



Ministère l'enseignement supérieur et de la recherche

ECOLE PRATIQUE DES HAUTES ETUDES

Sciences de la Vie et la Terre

Mémoire

Présenté par

MEDERIC Myriam

Pour l'obtention du Diplôme de l'École Pratique des Hautes Études : Étude de l'épissage alternatif des gènes *mec-2* et *xpb-1* chez *Caenorhabditis elegans*.

Soutenu le : 12 décembre 2022

Devant le jury suivant :

Dr ROSSEL Murielle Pr DOIGNON François Dr Massé Karine Dr DUPUY Denis Dr CATHELINE Gwenaëlle Directrice d'études Professeur Maitre de conférences Chargé de Recherche Directrice d'études

Présidente Rapporteur Examinateur Tuteur scientifique Tutrice pédagogique

Mémoire préparé sous la direction de :

Dr DUPUY Denis - Inserm U1212 / CNRS UMR 5320 – ARNA A l'Institut Européen de Chimie et Biologie à Pessac (Gironde) Directrice IECB : Dr GABELICA Valérie

et de :

Dr CATHELINE Gwenaëlle INCIA UMR 5287, Université de Bordeaux EPHE-PSL (Sciences de la Vie et de la Terre) Groupe de Recherche et d'Enseignement Thématiques de l'EPHE-PSL : TransNeuro

Remerciements

Mes premiers remerciements reviennent aux membres de mon jury Dr Rossel Mireille, Dr Doignon François, Dr Massé Karine Dr Denis Dupuy et Dr Gwenaëlle Catheline de m'avoir fait l'honneur d'accepter de juger mon travail pour le diplôme de l'Ecole Pratique des Hautes Etudes. Merci pour votre temps et pour votre implication.

Un grand merci à mon tuteur scientifique, Denis Dupuy, d'avoir accepté de me suivre tout au long de ces quatre années, de m'avoir accueillie dans son laboratoire, de m'avoir encadrée, formée et encouragée durant mon stage. Merci pour ta patience ce n'était pas simple tous les jours, d'avoir cru en moi et de m'avoir donné l'opportunité de reprendre mes études (à mon âge) en parallèle de mon travail à l'institut.

Merci à ma tutrice pédagogique, Gwenaelle Catheline pour ton soutien durant ces années, pour avoir été une écoute lors des moments de stress, et pour m'avoir poussée à croire en moi quand le doute pointait le bout de son nez.

Je remercie tous les membres du laboratoire, Camille, Sarah, Adnam, vous êtes une belle équipe. Merci à toi particulièrement Delphine, cela fait quatre ans que tu m'encadres, que tu me conseilles scientifiquement quand je butais sur un problème. Ton expertise dans les manips, a été primordiale pour l'avancement de mon projet scientifique et pour la rédaction de mon mémoire. Florian, ex-membre, qui a été un ami, une écoute avant que je ne commence ma formation, je n'oublierai pas les moments de partage et de fou rire. C'est grâce à toi que j'ai pu réaliser mon stage dans ce laboratoire, ton laboratoire, vu que Denis me voyait si souvent les après-midis avec toi.

Merci à Jean-Michel mon responsable hiérarchique, de m'avoir autorisé à faire cette formation en parallèle de mon travail. Merci à Jean-Louis Mergny directeur au début de ma formation, d'avoir autorisé que je m'inscrive à ce diplôme et à Valérie Gabélica ma nouvelle directrice, d'avoir permis de continuer et d'avoir une nouvelle meilleure organisation journalière, de mes semaines entre mon travail à l'institut et ma formation.

Merci à tous les membres de l'institut qui ont été patients et compréhensifs vis-à-vis de cette nouvelle organisation de travail.

Merci à toi Stéphanie, soutien de la première heure, dès mon arrivée à l'institut tu m'as accueillie avec beaucoup de sympathie. C'est très agréable de t'avoir à mes côtés. Merci pour ton aide sur les manips et pour ton soutien psychologique.

Merci à Fanny, on peut toujours compter sur toi, tu m'encourages à progresser et à avancer, merci d'exister. J'aime nos moments passés ensemble et les échanges que nous avons.

Merci aux collègues de l'unité pour votre bienveillance et votre soutien.

Merci à vous quatre, Thierry, Aurore, Isabelle et Mélanie, vous avez été et vous restez des personnes exceptionnelles, sans vous à mes côtés c'était peine perdue. Maintes fois vous m'avez boostée, stimulée, secouée quand il le fallait et bien sûr aidée autant de fois qu'il le fallait. J'ai de la chance de vous avoir. Ce n'était pas facile de reprendre ses études à mon âge avec la vie familiale, la vie professionnelle, les joies et les peines, c'était un défi à relever et avec vous ça a été plus simple.

Je veux remercier spécialement mon collègue Thierry, devenu un ami au fil des huit années à l'institut. Big force à toi pour ton implication, tu m'as portée et je te dois de ne pas lâcher.

Merci à mes enfants qui m'ont libérée du temps, m'ont soutenue. Merci à mon mari qui a été très patient, qui a accepté mes absences pour mes semaines de formation et qui m'a encouragée.

Par-dessus, tout je suis très reconnaissante à Dieu pour toutes ces personnes qu'il a mis sur ma route et autour de moi qui m'ont permise d'être là où j'en suis.

Table des matières

Liste d	des abréviations	7
Nome	enclature	8
Liste d	des figures	9
Liste d	des Tableaux	12
INTRO	DDUCTION	15
I.	Du matériel génétique à l'expression génétique	
Α.	Le matériel génétique	
В.	L'expression génétique	
	B.1. La transcription	
	B.2. Les modifications post-transcriptionnelle : La maturation des ARN pré-messagers	21
	B.2.a) L'épissage constitutif	21
	B.2.b) Le trans-épissage	22
	B.3. La Régulation post-transcriptionnelle	23
II.	Le nématode Caenorhabditis elegans	
III.	Étude de l'épissage de deux gènes présents chez C. elegans	30
Α.	Le gène de la méchanosensation mec-2	30
В.	Le gène x-box binding protein -1 (xbp-1)	
IV.	Objectif du projet	
MATE	ERIELS et METHODES	40
V.	Caenorhabditis elegans	41
Α.	Maintenance du ver <i>C. elegans</i>	41
В.	La décontamination et la synchronisation du ver C. elegans	41
C.	Transgenèse par micro-injection du ver C. elegans	42
D.	Observations microscopiques	43
VI.	Biologie moléculaire	44
Α.	PCR (Polymerase Chain Reaction) - classique	
В.	PCR sur colonie	44
C.	RT-PCR (Reverse Transcriptase PCR)	44
D.	RT-PCR direct sur <i>C. elegans</i> après mutagenèse	46
E.	Électrophorèse en gel d'agarose	46
F.	Purification de produit de PCR	46
G.	Technique de clonage, AQUA cloning (Advanced Quick Assembly cloning)	
Н.	Transformation bactérienne par choc thermique	47
I.	Extraction et purification de plasmide (miniprep)	
J.	Vérification des plasmides par digestion enzymatique	48
К.	Séquençage	
L.	Vérification de la synthèse protéique de la protéine XBP-1::eGFP par Western-blot	
VII.	Les milieux de culture et les solutions	50

Α.	Le	s milieux de culture	. 50	
В.	Le	s solutions	. 50	
RESULTATS				
VIII.	Étude	de l'épissage alternatif du gène mec-2	. 52	
Α.	Co	nstruction du plasmide pmec-2-a-b::MEC-2a-b E9-E10GGS6::fss	. 52	
	A.1.	Préparation du plasmide	. 52	
	A.2.	Observations microscopiques	. 57	
В.	Со	nstruction du plasmide pmec-2-c::MEC-2E9-E10-GGS6::FSS	. 60	
	B.1.	Préparation du plasmide pmec-2-c::MEC-2a-b E9-E10GGS6::FSS	. 60	
	B.2.	Préparation des mutagenèses dirigées	. 63	
IX.	Étude	de l'épissage du gène <i>xbp-1</i>	. 65	
Α.	Lig	née pxbp-1::xbp-1::FSS « Matrice Extrachromosomale»	. 65	
	A.1.	Préparation du plasmide	. 65	
	A.2.	Observations microscopiques	. 67	
В.	Lig	née pmyo3::xbp-1::fss « matrice extra chromosomale »	. 69	
	B.1.	Design du plasmide	. 69	
	B.2.	Observations microscopiques	. 71	
	В.З.	Vérification de l'épissage	. 75	
C.	Lig	née intégrée xbp-1(Cer125)	. 77	
	C.1.	Préparation pour l'intégration du système rapporteur dans génome	. 77	
	C.2.	Observations microscopiques	. 79	
	C.3.	Vérification de la transcription de <i>xbp-1::fss</i>	. 80	
	C.4.	Vérification de la séquence du gène xbp-1::FSS	. 82	
	C.5.	Vérification de l'épissage	. 82	
	C.6.	Vérification de synthèse de la protéine XBP-1::eGFP par la technique du Western blot	. 83	
DISCU	SSION	et PERSPECTIVES	. 87	
Référe	ences		. 94	
Annex	es		. 98	

Liste des abréviations

ADN : acide désoxyribonucléique ARNm: acide ribonucléique messager ARNsn: acide ribonucléique small nuclear C. elegans: Caenorhabditis elegans dNTP: déoxynucléotide triphosphate EA: épissage alternatif egl-15: egg laying defective E. coli: Escherichia coli ESE/ESS: exonic splicing enhancer/silencer FSS: Frame Switch Sensor GFP: Green Fluorescent Protein hnRNPs: heterogeneous nuclear ribonucleoproteins ISE/ISS: intronic splicing enhancer/silencer Kb: kilobases LB: Luria Broth L1, L2, L3, L4: stades larvaires 1, 2, 3, 4 mec-2/MEC-2: mecanosensory abdormal phenotype 2 mec-8/MEC-8: mecanosensory abdormal phenotype 8 mCherry: NGM: nematode growth medium NLS: nuclear localization sequence NMD: non-sens mediated decay pb : paire de base Pmec-2 : promoteur mec-2 PCR : polymerase chain reaction pré-ARNm : acide ribonucléique pré-messager PVM: posterior ventral microtubule RRM: recognition RNA motif SL: splice leader SR: Résidue serine et arginine snRNP: small nuclear ribonucleoprotein TRNs : Touch receptor neurons UTR : untranscribed region WT : Wild Type

Nomenclature

- Les gènes sont désignés par 3 ou 4 lettres minuscules, un tiret et un chiffre (tous en italique), ex : mec-2 et xbp-1.
- Les ARNs sont désignés par le nom du gène en italique ex : mec-2
- Les protéines sont désignées par le nom du gène en majuscules et non en italique ex : MEC-8.
- Deux lettres majuscules et un chiffre, désignent une souche contenant une ou plusieurs différences génétiques. Les lettres indiquent le laboratoire qui a construit la souche.
- Pour distinguer les allèles d'un même gène, les noms des allèles sont placés entre parenthèses, sans espace entre le gène et le nom de l'allèle. Le nom du gène et celui de l'allèle, par exemple mec-7(e1343) et mec-7(e1506).
- La région promotrice en amont d'un gène est indiquée par le nom du gène suivi d'un "p".
- Les noms du promoteur et du codage de la protéine sont séparés par deux points, comme le sont les parties des protéines de fusion. Ainsi, un gène codant pour une fusion translationnelle MEC-7::GFP pilotée par le promoteur mec-7 sera mec-7::gfp, tandis qu'une fusion transcriptionnelle mab-5 sera mab-5p::gfp.
- Les phénotypes mutants sont désignés par le nom du gène non en italique et par la première lettre en majuscule par exemple, les animaux Unc ou Sma sont non coordonnés ou de petite taille.

Liste des figures

Figure 1: Schéma d'une cellule eucaryote et procaryote montrant les différences principales	17
Figure 2: Schéma de la cellule eucaryote à l'expression des gènes	17
Figure 3: Le Nucléoside et l'ADN forme B	18
Figure 4 : schéma du mécanisme de transcription dans la cellule eucaryote	19
Figure 5: Unité de transcription eucaryote et les régions cis se trouvant dans le site du promoteur.	19
Figure 6: Formation du complexe de pré-initiation	20
Figure 7: Structure de l'ARNm eucaryote mature prêt à être traduit	21
Figure 8: Schéma des différentes étapes de l'épissage	22
Figure 9: L'épissage alternatif	24
Figure 10: Représentation schématique des différents modes d'épissage alternatif.	24
Figure 11: Les séquences cis et trans régulatrices de l'épissage	25
Figure 12: Anatomie de C. elegans	26
Figure 13: Cycle de vie du nématode C. elegans à t° 20°C	27
Figure 14: Système nerveux de C. elegans	28
Figure 15: Observation de C. elegans.	29
Figure 16 : Complexe de mécano-transduction proposé dans les récepteurs tactiles corporels de C. elegans	30
Figure 17: Représentation des NRT chez C. elegans.	31
Figure 18 : Mode d'action supposé de la protéine MEC-8 sur pré-ARNm mec-2	32
Figure 19: Schéma de l'épissage alternatif du gène mec-2.	32
Figure 20: Les différents isoformes de la protéine MEC-2 sous contrôle du promoteur distal	33
Figure 21: Schématisation des voies de signalisation de l'UPR.	34
Figure 22 : Représentation des variantes de la protéine XBP1 (XBP1U et XBP1S)	35
Figure 23 : Schématisation de la voie IRE-1 pour l'épissage de l'ARNm de xbp-1	36
Figure 24 : Eléments d'action cis de XBP-1s	36
Figure 25 : Photographie de la lignée transgénique xbp-1::GFP	37
Figure 26 : Photo de la méduse Aequorea victoria et du corail Discosama	38
Figure 27: Epissage du gène egl-15	39
Figure 28: Principe du clonage AQUA cloning.	47
Figure 29: Schéma de la construction du plasmide pmec-2-a-b::MEC-2a-b E9-E10GGS6::FSS.	52
Figure 30: Représentation de l'hybridation des amorces 1158 et 1159 pour la linéarisation du plasmide pmec-2 b::MEC-2E9-E10-GGS6::eGFP	-a- 53
Figure 31: Représentation de l'hybridation des amorces 1146 et 1147 pour l'extraction du fragment comportan mCherry dans le plasmide pmec-2-a-b::MEC-2E9-E9'-GGS6::mCherry	ıt la 53

Figure 32: : migration sur gel d'agarose des produits d'amplification de la mCherry et du plasmide pmec-2-a- b::MEC-2E9-E10-GGS6::eGFP
Figure 33: Migration de pmec-2-a-b::MEC-2E9-E10-GGS6::FSS
Figure 34 : Profil de l'alignement des séquences dans le logiciel « Clustal oméga » après séquençage 56
Figure 35 : Schéma des protéines fluorescentes produites en l'absence de MEC-8
Figure 36 : Schéma des protéines fluorescentes produites en présence de MEC-8
Figure 37: Photos d'un ver L4 en microscopie à fluorescence grossissement x20 58
Figure 38: Photos avec les différents filtres, localisation des protéines MEC-2A et MEC-2B dans les neurones. Grossissement x40
Figure 39: Représentation de l'hybridation des amorces 1128 et 1129 pour l'extraction du promoteur mec-2p-c dans le plasmide pmec-2-c::MEC-2E9-E10-GGS6::eGFP
Figure 40: Représentation de l'hybridation des amorces 1131 et 1132 pour l'ouverture du plasmide pmec-2-a- b::MEC-2E9-E10-GGS6::FSS ôtant le promoteur mec-2p-a-b
Figure 41: Migration des produits de PCRs
Figure 42 : Migration du plasmide pmec-2-c::MEC-2E9-E10-GGS6::fss digéré
Figure 43: Représentation graphique des 4 groupements GCAC sur le plasmide pmec-2-a-b::MEC-2a-bE9- E10GGS6::FSS
Figure 44: Schéma de l'épissage cytoplasmique alternatif du gène xbp-1 avec les 2 fluorochromes eGFP et DsRed
Figure 45: Schéma du plasmide pxbp-1::xbp-1::fss
Figure 46: Observation microscopique d'un ver adulte transgénique issue de la lignée pxbp-1::xbp-1::FSS
Figure 48: Photographie d'un ver adulte de la lignée pxbp-1::xbp-1::FSS67
Figure 49: Schéma du plasmide pmyo-3::xbp-1::FSS
Figure 50: Profil de migration de la digestion du plasmide pmyo-3::xbp-1::FSS2 par l'enzyme BamH-1HF 69
Figure 52: Photographie des vers de la lignée pmyo-3::xbp-1 ::FSS
Figure 51: Photographies de vers de la lignée pmyo-3::xbp-1::fss, au stade de développement L1, L3, L4 et adulte. 74
Figure 53: Schéma du positionnement des 2 amorces choisies au plus près de l'intron non conventionnel pour étudier l'épissage
Figure 54: Séquence du plasmide avec la représentation de l'hybridation des amorces 957 et 1300 servants pour l'amplification de la zone autour de l'intron alternatif
Figure 55: Gel agarose 1.5% des PCR et Gel bis-acrylamide 8%76
Figure 56: Représentation du processus insertion dans le génome du système rapporteur fluorescent selon la technique CRISPR-Cas9 dans le gène xbp-1
Figure 57: Montage de photos de l'abdomen du ver C. elegans de la lignée xbp-1 vue au microscope à fluorescence avec les filtres gfp et dsRed prise par le laboratoire du Dr Ceron
Figure 58: Photographies d'un ver adulte de la lignée xbp-1 (Cer125) au microscope à épi fluorescence et confocale prise au laboratoire
Figure 59: Profil de migration d'ARN totaux et de Rt-Pcr du gène F46F2.2 80
Figure 60: Tableau des amorces de vérification de l'insertion du système rapporteur dans le génome

Figure 61: Représentation schématique de l'ADNc de xbp-1(Cer125) après la RT-PCR sur les ARN totaux 8	31
Figure 62: Migration sur gel agarose 1 % des PCR réalisées sur le gène xbp-1(Cer125). Les bandes obtenues montrent l'insertion du rapporteur	32
Figure 63: Migration sur gel agarose 1 % et gel de bis-acrylamide 8% des PCR réalisées sur l'ADNc du gène xbp- 1(Cer125)	33
Figure 64: Zone d'intérêt du plasmide myo-2p::eGFP	33
Figure 65: Résultat Western-Blot sur la lignée myo-2p::eGFP ; pmec-2::Mec-2ab::FSS	34
Figure 66: Résultat Western-Blot sur la lignée xbp-1(Cer125)	35
Figure 67 : Modèle de localisation et de traitement de l'ARNm xbp-1 à la membrane du RE dans une cellule	91
Figure 68 : Les différentes interactions protéiques d'XBP- 1s	91
Figure 69: Profil des ARNms de xbp-1 avant l'insertion du FSS et après l'insertion du FSS, L'egfp dans l'exon 1 et la wrmScarlet dans l'exon	93

Liste des Tableaux

Tableau 1 : Programme sur Thermocycleur pour PCR	44
Tableau 2 : Concentrations obtenues du vecteur et de l'insert	54
Tableau 3 : Préparation du mélange de clonage Vecteur/ Insert	55
Tableau 4 : La préparation du mix d'injection	56
Tableau 5: Concentration des l'insert et du vecteur pour le clonage de pmec-2-c	62
Tableau 6 : Préparation du mélange Vecteur / Insert pour le clonage	62
Tableau 7: Tableau des amorces pour la mutation des sites GCAC	63
Tableau 8: Tableau des mutations réalisées sur les différents site GCAC du plasmide pmec2-a-b::MEC2a-bE9- E10GGS6::FSS	64
Tableau 10 : Tableau préparation du mix d'injection	70
Tableau 11: Résultats de l'observation visuelle sous des lumières fluorescentes, GFP et dsRed	74
Tableau 12: Tableaux des préparation des tubes pour dépôt sur gel bis-acrylamide	84
Tableau 13 : Préparation des échantillons pour le dépôt sur gel de bis-acrylamide	85

Liste des annexes

Annexe 1: Alignement des séquences après séquençage du plasmide construit pmec-2-a-b::MEC-2E9-E10- GGS6::FSS.	. 99
Annexe 2: Alignement des séquences après séquençage par les différents amorces les ADNc de l'ARNm de xbp-	
1(Cer125)	101
Annexe 3 : Protocole simplifié western-blot sur les vers C. elegans	102
Annexe 4: Protocole simplifié de la modification du génome par CRISP-Cas91	102

INTRODUCTION

I. Du matériel génétique à l'expression génétique

A. Le matériel génétique

La découverte de l'existence du matériel génétique vient de ces expériences réalisées depuis la fin du 19ème siècle :

En 1869 : Fredrich Micher utilisa une protéase (pepsine : découverte en 1836) pour séparer le noyau du cytoplasme. Il a observé une petite tache de poudre grise qui s'était détachée d'un liquide claire jaunâtre (cytoplasme). Il l'appela nucléine.

En 1928 : L'expérience de Griffith a montré que des souris ayant reçus des pneumocoques de souche S (smooth, encapsulés et virulents) tuée ne tue pas la souris, mais lorsqu'elle est mélangée à des pneumocoques avec la souche R (rough, non virulents) vivante, la souris injectée meurt. Comme conclusion il a établi que la souche R est transformée en souche S.

En 1944 : Avery O.T., Macleod C. M. et Mc Carty M. (Rockfeller institute - New York) ont repris les expériences effectuées en 1928 par Fred Griffith en utilisant les extraits de différentes biomolécules de la cellule S traités ou non par des enzymes spécifiques. Les résultats obtenus montrent qu'il y a transformation d'un nombre de souches R en souche S lorsque l'ADN purifié est utilisé (selon la fréquence de transformation). Les mêmes résultats ont été obtenus en dégradant les protéines et les ARN par des enzymes spécifiques. En conclusion, l'ADN est le facteur transformant et le support de l'information génétique.

En 1950 : Chargaff a fait l'équivalence en bases dans l'ADN des animaux, végétaux, bactéries et phages. En conclusion il a trouvé que :

- Le rapport A + G / T + C \approx 1 (purines / pyrimidines \approx 1) chez tous les organismes testés.
- Le rapport AT / GC est variable selon les espèces.

En 1952-1953 : Franklin a obtenu une excellente photographie d'ADN par diffraction des rayons X. La même année Watson et Crick utilisant les données de Chargaff et l'image obtenue par Franklin ont établi le modèle de la double hélice (la forme B).

Ces différents travaux ont démontré et confirmé que l'ADN est le support de l'information génétique.

On sait maintenant que l'ensemble de l'information génétique d'un individu s'appelle le génome, qui est contenu dans toutes les cellules sous la forme de chromosome et est supporté par l'ADN. Le génome humain est composé de 46 chromosomes arrangés en 23 paires, 22 paires d'autosomes et 1 paire de gonosomes (chromosomes sexuels).

Le chromosome est constitué d'une molécule d'ADN condensée autour de protéines, dont les principales sont appelées les histones. Les 46 chromosomes se trouvent dans le noyau pour les cellules eucaryotes et pour les cellules procaryotes qui ne possèdent qu'un chromosome circulaire, il se trouve dans le cytoplasme et est appelé le nucléoïde (*Figure 1*).

Un gène est un petit morceau d'ADN qui correspond à une information génétique spécifique qui est donnée par une succession dans un ordre précis des 4 nucléotides, Adénine (A), Guanine (G), Thymine (T), Cytosine (C). (*Figure 2*)

Les gènes codent pour la synthèse de protéines particulières qui exercent des rôles déterminés dans la cellule. L'ensemble des informations génétiques d'un individu (entre 20 000 et 30 000 gènes) est unique et est appelé le génotype.



Figure 1: Schéma d'une cellule eucaryote et procaryote montrant les différences principales

Chez tous les êtres vivants, l'ADN possède une structure identique. Elle est constituée de deux chaînes de nucléosides liés chacun par une liaison ester entre son carbone 3' (alcool secondaire) et le carbone 5' (alcool primaire) du nucléotide suivant, qui sont enroulées en hélice (*Figure 3*). Chaque nucléotide est constitué : d'un sucre, le désoxyribose, d'un acide phosphorique et d'une base azotée. Il existe 4 bases azotées : l'adénine (A), la thymine (T), la cytosine (C) et la guanine (G).

Ces deux chaînes de nucléotides sont unies entre elles par des liaisons hydrogènes pour former un hybride en forme de double hélice (modèle de J.D. Watson et F.H.C. Crick, 1953) dont les deux brins sont :

- ➔ Antiparallèles : l'un est constitué d'un enchaînement commençant à gauche et se poursuivant vers la droite, l'autre commençant à droite et se poursuivant vers la gauche ;
- → Complémentaires : chaque adénine (A) d'un des deux brins est liée par deux liaisons hydrogène avec une thymine (T) de l'autre brin, et chaque guanine (G) d'un brin est liée par trois liaisons hydrogène avec une cytosine (C) de l'autre brin.



Figure 2: Schéma de la cellule eucaryote à l'expression des gènes

En passant par le chromosome et finir par la structure en hélice de l'ADN. Le dépliement du chromosome permet de voir la double hélice d'ADN et dans le détail l'appariement des bases A, T, G et C.



