAAGCTCTAAAACCACTCAACTCGATGACAGGATAGAAGACATGACAATTCATCAGATGTGATCGATATCAGAAGGATCCAAAATGTCTTTTACCAGCAGTCAAGTGACACACAACGATAAAGGAGACTCACCAGGAGAGATTGGTGTATATGACAGTGTGAAGAAGATGCCAGTTAGATGCTACGGCCACAACCTCTCGCTCTGAAGAGCTTCCAGAAGGATTCATTGGGGAGTGGTCTGGGGAATGATCGTGGGCTTCCCTCCCTGGCCGGCATCATGTACCAAATAAAAGACAAGTTCATCCACCAGGCATGATAATGGTGGAGTGGTATGGACAAAAATGTGTTTCCAGGTGATTTGAAGACCTCGGCCATTTGCTGAGTTTGGCCTGCATTCAGTGCCAATTAATTCGCCACCTTCATACCTACAGAGAGCATCTACATCACTCAGGAAGCAGCACTTCGCTGTAATAAGAAGTTTCTACAGAGTGGAAAGATGGTGAAGAGCTGCTGAGAACTACCTGTTAGACTGGGCTTTGAGAGTTTCAACCATCTGGTCCAGCTGGATTCAAACAGCCAATGGTTAAGAAATGACCAGGCCAGTGGTGAAGACGAAACTCGTCGGGAAAGTTGACTTGTTTACCAACCAGTCTGACCCAATGAAGAAC AAGACCAAACTCAGCTCTGAGATGAACAACGCAGCTGCTGTGCCACTGAACAACCTGTCTGTACTTCTTCCAGTCGAAAAC TGCAATAGAAAACAAATTTGATCCTGAAAAAAAACGCTCCAAGTCACTGGAAAAGAAAAGGGTCCAGGGGCACATTAAGTACAGATGGAGAAAACATGGAAGTCTGGAGGAGGATGAGATGAGTGGCCAAAGAGTGGGAAAGATGAAGCTGCTGCGGGAAGGAGATCAAGCAGAGGAAGTACGAGAGTAGGGAGAAGTCTTAAAAAGACATTTGGGCCATTTTATGGATTAGACGACGACATGTCCGCCAGACTTTGTTCCATCCAAGAAAAAGAACTTTGTCAAAAACCTACAATAAAGGGAGTCCAGGCAACCAGTGTGTAACACAGCCTGACCAAAAGCTCAGAGAGGACCATCAAGTATGAGTCACTGAACTGGAATGAAGGCTTCTGTTTGTGTTGAGAACTGAGAACATCGAGATTTCCGACCCTTTTTTCAAGGCACCTCTGCTTAAAGTCAAAAGATGAACATTCACAGAACTCTATATCGCTACGATGAAGATGGATACCAGTCTACTGTACTATTTGCTGCTACGGGCTTGAGGTCATACTGTGTGGAAATGATAGCTGTGTCAGGCTCTATTGTGAGGACTGCCTGAACATCTTGGTTGGACCAGGGACATTTGAGTCACTTAAGCACCTTGACCCATGGATTTGTACTGTGCAACCTCATCTGCCACCGTGCCTGGTTCCTAGAGTGGAGTGGAGTGTCAAGTCCAGAAATGTTGTTGTAACAACAGTGCATGGAGTTTGAACCTCTGTTTACCATCTATTCTGCCTACCTGCAGCAGCCATTCGATGCTTCGATGCTTCGATGGCATTGGACAGCTTGGTGTGAAAAGATCTCGGCTTTAAAGTGGAGAAAGTATGTTGCTCTGAGATTTGCGAAGATTCGGTTTGTGTGTCGACCATCAACCATGACGGGAAGGTATCCCATGTTGGTGACGTTCCGATTCATCCCAGGAAAAATCTTGAGAAGTGGGGCCATTTGATCTGTTGGTTGGCGGTAGCCCTGTAATGACCTCTCCATCGTAAAATCCAATTCGAAAGGTTCTCTATGAAGGCACTGGCAGGCTCTTCTTTGAGTATTTTCGCATCCTGCAGCTACTGAAGCCAAAGAGCAACCCCTGCTTTCTTCTGGCTTTGAGAATGTGGATGTTGCATGAACACCCAGCAAAAGTCAACTGATGTGAGGAACCTGCTTTCTTGAGTGAACCCAGTACTTGTGGATGCAGTCAAAGTGAAGCCGGCACACAGAGCTAGATACTTCTGGGGAAACATACCCGGGATGAGCAGGCCATCATCGCCTCCAGAATGACAACTGAGTCTCCAGGAATGTTTAGAGATAGGTTCGAGAGGCCAGGGTAAACAAGGTGAGAACGATAACCACCAACCAAACTCACTGAAGCAGGGTAAGGACGTGTCCTGCTGCCGCTCTGCAGAATGAAAGTGAAGACAACCTCTGGATCACGGAGCTGGAAAAAATCTTTGGATTTCCEAAACATTACACTGATGTGAGGAACCTGAAACCGACAACAGAGGCAGAAGCTGCTGGGAAAGCGTGGAGCGTTCCCGTCACTCCGTCACCTTTTTGCCCCCTTGAAGGAGTAC TTTGCTGTGAGGAGCTTCTCTCTGACCGTCCCAAGCACGTCCAGCTCCGCTCTCTCCCCACTGCCACACTCTGCAGACC TCCACCAGCTTCCAGCAGTAGGGGAGTCTCCATCTCCATAGTCTGACAGAATGTTTAGTGTTAGTTCATTTTCAGATTATTTTTCTTGAATCTTTTATACTTTTAAAGACACTACAAAATACACACACACACA

> N. furzeri DNMT3a

GCGCATCTGGAAGAGGGCTGGCCGGCTGGGGTGGCGGAGGACGGGGAGAGAAGAGGAAGGAGGCTCCGGGAGCAAGCGAGACAAAAAAGGGTCAAGGCTGGGACTTGGGAAAATCGGGAGAGGAGAGAAGAGGATCATCTCTAATTTCTTATCATTTTGGATTCGTTTTCTACTGGAACCTTCCCCTACTTCCACCATCAGGCATTACCCATAAGCCCGTAAGAATCTAAAAGATGCCGTCCAACTACTGCGCAACACCAATACCACCTCATTCAAAGCTCAGACTTGCAGATTCACCTGGAACACAGACTGGGTGAGAGCACAATACTGGACATGGACATGGAGTGGAGGAGAACTGGCTCTAAGTAGTGCCTTGTGAAAAGCTGGAAAAGCTGGGAGAAAGCGCAAGCAACTTCCAGTGGAGAGCTGCAGCTCGCTCGCAGGCAGTCTGAACGTCAGCAAAAAGTTACTCTGGGGTGCAGGCCATT CATGGCACAAGTCCCAATGGAGAGATGGACACACATGAAGACAAGATGCCTGACGCTGCTTCCCAACAGGAGAGGAG GGAGGTGGAGAATGGATCTCTTGGAGCTCTCTCTGCCAGCTGAGGACAGCTCCAGGATCGCAGCTGAGTCCCCTCC AATGAGAACCCTTCTGTCCAGCCAGGATGAGCAGGACTCGGAGAAAAGACAAGGATGACGGACTGATGACGCGCAAG AAAAGAGGAAGACGGAAACTGGAGCGCCAACGAAATATGTGGAGCAAAAAGGAGGACGAAGGGTGTGATGCAATCAAGAC

TGAGGGAGGTCGAGGCAGGCAACGAGGGGGAGTTGGTTGGGAGATCAGCCTTCGACAAAAGGCCCATGCCAGAGTAACCTT
 CCAGGCTGGTGACCCATATTACATTAGCAAGCGGACCAGAGAGGAGTGGATGGCCAAGTGGAAAGATGGAGGCAGAGAAAAA
 GGCCAAAGTAATGTCAGTGTGAACACAATGAAGGACCTAGAGCAGAACCAAAACAGACTCAGGAGGGAGGGTCTCAGT
 AATCAAACACCCACAGCTTCAACGCCACAGCAGGACCTGAGCACTCAGCCGCACAGTCAGACTGTGTCCCGCTCCGAT
 CAACAGCAATCCCAGCCGTGTCTGCCAGCAGCAGCAGCAGCAGCCACTGAGCCAGCTTCCCCACGGTGGCAACCACCC
 CCGACCCCGTCCGCAATTGGAGTGGAAAGACAAAGCCTCCCCAAGTCTGCAGACGCTGAGCCTGAATATGAGGACGGTCTGGT
 TTTTGGCATCGGCGACCTGGTTTGGGGGAAGCTGCGAGGATTCTCGTGGTGGCCAGGCCCATCGTCTCTGGTGGATGACGG
 GCCGGAGCAGAGCTGCTGAGGGAACAGATGGGTCTATGCGTTTGGAGATGGAAAATTCTCAGTGGTCTGTGTAGAGAAAT
 CATGCCTTTAAGCTCCTTCAGTAATGCCTTTCACCAACCAACTTACAACAAAACAGCCCATGTACAAAAAAGCCATCTATGAAG
 TGCTGCAGGTGGCCAGCAGCCGGGACGAAAATCCTTCATAACCTGCCCTGAAAGCGACGAAACAGAAACCTCCAAGTCGGT
 GGAAATGTTGAACAAGCAGATGATCGAGTGGGCCATGACCGGGTTTACGCCACCGGACCCAGGGGCTTACAGCCTCCCGAG
 GAGGACGCTAAGTCCGAGGTTTATCTGAAATGTTGGTGGAAACCGGAAGCTGCAGCCTACACCCTCTCTGCAAAAAAAC
 CTCGCAAAAGCACAGCTGAGAAGCCCAAGGTCAAGGATATCATCGATGAGAGAACCGGAGAGCGACTTGTCTATGAAGTGA
 GACAAAAGTGCCGACGATCGAGGATATCTGCATCTTGTGGGGTGTGAATGTTGGCCTGGGGCATCTCTGTTTGTGGA
 GGAATGTGTCAGGGCTGCAAGAAGTCTTCTAGAAATGTGCTTATCAGTACGACGACGATGGGTATCAGTCTACTGCACAAAT
 CTGCTGCCGCGGGCGAGGTTTCTCATGTGTGAAACAACAAGTGTGCGGTGCTTTTGTGTCGAGTGTGTGGACCTCTCG
 CCGTCAAGGACCGCCACCGCCATCAAGAAAGACCTGGAAGTCTACATGTGTTGTTCACAAAAGGGGCTTGTGTTGCT
 GCTGGAGCGCCGATCAGACTGGCCAGCCGCTGCAGCACTTCTTGTCAAATAATCAGATCAAGATTTTGTATCCACCGAAGC
 TTTACCCCTGTCTATGGTGGAGAAGAGGAAGCCTGTTCGTGCTCTGCTACTATTGATGGCATAGCAACAGGACTCCTTGTG
 TTAAGAGAGCTTGGCATCCAAGTGGATCGTACGTAGCGTCTGAAGTGTGTAAGACTCCATCACAGTGGCATCGTCCGGCAT
 CAGGACGCTATCAGTCCGCGATGTGAGGATGTGAGGATGTACCCCGAAGCATATTCAGGAGTGGGCTCTTTGACGCTTAT
 TGGTGAAGCCCTGCAATGACCTGTCCATAGTAAATCCTGCTAGGAAGGGTCTTTATGAAGGACCGCCGCTGTGTTG
 AGTTTACCAGGCTGCTCCATGAGGCTCGGCCAAAGGAGGGGGATGACCGCCTTCTTCTGGCTGTTTGAAGACGTTGGTGGCC
 ATGGGTGTCAGCGACAAGAGGCATATCTCGCTTCTAGAGTGAACCCAGTGTGATGATGCAAGGAGGTTGCTGCTG
 CCCACCGTCTGTTACTTCTGGGTAATCTTCTGGTGAACAGCCCTTATCTGCGATGTTCTACTGACAGGCTGGATCTCC
 AGGACTGTTGGAACACCGCAGAACAGCTAAGTTTGGCAAAAGTGAAGGACCATCACCACAGGTCTAATTCATCAAGCAGGG
 TAAGGACCAACACTTCCCCGTATACATGAACGAGAAGGAAGACATCCTCTGGTGCAGTGTGAGAGGGGTGTTGGTTTT
 CCCGTCCACTACAGGACGTGTCCAACATGAGCCGACTGGCAAGGCAGAGGCTCTGGGCAGATCCTGGAGCGTCCCCGTTA
 TCCGCACCTGTTTGCACCCTCAAAGATTACTTCCGCTGTGTTTGAAGAGCAGGAGATGTTTGGTTCAGCCATTTAGAAAC
 TACAGTGTATTTTTGGTTAGAAAATAAGAGCCTTGAACCTGAAATTTACTGTAGTATAAATTTGAAAGGCAATCGTTGACGT
 AGATTTACATTTACTTGTATTTAAATTTGCGACTTATTCTATTTAAGTTTCAAGTTCAGTTCGACCTTTTACAGTGTATTGACTTG
 TTTATTCGTCTGTGAAGCTATGCTTCTCCTCAAGACTGAGACGTAATAAATGCTAAACGTGTAAGAGAGACGGTAAACA
 TCAGAAAGTCACTCTTCAAGCGTATGTAAATTTGTGAGCGTACTGTACGTCGGGCTGCCGGCTGCCACCCGACGCGCA
 CGCGACTCCAGCTACTGAGCTCACAGACGTGGAGAAAGAAAGAGGAGTCTTCTCGTGGTCAAGGAGCCGTTCTCTG
 AAAATCCAAAACACCAGATCACATACCTCTATGTTTACAGTCTTATACACAGCTGGCGGTAATAAAGGTTACTACTGTAAT
 ATGGTTACAACACCATGGTGTCTCAAGATGATCAGAAGCATATAGCCATGTGGCACTAAAACCTTTGGATGATGCAGAGCT
 CCTAATCTGTTGTGGAAGATTTCTGTTAGATATTTTTACATTTTATGATGATAAAAACATTTTCTGTGCTGTTTCTTTAT
 TGAACAGAAACACAGGACGACAGCTTACAGGCTGTCTCTTTTACAGCTTTTACTGTTTCTACTGGTGTCTATTCTTTG
 TCTCAGCTTGTGCTCCAGGCTCACAGCTGGGTCTGCACAAAACAGAGAGTGAAGCTAGCCACTGGGGATGATGTTTCT
 GTCGTTTTAAACGGGGGGAAAAAAACTTTTAAAATCAAAGTCTGGTGTGACTACTTTTCTAGTTTTCATCCAGATGAATATT
 ACTGTACAAGTGTAGTAAGCGCTGTGCAGGAGGAAGGCTAGCGTCTCTCTCAGCATCCGCTCTCAGATAACTTTTTCAG
 GTGCAAACTTCAACATCGTGGTGGGCTGACTTTTTATGCAATTTTGTGCTGTTTATCAACAATAAAGCACTTTTATA
 TGTATACTTTTTCCCTTTGGTTTTTTTTAAAGATTTTGGTCTTGTGAATCGTTTGTATGATGTGAAAGTATGCTTTGTA
 AACACATTTTGTATTCAGTAGGGCTGTGCCAGTAGACGTGTGGCTTTGATATGTGAAACTCCCCTTATTGGACCGTGTAGAGT
 TGGTAAGTTGGCATTGTTGATGTCGCAACTTCTCTTCCGCCATTTTTCGTAGTGGACTCAACTTTCAGCGCAATATGTCACC
 ATCCAAACTCAAAGAATGTCTTTTTAAGATTGTAACCTACCAGGCAGAGTTAAGAAGTGTGCTGTTTCTGTTATATTT
 TATCTGCCTTGTGAAGGAGATTTTCCGTTTTAAAGTGTCTCACATCCACAACTAAATTTCTCGATTGACGCTCAAGA
 TAAATCTGGAGGATTTACGCGTCATTTTGTCAAAAACAAAAACTTACAAAAGTGTAGACATAAAAGAGCCTAAAACAAAAGCT
 AATTATATCATCTGATTGTATCATATTTAAGATTAGAATTAGAAAATTAATTTCTAGTGTAGTTAATTTCTTAAAGTACCTTAT
 TGCCCTGTTTTTGTCTGATGGAGACGCGGCAAGCTCTGAGCTGTGATCGTTAACCAGACACTATCTCCAATAAAAA
 AAATCATCCCTTGTATCTTTGATCTTTTACATTTTGTAAATAGTTTCCAGTTGCTGTTTTCTTTACCTTACGACATGATACA
 AATTTAATCTCTGTTAATATATATCATCAGAAAAAGTCAAGTGGTGTGATTGAGAACTTGTCTTAAACATGTAGGTGTGAAA
 GGTCCAGGTCAGCTTACTGTTTGTGTTGTGTCAGAGATTGAGATTGATGTTGTGAGCTTAGCTGATGAACGCTTTAAGGG
 GAAGCTTACCAGCAGCCCTGCCAGTGTGGTTTTATTTTGTCTTTTGTAGCTTTTGTAAAGATAAAAAAGTGGGATTATAG
 GTGTTTTCAATGGTGTCAACATTTGATCTAAGAAAAATATTTGATTGGAAAAATAAATGAAATGCTTTGATGATGA
 CTGTTTTAAGTTCTTTTGTAAATGGAAAAAGACTTATTTTACTCGTACAAATATATTAACAAAGGCATTCTAACTTTGTAC
 TTGAAAAAATATGAAAAGTTTCTTATATGTTTGTAGACTAGATTGCTGTAGTTGGTGTCTTCTGTGGGAAAGTTATGTGAGC
 AAAAAACAGTAAACCTTTACTGTAGCATATAAAAAACTTCTTATCCTACATTCCACAACATATTATTCTCCACGCTACTTC
 TTTAAATGTTGCATTGAAATAATGCATTTTGTGTTTTGAAATTTTCAATGTTTTTATATTCACAAGTTTTTGTGTTGTTCAAGAC
 AGGTCAAATATTGAGTGTGCTGTCGACACTCTGACGTTTCTGTTTTTGTCTATAAATGTTTTATGGCCTATACGCTTTCTCT
 CAAAGACCAGTTTGTACAGCAATGTAGCAGCCATTTCCACTTTGACTCCCTCATATACTTCAGTTTGTGTTTGTCCAGAGAGGA
 GATAAAGGCTCTGTATGTAAGTGTGCTGGAGAACACTGAATAAATGTCTGTTTTTGTGCAAAAAACA

