



Fra-1 et Fra-2 dans les cancers du sein triple négatifs : mécanismes transcriptionnels et identification de cibles thérapeutiques potentielles

Claire Tolza

► To cite this version:

Claire Tolza. Fra-1 et Fra-2 dans les cancers du sein triple négatifs : mécanismes transcriptionnels et identification de cibles thérapeutiques potentielles. Médecine humaine et pathologie. Université Montpellier, 2016. Français. NNT : 2016MONTT094 . tel-01561011

HAL Id: tel-01561011

<https://theses.hal.science/tel-01561011>

Submitted on 12 Jul 2017

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

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

THÈSE

Pour obtenir le grade de Docteur

Délivré par l'**Université de Montpellier**

Préparée au sein de l'école doctorale **CBS2**
Sciences Chimiques et Biologiques pour la Santé

Et de l'unité de recherche **IGMM CNRS UMR5535**
Institut de Génétique Moléculaire de Montpellier

Spécialité **Biologie-Santé**

Présentée par **Claire TOLZA**

**Fra-1 et Fra-2 dans les cancers du sein triple
négatifs : mécanismes transcriptionnels et
identification de cibles thérapeutiques
potentielles**

Soutenue le **16 Décembre 2016** devant le jury composé de

Dr. Ruth RIMOKH, Directeur de Recherche,
INSERM, Lyon

Président/ Rapporteur

Dr. Karine BELGUISE, Chargée de Recherche,
CRNS, Toulouse

Rapporteur

Dr. Didier TROUCHE, Directeur de Recherche,
CNRS, Toulouse

Examineur

Dr. Dany CHALBOS, Directeur de Recherche,
INSERM, Montpellier

Examineur

Dr. Isabelle JARIEL-ENCONTRE
Chargée de Recherche, CNRS, Montpellier

Directeur de Thèse



REMERCIEMENTS

Tout d'abord, j'aimerais remercier les membres du jury d'avoir accepté d'évaluer mon travail de thèse : Ruth Rimokh et Karine Belguise en tant que rapporteurs, Didier Trouche et Dany Chalbos en tant qu'examinateurs.

Je tiens en premier lieu à remercier Isabelle Jariel-Encontre pour la qualité de son encadrement au cours de mes quatre années de thèse. Merci pour ta grande disponibilité, tes nombreux conseils, ton dynamisme, ton enthousiasme, ton soutien au quotidien et la confiance que tu m'as accordée. Merci pour ta grande gentillesse, grâce à toi ces quatre années sont passées très vite et je suis très heureuse de les avoir partagées avec toi. La réussite de cette thèse doit beaucoup à ton encadrement et je me considère chanceuse d'avoir pu en bénéficier.

Je remercie également tous ceux sans qui cette thèse ne serait pas ce qu'elle est, aussi bien par les discussions que j'ai eues la chance d'avoir avec eux que par leurs suggestions ou contributions. Je pense ici en particulier à Marc Piechaczyk et Charles Lecellier. Je remercie aussi Jean-Christophe Andreau, Ruth Rimokh et Charles Theillet pour les conseils qu'ils m'ont prodigués toutes ces années comme membres de mon comité de thèse.

Merci à Marc qui m'a ouvert les portes de son laboratoire et à Mireia avec qui j'ai fait mes débuts dans le labo, ce stage de Master 2 a été un élément décisif dans mon choix de poursuivre en thèse. Je tiens également à exprimer toute ma reconnaissance à tous les autres membres et ex-membres de l'équipe « Oncogenèse et Immunothérapie », merci pour la bonne ambiance que j'ai pu vivre au quotidien. Les nombreuses discussions que j'ai pu avoir avec chacun d'entre vous m'ont beaucoup apporté et ont hautement contribué à ce que ces quatre années soient à la fois formatrices, amicales, inoubliables et drôles parfois ! Un clin d'œil particulier à Fabienne pour notre collaboration qui je l'espère sera des plus fructueuses...

Je tiens aussi à mentionner le plaisir que j'ai eu à travailler au sein de l'IGMM et j'en remercie ici tous les membres, scientifiques comme non-scientifiques. Merci à tous les doctorants et post-doctorants du campus IGMM-CRBM-CPBS que j'ai pu côtoyer pendant ces quatre années. Une pensée particulière pour Tamara, Hayeon, Sophie et Marta, vous savez pourquoi ! Sans oublier tous ceux que je ne nomme pas, mais qui se reconnaîtront.

Mes remerciements vont aussi à celles et ceux qui me sont chers et sans lesquels ma vie hors labo n'aurait jamais été la même. Leurs attentions et encouragements m'ont accompagnée tout au long de ces années.

Enfin, je remercie profondément ma famille pour sa présence permanente et ses encouragements pendant les moments d'incertitude. Et surtout, un grand merci à mes parents et ma sœur qui ont toujours respecté mes choix. Merci pour votre affection et votre soutien indéfectible.

Merci à tous !

TABLE DES MATIERES

TABLE DES MATIERES	5
TABLE DES ILLUSTRATIONS	8
TABLE DES TABLEAUX.....	12
TABLE DES ANNEXES.....	13

AVANT PROPOS.....15

INTRODUCTION 19

I. LE CANCER DU SEIN	19
I.1. Quelques chiffres.....	19
I.2. Histologie du sein.....	19
I.3. Le cancer du sein en général	22
I.4. Classification des cancers mammaires	25
I.4.a. Classification sur critères anatomopathologiques.....	25
I.4.b. Classification sur critères moléculaires	27
I.4.c. Classification sur critères transcriptomiques	28
I.5. Le cancer du sein triple négatif (TNBC).....	32
II. LES FACTEURS DE TRANSCRIPTION AP-1	36
II.1. AP-1, une famille de complexes transcriptionnels.....	36
II.2. Régulation de l'activité du complexe AP-1	38
II.2.a. Regulation de l'abondance des membres AP-1 (FOS et JUN)	38
II.2.b. Dimérisation des protéines AP-1	40
II.2.c. Régulation de l'activité des protéines AP-1 par modifications post-traductionnelles.....	40
II.2.d. Interaction avec d'autres facteurs de transcription	43
II.3. AP-1 et mécanismes de régulation de la transcription des gènes cibles	44
II.3.a. Modifications de la structure chromatinienne par AP-1	45
II.3.b. Recrutement de modificateurs des histones par AP-1	47
II.3.c. Recrutement de co-facteurs transcriptionnels par AP-1	47
II.4. AP-1 et tumorigenèse.....	48
II.5. Fra-1, Fra-2 et cancer du sein.....	50
III. RÉGULATION DE LA TRANSCRIPTION : mécanismes généraux et nouveaux concepts	53
III.1. Structure tridimensionnelle de la chromatine	53
III.1.a. Organisation hiérarchique de la chromatine	53
III.1.b. Territoires chromosomiques	55
III.1.c. Domaines chromosomiques (compartiments A et B)	57
III.1.d. Domaines topologiquement associés (TAD)	57
III.1.e. Boucles de chromatine (DNA looping)	58
III.2. Dynamique de la structure chromatinienne et régulation de la transcription	59
III.2.a. Régions dépourvues de nucléosomes ou NFR.....	59
III.2.b. Modifications post-traductionnelles d'histones	60
III.3. Les différentes étapes de la transcription.....	63
III.3.a. Transcription basale et induite	63
III.3.b. Formation du complexe de pré-initiation.....	65
III.3.c. Initiation de la transcription.....	66
III.3.d. Transition entre initiation et élongation précoce	67
III.3.e. Élongation et notion de pause transcriptionnelle	67
III.4. Régulation de la transcription par l'ARN Pol II	68
III.4.a. Domaine C-terminal (CTD) de l'ARN Pol II et phosphorylations.....	68

III.4.b. Phosphorylation de la sérine-5 du CTD de RPB1 (Pol II P-S5)	69
III.4.c. Phosphorylation de la sérine-2 du CTD de RPB1 (Pol II P-S2)	69
III.4.d. Valeur informative du profil Pol II P-S5/P-S2	70
III.4.e. Classement des gènes suivant l'état d'activité de l'ARN Pol II	71
III.5. Concepts émergents	73
III.5.a. Enhancers et promoteurs : une classe unique d'éléments fonctionnels	73
III.5.b. Communication entre enhancers et promoteurs	76
RÉSULTATS	81
ÉTUDE I	85
I. MODÈLE EXPÉRIMENTAL	85
I.1. Caractérisation du système expérimental	85
I.1.a. Mesure des abondances relatives des ARNm AP-1 dans les différentes lignées cellulaires d'intérêt	85
I.1.b. Comparaison de l'abondance relative des membres AP-1 dans les différentes lignées cellulaires d'intérêt	87
I.2. Impact de l'inhibition par RNAi de Fra-1 et/ou Fra-2 sur les autres membres AP-1	89
II. IDENTIFICATION DES TRANSCRIPTOMES GOUVERNÉS PAR FRA-1 ET/OU FRA-2 DANS LA LIGNÉE MDA-MB231	91
II.1. Première étude transcriptomique : identification des gènes cibles de Fra-1	92
II.1.a. Résultats et confirmation par RT-qPCR	93
II.1.b. Discussion sur les biais liés aux siRNA	96
II.2. Deuxième étude transcriptomique : identification des transcriptomes gouvernés par Fra-1 et/ou Fra-2	98
II.2.a. Modification du protocole expérimental	98
II.2.b. Analyse des données brutes	99
II.2.c. Analyse des données transcriptomiques et « nettoyage des listes »	100
II.3. Résultats	104
II.4. Conclusion sur les approches transcriptomiques	106
III. IDENTIFICATION ET CARACTÉRISATION DES SITES DE FIXATION DE FRA-1 ET FRA-2 SUR LA CHROMATINE DE LA LIGNÉE MDA-MB231	109
III.1. Identification des sites de fixation de Fra-1 et Fra-2	109
III.2. Caractérisation des sites de fixation de Fra-1 et Fra-2	111
III.2.a. Identification des régions enhancers dans la lignée MDA-MB231	112
III.2.b. Caractérisation des régions régulatrices liées par Fra-1 et/ou Fra-2 dans la lignée MDA-MB231	113
III.2.c. Identification des gènes modulés par Fra-1 et/ou Fra-2 via leurs régions enhancers et/ou promoteurs dans la lignée MDA-MB231	115
III.3. Conclusion sur les approches ChIP-seq et FANTOM5	117
IV. SÉLECTION DES GÈNES CANDIDATS POUR POURSUIVRE LES ÉTUDES FONCTIONNELLES ET MÉCANISTIQUES	120
IV.1. Croisement des données	120
IV.1.a. Croisement des données transcriptomiques et ChIP-seq	120
IV.1.b. Croisement avec deux banques de données de cancers du sein	121
IV.2. Validation par RT-qPCR et extension des observations à deux autres lignées TNBC	125
Étude II	131
I. HMGA1 : Introduction	132
I.1. Structure et fonctions moléculaires de HMGA1	132
I.1.a. Les protéines HMG	132
I.1.b. Caractéristiques structurales des HMGA	133
I.1.c. Mécanismes d'action des protéines HMGA1	134

I.2.	Fonctions physiologiques de HMGA1	137
I.3.	Mécanismes connus de la régulation de l'expression de HMGA1	139
I.3.a.	Mécanismes transcriptionnels	139
I.3.b.	Mécanismes post-transcriptionnels	140
II.	HMGA1 : Résultats	141
II.1.	Fra-1 et Fra-2 régulent la transcription du gène HMGA1	141
II.2.	Identification des mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour la transcription du gène HMGA1	142
II.2.a.	Données de CHIP-seq pour le gène HMGA1	142
II.2.b.	Analyse de certaines modifications d'histones en présence et en absence de Fra-1 et/ou Fra-2 au niveau du gène HMGA1	144
II.2.c.	Analyse du recrutement de l'ARN polymérase Pol II et de son activité	147
II.2.d.	Analyse du recrutement de P300 au niveau des régions enhanceurs du gène HMGA1	151
III.	HMGA1 : Conclusion	152
Étude III		157
I.	CD68 : Introduction	157
I.1.	Structure et expression de CD68	157
I.2.	Fonction de CD68	159
II.	CD68 : Résultats	161
II.1.	Fra-1 et Fra-2 régulent la transcription du gène CD68	161
II.2.	Modulation de l'expression de CD68 <i>in vitro</i>	162
II.2.a.	Analyse de l'abondance de CD68 dans les 3 lignées TNBC versus la lignée MCF7	162
II.2.b.	Mesure de l'inhibition de l'expression CD68 par transfection de siRNA	163
II.3.	Impact de l'absence de CD68 sur la prolifération, la viabilité et le cycle cellulaire dans la lignée MDA-MB231	164
II.3.a.	Impact de l'absence de CD68 sur la prolifération	164
II.3.b.	Impact de l'absence de CD68 sur la viabilité	166
II.3.c.	Impact de l'absence de CD68 sur le cycle cellulaire	167
II.4.	Impact de l'absence de CD68 sur la motilité, l'invasion et la capacité à former des mammosphères	169
II.4.a.	Impact de l'absence de CD68 sur la motilité en 2D	169
II.4.b.	Impact de l'absence de CD68 sur l'invasion et la capacité à former des mammosphère en 3 dimensions	171
II.5.	Rôle potentiel de CD68 dans l'autophagie ?	174
II.5.a.	Localisation de CD68 dans la lignée MDA-MB231	174
II.5.b.	Le processus autophagique	178
II.5.c.	Modulation de la voie autophagique et impact sur la localisation de CD68 dans la lignée MDA-MB231	182
III.	CD68 : Conclusion	185
DISCUSSION GÉNÉRALE		191
ANNEXES		195
MATÉRIELS ET MÉTHODES		207
BIBLIOGRAPHIE		217