Passarge, 2007

Figure 3: Le Nucléoside et l'ADN forme B

a) Le Nucléoside Le sucre se lie à la base azotée par une liaison impliquant un des azotes (l'azote n°1 des pyrimidines ou azote n°9 des purines) et le carbone 1' de l'ose (carbone réducteur ou fonction semiacétalique). C'est une liaison Bêta-N-Glycosidique. b) ADN forme B. Toutes les bases ont la même conformation par rapport à l'ose et sont à l'intérieur de l'hélice. Les groupements phosphates et les désoxyriboses sont à l'extérieur (Passarge, 2007).

Il existe six configurations à la double hélice d'ADN mais la forme B est la forme biologique la plus importante, c'est celle décrite par Watson et Crick en 1953.

B. L'expression génétique

B.1. La transcription

A partir de l'ADN on obtient un Acide ribonucléique (ARN) par le mécanisme de la transcription, c'est la première étape de l'expression génétique (*Figure 4*).

Il existe trois principaux types d'ARN, l'ARN messager (ARNm), l'ARN de transfert (ARNt) et l'ARN ribosomique (ARNr). Chaque type ARN a une fonction précise. L'ARNm est porteur de l'expression génétique et est décodé en acide-aminés (aa) pour donner des protéines. L'ARNt est porteur de l'anticodon nécessaire pour la traduction de l'ARNm et l'ARNr quant à lui est le constituant des ribosomes utile au moment de la traduction.

La transcription est le premier phénomène de régulation qu'utilise l'individu pour contrôler et faciliter l'expression génétique ainsi que tous les mécanismes de métabolismes cellulaires.

Des différences existent là encore entre cellules eucaryotes et cellules procaryotes. Pour initier la transcription de l'ADN en ARN, chez les procaryotes une seule enzyme appelée ARN polymérase est nécessaire et l'ARNm produite est dite polycistronique.

La transcription chez les eucaryotes se fait dans le noyau et l'ARN polymérase ne transcrit qu'un seul gène à la fois, l'ARNm est dite monocistronique. Lors de la transcription la thymine est changée en Uracile (U) dans le transcrit.

Chez les eucaryotes 3 ARN polymérase différentes existent,

L'ARN polymérase I pour la synthèse de l'ARNr (28S, 18S et 5.8S), il transcrit 90% des ARN totaux de la cellule (Drygin et *al*.2010),

L'ARN polymérase II pour la synthèse des ARN pré-messagers, ce sont les précurseurs de l'ARNm. Pour la synthèse également de la majorité des micro-ARN et la majorité des petits ARN nucléaires (ARNsn) (Tamura, 1996)

L'ARN polymérase III pour la synthèse de petits ARN : ARNt, l'ARN ribosomique 5S, l'ARNsn U6, des micro-ARN (Dieci et *al*.2007).

La transcription par l'ARN polymérase II permet la synthèse d'abord d'un pré-ARNm qui après maturation donne un ou plusieurs ARNm selon les modifications post-transcriptionnelles. Chez les eucaryotes, il y a la formation d'un complexe d'initiation composé de plusieurs facteurs qui se complexe à l'ARN polymérase II pour initier la transcription. L'ARN pol II est régulée par des éléments en *cis* et des facteurs *trans* qui vont venir aider l'ARN pol II à se fixer au promoteur, à s'engager dans l'élongation et à également se détacher de l'ADN.



Le mécanisme de transcription se fait en 3 étapes distinctes :

Figure 4 : schéma du mécanisme de transcription dans la cellule eucaryote

• L'initiation

L'ARN polymérase se fixe en amont du gène au niveau de différents promoteurs dont le plus connu la boite TATA (TATA ou une variante) est situé -25 nucléotides en amont du site de début de la transcription. D'autres promoteurs sont également connus la boite GC (GGGCGG) à -50, promoteur des gènes domestiques ou de ménage (« housekeeping gènes ») et la boite CAAT (CCAAT) à -80 du début du site de la transcription. Ces promoteurs font partie des séquences dites actives en *cis*.



Figure 5: Unité de transcription eucaryote et les régions cis se trouvant dans le site du promoteur.

La TATA box permet la fixation de la sous-unité TBP (TATA Box-binding Protein) du facteur de transcription TFIID (Transcription Factor of ARN polymérase II). La CAAT box permet la fixation du facteur de transcription CTF (Catbox Transcription Factor) et la séquence GC pas présente sur tous les gènes permet la fixation du facteur de transcription SP1



Figure 6: Formation du complexe de pré-initiation.

• L'élongation

Dès que la formation de l'hybride ADN/ARN est stable l'élongation peut progresser. L'ARN polymérase se déplace ensuite le long du brin d'ADN séparant au fur et à mesure les deux brins de la double hélice et catalysant l'addition de ribonucléotides à l'extrémité 3[']OH de la chaine d'ARN en formation d'après la séquence de la matrice d'ADN. Une fois le site d'initiation de la transcription libérée une nouvelle molécule d'ARN polymérase vient s'y fixer et amorcer la synthèse d'un nouveau transcrit. Plusieurs ARN polymérase transcrivent donc simultanément l'ADN d'un gène en se déplaçant vers l'aval l'une derrière l'autre réalisant ainsi la synthèse de plusieurs molécules d'ARN à partir d'un seul gène.

• La terminaison

La terminaison de la transcription chez les eucaryotes suit deux principales voies alternatives. La terminaison est provoquée par le recrutement de complexes de terminaison qui se fixent à la fois à des séquences spécifiques sur le transcrit naissant et à l'ARN pol II.

Elle se fait très en aval du début de la polymérisation de la molécule ARN. Il semble que l'arrêt de la transcription se fait au niveau de région en épingle à cheveux riche en GC suivie par des régions riche en AT. A la fin de la transcription la polymérisation la polymérase se détache de la matrice ADN et une molécule ARN est libérée. Cette molécule ARN est de grande taille et contient aussi bien les séquences correspondant aux exons que celles correspondant aux introns. Elle est appelée transcrit primaire.

B.2. Les modifications post-transcriptionnelle : La maturation des ARN prémessagers

La maturation des ARNm se fait au niveau du noyau de la cellule eucaryote et se fait lors la phase d'élongation de l'ARN pendant la transcription. Elle se déroule en plusieurs étapes :

- L'ajout de la coiffe qui consiste à l'ajout d'une 7-métylguanosine triphosphate, à l'extrémité 5' du pré-ARNm (Banerjee 1980 ; Shuman 2001). Elle joue un rôle protecteur contre la dégradation de l'ARN et permet le recrutement des ribosomes pour l'initiation de la traduction
- L'ajout de la queue polyA à l'extrémité 3'UTR. (Edmonds et al. 1971 ; Darnell, Wall and Tushinski, 1971). La fin de la transcription est signalée par la séquence AAUAAA. Quand l'ARNpol II rencontre cette séquence, un facteur spécifique pour la coupure et la polyadénylation est recruté (Cleave and polyadenylation specificity factor : CPSF) (Barabino et *al.* 1997). Ce CPSF va couper l'ARNm juste en aval de la séquence signal de polyadénylation et va recruter une polymérase polyadénine qui va ajouter environ 250 adénines monophosphate à l'extrémité 3'UTR.
- L'épissage du pré-ARNm, qui est une excision des introns et un assemblage des exons, produit un ARNm qui ne contient que la partie codante du gène (Gilbert 1978). Un autre type d'épissage existe, c'est l'épissage alternatif (EA) où non seulement les introns sont excisés mais également certains exons du pré-ARNm. Cela génère plusieurs ARNm possibles et par conséquent plusieurs protéines.



Figure 7: Structure de l'ARNm eucaryote mature prêt à être traduit. Coiffe ajoutée, la queue polyA ajoutée et les introns éliminés.

B.2.a) L'épissage constitutif

L'étape de l'épissage du pré-ARNm est primordiale pour les cellules eucaryotes. En 1978 est découvert que le gène est une structure discontinue composé de parties codantes, exons (expressed region) et de parties non codantes, introns (intragenic region) (Breathnach, Mandel et Chambon, 1977) (Berget et Sharp, 1977).

Il existe plusieurs classes d'intron, comme GA-AG qui commence par GA et se termine par AG. Également (AT-AC, les groupes I, II et II...) mais la plus fréquente est la classe GT-AG ou GU-AG dans le transcrit ARNm.



Figure 8: Schéma des différentes étapes de l'épissage

On retrouve l'intervention des snRNP U1, U2, U5, U4/U6 et la formation d'un produit en forme de lasso (résidu adénylate) qui sera dégradé par la suite.

La machinerie d'épissage appelée le splicéosome, reconnaît les exons et élimine les introns durant la transcription du pré-ARNm dans le noyau (*Figure 8*). Le splicéosome est un complexe ribonucléoprotéique (RNP) composé de petits ARN nucléaires riches en uridine appelés ARNsn (*small nuclear* ARN) U1, U2, U4, U5 et U6 et d'une centaine de facteurs protéiques, SR protéines, associés au complexe ou liés au pré-ARNm, ce sont les facteurs trans de l'épissage.

Les étapes d'épissage sont les suivantes, au site d'épissage 5' de l'ARNm il y a le recrutement de la sous-unité U1 puis d'U2AF. Le point de branchement, qui est une adénine (représenté par un A encerclé) permet quant à elle le recrutement de la sous-unité U2, cet assemblage protéique va plier la molécule d'ARNm au niveau de l'A. Après ces étapes le reste du splicéosome, le complexe U4/U6 et U5 seront recrutés sur le pré-ARNm. Le splicéosome est ainsi activé quand le complexe U4/U6 se sépare et U4 est relargué.

La première réaction de trans-estérification a lieu entre le 2'OH de l'adénine du point de branchement et le phosphate du nucléotide au niveau du site d'épissage 5'. Ceci permet la formation d'une boucle au niveau de l'intron, et la formation d'une extrémité 3'OH libre, à l'extrémité de l'exon 1, qui pourra alors initier une nouvelle réaction de trans-estérification entre ce 3'OH libre et le phosphate au niveau du site d'épissage en 3'. Cette réaction permet la ligation des deux exons, mais aussi l'élimination de l'intron sous forme d'un lasso qui pourra être ensuite dégradé dans le noyau. Les différents snRNP sont recyclés pour un nouveau splicéosome.

B.2.b) Le trans-épissage

Chez C. elegans, une autre forme d'épissage en trans coexiste avec l'épissage en cis. Le transépissage c'est un processus par lequel le 5' UTR (untranscribed region) des pré-ARNm est clivé et remplacé par une séquence SL (spliced leader) qui provient d'un petit ARN nucléaire de 22 nucléotides non codant. A l'extrémité 3'de ce SL il y a ensuite l'ajout d'une coiffe de triméthylguanosine (TMG). (Thomas *et al.* 1988 ; Liou et Blumenthal 1990).

Au moins 84% des gènes sont concernés (Tourasse et al. 2017).

SL1 et SL2 à l'extrémité 5' de l'ARNm. La séquence leader est censée aider à l'initiation de la traduction. Comme les séquences SL1 et SL2 sont connues, elles peuvent être utilisées expérimentalement pour identifier la séquence à l'extrémité 5' des ARNm.

Certains ARNm de *C. elegans* sont formés de transcrits multigéniques, le premier ARNm étant épissé en SL1 et les ARNm suivants en SL2. Les gènes qui codent pour ces transcrits sont très proches les uns des autres en tandem et sont transcrits sous le contrôle d'un seul gène. Ces transcrits sont similaires à ceux produits par les opérons bactériens et codent pour des produits génétiques qui sont coexprimés (Blumenthal 2005).

Ils diffèrent cependant en ce que les transcriptions traitées chez *C. elegans* génèrent de multiples ARNms. Des expériences antérieures ont indiqué que l'ADN n'est pas méthylé chez *C. elegans*, mais des études récentes à plus haute résolution ont suggéré qu'une certaine méthylation se produisait (Hu *et al.* 2015). Les aspects de la régulation des gènes, tels que la transcription, la traduction, le remodelage de la chromatine et les modifications post-transcriptionnelles (ubiquitination, phosphorylation, méthylation des histones, glycosylation) ont tous été étudiés à l'aide des outils génétiques de *C. elegans*.

B.3. La Régulation post-transcriptionnelle

Dans la cellule eucaryote, un pré-ARNm peut donner naissance à plusieurs molécules d'ARNm matures de ce fait à plusieurs isoformes de la protéine.

En 1977, en étudiant les ARN des adénovirus, Berget et Sharp, ont observé que l'ARN primaire pouvait être épissé de différentes manières, ce qui permettait à l'ARNm de coder pour différentes protéines virales. Ce résultat est dû au phénomène appelé épissage alternatif (EA).

C'est un processus qui introduit des variants d'ARNm à partir d'un même pré-ARNm (*figure 9*). Cela a pour conséquence une augmentation importante de la diversité protéique au sein de l'organisme, démultipliant la capacité codante pour un génome de taille relativement faible. L'EA peut modifier les propriétés des protéines en altérant ou en modulant le niveau d'expression du transcrit.

L'EA est un mécanisme très important qui touche un peu plus de70 % des gènes chez l'humain (Pan *et al.* 2008), il n'a pas pour seul but d'introduire de la diversité génétique. Il permet également la régulation de l'expression génique, un nombre important des ARNm ayant subi un EA ne sont jamais traduits en protéine mais sont dirigés vers la voie NMD (Non-sense RNA Mediated Decay) pour être dégradés (Nagy and Maquat, 1998) (Stamm *et al.* 2005).

L'EA joue un rôle dans la stabilité des ARN en produisant des isoformes avec un codon stop prématuré qui devient alors une cible pour la dégradation médiée par le NMD. La coordination entre l'EA et le NMD a été mise en évidence chez *C. elegans*. Chez les vertébrés, un tiers des évènements d'épissage alternatif aboutissent ainsi à la production d'isoformes qui sont des substrats pour le NMD



Figure 9: L'épissage alternatif.

A partir du pré-ARNm il y a un épissage alternatif qui donne 2 ARNm matures avec 1 exon de différent et 2 protéines de structures et poids moléculaires différents

Lorsque le transcrit est concerné par l'EA, différentes combinaisons d'exons sont possibles grâce à l'utilisation d'exons cassettes ou de sites donneurs ou accepteurs alternatifs parmi d'autres exemples aboutissant à des variants d'ARNm matures



Figure 10: Représentation schématique des différents modes d'épissage alternatif.

En bleu et violet les exons pouvant être épissés. En bleu les exons constitutifs, en violet les exons alternatifs, et en traits pleins les introns. Différentes combinaisons d'assemblage des exons, voire de rétention d'intron sont possibles. En outre, l'utilisation de promoteurs transcriptionnels peut dicter les choix d'épissage. Enfin, l'épissage alternatif peut conduire à l'utilisation de signaux de sites de polyadénylation distincts.

Des séquences *cis* régulatrices de l'ARNm (séquences courtes 6-12nt) sont reconnues par des facteurs d'épissage alternatif, dont les classes principales sont les protéines de la famille SR (riche en sérines et arginines) et les hnRNPs (*heterogeneous nuclear ribonucleoproteins*). Ces séquences peuvent être promotrices ou inhibitrices de l'épissage, on les trouve aussi bien dans les exons que dans les introns. (*Figure 11*).

On distingue donc quatre types de séquences cis selon leurs spécificités :

- Les ESE (Exonic Splicing Enhancer) sont des régions présentes dans la plupart voire la totalité des exons chez l'humain (Wang et Burge, 2008).
- Les ESS (Exonic Splicing Silencer) sont comme les ESE très répandus dans les séquences exoniques. Elles recrutent le plus souvent des protéines de la famille des hnRNP qui interagissent par le biais de leurs domaines RRM avec ces séquences régulatrices (Pozzoli et Sironi, 2005).
- Les ISE (Intronic Splicing Enhancer)
- Les ISS (Intronic Splicing Silencer) sont eux principalement régulateurs de l'épissage alternatif. Ils se retrouvent dans les introns.

Les protéines SR se lient plutôt sur les ESE et ISE pour stimuler l'épissage alors que les hnRNPs se lient plutôt sur les ESS et ISS afin d'inhiber la reconnaissance des exons par le splicéosome. Les domaines de liaison à l'ARN de ces protéines peuvent reconnaître leurs cibles en fonction de leurs structures comme les domaines de liaison à ARN simple brin ou de leurs séquences nucléotidiques comme les RRM (*Recognition RNA Motif*).



Figure 11: Les séquences cis et trans régulatrices de l'épissage

En plus de cette régulation post-transcriptionnelle, il y a des régulations qui ont lieu dans les régions non transcrites de l'ARNm, comme la région 3'UTR par des microARNs.

Elles sont mises en évidence la première fois dans les plantes (pétunia) dans les années 1990.

II. Le nématode Caenorhabditis elegans

Caenorhabditis elegans (C. elegans) a pour Règne : Metazoa, Embranchement : Nematoda, Classe : Secernentea, Ordre : Rhabditida, Famille : Rhabditiae (*Figure 12*)

Il est transparent et fait environ 1,2mm de longueur et 70 μ m de diamètre à l'âge adulte. Ce ver vit généralement dans le sol, où il peut vivre jusqu'à 3 semaines. C'est un organisme bactérivore, c'est-à-dire qu'il se nourrit essentiellement des bactéries qu'il trouve dans son environnement. (Byerly *et al*, 1976). Le développement de *C. elegans* peut se produire dans une large gamme de température (de 8°C à 27°C). Il se retrouve dans des pays chauds ou dans des pays ayant des températures froides, c'est une espèce ubiquiste (Félix et Braendle, 2010). *C. elegans* est le plus souvent retrouvé dans des habitats anthropogéniques (e.g. compost, sols de jardin) (Kiontke & Sudhaus, 2006).

L'hermaphrodite adulte possède un nombre fixe de cellules : 1090 cellules sont produites et 131 sont des cellules apoptotiques (Altun et Hall, 2009). Chez un adulte mâle on retrouve 1031 cellules (Lints et Hall, 2009), cela a permis d'établir une cartographie de tous les lignages cellulaires de la fertilisation de l'œuf au stade adulte. (Sulston and Horvitz 1977; Sulston, Horvitz et Kimble 1988). Il possède 302 neurones.

La taille moyenne des gènes est 5kb. Ces caractéristiques font du nématode un modèle très puissant pour l'étude de l'épissage alternatif et de sa régulation *in vivo*. (Zahler, 2018)



Figure 12: Anatomie de C. elegans

Principales caractéristiques anatomiques d'un hermaphrodite (A) et d'un mâle (B) vus latéralement. (A) Le cordon nerveux dorsal (DNC) et le cordon nerveux ventral (VNC) s'étendent sur toute la longueur de l'animal à partir de l'anneau nerveux. Deux des quatre quadrants des muscles de la paroi corporelle sont représentés. (B) Le système nerveux et les muscles sont omis dans cette vue, révélant plus clairement le pharynx et l'intestin. (C) Coupe transversale à travers la région antérieure de l'hermaphrodite

Les cribles génétiques sont facilités par l'existence et le maintien possible de son état hermaphrodite, par l'invariance de son lignage cellulaire. Les mutations qui donnent lieu à des défauts de développement et de comportement sont facilement identifiées dans les cribles génétiques. *C. elegans* est le premier organisme multicellulaire avec une séquence complète du génome. (*C. elegans* Sequencing Consortium, 1998)

Concernant la régulation post-transcriptionnelle, 25% des gènes sont soumis à l'épissage alternatif et 30% de ces événements sont régulés de manière différentielle selon les stades du développement (Ramani *et al.* 2011).

Les nématodes *C. elegans*, sont construits sur un plan simple, avec l'intestin et la paroi du corps étant des tubes concentriques. Le nématode typique possède une bouche (= ouverture buccale) menant à un stoma (= cavité buccale), relié à la lumière d'un pharynx (= œsophage) musculaire et/ou glandulaire qui mène à un intestin tubulaire, un rectum et un anus postérieur. Le corps est cylindrique ou en forme de fuseau et la queue au-delà de l'anus varie énormément en longueur.

L'ingestion de bactéries se fait via une contraction rythmique du pharynx. Il s'agit d'une pompe neuromusculaire, joignant la bouche à l'intestin, qui transporte la nourriture en deux étapes : pompage de la nourriture et vague de contraction du muscle postérieur de l'isthme (péristaltisme) générant une forte pression forçant les bactéries à entrer dans l'intestin et par conséquent les disloquant (Avery et Shtonda, 2003 ; Avery et You, 2012 ; Fang-Yen *et al.* 2009).

C. elegans a plusieurs avantages au laboratoire dont sa transparence. Il a un cycle de croissance court, d'environ 3 jours à 20°C et un grand nombre d'œufs pondus, environ 300 œufs, cela fait de lui un bon modèle pour l'observation rapide de la synthèse protéique. Sa cuticule transparente tout au long de sa vie permet de suivre au microscope l'expression de différents gènes fusionnés à des rapporteurs fluorescents (Brenner *et al.* 1974).



Figure 13: Cycle de vie du nématode C. elegans à t° 20°C.

La durée de chacun des stades est indiquée entre parenthèses. La fertilisation intervient à 0 min. (Altun et Hall, 2009)

L'embryogenèse de *C. elegans* dure environ 14 heures à 20°C. Après la fécondation, une coquille d'œuf pratiquement imperméable est fabriquée, permettant à l'embryon de se développer de manière totalement indépendante de la mère. Les embryons sont généralement conservés à l'intérieur de l'hermaphrodite jusqu'au stade de 24 cellules environ, moment auquel ils sont pondus. L'embryon hermaphrodite éclot avec 558 noyaux et devient une larve de premier stade (L1).

Les animaux commencent à manger et se développent à travers quatre stades larvaires (L1 à L4). Le stade L1 dure environ 12 heures, les stades L2 et L3 durent environ 8 heures et le stade L4 dure 10h (*Figure 13*).

Chaque stade se termine par une période d'inactivité ressemblant à un sommeil, appelée léthargie (Raizen *et al.* 2008), au cours de laquelle une nouvelle cuticule (couche externe de collagène) est formée. La léthargie se termine par la mue de l'ancienne cuticule. Environ 12 heures après la mue L4, les hermaphrodites adultes commencent à produire leur progéniture pendant une période de 2 à 3 jours jusqu'à ce qu'ils aient utilisé tout leur sperme autoproduit (Wightman, 2015).

Mais une progéniture supplémentaire peut être générée si l'hermaphrodite appauvri en sperme s'accouple avec un mâle. Après la période de reproduction, les hermaphrodites peuvent vivre encore plusieurs semaines avant de mourir de sénescence.

C. elegans possède un système chimiosensoriel très développé (Figure 14), la majorité de son système nerveux et près de 5% de son génome sont dédiés à la reconnaissance de facteurs chimiques environnementaux. (Bargmann, 2006). Cela lui permet de s'adapter aux changements de son environnement en modifiant son comportement, par exemple la régulation de son passage au stade Dauer (stade de résistance) ou dans la recherche de nourriture. Au stade Dauer, lorsque la source de nourriture manque les jeunes larves L2 bloquent leurs développements, jusqu'à ce qu'une nouvelle source de nourriture soit disponible. Au laboratoire, ils peuvent rester jusqu'à 1 mois à 20°C et jusqu'à 5 mois à 15°C dans ce stade Dauer.



Figure 14: Système nerveux de C. elegans.

Les caractéristiques principales comprennent un cerveau circumoral ou anneau nerveux, composé de processus axonaux et dendritiques, des groupes de corps cellulaires neuronaux céphaliques ('ganglions'), une corde nerveuse ventrale contenant des processus et des corps cellulaires de motoneurones, et une corde nerveuse dorsale et des commissures circonférentielles constituées exclusivement de processus. Le neuropil (comprenant l'anneau et les cordons nerveux) est représenté en rouge, les corps cellulaires neuronaux en jaune, les muscles en bleu, l'intestin en vert.

Au laboratoire *C. elegans* est élevé sur des boîtes de Pétri coulé avec un milieu NGM (Nématode Growth Medium) sur lequel on fait pousser des bactéries, *E. coli*, la variété OP50 (qui est une souche auxotrophe en uracile et qui ne forme pas de biofilm) (*Figure 15*).



Corsi et al, 2015

Figure 15: Observation de C. elegans.

(A) Les boîtes de Pétri posées sur la base d'un microscope à dissection. Le tapis bactérien d'E. coli est visible à la surface de la gélose à l'intérieur de la boîte de Pétri, mais les C. elegans sont trop petits pour être vus. (B) Deux adultes C. elegans vus à travers le microscope à dissection. Les pistes dans la plaque indiquent où les animaux ont voyagé sur le tapis bactérien.

La taille du génome est d'environ 97.10^6 bases, le nombre de gènes prédits est d'environ 20000 gènes. L'information génétique est supportée par 5 paires de chromosomes autosomes et 1 paire de chromosomes sexuels, XX pour l'hermaphrodite et XO pour le mâle (Araiz *et al.* 2008). Le mâle possède qu'un chromosome sexuel X et 0,1-0,2% de la progéniture sont des mâles en raison d'une rare non-disjonction méiotique du chromosome X.

