

Université de Montréal

Génération de lignées de poissons-zèbres par génie génétique dans le cadre de l'étude du gène
C9orf72

par Alexandre Emond

Département de neurosciences
Faculté de médecine

Mémoire présentée
en vue de l'obtention du grade de maîtrise
en neurosciences

Janvier 2019

© Alexandre Emond, 2019

Résumé

Une mutation intronique du gène *C9orf72* est une cause prépondérante de la sclérose latérale amyotrophique (SLA), une maladie neurodégénérative des motoneurons de l'adulte. Cependant, la fonction de *C9orf72* demeure incertaine. Heureusement, les modèles animaux sont une excellente façon d'en faire l'étude *in vivo* dans des systèmes complexes. Conséquemment, pour étudier la fonction et l'expression de *C9orf72*, nous avons tenté de développer 3 lignées de poissons-zèbres pour lesquelles des étiquettes 12xhistidines ou eGFP devaient être insérés par CRISPR/Cas9 près du N et C terminal de *C9orf72*. Surprenamment, ces modifications furent mal tolérées durant le développement, empêchant l'obtention de sujets F0 adultes en mesure de transmettre un tag inséré. Pour étudier l'effet d'une invalidation persistante de *C9orf72* chez le poisson, une lignée présentant une délétion près du N-terminal menant à un knock-out fonctionnel de *C9orf72* a été développée des poissons injectés pour l'insertion de 12xHistidine. Nos résultats préliminaires suggèrent un phénotype au niveau de la mobilité et de la mortalité. De plus, deux lignées de poissons transgéniques exprimant le gène *C9orf72* humain, seul ou marqué au eGFP, ont été développées par Tol2 pour étudier la localisation de *C9orf72* et l'effet de sa surexpression. Ces lignées, une fois validées et caractérisées, permettront d'ouvrir de nouvelles perspectives pour caractériser la fonction et l'expression de *C9orf72* ainsi que la modélisation de la SLA. Il semble donc que la perte d'un allèle ou une haplo-insuffisance du gène *C9orf72* est mal tolérée, mais que la transgenèse, par addition d'un troisième allèle, permettrait l'étude de *C9orf72* *in vivo*.

Mots-clés : Sclérose latérale amyotrophique, SLA, Poisson-zèbre, knock-out, transgénique, CRISPR/Cas9, Tol2, *C9orf72*

Abstract

An intronic mutation in the gene *C9orf72* is the most prevailing cause of Amyotrophic lateral sclerosis (ALS), an adult onset neurodegenerative disease of the motoneurons. Despite this, the function of *C9orf72* remains uncertain. As animal models are an excellent way to study gene function and expression *in vivo* in complex systems, by using the CRISPR/Cas9 system we attempted to generate 3 lines of zebrafishes for which 12xhistidines or eGFP tags were inserted near either the N or C terminal of *C9orf72*. Surprisingly, these modifications were ill tolerated during the development, preventing us from obtaining F0 specimens able to pass on the inserted tag for any of our lines. To study the effects of a persistent knock-out of *C9orf72* in the zebrafish, a line with deletions near the N-terminal, resulting in a functional knock-out of *C9orf72*, was obtained from the fishes injected for the 12xHistidine insertion. Preliminary results indicate a higher mortality rate and decrease mobility phenotypes. Additionally, two transgenic zebrafish lines expressing human *C9orf72* alone or with eGFP were generated, using Tol2 in order to study *C9orf72* localisation and the effect of its overexpression namely on morphology, development and the central nervous system. These lines, once validated and characterised will open new avenues to determine the role and subcellular localization of *C9orf72*, as well as for modeling ALS. Thus, it appears that the loss of an allele or haploinsufficiency of the *C9orf72* gene is ill tolerated, but that transgenesis, by addition of a third allele, may allow for *in vivo* study of *C9orf72*.

Keywords: Amyotrophic lateral sclerosis, ALS, Zebrafish, knockout, CRISPR/Cas9, Tol2, *C9orf72*.

Table des matières

Résumé.....	i
Abstract.....	ii
Table des matières.....	iii
Liste des tableaux.....	viii
Liste des figures.....	ix
Liste des sigles.....	xi
Liste des abréviations.....	xv
Remerciements.....	xvi
1. Introduction.....	1
1.1 Sclérose latérale amyotrophique.....	2
1.1.1 Description.....	2
1.1.2 Contexte historique.....	2
1.1.3 Diagnostic et Épidémiologie.....	4
1.1.4 Histoire naturelle.....	5
1.1.5 Traitements.....	6
1.1.6 Étiologie.....	7
1.2. <i>C9orf72</i>	11
1.2.1 Structure du gène <i>C9orf72</i>	11
1.2.2 Conservation de <i>C9orf72</i>	11
1.2.3 Épissage alternatif de <i>C9orf72</i>	11
1.2.4 Structure de la protéine <i>C9orf72</i>	12
1.2.5 Profil d'expression.....	14
1.2.6 Fonction de la protéine <i>C9ORF72</i>	15
1.2.6.1 Régulation trafic vésiculaire.....	17
1.2.6.2 Réponse immunitaire.....	18
1.2.6.3 Signalisation de mTOR et système lysosomal.....	18
1.2.6.4 Élimination d'agrégats de protéine par macroautophagie.....	19
1.2.6.5 Régulations du cytosquelette de l'actine.....	19
1.2.6.6 Régulation de l'expression génique.....	19

1.2.7	<i>C9orf72</i> dans la SLA	20
1.2.7.1	Mutation de <i>C9orf72</i>	20
1.2.7.2	Mécanismes pathologiques des ERH de <i>C9orf72</i>	21
1.2.7.2.1	Mécanismes de gain de fonction toxique	21
1.2.7.2.2	Mécanisme de perte de fonction	26
1.3	Le poisson-zèbre	28
1.3.1	Avantages et utilisation en recherche	28
1.3.2	<i>C9orf72</i> chez le poisson-zèbre	31
1.3.3	Transgénèse stable chez le poisson-zèbre	34
1.4	Objectifs	38
2.	Résultats	39
2.1	Observation du patron de l'expression de <i>zfc9</i> par hybridation <i>in situ</i>	40
2.1.1	Sondes et type de marquage pour l'hybridation <i>in situ</i>	40
2.1.2	Stades des embryons fixés et montés pour l'hybridation <i>in situ</i>	41
2.1.3	Profil d'expression du gène <i>zfc9</i> à 24, 36 et 48 hpf	42
2.2	Tentative de génération de lignées avec un tag dans <i>zfc9</i>	43
2.2.1	Constructions génétiques et micro-injections	45
2.2.2	Confirmation de l'insertion des marqueurs moléculaires	48
2.2.3	Sélection des poissons fondateurs F0	50
2.2.4	Élevage des poissons injectés F0	51
2.2.5	Criblage pour l'identification de poisson fondateurs F0	51
2.3	Génération d'une lignée <i>zfc9</i> knock-out stable	54
2.3.1	Criblage pour la transmission d'indels dans <i>zfc9</i>	55
2.3.2	croisement et élevage de poissons candidats pour la F1	57
2.3.3	Criblage et identification de poissons candidats pour la F1	58
2.3.4	Génotypage des spécimens portant des indels et caractérisation de celles-ci	59
2.3.5	Caractérisation de l'effet des délétions identifiées sur la protéine <i>zfc9</i>	62
2.3.6	Croisement pour l'obtention de la F2	63
2.3.7	Crible des spécimens de la F2	63
2.4	Résultats préliminaires pour la caractérisation de l'effet du KO de <i>zfc9</i>	64
2.4.1	Test statistique des proportions de génotypes pour la progéniture de la F1	64

2.4.2. Comparaison de l'activité de nage des WT, HT et HM pour le KO de <i>zfc9</i>	66
2.5 Génération de lignées transgéniques exprimant le gène <i>HC9</i>	68
2.5.1 Construction génétiques et micro-injections.....	69
2.5.2 Sélection des poissons fondateurs F0	71
2.5.3 Crible des spécimens F0 pour la transmission de chaque transgène	72
2.5.4 État de la lignée HC9tg.....	73
2.5.5 État de la lignée HC9-eGFPTg.....	74
3. Discussion.....	78
3.1 Profil d'expression du gène <i>zfc9</i> aux stades 24, 36 et 48hpf.....	79
3.1.1 Complications rencontrées et améliorations pour de futures expériences d'HIS	80
3.1.2 Perspectives futures pour la caractérisation du profil d'expression de <i>zfc9</i>	81
3.2 Échec de la génération de lignées avec un tag moléculaire dans <i>zfc9</i>	82
3.2.1 Efficacité d'insertion avec le système CRISPR/Cas9 chez le poisson-zèbre	83
3.2.2 Nombre limité de candidats F0 criblés	84
3.2.3 Mortalité importante des larves F0 sélectionnées	86
3.2.4 Perspectives pour le marquage de <i>zfc9</i>	88
3.3 Génération d'une lignée knockout pour le gène <i>zfc9</i>	89
3.3.1 Efficacité d'intégration des indels dans les cellules germinales.....	90
3.3.2 Nature des indels de la lignée <i>zfc9</i> ^{KO}	90
3.3.3 Caractérisation préliminaire des spécimens KO pour <i>zfc9</i>	93
3.3.3.1 Mortalité accrue des spécimens HT et HM KO pour <i>zfc9</i>	94
3.3.3.2 Déficiences dans l'activité de nage des HT et HM KO pour <i>zfc9</i>	95
3.3.4 Perspectives futures pour une lignée <i>zfc9</i> ^{KO}	96
3.4 Génération de lignées transgéniques exprimant le gène <i>HC9</i> avec ou sans eGFP.	98
3.4.1 Efficacité d'intégration des transgènes dans la lignée germinale	99
3.4.2 Possible instabilité et dégradation de la protéine de fusion HC9-eGFP	99
3.4.3 Perspectives futures pour les lignées transgéniques HC9tg et HC9-eGFPTg.....	102
4. Conclusion	103
5. Matériel et méthode	105
5.1 Matériel et méthodes communs à plusieurs expériences	106
5.1.1 Maintien des poissons zébrés.....	106

5.1.2 Préparation de milieux de culture Luria Bertani (LB) et de plaques LB agar	107
5.1.3 Gels d'agarose et migration par électrophorèse.....	107
5.1.4 Micro-injections.....	108
5.1.5 Réaction par amplification en chaîne (PCR)	109
5.1.7 Coupure de nageoires caudales des poisson-zèbres.....	111
5.1.8 Analyse HRM et amorces utilisées	111
5.1.9 Séquençage	112
5.2 Matériel et méthodes spécifiques à une expérience.....	113
5.2.1 Extraction d'ARN total et synthèse de l'ADN codant.....	113
5.2.2 Synthèse et clonage des sondes pour l'hybridation <i>in situ</i>	116
5.2.3 Retrait du chorion et fixation des embryons pour l'HIS.....	119
5.2.4 Hybridation <i>in situ</i>	119
5.2.5 Synthèse des ARNg pour CRISPR/Cas9	124
5.2.6 Test de l'efficacité de coupure des ARNg.....	127
5.2.7 Synthèse des matrices d'ADN pour CRISPR/Cas9	128
5.2.8 Micro-injections pour les lignées C9orf72 ^{NHis} , C9orf72 ^{NGFP} et C9orf72 ^{CGFP}	129
5.2.9 Sélection des poissons fondateurs F0 des lignées C9orf72 ^{NGFP} et C9orf72 ^{CGFP}	130
5.2.10 Sélection de la progéniture des F0 des lignées C9orf72 ^{NGFP} et C9orf72 ^{CGFP}	130
5.2.11 Sélection de la progéniture de la lignées C9orf72 ^{N12H}	130
5.2.11.1 C9orf72 ^{N12H} knockin.....	131
C9orf72 ^{N12H} knock-out	131
5.2.12 Analyse statistique	131
5.2.13 Tests de nage.....	132
5.2.14 Constructions géniques Tol2 des lignées UAS-HC9-eGFP et UAS-HC9.....	132
5.2.15 Injection Tol2 pour les lignées HC9-eGFPtg et HC9tg.....	133
5.2.16 Sélection des poissons fondateurs F0 des lignées HC9tg et HC9-eGFPtg	134
5.2.17 Sélection de la progéniture des F0 des lignées HC9tg et HC9-eGFPtg.....	134
5.2.18 Sélection de la progéniture des lignées HC9tg et HC9-eGFPtg	134
5.2.19 Visualisation de l'expression de HC9-eGFP en présence de Gal4 endogène.....	135
5.2.20 Confirmation de l'expression de Gal4 endogène.....	135
5.2.21 Injection Gal4 pour tester l'expression de HC9-eGFP	135

5.2.22 Visualisation de l'expression de HC9-eGFP en présence de Gal4 exogène.....	136
Bibliographie.....	138

Liste des tableaux

Tableau I. Sondes marquées à la DIG générées pour la détection de l'ARNm de <i>zfc9</i> par hybridation <i>in situ</i>	41
Tableau II. Résultats du crible des poissons candidats F0 adulte des lignées <i>C9orf72^{NHis}</i> , <i>C9orf72^{NGFP}</i> et <i>C9orf72^{CGFP}</i>	54
Tableau III. Résultats du criblage des poissons F0 <i>C9orf72^{NHis}</i> pour la transmission d'indels en N-terminal de <i>C9orf72</i>	56
Tableau IV. Résultats du criblage des poissons F1 pour l'intégration d'indels en N-terminal de <i>C9orf72</i>	58
Tableau V. Résultats du génotypage des poissons F2 de la lignée <i>zfc9^{KO}</i> et proportion des génotypes.....	64
Tableau VI. Résultats du test χ^2 d'adéquation permettant de tester si les proportions des génotypes identifiés est comparable à ceux prédits par la théorie	65
Tableau VII. Résultats du criblage des poissons F0 des lignées transgéniques pour la transmission du transgène.....	73
Tableau VIII. État de la lignée HC9tg au terme du projet de maîtrise	74
Tableau IX. État de la lignée HC9-eGFPTg au terme du projet de maîtrise.....	75

Liste des figures

Figure 1. Schéma de la séquence ADN du gène <i>C9orf72</i> et des trois transcrits alternatifs produisant les 2 isoformes de protéine.....	12
Figure 2. Structure de la protéine <i>C9orf72</i>	14
Figure 3. Alignement des séquences d'acides aminés des protéines <i>HC9orf72</i> et <i>zfC9</i>	31
Figure 4. Structure de la protéine <i>zfC9</i>	32
Figure 5. Schéma de la séquence ADN du gène <i>zfC9</i> et des trois transcrits alternatifs produisant 3 isoformes de la protéine.....	34
Figure 6. Observation du profil d'expression du gène <i>zfC9</i> par hybridation <i>in situ</i> chez des embryons de stade 24, 36 et 48 hpf.....	43
Figure 7. Test du taux d'efficacité d'un ARNg pour l'induction de coupure par le système CRISPR/Cas9	46
Figure 8. Lieux de coupure/insertion par CRISPR/Cas9 prévues dans la séquence de <i>zfC9</i> avec pour les ARNg <i>C9orf72Nt</i> et <i>C9orf72Ct</i>	47
Figure 9. Schéma des protéines <i>zfC9</i> de fusion résultante de l'insertion réussie dans des tags pour les lignées <i>C9orf72^{His}</i> , <i>C9orf72^{NGFP}</i> et <i>C9orf72^{CGFP}</i>	48
Figure 10. Confirmation de la faisabilité de l'insertion par CRISPR/Cas9 des étiquettes moléculaires pour l'obtention des lignées <i>C9orf72^{NGFP}</i> et <i>C9orf72^{CGFP}</i>	49
Figure 11. Confirmation de la faisabilité de l'insertion par CRISPR/Cas9 des étiquettes moléculaires pour l'obtention la lignée <i>C9orf72^{NHis}</i>	50

Figure 12. Identification de la présence/transmission d'indels par analyse HRM et confirmation par séquençage de la région ciblée	57
Figure 13. Profils HRM commun des croisement E4a/E4b et E4b/E4c et caractérisations de leurs indels communs par séquençage	61
Figure 14. Schéma de l'effets des délétions identifiées chez les poissons E4 F1 sur la séquence d'acides aminés de la protéine zfc9	62
Figure 15. Comparaison de la distance de nage totale des spécimens WT, HT et HM pour le KO de <i>zfc9</i>	68
Figure 16. Constructions génétiques pour la génération des lignées HC9tg et HC9-eGFP et expression constitutive du GFP démontrant leur insertion.	71
Figure 17. Détection de fluorescence dans des embryons de la lignée HC9-eGFPtg injectés avec une solution de Gal4	77

Liste des sigles

Liste des sigles et la liste des abréviations

♀ : Femelle

♂ : Mâle

α : Alpha

β : Béta

ADAR : Double-stranded RNA-specific adenosine deaminase (Désaminase de l'adénosine spécifique à l'ARN double brin)

ADN : Acide désoxyribonucléique

ADNc : ADN codant

ALSFRS-R : Amyotrophic lateral sclerosis functional rating scale-revised

ALYREF : Aly/REF export factor (Facteur d'exportation Aly/REF)

ARF6 : ADP-ribosylation factor 6 (Facteur de ribosylation d'ADP 6)

ARN : Acide ribonucléique

ARNm : ARN messenger

ARNg : ARN de guidage

BAG3 : BCL2 Associated Athanogene 3 (athanogène 3 associé à BCL2)

BCL2 : B-cell lymphoma 2 (Cellule B de lymphome 2)

C9 : C9orf72

C9L : Isoforme longue du C9 humain

C9orf72 : Chromosome 9 open reading frame 72

C13H9orf72 : Orthologue du poisson-zèbre du gène C9orf72 situé sur le chromosome 13

CAB : Chromosome bactérien artificiel

Cas9 : CRISPR-associated protein-9 (Protéine 9 associées à CRISPR)

CBD : Coupure double brin

cDENN : Région central du domaine DENN

CLEAR : Réseau coordonné d'expression et de régulation lysosomal

cmlc2 : cardiac myosin light chain 2

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats (Courtes répétitions palindromiques groupées et régulièrement espacées)

CSPi : cellules souches pluripotentes induites
DFT : Démence fronto-temporale
DENN : Domaine différenciellement exprimé dans les cellules normales et néoplasiques
dDENN : Région la plus en aval du domaine DENN
EDN : endothélines
eGFP : Enhanced green fluorescent protein
ERH : Expansions de répétitions hexanucléotidique
F0 : Génération fondatrice d'une lignée injectée
F1 : Seconde génération d'une lignée injectée
F2 : Troisième génération d'une lignée injectée
FDA : Food and Drug Administration
FUS: Fused in sarcoma
GA: Glycine-alanine
GDP: Guanosine diphosphate
GEF : Guanine exchange factor (Facteur d'échange guanine)
GFP : Green fluorescent protein
GP : glycine-proline
GR : Glycine-arginine
GS : Granules de stress
GTP : Guanosine triphosphate
HC9 : C9 humain
HIS : Hybridation *in situ*
HM : Homozygote
hpf : heures post fertilisation
HRM : High resolution melt (Courbes de fusion à haute résolution)
Hsc70 : Heat shock cognate 71 kDa protein
Hsp90 : Heat shock protein 90 (Protéine de choc thermique 90)
HT : Hétérozygote
JENH : Jonction d'extrémités non homologues
KI : Knock-in
KO : Knock-out

LMK1 : LIM Domain Kinase 1 (Kinase de domaine LIM1)
LMK2 : LIM Domain Kinase 2 (Kinase de domaine LIM1)
LB : Luria Bertani
M : Mole par litre
MAP : Mitogen-activated protein (Protéine kinase activée par un mitogène)
MI : Motoneurones inférieurs
mM : Millimole par litre
MS : Motoneurones supérieurs
mTORC1: Mammalian target of rapamycin complex 1
NF- κ B : Nuclear factor-kappa B (Facteur nucléaire kappa B)
OMA : Oligonucléotides morpholinos antisens
ODNsb : Oligodésoxyribonucléotide simple brin
PA : Proline-alanine
PAdS : Personne-année de suivi
PCR : Polymerase chain reaction (Réaction en chaîne par polymérase)
PLA : Protéines liant l'ARN (RBP en anglais)
PR : Proline-arginine
PRD : Protéines de répétitions di-peptidiques
RAC1 : Ras-related C3 botulinum toxin substrate 1 (Substrat de toxine botulique parent aux Ras C3)
RANT : Traduction indépendante de l'AUG associée aux répétitions
RDH : Réparation dirigée par homologie
r.p.m. : Rotations par minutes
siARN : Petit ARN interférents
SIR : Séquences inversées répétées
SLA : Sclérose latérale amyotrophique
SLAf : SLA familiale
SLAs : SLA sporadique
SMCT8 : Smith-Magenis syndrome chromosome region, candidate 8
SNARE : Soluble N-éthylmaleimide-sensitive-factor Attachment protein REceptor
SNC : Système nerveux central

SOD1 : Cu/Zn Superoxide dismutase 1

SOC : Super Optimal broth with Catabolite repression (Milieu super optimal avec repression des catabolites)

SRSF : Serine/arginine-rich protein-specific kinase (Kinase spécifique au protéines riche en sérine/arginine)

TALENs : Nucléases effectrices analogues à un activateur de transcription

TARDBP : Transactive response DNA-binding protein, gène humain

TBK1 : TANK-Binding Kinase1

TDP-43 : TAR DNA binding protein, protéine humaine

TFEB : Transcription factor EB

uDENN : Région la plus en amont du domaine DENN

UTR : Untranslated transcribed region (Région transcrite non traduite)

mm : millimètre

µm : Micromètre

WDR41 : WD-40 repeat 41

WT : Wild type

zfC9 : C9 orf72 du poisson-zèbre

zfC9La : Isoforme longue de 462 aa de zfC9

zfC9Lb : Isoforme longue de 460 aa de zfC9

zfC9C : Isoforme courte de 326 aa de zfC9

ZFNs : nucléases à doigt de zinc

Liste des abréviations

aa : Acides aminés

Ct : C-terminal

DIG : dioxygénine

G4C2 : 4 guanines 2 cystéines

h : heure

kDa : Kilodalton

min : Minute, aussi abrégé m

Nt : N-terminal

nt : nucléotide

OPTN : Optineurin

pb : Paires de bases

p.ex. : Par Exemple

sec : seconde, aussi abrégé s

SQSTM1 : Sequestosome 1

tg : Transgénique

Remerciements

Je tiens tout d'abord et en particulier à remercier mon directeur de recherche Pierre Drapeau. Non seulement pour m'avoir accueilli à bras ouverts dans son laboratoire et d'avoir placé sa confiance en mes aptitudes, en me confiant en partie la réalisation de ce qui était encore un tout nouveau projet pour le laboratoire, mais aussi pour son appui constant à tous les niveaux au cours du cheminement de ma maîtrise dans les moments aisés comme les plus difficiles. J'éprouve aussi une grande reconnaissance pour les nombreuses opportunités scientifiques, d'apprentissage, de réseautage et d'expériences personnelles auxquelles il m'a généreusement donné accès au cours de ma maîtrise, me permettant même de prendre part à des congrès scientifiques internationaux. Ainsi, je dois en bonne partie ce que j'ai pu accomplir au cours de cette maîtrise ainsi que les progrès que j'ai effectués comme étudiant gradué au cours de celle-ci à Pierre et sa guidance comme directeur de recherche, lequel m'a poussé à me dépasser. J'espère simplement avoir été à la hauteur et que tu aies apprécié de m'avoir comme étudiant à la maîtrise autant que j'ai apprécié de t'avoir comme directeur de recherche. Je voudrais aussi remercier mon collègue de laboratoire Abdelhamid Bekri qui m'a aidé à maintes reprises durant cette maîtrise avec ses conseils scientifiques judicieux, son assistance technique exemplaire et son habilité à me pousser constamment à être un meilleur expérimentateur avec ses questions judicieuses à l'égard de mes raisonnements ainsi que mes expériences en laboratoire. Je souhaite que nos discussions aient été aussi intéressantes et utiles pour toi qu'elles l'ont été pour moi. Je ne peux oublier de remercier aussi Meijiang Liao dont l'expertise phénoménale en biologie moléculaire et l'aptitude de gestion des réactifs du laboratoire est l'une des principales raisons pourquoi les expériences effectuées dans le laboratoire Drapeau, dont les miennes, parviennent à aboutir éventuellement. Sincèrement merci de ta patience et de ton oreille attentive Meijiang. Je me dois aussi de remercier les divers membres du laboratoire Drapeau dont Amrutha Awaminathan, Éric Samarut, Marina Drits, Leonid Pospelov, Uday Kundap, Poulomee Bose et Raphaëlle Riche qui se sont montrés fort accueillants ainsi que sympathiques durant ma maîtrise et qui pour plusieurs d'entre eux ont apporté une aide certaine durant celle-ci. Je les remercie aussi pour les discussions intéressantes, les distractions souvent bienvenues et les moments mémorables qu'ils m'ont permis d'avoir durant les 2 dernières années et ce fut définitivement un plaisir de faire ma maîtrise en si bonne compagnie.

Il m'est aussi nécessaire de remercier le Dr Gary Armstrong. Je te suis redevable pour les discussions rares, mais souvent intéressantes et des plus utiles à mon projet, que nous avons eues lorsque l'on se croisait au laboratoire en soirée ou la fin de semaine, et en tant que membre de mon comité de maîtrise. J'espère que nous aurons la chance de discuter à nouveau à l'avenir et que mon mémoire s'avère à la hauteur de tes attentes. Merci aussi au Dr Alex Parker, aussi un membre de mon comité de maîtrise, avec qui ce fut aussi un plaisir de discuter de mon projet de maîtrise et dont les conseils m'ont permis à plusieurs reprises de voir clair dans des situations compliquées concernant mon projet ; c'est un honneur de travailler avec toi dans un futur proche. Je tâcherai de me montrer digne de ta confiance. Je tiens aussi à remercier Susie Daigle du département de neuroscience dont les prouesses administratives et la patience incommensurable m'ont permis d'éviter bien des ennuis.

Enfin, je voudrais remercier ma famille et mes amis pour m'avoir toujours soutenu, même si je n'ai pu passer des moments avec eux aussi souvent que je l'aurais souhaité pour eux. Sans vous je ne pense pas que j'aurais eu le courage de me rendre jusqu'au bout. Papa sans ton appui inconditionnel je n'aurais jamais pu y arriver, je ne peux te dire assez merci.

1.Introduction

1.1 Sclérose latérale amyotrophique

1.1.1 Description

La sclérose latérale amyotrophique (SLA) est une maladie neurodégénérative rare caractérisée par une dégénérescence des motoneurones corticaux et spinaux, mais pouvant affecter d'autres types neuronaux [1]. Les fonctions sensorielles et autonomes sont typiquement épargnées, mais une minorité des patients présentent des déficiences modérées de ces systèmes [2,3] et près de 50% des cas démontrent des troubles cognitifs, du comportement ou des formes de démences [4, 5, 6, 7]. Typiquement, la SLA se manifeste entre 40-70 ans, mais elle peut être plus précoce ou tardive [8]. La pathologie se manifeste par une paralysie générale progressive des muscles squelettiques conduisant au décès par suffocation typiquement 3 à 5 ans après diagnostic. Seul des traitements palliatifs sont présentement disponibles et aucun traitement curatif n'a encore été établi. À ce jour, la SLA est inévitablement fatale [1].

1.1.2 Contexte historique

Durant la seconde moitié du 19^{ième}, une approche novatrice clinico-pathologique pour la description ainsi que la classification des maladies fait son apparition quand des médecins européens se mettent, alors qu'ils étudient des pathologies du système moteur, à lier les signes cliniques aux observations faites durant l'autopsie. Cette approche révolutionne la neurologie et les critères de classification des maladies au point que ces principes sont encore employés aujourd'hui. Tout changement de paradigme s'accompagne cependant de nouveaux défis et l'un de ceux-ci s'avèrera être la séparation des pathologies causées par une déficience du système nerveux moteur de celles dues à une déficience du système musculaire. Un second dilemme se révélera être l'incertitude dans l'attribution de certains syndromes à des variantes de mêmes maladies ou à une pathologie distincte. C'est en cherchant à résoudre ces défis que plusieurs maladies, de nos jours rassemblés sous le terme collectif de maladies des motoneurones, ont été décrites initialement. Notamment, l'atrophie musculaire progressive, la paralysie bulbaire

progressive, la sclérose latérale primaire et ce qui sera éventuellement décrit formellement pour la première fois par l'illustre neurologue français Jean-Martin Charcot comme étant la SLA [9].

C'est en 1869 que Charcot publie ses observations, effectuées aux côtés d'Alix Joffroy, où il y décrit des cas de pathologie présentant une atteinte simultanée des motoneurones supérieurs et inférieurs, sans défaillance sensorielle. Ainsi, la SLA est établie comme une maladie neurologique spécifique associée avec une manifestation distincte [9]. Ce ne sera cependant qu'en 1874 que Charcot complète puis publie ses observations et propose originellement le terme sclérose latérale amyotrophique [10]. Dans cet ouvrage, il donne aussi une description détaillée de l'histoire naturelle de la SLA qui, malgré de nombreuses découvertes moléculaires et génétiques depuis, demeure dans l'ensemble encore d'actualité. C'est en 1941 avec le décès due à la SLA du célèbre joueur de baseball des Yankees de New York Henri Louis Gehrig, surnommé Lou Gehrig, que la SLA devient une cause célèbre et fait son entrée dans l'imaginaire populaire [11]. Encore aujourd'hui, en Amérique du Nord la SLA est couramment désignée comme la maladie de Lou Gehrig dans le langage familial. Durant les années 50-60 la SLA est associée à l'aphasie, la maladie de Pick [12,13] et le syndrome de Guam, une démence parkinsonienne [14]. Ainsi, outre son entrée dans le discours populaire et l'association de symptômes cognitifs à la SLA, on voit peu de nouveautés pour la SLA pendant le prochain siècle.

Cette quasi-stagnation des progrès sur la SLA prend fin dans les années 1990 avec la publication d'études associant l'agrégation de protéines et la gliose à la SLA, deux aspects qui seront éventuellement acceptés comme caractéristiques de la maladie [15,16]. En 1993, la découverte de mutations associées à la SLA de type familiale (SLAf) dans le gène codant pour l'enzyme Cu/Zn superoxide dismutase 1 (SOD1) se révèle un point tournant dans la recherche sur la SLA [17]. Elle galvanise celle-ci en étant la première cause génétique identifiée de la maladie, fournissant un point de départ à la compréhension de la pathogénèse ainsi que de sa physiopathologie, et en permettant l'élaboration des premiers modèles animaux de la SLA. De cette découverte découlera l'identification du premier composé permettant un traitement palliatif avec une certaine efficacité, le riluzole [18]. La découverte

d'autres gènes impliqués dans la pathologie et la génération des modèles animaux correspondants s'enchaînent alors rapidement dans les années suivantes. Parmi ces gènes on retrouve tout particulièrement TAR DNA binding protein (*TARDBP*) [10,20], Fused in sarcoma (*FUS*) [21,22] et bien entendu Chromosome 9 open reading frame 72 (*C9orf72*) [23]. Plusieurs avancées ainsi que des découvertes notables ont été faites depuis, mais la pathologie moléculaire de la SLA demeure nébuleuse et son élucidation continue à être dépendante de notre capacité à comprendre la fonction ainsi que les perturbations causées par les mutations des gènes associés à la maladie.

1.1.3 Diagnostic et Épidémiologie

La SLA est la plus répandue des maladies des motoneurones, mais les similarités entre ces différentes pathologies ainsi que l'hétérogénéité de sa présentation et de sa progression rendent ardue l'établissement d'un diagnostic clair [1]. L'acquisition de données épidémiologiques pour la SLA en est par conséquent compliquée. Pour établir un consensus des critères de diagnostic en recherche, la fédération mondiale de neurologie établit en 1994 les critères de diagnostic El Escorial [24]. Ils sont révisés en 1998 pour mettre en valeur l'importance des tests en laboratoire et renommés critères Airlie House, la spécificité du diagnostic s'en retrouve augmentée [25]. En 2008, ils sont révisés à nouveau et rebaptisés critères Awaji-Shima afin d'inclure des critères neurophysiologiques dans le but d'accroître la précision du diagnostic et d'exclure les faux positifs [26]. En dépit de ces révisions et des avancées récentes, il n'existe à ce jour aucun test de diagnostic définitif pour la SLA. Le diagnostic est donc réalisé par exclusion en se basant sur les observations cliniques appuyées, selon les circonstances, par des tests génétiques, d'électromyographie, de neuro-imagerie et de sérologies [1]. En moyenne, 8 à 16 mois séparent encore l'apparition des premiers symptômes du diagnostic dont la nature et l'exactitude doivent être confirmées par des tests post-mortem [27].

Une méta-analyse récente révèle une incidence globale standardisée de 1.68 cas de SLA pour 100 000 personne-année de suivi (PAdS). L'incidence standardisée pour 100 000 PAdS est cependant hétérogène à travers le monde, étant en moyenne plus élevée dans les pays occidentaux (1.81) comparativement aux pays asiatiques (0.83) et africains (0.89). Ceci et

d'autres données semblent indiquer une possible influence de l'ethnicité sur l'incidence [28,29]. Au Canada, peu de données sont disponibles à l'échelle du pays, mais des études de l'Ontario, Terre-Neuve, Nouvelle-Écosse et Labrador révèlent une incidence de 2 cas pour 100 000 individus par année en moyenne et une prévalence de 5 à 7 cas pour 100 000 personnes [30]. Une étude plus récente en Colombie-Britannique révèle une incidence de la SLA nominale plus élevée de 3.29 cas par 100 000 individus par année [31]. Les hommes sont de 1.2 à 1.5 fois plus susceptibles de développer la maladie que les femmes, mais cette disparité s'amointrit avec l'âge [32]. En effet, réparti sur la vie d'un individu, le risque total d'être atteint de la SLA en Europe et aux États Unis est d'environ 1:350 pour les hommes et 1:400 pour les femmes [33]. Pareillement, l'âge affecte l'incidence et la prévalence de la SLA, étant rare avant 40 ans puis augmentant dramatiquement avec l'âge avant d'atteindre son maximum entre 70 à 79 ans [34,35].

1.1.4 Histoire naturelle

La pathologie est caractérisée par l'emplacement anatomique où se manifestent les premiers symptômes puisque les motoneurons sont affectés de manière séquentielle [36]. Une importante hétérogénéité existe dans la manifestation de la SLA, mais la forme classique, ~70% des cas, est normalement regroupée sous deux formes cliniques : la forme spinale débutant aux extrémités des membres (~66%) et la forme bulbaire se manifestant au niveau des zones d'innervation bulbaire (~33%) [37]. La forme spinale se manifeste par des faiblesses indolores asymétriques dans les membres, puis un examen clinique révèle habituellement une atrophie accompagnée d'une faiblesse des muscles, des fasciculations, une hyperréflexie et une hypertonie de modérée à sévère [38]. Les muscles de la langue étant les premiers affectés, la forme bulbaire est associée à un défaut d'élocution (dysarthrie) et de déglutition (dysphagie) [39]. De plus, jusqu'à 50% des patients atteints de la SLA présentent des signes de dysfonctions du comportement et/ou de subtiles perturbations des fonctions cognitives et 5 à 15% des cas présentent une forme classique de SLA conjuguée à des troubles cognitifs et/ou du comportement suffisant pour remplir les critères de diagnostic de la démence fronto-temporale (DFT, SLA/DFT) [37,40,41,42]. En raison de recouvrement de facteurs sur les plans génétique, pathologique ainsi que clinique, la SLA et

la DFT sont maintenant généralement considérées comme représentant un continuum d'une même maladie nommée le spectre SLA/DFT. Pour 3 à 5 % des cas, les symptômes initiaux sont respiratoires, soit une dyspnée ou une orthopnée dès le début de la maladie [43]. Le reste des diagnostics relèvent des cas bulbaires isolés où les symptômes demeurent isolés au bulbe (5%) de manière prolongée et des cas de SLA restreint (10%) incluant l'amyotrophie spinale ainsi que la sclérose latérale primaire [44]. Généralement la forme spinale de la SLA présente le meilleur pronostic, suivie de la forme bulbaire puis respiratoire [43]. Les formes qui présentent des atteintes cognitives subtiles ont un pronostic similaire aux formes classiques, mais il est pire pour les formes SLA/DFT ou avec composante comportementale importante [45]. Les symptômes se manifestent généralement de façon asymétrique et sont la conséquence de la dégénérescence des motoneurones supérieurs et inférieurs [36]. Parmi les symptômes de la SLA, les faiblesses musculaires, les crampes, l'atrophie musculaire et la dénervation sont associées aux motoneurones inférieurs alors que la spasticité ainsi que l'hyperréflexie relèvent des motoneurones supérieurs [46]. La progression des symptômes est inexorable, se manifestant par une paralysie générale progressive. La dénervation des muscles respiratoires marque le stade final de la pathologie, caractérisé par une insuffisance respiratoire menant à un décès du patient même en présence d'assistance respiratoire [1]. Bien que la dégénérescence de la SLA affecte la majorité des motoneurones, les nerfs oculomoteurs communs, trochléaire et abducens, impliqués dans les mouvements oculaires, ainsi que le nerf pudendal, qui innerve la région du périnée ne semblent pas affectés [47]. En effet, les mouvements oculaires, la continence urinaire et fécale ainsi que les organes génitaux sont généralement épargnés par la SLA jusqu'en phase très tardive de la maladie [48].

1.1.5 Traitements

Près de 25 ans depuis la découverte d'un gène impliqué dans la SLA [17] et après une multitude d'essais cliniques de phase II et III infructueux [49], seulement deux composés ont démontré une efficacité modérée dans le traitement de la SLA [1]. Le riluzole, approuvé en 1995 [50], confère une modeste extension de l'espérance de vie d'environ 3 mois en moyenne [51,52]. L'édaravone, approuvée en 2017 [53], ralentirait la vitesse de déclin du

score "Amyotrophic lateral sclerosis functional rating scale-revised" (ALSFERS-R) et aurait un effet similaire sur le déclin de l'indice de qualité de vie des patients [54]. Cependant, son effet sur l'espérance de vie est incertain et son efficacité semble restreinte aux patients dans les stades précoces de la SLA. Le reste des soins disponibles est principalement de nature palliative et consiste surtout en une prise en charge des symptômes causés par la dégénérescence des motoneurons et la prise en charge des fonctions respiratoires et nutritionnelles des patients [55]. L'effet thérapeutique limité de ces deux composés illustre l'urgence de la nécessité de découvrir des composés ainsi que des voies de traitement plus efficaces.

1.1.6 Étiologie

L'étiologie de la SLA est complexe, ses causes étant multiples, diverses et la majorité du temps inconnues. Tout d'abord, ~90% des cas de SLA surviennent sans antécédent familial apparent (SLA sporadique, SLAs) alors que seulement ~10% des cas ont une composante héréditaire (SLA familial, SLAf) [56]. Néanmoins, jusqu'à présent aucune différence n'a été décelée entre la SLAf et la SLAs au niveau de la présentation des symptômes, du diagnostic, de l'histoire naturelle ou de la réaction aux traitements [57]. De plus, l'ensemble des gènes mutés associés à la SLAs sont aussi retrouvés mutés pour des cas de SLAf [58].

Ensuite, la perspective qui prévaut actuellement est que la SLA résulterait d'une interaction complexe entre des facteurs génétiques et/ou environnementaux se combinant pour augmenter le risque de développer la pathologie [1,33]. Ainsi, plusieurs études ont cherché à identifier puis étudier les différents facteurs génétiques et environnementaux affectant l'incidence de la SLA. Présentement, l'âge avancé [33], être de sexe masculin [32] et fumer sont généralement acceptés comme des facteurs de risques de la SLA, mais d'autres facteurs ont été proposés dont un index de masse corporelle bas, une forme physique athlétique, un stress physique intense prolongé, les traumatismes crâniens répétés, l'exposition aux pesticides, l'exposition au plomb, certains types d'alimentation et l'exposition à certains virus. Cependant, leur implication dans la SLA reste incertaine et contestée. Effectivement, peu d'études existent pour ces facteurs et la plupart

souffrent d'un échantillonnage inadéquat, une méthodologie discutable ainsi que d'un manque de reproductibilité [8].

Au niveau génétique, les mutations associées à la SLAf démontrent souvent une transmission autosomique dominante avec une pénétrance partielle et souvent inconnue, moins fréquemment la transmission est autosomique récessive et très rarement elle s'effectue selon un modèle dominant lié au chromosome X [59]. La cooccurrence des mutations, la pénétrance partielle observée et l'impossibilité d'expliquer la distribution de tous les cas de SLA dans une famille présentant une mutation reconnue ont mené à la proposition que la SLA serait une maladie oligogénique [58]. Plusieurs mutations dans des gènes différents seraient alors nécessaires pour que la maladie soit exprimée pleinement, leur effet étant additif ou synergique.

En ce qui concerne les gènes eux-mêmes, plus de 30 ont été identifiés comme liés à l'étiologie de la SLA et ils sont impliqués dans divers processus cellulaires, dont notamment la dynamique du cytosquelette, le trafic cellulaire et l'homéostasie des protéines ou de l'ARN l'ARN. Collectivement les mutations retrouvées dans ces gènes sont estimées être responsables de ~70% des cas de SLAf et ~15% de SLAs, la vaste majorité des cas de SLA demeurent donc sans cause génétique répertoriée. Du nombre grandissant de gènes associés à la SLA, les gènes *SOD1*, *FUS*, *TARDBP* et *C9orf72* (*C9*) sont considérés actuellement comme les gènes causatifs principaux puisqu'ils sont responsables collectivement de près de 50 à 60% des cas de SLAf, dont jusqu'à 40-50% pour le *C9* [58].

En 1993 *SOD1* est le premier à être identifié, *SOD1* étant un puissant antioxydant protégeant l'organisme des espèces réactives oxygénés. Plus de 180 mutations variées de *SOD1* ont été identifiées et elles seraient responsables de 10 à 20% des cas de SLAf ainsi que ~3% des cas de SLAs [59]. Ces mutations variées se retrouvent à travers l'ensemble de la séquence codante de *SOD1* et sont susceptibles d'affecter l'activité enzymatique, la structure ainsi que la stabilité de la forme native de *SOD1*. Par contre, ces mutations sont dominantes plutôt que récessives, impliquant un gain de fonction toxique dans la SLA. Ceci provoquerait un repliement aberrant, la déstabilisation de l'homodimère standard et l'agrégation de la protéine générant ainsi des espèces neurotoxiques [60]. Malgré l'expression ubiquitaire de *SOD1*, ces mutations

provoquent une toxicité qui affecte particulièrement les motoneurones entraînant spécifiquement leur dégénération [61]. Plusieurs voies hétérogènes de pathologies cellulaires ont été proposées pour expliquer la dégénérescence des motoneurones dont l'excitotoxicité, la dérégulation des facteurs neurotrophiques et des protéines de guidance axonale, des défauts du transport axonal, des dysfonctions mitochondriales, des perturbations de l'homéostasie des protéines et du métabolisme de l'ADN ou de l'ARN [59]. L'hétérogénéité des systèmes possiblement impliqués dans la toxicité de SOD1 illustre la difficulté et la complexité à laquelle se heurte la recherche en tentant d'élucider la pathogénèse de la SLA.

Le lien entre la SLA et la protéine TDP-43, encodé par le gène *TARDBP*, est établi en 2006 lorsqu'elle s'avère être la principale composante des inclusions retrouvées dans les neurones et les cellules gliales des patients atteints de SLA et/ou de DFT [62]. Ces inclusions sont l'indicateur pathologique majeur dans ~95% des cas de SLA et ~50% de DFT, incluant ceux associés à la mutation de *C9* [64]. Depuis, plus de 50 mutations de *TARDBP* ont été associées à l'étiologie de la SLA [65] et elles seraient responsable de ~3% des cas de SLAf ainsi que ~1.5% des cas de SLAs [66]. La protéine ubiquitaire TDP-43 est impliquée dans le métabolisme de l'ARN et la réponse au stress, elle se retrouve de manière prédominante dans le noyau, mais elle est aussi moins présente dans le cytoplasme [65]. La toxicité de TDP-43 provient de sa perte de localisation au noyau pour s'agréger dans le cytosol, résultant en une perte de fonction nucléaire et la formation d'agrégats cytoplasmiques anormaux [67]. Les mécanismes de toxicité associés incluent une hydrophobicité accrue menant à la séquestration de composants cellulaires cruciaux, la génération de ROS, une dérégulation du métabolisme de l'ARN et une inhibition des protéasomes [44].

Pour *FUS*, c'est en 2009 que des mutations faux-sens autosomiques dominantes sont identifiées comme responsables de cas de SLAf chez plusieurs individus de familles atteintes [68]. Plus de 50 mutations de *FUS* ont été découvertes [69] et elles sont responsables de ~5% des cas de SLAf ainsi que ~1% des cas de SLAs [70]. La protéine FUS lie l'ARN et l'ADN comme TDP-43 et présente plusieurs similarités de fonctions avec cette dernière, elle est aussi retrouvée principalement dans le noyau tout en pouvant transiter au cytoplasme. La toxicité associée aux

mutations de *FUS* est généralement acceptée comme le résultat d'altérations du système de réparation de l'ADN ainsi que de plusieurs processus en lien avec le métabolisme de l'ARN dont la transcription, l'épissage, son incorporation dans des granules de stress et la localisation subcellulaire. On observe aussi l'agrégation de la protéine FUS formant éventuellement de larges inclusions cytoplasmiques dans les neurones qui sont similaires aux agrégats de TDP-43, mais distincts à plusieurs égards dans leur composition [71,72,73].

Enfin, en 2011 une expansion massive de répétitions hexanucléotidique (G4C2) anormales dans le premier intron du gène *C9* est identifiée comme l'anomalie génétique la plus commune chez des familles présentant une pathologie avec les phénotypes SLA, DFT et TDP-43 [74]. Ces répétitions, seule mutation répertoriée dans *C9*, s'avèrent responsables de 40 à 50 % des cas de SLA et de 5 à 10% des cas de SLAs [75,76] pour un total de ~10% des cas de SLA toutes formes confondues, *C9* étant donc présentement le gène principal associé à l'étiologie de la SLA [1]. Les répétitions (G4C2) de *C9* sont aussi la cause la plus commune des cas de DFT familiale (~25%) et des cas de SLA/DFT (~88%) [77] en plus d'être identifiées comme une cause rare de maladies neurodégénératives autres que la SLA et la DFT, dont la maladie de Parkinson et la maladie d'Alzheimer [78]. Cependant, plusieurs aspects de la protéine *C9* demeurent méconnus notamment son profil d'expression, les mécanismes exacts de toxicités des répétitions et sa fonction biologique normale. L'intérêt du laboratoire Drapeau pour ce gène a donc été motivé par le développement de modèles in vivo permettant l'étude de *C9* dans des conditions normales et son implication dans diverses pathologies avec une emphase particulière sur la SLA. Mon projet de maîtrise a donc comme point central le gène *C9*, pour mieux comprendre sa fonction ainsi que son expression, afin d'offrir des pistes additionnelles pour comprendre les voies pathologiques dans lesquelles il est impliqué et potentiellement de révéler de nouvelles avenues thérapeutiques.

1.2. C9orf72

1.2.1 Structure du gène *C9orf72*

Le gène *C9orf72* (NC_000009.12) est localisé chez l'humain au niveau du locus 9p21.2. Il est constitué de 11 exons, dont 2 premiers exons non-codants alternatifs (1a et 1b) et 10 exons codants (2 à 11). Le gène est conservé des eucaryotes aux mammifères [79] avec généralement une seule copie et des homologues du gène ont même été identifiés chez des espèces hors du règne animal. [80].

1.2.2 Conservation de *C9orf72*

Une étude bio-informatique comparative de C9 a été établie par PSI-Blast le degré (%) d'homologie de séquence entre la protéine C9 humaine et celles d'espèces animales de différents taxons : primates (99%), mammifères dont le rat et la souris (96-98%), marsupiaux (94-96%), oiseaux (93%), reptiles (92%), (72-77%) pour les poissons incluant le poisson-zèbre (76%) et 58% pour le ver *C. elegans*. Ce degré de similarité et de conservation de la séquence élevée entre des espèces évolutivement distante suggère une importante contrainte fonctionnelle pour la protéine C9. La préservation de résidus d'acides aminés tout au long de la séquence de la protéine à travers les espèces indiquerait que C9 agit comme une protéine à un seul domaine essentiel à sa fonction. L'isoforme de C9, composée de 481 acides aminés, est la plus exprimée à travers toutes les espèces considérées [79].

1.2.3 Épissage alternatif de *C9orf72*

Par épissage alternatif, le gène *C9orf72* humain est transcrit en 5 ARN messagers (ARNm) distincts (4808 pb, 3200 pb, 1873 pb, 862 pb et 777 pb). De ces variants d'épissage de C9, trois transcrits (3200 pb, 1873 pb et 777 pb) sont des variants avec une queue polyadénylée et sont traduites en protéines [79]. La traduction de ces trois transcrits produit seulement 2 isoformes distinctes de la protéine C9 dont l'expression a été validée à l'aide d'anticorps spécifiques : une isoforme courte (C9C, 222 aa, 24.8 kDa) produite par la traduction du variant 1 et une isoforme longue (C9L, 481 aa, 54.3 kDa) résultant de la traduction des variants 2 et 3. Les deux isoformes présentent une localisation subcellulaire distincte ainsi que des profils biochimiques différents

[81,82]. Les deux autres variants d'épissage sont des transcrits traités, mais sans cadre de lecture annoté et il est présentement supposé que ces transcrits ne sont pas traduits en un produit protéique [79].