> N. furzeri DNMT3a-like

CCTCGTGATAGAGGCGAGTCCGCGAGCCCGCAGCAGCGGGGAGAGAGGACCCGGTAGAACGCGAGGAGGGGGAGACA
 AGAGGGGAGAGAAGGAGTGAGCCATCGTCAGGGTCCGGTGGCTGAGGGAAGGACGATACAGAGCCAGGGAGGACAGAG
 TGGGACAGCAATCGTCCAGCTAGATTGAATGAAGCTGCACCAGGAGTTTCCGAGGGCCGTCCCATGTTTTGATTTGCTTAGT
 TTTTATTTATTTTTTTTCTTCTTTTCTTTTGGGGGGTGGGGGGTTGAAGTTGACCGTTCCACCATGCCGTCC
 AACAGTCTGCAGCCGCGAGCCACTGAGACGCGGAGAACGATGCTTCTAATGATGTGAGCGAAGACAGAGCAGATCAG
 GACAGTCCCGAGGAGGGGACGCCAGCAAGCCCGAAATAAGAGAAGAGTGGGTCTCCAGGAAAAAGCAACAGCTG
 CTTCTGTAGATGGCAACCTCTGACCCCTGTGCAAAATGCCTCCCCAACTCCGCCAAGGAGCCGAATGTTTCCCCCTCCCTCG
 CAAGAAGCGTAGGCGCCAGAAACCAGAGCAAAACCGAGAAGGATAAGGAAGATCCAGAAATGGACAGGAACCTGTGACTCCCC
 TCGAGAGGGTGTAGACCGGGAGGCTCCGGAGGAGGCGCTTCCAGGGTACTTTCCAGGCTGGAGATCCCTACTACATCAGC
 AGGACAGAAAAGAAAGACTGGCTCAGCCGCTGGAAGATGGAGTGGGAGACAACAGGGCAGAGACGGGCGTACCGGGA
 GGCAGAGATTAACATGATGGACGACCTATCGGAGGTCGACTTTCAAAAAGAGGAGGGGCCAGCCAGCCCGGCCCTCCCCC
 TCACAGCAACACACAGATCCCCTCTCCAACGGTTGCCGTGACACCAGAGCCTGTTGCAAGGGGAGACCAGGCCATGCCA

GTGAGATCGAGTACCAGGACGGCCGGGTTTTAGCATCGGGACTCTGGTGTGGGAAAGTGGCGAGGCTTTCTGGTGGCCC
GGCAGGATCGTCTCTGGTGGATGAGCGGCCGGAGTCGAGCTGCAGACGGCACTCGCTGGGTTCATGTGGTTCGGTGTATGGAA
AGTTCTCTGTTGTTGTGTGGAGAACTGATGCCCTGAGCTCTCTCTCATCCGCCTTCCACCAACCAACGTACAACAACAG
CCCTGTACCGGAAGGCTATCTTTGAAGCTTCGAGGTCAGCCAGTGCAGGAGGGGAGGCCATGTCGATGTGACACGA
GCGAAGACGCCGAAGCGGTGGAACCTCAGACCAGACAGATGATTGAGTGGGCAATGACCGGCTTCTGCCAATGGTCTCA
ATCACTGGATCCTCCAGAAGAGGAACAAAATCAGTTCAAAAGAGGGATACGTAGAAATGTGGCCGGAGCCCGAGGCAGCTTA
CACACCTCCACCTGCAAAGAAACCACGCAAGAATCGGCAGAGAAAGCAAAGATCAGAGAGGTGATCGACGAAGGAACCCAG
AGAGAGACTCATACATGAGATCAAGAAGAAAACCAGGAATATTGAAGACATCTGCATCTCCTGTGGAAAGTCTGAACGTCTCT
CTGGAGCATCCTCTCTCATAGGATCAATGTGTGACGGGTGCAAAAACTCCTCCTGGAGTGCCTACCAGTATGATGATGA
CGGCTACCAGTCTACTGCACCATCTGCTGTGGAGGGAGGGAAAGTGTCTATGTGTGGCAACAATAACTGCTGCAGGTGTTTCT
GTGTGGAATGTGTGGATCTCTGGTGGAGCCGGCTCAGCTGCAGCAGCCATTAAGAAGAGCCCTGGAACCTGCTACATGTGC
GGACCTCGGACACCTATGGTGTGCTGCGACGCGGACCGGACTGGCCATCAGACTACAGCATTCTTCGCCCAACAACCCAG
AGCAGGAATTCGAACCAGCCAAAATGTATCTCCGGTCGACAGGAAAGAGGAAACCCAGTACAGGATCCTCTCCCTATTTGA
TGGCATTGCCACAGGCTGCTGGTGTGAAAGATCTGGGCATCCAGGTGGACAAGTACATCGCATCCGAAGTGTGTGAAGAT
TCCATCACTGTGGCCCTGGTTCGACACCAGGGACGCATCATGTACGTGGGCGATGTACGCAACGTCACCTCGCAACATATTGA
AGAGTGGGGACATTTGACCTGGTTATTGGAGGAAAGCCCTGTAATGACCTCTATAGTCAACCCTGCAAGGAAAGGCCCTT
ATGAGGCAACTGGCGGTTGTTTTGAGTTTTATCGTCTGTCTGCTGACGAAAGAGGCTCGACCAAAACAGGAGACAGCGCCGTT
TTCTGGCTCTTCGAGAAGCTGGTGTGCTATGGGAGTGTGAGCGACAAGCGGGACATCTCCCGATTTTTAGAGTGTAAACCCAGTGT
GATCGACGCCAAGGAAGTCTCTGCTGCCACCGGCTCGCTATTTCTGGGAAACCTGCCAGGCATGTCCCGACCTTAACGC
CGATGGCCAAACGACAGGTGACTTACAGGAGTGCCTGCAACACGCGCCGACAGCTAAGTTTGGAAAGTGTGCTACGATAAC
GACAGCTCCAACTGTGTGAGCAGGGGAAAGACGAGCATTTCCCGCTTTCATGGACAACAAGGAGGACATCCTGTTGT
ACGGAGATGGAAAGGGTGTGTTGGTTCACAGTCCACTACACCGACGCTGTCCAACATGAGTCGTCTGGCTAGGCAACGGCTGCT
GGGACGTTCTGGAGCGTCCCGTTCATCAGGCACCTCTTCGCCCGCTCAAGGAGTACTTTGCTGCAACTAGTTACAACCTT
CTCTAAC
ATGACTCAGACCGGTCACACACATAAACACAAACACACACATACATAAAGTAAAGGACTGTCTCTACATCCCGTATAGAT
CACTGGTTGACCTTTGTCCCTGGTGTCTACCTTGTCTGATTTTAAACACACTGGTGCAGGGTTTGTACTGTAAGTACTTCT
ACACACACTTCTGTTCAACCTTGAACCTGATGTCTTCTGGCTTAATGTGCTGCAATAGCTGGAGTTTTAGCCTCCAGGGGG
CGCAATACTACACAAACGCCCAACTTTGTGTTTTATATGTCTCTGAACACATCAACACACTAGTACTCCTCTGCCATA
AAGCAATACTGAATTCATGTTAGTGACAGATCTACTGGTGTACTACAGTGAATGAACACAGGACATCCACACACACACAA
TTAATGCTTGTGGTCTAGTCAGGTAGTGACAGAGTCCAAATGGTGCATTTTCAATACTACATTTTTATAGTTTTGTGCTCT
TAGAGGCTCAATTAACCCGACTTTTACAAGCAGGCAGAAGCAAAAAACAGAGAGCAATGGTTTTTAAAAGTAAAAATTTA
AAGTGGATTTTGTTCAGAGTGACTTTACTAAGCTAGGTTTTAAGATTTTCATGCTTCTCAAGGAAGAGCAGCTCTGGTTC
ACAATGATCTCAGCATAATGGTTTAAACCTGTTGATTTAAATCAACCAGCTTGATCTAAAACTGGGATGGATGGTTTAT
GAATAAATATTAAGAATACTGCAGCTTTAGAGATTTGATAAAGAAAGTGGATAAATACCTCAATATTTGTGTTTAA
GGATAGATTTAGCTGTGCAAAAAACGACTAAACCTTAAAAATAACAGAGCATCCTCCTTAGCATTAGCTAGTGTCTCTAG
AACATAAACAGTACCTTTGATGAGCAAAATATGTGTATGTATCAGTATTTTCATGCTTCTAGAGCCTGTCTGCTGTCTACTG
GACCAACGGTTAGCATATGTCCAAAGCTTTTGGCATTAACTTTTGAAGTTCAGCCCGCTTAACTCACCTTTGTGAGATCC
TGCAGAGCAGCGCTTCTACAGAACTAAACTGAATCATTTAAAGACTTAGATCTGGCAACAGCAGTAAATGAACCCCTCT
CTTGAGATGCATCTTCTCTGTCCTCAACTCCAGCGTGGCCCTTTTTGTCTGGTGAACCATCCACCAGTCTGGTCCAT
AGAAATTCATTTGTGCTTCAGTTCCTGCAAGCAGCGACATCCACTGGAATCAGAGACAGACTTTTACTGCTGGTTTTCTGCT
TTCTACATCTGAAACGGATAACCCCTGGATAAATAACCCAGCAGTTCCTTATAAACAACCTTTTTTATTTCTGCCTCTGTGAC
AATAAACTCAGCTGTCTGACTTTAGCAGAGGGGGTGGCTTTTAAAGTGAAGGCAAGTTCATTAATTAACACACACACAC
TTGCACACACACACACATGCGCGCACACACACACACACTTGTAAAGAAACACACAGTCACTGATGTGTGTTTTACCTAC
ACAGGGCTGTCTTCTGGTTCATCATTTCAGTCTGGTGTGTTGTTCTTCAGACGACCATCTTTCCAGAAAGCCTTCCAATCT
CTGACCCTTCTCAGATACCAGCTTCTAACTTTACTCTGCTGTCTATGATGACACACACACACACACACACACAGCCTCTCAA
TTGAAACAATAAACAGCAGGTTCAAACCTTTTCCATCATGTTTACTACTGTTTCTACAACCAACCTTCAGGTGCTTTTTTAACT
GCTCTTTTTACAGTGTCTTTGAGGTCGACACACACACACACATATTAGCTGGAAACAGACTGATAAATAACTACTC
TGTTATGAAACATCACAATCACACATGCATACACACTGTGTTTACCAGGTCGCTGCTGCCATCAGGCTTCTACGGCGCTG
CAGAGTGTGACATCATCCCTGCGACGCTGGTGTCTCCATCCCCACATCCTTCTCTGACCCAAATCCCCTCTGAACACCA
CATCCCCAGGTGCTGCATGTGCTCAGTCACTGATCTGATTTCCATGTTAGATTAGGGAGTAAATGATAAATAAATAAT
ATGGTAATAGTAATTAATAATAAATCTGTTATCGTGTGTTGTTGACTGTTCCAGGTTGGTGTCTCAGTTCAGTCTCGGC
TACAGAGGGGAGAGAAGGAGGCTAGAGAAGGCATATATTACAGTTAATGTAGTGACTTGTCTATTGTATCCAAGAATGACA
AAAGAAACTGGAACAATGAGCTGCACACTGATCTCAGATCAGAAATGAGTTGAAACAGCTGCTGAGATACAAACCACATC
ACTGTGGTGTCTTGAACCGTGTGCTTTCCAGTCCAGTGCAAAACACTTAGGGTCTTAAATCTGTTAATGACTAGAAAAAG
TAAAGGTTTTCCCACTGCTGCTGCTTTGGTCCCTTAAACTGAACTTGGTCTGTTTTCCCTTGGTGTGTTTAAACACTGG
GCAGGCATGAAAATGACAATCATACCTGAGGAGGTCAGTTAGACTTCTGTATTACCTGAATTAAGTATCCATTTAACT
ATATAATTAATCATGGTGTATTTCTTAAAGTTTACTGTGTTTACTGTTAATTTGACCAAAATCAAATATTTAATCATCACAAC
TGTTCACTTTAACCTCAGTTAATGATAACTATATCGATGAACAAGAGATAGTTTGTTCACAATTTTATTTTACAAAAAA
ATCCACATAGACGTGTATGCTTGTATGAGGGCATTTGTTGCTTTTACGCTAATGAGGCGGTGCAAACTAGAGTAGGTGTCA
CCATTTTAAATGAACATGGGATCAGCAGTTTATGCAATTTCTTAGAGATGTTTTGTTTTAAAAACAGAACTAGTATAAATAAGCTA
ACTGGTCATTTTATTGACTTGGTTAACAATCATCATGAGAATTAAAAAAAATCTGTTCTGTATCCATTACGGGAATTTAAG
AGGCTTTAAACCAACAAGAAGTTTTCCGAACATTTGGGATCACTGACCGAAATGCACATCCGGTGTATCTAAAACGGTGCAA
CTTCCAAAACAGTTTGTCTCTCCAGCAGTTTTGTTTTGACACACTCTGCTGTGAGAGTGAAGCAGGACGAGCGATCGGGAGAG
GAGACATTTGGTGACGCAAGTGTGTTTCTGTTTTGCTTATCAACAGCTTTTAGGCTGAGTGTGCTGTTTTAAACACTGA
TACACTTTCTAAACAACACGTAACACTGCTCTAGACACACAGAGACACACGACGACCTACAGATACAATGCTGTGGCG
CATGCAGACACTCATCTATGGACTCAATGAAAGATGAGCATTAGGAAAAGGAGGACCTGTGGTGTTCAAAAGGTGCTA
AAATTTTACTGAAAATAAGAAAAAATATATTGTTAGGAGTATAAATCAAGTTTGTCTTACAATGACTGTATTTAAGTTTGT
GTAGCTGAAAAGGCACTATTATTGTTTACTCTGTATGAATAAGAACTAATTAATAAATAAATAAATAAATAAATAAATAA
ATGTAATAA

> N. furzeri DNMT3b-like

ATCTCGACTGCCCAACGACAGACAGAGATGGAGAAACATGGAAGTCTGGAGGAGGAGAAAGATGAGTGGGCCAAAGAGTGG
GAGAAGATGAAGCCTGTGTGCGGGAAGGAGATCAAGCAGGAGAAAGTACGAGAGTAGGGAGAAGTCTTAAAAAGACATTTGGG
CCATTTATGGATAGACGACATGTGCGCAGACTTGTCCATCAAGAAAAAGAACTTTGTCAAAAGAACTACAATAAAG
GAGTCAGGCAACCAAGTGTGTACACACAGCCTGACCAAAAAGCTCAGAGAGGAGACCATCATAGGATCATGAACCTGAAAG
GGACATTTGAAGGCTTCTGTTGTGTTGTGGAACCTGAGAATTCGAGATTTGCACCCCTTTTTTCAAGGACGCTCTGCTTAA

GTGCAAAGATAA CTTACAGAACTCTATATCGCTACGATGAAGATGGATAACCAGTCTACTGTACTATTTGCTGCTACGGGC
 TTGAGGTCATACTGTGTGAAATGATAGCTGTTGCAGGTCCTATTGTGAGGACTGCCTGAACATCTTGGTTGGACCAGGGACA
 TTTGAGTCACTTAAGCACCCTTGACCCATGGATTTGTTACCTGTGCCAACCTCATCTGCCCCACGGTGCCTGGTCTTAGAGTG
 GATTGGAGTTC AAGTCCAGGAATGTTTGTCTAACACACAGTGAATGGAGTTTGAACCTCATCGTGTTTACCATCTATTCTT
 GCCTACCTGCGCAGACCAGTTCGAGTCTTTCTCTGTTTCGATGGCATTGGGACAGGCTACTTGGTGTGAAAGATCTCGGCTTT
 AAAGTGGAGAAGTATGTTGCCTCTGAGATTTGCGAAGATTCCGTTTGTGTGTCGACCATCAACCATGACGGGAAGGTATCCCA
 TGTGGTGGACGTTTCGATTCATCACCCAGGAAAATCTTGAGAAGTGGGGCCATTGATCTGTTGGTTGGCGGTAGCCCTGTA
 ATGACCTCTCCATCGTAAATCCAATTCGGAAGGGTCTCTATGAAGGCACTGGCAGGCTCTCTTTGAGTATTTTCGCATCTGC
 AGCTACTGAAGCCCAAAGAGGACAACCCCTCTGCCTTTCTTCTGGCTGTTTGAAGAATGTGGTCTTCATGAACACCCACGACAAA
 GTCAACATCTGCCGTTTCTTGGAGTGAACCCAGTACTTGTGGATGCAGTCAAAGTGAAGCCCGCACACAGAGCTAGATACTT
 CTGGGGGAACATACCCGGGATGAGCAGGCCCATCATCGCTCCAGAAATGACAAAAGTACTCCAGGAATGTTTAGAGATA
 GGTCTGAGAGCCAGGTAACCAAGGTGAGAACGATAACCACCAACCCAACTCACTGAAGCAGGTAAGGACGTGTCCTG
 CTGCGGTTCTGCAGAATGGAAGTGAAGGACAACCTCTGGATCACGGAGCTGGAAAAAATCTTTGGATTTCCCAAAACATTACA
 CTGATGTGAGGAACCTGAACCGACAACAGAGGCAGAAGCTGCTGGGGAAGGCGTGGAGCGTTCCTGTCATCCGTCACCTTTT
 TGCCCCCTTGAAGGAGTACTTTGCTTGTGAGGAGCTTCTCTCTGACCGTCCCCAGCACGTCCAGCTCCGTCTCTCCCCACT
 GCCCCACACTCTGCAGACCTCCACCAGCTTCCAGCAGTAGGGGAGTCTCCATCTCCATAGTCTGACAGAATGTTTAGTGTGTTA
 GTTCATTTACGATATTTTCTTGAATCTTTTATACTTTTAAAGACACTACAAAATACACACACACACACACACACACACACA
 GAGACACG