TABLE DES ILLUSTRATIONS

Figure 1 : Représentation du sein.	20
Figure 2 ; Système lymphatique du sein.....	21
Figure 3 : Représentation schématique du développement de la glande mammaire.....	22
Figure 4 : Principales étapes de la formation d'une métastase.	23
Figure 5 : Caractéristiques émergentes et facteurs favorisants.....	24
Figure 6 : Chevauchement entre les cancers du sein TNBC, Basal-like et la voie BRCA1 mutée.	29
Figure 7 : Sous-typage des tumeurs TNBC d'après les données du TCGA.....	34
Figure 8 : Représentation schématique de l'homologie de séquence entre les différentes protéines FOS et JUN humaines.	37
Figure 9 : Structure tridimensionnelle de l'hétérodimère JUN/FOS.....	37
Figure 10 : Protéines impliquées dans la régulation transcriptionnelle de Fra-1.	38
Figure 11 : Différents niveaux de régulation de l'activité AP-1 par les voies MAPK dans la cellule.....	41
Figure 12 : Ouverture de la chromatine condensée par un facteur pionnier pour initier la transcription.....	46
Figure 13 : Les différents niveaux hiérarchiques d'organisation de la chromatine dans un noyau de mammifère.	54
Figure 14 : Organisation hiérarchique du génome mammifère.....	55
Figure 15 : Implications de CTCF, cohésine et Mediator à l'échelle des TAD..	58
Figure 16 : Modification d'histones caractéristiques de l'activité des séquences régulatrices enhancer et promoteur.	62
Figure 17 : Interactions régulatrices de la transcription.	64
Figure 18 : Régulation de l'activité P-TEFb par P300.....	68
Figure 19 : Phosphorylations du domaine C-terminal de l'ARN Pol II.	70
Figure 20 : Schéma représentant la comparaison de la pause transcriptionnelle de l'ARN Pol II à l'activité du gène.	72
Figure 21 : Les éléments de régulation actifs, enhancer comme promoteur, sont liés par l'ARN Pol II et transcrits.....	73
Figure 22 : La stabilité des ARN transcrits distingue les régions régulatrices enhancer et promoter.....	74
Figure 23 : Représentation schématique de la communication enhancer-promoteur.....	76
Figure 24 : Mesure des abondances relatives des ARNm des membres AP-1 dans les quatre lignées d'intérêt.	86
Figure 25 : Analyse de l'expression des membres AP-1 dans différentes lignées de cancer du sein.....	88

Figure 26 : Impact de l'inhibition de Fra-1 et/ou Fra-2 par RNAi sur les membres FOS, dans les trois lignées triple négatives d'intérêt.....	90
Figure 27 : Gènes différentiellement exprimés en absence de Fra-1, réprimé par RNAi.	94
Figure 28 : Exemple de biais générés par l'utilisation du siRNA contrôle.....	95
Figure 29 : Comparaison des modulations induites par deux siRNA contrôles commerciaux.	95
Figure 30 : Représentation schématique de la machinerie des ARN interférents et des biais générés par l'utilisation de siRNA.....	96
Figure 31 : Génération de siCTL spécifiques pour chacun des trois siFra-1, d'après Buehler et al. 2012.	97
Figure 32 : Analyse en composante principale (PCA).	100
Figure 33 : Croisement du transcriptome siFra-1 avec ses deux conditions contrôles associées pour l'élimination des faux positifs.....	101
Figure 34 : Croisement de chaque transcriptome siRNA spécifique avec ses deux conditions contrôles associées pour l'élimination des faux positifs.	103
Figure 35 : Croisement des listes siFra-1, siFra-2 et siFra-1+siFra-2.	103
Figure 36 : Bilan de l'analyse ontologique en utilisant le logiciel GeneGo metaCore.	105
Figure 37 : Intersection entre les pics Fra-1 et Fra-2 obtenus par ChIP-seq dans la lignée MDA-MB231.	110
Figure 38 : Intersection entre les pics Fra-1, Fra-2 et c-Jun obtenus par ChIP-seq dans la lignée MDA-MB231.	111
Figure 39 : Descriptif des enhanceurs identifiés dans la lignée MDA-MB231.....	113
Figure 40 : Distribution des marques H3K4me1 caractéristique des enhanceurs et H3K4me3 caractéristique des promoteurs autour des pics Fra-1 et/ou Fra-2 ($\pm$ 2kb du centre du pic).....	113
Figure 41 : Distribution des pics Fra-1 et/ou Fra-2 entre les promoteurs ($\pm$ 1,5kb des TSS) et les enhanceurs actifs et inactifs.	114
Figure 42 : Identification des gènes sous le contrôle de Fra-1 et/ou Fra-2 associés à des enhanceurs présentant des pics de fixation pour Fra-1 et/ou Fra-2 dans la lignée MDA-MB231.	116
Figure 43 : Classifications des cancers du sein utilisées par les banques de données.....	122
Figure 44 : Données des banques de cancers du sein CIT et TCGA concernant Fra-1, Fra-2, HMGA1 et CD68.....	124
Figure 45 : Analyse de l'expression des ARNm de HMGA1, CD68, Fra-1 et Fra-2 dans différentes lignées de cancer du sein. Mesure de l'abondance d'ARNm par RT-qPCR.....	125
Figure 46 : Modulation du gène HMGA1 et CD68 en absence de Fra-1 et/ou Fra-2 dans les trois lignées TNBC.	126
Figure 47 : Structure des protéines HMG.....	132
Figure 48 : Représentation schématique des gènes HMGA1 et HMGA2 humains et leurs produits protéiques.....	133
Figure 49 : Alignement des séquences des protéines HMGA1 et HMGA2.....	134
Figure 50 : Mécanismes d'action des protéines HMGA..	136

Figure 51 : Expression des protéines HMGA1 dans le tissu mammaire normal et cancéreux....	138
Figure 52 : Abondances relatives des ARN naissants HMGA1 en absence de Fra-1 et/ou Fra-2 dans les MDA-MB231.....	141
Figure 53 : Cartographie des données de ChIP-seq au niveau du locus de HMGA1 dans la lignée MDA-MB231 wild-type.....	143
Figure 54 : Structure du locus de HMGA1 et position des fragments amplifiés par ChIP-qPCR..	144
Figure 55 : Expression des protéines Fra-1 et/ou Fra-2 après transfection de siRNA dans la lignée MDA-MB231, pour les expériences de ChIP-qPCR.....	145
Figure 56 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour H3K4me1, H3K4me3 et H3K27ac, en présence et en absence de Fra-1 et/ou Fra-2 par transfection des siRNA contrôle ou spécifiques dans la lignée MDA-MB231..	146
Figure 57 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour RNA Pol II et ses formes phosphorylées sur les ser5 et ser2 du CTD, en présence et en absence de Fra-1 et/ou Fra-2 par transfection des siRNA contrôle ou spécifiques dans la lignée MDA-MB231.	148
Figure 58 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour H3K36me3, après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231.	149
Figure 59 : Expression de Pol II et ses formes phosphorylées après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231.	149
Figure 60 : Données de ChIP-qPCR au niveau du locus de S26 pour la forme initiatrice de Pol II et ses formes phosphorylées sur les ser-5 et ser-2 du CTD, après inhibition de Fra-1 ou Fra-2 par siRNA dans la lignée MDA-MB231.	150
Figure 61 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour P300, après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231.....	151
Figure 62 : Structure schématique de la protéine CD68 humaine.....	158
Figure 63 : Abondances relatives des ARN naissants CD68 en absence de Fra-1 et/ou Fra-2 dans les MDA-MB231.....	161
Figure 64 : Abondance de CD68 dans les lignées de TNBC MDA-MB231, MDA-MB436 et SUM159PT.....	163
Figure 65 : Inhibition de l'expression de CD68 dans les lignées de TNBC MDA-MB232 et MDA-MB436.	164
Figure 66 : Mesure de la prolifération des cellules MDA-MB231 après inhibition de CD68 ou Fra-1.....	165
Figure 67 : Mesure de la viabilité cellulaire des cellules MDA-MB231 non-transfectées par cytométrie de flux au moyen du marquage DiOC6(3)/Pi.....	166
Figure 68 : Pourcentage de cellules MDA-MB231 apoptotiques et nécrotiques mesuré par cytométrie de flux au moyen du marquage DiOC6(3)/Pi, après inhibition de CD68 ou Fra-1 par siRNA.	167
Figure 69 : Mesure de la distribution dans le cycle cellulaire des cellules MDA-MB231 non-transfectées par cytométrie de flux au moyen du marquage Pi.	168
Figure 70 : Distribution dans le cycle cellulaire des cellules MDA-MB231 par cytométrie de flux au moyen du marquage Pi, après inhibition de CD68 ou Fra-1 par siRNA..	169

Figure 71 : Analyse de la motilité des cellules MDA-MB231 par la méthode de wound-healing associé à la videomicroscopie, après inhibition de CD68 ou Fra-1 par siRNA...	170
Figure 72 : Évolution de la morphologie en 3D des mammosphères au cours du temps, après inhibition de CD68 ou Fra-1 par siRNA dans les MDA-MB231.....	172
Figure 73 : Suivi de l'abondance des protéines CD68 et Fra-1 jusqu'à 7 jours post-transfection..	173
Figure 74 : Localisation de CD68 dans la lignée MDA-MB231.....	176
Figure 75 : Expériences de co-localisation dans la lignée MDA-MB231, par immunofluorescence sur cellules fixées.....	177
Figure 76 : Représentation schématique du processus autophagique.....	179
Figure 77 : Systèmes de conjugaison Ath12-Atg5 et LC3.	180
Figure 78 : Représentation schématique de la dérégulation du flux autophagique.	181
Figure 79 : Détection des protéines CD68 et LC3, dans les cellules MDA-MB231.s.	184
Figure 80 : Fusion entre les différentes vésicules intracellulaires.	185

TABLE DES TABLEAUX

Tableau 1 : Classification des cancers du sein par stade d'après l'UICC.	26
Tableau 2 : Principales caractéristiques des sous-types moléculaires des cancers du sein.	30
Tableau 3 : Conclusions du rapport INCa-SFSPM de 2009.	32
Tableau 4 : Description du statut pCR/Non-pCR en fonction du sous-typage des tumeurs TNBC.	34
Tableau 5 : AP-1 dans le développement et la tumorigenèse chez la souris.	49
Tableau 6 : Compositions et fonctions des facteurs généraux de transcription chez l'Homme.	66
Tableau 7 : Bilan de la nouvelle analyse transcriptomique genome-wide.	104
Tableau 8 : Bilan du croisement entre nos données transcriptomiques et celles publiées par Zhao.	108
Tableau 9 : Première sélection de 14 gènes induits par Fra-1 et/ou Fra-2 dans les MDA-MB231.	121
Tableau 10 : L'expression de 8/14 gènes induits par Fra-1 et/ou Fra-2 dans la lignée modèle corrèle avec l'expression de Fra-1 et/ou Fra-2 dans les deux banques de cancers du sein.	122
Tableau 11 : Bilan des résultats pour les 8 gènes candidats sélectionnés.	126

TABLE DES ANNEXES

Annexe 1 : Liste des gènes identifiés comme étant induits par Fra-1 et/ou Fra-2 ($FC \leq 2$ dans au moins une condition et $FDR < 0,05$).	197
Annexe 2 : Liste des gènes identifiés comme étant réprimés par Fra-1 et/ou Fra-2 ($FC \geq 2$ dans au moins une condition et $FDR < 0,05$).	198
Annexe 3 : Liste des gènes identifiés comme étant modulés par Fra-1 et/ou Fra-2 ($FC \geq 1,5 / \leq -1,5$ dans au moins une condition et $FDR < 0,05$) et ayant des pics pour Fra-1 et/ou Fra-2 au niveau de leur(s) région(s) enhancer(s).....	199
Annexe 4 : Liste des gènes identifiés comme étant modulés par Fra-1 et/ou Fra-2 ($FC \geq 1,5 / \leq -1,5$ dans au moins une condition et $FDR < 0,05$) et ayant des pics pour Fra-1 et/ou Fra-2 au niveau de leur région promoteur.	200
Annexe 5 : Données des banques de cancers du sein CIT et TCGA concernant Fra-1, Fra-2, HMGA1 et CD68.....	201
Annexe 6 : Cartographie des données de ChIP-seq au niveau du locus de HMGA1 dans la lignée MDA-MB231 wild-type.	202
Annexe 7 : Détection des protéines CD68 et LC3, dans les cellules MDA-MB231... ..	203

AVANT PROPOS

Le cancer du sein est une maladie hétérogène présentant des évolutions et des réponses aux thérapies très variées. Malgré les avancées thérapeutiques qui permettent une meilleure prise en charge des patientes et conduisent à une rémission de certaines tumeurs, les cancers du sein agressifs et métastatiques « Triple Négatifs » ne bénéficient d’aucune thérapie ciblée et restent incurables. L’identification de nouvelles cibles moléculaires diagnostiques et surtout, thérapeutiques constitue donc un enjeu majeur pour le traitement de ces cancers.