Les chromosomes ne contiennent pas de centromères. Pendant la mitose ce sont des faisceaux de microtubule qui se fixent à plusieurs endroits le long du chromosome, ces fixations sont dites holocentriques ou polycentriques.

Il ne semble pas qu'une séquence spécifique soit nécessaire pour l'attachement puisque les transgènes contenant de l'ADN extra-chromosomique peuvent être hérités au cours de nombreuses divisions cellulaires.

III. Étude de l'épissage de deux gènes présents chez C.

elegans

A. Le gène de la méchanosensation mec-2

Les gènes *mec* (*mechanosensory abnormal*) sont impliqués dans la transduction du toucher (*Figure 16*). Ces gènes déclenchent une réponse à un stimulus mécanique délicat, délivré transversalement à l'axe antéro-postérieur du corps du nématode. Ainsi, les mutants *mec* ont une sensibilité au toucher altérée et ne répondent plus à ce type de stimuli.

Chez *C elegans*, le gène *mec-2* est l'un des nombreux gènes identifiés dans un dépistage de l'insensibilité au toucher et code pour une protéine membranaire intégrale.

Nous savons également que, la sensation du toucher nécessite des canaux ioniques de type DEG/ EnaC (Degenerin/Epithelial sodium channel) (Ben-Shahar, 2011). Ces canaux sont exprimés uniquement dans les six NRT (Neurones Récepteurs du Toucher) de *C. elegans* (*Figure17*). Toutes les protéines constituant ces canaux ne sont pas encore connues aujourd'hui. Nous savons que la famille des gènes *mec* code pour des composants de ces canaux, dont MEC-4, MEC-6 et MEC-10



Figure 16 : Complexe de mécano-transduction proposé dans les récepteurs tactiles corporels de C. elegans.

En son centre se trouve un canal ionique composé de MEC-4, 6 et 10, qui interagit avec la protéine intracellulaire MEC-2. MEC-7 et MEC-12 sont des protéines des microtubules nécessaires à la mécanosensation normale. MEC-1, MEC-5 est un collagène et MEC-9 sont des protéines extracellulaires dont les fonctions doivent encore être caractérisées.

Le gène *mec-2*, a des domaines homologues avec la protéine Stomatin-Like protein-3 (STOML3) humaine. C'est la région centrale de la protéine MEC-2 qui a une très grande similarité. La STOML3 est une protéine membranaire intégrale des globules rouges humains dont on pense qu'elle régule la conductance cationique. Elle est également exprimée dans les cils des neurones sensoriels olfactifs (Agostinelli *et al.* 2021). Elle est requise pour la fonction de nombreux mécanorécepteurs. Elle module les canaux de la mécano-transduction et les protéines des canaux ioniques sensibles aux acides

MEC-2 régule positivement l'activité du canal de transduction mécano sensorielle putatif qui est composé en partie de protéines codées par les deux gènes de la dégénérine mec-4 et mec-10. Elle présente une activité de liaison au cholestérol et est impliquée dans plusieurs processus. Elle se localise sur la membrane de protection des neurones.

MEC-2 fait la liaison entre le canal de la mécano- transduction et les 2 protéines du cytosquelette MEC-7 et MEC-12, (respectivement β -tubuline et α -tubuline), le déplacement de ces 2 protéines permet l'ouverture du canal (Huang *et al.* 1995).

En comparaison on retrouve chez les mammifères, les protéines ASIC qui sont des homologues des 2 protéines MEC-4 et MEC-10 et sont également impliquées dans le modèle d'attache de la mécano-transduction. Ces protéines ASIC interagissent avec les protéines associées aux canaux PICK1 et STOML3 et pourraient être régulées par des protéines de la matrice extracellulaire (par exemple, le collagène, la laminine et la fibronectine) ou des protéines du cytosquelette telles que l'actine et les microtubules (Tavernarakis et Driscoll, 1997).



Worm base

Figure 17: Représentation des NRT chez C. elegans.

Ils sont au nombre de six, on retrouve les ALMs (anterior lateral microtubule cell droit et gauche), l'AVM (anterior ventral microtubule cell), les PLMs (posterior lateral microtubule cells droite et gauche), et le PVM (posterior ventral microtubule cell)

Le gène *mec-2* est composé de 15 exons et est soumis à un épissage alternatif. En effet l'exon 9' possède un codon STOP. L'EA peut soit exclure l'exon 9' du transcrit mature, et ainsi terminer la transcription au niveau de l'exon 13 du gène *mec-2*, soit inclure l'exon 9' du transcrit mature, et donc le transcrit mature finira jusqu'à l'exon 9'.

Le facteur régulateur de cet EA est la protéine MEC-8 qui est une RNA-binding protein with multiple splicing (RBPMS), qui agit en *trans* sur le pré-ARNm de *mec-2*. Elle permet de réguler cet épissage en favorisant l'exclusion de l'exon 9' du transcrit mature. (*Figure18*). Mec-8 est exprimé de façon ubiquitaire dans les embryons et de façon extensive dans les larves.

MEC-8 est également le facteur *trans* d'épissage alternatif du pré-ARNm *unc-52*. (Calixto *et al.* 2010).



Figure 18 : Mode d'action supposé de la protéine MEC-8 sur pré-ARNm mec-2

La protéine MEC-8 joue un rôle d'activateur d'épissage pour excision de l'exon 9'. C'est un facteur trans qui reconnait les 4 sites GCAC qui sont des sites Exonic ou Intronic Splicing Enhancer.

Ce gène possède deux promoteurs alternatifs, un promoteur dit distal et un promoteur dit proximal. Par RT-PCR sur l'ARN de *C. elegans* il a pu être mis en évidence la présence de trois isoformes protéiques (*Figure 19*) :

- L'isoforme MEC-2A, la forme longue de la protéine. Elle est transcrite par le promoteur distal, *pmec-2ab* et s'arrête à l'exon 10. Elle contient un domaine N-terminal dont la fonction est l'oligomérisation de MEC-2 lors de la formation des canaux ioniques et une partie C-terminale dont la fonction est d'amplifier l'activité du canal ionique.
- L'isoforme MEC-2B, la forme courte de la protéine. Elle résulte aussi du promoteur distal *pmec-2ab* et s'arrête à l'exon 9', ne possède plus l'extrémité C-terminale.
- La dernière isoforme connue à ce jour est MEC-2C, elle est transcrite par le promoteur proximal *pmec-2c* et s'arrête à l'exon 9' qui est l'isoforme la plus courte qui ne contient plus ni le domaine C-terminal, ni le domaine N-terminal.



Figure 19: Schéma de l'épissage alternatif du gène mec-2.

Mec-2 possède 2 promoteurs alternatifs, un permettant l'initiation de la transcription à l'exon 1 (bleu foncé) et un autre permettant l'initiation de la transcription à l'exon 2' (orange). L'épissage de mec-2 consiste en l'inclusion ou l'exclusion de l'exon 9' dans le transcrit mature. Cet exon est porteur d'un site polyA qui permet la terminaison de la transcription. L'inclusion de l'exon 9' dans le transcrit mature lors de l'épissage alternatif peut induire la fin de la transcription et donc le raccourcissement du transcrit mature. Cet épissage alternatif et ce promoteur alternatif entraînent la formation de trois transcrits matures chez C. elegans, et donc trois protéines, MEC-2A (exon 1 à 13), MEC-2B (exon 1 à 9') et MEC-2C (exon 2' à 9'). Nous savons cependant que l'isoforme longue MEC-2A est présente dans les cellules où la protéine MEC-8 est également exprimée. Le facteur MEC-8 est majoritairement neuronal. Les autres formes sont exprimées en présence ou en absence de ce facteur indifféremment.

MEC-8 étant exprimé chez *C. elegans* dès le stade embryonnaire de 50 cellules, de manière ubiquitaire. Au cours du développement, son expression persiste principalement dans l'hypoderme, l'intestin, les six neurones récepteurs du toucher, les neurones chimiosensoriels et mécano sensoriels de la tête, la queue et la vulve.

Une étude récente montre que MEC-8 établit la longueur canonique des isoformes mec-2/ Stomatine dans les neurones tactiles, mais l'isoforme courte MEC2-B, non canonique prédomine dans d'autres neurones, y compris dans les neurones olfactifs, et que mec-2 est nécessaire pour l'olfaction (*Figure 20*).

L'expression spécifique de l'isoforme endogène forcée révèle que l'isoforme courte MEC2-B, fonctionne dans l'olfaction mais pas dans la mécano-sensation ; cependant l'isoforme longue MEC-2A est fonctionnelle dans les deux processus. (Liang *et al.* 2021).

Cette étude indique également, la présence d'une nouvelle isoforme extra-longue MEC2-E. Dans l'isoforme longue MEC2-A, un exon a été identifié, qui diversifie davantage *mec-2* en isoformes longues/extra-longues. Ni l'un ni l'autre n'est suffisant pour la mécanosensation. Les deux sont nécessaires simultanément, fonctionnant probablement comme des hétéromères pour médier la mécanosensation.



Figure 20: Les différents isoformes de la protéine MEC-2 sous contrôle du promoteur distal.

La forme courte MEC-B a une fonction dans les neurones olfactifs et la forme longue MEC2-A et extralongue MEC2-E agissent conjointement dans les neurones mécano-sensoriels après épissage par MEC8.

B. Le gène *x*-box binding protein -1 (*xbp*-1)

Dans la cellule eucaryote, le réticulum endoplasmique (RE) est impliqué dans la synthèse, les modifications, le repliement et la maturation des protéines sécrétées à l'extérieur de la cellule et des protéines et des lipides membranaires.

Au cours du processus de production des protéines, le RE permet le contrôle qualité des protéines, notamment grâce aux nombreuses protéines chaperonnes contenue dans la lumière du RE qui assurent le bon repliement des protéines naissantes, et leur transport vers leur destination finale. Les protéines incorrectement repliées sont orientées vers la voie de dégradation du protéasome. (Ellgaard et Helenius, 2003 ; Coe et Michalak, 2009).

En cas de perturbation de la maturation protéique, l'accumulation de protéines mal repliées dans la lumière du RE produit une situation stressante à la cellule appelée stress du RE.

Ces perturbations peuvent avoir différentes origines telles que les anomalies de statut redox, les troubles de l'homéostasie, du calcium, la carence en glucose, les infections virales, les anomalies de glycosylation des protéines induites par la tunicamycine. (Pallet *et al.* 2009)

La cellule en retour pour répondre à ce stress du RE déclenche la voie Unfolded Protein Response (UPR), dont l'objectif est de restaurer l'homéostasie protéique des cellules.

Chez *C. elegans*, la réponse UPR au stress du RE se manifeste par 3 voies de signalisation avec 3 transducteurs de l'UPR : IRE1 (Inositol-Requiring Enzyme 1), PERK (Protein kinase R-like ER kinase) et ATF6 (Activating Transcription Factor 6). Ce sont des protéines transmembranaires ayant des domaines luminaux et cytosoliques (*Figure 21*).

Dans les cellules non stressées, la protéine immunoglobuline se liant aux chaperons du RE (BiP, également connue sous le nom de GRP78) se lie au domaine luminal des protéines transducteurs, inhibant leur activité.



Figure 21: Schématisation des voies de signalisation de l'UPR.

Les voies de signalisation de l'UPR sont médiées par trois protéines transmembranaires résidant dans le RE : IRE1, PERK et ATF6.

- > IRE1 : possède des activités kinase et endoribonucléase (RNase).
- > PERK : possède un domaine sérine/thréonine kinase dans sa région cytoplasmique.
- ATF6 : subit une translocation dans l'appareil de Golgi et le clivage de son domaine Nterminal qui devient un facteur de transcription.

Le stress du RE induit la dissociation de BiP des 3 protéines IRE1, PERK et ATF6, ce qui active les transducteurs de l'UPR.

(A) <u>Voie IRE1</u> : IRE1 possède des domaines sérine/thréonine kinase et RNase dans la région cytoplasmique. En cas de stress du RE, IRE1 est activé par la formation d'oligomères et s'auto phosphoryle, ce qui entraîne une augmentation de l'activité RNase. L'IRE1 activé reconnaît la structure tige-boucle de l'ARNm d'*xbp1* et induit un épissage non conventionnel en clivant 23 nucléotides introniques. (*Figure 23*).

(B) <u>Voie PERK</u> : PERK possède un domaine sérine/thréonine kinase dans sa région cytoplasmique. Le stress du RE induit une PERK activation par oligomérisation et autophosphorylation dans le domaine kinase. La protéine PERK activée phosphoryle alors le résidu sérine 51 de eIF2 α (Eukaryotic translation Initiation Factor 2 α), ce qui entraîne une diminution du stress RE en atténuant la traduction des ARNms. En revanche, l'eIF2 α phosphorylé favorise sélectivement la traduction de l'ATF4 (Activating Transcription Factor 4), qui active la transcription de CHOP (C/EBP Homologous Protein) et de GADD34 (Growth Arrest and DNA Damage-inducible protein 34). Après la résolution du stress RE, GADD34 interagit avec PP1 pour induire la déphosphorylation d'eIF2 α , restaurant ainsi la traduction des ARNms en protéines. Cependant, si le stress du RE n'est pas résolu, CHOP induit l'apoptose.

(C) <u>Voie de l'ATF6</u> : ATF6 possède un domaine b-ZIP N-terminal dans la région cytoplasmique. Le stress du RE induit la translocation de l'ATF6 du RE vers l'appareil de Golgi, où l'ATF6 est clivé par 2 protéases S1P et S2P. Ce clivage protéolytique libère la région N-terminale de l'ATF6, appelée ATF6(N). L'ATF6(N) fonctionne comme un facteur de transcription actif et régule à la hausse les gènes cibles codant pour les protéines chaperons du RE, les composants de l'ERAD et de la protéine XBP1.

La voie qui nous concerne est celle de l'endonucléase IRE1.

Nous savons que la protéine XBP-1, quand elle est mutée entraîne une plus grande sensibilité aux processus inflammatoires qui peuvent contribuer à la maladie d'Alzheimer. Dans le côlon, les anomalies de XBP1 ont été liées à la maladie de Crohn.

On nomme la protéine XBP-1u pour unspliced qui est la protéine venant du préARNm non épissé et XBP-1s pour spliced qui est celle venant du préARNm épissé (*Figure 22*).



Byrd et Brewer, 2012

Figure 22 : Représentation des variantes de la protéine XBP1 (XBP1U et XBP1S)

Les chiffres indiquent les positions des acides aminés. Les domaines de base et de la fermeture éclair à leucine (ZIP) sont indiqués ainsi que l'ORF1 et l'ORF2 dans le domaine carboxyl-terminal de XBP1U et XBP1S, respectivement. Les domaines de la région hydrophobe (HR), de la pause traductionnelle (TP) et de la dégradation "sensible aux protéasomes" (PSD) de XBP1U sont également mis en évidence.



Figure 23 : Schématisation de la voie IRE-1 pour l'épissage de l'ARNm de xbp-1

La protéine XBP1s est un facteur de transcription avec un domaine bZIP qui se lie à deux motifs promoteurs agissant en cis, l'élément de réponse au stress ER (ERSE) et l'élément de réponse aux protéines non repliées (UPRE) et les active. Ce qui à son tour, induit l'expression de gènes cibles UPR en aval, nécessaires pour la réponse au stress du RE. (*Figure 24*)



Park et al. 2021

Figure 24 : Eléments d'action cis de XBP-1s

XBP1s est un facteur de transcription actif et joue un rôle important dans la régulation transcriptionnelle des gènes cibles de l'UPR codant pour les chaperons du RE, les enzymes de repliement et les composants de l'ERAD. XBP1s se lie à trois éléments en cis : ERSE, ERSE-II, et UPRE.

L'ARNm d'x*bp1* épissé est traduit par un décalage du cadre de lecture en un facteur de transcription actif, XBP1s. Ensuite, XBP1s régule à la hausse l'expression des gènes cibles de
l'UPR, les protéines chaperons du RE, les composants de l'ERAD (Endoplasmic Reticulum Asociated protein Degradation) et des enzymes de biosynthèse des lipides.

La protéine XBP-1u non épissée fait 34 kDa et XBP-1s épissée fait 55kDa. L'expression du gène xbp-1 se fait dans les neurales et certaines cellules.

Dans l'article de Reeces-Hoyes *et al.* 2007, ils ont construit un modèle avec le rapporteur transcriptionnel xbp-1::GFP et ont obtenu les observation suivante (*Figure 25*).



Reece-Hoyes et al. 2007

Figure 25 : Photographie de la lignée transgénique xbp-1::GFP

L'expression de *C. elegans* observée est uniquement après l'éclosion. La coloration est essentiellement nucléaire suggérant un changement nécessaire dans le modèle de gène. Le gène xbp-1 semble être exprimé dans tous les tissus, sauf dans la lignée germinale.

IV. Objectif du projet

Mon objectif est de caractériser l'expression protéique in vivo de 2 gènes *mec-2* et *xbp-1*. Pour cela nous allons utiliser un système de gènes rapporteurs formé de 2 fluorophores appelé le Frame Shift Sensor (FSS), qui nous permettra de visualiser in vivo l'expression des isoformes des protéines MEC-2 et XBP-1 au cours de l'épissage de ceux-ci.

Ensuite je vais analyser par ARNi l'identification des régulateurs de l'épissage non conventionnel.

Les protéines GFP (Green Fluorescent Protein) et dsRed utilisées sont naturellement fluorescentes, elles permettent un suivi sur cellules vivantes de leurs expressions. Leurs gènes peuvent être fusionnés in-vitro au gène d'une protéine qu'on souhaite étudier.

- La GFP fut découverte en 1962 de la méduse Aequorea victoria, la GFP est une petite protéine (238 acides aminés, 27 kDa) qui forme un tonneau béta contenant 11 brins, tous antiparallèles. La longueur d'onde d'excitation de la GFP est 488nm et d'émission 515nm
- La DsRed est une protéine tétramérique ayant la propriété d'émettre une fluorescence de couleur rouge. Issue d'un corail, Discosoma. La longueur d'onde d'excitation de dsRed 565nm et d'émission 590nm.



Figure 26 : Photo de la méduse Aequorea victoria et du corail Discosama

La longueur d'onde (LO) démission est supérieure à la LO d'excitation.

Dans les années 1990, Martin Chalfie a l'idée d'utiliser la GFP comme marqueur de gènes d'expression pour dresser la carte du *Caenorhabditis elegans*.

Les gènes rapporteurs ont été utilisés chez *C. elegans* pour étudier l'épissage alternatif par une équipe japonaise (Kuroyanagi *et al.* 2006 ; Kuroyanagi *et al.* 2007). Dans ce type d'analyse, les gènes rapporteurs sont inclus sous la forme d'exons parmi les différents exons du gène d'intérêt, selon le type d'épissage, l'un ou l'autre gène rapporteur est exprimé.

Cette équipe japonaise, a démontré l'efficacité d'un rapporteur *in vivo* de l'épissage alternatif du gène *egl-15*. Une cassette rapportrice avec l'exon 5A en phase avec la RFP et l'exon 5B en phase avec la GFP du gène *egl-15*, sous contrôle de promoteurs tissus-spécifiques ont permis de montrer l'inclusion différentielle des exons de ce gène en fonction des tissus.



Figure 27: Epissage du gène egl-15

Des systèmes rapporteurs sont utilisés pour étudier l'épissage alternatif du gène egl-15 à l'échelle cellulaire in vivo (Kuroyanagi et al. 2007). Lors de l'épissage de ce gène, les exons 5A ou 5B sont épissés de façon mutuellement exclusive, donnant lieu à deux formes de la protéine EGL-15 avec des fonctions différentes. Le mini gène rapporteur egl-15BGAR permet de visualiser la sélection de l'exon 5A ou 5B suivant la fluorescence. Lorsque l'exon 5A est inclus, c'est la protéine RFP qui s'exprime ; lorsque l'exon 5B est inclus, c'est la GFP qui s'exprime. Ainsi, les cellules contenant la forme 5A et la forme 5B d'EGL-15 ont pu être identifiées. Ici, les rectangles représentent les exons et les lignes droites les introns. Dans les images de fluorescence du ver, le vert représente la GFP, le magenta la RFP.

MATERIELS et METHODES

V. Caenorhabditis elegans

A. Maintenance du ver C. elegans

Les animaux sont maintenus sur des boîtes de Pétri sur le milieu de croissance des nématodes Nematode Growth Media (NGM) ensemencées avec des bactéries *E. coli* OP50 (auxotrophes pour l'uracile) comme source de nourriture.

Les lignées sont entretenues par repiquage sur les boîtes adaptées quand la source de nourriture est épuisée. Pour cela, selon les besoins, les individus sont soit repiqués individuellement à l'aide d'une anse de platine, soit déplacés sur un morceau d'agar (chunk) déposé sur la nouvelle boîte.

Le laboratoire de Denis DUPUY a développé la sélection de *C. elegans* sur milieu NGM additionné de G418. Les boîtes sont préparées comme décrit ci-dessus. L'antibiotique G418 (0,4g/L) est ajouté dans le milieu après l'autoclave, une fois le milieu suffisamment refroidi (65°C environ). Les boîtes sont ensemencées avec une souche d'OP50neoR, une souche possédant le plasmide de résistance au G418. Les bactéries OP50neoR sont cultivées en milieu LB additionné d'antibiotique : la kanamycine (50 µg/l) ou la neomycine (50 µg/l).

La sélection par l'antibiotique G418 à une concentration de 0,4g/l empêche le développement des animaux au-delà du stade L2. Les animaux déposés sur les boîtes après avoir dépassé le stade de développement L2 peuvent survivre et poursuivre leur croissance. Ainsi chaque lot de boîte est testé en déposant une dizaine d'individus non résistants au G418 sur des boîtes. Si aucun vers ne se développe le lot peut être utilisé.

Les techniques de maintenance classiques s'appliquent également sur les boîtes G418.

B. La décontamination et la synchronisation du ver C. elegans

Le traitement à l'hypochlorite de sodium peut être utilisé pour décontaminer ou synchroniser une population. La population est lavée dans du tampon M9 puis incubée dans une solution de soude (0,5 M) et d'hypochlorite de sodium (5 %). Le tout est agité sur un vortex à température ambiante. La solution permet de dégrader la cuticule des adultes tout en préservant les œufs insensibles au traitement. Bien homogénéiser la solution et ne pas laisser la solution agir plus de 6-7min.

Quand la moitié des animaux paraissent dégradés, la solution est centrifugée 3 min à 3000rpm afin de précipiter les œufs. Trois lavages avec du tampon M9 sont ensuite réalisés afin d'éliminer les restes de solution d'hypochlorite de sodium et de soude. Le culot est ensuite repris dans un tube de 15ml avec 5ml d'H2O distillée et 5ml d'une solution de sucrose 60%. Après centrifugation, 3-4ml du surnageant est récupéré et est complété par de l'H2O distillée avant d'être à nouveau centrifugé. Les œufs sont ensuite repris avec 1ml de solution M9 puis sont déposés sur une boîte fraîche NGM avec des OP50.

C. Transgenèse par micro-injection du ver C. elegans

Les vers transgéniques sont obtenus après une micro-injection d'un mix de plasmides dans l'une des 2 gonades du ver. Ce mélange d'injection pour un volume final de 20µl se composent de :

- 10ng/µl de plasmide myo2p::eGFP, servira de marqueur d'injection en exprimant l' eGFP dans le pharynx du ver.
- $60 \text{ ng/}\mu\text{l}$ de plasmide rapporteur.
- 10ng/µl de plasmide pRG02 exprimant l'antibiotique G418 (généticine).

Les plasmides injectés formeront un concatémère répétitif extra-chromosomique, celui-ci sera traité par le ver comme un chromosome et sera transmis à sa descendance. Les chromosomes de *C. elegans* sont dits holocentriques, ils n'ont pas de centromères, c'est pourquoi les concatémères répétitifs extra-chromosomiques sont considérés par l'animal comme un chromosome.

Les mélanges d'injection pour les insertions par CRISPR-Cas9 dans le génome se composent de :

- 30pmol de Cas9, enzyme qui génère la cassure double brin
- 90pmol de tracrRNA, le guide
- 90pmol de crRNA
- 25ng/µl de plasmide donneur (système rapporteur)

Avant toute utilisation les mélanges d'injection sont centrifugés pendant 10min à 14000g, pour culotter d'éventuelles impuretés. Ensuite 5µl du mix sont chargés dans l'aiguille de micro-injections Femtotip® de 0,1µm de diamètre qui est branché à un injecteur Femtojet Eppendorf®. La micro-injection est visualisée grâce à un microscope (Nikon Eclipse Ti).

Les vers sont immobilisés dans une goutte d'huile sur un pad d'injection (lame avec 1 goutte d'agarose 2%).

Les micro-injections sont réalisées dans les bras distaux d'une des 2 gonades de jeunes adultes de *C. elegans* de souche N2. Les vers injectés sont ensuite récoltés avec une solution de M9 dans une boîte fraîche de NGM agar.