> N. furzeri LOC107391197

CACGGCGAACAGCATTGTCATTTTGTCTCGGGCATCGTGGGGCCATCACTTAGGGTGGGTGCAGGGATTGTTAGCCTTTGGG
 GGTAGTGCCGTTTCGATGGGACTGCCTCCCGTTCGATCGCCCAATCCCGTGTGAAAGCACGTCGCGGGGGTGGTCTGATGT
 GCTGTGGGACGGATTGAGAGGAGCAGGATCTGTACTGTGGTGAACACAGCCGTTGGTGAAGATGGCAGGTTACAGAGGTT
 GTGTCTCCACGATCTCCCTTGATAAAGTCCAGTCTGTTGATGTGGTGGCCTGGACCAACAGGCTCCTGAAGACAGACTTAC
 AGATGTGCGAGAGATGGGCTCAGGTGCGTGTCACTGTGAGATGATGGACTACACTTTTCTGGATCTGTTGACTTAAGTAAAG
 TGAAATTTGATGCGCAAAGTGAAGGATGACTGAGGCATAAAGTCAAGTCTTCTCAGGAGGCTTCCAGCGCTGGGGTACC
 AGGACCTTCTGTTGATGAGTTCATAGAAAGATTTAAAGAGCAGCTTTGAGTCTTGAAGTGGTTCAAAGTCTTCTGACG
 TGCAAATGTAAATGACAAAAGATTACGACCTGTGAAAGCCCGTAAAGCCAAAGCCATTGACCCACTTCAGTCTGTTGCCAG
 TCACGAAGCTCTGACTCAGGATCCGACGTGGAGAAGAACGGAGAGCGTAAAACACATCCTTACTCATAAGTGAAGGAA
 AACTTCAAGTGGGCAAGCCCAAGTATCGGAGAGCAGTATACTACTGCTTGTACTGTAATAACAACCTTAAACACATTTAA
 GAACGGGGTCCGTTGAGCTTACCCGACATGCAACACAGTGAACACACAAAAGAAAGCTGCGGTTTGGCTTCTTGTAGTGT
 GCAGCAGTTCGGTTCATTAACAGTCACTGTTACACTGGATCCACTGAAAGAGAACAGGCGCCGGCTAGCTTTGACGCTGTA
 GCTGGGATCGCAGTACCCTGACGATATCGTATCCATTTGTCGACACACGCCCTACTGCGTTCAGTTTATGGAGGGTGGCAG
 TTGGAGAGGAAGATGCGGTTCTGTGATTCTCGTGGGTTATTTCAACGTGGAAATTTCTAGACACTGCATCCGATTTCTGGAT
 GCCCTCAGCTGGTGGACGGTGCAGGATCACGACGAGCAGCGGTTGGTGCAGACCTTGAAGAAATTTGATCTGTCCACAG
 ATAATCTTGTCTGCTGTTTATATCGGTAGTAATGGTGTGATGCCTCCGAGAGATTACTACATCTCAGGCAACACCA
 AACGTAGTGGCCTTAGGAGGTCTGTACACCATCGCTGATGCCGTTGTACGCTGGGCTTAAGGAGCTCCCTAATGGGCTCG
 AGAAATGATGGTGGATCTTACGACATTACTTCTCCACTCTACTAATAATGACAATCTCAAGGCAGCTTTGGCTCAGATA
 TCACCGAGAACGGCAAAGCATTACTCAACATTAGTTCGCTTCGATTTTGGCAGTTGGTGACAAAAATCTTGGAGGTGG
 ACAGATTTAATCTCGTACTGTTCTTGTCTACGAGGGGATGGGAAGCCAAAGTCGATCTGCTCCCAAGGATGAGGATCCCAA
 AGTCCAAGCCACATTTATGTTCTGACGAGGCCCTGAAGCCCTGCAAAGGTTTACAGTCAAATCCAGGCAGCTTCTTGGG
 ATAATATGTCCTTATCTAGAAGAAGCCAGCTCTTCTTGGGCAGTACAGTCTTACTTCTCAGCCTCAAGCTGTTATTG
 ACTACTCAAGAAACCCGATCCGGAAATCTAAAAAGCAAGGAAGCTCCAGCTGTCAATCCAAGATCTCAACTGGGAGGCAT
 GGAAGACTTCAAAAAGCGTCCGAGGCAACAGCTCTGTCGAGGTAGAAGAAGATGCATTTGCTTCTACTTGTGCACTC
 ACCGGCTGCTATTGTAGAGAAGTGCCTTTAAGCGACTTGGAGCTAAAAAGCGTTGGTCAAGTGTCTAACCCCTCAGAAAGGGC
 TGGACATGACAGCAGCAGCAGTAGAGAAGTGGGAACAAAGCTGGGGTCTGCAGTCCCCTGAGGAGGTCAAAGAGTCA
 CAGCAGAGTTTACAGGATATCAGTCTGACTGGAGAGGGAGCGGGGACAAGACAGCCCTGGTTTCTGGAGAAGCAGTGGAC
 GCAGACGTTAAAGACACCAGGACCCTCTGTCTCAGGAAGCTGTTTGGACCTGCTGCTTCCCCTGCTCTCCTCTAGA
 TTCTCAACGCGTCTTGTGAGGCTTAAAGGACGGAGACTCCGAGATCTTTTCTGAGACTGAAAGATACTGCAGAAAGGGACT
 CTGACGCCACTCTTACTCAGCTCTTAGAAACAGCGCTGAGGTGAAAGACTCCCAGGATCACACTCCAGAAACAAATGGCCT
 TGATGCGGCGGTGAAGCCTTGTGAAGTTCGCTCACAAAGATGAGTACTCAGTGGCGGTGTGTGTGGAGTTAATGGCGGT
 TCCAGAATGAGGGACAGTGGGAGGTTTGGCTGGGAGAGCAGCTTACGTCAGAAGCCACCGCCCGTACGGTGTTC
 GGTGGTGTGTTGCTCCTGGTCCAAACCCATAGCCTGATGCGGACCGCAGTAAACGGTCAACAAGTCCAAAGACGAGGCTCC
 CCTTTCAGAAAATCCACTCCAACCAAGGACAGCACAAACAGTACCAGGATGGGAAGGGATTCCAGATGGGGGAGCTGGTG
 TGGGGCAAACCGAAGGGCTTTTCCATGTGGCCTGGGATGGTGTGCTTGGAAAATAAGTCTCTCCCCAGGCATGAGAA
 GGGTGGAGTGGTTGGGGATGGCATGTTCTCAGAGATCTACTGACAAAAGTCTGCTGCGTTCAGCCGGTGTCTC
 TGAAAAAATTCCTTGGCAGCTGCCCTCTACAGGGAAGTCTCTACCAGATCATCGAGTTGGCAGTGAAGCATCGGGAA
 GTCTTTTGTCTATGCAGAAGGTAAGAAGCAGGAGCTGAAGCTGATGGTGGACTGGGCTTTTGAAGGATTTCTGCCTACTGGAC
 CTAAAGGACTCTTACCGGCTGACTCCTGTCCAGTTTACAGAGTCTGCTGACTCCTCCCTGTCTGACTCTCAGCCCGCACAAAAA
 GAAAAATGTCCTTAAAAGTAAAGCGGTGGCTGTGATGGTCTTCAACCGAGAAAAAGATCATAGAAAAATCTCAAAGAGAAAAA
 CAAAATATCAGAGATTTTGTGTTGCTGTTCTGCTGGCTGAGATCAACGTGGAACATCCTTTGTTGAGGGTGGATTATGTCT
 GAGATGCAAGGAGAACTTCTCAGAGACGCTATATCGCTACGATGAAGATGGCTATCAGTCTTATGCACCGTGTGCTGCGCG
 GGGTCAGAGGTATCCTGTGTGGGAATCCAGCTGAGTGCAGATGCTTCTGTAAGGATGCTGTAAGAAATCCTGGTGGCTGAAGG
 AACCTTTGAGAAGCTGAAAGACATCGACTCCCTGGAATCTTCTATGTGTAACCCGCTCAGTGTGAAGGAAACCTTGAAGCTG
 AGACTCAGACTGGCGTGTCAAGGTTCAAGACTTCTTGGCAACAACAGCCATGGAGTTTGAAGTCAAGAGGTTTACCCCTC
 GATCCAGCTGATCAGCGCAGACCAATCAGAGTCTGTCAATCTTTGATGGTATTGCAACGGGATATTTGGTGTGCGAGACC
 TCGGCTTAAAGGTGGAGCGATACATCGCTTACAGATCTGCGATGACTCGATCGCGTGGGAATGATCAAACACGAAGGGAA
 GATTGAATATGTTAACGATGTTGCAACATACCAGGAAACATCTGGCTGAGTGGGGTCCGTTTGTACTGCTGATTGGTGGCA
 GCCGTGCAACGACTGTCCATGGTCAACCTCTACGAAAAGGACTGTTTGAAGGCACCGAAGACTCTTTTGTAGTTTTAT
 CGTATTCTGAACCTTTTGAAGCCAAAGGAGGACGACCTGCTTCTTCTGGTTATTCGAGAAGCTGGTGTTCAGAAATC
 CCATGACAAATCAGACATCTGCCGTTTCTGGAGTGAATCCAATCTTATTGATGCAGTGAAGTCAAGTCCGGCTCATAGAG
 CTCGCTACTTTGGGGAAACCTACCTGGCATGAACAGACCTTCTCTCGTTCATTGATGACAAAAGTCAACCTTCAGGATGTT
 TGAAGTTGGGCGCAAGGCAATGTTTGAAGAAAGTTCGACAATCACGAAAGTCCAACTCCATAAGGCAAGGGAAGGTAG
 GCTCGTCCCCTACCATGAACGGCAAGGACTACTGTTGTCACCGAAATGGAACAGATCTTTGGCTTTCCCAAACA
 CTACACGGATGTGAACAACATGGGCCGGGGCAGAGGCAGAGAGTACTGGGTGCTCTGGAGCGTCCCGGTCATCCGTCAC
 CTGTTGCTCCGCTGAAGGATTATTTGAAATGTGAATCAGGGCAGTGAAGTACTCAGAGTGTGCTAACCCAAATGCCTTAT

ACCTCTAATGTGCAGATTCACCGCTCTTGTGCATCAAAAGACCTCAGTTAGTATCGTCAGCTCTTGAATCGTGTGTTTGTGTTT
TTCCTGTTTTTGTGTTTGTGTTTGTGAAATCATGCTGTGATGTCAGAGCACTGGCCTGCGTGCAGGAGTGAGAGGAGGCGGCCAG
CCCGCTCTTTGGGCTTTACTCGGGTTTTCCGGTGAACCTAACACCGGTTTCATGTTTACATTTGTTGGGATCACCATGGTAAATGGA
GAAAGCTAAAATTCAGAGGACAAAAGTTAGACAAAAGTTGATAAAATTCAGTTTATAGATTACATACACTTTTTAATCCTCC
CACTGTCTTCATGGGTGACCCGCCAGGAAAGTTGACCATTGAGCAGGATTGATGGTTT

Annexe 4 : Séquences nucléiques ADN utilisées pour l'alignement des séquences et la comparaison deux à deux dans Dialign Genomatix

>N. furzeri DNMT1 Gene ID:107378352

ATGCCATCCAGAACGTCCTGTCTCTGCCAGACGATGTCAGGAAAAGGTTGGAGCTGCTGGATGAGGATGGATGCTCAGAAG
AGGAACAAGTGAAGAAGAAGATCAAGTTGGTGCAGGACTTTCTGCATGATGATGCTCAGGACCAACTGACCAGTCTTGAAGA
AAAAATAAAAAAATCAGAGATCTCAAAGGAGGTTTACATCTCAAAGGTGAAAGCCTTGCTGGGCAAAGAATTCATCTTGAG
AATGGCTACATGTTGATGGCGTGGAGCGTAACGAAAAGAGAAAATGGCCTCTTAATGGCTCCATAAAGACGAAGAGGATG
GAGATGGTGTGTTTCCATGGATACCCAGGAAGAAGACACGTCATCATCTGTCAAGTACCAACCACCTCCAAAGGGAAGGG
GGCCCGCAGGAGCAAGGGAGCGCAGACACAAAAAATCACCAGCCAGCAGCAGAGTACAGGGAACAGTGTGAAGCAGC
CAACCATCTTGTCCATGTTTTCTAAAGTCCAAAAGCGCAAGTCTGATGAGGTGAACGGAGAACCTGTCAATGAAAACAAGTGA
ACCTAAAAAGAAGAGGCCGTTGATGAGTCTCGGAGGAGAAAACGCTGAAAAGTGGAGTGGATGAAAACACTGCTGCTGA
GGAAACAAAGCCTGAAGTTAACAGGTCTGCGTACGACAGCTAAAGCTCCGCCCCCAAAATGTCAGACTGCAGACATACTG
GATGACTCCGATCTGAAATTTCTCAAGGAGATCTGATAATGCGCTGGAAGAGCCTGAAATGTTGACGGATGAGCGTTTGTG
CCTGTTTGACGCAACAGGAGTGGGTTTGAAGCTACGATGATCTGCCGACGACAAGATCACAGTTTCAGTGTGTACGAC
AAGCAGGCCATCTCGCCCTGACTCTGGGCTGATTGAGAAGAACGTTGAGCTTACTTCAGTGTGTGCTTAAACCCAT
CTACACGACAACCCGTTGGACGGTGTCTGCTAAAAGTCTGGGCCCAATAACGCTGGTGGATCACTAGGTTTG
ATGGAGGTGAAAAGGCTTTGATTGGCTTACAACATCTTTTGGCGACTACATCTTGATGCAGCCCAGCGAGGAGTACGCTCCC
ATCTTCGCGCTCATGCAGGAGAAAATCTACATGAGCAAGATTGTGATGGAGTTTTTGCAGAAAAATCCCGACGCTACTTATGA
GGACCTTAAACAGAATAGAGACAACCGTTCGCGCTGCAGGACTCAACTTAACTGCTTACTGAAGACACGCTGTGCGCC
ACGCTTCTGACTGTGGGAAGTGTGCCGCTGCAAAAGACATGATCAAGTTTGGAGGAAAGCGGACAGCCATCTGTGACGCTG
CATGAGAGATCTGATCAAACCTCGTGGAGTCACTTATAGAAAAAGCATGCTGCTGTACTGGCGAGCCGCAAGAAGACAGGCC
ATCCGTCACCCACGAAGATAGAGAAGGACAACAAGGGCCGACAAAAGCCACCACCAGCTGGTCTATCAGATCTTTG
ATGCTTCTTCTCAGTTCAGATTGAGCAGAACGATAAGGAGGGTGGAGTCAAAAAGACAGCGCTGTGGAGTCTGTGAGGTGTG
CCAGTCTCTGACTGTGGGAAGTGTGCCGCTGCAAAAGACATGATCAAGTTTGGAGGAAAGCGGACAGCCATCTGTGACGCTG
AGGCAGAGAAGATGTCCAAACCTTGCTGTGAAGGAGGCGGAGGATGATGAGAATATTGAAGAAGACGTTCCAGAAGAGAAG
CTAAAAAGGCTCCAAACACTAAGAGGAAGAGCAGACCAGTCCAAACTGGCGTGGATCGAAGAGCCCGTTCAGACTGTG
GGGAAGAAGCAGTCTCAAGAAGGTTCTGTGTAATGACGAGGTTTGAAGTGGGAGACTGTGTTTCACTGTTCTGTATGA
TCCACTTCTCTTCTTCACTGGCAAGGATCAGTACTGTGGGAGGACAACAATGGGAAGATGTTCCAGCCCACTGGTTCC
TTCGCGGGATTACACGGTGTCTCGGAGAGTCTCCGATCCACTGGAGCTGGTCTGCTGAGCAGTGTGAGGACATGCTGCTC
AACTACGTCCAGGAAAAGTCAAGCTACGTCACGTATAAGGCCCGTCTAACAACCTGGTCTATGGAGGGAGGGGTGGACGTTGACC
TCAAGGTGATCGAGGAGATGGAAGAGTTTCTTCTACCAGTCTGGTATGATGTCAACTTTGCTCGATTGAGACTCCCGCC
GAGACCTTCCATCAGAAGATGCCGGTTCAGTTTCTGTGGCAGTGCACCTGTAAGAGAGGGAAGAGCAGGAGATCC
CTCGTCCATCTGAACCCCTGGAGAAGATAGAGTACGACACCAAAAGCTCTTATGGTTTGTAGTCTGTTTCAAGAGAGAGTTC
AGGGTGGGAGATGGTGTCTACCTGCCTCTGACGCTTCCATTTCAAGTGTGAAGCCTGCCAGTCCGCTGAAACGCTCCACAG
GAAAGAGGATGTGGATGAAGATCTGTATCTGAGTACTACAGGAAATCTCAGACTACATTAAGGGGTCCAACCTGGATGCT
CCAGAGCCGTTCCGATCGGCCGATCAAGGAGATCTTGTGTCACCGACGGAGCACCAGGGAAGTCCGACACCTCGGATGTCA
AACTGAGACTTACAAGTTTCTACAGGCTGAGAACACTCAAAAAGCGTCAAAAAGCGGTTCAAAAGCGGATCCATCAAGCT
GTACTGGAGCGAAGAGGAGGTGACCGTCAAGCTGTGCGACGTGCTCGGCCGCTGTCAGGTGGAGTACGCTGAAGACCTGAAT
GAATCCGTCAGGAGTATTCCAGCGGTGGGCCGGATCGTCTTATTTCTCGAGGCTACAATGTAAATCCAAGAGCTTTGA
AGATCTCTCGAACCATGCCCCCTCTCTGTCCATAAAGGAAAGGAAAAGGCAAAAGGAAAAGGAAAAGGAAAAGGAAAAGG
TGCGCAGAAACCCCTGACAGGCTCACACAGAACCTCCAAAGTGGCCAAATTCGCACACTGGATGTTGTTCCGGGATGGGA
GGACTCTCTGAGGGTTTCCACCAGGCGGGCATCTCCGAGACTCTCTGGCCATAGAGATGTGGAACTGACGACAGGCCT
TCAGGCTCAACAACCTGGCACCAGGTGTTACGGAGGACTGTAACGCTTACTGAAGTTGGTCAATGTCGGGGGAGAAGAC
CAATTTCTTTGGCAGAAGCTGCCTCAGAAGGGGGATGTTGAGATGCTGTGTGGTGGACCTCCCTGTCAAGGGTTTCAAGTGGGA
TGAACCGCTTCAATTTCCGGACTTACTCCAAGTTCAAGAACTCTCTTGGTCTCGTATCTGAGTACTGCGACTACTGACAGC
CAAAAGTTTTCTGCTGGAGAACGTGAGGAACCTCGTTTCAATCAAAGCTCCATGGTTCTGAACTACGCTGACGCTGCTC
GTGCGAATGGGCTACCAGTGCACCTTTGGGGTCTTCAAGCGGGTCAAGTACGGCGTCCGACAGCCGCTGACAGGGCGATCA
TCTGGCTGCTGCCCCAGGAGAGAAGCTGCCCGTACCCTGAACCCCTCCACGTTTGTCTCCAGAGCCTGCCAGCTGAGC
GTGGTGGTGGACAAAAGAAATATGTCAGCAATGCCACGGAGGCAACGGTGGGATCTATAGAACCATCACCGTCCGCGAC
ACCATGTCGATCTGCCGGATTTGTAACGGTGGCAGCTTTGGAGATTTCCCTACAACGGCAGCCTCAGTCTGTTTCCA
GAGGCAGATCAGAGGCACGAGTACCAGCCATCCTAAGAGACCACATCTGCAAGGACATGAGCGCTCTGGTTGAGGCCAG
AATGCGCCACATCCCCTGGCTCCAGGCTCGGACTGGAGAGATCTCCAAATATTGAGGTTCCGGATGAAAGATGGCAACACG
ACTAAGAAGTGCCTTACACGATCACGATAAAAAGAACGGCCGAGCAGCAGTGGTGCACCTAGAGGTGTCTGCACCTGTG
CAGGAGGAAAACCTTGCAGCCCTAGACAGGACTTAAACACCTGATTTCCCTGGTGTCTGCCTACACTGGAACCCGAC
AATCACTGGGCTGGACTTTATGGCAGACTGGAGTGGGATGGCTTCTCAGCACAACAGTCAACAACCTGAAACCCATGGGAA
AACAGGGTCTGTTTCTGCACGCGAGCAGCACAGAGTGGTGTGAGTGTGAGAGTGTGCACGCTCTCAGGGATTCCCTGACAC
CTACCGATTCTTTGAAACATCCTGGACAAGCAGACAGGTTGAAAATGCTGTTCTCTCTCTCTCCAGAGCCATCGGGC
TGGAGATTAAGAAATGTGTCATAGAAAGGATAAAGCAAGAAGCAGGAGAAGATAACATCAAGCAAGAGAAGATGGAAGCTT
GA