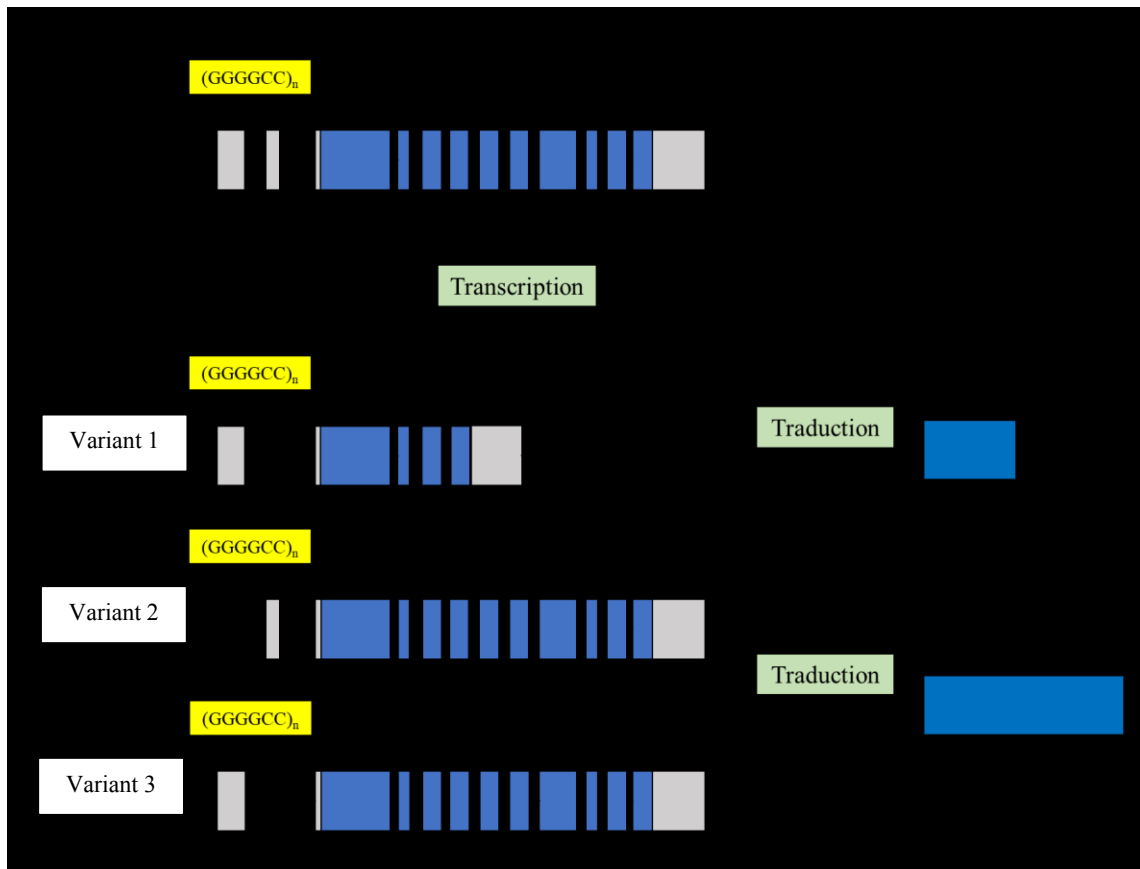


Figure 1. Schéma de la séquence ADN du gène *C9orf72* et des trois transcrits alternatifs produisant les 2 isoformes de protéine. En gris les exons non codants, en bleu les exons codants et la protéine résultante. Adapté de Balendra, R. and A. M. Isaacs, 2018 [126].

1.2.4 Structure de la protéine *C9orf72*

Des analyses bio-informatiques de haute sensibilité conjuguées à des méthodes d'analyses par homologie structurelle prédisent que C9 possède un domaine unique qui recouvre la quasi-totalité de sa séquence, appelé le domaine différenciellement exprimé dans les cellules normales

et néoplasiques (DENN) [83] (**figure 2**). C9 fait donc partie de la famille de protéines caractérisées par ce domaine, les protéines à domaine DENN. La sous-structure du domaine DENN était subdivisée en trois modules distincts (uDEEN en amont, cDENN au centre et dDEEN en aval). Cette subdivision est encore parfois employée aujourd’hui, mais des découvertes subséquentes révèlent plutôt deux lobes étroitement associés [84,85] (**Figure 2**). Le lobe N-terminal serait un domaine “longin” [84,85], un domaine associé à la modulation de l’activité des protéines SNARE impliquées dans le trafic membranaire par le contrôle de la fusion membranaire [86]. Le lobe C-terminal est un repliement complexe α - β de fonction inconnue et présentement associé à aucune autre famille de domaines [84,85]. Les protéines contenant le domaine DENN sont reconnues pour leur fonction de facteurs d’échange de nucléotides de guanine (GEF) pour les Rab GTPases [87]. Les Rab GTPase sont des protéines monomériques liant le GTP alternant entre deux conformations, une forme inactive liée au GDP et une forme active liée au GTP. Les GEF activent les Rab en catalysent la conversion des protéines Rab lié au GDP à leur forme liée au GTP. On associe les Rab GTPases à la régulation de plusieurs étapes du trafic membranaire dont la formation, la maturation, le transport, la fusion, l’amarrage et la fusion des vésicules. Elles sont aussi des déterminants clés de l’identité des organelles et, par le recrutement d’autres protéines, elles régulent pratiquement l’ensemble du trafic membranaire chez les eucaryotes [88]. Peu est donc connu de la structure de C9, mais de son domaine DENN il est prédit qu’elle ait un rôle dans la régulation du trafic membranaire.

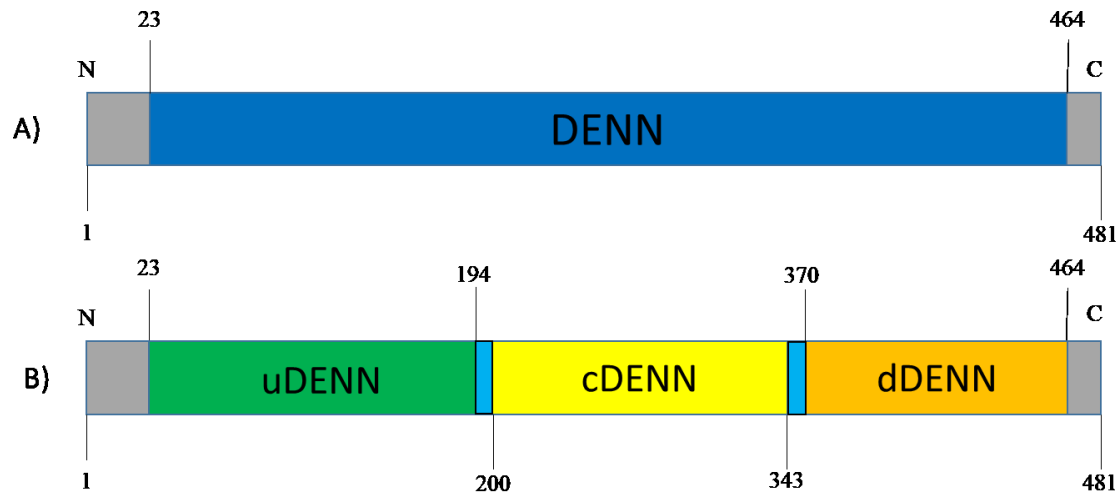


Figure 2. Structure de la protéine C9orf72 (A) Étendue du domaine fonctionnel DENN de C9 pour l'isoforme longue du 23^{ième} aa au 464^{ième} aa. (B) Subdivision en trois sous domaine encore parfois utilisé aujourd'hui (23-200 aa uDENN, 200-343 aa cDENN, 343-464 aa dDENN). Adapté des informations de la banque de données en ligne <https://prosite.expasy.org/>; (décembre ,2018).

1.2.5 Profil d'expression

Les isoformes de C9 sont exprimés à des niveaux variables dans l'ensemble des organes humains. Leur expression est relativement élevée dans les tissus associés au système immunitaire (rate) et particulièrement abondante dans le SNC, mais elle varie d'un transcrit à l'autre, ainsi que selon les régions du SNC avec l'expression la plus élevée dans le cervelet [73,89,90]. L'expression de C9 diffère aussi entre les types de cellules dans le SNC, étant particulièrement élevée dans les monocytes CD14+, les éosinophiles et les neutrophiles, mais comparativement peu élevée dans les autres types de cellules, incluant celles dérivées des cellules myéloïdes comme les macrophages ou les mastocytes. En termes d'expression absolue, C9 est faiblement exprimée relativement à d'autres gènes dans la plupart des cellules du corps humain incluant celles du SNC [91].

Dans la cellule, C9 se retrouve principalement dans le cytoplasme avec une distribution diffuse, mais on la retrouve aussi moindrement au noyau [89, 81, 92]. Elle est aussi retrouvée dans des vésicules du noyau et du cytoplasme, notamment au niveau des lysosomes, des autophagosomes,

et des endosomes, particulièrement ceux précoces [93,94,95]. On la retrouve également dans les granules de stress (GS), les p-body, l'axone, le cône de croissance et enrichie dans l'appareil de Golgi [92,96,97]. De plus, les isoformes de C9 ont une localisation cellulaire distincte à certains égards. C9L est cytoplasmique, peu présent au noyau et se retrouve aussi dans l'axone ainsi que l'élément présynaptique des neurones, alors que C9C se retrouve principalement à la membrane nucléaire et dans le noyau lui-même, particulièrement à la chromatine [81,90]. Par ailleurs, C9 a aussi été détectée dans le milieu conditionné de culture cellulaire, suggérant une sécrétion active de C9 [93].

1.2.6 Fonction de la protéine C9ORF72

Afin d'élucider les fonctions physiologiques encore incertaines de C9, plusieurs modèles cellulaires et animaux avec une expression réduite du gène *C9* (knockdown, KD), complètement abolie (knock-out, KO) ou surexprimée ont été générés.

Des cellules appauvries en *C9* montrent une dérégulation du transport endosomal marquée par une perturbation de l'initiation de l'autophagie accompagnée de l'accumulation d'agrégats cytoplasmiques SQSTM1/p62, un adaptateur et substrat pour la dégradation dans l'autophagie, ainsi que TDP-43 phosphorylée [93,98,]. On observe aussi une potentialisation de l'agrégation de protéines structurellement compromises, dont ataxin-2 avec des expansions polyglutamiques [98]. De plus, une déplétion de *C9* altère de manière significative la transcription de milliers de gènes, dont ceux codant pour des endothélines (EDN) et d'autres impliqués dans le cycle glutamate-glutamine, corrélant avec une signalisation EDN altérée ainsi qu'une accumulation du glutamate intracellulaire [99]. Le KD de *C9* entraîne en plus une abolition complète des granules de stress (GS), une inhibition de l'expression de protéines associées à la formation des GS et accélère la mort cellulaire [92]. Il a aussi été rapporté que des cellules KD pour *C9* présentent une prolifération des vacuoles, ainsi qu'une augmentation significative de la taille du noyau et de la cellule [99]. Enfin, le KD de *C9* provoque une diminution de la sécrétion des vésicules intracellulaires et extracellulaires accompagnée de dysfonctions du réseau trans-golgi [97]. Ensuite, les cellules KO pour *C9* présentent, en plus des précédents phénotypes, un élargissement de vésicules lysosomales, l'accumulation de lysosomes et de leurs enzymes dans les cellules ainsi qu'une perturbation de la signalisation de mammalian target of rapamycin

complexe 1 (mTORC1) lors de la réponse cellulaire qui dépend des lysosomes lors d'une carence en nutriments, ce qui entraîne une augmentation du flux de l'autophagie. En condition d'abondance de nutriments, les lysosomes sont retrouvés en périphérie de la cellule, ce qui est associé à une activation de mTORC1 alors que lorsque les nutriments sont limités, les lysosomes se regroupent dans la région périnucléaire facilitant la fusion autophagosome-lysosome qui entraîne le relâchement de nutriments pour répondre à la carence nutritionnelle [94,100,101,102]. Dans les cellules neuronales, le KO de *C9* provoque également une augmentation des récepteurs du glutamate et une sensibilité accrue à l'excitotoxicité [95]. Pour les cellules qui surexpriment *C9*, on remarque l'apparition d'agrégats nucléaires anormaux et de granules de stress cytoplasmique [93]. Finalement, un KD de *C9* chez les motoneurones provoque l'écourtement de l'axone et leur élongation significative si *C9* est surexprimée [96].

Pour les modèles animaux, un knock-out de l'homologue de *C9* chez le *C. elegans*, *alfa-1*, provoque des défauts de mobilité, qui progressent en une paralysie dépendante de l'âge, accompagnée d'une dégénérescence spécifique des motoneurones GABAergiques exacerbées par le stress osmotique [103]. Sa perte de fonction provoque aussi une dérégulation générale de l'homéostasie des lysosomes, menant à une diminution de la reformation des lysosomes ainsi qu'une dégradation défectueuse du contenu endosomal [104]. Chez les embryons du poisson-zèbre, un KD transitoire de *C9* mène à une arborisation anormale et un écourtement de l'axone des motoneurones accompagnés de déficits locomoteurs. Il affecte aussi négativement l'organisation et la densité des commissures formés par les axones des neurones du rhombencéphale [105]. Enfin, plusieurs souris KD et KO ont été développées. Les KD transitoires et les KO homozygotes conditionnels de *C9* dans les neurones et les glies n'entraînent pas de phénotype moteur ou comportemental [106,107]. Cependant, le KO ubiquitaire de *C9* provoque chez les souris des phénotypes de splénomégalie, d'adénopathie, de glomérulonéphrite, de leucocytose des neutrophiles, une augmentation des cytokines pro-inflammatoires, des événements néoplasiques et une réponse auto-immune sévère, accompagnée d'inflammation des organes périphériques ainsi que du SNC, entraînant un haut taux de mortalité [102,108,109,110]. Par contraste, l'haploinsuffisance de *C9* (lignées hétérozygotes pour le KO de *C9*) ne mène pas à des phénotypes aussi sévères [102].

Ces expériences illustrent que C9 joue un rôle important dans un vaste éventail de voies cellulaires. Notamment, au niveau du trafic vésiculaires incluant la régulation de l'autophagie, l'activité lysosoméale, des sécrétions extracellulaires, le transport vésiculaire du réseau trans-golgien et la protéostase. Elle serait aussi impliquée dans les voies de réponses au stress métabolique ainsi qu'environnemental et semblerait avoir un rôle dans la régulation de la transcription. Les modèles animaux indiquent aussi une implication de C9 dans le développement neuronal, la réponse immunitaire, la locomotion et l'espérance de vie. Les mécanismes et voies par lesquels C9 affecte ces systèmes restent encore nébuleux, mais plusieurs ont été proposés pour ses nombreux partenaires d'interaction identifiés.

1.2.6.1 Régulation trafic vésiculaire

Comme prévu par son domaine DENN, il a été démontré que C9 possède une activité GEF pour Rab5A, Rab7A et Rab11A [111] l'impliquant dans plusieurs voies du transport endosomal : (i) maturation des endosomes précoces aux endosomes tardifs (Rab5A, Rab7A) [112,113], (ii) régulation de l'exocytose des vésicules ainsi que recyclage des protéines de l'endocytose (Rab11A) [114] et (iii) régulation du trafic endosomes/lysosomes (Rab7A) [113]. De plus, ces Rab interviennent dans l'autophagie, faisant de C9 un régulateur de plusieurs voies de celle-ci : (i) initiation de l'autophagie (Rab5A), (ii) l'élongation des autophagosomes (Rab11A, Rab39b) et (iii) la maturation ainsi que fermeture des autophagosomes (Rab7A, Rab11A) [115]. Qui plus est, C9 affecte le transport vésiculaire en agissant en aval de l'activation de certaine Rab. En effet, C9 agirait comme effecteur pour Rab1a activée, régulant le trafic du complexe kinase ULK1 d'initiation au phagophore, entraînant la formation de l'autophagosome [116]. Elle serait aussi un effecteur de RAB7L1 activée, impliquée dans le transport vésiculaire, notamment au niveau du Golgi et la sécrétion de vésicules extracellulaires [93]. Additionnellement, C9 a été démontrée comme ayant des interactions, de natures encore inconnues, avec tous les membres de la famille de protéines RAB3, qui sont des Rabs abondantes dans les vésicules synaptiques et jouent un rôle clé dans la sécrétion de neurotransmetteurs [90].

De surcroît, C9 forme un complexe GEF, pouvant lier des Rab GTPases additionnelles quand elle interagit avec WD-40 repeat 41 (WDR41) et Smith-Magenis syndrome chromosome region, candidate 8 (SMCT8) [98,109,117]. Les fonctions et l'expression de WDR41 et SMCR8 sont

encore méconnues, mais SMCR8 possède aussi un domaine DENN et est impliquée dans l'autophagie, alors que WDR41 est enrichie au niveau de la membrane du cis-golgi [109]. En plus de se lier fortement, la stabilité de C9 et SMCR8 serait hautement dépendante de leur expression mutuelle. En effet, un KD de C9 mène à une diminution significative du niveau de protéine de SMCR8 et vice versa, [94,100]. Jusqu'à présent, le complexe C9-SMCR8-WDR41 possède une activité GEF pour Rab8a et Rab39b. Rab8a est associée à l'exocytose dérivée du golgi et au trafic du réseau trans golgien à la membrane plasmique, incluant le transport des récepteurs du glutamate [118]. Pour sa part, Rab39b régule le trafic intracellulaire, notamment au niveau des dendrites [119]. Pour l'interaction de C9 avec Rab39b, il est observé que TANK-Binding Kinase1 (TBK1) phosphoryle SMCR8 activant le complexe GEF permettant à Rab39b d'activer le complexe ULK1 par liaison à C9, provoquant l'initiation de la formation des autophagosomes.

1.2.6.2 Réponse immunitaire

Le rôle exact de C9 dans la réponse immunitaire est incertain, mais il est vraisemblablement aussi lié à son rôle dans l'autophagie. En effet, l'autophagie agit comme un effecteur important de la réponse immunitaire qui régule l'élimination des pathogènes. De plus, son rôle s'étend à la réponse immunitaire innée ainsi qu'acquise et inclut des fonctions dans la sélection thymique, la présentation d'antigène, la stimulation de l'homéostasie des lymphocytes ainsi que leur survie et la régulation de la production des cytokines [120].

1.2.6.3 Signalisation de mTOR et système lysosomal

En plus de son interaction avec les Rabs, la localisation du complexe C9/SMCR8 aux lysosomes sous conditions de carence nutritionnelle, ainsi que les nombreux phénotypes délétères des lysosomes associés au KD/KO de C9, indiquent un rôle dans le réseau coordonné d'expression et de régulation lysosomal (CLEAR) qui contrôle la plupart des gènes liés à la biogénèse des lysosomes et l'autophagie, améliorant les capacités de dégradation cellulaire [121,122]. En effet, en condition de famine et de KO pour C9, l'activité kinase de mTORC1 est inhibée, augmentant la localisation nucléaire de Transcription factor EB (TFEB), un facteur de transcription clé du réseau CLEAR, et accroît le flux autophagique [100,123]. Il est alors proposé que C9/SMCR8 se localisent aux lysosomes en conditions de stress métabolique où elles participent au processus

de détection des nutriments par mTORC1 en mitigeant son activité par l'entremise de partenaires encore inconnus. Ainsi, C9, seul ou en complexe avec SMCR8, agirait dans différentes phases de la voie mTORC1 dépendamment des conditions métaboliques ou de stress [94].

1.2.6.4 Élimination d'agrégats de protéine par macroautophagie

Ensuite, C9 agit dans la dégradation des agrégats de protéines par la macroautophagie sélective. Un processus impliquant l'acheminement spécifique d'organelles ou d'agrégats de protéines aux autophagosomes, médié par des protéines avec un domaine de liaison LC3, dont SQSTM1/p62 et OPTN. C9 seul ou avec SMCR8 régulerait la macroautophagie sélective en interagissant avec ces protéines adaptatrices et Rab39b, dirigeant aux autophagosomes les protéines polyubiquitinées et les organelles défectueuses à être dégradées [98,105,123]. En accord avec une telle fonction, C9 purifie aussi avec Hsc70, Hsp90, et Bag3, tous impliqués dans l'autophagie ainsi que les voies de dégradation des protéines, lors d'expériences de co-immunoprécipitation [98].

1.2.6.5 Régulations du cytosquelette de l'actine

C9 régule aussi la dynamique des filaments d'actine dans les motoneurones en inhibant l'activité de liaison au GTP de ARF6 et Rac1, ce qui inactive ces dernières. Ceci a pour résultat de réduire l'activité des kinases LIMK1 et LIMK2 qui sont responsables de la phosphorylation et l'inactivation de la cofiline, menant donc à l'activation de cette dernière qui lie l'actine pour médier le désassemblage des filaments [96]. La découverte que C9 agit aussi comme une GEF pour les Rho GTPases, centrales à la dynamique d'assemblage et du réarrangement du cytosquelette de l'actine [124], Cdc42 et RhoA, réaffirme une fonction de C9 dans la régulation du cytosquelette [111].

1.2.6.6 Régulation de l'expression génique

Il est apparent que le niveau d'expression de la protéine C9 altère l'expression génique, mais le mécanisme exact est encore inconnu. Une possibilité est que C9C fonctionne directement comme un facteur de transcription, telle que suggérée par sa localisation nucléaire, son association à la chromatine ainsi qu'à la région du promoteur de EDN1. La seconde possibilité,

non mutuellement exclusive, est que l'expression altérée de C9L et C9C affecte indirectement l'expression génique, tel que par la régulation indirecte de l'activité du facteur de transcription nucléaire NF- κ B [99]. En effet, l'expression réduite de C9 mène à des défauts de l'autophagie, résultant en une augmentation des agrégats p62 qui activent la signalisation de NF- κ B [125].

1.2.7 *C9orf72* dans la SLA

1.2.7.1 Mutation de *C9orf72*

La mutation de *C9* associée à la SLA/DFT se distingue des mutations des autres principaux gènes associés à la SLA/DFT par sa nature intronique, n'affectant pas la séquence codante du gène. Effectivement, les expansions de répétitions hexanucléotidique (ERH) G4C2 se retrouvent dans le premier intron des variants de transcrits 1 et 3 de *C9*, qui correspond à la région du promoteur de la variant 2 [126]. De plus, de 0 à 24 répétitions G4C2 dans *C9* sont communément présentes et bénignes dans la population générale. Comparativement, la taille des ERH retrouvées chez les patients atteints sont estimées entre plusieurs centaines à des milliers. Curieusement, les patients atteints de *C9*(SLA/DFT) ont une présentation clinique très hétérogène, incluant au sein d'une même famille [78,127]. De plus, des études de population suggèrent que de 0.2 à 0.6% de la population nord européenne porte les ERH pathologiques, une fréquence de porteur nettement en excès de l'incidence rapportée des pathologies des répétitions ERH de *C9* [128,129]. Additionnellement, l'âge de manifestation des pathologies liées aux ERH dans *C9* et le pronostic associé sont aussi remarquablement variables. Ainsi, il est devenu apparent qu'il existe une variabilité significative de la pénétrance des phénotypes associés avec la mutation de *C9*, suggérant la présence de facteurs génétiques agissant comme des modificateurs des pathologies liées à *C9*. Incidemment, des mutations modificatrices et des variantes d'autres gènes associés à la SLA/DFT sont souvent aussi présentes chez les patients atteints et positifs pour les ERH. Parmi les modificateurs géniques de *C9*(SLA/DFT), la longueur des ERH serait une candidate majeure selon les autres pathologies d'expansion de répétitions. En outre, les ERH particulièrement longues dans *C9* sont instables, pouvant mener à une mosaïque de leur longueur chez les cellules somatiques des patients *C9*(SLA/DFT), notamment dans différentes régions du cerveau. Elles peuvent aussi augmenter considérablement au fil des générations et un phénomène d'anticipation a été observé dans

certaine famille atteinte. Cependant, le nombre "critique" d'ERH pour provoquer la neurodégénérescence ou pour que celles-ci deviennent instables demeure incertain [78,126,127]. De plus, les résultats concernant une corrélation du nombre d'ERH avec le phénotype clinique ainsi que l'âge de manifestation de la maladie sont présentement conflictuels et aucunes preuves décisives sur la nature du rôle de modificateur génétique pour la longueur des ERH n'a encore été établie [58].

1.2.7.2 Mécanismes pathologiques des ERH de *C9orf72*

Trois mécanismes majeurs non-mutuellement exclusifs et possiblement additionnels dans leurs effets ont été proposés pour expliquer comment les ERH dans *C9* causent les pathologies qui leurs sont associées : 1) un gain de fonction toxique à partir des transcrits ARN des longues répétitions G4C2 sens et antisens; 2) un gain de fonction toxique par génération de protéines de répétitions di-peptidiques (PRD), dérivées de la traduction indépendante de l'AUG associée aux répétitions (RANT) des transcrits ARN des répétitions G4C2, 3) une perte de fonction ou haploinsuffisance de la protéine C9 causée par les ERH [126].

1.2.7.2.1 Mécanismes de gain de fonction toxique

Conformément à une hypothèse de toxicité de l'ARN, les structures répétitives d'ARN s'organisent en structures secondaires et tertiaires, formant des inclusions dans le noyau où elles peuvent lier et séquestrer des facteurs cellulaires, menant à une dérégulation secondaire de processus incluant l'épissage, la transcription et le transport de l'ARN [130]. Incidemment, on observe, en plus des agrégats TDP-43, des inclusions d'ARN sens et anti-sens des répétitions G4C2 de *C9* largement distribuées dans le SNC des patients atteints de C9(SLA/DFT). Ces inclusions caractéristiques des ERH de *C9* sont retrouvées majoritairement chez les neurones au niveau de l'hippocampe, du cortex moteur, du lobe frontal, du cervelet, de la moelle épinière et sporadiquement dans le cytoplasme [126,131]. Moins fréquemment, elles sont détectées dans les astrocytes, les microglies et les oligodendrocytes [126,132]. Les transcrits ARN incorporant les ERH forment des structures secondaires, incluant des structures en tige-boucle et des G-quadruplex. De plus, d'autres structures secondaires incluant des hétéroduplexs ADN-ARN, des duplex d'ARN et d'ADN "I-motif" pourraient émerger des séquences sens ainsi qu'antisens d'ADN/ARN des répétitions. In vivo, ces structures secondaires peuvent lier les protéines liant

l'ARN (PLA) et donc entraîner leur déplétion, fournissant ainsi une voie potentielle directe de la toxicité des transcrits d'ARN des ERH. Effectivement, la déplétion des PLA à un fort potentiel d'induction de neurotoxicité étant des protéines liant l'ARN double et simple brin, formant des complexes ribonucléoprotéines ayant des rôles majeurs dans plusieurs fonctions cellulaires dont les processus biologiques axonaux, mais principalement dans l'expression génique par le contrôle post traductionnel [133]. Parmi les PLA rapportée d'interagir avec les répétitions d'ARN, on retrouve ADARBP, hnRNP-H, SRSF1, hnRNP-A3, hnRNP-A1, Pur- α , SRSF2, hnRNP-F, ALYREF et nucléoline [134,135,136]. Cependant, malgré ces associations notamment à une dérégulation du Spliceosome et du transcriptome, l'implication exacte des transcrits d'ARN des ERH de *C9* et de leurs inclusions dans les pathologies SLA/DFT demeurent à être élucidées. [126,136].

Toujours dans l'optique d'un gain de fonction toxique, un autre trait distinctif des pathologies des ERH de *C9* est la production de PRD, dérivée de la RANT, pouvant interagir de manière délétère avec les domaines de faibles complexités dans les protéines cellulaires pour possiblement induire une toxicité [137]. En effet, bien que les répétitions soient dans une région non-codante, elles sont traduites de manière bidirectionnelle dans tous les cadres de lecture produisant 5 PRD distinctes : glycine-alanine (GA) et glycine-arginine (GR) des transcrits G4C2 sens; proline-arginine (PR) et proline-alanine (PA) des transcrits anti-sens; glycine-proline (GP) des transits sens et anti-sens [126,138]. Les PRD forment majoritairement des inclusions cytoplasmiques chez les neurones, mais elles forment aussi des inclusions névritiques et des pré-inclusions diffuses dans le cytosol. Des inclusions intra et para-nucléaires sont aussi parfois observées [139,140]. Cette pathologie des PRD est plus présente dans le cervelet, l'hippocampe ainsi que le néocortex, elle est moins fréquente dans les régions sous-corticales puis rarement observée dans le tronc cérébral et la moelle épinière [126,141]. Les poly-GA sont la forme la plus abondante de DPR associée aux ERH de *C9*, suivis par les poly-GP puis par les poly-GR. À l'opposé, les poly-PA et les poly-PR sont les formes de DPR les moins fréquentes [126,142]. Les PRD contenant de l'arginine (p.ex., poly-GR, poly-PR) semblent particulièrement toxiques, rapportées d'interférer avec la dynamique d'organelles sans membrane, incluant les granules de stress ainsi que les nucléoles et de perturber le transport nucléaire, induisent une mauvaise localisation des protéines altérant le traitement de l'ARN

[143,137]. Le rôle exact des PRD, particulièrement de chaque espèce, dans les pathologies SLA/DFT reste à être élucidé, mais le stress nucléaire, l'inhibition de la traduction, la suppression de l'ARN ribosomal, un transport nucléoplasmique dysfonctionnel, une signalisation de Notch anormale, une dysfonction du système ubiquitine-protéasome et des anomalies du spliceosome et des granules de stress ont tous été impliqués [137,144,145,146,147,148,149,150].

Afin de déterminer l'influence de l'ARN des ERH de *C9*, plusieurs modèles cellulaires et animaux exprimant les répétitions ont été générés. Dans les motoneurones la transfection de répétitions G4C2 entraîne l'apparition d'inclusions d'ARN et une réduction de leur survie. Aucune PRD n'a été détectée, mais coexprimer poly-PR avec les ERH a un effet négatif synergique sur la survie des neurones [143]. Chez la drosophile, l'expression de 30 ERH G4C2 avec une interruption de 6 pb (CTCGAG) au centre des ERH a été associée à la dégénérescence des motoneurones [151]. Curieusement, l'expression de ERH sens et antisens de longueur similaire à celle retrouvée chez les patients *C9*(SLA/DFT) dans la drosophile n'entraîne pas la dégénérescence des neurones adultes [152]. Cependant, un modèle de poisson-zèbre exprimant des ERH G4C2 durant le développement manifeste une axonopathie des motoneurones, suggérant une vulnérabilité des motoneurones durant le développement [153]. Un autre modèle du poisson-zèbre avec une expression stable d'ERH montrant des accumulations d'inclusion d'ARN ainsi que de PRD, présente des déficiences motrices, des perturbations cognitives, une atrophie musculaire, une perte de motoneurones et une mortalité accrue au début de l'âge adulte [154]. Cependant, dans la majorité de ces modèles, les PRD ne sont détectées que suite à une induction forte de l'expression des ERH et malgré ces observations, deux études chez la drosophile suggèrent un effet neurotoxique synergique dépendant des PRD plutôt que de l'ARN des ERH seul. Effectivement, la dégénérescence des neurones adultes est inhibée par l'insertion de codons stop dans la séquence des ERH exprimées, empêchant la production de PRD [155]. Ensuite, une surexpression ubiquitaire des répétitions a mené à la formation d'inclusion d'ARN dans les neurones et les glies sans production de PRD, aucune neurodégénérescence, réduction de la survie ou altération importante du transcriptome n'est observée. Une surexpression importante des ERH mène à la production de PRD et la réduction de l'espérance de vie [156].

Pour évaluer la contribution distincte des PRD à la pathologie C9(SLA/DFT), la séquence G4C2 a été altérée pour exprimer chaque DPR individuellement dans plusieurs modèles cellulaires et animaux. Des études dans les lignées cellulaires neuronales et non-neuronales indiquent que les poly-GR et les poly-PR sont les plus toxiques alors que poly-GA et poly-PA sont moins toxiques [126,146,157]. Conformément, les poly-GR et poly-PR sont neurotoxiques lorsqu'exprimées dans les yeux des drosophiles. De plus, les drosophiles les exprimant présentent des défaillances motrices et une réduction de l'espérance de vie [126,156,158]. D'ailleurs, plusieurs études démontrent que les poly-GA, poly-PA et les poly-GP ne semblent pas comparativement toxiques chez la Drosophile, bien qu'une réduction modérée de la survie ait été observée lorsque poly-GA est exprimée [156,158,159]. Aussi, l'expression des DPR contenant de l'arginine (GR/PR), mais pas celles contenant de l'alanine (GA/PA), dans le *C. elegans* causent un arrêt du développement, alors que l'expression post-développemental cause une paralysie dépendante de l'âge [160]. Ensuite, l'expression de chaque DPR chez le poisson-zèbre a un effet délétère sur le développement, l'activité locomotrice ainsi que la survie. L'intensité des phénotypes corrèle avec la longueur des répétitions et est plus importante pour l'expression de poly-GR [161]. Enfin, la surexpression de poly-GA dans le SNC de souris génère des inclusions de poly-GA, de la neurodégénérescence ainsi que des phénotypes moteurs, cognitifs et comportementaux. Cependant, ces effets ne sont pas observés si la séquence de ERH est mutée pour être incapable de s'agréger et les inclusions phospho-TDP-43 sont presque absentes, indiquant que d'autres PRD sont les principaux responsables de l'accumulation de ces dernières [145]. Malgré ces preuves d'un fort potentiel neurotoxique, des études post-mortem suggèrent que les PRD ne soient pas les principales responsables de la pathologie C9(SLA/DFT) [126]. En effet, la pathologie TDP-43 et la neurodégénérescence surviennent conjointement dans les régions affectées du SNC dans la DFT ainsi que la SLA [141,160], mais la pathologie des PRD ne coïncide pas avec celle de TDP-43 et n'est généralement pas retrouvée dans les mêmes neurones que les inclusion TDP-43 [163,164,165]. De plus, les inclusions de PRD sont rares dans la moelle épinière des cas de C9(SLA) alors que la pathologie TDP-43 y est commune [163,166]. Aussi, les inclusions de DPR sont fréquentes à travers plusieurs régions du cerveau, dont plusieurs considérées comme minimalement affectées dans la SLA/DFT, dont le cervelet ainsi que le lobe occipital et pariétal [126,166]. Néanmoins, une étude récente révèle que,

contrairement aux autres DPR, les inclusions poly-AGR corrélient avec les zones de neurodégénérescence dans la C9(SLA) et colocalisent avec la pathologie de TDP-43 [167].

Plusieurs modèles murins impliquant les mécanismes de gain de fonction par l'expression des ERH de C9 ont aussi été générés [126]. De ceux-ci, trois modèles transgéniques utilisant un chromosome bactérien artificiel (CAB) ont des phénotypes similaires [168]. Deux utilisent des CAB contenant l'exon 1-5 du gène C9 et 300-500 répétitions, le troisième exprime le gène complet avec 100-1000 répétitions. Ils en développent des inclusions d'ARN et de PRD dans le SNC, mais aucun n'a démontré une mort neuronale, une diminution de la survie ou la formation d'inclusions TDP-43 [138,169,170]. Un des modèles développe des troubles de mémoire, une perte de neurones de l'hippocampe et une augmentation des niveaux de TDP-43 phosphorylées, mais n'est pas entièrement récapitulatif de la SLA ou la DFT. De plus, on observe une forte corrélation entre la formation d'inclusions d'ARN et la longueur des répétitions [138]. Un autre modèle de souris transgénique CAB montre un phénotype flagrant de neurodégénérescence. La souris exprime le gène C9 complet avec différentes longueurs de répétitions dans différentes lignées. Trois lignées indépendantes, deux expriment ~500 ERH et une expriment fortement 36 ERH, développent des inclusions d'ARN, des agrégats de PRD, la pathologie TDP-43 et de la neurodégénérescence. Un sous-type de femelle développe un phénotype moteur aiguë avec paralysie, perte de poids et une diminution de la survie. Les autres souris manifestant une neurodégénérescence plus progressive. Ceci illustre que même de courtes ERH fortement exprimées sont toxiques, démontrant que de larges ERH ne sont pas cruciales dans ce modèle de souris [171]. Enfin, un modèle surexprimant 66 ERH par virus adénoassocié présente des phénotypes moteurs ainsi que comportementaux accompagnés d'inclusion d'ARN, de PRD, de phospho-TDP-43 et une perte neuronale [172].

Malgré ces modèles corroborant l'implication des deux mécanismes de gain de fonction dans la neurotoxicité, il demeure fortement débattu si l'ARN des ERH ou les DPR en résultant sont les espèces toxiques dans l'étiopathogénèse de C9(SLA/DFT) ou, si elles agissent de manière synergique, lesquelles sont les plus toxiques. Ceci est en partie parce que l'expression des ERH produit à la fois les ARN répétés ainsi que les PRD, rendant l'attribution des mécanismes de toxicité en aval ardu. De plus, malgré leur succès, les modèles *in vivo* et *in vitro* générés de gain de fonction présentent généralement certaines limitations notables pouvant les empêcher

potentiellement de récapituler pleinement les mécanismes des pathologies C9(SLA/DFT). En effet, souvent les répétitions exprimées ne sont pas aussi longues que celles retrouvées chez les patients atteints, les modèles de surexpression ne reflètent pas l'expression endogène dans les patients et les modèles exprimant les ERH de manière exogène ne récapitulent pas la diminution significative de l'expression de C9 qui est caractéristique des patients C9(SLA/DFT) [126].

1.2.7.2.2 Mécanisme de perte de fonction

Conformément à une hypothèse de toxicité par perte de fonction de C9, les niveaux d'une ou plusieurs des variants de transcrits de *C9* sont réduits de manière significative dans les lymphocytes, des cellules souches pluripotentes induites (CSPi) dérivées de neurones, le cortex frontal, le cervelet, le cortex moteur et la moelle épinière cervicale des porteurs des ERH de C9 comparativement aux contrôles [23,95,105,153,173,142,175]. De ce fait, le niveau de la protéine C9 est réduit de manière importante dans le cortex frontal et le cervelet des patients [81,90]. Par ailleurs, les transcrits en amont des répétitions sont plus abondants relativement à ceux en aval des ERH dans les lymphocytes, les tissus du cerveau et de la moelle épinière des porteurs des ERH [174]. Supportant aussi un effet de perte de fonction, dans les motoneurones de patients affectés de C9-SLA, l'isoforme C9C est délocalisée de la membrane nucléaire à la membrane plasmique [81,90]. Plusieurs mécanismes ont été proposés pour expliquer cette diminution d'expression de *C9*, dont la régulation épigénétique et l'instabilité transcriptionnelle. En effet, 20-40% des tissus analysés de patient atteints de C9(SLA/DFT) démontrent une hyperméthylation des îlots CpG dans la région d'un promoteur de *C9*, en aval des répétitions pathogéniques qui est associé à la transcription des variants 1 et 3. [177]. Cette hyperméthylation est associée avec une augmentation de la longueur des répétitions et la diminution de la transcription de *C9*. De plus, la région promotrice dans les ERH ou en aval de celles-ci associée à la transcription du variant 2 de *C9*, contient aussi moins de sites d'initiation de la transcription pour ceux qui sont porteurs des ERH [178] et il a été démontré que les répétitions elles-mêmes sont méthylées chez les patients C9(SLA/DFT) [179]. Additionnellement, des modifications des histones ont été identifiées dans le SNC de patients C9(SLA/DFT) porteurs des ERH, soit une augmentation des marqueurs de triméthylation répriment la transcription (H3K9me3, H3K27me3, H3K79me3 et H4K20me3). Incidemment, la réduction de l'expression de *C9* corrèle plus fortement avec ces marqueurs de répression de la chromatine qu'avec la méthylation

de l'ADN, suggérant que la conformation de la chromatine est plus immédiatement importante dans la détermination de l'expression du gène [173,180]. En ce qui concerne l'instabilité transcriptionnelle, la nature polymorphique de la structure des ERH de C9 pourrait aussi directement interférer avec la processivité des polymérase, menant les polymérase d'ARN à bloquer, glisser ou relâcher le brin d'ADN durant l'initiation de la transcription ou l'élongation. Cette perturbation de la processivité des polymérase provoque une réduction des niveaux des transcrits et/ou à des transcrits interrompus ou tronqués [181,182,183]. Enfin, la transcription bidirectionnelle des séquences riches en GC, dont les ERH de C9 pourrait produire de courts transcrits interrompus pouvant agir comme de courts ARN interférant qui inhiberait la transcription du gène [183].

Les fonctions probables de C9, décrits à la **section 1.2.3**, dont une implication dans les processus comme l'autophagie, la fonction immunitaire et la protéostase, supportent un rôle de sa perte de fonction dans la neurodégénérescence par l'implication reconnue de ces dernières dans les maladies neurovégétatives comme la SLA et la DFT [126]. Les phénotypes des modèles cellulaires et animaux de KD/KO de C9 décrits à cette même section appuient également une telle implication. Cependant, plusieurs résultats mettent en doute que la perte de fonction de C9 seule soit suffisante pour provoquer la neurodégénérescence. En effet, la réduction du niveau de protéine C9 est décrite dans des régions affectées ainsi qu'in affectées de patients porteurs de la mutation de C9 et les patient C9(SLA) démontrant une réduction >50% de l'ARNm de C9 présentent des phénotypes similaires à des patient avec une expression plus élevée [90,184,185]. De plus, aucune mutation codante pouvant mener à une altération ou une perte de fonction n'a encore été identifiée chez les patients atteints de C9(SLA/DFT) [126,184]. Cette interprétation est appuyée aussi par la présence de seulement des troubles moteurs et/ou comportementaux modérés ainsi qu'une réduction de l'espérance de vie chez les souris KO pour C9, sans qu'on observe une mort neuronale, ce qui est contrasté par la présence de phénotypes moteurs et neurologiques majeurs chez les modèles KO de C9 de *C. elegans* ainsi que les KD chez le poisson zébré [102,103,104,105,106,107,108]. L'ensemble de ces résultats s'accordent cependant avec un scénario stipulant que différentes cellules et/ou sous-populations neuronales aient des vulnérabilités distinctes à la réduction des niveaux de C9, qui pourrait provoquer la

neurodégénérescence lorsque combinée avec des stress additionnels et/ou comme une conséquence de mécanismes coopératifs de gain ainsi que perte de fonction des expansions des répétitions G4C2 [98,95]. Les différences entre les modèles animaux s'expliquent potentiellement par des vulnérabilités distinctes des populations neuronales au KO/KD de *C9* et aux différentes sources de stress/toxicités entre espèces. Ainsi l'haploinsuffisance de *C9* serait insuffisante pour précipiter les pathologies SLA/DFT, mais elle rendrait les cellules neurales plus vulnérables aux stress d'origine génétique ou environnemental, pouvant mener à la neurodégénérescence caractérisant les pathologies associées aux ERH de *C9*.

1.3 Le poisson-zèbre

1.3.1 Avantages et utilisation en recherche

Les modèles d'invertébrés dont le *C. elegans* et la drosophile ont de nombreux avantages biologiques pour la recherche, notamment la disponibilité de multiples outils génétiques, un cycle de vie court, une reproduction rapide et prolifique, un entretien peu coûteux ainsi qu'une excellente caractérisation de leur biologie et génétique. De plus, les processus cellulaires découverts chez ces modèles sont conservés évolutivement chez l'humain, justifiant leur usage privilégié en recherche biomédicale. Cependant, l'absence d'organes ou structures similaires à l'humain, dont une moëlle épinière, et une diversité moindre de tissus différenciés, limitent la capacité translationnelle de certaines découvertes faites dans ces modèles [186]. Les modèles de mammifères, dont la souris, présentent l'avantage d'être plus rapprochés évolutivement des humains et ils démontrent ainsi une similarité accrue au niveau de leurs processus biologiques avec ces derniers. Néanmoins, les manipulations génétiques y sont souvent plus complexes, laborieuses ainsi qu'onéreuses, tant en temps qu'en ressources, et les résultats peuvent varier de manière significative d'un fond génétique à l'autre même pour des modifications de nature similaire comme le démontrent les différents modèles murins pour l'étude du KO et l'expression des ERH de *C9* [126, 187].

Le poisson-zèbre (*Danio rerio*), décrit pour la première fois en 1822 par le Dr. Francis Buchanan-Hamilton, se présente alors comme un modèle intermédiaire intéressant entre les invertébrés et les rongeurs, [188]. En effet, ce poisson à l'apparence anodine, natif de l'inde,

membre de l'infra-classe des téléostéens et de la famille des Cyprinidés, a de nombreux avantages pour la recherche biomédicale. Effectivement, il se prête remarquablement à une gamme constamment croissante d'outils et techniques du génie génétique permettant des approches de génétique classique et inverse. De plus, diverses caractéristiques biologiques du poisson-zèbre rendent ces outils et techniques particulièrement puissants lorsque appliqués à ce modèle. Notamment, les poissons zébrés adultes, ne mesurant qu'environ 4-5 centimètres, peuvent être maintenus et élevés dans un espace restreint à des frais moindres. Également, une femelle peut produire 100 à 200 œufs par semaine, chacun d'un diamètre d'environ 1 millimètre, donnant lieu à un nombre important de spécimens avec un développement synchronisé. En sus, les œufs se développent rapidement *ex utero*, menant à une éclosion d'embryons transparents en moins de 72 heures sous conditions normales. Ensuite, leur transparence embryonnaire temporaire, qui peut être prolongée artificiellement, conjuguée au développement d'une vaste gamme de marqueurs fluorescents et de lignées transgéniques les exprimant, permettent l'identification de sous-populations cellulaires, rendant aisée l'observation *in vivo* de la physiologie, du comportement et de l'expression protéique chez les embryons [189, 190]. Il est donc possible de corréler des caractéristiques anatomiques et comportementales entre elles ainsi qu'avec l'expression protéique. La génération de plusieurs lignées sans pigmentation permet même d'étendre en partie cet avantage aux poissons adultes. Ainsi, le poisson-zèbre se porte efficacement et avec une aisance relative à plusieurs techniques d'imagerie moléculaire et immunohistochimique [189,191]. Ces avantages, qui récapitulent largement ceux des invertébrés, sont en plus complétés par une similarité considérable des structures et organes des poissons zébrés à ceux des mammifères, quoique parfois moins élaborés que leurs homologues. Qui plus est, les processus biologiques, et les voies génétiques du poisson-zèbre sont similaires à ceux des mammifères. En effet, la morphologie du cerveau des poissons-zèbres est remarquablement similaire à celle des rongeurs, incluant la macro-organisation générale et la morphologie cellulaire. De plus, les poissons zèbres présentent tous les systèmes neuromédiateurs majeurs retrouvés chez les humains et les rongeurs ainsi qu'un système endocrinien généralement homologue à ces derniers. Qui plus est, les 25 paires de chromosomes du poisson-zèbre contenant approximativement 26 000 gènes sont similaires en nombre de chromosomes ainsi que de gènes aux humains et présentent des orthologues évidents pour environ 70% des gènes de ces derniers. Additionnellement et à l'instar des mammifères, les

embryons exprimant des mutations embryo-létales peuvent être étudiés quelques jours après fertilisation sans avoir à se préoccuper de défaut d'implantation utérine ou de survie post-éclosion et les poissons zébrés se portent assez aisément aux criblages à haut débit de petites molécules [189, 192].

En plus de ces avantages, le poisson-zèbre a une histoire longue en tant que modèle pour la recherche. En effet, son usage comme modèle du développement et de la recherche embryologique remonte au moins aux années 1930 [193]. De plus, les années 1940 à 1970 voient plusieurs groupes de recherches utiliser le poisson-zèbre et démontrer son potentiel pour la recherche développementale, comportementale, cognitive, toxicologique, neurobiologique et oncologique [189, 194,195,196,197]. Néanmoins, c'est le développement de techniques de génétique moléculaire, dont le clonage [198], la mutagenèse [199], la transgénèse [200] et la cartographie génétique [201] des années 1980 qui permettent au poisson-zèbre d'être véritablement reconnu comme un modèle animal viable et au potentiel considérable pour la recherche. Son utilisation est maintenant reconnue dans de nombreux domaines de recherche biomédicale tel l'oncologie [202], l'immunologie [203], la biologie développementale des vertébrés [204] incluant le neurodéveloppement [205], la régénération [206], le vieillissement [207], la biologie cellulaire [208], la neurobiologie [209], les maladies psychiatriques [210] et neurodégénératives [211].

Enfin, pour l'étude de gènes comme *C9* associés à des maladies à forte composante motrice tel que la SLA, le poisson zèbre présente l'avantage notable de posséder un système neuromusculaire simplifié, mais similaire à celui des mammifères. L'embryon du poisson-zèbre développe rapidement un circuit moteur fonctionnel et son comportement moteur est extrêmement stéréotypé [212]. Ceci permet d'observer aisément ainsi qu'en un délai très court, soit en moins de 72 heures et possiblement seulement 24 heures, l'influence d'une perte, réduction ou augmentation de l'expression d'une protéine qui pourraient potentiellement avoir un effet délétère sur le système neuromusculaire du poisson-zèbre.

1.3.2 *C9orf72* chez le poisson-zèbre

Plusieurs caractéristiques génétiques du poisson-zèbre rendent possible et attrayant son usage pour l'étude du gène *C9*. En effet, le poisson-zèbre a un orthologue de *C9*, le gène *zgc:100846* (GCF_000002035.6) localisé au chromosome 13 (NC_007124.7) [83]. Ce seul orthologue de *C9*, référé subséquemment comme *zfC9* alors que le gène humain sera *HC9*, démontre une importante homologie de séquence (76%) [83] avec le gène humain et la séquence d'acides aminés totale de sa protéine (85%) incluant le domaine fonctionnel DENN (73%) (**Figure 3 et 4**). De plus, l'organisation de la protéine *zfC9* et *HC9* est quasi identique hormis un décalage de 10 aa (**Figure 2, 3 et 4**)



Figure 3. Aligment des séquences d'acides aminés des protéines HC9 et zfC9. Aligment peptidique de la protéine *C9orf72* humaine et du poisson-zèbre démontrant une forte homologie (85% total, 74% dans le domaine DENN.) Le rouge indique une position pour laquelle un résidu unique est pleinement conservé. Le bleu indique un changement de résidu, mais que les deux acides aminés ont des propriétés fortement similaires (>0.5 dans la matrice de Gonnet PAM 250). Le vert indique une substitution de résidu et qu'il y a une faible similarité entre les deux acides aminés. Le noir indique une substitution de résidu sans similarité significative. Adapté de Ciura, S., et al. (2013) [105].

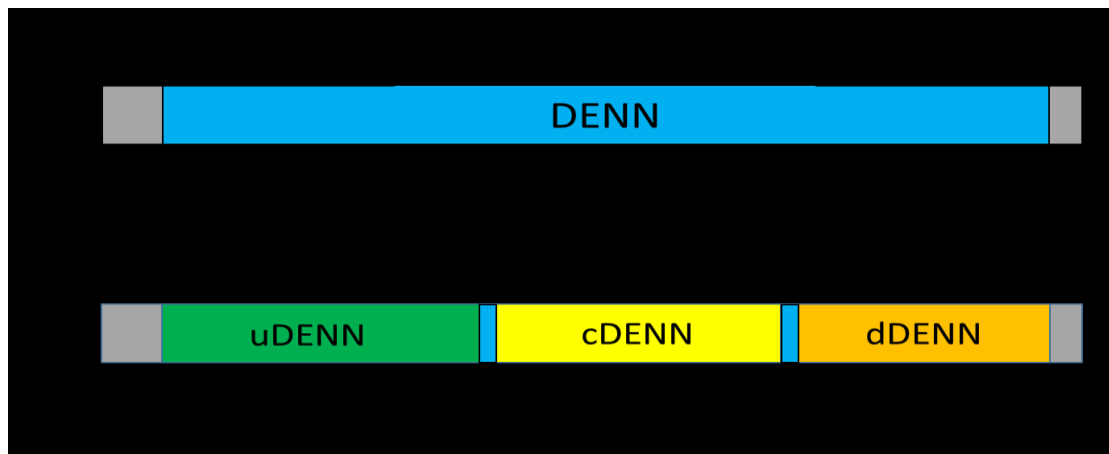


Figure 4. Structure de la protéine zfc9. (A) Étendue du domaine fonctionnel DENN de zfc9 pour zfc9La du 23^{ième} acide aminé 464^{ième}. (B) Subdivision en trois sous domaine encore parfois utilisé aujourd’hui (23-200 uDENN, 200-343 cDENN, 343-464 dDENN). Adapté des informations de la banque de données en ligne <https://prosite.expasy.org/>; (Décembre ,2018).