Les facteurs de transcription Fra-1 et Fra-2 sont des protéines de la famille multigénique FOS. Elles participent à la formation de la famille ubiquitaire de complexes transcriptionnels dimériques AP-1 (qui impliquent aussi les membres de la famille multigénique JUN) dont l’action est pléiotrope. Malgré plus de 10 000 références concernant AP-1, il n’existe encore que quelques données fragmentaires sur les mécanismes transcriptionnels mis en jeu par les protéines AP-1 pour gouverner l’expression de leurs gènes cibles. De la même manière, l’ensemble des gènes contrôlés par les différents membres AP-1 n’est pas caractérisé. Ceci peut s’expliquer par le fait que l’identification des gènes cibles ne peut se faire dans sa globalité et nécessite la caractérisation des gènes cibles pour chacun des membres AP-1 dans les différents contextes cellulaires. En effet, suivant sa composition, le type cellulaire et les voies de signalisation activées, AP-1 peut activer ou réprimer différentes catégories de gènes.

Fra-1 et Fra-2 sont co-surexprimés de façon aberrante dans de nombreuses tumeurs triple négatives où ils augmentent l’agressivité tumorale en altérant diverses fonctions cellulaires (prolifération, résistance à l’apoptose, invasion, motilité, etc.). Récemment, l’expression de Fra-1 également a été montrée cruciale pour l’établissement des métastases *in vivo*. Cependant, viser Fra-1 ou Fra-2 en tant que cible thérapeutique est exclu pour le moment, car il est impossible d’inhiber ces facteurs de transcription par des approches utilisables en clinique.

De façon intéressante, dans les tumeurs du sein triple négatives, les autres protéines FOS, à savoir c-Fos et FosB, sont peu ou pas exprimées. Ceci constitue un avantage pour rechercher spécifiquement les gènes cibles de Fra-1 et/ou Fra-2, dont certains pourraient constituer des cibles thérapeutiques potentielles, et pour entreprendre l’étude des mécanismes transcriptionnels contrôlés par Fra-1 et/ou Fra-2.

Dans ce contexte, l’objectif général de ma thèse s’est orienté autour de trois axes :

(i) Une étude transcriptomique sur une lignée modèle de cancer du sein triple négatif (MDA-MB231) pour identifier les gènes codants sous le contrôle de Fra-1 et/ou Fra-2. Cette étude a été associée à une caractérisation génomique des sites de fixation de Fra-1 et Fra-2 sur la chromatine pour identifier des cibles directes. Les résultats obtenus ont permis la sélection de gènes pour les études transcriptionnelles et fonctionnelles.

(ii) L'étude des mécanismes transcriptionnels gouvernés par Fra-1 et/ou Fra-2 pour la régulation de leurs gènes cibles. Mes efforts se sont concentrés sur l'un des gènes identifiés au cours de l'approche transcriptomique. Il s'agit du gène HMGA1, choisi en raison de son rôle maintenant bien établi dans l'agressivité des tumeurs épithéliales.

(iii) L'étude fonctionnelle a été focalisée sur l'une des cibles dont l'expression est régulée positivement et avec un fort différentiel par Fra-1 et Fra-2. Il s'agit de CD68, protéine transmembranaire dont le rôle n'est pas clairement établi et dont l'implication dans le phénotype tumoral n'a jamais été étudiée.

Le présent manuscrit est classiquement scindé en trois chapitres. L'introduction présente les éléments indispensables à la justification et à la compréhension facilitée des questions abordées. Les résultats acquis sont présentés sous la forme de trois études qui reflètent les trois axes développés au cours de mon travail de thèse. Chaque étude est suivie d'une courte discussion. Une discussion plus générale est présentée en conclusion du manuscrit.

Comme mentionné dans le manuscrit, mon travail de thèse a été retardé par différents aléas qui ont été chronophages et qu'il a fallu contourner. Ma première étude transcriptomique a fait apparaître que les biais liés au siRNA contrôle étaient importants, bien que non mentionnés dans les études transcriptomiques publiées. La résolution de ces biais a nécessité des mises au point, une deuxième étude transcriptomique et des analyses différentielles. Une partie de l'étude des mécanismes transcriptionnels et du séquençage à haut débit était basée sur l'utilisation d'anticorps polyclonaux de lapin. L'arrêt subit de la distribution de ces anticorps par Santa Cruz Biotechnology en fin 2015, suite à des problèmes sanitaires de leur animalerie, nous a contraint à devoir tester de nouveaux anticorps dont un bon nombre ne s'est pas révélé satisfaisant. Enfin, alors qu'une extinction transitoire mais complète de CD68 est observée par transfection de siRNA, l'extinction stable par transfection des shRNA correspondants n'a pu être obtenue et a constitué un obstacle aux études, menées en fin de thèse, nécessitant une extinction sur le long terme.

INTRODUCTION

INTRODUCTION

I. LE CANCER DU SEIN

I.1. Quelques chiffres

Avec 53 000 nouveaux cas estimés en France en 2011, le cancer du sein est le cancer le plus fréquent chez la femme (33% de l'ensemble des cancers féminins). De plus, avec 11 500 décès estimés, il est également la première cause de mortalité liée au cancer chez la femme (18,3% de l'ensemble des décès féminins par cancer). On estime à l'heure actuelle qu'une femme sur 9 développera un cancer du sein au cours de sa vie et qu'une sur 27 en mourra, ce qui souligne l'importance de cette pathologie en terme de santé publique. Il est à noter que les hommes peuvent également développer un cancer du sein. Ces cas sont cependant rares, puisqu'ils représentent seulement 1% de l'ensemble des cancers chez l'homme (*Institut National Du Cancer* : <http://www.e-cancer.fr>).

A l'heure actuelle, la mammographie est le meilleur moyen pour détecter le cancer du sein dans ses stades précoces. En moyenne, la tumeur peut être détectée 1,7 ans avant que la femme ne ressente une grosseur. Aux stades précoces d'une tumeur localisée, les chances de survie à 5 ans sont de 95%. Ces chances diminuent aux stades tardifs : elles sont inférieures à 50% quand la tumeur a disséminé dans les ganglions lymphatiques et inférieure à 20% quand elle a disséminé dans des organes distants. Le dépistage précoce des lésions cancéreuses, la chirurgie, avec l'ablation sélective de la tumeur, et les différentes thérapies (chimiothérapie, radiothérapie, hormonothérapie ou autres thérapies ciblées) ont contribué à diminuer sensiblement la mortalité due au cancer du sein. Cependant, malgré ces progrès certains types de cancer du sein agressifs et métastatiques sont difficiles à traiter et restent aujourd'hui incurables.

I.2. Histologie du sein

Le sein est une glande exocrine dont la fonction biologique est de produire du lait afin de nourrir le nouveau-né. La glande mammaire est constituée de deux compartiments cellulaires : le compartiment épithélial (essentiellement composé de graisse, de tissu conjonctif, des lobes et des canaux) et le compartiment mésenchymateux (vaisseaux sanguins et lymphatiques).

Le compartiment épithélial s'articule autour d'un réseau de lobes, chaque lobe étant constitué de plusieurs lobules, constitués eux-mêmes d'acini qui fabriquent le lait. Il existe entre quinze et vingt lobes dans le sein, qui se jettent dans les canaux galactophores, ces derniers venant aboutir au mamelon par cinq à neuf orifices (**Figure 1.A**). Les lobules et les canaux sont composés de deux types cellulaires différenciés, les cellules luminales et les cellules myoépithéliales, ainsi

qu'un faible nombre de cellules progénitrices (cellules souches et précurseurs immatures) (**Figure 1.B**). Les cellules luminales sont en contact avec la lumière des lobules et des canaux, elles expriment les cytokératines CK8/18 et CK19. Les cellules myoépithéliales entourent les cellules luminales et sont en contact avec la membrane basale et le stroma environnant, elles expriment les cytokératines CK5/6, CK14, CK17 et des marqueurs musculaires lisses comme l'actine musculaire lisse et p63.

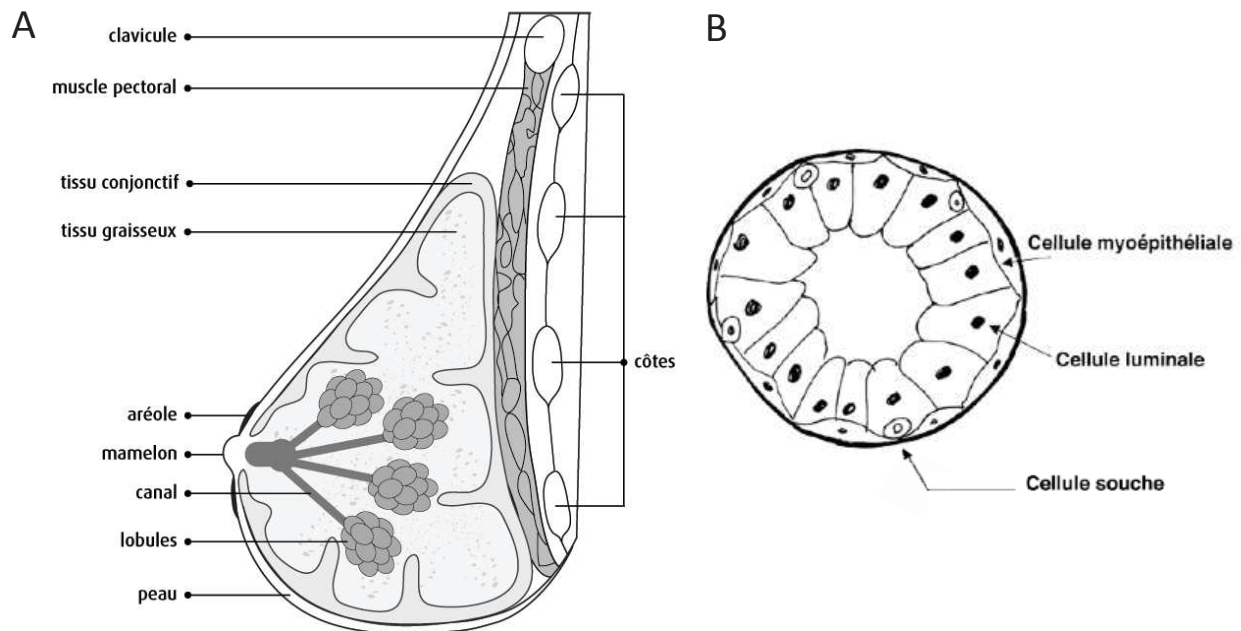


Figure 1 : Représentation du sein. A) Schéma de l'anatomie du sein. Le sein est formé de 15 à 20 lobes, qui sont séparés entre eux par du tissu graisseux et des bandes de tissu conjonctif dans lesquels courent les vaisseaux sanguins et les vaisseaux lymphatiques. Les lobes, qui ont pour fonction de sécréter le lait, sont reliés au mamelon par des canaux galactophores, qui conduisent le lait à l'extérieur du sein. Chaque lobe est formé de plusieurs lobules composés d'une multitude d'acini (partie sécrétrice de la glande) qui se regroupent de façon très dense autour d'un canal alvéolaire. Plusieurs canaux alvéolaires se réunissent et forment un canal lobulaire qui se déverse dans les canaux galactophores. C'est au niveau des cellules épithéliales qui tapissent l'intérieur de ces différents canaux que se développent la plupart des cancers du sein. B) Coupe transversale d'un canal de glande mammaire. Le schéma montre les différentes cellules épithéliales entourées de stroma. On distingue trois types de cellules : les cellules luminales en contact avec la lumière, les cellules myoépithéliales en contact avec la membrane basale et les cellules progénitrices (cellules souches et précurseurs immatures) en position basale ou supra-basale. *Source :* <http://www.cancer.ca>.