> N. furzeri DNMT3a Gene ID:107392003

ATGCCGTTCAACTGTGCGCAACACCAATACCACCTCATTCAAAAGCTCAGACTTTGCAGATTACCTGGAAACAGACTGGG
TGAGAGCAATACTGGACATGGACCTGATGGAGGAGAAGTCGGCTCTAAGTAGTGCCTCTTGTAAAGCTGAAAGCCTGGG
AGAAAGCGCAAGCAACTCCAGTGGAGAGTGCAGCTCGCAGGCAAGTGAACGTCAGCAAAAGTTACTCTGGGGTGC
GCAGGCCATTCATGGACAAGTCCACAATGGAGAGATGGACACACATGAAGACAAGATGCTGACGCTGCTTCCCAACA
GGAGAGGAGGGAGGTGGAGAATGGGATCTCTTGGAGCTCTCTCTGCCAGGCTGAGGACAGCTCCAGGATCGCAGCCTG
AGTCCCTCCATGAGAACGGCTTCTGTCCAGCAGGATGAGCAGGACTCGGAGAAAAGACAAGGATGACGGACTGATGATGC
AGCGAAGAAAAGGAGAAGACGGAAACTGGAGCGCCCAAGAAATATGTGGAGCAAAAAGGAGGAGAAGGGTGTGATGCA
ATCAAGACTGAGGAGAGAAGAAAAGGCAAAAGTATGTCAGTGTGATGACACAATAAGAGGACTGACGCAAGCAACCAAG
ACTCAGGAGGAGGAGGTCTCAGTAATCAAAACCCACAGCCTTCAACGCCACAGCAGGAGCCTGAACCTCAGCCGACAGT

AGACTCTGTCCCGCCTCCGTATCAACAGCAATCCCAGCCGTGTCTGCCCCAGCAGCAGCAGCAGCAGCCACTGAGCCAGC
 TTCCCCACGGTGGCAACCACCCCGACCCCGTCCGCAATTGGAGTGGAAAGACAAAGCCTCCCCAAGTCTGCAGACGCTGAG
 CCTGAATATGAGGACGGTCTGGTTTTGGCATCGGCGACCTGGTTTTGGGGGAAGCTGCGAGGATTTCTGTGGTGGCCAGGCC
 GCATCGTCTGGTGGATGACGGGCCGGAGCAGAGCTGCTGAGAGCCAGATGGGTCAATGGGTCAATGGGTATGGAAATTT
 CTCAGTGGTCTGTGTAGAGAACTCATGCCTTTAAGCTCCTTCAGTAATGCCTTTACCAACCAACTTACAACAAACAGCCCA
 TGTACAAAAAAGCCTATATGAAGTGTGTCAGGTGGCCAGCAGCCGGGCAGGAAAATCCTTCATAACCTGCCCTGAAAGCGA
 CGAAACAGAAACCTCAAGTCCGGTGGAAATGTTGAACAAGCAGATGATCGAGTGGGCCATGACCCGGTTTACGCCACCAGG
 CCCAGGGGTTACAGCCTCCGAGGAGGAGCGTAACCTCGAGGTTTATCTGAAATGTGGGTGGAACCCGGAAGCTGCAGCT
 ACACCCCTCCTCTGCAAAAAAACCTCGAAAAAGCACAGTGTAGAAGCCCAAGGTCAAGGATATCACCGATGAGAGAACC
 GAGAGCGACTTGTCTATGAAGTGTGACAAAAAGTGCAGCAGCATCGAGGATATCTGCATCTTGTGGGGGTTTGAATGTTGG
 CCTGGGGCATCCTCTGTTTGTGGAGGAATGTGTCAGGGTCAAGAACTGCTTCTAGAATGTGCTTATCAGTACGACGAGC
 ATGGGTATCAGTCTACTGCACAATCTGCTGCGGGGGGAGAGGTTTCTATGTGTGAAACAACAAGTGTGTCGGTGCCTT
 TGTGTGAGTGTGTGGACCTCTCTGTCGGTCAAGGAGCCGCCACGCCATCAAAGAAGACCCGTGGAACCTGCTACATGT
 GTTGTCCAAAGGGGGTGTGGTCTGCTGGAGCGCGATCAGACTGGCCAGCCGCTGCAGCACTTCTTTGCAAATAATCAC
 GATCAAGATTTTGTATCCACGAAGCTTTACCCCTGTGATGGTGGAGAAGAGGAAGCTGTTCTGTCCTGTCATTTGA
 TGGCATAGCAACAGGACTCCTTGTGTTAAGAGACTTGGCATCAAGTGGATCGCTACGTAGCGTCTGAAGTGTGTGAAGAC
 TCCATCACAGTGGGCATCGTCCGATCAGGACGCATCATGTATGTGCGCATGTGAGGAATGTCACCGATGAGAGAATTC
 AGGAGTGGGGTCTTTTACCTTGTATTGGTGGAAAGCCCTGCAATGACCTGTCCATAGTAAATCCTGCTAGGAAGGGTCTT
 TATGAAGGCACCGCCGCTGTCTTTGAGTTTTACCGCTGCTCCATGAGGCTCGGCCAAAGGAGGGGGATGACCCGCCCTT
 CTTCTGGCTGTTTGAACGTTGGTGGCCATGGGTGTGACGACAAAGAGGACATATCTGCTTTCTAGAGTGAACCCAGTGA
 TGATTGATCCAAGAGGTTGTCTGTGCCACCGTCTGTTACTTCTGGGTAATCTTCTGGGTAACAGGCTTTATCTG
 CGATGTTCACTGACAGGTTGATCTCCAGGACTGCTTGAACACCGCAGAACGTAAGTTTTGGCAAAGTGGAGGCATCAC
 CACCAGGTCTAATTCATCAAGCAGGTAAGGACCAACTTCCCCGTATACATGAACGAGAAGGAAGACATCCTCTGGTGC
 ACTGAGATGGAGAGGGTGTGGTTTTCCCGTCCACTACCGGACGTGTCCAACATGAGCCGACTGGCAAGGCAGAGGGCTCC
 TGGCAGATCCTGGAGCGTCCCGTTATCCGCCACTGTTTGCACCACTCAAAGATTACTTCGCTGTGTTGA

> N. furzeri DNMT3a-like Gene ID:10737776

ATGCCGTCCAACAGTCTGCAGCCGCGAGCCACCTGAGACGCCGGAGAACGATGCTTCTAATGATGTGAGCGAAGACAGAG
 CAGATCAGGACAGTCCCAGGAGGGGACGCCAGCAAGCCCCGAAATAAGAGAAGAGTGGGTCTCCAGCAGGAAAAGCA
 AACAGTGTCTCTGAGATGGCAACCTCTGACCCTGTGCAATGCCTCCCCAACTCCGCCAAGGAGCCGAATGTTTCCCC
 TCCCCTCGCAAGAAGCGTAGGCGCCAGAAACCAGAGCAAACCGAGAAGGATAAGGAAGATCCAGAAATGGACAGGAAGTGT
 GACTCCCCTCGAGAGGGTGTGACCCGGGAGGCTCCGGAGGAGCCCGTCCAGGGTACTTCCAGGCTGGAGATCCCTACT
 ACATCAGCAGGAGACAGAAAGAAGATGGCTCAGCCGCTGGAAGATGGAGGTGGGAGACAACAGGCGGAGAGACGGGGC
 TACCGGGAGGCAGAGATTAACATGATGGACGACCTATCGGAGGTCGACTTCAAAAAGAGGAGGGGGCCAGCCAGCCCGGCC
 CCTCCCCCTCACAGCAACACAGACTCCCGCTCTCAACGGTTGCGGTGACACCAGAGCCTGTTGCAAGGGGAGACCAGG
 CCATGCCAGTGAGATCGATACAGGACGGCCGGTGAACCTCAGACAGCCGAGATGATTGAGTGGGCAATGACCCGCTTCTGCCAAT
 TGGTGGCCCGCAGGATCGTCTCTGGTGGATGAGCGCCGGAGTTCGAGCTGCAGACGGCACTCGCTGGGTCTATGTGGTTCG
 GTGATGGAAGTCTCTGTTGTTGTGTGGAGAACTGATGCCCTGAGCTCTTCTCATCCGCCTTCCACCAACCAACGTAACA
 ACAACAGCCCATGTACCGGAAGGCTATCTTTGAAGCTCTGCAGGTAGCCAGCGTGCAGGCAAGGAGGGAGCCAGTTCGCTCAT
 TGACACGAGCGAAGACGCCAAGCGGTGGAACCTCAGACAGCCAGATGATTGAGTGGGCAATGACCCGCTTCTGCCAAT
 GGTCTCAATCACTGGATCTCCAGAAGAGGAACAAAATCAGTTCAAAGAGGGATACGTAGAAATGTGGCCGGAGCCCGAG
 GCAGTTACACACCTCCACCTGCAAGAAACCACGCAAGAACTCGGCAGAGAAAGCAAGATCAGAGAGGTGATCGACGAA
 GGAACAGAGAGAGACTCATACATGAGATCAAGAAGAAAACCAGGAATATTGAAGACATCTGCATCTCTGTGGAAGTCTG
 AACGTTCTCTGGAGCTCTCTTCTCATAGGATAAGTGTTCAGGGTTCGAAAGTGCAGGCAAAAACCTTCTCCAGCTACCAAGTA
 TGATGATGACGGCTACCAGTCTACTGCACCATCTGCTGTGGAGGGAGGGAAGTGTCTATGTGTGGAAACAATAACTGCTGC
 AGGTGTTTCTGTGTGGAATGTGTGGATCTTGGTGGAGCCGGCTCAGCTGCAGCAGCCATTAAGAAGACCCCTGGAAGTGC
 CTACATGTGCGGACCTCGGAGCCTATGGGTTGCTGCGACGACGGGACGACTGGCCAATCAGACTACAGCATTTCTTCGCCA
 ACAACCAGGACAGGAATTCGAACCAGCCAAATGTATCTCCGGTTCGACGAGAGAAGAGGAACCCGATCAGAGTCTCTCTC
 CCTATTTGATGGCATTCCAGCCAGGCTGTTGTTGCTGAAAGTCTGGGATCCAGGTGGACAAGTACATCGCATCCGAAGTGT
 GTGAAGATTCCATCACTGTGGGCTGGTTCGACACCAGGGACGCATCATGTACGTGGGCGATGTACGCAACGTCCTCGCAA
 ACATATTGAAGAGTGGGGACCATTTGACCTGGTATTGGAGGAAGCCCTGTAATGACCTCTCTATAGTCAACCCTGCAAGGA
 AAGGCTTTATGAGGAACTGGGCGTTGTTTTTGTAGTTTTATCGTGTCTGCACGAGGCTCGACAAAACAGCAGCAGGAG
 CGGCCGTTCTTGGCTTCGAGAACGTGCTGCTATGGAGTGTGACGACAAGCGGGACATCTCCGATTTTTAGAGTGTAA
 CCCAGTGTATGTCGACGCCAAGGAAGTCTGCTGCCACCGCGCTCGCTATTTCTGGGGAAACCTGCCAGGCATGTCCCGAC
 CTCTAACCGCGATGGCAACGACAGGCTGGACTTACAGGAGTGCCTCGAACACGCGCCGACAGCTAAGTTTGTGAAAGTTGCG
 TACGATAACGACAGCTCCAACCTGTGAAGCAGGGGAAAGCAGGACATTTCCCGTCTTATGGACAACAAGGAGGACATC
 CTCTGGTGTACGAGATGGAAGGGTGTGGTTTTCCAGTCCACTACCCGACGCTGTCCAACATGAGTCTGTGGCTAGGCA
 ACGGCTGCTGGGACGTTCTGGAGCGTCCCGTCTCAGGCACCTCTTCGCCCGCTCAAGGAGTACTTTGCCTGCAACTAG

> N. furzeri DNMT3b Gene ID:107390569

ATGGATCGCAGAACACCAGAAGAACCAGCAAGGAACCAATCACTACTAACCTCAGCCAAAATGTATCTGTGATATCTTTGA
 TATCAGAAGATGAAGAGGACATTTGTGTGTCTCTGGAGCAGACGGGAATGCTAAGGAGGATGAGTGTTTTTCAGAAACA
 GCAGAATCTGACGGCAGCCAAACCAGGTGTGGAGATGGAGAATGCTGGAGATCATATGGCCACTGTGATTTAGTGGGCTTT
 ATCAGAAGATGCTGTGGTACTCACGTCTACTCACTGATCCAAAAGAGATTGTGCAGGCTGTTTCCAACACGCCTTTCTGCCT
 GTACCTGTACGTGGGAGTGCAGGTGGATGAAGGGCAGCAGCACCAGCTTTGTGCTCATTAGTTATTTGACCAAAGCTCGGTA
 GGAGTGTAGTGAAGCTGTGCACACTCTGCAGATAAGCACCCGACCCGTTATCCACAGGTCTCCGGAGACACAGCAGCAGT
 CACGCTGTCAGTCTGTTCTGATGCTGCTTTGATTCGGACGCTAGAAAACTGGAGCTTCCACTTTCAACATGGCTGTGTT
 TACTGCAGCGAGCCCCAGCTGGGCGAGATGTGTGTCTGGGCTCCAGACGTTACGCCAAAATGTGATATCACTCTGCAGCA
 TCCCCGGAATTCAGCAAGAGCCTCCAGGCTGCACTCAGACTTCAATTTACTCAGTGGTTGACTTGTATCTGGACATCCAT
 CATCTTACTCCACTGCACTTCCATTAACGACAGCCTGAAGGAGGTTTGTGTAACACTCAGTCTACAACCCCTCTGTCCC
 ATCTAACCGAGTGTTCGTTCTGTCACACAGTGTGCAAGATGTTGCCCCGCTGGCAGGATCTGCTGGACTATTTTCAATC
 CCTGAAAACAAACGGAGGTTACTAACCAGATCAGAATGCGGCTCATGAGTGAAGATGTCAGACTGCATTTCTCTCTGCTGTC
 ACGCTCTGGAGCCCTCAAAGTCTCCAGGAGTGCAGGAGGCCGGCACAAAGACCTGGCTGTTGAGCTTACAGCTCACGTC
 CATTCTGCTCAGTCTATGACGCCAGCATCTCCGCTCTCTGCTGTGACGTGGCTTTCTGTGGAGGACGAGTTTGCAGCTGT
 CAGCAGTGAAGAGCCTCTGCTTTAGCCGAGGTGAACATGTTCTCAGACCAGACCTCATGTGGCCATGGGCTGTG
 AACGGAGTGAAGGGACATGATGGTTCTGTAAGCTGTCACAGCTTCTATAAAGCGCGCTTCCAGATTTGATGGAGA
 GTTTACCAGAAGCACTTGGGGATGTGACACTAAAGAATCAAGGACATTTGTGAAAGACCCAAAAAATTCACGGTCAGACA

TTTTCCAAGAAGGTCCTGTCCTCCAAATTGGTGTTCAGCTGGGTTTGTGTGCTGTTGGACCTACAGAGAACCAACTGGCATCCA
 GTTACAGCGCGTTCCCTCAGATGCTGAGCCAGGAATCGAACTGGAAGGAGTGTAAACAGCTGGGCCAGGTTTTGTCTCAGAT
 GAGACGGTGGCAAGTCTGCACAACTTCTTCTGACCTGCTGACCATGCCAGCTCTCTACACAGGACTCGGGTGTCTCTG
 AGGCCTACCCGAACCTCCGAATGCCTCACCAGGGGAGAAAACGCAGAAAACCTCCCTCTGCTGAAACCTAGGAGATCACG
 ACCTTCTGTGCTCAAAAAGGAGACCACCTTCCCTGGCCAGGCATGCAGGCCAGAGAAGAGAACAGGAGCTTGAGGAGACTTCA
 GGAGACAGTGACAGCACTGCAGCAAAGCCTTTAAGCTCTAAAACCACTCAACTCGATGACAGGATAGAAGACACTGACAATT
 CATCAGATGTGATCGATATTACAGAAGGATCCAAAATGTCTTTACCAGCAGTCAAGTGACACCACAACGTAAGGAAGTCA
 CCAGGGAGAGATTGGTCTGATATTGACGAGTGTGAAGAAGATGCCAGCTTAGATGCTACGGCCACAACCTCTCGCTGAA
 GAGCTTCCAGAAGGATTATTGGGGAGCTGGTCTGGGGAATGATCGTGGGCTTCCCTCCCTGGCCGGGCATCATCGTACAAA
 TAAAAGACAAGTTCATCCACCAGGCATGATAATGGTGGAGTGGTATGGACAAAAAATGTGTTCCAGGTCAGTTGAAGACC
 CTGCGCCCATTTGCTGAGTTTGGCCTGCACTTCAGTGCCAATTATTTGCCACCTTCATCACCTACAGAGAGAGCATCTACACA
 TCACCTCAG