D. Observations microscopiques

Les vers sont observés au quotidien, en direct sur leurs milieux de croissance dans les boîtes transparentes de NGM agar sous une loupe couplée à une lampe UV.

Afin de prendre des photos, les vers sont déposés dans une goutte d'azoture de sodium (AcrossOrganics) à 10mmol/l, qui est un inhibiteur musculaire, sur des lames recouvertes d'une fine couche d'agarose 2% (pads) puis recouverte d'une lamelle.

Les lames sont observées sous microscope à fluorescence (Axio Imager Z1, Zeiss) au grossissement 10X ou 40X par DIC (Differential Interference Contrast Microscopy), sous les canaux d'observation de la GFP (excitation à 395 nm) et de la DsRed (excitation à 558 nm). Les images sont ensuite traitées avec le logiciel Image J.

VI. Biologie moléculaire

A. PCR (Polymerase Chain Reaction) - classique

Le mix pour les amplifications par PCR contient 1 unité de la polymérase *Phusion* de chez NEB, une quantité de matrice de 1 à $20ng/\mu l$ d'ADN, $100\mu M$ de dNTPs, $0.5\mu M$ de chacune des amorces, $10\mu l$ de tampon Phusion HF (5X) et de l'eau ultrapure qsp 50 μ l de volume final.

Chaque amplification est réalisée en parallèle d'un contrôle négatif et d'un contrôle positif.

Nombre de cycles	Etapes	Température	Durée
1x	Dénaturation initiale	98°C	2min
	Dénaturation	98°C	30s
30x	Hybridation	Tm	30s
	Elongation	72°C	4min
1x	Elongation finale	10°C	∞

 Tableau 1 : Programme sur Thermocycleur pour PCR

B. PCR sur colonie

Les colonies bactériennes obtenues après sélection sur des boîtes LB agar + antibiotique sont vérifiées par PCR sur colonie directement dans le mix réactionnel. Durant le 1er cycle les vers seront lysés sous l'effet de la chaleur et l'ADN plasmidique sera libéré.

Pour cela chaque colonie est récupérée avec un cône stérile puis est plongée dans chaque mix PCR préparé, la colonie servant de matrice ADN. Le cône est ensuite utilisé pour une culture bactérienne, dans 5 ml de LB liquide + de 5 μ l d'antibiotique 1000X, dans les conditions de 37°C pendant 1 nuit.

Cette culture servira à la purification du plasmide de cette colonie si le résultat de la PCR est positif.

Le programme sur le thermocycleur est le même que pour les PCR classique sauf le temps de la 1^{ère} dénaturation qui est de 5min.

C. RT-PCR (Reverse Transcriptase PCR)

La rétro-transcription se fait sur 1µg d'ARN totaux pour l'obtention d'ADNc.

 4μ l d'ARN totaux à 250ng/µl auxquels on rajoute 2µl d'amorce random Hexamère à 50µM et 1µl de dNTP à 10mM, on incube 5min à 65°C, pour détruire les structures secondaires des ARN puis on met sur glace à 4°C. On rajoute 2µl de tampon Maxima 10x, 1 µl de la polymérase MAXIMA H, 9µl d'H₂O ultrapure et 1µl de RNAse Inhibitor. Le tout est incubé 30min à 50°C. La réaction est stoppée par une inactivation de l'enzyme en chauffant pendant 5min à 85°C.

Les ADNc sont stockés au congélateur à -20°C pour une utilisation prochaine sinon ils sont stockés à -80°C.

L'amplification est faite sur 1µl d'ADNc (mix RT) et les conditions pour la PCR avec les amorces spécifiques de la séquence d'ADNc à amplifier. La réaction d'amplification est réalisée dans un thermocycleur. Le produit de RT-PCR est conservé au réfrigérateur à 4°C.

D. RT-PCR direct sur C. elegans après mutagenèse

Dix vers du même âge sont récupérés dans un tube PCR de 200 μ l, auquel on ajoute 5 μ l de solution de lyse de vers (mélange de détergents et protéinase K) qui seront incubés 10min à 65°C puis inactivés 1min à 85°C.

L'échantillon est conservé sur la glace. On fait agir la DNaseI pour digérer les molécules d'ADN selon le protocole suivant : 5μ l de lysat en contact d' 1μ l de DNaseI, $0,7\mu$ l de tampon DNaseI et $0,3\mu$ l d'H₂O biomol. L'incubation dure 10min à 37°C et l'inactivation 10min à 75°C.

L'étape de rétrotranscription permettant d'obtenir les ADNc (ADN complémentaire) est réalisée selon la procédure suivante : 7μ l d'échantillon, 1μ l de dNTP, 1μ l de random hexamère et 1μ l d'H₂O ultrapure incubé 5min à 65°C (pour déduire les structures secondaires des ARN). Puis on rajoute 3μ l de tampon MAXIMA 5x, 1μ l de la polymérase MAXIMA H et 1μ l de RNAse Inhibitor. Le volume total est de 25 μ l et le programme sur le thermocycleur est de 10min à 25°C, 30min à 50°C et 5min à 85°C.

Sur les solutions d'ADNc obtenus on fait une PCR classique en prenant $1\mu l$ de solution comme matrice.

E. Électrophorèse en gel d'agarose

Pour cette étape, 500ng des réactions de PCR, RT-PCR, de plasmides digérés ou d'ARN totaux sont mélangés à du tampon de charge pour faciliter le chargement des échantillons dans les puits pour électrophorèse horizontale (purple loading Dye 6X NEB) et sont ensuite déposés dans les puits d'un gel d'agarose à 0,6%, 1% ou 1,8% contenant du BET (Bromure d'Ethidium 0,5µg/ml final). L'électrophorèse est réalisée à 120mV pendant 25 minutes dans une cuve de tampon TAE 0,5X. La révélation des gels se fait par à un trans-illuminateur UV.

Un marqueur de taille de 1Kb DNA ladder (NEB) ou de Low Molecular Weight DNA Ladder (NEB) est déposé à chaque migration en même temps, afin de vérifier la bonne taille des échantillons.

F. Purification de produit de PCR

Les PCR ayant migrées à la bonne taille sont purifiées à l'aide du kit PCR Purification (QIAGEN) en suivant les instructions données par le fournisseur. L'élution est réalisée dans 30μ L d'eau biomol. Les éluats d'ADN sont dosés par spectrophotométrie (NanoDrop-Thermoscientific). Une fois purifiés, les produits de PCR sont digérés par l'enzyme de restriction, la DpnI pendant 2h à 37°C sous agitation, afin de digérer la matrice d'ADN méthylé de départ ayant servi à initier la PCR.

G. Technique de clonage, AQUA cloning (Advanced Quick

Assembly cloning)

Pour la construction des plasmides par clonage, on utilise la technique de l'AQUA Cloning (*figure 28*).

L'AQUA Cloning, permet l'assemblage rapide et sans enzyme de plusieurs molécules d'ADN linéaires avec de courtes régions homologues flanquantes de 16 à 32bp en plasmides circulaires médiés par *E. coli*. (Beyer *et al.* 2015)

Pour un volume final de 6μ l, il faut mettre en contact les vecteurs et les inserts porteurs des régions homologues de 20 pb en suivant un ratio molaire de 3:1 (insert-vecteur). Puis ils sont laissés 1h à température ambiante. Enfin ces 6μ L de produit d'assemblage peuvent être utilisés pour transformer 50 μ L de bactéries thermo-compétentes *E. coli* DH5 α , par choc thermique.



Figure 28: Principe du clonage AQUA cloning.

a) Les parties d'ADN sont produites par PCR ou par digestion de restriction (ou les deux). Les oligonucléotides sont conçus pour apporter des régions homologues flanquantes à des fragments d'ADN adjacents d'une longueur optimale de 32 pb. Les parties d'ADN sont assemblées dans un plasmide circulaire par une directionnalité déterminée par la séquence. (b) Détails du clonage AQUA cloning (1) Les parties d'ADN sont générées par amplification PCR, ou issues d'une digestion enzymatique. (2) Ensuite, les parties d'ADN sont purifiées par électrophorèse sur gel et (3) mélangées et simplement incubées dans de l'eau avant transformation dans des E. Coli (Top10 pour un assemblage in vivo). (4) Enfin, les colonies obtenues sont confirmées pour un assemblage correct par des méthodes standard telles que la PCR analytique, la digestion de restriction ou le séquençage complet.

H. Transformation bactérienne par choc thermique

Le choc thermique est effectué en plaçant le mélange plasmides-bactéries à 4°C pendant 30min, puis à 42°C pendant 45s, puis 5min à 4°C et enfin une heure à 37°C après ajout de 250 μ L de milieu LB liquide.

Le mélange est étalé sur une boite de LB agar supplémenté avec l'antibiotique ampicilline à 100 μ g/ml. La croissance se fait à 37°C pendant la nuit.

I. Extraction et purification de plasmide (miniprep)

Les bactéries transformées présentent une résistance à l'antibiotique ce qui permet de les sélectionner après les avoir mises en culture une nuit à 37°C sur des boîtes LB agar contenant de l'antibiotique.

Une culture bactérienne de chaque colonie repiquée est faite dans 5 ml de LB liquide (supplémenté de 5 µl d'Ampicilline 1000X) pendant une nuit à 37°C et sous agitation à 180 rpm, dans le but de multiplier le nombre de bactéries puis d'extraire les plasmides.

L'extraction des plasmides est réalisée par lyse alcaline en utilisant le kit Plasmid Plus Purification (QIAGEN®).

Ensuite 5ml de culture bactérienne sont centrifugée à 6000g pendant 10min. Le surnageant est éliminé et le culot est resuspendu dans 250µl de tampon P1 (contenant de l'ARNase), puis 250µl de tampon P2 sont ajoutés, pour lyser les cellules. Enfin 350µl de tampon N3 sont rajoutés pour neutraliser la réaction de lyse. Le mélange est centrifugé pendant 10min à 16000g. Le surnageant est placé dans une colonne fournie dans le kit, puis est centrifugé 1min à 16000g. Un 1er lavage est réalisé avec 500µl de tampon PB et un 2nd avec 750µl de tampon PE. Une nouvelle centrifugation est faite afin d'éliminer tous résidus de tampon.

La colonne est ensuite placée sur un nouveau tube de 1,5ml pour une élution avec 30µl d'eau ultrapure.

Les plasmides sont dosés par spectrophotométrie (NanoDrop-Thermoscientific®) avant d'être stockés à 4°C.

J. Vérification des plasmides par digestion enzymatique

Pour un volume final de 50 μ l, une digestion enzymatique est réalisée avec 20 Unités d'enzymes PstI pour 300 à 500ng de plasmide, auxquels s'ajoutent 5 μ l de NEB Buffer 10X pendant 2h à 37°C. Les profils de digestion sont ensuite visualisés par électrophorèse (20min à 120V) en gel d'agarose de 0,6 à 1 %.

K. Séquençage

Les séquençages sont réalisés par l'entreprise Eurofins en utilisant des amorces permettant l'amplification de l'un des deux brins du plasmide à vérifier.

Le mix de séquençage est composé de 5µl de plasmide purifié ayant une concentration de 80 à $100ng/\mu l$ et de 5µl d'amorce à la concentration de 5µM.

L. Vérification de la synthèse protéique de la protéine XBP-

1::eGFP par Western-blot

Le western-blot est réalisé suivant le protocole de l'article « Western-blot analysis of *C. elegans* Proteins » (DE Jeong *et al.* 2018).

Le moyen d'extraction est un homogénéisateur. Le matériel utilisé est composé de deux boîtes NGM de *C. elegans* à la limite de la saturation avec des individus à plusieurs stades de développement, afin d'avoir une concentration détectable de protéine eGFP. Soit un total de 8000 vers triés ayant l'expression de la protéine MYO2::eGFP. Chaque boîte est traitée comme un échantillon de protéine total.

Les vers de la lignée *pmyo-2::pmec-2ab*::FSS servent de contrôle positif, car ils expriment la protéine MYO2::eGFP au niveau du pharynx. Ils ont été sélectionnés sur des boîtes NGM +

G418, puis elles ont été triées en fonction de leur fluorochrome eGFP par le cytomètre en flux Copas Profiler (union biometrica).

Le 1^{er} western-blot (WB) a permis la mise à l'échelle du niveau de protéine totale extrait du lysat de vers, nécessaire pour pouvoir détecter l'eGFP avec l'anticorps anti-GFP.

Le 2^{ème} WB est réalisé sur l'échantillon à analyser de vers transgéniques de la lignée xbp-1(Cer125).

Après 3 lavages des vers dans du tampon M9, le culot est récupéré puis est resuspendu dans 100µl de 1x PBS (Phosphate Buffer Saline) dans un tube de 1,5ml. L'ensemble est congelé à - 20°C jusqu'au lendemain.

Les tubes sont ensuite décongelés dans de la glace et un cocktail inhibiteur de protéases est rajouté. Chaque échantillon est broyé dans un homogénéisateur par 20 allers-retours. L'homogénéisât est récupéré dans un tube de 1,5ml et les tubes sont centrifugés à 13000g pendant 10min à 4°C.

Les différents surnageants sont récupérés et conservé sur la glace à 4°C.

On procède au dosage de la concentration en protéines totales par la technique de Bradford avec une gamme étalon faite avec la BSA.

On mélange 754 μ l d'eau ultrapure + 1 μ l d'extrait de protéines totales + 200 μ l de réactif de Bradford. Le tout est mis dans des cuvettes à usage unique de spectrophotométrie et la lecture est faite après 5min d'attente dans l'obscurité. La lecture se fait à la longueur d'onde de 595nm. Le tampon PBS sert de témoin pour faire le zéro.

Les différents échantillons sont mélangés avec du bleu Laemli 5x, avant d'être déposé sur un gel de bis-acrylamide 12 % pour une migration dans un tampon Tris-glycine en présence d'un marqueur de taille de protéines.

VII. Les milieux de culture et les solutions

A. Les milieux de culture

Milieu NGM (Nematode Growth Medium) agar :

3g de NaCl, 3,5g de peptone, 12g d'agar, 975ml H2O. Le milieu est autoclavé puis, lorsque la température atteint 37°C on rajoute 1ml de $CaCl_2$ (1M), 1ml de MgSO₄ (1M) et 25ml de tampon phosphate potassium (1M).

Le milieu NGM agar sert pour la culture des vers C. elegans en boite de pétri.

Le NGM est coulé sous PSM.

Les boîtes peuvent être ensemencées avec une culture fraîche de bactéries OP50neoR et seront utilisables après 24h. Les boîtes peuvent être stockées à 4°C plusieurs semaines avant ou après avoir été ensemencées avec les bactéries OP50neoR.

Milieu LB (Luria Broth) liquide et solide

10g de tryptone, 51 d'extrait de levure, 10g de NaCl et H₂O qsp 1L. Pour obtenir des boîtes de Pétri avec milieu LB agar on rajoute 12g d'agar pour 1L de milieu LB.

Milieu SOB (Super Optimal Broth)

20g de bactotryptone, 51 d'extrait de levure, 2ml de solution de NaCl à 5M, 2,5ml de solution de KCl à 1M, 10ml de solution de MgCl₂ à 1M, 10ml de solution de MgSO4 à 1M, H₂O qsp 1L. Ce milieu est utilisé pour l'obtention de bactéries chimio-compétentes.

B. Les solutions

Tampon M9 :

22mM KH₂PO₄, 22mM Na₂HPO₄, 85mM NaCl, 1mM MgSO₄, H2O qsp pour 1L. Ce tampon est utilisé pour toutes les manipulations avec les vers. Ce tampon s'utilise filtré.

Solutions de congélation_:

0,58g de NaCL, 0,68g de KH2PO4, 30g de glycérol, H2O qsp 50ml. Cette solution est pour la conservation des vers *C. elegans* au congélateur -80°C.

Solution de lyse pour ver :

0,25mM d'EDTA, 5mM de Tris pH8, 0,5% triton, 0,5% de Twen20 et PK (protéinase K) 2mg/ml.

Solution de synchronisation :

2,5ml de NaOH 2M, 2ml d'hypoclorite de soduim (NaClO) à 4%, H₂O qsp 5ml. Solution à préparer extemporanément.

Tampon Laemmli 5X :

Mélanger 5mL de Tris / HCl 0.5M pH 6.8, 2g de SDS, 4g de glycérol, 5mL de β -mercaptoéthanol, 1.25mL de Bleu de bromophénol à 1%, qsp 20mL d'eau Ultrapure.

RESULTATS

VIII. Étude de l'épissage alternatif du gène mec-2

A. Construction du plasmide pmec-2-a-b::MEC-2a-b E9-E10GGS6::fss

A.1. Préparation du plasmide

Pour l'étude du gène *mec-2*, un plasmide a été construit avec le promoteur *mec2-ab* et le système de gènes rapporteurs, *eGFP* et *mCherry* qui sera nommé FSS et qui sera en fusion avec le gène *mec-2a-b*. L'ARNm de l'*eGFP* est dans le même cadre de lecture lors de la transcription du gène *mec2*, forme A épissé (protéine MEC-2A) et l'ARNm de la mCherry est dans le même cadre de lecture lors de la transcription du gène *mec-2*, forme B non-épissé (protéine MEC-2B (*Figure 29*). De tel sorte que les vers lorsqu'ils exprimeront le gène *mec-2* non épissé ils présenteront une fluorescence rouge et lorsqu'il y a un épissage ils présenteront une fluorescence verte.



Figure 29: Schéma de la construction du plasmide pmec-2-a-b::MEC-2a-b E9-E10GGS6::FSS. Le fragment d'ADN de la mCherry est extraite pour être clonée dans le plasmide linéarisé pmec-2-ab::MEC-2E9-E10-GGS6::eGFP

Pour réaliser la construction de ce plasmide, 2 plasmides existants dans le laboratoire sont utilisés. Le plasmide *pmec-2-a-b*::MEC-2E9-E9'-GGS6::MCherry qui contient le brin d'ADN exonique 9 à 9' du gène *mec-2* et celui du gène de la *mCherry* et le plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP qui lui contient l'ADN exonique 9 à 10 et la séquence du gène de la *eGFP*. Dans les 2 plasmides les exons sont sous le contrôle du promoteur *mec-2ab*. L'objectif est de cloner le brin d'ADN contenant la mCherry dans le plasmide ouvert *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP.

Dans un 1^{er} temps, le plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP sera linéariser en faisant une amplification par PCR de la totalité de mon plasmide. Les amorces sens et anti-sens

s'hybrideront à la fin de l'exon 9' pour obtenir à la fin de la PCR un brin d'ADN de 8228pb, en utilisant les amorces suivantes (*Figure 30*) :

Sens (1159) : 5'gaacacttcacgaattcaaaattcattgaca 3'

Anti-sens (1158) : 5' aagaggaagacggtattccgcca 3'

Les amorces sont choisies avec le logiciel Snapgène.



Figure 30: Représentation de l'hybridation des amorces 1158 et 1159 pour la linéarisation du plasmide pmec-2-a-b::MEC-2E9-E10-GGS6::eGFP

Les amorces sens et anti-sens s'hybrideront de tel sorte que tout le plasmide soit amplifié et qu'après les 30 cycles de PCR nous obtenions une solution avec le plasmide ouvert.

Dans un 2nd temps, une extraction par PCR du fragment contenant le gène de la *mCherry* est faite à partir du vecteur *pmec-2-a-b*::MEC-2E9-E9'-GGS6::mCherry, avec des amorces ayant des séquences terminales d'homologies avec les séquences terminales du plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP linéarisé. En rouge dans les amorces, sont représentées les brin d'ADN d'homologie, de 20 nucléotides, au plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP linéarisé et en noir le fragment d'ADN d'hybridation pour l'extraction par PCR de la séquence de la *mCherry* (*Figure 31*) :

Sens (1146) : 5' cggaataccgtcttcctcttcaactttgtatacaaaagttgga 3' Anti-sens (1147) : 5' tttgaattcgtgaagtgttcctacttgtacagctcgtccatgc3'



Figure 31: Représentation de l'hybridation des amorces 1146 et 1147 pour l'extraction du fragment comportant la mCherry dans le plasmide pmec-2-a-b::MEC-2E9-E9'-GGS6::mCherry.

Chaque PCR est réalisée dans un volume final de 25µl. Pour l'extraction de ma *mCherry* dans le plasmide *pmec-2-a-b*::MEC-2E9-E9'-GGS6::MCherry la température d'hybridation programmée est 72°C et pour la linéarisation du plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP la température d'hybridation est de 61°C.

Les produits des PCRs sont vérifiés par électrophorèse sur gel d'agarose 1 % en présence d'un marquer de taille NEB 1kb plus. Pour cela 5µl de chaque PCR sont déposés après ajout de 1µl de gel loading (*Figure 32*). Les migrations des 2 produits de PCR indiquent la présence d'une bande à environ 822pb et d'une bande supérieure à 8228pb. La taille attendue pour la linéarisation de *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP est de 8228pb et la taille attendue pour l'amplification de la mCherry est 822pb, donc les 2 produits pour réaliser le clonage de la construction du nouveau plasmide sont bien présents.



Figure 32: : migration sur gel d'agarose des produits d'amplification de la mCherry et du plasmide pmec-2-ab::MEC-2E9-E10-GGS6::eGFP

Avant de réaliser le clonage, une digestion enzymatique par l'enzyme DpnI des 2 PCRs est faite pendant une nuit à 37°C, cela permettra de digérer les brins primaires méthylés qui ont permis de faire les PCRs. Puis les 2 produits de PCRs sont purifiés avec le kit Clean up PCR et dosés. Le volume final d'élution est de 20μ l.

Cmcherry: 108,9ng/µl et Cplasmide: 64,8ng/µl

Le clonage est ensuite réalisé selon la technique de l'AQUA-cloning (Beyer et *al.* 2015). Le plasmide linéarisé *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP sera le vecteur et l'ADN extraite contenant la *mCherry* sera l'insert. Les proportions en poids moléculaire, nécessaires pour la réalisation du clonage sont 1/3 de vecteur et 2/3 d'insert. Il est plus facile de convertis en mol/l la concentration de vecteur et de l'insert pour pouvoir faire le rapport de clonage (*Tableau 2*).

	Concentration massique	Concentration molaire	
Vecteur	64,4 ng/μl	12 nmol/μl	
Insert	108,9 ng/µl	239 nmol/μl	

Tableau 2 : Concentrations obtenues du vecteur et de l'insert

On prélève 5 μ l de vecteur soit 60nmol/ μ l.

	Proportion	Concentration	Volume de solution initiale	Volume H20 à ajuster	Volume final
Vecteur	1	60nM	5µl	ΟμΙ	5µl
Insert	3	180nM	3,76µl	1,24µl	5µl

C'est le volume nécessaire pour le clonage (Tableau 3).

Tableau 3 : Préparation du mélange de clonage Vecteur/ Insert

Après le clonage et la transformation dans les bactéries DH5 α , des colonies sont obtenues sur la gélose de LB agar contenant l'antibiotique ampicilline. Un échantillon de 6 colonies est choisi pour vérifier que la transformation a bien réussi.

Une culture dans 5ml de milieu LB liquide contenant de l'ampicilline est réalisée pour la sélection en vue de faire la purification du plasmide avec le kit de purification de miniprep du fabricant QIAGEN.

Les 6 différentes solutions de 15ml de plasmides purifiés sont faites et pour s'assurer que c'est le bon plasmide qui a été obtenu, une digestion enzymatique avec l'enzyme de restriction PstI est réalisée.

Le plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::FSS qui a été construit doit posséder 3 sites de coupure par l'enzyme de restriction PstI. Des sites de coupure PstI avaient été implantés de chaque côté du brin d'ADN du promoteur *pmec-2*, et un nouveau site a été rajouté, il est présent dans le brin d'ADN de la *mCherry* qui a été clonée.

Le profil de migration attendu pour chaque solution de plasmides digérés est de 3 bandes de taille respective 788pb, 2043pb et 6179pb.

Le résultat de la migration sur gel d'agarose indique que seules les solutions 1 et 6 présentent le profil de migration attendu (*Figure 33*). En effet on observe 3 fragments donc on a bien 3 sites de coupure à l'enzyme PstI.



Figure 33: Migration de pmec-2-a-b::MEC-2E9-E10-GGS6::FSS

Migration des digestions du plasmide cloné provenant des 6 colonies sur Gel d'agarose 0,6% après digestion par l'enzyme PstI. D'après la séquence du plasmide, nous devons obtenir 3 brins après la digestion.

Les 2 solutions de plasmides purifiées 1 et 6 ont donc été envoyés à séquencer afin de vérifier la bonne insertion du rapporteur mCherry et l'absence de mutations dans la séquence.

Les résultats du séquençage montrent que seul le plasmide de la solution 1 est correct. Le clonage de la mCherry dans le plasmide pmec-2-a-b::MEC-2E9-E10-GGS6::GFP est réussi. (*Figure 34*).