Comme pour l’humain, l’épissage alternatif de zfc9 est prédit de produire 3 variants de transcrits différents codants : variant 1 (2427 pb), variant 2 (2152 pb) et variant 3 (2389 pb) [79,105] (**figure 5**). Cependant, leur traduction résulte en deux isoformes longues (zfc9La, 462 aa; zfc9Lb 460 aa) très similaires et une isoforme courte (zfc9C, 326 aa) (**figure 5**). En l’absence d’un anticorps efficace dédié disponible en date de ce projet, les profils d’expression des isoformes de zfc9 demeurent méconnus et l’expression de chacune des isoformes reste à être confirmée expérimentalement. Néanmoins, l’expression de l’ARNm de *zfc9* a été confirmée par hybridation *in situ*. Elle est enrichie dans le système nerveux en développement et faiblement présente dans la moelle épinière des embryons et des larves de poissons zébrés. De plus, *zfc9* est exprimée dans la plupart des organes et tissus au stade adulte avec une expression notable dans le cerveau antérieur, médial, postérieur et dans la moelle épinière [105]. Des modèles de KD de *zfc9* ont été générés par l’usage d’oligonucléotides morpholinos antisens (OMA) différents, ciblant soit l’ATG commun, les sites spécifiques d’initiation [105] ou les régions en amont de ces sites d’initiation des transcrits codant de *zfc9* [213]. Ces modèles révèlent des déficiences motrices au niveau de l’enroulement spontané ainsi que la nage spontanée ou provoquée par stimulus tactile. Les déficiences motrices sont accompagnées d’une

axonopathie des motoneurones, caractérisée par des processus axonaux écourtés qui ne projettent pas complètement dans le myotome ventral et une arborisation perturbée résultant en des embranchements surnuméraires [105, 213]. De plus, le KD de *zfc9* affecte la formation des axones des neurones du rhombencéphale résultant en la perte d'organisation et de densité des structures à forme d'échelle des commissures. Enfin, le KD de *zfc9* entraîne une augmentation du signal activé d'apoptose caspase-3 dans le SNC, indiquant une apoptose neuronale accrue et ce particulièrement dans le proencéphale. De plus, l'injection de constructions artificielles de *zfc9* présentent des délétions de segments du motif DENN, récapitulant les phénotypes des OMA par un effet dominant négatif [213]. Dans tous les cas, les phénotypes sont secourus par l'injection d'ARN codant (ARNc) de *zfc9* et de *C9* humain. Ces KD sont transitoires et les phénotypes sont observés chez des embryons vieux de 2 à 4 jours [213,125]. Cependant, aucun modèle de KO stable de *zfc9* n'a encore été décrit dans la littérature, bien qu'un rapport préliminaire évoque qu'un KD de *zfc9* ne récapitulerait pas les phénotypes moteurs et axonaux des KD ([214]; Schmid, Hruscha, Haass, nonpublié). Ceci pourrait s'expliquer par l'une des controverses associées à l'utilisation oligonucléotides morpholinos (OM), soit la tendance aux phénotypes induits par les MO de ne pas être récapitulés chez les mutants correspondants, possiblement en raison du sauvetage du mutant zygotique par l'ARNm maternel dont la transcription est bloquée par les OM [215], des effets hors-cible des OM [216] ou une compensation génétique présente chez les mutants absente chez les poissons injectés avec les OM [217]. De ces doutes, la génération d'une lignée *zfc9* KO stable puis sa caractérisation semble donc être utile, voir même nécessaire.

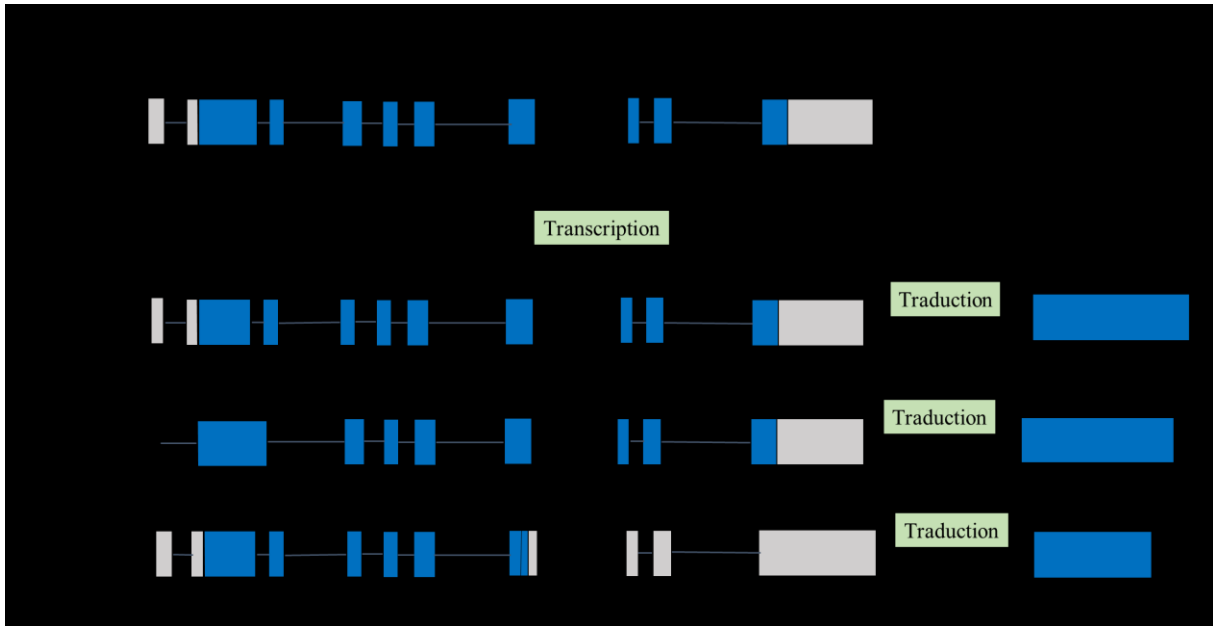


Figure 5. Schéma de la séquence ADN du gène *zfc9* et des trois transcrits alternatifs produisant 3 isoformes de la protéine. En gris les exons non codants, en bleu les exons codants et la protéine résultante. Variant 1 (Nom : zgc : 100846-001, ID : ENSDART 00000015127.6), variant 2 (nom : zgc : 100846-201, ID : ENSDART 00000125180.2), variant 3 (Nom : zgc : 100846-202, ID : ENSDART 00000126136.2). Assemblage génomique du poisson-zèbre (GRCz10).

1.3.3 Transgénèse stable chez le poisson-zèbre

L'utilisation de techniques de transgénèse transitoire dont la micro-injection d'ARNm et de plasmides pour exprimer des protéines exogènes ou l'injection d'OMA, de microARN ainsi que de petit ARN interférant pour contrôler l'expression est certainement possible et efficace chez les poissons zébrés. Cependant, la transgénèse transitoire comporte certaines limitations notables. En effet, les constructions injectées sont dégradées rapidement, ne permettant une observation du phénotype qu'au début du développement du poisson-zèbre. De plus, une quantité importante d'embryons doit être injectée à plusieurs reprises, par un ou plusieurs injecteurs, causant un risque de variation dans la quantité de construction injectée ainsi que la qualité de l'injection d'un groupe et d'un embryon à l'autre, due à des variations expérimentales de l'injecteur. Ceci en plus des phénomènes de compensation, de la toxicité inhérente des constructions injectées et des effets hors cibles possibles, rendent le développement de lignées

transgéniques stables une priorité, voir une nécessité pour l'étude d'un gène associé à une maladie comme la SLA dont les premiers symptômes se manifestent chez l'adulte [189]. Pour créer des transgènes stables, la construction correspondante doit être intégrée dans le génome du poisson-zèbre. À cette fin, plusieurs outils de transgénèse initialement disponibles chez d'autres modèles ont été adaptés avec succès chez le poisson-zèbre, dont plusieurs systèmes de transposon incluant "sleeping beauty", Tol1 et Tol2. En raison de sa simplicité d'utilisation ne nécessitant pas de structures ou de constructions préexistantes, sa fonction bien caractérisée et son efficacité à induire l'intégration génomique, Tol2 est le système d'intégration transgénique utilisé pour ce projet ainsi que celui le plus communément employé chez le poisson-zèbre [189,218].

Le transposon Tol2 a été identifié dans le génome du poisson Medaka en 1995 [219], et peut s'insérer dans le génome d'un hôte de la même espèce [220] ainsi que dans le génome d'autres vertébrés, dont notamment le poisson-zèbre [221]. En effet, les transposons sont des fragments d'ADN capables de se mouvoir d'un locus à l'autre, que ce soit dans leur génome d'origine ou un autre génome. Ces fragments d'ADN sont composés d'une séquence codant pour une transposase, flanquée à chaque extrémité par des séquences inversées répétées (SIR). Ces SIR sont nécessaires pour la réinsertion de la séquence dans le génome alors que l'enzyme transposase est responsable de catalyser sa propre transposition. La transposase peut aussi agir en trans et donc sur toutes séquences flanquées par les SIR [222]. L'élément Tol2 a une activité autonome contenant un gène codant une transposase complète et fonctionnelle capable d'identifier, exciser et réinsérer l'élément d'ADN défini par ses SIR ou d'autres éléments avec les mêmes SIR. Dans la méthode Tol2, un plasmide donneur-transposon et de l'ARNm synthétique codant pour sa transposase sont introduits par micro-injection dans des œufs le plus tôt possible après leur ponte et fertilisation. Les embryons de stade 1 à 4 cellules sont fréquemment utilisés pour toutes injections puisque à ce stade il n'y a pas de membranes séparant les cellules et le vitellus, permettant aux solutions injectées dans l'une des cellules ou le vitellus de se propager uniformément dans l'organisme [223,224]. La protéine transposase est alors traduite de l'ARNm injecté et catalyse l'excision de la construction transposon du plasmide donneur, menant à l'intégration stable et non-dirigé dans le génome de l'ADN excisé une fois la transposase dégradée [222]. En moyenne, une progéniture F1 transgénique est obtenue de

50% à 70% des poissons injectés, à une fréquence de 3% à 100% [223,225]. De plus, en associant le transgène d'intérêt avec divers promoteurs inductifs spécifiques, particulièrement en conjonction avec le système Gal4-UAS, il est possible de contrôler l'expression du gène tant spatialement (p.ex : le promoteur Hb9 exprimé spécifiquement dans les motoneurons et certains interneurons, le promoteur cardiaque *cmlc2*) que temporellement (p.ex le promoteur *hsp70* induit par différents stress dont le choc thermique) [226].

L'expression de transgènes est un outil puissant pour la modélisation génétique et pathologique, mais la modification du gène endogène, lorsque possible, est généralement préférable que ce soit pour la modélisation ou la recherche fondamentale. Ainsi, l'approche favorisée est la mutagenèse dirigée à un site spécifique, permettant entre autres la délétion et l'insertion d'éléments génétiques, pour laquelle plusieurs techniques sont aussi disponibles chez le poisson-zèbre. La première de ces techniques à être adaptée au poisson-zèbre a été les nucléases à doigt de zinc (ZFNs), suivie des nucléases effectrices analogues à un activateur de transcription (TALENs) et enfin plus récemment le système CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat) /Cas9 (CRISPR associated protein 9). Ce dernier est le plus favorisé en général ainsi que pour ce projet en raison de la facilité à programmer son domaine de liaison à l'ADN, son efficacité, sa précision et l'aisance relative de son emploi lorsque comparé aux ZFNs ou aux TALENs qui nécessitent comparativement un temps et une expertise considérables pour être utilisées efficacement [218,227].

Le système CRISPR/Cas9 est formé d'une famille diverse d'endonucléases retrouvées dans les archées et les génomes bactériens qui a été adaptée pour l'édition des génomes d'eucaryotes depuis 2013 [228,229]. Ce système impliqué dans la réponse immunitaire adaptative des procaryotes emploie des ARN non-codants pour guider la nucléase Cas9, afin d'induire une coupure double brin (CDB) de l'ADN à un site spécifique. En effet, les systèmes CRISPR adaptés pour l'édition du génome ont deux composantes : un ARN de guidage (ARNg) et une endonucléase Cas9. L'ARNg est un ARN court composé d'une séquence d'échafaud nécessaire pour la liaison de la protéine CAS et un espaceur de ~20 nucléotides déterminé par l'expérimentateur qui définit la cible génomique à être modifiée. Cependant, la protéine Cas9 possède un domaine qui lie de courte séquence d'ADN de 2-6 nucléotides nommé PAM

(Protospacer adjacent motif) et l'endonucléase clive la séquence définie par l'espaceur de l'ARNg seulement si une séquence PAM correspondante est adjacente à cette dernière. Ainsi, par la spécificité des 20 nucléotides de l'ARNg en plus de la spécificité des 2-6 nucléotides fournie par le site PAM, l'endonucléase est assez spécifique pour cliver le génome d'un métazoaire à un site potentiellement unique. CRIPR/Cas9 permet donc de cibler n'importe quel site d'un génome en modifiant simplement l'espaceur de l'ARNg tant qu'il est flanqué d'une séquence PAM.

La génération de CDB dans l'ADN génomique étant toxique, ils sont rapidement réparés par les mécanismes endogènes de réparation des CDB. Un de ces systèmes est la voie de réparation par jonction d'extrémités non homologues (JENH) qui entraîne de petites insertions ou délétions au site de CDB, les allèles résultantes sont généralement des allèles KO prédites de mener à une ablation de la fonction du gène permettant l'utilisation du système CRIPR/Cas9 pour générer des KO ciblé. Un second mécanisme de réparation moins fréquent des CBD est la réparation dirigée par homologie (RDH) qui emploie une matrice d'ADN pour guider la réparation et qui peut être fournie par l'expérimentateur ou la chromatide sœur pour générer un allèle de la matrice d'ADN. En générant des matrices d'ADN avec des bras d'homologie flanquant une séquence non-homologue, il est ainsi possible d'employer le système CRIPR/Cas9 et la RDH pour effectuer une insertion de matériel génétique spécifique à un site ciblé; on nomme une telle insertion un knock-in (KI). Chez le poisson-zèbre, l'ARNg est directement injecté dans le zygote avec de l'ARNm de la protéine Cas9 avec une séquence de localisation nucléaire, les KI nécessitant simplement l'ajout d'un ARNg à simple brin ou dans un plasmide au mélange injecté pour servir de matrice d'ADN. Des réparations de CDB à faible fidélité surviennent alors au site ciblé dans chaque cellule diploïde. Ceci résulte en la génération de plusieurs allèles distinctes dans un même individu injecté, menant à des animaux F0 adultes hautement mosaïques. Les adultes mosaïques peuvent alors être croisés avec une lignée de type sauvage (croisement dit de type *outcross*) pour une progéniture, uniforme au niveau des allèles des cellules somatiques, contenant des hétérozygotes. Ces hétérozygotes peuvent alors être croisés entre individus de même génotypes (croisement de type *incross*) de sorte à obtenir des mutants homozygotes [230].

1.4 Objectifs

Malgré de récentes avancées considérables et en dépit de la notoriété que lui vaut son rôle comme l'un des gènes majeurs associés à l'étiologie de la SLA, notre compréhension de la fonction et du profil d'expression normale de C9 demeure encore incomplète. En l'absence d'anticorps dédiés pour la détection de zfC9, nous avons donc cherché à générer des lignées de poissons-zèbres génétiquement modifiées par CRISPR/CAS9, de sorte à marquer le N ou le C-terminal du gène *zfC9* avec des tags moléculaires composés de 12 histidines (12xHis) et de la protéine fluorescente verte améliorée (eGFP,) respectivement, permettant notamment l'étude de sa protéine au niveau de son profil d'expression, sa localisation et ses interactions protéiques autant in vivo qu'in vitro. Additionnellement, des anticorps efficaces existant pour le C9 humain, nous avons cherché à produire par système Tol2 deux lignées transgéniques de poissons zébrés exprimant soit le gène *C9* humain natif ou fusionné à un tag moléculaire eGFP. Ces lignées permettront de générer des modèles de surexpression stable de *C9* et d'étudier la localisation cellulaire de *C9* in vivo et in vitro. Des modèles transitoires de KD de *zfC9* par AMO montrant des phénotypes moteurs ainsi que neurologiques, nous avons aussi cherché à générer par CRISPR/CAS9 une première lignée knock-out stable de *zfC9*. Sa caractérisation nous permettant de décrire un nouveau modèle de perte de fonction de C9 pour la SLA, vérifier les observations des modèles knockdown transitoire de *zfC9* et déterminer si l'ablation persistante de *zfC9* se traduit en d'autres phénotypes au cours du cycle de vie des poissons, nous informant possiblement sur la fonction biologique normale de C9.

Les objectifs de ce projet de maîtrise étant donc :

- 1) Générer des lignées de poissons zébrés génétiquement modifiées par CRISPR-CAS9, de sorte à marquer avec des tags moléculaires le N et le C-terminal du gène *zfC9*.
- 2) Produire par système Tol2 deux lignées transgéniques poissons-zèbres exprimant soit le gène *C9* humain natif ou fusionné à un tag moléculaire eGFP.
- 3) Concevoir par CRISPR/CAS9 une lignée knock-out stable de *zfC9*.

2. Résultats

2.1 Observation du patron de l'expression de *zfC9* par hybridation *in situ*

Puisque la génération de nos lignées marquées pour C9 permettant son observation ayant été prévue au début de ce projet nécessiterait plusieurs mois dans les meilleures circonstances, nous avons entrepris en attendant d'observer le profil d'expression du gène chez le poisson-zèbre dans des stades de développement précoce. Ceci nous permettant d'anticiper le profil d'expression de *zfC9* attendu pour ces lignées et de guider nos observations une fois la lignée KO obtenue. Le profil d'expression temporel et spatial d'une protéine est un outil précieux dans l'étude de leur fonction biologique normale et de leur implication dans une pathologie à laquelle une de leur mutation a été associée. La connaissance d'un tel profil est aussi fort utile pour guider les expériences avec des modèles de perte de fonction ou d'expression transgénique d'homologues de la protéine, en plus de pouvoir guider l'interprétation les résultats obtenus. Pour y arriver, nous avons employé la technique d'hybridation *in situ* (HIS) qui permet la détection et la localisation avec une grande sensibilité de séquences monobrin d'ADN ou d'ARN à l'aide de l'hybridation spécifique de sondes marquées complémentaires au monobrin ciblé. Cette technique peut donc être utilisée pour détecter l'expression d'un ARNm d'un gène spécifique dans des cellules, des sections de tissus, des tissus complet et, d'un intérêt particulier pour ce projet, dans des embryons de poisson-zèbre complets fixés puis perméabilisés [231].

2.1.1 Sondes et type de marquage pour l'hybridation *in situ*

Le type de marquage des sondes employées a été un marquage à la dioxygénine (DIG). Sa détection est possible par immunohistochimie à l'aide d'un anticorps liant spécifiquement DIG, lui-même lié à une enzyme (phosphatase alcaline) qui, exposée à une solution contenant un substrat chromatique (BM purple) de l'enzyme, provoque une réaction qui induit une coloration (mauve/violet) permanente et spécifique au site de liaison de l'anticorps. Cette approche a été utilisée puisque la sensibilité de ce marquage avait déjà été éprouvée préalablement par d'autres membres du laboratoire et l'expertise ainsi que le matériel pour ce marquage étaient déjà disponibles en plus d'être aisément accessibles. Quatre sondes antisens d'ARN simple brin marquées à la DIG de tailles différentes ciblant des régions distinctes de l'ADN codant (ADNc) de *zfC9*, retrouvées dans l'ARN messenger (ARNm) correspondant, ont été élaborées pour la

détection par HIS de l'ARNm résultant de l'expression du gène *zfc9*. Les sondes d'ARN pour l'HIS devant être d'une longueur minimale de 200 pb, l'ensemble des 4 sondes sélectionnées étaient prévues de détecter de manière non spécifique les trois transcrits d'ARNm prédit de *zfc9* (**tableau 1**). De celles-ci, les sondes *zfc9HIS1* et *zfc9HIS4* ont produit un marquage de force adéquate pour l'observation de l'expression de *zfc9*, suite à une HIS test sur un nombre d'embryons limité. Cependant, l'intensité du signal plus élevé pour la sonde *zfc9HIS4* a mené à son utilisation pour les véritables expériences d'HIS.

Tableau I. **Sondes marquées à la DIG générées pour la détection de l'ARNm de *zfc9* par hybridation *in situ*.** Tableau récapitulatif des paramètres des sondes d'ARN antisens générées pour la détection de l'ARNm de *zfc9* par hybridation *in situ*. Taille indique la longueur de la sonde en paire de bases (pb), Exon ciblé indique le numéro de l'exon ciblé en utilisant le transcrit *zfc9La* comme référence. La force du signal est l'évaluation qualitative de la force du signal relative après avoir testé chaque sonde lors d'HIS sur des embryons 24 heures post fertilisation (hpf)

Sonde	Taille (pb)	Exon ciblé	Force du signal
<i>zfc9-HIS1</i>	437	2	Forte
<i>zfc9-HIS2</i>	469	10	Très faible
<i>zfc9-HIS3</i>	836	10	Faible
<i>zfc9-HIS4</i>	793	10	Très Forte

2.1.2 Stades des embryons fixés et montés pour l'hybridation *in situ*

Des embryons de type sauvage (WT) de 24, 36 et 48 heures post fertilisation (hpf) ont été sélectionnés pour la détection de l'expression du gène *zfc9*. Ces stades étant choisis puisque, durant cette période le cerveau du poisson-zèbre est différencié en 5 lobes distincts, la morphologie de ces lobes se précise progressivement, la notochorde est bien développée, le système nerveux est encore en développement et par 48hpf l'ensemble des organes principaux du poisson-zèbres sont formés bien qu'encore à un stade immature [232]. De plus, lorsque Ciura et al [105] ont observé l'expression du gène *zfc9* par *in situ*, ils ont utilisé des embryons de 24 hpf, nous permettant de vérifier leurs observations et passé 48hpf, les sondes pénètrent difficilement dans les tissus profond, rendant difficile les tests sur des embryons complets [232].

Les embryons ont été préparés pour l'HIS en les retirant de leur chorion puis en les fixant par traitement au paraformaldéhyde avant de les déshydrater dans le méthanol pour être conservés à -20 °C en attendant leur usage pour l'HIS.

2.1.3 Profil d'expression du gène *zfC9* à 24, 36 et 48 hpf

Les profils d'expression du gène *zfC9* à 24, 36 et 48 hpf ont été observés à l'aide d'un microscope confocale Quorum Technologies sous éclairage normal. Pour l'ensemble des trois stades, on a observé une expression élevée dans les yeux et l'encéphale, dont au niveau du proencéphale, du mésencéphale et du rhombencéphale (**Figure 6**). Une expression forte du gène a aussi été observée dans la moelle épinière au niveau de ce qui apparaît être la notochorde au stade 24hpf (**Figure 6.A**). Cette expression était présente aussi aux stades 36 et 48 hpf, mais elle était moins prononcée (**Figure 6.B, C**). À 24 hpf, l'expression dans les tissus ne faisant pas partie du CNS, l'œil en raison de la rétine est ici inclus dans les tissus du CNS, était soit absente ou faible au point qu'elle devient difficilement distinguable du bruit de fond (**Figure 6.A**) alors qu'une expression diffuse commence à être observable dans ces tissus au stade 36hpf (**Figure 6.B**) et devient plus prononcée bien que toujours diffuse au stade 48hpf (**Figure C**). Ceci était observé notamment au niveau des tissus des muscles autour de la notochorde. Au stade 48hpf, certaines structures bien définies expriment fortement *zf^{C9}* pouvaient être identifiées dans le CNS, dont le nerf optique, le bulbe olfactif, le ganglion trigéminal et le circuit habenulaire (**Figure 6C**).

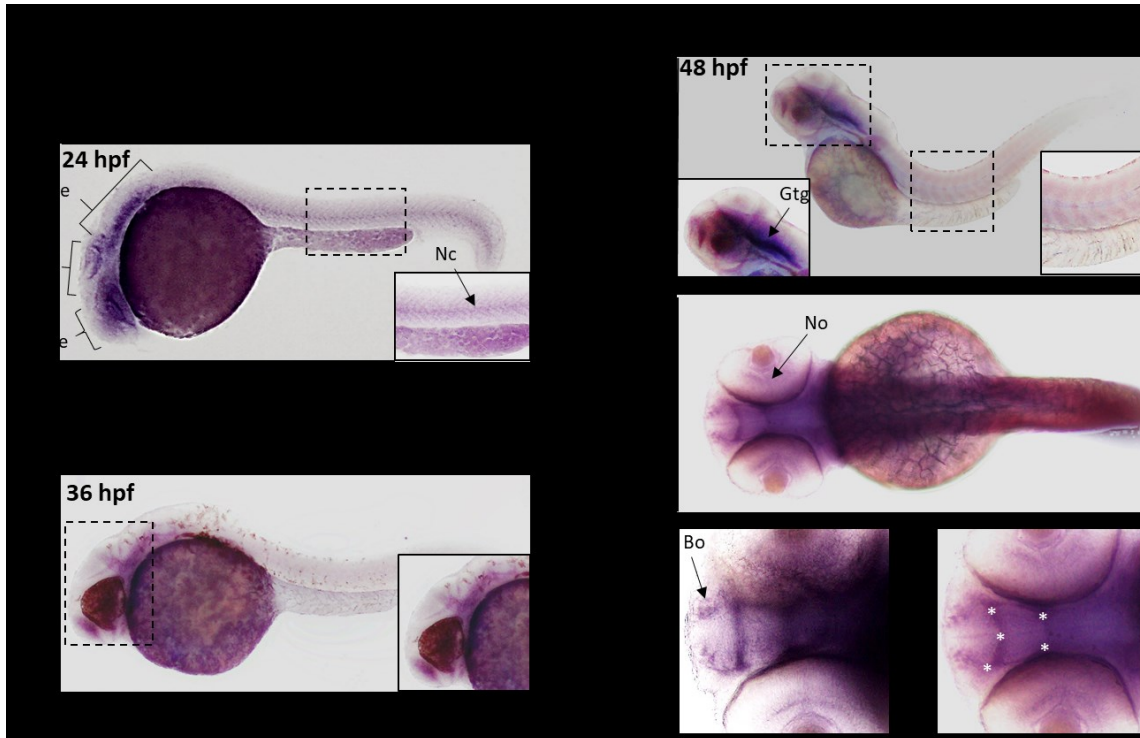


Figure 6. Observation du profil d'expression du gène *zfc9* par hybridation *in situ* chez des embryons de stade 24, 36 et 48 hpf. (A) Une vue latérale d'embryons 24hpf révèle une expression forte de *zfc9* dans les yeux et l'encéphale (Pe : proencéphale, Mes : mésencéphale, Re : (rhombencéphale). Son expression est importante aussi dans la moelle épinière (Nc : notochorde). Dans les tissus ne faisant pas partie de SNC, le signal est faible jusqu'à être indistinguable du bruit de fond. (B) Une vue latérale d'embryons 36hpf présente une expression forte dans les yeux ainsi que les régions de l'encéphale et on observe une expression plus modérée dans la moelle épinière au niveau de la notochorde. Une expression diffuse, mais distincte du bruit de fond est observable dans le reste des tissus. (C) Des vues ventrale et dorsale d'embryons au stade 48hpf révèlent une forte expression du gène dans l'encéphale, incluant dans plusieurs structures identifiables (No : nerf optique, Bo : bulbe olfactif, Gtg : ganglion trigéminal, * blanche : circuit habenulaire). Une expression modérément élevée de *zfc9* est aussi observée dans la moelle épinière au niveau de la notochorde et une expression modérée ainsi que diffuse du gène est observée dans le reste des tissus. L'échelle (barres noires) est de 100 μ m.

2.2 Tentative de génération de lignées avec un tag dans *zfc9*

Afin de rendre réalisable des tests par immunologie et/ou imagerie permettant notamment de caractériser la localisation, l'expression ainsi que l'interactome de la protéine C9 du poisson-zèbre (*zfc9*), il était prévu, en l'absence d'un anticorps dédié et efficace chez le poisson-zèbre,

de générer à l'aide du système CRISPR/Cas9 des lignées pour lesquelles zfC9 endogène serait marqué à l'aide de tags moléculaires. Les tags sélectionnés étant la "enhanced green fluorescent protein" (eGFP) et une étiquette de 36 paires de bases (pb) codant pour 12 histidines (12xHis). De plus, initialement il avait été proposé de générer quatre lignées différentes, une pour chaque tag inséré à l'extrémité N ou C-terminal (Nt, Ct). Cependant, comme décrit plus bas, il ne nous a pas été possible de tenter une insertion directement au Nt ou au Ct de C9. Néanmoins, afin de faciliter la distinction ainsi que le repérage pour les étiquettes, mutations et les lignées associées qui sont mentionnées dans cet ouvrage, la terminologie N-terminal et C-terminal sera utilisée pour référer à un positionnement qui est plus rapproché de l'extrémité correspondante, plutôt que littéralement à cette l'extrémité comme le veut l'usage normal de cette terminologie. L'approche à quatre voies a été motivée par le peu d'informations disponibles sur le domaine fonctionnel et la structure de C9 humain (HC9) et de zfC9, une stratégie à plusieurs approches nous permettant d'augmenter nos chances d'obtenir au moins une lignée marquée pour zfC9 pour laquelle la fonction, la stabilité et la localisation de la protéine marquée ne seraient peut-être pas compromises. Le tag eGFP a été sélectionné puisque les protéines de fusion l'incorporant sont généralement non toxiques pour les cellules, il permet l'observation simple in vivo de l'expression de ses protéines de fusions, sa détection ne nécessite pas de fixation ou de perméabilisation des embryons ainsi que des cellules et plusieurs outils moléculaires abordables pouvant le détecter de manière efficace sont disponibles [233]. De plus, sa fluorescence permet un criblage simple et rapide des poissons l'ayant intégré. Il comporte néanmoins des désavantages qui ont motivés l'usage du second tag 12xHis. Notamment, la taille considérable de eGFP, avec une séquence ADN de 717 paires de base résultant en une protéine de 238 acides aminés (aa) de 26.9 kDa, peut potentiellement perturber, altérer ou abolir la fonction native de la protéine de fusion d'intérêt. Additionnellement, sa taille importante peut affecter négativement le taux d'insertion par RDH puisque ces dernières tendent à décroître en efficacité plus la taille de l'insert est importante [233, 234]. Ainsi, bien qu'un tag 12xHis ne possède pas l'avantage de permettre une présélection par fluorescence ou d'effectuer de l'imagerie par fluorescence in vivo, sa petite taille le rend relativement peu obstrusif lui conférant une efficacité d'insertion vraisemblablement supérieur au eGFP ainsi qu'un risque moindre d'affecter la fonction de la protéine. De plus, les tags poly-His sont aussi robustement liés/détectés par une vaste gamme d'outils moléculaires accessible, abordables et éprouvés

[235]. La charge du travail de cet aspect du projet a initialement été partagé entre le présent rédacteur et une étudiante au postdoctorat du même laboratoire jusqu'à un peu après le début du criblage des premières cohortes injectés.

2.2.1 Constructions génétiques et micro-injections

Au terme de ce projet, il a été tenté de générer 3 des 4 lignées initialement envisagées. La première, C9orf72^{NHis}, aurait intégré le 12xHis tag au Nt de *zfC9* près du codon d'initiation commun trois variants de *zfC9*. La seconde lignée, C9orf72^{NGFP}, devait présenter l'intégration d'une étiquette composée eGFP au Nt de *zfC9* près de ce même l'ATG. Enfin, C9orf72^{CGFP} aurait aussi incorporer un tag eGFP dans *zfC9*, mais cette fois-ci au Ct près des codons d'arrêts des deux transcrits long de *zfC9* (*zfC9La*, *zfC9Lb*). La conception des ARNg s'est faite à l'aide d'algorithmes disponibles en ligne. Les ARNg, avec le score prédit le plus élevé et avec une plus grande proximité à l'ATG pour le Nt et au codon stop pour le Ct, ont été favorisés afin de permettre des insertions dans le cadre de lecture de *zfC9* avec un risque moindre d'affecter sa structure et sa fonction. L'objectif initial était d'éviter de perturber le domaine fonctionnel DENN de *zfC9*. Cependant, peu d'ARNg avec un score favorable étaient prédits à proximité de l'ATG ou du codon d'arrêt lorsque les banques de données ont été consultées, alors que l'assemblage génomique GRCz10 était encore utilisé (2017), nous poussant à considérer des ARNg plus loin en aval de l'ATG ou en amont du codon stop au risque de devoir insérer nos marqueurs dans le domaine fonctionnel de la protéine *zfC9* (**Figure 5**). Les ARNg élaborés ont été testés pour leur taux d'efficacité pour l'induction de CDB ciblés avec le système CRIPRS/Cas9, en les injectant avec l'ARNm codant pour une Cas9 à localisation nucléaire dans des embryons à 1 cellule, puis en évaluant les proportions d'embryons avec des insertions ou délétions (indels) dans leur génome à proximité du site de coupure correspondant par l'usage de la technique d'analyse des courbes de fusion à haute résolution (HRM, High Resolution Melt) (**Figure 7**).

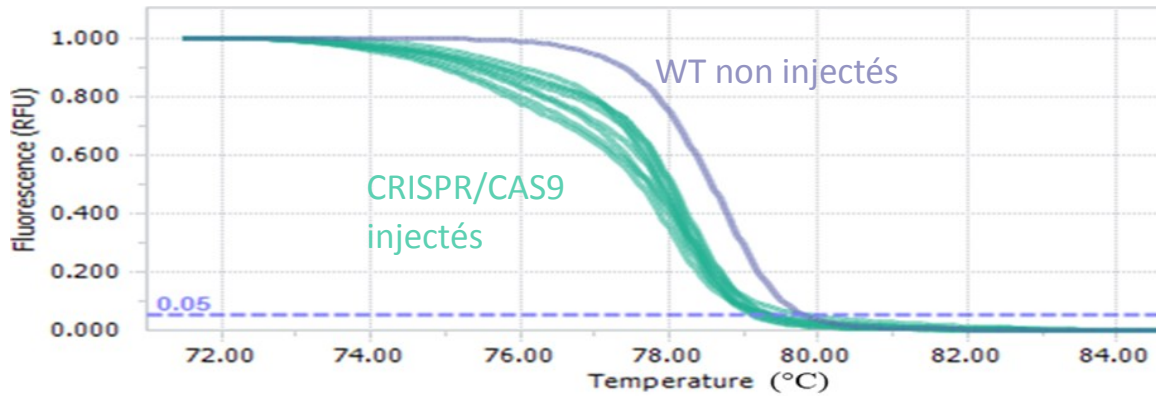


Figure 7. Test du taux d'efficacité d'un ARNg pour l'induction de coupure par le système CRISPR/Cas9. Test de l'efficacité de l'ARNg C9orf72Ct par analyse HRM effectué sur l'ADN génomique d'embryons TL WT injectés avec C9orf72Ct et Cas9. La présence d'indels est déterminée par le décalage du profil de la courbe HRM du spécimen comparé à celui de la courbe de WT pas injectés. Les courbes HRM sont normalisées et la fluorescence rapportée en % d'unité de fluorescence relatives (RFU). Les profils en verts indiquent la présence d'indels (13/16) et le bleu la séquence WT. L'efficacité de C9orf72Ct est donc de 81.3%. Les extractions ont été faites à 8hpf sur 16 embryons TL WT injectés et 4 non injectés.

L'analyse HRM consiste en l'amplification par PCR de la région de l'ADN d'intérêt en présence d'un des colorants intercalants qui ont la propriété de devenir plus fluorescents lorsqu'ils sont fixés à de l'ADN double brin. L'amplicon est alors chauffé progressivement et précisément d'environ 50°C à 95°C et la diminution de la fluorescence est enregistrée en temps réel alors que l'augmentation de la température provoque une séparation croissante des brins d'ADN [236]. Le profil des courbes HRM étant très sensible aux modifications de la séquence ADN amplifiée, leur analyse par comparaison avec une courbe WT de référence permet la détection rapide dans une séquence ADN d'indels aussi petites qu'un seul nucléotide ainsi que les polymorphismes nucléotidiques. Elle permet aussi d'identifier le génotype soit WT, hétérozygote (HT) ou homozygote (HM) d'un spécimen pour une allèle particulière. L'analyse HRM est aussi plus simple, rapide et moins onéreuse lorsque comparée aux autres techniques de génotypage employées chez le poisson-zèbre. En effet ces dernières impliquent communément l'amplification spécifique d'un locus par PCR suivi de la digestion par une enzyme de restriction spécifique ou directement le séquençage de la région d'intérêt, ce qui requiert un temps considérable et une qualité d'ADN génomique adéquate pouvant nécessiter des extractions génomiques plus élaborées [237].

Le seuil d'efficacité minimal pour un ARNg a été établi à $\geq 80\%$ pour ce projet, une induction importante de CDB augmentant les chances qu'une insertion survienne par RHD. Les ARNg ne montrant pas une efficacité adéquate à une concentration de 30 ng/ μ L ont été testés à nouveau avec des concentrations de 50 puis finalement 100 ng/ μ L avant d'être abandonnés. Ainsi au terme de nos analyses HRM et avec les ARNg sélectionnés pour procéder aux knock-in (KI) ont été C9orf72Nt pour le Nt et C9orf72Ct pour le Ct (**Figure 8**).

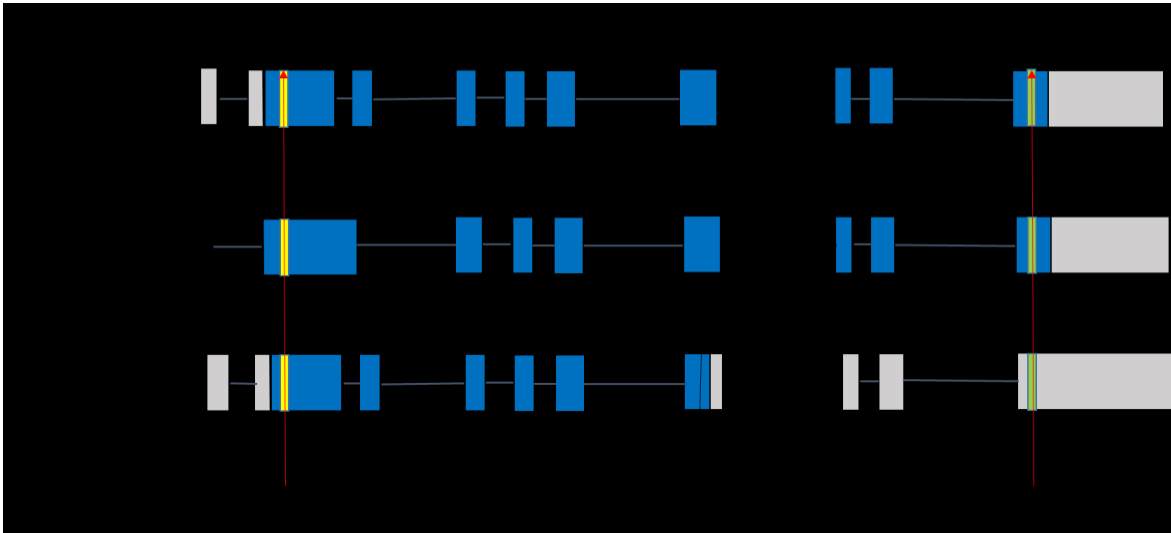


Figure 8. Lieux de coupe/insertion par CRISPR/Cas9 prévues dans la séquence de *zfc9* avec pour les ARNg C9orf72Nt et C9orf72Ct. (A) Site de coupe de C9orf72Nt, 231-232 pb en aval de l'ATG des variants 1 et 2. (B) Site de coupe de C9orf72Nt, 81-82 pb en amont du codon stop des variants 1 et 2. Le gris indique les exons non codants et le bleu les exons codant de *zfc9*.

Les matrices d'ADN de type Oligodésoxyribonucléotide simple brin (ODNsb) pour guider la réparation et l'insertion des tags ont été conçues de sorte à intégrer leur tag respectif dans le cadre de lecture principal de C9 et perturber le site de reconnaissance des ARNg correspondant, pour que ces derniers ne puissent guider à nouveau Cas9 pour couper une seconde fois après l'insertion. Les sites de coupe en Nt et Ct des ARNg sélectionnés étaient prédits de couper à 231-232 pb de l'ATG et 81-82 pb du codon d'arrêt de *zfc9* respectivement. Avec les matrices d'ADN conçues pour l'insertion de leur tag, l'insertion en phase de lecture était attendue de s'effectuer entre les acides aminés (aa) 78-79 de l'isoforme *zfc9*^L pour C9orf72^{NHis} ainsi que

C9orf72^{NGFP} et entre les aa 435-436 pour C9orf72^{CGFP}. Les protéines zfC9 de fusions résultantes auraient une longueur de 474 aa pour C9orf72^{NHis} alors que celles de C9orf72^{NGFP} et C9orf72^{CGFP} auraient toutes deux une longueur de 700 aa (**Figure 9**).

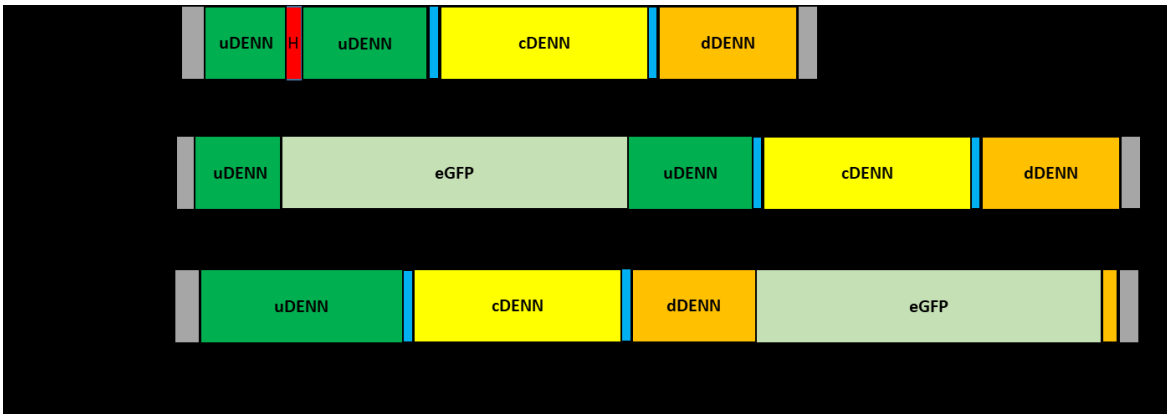


Figure 9. Schéma des protéines zfC9 de fusion résultante de l’insertion réussie dans des tags pour les lignées C9orf72^{NHis}, C9orf72^{NGFP} et C9orf72^{CGFP}. C9orf72^{NHis} : l’insertion au Nt du tag 12xHis (36 pb,12 aa) dans le cadre de lecture de zfC9 sépare le cDENN et résulte en une protéine de fusion de 474 aa. C9orf72^{NGFP} : l’intégration au Nt du tag eGFP (717 pb,238 aa) dans le cadre de lecture de zfC9 divise le cDENN et résulte en une protéine de fusion de 700 aa. C9orf72^{CGFP} : l’incorporation d’eGFP au Ct scinde le dDENN et produit une protéine de fusion de 700 aa. **H** indique le 12xHis.

Enfin, pour générer les lignées avec un marquage au Nt l’ARNm de Cas9, C9orf72Nt puis la matrice d’ADN C9Nt-12xHis pour C9orf72^{NHis} et C9Nt-eGFP pour C9orf72^{NGFP} ont été injectés chez des embryons de poisson-zèbre Tubingen long fin (TL) WT au stade 1 à 2 cellules. La génération de la lignée C9orf72^{CGFP} s’est effectuée de manière similaire en substituant l’ARNg pour C9orf72Ct et en employant la matrice de réparation de l’ADN C9Ct-eGFP.

2.2.2 Confirmation de l’insertion des marqueurs moléculaires

La faisabilité des insertions des tags avec les couples d’ARNg et leur matrice ADN de réparation respective a été vérifiée expérimentalement par expression mosaïque des plasmides pour chaque lignée avant de poursuivre. Suite aux micro-injections, les embryons des lignées marquées au eGFP ont été observés entre 24 à 72 heures post-fertilisation (hpf) sous stéréomicroscope à

épifluorescence Olympus SZX7. La présence d'une mosaïque de cellules fluorescentes verte chez les embryons injectés pour les lignées $C9orf72^{NGFP}$ et $C9orf72^{CGFP}$ a confirmé que ces insertions du tag eGFP pouvaient fonctionner (**Figure 10**).

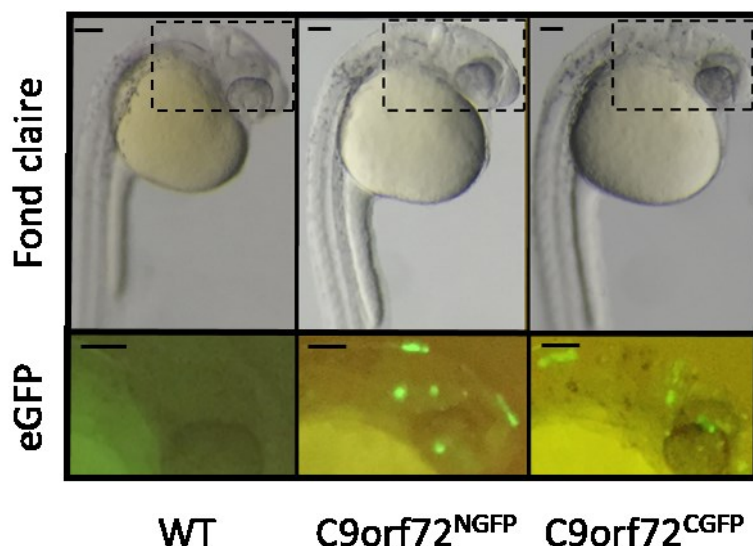


Figure 10. Confirmation de la faisabilité de l'insertion par CRISPR/Cas9 des étiquettes moléculaires pour l'obtention des lignées $C9orf72^{NGFP}$ et $C9orf72^{CGFP}$. La faisabilité de l'insertion du tag eGFP au Nt ou Ct par l'usage du système CRISPR/Cas9 avec nos constructions moléculaires a été confirmée pour les lignées $C9orf72^{NGFP}$ et $C9orf72^{CGFP}$ par l'observation sous microscope à épifluorescence d'embryons injectés. Effectivement, des embryons injectés pour les lignées $C9orf72^{NGFP}$ et $C9orf72^{CGFP}$ présentent des sites de fluorescences de couleur verte avec une intensité et un patron de signal distinct de la faible autofluorescence verte observée chez les embryons 24hpf WT au niveau de leur vitellus. Les embryons observés sont 24hpf; l'échelle (barres noires) est de 50 μ m.

Pour la lignée $C9orf72^{NHIS}$ la confirmation de l'intégration du tag 12xHis dans zfc9, qui n'émet pas de fluorescence, a été faite par amplification PCR à partir d'ADN génomique extraite de larves 96 heures après les micros-injections et l'utilisation d'amorces spécifiques à l'insertion du tag 12xHis. Le succès de l'amplification PCR a été déterminé par l'observation d'une bande de 245 pb suite à une migration sur gel d'agarose (**Figure 11**). L'extraction a été faite à 96 hpf de sorte à écarter toute possibilité que l'amplification PCR puisse être attribuée à de l'ADN

matriciel du tag 12xHis persistant dans les embryons injectés. De plus, pour accélérer la procédure en cas de faible taux d'insertion et puisque nous ne cherchions qu'à confirmer sa faisabilité, les amplifications ont été effectuées sur des solution rassemblant l'ADN génomique extraite de 5 embryons injectés.



Figure 11. Confirmation de la faisabilité de l'insertion par CRISPR/Cas9 des étiquettes moléculaires pour l'obtention la lignée C9orf72^{NHis}. La faisabilité de l'insertion du tag 12xHis au Nt ou Ct par l'usage du système CRISPR/Cas9 avec nos constructions moléculaires est confirmée pour la lignée C9orf72^{NHis}. En effet, l'ADN génomiques de certains embryons injectés pour la lignée C9orf72^{NHis} permet l'amplification d'une séquence ADN d'une longueur de 245 pb avec un couple d'amorces PCR ne fonctionnant que si le tag 12xHis a été inséré au site Nt ciblé de zfC9. Cette amplification est observée par la présence d'une bande de 245 pb suite à la migration sur gel des produits PCR visualisé sous UV. La bande est présente pour des rassemblements d'ADN génomique de nos embryons WT injectés et absente pour les WT non injectés. Gel 1%, migration à 80 V, extraction faite à 96 hpf, amplification sur des extraction d'ADN de 5 embryons à la fois pour les WT injectés et non injectés. A, B et C dénotent des groupes d'extraction différents.

2.2.3 Sélection des poissons fondateurs F0

Pour les lignées C9orf72^{NGFP} et C9orf72^{CGFP} seuls les embryons présentant de la fluorescence évidente dans plusieurs cellules, lorsque observés sous microscope à épifluorescence entre 24 et 72 hpf, ont été sélectionnés pour être élevés jusqu'à l'âge adulte. Ainsi on s'assurait que l'intégration de l'eGFP soit survenue au moins à quelques reprises dans l'embryon, indiquant une probabilité non nulle qu'elle soit survenue dans les lignées germinales de ces embryons et qu'ils puissent ainsi la transmettre à leur progéniture pour former la F1 de ces lignées. N'ayant pas de marqueur d'insertion pouvant être simplement visualisé, l'ensemble des embryons

injectés pour former la F0 de la lignée C9orf72^{NHis} ont été conservés. Pour l'ensemble des lignées, les embryons monstrueux et difformes non-viables ont été retirés progressivement alors que leur état se manifestait jusqu'au cinquième jour après injection (éclosion après 2 jours).

2.2.4 Élevage des poissons injectés F0

Les injections pour la lignée C9orf72^{NHis} ont eu lieu sur une période de deux semaines pour chaque cohorte et étaient répétées jusqu'à l'obtention de (~250) embryons injectés ne présentant pas de défauts évidents de morphologie 5 jours après l'injection. Quatre cohortes au total ont ainsi été élevées séparément durant ce projet pour la lignée C9orf72^{NHis}. Cependant, seulement 86 des embryons sélectionnés ont survécu pour atteindre l'âge adulte et 81 ont survécu pour être criblés pour l'insertion et la transmission du tag 12xHis (**Tableau 2**).

Suivant un modèle d'injection similaire pour la lignée C9orf72^{NGFP}, deux cohortes de ~120 embryons sélectionnés pour leur fluorescence et leur morphologie normale à 72hpf ont été élevées. Des deux cohortes d'embryons sélectionnés, un seul des embryons a survécu jusqu'à l'âge adulte afin d'être criblé pour la transmission du tag eGFP en Nt (**Tableau 2**).