> N. furzeri DNMT3b-like Gene ID:107372626

TTCTGTTTGTGTGTGGAACCTGAGAACATCGAGATTCGCACCCTCTTTTCAAGGCAGCCTCTGCTTAAAGTGCAAAGATAAC
 TTCACAGAACTCTATATCGCTACGATGAAGATGGATACCAGTCTACTGTACTATTTGCTGCTACGGGCTTGAGGTCATACT
 GTGTGAAATGATAGCTGTTGCAGGTCCTATTGTGAGGACTGCCTGAACATCTTGGTTGGACCAGGGACATTTGAGTCACTTA
 AGCACCTTGACCCATGGATTGTTACCTGTGCCAACCTCATCTGCCCCAGGTGCGCTGGTTCCTAGAGTGGATTGGAGTGTTC
 AAGTCCAAAGAAATTTGCTAACACAGTGAATGGAGTTGAACCTCATCTGTTTACCATCTATTCTGCTACCTGCAGTACTGAA
 GACCAGTTCGAGTCTTCTCTGTTTCGATGGCATTGGGACAGGCTACTTGGTGTGAAAAGATCTCGGCTTTAAAGTGGAGAAG
 TATGTTGCCTCTGAGATTTGCCAAGATTCGGTTTGTGTGTCACCATCAACCATGACGGGAAGGTATCCCATGTTGGTGACGT
 TCGATTATCACCCAGGAAAATCTTGAGAAGTGGGGCCATTTGATCTGTTGGTTGGCGGTAGCCCTGTAATGACCTTCCA
 TCGTAAATCCAATTTGGAAGGGTCTCTATGAAGGCACTGGCAGGCTCTTTGAGTATTTTCGCATCCTGCAGTACTGAA
 CCCAAAGAGGACAACCCTCTGCCTTTCTTCTGGCTGTTTGAAGAATGTGGTCTTCATGAACACCCACGACAAAGTCAACATCTG
 CCGTTTCTTGTAGTGAACCCAGTACTTGTGGATGCAAGTGAAGTGAAGCCGGCACACAGAGCTAGATACTTCTGGGGGAAC
 ATACCCGGGATGAGCAGGCCATCATCGCTCCCAAGATGACAACTGAGTCTCCAGGAATGTTAGAGATAGGTCGAGAGG
 CCAGGTAACCAAGGTGAGAACGATAACCAACCAAACTACTGAAAGCAGGGTAAGGACGTGTCCCTGCTGCCGGTCT
 GCAGAATGGAAGTGAGGACAACCTCTGGATCACGGAGCTGAAAA

> N. furzeri LOC107391197 Gene ID:107391197

ATGGCAGGTTTCAGAGGTTGTGTCTCCAGATCCTCCCTTGATAAGTCCAGTCGTTTGTATGTGGTGGCCTGGACCAACAGGCT
 CCTGAAGACAGACTTCACAGATGTGCGAGAGATGGGCTCAGGTGCGTGTCACTGTCAAGATGATGGACTACACTTTTCTGGAT
 CTGTTGACTTAAAGTAAAGTAAATTTGATGCGCAAAGTGAAGATGACTGCAGGCATAACTACAGTCTTCTCAGGAGGCTTC
 ACCGCGCTCGGGTACCAGGACCATCTCTGTTGATGAGCTCATAGAAAGAGATTTAAAGAGCAGCTTTGAGTCTTGAAGT
 GTTCAAAGTTTCTACGCTGCAAATGTAATGACAAAGATTACGACCCTGTGAAAGCCCGTAACGCCAAGCCATTGACCCA
 CTTCAAGTGTGCTGCCAGTACGAAAGCTGTGACTAGGATCCGACGTGGAGAAGAACGGAGAGCGTAAAACACATCCTTACA
 CTCATAAGTGAAGGAAAACCTCAAGTGGGCAAGCCCAAGTGTGATGCGAGAGCAGTATACCTACTGTACTGTAATAA
 CAACTTAAACACATTTAAGAACGGGGTCTGAGCTTACCGCATGCAAAACACGATGAAAACACGAAAGAAAGTGGGTT
 TGGCTTCTTGTAGTGTGACAGCAGTTCGGTTCATTAACAGTCACTGTTACACTGGATCCACTGAAAGAGAACAGGCGCCGGC
 TAGTGTGACCGTGAAGTGGGATCGCAGTACCCTGACGATATCGTATCCATTGTGACACACGCGCTACTGCGTTACGT
 TTATGGAGGGTGGCAGTTGGAGAGGAAGATCGGCTCTGTGATTTCTGTTGGTTATTTCAACGTGGAATTTCTAGACACT
 GCATCCGATTTCTGGATGCCCTTACGCTGGTGGAGCGTGGGAAATCAGCAGCAGCAGCGGTTGGTGGCAGACTTGAAGAA
 ATTTGATCTGTCCACAGATAATCTTGTCTGCTGTTTATATCGGTAGTAATGGTGTGCTCCGAGCAGATTTACTCACATCT
 CAGGGAACCAACCAACGATGAGCCTTAGGAGTCTGTACACCATCGTGTATGCCGCTTGTACGCTGGGCTTAAAGGAG
 CTCCTAATTTGGGCTCGAGAAATGATGGTGGATCTTACGCACTTACTTCCACTCTACTAATAATGACAATCTCAAGGC
 ACTGTTTGGCTCAGATATCACCGGAGAACGGCAAGCATTTTACCTCAACATTAAGTTGCCCTTCGATTTTGGCAGTTGGCA
 AAATCTTGGAGGTGTGGACAGATTTAATCTCGTACTTCAAGTCTTGTACAGAGGGCGATGGGAAAGCCAAGTGCATCTGCTCC
 CAGTGAAGGATCCCAAAGTCCAAGCCACATTTATGTTCTGACAGCAGGCCCTGAAGCCCCTGCAAAGGTTTACAGTCAAAT
 CCAGGCAGCTTCTGGGATAATATGTCCCTTATCTAGAAGAAGCCAGCTCTTCTTGGGACGTCACAGTCTTACTTCTTCA
 GCCTCAAGCTGTTATTGACTACCTCAAGAAACCCGATCCGGAATCCTAAAAAGCAAGGAACCTCAGCTGTCAATTCGAAGAT
 CTCAACCTGGGAGGCATGGAAGACTTCAAAAAGCGTCCGAGGCAACAGCAGTCTGTGCGAGGTAGAAGAAGATGCATTGT
 CCTTCTATGTTGCACTACCCGGTGCATTGTAGAGAAGCTGCCTTAAAGCGACTTGGAGCTAAAAAGCGTTGGTTCAGCTGCTT
 AACCTCAGAAAGGGCTGGACATGACAGCAGCAGCAGTAGAGAAGCTGGGAACAAAGCTGGGGTCTGAGCTCCCCTGAG
 GAGTCAAAGAGCTCACAGCAGAGTTTACAGATATCAGTCACTGGAGAGGGAGCGGGGCAAGCAGCCCTGGTTTCTC
 TGGAGAAGCAGTGGACGAGCGCTTAAAGACACCAGACCCTTGTCTTTCAGGAAGCTTGTGTTGACCCTGCTGTCTGCTC
 CCGTGTCTCTCTAGATTCTCAACGCGTCTTGTGAGGCTTAAAGGACGGAGACTCCGAGATCTTTCTGAGACTGAAGA
 TACTGCAGAAAGGGACTGTGACGCCACCTTACTCAGTCTTAGAAACAGCGCTGAGGTGAAAGACTCCAGGATCACACT
 CCAGAAACAAATGGCCTTGTGCGGCGGTGAAGCCTTGTGAAGTTCAGCTCACAAAGATGAGTATCTCAGTGGCGGTGTG
 GTGGAGTTAATGGCGGTTCCAGAATGAGGGGACCAAGTGGAGGTTTGGCTGGGAGAGCAGCTTACGTCAGAAAGCACC
 GGCCCGTACGGTGTTCAGGCTGGTGTGGCTCCTGGTCCAAACCCATAGGCCTCGATGCGGACCGCAGTAACGGTACAAGT
 CCAAAGACGAGGCGTCCCCTTTCAGAAAATCCACTCCAACAAAGGACAGCACAAACAGTACCAGGATGGGAAGGGATTCC
 AGATGGGGGAGCTGGTGTGGGGCAAACCGAAGGGCTTTTCCATGTGGCCTGGGATGGTGTGCGTTGGAAAACCTAAGTCTC
 TCCCAGGCATGAGAAGGTTGGAGTGGTTTGGGGATGGCATGTTCTCAGAGATCTACACTGACAAAACCTGCTGCCGTTTGTG
 CCTTACGCCGGTGTCTTGCAAAAATTCCTTGGCAGCCTGCCCTCTACAGGGAAGCTCTCTACCAGATCATCGAGTTGGCA
 GCTGAGCGATGCGGGAAGTCTTTGCTCATGCAGAAGGTAAGAAACAGGAGCTGAAGCTGATGGTGGACTGGGCTTTTGAAG
 GATTTCTGCTACTGACGAAATAAGGACTTACCCTGACTCCTGCTTCCAGTCTGCTGACTCTCCCTGCTGACT
 CTCAGCCCGCACAAAATAATATGCTTTAAAAGTAAAGCGTGGCTGTGATGGTCTTCAACCGAAGAAAAGATCAGATA
 AAATCTCAAAGAGAAAAACAAAATATCAGAGATTTTGTGTTGCTGTTGCTGCGCTGAGATCAACGTGGAACATCCTTTGT
 TTGAGGGTGGATTATGTCTGAGATGCAAGGAGAATTTCTCAGAGACGCTATATCGCTACGATGAAGATGGCTATCAGTCTAT
 TGCACCGTGTGCTGCGCGGGTTCAGAGGTCATCTGTGTGGAACTCCAGCTGCTGCAGGATATTTGGTGTGCGAGACCTCG
 GCTTAAAGGTGGAGCATATACCTGCTCAGAGATCGATGACTCGATGCGCGTGGGAATGATCAACACGAAGGAAGAT
 TGAATATGTTAACGATGTTTCGCAACATTACAGGAAACATCTGGCTGAGTGGGGTCCGTTTGTATCTGCTGATTTGGTGGCAGCC
 CGTGAACGACCTGTCCATGGTCAACCTCTACGAAAAGGACTGTTTGGGGCACCGGAAGACTCTTTTGTAGTTTATCGT
 ATTCTGAACCTTTTGAAGCCAAAGGAGGGAGACGACCCTTCTTCTGTTTATTCGAGAACGTGGTTTTCATGAGTGCCCA
 TGACAAATCAGACATCTCCCGTCTTCTGGAGTGAATCCAATCTTATTGATGCAAGTCAAGTCCGGCTCATAGAGCTC
 GCTACTTTTGGGAAACCTACCTGGCATGAACAGACTCTTCTCTGTTTATTGATGACAAAAGTCAACCTCAGGATTTGTTGG
 AAGTGGGGCGCAAGGCAATGTTTGAAGAAGTTCGACAATCACGACAAAGTCCAACCTCATAAGGCAAGGGAAGGTAGGCT

KSSDYIKGSNLDAPQPPFRIGRIKEIFCNKRSNGKPDTEIKLRL YKFYRPENTHKGPKGAYHSDINQLYWSDEEATVSMTEVLTRCR
VEYAEDELVESVDYNSKGPDRFYFLEAYNAKTKSFEDPPNHARS AVNKGKGGKGGKGGKGAAPQEPDQEAQEP AVPKLRTL
DVFSGCGGLSEGFHQAGISETNWAIE MWDPAAQAFRLNPNPGTTVFTEDCNVLLKLVMSGEK TNSLGQKLPQKGDVEMLCGGPPC
QVFGSMNRFNSRTYSKFKNSLVVSYLSYCDYRYPKFFLEENVRNFVSKFRSMVLKLT LRCLVRMGYQCTFGVLQAGQYGAQTR
RRAIILAAAPGEKLP RYPEPLHVFAPRACLSVA VDEKKYVSNVTRGNGGIYRTITVRDTMSDLPEIRNGAAALEISYNGEPQSWFQ
RQIRGSQYQPILRDHICKDMSALVAARMRHIPLAPGSDWRDLPNIEVRLDGT TTKLRYTHSDKKNRSGTGALRGVCSCEGKQ
CDPADRQFNTLIPWCLPHTGNRHNHWAGLYGRLEWDGFFSTTVTNPEPMGKQGRVLHPEQHRVSVVRECARSQGFPD TYRFFGN
VLDKHRQVGN AVPPPLSKAIGLEVKKCVLEKMRENATEPVKQEKMELSD

>Anolis carolinensis DNMT1

MVPSSAGPAPPSSPRCARNCSRPPSRSLPTAASLPVCLAGPEMPARAAPLTPEELRWRLKDLERNEDGLGERECVKEKLSLMHGF
LQADVQNQLKDLETKFHKEELSEERYLAKVKALLHKELSFENGDNLELGQKTNGCAENCANGSDEEPMGNKTEEDSGMEIEEA
ASSSSVPLASPVLKARKGRKSKSNSESRTKTPSSSRVTRSSGQOPTVAMFSKGSNKRKSEE VNGKLLQETNLENEEEEEEEQFGEK
EPDVKRFKTEKTEGLERKEEATQIKTGTPVKTTTPKCAECRQYLDLDPDLKFFQGDNDALDEPEMLTDERLSIFDANEDGFESYDD
LPQHKVTSFSVYDRKGHLCPFDGTGLIERNIELYFSGVVKPIYDDNCLDGGVRRARKMGPINAWWITGFDGGEKALIGFTTAFADYIL
MEPSEEYAPIFALMQEKIYMSKIVVEFLQNNPDVSYEDLLNKIETT VPPAGLNFNRFTEDSLLRHAQFVVEQVESYDEAGDSDEPPV
LITPCMRDLIKLAGVTLGKRRRAARRQAIRHPTKIDKDKGPTKATTTKLVYLFIDTFFSEQIEKTEKEEDNENVSKRRRCVCEVCCQ
PECGKCRACQDMKIFGSGKTKQACFQRRCPNLA VKEADDDEEVDDNIPEVSPKMLQGRKKKQSKSRISVWGHVPMKSDGKGD
YYQKVCIDSETLQIGDCVSVSPDEPTKALYLARITAMWEDVNSSEPMFHVHWFRCRGTDTVLGATSDPLEFLVDECPDLEFLVSIHG
KVNVLKAPSENWSLEGGDLVEIKMVEDDGRTYFYQMWYDQYARFETPPKIQPTEDNIHKFCMSCIRLDEVRQKEIPRVLEPQEE
SEGKMFYSMAMKNGVQYRVGDGVFLLPDAFSLRLASPIKQKKEAVNEELYPEHYRKYSEYIKGSNQDAPEYRIGRIKEIYCNI
RSSGRPNEAEIKLRVYKLYRPENTHKS VKASYHADINLLYWSDEEASVEFKDVQGRCTVVYGEDL TENIQDYSASSDRFYFLEAY
NAKTKFEDPPTV HARSSGNKGGKGGKGGKSGKSS TASEHSEAAKMKPPKLRSLDVFSGCGGLSEGFHQAEVSETLWAIEM
WEPAAQAFRLNPNPGTTVFTEDCNVLLKLVMSGEK TNSLGQKLPQKGDVEMLCGGPPCQVFGSMNRFNSRTYSKFKNSLVVSYLSY
CDYRPRFFLEENVRNFVSKFRSMVLKLT LRCLVRMGYQCTFGVLQAGQYGAQTRRRAIVLAAAPGEKLPMPPEPLHVFAPRAC
QLSVVDDKFKVSNITRYSGPFRITITVRDTMSDLPEIRNGASALEISYNGEPQSWFQRQIRGSQYQPILRDHICKDMSALVAARMR
HIPLAPGSDWRDLPNIEVRLDGT TTKLRYTHHEKKNRSGTGALRGVCSCEGKACDPADRQFNTLIPWCLPHTGNRHNHWAGL
YGRLEWDGFFSTTVTNPEPMGKQGRVLHPEQHRVSVVRECARSQGFPD TYRFFGNVLDKHRQVGN AVPPPLAKAIGLEIKSCVSA
KIKEGATDTIKREKMDTSD

>Xenopus tropicalis dnmt1

LTVSPLLYTAMPAQASALPADVRKRLKDLERDQDGMTEKECVQKLSLVLGFLFLEADARNKLSDESLSSEELSEEGLTKV
KALLRKLQSCENGDSLNGETNGCSTNGTCGSD EEDVEMSESNTAGVKNRKRTRKSKTNGENKTESPARARSSRSTTGKQATILSMF
SKGSSDLAMHQHSCNITFDNLITSDGIRRRCSIPFQHIMLEYQFQVDKRPSTSSVKFQWSLALTSNSTESSVLSILKASPFYITISGLLR
SNKRKSSDEEKD TDVSADADQPEEKEKEEKRIKIEVNESENERSTAEESKVKVPVLPKTPPPKCMDCRQYLD DDPDLKYFQGD PD
DALDEPEMLTDERLSLFEANEDGFESYDDLPHQKVTCSFVYDKRGHLCPFD SGLIEKNVELYFSGYVVKPIYDDNPSLDDGGVRAKLL
GPNALWITGFDGGEKALIGFTTAFADYILMDPSEYSSIFALMQEKIYMSKIVVEFLQNNPDVSYEDLLNKIETT VPPSGLNRRFT
EDSLLRHAQFVVEQVESYDEAGDSDEQPVIVTPCMERLILLAGVTLGKRRRAARRQAIRHPTKIEKDNKGPTKATTTKLVYQIFD TFF
SEQIEKDADKENGIRRRRCVCEVCCQPCGQCKACQDMLKFGGAGRTKQACTQRRCPNLA VKEADEDEEVEDVVPMPSPKKI
LQGGKKLEKKNRISVWGDPIKTEGKKDFYLYKVSIDSEILEVGDVSVSPDNPTPELYLARITSMWEDACGPMFHAHWFC LGTDTV
LGATSDPLEFLVDECEDMQLSYIHGKVKVIYKAPSDNWFMEGGTHTDIK VVEDDGTSTYFYQLWYDPEYARFETPTPTQPTEDNK
YKFTCSARLAEIRQEMPRVSCPVEDLDSKVCYSTAIKNDVQYKVG DGVLLLPDAFSSVKLGSPMKRPQKDDVDEDLYPEYY
RKSSDYIKGSNLDAPPEYRLGRIKEIFCNKRSNGKANESDIKLR IYKFYRPENTHKGVKASYHSDVNMVYWSDEEAVVDFKAVQG
HCTVEYGEDLTETIQEYSAAGSDRFYFLEAYSAKTKSFEDPPNHARS AVNKGKGGKGGKGGRTQTNSKGGKASKSENEQQNCGDK
LPLKRLTLDVFSGCGGLSEGFHQAGISETNWAIE MWEPAAQAFRLNPNPGTTVFTEDCNVLLKLVMSGEK TNSLGQRLPQKGDVEML
CGGPPCQVFGSMNRFNSRTYSKFKNSLVVSYLSYCDYRYPKFFLEENVRNFVSKFRSMVLKLT LRCLVRMGYQCTFGVLQAGQY
GVAQTRRRAIVLAAAPGEKLPMPPEPLHVFAPRACLSV VVDEKKYVSNITR TNSLFRITITVRDTMSDLPEIRNGASALEISYNGEP
QSWFQRQIRGSQYQPILRDHICKDMSALVAARMRHIPLAPGSDWRDLPNIEVRLDGT TTKLRYSHHDKKNRSGTGALRGVCS
CEGKPCDPADRQFNTLIPWCLPHTGNRHNHWAGLYGRLEWDGFFSTTVTNPEPMGKQGRVLHPEQHRVSVVRECARSQGFPD
TYRFFGNVLDKHRQVGN AVPPPLSKAIGSEIKKCVLSREKENG TETVKA EKMET