<i>mec-2</i> 1143	caaattgcagTGATAACAATTTTGCATTTCCAACGCACAAAGCTGGCGGAATACCGTCTT CAAATTGCAGTGATAACAATTTTGCATTTCCAACGCACAAAGCTGGCGGAATACCGTCTT ********************************	360 342
<i>mec-2</i>	CCTCTTGTGAGCAAGGGCGAGGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTC	420
1143	CCTCTTGTGAGCAAGGGCGAGGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTC	402

Figure 34 : Profil de l'alignement des séquences dans le logiciel « Clustal oméga » après séquençage En violet l'exon 9' et en rouge le début de la mCherry. Le résultat de séquençage correspond parfaitement à la séquence théorique attendue.

Ce plasmide sera ensuite injecté dans les vers N2 WT (wild type), afin d'obtenir une lignée de vers transgéniques avec une matrice extra chromosomique.

Le mix d'injection de plusieurs plasmides préparé se compose du plasmide pmec-2-a-b::MEC-2E9-E10-GGS6::FSS qui est notre plasmide d'intérêt, du plasmide myo2p::eGFP qui est un rapporteur transcriptionnel et sert de marqueur d'injection et du plasmide pRG02 qui permettra de faire une sélection, car ce plasmide donnera aux vers transgéniques une résistance à l'antibiotique Néomycine. Myo-2 signifie myosine-2, protéine intra cytoplasmique qui joue un rôle dans la contraction musculaire, présente dans les cellules du pharynx.

	C initiale (ng/µl)	C finale (ng/µl)	V à prélever (µl)	V final (µl)
<i>pmec-2-a-b</i> ::MEC- 2E9-E10-GGS6::FSS.	208,5	50	4,80	
myo2p::eGFP	196	10	1,02	20
pRG02	385	10	0,52	20
		Vol tot échantillon	6,34	
		Vol (H2O) à ajuster	13,66	

Tableau 4 : La préparation du mix d'injection

Après plusieurs essais d'injection de ce plasmide, je n'ai pas obtenu de lignée de vers. Les observations microscopiques sont sur la lignée de vers obtenues après l'injection des 2 plasmides de départ *pmec-2-a-b*::MEC-2E9-E9'-GGS6::Mcherry et *pmec-2-a-b*::MEC-2E9-E10-GGS6::eGFP

A.2. Observations microscopiques



Figure 35 : Schéma des protéines fluorescentes visualisées en l'absence de MEC-8



Figure 36 : Schéma des protéines fluorescentes produites en présence de MEC-8



Figure 37: Photos d'un ver L4 en microscopie à fluorescence grossissement x20.

Les images sont prises avec les filtres GFP et dsRed. La dernière image est une superposition des 2 ou 3 premières. En vert c'est l'expression de MEC-2B::eGFP et en rouge l'expression de MEC-2A::mCherry.



zone centrale

Figure 38: Photos avec les différents filtres, localisation des protéines MEC-2A et MEC-2B dans les neurones. Grossissement x40.

Les fluorescences vertes et rouge sont observées dans les 6 neurones récepteurs du toucher (*Figure 37*).

La coloration verte du pharynx correspond à l'expression de la protéine eGFP dans le pharynx du ver. Là où le promoteur *myo-2* est présent il y a l'expression de l'eGFP, c'est le cas dans le pharynx du ver. C'est la preuve que l'injection a fonctionné.

D'après la construction du plasmide pmec-2-a-b::MEC-2E9-E10-GGS6::FSS, la fluorescence rouge est l'expression de la protéine MEC-2B::mCherry, qui correspond à la forme courte de la protéine MEC-2. La fluorescence verte est l'expression de la protéine MEC-2A::eGFP, qui correspond à la forme longue de la protéine MEC-2 après épissage par MEC-8.

On constate que les 2 isoformes sont présentes en même temps dans certains neurones. Il y a une colocalisation dans certains neurones d'où la fluorescence jaune observée lors de la superposition des images (*Figure 38*).

B. Construction du plasmide *pmec-2-c::MEC-2E9-E10-*GGS6::FSS

B.1. Préparation du plasmide *pmec-2-c::MEC-2a-b E9-E10GGS6::FSS*

Le gène *mec-2* a un $2^{\text{ème}}$ promoteur proximal qui donne une $3^{\text{ème}}$ isoforme de la protéine, c'est MEC-2c.

Nous allons également étudier son épissage avec le système de rapporteurs fluorescents, GFP et mCherry.

Pour la construction du plasmide *pmec-2-c*::MEC-2E9-E10-GGS6::FSS, le promoteur *mec-2p-a-b* du nouveau plasmide construit, *pmec-2-a-b*::MEC-2E9-E10-GGS::FSS, est extrait pour être remplacer par le promoteur *mec-2p-c* qui est récupéré dans l'ADN génomique de *C. elegans*. Les deux extractions sont réalisées en faisant 2 PCRs.

Pour l'extraction et l'amplification du promoteur *mec-2p-c* les amorces 1128 et 1129sont utilisées (*Figure 39*). L'amplification est faite avec un Tm de 60° C pour 35 cycles.

Sens (1128) : ₅·gattacgccaagctctgcagatcttaacaaaccgcaaaaa ₃·

Anti-sens (1129) : ⁵'aaaacaatttaattctgcagtcagtccggaataagagttc ³'



Figure 39: Représentation de l'hybridation des amorces 1128 et 1129 pour l'extraction du promoteur mec-2p-c dans le plasmide pmec-2-c::MEC-2E9-E10-GGS6::eGFP.

Les amorces 1131 et 1132 sont utilisées pour ouvrir le plasmide précédemment construit, *pmec-2-a-b*::MEC-2E9-E10-GGS::FSS (*Figure 40*). L'amplification se fait à un Tm de 58°C.

Sens (1131) : 5'aattaaattgttttttttcagttgg 3'

Anti-sens (1132) : 5[°]ctgcagagcttggcgtaatcatggtc 3[°]



Figure 40: Représentation de l'hybridation des amorces 1131 et 1132 pour l'ouverture du plasmide pmec-2a-b::MEC-2E9-E10-GGS6::FSS ôtant le promoteur mec-2p-a-b

Les produits de chaque PCR ont été vérifiés. Les migrations sur gel d'agarose 0.6% des produits de PCR indiquent la présence de *mec-2p-c* à 2065pb et p-MEC-2E9-E10-GGS6::FSS à 6973pb (*Figure 41*).



Figure 41: Migration des produits de PCRs

Migration des produits de PCR de p-MEC-2E9-E10-GGS6::FSS ouvert et du promoteur mec-2p-c sur gel d'agarose 0.6% en préparation au clonage. Les tailles obtenues sont celles attendues, environ 6973pb et 2065pb.

A partir de ces 2 produits de PCR une digestion avec l'enzyme de restriction DpnI est faite afin de digérer la matrice primaire méthylée mise au départ. La digestion se fait à 37°C pendant une nuit.

Après, pour se débarrasser des résidus de tampon, d'enzyme et de sel qui risquent t'interférer à l'étape de clonage un clean-up est réalisé sur les produits de PCR. Le volume final d'élution après le clean-up de chaque PCR, est de 20μ l. On procède au dosage des produits avant clonage (*Tableau 5*).

		Concentration en ng/μl	Concentration en nM
Insert	pmec-2-c	145,5	108,57
Vecteur	<i>p-MEC-2E9-E10-GGS6::FSS</i>	157,7	32,57

Tableau 5: Concentration des l'insert et du vecteur pour le clonage de pmec-2-c

	Proportion	Concentration	Volume de solution initiale
Vecteur	1	97,71nM	3µl
Insert	3	375,71nM	3µl

Tableau 6 : Préparation du mélange Vecteur / Insert pour le clonage

L'AQUA-cloning est fait en mélangeant 3μ l de l'insert et 3μ l de vecteur. Je laisse 1h à température ambiante avant de faire une transformation dans des bactéries chimio-compétentes *DH5a*. Le produit de la transformation bactérienne a été ensemencé sur une boîte de LB agar supplémentée par l'antibiotique Ampicilline.

Des colonies isolées sont obtenues sur la boîte, après une nuit à l'étuve à 37°C. Puis 5 colonies sont prélevées pour une culture dans 5ml du milieu LB ampicilline afin d'extraire le plasmide cloné.

Les résultats de la digestion des plasmides des 5 colonies montrent que seuls les échantillons 1,2 et 3 sont digérés en 3 brins d'ADN (*figure 42*).



Figure 42 : Migration du plasmide pmec-2-c::MEC-2E9-E10-GGS6::fss digéré

Migration sur gel d'agarose 0,6 % des produits de digestions par l'enzyme PstI, de 5 plasmides purifiés venant de colonies repiquées sur la boîte LB agar après l'AQUA-cloning et la transformation dans des bactéries E. coli chimio compétentes.

B.2. Préparation des mutagenèses dirigées

Afin de vérifier le degré d'affinité pour la fixation de la protéine MEC-8 sur le gène de mec-2, des mutagenèses dirigées sont réalisées sur le plasmide *pmec2-a-b*::MEC2a-bE9-E10GGS6::FSS construit précédemment, en faisant varier un ou plusieurs nucléotides du site de fixation, GCAC, de la protéine MEC-8 sur le pré-ARNm *mec-2*.

Il existe 4 sites GCAC sur le plasmide pmec2-a-b::MEC-2a-bE9-E10GGS6::FSS (Figure 43).



Position des GCAC sur le plasmide :

- Site 1 : 2349-2352 nt
- Site 2 : 2400-2403 nt
- Site 3 : 3800-3803 nt
- Site 4 : 4609-4612 nt

Figure 43: Représentation graphique des 4 groupements GCAC sur le plasmide pmec-2-a-b::MEC-2a-bE9-E10GGS6::FSS.

Tout d'abord c'est l'adénine (A) qui est muté en une guanine (G) par mutagenèse dirigée par la technique de la PCR avec des amorces spécifiques de chaque site. Les 4 constructions obtenues ont chacune un site muté. Au départ on a la séquence GCAC et on arrive à la séquence GCGC.

Puis à partir de ces 4 constructions mutées, on réalise une seconde mutation sur l'un des 3 sites restants puis une troisième mutation.

Nom des amorces	Séquences	Tm
Mut 2346-2352 (Fow)	tcgacaaaaagcgcaatctattcctatcaaattg	59,3
Mut 2346-2352 (Rev)	ggaatagattgcgctttttgtcgaatactctgtg	60,6
Mut 2400-2403 (Fow)	catttccaactcatctggcggaatacctc	63,2
Mut 2400-2403 (Rev)	cgccagctttatgagttggaaatgcaaaattg	60,7
Mut 3800-2803 (Fow)	catgttgttctcgcgcaaagagtatcgtatgttgc	63,3
Mut 3800-3803 (Rev)	gatactctttgcgcgagaacaacatgtattaccc	61,6
Mut 4609-4612 (Fow)	cgcttcttaagcgcaaaaatacgaaaaaaaatcgt	60,4
Mut 4609-4612 (Rev)	cgtatttttgcgcttaagaagcgtaaaaagtag	58,5

Tableau 7: Tableau des amorces pour la mutation des sites GCAC

		Mutations appliquées			
		Site 1	Site 2	Site 3	Site 4
	Site 1		Site 1-2	Site 1-3	Site 1-4
Mutations	Site 2			Site 2-3	Site 2-4
primaires	Site 3				Site 3-4
	Site 4				
		Mutations appliquées			
		Site 1	Site 2	Site 3	Site 4
	Site 1-2			Site 1-2-3	Site 1-2-4
	Site 1-3				Site 1-3-4
Mutations secondaires	Site 1-4				
	Site 2-3				
	Site 2-4				
	Site 3-4		Site 2-3-4		

Tableau 8: Tableau des mutations réalisées sur les différents site GCAC du plasmide pmec2-a-b::MEC2a-bE9-E10GGS6::FSS

Après chaque PCR, une digestion du plasmide avec l'enzyme DpnI est faite puis un PCR cleanup avant de faire les transformations respectives dans des bactéries chimio compétentes, DH5 α pour pouvoir obtenir un plasmide purifié.

Après purification des clones obtenus pour chaque site de mutations, je les envoie à séquencer.

Les mutations ne sont pas encore injectées.

IX. Étude de l'épissage du gène xbp-1

Comme précisé précédemment, chez *C. elegans* le gène xbp-1 est épissé de manière nonconventionnelle d'un intron de 23 nucléotides dans le cytoplasme, dû à la réponse au stress du RE. La protéine fonctionnelle qui en résulte est nommé XBP-1s (s pour spliced). Lorsqu'elle est traduite sous sa forme non épissé elle se nomme XBP-1u (u pour unspliced)

A. Lignée pxbp-1::xbp-1::FSS « Matrice Extrachromosomale»

A.1. Préparation du plasmide

Pour l'étude de l'épissage non conventionnel du gène *xbp-1*, la construction du plasmide *pxbp-1*::xbp-1::FSS était déjà disponible dans le laboratoire. Le promoteur *xbp-1* et le gène ont été pris à partir de l'ADN génomique du ver *C. elegans* de la lignée N2. Ils ont été clonés dans le vecteur pBS 3'UTRunc-119, par la technique de l'AQUA-cloning.

Le système rapporteur utilisé cette fois est composé de l'eGFP et de la dsRed (*Figure 44*). La protéine non épissée, XBP-1u est en fusion avec la GFP et la protéine XBP-1s suite à l'EA est fusionnée à la DsRed. Il a été cloné à la suite du gène avec aucun décalage pour l'*egfp* et de +2nt pour la *dsRed* afin de respecter les cadres de lecture de chaque ARNm lors de la transcription du gène *xbp-1*.



Figure 44: Schéma de l'épissage cytoplasmique alternatif du gène xbp-1 avec les 2 fluorochromes eGFP et DsRed.

Deux isoformes sont possible, XBP-1u::eGFP et XPB-1s::dsRed à cause du changement de cadre de lecture. Les exons E3 et E4 dans XBP-1::eGFP ne sont pas les mêmes dans la protéine XBP-1::dsRed. Cette lignée a été préparée comme suivant, après le clonage du rapporteur fluorescent FSS et la transformation dans les bactéries $Dh5\alpha$, quelques colonies de bactéries obtenues sont mises en culture dans 5ml de milieu LB avec l'antibiotique ampicilline (1000x), afin de faire par la suite une miniprep.

Afin de vérifier la bonne insertion du gène dans le vecteur pBS, on réalise une digestion enzymatique du plasmide purifié (*Figure 45*).

La digestion du plasmide se fait avec l'enzyme BamHI. Suite au clonage le plasmide possède maintenant 2 sites de coupure, un sur le vecteur et l'autre au niveau de l'exon 4.

Il est obtenu après l'électrophorèse du produit de la digestion sur gel d'agarose 1 %, 2 bandes à 5225pb et à 1757pb.



Figure 45: Schéma du plasmide pxbp-1::xbp-1::fss. L'enzyme BamHI coupe le plasmide en 2 bandes de 5225pb et 1757pb.

A.2. Observations microscopiques

Au laboratoire, il a été observé une expression nucléaire des 2 isoformes du gène *xbp-1* dans les neurones et dans l'intestin du ver dans des conditions normales (*Figure 46*).



Figure 46: Observation microscopique d'un ver adulte transgénique issue de la lignée pxbp-1::xbp-1::FSS. Image du ver adulte photographié dans le filtre UV dsRed, gfp et champ clair.



Figure 47: Photographie d'un ver adulte de la lignée pxbp-1::xbp-1::FSS. Observation d'un embryon dans les gonades d'un ver hermaphrodite, avec l'expression embryonnaire nucléaire de XBP-1u::eGFP.

Dans la lignée obtenue, le changement de coloration qui indique l'épissage de l'intron non conventionnel intervient en fonction de l'âge du ver et non en fonction d'un état de stress du RE. Les individus au stade larvaire présentent une fluorescence majoritairement verte. Après l'éclosion les jeunes larves présentent dans quelques cellules du corps une coloration soit verte ou rouge. En grandissant à partir du stade L3 les vers font une transition de leur fluorescence, ils passent de vert à une fluorescence majoritairement rouge. Chez le ver adulte la fluorescence rouge est principalement dans les cellules neuronales et plus il vieillit plus la fluorescence rouge se localise dans les noyaux des cellules de l'intestin.

XBP-1S::dsRed



Figure 48 : vers obtenus après injection du plasmide pxbp-1::xbp-1::FSS

On obtient plusieurs profils de l'expression de la fluorescence du rapporteur. Aucune lignée n'a été obtenue. Les vers n'ont pas dépassé le stade L3.

B. Lignée pmyo3::xbp-1::fss « matrice extra chromosomale »

Pour prouver que le gène *xbp-1* s'exprime exactement où le promoteur pxbp-1 est présent et pour s'assurer que la visualisation obtenue après l'injection du plasmide *pxbp-1*::xbp-1::FSS est correct, on décide de changer de promoteur et de mettre le gène sous le contrôle du promoteur endogène *myo-3 (myosine-3)*. Le promoteur *myo-3* est présent dans les cellules musculaires du corps du nématode, on s'attend à ce que l'ensemble du muscle qui compose la paroi extérieure du ver exprime la fluorescence.

Cela permettra de quantifier l'expression de la fluorescence du rapporteur sachant que l'expression musculaire sera plus grande.

B.1. Design du plasmide

Le plasmide est construit en clonant le promoteur pmyo-3 dans le précédent plasmide à la place du promoteur pxbp-1. Le plasmide construit fait 8161pb.



Figure 49: Schéma du plasmide pmyo-3::xbp-1::FSS.

Le clonage du promoteur se fait comme précédemment par la technique de l'AQUA-cloning. Après purification, les 2 solutions de plasmides ont une concentration de 256ng/µl et 208ng/µl.

Afin de vérifier la bonne insertion du promoteur, on réalise une digestion enzymatique avec l'enzyme de restriction BAMHI. Cette enzyme digère le plasmide en 2 brins de 5225pb et 2936pb.

Nous observons sur le gel que le plasmide est effectivement digéré en 2 fragments d'ADN (*Figure 50*).



Figure 50: Profil de migration de la digestion du plasmide pmyo-3::xbp-1::FSS2 par l'enzyme BamH-1HF. On obtient 2 bandes de 5225pb et de 2936pb qui démontre que le plasmide est bien construit.

Le plasmide est préparé pour l'injection dans les vers N2, avec un marqueur de co-injection pRGO2.

	C initiale (ng/µl)	C finale (ng/µl)	V à prélever (µl)	V final (µl)
Pmyo-3::xbp-1::FSS.	256	50	3,90	
pRG02	385	10	0,52	20
		Vol tot échantillon	4,43	
		Vol (H2O) à ajuster	15,57	

Tableau 9 : Tableau préparation du mix d'injection

B.2. Observations microscopiques



Stade L1







71

Stade L4



Adulte





Figure 51: Photographie des vers de la lignée pmyo-3::xbp-1 ::FSS. Images prises au grossissement x20. C'est un montage fait à partir des images des 3 filtres.
Précédemment dans le laboratoire, des observations avaient été faites sur cette lignée.

Des membres de l'équipe ont observé que lors de l'expression de *xbp-1* dans cette souche, ils ont un comportement inattendu. En effet, les larves L1 présentent une fluorescence verte dans les cellules musculaires mais, à mesure que l'animal grandit, ils ont pu voir un passage à la fluorescence rouge sans qu'aucune fluorescence verte ne puisse plus être observée chez les adultes. Cela suggère que l'épissage non conventionnel *xbp-1* se produit sans induction chimique du stress RE et que dans cette lignée, XBP-1u::eGFP n'est exprimé prioritairement que chez les jeunes vers et XBP-1S::dsRed uniquement chez les adultes.

Par conséquent, ils ont essayé d'étudier quel phénomène est à l'œuvre ici et s'il est toujours régulé de la même manière que la copie endogène de *xbp-1*.

Tout d'abord, ils ont essayé de déterminer avec précision quand le passage de la fluorescence verte à la fluorescence rouge se produit dans le développement. Plusieurs adultes porteurs d'œufs fluorescents ont été placés sur une plaque NGM et ils les ont laissés pondre pendant 2h avant de les retirer, ce qui a permis de savoir précisément quel âge avaient les individus dans la boîte. Ensuite, nous avons observé les plaques toutes les deux heures, pendant la journée, pendant trois jours. Comme le début de la transition de la fluorescence verte à la fluorescence rouge s'est produit pendant la nuit, nous l'avons établi entre 28H et 40H de développement (respectivement la dernière observation le jour 1 et la première observation le jour 2).

Connaissant ce délai, ils ont repensé l'expérience afin de pouvoir suivre la transition avec plus de précision. Ils ont commencé par synchroniser une grande population de vers afin d'obtenir plus de résultats quantitatifs. Ils ont réalisé une préparation d'œufs en exposant des adultes gravides à une solution d'eau de Javel (tampon de lyse de ver). Ce processus tue les adultes mais les œufs sont protégés par la coquille. Après avoir récupéré les œufs, ils les ont laissés éclore pendant la nuit en culture liquide M9 sans aucune source de nourriture afin que tous les vers s'arrêtent au stade L1 (temps de développement : 10H). Ensuite, ils les remettent sur des plaques NGM avec *E. coli* OP50 comme source de nourriture et les laissent pousser pendant un certain temps avant de commencer les observations.

Ils ont étudié deux populations de vers et, après la synchronisation L1, les vers ont été divisés sur des plaques différentes. Ils ont choisi de procéder de cette façon afin de pouvoir également effectuer une analyse COPAS (Complex Object Parametric Analyzer and Sorter) car les vers ne peuvent pas être récupérés après avoir été analysés. Par conséquent, nous avons d'abord effectué les observations sous lumière fluorescente, puis analysé l'échantillon avec le COPAS. Les observations ont commencé lorsque les vers ont atteint 34H de développement.

Les résultats de l'observation visuelle sous des lumières fluorescentes sont inclus dans la (*Tableau 11*), mais les données de l'analyse COPAS n'étaient pas exploitables en raison d'un niveau de fluorescence inférieur au seuil de détection de l'instrument.



Tableau 10: Résultats de l'observation visuelle sous des lumières fluorescentes, GFP et dsRedA) Tableau des comptages faits à différentes et filtres sur le microscope à partir de vers synchronisés austade L1. B) Représentation graphique de l'état fluorescent d'une population de vers en fonction du temps.

Les deux expériences temporelles nous ont données les mêmes résultats, confirmant la validité des résultats présentés ici : la transition commence à 35H de développement et se termine à 60H. Lorsqu'ils ont comparé ces résultats au cycle de développement de *C. elegans* à 20°C, le début de la transition correspond à la transition entre les stades L2 et L3 (deuxième mue) et se termine après que les animaux deviennent complètement adultes, juste avant qu'ils ne commencent à pondre des œufs.

Voici la présentation des vers observés au stade L1, L3, L4 et adulte (Figure 52).



Figure 52: Photographies de vers de la lignée pmyo-3::xbp-1::fss, au stade de développement L1, L3, L4 et adulte.

Après cela, les animaux ont subi un choc thermique de 2h à 37°C pour savoir si le changement de fluorescence au cours du temps est impacté par le stress de RE. Les résultats ne montrent pas de différences par rapport aux résultats microscopiques dans des conditions normales. Pour aller plus loin, une vérification à l'échelle moléculaire est faite chez les vers de cette lignée.

B.3. Vérification de l'épissage

Pour cela, une RT-PCR est réalisée sur 10 vers transgéniques de chaque lignées N2, *pmyo-3::xbp-1::FSS* aux stades L1 et L2. La PCR est faite sur les ADNc de différentes lignées obtenues, avec des amorces proches du site épissage dans l'exon 2 (*figure 53*). On attend une amplification de 293pb pour un ADNc ne provenant pas d'un ARNm épissé. Dans le cas contraire l'ADNc fera 270pb.



Figure 53: Schéma du positionnement des 2 amorces choisies au plus près de l'intron non conventionnel pour étudier l'épissage

Les amorces utilisées sont les suivantes :

Sens (957) : $_{5'}$ aaaacgaacgtcttcgccgt $_{3'}$

Anti-sens (1300) : 5' atccatacgacggagttggttg 3'

	957 AAAACGAACGTCTTCGCCGT
	GCGCGATCTGGTGGAGGAGGAGCACCGCCGGCTCCGCGCTGAAAACGAACG
•••	CGCGCTAGACCACCTCCTCTTGGCGGCCGAGGCGCGACTTTTGCTTGC
	exon2
•••	CGATCCACCTCCATCAACAACAACATCAGCAACCAACTCCGTCGTATGGATTCCAAGAAGAACAACAACAACAAGTGTGGA
	exon 3
	GTTGGTTGAGGCAGCATACCTA 1300

Figure 54: Séquence du plasmide avec la représentation de l'hybridation des amorces 957 et 1300 servants pour l'amplification de la zone autour de l'intron alternatif.

Après la PCR, une migration sur un gel d'agarose est faite pour vérifier le résultat de l'amplification. La migration sur gel, nous montre que l'on obtient bien une bande à 293pb (*Figure 55*). Pour affiner se résultats, on réalise la migration des mêmes PCRs sur un gel de bis-

acrylamides 8%. La migration se fait dans un tampon TBS et la révélation est obtenue avec la visualisation sur un Geldoc après une incubation de 15 min du gel dans un bain de sybrgreen (un intercalant des acides nucléiques) (*Figure 55*). La migration sur un gel de bis-acrylamide est plus précise. Je pourrai vérifier ma différence entre des brins d'ADN de 293pb et 270pb.