En ce qui concerne la lignée C9orf72^{CGFP}, quatre cohortes de ~200 embryons ont été sélectionnés et élevés de manière identique à ceux de la lignée C9orf72^{NGFP}. Au terme de ce projet 78 spécimens ont atteint l'âge adulte et 3 de ces spécimens adultes ont décédé avant que leur criblage ne soit complété, laissant un total de 75 disponibles pour être criblés (**Tableau 2**).

2.2.5 Criblage pour l'identification de poisson fondateurs F0

Afin d'identifier les spécimens injectés en mesure de transmettre les tags correspondant pour chacune des lignées C9orf72^{NHis}, C9orf72^{NGFP} et C9orf72^{CGFP}, les spécimens injectés ayant survécu ont été croisés avec des poisson TL WT. Chaque poisson F0 étant croisé avec au moins 2 spécimens WT de sexe opposé (croisement de type *outcross*). De ces croisements nous espérons obtenir des spécimens HT pour l'intégration des différents tags moléculaires. Les croisements ont été tentés aussitôt que les poissons injectés arrivaient à maturité sexuelle (environ 10-12 semaines [232]), que leur taille ait été jugée adéquate pour qu'il ne risque pas de fuir par les interstices des aquariums du système et qu'elle soit comparable à leurs partenaires

de croisement WT. Une fois ces critères rencontrés, les poissons à être criblés étaient placés dans des aquariums individuels de sorte à pouvoir clairement établir quel poisson était responsable de quelle progéniture à être testé et donc auquel était attribué un résultat négatif ou positif. De 4 à 5 jours au minimum étaient laissés aux spécimens pour se reposer et maturer en isolation avant de tenter de les croiser, particulièrement pour les femelles pour lesquelles il était attendu que leur abdomen commence à s'enfler de manière visible de par l'accumulation des œufs avant de commencer à les croiser. Cette approche a été répétée pour chaque génération à être criblée et ce pour chaque lignée utilisée durant ce projet, que ce soit pour une insertion d'un tag, un KO ou l'intégration d'un transgène. Notons aussi que pour toute modification génique ayant une intégration mosaïque (CRISPR/Cas9, Tol2) dans les cellules des poissons, la transmission de telles modifications, même si réussie dans des cellules germinales, ne suivra pas les proportions mendéliennes.

Pour la lignée C9orf72^{NHis}, le criblage pour la transmission du tag 12xHis s'est effectué de manière identique à la confirmation de la faisabilité de son insertion avec nos constructions génétiques tel que décrit à la **section 2.2.2**. Les seules distinctions étant que l'extraction d'ADN génomique a été effectuée à partir d'embryons 8hpf résultant du outcross C9orf72^{NHis} x WT plutôt que de larves 96hpf injectées. Effectivement, ne cherchant qu'à détecter initialement la capacité pour chaque poisson de survivre à transmettre le tag 12xHis, l'amplification par PCR spécifique au tag 12xHis a encore été effectuée sur des extractions d'ADN génomique à partir de 5 embryons en même temps. Pour qu'un poisson soit déclassé et considéré incapable de transmettre l'intégration du 12xHis tag, l'ADN génomique d'au moins 80 embryons obtenus par poisson devaient se révéler incapable de permettre une amplification par PCR en utilisant les amorces spécifiques à l'intégration de 12xHis. En cas d'amplification, une seconde confirmation d'un différend croisement avec le même poisson injecté était nécessaire pour considérer ce dernier comme un candidat fondateur F0 pour la lignée C9orf72^{NHis}. Après l'identification de tel candidats, ceux-ci auraient été croisés en outcross à nouveau et leur progéniture aurait été criblée pour l'insertion du tag 12xHis, leur ADN génomique étant extraite par la coupe non-léthale d'un fragment de la nageoire caudale des poissons de taille adéquate. Ce criblage aurait été effectué par analyse HRM, en raison de la grande sensibilité de la technique, avec des amorces spécifiques à l'insertion du tag 12xHis. L'ADN génomique des

poissons permettant l'amplification aurait été alors séquencé pour vérifier l'intégration dans le cadre de lecture du gène *zfc9*, la difficulté d'analyse des chromatogrammes HT nécessitant peut-être l'attente de l'obtention de spécimens HM pour une confirmation finale de la nature de l'insertion. Enfin, l'expression de la protéine *zfc9* marquée à l'histidine aurait été confirmée à l'aide de la technique du western blot, effectuée avec un anticorps spécifique au tag poly-histidine et une extraction protéique de la progéniture de ces poissons HT pour l'intégration du tag dans l'ensemble de leurs cellules. Par une succession de croisements entre les poissons exprimant la protéine *zfc9* marquée à l'histidine et le génotypage de leur progéniture par analyse HRM, une lignée HM *C9orf72^{NHis}* stable aurait éventuellement été générée. Cependant, après le criblage des 81 poissons adultes survivant ayant le potentiel d'être des fondateurs pour la lignée *C9orf72^{NHis}*, aucun n'a démontré la capacité de transmettre le tag 12xHis (**Tableau 2**).

Le criblage pour l'intégration et la transmission du tag eGFP des lignées *C9orf72^{NGFP}* et *C9orf72^{CGFP}* s'est effectué simplement en observant sous microscope à épifluorescence les embryons résultant des croisements des poissons injectés avec les WT et en recherchant les spécimens fluorescents. La recherche de fluorescence était effectuée à 4 reprises pour chaque portée d'embryons, soit à 24hpf, 36hpf, 48hpf et 72hpf. Passé 72hpf, il était considéré qu'aucune fluorescence ne se développerait chez les embryons en question. L'ARNm de *zfc9* étant exprimé à 24hpf dans l'ensemble des organes majeurs du poisson-zèbre et particulièrement dans le CNS [105] (**figure 6**), un patron de fluorescence similaire était attendu pour les embryons exprimant la protéine de *zfc9* avec un tag eGFP en N ou Ct. Un poisson injecté des lignées *C9orf72^{NGFP}* ou *C9orf72^{CGFP}* était déclassé et considéré incapable de transmettre le tag eGFP intégré après qu'au moins 120 embryons aient été observés sans fluorescence à 72hpf. En cas d'identification d'embryons fluorescents, le croisement devait être répété avec le poisson candidat jusqu'à une seconde obtention d'embryons fluorescents. Suite à l'identification d'un poisson confirmé de transmettre le tag GFP, les membres de sa progéniture démontrant de la fluorescence auraient été élevés et leur site d'insertion prévu en N ou Ct aurait été séquencé pour confirmer l'intégration du GFP dans le cadre de lecture de *zfc9* tel qu'attendu en raison de la fluorescence observée. Des lignées HM *C9orf72^{NGFP}* ou *C9orf72^{CGFP}* auraient ensuite été produites par une série de croisements, par une méthode similaire à celle décrite pour la lignée *C9orf72^{NHis}*, mais adaptée à l'insertion du tag eGFP. Malheureusement, l'unique poisson injecté

de la lignée C9orf72^{NGFP} ayant survécu jusqu'à l'âge adulte s'est avéré incapable de transmettre le tag GFP inséré en Nt de zfc9 tout comme les 75 poissons injectés survivant de la lignée C9orf72^{CGFP} se sont montrés incapables de transmettre le tag GFP inséré en Ct (**Tableau 2**).

Tableau II. **Résultats du crible des poissons candidats F0 adulte des lignées C9orf72^{NHis}, C9orf72^{NGFP} et C9orf72^{CGFP}.** Tableau récapitulatif de l'élevage et du criblage des poissons injectés et sélectionnés pour la transmission des tags moléculaires insérés dans zfc9 pour les lignées C9orf72^{NHis}, C9orf72^{NGFP} et C9orf72^{CGFP}. F0 dénote les poissons potentiels et non de véritables fondateurs pour ces lignées.

	C9orf72 ^{NHis}	C9orf72 ^{NGFP}	C9orf72 ^{CGFP}
F0 Adultes	86	78	1
Décès adultes	5	3	0
F0 criblés	81	75	1
F0 positif	0	0	0

2.3 Génération d'une lignée zfc9 knock-out stable

Tel que mentionné à la **section 1.3.2**, aucun modèle de KO du gène zfc9 n'a encore été rapporté dans la littérature. En effet, seul des KD transitoires de zfc9 résultant en des défaillances au niveau du système moteur, une perturbation des prolongements des axones des motoneurons et un signal d'apoptose accrue dans le SNC ont été décrits jusqu'à présent [105,213]. Une lignée poissons-zèbres KO pour zfc9 permettrait de confirmer ou d'infirmer ces observations en plus de permettre de caractériser l'effet de ce KO tout au long du cycle de vie du poisson-zèbre plutôt que seulement dans les stades du développement précoce. Une telle lignée serait ainsi potentiellement un outil puissant pour l'étude de la fonction normale du gène zfc9, offrant un modèle intermédiaire entre les modèles de vers *C. elegans* et les modèles murins dont les phénotypes pour le KO de leur HM C9 diffèrent. De plus cette lignée offrirait une avenue potentielle pour modéliser le mécanisme de toxicité de perte de fonction de C9 SLA chez le poisson-zèbre et potentiellement de récapituler les trois mécanismes de toxicité proposés dans un même modèle en effectuant notamment des croisements avec une des lignées récemment

développées expriment de manière stable un certain nombre d'ERH G4C2 [154]. Nous avons donc tenté générer une telle lignée à l'aide du système CRIPSR/Cas9.

2.3.1 Criblage pour la transmission d'indels dans *zfC9*

Les réparations de l'ADN par JENH susceptibles d'induire des délétions ou des insertion aléatoires de nucléotides (indels) sont considérablement plus fréquentes que les RDH, suite à l'induction de CBD et ce même lors de l'utilisation optimale du système CRIPR/Cas9 avec une matrice d'ADN. Les modifications aléatoires sont donc plus fréquemment retrouvées et transmises que les insertions ciblées. De plus, ces modifications aléatoires peuvent potentiellement et avec une fréquence élevée induire un changement du cadre de lecture de la séquence menant à un KO du gène ciblé et ce particulièrement si elles surviennent au Nt près du codon d'initiation [230]. Ainsi, les poissons injectés initialement pour élaborer la lignée C9orf72^{NHis} avaient une probabilité relativement importante d'avoir intégré différentes insertions ou délétions menant au KO de *zfC9* en plus d'avoir possiblement incorporer le tag 12xHIS dans certaines de leurs cellules germinales. Ils ont donc été simultanément criblés pour la transmission d'indels et de l'insertion du tag 12xHis afin d'identifier des poissons susceptibles de former la F0 d'une lignée *zfC9*^{KO}.

Le criblage pour la transmission d'indels s'est effectué par analyse HRM de manière quasi identique à l'évaluation de l'efficacité de coupure des ARNg utilisés pour les insertions par système CRISPR/CAS9. Les principales différences étant que l'extraction d'ADN génomique était effectuée sur des embryons résultant de croisement entre des poissons F0 C9orf72^{NHis} et des TL WT. En raison de l'incidence relativement élevé d'indels aléatoires dans la séquence d'ADN ciblée lors de l'usage du système CRISPR/Cas9 et donc leurs transmissions relativement plus fréquentes, le seuil pour déclasser un poisson et le considérer négatif au niveau des indels a été plus bas que celui pour l'insertion du tag 12xHis. Le criblage s'est effectué en deux temps, 24 embryons étaient initialement testés pour chaque poisson par analyse HRM et, si aucune des courbes de l'analyse HRM obtenues ne différaient du profil de la courbe des contrôle WT, le poisson géniteur de ces embryons était déclassé une première fois. Les poissons déclassés une fois étaient ensuite croisés entre eux de sorte à accroître les probabilités de repérer une dernière fois des indels au niveau du site de coupure de l'ARNg C9orf72Nt et couvrir plus de poissons

plus rapidement. Si aucune courbe de l'analyse HRM distincte des contrôles WT était observée après un criblage de 24 embryons additionnels, alors les poissons étaient définitivement déclassés et jugés incapables de transmettre d'indels à leur progéniture. Un total de 34 poissons a été criblé pour au moins 24 embryons, et de ceux-ci 2 poissons ont été identifiés comme en mesure de transmettre des indels à leur progéniture, 17 poissons ont été déclassés une seule fois et 14 ont été définitivement exclus après un second tour d'analyse par HRM (**Tableau 3**).

Tableau III. **Résultats du criblage des poissons F0 C9orf72^{NHis} pour la transmission d'indels en N-terminal de C9orf72.** C9orf72^{NHis} criblés : Résultats du criblage des poissons F0 C9orf72^{NHis} pour des spécimens aptes à transmettre des indels au Nt de *zfc9* et ainsi devenir candidats F0 pour la lignée C9orf72^{KO}. Total est le nombre de poissons criblés par analyse HRM au moins 1 fois avec 24 embryons; (+) indique un poisson apte à transférer une modification de la séquence de *zfc9* en Nt, ♀ signifie un poisson femelle et ♂ dénote un poisson mâle. ♀+/♂+ Criblés : Résultats des analyses HRM sur les embryons de E4 et O2. Tests est le nombre d'embryons dont l'analyse HRM a fonctionné, Indels indique le nombre d'embryons analysés avec un profil HRM indiquant des indels et % indique le taux de transmission en pourcentage.

<u>C9orf72^{NHis} criblés</u>						<u>♀+/♂+ Criblés</u>					
Total	Déclassé 1x	Déclassé 2x	+	♀+	♂+	O2 (♀)			E4 (♂)		
						Tests	Indels	%	Tests	Indels	%
34	10	24	2	1	1	64	11	17.1	67	5	7.4

Les poissons positifs pour la transmission d'altérations au Nt de leur séquence de *zfc9*, soit le mâle E4 et la femelle O2, ont été reconfirmés avec un HRM sur 48 de leurs embryons. Afin d'être absolument certains que ces courbes indiquent des indels, la séquence d'ADN de *zfc9* flanquant le site de coupure au Nt a été amplifiée par PCR à partir de l'ADN génomique des embryons ayant produit les courbes d'intérêt puis le résultat a été séquencé. Suite au séquençage, la présence d'indels a été confirmée pour le mâle E4 et la femelle O2 (**Figure 12**). Le taux de transmission d'indels, évalué du nombre d'embryons total testés avec succès par HRM pour chacun des deux poissons positifs, s'est avéré plus élevé pour le mâle E4 (17.1%) que chez la femelle O2 (7.4%) (**Tableau 3**).

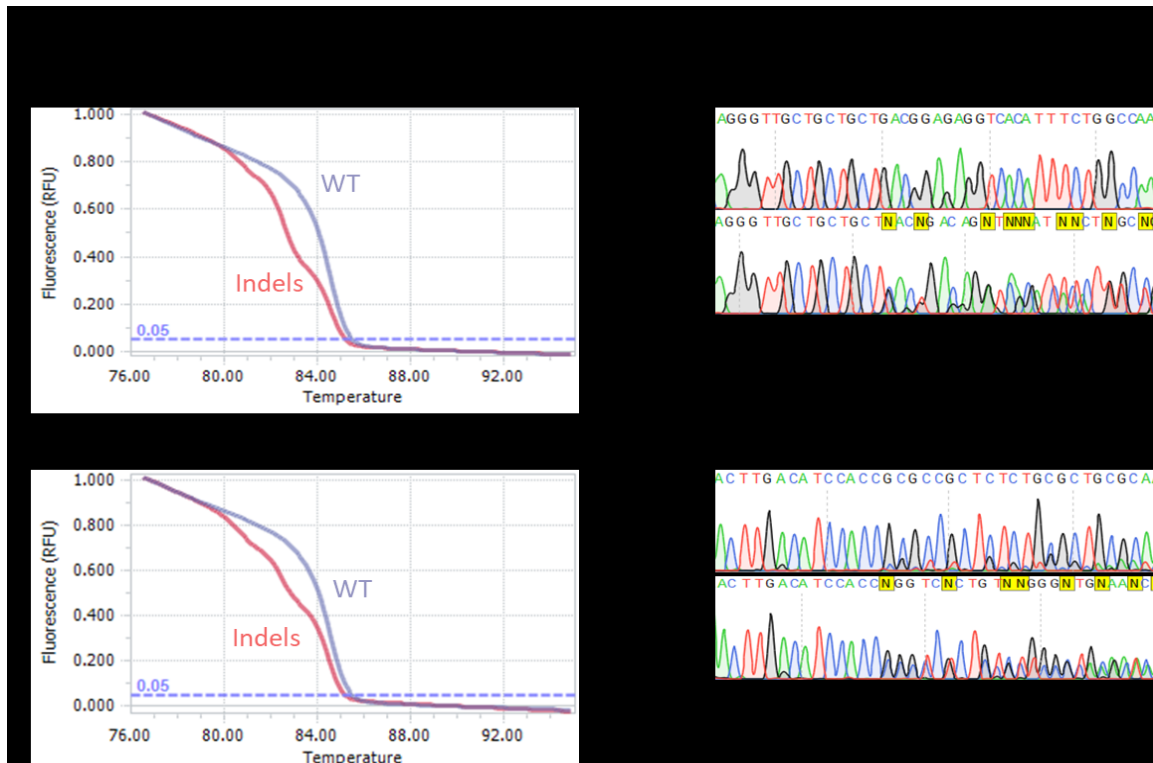


Figure 12. Identification de la présence/transmission d'indels par analyse HRM et confirmation par séquençage de la région ciblée. Profil de courbes d'analyse HRM identifiées comme susceptibles d'indiquer une transmission d'indels dans la séquence ADN de *zfc9* et leur confirmation par amplification PCR de la région analysée par HRM incorporant le site de coupe de l'ARNg suivi de son séquençage. (i) Profil HRM d'un des descendants de E4 (8hpf) dont le décalage avec la courbe WT suggère la présence d'indels, confirmé par la perturbation de la séquence du chromatogramme correspondant comparé au WT. (ii) Profil HRM d'un des descendants de O2 (8hpf) dont le décalage avec la courbe WT indique aussi la présence d'indels, confirmé par la perturbation de la séquence du chromatogramme correspondant comparé au WT. Les courbes HRM sont normalisées et la fluorescence rapportée en % d'unité de fluorescence relatives (RFU). Les profils rouge indiquent la présence d'indels et les bleus le profil sont ceux du contrôle WT.

2.3.2 croisement et élevage de poissons candidats pour la F1

Aussitôt qu'identifiés, les poissons E4 et O2 ont été croisés avec des poissons WT TL, de sorte à obtenir une progéniture pouvant être criblée pour obtenir des individus HT au niveau de l'intégration d'indels au Nt du gène *zfc9*. Les croisements, la sélection et l'élevage de cette

progéniture ont été exécutés de manière identique à ce qui a été décrit pour les poissons F0 de C9orf72^{NHis} aux **sections 2.2.3 à 2.2.4**

2.3.3 Criblage et identification de poissons candidats pour la F1

Une fois atteint une taille adéquate, coïncidant généralement avec la fin du stade juvénile (6 à 12 semaines après fertilisation [232]), les poissons susceptibles de pouvoir être candidats pour la F1 de la lignée zfc9^{KO} ont été génotypés par analyse HRM. Cependant, l'extraction s'est faite cette fois à partir d'un fragment de nageoire caudale sectionnée (fin clip) de sorte à être non-léthale, alors que le spécimen était anesthésié et rendu temporairement immobile par l'ajout d'une solution de tricaine dans l'eau ou il baignait avant l'incision. Lorsque bien exécuté, les spécimens retrouvent rapidement leur mobilité une fois retourné à un aquarium rempli d'eau du système et ils récupèrent entièrement, régénérant même le segment de nageoire sectionné au terme de quelques semaines [232]. Ainsi 33 spécimens, issus du mâle E4 et les plus avancés dans leur développement, ont été criblés en premier par analyse HRM, révélant que 3 d'entre eux étaient HT pour des indels au Nt de *zfc9*, E4a, E4b et E5c, alors que les 29 restants étaient WT (**Tableau 4**). L'ensemble des 27 descendants de O2 finalement criblés se sont avérés avoir un profil HRM WT. Le criblage de poissons additionnels s'est arrêté une fois que nous avons identifiés des poissons HT pour des indels dans *zfc9* menant à un KO du gène (**Tableau 4**). Ainsi, plusieurs E4 et O2 F1 n'ayant pas été criblés demeurent encore en réserve.

Tableau IV. **Résultats du criblage des poissons F1 pour l'intégration d'indels en N-terminal de *zfc9*.** Résultats du criblage des poissons F1 juvéniles de E4 et O2 ayant possiblement intégré des indels au Nt de *zfc9* entraînant un KO du gène. Total est le nombre de poissons F0 criblés par analyse HRM à partir de l'ADN génomique extrait de fragment de nageoire caudale ; (+) indique des poissons ayant intégrés une modification de la séquence de *zfc9* en Nt, ♀ signifie des poissons femelles adultes et ♂ dénote un poissons mâles adultes.

<u>E4 F1 Criblés</u>				<u>O2 F1 criblé</u>			
Total	(+)	♀+	♂+	Total	(+)	♀+	♂+
33	3	1	2	27	0	0	0

2.3.4 Génotypage des spécimens portant des indels et caractérisation de celles-ci

Les réparations imparfaites de l'ADN, suite aux CBD induites par l'usage du système CRISPR/Cas9 qui résultant en des délétions ou des insertions, sont aléatoires et même avec des injections optimales les sujets injectés présentent généralement une distribution mosaïque de différentes altérations à leur ADN à travers leur tissu, incluant leurs cellules germinales [227]. Ainsi, il n'est pas certains qu'un spécimen en mesure de transmettre des indels à sa progéniture ne va transmettre qu'un seul type d'indels à ses descendants. Une solution à ce problème est le séquençage des poissons HT F1, afin de tenter d'identifier la nature des indels pour un poisson donné, en comparant la séquence obtenue avec celle d'un WT, nous permettant directement de déterminer si ces indels sont de mêmes types et, plus important encore, si ceux-ci entraînent le KO de *zfc9*. Notons que l'interprétation des chromatogrammes de spécimens HT pour des indels inconnues et aléatoires peut s'avérer laborieux et sujet à des erreurs d'interprétation dépendamment de la nature des indels, dont notamment dans le cas où plusieurs indels distinctes sont survenues à différents sites dans la même région séquencée. Une façon de vérifier avec certitude si les indels sont de même type entre spécimens, si les circonstances sont propices, est de croiser entre-eux les spécimens obtenus ayant des profils de courbes HRM similaires et de cribler leur progéniture par analyse HRM. En effet, si les indels sont différents entre le mâle et la femelle croisés, quatre profils différents pour les courbes HRM seront obtenus de ces croisements, soit un profil WT, un profil Ht pour le premier type d'indels, un profil pour les Ht de second type d'indels et finalement un profil distinct pour les embryons Ht pour les deux types d'indels. Si les indels sont de même nature, alors 3 profils clairs seront observés, soit un profil WT, un seul profil Ht et un profil ressemblant au WT, mais décalé de celui-ci selon l'axe de température. Ces observations par analyse HRM peuvent ensuite être confirmées assez rapidement par séquençage de la section amplifiée.

Ainsi, comme les profils des courbes de l'analyse HRM des trois spécimens F1 était très similaires, chacun étant en âge d'être croisé et ayant 1 femelle (E4b) ainsi que deux mâles (E4a et E4c) de disponible, nous avons donc tenté de croiser nos spécimens ensemble, un mâle à la fois. Une fois des embryons des croisement E4a/E4b et E4b/E4c obtenus, nous avons tenté de

déterminer par analyse HRM sur ces embryons si les indels d'au moins un mâle étaient de même nature que celles de la femelle. Pour les croisements E4a/E4b et E4b/E4c les trois mêmes profils de courbes assumés de correspondre au spécimens WT, HT et WT des croisements ont été observés par analyse HRM (**Figure 13 A**). Afin de confirmer ces observations et potentiellement caractériser la nature de ces indels, plusieurs ADN génomiques d'embryons avec un profil HM pour les indels ont été envoyés pour être séquencés. Le séquençage a révélé que les indels étaient bien identiques pour les spécimens HM comme le démontrent les chromatogrammes obtenus (**Figure 13 B**). En effet, les pics des chromatogrammes dans les régions où sont retrouvés les indels sont clairs, bien définis et ne se superposent pas, contrairement à ce qui serait attendu pour des poissons HT, pour lesquels la présence de différentes séquences sur chaque copie sœur du même chromosome dans la région séquencée, entraînerait une superposition de pics associés à différents nucléotides comme montré à la **figure 12**. La nature des indels est aussi révélée de ce séquençage, s'avérant être deux délétions à des sites distincts du Nt de zfc9. La première est une délétion symétrique de 4 pb localisée 73 pb en amont du site de coupure prévu avec l'ARNg C9orf72Nt (**figure 13 B1**) et la seconde est une délétion asymétrique de 7 pb au site de coupure prévu pour le même l'ARNg (**figure 13 B2**). Aucun autre indel n'était présent dans zfc9 jusqu'à 120 pb en amont de la première délétion et 140 en aval de la seconde.

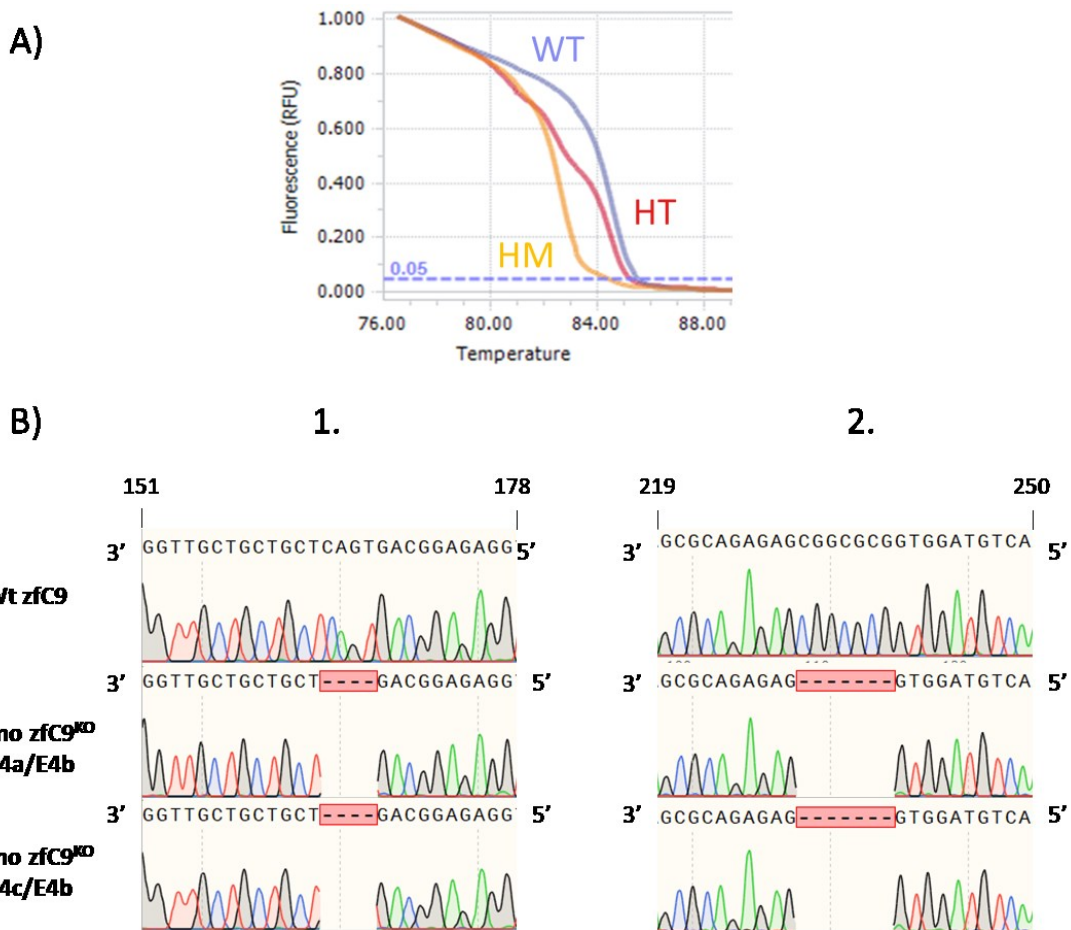


Figure 13. Profils HRM commun des croisement E4a/E4b et E4b/E4c et caractérisations de leurs indels communs par séquençage. (A) 1. L'analyse HRM à partir de l'ADN génomique de la progéniture des croisement E4a/E4b et E4b/E4c démontre les mêmes 3 profils de courbes HRM. Ces trois profils sont caractéristiques de spécimens WT, HT et HM pour une séquence donnée avec un seul type de mutations. (B) Confirmation de l'homologie des indels et caractérisation de celles-ci par séquençage des séquences des HM issues de chacun des deux croisements, illustrées par la superposition des chromatogrammes des régions avec des indels qui sont comparés à une séquence WT. (1) Indique une première délétion symétrique de 4 pb localisé 73 pb en amont du site de coupe prévu avec l'ARNg C9orf72Nt. (2) Indique une seconde délétion asymétrique et de 7 pb au site de coupe prévu. Les courbes HRM sont normalisées ainsi que les valeurs de variation de fluorescence en fonction de la température, la fluorescence rapportée en % d'unité de fluorescence relatives (RFU). Les profils rouges indiquent un profil HT pour les indels, les bleus représentent le contrôle WT et l'orange les profils HM.

2.3.5 Caractérisation de l'effet des délétions identifiées sur la protéine zfC9

Pour déterminer si les délétions identifiées entraînent un KO fonctionnel de zfC9, la séquence d'ADN codant du gène de nos poissons avec les délétions identifiées a été traduite dans le cadre de lecture principale du gène, soit le seul à produire les trois isoformes prédites pour *zfC9*. Nous observons alors qu'un changement de cadre de lecture a lieu à partir de la première délétion altérant complètement la séquence d'aa en aval de cette dernière et provoquant l'insertion d'un codon d'arrêt (TGA) prématuré dans le cadre de lecture de la protéine zfC9 avant même que la seconde délétion puisse affecter la protéine (Figure 14).

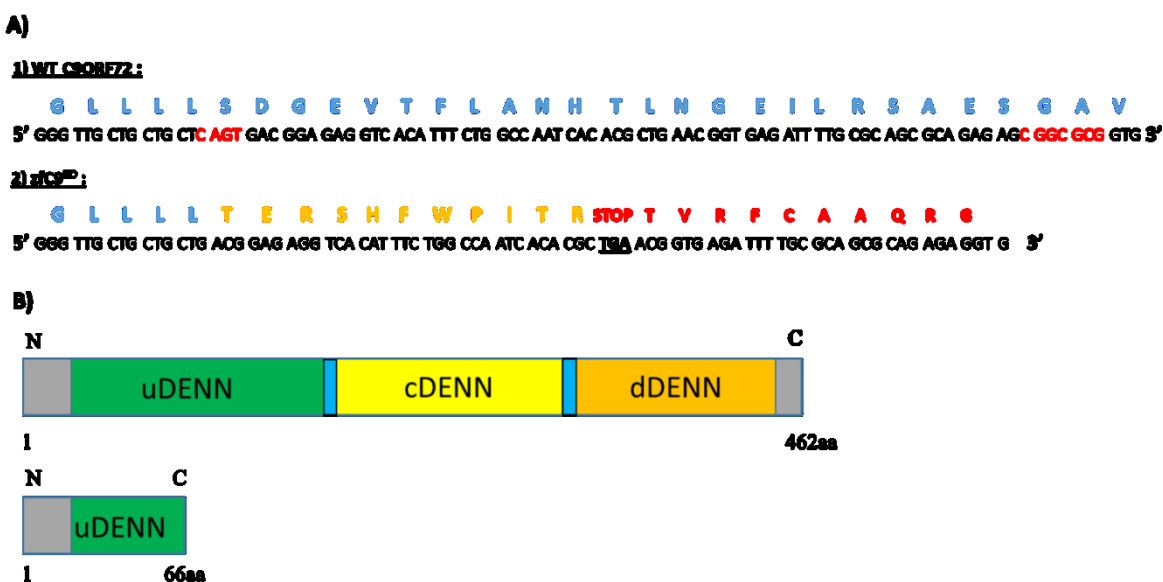


Figure 14. Schéma de l'effets des délétions identifiées chez les poissons E4 F1 sur la séquence d'acides aminés de la protéine zfC9. (A) L'alignement des séquences d'acides aminés incluant ses codons correspondant pour zfC9 WT (1) avec la séquence incorporant les délétions (2) démontre qu'elles provoquent un changement de cadre de lecture. Il en résulte la perturbation de la séquence de zfC9 au 55^{ième} aa et l'insertion d'un codon d'arrêt TGA dans le cadre de lecture qui tronque la protéine au 66^{ième} aa. Les aa communs à la séquence mutée et WT sont indiqués en bleu, les acides aminés qui diffèrent sont en jaune et celles absentes en raison du codon d'arrêt sont en rouge. L'alignement de séquence est du 51^{ième} aa de zfC9 au 81^{ième}. (B) Comparaison de la structure générale de la séquence d'aa WT de zfC9 avec celles incorporant les délétions. L'insertion du codon d'arrêt prématuré dans le cadre de lecture de zfC9 entraîne la perte de 85,7% de la séquence de sa protéine (66 vs 462 aa), incluant la perte complète de deux des trois sections de son domaine DEEN et la perte de 71.7% de la première section (43 vs152 aa, Figure 4).

La protéine tronquée résultante est longue de 66 acides aminés sur les 462 du transcrit le plus long de zfc9, elle est identique à la séquence d'aa des trois isoformes WT de zfc9 pour ses 55 premiers acides aminés, puis présente une séquence distincte de 11 aa avant que la traduction ne prenne fin par la lecture du codon TGA (**Figure 14**). De la perte de 85,7% de la séquence totale de la protéine et de la majorité des sections prédites de composer son domaine fonctionnelle DENN, il a été conclu que les indels des poissons E4 a, b et c entraînent un KO fonctionnel des trois isoformes prédites de zfc9. Il a donc été possible de lancer une lignée zfc9^{KO} à partir de ces poissons F1.

2.3.6 Croisement pour l'obtention de la F2

Leurs indels étant communs et confirmés de mener à un KO de zfc9, les poissons F1 (E4a, E4b, E4c) ont été croisés ensemble de sorte à obtenir la F2 de la lignée zfc9^{KO} pour laquelle il serait finalement possible d'obtenir des spécimens HM pour le KO de zfc9.

2.3.7 Crible des spécimens de la F2

Afin d'identifier des poissons HM pour les délétions entraînant le KO de zfc9, les spécimens résultant du croisement des individus de la F1 ayant survécus jusqu'à la taille appropriée ont été génotypés par analyse HRM, suite au prélèvement de nageoire caudale tel que décrit précédemment. En date de la rédaction de ce mémoire, 57 descendants des poissons F1 ont été criblés. De ceux-ci, 23 étaient WT (40.35%), 24 se sont révélés HT (42.11%) et 10 (17.51%) d'entre eux se sont avérés HM pour les délétions. Parmi les 10 spécimens HM, 7 se sont révélés être des femelles et 3 des mâles (**Tableau 5**). L'obtention de la F3 et donc d'une lignée HM stable zfc9^{KO} sera complétée une fois qu'un croisement entre les HM mâles et femelles de la F2 engendrera une progéniture entièrement HM apte à atteindre l'âge de reproduction.

Tableau V. **Résultats du génotypage des poissons F2 de la lignée $zfC9^{KO}$ et proportion des génotypes.** Récapitulatif des résultats du génotypage par analyse HRM sur l'ADN génomique des poissons F2 juvéniles de la lignée $zfC9^{KO}$ qui ont été obtenus du croisement des poissons F1 E4(a,b,c). Total est le nombre de poissons F2 génotypé par analyse HRM. %WT, %HT et %HM sont la proportion en pourcentage des spécimens de chacun des génotypes correspondants.

Total	WT	%WT	HT	%HT	HM	%HM
57	23	40.35	24	42.11	10	17.54

2.4 Résultats préliminaires pour la caractérisation de l'effet du KO de $zfC9$

Nous n'avons pas été en mesure d'obtenir une lignée HM $zfC9^{KO}$ stable avant l'échéance de cette maîtrise, mais nous avons tout de même cherché à obtenir des résultats préliminaires concernant un possible phénotype du KO de $zfC9$, de sorte à guider nos futures expériences une fois ceux-ci obtenus. En l'absence d'une source stable de spécimens uniquement HM (croisement HM/HM) et de spécimens strictement HT (croisement WT/HM), nous avons tenté d'investiguer un possible phénotype de mortalité accrue et de déficits de fonctions motrices associés au KO de $zfC9$ à l'aide de test préliminaires sur la progéniture de la F1 de la lignée $zfC9^{KO}$.

2.4.1 Test statistique des proportions de génotypes pour la progéniture de la F1

Les spécimens F1 de la lignée $zfC9^{KO}$ ont été croisés à mainte reprise de sorte à nous assurer des chances raisonnables d'y identifier là des spécimens HM pour le KO de $zfC9$. Cependant, une mortalité considérable a été constatée avec une consistance intrigante et problématique. Nous avons aussi observé que les proportions des génotypes identifiés par le criblage de la progéniture adulte de la F1 de la lignée $zfC9^{KO}$ différaient des ratios de 25% WT, 50% HT et 25% HM attendus selon les lois de Mendel [238] sur la transmission des gènes entre les générations. Nous avons alors commencé à considérer que le KO de $zfC9$ puisse possiblement être mal toléré par les larves du poisson-zèbre durant leur développement, menant à la mort

durant cette période d'un nombre disproportionné des spécimens HM et HT. Pour obtenir des résultats préliminaires sur cette hypothèse, nous avons tenté de déterminer si la disparité entre les proportions de génotypes obtenues et ceux attendues était statistiquement significatives. Pour ce faire, nous avons posé l'hypothèse nulle selon laquelle les proportions des génotypes WT, HT et HM obtenues ne différaient pas des proportions attendues puis nous avons testé cette dernière à l'aide d'un test statistique du χ^2 d'adéquation. Nous en avons obtenus une valeur-p de 0.025 nous permettant de rejeter l'hypothèse nulle avec un intervalle de confiance supérieur à 95% (**Tableau 6**).

Tableau VI. **Résultats du test χ^2 d'adéquation permettant de tester si les proportions des génotypes identifiés est comparable à ceux prédits par la théorie.** Le nombre de spécimens de chaque génotypes (WT, HT et HM) obtenus pour la progéniture d'un croisement HET/HET *zfc9*^{KO} est comparé aux valeurs prédites par les lois de Mendel à l'aide du test χ^2 d'adéquation pour tester l'hypothèse nulle que les valeurs obtenues sont comparables aux valeurs attendues. Une valeur-p ≤ 0.95 nous permettant de réfuter l'hypothèse nulle avec confiance. WT : type sauvage, HT : hétérozygote, HM : homozygote, Vobt : valeurs obtenues, Vatt : valeur attendue, $\chi^2 : \sum (Vobt-Vatt)^2/Vatt$, ddl : degré de liberté.

Catégories	Vobt	Vatt	Vobt-Vatt	(Vobt-Vatt) ²	(Vobt-Vatt) ² /Vatt	χ^2	ddl	Valeur-p
WT	23	14.25	8,75	76.56	5.37	7.35	2	0.025
HT	24	28.5	-4,5	20.25	0.11			
HM	10	14.25	-4,25	18.06	1.27			

Le rejet de l'hypothèse nulle signifie que les proportions de spécimens des trois génotypes possibles, obtenues suite au criblage de la progéniture des poissons HT/HT KO pour *zfc9*, diffèrent de manière statistiquement significative de celles attendues; les spécimens WT étant surreprésentés alors que les spécimens HT et HM sont sous représentés. Malgré les nombreuses limitations associées à notre emploi de ce test statistique, ce résultat préliminaire suggère une certaine mortalité accrue des spécimens avec un KO partiel ou complet de *zfc9*.

2.4.2. Comparaison de l'activité de nage des WT, HT et HM pour le KO de *zfc9*

Nous avons aussi observé qualitativement chez la progéniture de croisement HT/HT KO pour *zfc9* que plusieurs des larves semblaient se distinguer du reste par une nage moins active lorsque observée dans les aquariums d'élevage dans les jours précédant les phases de mortalité importante (12 à 20 jours). Ceci, avec les observations rapportées que le KD transitoire de *zfc9*, s'accompagne de déficits moteur chez les larves de poisson-zèbre [105,213], ce qui nous a mené à tenter de tester de manière préliminaire si les larves HT et HM pour le KO persistant de *zfc9* présentent elles aussi des déficiences motrices. Pour ce faire, nous avons cherché à quantifier puis comparer la distance de nage à intervalle régulier pour une période de temps étendue après un stimulus pour des larves des trois génotypes différents. Ne possédant pas de spécimens HM en mesure de se reproduire de sorte à fournir des larves strictement HT ou HM pour le KO de *zfc9* au moment d'exécuter cette expérience, nous avons travaillé à nouveau avec la progéniture d'un croisement HT/HT de spécimens KO pour *zfc9*. Il a donc été nécessaire de génotyper les larves individuellement après chaque test de nage afin de pouvoir attribuer les résultats obtenus aux spécimens correspondants. Pour y arriver nous avons employé l'analyse HRM sur l'ADN génomique extrait de larve entière comme décrit précédemment pour les embryons. Le profil de nage pour chaque larve a été obtenu à l'aide de l'appareil DaniVision (Nodus) qui permet l'enregistrement puis l'interprétation de multiples profils de nage de larves de poisson-zèbre dans des arènes individuels. Les larves ont été testées à 10 jours après leur naissance, soit quelques jours avant la période typique de mortalité importante observée pour la progéniture de la F1 lignée *zfc9*^{KO} (12 à 20 jours). Ceci a été fait dans l'espoir d'accroître nos chances d'observer une différence significative entre les larves WT, HT et HM, puisque l'observation de larves avec une nage plus amorphe dans ces cohortes devenait apparente seulement peu avant que nous commencions à observer la mortalité typique de ce croisement. Passé 10 jours, la mort des larves durant le test devient plus probable, ce qui affecte les résultats du test de nage et nécessite de les retirer pour ne pas biaiser les données à la baisse. Le profil de test de base employé débutait par une période de repos dans l'obscurité suivi d'un stimulus par l'éclairage avec une lumière blanche persistante jusqu'à la fin du test, la période dans l'obscurité servant de point zéro/repos pour normaliser les comportements de nage avant le début du test, par suite

du stress infligé aux larves durant leur transfert. Du test avec le plus grand nombre de spécimens et la meilleure répartition des 3 génotypes, nous n'avons observé en condition de repos dans l'obscurité aucune différence statistiquement significative pour la distance de nage, mais une fois l'éclairage activée, on observe une période relativement courte d'activité élevée (environ 30 min, **Figure 15**) pour laquelle les larves de génotype WT se distinguent de manière évidente et statistiquement significative des larves de génotype HT ou HM pour le KO de *zfc9* (**Figure 15**). Une fois cette période de haute activité passée, la nage devient moins active pour les spécimens de chacun des 3 génotypes puis l'écart d'activité entre les WT et les HT/MH s'amenuise jusqu'à cesser d'être statistiquement significatif pour la vaste majorité des périodes de temps considérées. (**Figure 15**). La distance de nage des HT et des HM pour le KO de *zfc9* est comparable statistiquement, bien que de manière générale on semble observer une nage moins active pour les spécimens HM que les HT.

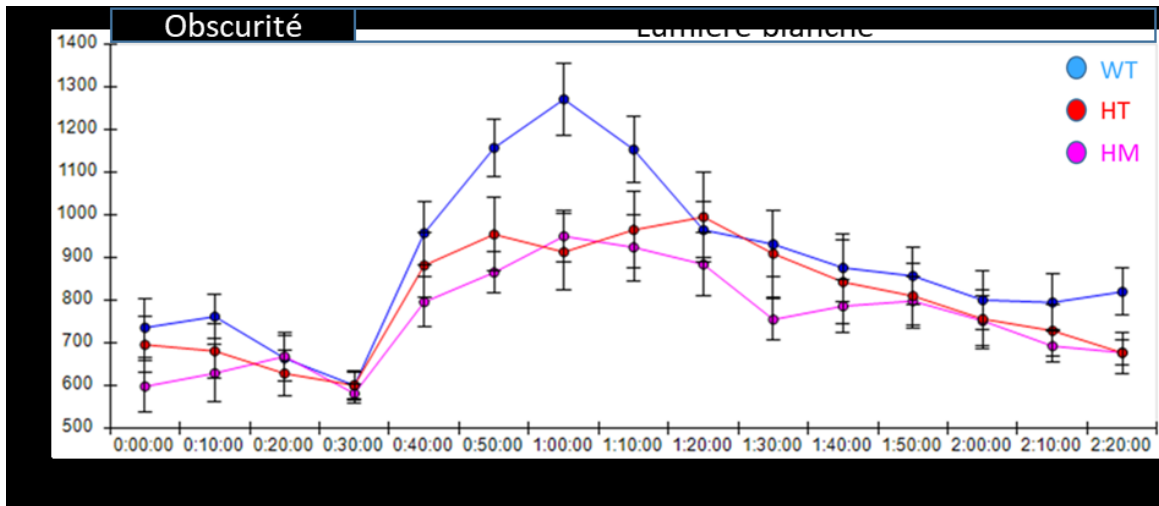


Figure 15. Comparaison de la distance de nage totale des spécimens WT, HT et HM pour le KO de *zfc9*. L'activité de nage de larve WT, HT et HM pour le KO de *zfc9* a été mesurée par l'analyse de la distance totale parcourue avec comme référence le centre des puits dans lesquels reposent les larves et ce à des intervalles réguliers de 10 min, répartis sur une période d'obscurité de 30 min, suivie d'un éclairage persistant à la lumière blanche pendant 120 min. L'activité de nage pour les larves des trois génotypes ne diffère pas de manière statistiquement significative durant les 30 min d'obscurité initiale, où leur activité se normalise à la baisse alors qu'ils entrent dans une phase de repos. Passé 30 min, la lumière est activée et les larves des trois génotypes présentent une activité accrue de nage en réponse au stimulus, atteignant un plateau après environ 30 min, avant de décroître progressivement jusqu'à se stabiliser vers la fin de la période de test. La différence entre l'activité des WT et les HT/HM devient statistiquement significative pour 4 périodes de 10 min suivant l'activation de la lumière (0 :50 :00 à 01 :10 :00 et 2 :20 :00). L'activité des HM et HT ne diffère de manière statistiquement significative à aucun point. De manière générale, les valeurs d'activités des WT sont les plus élevées, suivies par celles des HT, alors que celles des HM tendent à être les plus basses. Les barres indiquent l'erreur standard; la distance est en millimètres (mm); les périodes de temps sont notées heures (h): minutes (m) : secondes (s). Le nombre de spécimens testé pour chaque génotype est : 18 HM, 22 HT et 17 WT. Barre noire : période d'obscurité; Barre blanche : période d'éclairage blanc.

2.5 Génération de lignées transgéniques exprimant le gène *HC9*

Diverses considérations nous ont amenées à chercher à générer des lignées transgéniques de poissons zébrés en mesure d'exprimer le *HC9*. Notamment, il a été rapporté que l'expression de *HC9* exogène dans des cellules humaines ainsi que dans des modèles animaux permet de secourir les phénotypes délétères associés au KD/KO de *C9orf72* [95,102,105,213].

Additionnellement, comme nous cherchions à générer une lignée $zfC9^{KO}$ et avec la possibilité, que même si réussi le marquage de $zfC9$ pourrait entraîner une perturbation de sa fonction, il a été considéré intéressant d'être en mesure de vérifier si l'expression persistante de $HC9$ pourrait secourir de manière durable les possibles effets délétères d'un KO de $zfC9$. De plus, bien que sa surexpression transitoire semble bien tolérée, peu d'information est disponible actuellement sur l'effet de la surexpression persistante de $HC9$ ou de ses orthologues. L'étude de potentielles conséquences de la surexpression des orthologues de $C9$ s'avérerait utile aussi pour la détermination de sa fonction et pour baliser la recherche, qui considère le potentiel thérapeutique de rétablir ou d'accroître le niveau d'activité de $C9$ ou des voies dans lesquelles le gène est impliqué. L'utilisation du transgène humain a été motivée par l'existence et la disponibilité de plusieurs anticorps commerciaux dont l'efficacité de détection a été établie, permettant ainsi sa détection et sa quantification relativement aisées dans différents modèles de poissons zébrés qui l'expriment sans nécessité de risquer de le marquer avec un tag moléculaire [82].

2.5.1 Construction génétiques et micro-injections

Nous avons donc cherché à générer deux lignées transgéniques avec l'aide du système de transposon Tol2. La première lignée, HC9tg, était prévue d'exprimer le gène $C9orf72$ humain WT, tandis que la deuxième lignée, HC9-eGFPtg, exprimerait une protéine de fusion formée du gène humain $C9orf72$ WT et le tag fluorescent eGFP, le tag eGFP étant employé pour ses nombreux avantages décrits précédemment à la **section 2.2**, dont au niveau de l'imagerie in vivo. Le choix de générer une lignée transgénique exprimant $HC9$ WT étant lié aux avantages de la disponibilité d'anticorps pour sa détection et du risque, en insérant un tag ou en l'exprimant sous forme de protéine de fusion, d'affecter sa fonction ou sa stabilité de manière difficile à prédire en raison du peu d'information sur la structure et le domaine fonctionnelle DENN de sa protéine. L'expression des transgènes a été placée sous le contrôle du système Gal4-UAS, permettant potentiellement de contrôler dans quels tissus et à quel moment le transgène est exprimé. En effet, il existe plusieurs lignées poissons-zèbres exprimant le promoteur Gal4 sous le contrôle de différents promoteurs spécifiques à certaines populations de cellules dont la lignée Gal9, disponible dans notre laboratoire, pour laquelle l'expression de Gal4 est sous le contrôle du facteur de transcription des motoneurones MNX1, aussi connus sous le nom de HB9, permettant que Gal4 ne soit exprimé que dans les motoneurones. D'autres lignées encore placent

le contrôle de l'expression de Gal4 sous des promoteurs inductibles, comme le promoteur inductible de choc thermique *hsp70*, permettant de contrôler le moment du développement ou de la vie adulte du poisson-zèbre à lequel Gal4 est exprimé. De plus, des lignées exprimant de manière ubiquitaire Gal4 sont disponibles si une expression généralisée du transgène est désirée. Nous avons décidé d'insérer initialement ces transgènes dans des lignées TL WT n'exprimant pas Gal4 et donc pour lesquelles le transgène ne serait pas exprimé. Si cette décision rendait plus longue la génération de lignées exprimant le transgène, elle ne l'a pas été sans raisons. En effet, de cette façon une fois générée, les lignées HC9tg et HC9-eGFPtg pourront être croisées rapidement et sans contraintes avec différentes lignées exprimant Gal4 dans différentes populations cellulaires ou sous différentes conditions selon les expériences envisagées. De plus, la toxicité de la surexpression du transgène étant inconnue, ceci nous permet de nous assurer qu'une potentielle toxicité de l'expression des transgènes complique l'obtention des lignées. Enfin, chaque construction pour nos lignées transgéniques contient aussi un deuxième gène servant de gène rapporteur, permettant d'identifier aisément rapidement les embryons transgéniques. Le gène rapporteur *GFP* associé au promoteur *cmlc2* (cardiac myosin light chain 2), entraîne une expression constitutive de GFP dans le cœur des poissons transgéniques qui est observable par stéréomicroscope à épifluorescence aussitôt qu'à la 27^{ième} heure après la fertilisation des embryons (**figure 16**). Les constructions pTOL2-UAS-hC9orf72 (HC9tg) et pTOL2-UAS-hC9orf72-gfp (HC9-eGFPtg) ont été co-injectées avec l'ARNm de la transposase Tol2 dans des embryons fertilisés alors qu'ils étaient au stade 1 à 2 cellules.