>Anolis carolinensis DNMT3A

MNAVEEALPRSEPK EEEASPPASQPTDPASPNVATTPEPLGPAADKGM SKTADDEPEYEDGRGFGIGELVWGKLRGFSW WPG
RIVSWWMTGRSRAAEGTRVWMWFGDGKFSVVCVEKLLPLSSFANAFHQATY NKQPMYRKAIYEV LQVASSRSGKIFPSCPENDE
TDTSKAVEIQNKQMI EWALGGFPQSGPKGLEPPEEERNPYKEYVTEMWVEPAAA YAPPPPAKPRKSTAEKPKVKEIIDERTRERL
VYEVQRKCRNIEDICISGSLNVTLEHPLFIGGMCQCNCKFLECA YQYDDDGYSYCTICCGGRELVMCGN NCCRCFCVECV
LVGPAAQAIAIKEDPWCYMCYMGHKG IYGLLRRRDDWPSRLQMFANNH DQEFDPKVPYPPVAERRKPIRVL SLDGIATGLLVL
KDLGIQVDRYIASEVCEDSITVGMVRHQKIMYVGDVRNVTQKHIQEWGPF DLVIGGSPCNDLSIVN PARKGLYEGTGRLFFEFYR
LLHEARPKEGDERPFFWLFENVVAMGVSDKRDISRFLSNPVMIDAKEVSA AHRARYFWGNLPGMNRPLASTVNDKLELQECLEH
GRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYT DVSNSRLARQRLGRSWSVPVIRHLFAPLKEY
FACV

>Homo sapiens DNMT3A

MPAMPSSGPGDTSSSAAEREEDRKDGEEQE EPRGKEERQEPSTTARKVGRPGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASE
LLPNGDLEKRSEPP EEGSPAGGQKGGAPAE GEGAAETLPEASRAVENGCC TPKEGRGAPAEAGKEQKETNIESMKMEGSRGRLR
GGLGWESSLRQRPMPRLTFQAGDPYI SKRKRDEWLARWKREAEKKAVIAGMNAVEENQGPGESQKVEEASPPAVQPTDPAS
PTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKLRGFSW WPGRIVSWWMTGRSRAAEGTRVWMWFGDGKFS
VVCVEKLMPLSSFCFAHQATY NKQPMYRKAIYEV LQVASSRAGKLPVCHDSDESDTAKAVEVQNKPMIEWALGGFPQSGPKGL
EPPEEENPYKEYVYTDMMWVEPAAA YAPPPPAKPRKSTAEKPKVKEIIDERTRERL VYEVQRKCRNIEDICISGSLNVTLEHPLFV
GGMQCNCKNCFLECA YQYDDDGYSYCTICCGGRELVMCGN NCCRCFCVECV DILLVGPAAQAIAIKEDPWCYMCYMGHKG IY
GLLRRREDWPSRLQMFANNH DQEFDPKVPYPPVAEKRRKPIRVL SLDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQ
KIMYVGDVRSVTQKHIQEWGPF DLVIGGSPCNDLSIVN PARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGV
DKRDISRFLSNPVMIDAKEVSA AHRARYFWGNLPGMNRPLASTVNDKLELQECLEH GRIAKFSKVRTITTRSNSIKQGKDQHFPV
MNEKEDILWCTEMERVFGFPVHYT DVSNSRLARQRLGRSWSVPVIRHLFAPLKEYFACV

>Mus musculus Dnmt3a

MPSSGPGDTSSSLE REDRKEE EENRGKEERQEPSATARKVGRPGRKRKHPPVESDTPKDPAVTTKSQPM AQDSGSPDLLP
NGDLEKRSEPP EEGSPAAGQKGGAPAE GEGTETPEASRAVENGCCVTKEGRGASAGEGKEQKQTNIESMKMEGSRGRLRGGLG
WESSLRQRPMPRLTFQAGDPYI SKRKRDEWLARWKREAEKKAVI VMNAVEENQASGESQKVEEASPPAVQPTDPASPTVA

TTPEPVGGDAGDKNATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRVWVWFGDGKFSVVCV
EKLMLPSSFCFAHQATYQNKQPMYRKAIYEVLVASSRAGKLPACHSDSESDSGKAVEVQNKQMIWALGGFQPSGPKGLEPPE
EENKPYKEVYTDWVVEPEAAAAPPPAKKPRKSTTEKPKVKEIIDERTRELRVYEVQRKCRNIEDICISCGSLNVTLEHPLFIGGM
CQCKNCFLECAQYDDDDGYQSYCTICCGGREVLMECGNNCCRCFCVCEVDLLVGPAAQAIAIKEDPWNCYMCCHKGTGYLLR
RREDWPSRLQMFANNHDQEFDPKPYPPVPAEKRPKIRVLSLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIM
YVGDVRSVTQKHQEWGPFDLVIGGSPCNDLSIVNPARGLYEGTGRLLFFEFYRLLHDARPKEGDDRRPFFWLFENVVAMGVSDKR
DISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVPMN
EKEDILWCTEMERVFGFPVHYTDVSNMSRLARQLLGRSWSVVPVIRHLFAPLKEYFACV

>Gallus gallus DNMT3A

MVESDTPKDTAAVPCPPCPPEASPAEPLPNGDLEADGAQWKGTEEGGASPKSGRPEEDETESLADGETGRALENGRCTPKEGLD
APADEGELAPSDPQKRRGRKLEATEKSKEEKEENNFDLSKMEGSRGRLRGGWESSLRQRPQRHTFQAGDPYYISKRRKDE
WLARWKREAEKAKVIAVMNVVEETPRAEPQKEEASPPASQOPTDASPNAVATPEPVVADAVDKNTSKSADDEPEYEDGRGL
GIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRVWVWFGDGKFSVVCVEKLLPLSSFFSAHQATYQNKQPMYRKAIYEV
QVASSRAGKIFPACPENDETDTSKVEIQNKQMIWALGGFQPSGPKLEPPEERNPYKEVYTEMWVEPEAAAAPPPAKKPRK
STTEKPKVKEIIDERTRELRVYEVQRKCRNIEDICISCGSLNVTLEHPLFIGGMQCKNCFLECAQYDDDDGYQSYCTICCGGREVL
MCGNNCCRCFCVCEVDLLVGPAAQAIAIKEDPWNCYMCCHKGIYGLLRREDWPSRLQMFANNHDQEFDPKPYPPVPAEK
RKPIRVLFLFDGIATGLLVKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVNRNVTQKHQEWGPFDLVIGGSPCNDLSIVN
PARKGLYEGTGRLLFFEFYRLLHEARPKEGDDRRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRARYFWGNLPGMNR
RPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVPMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQLL
GRSWSVVPVIRHLFAPLKEYFACV

>Xenopus tropicalis dnmt3a

QAESATVSKDIASVPKCHTSPNSARLDDPIHNGEMDSRQHKAKRPTQDRPAPELSPRDTKTCDSPEEVRPDSVTPVVENGCCTPSDD
IDPQDMKEDCVSVISTPKRGRRLSSSEQSKEEEREDDNDSPKEEGYRGLRGGWESSLRQRPQRHTFQAGDPYYISKRRKRD
EWLARWKREAEKAKVIAVINAMEETVSREPQKEEPCPPASQOPTDASPNAVATPEPVVADAVDKNTSKSADDEAEYEDGRGF
GIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRVWVWFGDGKFSVVCVEKLLPLSSFFAGTFHQATYQNKQPMYRKAIYEV
QVASTRARKVFPSCPENDESSTKAVEIQNKQMIWALGGFQPSGPKLEPPEERNPYKEVYTEMWVEPEAAAAPPPAKKPRK
STTEKPKVKEIIDERTRELRVYEVQRKCRNIEDICISCGSLNVTLEHPLFIGGMQCKNCFLECAQYDDDDGYQSYCTICCGGREVL
MCGNNCCRCFCVCEVDLLVGPAAQAIAIKEDPWNCYMCCHKGIYGLLRREDWPSRLQMFANNHDQEFDPKPYPPVPAEK
RKPIRVLFLFDGIATGLLVKDLGIHIERYIASEVCEDSITVGMVRHQGKIMYVGDVNRNVTQKHQEWGPFDLVIGGSPCNDLSIVN
PARKGLYEGTGRLLFFEFYRLLHEARPKEGDDRRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAREVSAHRARYFWGNLPGMNR
RPLASTVNDKLELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFVPMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLARQLL
GRSWSVVPVIRHLFAPLKEYFACV

>Danio rerio dnmt3ab

MPSNTVATAAAAFKSSDFNCDTNGNQMNESMETMIIDEMPSLCTPVRKVGRPRGRKHLHVTCYGPLNQLMAKQHKAELOSS
LQSQNGEVEHRENKSIIEGRHSTLDRKRVENGLFCQSGCKAESTSDSSLSPLSENGFVSSQDSEKCEETLPTLPRKGRGRKLERPT
KYVEHKEDESDGPKNEGARGRLRGVWEVSLRQRPMPRVTFQAGDPYYISRRTRELLAKWKLEAEKAKVIAVMSVMSMEDH
GEVESEPKKDVITVPLLPMPPTPTTPMQQPPQQAQHQTNLKPQSPQQPPPPFQQQQQPPQPTDASPNAVATPEPVVADAVDKNTSKS
ADDEAEYEDGRGF GIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRVWVWFGDGKFSVVCVEKLLPLSSFFHNAFHQP
TSPKSTESEYEDGRGFINELVWGKLRGYSWWPGRVVSWRMTRRSPAAGTRWVTFWFGDGKFSVVCVEKLLPLSSFFHNAFHQP
TYNKQPMYKKAIFEVLQVASTRAGKTFLSCPASDDTSSSKLVDVENKMIWAMSGFQPTGPKGLDPPESPEYEMCEWVEPEAAS
YLPPAKKPRKSTTEKPKVKEIIDERTRELRVFDVQRKCRNIEDICISCGSLNVTLEHPLFIGGMQSCNTCFLECAQYDDDDGYQ
SYCTICCGGREVLMECGNNCCRCFCVCEVDLLVGPAAQAIAIKEDPWNCYMCCHKGIYGLLRREDWPSRLQMFANNHDQEF
EPPRLYPPVLAENRKPPIRVLFLFDGIATGLLVRELGIQVERVYASEVCEDSITVGIVRHQGRIMYVGDVRLTRKNIQEWGPFDLVI
GGSPCNDLSIVNPARGLYEGTGRLLFFEFYRLLHEARPKEGDTRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEVSAHRA
RYFWGNLPGMNRPLSAMCTDKLDLQDCLEHGRTAKFKGVRTITTRSNSIKQGKDQHFVPMNDKEDILWCTEMERVFGFPVHYT
DVSNSRLARQLLGRSWSVVPVIRHLFAPLKEYFACVGT

>Homo sapiens DNMT3B

MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDPPILEAIRTEIPRGRSSSRLSKREVSSLLSYTQDLTGDDGEDGDGSDTPVM
PKLFRRETRRSESPAVRTRNNNSVSSRERHRSPRSTRGRQGRNHVDESPVEFPATRSLRRRATASAGTPWSPSSYLTIDLTDDE
DTHGTQSSSTPYARLAQDSQQGGMESQPVEADSGDSEYQDGKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMSGM
RWVWFGDGKFESEADKLVALGLFSQHFNLAFTNKLVSYRKAMYHTLEKARVRAGKTFSSPGDSELDLQKPMLEWAHGGFK
PTGIEGLKPNNTQPVVNSKSVRRRAGSRKLESRYENKTRRRRTADDSATSDYCPAPKRLKTCYNNNGKDRGDEDQSRQMASDVA
NNKSLEDGCLSCGRKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDDGYQSYCTVCCGRELLECSNTSCRCFCVCELEVLVGTG
TAAEAKLQEPWSCYMCLPQRCHGVLRRRKDWNVRLQAFSTSDTGLEEAPKLYAIPAARRRPIRVLFLFDGIATGYLVKELGKIG
YVYASEVCEESIAVGTVKHEGNIKVYNDVRNITKNIEEWGPFDLVIGGSPCNDLSIVNPARGLYEGTGRLLFFEFYHLLNYSR
PKEGDDRRPFFWLFENVVAMKVGDKRDISRFLESNPVMIDAIKVSAAHRARYFWGNLPGMNRPIASKNDKLELQDCLEYNRIAKL
KKVQTITTKSNSIKQGKNQLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGRGARQKLLGRSWSVVPVIRHLFAPLKYDFACE

>Mus musculus Dnmt3b

MKGDSTRHLNEEGASGYEECIIVNGNFSQSSDTPKAPSPVLEAICTEPTCTPETRGRSSSRLSKREVSSLLNYTQDMTGDGDRD
DEVDDGNSDILMPKLTRETKDTRRSESPAVRTRHSNGTSSLERQASPRITRGRQRHVVQEYVPEFPATRSLRRRASSASTPW
SSPASVDFMEEVTPKSVSTPSVDLSQDGDQEGMDTTQVDAESRDGDSTEYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSK
RQAMPGMRWVWFGDGKFESEADKLVALGLFSQHFNLAFTNKLVSYRKAMYHTLEKARVRAGKTFSSPGDSELDLQKPMLEW
AHGGFKPTGIEGLKPNKQPVVNSKSVRRSRSRNLPRRRRNRKSRRTTNDSSAASESPPKRLKTNYSYGGKDRGEDEESRERMASE
VTNNKGNLEDRCLSCGKKNPVSFHPLFEGGLCQSCRDRFLELFYMYDEDGYQSYCTVCCGRELLECSNTSCRCFCVCELEVLVGTG
AGTAEDAKLQEPWSCYMCLPQRCHGVLRRRKDWNVRLQAFSTSDTGLEEAPKLYAIPAARRRPIRVLFLFDGIATGYLVKEL
GIKVEKYIASEVCAESIAVGTVKHEGNIKVYNDVRKIKKNIEEWGPFDLVIGGSPCNDLSIVNPARGLYEGTGRLLFFEFYHLLNY
TRPKEDNRPFWMFENVVAMKVNDKDISRFACNPMIDAIKVSAAHRARYFWGNLPGMNRPMASKNDKLELQDCLEFSRT
AKLKKVQTITTKSNSIRQGNQLFPVVMNGKDDVLWCTELERIFGFPVHYTDVSNMGRGARQKLLGRSWSVVPVIRHLFAPLKYDF
ACE

>Danio rerio dnmt3ba

MATNVSLEPNPDCKSRYEVLGWINETLQNTFTQVEQCRSGACFCQLIDLFPGTINLKKVKFESQKRSDFMQNYSLQAADRDL
EVTPEPVNELLSGKFRPNFTYLKWFKFFYANVQKERRVYNAFAEARDQEQIYVDDVMKSPKALKSSYESGREAESDMEINGGR
RSATYDPKWRNLKWRASDMGDNYAYCTCDYNIERLAGFHLKRDQLTQNHMKHETGRNLPGRKQIEESISCESTMILLFIQSH
CLSSLPSRINRVSQRTARCILGLKYPNDIVSACKLNPYCIYIGQVDLKVTKGDKTNCHEVVLVLAGFFEEKQARCYRFLDVFQPEDSAS
SVSGGLFSILKKEIPASNMAVYINDHELTSVVSQIRELNPQVIDLGLYSIPTACSAQLQTHSVQVEELIANIYRHFTGSTSND

AAHRARYFWGNLPGMNRPLCASGMDKLELQDCLEHGRVAKFGKVRTITTRSNSIKQGGKQDHPVPLMNGKEDILWCTELERIFGFP
VHYTDVSNMGRGARQRLGRSWSVPVIRHLFAPLKDYFACE