Nous observons, une double bande pour les échantillons N2 (WT), *pmyo-3::xbp-1::FSS* (L1) et *pmyo-3::xbp-1::FSS* (L4). Cela laisse à penser qu'il y a la présence des 2 formes d'ARNm de XBP-1 : XBP1u et XBP1s et qu'il y a eu un épissage alternatif de 23pb du gène *xbp-1* dans toutes ces lignées.



Figure 55: Gel agarose 1.5% des PCR et Gel bis-acrylamide 8%

a) Gel agarose 1.5% des PCR d'ADNc des lignées N2 et pmyo3::XBP1::FSS aux stades L1 et L4. La RT-PCR est réalisée après la lyse de 10 vers de chaque lignés. La taille attendue est 293 pb. La PCR est faite avec un Tm de 55°C. Il y a la présence de la bande attendue à 293pb environ. b) Gel bis-acrylamide 8% sur les mêmes tubes de PCR que précédemment. Dans les puits N2, des vers Fluo rouge et Fluo vert, il y a 2 bandes qui laissent penser à la présence des 2 ARNm de XBP-1 de taille différentes : XBP1u et XBP1s.

On pourrait dire que la perte de cet intron alternatif non conventionnel de 23pb n'est pas seulement liée à un stress du RE parce que nous observons que dans des conditions normales de culture à 20°C, les vers N2 WT subissent une perte de cet intron également.

Cela pourrait expliquer le changement de fluorescence des vers en fonction de leur âge et pas seulement en fonction du stress du RE.

Afin d'éliminer le fait qu'une matrice extra-chromosomale peut conduire à la sur-expression du transgène dans le noyau cellulaire et affecter certains aspects de la régulation des gènes. Afin de nous débarrasser d'un éventuel artefact, nous avons décidé d'obtenir une lignée transgénique par une insertion génomique à copie unique.

C. Lignée intégrée xbp-1(Cer125)

C.1. Préparation pour l'intégration du système rapporteur dans génome

Cette lignée de vers transgénique est le résultat d'une collaboration avec le laboratoire du Dr Julian CERON en Espagne. Ils ont fait l'intégration du système rapporteur FSS en utilisant la technique de CRISPR-Cas9 par une méthode développée dans leur laboratoire. Cette technique permet l'insertion de longs fragments comme le système rapporteur (1500nt), en réalisant deux insertions séquentielles au niveau du gène d'intérêt.

La lère étape permet l'insertion d'un petit fragment de la egfp et dsRed, contenant des bras d'homologie de 36pb aux extrémités 5' et 3' avec le gène xbp-1. Cette insertion est facilitée parce que les bras d'homologie utilisés sont courts, cela permettra la réparation dirigée par homologie (HDR) au niveau du gène xbp-1. Après la 1ere insertion, une 2ième cassure est faite par la méthode CRISPR-Cas9 à la jonction egfp/dsRed et une nouvelle insertion plus longue du reste du système rapporteur est réalisée avec toujours des bras d'homologie aux extrémités 5' et 3'. Une nouvelle insertion par HDR est faite (*Figure 56*).



Figure 56: Représentation du processus insertion dans le génome du système rapporteur fluorescent selon la technique CRISPR-Cas9 dans le gène xbp-1.

L'insertion du système de 2 rapporteurs se fait en 2 étapes avec 2 cassures de l'ADN double brin par l'enzyme Cas9.

L'insertion est faite selon les étapes suivantes :

crRNA(1) et PAM(1) en gras est dans l'orientation anti-sens : crRNA de l'étape 1 : 5'-TCATCGACGGGGACATGGCT AGG

Pas de hors-cibles avec le NGG PAM, l'enzyme coupe 1 nt après le codon stop de l'isoforme non épissée.

Deux nucléotides (en orange) ont été retirés du gène *xbp-1* afin de préserver le cadre. Les séquences protospacer et PAM de l'étape 1 sont soulignées dans la séquence du gène.

La séquence de la zone du gène xbp-1 au départ ou va se faire l'insertion du transgène: 5'CTTCAGAAACGCGTGAAACACTTTCGCCGAATTC**CCT**AGCCATGTCCCCGATG AGCAGCTCGAGCACTGAC₃'

La séquence du brin d'ADN après 1ere insertion :

5[,]CTTCAGAAACGCGTGAAACACTTTCGCCGAATTCCCgtgagcaagggcgaggagctgttcaccgg<u>gg</u> <u>tggtgcccatcctggtcg</u>aggacgtcactccaccggaggaatggacgagctctacaagGCCATGTCCCCGTCGATGA GCAGCTCGAGCACTGAC₃,

En violet : bras droit d'homologie (36) En vert : fragment de GFP 5' (47 nt) En rouge : fragment de wrmScarlet 3' (43 nt) En bleu : bras droit d'homologie (36 nt)

En italique = crRNA étape 2, la séquence du protospacer est surlignée est la séquence du PAM est mis en gras.

crRNA de l'étape 2 : 5' ggtggtgcccatcctggtcg agg 3'

La séquence du brin d'AND après la 2ème insertion :

CTTCAGAAACGCGTGAAACACTTTCGCCGAATTCCCgtgagcaagggcgaggagctgttcaccggggtggt gcccatcctggtcgagctggacggcgacgtaaacggccacaagttcagcgtgtccggcgagggcgagggcgatgccacctacgg caagetgaccetgaagtteatetgeaceaeeggeaagetgeeegtgeeetggeeeaeeetggeeetggeeetggeeetggeetgaeetgaeetgaeetgeegtgea qtqcttcaqccqctaccccqaccacatqaaqcaqcacqacttcttcaaqtccqccatqcccqaaqqctacqtccaqqaqcqcacca tcttcttcaaggacgacgacgacaactacaagacccgcgccgaggtgaagttcgagggcgacacccctggtgaaccgcatcgagctgaa gggcatcgacttcaaggaggacggcaacatcctggggcacaagctggagtacaactacaacagccacaacgtctatatcatggcc gacaagcagaagaacggcatcaaggtgaacttcaagatccgccacaacatcgaggacggcagcgtgcagctcgccgaccacta ccagcagaacacccccatcggcgacggccccgtgctgccccgacaaccactacctgagcacccagtccgccctgagcaaaga ccccaacgagaagcgcgatcacatggtcctgctggagttcgtgaccgccgccggggatcactctccggcatggacgagctgtacaagt aattgtcagcaagggagaggcagttatcaaggagttcatgcgtttcaaggtccacatggagggatccatgaacggacacgagttcga gatcgagggagaggggaggggacgtccatacgagggaaccccaaaccgccaagctcaaggtcaccaagggaggaccactccc atteteetgggacateeteeceacaatteatgtacggateeegtgeetteaceaageaceeageegacateecagactaetaeaage aatcetteecagagggatteaagtgggagegtgteatgaacttegaggaeggaggageegteacegteacecaagaeaceteecte gaggacggaaccctcatctacaaggtcaagctccgtggaaccaacttcccaccagacggaccagtcatgcaaaagaagaccatg ggatgggaggcctccaccgagcgtctctacccagaggacggagtcctcaagggagacatcaagatggccctccgtctcaaggacg gaggacgttacctcgccgacttcaagaccacctacaaggccaagaagccagtccaaatgccaggagcctacaacgtcgaccgta agetegacateacetececacaacgaggactacaccgtcgtcgagcaatacgagcgttccgagggacgtcactecaccggaggaat ggacgagctctacaagGCCATGTCCCCGTCGATGAGCAGCTCGAGCACTGAC

Après la vérification que l'insertion du système rapporteur FSS s'est bien faite, les observations microscopiques du laboratoire Ceron sur l'expression des deux isoformes, n'a pu détecter aucun signal dans le canal GFP ou dsRed.

C.2. Observations microscopiques

L'équipe du Dr CERON a obtenu les photos suivantes :

Les vers transgéniques obtenus n'émettent pas de fluorescence sous excitation UV avec aucun des filtres, GFP ou dsRed avec un microscope à fluorescence.

Avec un microscope à épi fluorescence, nous observons une fluorescence verte avec le grossissement 10X, c'est de l'auto-fluorescence (*Figure 57*).



Figure 57: Montage de photos de l'abdomen du ver C. elegans de la lignée xbp-1 vue au microscope à fluorescence avec les filtres gfp et dsRed prise par le laboratoire du Dr Ceron.





Figure 58: Photographies d'un ver adulte de la lignée xbp-1 (Cer125) au microscope à épi fluorescence et confocale prise au laboratoire

A) Les photos prisent avec un microscope à épi fluorescence. Grossissement x10. B) Grossissement dans la zone du pharynx, vue à x63 au microscope confocale, dans les filtres dsRed et GFP. Nous observons seulement de l'auto-fluorescence.

Nous décidons de vérifier au niveau moléculaire si l'ARNm de *xbp-1u*::GFP et *xbp-1*::WrmScarlet sont bien transcrits.

C.3. Vérification de la transcription de *xbp-1::fss*

On réalise donc une RT-PCR afin de synthétiser de l'ADNc des différents ARNm transcrits et vérifier dans le transcriptome du ver s'il y a la présence de l'ADNc de *xbp-1* fusionné avec le FSS.

Pour obtenir les ADNc de la lignée, 10 vers adultes gravides sont récupérés et après avoir été lysés, une extraction d'ARN totaux est réalisée (*Figure 59a*).

Suite à une rétrotranscriptase avec l'enzyme Maxima, la solution d'ADNc est conservée à -20°C. A partir de ces ADNc sont réalisés une PCR en utilisant des amorces spécifiques pour l'amplification du gène de ménage F46F2.2, toujours transcrit afin de vérifier que la rétrotranscriptase a bien fonctionné (*Figure 59b*). Le gène F46F2.2 a migré à la bonne taille dans les 2 lignées xbp-1(Cer125) et N2 donc la RT-PCR a bien fonctionnée.



Figure 59: Profil de migration d'ARN totaux et de Rt-Pcr du gène F46F2.2

a) Profil de migration sur gel d'agarose 1 % de la purification d'ARN totaux de 10 vers C. elegans des lignées xbp-1(cer125) - N2. b) Profil de migration des Rt-Pcr du gène F46F2.2 (ubiquitaire du génome de C. elegans) sur ADNc des lignées xbp-1(cer125) - N2.

Pour amplifier spécifiquement l'ADNc du gène d'intérêt xbp-1(Cer125), différents couples d'amorces (*Figure 60*), pouvant s'hybrider à différents endroits de l'ADNc d'xbp-1(Cer125) sont utilisés pour prouver la présence ou pas de l'ADNc d'xbp-1(Cer125) avec la eGFP ou la wrmScarlet en fusion.

Numéro	Séquence (5'-3')
810	Fw : CACTAGATCCATCTAGAAATATGAGCAACTATCCAAAACG
957	Fw : AAAACGAACGTCTTCGCCGT
944	Rev : GTCAGTGCTCGAGCTGCTCATCGACGGGGACATTACTTGTACAGCTCGTC
946	Rev : GGGAAATTCACTAGGACACTACAATAACGATATTACAGGAACAGGTGGTG

Figure 60: Tableau des amorces de vérification de l'insertion du système rapporteur dans le génome



Figure 61: Représentation schématique de l'ADNc de xbp-1(Cer125) après la RT-PCR sur les ARN totaux. Représentation des PCRs de vérification de la transcription du système rapporteur FSS avec 4 couples d'amorces spécifiques.

	810	957
946	X1	Х3
944	Х2	X4

Tableau 11 : tableau récapitulatif des amorces pour les PCRs réalisées

Nous obtenons les bandes attendues pour les différentes PCRs réalisées. Pour X1 une bande à environ 2356pb, pour X2 une bande à 1690pb, pour X3 c'est très faible mais il y a une bande à 1990pb et pour X4 une bande à 1614pb. Cela sous-entend que l'ADNc du FSS est transcrit avec la *xbp-1* (*Figure 62*).



Figure 62: Migration sur gel agarose 1 % des PCR réalisées sur le gène xbp-1(Cer125). Les bandes obtenues montrent l'insertion du rapporteur.

Les bandes obtenues sont celles attendues pour les PCRs. Chaque PCR sont nommés X1, X2, X3 et X4 pour faire référence à xbp-1. Elles sont réalisées sur le même mix de départ et ensuite séparées dans 4 tubes avec 4 couples d'amorces spécifiques. Nous avons obtenu respectivement les bandes à environ 2356pb, 1690pb, 1990pb (bande très faible) et 1334pb.

Chaque produit de PCR X1, X2, X3 et X4 sont purifiées avec le kit de PCR clean-up.

C.4. Vérification de la séquence du gène xbp-1::FSS

Afin de vérifier si la séquence d'ARNm du gène *xbp-1* ne possède pas de mutation et de codon stop, j'envoie à séquencer les produits de PCR obtenus X1, X2 et X3 avec respectivement les amorces (810 et 946), (810 et 944) et (957 et 944), soit 6 tubes.

La séquence nucléotidique obtenue après séquençage est conforme à la séquence théorique attendue. (Les différents alignements sont en annexe)

C.5. Vérification de l'épissage

Pour aller plus loin, nous cherchons à vérifier si l'épissage à bien lieu dans cette lignée de vers transgéniques. Nous réalisons des PCRs avec les amorces qui entourent la zone de l'intron alternatif.

Nous constatons sur le gel non dénaturant de bis-acrylamide une double bande pour le ver N2(WT) et le ver de la lignée *xbp-1*(Cer125) (*Figure 63*). Nous constatons ainsi qu'il y a bien un épissage qui a lieu dans ces lignées même s'il n'y a pas de fluorescence verte ou rouge émise au microscope à fluorescence.





Figure 63: Migration sur gel agarose 1 % et gel de bis-acrylamide 8% des PCR réalisées sur l'ADNc du gène xbp-1(Cer125)

a) Migration sur gel agarose 1 % des PCR réalisées sur l'ADNc du gène xbp-1(Cer125). Les amorces utilisées sont spécifiques pour amplifier la zone de l'intron inconventionnel et vérifier l'épissage des 23 nucléotides. b) Gel de bis-acrylamide 8% des PCRs des vers de la lignée N2 négatif qui sert de témoin et de la lignée xbp-1(cer125).

Pour poursuivre l'investigation, on a décidé de vérifier si la protéine eGFP est traduite dans cette lignée de vers, par la technique du western-blot avec un anticorps anti-GFP.

C.6. Vérification de synthèse de la protéine XBP-1::eGFP par la technique du Western blot

Pour mettre au point la technique du western-blot sur cette lignée de vers, un test est d'abord fait sur la lignée myo-2p::eGFP ; pmec-2::mec-2ab::FSS, parce que cette lignée exprime l'eGFP dans le pharynx et dans certains neurones du toucher. La fluorescence est visible au microscope à contrairement à la lignée xbp-1(Cer125). Le plasmide myo-2p::eGFP est le marqueur d'injection souvent utilisé, il sert à l'expression du promoteur transcriptionnel du gène myo-2 par l'eGFP (*Figure 64*).

Par ce Western-Blot, l'objectif est de détecter un seuil de révélation de la protéine eGFP.

Taille estimée de l'eGFP =32kDa



Figure 64: Zone d'intérêt du plasmide myo-2p::eGFP

A partir des plaques de NGM saturées sont récoltées les vers de la lignée N2(WT) et de la lignée *myo-2p::eGFP*; *pmec-2::mec-2ab::FSS*. La lignée N2(WT) servira de témoin négatif.

Les vers fluorescents verts sont triés à l'aide du cytomètre de flux COPAS.

Chaque lignée de vers récoltés est placée dans le tampon PBS pour être lysée par broyage. Ils seront ensuite centrifugés à 14000g pendant 10min à +4°C. Les différents surnageants sont récoltés puis dosés par la technique de Bradford.

Les différentes concentrations de protéines totales de vers ont permis de faire une gamme étalon (*Tableau 12*).

Tous les échantillons sont soumis à une migration sur un gel SDS de bis-acrylamide à 15, sont transférés sur la membrane de nitrocellulose pendant une nuit à 30V et à +4°C. Après le transfert la membrane colorée au rouge ponceau pour s'assurer que le transfert à bien eu lieu. Puis est rincée au PBS-T et est saturée avec une solution à 5% de lait écrémée.

L'anticorps primaire, anti-GFP est déposé sur la membrane à une dilution de 1/10000 dans une solution de TBS-T-lait, pendant une nuit à $+4^{\circ}$ C. Puis après plusieurs rinçages l'anticorps secondaire est mis en contact avec la membrane. Cet anticorps secondaire anti-anticorps est spécifique de l'anticorps primaire.

La révélation se fait par électro-chimiluminescence (Figure 65).

Chaque tube est composé de la solution de protéines totales à laquelle on ajoute le tampon de charge Laemmli 5X et on ajuste avec un volume d'eau nécessaire. Les vers triés sont environ 7800 et les non triés proviennent de 2 boîtes NGM pleines de vers.

	myo-2p::eGFP ; pmec- 2::mec-2ab::FSS (triés)	myo-2	; eGFP: 2ab:: (non t	pmec-2: FSS r riés)	:mec-	Ladder
	(soit 7800 vers)					
Qté de protéine totale	19.5µg	50µg	30µg	50µg	100µg	
Vol d'échantillon	35 μl	20.5 μl	8 µl	14 µl	28 µl	4 μl
Vol d'H2O	0 μΙ	7.5 μl	20 µl	14 µl	0 μΙ	
Vol tampon Laemmli 5X	8.8µl	7 μl	7 µl	7 µl	7 μl	

p*myo-2*::eGFP ; p*mec-2*::mec-2-a-b::FSS



Figure 65: Résultat Western-Blot sur la lignée myo-2p::eGFP ; pmec-2::Mec-2ab::FSS On observe une détection de la eGFP à environ 32Kd. La quantité de protéines est observable dès 19.5 μg protéines totales lorsque les vers sont triés et dès 30μg pour les vers non triés. Sur le gel est détecté la présence d'eGFP à environ 32Kd, c'est la taille qui était attendue. Il y a une révélation par l'anti-GFP pour environ 19,5 μ g de protéines totales extraites des vers triés et dès 30 μ g de protéines totales pour les vers non triés.

Au vue de ces résultats on peut effectuer un Western-Blot sur notre lignée xbp-1(Cer125) en se servant de la lignée pmec-2::mec-2ab::FSS comme témoin positif. La stratégie sera d'avoir un très grand nombre de vers xbp-1(Cer125) pour espérer détecter la présence de la eGFP en fusion avec le gène xbp-1.

Des extractions de protéines totales de nos différentes lignées sont réalisées. Pour la lignée xbp-1::FSS(Cer125) 2 grandes boîtes NGM saturées de vers non synchronisés sont récoltées, afin d'avoir une très grande quantité de protéines totales. Pour le témoin (+) les vers sont triés, soit environ 4000 vers. L'échantillon pour le témoin (-) est une boîte NGM saturée de vers N2.

Après l'extraction et le dosage de la quantité de protéines totales, les échantillons sont préparés pour le dépôt sur gel (*Tableau 13*). Ils sont concentrés avant d'être mis dans les puits du gel. La migration se fait sur un gel SDS à 15% bis-acrylamide. Une coloration au rouge ponceau après le transfert sur la membrane a été faite et prouve le dépôt et le transfert des échantillons.

	<i>myo-2p</i> ::GFP ; pmec- 2::mec-2ab::FSS (triés) Témoin (+)	<i>xbp-1</i> ::FSS intégré au génome	N2 Témoin (-)	Ladder
Qté de protéine totale	28 µg	183.15 μg	182.4 μg	
Vol de Protéine	50 μl	55 μl	96 µl	4 μl
Vol tampon Laemmli 5X	12.55	13.8 µl	24 μl	

Tableau 13 : Préparation des échantillons pour le dépôt sur gel de bis-acrylamide.







Figure 66 : Résultat Western-Blot sur la lignée xbp-1(Cer125) et coloration rouge ponceau du transfert des proteines sur la membrane.

Les résultats du western-blot nous montrent une absence de bandes dans le puit de la lignée xbp-1(Cer125). S'il n'y a pas de révélation par l'anticorps secondaire on suppose qu'il n'y a pas la présence d'eGFP dans cette lignée. Le T (+) révèle bien la présence d'eGFP à la bonne taille et le T(-) est effectivement négatif. Une coloration au rouge ponceau après le transfert sur la membrane a été faite et prouve le dépôt et le transfert des échantillons.

DISCUSSION et PERSPECTIVES

L'objectif principal du stage était de visualiser in vivo l'expression des gènes *mec-2* et *xbp-1* grâce à un système de 2 rapporteurs fluorescents, frame shift sensor (FSS), chez le nématode *C. elegans*. Pour cela nous avons mis en place des lignées transgéniques de *C. elegans* permettant la visualisation de ce FSS et de distinguer les différentes formes d'épissage du gène *mec-2* soit MEC-2A, MEC-2B et MEC-2C et du gène *xbp-1* soit XBP-1U et XBP-1S.

Étude de l'épissage alternatif du gène mec-2

Nous avons construit un transgène contenant les exons 9, 9' et 10 sous le contrôle du promoteur *mec-2-ab* avec un système rapporteur de 2 gènes rapporteurs fluorescents, eGFP et mCherry. C'est le plasmide *pmec-2-a-b*::MEC-2E9-E10-GGS6::FSS. La visualisation des 2 isoformes issues du promoteur mec-2ab est la couleur verte pour MEC-2::A et la couleur rouge pour MEC-2::B. Une transfection de ce transgène est faite dans les vers N2. Nous avons constaté que les 2 isoformes de la protéine MEC-2 s'expriment dans les neurones de *C. elegans*.

Nous avons effectivement observé que MEC-2A::eGFP (la forme longue de la protéine MEC-2) s'exprime dans les neurones du toucher de *C. elegans* et que MEC-2B::MCherry (la forme courte de MEC-2) s'exprime également dans les neurones du toucher et ainsi que dans d'autres neurones.

De plus, l'auteur dans l'article de Calixto et *al.* 2010, a démontré que la protéine MEC-8 est nécessaire à l'élimination du neuvième intron de *mec-2* dans les neurones du toucher ce qui permet l'épissage du gène et l'expression de la forme longue. L'expression de *mec-8* après le stade embryonnaire persiste dans l'hypoderme, l'intestin, les six neurones récepteurs tactiles (TRN), les neurones chimiosensoriels de la tête et de la queue et dans les noyaux vulvaires.

D'après le récent article de Liang *et al.* 2022, c'est grâce à la protéine MEC-8 que l'isoforme MEC-2A, longue canonique dans les neurones du toucher est possible suite à l'épissage alternatif du préARNm. Ils démontrent par la suite que l'isoforme courte MEC-2B, non canonique prédomine dans d'autres neurones, y compris les neurones olfactifs, et MEC-2B est nécessaire à l'olfaction. Ils ont découvert à l'intérieur de l'isoforme longue un exon de cassette qui est une isoforme extra-longue, MEC-2::E. Ni l'un ni l'autre n'est suffisant pour la mécanosensation mais que les deux sont nécessaires simultanément en fonctionnant probablement comme des hétéromères pour médier la mécanosensation.

Nous avons également construit un transgène avec le promoteur distal *mec-2-c*, pour observer son profil d'expression. Les injections de ce transgène seront réalisées ultérieurement par la personne qui poursuivra le projet au sein de l'équipe.

Les différentes mutations du site de fixation de la protéine MEC-8, réalisées sur le transgène *pmec-2ab*::mec-2- E9-E10GGS6::FSS, n'ont pas encore été injectées.

Perspective

Pour poursuivre ce projet, il serait très intéressant de réaliser ces différentes injections pour ensuite rechercher par la technique des RNAi les différents régulateurs autres que MEC-8 de l'épissage alternatif de ce gène.

Étude de l'épissage du gène xbp-1

Dans la lignée *pxbp-1*::xbp-1:: fss obtenue, le changement de coloration du vert au rouge qui indique l'épissage de l'intron non conventionnel de 23nt du gène, intervient en fonction de l'âge du ver et non pas en fonction d'un état de stress du RE. La fluorescence observée varie en intensité et en localisation d'un ver à l'autre.

Les individus au stade larvaire présentent une fluorescence majoritairement verte. Après l'éclosion les jeunes larves présentent dans quelques cellules du corps une coloration soit verte ou rouge dans quelques cellules le long du corps. En grandissant à partir du stade L2, ils passent de la fluorescence verte à une fluorescence majoritairement rouge. Chez le ver adulte la fluorescence rouge est principalement dans les noyaux des cellules neurales situées au niveau du pharynx et plus il vieillit et plus la fluorescence rouge se localise dans les noyaux des cellules de l'intestin, on ne distingue plus de fluorescence verte.