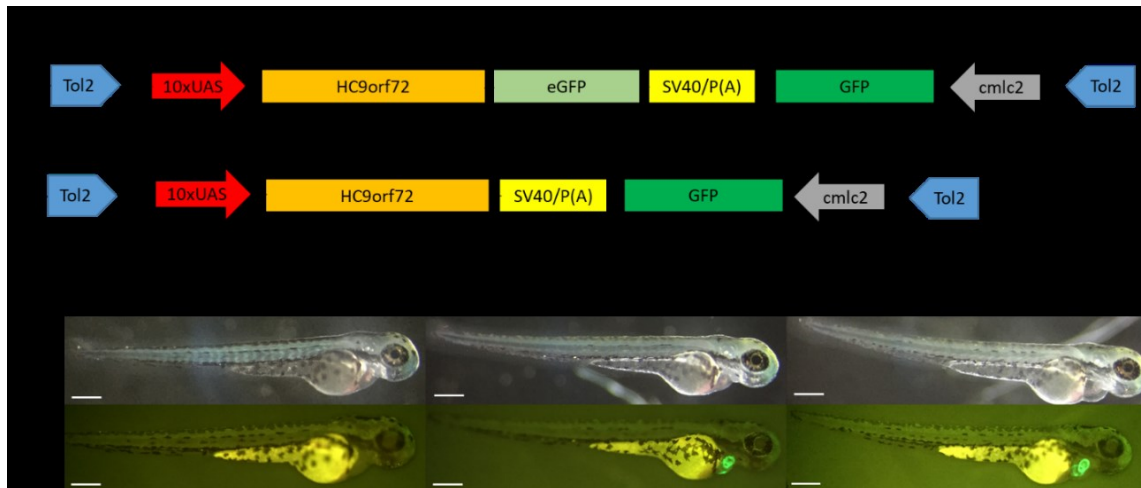


Figure 16. Constructions génétiques pour la génération des lignées HC9tg et HC9-eGFP et expression constitutive du GFP démontrant leur insertion. (A) Constructions Tol2 injectées avec la transposase Tol2 pour la lignée HC9-eGFPtg (1) et HC9tg (2) respectivement. Les gènes humains *C9orf72* WT conjugués à eGFP ou sans sont liés à 10 répétitions séquences d'activation en amont (upstream activator sequence, UAS) auxquels peuvent se lier le facteur de transcription Gal4 pour initier la transcription des transgènes. Le gène rapporteur GFP est sous le contrôle du promoteur cardiaque *cmlc2*. Chaque transgène est achevé par un signal de polyadénylation SV40 (SV40/P(A)). (B) L'expression constitutive de GFP au niveau du cœur indique l'insertion du transgène chez un embryon pour les lignées HC9tg et HC9-eGFPtg. Échelle (barres blanches) : 250 μ m. Embryons de 48hpf.

2.5.2 Sélection des poissons fondateurs F0

Après les micro-injections, les embryons de 30 à 72 hpf ont été observés le matin et le soir sous microscope à épifluorescence pour repérer une fluorescence au niveau du cœur. Comme le gène rapporteur *GFP* est présent sur le même plasmide pour chaque gène d'intérêt, la présence de fluorescence au niveau du cœur indique l'intégration du gène rapporteur et par conséquent des transgènes *HC9* ou *HC9-eGFP* dans certaines cellules cardiaques de façon mosaïque. Les embryons avec une forte mosaïcité dans le cœur ont été sélectionnés, puisque celle-ci suggère une plus forte possibilité d'une mosaïcité d'insertion du transgène correspondant dans les cellules germinales, une insertion qui sera nécessaire encore une fois pour engendrer des générations subséquentes ayant au moins une insertion stable dans l'ensemble de leurs cellules. Ces embryons exprimant GFP au niveau du cœur ont donc été sélectionnés pour former la génération F0 des lignées HC9tg ou HC9-eGFPtg, puis ils ont été élevés jusqu'à leur maturation

sexuelle tel que décrit précédemment à la section **2.2.5**. Les injections étaient étalées sur deux semaines pour chaque cohorte et s'arrêtaient une fois obtenu 100 embryons avec de la fluorescence sélectionnés pour chaque lignée transgénique. Au total, 2 cohortes ont été sélectionnées pour nos deux lignées transgéniques.

2.5.3 Crible des spécimens F0 pour la transmission de chaque transgène

Sur les 200 embryons sélectionnés au total pour chacune de nos cohortes, ~105 ont atteint l'âge adulte pour la lignée HC9tg et ~120 pour la lignée HC9-GFPtg. Contrairement aux autres lignées que nous avons générées, ces nombres demeurent approximatifs puisqu'il n'a pas été nécessaire de retirer des aquariums communs de chaque cohorte tous les spécimens survivants pour identifier des progéniteurs pour la F1 de nos deux lignées transgéniques et leur décompte dans ces aquariums est laborieux, nécessitant de retirer chaque individu un à un. Le criblage pour la transmission du transgène s'est effectué de manière identique pour nos deux lignées transgéniques. Chaque poisson F0 ayant survécu et rencontrant les critères minimaux décrits précédemment (**section 2.2.5**) était croisé avec deux spécimens de TL WT de sexe opposé, la transmission des transgènes était évaluée de la même manière que pour la sélection des embryons suite aux injections, à l'exception qu'une fluorescence complète du cœur était attendue. Pour chaque poisson, un minimum de 120 embryons devait être observé sans fluorescence passée les 72hpf pour être déclassé et considéré incapable de transmettre le transgène ou d'avoir une efficacité de transmission trop faible pour être pratique. Lorsqu'un embryon avec de la fluorescence était identifié, le poisson F0 progéniteur était recroisé jusqu'à réobtenir un autre embryon au cœur fluorescent ou déclassé après une observation de 240 embryons additionnels sans fluorescence.

Pour les candidats de la F0 de la lignée HC9tg, à l'échéance de ce projet de maîtrise 44 spécimens ont été déclassés complètement, 13 sont en attente d'être recroisés après avoir criblé plus de 60 embryons sans trouver de fluorescence et deux femelles ont été identifiées comme aptes à transmettre une fluorescence du cœur et donc du transgène à leur progéniture (**Tableau 7**).

Dans le cas de la lignée HC9-eGFPtg, au terme de ce projet de maîtrise, 39 spécimens ont été déclassés complètement, 17 sont en attente d’être recroisés après avoir criblé plus de 60 de leurs embryons sans trouver de fluorescence et deux spécimens de sexe différent ont été identifiés comme apte à transmettre une fluorescence du cœur et donc du transgène à leur progéniture (**Tableau 7**).

Tableau VII. **Résultats du criblage des poissons F0 des lignées transgéniques pour la transmission du transgène.** Résultats du criblage des poissons F0 adultes des lignées transgéniques HC9tg et HC9-eGFPtg pour la transmission de leur transgène respectif. Déclassé : nombre de spécimen F0 criblé pour 120 embryons sans fluorescence ; ½déclassé : nombre de poissons criblés pour au moins 60 embryons sans fluorescence sans avoir encore été déclassés; + : poissons en mesure de transmettre une fluorescence au niveau du cœur à et donc le transgène à sa progéniture; ♀ : poissons femelles adultes ; ♂ : poissons adultes mâles.

<u>HC9tg</u>						<u>HC9-eGFPtg</u>					
Déclassé	½Déclassé	Reste	+	♀+	♂+	Déclassé	½Déclassé	Reste	+	♀+	♂+
44	13	~48	2	2	0	39	17	~64	2	1	1

2.5.4 État de la lignée HC9tg

Pour la lignée HC9tg, la première des femelles (♀HC9-1) a montré un taux de transmission d’approximativement 1 embryon fluorescent par 30 embryons observés (~3.33%) lorsque croisée avec des TL WT. La seconde femelle (♀HC9-2) a démontré un taux plus faible de transmission d’environ 1 embryon fluorescent tous les 120 embryons observés (~0.83%). ♀HC9-1 a donc été initialement sélectionnée pour engendrer la F1 de la lignée HC9tg puisque l’obtention d’embryon fluorescent de la seconde était irrégulière dans le meilleur des cas. Cependant, ♀HC9-1 a été retrouvée morte dans son aquarium un matin après que seulement 9 de ses embryons aient pu être sélectionnés (**Tableau 8**). Nous avons alors entrepris de recommencer à croiser la ♀HC9-2 dans l’espoir de pouvoir obtenir au moins quelques spécimens adultes pour lancer la F2 de la lignée HC9tg si les spécimens obtenus de ♀HC9-1 venaient tous à décéder.

Au terme de cette maîtrise, 4 des 9 embryons de ♀HC9-1 ont atteint l'âge adulte dont 3 femelles et 1 mâle alors que 4 des 6 spécimens sélectionnés de ♀HC9-2 ont atteint le stade juvénile (**Tableau 8**). Les 4 spécimens adultes de ♀HC9-1 ont été confirmés d'être en mesure de transmettre, tel qu'attendue, la fluorescence du cœur à leur progéniture dans des proportions mendélienne lorsque croisés avec des TL WT. Jusqu'à présent, 112 embryons nés du croisement des femelles F1 obtenus et du seul mâle ont été sélectionnés pour leur fluorescence du cœur afin de s'assurer de la subsistance de la lignée (**Tableau 8**).

Tableau VIII. **État de la lignée HC9tg au terme du projet de maîtrise.** Nombre de poissons vivants adultes et en développement des différentes générations F0, F1 et F2 de la lignée HC9-eGFPTg qui ont intégrés le transgène et sont aptes à le transmettre à leur progéniture à la date d'échéance de ce projet de maîtrise. ♀(+) et ♂(+) respectivement : poissons adultes femelles et poissons adultes mâles apte à transmettre le transgène dans des proportions non mendéliennes; ♀+ et ♂+ : respectivement : poissons adultes femelles et poissons adultes mâles aptes à transmettre le transgène dans des proportions mendéliennes; juvénile : poisson au stade juvénile en élevage; larves : poissons au stade larvaire en élevage.

F0		F1♀ HC9-1		F1♀ HC9-2	F2 ♀HC9-1
♀(+)	♂(+)	♀+	♂+	Juvéniles	Larves
1	0	3	1	4	112

2.5.5 État de la lignée HC9-eGFPTg

En ce qui concerne la lignée HC9-GFPTg, malgré l'identification de deux spécimens de sexe différent en mesure de transmettre la fluorescence à leur progéniture, seulement la femelle a pu être employée pour engendrer la F1 puisque le mâle est décédé lors d'un transfert d'aquarium pour le croiser (**Tableau 9**). La femelle (HC9-eGFPTg-1), lorsque croisée avec des WT TL, présentait un taux de transmission d'environ 2 spécimens par 30 embryons observés (~6.66%), elle a donc été croisée à plusieurs reprises pour tenter d'obtenir le plus d'embryons fluorescents possibles ayant intégrés le transgène pour être en mesure d'établir une F2 stable de la lignée HC9-GFPTg. Cependant, les croisements ont été laborieux et rarement productifs puisque cette femelle souffre de difformités au niveau de la colonne vertébrale, limitant son activité durant

les croisements et forçant une supervision active durant ceux-ci pour éviter qu'elle ne soit blessée par les mâles.

En date de la fin de ce projet de maîtrise, 31 embryons fluorescents avaient été sélectionnés puis élevés de la femelle F0 de la lignée HC9-GFPtg. De ceux-ci 14 sont décédés, 6 sont encore au stade juvénile et 11 ont atteints l'âge adulte, dont 4 mâles et 7 femelles (**Tableau 9**). L'ensemble des 11 spécimens adultes survivants ont été confirmés comme transmettant la fluorescence au niveau du cœur à leur descendant dans des proportions mendéliennes. Du croisement de ces spécimens adultes forment la F1 de la lignée HC9-GFPtg, 143 embryons avec de la fluorescence au niveau du cœur ont été sélectionnés pour être élevés et assurer la subsistance de la lignée (**Tableau 9**).

Tableau IX. **État de la lignée HC9-eGFPtg au terme du projet de maîtrise.** Nombre de poissons vivants adultes et en développement des différentes générations F0, F1 et F2 de la lignée HC9tg qui ont intégrés le transgène et sont aptes à le transmettre à leur progéniture à la date d'échéance de ce projet de maîtrise. ♀(+) et ♂(+) respectivement : poissons adultes femelles et poissons adultes mâles aptes à transmettre le transgène dans des proportions non mendéliennes; ♀+ et ♂+ respectivement : poissons adultes femelles et poissons adultes mâles aptes à transmettre le transgène dans des proportions mendéliennes; juvénile : poisson au stade juvénile en élevage; larves : poissons au stade larvaire en élevage.

<u>F0</u>		<u>F1 ♀ HC9-eGFPtg-1</u>			<u>F2 HC9tg-1</u>
♀(+)	♂(+)	♀+	♂+	Juvéniles	Larves
1	0	7	4	6	143

Puisque le transgène de la lignée HC9-eGFPtg permet l'expression de la protéine de fusion HC9-eGFP en présence de Gal4, il était assumé qu'il serait aisé de vérifier son expression visuellement chez des embryons résultant d'un croisement de spécimens F1 de cette lignée avec des spécimens d'une lignée exprimant Gal4, puis en repérant le patron de fluorescence attendue des tissus exprimant Gal4 dans la lignée utilisée. Nous avons donc croisé les spécimens F1 avec des spécimens d'une lignée exprimant Gal4 de manière constitutive pour confirmer rapidement et sans équivoque l'expression de HC9-eGFP dans notre lignée transgénique. Cependant, de la

progéniture obtenue de ces croisements, aucune fluorescence décelable sous microscope à épifluorescence n'a été observée à 24, 36 et 72 hpf autre que la fluorescence au niveau du cœur des embryons sélectionnés. Pour vérifier que la lignée exprimant Gal4 utilisée permettait bien son expression ubiquitaire, elle a été croisée avec une autre lignée dont l'expression de GFP était sous le contrôle du système Gal4-UAS sans exprimer Gal4 elle-même. La présence de fluorescence dans la progéniture de ce croisement a confirmé que la lignée Gal4 exprimait bien le facteur de transcription. Supposant une possible dégradation rapide de la protéine de fusion comme étant la raison pour laquelle nous ne pouvions observer la fluorescence associée à son expression, nous avons injecté directement une solution d'ARNm de Gal4 dans des embryons au stade 1-2 cellules provenant du croisement entre les mâles et les femelles F1 de la lignée HC9-eGFPtg. Notre raisonnement étant que la protéine de fusion s'était possiblement dégradée trop rapidement pour que son accumulation produise un signal de fluorescence décelable, une surexpression du transgène par une injection d'une concentration significativement plus élevée de Gal4 que ce qui est retrouvé normalement in vivo dans nos lignés Gal4 permettrait alors peut-être l'observation de son expression pendant un court laps de temps. Les embryons injectés ont donc été observés sous microscope à épifluorescence à 20 hpf, pour éviter que le Gal4 injecté soit entièrement dégradé, tout en permettant à l'embryon d'être suffisamment développé pour qu'on puisse observer son expression dans au moins quelques structures différenciées et une fluorescence à peine perceptible a été observée. Pour confirmer cette observation, des embryons présentant cette faible fluorescence ont été observés avec un microscope confocale à balayage laser (Quorum Technologies) à une longueur d'onde appropriée pour l'observation de la fluorescence d'eGFP (491nm). Cette observation a révélé la présence d'un faible signal de fluorescence chez les embryons provenant du croisement des mâles et des femelles F1 de la lignée HC9-eGFPtg (**Figure 17**). Le signal observé semble plus intense au niveau des yeux et de la tête (encéphale), mais on détecte une fluorescence présente, bien que plus diffuse, dans le reste de l'embryon (**Figure 17**). Cette fluorescence est aussi distincte du signal du gène GFP rapporteur de l'insertion du transgène qui est exprimé dans le cœur et qui généralement ne devient visible qu'environ 27hpf. Enfin, aucun signal de fluorescence n'a été observé chez des embryons TL WT de 20hpf injectés avec la même solution Gal4 au stade 1-2 cellules (**Figure 17**). La présence d'un tel signal de fluorescence tend à confirmer que le transgène est bien

intégré dans les spécimens de la lignée HC9-eGFPTg et que la protéine de fusion HC9-eGFP peut être exprimé chez des poissons surexprimant Gal4.

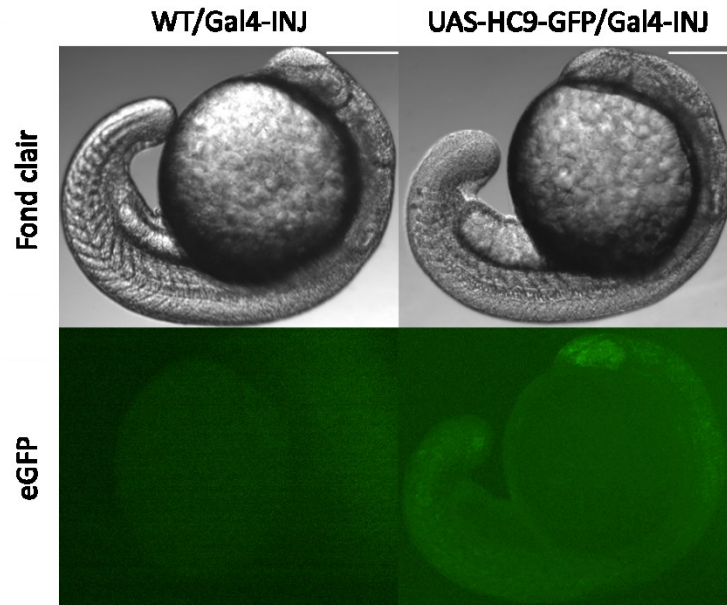


Figure 17. Détection de fluorescence dans des embryons de la lignée HC9-eGFPTg injectés avec une solution de Gal4. Un faible signal de fluorescence est détecté chez des embryons 20hpf obtenus du croisement de mâles et de femelles F1 de la lignée HC9-eGFPTg qui ont été injectés au stade 1-2 cellules avec ~1nl d'une solution de 10ng/ μ l d'ARNm de Gal4. Ce signal de fluorescence, distinct de celui attendu du gène GFP rapporteur exprimé dans le cœur, est principalement retrouvé à la tête, mais il est aussi faiblement présent dans le reste de l'embryon. Aucun signal n'est observé chez les embryons TL WT 20 hpf injectés de manière identique avec la même solution d'ARNm de Gal4. L'échelle (barres blanches) est de 100 μ m; le laser pour l'observation d'eGFP a une longueur d'onde de 491nm; les embryons sont retirés de leur chorion et fixé dans de l'agarose "Low Melting Point" avant d'être observés.

3. Discussion

Au cours de ce projet de maîtrise nous avons généré des sondes monobrins d'ARN antisens nous permettant d'observer par hybridation *in situ* (HIS) le profil d'expression du gène *zfc9* dans des embryons à différents stades du développement incluant 24, 36 et 48 hpf. Il a aussi été tenté sans succès de générer trois lignées différentes intégrant un tag moléculaire à l'extrémité Nt ou Ct par système d'édition du génome CRISPR/Cas9, soit la lignée C9orf72^{NHis} devant intégrer un tag 12xHis au Nt, la lignée C9orf72^{NGFP} qui aurait intégré une étiquette moléculaire eGFP au Nt et enfin la lignée C9orf72^{CGFP} supposée intégrer un tag eGFP au Ct. De plus, nous sommes parvenus à identifier un progéniteur capable de transmettre à sa descendance des indels dans la région Nt de *zfc9* prédites de provoquer un KO fonctionnel des trois isoformes de la protéine du gène *zfc9*. De ce progéniteur, il a été obtenu jusqu'à la F2 une lignée de poissons-zèbres KO pour *zfc9*. De plus, des tests préliminaires effectués pour la lignée *zfc9*^{KO} suggèrent que le KO persistant de *zfc9* s'accompagnerait d'un déficit moteur caractérisé par une activité de nage diminuée chez les larves HT ou HM, en plus de provoquer une mortalité accrue des spécimens de ces deux même génotypes. Additionnellement, nous avons été en mesure d'obtenir des spécimens adultes HT exprimant un gène rapporteur GFP indiquant l'intégration par système Tol2 des constructions transgéniques pour nos lignées HC9tg et HC9-eGFPtg, qui devraient permettre une expression sous le contrôle du système Gal4-UAS du gène *HC9* WT et HC9 marqué avec eGFP. Enfin, nous avons confirmé par injections d'ARNm de Gal4 dans des embryons HC9-eGFPtg la capacité d'expression de de HC9-eGFP bien qu'elle semble dégradée.

3.1 Profil d'expression du gène *zfc9* aux stades 24, 36 et 48hpf

Les profils d'expression du gène *zfc9* obtenu de nos sondes monobrin d'ARN antisens pour les embryons de stade 24hpf, soit une expression relativement élevée au niveau des tissus du SNC incluant le proencéphale, le mésencéphale, le rhombencéphale et la moelle épinière au niveau de la notochorde, récapitulent dans l'ensemble les observations faites par Ciura et al., 2013 à ce stade [105]. De plus, comme eux, nous n'avons pas observé de signal décisivement distinct du bruit de fond dans les tissus hors du SNC à ce stade. Comme attendu, nos observations des profils d'expression de *zfc9* aux stades 36hpf et 48hpf montrent aussi une expression relativement forte et soutenue dans la plupart des tissus du SNC, incluant plusieurs structures distinctes qui peuvent être identifiées au stade 48hpf, dont le nerf optique, le bulbe olfactif, le ganglion trigéminal et le circuit habénulaire. Les éléments plus intéressants cependant sont la

force du signal observé dans le tronc au niveau de la notochorde qui, bien qu'il demeure relativement fort, diminue en intensité relative du stade 24hpf aux stades 36-48hpf et l'apparition d'un signal diffus dans les tissus hors du SNC qui semble se renforcer du stade 36hpf au stade 48hpf. Ces observations indiquent que le profil d'expression du gène *zfc9* varie en intensité et possiblement dans sa distribution spatiale durant cette période du développement du poisson-zèbre, particulièrement au niveau de la moelle épinière. Ceci revêt un certain intérêt lorsque l'on considère que les phénotypes délétères rapportés pour les KD transitoires de *zfc9*, durant les phases de développement précoce, lui suggèrent un rôle possible dans la motilité et le neurodéveloppement, dont notamment au niveau des motoneurones [105,213]. De plus, ces observations s'accordent avec les observations, chez le modèle murin, que l'expression des isoformes de l'orthologue de *HC9* de la souris changent en intensité à plusieurs reprises au cours du développement et de la phase de développement au stade adulte [239].

3.1.1 Complications rencontrées et améliorations pour de futures expériences d'HIS

Un problème rencontré au cours des expériences d'HIS pour observer le profil d'expression de *zfc9* a été la faiblesse relative du signal obtenu par chacune de nos sondes, qui était talonné par le signal du bruit de fond. En effet, la période de révélation pour nos sondes s'étalait souvent sur près de 4 heures alors que la majorité des gènes exprimés ne nécessitent qu'une période d'exposition à la solution de révélation de 1 à 1.5 heure, comme cela fut le cas pour nos contrôles positifs. Cette faiblesse du signal peut vraisemblablement être attribuée à la faible expression rapportée du gène *HC9* [91], qui est semblerait-il récapitulé pour *zfc9* [240] et ce même dans les régions où son expression est relativement élevée. Puisque la révélation a été stoppée lorsque le signal de bruit de fond risquait de compromettre la qualité et l'interprétation du marquage HIS, il est possible qu'il y ait une expression de *zfc9* dans les tissus hors du SNC, même au stade 24hpf, mais que l'expression soit si faible que sa distinction du bruit de fond ne puisse se faire clairement avec la méthode que nous avons employée. La génération de sondes marquées à la DIG plus longues, des sondes d'HIS de jusqu'à 1.5 kilo bases ayant été utilisées avec succès, permettrait non-seulement d'accroître la spécificité du signal, mais aussi de renforcer celui-ci, puisque des sondes plus longues comportent plus de sites marqués à la DIG pour lier l'anticorps et ce pour la liaison d'une même copie de l'ARNm de *zfc9* [231]. Cependant, la pénétrance des

sondes diminue avec leur taille, au risque de résulter en un marquage incomplet ou simplement absent si leur longueur est trop importante [231]. C'est pourquoi nos sondes ont été conçues avec des longueurs oscillants entre les deux extrêmes. L'optimisation de la perméabilisation à la protéinase K (concentration/temps d'incubation) est aussi liée à l'intensité du signal et du bruit de fond. Plus la perméabilisation est extensive, plus le signal sera clair et moins le bruit de fond sera important, mais le risque de compromettre la structure des tissus s'accroît avec le degré de perméabilisation [241]. Une optimisation de la perméabilisation pour chaque stade observé serait alors une avenue pour réduire l'intensité du bruit de fond lors d'expériences futures d'HIS pour l'observation du profil d'expression de *zfc9*. L'utilisation de sondes fluorescentes serait une autre alternative pour la détection efficace de l'expression *zfc9* qui est faiblement exprimée. En effet, malgré certains désavantages lorsque comparé aux sondes colorimétriques, dont au niveau du coût et de la complexité relative de leur utilisation, ces sondes sont rapportées comme ayant une sensibilité permettant la détection de gène exprimés à moins de 50 copies par cellules et ont à leur disposition plusieurs options permettant l'amplification spécifique du signal sans augmenter le bruit de fond [242,243].

3.1.2 Perspectives futures pour la caractérisation du profil d'expression de *zfc9*.

Des changements dans l'expression génique sont attendus au cours du développement pour une multitudes de gènes alors que les besoins, structures et activités de l'organisme changent [244], mais aussi potentiellement avec l'âge même une fois le développement achevé et le stade adulte dépassé [245]. Une caractérisation précise du moment, de la distribution spatiale et de l'amplitude des changements dans l'expression d'un gène comme *zfc9* à travers différents stades du cycle de vie d'un organisme modèle comme le poisson-zèbre, même passé les 48hpf, s'avérait donc un outil fort utile à élaborer à l'avenir pour l'étude de la fonction normale de la protéine associée et de l'interprétation des phénotypes de modèles de KD ou KO du gène en question. Cependant, l'observation du profil d'expression d'un gène pour des embryons complets à l'aide de HIS se complique progressivement passés 48hpf. Il sera alors presque nécessaire pour de futures expériences d'HIS d'effectuer celles-ci sur des coupes histologiques pour les embryons âgés de 5 jours ou plus et, puisque ce problème est aggravé pour les gènes à faible expression comme *zfc9*, il pourrait même s'avérer plus efficace de passer à l'analyse de

coupes histologiques dès les embryons de 72 hpf. La quantification du niveau d'expression de *zfc9* dans différents tissus, notamment au niveau de la moelle épinière et les différentes régions du SNC associées au système moteur, à des stades distincts du cycle de vie du poisson-zèbre avec la technique éprouvée du PCR quantitatif, dont l'exécution est moins laborieuse que l'HIS, s'avèrerait aussi un outil utile à développer pour l'étude du gène *zfc9* [246].

L'étude directe du profil d'expression spatiale et temporelle des différentes isoformes de protéine encodée par *zfc9* demeurerait néanmoins l'approche la plus puissante pour son étude. Notre échec à obtenir des lignées marquées pour le gène *zfc9* et l'absence d'anticorps développé spécifiquement pour la détection de *zfc9* à ce jour rend cette approche inaccessible. Cependant, le besoin de générer des anticorps ou d'autres méthodes permettant de détecter l'expression spécifique des différentes isoformes de *zfc9* demeure, ces profils étant vraisemblablement distincts d'une isoforme à l'autre comme ils le sont pour les isoformes de l'orthologue murin de HC9 [239].

3.2 Échec de la génération de lignées avec un tag moléculaire dans *zfc9*

Malgré nos efforts soutenus et les injections répétées, nous avons été incapables d'obtenir même une seule des trois lignées différentes prévues étant supposées intégrer un tag moléculaire à l'extrémité Nt ou Ct de *zfc9* par système d'édition du génome CRISPR/Cas9. En effet, pas un seul poisson apte à transmettre le tag correspondant pour chaque lignée, soit 12xHis au Nt pour la lignée C9orf72^{NHis}, eGFP au Nt dans le cas C9orf72^{NGFP} et eGFP au Ct pour C9orf72^{CGFP}, n'a pu être obtenu et confirmé. Plusieurs explications peuvent être fournies pour expliquer ceci, dont notamment la faible efficacité d'intégration puis de transmission associée à la technique de KI dirigée avec par système CRISPR/Cas9 et une potentielle mortalité accrue chez les poissons injectés résultant de la perturbation de la fonction normale de *zfc9* exacerbée par le site de coupure sous-optimal de nos ARNg.

3.2.1 Efficacité d'insertion avec le système CRISPR/Cas9 chez le poisson-zèbre

Une explication pour notre échec à identifier des poissons fondateurs pour les lignées C9orf72^{NHis} ou C9orf72^{CGFP}, malgré le criblage de 81 et 75 embryons respectivement, est la faible efficacité pour la technique d'insertion de séquences précises à un site spécifique utilisant le système CRISPR/Cas9. En effet, pour l'utilisation du système CRISPR/Cas9, l'efficacité maximale des RDH guidées avec une matrice d'ADN ODNsb a été démontrée dans une étude récente sur le poisson-zèbre comme oscillant entre 4-8% lorsque optimisée [247]. De plus, si l'on ne considère que les réparations dites parfaites permettant l'insertion au site exact désiré sans l'intégration d'indels indésirables, ce qui est nécessaire pour l'insertion de tags qui doivent être dans le cadre de lecture du gène ciblé afin d'être exprimés et détectés par nos techniques de criblage, alors les taux maximaux d'efficacité d'insertion observés sont réduits à des valeurs encore plus basses de seulement 1-4% [247]. Notons que ces taux sont rapportés pour des KI de séquences ne différant des séquences originales des gènes ciblés que par des substitutions non contiguës de 5 nucléotides, dont l'intégration est vraisemblablement plus efficace que pour nos insertions de 36 pb (12xHis) et 717 pb (eGFP) de taille considérablement plus élevée, dont la séquence hors des bras d'homologie de leur matrice ne présentent pas d'homologie avec la séquence cible [233, 234]. Qui plus est, une étude chez le poisson-zèbre pour laquelle a été effectué l'insertion d'une séquence résultant en la substitution de 5 nucléotides non contiguës, à l'aide du même système CRISPR/Cas9 que nous avons employé et de matrices d'ADN ODNsb, a rapporté que le taux d'efficacité le plus élevé obtenu d'intégration dans la lignée germinale, permettant la transmission du KI à une génération subséquente, a été de 4% [248,249]. Enfin, même lorsqu'il y a intégration dans la lignée germinale, le taux de transmission des intégrations sans erreur peut s'avérer aussi relativement bas avec des valeurs rapportées récemment allant de moins de 0.1 % à 23%. [247]. Ainsi, bien qu'un nombre de poissons F0 injectés prévu d'être statistiquement adéquat pour avoir une probabilité raisonnable d'obtenir au moins un spécimen en mesure de transmettre l'intégration du tag désiré sans erreur a été criblés pour nos lignées C9orf72^{NHis} et C9orf72^{CGFP}, notre échec à identifier un poisson positif pour les deux lignés n'est pas un résultat improbable en raison de la faible efficacité

inhérente de la technique d'induction de KI de séquences spécifiques avec le système CRISPR/Cas9.

3.2.2 Nombre limité de candidats F0 criblés

Au-delà de l'efficacité de la technique elle-même, un nombre plus important de poissons F0 criblés puis élevés aurait indéniablement accru nos probabilités d'obtenir des poissons F0 en mesure de transmettre l'intégration sans erreur des tags pour les lignées C9orf72^{NHis}, C9orf72^{CGFP} et particulièrement C9orf72^{NGFP} pour laquelle un seul spécimen F0 sélectionné a pu être criblé, particulièrement si l'on considère que la faisabilité des insertions a été confirmée pour chacune des lignées que nous avons tenté de générer. Cependant, deux limitations importantes nous ont empêché d'effectuer un criblage avec une nombre plus élevé de spécimens F0 dans le temps alloué à ce projet.

La première concerne directement notre aptitude à obtenir un nombre important de spécimens adultes des embryons sélectionnés suite aux injections pour chacune des lignées. En effet, malgré les injections et la sélection subséquente d'un nombre considérable d'embryons prometteurs à plusieurs reprises pour chacune des lignées devant insérer un tag dans zfC9, une forte mortalité est survenue de manière consistante pour chacune des cohortes de nos 3 lignées et ce particulièrement pour la lignée C9orf72^{NGFP}. Les cohortes sélectionnées pour chacune de nos trois lignées présentaient typiquement une faible mortalité jusqu'à environ 11 jours après les injections, puis la majorité des larves de chacune de nos cohortes décédaient alors au cours d'une période de 12 à 20 jours après les injections. Plusieurs approches ont été tentées pour réduire la mortalité observée pour les différentes cohortes. L'assistance de collègues avec une longue expérience pour les micro-injections dans les embryons de poisson-zèbre a été sollicitée pour certaines cohortes, le nombre de poissons élevés par aquarium a été réduit et les proportions de nourriture ont été augmenté ou diminué pour différentes cohortes. De ces tentatives, seule la réduction du nombre de spécimens élevés dans un même aquarium et l'augmentation de la quantité de nourriture distribuée chaque jour ont semblé réduire de manière modeste la mortalité observée.

La seconde limitation expliquant pourquoi il n'a pas été criblé plus de spécimens F0 pour chacune des trois lignées est liée au temps considérable et l'espace non-négligeable nécessaire pour l'élevage puis le criblage des spécimens de chaque lignée. En effet, bien que les tentatives de croisement puissent débuter dès l'atteinte de la maturité sexuelle des poissons zébrés, qui survient généralement après 10 à 12 semaines, les croisements de spécimens aussi jeunes sont souvent laborieux, nécessitant typiquement plusieurs essais avant de produire des embryons fertilisés qui sont généralement peu abondants et dont la viabilité est souvent moindre initialement [232]. La production vantée de plus d'une centaine d'œufs par couvée par femelle, pour chaque croisement du poisson-zèbre, ne survient généralement qu'à un stade plus avancé s'approchant de la période d'âge entre 7 à 18 mois, qui est typiquement optimale pour la production d'œufs fertilisés [232]. De plus, les croisements pour un même spécimen ne peuvent survenir que tous les 2-3 jours pour assurer une production adéquate d'œufs fertilisés pour chaque croisement et éviter de surmener les spécimens, 1-2 journées de repos étant aussi nécessaires avant de tenter un autre croisement même lorsqu'aucun œuf n'a été produit [232]. Par conséquent, bien que le temps d'une génération pour le poisson-zèbre puisse être aussi court que 3 à 4 mois, la nécessité d'un nombre important d'embryons pour effectuer des criblages efficaces d'insertion peut certainement nécessiter une période de temps plus étendue. Additionnellement, bien que les poissons zébrés puissent être élevés en nombre important dans des aquariums communs jusqu'à environ la fin de leur state juvénile (10 à 12 semaine), il est nécessaire de les transférer dans des aquariums individuels pour permettre le criblage de spécimens spécifiques et de leur assurer une croissance optimale permettant l'atteinte rapide du stade adulte durant lequel il est possible de les croiser et donc de les cribler [232]. Les poissons zébrés demeurent dans ces aquariums individuels jusqu'à la fin de leur criblage, ce qui nécessite un espace qui peut s'avérer conséquent. Il en résulte donc que pour un espace donné seulement une quantité limitée de poissons peuvent être croisés pour être criblés en même temps. De ces considérations ainsi que du fait que le présent rédacteur a été le seul responsable des injections additionnelles et du criblage de tous les lignés associées à ce projet à partir de la fin de la première année de cette de maîtrise, il aurait été ardu de procéder au crible d'un nombre considérablement plus important de poissons F0 que ce qui a été fait.

3.2.3 Mortalité importante des larves F0 sélectionnées

Comme mentionné précédemment, une forte mortalité survenant de manière consistante entre 12 et 20 jours suite aux injections, pour chacune des cohortes de nos 3 lignées $C9orf72^{NHIS}$, $C9orf72^{NGFP}$ et $C9orf72^{CGFP}$ a été observée. Nous proposons que cette mortalité puisse avoir contribué à notre échec dans l'identification des spécimens F0 en mesure de transmettre le tag inséré dans *zfC9* pour chacune des trois lignées. Cependant nous avançons l'hypothèse que cette mortalité n'a pas que simplement limité le nombre de poissons adultes F0 qui ont pu être criblés, mais qu'en plus les poissons ayant intégré efficacement les tags ou des indels dans la séquence de *zfC9* ont été affectés de manière disproportionnée.

Les constatations qui nous ont menées initialement à cette hypothèse sont les suivantes. Premièrement, une certaine mortalité accrue de base chez les embryons injectés avec les composantes du système CRISPR/Cas9 comparativement aux WT non injectés est attendue. Cela est attribuable aux stress infligés aux embryons lors de l'injection, notamment par leur perforation par l'aiguille, la toxicité non nulle des différentes composantes de la solution injectée et le risque de coupure hors cibles délétères même avec des ARNg optimisés [250]. Cependant, la mortalité attribuable directement à la toxicité des injections est généralement rapportée de se manifester principalement au cours des premières 72hpf chez le poisson-zèbre, résultant habituellement en la mort d'embryons affectés avant leur éclosion de leur chorion et l'apparition de difformités évidentes, entraînant le décès rarement plus tard que 96hpf [250]. Ainsi, puisque seulement une faible mortalité était constatée pour les cohortes d'embryons sélectionnés pour chacune de nos trois lignées jusqu'à atteindre environ 12 jours après les injections et qu'ils ne développaient pas de défauts morphologiques apparents, nous en sommes venus à considérer qu'il est peu probable que la mortalité observée ait résulté seulement de traumatismes et de la toxicité liés aux injections. Nous menant alors à postuler que l'insertion d'indels et possiblement même l'insertion sans erreur des tags dans la séquence *zfC9* puisse résulter en la perte de fonction de la protéine potentiellement mono et bi allélique dans plusieurs tissus des spécimens affectés. Une perte de fonction qui serait délétère et rendrait ces spécimens affectés plus susceptibles de décéder durant la période critique de 12 à 20 jours après les injections.

Un premier point appuyant cette hypothèse est la possibilité de détecter des manifestations partielles à complètes de phénotypes, associés à la perte de fonction de gène ciblé dès la F0 injectée chez le poisson-zèbre, en raison de la haute efficacité de la mutagenèse dans les cellules somatiques par insertion d'indels aléatoires à l'aide du système CRISPR/Cas9 [251,252]. Ensuite, le degré de similarité et de conservation de la séquence élevée sur toute la longueur entre des espèces évolutivement distante suggère une importante contrainte fonctionnelle pour la protéine C9, qui laisse supposer avec un degré de confiance élevé que des perturbations même mineures de son domaine fonctionnel DENN pourraient être délétères et entraîner une perte de fonction de la protéine [79]. Ainsi, de leur insertion dans le domaine DENN de *zfC9*, même s'ils sont intégrés parfaitement dans le cadre de lecture de *zfC9*, les tags utilisés auraient pu entraîner par leur insertion une perte de fonction de la protéine. Ceci est appuyé par la récente étude chez le poisson-zèbre pour laquelle il a été démontré que l'entièreté du domaine DENN est indispensable à la fonction de la protéine, l'injection de construction exprimant des versions partielles du domaine DENN ayant été suffisante pour récapituler par un effet dominant négatif les phénotypes délétères observés lors du KD de *zfC9* [213]. Le fait que la mortalité la plus importante a été observée pour les cohortes F0 de la lignée C9orf72^{NGFP}, pour laquelle l'insertion du tag eGFP dans la section Nt du domaine DENN de *zfC9* serait le plus susceptible d'induire la perte de fonction de *zfC9* par son insertion volumineuse dans la première section du domaine DENN, est aussi un indicateur en faveur d'une telle hypothèse. En ce qui concerne la possibilité qu'une perte de fonction de *zfC9* entraîne une augmentation de la mortalité chez les poissons injectés, une augmentation de la mortalité considérable chez les vers *C. elegans* KO pour leur orthologue de *C9* qui est exacerbée par certains facteurs de stress environnementaux [103], ainsi que la mortalité accrue de manière significative des spécimens murins HT et HM KO pour leur orthologue de *C9* [253], tend à appuyer une telle hypothèse. Enfin, nos résultats préliminaires sur les proportions des génotypes obtenus pour la progéniture d'un croisement HT/HT pour le KO de *zfC9* indiquent qu'il est possible que moins de spécimens HT et HM pour le KO de *zfC9* survivent pour atteindre l'âge adulte comparativement aux WT. Notons cependant que le KO de *zfC9* n'est pas systématiquement léthal chez le poisson-zèbre comme en atteste notre obtention de spécimens adultes HM pour son KO.

Notre hypothèse pour expliquer le phénomène de mortalité et donc en partie nos difficultés à identifier des spécimens F0 en mesure de transmettre un tag intégré dans la séquence de *zfC9* à

leur progéniture est donc la suivante. Les embryons injectés puis sélectionnés pour les F0 de nos trois différentes lignées seraient partiellement KO pour *zfC9* et les poissons les plus susceptibles d'avoir intégré les tags seraient les plus à risques d'avoir un KO partiel ou complet de *zfC9* répartis à travers leurs tissus. Si les phénotypes délétères rapportés du KO/KD des différents orthologues de C9 sont récapitulés chez le poisson-zèbre, dont ceux au niveau du système immunitaire (souris) [253] et particulièrement le système moteur (poisson-zèbre) [105], ces spécimens seraient alors plus vulnérables au cours d'une période critique de leur croissance et ce particulièrement suite au stress infligé par les injections. Cette période de croissance serait particulièrement critique puisque vers le 5^{ème} jour le vitellus est complètement absorbé et cesse de subvenir aux besoins nutritionnels des larves, elles doivent alors se nourrir d'elles-mêmes et dépendent entièrement de la nourriture ajoutée à leur aquarium [232]. Une des causes principales de leur mort pourrait alors être leur incapacité à se nourrir adéquatement en raison de possibles déficiences motrices associées au KD/KO de *zfC9* et ce particulièrement en présence de poissons moins affectés avec lesquels ils doivent compétitionner. De ces carences, les poissons les plus affectés seraient laissés amoindris, entraînant une situation selon laquelle moins ils parviennent à se nourrir, plus ils sont faibles et plus ils sont faibles moins ils sont aptes à se nourrir jusqu'à ce qu'ils en décèdent.

3.2.4 Perspectives pour le marquage de *zfC9*

Il demeure plusieurs justifications valables et des utilités considérables associées à la génération de tel lignées. En effet, l'insertion d'un tag fluorescent comme eGFP continue de fournir l'avantage notable et unique de permettre l'observation de l'expression de la protéine *in vivo* dans un état normal. De plus, l'insertion d'un tag non-fluorescent comme 12xHis avec ses méthodes de détection éprouvées de grande sensibilité et avec une excellente spécificité, ainsi que largement disponibles, s'avérait moins dispendieux d'usage qu'un anticorps ciblé pour une protéine en particulier. De plus, étant originellement élaboré pour détecter une séquence peptidique dans le centre de la protéine HC9, il est possible, comme il a été rapporté pour plusieurs anticorps commerciaux détectant HC9 [252], que la sensibilité et la spécificité de l'anticorps détectant *zfC9* ne soient pas adéquates pour la détection précise de *zfC9* par immunohistochimie au niveau cellulaire.

Plusieurs considérations devraient néanmoins être tenues en compte pour procéder à la génération efficace de tel lignées après notre propre échec à le faire. Tout d'abord, la sensibilité apparente et prédite de la protéine *zfc9* aux modifications de sa séquence devra se refléter dans la stratégie d'approche. Contrairement à ce que nous avons tenté avec nos ARNg, les insertions de tags moléculaires devraient s'effectuer au moins hors des sections de séquences appartenant au domaine fonctionnel DENN de *zfc9* et optimalement directement suivant l'ATG ou le codon d'arrêt, puisque même les séquences hors du domaine fonctionnel de *zfc9* sont conservées de manière importante entre des espèces évolutivement distantes, leur suggérant une possible importance fonctionnelle [79]. Pour y arriver, plus d'un ARNg avec un site de coupure à quelques pb de l'ATG de *zfc9*, avec un score élevé et prédit de ne pas avoir d'effet hors cible, sont maintenant proposés par plusieurs algorithmes depuis le passage à l'assemblage le plus récent de la séquence du génome du poisson-zèbre (GRCz11). Malheureusement aucun ARNg, de score élevé et sans effet hors cible, apte à couper à proximité du codon d'arrêt de *zfc9*, ne semble être présentement proposé par les algorithmes consultés et ce même avec le nouvel assemblage. Cependant, les algorithmes n'étant pas infailibles, nous recommanderions néanmoins de tester certains des nouveaux ARNg prédits permettant de couper à proximité du codon d'arrêt de *zfc9*, l'un d'entre eux pourrait s'avérer adéquat pour l'insertion d'un tag moléculaire malgré les prévisions. Enfin, plusieurs raffinements de la technique d'insertion de CRISPR/Cas9, rapportée d'accroître l'efficacité d'intégration, pourrait aussi être appliqués pour l'obtention future de telle lignée. Notamment, il a été rapporté que l'utilisation d'une nucléase CRISPR/Cas Cpf1 alternative dérivée de *Lachnospiraceae* bacterium ND2006 (LbCpf1) permet d'augmenter de plus de 4 fois l'efficacité d'intégration de séquences simple brin d'ADN comparativement à Cas9 une fois optimisée [254]. L'usage de composés agissant comme des antagonistes au processus de réparations par JENH (*scr7*) ou d'agonistes pour les RDH (RS-1) a aussi été proposé d'accroître l'efficacité d'intégration bien que leur effet soit rapporté d'être modeste [249].

3.3 Génération d'une lignée knockout pour le gène *zfc9*

Nous sommes parvenus à partir des poissons injectés pour la conception de la lignée *C9orf72^{NHis}*, d'identifier un progéniteur capable de transmettre à sa descendance des indels dans la région Nt de *zfc9*, dont la nature a été déterminée comme étant deux délétions distinctes de 4

et 7 pbs, prédites de provoquer un KO fonctionnel des trois isoformes de la protéine du gène *zfC9*, en entraînant un changement de cadre de lecture résultant en leur troncation. De ce progéniteur, il a été obtenu jusqu'à la F2 d'une lignée poissons-zèbres KO pour *zfC9*, de laquelle des spécimens adultes HM mâles et femelles disposés à engendrer prochainement une F3 composée entièrement de spécimens HM pour le KO de *zfC9* ont été identifiés.

3.3.1 Efficacité d'intégration des indels dans les cellules germinales

L'efficacité de mutagenèse par insertion d'indels aléatoires dans une séquence spécifique, résultant des réparations JENH des CDB résultant du système CRISPR/Cas9, est considérablement plus élevée que l'efficacité d'insertion par RDH. Effectivement, selon l'efficacité de coupure de l'ARNg et la qualité des injections, jusqu'à ~99% des embryons injectés peuvent intégrer au moins de manière mosaïque des indels dans la séquence d'un gène ciblé [255]. En raison de l'efficacité élevée d'insertion, le taux de transmission d'indels est aussi considérablement plus important que pour les KI dirigés, les taux de transmission étant rapportés, par une étude de mutagenèse par CRISPR/Cs9 à haut débit sur plusieurs centaines de gènes, d'être en moyenne de 28% pour les poissons ayant intégré des indels [255]. De l'efficacité d'insertion d'indels pour 87,5% des embryons lors des tests d'efficacité de l'ARNg C9orf72Nt utilisé pour la génération de cette lignée, notre identification de seulement 2 spécimens sur 34 (6%) en mesure de transmettre des indels à leur progéniture, est significativement inférieure à ce qui était attendu. Un tel résultat pourrait s'expliquer par et appuyer notre hypothèse, décrite précédemment à la **section 3.2.3**, selon laquelle les poissons ayant intégré avec une plus grande efficacité des indels dans *zfC9*, susceptibles d'entraîner sa perte de fonction, ont un taux de survie moins élevé et donc atteignent l'âge adulte dans des proportions moindres.

3.3.2 Nature des indels de la lignée *zfC9*^{KO}

Les spécimens de la lignée *zfC9*^{KO} intègrent deux délétions de plusieurs paires de bases à deux différents emplacements de la région Nt du gène *zfC9*. La première est une délétion de 4 pb située 164 pb en aval de l'ATG des trois transcrits de *zfC9*, elle résulte en un changement du cadre de lecture du gène provoquant l'insertion d'un codon-stop prématuré. Il est prédit d'en

résulter la production d'une protéine zfc9 tronquée de 66 aa dont seulement 55 seraient identiques à la séquence d'origine de 462 aa pour l'isoforme zfc9La. La seconde est une délétion de 7 pb à l'emplacement de coupure prévu situé 233 pb en aval de l'ATG commun des trois transcrits de zfc9. Sans la présence de la première délétion, elle induirait par elle-même un changement de cadre de lecture provoquant aussi l'insertion d'un codon de terminaison prématuré dans le cadre de lecture de *zfc9*, dont il est prédit qu'il résulterait une protéine zfc9 tronquée de 95 acides aminés, dont que 77 seraient identiques à la séquence WT de l'isoforme zfc9La. La première délétion est cependant suffisante pour induire la perte de plus de 85% de la séquence d'aa des deux isoformes longues de zfc9 ainsi que près de 80% de la séquence de l'isoforme courte. De plus, même si ces délétions ne provoquaient pas l'insertion de codons d'arrêt prématurés, la présence de la seconde délétion ne restaurerait pas le cadre de lecture pour la séquence en aval de cette dernière, résultant en une séquence d'aa sans homologie significative avec la séquence de zfc9. Ainsi, nous sommes confiants d'affirmer que l'intégration de ces délétions dans la séquence du gène zfc9 des spécimens de notre lignée zfc9^{KO} résulte en une perte de fonctions des différentes isoformes de protéines associées. La présence de deux sites de délétion, bien que fortuite pour le KO de *zfc9*, est cependant un résultat inattendu. En effet, la seconde délétion est exactement au site de coupeur prévu pour l'ARNg (C9orf72Nt) utilisée et donc attendue, mais la première délétion survient inexplicablement 73 pb en amont de ce site. À partir des prévisions des algorithmes employés pour sélectionner initialement nos ARNg, aucun effet hors cible n'était prédit pour l'ARNg C9orf72Nt et, bien que la première délétion survienne en amont d'une séquence PAM identique à celle retrouvée pour C9orf72Nt, la séquence qui la précède n'a aucune similarité notable avec notre ARNg lorsque comparé avec des logiciels d'alignement de séquences. De plus, aucun autre indel n'a été détecté hormis ces deux délétions dans la séquence de zfc9 avoisinant le site de coupure prédit de C9orf72Nt et l'intégration des indels au site de coupure aurait perturbé la séquence de reconnaissance de l'ARNg, empêchant des coupures additionnelles à cet emplacement. Au-delà d'une erreur inhabituelle des systèmes de réparation des CDB par JENH ou d'un effet hors cible de notre ARNg, nous n'avons pas d'explication pour ce phénomène. Cependant, un effet hors cible imprévu de notre ARNg utilisé pour la génération des lignées étiquetées en Nt pour *zfc9*, bien qu'improbable en raison des vérifications préalables effectuées, pourrait être une autre raison de notre échec à établir ces lignées.