Au sein de toutes ces structures circulent les vaisseaux sanguins (artères et veines) et les vaisseaux lymphatiques, ils constituent le compartiment mésenchymateux. Les vaisseaux lymphatiques recueillent et transportent la lymphe jusqu'aux ganglions lymphatiques. Ils drainent les excès de liquide présents dans les lésions et participent à la détoxification des organes. Ils participent également à la circulation des nutriments et des globules blancs. Ce sont des éléments essentiels du système immunitaire. Dans les cas de cancer du sein, les cellules tumorales issues du compartiment épithélial et à potentiel métastatique peuvent se détacher de la tumeur primaire et pénétrer dans les vaisseaux lymphatiques. Elles sont alors véhiculées par le flux liquidien jusqu'aux ganglions lymphatiques (**Figure 2**). L'envahissement des ganglions par

les cellules cancéreuses est la première étape vers le développement de métastases. La présence ou l'absence de cellules malignes au niveau du ganglion sentinelle (premier ganglion drainant les vaisseaux lymphatiques) conditionne le pronostic du cancer et son traitement.

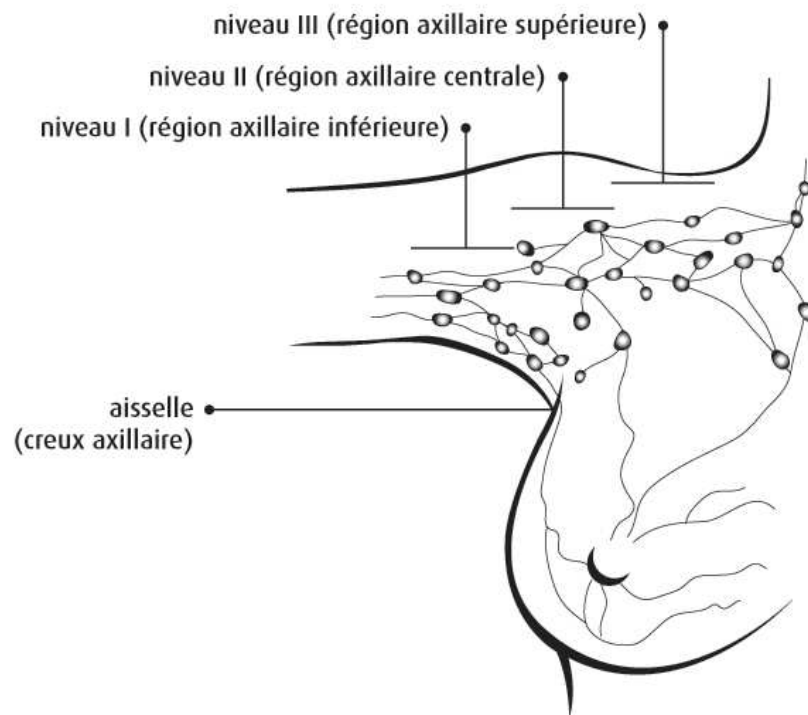


Figure 2 ; Système lymphatique du sein. Les ganglions lymphatiques du sein sont surtout situés : à l'axilla (ganglions axillaires), au-dessus de la clavicule (ganglions sus-claviculaires), sous la clavicule (ganglions sous-claviculaires) ou autour du sternum (ganglions mammaires internes). L'axilla contient 30 à 50 ganglions lymphatiques. Ils sont répartis en 3 niveaux selon leur proximité au muscle pectoral situé dans le thorax, soit : niveau I (située dans la partie la plus basse ou inférieure de l'axilla), niveau II (située dans la partie centrale de l'axilla) et niveau III (située sous la partie centrale de la clavicule et tout près d'elle). *Source : <http://www.cancer.ca>.*

Contrairement aux autres organes qui sont formés durant l'embryogenèse et qui conservent ensuite leur architecture tout le long de la vie adulte, la glande mammaire est en évolution permanente tout au long de la vie de la femme : du stade foetal à la ménopause, en passant par le stade pré-pubère, pubère, la grossesse et la lactation (**Figure 3**). Tous ces stades sont connectés au développement sexuel et à la reproduction. Ces changements ont lieu sous l'influence des hormones sexuelles d'origine ovarienne (oestrogène et progestérone) et d'un certain nombre de facteurs de croissance.

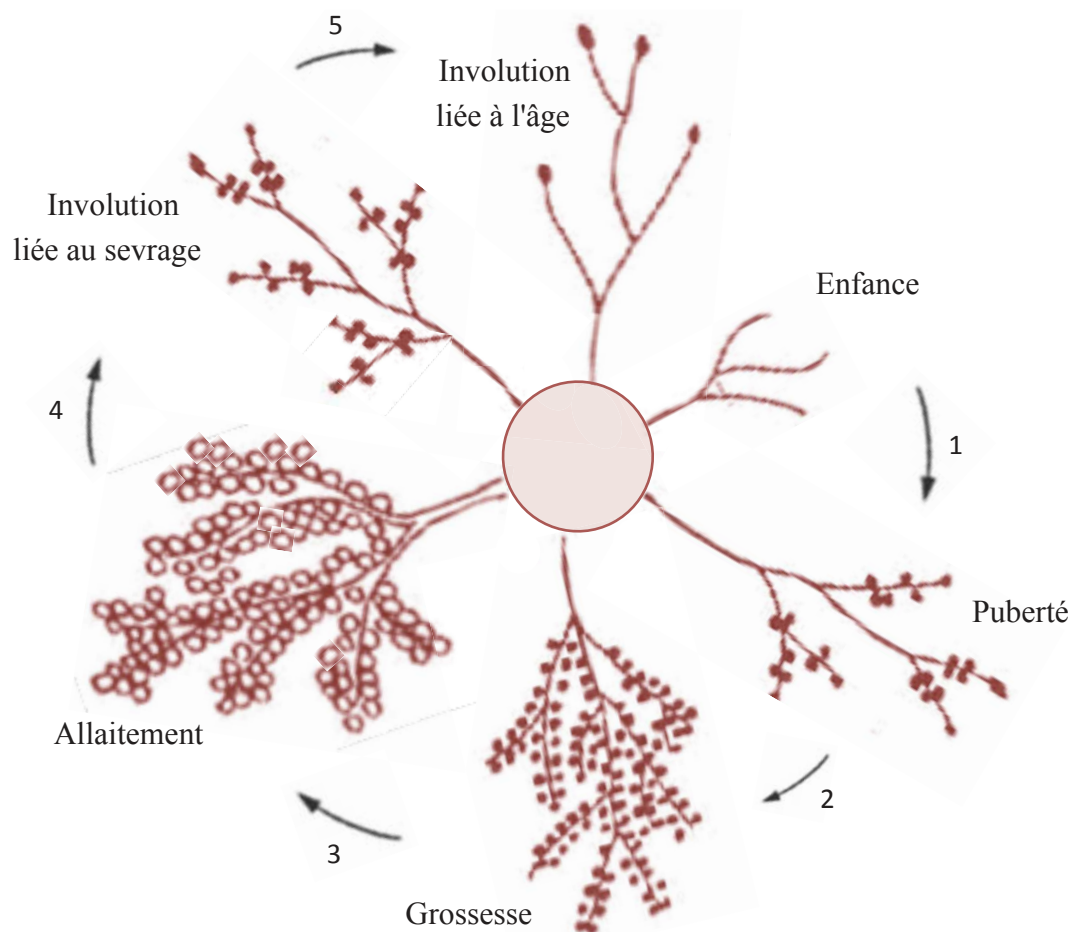


Figure 3 : Représentation schématique du développement de la glande mammaire. (1) Durant la puberté, l'augmentation des taux d'oestrogènes, d'hormones de croissance et d'IGF-I provoque la prolifération de l'épithélium mammaire qui se développe en plusieurs canaux. Cette prolifération continue donne naissance à des alvéoles. (2) Pendant la grossesse les taux élevés d'oestrogène et de progestérone stimule la prolifération épithéliale des cellules de l'alvéole et cela entraîne la formation d'environ 80 lobules au niveau de chaque alvéole. (3) Un fois les cellules épithéliales lobulaires complètement différenciées, elles sont capables de produire du lait. (4) À l'arrêt de la lactation la glande mammaire régresse, on parle d'involution liée au sevrage. (5) Enfin, durant la ménopause les cellules épithéliales s'atrophient d'avantage par manque de soutien hormonal, on parle alors d'involution liée à l'âge. D'après <http://www.vetopsy.fr/reproduction/lactation/mammogenese-puberte.php>.

I.3. Le cancer du sein en général

Comme décrit précédemment, la glande mammaire est en remodelage permanent tout au long de la vie de la femme sous l'influence d'hormones et de facteurs de croissance. Les cancers du sein les plus fréquents (95%) sont des adénocarcinomes qui se développent après la transformation des cellules épithéliales de la glande mammaire. 75% d'entre eux se développent au niveau des canaux galactophores (cancer canalaire) et 25% au niveau des lobes (cancer lobulaire).

Il existe deux grandes catégories de cancer du sein selon l'étendue de la tumeur. Les cancers *in situ*, qui restent confinés à l'intérieur des canaux galactophores ou des lobes sans envahissement des tissus voisins. Ces tumeurs représentent 25% des cas de cancer du sein, ils

sont généralement à bon pronostic et essentiellement traités par une intervention chirurgicale. Dans 75% des cas, les cellules cancéreuses envahissent les tissus voisins, on parle alors de cancers *invasifs* ou *infiltrants*. Les cellules cancéreuses peuvent s'échapper de la tumeur primaire et passer par les vaisseaux sanguins et lymphatiques pour s'accumuler dans les ganglions les plus proches (ganglions auxiliaires) et atteindre des organes distaux, qu'elles peuvent coloniser pour former des tumeurs secondaires ou métastases (**Figure 4**). C'est cette capacité de s'étendre à d'autres tissus et organes qui fait du cancer une maladie potentiellement mortelle.

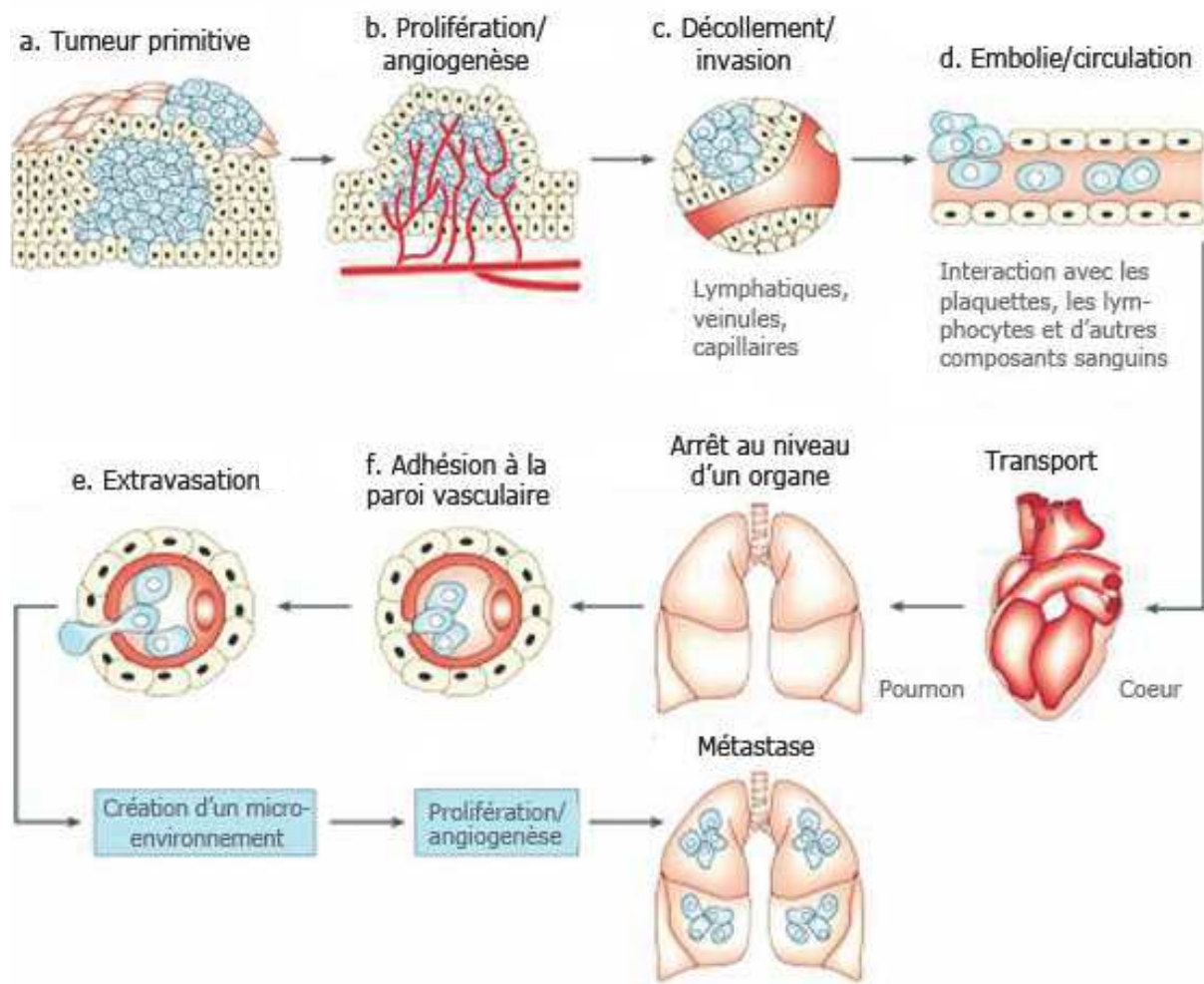


Figure 4 : Principales étapes de la formation d'une métastase. Source : <http://reflexions.ulg.ac.be/>.