En ce qui concerne la lignée *pxbp-1*::xbp-1::FSS avec la délétion des exons 3 et 4, nous observons une fluorescence majoritairement rouge et localisée dans les noyaux des cellules intestinales. En comparant la lignée des transgéniques avec la délétion avec la lignée sans délétion, il y a une perte de la fluorescence dans les noyaux des cellules neurales des transgéniques avec délétions. Ce qui nous amène à penser que les exons 3 et 4 jouent un rôle dans l'expression du facteur de transcription XBP-1, au sein des cellules neurales. L'expression des 2 isoformes sont similaires quant à leurs localisations, dans les noyaux des cellules du ver.

L'observation de la lignée *pxbp-1*::xbp-1::FSS avec cette fois une délétion de la partie Nterminale, c'est-à-dire des exons 1 et 2, montre une fluorescence verte exclusivement localisé dans le cytoplasme des cellules du pharynx ou dans les 1^{er} cellules intestinales de l'intestin, mais pas dans toutes les cellules intestinales. Nous observons l'expression que de la forme XBP-1U du gène xbp-1 qui demeure dans le cytoplasme et qui n'est pas épissée à cause de la délétion. Cela suggère que cette région des 2 premiers exons est importante pour une expression nucléaire et une expression dans les neurones. En outre, l'absence d'expression de la forme épissée XBP-1s indique que les 2 premiers exons sont également nécessaires pour qu'il y ait l'épissage non conventionnel de l'ARNm de *xbp-1*.

Au sein des cellules de mammifères, la protéine issue de l'ARNm non épissé contient un domaine d'exclusion du noyau (NES) dans sa partie C-terminale. Les protéines issues de l'ARNm épissé et non-épissé ont en commun leurs régions N-terminale composées entre autres d'une séquence d'adressage au noyau. De ce fait la protéine XBP-1u peut se localiser dans le noyau et dans le cytoplasme alors que la protéine épissée XBP-1s ne se retrouve que dans le noyau. La séquence d'adressage est effectivement dans la partie N-terminale et que l'expression XBP-1U::eGFP verte dans cette lignée se fait exclusivement dans le cytoplasme.

La dernière lignée transgénique générée est celle avec le rapporteur transcriptionnel, *pxbp-1*::eGFP. Il permettra d'observer la localisation de l'expression du promoteur *xbp-1*. Nous observons une fluorescence verte au niveau des cellules neurales de la tête de la queue et de la vulve des vers et au niveau des premières cellules intestinales. Ce qui est cohérent avec l'observation de la fluorescence verte de la lignée exprimant exclusivement XBP-1U.

Une précédente étude dans l'article de Shim et *al.* 2004, où ils ont cherché à déterminer l'expression et la localisation de la protéine XBP-1, les a conduits à générer un transgène, *xbp-1*(1-60)::GFP, qui contient le promoteur et le premier exon de XBP-1 (nucléotides 1-180) qui sont fusionnés à la GFP et dans le même cadre de lecture. Ils ont fait les transfections dans les

nématodes et ils ont constaté que XBP-1(1-60)::GFP est exprimé dans toutes les cellules des embryons et des larves adultes. Ils ont constaté que l'expression de XBP-1(1-60)::GFP était plus élevée dans les neurones que dans les cellules non neurales. Ces observations sont différentes des lignées générées au laboratoire. La lignée avec le promoteur transcriptionnel *pxbp-1*::eGFP ne montre pas dans le vers adulte la présence d'XBP-1 dans toutes les cellules.

Dans un autre article de Reeces-Hoyes et *al.* en 2007, les auteurs disent que le promoteur transcriptionnel indique la présence du transcrit *xbp-1* dans tous les tissus après l'éclosion.

La présence de la forme active épissée XBP-1s, en dehors des conditions de stress du RE trouve une partie d'explication dans l'article de Shim et *al.* Après analyse des rapporteurs traductionnels et à la suite de RT-PCR, a montré un niveau basal de la protéine XBP-1s, la forme épissée. Cela pourrait expliquer en partie le changement de couleur qui traduit la présence de la forme XBP-1s de la protéine. Nous l'avons constaté en faisant nous aussi une RT-PCR également sur des vers *C. elegans* N2, élevés dans des conditions normales, il y avait effectivement la présence d'ADNc de la forme XBP-1s. Cela montre que les 2 forment sont présentes même sans stress du RE.

Mais dans nos lignées nous avons en plus très clairement un changement de fluorescence en fonction de l'âge. Si nous prenons le cas de la lignée transgénique *pmyo-3*::xbp-1::fss, les embryons et les larves L1 expriment exclusivement l'isoforme XBP-1::eGFP et dès le stade L2 commencent à exprimer l'isoforme XBP-1::dsRed, jusqu'à l 'exprimer exclusivement à l'âge adulte. Selon l'article de Yoshida en 2001, lors du stress du RE la protéine XBP-1s peut ellemême activer son propre promoteur et activer la traduction de son gène. Cela pourrait expliquer pourquoi on a l'apparition d'une fluorescence rouge pour les vers adultes dans certaines lignées au fur et à mesure du vieillissement.

Dans l'article de Taylor et Dillin en 2013, ils ont cherché à analyser le rôle de la réponse UPR du réticulum endoplasmique dans le vieillissement de *C. elegans* et ils ont découvert que la perte de protéostasie du réticulum liée à l'âge pouvait être inversée par l'expression d'une forme constitutivement activé de XBP-1s. L'expression de *xbp-1s* dans les neurones initie une réponse cellulaire non autonome dans laquelle l'UPR est activé dans les tissus non neuronaux, protégeant ainsi contre le stress du réticulum et prolongeant la durée de vie de l'organisme entier. Cela suggère que l'expression de XBP-1s::dsRed dans nos lignées sont peut-être due à l'épissage alternatif du gène par l'activation de manière constitutive de XBP-1s.

Une autre hypothèse serait que les constructions au laboratoire aient perturbé le phénomène de pause dont parle Yanagitani *et al.* 2011, dans son article, par l'ajout du système des 2 gènes rapporteurs dans la région N-terminale du gène. Cela a pu entrainer une perturbation lors de la traduction du gène, d'où un épissage du gène accentué (*Figure 67*). En effet dans cet article d'Yanagitani *et al.* rapporte que la traduction de l'ARNm de xbp-1u qui ne se passe pas loin de la membrane du RE, subit une brève pause de la traduction pour stabiliser le complexe ARNm-ribosome-chaîne polypeptidique naissante. La pause ribosomale permettrait de favoriser l'épissage de l'ARNm de xbp-1u. C'est un module peptidique conservé à l'extrémité Cterminale qui est responsable de cette pause.



Byrd et Brewer 2012

Figure 67 : Modèle de localisation et de traitement de l'ARNm xbp-1 à la membrane du RE dans une cellule. Le domaine HR de localisation au RE, le domaine TP et les régions de ciblage des microARN (MTR) sont mis en évidence. L'intron de 26 nt épissé à partir de l'ARNm XBP1u par IRE1α est représenté en rouge. (c)

L'analyse des lignées transgéniques, a permis de comprendre que comparé à la littérature il y a une réelle complexité de la régulation de l'expression de *xbp-1*.

Des études récentes ont indiqué que la protéine XBP-1s interagit avec diverses autres protéines, et que ces mécanismes moléculaires sont importants pour réguler la stabilité de la protéine et l'activité transcriptionnelle de XBP1s (*Figure 68*).



Park, Kang et So, 2021

Figure 68 : Les différentes interactions protéiques d'XBP- 1s.

XBP-1s réagit avec de nombreuses autres protéines, ce qui permet la régulation de la stabilité de la protéine et de l'activité transcriptionnelle de XBP-1s.

Perspectives :

Pour poursuivre l'étude de la lignée xbp-1(Cer125), on pourrait faire une purification des polysomes afin de tester la présence de l'ARNm de la GFP et de la wrmSc qui est fusionné à xbp-1, pendant la traduction.

On aura la certitude que le l'ARNm de *xbp-1*::FSS est traduit et que le fait de ne pas voir de fluorescence a une autre explication que la non traduction.





C'est un ensemble de ribosome fixé en même temps sur un ARNm.

Pour poursuivre l'étude du gène xbp-1, on pourrait envisager, de changer de construction du système rapporteur toujours par une insertion dans le génome, pour créer une lignée où l'eGFP et la wrmScarlett ne seraient pas insérées au même endroit.

On ferait l'insertion de la séquence de l'egfp au niveau de l'exon 1 puis dans un deuxième temps faire l'insertion de la séquence de la wrmScarlet dans l'exon 3 après l'intron non conventionnel.

Nous obtiendrons une lignée de vers qui exprimeront l'eGFP dans des conditions normales de culture et en dans des conditions de stress du RE, il y aura épissage et ils exprimeront la wrmScarlet.

Pour faire ces insertions dans le génome du ver, nous pourrions utiliser la méthode tirée de l'article de Ghanta et *al.* (2021). Dans cette article l'auteur démontre que lorsque le L'ADN double brin donneur est refusionné, l'efficacité de l'insertion génomique est meilleure, surtout pour les longs fragments >1kb.

Dans notre cas, on pourrait appliquer cette méthode même si nous devons insérer moins d'un kb mais en augmentant l'efficacité de la reconnaissance par homologie du donneur double brin d'ADN.



Figure 70: Profil des ARNms de xbp-1 avant l'insertion du FSS et après l'insertion du FSS, L'egfp dans l'exon 1 et la wrmScarlet dans l'exon

Références

- Agostinelli, Emilio et al. 2021. « A role for STOML3 in olfactory sensory transduction ». *Eneuro* 8(2).
- ALTUN, ZF, et DH HALL. 2009. « Hermaphrodite life cycle ».
- Araiz, C., M.-T. Château, S. Descamps, et S. Galas. 2008. « Génomique quantitative chez *Caenorhabditis elegans* : stratégies pour l'identification de nouvelles cibles thérapeutiques et de nouveaux mécanismes moléculaires chez l'homme ». *IRBM* 29(5): 289-96.
- Avery, Leon, et Boris B. Shtonda. 2003. «Food transport in the *C. elegans* pharynx ». *Journal of Experimental Biology* 206(14): 2441-57.
- Avery, Oswald T., Colin M. MacLeod, et Maclyn McCarty. 1944. « Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III ». *Die Entdeckung der Doppelhelix*: 97.
- Banerjee, Amiya K. 1980. « 5'-terminal cap structure in eucaryotic messenger ribonucleic acids ». *Microbiological reviews* 44(2): 175-205.
- Barabino, S. M. et al. 1997. « The 30-kD subunit of mammalian cleavage and polyadenylation specificity factor and its yeast homolog are RNA-binding zinc finger proteins. » *Genes & development* 11(13): 1703-16.
- Bargmann, Cornelia I. 2006. « Chemosensation in *C. elegans* ». WormBook: The online review of *C. elegans biology* [Internet].
- Ben-Shahar, Yehuda. 2011. « Sensory functions for degenerin/epithelial sodium channels (DEG/ENaC) ». Advances in genetics 76: 1-26.
- Berget, Susan M., Claire Moore, et Phillip A. Sharp. 1977. « Spliced segments at the 5' terminus of adenovirus 2 late mRNA ». *Proceedings of the National Academy of Sciences* 74(8): 3171-75.
- Beyer, Hannes M. et al. 2015. « AQUA cloning: a versatile and simple enzyme-free cloning approach ». *PloS one* 10(9): e0137652.
- Breathnach, Rory, Jean-Louis Mandel, et Pierre Chambon. 1977. « Ovalbumin gene is split in chicken DNA ». *Nature* 270(5635): 314-19.
- Brenner, Sydney. 1974. « The genetics of Caenorhabditis elegans ». Genetics 77(1): 71-94.
- Byerly, L., R. C. Cassada, et R. L. Russell. 1976. « The life cycle of the nematode *Caenorhabditis elegans* : I. Wild-type growth and reproduction ». *Developmental biology* 51(1): 23-33.
- Calixto, Andrea, Charles Ma, et Martin Chalfie. 2010. « Conditional gene expression and RNAi using MEC-8–dependent splicing in C. elegans ». *Nature methods* 7(5): 407-11.
- Chalfie, Martin et al. 1994. «Green fluorescent protein as a marker for gene expression ». *Science* 263(5148): 802-5.
- Coe, Helen, et Marek Michalak. 2009. « Calcium binding chaperones of the endoplasmic reticulum ». *Gen Physiol Biophys* 28(Spec No Focus): F96-103.
- Corcos, Laurent, et Stéphanie Solier. 2005. « Épissage alternatif, pathologie et thérapeutique moléculaire ». *médecine/sciences* 21(3): 253-60.
- Corsi, Ann K., Bruce Wightman, et Martin Chalfie. 2015. « A transparent window into biology: a primer on Caenorhabditis elegans ». *Genetics* 200(2): 387-407.
- Darnell, J. E., R. Wall, et R. J. Tushinski. 1971. « An adenylic acid-rich sequence in messenger RNA of HeLa cells and its possible relationship to reiterated sites in DNA ». *Proceedings of the National Academy of Sciences* 68(6): 1321-25.

- Dieci, Giorgio et al. 2007. « The expanding RNA polymerase III transcriptome ». *TRENDS in Genetics* 23(12): 614-22.
- Drygin, Denis, William G. Rice, et Ingrid Grummt. 2010. « The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer ». *Annual review of pharmacology and toxicology* 50(1): 131-56.
- Edmonds, Mary, Maurice H. Vaughan Jr, et Hiroshi Nakazato. 1971. « Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship ». *Proceedings of the National Academy of Sciences* 68(6): 1336-40.
- Ellgaard, Lars, et Ari Helenius. 2003. «Quality control in the endoplasmic reticulum ». *Nature reviews Molecular cell biology* 4(3): 181-91.
- Fang-Yen, Christopher, Leon Avery, et Aravinthan DT Samuel. 2009. «Two size-selective mechanisms specifically trap bacteria-sized food particles in Caenorhabditis elegans ». *Proceedings of the National Academy of Sciences* 106(47): 20093-96.
- Félix, Marie-Anne, et Christian Braendle. 2010. « The natural history of Caenorhabditis elegans ». *Current biology* 20(22): R965-69.
- Ghanta, Krishna S., Takao Ishidate, et Craig C. Mello. 2021. « Microinjection for Precision Genome Editing in Caenorhabditis Elegans ». *STAR Protocols* 2(3): 100748.
- Gilbert, Walter. 1978. « Why genes in pieces? » Nature 271(5645): 501-501.
- Horvitz, H. Robert, et John E. Sulston. 1990. « "joy of the Worm" ». Genetics 126(2): 287.
- Huang, Mingxia, Guoqiang Gu, Edwin L. Ferguson, et Martin Chalfie. 1995. « A stomatin-like protein necessary for mechanosensation in C. elegans ». *Nature* 378(6554): 292-95.
- Kiontke, Karin, et Walter Sudhaus. 2006. « Ecology of Caenorhabditis species ». WormBook 9: 1-14.
- Kloosterman, Wigard P. et al. 2012. « Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms ». *Cell reports* 1(6): 648-55.
- Kuroyanagi, Hidehito, Tetsuo Kobayashi, Shohei Mitani, et Masatoshi Hagiwara. 2006. « Transgenic alternative-splicing reporters reveal tissue-specific expression profiles and regulation mechanisms in vivo ». *Nature methods* 3(11): 909-15.
- Kuroyanagi, Hidehito, Genta Ohno, Shohei Mitani, et Masatoshi Hagiwara. 2007. « The Fox-1 family and SUP-12 coordinately regulate tissue-specific alternative splicing in vivo ». *Molecular and cellular biology* 27(24): 8612-21.
- Lamm, Ehud, Oren Harman, et Sophie Juliane Veigl. 2020. « Before watson and crick in 1953 came friedrich miescher in 1869 ». *Genetics* 215(2): 291-96.
- Liang, Xiaoyu, Canyon Calovich-Benne, et Adam Norris. 2022a. « Sensory neuron transcriptomes reveal complex neuron-specific function and regulation of mec-2/Stomatin splicing ». *Nucleic acids research* 50(5): 2401-16.
- Nagy, Eszter, et Lynne E. Maquat. 1998. « A rule for termination-codon position within introncontaining genes: when nonsense affects RNA abundance ». *Trends in biochemical sciences* 23(6): 198-99.
- Pallet, Nicolas et al. 2009. « Endoplasmic reticulum stress: an unrecognized actor in solid organ transplantation ». *Transplantation* 88(5): 605-13.
- Park, Sung-Min, Tae-II Kang, et Jae-Seon So. 2021. « Roles of XBP1s in transcriptional regulation of target genes ». *Biomedicines* 9(7): 791.

- Pozzoli, U., et M. Sironi. 2005. « Silencers regulate both constitutive and alternative splicing events in mammals ». *Cellular and Molecular Life Sciences CMLS* 62(14): 1579-1604.
- Raizen, David M. et al. 2008. « Lethargus is a Caenorhabditis elegans sleep-like state ». *Nature* 451(7178): 569-72.
- Ramani, Arun K. et al. 2011. « Genome-wide analysis of alternative splicing in *Caenorhabditis* elegans ». Genome research 21(2): 342-48.
- Schafer, William. 2016. « Nematode nervous systems ». Current Biology 26(20): R955-59.
- Shuman, S. 2001. «The mRNA capping apparatus as drug target and guide to eukaryotic phylogeny ». In *Cold Spring Harbor symposia on quantitative biology*, Cold Spring Harbor Laboratory Press, 301-12.
- Stamm, Stefan et al. 2005a. « Function of alternative splicing ». Gene 344: 1-20.
- Sulston, John E., et H. Robert Horvitz. 1977. « Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans* ». *Developmental biology* 56(1): 110-56.
- Tavernarakis and, Nektarios, et Monica Driscoll. 1997. «Molecular modeling of mechanotransduction in the nematode Caenorhabditis elegans ». *Annual Review of Physiology* 59(1): 659-89.
- Uemura, Aya et al. 2013. « UBC9 regulates the stability of XBP1, a key transcription factor controlling the ER stress response ». *Cell structure and function*: 12026.
- Wakabayashi, Sadao, et Hiderou Yoshida. 2013. « The essential biology of the endoplasmic reticulum stress response for structural and computational biologists ». *Computational and structural biotechnology journal* 6(7): e201303010.
- Wang, Eric T. et al. 2008. « Alternative isoform regulation in human tissue transcriptomes ». *Nature* 456(7221): 470-76.
- Wang, Zefeng, et Christopher B. Burge. 2008. « Splicing regulation: from a parts list of regulatory elements to an integrated splicing code ». *Rna* 14(5): 802-13.
- Watson, James D., et Francis HC Crick. 1953. « Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid ». *Nature* 171(4356): 737-38.
- Yanagitani, Kota, Yukio Kimata, Hiroshi Kadokura, et Kenji Kohno. 2011. « Translational pausing ensures membrane targeting and cytoplasmic splicing of XBP1u mRNA ». *Science* 331(6017): 586-89.
- Yoshida, Hiderou et al. 2001. « XBP1 MRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor ». *Cell* 107(7): 881-91.
- You, Young-Jai, et Leon Avery. 2012. « Appetite control: worm's-eye-view ». Animal cells and systems 16(5): 351-56.
- Zahler, Alan M. 2005. « Alternative splicing in C. elegans ». WormBook: The Online Review of C. elegans Biology [Internet].
- 2018. « Pre-mRNA splicing and its regulation in *Caenorhabditis elegans* ». WormBook: The Online Review of C. elegans Biology

Annexes

Annexe 1: Alignement des séquences après séquençage du plasmide construit pmec-2-ab::MEC-2E9-E10-GGS6::FSS.

CLUSTAL 0(1.2.4) multiple séquence alignement

mec2 1143	CTGCAGaattaaattgtttttttttcagttgggaaacactttgctcactccgtagcagcc CATTTTTTTCGGTTGGGACACTTTGCTCACTCCGTAGCAGCC	60 42
	***** * ********************	
mec2	aGGGATTGGCCAAAGGACCCAAAGgtatgtttcgaatgatactaacataacatagaacat	120
1143	AGGGATTGGCCAAAGGACCCAAAGGTATGTTTCGAATGATACTAACATAACATAGAACAT **********************************	102
mec2	tttcagGAGGACCCTTGGCTCACAAGTTTGTACAAAAAAGCAGGCTccaccATGGCTGAA	180
1143	TTTCAGGAGGACCCTTGGCTCACAAGTTTGTACAAAAAAGCAGGCTCCACCATGGCTGAA ***********************************	162
mec2	AAGAACTCAACGATTATATTCCCGTTCCCAATTGATCTTCTCAGTGCATTCCTCCAACGA	240
1143	AAGAACTCAACGATTATATTCCCGTTCCCAATTGATCTTCTCAGTGCATTCCTCCAACGA ********************************	222
mec2	ACACCGCCTAAAGTgtaagttttcacagagtattcgacaaaaagcacaatctattcctat	300
1143	ACACCGCCTAAAGTGTAAGTTTTCACAGAGTATTCGACAAAAAGCACAATCTATTCCTAT ***********************	282
mec2	caaattgcagTGATAACAATTTTGCATTTCCAACGCACAAAGCTGGCGGAATACCGTCTT	360
1143	CAAATTGCAGTGATAACAATTTTGCATTTCCAACGCACAAAGCTGGCGGAATACCGTCTT ********************************	342
mec2	CCTCTTGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTC	420
1143	CCTCTTGTGAGCAAGGGCGAGGAGGAGAACATGGCCATCATCAAGGAGTTCATGCGCTTC *******************************	402
mec2	AAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAG	480
1143	AAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAG ***********************	462
mec2	GGCCGCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTG	540
1143	GGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTG **************************	522
mec2	CCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAG	600
1143	CCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAG *********************************	582
mec2	CACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAG	660
1143	CACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAG ******************************	642
mec2	CGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAG	720
1143	CGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAG ***********************************	702
mec2	GACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCC	780
1143	GACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGCGCACCAACTTCCCCTCCGACGGCCCC ****************************	762
mec2	GTAATGCAGAAGAAGACCATGGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCCGAGGAC	840

1143	GTAATGCAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGAC ****************************	822
mec2 1143	GGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGC-G-GCCACTACG GGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAAGCTGAAGGACGGCGGCCACTTACG ************************************	898 882
mec2 1143	ACGCTGAGGTCAAGAC-CACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTAC ACGCTGAGGTCAAGACCAACTAACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTAC **********************************	957 942
mec2 1143	AACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGAACGTCAACATCAAGTTGGACATCACCTCCACCAACGAGGACTACACCATCGTGGAACTA***********************************	1017 1002
mec2 1143	TACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGC-ATGGACGAGCTGTACAAGTATTACGAACGCCCCGAAGGGCCCGGCACCTCCACCGGGCGCATGGAACAAGCTGTAAAAGTT** </td <td>1073 1062</td>	1073 1062
mec2 1143	GGAACACTTCACGAATTCAAAATTCATTGACATGCGTGTGATCAGCCAATTTCATTT AGGAACACTTACCAATCATAAATCATTGGACAGTGCGGTGATATCGTGCCACTTTTATTT * * * ***** *** **** * **** ***	1130 1122
mec2 1143	T-TCCACATCGCTTTGAGTGACCTCACACCCACT TCCCACCGTCGCGTTGAAGGGCCTCACCAACCTCCCGTAAAATAATTGTGTCTTATATGG * * * **** **** * **** *	1163 1182
mec2 1143	1163 CTTATGTC 1190	

Annexe 2: Alignement des séquences après séquençage par les différents amorces les ADNc de l'ARNm de xbp-1(Cer125).