Il est à noter que ce KO de *zfc9* pourrait potentiellement mener à l'expression de formes tronquées des différentes isoformes de la protéine incorporant une portion de la section uDENN du domaine DENN et que, comme il a été démontré que l'expression de constructions partielles de ce domaine entraîne un effet dominant négatif, nos spécimens HT et HM pour le KO de *zfc9* pourraient présenter aussi cet effet [213]. S'il s'avérait que ce soit le cas, nous devrions nous attendre à ce que les spécimens HT pour le KO de *zfc9* présentent une diminution de l'activité de la protéine plus prononcée qu'un KO HT du gène entraînerait normalement et donc, si les effets de la perte de fonction de *zfc9* sont dépendants du degré de cette perte, on devra s'attendre à des phénotypes pour les HT de cette lignée qui se rapprocheraient de ceux obtenus pour les HM. De plus, les tentatives de sauvetage des phénotypes observés pour le KO de *zfc9* par l'expression exogène de *zfc9* ou de *HC9* pourraient être compromises s'il s'agit bien d'un effet dominant négatif. Cependant, il est assez probable que l'ARNm codant pour les possibles versions tronquées de *zfc9* ne soient pas traduits et s'ils le sont il est probable qu'ils soient instables, menant à leur dégradation rapide. De plus, même s'ils sont exprimés, ils seraient considérablement plus courts que les constructions de délétion partielle du domaine DENN pour lesquelles l'effet dominant négatif a été rapporté et donc possible qu'ils ne puissent avoir un tel effet [213].

Enfin, le KO de *zfc9* par la présence de ces délétions devra néanmoins être confirmé définitivement avant que la génération d'une lignée KO stable soit certaine et que sa caractérisation puisse véritablement débuter. En effet, un simple test par buvardage de western sur des extractions de protéines provenant de poissons de la lignée *zfc9*^{KO} HM pour l'intégration des délétions en utilisant, une fois disponible, un anticorps détectant *zfc9* nous permettra, avec les contrôles appropriés, de rapidement déterminer si *zfc9* est encore exprimée ou non chez ces spécimens. Même si l'obtention de spécimens F3 venait à tarder en raison des difficultés rencontrées pour le croisement des spécimens F2 HM identifiés, une extraction protéique effectuée à partir d'un fragment de nageoire de ces spécimens devrait permettre d'effectuer le test par buvardage de Western, en supposant que l'expression de *zfc9* dans ces tissus soit suffisante pour être détectée clairement par l'anticorps utilisé, il sera donc particulièrement important d'avoir des contrôles positifs et négatifs adéquats pour l'expression

de *zfc9* lorsque ce test sera effectué. Des approches complémentaires pour vérifier le KO de *zfc9* seraient la vérification par RT-PCR et par HIS de l'absence de l'ARNm de *zfc9* chez les spécimens KO, la présence des délétions pouvant entraîner sa dégradation. Notons cependant que l'absence de détection de l'ARNm *zfc9* confirmerait son KO, mais que sa détection n'invaliderait pas nécessairement notre lignée puisque les délétions dans la séquence ADN de *zfc9* peuvent engendrer le KO fonctionnel de sa protéine sans nécessairement mener à la dégradation de l'ARNm.

3.3.3 Caractérisation préliminaire des spécimens KO pour *zfc9*

Nos tests préliminaires sur la progéniture d'un croisement HT/HT des spécimens F1 de la lignée *zfc9*^{KO} indiquent une mortalité plus élevée et une déficience de l'activité de nage chez les spécimens HT ainsi que HM pour le KO de *zfc9* comparativement aux WT. Ces résultats sont en accord avec ce qui a été prédit par les phénotypes observés suite au KD ou KO des orthologues de *HC9* dans différents modèles animaux, incluant le poisson-zèbre. De plus, ce sont ces mêmes phénotypes qui ont initialement guidés, avec nos observations faites lors de l'élevage des différentes générations de cette lignée, notre choix de la nature de ces tests préliminaires supposant que le KO de *zfc9* pourrait entraîner des phénotypes similaires chez le poisson-zèbre. En effet, les souris HM ainsi que HT KO pour leur orthologue de *HC9* décèdent en plus grand nombre ainsi que de manière plus précoce comparativement aux spécimens WT d'une même portée [253] et ceci est récapitulé pour les vers *C. elegans* HM pour le KO de leur orthologue de *HC9* [103]. De plus, le KD transitoire de *zfc9* provoque des déficiences motrices chez des larves de 48-96hpf, se manifestant par un niveau d'activité amoindri pour la nage spontanée et celle provoquée comparativement aux spécimens WT [105, 213]. Cependant, nous devons mettre l'accent sur la nature préliminaire et imparfaite des paramètres des tests utilisés pour obtenir nos résultats pour les possibles phénotypes liés au KO de *zfc9*, principalement en raison de l'accès au cours de ce projet à seulement la progéniture limitée de croisement HT/HT des spécimens F1 de la lignée *zfc9*^{KO}, pour obtenir des spécimens des trois génotypes possibles pour le KO de *zfc9* afin d'effectuer nos tests préliminaires.

3.3.3.1 Mortalité accrue des spécimens HT et HM KO pour *zfc9*

Nous avons démontré à l'aide d'un test statistique du χ^2 que les proportions de spécimens de génotypes WT, HET et HM obtenus du croisement de spécimens HT/HT pour le KO de *zfc9* diffèrent de manière statistiquement significative (valeur- $p \leq 0.05$) des proportions prédites si chacun des génotypes est aussi susceptible d'atteindre l'âge adulte que les autres. Les spécimens HT (42% vs 50%) et HM (17.5% vs 25%) étant sous-représentés et les spécimens WT (40% vs 25%) sur-représentés indique que le KO de *zfc9* même partiel diminuerait le taux de survie jusqu'à l'âge adulte des spécimens affectés. Ces écarts modestes, bien que significatifs avec les proportions attendues pour chacun des génotypes, ne sont pas aussi élevés que ce que la mortalité observée suite aux injections nous aurait laissé supposer selon notre hypothèse, mais ceci pourrait s'expliquer par l'absence de stress dû aux injections pour la progéniture de la F1 et de leur élevage en cohortes moins nombreuses due à la faible productivité de leur progéniture, résultant en une mortalité moins élevée. Cependant, bien que ce résultat en soit un qui était attendu, certaines limitations importantes de la méthode et des paramètres du test doivent être mentionnées. Tout d'abord, notre conclusion préliminaire d'une mortalité plus abondante des poissons HT ou HM pour le KO de *zfc9* repose sur la supposition que certains des génotypes n'étaient pas initialement sur-représentés ou sous-représentés de manière importante dans la cohorte d'embryons élevés des croisements qui ont engendré les 57 spécimens génotypés. Malheureusement, les proportions obtenues initialement d'un croisement donné peuvent différer de manière non négligeable de celles prévues (de 25% WT, 25% HM et 50% HT) par les lois de Mendel [238] sur la transmission des allèles pour un croisement HT/HT et ce particulièrement lorsqu'un nombre restreint de spécimens est considéré, comme c'est le cas pour nos 57 individus survivants génotypés. Une autre faiblesse des tests statistiques à partir d'un nombre limité de spécimens est que toute variation, même que d'un nombre modeste de spécimens pour chacune des catégories, a un effet disproportionné sur les résultats [256]. Ce qui est évidemment problématique pour la robustesse de notre conclusion et ce particulièrement puisque les différences entre les quantités de spécimens obtenus de chaque génotype et celles attendues, si la mortalité ne diffère pas entre eux, est modeste pour chaque génotypes (WT : +8.75, HT : -4.5, HM ; -4.25). Une différence de quelques spécimens de chaque génotype suffirait donc à ce que les différences des proportions avec celles attendues cessent d'être statistiquement

significatives. Le nombre de spécimens génotypés (57) est donc trop bas pour en tirer avec une confiance élevée des conclusions définitives concernant les taux de mortalité relatif aux différents génotypes pour le KO de *zfc9*. Optimalement, pour vérifier rigoureusement notre hypothèse plusieurs tests devraient être effectués à plusieurs reprises à partir des proportions des génotypes obtenues de cohortes différentes composées d'un nombre plus élevés de spécimens provenant de croisement HT/HT d'une même paire de progéniteurs [256]. Il sera donc nécessaire de confirmer cette tendance de mortalité accrue des spécimens KO HT et HM pour *zfc9* à l'aide de tests additionnels. L'accès à des spécimens HM adultes en mesure de produire des cohortes abondantes de spécimens strictement du génotype HT ou HM pour le KO de *zfc9*, dont nous disposerons prochainement, sera certainement un atout pour une telle caractérisation.

3.3.3.2 Déficiences dans l'activité de nage des HT et HM KO pour *zfc9*

Les déficiences dans l'activité de nage observées pour les spécimens HT ou HM, pour le KO de *zfc9* comparativement au WT, n'ont été statistiquement significatives que durant une période de temps limitée suite à l'application d'un stimulus (lumière) continu. Ceci est intrigant puisque qu'il a été rapporté que le KD de *zfc9* entraîne des déficiences motrices significatives chez les larves testées autant pour la nage spontanée qu'induite [105]. Plusieurs phénomènes pourraient expliquer en partie cette disparité entre les résultats pour le KO stable de *zfc9* et son KD transitoire. Premièrement, il est possible que, comme rapporté à plusieurs reprises pour des KD transitoire en employant des OMA, un effet compensatoire a lieu chez les spécimens avec un KO stable de *zfc9*, menant à un phénotype moindre que pour les spécimens injectés avec des OMA pour lesquelles ce phénomène de compensation n'a pas lieu [215,217]. Ensuite, il est possible que les déficiences motrices observées suite au KD de *zfc9* à 48-96hpf, probablement liées aux perturbations du développement des motoneurones aussi rapportés [105], soient compensées quelque peu au cours du développement, menant à une manifestation d'un phénotype moteur moins prononcé à des stades plus avancés. Le fait que les quelques spécimens HM adultes pour le KO de *zfc9* que nous avons réussis à obtenir ne semblent pas montrer de défauts flagrants au niveau de leur aptitude à nager, ou du moins pas à un degré comparable à ce qui été rapporté pour les larves de 48-96hpf tend à appuyer une telle possibilité [105]. Le fait que l'activité de nage des spécimens HT et HM tend à être inférieure, bien que pas de manière

significative, au WT même dans les périodes d'activité réduite, pourrait pour sa part indiquer que les déficiences motrices sont plus apparentes dans une période d'activité élevée suite à un stimulus, mais qu'elles pourraient être présentes même durant les phases d'activités faibles. L'activité de nage basse durant la période d'obscurité provenant de l'entrée dans une phase de repos des poissons zébrés et la diminution progressive de l'activité après une période étendue d'exposition au stimulus, s'expliquant par une acclimatation des larves à ce dernier, qui mène à leur désintérêt/ennui qui se traduit par une activité réduite [257]. De plus, notre incapacité à identifier des écarts de valeur d'activités de nage entre les spécimens des différents génotypes dans ces périodes de faible activité pourrait être attribuable à l'une des limitations des paramètres de ce tests de nage. En effet, peu de spécimens pour chaque génotype ont pu être testés en même temps (18 HM, 22 HT ,17 WT) et donc les intervalles d'erreur, pour les valeurs moyennes de chaque groupe, en sont plus importantes, ce qui entraîne la nécessité d'écarts d'activité entre les groupes particulièrement importants, comme suite à l'application d'un stimulus, pour qu'elles soient significatives [256]. Ceci pourrait aussi expliquer pourquoi aucune différence significative n'a été observée pour l'activité de nage des spécimens HM et HT pour le KO de *zfc9*, bien qu'il soit aussi plausible qu'un KO total de *zfc9* n'entraîne pas de phénotypes d'une sévérité significativement plus sévère qu'un KO partiel, malgré les observations chez les modèles murins qui indiquent que ce serait le contraire [253]. Ainsi, pour confirmer que les spécimens KO pour *zfc9* ont des déficiences dans leur activité de nage et afin d'en caractériser la sévérité, des tests additionnels avec un nombre de spécimens plus élevé et consistant pour chaque génotype du KO de *zfc9* seront donc nécessaires. Enfin, des tests sur des spécimens à différents stades du cycle de vie du poisson-zèbre pourront aussi déterminer si la sévérité du phénotype du Ko de *zfc9* varie avec l'âge des spécimens.

3.3.4 Perspectives futures pour une lignée *zfc9*^{KO}

La génération d'une lignée stable pour le KO de *zfc9*, une fois confirmée par buvardage de western, permettra par sa caractérisation à la fois de nous informer sur la fonction de *zfc9* et de modéliser dans le poisson-zèbre le mécanisme de perte de fonction de C9 associé à la SLA. Cette caractérisation débutera bien entendu avec la vérification de si les spécimens de cette lignée récapitulent les phénotypes observés pour les KD de *zfc9*, soit les défauts de motilité et la perturbation du neurodéveloppement, notamment au niveau des motoneurones [105, 213].

Comme mentionné plus haut, des tests des paramètres de nage spontanée et celle provoquée de spécimens *zfc9*^{KO} à divers stades du cycle de leur vie permettra d'identifier si le KO de *zfc9* entraîne des perturbations du système moteur lorsque comparé aux spécimens WT. L'effet du KO de *zfc9* sur les motoneurons pour sa part pourra être observé à l'aide de diverses approches. Notamment, elle pourra être faite *in vivo* durant les premiers jours du développement à l'aide de croisement avec des lignées exprimant un marqueur fluorescent seulement dans les motoneurons, permettant l'observation aisée sous microscope pour vérifier si leur développement, morphologie et organisation axonale sont perturbés [258]. Il sera aussi possible d'effectuer des études d'électrophysiologie sur différentes populations de motoneurons de poissons KO pour *zfc9* à différents stades et âges afin de vérifier si leur activité est altérée [259]. Des extractions de motoneurons de poissons KO pour *zfc9* pourraient aussi être faites à partir d'embryons pour en faire l'étude en culture [260] et l'étude de l'état des motoneurons adultes pourra se faire à l'aide d'expériences d'immunohistochimie sur des coupes histologiques des tissus nerveux du système moteur. Tout phénotype délétère significatif des motoneurons ainsi que de déficiences motrices, particulièrement s'ils s'aggravent avec l'âge, augurerait bien pour le potentiel du poisson-zèbre comme modèle génétique du mécanisme de perte de fonction de *HC9* dans la SLA, permettant d'identifier par criblage à haut débit des molécules thérapeutiques potentielles pour le traitement de la SLA, dont particulièrement pour les cas associés aux ERH de *HC9*. Il sera aussi vérifié si les poissons KO pour *zfc9* récapitulent les phénotypes observés chez les autres modèles de KO de *C9*, dont des signes de maladie auto-immune, un taux de mortalité accrue et une neurodégénérescence, particulièrement au niveau des motoneurons, exacerbée par des stress environnementaux [103,253]. De plus, il sera intéressant de vérifier si le KO de *zfc9* a des effets au niveau de la taille ainsi que la morphologie des spécimens, en plus de caractériser ses effets sur le transcriptome du poisson-zèbre à l'aide de séquençage aléatoire du transcriptome entier (RNA-Seq), susceptible de fournir des indices sur des voies additionnelles impliquant *zfc9* et pouvant aussi servir de piste pour d'éventuelles cibles thérapeutiques pour le traitement de la SLA.

La caractérisation des spécimens HM pour le KO de *zfc9* sera vraisemblablement priorisée pour la caractérisation du KO, en raison des phénotypes plus sévères rapportés dans le modèle murin pour les spécimens HM [253], mais les spécimens HT pourraient potentiellement mieux

récapituler l'effet des ERH dans *HC9* chez les patients atteints de SLA qui présentent une haploinsuffisance du gène plutôt qu'une perte de fonction totale [126]. Enfin, il serait aussi concevable, à partir de notre lignée *zfC9^{KO}*, de récapituler dans un même modèle de poisson-zèbre jusqu'à l'ensemble des trois mécanismes proposés par lesquels les ERH de *HC9* entraîneraient la manifestation de la SLA et ce, sans recourir à l'insertion directe d'ERH dans le premier intron de *zfC9* par édition du génome. Effectivement, par croisement de spécimens de notre lignée *zfC9^{KO}* avec ceux d'une lignée transgénique exprimant une ou plusieurs des DPR produites de la RANT des transcrits ARN de *C9* avec des ERH, il serait possible de récapituler à la fois la perte de fonction de *C9* et l'expression de DPR toxique. De telles lignées transgéniques exprimant 100 répétitions des DPR toxiques GR de manière ubiquitaire, seulement dans les neurones ou spécifiquement dans les motoneurones, ont été obtenues et caractérisées précédemment dans notre laboratoire [161]. De plus, en croisant nos spécimens KO pour *zfC9* avec une lignée transgénique exprimant un certain nombre de copie d'ERH, il serait possible de récapituler dans un même modèle les 3 principaux mécanismes de toxicité par lesquelles les ERH de *C9* provoqueraient la SLA, soit l'expression d'ARNm mutant toxique, l'expression de DPR toxique et la perte de fonction de *C9*. Une telle lignée pourrait être générée par nous même à l'aide du système Tol2 de manière similaire aux lignées transgéniques obtenues pour ce projet ou préférablement des spécimens pourraient être obtenus d'un laboratoire qui a déjà publié leurs résultats pour la génération et la caractérisation d'une telle lignée ayant démontré sa capacité à modéliser certains aspect clés associés à la *C9(SLA/DFT)* [161].

3.4 Génération de lignées transgéniques expriment le gène *HC9* avec ou sans eGFP.

Nous avons été en mesure d'obtenir des spécimens F1 adultes HT exprimant un gène rapporteur GFP, indiquant l'intégration par système Tol2 des constructions transgéniques pour nos lignées *HC9tg* et *HC9-eGFPTg*, qui devraient permettre une expression sous le contrôle du système Gal4-UAS du gène *HC9* WT et *HC9* marqué avec eGFP. Enfin, nous avons confirmé par injections d'ARNm de Gal4 dans des embryons *HC9-eGFPTg* la capacité d'expression de de *HC9-eGFP* bien qu'elle semble dégradée.

3.4.1 Efficacité d'intégration des transgènes dans la lignée germinale

L'efficacité d'intégration d'un transgène dans la lignée germinale à l'aide du système Tol2 est rapportée être très efficace puisqu'elle oscille généralement entre 50 à 70% pour les poissons F0 injectés dont l'intégration a été confirmée de s'être faite au moins de manière mosaïque dans les cellules somatiques. De plus, les progéniteurs F0 transgéniques produisent en général des descendants F1 transgéniques dans des proportions pouvant varier énormément, mais habituellement rapportées d'osciller de 3 à 100% [223,225]. Cependant, l'efficacité d'intégration dans la lignée germinale, obtenue par le criblage des spécimens F0 pour nos deux lignées transgéniques, a été nettement inférieure avec un taux de 4,5% pour la lignée HC9tg et de 5,1% pour la lignée HC9-eGFPTg. Une explication possible pour ce taux d'efficacité bas étant que l'ARNm de transposase utilisé pour les injections était partiellement dégradé, menant à une efficacité d'insertion du transgène dans la lignée germinale considérablement moindre que ce qui est habituellement attendue. Autrement, nous n'avons pas d'explications satisfaisantes pour un taux d'efficacité d'insertion aussi bas de nos transgènes dans la lignée germinale avec la technique Tol2. Effectivement, une mortalité élevée des spécimens injectés ayant intégré de manière efficace le transgène ne peut expliquer ici ce résultat pour plusieurs raisons. Notamment parce que les spécimens des deux lignées sélectionnées, pour l'élevage en fonction de leur expression mosaïque du gène rapporteur GFP sous contrôle du promoteur cardiaque *cmlc2*, ont montré un taux de mortalité relativement modeste avec près de 60% des spécimens sélectionnés pour chaque cohorte survivants jusqu'à l'âge adulte. De plus l'expression de GFP n'est généralement pas toxique à moins d'atteindre des niveaux d'expression improbables [233] et même si l'expression de HC9 ou la surexpression des orthologues de C9 en général s'avérait toxique, en l'absence d'expression du promoteur Gal4 il n'y a aucune chance que HC9 ait été exprimée dans les spécimens FF0 injectés pour nos lignées transgéniques.

3.4.2 Possible instabilité et dégradation de la protéine de fusion HC9-eGFP

Le transgène de la lignée HC9-eGFPTg permettant l'expression de la protéine de fusion HC9-eGFP en présence de Gal4, la vérification de son expression devait pouvoir se faire par le repérage d'un signal de fluorescence chez des embryons résultant d'un croisement de spécimens F1 de cette lignée avec des spécimens d'une lignée exprimant Gal4 de manière ubiquitaire.

Cependant, de la progéniture obtenue de ces croisements, aucune fluorescence décelable sous microscope à épifluorescence n'a été observée à 24, 36 et 72 hpf autre que la fluorescence attendue au niveau du cœur des embryons sélectionnés. L'expression ubiquitaire de Gal4 de la lignée utilisée a été confirmée par croisement avec une autre lignée dont l'expression de GFP était sous le contrôle du UAS, afin d'écartier la possibilité d'une expression absente ou trop faible de Gal4 pour expliquer l'absence de l'expression du transgène. L'intégration et l'expression du gène rapporteur sans l'intégration du reste du transgène est un phénomène particulièrement rare lors de l'insertion d'un transgène à l'aide de Tol2 [223,225]. Ainsi, nous avons tenté d'injecter de l'ARNm de Gal4 dans des embryons à 1-2 cellules provenant du croisement de spécimens F1 de la lignée HC9-eGFP, ceci menant à une surexpression du promoteur Gal4 qui devrait à son tour entraîner une expression élevée de C9H-eGFP, si le transgène a bien été intégré, nous permettant de vérifier si les spécimens F1 étaient véritablement incapables de l'exprimer ou si son abondance était simplement trop faible pour être détectée. Un faible signal diffus de fluorescence dans l'ensemble de l'embryon a été observé à 20hpf chez les embryons injectés, indiquant qu'il y a bien intégration du transgène permettant l'expression de HC9-eGFP chez les descendants de la lignée C9H-eGFPtg lorsqu'il y a présence de Gal4. La faiblesse de l'expression de HC9-eGFP est cependant surprenante et problématique pour l'utilisation future de la lignée. En effet, l'injection de l'ARNm de Gal4 à la concentration utilisée est habituellement plus qu'amplement suffisante pour induire un signal de fluorescence de forte intensité dans les autres lignées de notre laboratoire ayant intégrées un transgène eGFP, dont l'expression est sous le contrôle du promoteur Gal4. De plus, l'intensité du signal de fluorescence du gène rapporteur étant forte au niveau du cœur pour les spécimens de la lignée C9H-eGFPtg, l'intégration du transgène impliquant l'insertion du même nombre de copies des séquences pour les deux gènes et l'injection de d'ARNm de Gal4 entraînant une forte expression de ce promoteur, il était escompté que le signal de fluorescence de l'expression de HC9-eGFP, suite aux injections, s'approche au moins en intensité de celui du gène rapporteur.

Nous avançons trois hypothèses possibles pour expliquer nos observations et celles-ci ne sont pas toutes nécessairement mutuellement exclusives. La première est que la fusion de zfc9 avec eGFP est instable possiblement en raison de forte contrainte fonctionnelle sur la structure et le bon repliement de la protéine avec lesquelles la taille considérable de eGFP interférerait lorsque

fixé au Ct de *zfc9*. Il en résulterait la dégradation rapide de la protéine qui, sauf lorsque largement surexprimée, aurait alors une abondance inadéquate pour produire un signal de fluorescence pouvant être distingué. Cependant, la détection de cellules avec un signal de fluorescence aisément repérable, suite aux injections pour la génération de la lignée *C9orf72^{CFP}* par l'insertion d'un tag eGFP dans la région Ct de *zfc9*, tendrait à indiquer que cette hypothèse est moins probable. Néanmoins, en raison du peu de connaissances à propos du repliement et de la structure tridimensionnelle des orthologues de la protéine HC9, il n'est pas entièrement impossible qu'une insertion de eGFP dans le Ct de *zfc9* plutôt qu'à sa suite produise une protéine de fusion moins instable, permettant son accumulation à un niveau auquel l'observation de son signal de fluorescence est possible.

Nos seconde et troisième hypothèses s'appuient sur la faible expression endogène détectée de HC9 ainsi que de ses orthologues et ce même dans les tissus ainsi que les cellules dans lesquelles ils sont rapportés être comparativement plus hautement exprimés [91,239]. Elles s'appuient aussi sur la possibilité que HC9 puisse agir et être reconnue dans l'ensemble comme équivalente à *zfc9* chez le poisson-zèbre, une supposition qui est appuyée par la capacité de son expression de secourir les phénotypes résultants du KD de ce dernier [105]. Nous proposons donc qu'il est possible, bien que peu probable, que le niveau d'expression de *zfc9* soit strictement régulé et qu'une détection de l'augmentation persistante de son abondance au-delà des niveaux standards provoque une augmentation du rythme de sa dégradation et par ce fait même celle de HC9 par les divers processus cellulaires responsables de la protéostase. Enfin, notre dernière hypothèse similaire à la précédente proposerait plutôt de l'implication de HC9 dans des processus impliquant le trafic endosomal, dont notamment au niveau de l'autophagie et du système lysosomal [93,108], que le cycle de renouvellement/dégradation de la protéine (turnover) soit très actif malgré la faible expression relative rapportée de son ARNm. Pour ces deux hypothèses, le niveau d'expression constant de notre transgène lorsque Gal4 est exprimé ne serait donc pas suffisant pour devancer la vitesse à laquelle il s'est dégradé, empêchant son accumulation à un niveau auquel un signal de fluorescence peut être détecté. Si l'une de nos deux dernières hypothèses était avérées, malgré leur improbabilité, ceci pourrait être problématique aussi pour la lignée transgénique HC9tg puisqu'une abondance de HC9 trop faible pourrait rendre plus difficile ou même empêcher sa détection à l'aide d'un anticorps et compromettre plusieurs des

expériences pour laquelle cette lignée aurait été utile. D'une manière ou d'une autre, le fait que la lignée HC9-eGFP ne nous permet pas d'observer l'expression de HC9-GFP in vivo sans recourir à une surexpression exogène importante de Gal4 signifie que son utilité comme outil de recherche est sévèrement compromise.

3.4.3 Perspectives futures pour les lignées transgéniques HC9tg et HC9-eGFPtg

Si l'expression de HC9 en présence de Gal4 est confirmée par buvardage de Western pour les lignées HC9tg et HC9-eGFPtg, les expériences subséquentes s'effectueront avec la première de celles-ci, puisque HC9-eGFP ne permet pas la détection in vivo de l'expression de sa protéine de fusion transgénique, tel qu'était son utilité première et la possibilité que ce soit dû à une instabilité due à l'intégration d'eGFP qui pourrait aussi affecter la fonction native de HC9, rendant ainsi son usage superflu. La lignée HC9tg pourra alors servir de base pour des croisements avec différentes lignées exprimant Gal4 avec des profils différents ou avec un promoteur permettant de contrôler temporellement et même spatialement l'expression du transgène. Il sera ainsi possible de générer des lignées exprimant HC9 de manière soit ubiquitaire, spécifique à un tissu, ou spécifique à des populations de cellules et ce seulement suite à un stimulus comme un choc thermique si nous le désirons. De telles lignées permettront entre autres de vérifier si la surexpression stable, persistante, ubiquitaire ou ciblée de HC9 peut être délétère à long ou à court terme, cette caractérisation s'effectuant selon un modèle similaire à celui décrit pour la caractérisation de la lignées zfc9^{KO} décrit à la **section 3.3.4**. Elles permettront aussi de vérifier si l'expression de HC9 permet de secourir les phénotypes possibles résultant du KO stable de zfc9, que ce soit avec une expression stable depuis la conception ou seulement plus tardivement, de sorte à récapituler les conditions d'un traitement possible pour la SLA. Ces observations s'avéreront utiles pour l'étude de la fonction de HC9 ainsi que ses orthologues et permettront de baliser certains aspects de la recherche de traitements pour la C9(SLA/DFT) qui envisageraient la modulation de l'expression de HC9 ou des voies cellulaires dans lesquelles elle est impliquée comme avenues potentielles de traitement.

4. Conclusion

Lors de cette maîtrise, nous avons tenté sans succès de générer trois lignées exprimant le gène *zfc9* avec un tag moléculaire de 12 Histidine à son Nt ($C9orf72^{NHIS}$), une étiquette eGFP à son Nt ($C9orf72^{NGFP}$) et un tag à eGFP à son Ct ($C9orf72^{NGFP}$). Cependant, il a été possible de générer jusqu'à la F2 d'une lignée intégrant des délétions au Nt de *zfc9* prédit d'entraîner sa perte de fonction ($zfc9^{KO}$) et de montrer à l'aide de tests préliminaires que le Knockout de ce gène induirait des déficiences au niveau du système moteur ainsi qu'une augmentation de la mortalité des spécimens HM et HT pour la perte de fonction de *zfc9*. De plus, nous avons obtenus des spécimens adultes F1 HT exprimant un gène rapporteur GFP indiquant l'intégration de constructions géniques, permettent la génération de lignée transgénique en mesure d'exprimer sous contrôle du système Gal4-UAS le gène Humain *C9orf72* WT (HC9tg) et marqué avec eGFP (HC9-eGFPtg), bien que la florescence associée à l'expression de ce dernier soit indétectable sans une surexpression exogène de Gal4. Une majorité du travail réalisé impliquant les tentatives d'injections puis de criblage pour identifier des spécimens F0 en mesure de transmettre les traits désirés à leur progéniture, pour chacune des différentes lignées que nous avons cherché à générer, puis l'élevage et le croisement de cette progéniture jusqu'à l'obtention des générations permettant la caractérisation ainsi que l'usage des lignées obtenues. Une part un peu moins significative du temps alloué à cette maîtrise a été allouée à la caractérisation des indels transmis pour la génération de lignée $zfc9^{KO}$, aux tests préliminaires de possibles phénotypes de nage y étant associés, aux tests de l'expression du transgène de la lignée HC9-eGFPtg et pour la caractérisation par hybridation *in situ* de l'expression de l'ARNm de *zfc9* de 24 à 48 hpf chez le poisson-zèbre. Les prochaines étapes immédiates de ce projet étant l'établissement de générations stables ainsi que la confirmation par buvardage de Western de leur trait d'intérêt pour les lignées $zfc9^{KO}$, HC9tg et HC9-eGFPtg, suivi de leur caractérisation.

Cette maîtrise a permis d'établir les bases pour une étude génétique de la sclérose latérale amyotrophique et de la fonction du gène *zfc9* en utilisant le poisson-zèbre comme modèle. Ces lignées, une fois confirmées ainsi que caractérisées, seront des outils qui pourront être employés non seulement pour obtenir une meilleure compréhension des défauts biochimiques de la perte

de fonction ou de la surexpression de zfC9, nous informant sur sa fonction normale et la pathogénèse de la SLA, mais qui permettront aussi l'étude des réseaux génétiques de la SLA et de trouver potentiellement de nouvelles approches thérapeutiques en utilisant le poisson-zèbre comme système de criblage in vivo de petites molécules. Ces lignées ont donc un potentiel notable comme modèles génétiques in vivo de la SLA associée à la mutation du gène *C9orf72*.

5. Matériel et méthode

5.1 Matériel et méthodes commun à plusieurs expériences

5.1.1 Maintien des poissons zébrés

Les poissons zébrés (*Danio rerio*) étaient maintenus à 28°C dans des aquariums d'eau douce filtrée, dans une salle ayant un cycle nuit-jour de 12 heures. La veille du jour où des embryons étaient nécessaires, les poissons adultes des différentes lignées utilisées étaient, selon la nécessité et le croisement désiré, placé dans des aquariums de reproduction avec des poissons de sexe opposé, soit de la même lignée transgénique (incross), soit de type sauvage (outcross), et gardés séparés jusqu'au lendemain matin. À l'allumage des lumières de la salle au matin, les séparateurs étaient enlevés et les poissons étaient laissés en contact jusqu'à la production d'œufs. Les embryons étaient alors récoltés et les poissons adultes séparés et remis dans leurs aquariums respectifs. Les croisements ne débutaient pas avant que les poissons aient atteint l'âge de maturité sexuelle (10-12 semaines [232]) ainsi qu'une taille comparable à leur partenaire de croisement

Les embryons récoltés étaient conservés à 28,5°C dans de l'eau du système contenant 0,02% de bleu de méthylène, jusqu'à 1 semaine d'âge. Ils étaient ensuite gardés dans de l'eau sans bleu de méthylène. Chaque jour, les embryons morts étaient retirés et l'eau changé au matin et en soirée jusqu'à leur transfert. Les embryons gardés pour élevage des nouvelles générations étaient transférés dans des aquariums appropriés vers 8 jours post fertilisation et nourris selon les recommandations de *The Zebrafish Book* (Westerfield, 2007) [232]. Les aquariums étaient inspectés de manière journalière pour en assurer la propreté et le retrait des poissons décédés si nécessaire.

Toutes les manipulations ont été réalisées conformément aux normes du code de déontologie des études sur les animaux (CDEA) de l'Université de Montréal.

5.1.2 Préparation de milieux de culture Luria Bertani (LB) et de plaques LB agar

Le milieu de culture LB et les plaque LB agar utilisées pour l'élevage de colonies de bactéries transformées sur plaque ainsi que leur culture après repiquage a été préparées selon les spécification décrites par Green, Rachel et al [261]. Les plaques LB étaient préparées de sorte à contenir 50 µg/mL d'ampicilline pour la culture sélective de bactéries intégrant nos plasmides utilisés qui conféraient une résistance à l'ampicilline.

Préparation de plaque LB agar et milieu LB : 1. Mélanger dans un contenant de verre allant au four micro-ondes et adapté à l'autoclave : 2.5 g d'extrait de levure, 5 g de typtone, 5 g de NaCl et 7.5 g d'agar-agar dans 5L d'H₂O stérile. 2. Chauffer au micro-ondes jusqu'à ébullition et dissolution complète des composantes en agitant de manière périodique. 3. Traité à l'autoclave à 121 °C (15-30 min). 4. Refroidir la solution à 50 °C et nettoyer la zone de travail avec de l'éthanol 70% pour éviter la contamination. 5. Allumer la flamme d'un bec Bunsen pour protéger la zone de travail des contaminants portés dans l'air. 6. Ouvrir le contenant avec la solution LB-agar près de la flamme et ajouter de l'ampicilline de sorte à avoir une concentration finale de 50 µg/mL. 7. Verser sous la flamme 10-15 ml de la solution bien agitée dans chaque plaque de 100 mm x 15 mm. 8. Refermer les plaques puis laisser solidifier et sécher au cours de la nuit. 9. Placer les plaques dans un sac de plastique et placer à 4 °C jusqu'à utilisation. Pour la préparation du milieu LB suivre les étapes 1 à 4 sans ajouter d'agar-agar, simplement s'assurer de travailler près de la flamme lors de l'ouverture du contenant pour ne pas contaminer le milieu LB.

5.1.3 Gels d'agarose et migration par électrophorèse

L'ensemble des migrations sur gel décrites pour les diverses expériences de ce mémoire ont été effectuées sur gel d'agarose (Wisent Bio) coloré (Fluo-DNA/RNA gel staining solution, Zmtech scitific) pour visualisation sous UV (RED, Alpha innotech). L'échelle d'ADN utilisé étant le GeneRuler 100 pb Plus DNA Ladder (Invitrogen). La préparation et la migration de ces gels ont été effectués tel que décrit par Lee, Pei Yun et al [262]. L'appareil employé pour fournir le courant nécessaire à la migration étant un PowerPac™ Basic Power Supply (BIO RAD).

Préparation d'un gel agarose 1% typique pour cette maîtrise :1. Mesurer 0.30 g d'agarose (Wisent Bio) 2. Mélanger avec 33mL (pour compenser pour l'évaporation durant le chauffage de la solution) de tampon TAE 1X (TrisBase 40 mM (Sigma), acide acétique (Sigma) 20 mM et EDTA (Eurobio) 1mM) dans une flasque d'au moins 100 ml allant au four à micro-ondes. 3. Chauffer au micro-ondes jusqu'à ce que l'agarose soit entièrement dissoute (45 sec à 1min) en agitant périodiquement le mélange et en prenant garde de ne pas laisser déborder celui-ci. 4. Laisser refroidir la mixture jusqu'à 50 °C (Relativement confortable au touché) et ajouter 2 µl de Fluo-DNA/RNA gel staining solution (Zmtech scitific). 5. Agiter de sorte à bien mélanger puis verser lentement dans un moule à gel de taille adéquate (30ml) avec un peigne inséré. 6. Laisser reposer jusqu'à solidification complète à température ambiante (20-30 min). Pour préparer un gel à 2% nous avons simplement doublé la quantité d'agarose sans ajuster le volume de tampon TAE.

Migration par électrophorèse typique : 1. Un volume adéquat de tampon de chargement DNA Gel Loading Dye (6X) (Thermo scientific) est ajouté aux échantillons selon le volume à faire migrer de sorte à assister au chargement, permettre de repérer le progrès de la migration et augmenter avec sa concentration en glycérol la densité des échantillons et facilite leur dépôt au fond des puits plutôt que de diffuser dans le tampon. 2. Remplir la boîte de support pour le moule à gel de tampon TAE 1X. 3. L'échelle d'ADN GeneRuler 100 pb Plus DNA Ladder (Invitrogen) est chargée avec attention (2 µl) dans le premier puit du gel. 4. Les échantillons sont prudemment chargés dans le reste des puits du gel. Faire migrer à 80-150 Volts jusqu'à ce que la ligne du colorant atteigne 75-80% du gel ou la distance désirée.

5.1.4 Micro-injections

Les micro-injections dans les embryons 1-2 cellules pour l'ensemble de nos lignées ont été effectuées selon les principes décrits par Rosen, Jonathan N et al [263]. Seule la nature de la solution injectée étant différente selon l'expérience et la lignée concernée. La solution était placée par capillarité dans des aiguilles formées à partir de capillaires de borosilicate à épaisseur fine de 1.0 mm (Sutter Instrument) ayant été préalablement étirés jusqu'à séparation pour former deux aiguillées distinctes à l'aide du programme #9 de l'étireur de micropipette horizontale Flaming/Brown model-87 (Sutter instrument). La solution était ensuite injectée dans les cellules

en utilisant un injecteur à pression Picospritzer III (General Valve). De préférence l'injection se faisait à l'intersection du vitellus et de la cellule de l'embryon en passant par le vitellus pour ne pas endommager la cellule.

5.1.5 Réaction par amplification en chaîne (PCR)

Les amorces PCR ont été conçues à l'aide du logiciel SnapGene version 4.2.1.1 associé à l'outil en ligne Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) et suivant les directives générales établies par Chuang, L. Y. et al [264]. Sauf mention du contraire, les réactions PCR ont été faites avec 0.5 µL de dNTP (10 µM), 0.5 µL de chaque amorce (10 µM), 2.5 µL de tampon PCR 10x, 0.125 µL de Taq DNA polymérase (0.625 unités), 1 µL d'ADN génomique et de l'H₂O doublement distillée jusqu'à un volume de 25 µL. Selon les nécessités, des PCR dans des volumes moindres ou supérieurs ont été effectués en ajustant de manière proportionnelle les quantités de chaque réactif au nouveau volume final. L'appareil réemployé pour les PCR a été un t100 thermal cycler (BioRad). Tous les réactifs excluant l'eau et l'ADN étant du kit taq polymérase de GenedireX.

Amorces pour l'amplification de la sonde zfC9-HIS4 pour les HIS.

zfC9-HIS4-F (sens) : 5'-GGCGACTTGAATATCATCATGGCG -3'

zfC9-HIS4-R (antisens): 5'-CGGTAGATACTCTAGTA TCGGGTAC-3'

Programme PCR : 95 °C pour 5 min, suivi de 34 cycles de : 95 °C pour 30 s, 59 °C pour 60 s, 72 °C pour 45s et finalement 72 °C pour 10 min.

Amorces spécifiques au tag 12xHIS.

C13h9orf72-HIS-F (sens) : 5'-GAGCCATCATCACCATCACCACC-3'

C13h9orf72-RP-1 (antisens) : 5'-CTGCATGCAGATGCGTCCCTTG-3'

Programme PCR : 95 °C pour 5 min, suivi de 34 cycles de : 95 °C pour 30 s, 59 °C pour 30 s, 72 °C pour 45s et finalement 72 °C pour 10 min.

Amorces pour l'amplification de la matrice C9Nt-eGFP.

C9orf72-gfp-F1 (sens) : 5'-GAGGTCACATTTCTGGCCAATCACACGCTGAACGGTGA
GATTTTGC GCAGCGCA GAGAGCG **TGAGCAAGGGCGAGGAGCTG**-3'

C9orf72-gfp-R1 (antisens) : 5'-CAGAGAGACGATGATGACTCCCTTTTCAGCCAAGAC
GAAGAACTTGACATCCACCGCGCCA **AGATACATTGATGAGTTGGACAAACCAC**-
3'

Programme PCR : 95 °C pour 5 min, suivi de 34 cycles de : 95 °C pour 30 s, 60 °C pour 60 s, 72 °C pour 45s et finalement 72 °C pour 10 min.

Amorces pour l'amplification de la matrice d'ADN C9Ct-eGFP.

C9orf72-gfp-F2 (sens) : 5'-TGCGCAATCTGAAAACCGATTTGGACCTGACGGTGGAG
GGCGACTTGAATATCATCATGGCGATGGCGGAGAAG **GTGAGCAAGGGCGAGGAG**
CTG-3'

C9orf72-gfp-R2 (antisens) : 5'-GGAGGTCGCGCTCCTGCACGCTGGTCAGGAAAGAC
TTCCCGAAAACAAAAGAGTGAAGGCCGGCCCGCAG **AAGATACATTGATGAGTTG**
GACAAACCAC-3'

Programme PCR : 95 °C pour 5 min, suivi de 34 cycles de : 95 °C pour 30 s, 60 °C pour 60 s, 72 °C pour 45s et finalement 72 °C pour 10 min.

Les paires de bases surlignées en **vert** pour les amorces servant à l'amplification des matrices d'ADN eGFP Nt et Ct indiquent que ces bases sont celles qui se lient à la séquence de eGFP

dans le plasmide pEGFP-N1, alors que le reste des bases qui ne sont pas surlignées représentent les bras d'homologie pour l'intégration.

5.1.6 Extraction d'ADN génomique

L'extraction génomique de l'ADN a été effectuée dans 20 µL de NaOH 50 mM à partir de soit le segment de nageoire caudale coupée de poisson, d'embryons complets encore dans leur chorion ou de larves à des stades plus avancées. Les échantillons ont été bouillis pour 10 min et 1/10 du volume de Tris-HCL 100 mM a ensuite été ajouté pour tamponner la réaction tel que décrit par Meeker, N. D., et al [265]. Le résultat de l'extraction a ensuite été dilué 1/10 dans une solution d'H₂O deux fois distillée pour les extractions faites sur des embryons de plus de 24hpf, pour celles effectuées sur des regroupements de 5 embryons et pour celles effectuées sur les nageoires caudales de spécimens adultes.

5.1.7 Coupure de nageoires caudales des poisson-zèbres

Les poissons ayant atteint au moins le stade juvénile (4 semaines après fertilisation) et montrant une taille adéquate ont été anesthésiés dans la tricaine méthanesulfonate (MS222, Sigma-Aldrich) à une concentration finale de 160 mg/L jusqu'à ce que les mouvements cessent, le spécimen était alors placé sur une surface de plastique translucide et une petite section de la nageoire caudale était alors coupée avec un scalpel (fisher scientifique) puis placée dans un tube à PCR contenant une solution NaOH pour l'extraction d'ADN génomique. Les poissons étaient immédiatement replacés dans des aquariums remplis d'eau fraîche du système après l'opération terminée.

5.1.8 Analyse HRM et amorces utilisés

Les réactions PCR pour l'HRM ont été effectuées dans 2.5 µL de solution Precision Melt Supermix for HRM analysis (Bio-Rad #172-5112), 0.25 µL de chaque amorce (10 µM) et 1 µL d'ADN génomique et de l'eau doublement distillée jusqu'à un volume de 5 µL. La PCR a été effectuée dans un système LightCycler 96 (Roche) utilisant des plaques blanches à 96 puits ou des barrettes de 8 tubes selon le nombre d'échantillons. Le protocole PCR utilisé a été le Two-step Evagreen reaction protocol qui consiste en un programme de 95 °C pour 2 minutes (min), puis 45 cycles de 95 °C pour 10 secondes (s) et 60 °C pour 30 s, suivie de 95 °C pour 30 s et 60

°C for 60 s, la température augmentant de 0.02 °C/s jusqu'à 95 °C pour 10 s, puis refroidie à 40 °C. Les analyses des courbes de fusion ont été effectuées à l'aide du logiciel LightCycler 96 de la version 1.1 (Roche). Les amorces HRM ont été conçues à l'aide du logiciel SnapGene version 4.2.1.1 associé à l'outil en ligne Primer-BLAST (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) et suivant les directives générales établit par Chuang, L. Y. et al [264].

Amorces HRM pour le test d'efficacité de l'ARNg C9orf72Nt, l'identification d'indels et génotypage des spécimens de la lignée zfc9^{Ko}.

C9orf72Nt-FW (sens) : 5'-CGGAGAGGTCACATTTCTGGCC-3'

C9orf72Nt-RV (antisens) : 5'-GGATGTCAAGTTCTTCGTCTTGGC-3'

Amorces HRM pour le test d'efficacité de l'ARNg C9orf72Ct.

C9orf72Ct-FW (sens) : 5'-ATCTGAAAACCGATTTGGACC-3'

C9orf72Ct-RV (antisens) : 5'-GAAGTCTTTCCTGACCAGCG-3'

5.1.9 Séquençage

L'ensemble des échantillons à être séquencés au courant de ce projet l'ont été par le centre Genome Quebec/McGill utilisant la technologie Applied Biosystems 3730xl DNA Analyzer. Seuls les amorces et les échantillons d'ADN à être séquencés ont été préparés et envoyés selon leurs directives (<http://gqinnovationcenter.com/services/nanuq.aspx>). Les chromatogrammes résultant des séquençages ont été visualisés et analysés à l'aide du logiciel SnapGene version 4.2.1.1.

Amorce pour le séquençage du site des indels de la lignée zfc9^{KO} :

zfc9^{KO}-seq (sens) : 5'-CGCTTACTGGGACAACATCCTC-3'

Amorces pour le séquençage servant à vérifier l'intégration de HC9 et HC9-GFP dans les vecteurs recombinant des lignées transgéniques :

HC9orf72-R (antisens) : 5'-HGGGCTCCAAAGACAGAACAGG-3'

eGFP-R (antisens) : 5'-GCTTCTCGTTGGGGTCTTTGC-3'

Amorces pour la vérification du sens de l'intégration de la matrice d'ADN pour la sonde zfc9-HIS4 :

zfc9-HIS4-F (sens) : 5'-GGCGACTTGAATATCATCATGGCG-3'

5.2 Matériel et méthodes spécifiques à une expérience

5.2.1 Extraction d'ARN total et synthèse de l'ADN codant

L'ARN total d'une cohorte de 50 embryons de poissons zébrés de 24hpf a été extrait puis purifié de ses impuretés utilisant le kit TRIzol™ Plus RNA et le kit RNeasy Mini RNase-Free (DNase Set, QIAGEN, # de catalogue 79254), ensuite la synthèse de l'ADN codant (ADNc) à partir de l'ARN total extrait s'est fait l'aide du kit SuperScript First-Strand System for RT-PCR (Invitrogen, # de catalogue 11904018) et tout ceci selon la procédure décrite pour chaque kit utilisé et rassemblé sous le protocole décrit par Peterson, Samuel M et al [266].

Protocole de l'extraction de l'ARN total et synthèse de l'ADNc

Pour obtenir une quantité suffisante d'ARN, 50 embryons poissons-zèbres de 24 hpf ont été rassemblés dans un tube microfuge de 1.5 ml dans un minimum possible de volume d'H2O restant. Ensuite, 250 µl de réactif TRIzol ont été immédiatement ajoutés, sous la hotte chimique, au tube contenant les embryons puis les embryons ont été lysés et homogénéisés avec un pilon à pastille (pellet pestle, Sigma) jusqu'à perturbation jugée suffisante des tissus (environ 20 mouvements de va et vient). Une fois les cellules suffisamment homogénéisées, 750 µl de réactif TRIzol ont été ajoutés pour un volume total de 1ml puis, pour permettre la dissociation des complexes de nucléoprotéines, le mélange homogénéisé a été incubé pour 5 min. à la température de la pièce. Après, 0,2 ml de chloroforme est ajouté et le tube est agité vigoureusement 15 sec. pour mélanger et incubé 2 min. à la température de la pièce avant d'être

centrifugé (Centrifuge 5425 R, Eppendorf) à 12 000 g pour 15 min à 4°C. La mixture s'est alors séparée en une phase inférieure de phénol-chloroforme, une interphase et une phase supérieure translucide contenant l'ARN. Cette phase supérieure a été alors transférée dans un nouveau tube microfuge 1.5 ml à l'aide d'une pipette 1000 µl en faisant attention de ne pas aspirer de l'interphase. Pour précipiter l'ARN 0.5 ml d'isopropanol (Sigma) a été ajouté avant de laisser la solution reposer 10 min. à température pièce avant de centrifuger le mélange à 12 000 g pour 10 min. à 4°C, pour que l'ARN forme une pastille translucide au fond du tube. Le surnageant a été retiré doucement à l'aide d'une pipette pour ne pas perturber la pastille d'ARN, puis la pastille a été lavée dans 1 ml d'éthanol 75%. Le tout a alors été mélangé par de délicates inversions du tube qui a été centrifugé à 7 500g pour 5 min à 4°C. Après la centrifugation, l'éthanol a été retiré puis la pastille a été séché à l'air ambiant pendant 10 min. alors que le tube était inversé. La pastille d'ARN a ensuite été re-suspendue dans 100 µl de H₂O sans ARNase et incubée à 55 °C avec des agitations par vertex périodique du tube pour aider la réhydratation de l'ARN et la qualité de l'ARN obtenue été vérifiée à l'aide du ratio d'absorbance 260/280 nm (un ratio d'environ 2.0 étant considéré pur pour l'ARN) à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf). L'ARNm extrait a ensuite été traité à l'ADNase et purifiée sur colonne à l'aide et selon les instructions du kit RNeasy Mini RNase-Free (DNase Set, QIAGEN, # de catalogue 79254).