De manière générale et simplifiée, la tumorigenèse est un processus multi-étape issu d'une succession d'altérations génétiques qui conduisent à la transformation cellulaire. La progression vers un phénotype agressif implique l'acquisition d'avantages sélectifs permettant la survie, la prolifération et la dissémination (Hanahan and Weinberg, 2000). Les cellules à potentiel métastatique ont acquis, au fil des modifications génétiques et de leur sélection, des capacités fonctionnelles leur permettant de ne pas dépendre de signaux de prolifération pour se diviser, d'être indépendantes des signaux de survie, de ne plus répondre aux signaux d'arrêt, d'être résistantes à l'apoptose, de proliférer de manière illimitée, d'envahir les tissus environnant, de

former des métastases et de promouvoir l'angiogenèse pour l'irrigation des tumeurs. Ces gains de fonction dépendent de l'acquisition par la cellule de caractéristiques permissives, à savoir l'instabilité génétique et la capacité à induire un processus inflammatoire (Hanahan and Weinberg, 2000, 2011). Le processus inflammatoire est issu de l'infiltration des tumeurs par les cellules du système immunitaire qui, de façon paradoxale, peuvent contribuer à favoriser la progression tumorale par l'apport dans le microenvironnement de la tumeur de cytokines, de facteurs de croissance, de facteurs angiogéniques ainsi que d'espèces oxygénées réactives qui sont mutagènes pour les cellules cancéreuses avoisinantes. Récemment, deux nouvelles caractéristiques des cellules cancéreuses ont été mises en évidence : un changement important dans leur métabolisme énergétique leur permettant de supporter la croissance et la prolifération cellulaire, ainsi que l'acquisition de capacités à échapper à la reconnaissance et à la destruction par le système immunitaire (**Figure 5**).

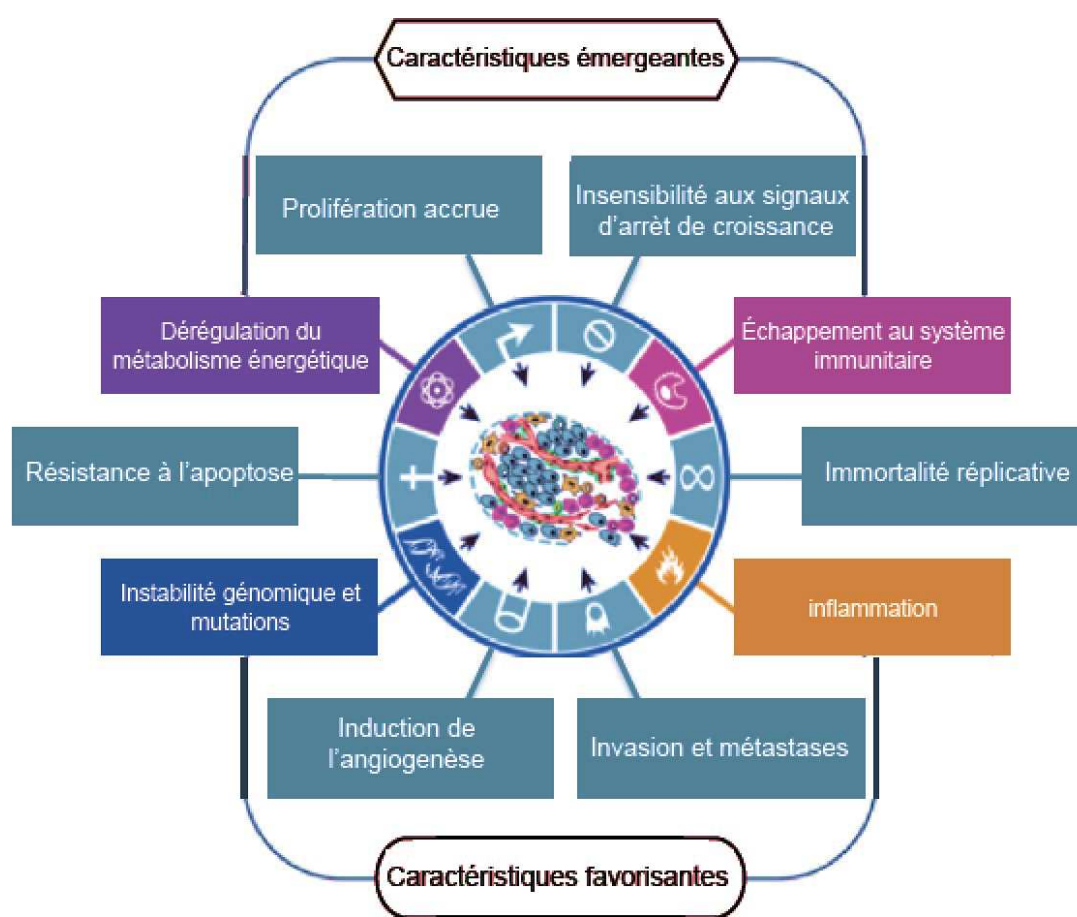


Figure 5 : Caractéristiques émergentes et facteurs favorisants. Cette illustration englobe les six caractéristiques du cancer (Hallmarks of cancer) (en gris) initialement proposées par Hanahan et Weinberg (2000), auxquelles s'ajoutent de nouvelles caractéristiques émergentes et favorisantes qui interviendraient dans la pathogenèse des cancers. Ces dernières impliquent (i) la capacité à modifier ou à reprogrammer le métabolisme énergétique de la cellule, afin de supporter la prolifération néoplasique, (ii) la capacité à éviter la destruction par le système immunitaire, (iii) l'instabilité génétique, qui dote la cellule cancéreuse d'altérations géniques favorisant la progression tumorale, ainsi que (vi) l'inflammation induite par les cellules du système immunitaire inné qui confèrent un terrain favorable à la progression tumorale. D'après Hanahan and Weinberg, (2000, 2011).

On trouve quatre sites préférentiels de métastases pour le cancer du sein : le poumon, le foie, l'os et le cerveau. Un des mécanismes rapportés pour être mis en jeu dans ce tropisme est la production par les cellules tumorales du sein du récepteur de chimiokine CXCR4. En effet, CXCR4 n'est pas exprimé dans le tissu épithélial mammaire sain alors qu'il est fortement détecté dans les cellules métastatiques. Qui plus est CXCL12 (SDF1), le ligand de CXCR4 exprimé par de très nombreux tissus, l'est à des taux particulièrement élevés dans trois des quatre tissus cibles des métastases de cancer du sein (poumon, foie et moelle osseuse) ainsi que dans les ganglions lymphatiques, passage obligé lors de la dissémination par voie lymphatique (Borsig et al., 2014; Müller et al., 2001).

I.4. Classification des cancers mammaires

Une fois le diagnostic de cancer confirmé, la stratégie thérapeutique est guidée par le facteur pronostique qui dépend du stade et du grade de la tumeur ainsi que du statut hormonal et du statut HER2 des cellules tumorales.

I.4.a. Classification sur critères anatomopathologiques

Le système de classification couramment utilisé pour déterminer le stade de la tumeur est la classification TNM, de l'Union Internationale Contre le Cancer (UICC) et de l'American Joint Committee on Cancer (AJCC), 7ème édition. TNM signifie en anglais « Tumor, Nodes, Metastasis » soit « tumeur, ganglions, métastases ». Comme son nom l'indique, cette classification tient compte de la taille de la tumeur primaire, de l'absence ou présence d'envahissement ganglionnaire ainsi que de leur emplacement et du caractère infiltrant de la tumeur via la présence ou non de métastases à distance (**Tableau 1**).

L'agressivité des tumeurs est ensuite évaluée par analyse histologique d'un échantillon de tumeur. Cet examen permet de définir le type de cancer et son grade. Le système de classification couramment utilisé pour déterminer le grade de la tumeur est la classification SBR (Scarff, Bloom et Richardson). Elle est basée sur trois paramètres morphologiques : le degré de différenciation des cellules tumorales (architecture cellulaire), la forme et la taille du noyau de la cellule, ainsi que le nombre de cellules en division (index mitotique) qui reflète la vitesse à laquelle les cellules cancéreuses se multiplient. La somme des notes obtenues pour chacun de ces trois critères permet ensuite de déterminer le grade de la tumeur évalué de I à III. Le grade de la tumeur est la fois un marqueur pronostique, car plus le grade est élevé plus la tumeur est agressive et de mauvais pronostic, et un marqueur prédictif, car les tumeurs différenciées répondent mieux à l'hormonothérapie tandis que les tumeurs moins différenciées répondent mieux à la chimiothérapie.

Stadification de l'UICC – Cancer du sein

Stade UICC	TNM		Explication
Stade 0	Tis	N0	M0 Cancer in situ – Cancer confiné aux canaux, aux lobules ou au mamelon, qui ne s'est pas propagé au tissu mammaire voisin. Le cancer ne s'est pas propagé aux ganglions lymphatiques. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade I	T1	N0	M0 La tumeur mesure 2 cm ou moins de diamètre. Le cancer ne s'est pas propagé aux ganglions lymphatiques. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade IIA	T0	N1	M0 Absence de tumeur dans le sein Le cancer est présent dans 1 à 3 ganglions axillaires. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T1	N1	M0 La tumeur mesure 2 cm ou moins de diamètre. Le cancer s'est propagé dans 1 à 3 ganglions axillaires, dans les ganglions mammaires internes ou bien dans les deux. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T2	N0	M0 La tumeur mesure plus de 2 cm mais pas plus de 5 cm de diamètre. Le cancer ne s'est pas propagé aux ganglions lymphatiques. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade IIB	T2	N1	M0 La tumeur mesure plus de 2 cm mais pas plus de 5 cm de diamètre. Le cancer s'est propagé dans 1 à 3 ganglions axillaires, dans les ganglions mammaires internes ou bien dans les deux. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T3	N0	M0 La tumeur mesure plus de 5 cm de diamètre. Le cancer ne s'est pas propagé aux ganglions lymphatiques. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade IIIA	T0	N2	M0 Absence de tumeur dans le sein Le cancer est présent dans 4 à 9 ganglions axillaires ou dans les ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T1	N2	M0 La tumeur mesure 2 cm ou moins de diamètre. Le cancer s'est propagé dans 4 à 9 ganglions axillaires ou dans les ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T2	N2	M0 La tumeur mesure plus de 2 cm mais pas plus de 5 cm de diamètre. Le cancer s'est propagé dans 4 à 9 ganglions axillaires ou dans les ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T3	N1, N2	M0 La tumeur mesure plus de 5 cm de diamètre. Le cancer s'est propagé dans 1 à 9 ganglions axillaires ou dans les ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade IIIB	T4	N0, N1, N2	M0 La tumeur s'est propagée à la paroi thoracique ou à la peau. L'un des énoncés suivants s'applique : - Le cancer ne s'est propagé dans aucun ganglion lymphatique. - Le cancer s'est propagé dans 1 à 9 ganglions axillaires. - Le cancer peut ou non s'être propagé dans les ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
Stade IIIC	tout T	N3	M0 La tumeur est de n'importe quelle taille. L'un des énoncés suivants s'applique : - Le cancer s'est propagé dans au moins 10 ganglions axillaires. - Le cancer s'est propagé dans au moins 1 ganglion infra-claviculaire ou sus-claviculaire. - Le cancer s'est propagé dans plus de 3 ganglions axillaires et aux ganglions mammaires internes. Le cancer ne s'est pas propagé vers des emplacements éloignés.
	T4d		Le cancer inflammatoire du sein est de stade III (IIB ou IIIC), sauf s'il s'est propagé vers des emplacements éloignés ou dans des ganglions lymphatiques situés loin du sein, auquel cas il est de stade IV.
Stade IV	tout T	tout N	M1 Tumeur de n'importe quelle taille. N'importe quel degré d'atteinte ganglionnaire. Le cancer s'est propagé vers des emplacements éloignés comme les os, le foie, les poumons, le cerveau ou des ganglions lymphatiques situés loin du sein.