>Oryzias latipes DNMT3b 1443aa

MAGTEAVLVKWLNNLLKTNFQEVQEMSSGACYCQIMDCPTIPGSIDLAKVKFDAQCEDDYKHNFLLQEAFKKQGITEGIPVEELIQ
GDIKTNSEFLKWKIFHLTSKNQKAVKNNRDISPNPASPNNRSDQLGTEKNNLVAATGKGSLYTEKWKETFDWAERSTLGGDYTF
CTSCQSNLHTINKGLVELQRHSETKHKHKGTLNLYEGNNSQNKTKLPCSDAALRFIRSRSNFAPMNGEKPSKHFLYRTLGRRYPKDI
ASVCRHTPYCVVYVEEVSVGKDQHVSVMLVGFYDLQTAHRSVFLDAFESESGDQTASYVFTLEKFEPAENLAAVYVSGDSQL
SAQICSELRLKLPILISIGPLYSIADAACCAGVKKLSGLVQELLSALEEQYFSSSIQSDLEALLDSVYDEKSSFLDIGCLKFCRLISKI
LEMWTDLTRHFSRNDEEAELICSQLKDPKVKATFLFLRKLALKPLQKQFQTQTLSDSWTSMILLIKEASSLLSTYTSCFLQSSAAARY
LEEHDARILRSKSLSSPDLDLGGNEVEDALRACDEPAALPHFCEQALSFYAALTGCIAEELPLSVKALRSIAQLLSPQSWLKVTS
QVKELGTQLGVCNSTDEIRQLTEEFLECKPTREVGPAGTTVDLSLQHWRSRLLDGAKPASLFRRLVLTLLSLPCPPDPKQVLSRVL
ETTEASSDEESTDSCDSTCSRKDISIHPHVQNPQKLDLLKALRPEVRLIKIDCLNIDGYENFISERNHRGSLGCESSLRKPPAR
TVFQAGVSTWCKPVGFDDKAAPELRDEAGEERPHSGKHTPTKVRKHNYQDGGKGFVGDVWGRVKSLSLWPGMVPWKS
PVGMRVVEWFGDGMISSETNTEGLLPFSTFSKSFCKNSYASLPMYKKAIEILELAAERSRKSFADEAGDKEKELKMLLEWAFEGFL
PTGPEGFAPPSVQKEPSEISEDYPPKRRKYLKSNRSTSAVTNRESLIEMIKEPGKTIEDFCLSCSSTEVVDVQHPLFEGGLCLRCKE
NFSETLRYRDEGDGYQSYCTVCCAGLEVVLCDNVNCCRCFCCKDCLNILIEDGTFEKLRDTEWNCFCMTPSQCKGNLQRPDWSVR
VQDFFANSLAEFEPHRVYPSIPVAERRPLKVLFLDGIATGYLVKHLGFKIERYIASEICDDSIAGVMKHEGKIEYVNDVTRITRK
HLAEWGPFLLIGGSPCNDLATVNPNRKGLYEGTGRLLFFEFYRILNLLRPREEDNRPFVWFENVVMNSRDKTDICRFLCNPVLI
DAIDVSPAHRARYFWGNLPGMSRRVATSLDDKVLNQCLEIGRVAKFDKVRTITTKTNCMRQGGKGTGQLPVMINGKDDSLWCTEM
ELIFGFPKHYTDVNNMSRMQRQRVLRGWSVVPVIRHLLAPLKDYACEERV

>Takifugu rubripes DNMT1 1493aa

MPTRTSLSLPEDVKKRLQVLDDEDGSTEEDHVKEKLLVQKFLHIDAQDQLTSLEESMKNSEISMEVYISKVKALLGKELHLENGNS
NGYTNNGSHKDEDGEEASMDVQEGEETMKSPTASKGKGGRRSKANSDTKKSPASTRVTRNSGKOPTIMSMFSKVQKRKSEDLNGD
STNGQNEDEKEDDVEEAAREEKRKLVESNENESTEEIKPETKVPVPTAKTPPPKQDCRQYLDSDLKFFQGDPDNALDEPEMLTDE
RLSLFDSNEDGFESYDLPQHKITNFSYDKRGHLCPFDNSGLIEKNVELYFSCVVKPIYDDNPLDGGVPAKLLGPINAWHITGFDG
GEKALIGFTTAFADYILMHSSDYAPIFALMQUEIKYMSKIVVEFLQKNREATAIEDLLNKIETTVPAGLNFNCTEDTLRHAQFVVE
QVESYDEAGDSDEERIVTPCMRDILLAGVTLGKRRARRQAIRHPTKIEKDSKGPATKATTKLVCQIFDFFSDQIEQNDKEVVG
VKRHRGCVCEVCQSPDCGKCTACKDMIKFGGSGKSKQACKQRRCPNLAVKEAEDDENVEEDDAPVEKTKKVSASHAKRKKQTECN
LTDWIGEPVLTGKRNYYKQVINSIDVLEVGDCVSVSSEDDPSVPLILARITSLWEDTHGKMFHAHWFMRGIHTVLGECSDPLELVLV
DDCEDMQLNYVQKVNIMYKAPSNNWFMEGLDVEDIKVIADDDGKSFYQFQWYDPEFARFETPPKTPPSEDCKFKFCSSCVRTTEQ
KEQEIPHFVLELDESDSKALFALACFKGEQFKVGDVSYLPPDAFNFSVKPASPMPKRSRHRKDDVDEELYPEYRKSDDYIKGSNLD
APDPFRIGRIKEIFCHKRSNGKPDTEVVKLRLYKPYRPNETHKGVKATYHTDINQLYWSDEEVTVDMAQVLRGRCQVEYGEDLNEPV
QEYSSGGPDRFYFLEAYNAKTSFEDPPNHARSAAHKGKGGKGGKGGKAAQDSDHTPNPQMPKVPKYRTRLDVFSGCGGLS
EGFHQAGISELWAIEMWEPAAQAFRLNNGPATTFTEDCNILLKVMGSEKTNLSGQKLPQKGDVEMLCGGPQSGSMNRNFS
RTYSAFKNSLVVSYLSYCDYRPPKFFLLENVRFVSKSSMVLKLTLRCLVRMGYQCTFGVLQAGQYGAQTRRRAILAAAPGEE
LPRYPELHVAFAPRACSLNVVDDKFKVSNVTRGNGGIYRTITVRDMSDLPEIRNGASALEISYNGEPQSWFQRQIRGTQYQPILR
DHICKDMSALVEGRMRHIPLAGSDWRDLNIEVRLRDGMTKLLRYTHLDKKNRSGSTGALRGVCTCAGGTPCDPADRQFNLI
PWCLPHTGNRHNHWAGLYGRLEWDGFFSTVTNPEPMGKQGRVLRHPEQHRVVSVRECARSQGFDPDYRFFGNILDKHRQVGNV
PPPLSRAIGLEIKCVTMRIKQEQAENVKHEKMELSE

>Takifugu rubripes DNMT3a-like 983aa

MPSNTVAGTFAFKGSDYTDSPGRMGESTIMDVLMTEDKSASASYKSGKPGRRKRKHPGEGSCGPFKDAVGGGHSYPMGRSLQ
VYNGDADGHRERMSDADCFPKQERRDAENGIPRDSDFCRAEDSPQGSWSVPQENGLQPSDEQDEKDKDSSLTLRKRKRGRRL
ERPTKYVEHKEEDGVDVYKTEGDRGRLRGGVWWEISLRQRPMPRITTFQAGDPYYSKRRTREEWLAKWKLEAENKARQMSAMSAM
KDHKESETESREEEVSIKHPPPSKLLQPPQPHSQPPQPHHQQSLPQQQQHQPPQQQQQPTDPASPTVATTPEVPAIGGAEKTS
PNPADTEPEYDDGRGFGIGELVWGLKRGFSWVWGRIVSWWMTGRSRAAEGRVWVWVWFGDGKFSVVCVEKLLPLSSFNAFHQPT
YKQMYKKAIEYVQLVQASRRAGKAFMACPDSDETESKSVELNKQIMIDWAMTGFQPTGPKALEPPEERNPYTEVPEMWPVE
ETAAYTPPAKPKRSTAEKPKVKDIIDERTRELRHIDVRQKCRNLEDICISGSLNVSLHPLFSGGMCQSKNCFLECYQYDDD
GYQSYCTICCGGREVLVLMCGNCCRCFCVCEVDLLVGPAAHAAIKEDPWNCYMCQKGVFGLLARRIDWPSRLQHFANNHEQ
DFETPKVYPPVMAEKRPVRLSLFDGIATGLLVLKELGIEVDRYVASEVCEDSITLGIVRHQGRIMYVGDVRSVTRKHQIWEWGPFD
LVIGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRILHARPKEGDDRPVWFENVAMGVSDKRDISRFLCNPVMDAKEVSA
HRARYFWGNLPGMNRPLVCEWPLSSTNTDKLDLQCELEHGRVAKFEKVRTITTRSNSIKQGGKQDHPVPLMNGKEDILWCTELERIFG
FPVHYTDVSNMSRLARQRLGRSWSVVPVIRHLFAPLKDYFACE

>Kryptolebias marmoratus DNMT1 1508aa

MPSRTSLSLPPDVRKLEKLDDEDGSEEDQVKEKLLVQDFLHDDAQDQLTTLEEMKKESEISKEVYISKVKALLGKELHLENGSH
VDGAEQNGKANGFTNGSHKDEKEDGDVAMDTQEEEAASAAVVKSPASRGGKGGRRSKGSDTDTKSPASSRVTRNNGKQPTILA
MFSKVQKRKSDLENEAVNGTDELKEDDDKDESREEKRLKVDSEKPAPEELKTEVTKPVSTAKTPPPKCPDCRQYLDSDLK
FFQGDPDNALDEPEMLTDERLSLFDNEDGFESYDNLPHRITNFSYDVKHGLCPFDNGLIEKNVELYFSCVVKPIYDDNPLCEGG
VFAKLLGPINAWWITGFDGGEKALIGFTTAFADYILMQPSEYAPIFALMQUEIKYMSKIVVEFLQKNPDASYEDLLNKIETTVPAGL
NFCNCTEDTLRHAQFVVEQVESYDEAGDSDEQPIIVTPCMRDILLAGVTLGKRRARRQAIRHPTKIEKDSKGPATKATTTQLVYQ
IFDAFFADQIEQIDKEAGLKRQRCGVCEVCQSPDCGKCAACKDMIKFGGSGRNRKQACKQRRCPNLAVKEAEDDENIEEEDVQEEKP
KKVNLAKRKKQTQSKLAWIGESIQTVGRKQYKYSVNDVLEVGDCVSVSDDPSIPLYLARITSLWEDNNGKMFHAHWFRLGI
HTVLGESSDPLEIVVDDCEDMLLTYVQKVNITYKPPSNNWCMEGGVVDVVKVIEDDGKSFYQFQWYDLDFARFETPPKTCPTTE
CKFRFCGSCARTKVEREDQELPQFEPLDNEGDGDKVYFYSACFRGQFRVGDVSYMPPGAFHFSVKPASPVRKSRHRKEDVDELYP
EYRKSDDYIKGSNLDAPQFRIGRIKEIYCLRRSNGKPDTEVVKLRLYKPYRPNETHKGVKASYHTDNLVQYWSDEEVAVCLSDV
LGCCQVEYAEADLNEIQDYSSGGPDRFYFLEAYNAKSKSFEDPPNHARSVHKGKGGKGGKGGKGAQVEPQDSHIEPPKVKY
RTLDFVSGCGGLSEGHHQAGISELWAIEMWEPAAQAFRLNNGPATTFTEDCNVLLKVMGSEKTNLSGQKLPQKGDVEMLCGGP
PCQGFSGMNRNFSRTYSKFKNSLVVSYLSYCDYRPPKFFLLENVRFVSKNSMVLKLTLRCLVRMGYQCTFGVLQAGQYGAQ
TRRAIVLAAAPGKLPKPEPLHVAFAPRACSLNVVDDKFKVSNVTRGNGGIYRTITVRDMSDLPEIRNGASALEISYNGEPQSW
FQRQIRGTQYQPILRDHICKDMSALVEARMRHIPLAGSDWRDLNIEVRLRDGMTKLLRYTHLDKKNRSGTGALRGVCTCA
GGKPCDSADRQFNLIWCLPHTGNRHNHWAGLYGRLEWDGFFSTVTNPEPMGKQGRVLRHPEQHRVVSVRECARSQGFDPDYR
FFGNILDKHRQVGNVPPPLSRAIGLELKKCVLERMKEEAENIKQEKEVSD

>Kryptolebias marmoratus DNMT3a 985aa

MPSNTVTNTTAFKGFADSPGNRMGESTILDMDLMEEKSTLNSASCKVGKPGRRKRKPPVSCGSLTDSRSVGNYSYGVCRP
MTQVHNGDMLDGCDFPKQEKREVENGIPRDLSSQAEDSSQPSLSPSMENGLSREEQDLEKERDDGVMMPRKKRGRKLE
RPTKYVEHKEDEGCAIKTEGGRGRQRGGVWWEISLRQRPMPRITTFQAGDPYYSKRRTREEWMAKWKMEAEKAKLMSAMNTM
KDNKNETESEKKEEVSIIKHPPPSKLLQPPQPHSQPPQPHHQQSLPQQQQHQPPQQQQQPTDPASPTVATTPEV

TIGGEDKTSPKSADAEPYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLLPLSS
FNNAFHQPTYNKQPMYKKAIEYVLQVASSRAGKTFMACPDSDETTETSKSVEMLNKQMIWAKTGFQPTGPRGLEPPEEERSNSEV
YEMWVEPEAAAAYTPPAKPKRKSTVEKPKVKEIVDERTRERLVYEVGQAQRSIEDICISCGGLNVSLEHPLFGGGMCCQSCNCFLE
CAYQYDDDGYSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPAAALAAIKEDPWNCYMCCKPQVDFGLERRSDTRQMIW
HFFANNHDQDFDPPKLYPPVMVEKRPVLSLFDGIATGLLVKELGIQVDRYVASEVCEDESITVIVRHHGRIMYVGDVRNVTR
KHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLHEARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLECNPV
MIDAKEVSAHRARYFWGNLPGMNRPLTAMCTDRLLDQDCLEHGRTAKFGKVRTITTRSNSIKQGGKQDQHPVYMNEKEDILWCT
EMERVFGFPVHYTVDVSNMSRLARQRLGRSWSVPVIRHLFAPLKDYIYACV

>Kryptolebias marmoratus DNMT3a

MPSNSPAAAEPPQTLENDASNDVSEDGADQDSPEEGTPASLRNKRVRGRPSRKSQQLPEMATSDPCANAPPTPEEPEPSPRKK
RRHQKTEQTEKNKDEPETDRNCDSPREGETGRLRRRPVPRVTFQAGDPYYISRRQKEEWSRWMKMEARRAYREAEINMDDLSL
ADFQKEEPEASPAPPPSQHTDPASPTVAVTPEPVARGDQAMPSEIECQDGRGFSIGTLVFGKLRGFSWWPGRIVSWWMSGRSRAA
DGTRVHQQGRIMYVGDVRNVTRKHIEEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLHEARPKEGDDRPFFWLFEN
VAMGVSDKRDISRFLECNPV
MIDAKEVSAHRARYFWGNLPGMNRPLTAMCTDRLLDQDCLEHGRTAKFEKLRITITTRSNS
VKQGGKDEHFPVYMDNKEDILWCTEMERVFGFPVHYTVDVSNMSRLARQRLGRSWSVPVIRHLFAPLKDYIYACV

>Kryptolebias marmoratus DNMT3b 803aa

MPFNKYSPAMLEESSNVATAAVNGDTPGAENDSSVELTNDNSPLTAAEPPSPFLSKQNGDAASPQDGNQCSGRSRKRGRKRPEE
EPAWDSYCNKSVGAGQLSLRQRPRRTFYQAGLTTNGHNKPRRQRNQLSVLLYVSGLQGAAGSPVAPETSCPDLMEQDS
KDSAQSSSTMSSSETQPEYSDNKGFGIGELVWGKIKGFSWWPGRIVTWRATGKRQASHGMRWLQWFGDGKFSVVSADKLDLTA
FPKFFNRASYTKLASYRAIFQALEMASVRAEKTFFPCESDPEQVKPMLDWANGGFLPKGEEGLKPAANSANPLDHLQNLVSL
HEYFAGAKRAKPNFCKIAAPEESCSREQMVNEVRKKKQSIIEFCLSCGKMRAATFHPLFEGGLCQTKCDVYLEMSYMYDDDG
QSYCTVCCGGREVLICGNANCCRCFCVDCLDILDPGASNSQMYRQYQDPWRCYMCQPLLQYGVLRKRHDWSMKLQEFFANDGQEF
EKTNIYPAVPAEQRRPIRVLSLFDGIATGYLVRDLGFKVDLYVASEVCEDESISVGVVRHEGKIQYVHDVRNITKNIQEWGPFDLV
GGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLSEAKPKEGEDRPFFWLFENVVAMGVNDKRDISRFLECNPV
MIDAKEVSAHRARYFWGNLPGMNRPLCASGMDKLELQDCLEHGRTAKFEKLRITITTRSNSIKQGGKQDQHPVLMNGKEDVLWCTELERIFGFPVHYT
LSNMGRGARQLLGRSWSVPVIRHLFAPLKDYIYACV

>Kryptolebias marmoratus DNMT3b-like 1441aa

MDCDTEEEPTATKTHEKNKSDVIFLISEDEEEKADQMEIVVKDEPQPHQTAEAAVKTQSAGGDLVTFIKKHCCNNIYSAHPKDLV
HVSNTFPFLYLTVGVPVGGQSTSFVLIGYFDRASDDRVKLVHTLQMNVDTADPHNSAGVTTDELTHAAYSACLLETILNKL
ELPLSNLAFVYCSAPHLESQLLVPKLQTFNPKMISLNCISGMAGRACQALIASFSHVVDLIRDIHHYSTCSPVDDSLKXVFAEAG
SYNPLCPVSAQCLFTHSVQNMMAACWQDLLEYFKLLTHSEATDQIRFQMLDDKVKLHFLFLSHSLEPLRALQDLQSGSASLIVELQ
LVSILLQSYATSILRSSAVDNFLLKRDMLLHSEKELLPTAEVKVGSRASDFLWATSLNRRQQVRNGFLKDAQSFYIAALQSLVKS
I
PGPLVDVOTQRNIGPLLDKPNNTDKSFSWKVLSVICISMGLCPDEASEMQLAQNYHIFIKRLKEKQNSHNSWAKILHRMRPWT
LHLQFLTLALPSSLQRTLVSFDKAYENIKIPHDREETPKASNPDWARRPSIRRRKLPVQKTEEGGESSTEGSDSHVASPRTYNNTK
HCGSVDNTDYTDNSSDVIDITEESIIFPNNEQVANGDEMQAQAEYIFIVTSEENPSLSDSPVSTSGSVESSEKTIGELVWQIEGFSWP
GIIVPYKRLRLPEKRMVWYQGRMCSQVSLQTLQPYSEFDRHFCAKSFATLVTYREAFSSQLQEAALRCKKQFSPLVEDSEDKLKE
MLDWAFGGQFQSPGNGLKPTNGVFKKTRPKVKKLKKKATSFSDPSNPLKNNGNIHTEIKDVAVSLNRLSCKSIYVSSQPKTAVK
KRSKQMEERHQITITAGMRKRQEEWEGEKVKLVRKWNMKSQWGRKRKKEGKGVKKNYWFYGLDNRSPDFVPHKRSY
RAYNKVNQMTSVYTPDQKLRQETINRIMDLNINIEAYCLCCGVEDIEFHPLFKGSLCLKCKNNFTETLFRYDEDGYQSYCTICCY
GMEVILCGNDSCCRSYCEDCLNILAGPGTFDSLKLLDPWICYLQPHQPHGALVPREDWSIRVQGLFANNSGIEFEPHRVYPSIPASL
RRPLRVLSLFDGIGTYLVLKDLGFKVDKYVASEVCEDESIAVSTINHDKIVHVGDRFITQEHINDWGPFDLMIGGSPCNDLSIVNP
LRKGLYEGTGRLLFFEFYRLLQKPKKEEDRPFFWLFENVVFNMTHDKVNICRFFECNPVLVDAVKVSPAHRARYFWGNIPGMSRP
I
IASQTDKLNLDQCYELGREAKVTKVRTITNPNLSKQGGKQDQHPVLMNGKEDVLWCTELERIFGFPVHYT
KAWSVVIRHLFAPLKDYIYACV
EELPPLSTACMPACSSNTSSSPATPDLQQLR