810-X1 810-x2	ATTTTGGGCCTCCCATAACCGACGCCTACTTTATGTGGCTCCCGCACGCCACGCTGGCAG
946-v1	
944-X2	
957-X1	
VDD 1	
944-x4	AGCATTTTTTTTTTTTTTTTTTTTTTTTTTTCCACACCACG
810-X1	CGTCCACAGCCTCGGAGAATGGCTCCCAAGCGTGCACTTCCAACAAAACAAGTTGTCGCA
810-x2	GCGCCACAGCCTCAGAGAATGGCTCCCAAGCGTGCACTTCCAACAGAACAAGTTGTCGCA
946-x1	
944-X2	
957-X4	
VDD_1	
944-x4	
810-X1	CAACTTCTTGGCGATGATATGGGACCATCTGGGCCACGGAAAATAGAACGACTGAATCAT
810-x2	CAACTTCTTGGCGATGATATGGGACCATCTGGGCCACGCAAAAGAGAACGACTGAATCAT
946-x1	
944-X2	GATGCTCTTCGCAATGTTTTCTTTTTTCCCTTTC
957-X4	
XBP-1	CAACTTCTTGGCGATGATATGGGACCATCTGGGCCACGCAAAAGAGAACGACTGAATCAT
944-x4	
810-X1	
810-x2	
916-v1	
011 V2	
944-XZ	IIIICIIAICACAGGAAIAIIAIIIGICGGGGGAIAAIGIICIAAAAAAAIAIIGIIGGCG
937-X4	
XBP-1 944-x4	TTGAGTCAGGAGGAGAAAATGGATCGTCGGAAACTTAAAAATCGAGTCGCAGCC
010 1/1	
810-X1	CAAACAIGICIACGAGACAAAAAGGAAGGIIGAIGCIGAAIGIIIAICGAACCCIGAAAAA
810-x2	CAAAATGCTAGAGACAAAAAGAAGGTTAGTGCTGAATGTTTATCGAACCCTGAAAAT
946-x1	
944-X2	CATATTCGACACAAAATATTTGTTTTTTTTTT
957-X4	CCCGTACAT
XBP-1 944-x4	CAAAATGCTAGAGACAAAAAGAAGGTTAGTGCTGAATGTTTATCGAACCCTGAAAAT CCCCCAAATCT
810-X1	AAGTTTAATTCTTGTAGGAAAGATCAACAAAGATCGAAGATGTGATGCCCGATCTGGTAG
810-x2	AAGTTTAATTCTTGTAGGAAAGATCAGCAAAGATCGAGGATGTGATGCGCGGATCTGGTGG
946-x1	
944-X2	TTTATATTTTTCGAAAATCCAAAGATGGGAGAGTGTCCCCGTTTTTCC
957-X4	AGTCATATTGCGAAGCACAGCA
XBP-1	AAGTTTAATTCTTGTAGGAAAGATCAGCAAAGATCGAGGATGTGATGCGCGATCTGGTGG
944-x4	AAAATTTATACCCGAAG
810-21	ΔΩΩΔΩΔΔΩΩΩΩΟΤΩΩΩΩΩΔΑΣΔΩΩΩΩΔΑΣΔΩΩΩΩΩΩΔΑΣΔΩΩΩΩΩΔΑΣΔΩΩΩΩΔΑΣΔΩΩΩΩΔΑ
810-22	
916-21	
911_V?	
957-V1	
XBP-1	AGGAGAACCGCCGGCTCCGCGCTGAAAACGAACGTCTTCGCCGTCAAAATAAAAATCTTA

944-x4	
810-X1	TGAACCAGCATAACGAGTCCGTCATGTATATGGAAGAGAACAACTAAAACTTGATGAACA
810-x2	TGAACCAGCAGAACGAGTCCGTCATGTATATGGAAGAAAACAACGAAAAACTTGATGAACA
946-x1	
944-X2	CCACCAAAACAATTCGTCTTGTTTATGAAGAAACATTGAAAAATTAGTGACT-
957-X4	CGAGAAGACTTGATGAACA
XBP-1	TGAACCAGCAGAACGAGTCCGTCATGTATATGGAAGAAAACAACGAAAACTTGATGAACA
944-x4	GAAAATTTTGAAACAAGAAAATTTTGAAAC
010 V1	
010-X1	
010-XZ	GCAAIGAIGCAIGCAICIACCAGAACGICGICIACGAAGAAGAAGICGICGGIGAGGIIG
940-X1 044 V2	
944-AZ	
997-A4	
XBP-1	
944-x4	ACAATGAT-CCTGCTTTTCCCAGAACTTTTTCT-CGAAGAAGAAGTCGTCGGTGAGGTTG
810-X1	CACCAGTTGTCGTCGTCCGAGGAGAGGATCGCCGTGCCTTTGAATCATCAGCATTCATT
810-x2	CACCAGTTGTCGTCGTCGGAGGAGGAGGAGCGTCGCCGTGCCTTTGAATCAGCAGCATTCATT
946-x1	
944-X2	CACCAGTGATC-TTGTCGAAGAGAGATTGCCGTTCC-TTGGATCAGCAGCATTCATTA
957-X4	CACCAGTTGTCGTCGTCGGAGGAGGAGGATCGCCGTGCCTTTGAATCAGCAGCATTCATT
XBP-1	CACCAGTTGTCGTCGTCGGAGGAGGAGGATCGCCGTGCCTTTGAATCAGCAGCATTCATT
944-x4	CACCAGTTGTCGTCGTCGGAGGAGGAGGATCGCCGTGCCTTTGAATCAGCAGCATTCATT
810-X1	ATGAGCCTCATCAGC-GGAACAGGCCCGATCCACCTCCATCAAGCAACAACATCAGCA
810-x2	ATGAGCCTCAGCAGTGGGAACAGGCCCGATCCACCTCCATCAA-CAACAACATCAGCA
946-x1	CCCCC
944-X2	CAGATCCTCAGCAGTGGGAACAGGCTCGATCCACTTCCATCAA-CAACAACATCAGCA
9.57-X4	ATGAGCCTCAGCAGTGGGAACAGGCCCGATCCACCTCCATCAA-CAACAACATCAGCA
XBP-1	ATGAGCCTCAGCAGTGGGAACAGGCCCGATCCACCTCCATCAA-CAACACATCAGCA
944-x4	ATGAGCCTCAGCAGTGGGAACAGGCCCGATCCACCTTTCCTTCAACACAACATCAGCA ** *
810-X1	ACCAACTCCGTCGTATGGATTCCAAGAAGAACAACAACAATCACTGTGGATATGTATCTAA
810-x2	ACCAACTCCGTCGTATGGATTCCAAGAAGAACAACAACAATCAGTGTGGATATGTATCTAA
946-x1	<i>TACAAAAAGAGAAATACAAAAAGAGAAATAA</i>
944-X2	ACCAATTCCGTCGTATGGATTCCAAGAAGAACAACAACAATCAGTGTGGATATGTATCTAA
957-X4	ACCAACTCCGTCGTATGGATTCCAAGAAGAACAACAACAATCAGTGTGGATATGTATCTAA
XBP-1	ACCAACTCCGTCGTATGGATTCCAAGAAGAAGAACAACACAATCAGTGTGGATATGTATCTAA
944-x4	
	* *** **** **
810-X1	CTATCATCTCGATTCTATGCAA-CCACATGGATCGCAACAAGAAGATGGACACCTCGAAC
810-x2	CTATCATCTCGATTCTATGCAA-CCACATGGATCGCAACAAGAAGATGGACACCTCGAAC
946-v1	
944-X2	
957_ <i>Y</i> /	
ידע 1	
944-x4	* * * * * * * * * * * * * *
810-X1	AAGTCCTCGAACATCTCAAGGTAATATAAGTAGAGACTTCCTAATGTTTATGTATTAA
810-x2	AAATCCTCGAACATCTCAAGGTAAGATAATTAGAAACTTCCTAATGTTTATGTATTAA
946-x1	GGAAACCCCCCCCCCCCATAAAATTTTTTTTTTT
944-X2	AAATCCTCGAACATCTCAAGGTAAGATAATTAGAAACTTCCTAATGTTTATGTATTAA
957-X4	AAATCCTCGAACATCTCAAGGTAAGATAATTAGAAACTTCCTAATGTTTATGTATTAA
XBP-1	AAATCCTCGAACATCTCAAGGTAAGATAATTAGAAACTTCCTAATGTTTATGTATTAA
944-x4	
~	** ** * * * * ** *** ** ** ** ** ** **

810-X1	CTTCATAAATTTCCAGAGCCCCAAGCGGAGAGTTCGATCGA
810-x2	CTTCATAAATTTCCAGAGCCCAAGCGGAGAGTTCGATCGA
946-x1	CCCGGGAGGGTGTTTTTTTGGTGTGT
944-X2	CTTCATAAATTTCCAGAGCCCAAGCGGAGAGTTCGATCGA
957-X4	CTTCATAAATTTCCACACCCCAACCCCACACTTCCATCCATTCCTTCCTCC
עסע 1	
ABP-1	CITCATAAATITCCAGAGCCCAAGCGGAGGGTCGATCGATCGTGCGGCTACATIGA
944-x4	CTTCATAAATTTCCAGAGCCCAAGCGGAGAGTTCGATCGA
	** * * * *** * * **
810-X1	GGAAAGGAGCAGACGGTTATGCATCGTCTTGTTCAAGCAGATCCATGTACACATCTTCAG
810-x2	GG-AAGGAGCAGACGGTTATGCAGCGTCTTGTTCAAGCGGATCCATGTACACATCTTCAG
916_ <u>v</u> 1	
944-XZ	GG-AAGGAGCAGACGGTTATGCAGCGTCTTGTTCAAGCGGATCCATGTACACATCTTCAG
957-X4	GG-AAGGAGCAGACGGTTATGCAGCGTCTTGTTCAAGCGGATCCATGTACACATCTTCAG
XBP-1	GG-AAGGAGCAGACGGTTATGCAGCGTCTTGTTCAAGCGGATCCATGTACACATCTTCAG
944-x4	<i>GG-AAGGAGCAGACGGTTATGCAGCGTCTTGTTCAAGCGGATCCATGTACACATCTTCAG</i>
	**** * *** ** ** ** * ** * *** *
810-X1	AGACGCGTGAAACACTTTGCCGGCCGCAATGAGCAAGGGGCGAGGAGCTGTTCACCGGGG
810-22	
010 22	
946-X1	AGGAAAACIIIG-GGCCCAGAAA-AGGGGGGGIIIICC-CGG
944-X2	AAACGCGTGAAACACTTTG-CGGCCGCAGTGAGCAA-GGGCGAGGAGCTGTTCACC-GGG
957-X4	AAACGCGTGAAACACTTTG-CGGCCGCAGTGAGCAA-GGGCGAGGAGCTGTTCACC-GGG
XBP-1	AAACGCGTGAAACACTTTG-CGGCCGCAGTGAGCAA-GGGCGAGGAGCTGTTCACC-GGG
944-x4	AAACGCGTGAAACACTTTG-CGGCCGCAGTGAGCAA-GGGCGAGGAGCTGTTCACC-GGG
	* * * * * * * * * * * * * * * * * * * *
010 21	
810-X1	GIGGIGCGCAICCIGAICIAGCIGGACIGCCIACGIAIACGGGCCACAAGIIICAGCGIG
810-x2	GTGGTGCCCATCCTGGTCGAGCTGGACGG-CGACGTAAAC-GGCCACAAG-TTCAGCGTG
946-x1	GGGGCCCCCTGGGGGTGGAG-GGGAGTAAACGCCCCAAA-TTTAAGGGT
944-X2	GTGGTGCCCATCCTGGTCGAGCTGGACGG-CGACGTAAAC-GGCCACAAG-TTCAGCGTG
957-X4	GTGGTGCCCATCCTGGTCGAGCTGGACGG-CGACGTAAAC-GGCCACAAG-TTCAGCGTG
XBP-1	GTGGTGCCCATCCTGGTCGAGCTGGACGG-CGACGTAAAC-GGCCACAAG-TTCAGCGTG
9/1 - y/1	
<u>944</u> -X4	* ** * * * * * * **** * *** *** *** **
810-X1	TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAATTTCATCTGCAC
810-x2	TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG-TTCATCTGCA-
946-x1	TCCGGGAGGGGAGGCGAT-CCACTT-CGGCAAG-TGA-CCTGAAG-TTCAT-TGCCC
944-X2	TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAG-TTCATCTGCAC
957-X4	TCCGCCGAGGGCGAGGGCGATGCCCACCTACGGCAGCCTGACCCTGAAG-TTCATCTGCAC
VDD 1	
ABF = I	
944-X4	1CCGGCGAGGGCGAGGGCGAIGCCACCIACGGCAAGCIGACCCIGAAG-IICAICIGCAC ***** * * * ****** **** * ***** *******
810-X1	CACCGGCAAGCTGCCCGTGGCCCTGGCCCCACCCTCGTGGACCACCCTGAACTTACGGGC
810-x2	CACCGGCCAGCTGCCCGT-GCCCTGG-CCCACCCTCGT-GA-CACCCTG-AACTAC-GGG
946-x1	CA-CGGCAAGTT-CCCGT-G-CCTGG-CCCA-CCTTGG-GACCA-CCTG-ACTTAGGG
944-X2	CACCERCAAGCTECCCET_ECCCTEE_CCCACCCTCET_GACCACCCTE_ACCTAC_EEC
0E7 V/	
937-X4	
XBP-1	CACCGGCAAGCTGCCCGT-GCCCTGG-CCCACCCTCGT-GACCACCCTG-ACCTAC-GGC
944-x4	CACCGGCAAGCTGCCCGT-GCCCTGG-CCCACCCTCGT-GACCACCCTG-ACCTAC-GGC ** **** ** * ***** * ***** **** *** **
810-X1	TTGCAAGGGCTTTCAGCCGGCTACCCCCGACCCATTGAAGACAACCCGATTTCTTTC
810-x2	GGGAAAGGTTAAACC-GCTACCCGACC-CATGAAG-CC-CCCGACTTCTTC-AG-
946-x1	GTGCAGGGTTCAGCC-GTTACCCCACCACA-CAAC-CACCACCACCAC-TCTTCAAC-
911 AI 911_V2	
244-AZ	
95/-X4	GTGCAGTGCTTCAGCC-GCTA-CCCCGACCACATGAAG-CAGCACGACTTCTTCAAG-
XBP-1	GTGCAGTGCTTCAGCC-GCTA-CCCCGACCACATGAAG-CAGCACGACTTCTTCAAG-
944-x4	GTGCAGTGCTTCAGCC-GCTA-CCCCGACCACATGAAG-CAGCACGACTTCTTCAAG-
	* * * ** * ** ** ****** **** * *** **** ***
810-X1	TTCCCCCATGGCCCAAAGGTTTCTTTCCAGAGAGGGCCCCCTTTTTTTAAGGGCAAAA
810-22	
916-21	
ショロースエ	-ICCGCCAI-GCCCGAAGGCIACGICCAGGAGCGCACCAICIICIICAAGGACG

944-X2	-TCCGCCAT-GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG
957-X4	-TCCGCCAT-GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG
XRP-1	
944-X4	-ICCGCCAI-GCCCGAAGGCIACGICCAGGAGCGCACCAICIICIICAAGGACG
	*** *** * ** * * * * * * * * * * * * * *
010 V1	
810-X1	AGGGCGAATTAAAAAAACGCGCCCCCAAGGGGGATTTTCTCGGGGGCCCCACCCCTGGGG
810-x2	ACGGGAAAAAAAACCCGGCCCAAGGGGGGTTGCGGGGGGACCCTGGGGA
946-x1	ACGGCAAATACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA
944-X2	ACGGCAACTACAAGACCCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA
057 V/	
937-X4	ACGCAACIACAAGACCCGCGCCGAGGIGAAGIICGAGGGCGACACCCIGGIGA
XBP-1	ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA
944-x4	ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGA
	* * * * * * * * * * * * * * * * * * * *
810-X1	GGACACCCCCCTCTTCGCTGAAAGGAGGTCCCCTTTTTCTTGGGGGGGAGGGGGGAAATCT
810-x2	ATC-CTTTAAAGTGGAGGGTTCTTTTTTGGGGGAGGGGAT
946-x1	ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT
944-X2	AAAGGAGGACTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT
057 V/	
957-74	ACGCAICGAGCIGAAGGGCAICGACIICAAGGAGGACGGCACAI
XBP-1	ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT
944-x4	ACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACAT
	* * * * * * * * * * * * * *
810-X1	CTCGTGGGCACAAACGTGTGAAAAAATTATAAACACCCCCCACAATTTTTTTT
810-x2	CCCG-GGCCACCGGG-TAATTCCCCCCCCCC
946-x1	
944-XZ	CCTG-GGGCACAAGC-TGGAGTACAACTAC-AACAGCCACAACGTCTATATCATGGCCGA
957-X4	CCTG-GGGCACAAGC-TGGAGTACAACTAC-AACAGCCACAACGTCTATATCATGGCCGA
XBP-1	CCTG-GGGCACAAGC-TGGAGTACAACTAC-AACAGCCACAACGTCTATATCATGGCCGA
944 - x4	CCTG-GGGCACAAGC-TGGAGTACAACTAC-AACAGCCACAACGTCTATATCATGGCCGA
<i>y i i i i</i>	* * ** ** * * ** ** * * * * * *
010 V1	
010-X1	AAAAAAAAAAAAGGGCGGGGAIIIIIIAICICCCCCCCCCGAGGGGGGGGGG
810-x2	
946-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-x2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-X2 957 X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-X2 957-X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-X2 957-X4 XBP-1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG
946-x1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** ** * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4	MARINAMINACCOLOG COCOTA MARINAMINCCCCCCCC CONCOCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x2 946-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG cAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x2 946-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x2 946-x1 944-x2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x2 946-x1 944-X2 957-X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-X2 946-x1 944-X2 957-X4 YBP 1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCGGG CCGCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGGCCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG CCCGACAACCACTACCTGGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAACGG GCCCGACAACCACTACCTGGCCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG ** ** ** * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X1 810-X1 810-X1 810-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG ** * ** * * * * * * * * * * * * * * *
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-x2 946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X1 810-X2 946-x1	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCCCCCCAAGGCAGAACACCCCCCACGGCGG
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X1 810-X2 946-x1 944-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCCAAGATCCGCCACAACATCGAGGACGGCCGGC
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-X2	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCCAAGATCCGCCACAACATCGAGGACGGCGGC CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCTCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGCAGACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGCAGCACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCCCGACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCCAACGAGAAGCG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCG CCCGACAACCACTACCTGGGGCCCCGGCCTGGGCACAAGACCCCAACGACAAGCC CCGACCACCACTGCTGCGCGCGCGGGATCACTCTCGGCATGGACG CCGACCACCACTGCCGCCGCGCGGGATCACTCTCGGGCACGGCCCGGCATCCCCCCGCGCGGATCACCTCCGGCCCGGCGCGGATCACTCTCGGCCCCCCCGCGCGCG
946-x1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-X2 957-X4 XBP-1 944-X2 957-X4 XBP-1 944-x4 810-X1 810-X2 946-x1 944-x4 810-X1 810-X1 810-X2 946-x1 944-x4	CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAG CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCGGC CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCGGC CAAGCAGAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCGGCG CAAGCAGAGAACGGCCACTACAAGGAACACCCCCCACCGCGCGGCGCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCT CGCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG GCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG CCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGG CCCGACAACCACTACCTGGCGCCAGTCGCCCTGAGCAAAGACCCCAACGAGAAGG CCCGACAACCACTACCTGGCGCCGCGCGGGACCACTCTCGGGAAGGG CCGACAACCACTACCTGGCGCGGGGCCGCGGGACCACTCTCGGGCAAGGACG CCGACCACTGCTGCTGCTGGAGTTCGTG-ACCGCCGCGCGGGATCACTCTCGGGACTGGC CGATCACATGGTCCTGCTGCGGCGCGCGCGGGATCACTCTCGGC

944-x4	CGATCACATGGTCCTGCTGGAGTTCGTGTACCGCCGCGGGATCACTCTCGGTGC
810-X1	
810-x2	
946-x1	AGCTGTACAAGTAATTGCCTCCTCCGAGAACGTCATCACCGAGTTCATGCGCTTCAAGGT
944-X2	
957-X4	
XBP-1	AGCTGTACAAGTAATTGCCTCCTCCGAGAACGTCATCACCGAGTTCATGCGCTTCAAGGT
944-x4	
810-X1	
810-x2	
946-x1	GCGCATGGAGGGCACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCG
944-X2	GCATGGAGCACATA
957-X4	ТGAAAATGATATCATCTGCGCCCCAAAAGAAAAAAAAAAA
XBP-1	GCGCATGGAGGGCACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCG
944-x4	CATGGATGA
010 V1	
810-x2	
810-X2	
946-X1	CCCCTACGAGGGCCACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTT
944-X2	
957-X4	ААААААСАТАААААААААААААААААААТТТАААААААА
XBP-1	CCCCTACGAGGGCCACAACACCGTGAAGCTGAAGGTGACCAAGGGCGGCCCCCTGCCCTT
944-x4	
810-X1	
810-x2	
946-v1	
011 V2	
944-XZ	
937-A4	
XBP-1 944-x4	
810-X1	
810-x2	
946-x1	CGCCGACATCCCCGACTACAAGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGT
944-X2	
9.57 - X4	
XRP-1	CCCCGACATCCCCGACTACAAGAAGCTCTTCCCCGAGGGCTTCAAGTGGGAGCGCGT
944-x4	
010 11	
OIU-XI	
810-x2	
946-x1	GATGAACTTCGAGGACGGCGGCGTGGCGACCGTGACCCAGGACTCCTCCCTGCAGGACGG
944-X2	
957-X4	
XBP-1	GATGAACTTCGAGGACGGCGGCGTGGCGACCGTGACCCAGGACTCCTCCCTGCAGGACGG
944-x4	
810-81	
810_**?	·
$O \perp U = X \angle$	
946-X1	UTGUTTUATUTACAAGGTGAAGTTCATCGGCGTGAACTTCCCCTCCGACGGCCCCGTGAT
944-X2	
957-X4	
XBP-1	CTGCTTCATCTACAAGGTGAAGTTCATCGGCGTGAACTTCCCCTCCGACGGCCCCGTGAT
944-x4	

810-X1	
810-x2	
946-x1	GCAGAAGAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGT
944-X2	
957-X4	AAAAAAAAAAAAT
XBP-1	GCAGAAGAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGT
944-x4	
810-X1	
810-x2	
946-x1	GCTGAAGGGCGAGACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGA
944-X2	
957-X4	
XBP-1	GCTGAAGGGCGAGACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGA
944-x4	
810-X1	
810-x2	
946-x1	GTTCAAGTCCATCTACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACGTGGA
944-X2	
957-X4	
XBP-1	GTTCAAGTCCATCTACATGGCCAAGAAGCCCGTGCAGCTGCCCGGCTACTACTACGTGGA
944-x4	
810-X1	
810-x2	
946-x1	CGCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAGCAGTACGAGCG
944-X2	
957 - X4	
XRP-1	CGCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAGCAGTACGAGCG
944-x4	
010 V1	
010-A1	
$O \perp U = X \angle$	
940-X1	CALLUGAGGGUUGUGA
944-XZ	
<i>93/-X4</i>	
XBP-I	LALLGAGGGUUGUUAUUAUUTGTTUUTGTAA
944-X4	



Annexe 4: Protocole simplifié de la modification du génome par CRISP-Cas9

Selon le protocole de Krishna S Ghanta *et al.* 2021 ; « Microinjection for precision genome editing in *Caenorhabditis elegans* »






ÉCOLE PRATIQUE DES HAUTES ÉTUDES SCIENCES DE LA VIE ET DE LA TERRE

Etude de l'épissage alternatif des gènes mec-2 et xbp-1

MEDERIC Myriam

12 décembre 2022 RÉSUMÉ

L'épissage alternatif (EA) est un mécanisme très important qui touche un peu plus de 70 % des gènes chez l'humain, il n'a pas pour seul but d'introduire de la diversité génétique. Il permet également la régulation de l'expression génique, un nombre important des ARNm ayant subi un EA ne sont jamais traduits en protéine mais sont dirigés vers la voie NMD (Non-sense RNA Mediated Decay) pour être dégradés.

L'EA joue un rôle dans la stabilité des ARNs en produisant des isoformes avec un codon stop prématuré qui devient alors une cible pour la dégradation médiée par le NMD.

Ici nous avons choisis d'étudier l'EA de 2 gènes, le gène de la mecanosensation du toucher, mec-2 et le gène x-box binding protein-1 (xbp-1) qui est un des gènes qui intervient dans la réponse aux stress du réticulum endoplasmique (UPR) pour rétablir l'homéostasie protéique dans le réticulum.

Ces 2 gènes ont en commun, en dehors du fait qu'ils subissent un épissage alternatif, c'est que ce sont des homologues de gènes chez l'humain.

Le gène mec-2 possède 2 promoteurs connus, quand il est transcrit par son promoteur proximal pmec2-ab, il produit 2 isoformes, MEC-2A et MEC-2B après épissage par la protéine MEC-8. Le 2ème promoteur dit distal, *pmec-2c*, se traduit par une protéine de forme courte MEC-2C qui ne résulte pas d'un épissage.

Nous avons transfecté les vers C. elegans avec une construction du gène mec-2 fusionnée à un système de gènes rapporteurs afin de visualiser le profil d'expression de ces 2 isoformes dans le ver. Elles sont visibles dans les neurones du toucher pour les 2 formes longues MEC-2A et courtes MEC-2B, cette dernière se retrouve également dans d'autres neurones de C. elegans.

Le gène *xbp-1* a un promoteur connu et est épissé de manière non conventionnelle au moment du stress du réticulum endoplasmique. Dans le cytoplasme son préARNm est épissé de 23 bases par l'endoribonucléase IRE-1, puis les 2 bouts sont ligés par la ligase RtcB. La traduction de cet ARNm épissé donne naissance à une protéine, facteur de transcription XBP-1s actif qui joue un rôle primordial dans la réponse aux stress du RE.

Nous avons développé différentes lignées de vers dans le but d'étudier les différents facteurs de régulation de ce gène. L'expression après transfection dans le ver montre que la protéine XBP-1 exogène donne des résultats inattendus et différents de la littérature.

Nous avons constaté au vue des résultats, que l'expression du gène xbp-1 est médié par plusieurs facteurs qui changent les profils d'expression de XBP-1. Aucune des lignées construites n'a permise de poursuivre l'étude pour rechercher les facteurs régulant l'épissage de xbp-1.

Mots-clés : Epissage alternatif, C. elegans, xbp-1, mec-2, système rapporteur, GFP, mCherry régulation, transgène.