Tableau 1 : Classification des cancers du sein par stade d'après l'UICC. Source : <http://www.cancer.ca/fr-ca>.

I.4.b. Classification sur critères moléculaires

En complément de ces critères anatomo et histopathologiques, les oncologues utilisent une classification moléculaire pour mieux caractériser les sous-types tumoraux, affiner le pronostic et adapter au mieux les traitements. Cette classification repose sur l'analyse de trois marqueurs moléculaires : deux récepteurs hormonaux qui permettent de définir le statut hormonal de la tumeur ainsi que le récepteur de type 2 à l'EGF (HER2).

Le statut hormonal (RH) est basé sur la présence ou l'absence des récepteurs aux oestrogènes (ER) et à la progestérone (PR) à la surface des cellules cancéreuses. L'intérêt de la stratification RH a été démontré depuis de nombreuses années du fait de son implication thérapeutique. Environ 70% des tumeurs mammaires diagnostiquées sont ER+. La tumeur est considérée ER+ quand plus de 10% des cellules cancéreuses expriment les récepteurs aux oestrogènes. L'utilisation de traitements anti-hormonaux à base de Tamoxifène (molécule anti-oestrogénique) depuis 1970, puis des inhibiteurs d'aromatase et des SERM (Selective Estrogen Receptor Modulators) a révolutionné le traitement hormonal du cancer du sein qui, au préalable, reposait essentiellement sur la suppression ovarienne par chirurgie ou radiothérapie (Obiorah and Jordan, 2011). Bien qu'il n'existe aucun anti-progestatif utilisé en thérapie, l'analyse de l'expression du PR joue un rôle important pour le pronostic. Le PR est considéré comme un marqueur de la fonctionnalité des ER, car l'expression des PR est régulée par les ER. Ainsi il est rarement observé chez les tumeurs n'exprimant pas ER (Hopp et al., 2004).

Le statut HER2 est basé sur la surexpression ou non du récepteur de type 2 à l'EGF (Epidermal Growth Factor) appelé HER2. Bien qu'aucun ligand n'ait été identifié pour HER2, il joue un rôle central en étant le partenaire de dimérisation privilégié des autres récepteurs à l'EGF ayant fixé leur ligand (Holbro et al., 2003). Dans les tumeurs dites HER2+, il y a une forte surexpression de HER2 liée à l'amplification du gène. Cette forte concentration en HER2 entraîne sa dimérisation spontanée et l'activation accrue des cascades de signalisation situées en aval et impliquées dans la prolifération et la survie cellulaire. Environ 15% des cancers du sein surexpriment HER2 et définissent un sous-type tumoral à pronostic défavorable. Le traitement au Trastuzumab (Herceptin) est le traitement clé chez les patientes HER2+. Cet anticorps est conçu pour se fixer de façon spécifique aux protéines HER2 et bloquer leur fonctionnement, ce qui a pour effet d'inhiber la croissance des tumeurs.

La classification moléculaire, basée sur le statut RH et le statut HER2, conduit à considérer quatre sous-groupes : RH+/HER2+, RH+/HER2-, RH-/HER2+ et RH-/HER2-. Trois groupes sur quatre peuvent donc bénéficier de l'hormonothérapie et/ou du traitement au Trastuzumab. En revanche, le dernier groupe RH-/HER2-, aussi appelé « Triple Négatif » ou TNBC » car ER-/PR-/HER2-, ne bénéficie actuellement d'aucune thérapie ciblée. Les tumeurs triple négatives représentent 15 à 20% des tumeurs mammaires et elles sont agressives et hautement

métastatiques. Malgré de nouveaux schémas de chimiothérapie, ces cancers au pronostic très sombre restent incurables. L'identification de nouvelles cibles thérapeutiques constitue donc un enjeu majeur pour le traitement de ces cancers.

1.4.c. Classification sur critères transcriptomiques

L'utilisation des marqueurs histologiques et moléculaires précédemment cités permet de classer les tumeurs selon leur agressivité et l'expression de certains marqueurs, afin de choisir les traitements les mieux adaptés. Cependant, ces marqueurs ont des limitations et ne reflètent pas l'hétérogénéité moléculaire et cellulaire de la maladie, ni les altérations génétiques sous-jacentes.

Au début des années 2000 et grâce au développement des approches transcriptomiques à grande échelle, plusieurs études se sont basées sur l'analyse des profils d'expression des gènes des tumeurs pour essayer d'obtenir une meilleure stratification du risque des patientes de manière à optimiser les traitements (rapport bénéfice/toxicité) (Perou et al., 2000; Sorlie et al., 2001). L'ensemble de ces études a conduit à une classification en cinq sous-groupes sur la base de l'expression d'un panel de gènes sélectionnés. Les cinq sous-groupes sont : luminal A, luminal B, basal-like, HER2+ et normal breast-like (**Tableau 2**). Ces études ont également montré que ces différents sous-groupes ont des devenir cliniques différents, les sous-groupes basal-like et HER2+ étant associés aux plus mauvais pronostics. Plusieurs autres études ont renforcé ce classement en sous-groupes moléculaires ainsi que la corrélation entre sous-groupe et temps de survie sans rechute (Yersal and Barutca, 2014).

Les différents sous-groupes de cancers du sein sont présentés ci-dessous de façon simplifiée :

** Profil luminal :* Ces tumeurs regroupent approximativement 75% des cancers du sein et peuvent être divisées en deux groupes. Le sous-type luminal A est le plus répandu et représente 50 à 60% de tous les cancers. Ces tumeurs sont généralement de bon pronostic et caractérisées par un grade histologique faible, l'expression d'un haut niveau de ER et des cytokératines de la lumière épithéliale (CK8 et 18). Le sous-type luminal B représente 15 à 20% des cancers du sein. Il est caractérisé par un phénotype plus agressif et un pronostic moins favorable que le luminal A. La différence majeure entre ces deux sous-types est l'expression plus grande des gènes impliqués dans la prolifération (Myb, CCNE1, ect.) chez les luminal B. Cependant, cet aspect est difficilement utilisable en clinique car l'expression des gènes de prolifération forme un continuum entre luminal A et luminal B et il est difficile de choisir où placer le seuil. A l'heure actuelle, l'index Ki67 est proposé comme un marqueur de prolifération qui permettrait de différencier les tumeurs lumineales A des lumineales B.

* Profil basal-like : Ce profil, caractérisé par un haut grade histologique, des index mitotiques et prolifératifs élevés, une grande instabilité génétique et un taux élevé de mutation p53, est retrouvé dans 10 à 20% des cancers du sein. Ces tumeurs présentent des réponses variables aux chimiothérapies, elles sont agressives et de mauvais pronostic, avec un taux important de métastatisation dans les poumons et le cerveau. Les tumeurs appartenant au groupe basal-like n'expriment pas les récepteurs ER et PR et ne surexpriment pas HER2, en revanche elles expriment des lamines, EGFR et de hauts niveaux des marqueurs des cellules myoépithéliales, tels que les cytokératines 5/6, 14 et 17. Les appellations basal-like et TNBC sont souvent considérées comme synonymes mais elles ne le sont pas complètement. Les données cliniques, de transcriptome et d'immunohistochimie ont montré que bien qu'il existe un grand recouvrement entre tumeurs TNBC et basal-like, il n'est pas total : plus de 20% des tumeurs basal-like ne sont pas des TNBC. Les cancers basal-like et TNBC sont donc des sous-types distincts. Par ailleurs, 60 à 70% des cancers du sein de patientes présentant une mutation familiale du gène BRCA1 ont un aspect morphologique de type basal-like ou TNBC (**Figure 6**) (Foulkes et al., 2010; Kreike et al., 2007; Oakman et al., 2010; Stover et al., 2016).

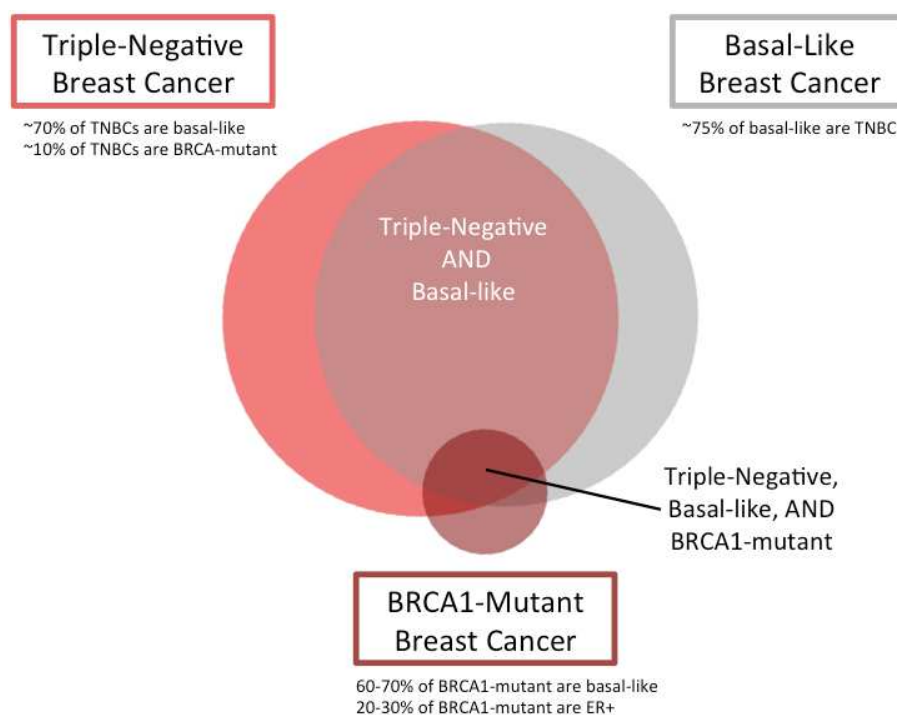


Figure 6 : Chevauchement entre les cancers du sein TNBC, Basal-like et la voie BRCA1 mutée. *Source : Stover et al. (2016).*

* Profil HER2+ : Ces tumeurs, qui surexpriment à un haut niveau HER2, représentent 10 à 20% des cancers du sein. Ce profil définit un groupe de tumeurs hautement prolifératives avec un haut grade histologique et un mauvais pronostic. Il est à noter qu'environ 50% des tumeurs HER2+ expriment également ER, mais à un bas niveau.

** Profil normal breast-like :* Ces tumeurs représentent 5 à 10% des carcinomes mammaires et expriment des gènes caractéristiques du tissu adipeux. Le phénotype de ces tumeurs est mal caractérisé et leur pronostic est généralement intermédiaire entre les cancers luminaux et basal-like. De plus, leur existence a été remise en question : il a été proposé qu'elles puissent résulter d'un artéfact lié à la contamination des cellules tumorales par le tissu normal environnant.

Sous-types Moléculaires des cancers du sein généralement acceptés				
Sous-type moléculaire	Gènes représentatifs du sous-type	Marqueurs caractéristiques du sous-type	Grade histologique	Pronostique associé
Luminal A	Récepteur aux œstrogènes (ESR1)	ER ⁺ et/ou PR ⁺ , HER2 ⁻ , Ki-67 peu abondant.	Bas grade	Sensible à l'hormonothérapie. Réponse à la chimiothérapie variable. Généralement de bon pronostic.
Luminal B	Récepteur aux œstrogènes (ESR1)	ER ⁺ et/ou PR ⁺ , HER2 ⁺ .	Grade plus élevé que le sous-type Luminal A	Sensible à l'hormonothérapie. Réponse à la chimiothérapie variable. De moins bon pronostic que le sous type Luminal A.
Basal-like	Kératine 5 (KRT5) Kératine 17 (KRT17) Lumière, γ 2 (LAMC2)	ER ⁻ , PR ⁻ , HER2 ⁻ , CK 5/6 ⁺ et/ou EGFR ⁺ .	Grade élevé	Insensible à l'hormonothérapie. Réponse à la chimiothérapie variable. Généralement de mauvais pronostic.
HER2	HER2 (<i>ERBB2</i>), amplification du locus <i>ERBB</i>	ER ⁻ , PR ⁻ , HER2 ⁺ .	Grade élevé	Répond à la thérapie ciblée (trastuzumab). Réponse à la chimiothérapie variable. Généralement de mauvais pronostic

Tableau 2 : Principales caractéristiques des sous-types moléculaires des cancers du sein. Adaptées d'après Gruver *et al.* (2011).