>Danio rerio DNMT1 1500aa

MPTKTSLSLPEVDKRLVLEDEGGDSLSDEECVKEKLRLLQEFLLADTQDQLKNEDEKLSSELSTEYVMSEVKA VLKALGVK
EGDGEVQNGHSNGFSENGSHKDNGEQEGAMDTQEEGDAIKSPSAPKGRGRRSKADSEPKKSPASSRVRTNTGKQQTIVSMFSRV
PKRKSDELNGEPANGDTEIKTEETITEEVREEKRLKTEDEKPEAENANLKPVSTAKTPPKCPDCRQYLDSDLKFFQGDPPDLD
EPEMLTDERLSLFDSDNEDGFESYEDLPQHKITNFSYDKRGHLCPFDGLIEKNVELYFCAVKPIYDDNCPMDGGVPAKLLGPINA
WWITGFDGGEKALIGFTTAFADYILMDPSEEYSAIFALMQUEKIYMSKIVVEFLQKNQDATYEDLLNKIETTVPAGLNFNRFTEDTL
LRHAQFVVEQVESYDEAGSDQPIITPCMRDLIKLAGVTLGKRRRAARRQA VRHPTKIEKDNKPTKATTTKLVLYLIDTFDSDQID
QNNKDDGGVQRQCGVCEVCQAPDCGKCSACKMDKIFGSGRSKQCAQKRRCPNLAVKEAEDDENMDEEDVLVVKDTKKMSQT
K
KKKQTKNKISWVGEPLKTEGKKEYMKVRVENEVLEVGDCVSVSPDDPSHPLYLARITALWDDGEKMFHAHWFCRGTDTVLG
ESSDPLELFLVDECEDEMQLSFIHGKVNVFYKAPSENWYMEGGMDEDIKVIDDDGSEFFYQLHYEGECARFETPPKVTPSEDCKYKF
CASCTRNKERAEASVPHAYEPELEDESISKVYFGLVNYKGEQYKVGDSVYLPPEAFNFVKAASPVKRSRHRKDDVDEDLPEYR
KSSDYKGSNLDAPQFRIGRIKEIFCNKRSNGKPDTEIKLRLYKFRYPENTHKGPKGAYHSDINQLYWSDEEATVSMTEVLTRCV
EYAEADVESVQDYSNKGPRDFYFLEAYNAKTKSFEDPPNHARSANVKGKGGKGGKGGKGAAPQEPQDQEPQEPVVPKLRITLD
V
VFSGCCGLSEGFHQAQISETHWAIEMWDPAAQAFRLNPNPVTFTFEDCNVLLKLVMSGEKNSLQKLPQKGDVEMLCGGPPCQ
GFGSMNRNFSRTYSKFNLSVLSYSCDYRPPKFFLEENVRNFSFKRSMVLLKTLRCLVRMGYQCTFGVLELQAGQYGAQTRR
RAILAAAPGEKLPRIYEPHFVAPRACLSVADEKKYVSNVTRNGNGIYRITVTRDMSDLPEIRNGAAALEISYNGEPQSWFOR
QIRGSQYQILRDHICKDMSALVAARMRNIPLAGSDWRDLNPFVRLRDGTTTKLRYTHPKDKNGRSGTGALRGVCSCEGKPC
DPADRQNTLIPWCLPHTGNRHNHWAGLYGRLEWDGFFSTVTNPEPMGKQGRVLRHPEQHRVSVRECARSQGFPDTRFFGNV
LDKHRQVGNVPPPLSKAIGLEVKKCVLEKMRNATEPVKQEKMELSD

>Danio rerio DNMT3ab 978aa

MPSNTVATAAAAFKSSDFNCDTNGNQMNESMETMIIDEMPSLCTPVKRVGRPGRKRKHLHVTCYGPLNQLMAKQKHEAELQSS
LQSQNGEVERHENKSIIEGRHSTLDRKRVENGLFCQSGCKAESTSDSSLSPLENGFVSSQDSEKCEETLPTLPRKRRGRKLERPT
KYVEHKEDESGDPSKNEGARGRLRGVGVWEVSLRQRPMRPTVTFQAGDPYYISRRTRELLAKWKLEAEKAKLVSMNMSMEDH
GEVESEPKEDYIVPLRFPMPPTPTPPMQQQPPQQAQHTNLKQPSPQPPPPFQQQQQPPQDPAAPTVAATPEPVSIGDGD
K
TSPKSTESYEDVGRFGFNLVWGLKRGYSWWPGRVSVWRMTRRSPAAEGTRWVTFWFGDGKFSVVCVEKLLPLSSFHNAFHQP
TYNKQPMYKKAIEYVLQVASTRAGKTFLSCPASDDTSSKLVNENKMIWAMSGFQPTGPKGLDPPESEPEYEMCEWVEPEAAS
YLPPAKKPRKSTVTEKPKVKEMIDERTRERLVFDRQKCRNIEDICISGSLYVSLHPLFIGGMCQSCNTCFLECAQYQYDDDG
QY
SYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPAAALAAIKEDPWNCYMCGLKTYGILLERRADWPCRLOHFFANNHDQEF

EPRLYPVLAENRKP... GGSPCNDLSIVNPARKGLYEGTGRLLFFEFYRLLHEARPK... RYFWGNLPGMNRPMASAMCTDKLDLQDCLEHGRTAKFGKVRTITTRSNSIKQGGKDQHPFVFMNDKEDILWCTEMERVFGFPVHYT

>Danio rerio DNMT3aa 1010aa

MAKQSPKQEAQEVECEVSKEEKVTDTPTRRVGRPARKRKAPAVNSCEVVTKSSGCVKQSECNGEADSRVAKKSLQRNDSSPESH... QNGADTPPAGEDSAPSTPRKKRGRKLERTERDIEEDTSSDTSRGESDGRHRGRQGWDSLRRRPVQRETFQAGDPYHISRREKEE

>Danio rerio DNMT3bb 816aa

MRKEEIKKSTEIVMPSNKYPSAESDKMTATAAMNRDTSVGDGLSENDSGLEMTSENSPLTPAEPSPFCPKQNGGAASPADESNSI... RRKRSRKRSDTEEDSAWSSYSEEKAEVSGGCETGLRQRPRRTIFQAGLTAHSKPRSRERGHSHKEDHSDLVASVPEGPALELMEQ

>Danio rerio DNMT3b 1448aa

MVADVKIGDDKQSLCELLDWNGLLQATFSQVEDTCSGA AFCQLMDIIPGSDVTKVNFTAENLDILNLYNLLQEAFSKAQIQK... ELETLLVNGDIMTTCDLLTWFKDMYDHNFAKQKCNQVAFIKPEVSLKSSREFETIEKENVSSLYNTEETSSNQKTOHVEKTSQE

>Danio rerio DNMT3ba 1500aa

MATNVSLEPNPDDKCSRYEVLGWINETLQNTFTQVEQCRSGACFCQLIDLFPGTINLKKVKFESQKRSDFMQNYGLLQAAFRDL... EVTEPVPVNELLSGKFRPNFTYKWFKFFYANVKQERVYNAFEARDGQEI VVDDVMKSPKALKSSYESGRAGEESDMEINGGR

>Fundulus heteroclitus DNMT1 1584aa

MPSKTSLSLDDVVRKRLQMLDDDDGSPPEEDQVKEKLLVLDLFLHDDAQDQLTSLEEKMKKSEISKEYISKVKVLLGKELHLENG... SLVDGAEQNGKANGFANGSHKDEEDVEMAAEREEAETKSPAALRGKGGRRSKANSETKKSPTSTRVTRNSGKOPTILSMFSKV

SDYIKGSNLDAPEPFRVGRKEIFYHRRSNGKADTSEVKLRLYKFYRPENTHRGVKASYHTD VNQLYWSDEEVTVMNSEVLGRCC
VEY AEDLNESVHDYSSGGPDRFYFLEAYSAKSKSFEDPPNHARSSVHKGKGGKGGKGGKASQDSQESRTEPQTLKVPKYRTL
DVFSGCGGLESEGFHQAGISETLWAIEMWEPAAQAFRLNPNPPTVTFEDCNVLLKLVMGSEKTNLSGQKLPQKGDVEMLCGGPPCQ
GFSGMNRFNSRYSKFKNSLVVSYLSYCDYYRPFLENNVRNFVSKNSMVLKLTLRCLVRMGYQCTFGVLQAGQYVGAQTRR
RAIILAAAPGKLPYCPSEPLHVFAPRACSLTVAVDEKKYVSNVTRGNNGGIYRTITVRDMSDLPEIRNGAAALEISYNGEPQSWFQR
QIRGTQYQPILRDHICKDMSALVEARMRHIPLAPGSDWRDLNIEVRLKDGTMTKKLR YTHHDKKNRSGSTGALRGVCTCAGGKT
CDSADRQFNTLIPWCLPHTGNRHNNHWAGLYGRLEWDGFFSTVTNPEPMGKQGRVLHPEQHRVSVVRECARSQGFPTDYRFFGN
VLDKHRQVGNVPPPLSRAIGLEIKKCVADRMKEEQATADIKQXXXXXXXXXXXXXXXXXXXXFPFHVLMLE YFLSDFLTFSCHQVA
TMYVVIIPIYCVLMVFIHAVKCSMWSKLYVDFICNISNKVLLKL

>Fundulus heteroclitus DNMT3b 1496aa

MADPLVVIPNASQDKSSRFYVVSWINLKLKTNFKDVREMGSGACQCQMMMDCTFPGSVDLTKVKFDAQSTEDFRHNFSLQDAFK
QKDIKTIPIEELTKGDFQTCYEFMKWFKVFHTANAISKDYDPLKARNNKDITPISISSESGNSDRETNNGTTPTKGFPPYTETWKEDF
DWADRSPLGERHTYCSCHTNLSTYAKGVVELQRHMETNKHKKRAKHLQSSQRKSQLTKPFPCSDAAIRFIYNRCYTGSAKGEQV
PRHFARCKLGLKYPKDLIAVSKHTPYCVVYVEGLTAGVDDTIVTVLLRFFDSESSTHRIRFLDAFAQAADGDGDQTA AAAV VETLKK
FELSTYNLAACYVGEGEASEQIRSRRLRELPNIIAFGGLYAIADAA YHAGVTELSSQVQELIVDLYAHYSSCSAKNDNLKALFGPDIT
KHSSPFYLSISCLRHFHQLITKMLEMWSLDLILYHSSCPEDCKKAKSLCSQLKDPKVKATIMFLEQALKPLQKQKQIQESSWTDMPIL
EASALLGTYTSHFLQNAAXXX
DDFNLGGSAVRDFLQGEAAESVPLQEEVLPFYVA VGTGICAEELPLSDTVLRSIAQVLPNPOSSLTVAKAAVEELGKLGICNTPEE
VRQLKKEYSEYLLAEEGKSKEGVNEDTLETPLARVLKSPGSAFLFRKLVLTLLSFPCPPLGAHHVLSQALDNDCCSETDTTFVESEC
EALSDGVL TNNGSKNLQVLANGLNGLFLKPCVRLTKINDLCDIVNGLSNEGSRWGGHGWESSLRKPPARKVVFQAGVCTWSK
PIGKNDPELNEVLEGMSPSRKGTSPRARHSQYQDGKAPV GELVWGKLSGMSMWPGLVMDWKS KHAPPGVRRVEWFGDMFS
EISTERLKPFSFSLCFCKNSFASLPVYKEAIYHICELAAERCGTFDKAKGDKEKELKLMVEWASEGFLPTGPRGFAVPDPSSNQDS
NDSASPDYQPPAKRKHAKLNREGGHATTHNRDSLKEKVKKEGKTIEHFCLSCASPEIEVQHPLFEGGLCHKCKENFTETLYRYDED
GYQSYCTICCGGLEVLICENSSCCRCFCCKDLEILVAEGTFEKLKLVDPWICFMCEPSKCDGHLKVRPDSVVKVQDLFANNNAMEF
ESHRVYPSIPADRRKPLKVLSLFDGIATGYLVLRLDGFKIERYIAEICDDSI VGMVKHEGKIEYVNDVRTITRKHAEWGPFDLLI
GGSPCNLDMVNPLRKGLYEGTGRLEFFFYRILTLRLPKEEDDRPFDFKAGFENVVFMAHDKNDICRFLCNPVLVDAIKVSPAHR
RYFVGNIPGMSRPLSTSLDDKVTLDQCLEAGRMKFEKVRTITTKSNSIRQGGKTGPLVLMNGKEDYLWCTEMEQIFGFPKHYTD
VNNMGRMQRQKVLGRSWSVPVIRHLLAPLKDYFECE

>Fundulus heteroclitus DNMT3a 989aa

MPSNTVTNTNSTSCKGSDFADSPGHRMGESTILGTDLMEEQSAFSSASFKAGKPRKRKQFPAESCGSITDCAGVCNSYPGLCRPT
VSQVHNGDVDAHGDKTPDAACLKQERREVENGI PRCLSSSSSYQEEEESSQGRSPSPCQENGFP SERGPESDKDDGLMMPRKRKRG
RRKLERPTKYVEHKEDDGC DTIKTEGGRGRLRGGVGVWEISLRQRPMRITFQAGDPYIYISKRTEELMAKWKMEAEKKAKLMSEI
NIMKDFEKNETETQKEEVSIIKHPQPSKQPPQPEPLSQVYQQQSQPSLPQQQQKQPPQQQQQQQQQQQPTDPASPPVATTPPEV
AIGDEEKSSPKSGDTEPEYEDGRGFGIGELVWGKLRGFSWVWGRIVSWVVTGRSRAAEGTRVWVWFGDGKFSVVCVCKLLPLSS
FNNAFHQPTYNKQPMYKAIYEVLVQVASSRAGKAFIACPDSESETSKSVDQLNKQMI EWAMAGFQPTGPKGLEPAEEERNSEVY
PETWVEPEAAA YTPPAKPRKSTVEKPKVKDIIDEGTRERLVYEVROKCRSLEDICISCGSLNVSLEHPLFAGGMCQSCKNCFLEC
AYQYDDDGYSYCTICCGGREVLMCGNNNCCRCFCVECVLLVKGQAAVA AIKEDPWNCFCMCCPKGVFGLLERRADWP SRLQT
FFANNHDQDFDPKLYAPVLAEKRPVRLSLFDGIATGLVLLKELGIQVDRYVASEVCEDSITVIVRQVGMVGDVVRNVT RSK
HIQEWGPFDLVIGGSPCNL SIVN PARKGLYEGTGRLEFFEFYRLLHEARPKGEGDNRPFVLFENVAMGVSDKRDISRFLCNPVMI
DAKEYSAAHRARYFVGNLPGMNRPLSAMCSDKLDLQDCLHGR TAFEKVRTITTRSNSIKQGGKQDHPVYMNKEDILWCTEM
ERVFGFPVHYTDVSNMSRLARQLLGRSWSVPVIRHLFAPLKDYFACV

>Fundulus heteroclitus DNMT3a-like

MPSNSPAAAEPPQTPENDGAKNLGEDGADQDSPEEGTPASPHNKRVRGRPSRKRKQLLEMAASDPCANAPPTPEEPEPSPSSQKK
RRRRKLEQSERNKDETDERNCDSPREIETGRLRRRPVPRVTFQAGDPFYISRRQKEEWLSRWKMEAEARRAYREAEINMDDLSEN
GFQKEEPPASPAPPSQQHTDPASPTVA VTPDPVARGDQAMPNEIEYQDGRGFGIGTLVFGKLRGFSWVWGRIVSWVWMSGRSRAA
DGTRVWVWFGDGKFSVVCVEKLMPLSSFSFAHQPTYNKQSMYRKAIFEALQVAVRAGRVPACDASDEAEGLVQTRQMI EW
AMTGFPLTGPQSLDPPEEEQNQFKEGYLEMLPEPEAA YTPPAKPRKNTAEKAKIREVIDEGTRERL IHEIKKTRNIEDICISCGSL
NVSLEHPLFMGAMCQCGKNSFLECA YQYDDDGYSYCTICCGGREVLMCGNNNCCRCFCVECVLLV GAGSAAA AIKEDPWNC
YMCGRSTYGLRRRDDWPRLQHFFANNHEQEFEPAKLYPPVA AEKRPVRVLSLFDGIATGLVLLKDLGIQVDKYVASEVCEDS
ITVGMVRHQGRIMYVGDVNRVTRKHIDEWGPFDLVIGGSPCNL SIVN PARKGLYEGTGRLEFFEFYRLLHEARPKGDERPFFWLF
ENVVAMGVSDKRDISRFLCNPVMIDAKEVSAHRARYFVGNLPGMNRPLTPMANDKLDLQECLEHGR TAPRSICGSRASCPC
ALSWSRQGTGAPAEEDGVGIHKDRRIMAA NMGSMRILIKGGKVVND DFTQEADVYIENGIIQQVGKELMIPGGAKVIDASGKLV
LPGGIDTCVHLQDTFMNASIQDDFYSGTKAALMGTTVMALVLPQHC SLVDAY EKCRALADAKACCDYALHVGVTVWVWGP
VHNEMESLVREHGVSFQMFMA YKDMMLLRDSEL YQSLQCTCKDIGAIARVHAENGELVAEGAKEALDLGISGPEGIEISRPEELES
EATHRAITIANRARCPIYLVNVSMLAGDMIAAAKMQGKVVAHAE TTVAHAVLNGNYYHHWAAHAAHVIVPPLRLDPNTPSYL
MGLLGNDTLSVVA SEHRPFSTKQRALGKEDFSKIPHG VAGVQDRMSVIW ERGVTGKMDENRFVA VTSNSAAKIYNLYPRKGRII
PGADADV VVWDPDAARTISVSTQVQGGDFNL YEGMRCHGVPLVTISRGRVLCENG VFMCAEGSGKFYQRTFPDYLYK KMVQRE
KNQAYKGVVERDPYSGDVAKVANTMKKELGLPIDGDSPLKPPVRVHQGV RDLHESSFSLSGSQVDDHIPKRSSARILAPPGRRSSG
IW