Préalablement à la synthèse de l'ADNc, l'ensemble de la zone de travail et les outils utilisés au cours de l'expérience ont été nettoyés avec de l'eau double distillée passée à l'autoclave (DDaH₂O) et rincés avec une solution (éthanol 70%/ DDaH₂O 30%), puis traités avec une solution dégradant l'ARNase (RNAaseZap, Invitrogen). Ceci était fait pour éviter toute contamination pouvant introduire de l'ARNase apte à dégrader l'ARN synthétisée. La synthèse de l'ADNc à partir de l'ARN total extrait s'est fait avec l'ARN total extrait et les réactifs du kit SuperScript First-Strand System for RT-PCR (Invitrogen, # de catalogue 11904018). Chaque réactif a été centrifugé avant de commencer puis le mélange ARN/ Amorce a été préparé dans un tube microfuge 0.5 ml stérile selon les indications suivantes : jusqu'à 8 µl d'ARN totale contenant jusqu'à 5 µg d'ARN, 1 µl de mix dNTP 10 mM, 1 µl d'Oligo(dT)₁₂₋₁₈ (0.5 µg/ µl) et un volume x d'H₂O ultrapure traitée au DEPC (Invitrogen) pour un volume total de solution de 10 µl. Le mélange a ensuite été incubé 5 min. à 65°C puis transféré immédiatement sur glace

pour 10 min. Une solution maître a été préparée durant la période de refroidissement contenant 2 µl de tampon RT 10X, 4 µl de MgCl₂ 25 mM, 2 µl de DTT 0.1 mM et 1 µl de RNaseOUT. Après le refroidissement de la première solution, 9 µl de la solution maître ont été ajoutés à la solution ARN/amorce puis le tout a été agité délicatement et rassemblé par centrifugation brève (quick spin). La solution totale de 19 µl a été incubée à 42°C puis 1 µl (50 unités) de SuperScript II RT a été ajouté à la solution qui a subi une autre incubation à 42°C pour 60 min. La réaction a été interrompue par une incubation à 70°C pour 15 min. suivi d'un refroidissement à 4 °C. Une fois la solution refroidie à 4 °C, 1 µl de RNase H a été ajouté au tube sur glace pour dégrader la matrice d'ARN ne laissant que le produit ADNc. Pour isoler et précipiter le produit ADNc des tube Phase Lock Gel (Quantabio) de 1.5 ml ont été utilisés. Ceux-ci ont d'abord été préparés par centrifugation à 12 000 g pour 2 min. pour que le gel se retrouve entièrement au fond du tube. Ensuite 81.5 µl de phénol (saturé en tris avec un pH 8, Sigma), 81.5 µl de 24 :1 chloroforme /alcool isoamylique et le produit d'ADNc (21 µl) ont été versés dans le tube puis mélangés plusieurs fois par de délicates inversions du tube. Le tube a ensuite été centrifugé 12 000 g pour 5 min. à la température de la pièce. L'ADNc devant se retrouver dans la phase aqueuse supérieure, celle-ci (~20µl) a été transférée dans un tube microfuge 1.5 ml stérile puis, 1 fois le volume récupéré a été ajouté en NaOAc 3M (Thermo Scientific) avant d'agiter le tout délicatement. Ensuite, 7 µl de glycogène 5 mg/ml ont été ajoutés avant d'agiter la solution délicatement puis d'ajouter 1 fois le volume de la solution (~30µl) en isopropanol et mélanger délicatement à nouveau. La solution a été incubée à la température de la pièce pour 10 min. et centrifugée à 12 000g pour 20 min. encore à la température de la pièce, de sorte qu'une petite plaque d'ADNc translucide se forme au fond du tube. Le surnageant a alors été prudemment retiré avec une pipette avant d'ajouter 500 µl d'éthanol 70% puis de mélanger la solution par inversion. L'échantillon a ensuite été centrifugé à 12 000 RPM pour 5 min. puis le surnageant a été retiré avec une pipette. De l'éthanol 70% a été ajouté à nouveau et cette étape a été répétée. Après centrifugation tout le liquide a été retiré du tube avec une pipette et la plaque d'ADNc a été séchée à l'air ambiant après avoir inversé le tube. Enfin, 20 µl d'H₂O sans nucléase ont été ajoutés au tube pour réhydrater la plaque d'ADNc et la concentration de cette dernière fut obtenue à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf). L'ADNc obtenue a été conservée à -20 °C.

5.2.2 Synthèse et clonage des sondes pour l'hybridation *in situ*

Les matrices d'ADN des sondes spécifiques à *zfc9* dont *zfc9*-HIS4 employées pour générer nos résultats d'HIS correspondant à la section 5' de la séquence codante et le dixième exon du transcrit *zfc9*La (793 pb amplifié à partir de l'ADNc d'embryons de 24hpf avec les amorces sens *zfc9*-HIS4-F et antisens *zfc9*-HIS4-R), ont été amplifiées par PCR de l'ADNc d'embryons 24hpf. L'amplification s'est faite dans un volume de 50 µl selon les paramètres décrits à la **section 5.1.5** avec un volume de solution d'ADNc contenant 100 ng d'ADNc et selon le programme PCR associé au pair d'amorces de chaque sonde. L'amplification a été confirmée par migration sur gel 1% de 3 µl du produit PCR et l'observation d'une seule bande à la hauteur correspondant à la taille attendue. La matrice d'ADN de sonde a ensuite été insérée dans le vecteur pCRTMII-TOPO® en utilisant les réactifs fournis et la procédure décrite dans le protocole associé au TOPO TA cloning kit (# catalogue : K4575J10, Invitrogen). La réaction étant faite d'un mix de 4 µL du produit PCR, 1 µL de solution de sel (200 mM NaCl; 10 mM MgCl₂) et 1 µL de vecteur TOPO® pour une solution de volume final de 6 µL. La solution a ensuite été mélangée délicatement et incubée 10 min. à température ambiante avant d'être placée sur glace.

Le produit du clonage TOPO a été ensuite transformé dans des bactéries *E. coli* compétentes One Shot® TOP10F (Invitrogen). Préalablement à la transformation : une fiole de milieu SOC a été chauffée à la température de la pièce, les plaques LB agar à être utilisées ont été préchauffées à 37°C, 40 µl de X-gal 40 mg/ml et 40 µl d'IPTG 100 mM ont été étendus avec des tiges en L stériles (Fisherbrand) sur chaque plaque LB-agar à proximité de la flamme, puis elles ont été incubées à 37°C jusqu'au moment de leur utilisation et enfin une fiole de cellules One Shot® a été décongelée sur glace. Pour la réaction de transformation elle-même, 2 µl de la réaction de clonage TOPO® a été ajoutés dans la fiole de bactérie *E. coli* chimiquement compétente et le tout a été doucement mélangé puis incubé sur glace 10 min. avant d'être soumis à un choc thermique. Le choc thermique a été provoqué en incubant les cellules 30 sec. à 42°C puis en transférant immédiatement le tube sur glace, après quoi 250 µl de milieu SOC à température pièce ont été ajoutés. Ensuite, le tube a été fermé de manière sûre et a été incubé 1 h à 37°C avec agitation horizontale 200 r.p.m., après quoi 50 µL du produit de transformation étaient étendus

en quantités différentes, pour assurer au moins une plaque avec des colonies bien espacées, sur au moins 2 des plaques LB agar préchauffées à proximité de la flamme d'un bec bunsen. Un volume de 20 µL de milieu SOC était ajouté à chaque plaque LB pour assurer un étalage uniforme des bactéries transformées et les plaques étaient refermées puis incubées durant la nuit à 37°C à l'envers pour éviter que la condensation n'affecte les bactéries. Les plaques LB ayant été enduites préalablement à l'application des bactéries transformées d'IPTG et de X-gal, il était possible de faire une sélection blanche/bleu des colonies bactériennes obtenues le lendemain. Les bactéries contenant un plasmide ayant intégrées la séquence de la sonde considérée ne devaient pas avoir d'activité β-galactosidase et étaient donc blanches ou bleues pâles. Des colonies blanches différentes ont donc ensuite été repiquées (4 à 8) avec des embouts de pipette stériles, chacune dans des tubes de plastique de culture contenant 4 mL de milieu LB auquel de l'ampicilline avait été ajouté de sorte à avoir une concentration finale d'ampicilline de 50 µg/mL. Ces tubes ont ensuite été incubés toute la nuit à 37 °C dans un incubateur permettant l'agitation à 200 rpm. Les cultures bactériennes obtenues au lendemain étaient alors utilisées pour extraire et purifier à l'aide du kit EZ-10 Spin Column Plasmid DNA Miniprep (Bio Basic, # de catalogue BS414) l'ADN plasmidique de la sonde concernée. La concentration d'ADN de la matrice pour chaque sonde était ensuite testée à partir de 1 µL de la solution d'extraction à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf). L'orientation d'insertion et la confirmation de présence de la séquence s'est faite par séquençage à l'aide des amorces sens respectives pour chaque sonde.

Les vecteurs étaient ensuite linéarisés à deux reprises avec des enzymes de restriction différentes, une fois de sorte à produire la sonde antisens et la seconde pour produire la sonde de contrôle négatif sens selon l'utilisation de la polymérase d'ARN sp6 ou T7. L'ADN plasmidique pour la sonde zfc9-HIS4 a été linéarisé à l'aide de NotI-HF® (New England Biolabs) pour la production de la sonde antisens et avec KpnI-HF® (New England Biolabs) pour la sonde sens, dans les deux cas sur 5µg d'ADN. L'ADN linéarisé a été purifié par extraction phénol/chloroforme. Ceci impliquait d'ajouter d'abord un volume égal à la solution de matrice d'ADN linéarisé de phénol-chloroforme puis de d'agiter vigoureusement par vortex 30 durant 30 sec la solution avant de centrifuger (Centrifuge 5425 R, Eppendorf) celle-ci 3 min à 13 000 r.p.m. à 4 °C. Le surnageant aqueux a ensuite été transféré dans un tube Eppendorf 1.5

mL stérile puis 10% du volume a été ajouté en acétate de sodium 3M ainsi que 2.5 X le volume en éthanol 100% pour précipiter l'ADN avant de centrifuger le tout à 13 000 r.p.m. 15 min à 4 °C. Après l'éthanol a été retiré puis la plaque d'ADN lavée avec 500 µl d'éthanol 75% puis centrifugé à 10 000 r.p.m. à 4 °C, puis le surnageant a été retiré prudemment et la plaque d'ADN séchée pour 5 min. à la température de la pièce en inversant le tube. La plaque d'ADN a été re-suspendue dans 10 µl d'eau traitée au DEPC et incubée 20 min. à 55 °C. Sa concentration a été évaluée à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf) et sa pureté évaluée par migration sur gel d'agarose 1%. Enfin la réaction de synthèse et le marquage des sonde sens ou antisens se sont effectués à l'aide du kit DIG RNA Labeling (SP6/T7) (Roche), la polymérase d'ARN sp6 ayant été utilisée pour produire la sonde zfC9-HIS4 sens alors que T7 a été utilisé pour produire la sonde zfC9-HIS4 antisens. La réaction de synthèse et marquage des sondes ARN s'est effectuée dans des tubes Eppendorf 200 µl stériles sans ARNase dans lesquelles 1-5 µg d'ADN de matrice de sonde purifié ont été ajoutés à 13 µl d'H₂O double distillés sans ARNase, et traité au DEPC (Invitrogen). Le tube de réaction a été refroidi sur glace avant d'y ajouter les réactifs suivants : 2 µl de mixture de marquage NTP 10X, 2 µl de tampon de transcription 10X, 1 µl d'inhibiteur protector ARNase et 2 µl de polymérase ARN SP6 ou T7 pour un volume final de 20 µl. Les réactifs ont ensuite été doucement agités puis rassemblés au fond du tube par une centrifugation à 2000 r.p.m. de 30 sec à température pièce puis la solution a été incubée 2 h à 37 °C. La matrice d'ADN a été retirée en ajoutant 2 µl d'ADNase I sans ARNase et en incubant la solution 30 min. à 37 °C. La solution a été arrêtée par l'ajout de 2 µl de EDTA 0.2 M à pH 8.0. L'ARN de la sonde a ensuite été précipitée en ajoutant 2.5 µl de LiCl 4 M et 75 µl d'éthanol mélangés et placés à -20 °C pour 45 min pour toute la nuit. Le lendemain, la solution a été centrifugée 30 min à 13 000 r.p.m. à 4 °C puis la plaque d'ARN obtenue lavée avec 500 µl d'éthanol 70% conservés à -20 °C, le tout a été centrifugé pour une période de 10 min. additionnelles à 13 000 r.p.m., le surnageant retiré et la plaque d'ARN séchée à l'air ambiant pendant 5 min. en inversant le tube. L'ARN a alors été re-suspendue dans 50 µl d'H₂O double distillée sans ARNase traitée au DEPC pour être incubée 30 min. à 37 °C. L'intégrité de la sonde et sa qualité ont été vérifiées en faisant migrer 2 µl du produit obtenu sur un gel d'agarose 1%, la présence d'une seule bande à la taille attendue confirmant le succès de la synthèse de la sonde. La sonde est finalement quantifiée à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf). L'ensemble des solutions et des manipulations pour chaque

étape étant effectué selon les indications de chaque kit et tel que décrit en un seul protocole de synthèse de sonde DIG pour l'HIS par Chitramuthu, Babykumari P et al [267].

5.2.3 Retrait du chorion et fixation des embryons pour l'HIS

Les embryons des stades 24,36 et 48 hpf à être utilisés pour les HIS ont été retirés de leur chorion manuellement à l'aide de pince aiguës (Dumont Watchmakers Forceps no. 5, Fine Science Tools). Pour éviter l'apparition de pigmentation chez les embryons, une concentration de 0.0045% de solution de PTU était maintenue dans leur eau. Les embryons ont ensuite été fixés au stade désiré dans une solution 4% (wt/vol) paraformaldéhyde et 1x PBS à 4°C durant toute la nuit dans des boîtes de pétri. Les embryons étaient ensuite lavés 3X 5 mins avec du PBS 1X déshydratés à la température de la pièce et progressivement par trempages successifs de 5 minutes dans des solutions PBS/méthanol de concentration croissante jusqu'à atteindre 100% méthanol. (PBS 1X/éthanol : 75/25%, 50/50%, 25/75%, 0/100%). Ils étaient ensuite conservés à -20°C jusqu'à leur utilisation.

5.2.4 Hybridation *in situ*

Les hybridation *in situ*, la préparation des solutions utilisées pour celles-ci et le montage des embryons au terme de l'HIS dans une solution 100% glycérol pour l'observation sous microscope ont été largement effectués selon le protocole décrit par Thisse, C. et al [231]. Les principales exceptions étant que 100 ng de sonde dans 200µl de solution d'hybridation ont été utilisés pour les périodes d'hybridation plutôt que 30 à 50 ng. Les sondes antisens servant à détecter l'expression du gène *zfc9*, les sondes sens étant utilisées comme contrôle négatif.

Protocole d'HIS complet :

Préalablement à l'expérience et au début du deuxième jour des manipulations, l'ensemble de la zone de travail et les outils utilisés au cours de l'expérience d'HIS ont été nettoyés avec de l'eau double distillée passé à l'autoclave (DDaH₂O) et rincés avec une solution (éthanol 70%/DDaH₂O 30%), puis traités avec une solution dégradant l'ARNase (RNAaseZap) ou autoclavé selon les contraintes de chaque objet. Ceci était fait pour éviter toute contamination pouvant introduire de l'ARNase apte à dégrader les sondes d'ARN utilisées pour l'HIS et donc de

compromettre l'expérience. Un volume de 1000 ml de DDaH₂O était préparé avant chaque *in situ* de sorte à disposer de solution H₂O stérile raisonnablement dépourvue d'ARNase.

Solutions et réactifs utilisés :

Albumine de sérum bovin, 99% (Sigma).

Anticorps : Anti-digoxigenin-AP Fab Fragment. (Roche, # de catalogue : 11093274910).

ARNt de wheat germ type V, poudre lyophilisé (Sigma) : La poudre lyophilisé a été re-suspendue dans de l'eau double distillée stérile à une concentration de 50 mg/ml, puis les protéines ont été retirées après plusieurs extractions phénol/chloroforme.

Glycérol $\geq 99\%$ (Sigma).

Héparine sel de sodium (Sigma).

Acide éthylènediaminetétraacétique (EDTA) pH 8 (Eurobio).

MeOH 99.8% (Sigma).

N,N-Diméthyleformamide anhydre (formamide déionisé), 99.8% (Sigma).

Proteinase K 10 mg/ml (Roche Diagnostics).

RNAaseZap (Invitrogen)

Sérum de mouton (Interchim SA).

Solution 20X SSC (Invitrogen) : solution avec des concentrations de 3 M de chlorure de Sodium et 0.3 M de citrate de sodium.

175.3 g de NaCl et 88.2g acide citrique sel de trisodium dihydraté dissous dans 1 L d'eau double distillée stérile.

Tampon phosphate salin (PBS) 1X (Sigma) : NaCl 137 mM, KCl 2.7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1.8 mM.

Solution de révélation : BM purple (Roche, # de catalogue : 11442074).

Tween 20 (Sigma).

PBT: 1X PBS, 0.1% Tween 20 vol/vol.

Mélange d'hybridation (MH) : 50% formamide déionisée, 5X SSC, 0.1% Tween 20, 50 mg/ml héparine, 500 mg/ml ARNt sans ARNase, pH 6.

Tampon Tris alcalin: 100 mM Tris HCl, pH 9.5, 50 mM MgCl₂, 100mM NaCl and 0.1% Tween 20 (vol/vol).

Solution d'arrêt : 1X PBS, pH 5.5, EDTA 1 mM, 0.1% Tween 20 (vol/vol).

Solution de blocage: 1X PBT, 2% sérum de mouton (vol/vol), 2 mg/ml albumine de sérum bovin (Sigma).

Jour 1

Les embryons fixés et déshydratés de 24,36 et 48 hpf, conservés à -20 °C, 20 de chaque stade par expérience, séparés tout au long du protocole en groupe selon leur stade de développement, ont premièrement été réhydratés. Pour y arriver, ils ont été d'abord transférés en groupe selon leur stade de développement dans des nacelles faites de mailles, complètement submergées dans des puits remplis d'une solution 100% MeOH d'une plaque de plastique à 24 puits reposant sur glace (afin de limiter l'évaporation du MeOH). Les mailles des nacelles étant de taille assez petite pour empêcher les embryons de quitter la nacelle, mais assez grande pour laisser les solutions pénétrer aisément leur intérieur. La réhydratation des embryons s'étant effectuée par trois trempages successifs de 5 mins dans des solutions progressivement moins concentrées, afin de ne pas endommager les embryons par une réentrée d'H₂O trop subite dans les tissus, en méthanol et plus concentrée en PBT 1X (MeOH/PBT 1X : 75/25%, 50/50%, 25/75%, 0/100%). Les solutions sont placées dans les puits de la plaque de manière ordonnée, de sorte à permettre un transfert rapide des nacelles à l'aide de pinces. Les embryons ont ensuite été lavés 4x 5mins dans des puits contenant une solution 100% PBT 1X, puis ils ont été perméabilisés à température pièce par incubation dans des puits contenant une solution de PBT 1X avec une concentration de protéinase K de 10 µg/ml. Le temps d'incubation pour une perméabilisation efficace sélectionnée étant de 15 mins pour les embryons de 24hpf, 20 mins pour ceux de 36 hpf et 30 mins pour les 48hpf. La digestion à la protéinase K est arrêtée en transférant les nacelles de chaque groupe d'embryon dans des puits contenant une solution PBT 1X % (wt/vol) paraformaldéhyde, une fois pour un rinçage rapide de sorte à retirer la plupart de la protéinase K et une seconde fois pour une incubation prolongée de 20 mins. Pour laver la protéinase K résiduelle, les embryons ont ensuite été transférés de puits 5x pour des lavages de 5mins dans du PBT 1X. Les embryons sont transférés de leur nacelle dans des tube Eppendorf 1.5ml stériles contenant 700 µl de milieu d'hybridation (MH). Le transfert s'est effectué à l'aide de pipettes 1000 µl (Rainin Classic Pipette PR-1000) avec des embouts stériles (autoclavés) dont

l'extrémité a été coupée (~ 2 mm d'ouverture), les embryons ont été délicatement aspirés du fond de leur nacelle et transférés en laissant ceux-ci descendre par gravité au bas de l'embout et en touchant délicatement la surface de la solution contenue dans le nouveau réceptacle. Cette procédure de transfert est répétée pour tous les transferts des embryons d'un contenant à l'autre pour l'ensemble du protocole d'HIS. Les embryons sont pré-hybridés par incubation dans le HM pendant 2 h à 70 °C, ceci afin de réduire le bruit de fond généré par la liaison non spécifique des sondes en saturant les sections de tissus pouvant autrement lier des acides nucléiques de manière aspécifique. Le MH est ensuite retiré presque complètement des tubes Eppendorf puis est remplacé avec 200 µl de MH frais contenant 100ng de sonde ARN marqué à la DIG, les embryons sont ainsi hybridés à 70 °C pendant la nuit (~16 h).

Jour 2

Toutes les solutions à partir de cette étape ont été préchauffées avant utilisation et maintenues à 70 °C jusqu'à mention du contraire, les plaques dont les puits contenaient les solutions et temporairement les nacelles avec les embryons étant conservées dans un incubateur sec à cette température, puisqu'une température élevée est nécessaire pour une forte stringence de l'hybridation de la sonde afin d'éviter un signal aspécifique. Les embryons ont été transférés à nouveau dans des nacelles submergées dans des puits contenant du MH sans tRNA et héparine (MH-). Le MH- a progressivement été changé pour une solution SSC 2X avec PBT 1X par une série de lavage de 10 mins chacun dans des puits contenant des solutions progressivement moins concentrées en MH et plus concentrées en SSC 2X avec PBT 1X (MH /SSC 2x PBT 1X : 75/25%, 50/50%, 25/75%, 0/100%). Les embryons ont été ensuite lavés 2 fois 30 mins avec des solution SSC 0.2X avec PBT 1X. Les solutions utilisées à partir de ce point n'étaient plus requises d'être maintenues à 70 °C, les embryons n'étant plus en solution permettant l'hybridation aspécifique.

La solution de SSC 0.2 X avec PBT 1X a été progressivement changée pour une solution PBT 1X pure, par une série de lavage de 10 mins par le transfert des embryons dans des tubes Eppendorf 1.5 ml contenant des solutions progressivement moins concentrées en SSC et plus concentrées en PBT 1X (SSC 0.2X / PBT 1X : 75/25%, 50/50%, 25/75%, 0/100%), chaque

lavage étant effectué à 40 r.p.m. en plaçant les tubes sur un agitateur horizontal. Les embryons ont ensuite été transférés dans des tubes Eppendorf 1.5 ml contenant 1.2 ml de solution de blocage et incubé 3h sous agitation pour saturer les sites de liaison aspécifiques de l'anticorps. L'anticorps détectant la DIG est ajouté à la solution de blocage dans chaque tube des différents groupes d'embryons de sorte à avoir une dilution finale de 1/5000, les embryons sont incubés pendant la nuit (~16 h) à 4 °C avec agitation à 40 r.p.m.

Jour 3

À la température de la pièce, la solution avec l'anticorps a ensuite été retirée quasi entièrement, tout en laissant une quantité suffisante pour empêcher le contact des embryons à l'air ambiant, puis 1.2 ml d'une solution de PBT 1X est ajouté aux tubes des embryons pour un lavage de 15 mins avec agitation à 40 r.p.m. Ce lavage a été répété un total de 6 fois avec une solution fraîche de PBT 1X pour chaque lavage. La solution de PBT 1X a ensuite été retirée pour permettre l'ajout de 1.2 ml d'une solution tampon Tris alcalin dans laquelle les embryons sont incubés 3 fois 5 mins par lavage avec agitation, chaque lavage s'est fait avec un volume de solution Tris alcaline fraîche. Les embryons ont ensuite été transférés dans les puits d'une plaque de céramique contenant 700 µl de solution de révélation BM purple (Roche), un substrat chromogénique de la phosphatase alcaline présente avec l'anticorps dont la réaction provoque une coloration (mauve/violet) permanente et spécifique au site de liaison de l'anticorps. La révélation a été suivie périodiquement à l'aide d'un stéréomicroscope Olympus SZX7 sous éclairage normal, entre les brèves observations les embryons étaient conservés dans l'obscurité puisque la lumière peut affecter la coloration. La révélation était suivie. La révélation s'effectuant à température pièce jusqu'à l'atteinte d'une coloration jugée adéquate, pour nos expériences ceci impliquait une révélation de 4 h en général. La réaction était arrêtée en transférant les embryons dans des tubes Eppendorf 1.5 ml puis en retirant un maximum de solution de révélation, puis lavant 3 fois 15 mins à température pièce avec agitation (40 r.p.m.) avec 1.2 ml de solution d'arrêt. Les embryons étaient alors transférés dans des plaques à 6 puits contenant une solution 100% glycérol avec un minimum de solution d'arrêt, la plaque a été placée sur un agitateur à tube à essai et agitée doucement à température pièce dans l'obscurité pendant la nuit.

Jour 4

Le lendemain les embryons sont monté dans du glycérol 100% afin d'être observés sous microscope. Pour monter les embryons, 4 lamelles 18mm x 18 mm (épaisseur 1.5) ont été collées ensemble avec une goutte de cyanoacrylate pour former un pont. Une goutte de glycérol contenant un embryon est placée dans le milieu de la lame entre deux ponts de lamelles et couvert par une lamelle plus large de 24 mm x 40 mm (épaisseur 1). Ce système a été utilisé pour rouler, à l'aide de la lamelle du dessus, les embryons dans la positions désirée.

Les embryons ont été observés à l'aide d'un microscope confocale Quorum Technologies (fixed stage Olympus microscope BX61 WI) sous éclairage normal à l'aide de la caméra intégrée et du logiciel d'analyse d'image Volocity® version 4.0 (Quorum Technologie) qui a aussi été utilisé pour le traitement de ces images.

5.2.5 Synthèse des ARNg pour CRISPR/Cas9

La synthèse des ARNg et de l'ARNm de la Cas9 avec un signal de localisation nucléaire utilisée a été effectué suivant le protocole décrit par Moreno-Mateos, Miguel A et al [268]. La purification du produit PCR servant de matrice pour la synthèse des ARNg se faisant sur colonne à l'aide du kit EZ-10 Spin Column PCR Products Purification (BIO Basic, # de catalogue BS364) et la transcription pour la production de l'ARNg se faisant à l'aide du kit AmpliScribe-T7-Flash transcription kit (Epicentre).

Protocole de la synthèse des ARNg

Les matrices d'ADN des ARNg ont été générées par PCR de remplissage (fill-in PCR). Un oligomère de 55-54 nt contenant un promoteur T7, la séquence de ~20 nt de chaque ARNg liant la séquence d'ADN du gène ciblé (*zfc9*) et une queue constante de 15 nt pour le recuit (annealing), ont été utilisés avec un oligonucléotide antisens universel de 80nt pour ajouter la queue 3' invariable des ARNg. La réaction PCR utilisée a été la suivante : 2.5 µL de tampon PCR II (10×) (AmpliTaq DNA Polymerase ki, Applied Biosystems), 0.5 µL de mix dNTP (10 mM par nt), 1.5 µL MgCl₂ (25 mM) (AmpliTaq DNA Polymerase kit), 1 µL d'amorce universelle (10 µM), 1 µL de l'amorce oligomère 55-54 nt spécifique de l'ARNg synthétisé (10 µM), 0.2 µl de la polymérase AmpliTaq (5 U/µL) et 18.3 µL d'H₂O sans ARNase/ADNase

pour un volume total de 25 μ L. Un produit de PCR de \sim 120 pb a été généré (selon l'ARNg) avec le programme PCR suivant : 3 min à 95°C, 30 cycles de 30 sec à 95°C, 30 sec à 45°C et 30 sec à 72°C, puis une étape finale à 72°C pour 7 min. Une partie de produit PCR a été visualisé sur gel d'agarose 2% pour confirmer l'amplification et le reste du produit a été purifié sur colonne à l'aide du kit EZ-10 Spin Column PCR Products Purification (BIO Basic, # de catalogue BS364), puis \sim 120-150 ng ont été utilisé comme matrice pour un transcription in vitro avec la polymérase d'ARN T7 à l'aide et selon les directives du kit AmpliScribe-T7-Flash transcription kit (Epicentre). La réaction s'est faite à 37°C. avec un mix de 6.3 μ L de matrice d'ADN d'ARNg, 1.8 μ L d'ATP (100 mM; du kit AmpliScribe), 1.8 μ L de CTP (100 mM; du kit AmpliScribe), 1.8 μ L de GTP (100 mM; du kit AmpliScribe), 1.8 μ L d'UTP (100 mM; du kit AmpliScribe), 2 μ L de DTT (100 mM) (kit AmpliScribe), 2 μ L de tampon de réaction AmpliScribe-T7 Flash 10X, 0.5 μ L d'inhibiteur d'ARNase RiboGuard (AmpliScribe kit) et 2 μ L d'enzyme AmpliScribe T7-Flash pour un volume de solution total de 20 μ L. Les ARNg obtenus ont été traités à la ADNase (pour retirer la matrice d'ADN) par l'ajout de 1 μ L d'ADNase 2 U/ μ L (TURBO DNase, Invitrogen) puis incubé 20 min à 37°C. Pour précipiter l'ARN, 80 μ L d'H₂O sans ARNase/ADNase ont été ajoutés à la solution d'ARNg traitée à la l'ADNase avec 10 μ L d'acétate de sodium 3 M pH 5.2, puis le tout a été mélangé par vortex. 300 μ L d'éthanol de grade ARN 95-100% ont été ajoutés ensuite puis le tout mélangé par vortex. La solution a été incubée 1h à -80 °C puis centrifugée (Centrifuge 5425 R, Eppendorf) à 16,100g pour 30 min à 4 °C. Le surnageant a ensuite été retiré délicatement puis 750 μ L ont été ajoutés pour rincer le précipité translucide avant de centrifuger à nouveau à 16,100g pour 5 min à 4 °C. Ces étapes ont été répétées puis le surnageant a été retiré délicatement avant de sécher la pastille translucide 5 min pour évaporer l'éthanol. La pastille a été re-suspendue dans 100 μ L d'H₂O sans ARNase/ADNase. 5 μ L ont été dilués dans 25 μ L d'H₂O sans ARNase/ADNase puis 1 μ L de cette dilution a été utilisé pour quantifier l'ARNg avec un spectrophotomètre BioPhotometer Plus (Eppendorf).

Amorces d'oligomère pour la synthèse des ARNg

Les ARNg ont été conçus utilisant les outils en ligne CRISPRscan (<http://www.crisprscan.org/>) et CHOPCHOP (<http://chopchop.cbu.uib.no/>).

Promoteur T7/ séquence spécifique ARNg/15 nt queue pour l'annealing avec l'amorce universel.

C9orf72Nt : 5'-taatacgaactactataGCGCAGCGCAGAGAGCGGCGCGGgttttagagctagaa-3'

C9orf72Ct : 5'-taatacgaactactataTGGCGATGGCGGAGAAAGCTGCGGgttttagagctaga-3'

(Les séquences PAM sont indiquées en rouge et le site de coupure prédit en orange)

Amorce oligomère universel

5'-AAAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACT
AGCCTTATTTAACTTGCTATTTCTAGCTCTAAAAC-3'

5.2.6 Synthèse de l'ARNm de Cas9

L'ARNm de Cas9 a été produit du plasmide pT3TS-nCas9n (#2656) codant pour une protéine Cas9 optimisée pour l'édition du génome chez le poisson-zèbre. Le produit linéarisé de la digestion du vecteur pT3TS-nCas9n par enzyme de restriction XbaI (New England Biolabs) a été extrait du gel agarose 1% après migration et purifié sur colonne à l'aide du kit EZ-10 Spin Column DNA Gel Extraction (BIO Basic, # de catalogue BS354), la transcription de l'ARNm de Cas9 a été faite à l'aide kit mMESSAGING mMACHINE™ T3 Transcription (Invitrogen) et la purification de l'ARNm a été effectuée à l'aide du kit RNeasy Mini (QIAGEN).

Protocole de synthèse de l'ARNm de Cas 9

L'ARNm de Cas9 a été produit du plasmide pT3TS-nCas9n (#2656) codant pour une protéine Cas9 optimisée pour l'édition du génome chez le poisson-zèbre. Pour fournir une matrice d'ADN pour la transcription de l'ARNm de Cas9, le plasmide pT3TS-nCas9n a d'abord été linéarisé à l'aide d'une incubation de 2 h à 37°C d'un mélange réactionnel contenant 2 µg de pT3TSnCas9n, 3 µL de tampon CutSmart (New England BIOLabs), 1 µL d'enzyme de restriction XbaI (10 U, New England BIOLabs) et un nombre de µL H2O sans ARNase/ADNase

pour un volume final de 30 μ L. Le produit linéarisé de la digestion du vecteur pT3TS-nCas9n par enzyme de restriction a été migré sur un gel agarose 1%, puis visualisé sous UV à longue longueur d'onde et extrait avec un scalpel pour être purifié sur colonne à l'aide et selon les indications du kit EZ-10 Spin Column DNA Gel Extraction (BIO Basic). La transcription in vitro de l'ARNm de Cas9, à partir de la matrice d'ADN obtenue, a été faite à l'aide kit mMESSAGE mMACHINE™ T3 Transcription (Ambion). Tout d'abord, le tampon de réaction 10X et le 2 \times NTP/CAP (ribonucléotides) ont été bien agités par vortex. Le 2 \times NTP/CAP a été gardé à -20°C dans une glacière de banc de laboratoire (benchtop cooler) comme pour tous les enzymes ou solutions sensibles utilisées dans les protocoles de ce projet de maîtrise, mais le tampon de réaction a été conservé à la température de la pièce lors de la préparation de la réaction puisque la spermidine dans le tampon peut précipiter à basse température et faire échouer la réaction. Pour la réaction de transcription les réactifs ont été ajoutés dans l'ordre suivant et le tout a été incubé à : 6 μ L de plasmide linéarisé (~200–300 ng d'ADN), 10 μ L de NTP/CAP (2X), 2 μ L de tampon de réaction (10X), 2 μ L de mélange d'enzyme-enzyme pour un volume total de 20 μ L. Le produit de réaction été traité avec à la ADNase (pour retirer la matrice d'ADN) par l'ajout de 1 μ L d'ADNase 2 U/ μ L (TURBO DNase, Invitrogen) puis incubé 20 min à 37°C. L'ARNm de Cas9 a ensuite été purifiée sur colonne à l'aide et selon les instructions du kit RNeasy Mini (QIAGEN). Des aliquotes de 5 μ L ont été préparées de l'ARNm de Cas9 purifiée et la concentration de l'ARNm a été obtenue un utilisant 1 μ L de cette dernière avec un spectrophotomètre BioPhotometer Plus (Eppendorf).

5.2.7 Test de l'efficacité de coupure des ARNg

Des poissons zébrés Tubingen long fin (TL) de type sauvage (WT) ont été collectés pour les micro-injections avec chaque ARNg afin de tester leur efficacité. L'efficacité des ARNg était testée pour trois concentrations pour chaque ARNg (30, 50 et 100ng/ μ L). Pour chaque concentration d'ARNg, en partant de la plus basse, 32 embryons récoltés étaient injectés au stade 1-2 cellules avec une goutte de ~1 nl d'un mix d'H₂O deux fois distillés sans ARNase, de colorant phénol rouge 10% v/v (Sigma-Aldrich) et de l'ARNm de l'ARNg à la concentration testée. Le phénol rouge étant présent pour permettre la visualisation des injections. L'ADN génomique des 16 meilleurs embryons injectés était alors extrait puis à partir de celui-ci la présence d'indels était constatée par analyse HRM et comparaison des profils de courbes

obtenues avec celle d'un embryon non injecté WT. Le taux d'efficacité étant établi comme le ratio d'embryons injectés analysés avec un profil de courbe HRM indiquant l'intégration d'indels sur le nombre d'embryons injectés analysés avec un profil WT. Si le taux était sous 80% alors les injections étaient répétées avec une concentration plus élevée jusqu'à un maximum de 100ng/μL après quoi, si l'efficacité demeurait sous 80%, l'ARNg était déclassé.

5.2.8 Synthèse des matrices d'ADN pour CRISPR/Cas9

La matrice d'ADN pour le tag 12x histidine a directement été synthétisée par les services de la compagnie invitrogen™. Les matrices eGFP Nt et Ct ont été générées par amplification PCR sur la séquence eGFP du plasmide pEGFP-N1 à l'aide de paires d'amorces composées d'une section homologue à la séquence d'eGFP, permettant sa liaison et une section composée d'une séquence homologue (bras d'homologie) au site d'insertion ciblé de *zfc9*, permettant l'insertion dirigée de eGFP au Nt ou au Ct. Par complétion de l'élongation on se retrouvait avec des matrices d'ADN incorporant les deux bras homologues flanquant la séquence de eGFP.

Matrice d'ADN pour le tag 12 x histidine Nt (C9Nt-12xHis)

Bras d'homologie 5' : AACGGTGAGATTTTGC GCAGCGCAGAGAGC

Bras d'homologie 3' : GGCGCGGTGGATGTCAAGTTCTTCGTCTTG

Séquence 12x histidine : GAGCCATCATCACCATCACCACCATCATCACCATCACCACG
G

Matrice d'ADN pour le eGFP Nt (C9Nt-eGFP)

Bras d'homologie 5' : GAGGTCACATTTCTGGCCAATCACACGCTGAACGGTGAGATT
TTGCG CAGCGCAGAGAGC

Bras d'homologie 3' : CAGAGAGACGATGATGACTCCCTTTTCAGCCAAGACGAAGA
ACTTGA CATCCACCGCGCC

Séquence eGFP amplifiée par PCR à partir de la séquence eGFP du vecteur pEGFP-N1 avec les amorces C9orf72-gfp-F1 et C9orf72-gfp-R1(**section 5.1.5**)

Matrice d'ADN pour le eGFP Ct (C9Ct-eGFP)

Bras d'homologie 5' : CTGCGCAATCTGAAAACCGATTTGGACCTGACGGTGGAGGG
CGACTTGAATATCATCATGGCGATGGCGGAGAAG

Bras d'homologie 3' : GGAGGTCGCGCTCCTGCACGCTGGTCAGGAAAGACTTCCCGA
AAACAAAAGAGTGAAGGCCGGCCCGCAG

Séquence eGFP amplifiée par PCR à partir de la séquence eGFP du vecteur pEGFP-N1 avec les amorces C9orf72-gfp-F2 et C9orf72-gfp-R2(**section 5.1.5**)

5.2.9 Micro-injections pour les lignées C9orf72^{NHis}, C9orf72^{NGFP} et C9orf72^{CGFP}

Des embryons de poissons zébrés Tubingen long fin (TL) de type sauvage (WT) ont été collectés pour les micro-injections. Pour la lignée C9orf72^{NHis} une goutte de ~1 nl d'un mix d'H₂O deux fois distillés sans ARNase, de 100ng/μL de l'ARNm de Cas9 de localisation nucléaire, 100ng/μL de l'ARNm de l'ARNg C9orf72Nt, 30 ng/μL de la matrice d'ADN C9Nt-12xHis et de colorant phénol rouge 10% v/v (Sigma-Aldrich) a été injectée dans des embryons au stade monocellulaire. Le phénol rouge étant présent pour permettre la visualisation des injections. Pour C9orf72^{NGFP} la démarche a été identique, sauf qu'on remplace 30 ng/μL de la matrice d'ADN C9Nt-12xHis par 30 ng/μL de la matrice d'ADN C9Nt-eGFP. Enfin, le tout est identique pour C9orf72^{CGFP}, sauf que l'ARNg est remplacé par C9orf72Ct concentré à 30ng/μL et la matrice d'ADN utilisée est C9Ct-eGFP concentrée à 30 ng/μL.

5.2.10 Sélection des poissons fondateur F0 des lignées C9orf72^{NGFP} et C9orf72^{CGFP}

Les embryons injectés avec les constructions génétiques C9Nt-eGFP et par C9Ct-eGFP ont été examinés après 24 heures post-fertilisation (hpf) en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Seulement les embryons exprimant une fluorescence aisément observable de la protéine eGFP ont été sélectionnés et élevés comme poissons fondateurs F0. L'observation était répétée matin et soir, chaque jour à partir de 24 hpf jusqu'à 72 hpf.

5.2.11 Sélection de la progéniture des F0 des lignées C9orf72^{NGFP} et C9orf72^{CGFP}

Les embryons résultant du croisement de fondateur F0 des lignées C9orf72^{NGFP} ou C9orf72^{CGFP} avec des TL WT étaient récoltés. Chaque croisement impliquant 1 spécimen F0 et 2 TL WT de sexe opposé au candidat F0. Les embryons obtenus de ces croisements ont été examinés après 24 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Seulement les embryons exprimant une fluorescence aisément observable de la protéine eGFP auraient été sélectionnés et élevés comme poissons F1 alors que leur progéniture aurait été identifiée comme un fondateur confirmé pour sa lignée. L'observation était répétée matin et soir, chaque jour à partir de 24 hpf jusqu'à 72 hpf. Après 120 embryons sans fluorescence un spécimen fondateur F0 était déclassé.

5.2.12 Sélection de la progéniture de la lignées C9orf72^{N12H}

Tous les spécimens injectés avec la construction C9Nt-12xHis survivants étaient considérés comme des fondateurs F0. Pour identifier les individus injectés ayant intégré le tag 12xHis et en mesure de transmettre cette intégration à leur descendance, des croisements des poissons injectés F0 avec des poissons TL WT ont été effectués. Chaque croisement impliquant 1 spécimen F0 et 2 TL WT de sexe opposé au candidat F0.

5.2.11.1 C9orf72^{N12H} knockin

Les embryons récoltés des croisements ont été rassemblés en cohortes de 5 pour en extraire à 8hpf l'ADN génomique. Une amplification par PCR était ensuite effectuée sur l'ADN génomique extrait et ce à l'aide des amorces spécifiques à l'intégration du tag 12xHis au Nt de *zfC9* (section 5.1.5). Le produit était ensuite migré sur gel d'agarose 1%. L'amplification réussie (présence de bandes) indiquant qu'il y aurait eu une transmission de l'insertion de 12xHis de la F0 à la F1. Les spécimens, dont 120 embryons étaient criblés sans identifier d'insertion, étaient déclassés.

C9orf72^{N12H} knock-out

Pour identifier les spécimens F0 en mesure de transmettre des indels, l'ADN génomique était extrait individuellement d'embryons 8hpf résultant des croisements. Une analyse par HRM était alors effectuée pour génotyper l'embryon en comparant le profil de courbe à celui d'un WT de référence, un profil distinct du WT indiquant la transmission d'indels au Nt de *zfC9*. Pour le criblage des différents spécimens des générations F1 et F2 de la lignée *zfC9*^{KO} la méthode est restée la même, excepté que l'ADN génomique a été extrait de fragments de nageoires caudales d'adultes plutôt que d'embryons entiers et pour les spécimens F2 la comparaison avec les profils connus de courbe HRM des 3 génotypes (WT, HT et HM) permettait le génotypage directe des spécimens une fois que les indels ont été caractérisées par séquençage. Pour la F2 seuls les spécimens avec un profil HT ou HM ont été conservés. Les spécimens F2 ont été obtenus du incross de spécimens F1 transmettant des indels résultant en le KO de *zfC9*.

5.2.13 Analyse statistique

Les données du nombre de spécimens identifiés pour chaque génotypes (WT, HT, HM) de la progéniture d'un croisement de spécimen F1 HTxHT de la lignée *zfC9*^{KO} ont été analysées à l'aide du test du χ^2 . Les proportions de génotype pour les adultes utilisés comme références pour tester notre hypothèse nulle ont été 25% WT, 25%HM et 50% HT, tel que prévu par les lois de Mendel sur la transmission des allèles lors d'un croisement HTxHT [238] si certains génotypes ne sont pas plus susceptibles de décéder. L'hypothèse nulle étant que les proportions des génotypes WT, HT et HM obtenues ne différaient pas des proportions attendues. Une valeur de $P < 0,05$ a été utilisée comme seuil de rejet de l'hypothèse nulle.

5.2.14 Tests de nage

Les tests de nage ont été effectués sur des larves âgées de 10 jours placées individuellement dans des plaques de 96 puits (Nunc™ MicroWell™ 96-Well Optical-Bottom Plates with Coverglass Base) remplies avec 200 µl d'eau du système à l'aide du système DanoVision (Noldus). Le programme de test durant lequel la nage était enregistrée a été composé d'une période de 30 minutes dans l'obscurité suivi de l'activation d'un éclairage à la lumière blanche maintenue durant 2 heures, après quoi le test de nage et l'enregistrement prenait fin. Les données acquises ont été traitées et analysées à l'aide du logiciel EthoVision XT 12. Les géotypes de chaque embryon testé ont été obtenus par analyse HRM, suite à l'extraction de leur ADN génomique à partir des larves entières. Les données montrées sont représentatives de la moyenne de la distance totale parcourue, pour chaque période de 10 minutes, à partir du milieu du puits pour chaque individu d'un même géotype (WT, HT, HM). Les barres d'erreurs indiquent l'erreur standard et les différences de valeur de nage ont été déterminées comme significative, si les intervalles d'erreurs pour une valeur donnée de distance de nage totale à une période de temps précise ne se superposaient pas avec les intervalles d'erreurs pour une valeur de distance de nage d'un autre géotype.

5.2.15 Constructions géniques Tol2 des lignée UAS-HC9-eGFP et UAS-HC9

Le tol2kit basé sur la technologie "multisite Gateway" a été utilisé pour assembler les constructions pour générer les lignées transgéniques de poissons zébrés HC9tg et HC9-eGFPtg selon le protocole décrit par Kwan K.M. et al [269]. Le "middle donor vector" pDONR221 contenant le gène humain C9orf72 a été généreusement fourni par le Dr Edor Kabashi qui l'avait employé lors de ses expériences visant à caractériser l'effet du KD de *zfc9* [105]. Pour produire la construction génique permettant l'obtention de la lignée HC9tg (pTOL2-UAS-C9orf72), ceci a été utilisé avec le "5' entry clone" p5E-UAS et le "3' entry clone p3E-polyA" pour insérer ces éléments dans le vecteur de destination pDestTol2CG2. La génération de la construction génique pour la lignée HC9-eGFPtg (pTOL2-UAS-C9orf72-eGFP) s'est effectuée de manière identique, mais avec l'addition du vecteur "3' entry clone p3E-EGFP". La réaction de recombinaison s'est faite avec 10 fmol de chacun des vecteurs d'entré/donneur et 20 fmol du vecteur de destination qui ont été mélangés avec le LR Clonase II Plus enzyme mix (Invitrogen)

et incubé à 25°C pour 16 h. Avant de transformer le produit de réaction dans des bactéries *E. coli* compétentes OneShot TOP10 cells (Invitrogen), 1 µl de solution protéinase K (Invitrogen) 2 µg/µl a été ajouté dans le mélange incubé à 37°C. À partir de l'étape de transformation de bactéries compétentes avec le produit de la réaction de recombinaison traitées à la protéinase K, le même protocole de transformation de bactérie que celui utilisé pour le clonage des vecteurs ayant intégrés les séquences de nos sondes pour l'HIS (**section 5.2.2**), a été utilisé pour les spécifications du matériel ainsi que de la méthode du clonage des vecteurs pTOL2-UAS-C9orf72 et pTOL2-UAS-C9orf72-eGFP. Cependant, comme ces vecteurs n'avaient pas d'activité β-galactosidase à perturber pour indiquer la recombinaison réussie, les plaques LB-agar n'ont pas été enduites d'IPTG et de X-gal préalablement à l'application sur plaque des bactéries transformées avec nos deux vecteurs. Des cultures de bactéries obtenues sur plaque, nous avons sélectionné les colonies bactériennes translucides pour être repiquées, puisqu'elles sont rapportées d'intégrer le clone désiré du vecteur recombinant 99% du temps [269]. Encore comme décrit à la **section 5.2.2**, les colonies sélectionnées ont été repiquées dans 4mL de milieu LB, puis placé toute la nuit à 37 °C dans un incubateur permettant l'agitation à 200 rpm. Les cultures bactériennes obtenues au lendemain étaient alors utilisées pour extraire et purifier à l'aide du kit EZ-10 Spin Column Plasmid DNA Miniprep (Bio Basic) l'ADN plasmidique pour chacun des vecteurs pTOL2-UAS-C9orf72 et pTOL2-UAS-C9orf72-eGF. La concentration d'ADN des vecteurs était ensuite testée à l'aide d'un spectrophotomètre BioPhotometer Plus (Eppendorf) et leur insertion confirmée par séquençage pTOL2-UAS-C9orf72 et pTOL2-UAS-C9orf72-eGFP. L'amorce HC9orf72-R reconnaissant spécifiquement la séquence de C9orf72 humain étant utilisée pour confirmer l'intégration de *C9orf72* dans les deux vecteurs et l'amorce eGFP-R, reconnaissant spécifiquement la séquence de eGFP, a été utilisée uniquement pour confirmer l'intégration d'eGFP dans le vecteur pTOL2-UAS-C9orf72-eGFP. L'ensemble des solutions et des manipulations pour chaque étape de cette section étant effectué selon les indications de chaque kit utilisé outre les exceptions mentionnées.

5.2.16 Injection Tol2 pour les lignées HC9-eGFptg et HC9tg

Des embryons de poissons zébrés TL de type sauvage (WT) ont été collectés pour les micro-injections. Une solution d'H₂O deux fois distillée sans ARNase, contenant 40 ng/µl du vecteur de HC9 pTOL2-UAS-C9orf72 (HC9) ou pTOL2-UAS-C9orf72-eGFP (HC9-eGFP) selon la

lignée à générer, 40 ng/μl d'ARNm de transposase et du colorant phénol rouge 10% v/v (Sigma-Aldrich), a été injectée dans des embryons au stade monocellulaire (~1 nl par embryons). Le phénol rouge étant présent pour permettre la visualisation des injections

5.2.17 Sélection des poissons fondateurs F0 des lignées HC9tg et HC9-eGFPTg

Les embryons injectés avec les constructions génétiques pTOL2-UAS-C9orf72 et pTOL2-UAS-C9orf72-eGFP ont été examinés après 27 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Seulement les embryons exprimant une fluorescence aisément observable de la protéine eGFP dans le cœur ont été sélectionnés et élevés comme poissons fondateurs F0. L'observation était répétée matin et soir, chaque jour à partir de 27hpf jusqu'à 72 hpf.