Au cours de ces dernières années, de nouveaux sous-groupes sont venus enrichir et complexifier cette classification. En 2012, Guedj *et al* ont analysé le transcriptome de 537 tumeurs primaires sur des puces Affimetrix et ont par la suite défini une loi de classification qui a ainsi permis d'identifier six sous-groupes moléculaires, qui exclue le sous-groupe HER2⁺ et fait apparaître le sous-groupe mApo (molecular-Apocrine) comme décrit précédemment par Farmer *et al.* Les tumeurs mApo sont caractérisées par l'absence des récepteurs aux oestrogènes et par l'expression du récepteur aux androgènes. De plus, le groupe mApo est plus riche que les autres en tumeurs surexprimant HER2 (Farmer *et al.*, 2005; Guedj *et al.*, 2012). Un autre sous-groupe appelé claudin-low, car exprimant de très faibles niveaux des gènes impliqués dans les jonctions serrées et le contact cellule-cellule (notamment claudines 3/4/7 et E-Cadhérine), a été mis en évidence. Les tumeurs de type claudin-low se caractérisent par une forte expression des gènes impliqués dans la transition épithélio-mésenchymateuse. Ce sous-groupe récent n'est pas admis par tous et est encore largement discuté (Perou, 2011; Prat *et al.*, 2010).

Les taxonomies moléculaires établies par études transcriptomiques sont importantes pour caractériser les altérations des programmes génétiques des tumeurs et leur complexité, pour mieux comprendre la biologie de ces tumeurs et pour identifier de nouvelles cibles thérapeutiques potentielles. Cependant, elles présentent plusieurs limitations qui ne les rendent

pas, pour l'instant, directement applicables en routine au niveau clinique. Ces limitations sont notamment liées au choix et au nombre de gènes considérés dans chaque analyse qui varient en fonction des études, entraînant une reproductibilité inter-laboratoires qui n'est pas parfaite. Par ailleurs, si ces classements sont intéressants au niveau statistique du fait du nombre de gènes analysés, ils ne fonctionnent que sur des cohortes importantes de patientes et ne sont pas suffisamment fiables ou discriminants au niveau individuel.

En pratique, au niveau clinique, la taxonomie moléculaire utilisée pour orienter les choix thérapeutiques repose essentiellement sur les statuts ER, PR, HER2 et Ki67 des tumeurs. Cependant, pour promouvoir la désescalade thérapeutique et éviter les chimiothérapies adjuvantes quand elles ne sont pas nécessaires, des tests cliniques basés sur l'expression de panels restreints de gènes ont vu le jour. Ces tests proposent des classifications pronostiques et concernent essentiellement les patientes atteintes de cancers ER+ sans invasion ganglionnaire, car elles représentent un groupe hétérogène de réponses aux hormonothérapies et aux chimiothérapies. Deux de ces tests « MammaPrint » et « Oncotype DX » ont été respectivement approuvés par la FDA (Food and Drug Administration) ou recommandés par l'ASCO (American Society of Clinical Oncology) et la NCCN (National Comprehensive Cancer Network). Par ailleurs, le dosage de uPA et de son inhibiteur PAI-1, qui jouent un rôle important dans l'invasion tumorale via la prolifération, la dégradation de la matrice extracellulaire, l'adhésion et la migration cellulaires, a été validé pour avoir une valeur pronostique sur les risques de rechute.

Cependant, ces tests n'ont pour l'instant pas obtenu l'autorisation de mise sur le marché (AMM) en France. En 2009, l'INCa (Institut National du Cancer) et la SFSPM (Société Française de Sénologie et de Pathologie Mammaire) se sont associés pour évaluer les biomarqueurs tissulaires MammaPrint®, Oncotype DX™ et uPA/PAI-1. Les conclusions faites dans le rapport indiquent que les études publiées ne sont pas considérées comme suffisamment robustes pour guider la décision d'une chimiothérapie adjuvante (**Tableau 3**). La mise à jour parue en 2013, incluant les études publiées entre 2009 et 2013 n'a pas modifié les conclusions faites en 2009.

L'identification de nouveaux marqueurs permettant une meilleure stratification du pronostic individuel est donc toujours d'actualité, ainsi que l'identification de nouvelles cibles thérapeutiques.

	Valeur pronostique	Valeur prédictive	Technique
uPA/PAI-1	Niveau de preuve élevé (LOE I)	Niveau de preuve intermédiaire (LOE II)	Test biochimique ELISA sur tissu congelé, marquage CE, approuvé FDA
Oncotype DX™	Niveau de preuve intermédiaire (LOE II)	Niveau de preuve intermédiaire (LOE II)	Analyses effectuées par RT-PCR (ARN puis ADNc) sur l'expression de 21 gènes au niveau d'une seule plate-forme à partir de blocs fixés et inclus en paraffine
MammaPrint®	Niveau de preuve bas (LOE III).	Il n'y a pas d'études cliniques permettant de conclure à une valeur prédictive	Analyses effectuées par PCR (ARN puis ADNc) sur l'expression de 70 gènes au niveau d'un seul laboratoire. Les échantillons sont issus de tissus cryopréservés ou conservés dans une solution permettant la stabilisation de l'ARN

Tableau 3 : Conclusions du rapport INCa-SFSPM de 2009. Rapport intégral : © uPA/PAI-1, Oncotype DX™, MammaPrint® - Valeurs pronostique et prédictive pour une utilité clinique dans la prise en charge du cancer du sein, collection état des lieux et des connaissances, ouvrage collectif édité par l'INCa, Boulogne-Billancourt, décembre 2013. *Source : www.e-cancer.fr.*

I.5. Le cancer du sein triple négatif (TNBC)

Comme mentionné précédemment, 15 à 20% des tumeurs mammaires diagnostiquées sont classées dans le sous-groupe triple négatif (TNBC). Ce sont des cancers agressifs, métastatiques et incurables à l'heure actuelle. Ces cancers sont caractérisés par l'absence des récepteurs aux oestrogènes et à la progestérone, ainsi que par l'absence d'amplification et de surexpression de HER2. Ces tumeurs ne bénéficient d'aucune thérapie ciblée utilisable en clinique à l'heure actuelle et la majorité des femmes atteintes de ce type de cancer finit par décéder malgré les chimiothérapies. Moins de 30% des femmes avec un TNBC métastatique survivent à 5 ans (Brouckaert et al., 2012; Foulkes et al., 2010).

Les tumeurs TNBC peuvent échapper au dépistage par mammographie. En effet, se développant rapidement, il s'agit souvent d'un cancer d'intervalle (qui survient entre deux mammographies) ou qui touche des femmes jeunes pas encore concernées par ce dispositif. Au diagnostic, la taille tumorale est souvent déjà importante et l'envahissement ganglionnaire plus fréquent. Sur le plan anatomopathologique, les TNBC sont le plus souvent des carcinomes canaux de grade élevé, avec un index mitotique fort et de nombreuses atypies cellulaires.

A l'heure actuelle, les patientes atteintes d'un cancer TNBC ont des options de traitements limitées. La chimiothérapie est le seul traitement systémique validé. Elle comprend des anthracyclines (epirubicine ou adriamycine) associées à des taxanes (paclitaxel ou docetaxel). Du fait de leur haut niveau prolifératif, ces tumeurs sont initialement plus chimiosensibles que les autres sous-types de cancer du sein. Cependant, leur forte instabilité génétique génère rapidement des mécanismes de résistance aux traitements. L'identification de nouvelles cibles

thérapeutiques constitue donc un enjeu majeur pour l'amélioration du traitement de ces cancers et c'est un champ d'investigation actif.

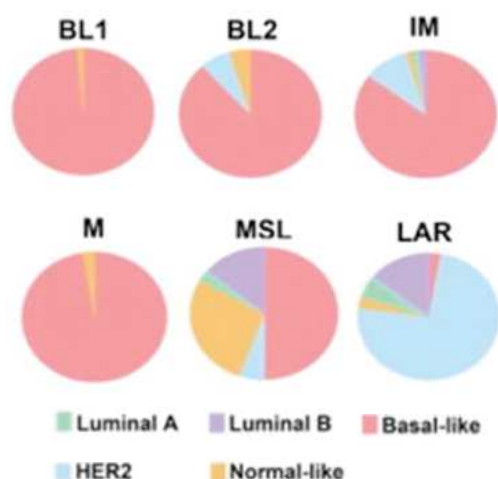
Cependant, une des difficultés majeures pour l'identification de nouvelles cibles est sans doute liée à l'hétérogénéité des tumeurs TNBC. Les travaux de Lehmann *et al* (Lehmann et al., 2011; Saha and Nanda, 2016) ont montré que les TNBC peuvent être classés en six sous-types moléculaires sur la base de leur transcriptome :

- * Basal-like 1 (BL1), caractérisé par une expression élevée des gènes impliqués dans le cycle cellulaire et la réponse aux dommages à l'ADN.
- * Basal-like 2 (BL2), enrichi en marqueurs myoépithéliaux et en produits des gènes impliqués dans la signalisation des facteurs de croissance.
- * Mésenchymal (M), exprimant les gènes impliqués dans la transition épithélio-mésenchymateuse, dans la motilité et la voie des facteurs de croissance.
- * Mesenchymal stem-like (MSL), proche du sous-type M. Il s'en différencie par une faible expression des gènes impliqués dans la prolifération et par l'expression de gènes impliqués dans l'angiogenèse, ainsi que ceux associés aux cellules souches.
- * Immunomodulatory (IM), caractérisé par l'expression de gènes de cytokines et de gènes impliqués dans la transduction du signal immun.
- * Luminal (LAR), exprimant des gènes luminaux sous le contrôle de la signalisation des récepteurs aux androgènes.

Le croisement des six sous-types de TNBC identifiés par Lehmann *et al* avec les cinq sous-groupes de cancer du sein initialement proposés par Perou *et al* indique que (i) les TNBC des sous-types BL1, BL2, IM et M appartiennent au groupe basal-like, (ii) les TNBC du sous-type LAR clusterisent fortement avec le groupe HER2 et (iii) les TNBC du sous-type MSL appartiennent aux groupes basal-like et normal breast-like (**Figure 7.A**) (Mayer et al., 2014).

De plus, l'analyse rétrospective de 163 tumeurs TNBC présentes dans la banque de données publique de cancer du sein du TCGA (The Cancer Genome Atlas) a mis en évidence sur la base de ce sous-typage des TNBC que les patientes BL1, IM et MSL avaient une survie médiane (OS) et une survie sans rechute (DFS) environ deux fois plus longues que les patientes avec des tumeurs BL2, LAR et M. Les patientes de sous-type IM ayant le meilleur pronostic (**Figure 7.B**) (Mayer et al., 2014; Saha and Nanda, 2016).

A



B

TNBC subtype ^a	# Samples (percentage)	Median DFS (mo) ^b	Median OS (mo) ^b
BL1	27 (17%)	20.1	21.1
BL2	12 (7%)	12.5	8.4
IM	30 (18%)	22.7	24.8
M	39 (24%)	9.1	9.5
MSL	10 (6%)	13.9	20.9
LAR	14 (9%)	4.4	5.7
UNC	31 (19%)	22.0	24.9
All TNBC	163 (100%)	11.8	15.2

Abbreviations: DFS, disease-free survival; IM, immunomodulatory; M, mesenchymal.

Figure 7 : Sous-typage des tumeurs TNBC d'après les données du TCGA. Source : Mayer et al. (2014).

Par ailleurs, sur la base des sous-types de TNBC Masuda *et al* ont exploré la réponse à la chimiothérapie néoadjuvante, donnée avant la chirurgie pour faire régresser la tumeur, de manière rétrospective à partir de bases de données publiques. Leurs résultats ont fait apparaître que la réponse complète (pCR), définie comme l'absence de détection de carcinome invasif et d'envahissement ganglionnaire lors de la chirurgie, est corrélée au sous-type de TNBC : les BL1 ayant le pCR le plus fort (52%) et les BL2 et LAR les plus faibles (0 et 10% respectivement) (**Tableau 4**) (Masuda et al., 2013). Ces études devront être complétées pour déterminer quelles sont les différences à long terme sur le devenir des patientes.

	pCR	Non-pCR	pCR rate	95% CI	P value
BL1	11	10	0.52	0.31–0.73	P=0.043
BL2	0	8	0.00	0.00–0.00	
M	8	18	0.31	0.13–0.48	
IM	8	19	0.30	0.12–0.46	
MSL	3	10	0.23	0.0001–0.45	
LAR	2	18	0.10	0.03–0.23	
UNS	5	10	0.33	0.09–0.57	

Tableau 4 : Description du statut pCR/Non-pCR en fonction du sous-typage des tumeurs TNBC. Source : Masuda et al. (2013).

Ces études basées sur les sous-types de TNBC suggèrent qu'il est sans doute sub-optimal de traiter les différents types de TNBC par les mêmes approches thérapeutiques. Ainsi, il est proposé que, suivant les anomalies génétiques des différents sous-types, différentes stratégies thérapeutiques pourraient être développées. Par exemple, les tumeurs du sous-type LAR pourraient bénéficier de traitements visant les récepteurs aux androgènes (Lehmann et al., 2011; Saha and Nanda, 2016).