5.2.18 Sélection de la progéniture des F0 des lignées HC9tg et HC9-eGFPTg

Les embryons résultant du croisement de fondateur F0 des lignées HC9tg et HC9-eGFPTg avec des TL WT était récoltés. Chaque croisement impliquant 1 spécimen F0 et 2 TL WT de sexe opposé au candidat F0. Les embryons obtenus de ces croisements ont été examinés après 27 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Seulement les embryons exprimant une fluorescence aisément observable de la protéine eGFP ont été sélectionnés et élevés comme poissons F1 alors que leur progéniture a été identifiée comme un fondateur confirmé pour sa lignée correspondante. L'observation était répétée matin et soir, chaque jour à partir de 27hpf jusqu'à 72 hpf. Après 120 embryons sans fluorescence, un spécimen fondateur F0 était déclassé.

5.2.19 Sélection de la progéniture des lignées HC9tg et HC9-eGFPTg

Les embryons résultant d'incross des F1 des lignées HC9tg ou HC9-eGFPTg était récoltés. Chaque croisement impliquant 1 spécimen de chaque sexe. Les embryons obtenus de ces croisements ont été examinés après 27 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Seulement les embryons exprimant une fluorescence aisément observable de la protéine eGFP ont été sélectionnés et élevés comme poissons F2. L'observation était répétée matin et soir, chaque jour à partir de 27hpf jusqu'à 72 hpf.

5.2.20 Visualisation de l'expression de HC9-eGFP en présence de Gal4 endogène

Des embryons ont été récoltés du croisement poissons-zèbres F1 de la lignée HC9-eGFPtg avec des spécimens exprimant Gal4 de manière ubiquitaire et endogène. Les embryons obtenus de ce croisement ont été examinés après 24 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Ceci a été fait matin et soir chaque jour à partir de 24hpf jusqu'à 72 hpf. La présence de fluorescence ubiquitaire au-delà de la faible fluorescence observée naturellement chez les embryons poissons-zèbres à ce stade de développement signifiant une expression du transgène HC9-egGFP.

5.2.21 Confirmation de l'expression de Gal4 endogène

Des embryons ont été récoltés du croisement poissons-zèbres de la lignée exprimant Gal4 de manière endogène et ubiquitaire avec des spécimens exprimant eGFP de manière ubiquitaire en présence de Gal4. Les embryons obtenus de ce croisement ont été examinés après 27 hpf en utilisant le stéréomicroscope à épifluorescence Olympus SZX7 (Quorum Technologies). Ceci a été fait matin et soir chaque jour à partir de 24 hpf jusqu'à 72 hpf. La présence de fluorescence ubiquitaire au-delà de la faible fluorescence observée naturellement chez les embryons poissons-zèbres à ce stade de développement confirmant une expression de Gal4 dans la lignée utilisée pour le croisement avec HC9-eGFPtg pour visualiser l'expression de HC9-eGFP.

5.2.22 Injection Gal4 pour tester l'expression de HC9-eGFP

Des embryons issus d'incross de poissons zèbres F1 de la lignée HC9-eGFPtg ont été collectés pour les micro-injections. Une solution d'H2O deux fois distillée sans ARNase contenant 10 ng/µl d'ARNm et du phénol rouge 10% v/v (Sigma-Aldrich) a été injectée dans des embryons au stade monocellulaire (~1 nl par embryons). Le phénol rouge étant présent pour permettre la visualisation des injections. L'ARNm Gal4 utilisée avait déjà été synthétisé par Abdelhamid Bekri, un membre du laboratoire Drapeau, préalablement dans le cadre de son propre projet de recherche et il a gracieusement offert un échantillon.

5.2.23 Visualisation de l'expression de HC9-eGFP en présence de Gal4 exogène

Les embryons HC9-eGFP injectés avec Gal4 ont été initialement observés à 20hpf sous microscope à épifluorescence Olympus SZX7 (Quorum Technologies). Ceux pour lesquels un signal de fluorescence était détecté ont été fixés dans une goutte de 200 µl d'un gel 1% d'agarose à faible température de fusion (UltraPure™ Low Melting Point Agarose, Invitrogen) selon la méthode décrite par Keller, Philipp J [270]. La solution 1% agarose étant fondue à 80 °C, on attendait que le tube Eppendorf 1.5 ml le contenant ne provoque pas de douleur au touché avant de placer la goutte dans une boîte de pétri sur une lamelle et y transférer l'embryon. Il était alors orienté et maintenu avec un filament de plastique dans la position désirée jusqu'à solidification complète de gel. L'observation des embryons fixés s'est faite ensuite avec un microscope confocale Quorum Technologies (fixed stage Olympus microscope BX61WI) sous éclairage laser à 491nm pour l'observation du signal de eGFP, par l'entremise de la caméra intégrée au microscope et du logiciel d'analyse d'image Volocity® version 4.0 (Quorum Technologie) qui a aussi été utilisé pour le traitement des image obtenues.

Bibliographie

1. Oskarsson, B., et al. (2018). "Amyotrophic Lateral Sclerosis: An Update for 2018." *Mayo Clin Proc* 93(11): 1617-1628.
2. Isaacs, J. D., et al. (2007). "Amyotrophic lateral sclerosis with sensory neuropathy: part of a multisystem disorder?" *J Neurol Neurosurg Psychiatry* 78(7): 750-753.
3. Piccione, Ezequiel A et al. (2015). "Autonomic system and amyotrophic lateral sclerosis" *Muscle & nerve* vol. 51,5 : 676-9.
4. Grossman, A. B., et al. (2007). "Detecting neurobehavioral changes in amyotrophic lateral sclerosis." *Amyotroph Lateral Scler* 8(1): 56-61.
5. Woolley, S. C. and S. K. Jonathan (2008). "Cognitive and behavioral impairment in amyotrophic lateral sclerosis." *Phys Med Rehabil Clin N Am* 19(3): 607-617, xi.
6. Murphy, Jennifer et al. (2016) "Cognitive-behavioral screening reveals prevalent impairment in a large multicenter ALS cohort" *Neurology* vol. 86,9: 813-20.
7. Ferrari, R et al. (2011). "FTD and ALS: a tale of two diseases" *Current Alzheimer research* vol. 8,3 : 273-94.
8. Ingre, Caroline et al. (2015). "Risk factors for amyotrophic lateral sclerosis" *Clinical epidemiology* vol. 7 181-93. 12 Feb., doi:10.2147/CLEP.S37505.
9. Charcot, J. M. & Joffroy, A. (1869). "Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere". *Arch. Physiol. Neurol. Pathol.* 2, 744–754.
10. Charcot J-M. (1874). "De la sclérose latérale amyotrophique". *Prog Med*;2:325-327, 341-342, 453-455.
11. Shampo, Marc A. et al. (1993). "Lou Gehrig—Amyotrophic Lateral Sclerosis" *Mayo Clinic Proceedings* , Volume 68 , Issue 9 , 929.
12. Michaux L, Sampson M, Harl JM, et al. (1955). "Development of dementia in two cases of amyotrophic lateral sclerosis with aphasia". *Rev Neurol (Paris)*;92:357-67.
13. Lechelle P, Buge A, Leroy R.(1954). "Pick's disease; terminal amyotrophic lateral sclerosis. " *Bull Mem Soc Med Hop Paris*;70: 1090-4.
14. Plato CC, Reed DM, Elizan TS, et al. (1967). "Amyotrophic lateral sclerosis-Parkinsonism-dementia complex of Guam. IV. Familial and genetic investigations. " *Am J Hum Genet*;19:617-32.
15. Leigh, P. N., et al. (1991). "Ubiquitin-immunoreactive intraneuronal inclusions in amyotrophic lateral sclerosis. Morphology, distribution, and specificity." *Brain* 114 (Pt 2): 775-788.

16. Kawamata, T., et al. (1992). "Immunologic reactions in amyotrophic lateral sclerosis brain and spinal cord tissue." *Am J Pathol* 140(3): 691-707.
17. Rosen, D. R., et al. (1993). "Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis." *Nature* 362(6415): 59-62.
18. Lacomblez, L., et al. (1996). "Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II." *Lancet* 347(9013): 1425-1431.
19. Arai, Tetsuaki et al. "TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis." *Biochemical and biophysical research communications* 351 3 (2006): 602-11.
20. Neumann, M., et al. (2006). "Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis." *Science* 314(5796): 130-133.
21. Vance, C., et al. (2009). "Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6." *Science* 323(5918): 1208-1211.
22. Kwiatkowski, T. J., Jr., et al. (2009). "Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis." *Science* 323(5918): 1205-1208.
23. DeJesus-Hernandez, M., et al. (2011). "Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS." *Neuron* 72(2): 245-256.
24. Brooks BR. (1994). "El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. " *Journal of the neurological sciences*. Jul;124 Suppl:96-107.
25. Brooks, B. R., et al. (2000). "El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis." *Amyotroph Lateral Scler Other Motor Neuron Disord* 1(5): 293-299.
26. de Carvalho M, Dengler R, Eisen A, England JD, Kaji R, Kimura J, et al. (2008). "Electrodiagnostic criteria for diagnosis of ALS". *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*;119(3):497–503.
27. Cellura, E., et al. (2012). "Factors affecting the diagnostic delay in amyotrophic lateral sclerosis." *Clin Neurol Neurosurg* 114(6): 550-554.
28. Marin, B., et al. (2017). "Variation in worldwide incidence of amyotrophic lateral sclerosis: a meta-analysis." *Int J Epidemiol* 46(1): 57-74.

29. Zaldivar T, Gutierrez J, Lara G, Carbonara M, Logroscino G, Hardiman O. (2009). "Reduced frequency of ALS in an ethnically mixed population: a population-based mortality study. " *Neurology*;72(19):1640-5.
30. Wolfson, C., et al. (2009). "Incidence and prevalence of amyotrophic lateral sclerosis in Canada: a systematic review of the literature." *Neuroepidemiology* 33(2): 79-88.
31. Golby, R., et al. (2016). "Five-Year Incidence of Amyotrophic Lateral Sclerosis in British Columbia (2010-2015)." *Can J Neurol Sci* 43(6): 791-795.
32. Manjaly, Z. R., et al. (2010). "The sex ratio in amyotrophic lateral sclerosis: A population based study." *Amyotroph Lateral Scler* 11(5): 439-442.
33. Al-Chalabi, A. and O. Hardiman (2013). "The epidemiology of ALS: a conspiracy of genes, environment and time." *Nat Rev Neurol* 9(11): 617-628.
34. Huisman MH, de Jong SW, van Doormaal PT, et al. (2011) "Population based epidemiology of amyotrophic lateral sclerosis using capture-recapture methodology." *J Neurol Neurosurg Psychiatry*. 82(10):1165–1170.
35. Fang F, Valdimarsdottir U, Bellocco R, et al. (2009). "Amyotrophic lateral sclerosis in Sweden, 1991–2005. " *Arch Neurol*. 66(4):515–519.

Histoire naturelle

36. Ravits, J. M. and A. R. La Spada (2009). "ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration." *Neurology* 73(10): 805-811.
37. Swinnen, B. and W. Robberecht (2014). "The phenotypic variability of amyotrophic lateral sclerosis." *Nat Rev Neurol* 10(11): 661-670.
38. Rowland, L. P. and N. A. Shneider (2001). "Amyotrophic lateral sclerosis." *N Engl J Med* 344(22): 1688-1700.
39. Amato, A. & Russell, J. *Neuromuscular Disorders* 97–113 (McGraw-Hill, 2008)
40. Watermeyer, T. J., et al. (2015). "Impact of disease, cognitive and behavioural factors on caregiver outcome in amyotrophic lateral sclerosis." *Amyotroph Lateral Scler Frontotemporal Degener* 16(5-6): 316-323.
41. Ferrari, R., et al. (2011). "FTD and ALS: a tale of two diseases." *Curr Alzheimer Res* 8(3): 273-294.
42. Robberecht, W. and T. Philips (2013). "The changing scene of amyotrophic lateral sclerosis." *Nat Rev Neurosci* 14(4): 248-264.
43. Chio, A., et al. (2011). "Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study." *J Neurol Neurosurg Psychiatry* 82(7): 740-746.

44. van Es, M. A., et al. (2017). "Amyotrophic lateral sclerosis." *Lancet* 390(10107): 2084-2098.
45. Phukan, J., et al. (2007). "Cognitive impairment in amyotrophic lateral sclerosis." *Lancet Neurol* 6(11): 994-1003.
46. Kiernan, M. C., et al. (2011). "Amyotrophic lateral sclerosis." *Lancet* 377(9769): 942-955.
47. Picher-Martel, V., et al. (2016). "From animal models to human disease: a genetic approach for personalized medicine in ALS." *Acta Neuropathol Commun* 4(1): 70.
48. Brown, R. H. and A. Al-Chalabi (2017). "Amyotrophic Lateral Sclerosis." *N Engl J Med* 377(2): 162-172.
49. Petrov, Dmitry et al. "ALS Clinical Trials Review: 20 Years of Failure. Are We Any Closer to Registering a New Treatment?" *Frontiers in aging neuroscience* vol. 9 68. 22 Mar. 2017, doi:10.3389/fnagi.2017.00068
50. Bensimon, G., et al. (1994). "A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group." *N Engl J Med* 330(9): 585-591.
51. Miller, R. G., et al. (2007). "Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND)." *Cochrane Database Syst Rev*(1): CD001447.
52. Lacomblez, L., et al. (1996). "Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II." *Lancet* 347(9013): 1425-1431.
53. Al-Chalabi, A., et al. (2017). "July 2017 ENCALS statement on edaravone." *Amyotroph Lateral Scler Frontotemporal Degener* 18(7-8): 471-474.
54. Writing, G. and A. L. S. S. G. Edaravone (2017). "Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial." *Lancet Neurol* 16(7): 505-512.
55. Brown, R. H. and A. Al-Chalabi (2017). "Amyotrophic Lateral Sclerosis." *N Engl J Med* 377(2): 162-172.
56. Taylor, J. P., et al. (2016). "Decoding ALS: from genes to mechanism." *Nature* 539(7628): 197-206.
57. Talbot, K. (2011). "Familial versus sporadic amyotrophic lateral sclerosis--a false dichotomy?" *Brain* 134(Pt 12): 3429-3431.
58. Nguyen, H. P., et al. (2018). "ALS Genes in the Genomic Era and their Implications for FTD." *Trends Genet* 34(6): 404-423.
59. Chen, S., et al. (2013). "Genetics of amyotrophic lateral sclerosis: an update." *Mol Neurodegener* 8: 28.

60. Bunton-Stasyshyn, R. K., et al. (2015). "SOD1 Function and Its Implications for Amyotrophic Lateral Sclerosis Pathology: New and Renascent Themes." *Neuroscientist* 21(5): 519-529.
61. Hayashi, Y., et al. (2016). "SOD1 in neurotoxicity and its controversial roles in SOD1 mutation-negative ALS." *Adv Biol Regul* 60: 95-104.
62. Pansarasa, Orietta et al. "SOD1 in Amyotrophic Lateral Sclerosis: "Ambivalent" Behavior Connected to the Disease" *International journal of molecular sciences* vol. 19,5 1345. 3 May. 2018, doi:10.3390/ijms19051345
63. Neumann, M., et al. (2006). "Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis." *Science* 314(5796): 130-133.
64. Majumder, Vivek et al. "TDP-43 as a potential biomarker for amyotrophic lateral sclerosis: a systematic review and meta-analysis" *BMC neurology* vol. 18,1 90. 28 Jun. 2018, doi:10.1186/s12883-018-1091-7
65. Buratti, E. (2015). "Functional Significance of TDP-43 Mutations in Disease." *Adv Genet* 91: 1-53.
66. Lattante, S., et al. (2013). "TARDBP and FUS mutations associated with amyotrophic lateral sclerosis: summary and update." *Hum Mutat* 34(6): 812-826.
67. Martin, Sarah et al. "What causes amyotrophic lateral sclerosis?" *F1000Research* vol. 6 371. 28 Mar. 2017, doi:10.12688/f1000research.10476.1
68. Vance, C., et al. (2009). "Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6." *Science* 323(5918): 1208-1211.
69. Deng, H., et al. (2014). "The role of FUS gene variants in neurodegenerative diseases." *Nat Rev Neurol* 10(6): 337-348.
70. Kwiatkowski, T. J., Jr., et al. (2009). "Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis." *Science* 323(5918): 1205-1208.
71. Shang, Yulei and Eric J Huang. "Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis" *Brain research* vol. 1647 (2016): 65-78.
72. Deng, H. X., et al. (2010). "FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis." *Ann Neurol* 67(6): 739-748.
73. Farrarwell, Natalie E et al. "Distinct partitioning of ALS associated TDP-43, FUS and SOD1 mutants into cellular inclusions" *Scientific reports* vol. 5 13416. 21 Aug. 2015, doi:10.1038/srep13416
74. DeJesus-Hernandez, M., et al. (2011). "Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS." *Neuron* 72(2): 245-256.

75. Majounie, E., et al. (2012). "Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study." *Lancet Neurol* 11(4): 323-330.
76. Byrne, S., et al. (2012). "Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study." *Lancet Neurol* 11(3): 232-240.
77. Renton, A. E., et al. (2014). "State of play in amyotrophic lateral sclerosis genetics." *Nat Neurosci* 17(1): 17-23.
78. Van Mossevelde S., van der Zee J., Cruts M., Van Broeckhoven C. (2017) "Relationship between C9orf72 repeat size and clinical phenotype." *Curr. Opin. Genet. Dev.* 2017;44:117–124. doi: 10.1016/j.gde.2017.02.008.
79. Iyer, S., et al. (2018). "A comparative bioinformatic analysis of C9orf72." *PeerJ* 6: e4391.
80. Zhang, D., et al. (2012). "Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease." *Front Genet* 3: 283.
81. Xiao, S., et al. (2015). "Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis." *Ann Neurol* 78(4): 568-583.
82. Davidson, Y. S., et al. (2018). "Immunohistochemical detection of C9orf72 protein in frontotemporal lobar degeneration and motor neurone disease: patterns of immunostaining and an evaluation of commercial antibodies." *Amyotroph Lateral Scler Frontotemporal Degener* 19(1-2): 102-111.
83. Levine et al. (2013) Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. "The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs." *Bioinformatics.* 2013;29:499–503.
84. Wu, Xudong et al. (2011). "Insights regarding guanine nucleotide exchange from the structure of a DENN-domain protein complexed with its Rab GTPase substrate" *Proceedings of the National Academy of Sciences of the United States of America* vol. 108,46 (2011): 18672-7.
85. Levine, Timothy P et al. (2013). "Discovery of new Longin and Roadblock domains that form platforms for small GTPases in Ragulator and TRAPP-II" *Small GTPases* vol. 4,2 (2013): 62-9.
86. Daste, F., et al. (2015). "Structure and function of longin SNAREs." *J Cell Sci* 128(23): 4263-4272.
87. Marat, A. L., et al. (2011). "DENN domain proteins: regulators of Rab GTPases." *J Biol Chem* 286(16): 13791-13800.

88. Hutagalung, A. H. and P. J. Novick (2011). "Role of Rab GTPases in membrane traffic and cell physiology." *Physiol Rev* 91(1): 119-149.
89. Renton, Alan E et al. "A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD" *Neuron* vol. 72,2 (2011): 257-68.
90. Frick, Petra et al. "Novel antibodies reveal presynaptic localization of C9orf72 protein and reduced protein levels in C9orf72 mutation carriers" *Acta neuropathologica communications* vol. 6,1 72. 3 Aug. 2018, doi:10.1186/s40478-018-0579-0
91. Rizzu, Patrizia et al. "C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers" *Acta neuropathologica communications* vol. 4,1 37. 14 Apr. 2016, doi:10.1186/s40478-016-0306-7
92. Maharjan, N., et al. (2017). "C9ORF72 Regulates Stress Granule Formation and Its Deficiency Impairs Stress Granule Assembly, Hypersensitizing Cells to Stress." *Mol Neurobiol* 54(4): 3062-3077.
93. Farg, M. A., Sundaramoorthy, V., Sultana, J. M., Yang, S., Atkinson, R. A. K., Levina, V., Halloran, M. A., Gleeson, P. A., Blair, I. P., Soo, K. Y., King, A. E., Atkin, J. D. (2014). "C9ORF72, implicated in amyotrophic (sic) lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking." *Hum. Molec. Genet.* 23: 3579-3595, 2014.
94. Amick, J., et al. (2016). "C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling." *Mol Biol Cell* 27(20): 3040-3051.
95. Shi, Yingxiao et al. (2018). "Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons" *Nature medicine* vol. 24,3 (2018): 313-325.
96. Sivadasan, R., et al. (2016). "C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons." *Nat Neurosci* 19(12): 1610-1618.
97. Aoki, Y., et al. (2017). "C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia." *Brain* 140(4): 887-897.
98. Sellier, C., et al. (2016). "Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death." *EMBO J* 35(12): 1276-1297.
99. Fomin, V., et al. (2018). "The C9ORF72 Gene, Implicated in Amyotrophic Lateral Sclerosis and Frontotemporal Dementia, Encodes a Protein That Functions in Control of Endothelin and Glutamate Signaling." *Mol Cell Biol* 38(22).
100. Ugolino J., Ji Y.J., Conchina K., Chu J., Nirujogi R.S., Pandey A. et al. (2016) Loss of C9orf72 enhances autophagic activity via deregulated mTOR and TFEB signaling. *PLoS Genet.* 12, e1006443.
101. Zhang Y, et al. (2018). "The C9orf72-interacting protein Smcr8 is a negative regulator of autoimmunity and lysosomal exocytosis." *Genes Dev* 32:929–943.

102. O'Rourke, J G et al. (2016). "C9orf72 is required for proper macrophage and microglial function in mice" *Science (New York, N.Y.)* vol. 351,6279 (2016): 1324-9.
103. Therrien, Martine et al. (2013). "Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*" *PloS one* vol. 8,12 e83450. 12 Dec. 2013, doi:10.1371/journal.pone.0083450.
104. Corrionero, A. and H. R. Horvitz (2018). "A C9orf72 ALS/FTD Ortholog Acts in Endolysosomal Degradation and Lysosomal Homeostasis." *Curr Biol* 28(10): 1522-1535 e1525.
105. Ciura, S., et al. (2013). "Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis." *Ann Neurol* 74(2): 180-187.
106. Lagier-Tourenne, Clotilde et al. (2013). "Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration" *Proceedings of the National Academy of Sciences of the United States of America* vol. 110,47 (2013): E4530-9.
107. Koppers, M. et al. (2015). "C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits". *Ann. Neurol.* 78, 426–438 (2015).
108. Sullivan, P. M., et al. (2016). "The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway." *Acta Neuropathol Commun* 4(1): 51.
109. Sudria-Lopez, Emma et al. "Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects" *Acta neuropathologica* vol. 132,1 (2016): 145-7.
110. Atanasio, Amanda et al. "C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice" *Scientific reports* vol. 6 23204. 16 Mar. 2016, doi:10.1038/srep23204
111. Iyer, Shalini et al. "C9orf72, a protein associated with amyotrophic lateral sclerosis (AL) is a guanine nucleotide exchange factor" *PeerJ* vol. 6 e5815. 17 Oct. 2018, doi:10.7717/peerj.5815
112. Huotari, Jatta and Ari Helenius. (2011). "Endosome maturation" *EMBO journal* vol. 30,17 3481-500. doi:10.1038/emboj.2011.286
113. Vanlandingham, Phillip A and Brian P Ceresa. "Rab7 regulates late endocytic trafficking downstream of multivesicular body biogenesis and cargo sequestration" *Journal of biological chemistry* vol. 284,18 (2009): 12110-24.
114. Takahashi, S., et al. (2012). "Rab11 regulates exocytosis of recycling vesicles at the plasma membrane." *J Cell Sci* 125(Pt 17): 4049-4057.

115. Ao, X., et al. (2014). "Regulation of autophagy by the Rab GTPase network." *Cell Death Differ* 21(3): 348-358.
116. Webster, C. P., et al. (2016). "The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy." *EMBO J* 35(15): 1656-1676.
117. Yang, M., et al. (2016). "A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy." *Sci Adv* 2(9): e1601167.
118. Esseltine, J. L., et al. (2012). "Rab8 modulates metabotropic glutamate receptor subtype 1 intracellular trafficking and signaling in a protein kinase C-dependent manner." *J Neurosci* 32(47): 16933-16942a.
119. Giannandrea, M., et al. (2010). "Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly." *Am J Hum Genet* 86(2): 185-195.
120. Deretic, Vojo et al. (2013) "Autophagy in infection, inflammation and immunity" *Nature reviews. Immunology* vol. 13,10 (2013): 722-37.
121. Sardiello, M., et al. (2009). "A gene network regulating lysosomal biogenesis and function." *Science* 325(5939): 473-477.
122. Settembre, C. and A. Ballabio (2011). "TFEB regulates autophagy: an integrated coordination of cellular degradation and recycling processes." *Autophagy* 7(11): 1379-1381.
123. Nassif, M., et al. (2017). "The Enigmatic Role of C9ORF72 in Autophagy." *Front Neurosci* 11:442.
124. Sit ST, Manser E. (2011). "Rho GTPases and their role in organizing the actin cytoskeleton." *J Cell Sci.* 2011 Mar 1;124(Pt 5):679-83. doi: 10.1242/jcs.064964.
125. Moscat, Jorge and Maria T Diaz-Meco. (2009) "p62 at the crossroads of autophagy, apoptosis, and cancer" *Cell* vol. 137,6 (2009): 1001-4.
126. Balendra, R. and A. M. Isaacs (2018). "C9orf72-mediated ALS and FTD: multiple pathways to disease." *Nat Rev Neurol* 14(9): 544-558.
127. Ng, Adeline S L and Eng-King Tan. (2017) "Intermediate C9orf72 alleles in neurological disorders: does size really matter?" *Journal of medical genetics* vol. 54,9 (2017): 591-597.
128. Cruts M., Gijssels I., Van Langenhove T., Van Der Zee J., Van Broeckhoven C. (2013). "Current insights into the C9orf72 repeat expansion diseases of the FTL/ALS spectrum." *Trends Neurosci.* 2013;36:450-459.
129. Beck, Jon et al. (2013). "Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population" *American journal of human genetics* vol. 92,3 (2013): 345-53.

130. Ciesiolka, Adam et al. (2017). "Structural Characteristics of Simple RNA Repeats Associated with Disease and their Deleterious Protein Interactions" *Frontiers in cellular neuroscience* vol. 11 97. 11 Apr. 2017, doi:10.3389/fncel.2017.00097.
131. DeJesus-Hernandez, Mariely et al. "In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers." *Acta neuropathologica* vol. 134,2 (2017): 255-269.
132. Lagier-Tourenne, Clotilde et al. "Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration." *Proceedings of the National Academy of Sciences of the United States of America* vol. 110,47 (2013): E4530-9.
133. Liu EY, Cali CP, Lee EB. "RNA metabolism in neurodegenerative disease". *Dis Model Mech.* 2017;10(5):509-18.
134. Lee, Youn-Bok et al. "Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic" *Cell reports* vol. 5,5 (2013): 1178-86.
135. Cooper-Knock, J., et al. (2014). "Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions." *Brain* 137(Pt 7): 2040-2051.
136. Douglas, A. G. L. (2018). "Non-coding RNA in C9orf72-related amyotrophic lateral sclerosis and frontotemporal dementia: A perfect storm of dysfunction." *Noncoding RNA Res* 3(4): 178-187.
137. Lee, K. H., et al. (2016). "C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles." *Cell* 167(3): 774-788 e717.
138. Jiang, J., et al. (2016). "Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs." *Neuron* 90(3): 535-550.
139. Gendron, Tania F et al. "Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS" *Acta neuropathologica* vol. 126,6 (2013): 829-44.
140. Schludi, Martin H et al. "Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing." *Acta neuropathologica* vol. 130,4 (2015): 537-55.
141. Mackenzie, I. R. et al. (2015). "Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers." *Acta Neuropathol.* 130, 845–861.
142. Gomez-Deza, Jorge et al. (2015). "Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration" *Acta neuropathologica communications* vol. 3 38. 25 Jun. 2015, doi:10.1186/s40478-015-0218-y
143. Gitler AD, Tsuiji H. (2016). "There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS." *Brain Res.* 2016 Sep 15; 1647(0):19-29.

144. Wen, Xinmei et al. (2014). "Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death" *Neuron* vol. 84,6 (2014): 1213-25.
145. Kanekura, Kohsuke et al. (2014). "Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation" *Human molecular genetics* vol. 25,9 (2016): 1803-13.
146. Jovicic, A. et al. (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neurosci.* 18, 1226–1229 (2015).
147. Kwon, Ilmin et al. (2014). "Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells" *Science (New York, N.Y.)* vol. 345,6201 (2014): 1139-45.
148. Tao, Z., et al. (2015). "Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity." *Hum Mol Genet* 24(9): 2426-2441.
149. Yin, Shanye et al. (2017). "Evidence that C9ORF72 Dipeptide Repeat Proteins Associate with U2 snRNP to Cause Mis-splicing in ALS/FTD Patients" *Cell reports* vol. 19,11 (2017): 2244-2256.
150. Yamakawa, M., et al. (2015). "Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS." *Hum Mol Genet* 24(6): 1630-1645.
151. Zhang, Ke et al. (2015). "The C9orf72 repeat expansion disrupts nucleocytoplasmic transport" *Nature* vol. 525,7567 (2015): 56-61.
152. Moens, T. G. et al. (2018). "Sense and antisense RNA are not toxic in *Drosophila* models of C9orf72-associated ALS/FTD". *Acta Neuropathol.* 135, 445–457 (2018).
153. Swinnen, B. et al. (2018). "A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism". *Acta Neuropathol.* 135, 427–443 (2018).
154. Shaw, M. P., et al. (2018). "Stable transgenic C9orf72 zebrafish model key aspects of the ALS/FTD phenotype and reveal novel pathological features." *Acta Neuropathol Commun* 6(1): 125.
155. Tran, Helene et al. (2015). "Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a *Drosophila* Model of C9ORF72 FTD/ALS" *Neuron* vol. 87,6 (2015): 1207-1214.
156. Mizielińska, S. et al. (2014). "C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins." *Science* 345, 1192–1194 (2014).
157. May, Stephanie et al. (2014). "C9orf72 FTL/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration" *Acta neuropathologica* vol. 128,4 (2014): 485-503.
158. Yang, D. et al. (2015) "FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. " *Acta Neuropathol.* 130, 525–535 (2015).
159. Xu, Zihui et al. (2013). "Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration" *Proceedings of the National Academy of Sciences of the United States of America* vol. 110,19 (2013): 7778-83.
160. Rudich, Paige et al. (2017). "Nuclear localized C9orf72-associated arginine-containing dipeptides exhibit age-dependent toxicity in *C. elegans*" *Human molecular genetics* vol. 26,24 (2017): 4916-4928.

161. Swaminathan, A. et al. (2018). "Expression of C9orf72-related dipeptides impairs motor function in a vertebrate model." *Hum. Mol. Genet.* 27, 1754–1762 (2018).
162. Gendron, T. F. et al. (2015). "Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of C9ORF72 repeat expansion carriers." *Acta Neuropathol.* 130, 559–573 (2015).
163. Mackenzie, I. R. et al. (2015). "Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers." *Acta Neuropathol.* 130, 845–861 (2015).
164. Mackenzie, I. R. et al. (2013). "Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations." *Acta Neuropathol.* 126, 859–879 (2013).
165. Mann, D. M. et al. (2013). "Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72." *Acta Neuropathol. Commun.* 1, 68 (2013).
166. Davidson, Y. S. et al. (2014). "Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72." *Acta Neuropathol. Commun.* 2, 70 (2014).
167. Saberi, Shahram et al. (2017). "Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis" *Acta neuropathologica* vol. 135,3 (2017): 459-474.
168. Zhang, Yong-Jie et al. (2014). "Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress" *Acta neuropathologica* vol. 128,4 (2014): 505-24.
169. Ohki, Yu et al. (2017). "Glycine-alanine dipeptide repeat protein contributes to toxicity in a zebrafish model of C9orf72 associated neurodegeneration" *Molecular neurodegeneration* vol. 12,1 6. 14 Jan. 2017, doi:10.1186/s13024-016-0146-8.
170. O'Rourke, Jacqueline G et al. (2015). "C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD" *Neuron* vol. 88,5 (2015): 892-901.
171. Liu, Y. et al. (2016). "C9orf72 BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD." *Neuron* 90, 521–534 (2016).
172. Chew, J. et al. (2015). "C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits." *Science* 348, 1151–1154 (2015).
173. Belzil, Veronique V et al. (2013). "Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood" *Acta neuropathologica* vol. 126,6 (2013): 895-905.
174. Waite, A. J., et al. (2014). "Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion." *Neurobiol Aging* 35(7): 1779 e1775-1779 e1713.
175. Donnelly, Christopher J et al. (2013). "RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention" *Neuron* vol. 80,2 (2013): 415-28.

176. Haeusler, A. R., et al. (2014). "C9orf72 nucleotide repeat structures initiate molecular cascades of disease." *Nature* 507(7491): 195-200.
177. Gijssels, I., et al. (2016). "The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter." *Mol Psychiatry* 21(8): 1112-1124.
178. Sareen, D., et al. (2013). "Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion." *Sci Transl Med* 5(208): 208ra149.
179. Xi, Z. et al. (2015). "The C9orf72 repeat expansion itself is methylated in ALS and FTLN patients." *Acta Neuropathol* 129, 715–727 (2015).
180. Belzil, Veronique V et al. (2014). "Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients" *Brain research* vol. 1584 (2014): 15-21.
181. Van Blitterswijk, M., et al. (2015). "Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72." *Acta Neuropathol* 130(6): 863-876.
182. Li, Yanjie et al. (2015). "Expanded GAA repeats impede transcription elongation through the FXN gene and induce transcriptional silencing that is restricted to the FXN locus" *Human molecular genetics* vol. 24,24 (2015): 6932-43.
183. Haeusler, A. R., et al. (2016). "The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease." *Nat Rev Neurosci* 17(6): 383-395.
184. Harms, M. B., et al. (2013). "Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis." *Neurobiol Aging* 34(9): 2234 e2213-2239.
185. Fratta, P., et al. (2013). "Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia." *Acta Neuropathol* 126(3): 401-409.
186. Wilson-Sanders, S. E. (2011). "Invertebrate models for biomedical research, testing, and education." *ILAR J* 52(2): 126-152.
187. Vandamme, Thierry F. (2014). "Use of rodents as models of human diseases." *Journal of pharmacy & bioallied sciences* vol. 6,1 (2014): 2-9.
188. Hamilton, F. 1822. *An account of the fishes found in the river Ganges and its branches*. Edinburgh: A. Constable and Company.
189. Meyers, J. R. (2018). "Zebrafish: Development of a vertebrate model organism." *Current Protocols Essential Laboratory Techniques*, e19. doi: 10.1002/cpet.19.
190. Perry, S., et al. (2010). "Fish Physiology: Zebrafish." Academic Press, 23 juin 2010 - 468 pages
191. Cheng, K. C., et al. (2011). "Whole-animal imaging, gene function, and the Zebrafish Phenome Project." *Curr Opin Genet Dev* 21(5): 620-629.

192. Kalueff, Allan V et al. (2014). "Zebrafish as an emerging model for studying complex brain disorders" *Trends in pharmacological sciences* vol. 35,2 (2014): 63-75.
193. Creaser, C. W. (1934). "The technic of handling the zebra fish (*Brachydanio rerio*) for the production of eggs which are favorable for embryological research and are available at any specified time throughout the year". *Copeia*, 1934, 159. doi: 10.2307/1435845.
194. Urbani, E. (1948). "Studi di fisiologia embrionale dei teleostei; l'assunzione di ossigeno dell'uovo vergine fecondato e durante i primi stadi di sviluppo del *Brachydanio rerio*". *La Ricerca scientifica*, 18, 1047–1054.
195. Battle, H. I., & Hisaoka, K. K. (1952). "Effects of ethyl carbamate (Urethan) on the early development of the teleost *brachydanio rerio*". *Cancer Research*, 12, 334–340.
196. Stanton, M. F. (1965). "Diethylnitrosamine- induced hepatic degeneration and neoplasia in the aquarium fish, *brachydanio rerio*". *Journal of the National Cancer Institute*, 34, 117–130. doi: 10.1093/jnci/34.1.117.
197. Endo, A., & Ingalls, T. H. (1968). "Chromosomes of the zebra fish. A model for cytogenetic, embryologic, and ecologic study. *The Journal of heredity*", 59, 382–384. doi: 10.1093/oxfordjournals.jhered.a107755.
198. Streisinger G, Walker C, Dower N, Knauber D, Singer F. (1981). "Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*).". *Nature*. 1981 May 28;291(5813):293-6.
199. Walker C, Streisinger G. (1983). "Induction of Mutations by gamma-Rays in Pregonial Germ Cells of Zebrafish Embryos. *Genetics*. " 1983 Jan;103(1):125-36.
200. Stuart GW, McMurray JV, Westerfield M. (1988). "Replication, integration and stable germline transmission of foreign sequences injected into early zebrafish embryos. " *Development*. 1988 Jun;103(2):403-12.
201. Streisinger G, Singer F, Walker C, Knauber D, Dower N. (1986). "Segregation analyses and genecentromere distances in zebrafish. *Genetics*. " 1986 Feb;112(2):311-9.
202. Kirchberger, S., et al. (2017). "Quo natus, *Danio*?-Recent Progress in Modeling Cancer in Zebrafish." *Front Oncol* 7: 186.
203. Novoa, B. and A. Figueras (2012). "Zebrafish: model for the study of inflammation and the innate immune response to infectious diseases." *Adv Exp Med Biol* 946: 253-275.

204. Murray, P. J., et al. (2013). "Modelling Delta-Notch perturbations during zebrafish somitogenesis." *Dev Biol* 373(2): 407-421.
205. Sakai, C., et al. (2018). "Zebrafish Models of Neurodevelopmental Disorders: Past, Present, and Future." *Front Mol Neurosci* 11: 294.
206. Gemberling, Matthew et al. (2013). "The zebrafish as a model for complex tissue regeneration" *Trends in genetics : TIG* vol. 29,11 (2013): 611-20.
207. Van houcke, J., De Groef, L., Dekeyster, E., & Moons, L. (2015). "The zebrafish as a gerontology model in nervous system aging, disease, and repair." *Ageing Research Reviews*, 24, 358–368. doi: 10.1016/j.arr.2015.10.004.
208. Zorraca, V. and S. Mostowy (2018). "Zebrafish Infection: From Pathogenesis to Cell Biology." *Trends Cell Biol* 28(2): 143-156.
209. Kim, D. H., et al. (2017). "Pan-neuronal calcium imaging with cellular resolution in freely swimming zebrafish." *Nat Methods* 14(11): 1107-1114.
210. Psychéatrique : Brennan, C. H. (2011). "Zebrafish behavioural assays of translational relevance for the study of psychiatric disease." *Rev Neurosci* 22(1): 37-48.
211. Yvonne M. Bradford, et al . (2017). "Zebrafish Models of Human Disease: Gaining Insight into Human Disease at ZFIN." *ILAR Journal*, Volume 58, Issue 1, 1 July 2017, Pages 4–16, <https://doi.org/10.1093/ilar/ilw040>
212. Babin, P. J., et al. (2014). "Zebrafish models of human motor neuron diseases: advantages and limitations." *Prog Neurobiol* 118: 36-58.
213. eh, Tu-Hsueh & Liu, Han-Fang & Li, Yu-Wen & Lu, Chin-Song & Shih, Hung-Yu & Chiu, Ching-Chi & Lin, Sheng-Jia & Huang, Yin-Cheng & Cheng, Yi-Chuan. (2018). "C9orf72 is essential for neurodevelopment and motility mediated by Cyclin G1." *Experimental Neurology*. 304. 10.1016/j.expneurol.2018.03.002.
214. Stepto, A., et al. (2014). "Modelling C9ORF72 hexanucleotide repeat expansion in amyotrophic lateral sclerosis and frontotemporal dementia." *Acta Neuropathol* 127(3): 377-389.
215. Stainier, Didier Y R et al. (2017). "Guidelines for morpholino use in zebrafish" *PLoS genetics* vol. 13,10 e1007000. 19 Oct. 2017, doi:10.1371/journal.pgen.1007000

216. Kok, F O et al. (2014). "Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish" *Developmental cell* vol. 32,1 (2014): 97-108.
217. Rossi A, Kontarakis Z, Gerri C, Nolte H, Hölper S, Krüger M, et al. (2015). "Genetic compensation induced by deleterious mutations but not gene knockdowns." *Nature*. 2015;524:230–3. doi: 10.1038/nature14580 .
218. Sassen WA, Köster R. (2015). "A molecular toolbox for genetic manipulation of zebrafish." *Advances in Genomics and Genetics*, Volume 2015:5 Pages 151—163.
219. Koga A, Inagaki H, Bessho Y, Hori H. (1995). "Insertion of a novel transposable element in the tyrosinase gene is responsible for an albino mutation in the medaka fish.", *Oryzias latipes*. *Mol Gen Genet*. 1995 Dec 10;249(4):400-5.
220. Koga A, Suzuki M, Inagaki H, Bessho Y, Hori H. (1996). "Transposable element in fish." *Nature*. 1996 Sep 5;383(6595):30.
221. Kawakami K, Koga A, Hori H, Shima A. "Excision of the tol2 transposable element of the medaka fish, *Oryzias latipes*, in zebrafish, *Danio rerio*." *Gene*. 1998 Dec 28;225(1-2):17-22.
222. Grabher C, Wittbrodt J. "Meganuclease and transposon mediated transgenesis in medaka." *Genome Biol*. 2007;8 Suppl 1:S10.
223. Kawakami, Koichi. (2007). "Tol2: a versatile gene transfer vector in vertebrates" *Genome biology* vol. 8 Suppl 1, Suppl 1 (2007): S7.
224. Rosen, J. N., et al. (2009). "Microinjection of zebrafish embryos to analyze gene function." *J Vis Exp*(25).
225. Urasaki, A., et al. (2006). "Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition." *Genetics* 174(2): 639-649.
226. Asakawa, K. and K. Kawakami (2008). "Targeted gene expression by the Gal4-UAS system in zebrafish." *Dev Growth Differ* 50(6): 391-399.
227. Li, Mingyu et al. (2016) "Zebrafish Genome Engineering Using the CRISPR-Cas9 System" *Trends in genetics : TIG* vol. 32,12 (2016): 815-827.
228. van der Oost, J., et al. (2009). "CRISPR-based adaptive and heritable immunity in prokaryotes." *Trends Biochem Sci* 34(8): 401-407.

229. Cong, Le et al. (2013). "Multiplex genome engineering using CRISPR/Cas systems." *Science* (New York, N.Y.) vol. 339,6121 (2013): 819-23.

230. Li, Mingyu et al. (2016). "Zebrafish Genome Engineering Using the CRISPR-Cas9 System." *Trends in genetics : TIG* vol. 32,12 (2016): 815-827.

Résultats

231. Thisse, C. and B. Thisse (2008). "High-resolution in situ hybridization to whole-mount zebrafish embryos." *Nat Protoc* 3(1): 59-69.

232. Westerfield, M. (2007) "THE ZEBRAFISH BOOK, 5th Edition; A guide for the laboratory use of zebrafish (*Danio rerio*)", Eugene, University of Oregon Press. Paperback.

233. Snapp, Erik. "Design and use of fluorescent fusion proteins in cell biology" *Current protocols in cell biology* vol. Chapter 21 (2005): 21.4.1-21.4.13.

234. Canver, Matthew C et al. "Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells" *Journal of biological chemistry* vol. 289,31 (2014): 21312-24.

235. Loughran, S. T. and D. Walls (2011). "Purification of poly-histidine-tagged proteins." *Methods Mol Biol* 681: 311-335.

236. Samarut, Éric et al. (2016) "A simplified method for identifying early CRISPR-induced indels in zebrafish embryos using High Resolution Melting analysis" *BMC genomics* vol. 17 547. 4 Aug. 2016, doi:10.1186/s12864-016-2881-1

237. Wittwer, C. T. (2009). "High-resolution DNA melting analysis: advancements and limitations." *Hum Mutat* 30(6): 857-859.

238. Castle, W. E. (1903). "Mendel's Law of Heredity." *Science* 18(456): 396-406.

Discussion (current)

239. Atkinson, R. A., et al. (2015). "C9ORF72 expression and cellular localization over mouse development." *Acta Neuropathol Commun* 3: 59.

240. Hruscha, A., et al. (2013). "Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish." *Development* 140(24): 4982-4987.

241. Chitramuthu, Babykumari P and Hugh P J Bennett. "High resolution whole mount in situ hybridization within zebrafish embryos to study gene expression and function" *Journal of visualized experiments : JoVE* ,80 e50644. 19 Oct. 2013, doi:10.3791/50644

242. Oka, Y. and T. N. Sato (2015). "Whole-mount single molecule FISH method for zebrafish embryo." *Sci Rep* 5: 8571.

243. Choi, H. M. T., et al. (2018). "Third-generation in situ hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust." *Development* 145(12).
244. Hasegawa, Y., et al. (2015). "Variability of Gene Expression Identifies Transcriptional Regulators of Early Human Embryonic Development." *PLoS Genet* 11(8): e1005428.
245. Arslan-Ergul, Ayca and Michelle M Adams. (2014) "Gene expression changes in aging zebrafish (*Danio rerio*) brains are sexually dimorphic" *BMC neuroscience* vol. 15 29. 18 Feb. 2014, doi:10.1186/1471-2202-15-29
246. Wagner, E. M. (2013). "Monitoring gene expression: quantitative real-time rt-PCR." *Methods Mol Biol* 1027: 19-45.
247. Boel, A., et al. (2018). "CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments." *Dis Model Mech* 11(10).
248. Armstrong, Gary Alan Barclay et al. (2016). "Homology Directed Knockin of Point Mutations in the Zebrafish *tardbp* and *fus* Genes in ALS Using the CRISPR/Cas9 System" *PloS one* vol. 11,3 e0150188. 1 Mar. 2016, doi:10.1371/journal.pone.0150188.
249. Cornet, Carles et al. (2018). "Combining Zebrafish and CRISPR/Cas9: Toward a More Efficient Drug Discovery Pipeline" *Frontiers in pharmacology* vol. 9 703. 3 Jul. 2018, doi:10.3389/fphar.2018.00703
250. Trubiroha, A., et al. (2018). "A Rapid CRISPR/Cas-based Mutagenesis Assay in Zebrafish for Identification of Genes Involved in Thyroid Morphogenesis and Function." *Sci Rep* 8(1): 5647.
251. Shankaran, S. S., et al. (2017). "CRISPR/Cas9-Directed Gene Editing for the Generation of Loss-of-Function Mutants in High-Throughput Zebrafish F0 Screens." *Curr Protoc Mol Biol* 119: 31 39 31-31 39 22.
252. Burger, A., et al. (2016). "Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes." *Development* 143(11): 2025-2037.
253. Burberry, Aaron et al. (2016) "Loss-of-function mutations in the *C9ORF72* mouse ortholog cause fatal autoimmune disease" *Science translational medicine* vol. 8,347 (2016): 347ra93.

254. Moreno-Mateos, Miguel A et al. (2017) "CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing" *Nature communications* vol. 8,1 2024. 8 Dec. 2017, doi:10.1038/s41467-017-01836-2.
255. Varshney, Gaurav K et al. (2015) "High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9" *Genome research* vol. 25,7 (2015): 1030-42.
256. Ali, Zulfiqar and S Bala Bhaskar. (2016) "Basic statistical tools in research and data analysis" *Indian journal of anaesthesia* vol. 60,9 (2016): 662-669.
257. Kalueff, Allan V et al. (2013) "Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond" *Zebrafish* vol. 10,1 (2013): 70-86.
258. Stil, A. and P. Drapeau (2016) "Neuronal labeling patterns in the spinal cord of adult transgenic Zebrafish." *Dev Neurobiol* 76(6): 642-660.
259. Wen, Hua and Paul Brehm. (2010) "Paired patch clamp recordings from motor-neuron and target skeletal muscle in zebrafish" *Journal of visualized experiments: JoVE* ,45 2351. 20 Nov. 2010, doi:10.3791/2351.
260. Chen, Zheyang et al. (2013) "Primary neuron culture for nerve growth and axon guidance studies in zebrafish (*Danio rerio*)" *PloS one* vol. 8,3 (2013): e57539.
261. Green, Rachel and Elizabeth J Rogers. (2013) "Transformation of chemically competent *E. coli*" *Methods in enzymology* vol. 529 (2013): 329-36.
262. Lee, Pei Yun et al. "Agarose gel electrophoresis for the separation of DNA fragments" *Journal of visualized experiments : JoVE* ,62 3923. 20 Apr. 2012, doi:10.3791/3923
263. Rosen, Jonathan N et al. "Microinjection of zebrafish embryos to analyze gene function" *Journal of visualized experiments : JoVE* ,25 1115. 9 Mar. 2009, doi:10.3791/1115
264. Chuang, L. Y., et al. (2013). "Specific primer design for the polymerase chain reaction." *Biotechnol Lett* 35(10): 1541-1549.
265. Meeker, N. D., et al. (2007). "Method for isolation of PCR-ready genomic DNA from zebrafish tissues." *Biotechniques* 43(5): 610, 612, 614.

266. Peterson, Samuel M and Jennifer L Freeman. "RNA isolation from embryonic zebrafish and cDNA synthesis for gene expression analysis" *Journal of visualized experiments: JoVE* ,30 1470. 7 Aug. 2009, doi:10.3791/1470.
267. Chitramuthu, Babykumari P and Hugh P J Bennett. (2013). "High resolution whole mount in situ hybridization within zebrafish embryos to study gene expression and function" *Journal of visualized experiments : JoVE* ,80 e50644. 19 Oct. 2013, doi:10.3791/50644.
268. Vejnar, C. E., et al. (2016). "Optimized CRISPR-Cas9 System for Genome Editing in Zebrafish." *Cold Spring Harb Protoc* 2016(10).
269. Kwan K.M. , Fujimoto E., Grabher C., Mangum B.D., Hardy M.E., Campbell D.S., Parant J.M., Yost H.J., Kanki J.P., Chien C.B. (2007) The Tol2kit: a multisite gateway-based construction kit for Tol2 transposon transgenesis constructs. *Dev. Dyn.* , 236, 3088–3099.
270. Keller, Philipp J. "In vivo imaging of zebrafish embryogenesis" *Methods (San Diego, Calif.)* vol. 62,3 (2013): 268-78.