L'identification de nouvelles cibles thérapeutiques constitue un enjeu majeur pour le traitement des cancers TNBC. De nombreuses pistes sont explorées mais pour l'instant sans succès avéré. Parmi les plus prometteuses on peut citer : les sels de platine et les inhibiteurs de PARP (Poly-ADP Ribose Polymérase), les inhibiteurs des récepteurs des androgènes ou encore l'immunothérapie anti-PD1 (Programmed Cell Death 1) qui vise à stimuler la réponse anti-tumorale des lymphocytes T (Collignon et al., 2016; Rapoport et al., 2014). Cependant, étant donné la complexité des TNBC ainsi que la redondance et les « cross-talk » dans les voies de signalisation intracellulaire, il est vraisemblable que la combinaison de deux ou plusieurs thérapies ciblées constituerait des approches optimales pour éviter les échappements ou les résistances acquises aux traitements.

Un des points délicats sera également la stratification des patientes en sous-types pour le choix des traitements futurs. En effet, l'utilité clinique de ses sous-types nécessite l'identification de marqueurs validés et robustes pour chaque sous-type de TNBC de manière à poser le pronostic et guider les choix thérapeutiques de façon plus optimale.

II. LES FACTEURS DE TRANSCRIPTION AP-1

La tumorigenèse est un mécanisme en plusieurs étapes qui est issue de mutations (héréditaires, sporadiques ou induites) dans le génome. Les altérations de l'ADN conduisent très souvent, par des mécanismes divers (mutations ponctuelles, amplifications, délétions ou remaniements de parties de chromosomes), à l'activation constitutive de voies de signalisation. L'activation exacerbée de ces dernières altère l'activité des facteurs de transcription en aval, ce qui conduit à une modification des programmes d'expression génique et contribue au développement tumoral. L'une des familles de facteurs de transcription fréquemment dérégulée dans les cancers est celle des facteurs de transcription AP-1. L'activité de ces facteurs est induite par une grande variété de stimuli extracellulaires (facteurs de croissance, signaux mitogènes, cytokines, chocs osmotiques ou thermiques...) et permet de transformer des signaux à court terme en réponses à long terme en enclenchant différents programmes transcriptionnels. Les facteurs de transcription AP-1 constituent donc des intégrateurs de signaux qui médient la transduction de différents stimuli en étant les cibles de nombreuses cascades de signalisation et en étant des effecteurs nécessaires à l'activité de ces dernières (Ozanne et al., 2006).

II.1. AP-1, une famille de complexes transcriptionnels

Le complexe AP-1 (Activating Protein-1) est un terme générique définissant une famille ubiquitaire de complexes transcriptionnels dimériques appartenant à la famille des protéines à b-zip (basic domain-leucine zipper). Le domaine b-zip est composé d'une région riche en acides aminés basiques, permettant la liaison à l'ADN de manière séquence-spécifique, située en amont d'une région contenant une glissière à leucine (répétition de leucines tous les 7 acides aminés) permettant la dimérisation (Vinson et al., 2002). Les domaines de dimérisation et de liaison à l'ADN sont hautement conservés chez toutes les protéines AP-1 (**Figure 8**). Les membres AP-1 les plus connus sont les membres des familles JUN (c-Jun, JunB et JunD) et FOS (c-Fos, FosB, Fra-1 et Fra-2). Les protéines JUN peuvent former des homodimères ou des hétérodimères avec les protéines FOS, alors que les protéines FOS ne peuvent qu'hétérodimériser avec les protéines JUN. Les hétérodimères JUN/FOS sont favorisés car plus stables que les homodimères JUN/JUN (Hess et al., 2004a). Les dimères AP-1 se fixent sur des séquences ADN spécifiques appelées motifs AP-1, qui peuvent être des sites TRE (TPA-Responsive Element) « 5' TGA(G/C)TCA 3' » ou CRE (Cyclic-AMP responsive Element) « 5' TGAGCTCA 3' » (**Figure 9**) (Karin et al., 1997). Néanmoins, *in vivo* les séquences reconnues peuvent dévier sensiblement de ces séquences consensus.

D'autres familles de protéines peuvent entrer dans la composition des facteurs de transcription AP-1. On trouve la famille ATF (Activating Transcription Factor), la famille MAF (de l'oncogène viral Musculo-Aponeurotic Fibrosarcoma), la famille JDP (JUN Dimerizing Protein) ou encore la famille C/EBP (CAAT/Enhancer Binding Protein).

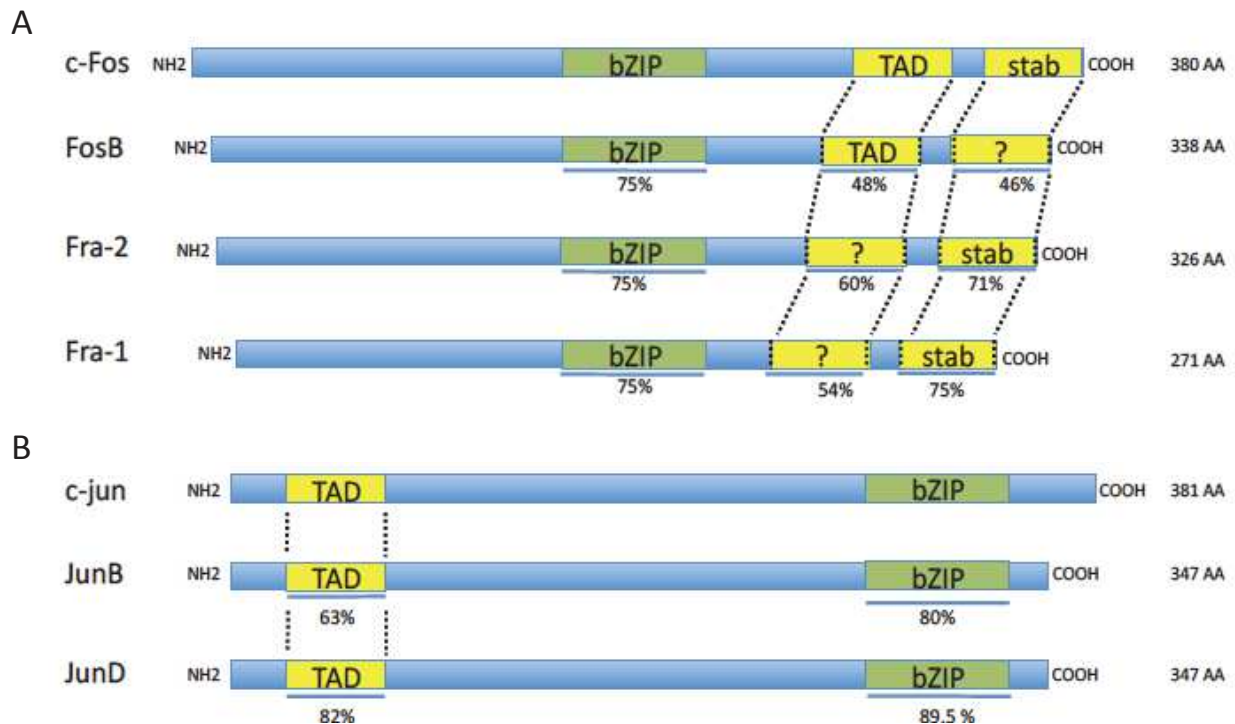


Figure 8 : Représentation schématique de l'homologie de séquence entre les différentes protéines FOS et JUN humaines. A) La comparaison des séquences des protéines FosB, Fra-1 et Fra-2 à celle de c-Fos montre une grande conservation du domaine de dimérisation et de liaison à l'ADN (b-zip). Les domaines de transactivation (TAD) et le déstabilisateur C-terminal (stab) sont plus ou moins conservés. B) La comparaison des séquences des protéines JunB, JunD à celle de c-Jun montre une grande conservation du domaine de dimérisation et de liaison à l'ADN (b-zip) et une conservation moindre du domaine de transactivation. *Comparaison faite sur Genius : <http://www.geneious.com>.*

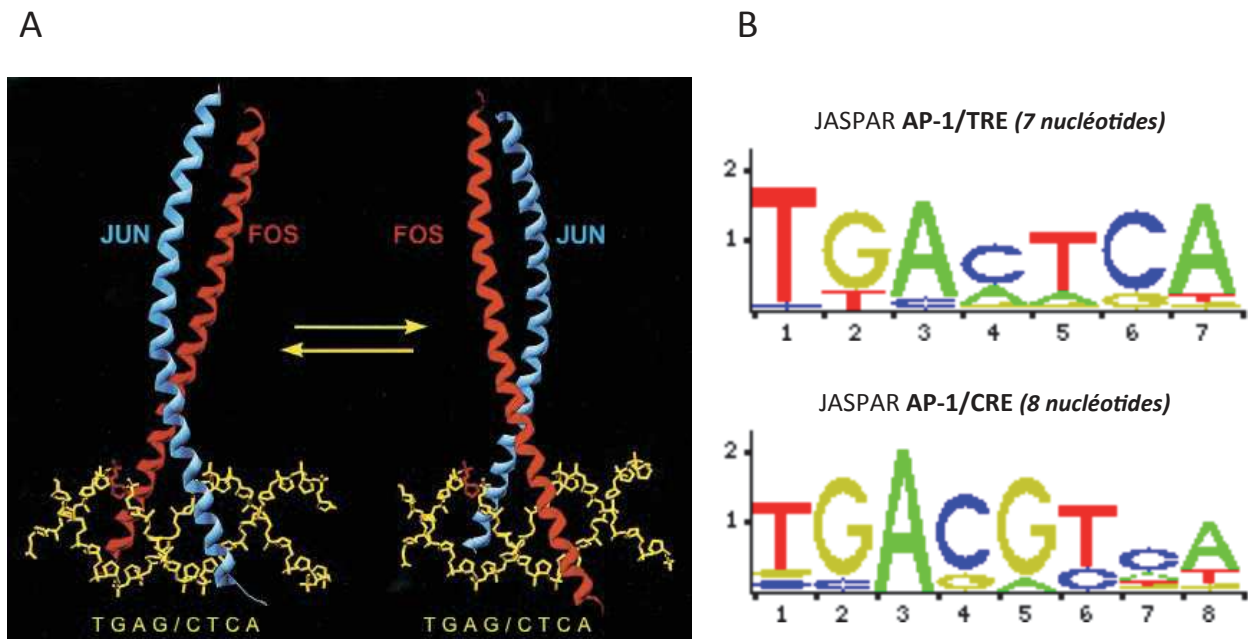


Figure 9 : Structure tridimensionnelle de l'hétérodimère JUN/FOS. A) Les domaines b-zip de JUN et FOS dimérisés et fixés sur un motif AP-1 sont représentés. Le dimère peut se fixer dans les deux orientations sur la séquence cible. B) Matrice PWN (position weight matrix) JASPAR pour les motifs AP-1 reconnus par les dimères FOS/JUN. *Source : Hess et al. (2004).*

II.2. Régulation de l'activité du complexe AP-1

La régulation de l'activité AP-1 est très complexe et fait intervenir plusieurs niveaux de régulation, ce qui permet la formation de différents dimères d'activités transcriptionnelles différentes. Nous nous intéresserons ici principalement aux membres FOS et JUN.

II.2.a. Régulation de l'abondance des membres AP-1 (FOS et JUN)

Le niveau d'expression des différents membres AP-1 est régulé par des mécanismes transcriptionnels, post-transcriptionnels, traductionnels et post-traductionnels.

* Régulation transcriptionnelle des membres AP-1. Les gènes codant pour la plupart des protéines AP-1 sont des gènes précoces immédiats. Ils sont exprimés rapidement et transitoirement suite à différents stimuli (tels que les facteurs de croissance, les cytokines, les infections bactériennes ou virales, et toute une variété de stress chimiques ou physiques) qui activent différentes cascades de signalisation dont la voie ERK1/2 connue pour être impliquée dans l'expression des protéines FOS (Piechaczyk and Blanchard, 1994). Certains gènes codant pour des membres AP-1 (Fra-1, Fra-2 et c-Jun) possèdent des sites de liaison pour AP-1 dans leurs régions régulatrices. Leur expression dépend donc, en partie, de l'expression préalable d'autres membres AP-1 et ils peuvent bénéficier d'une boucle d'auto-amplification conduisant à leur expression sur le plus long terme (Adiseshaiah et al., 2008; Murakami et al., 1999; Sonobe et al., 1995; Wong et al., 1999; Young and Colburn, 2006). C'est le cas en particulier pour Fra-1 comme cela est illustré dans la **Figure 10** (Young and Colburn, 2006).

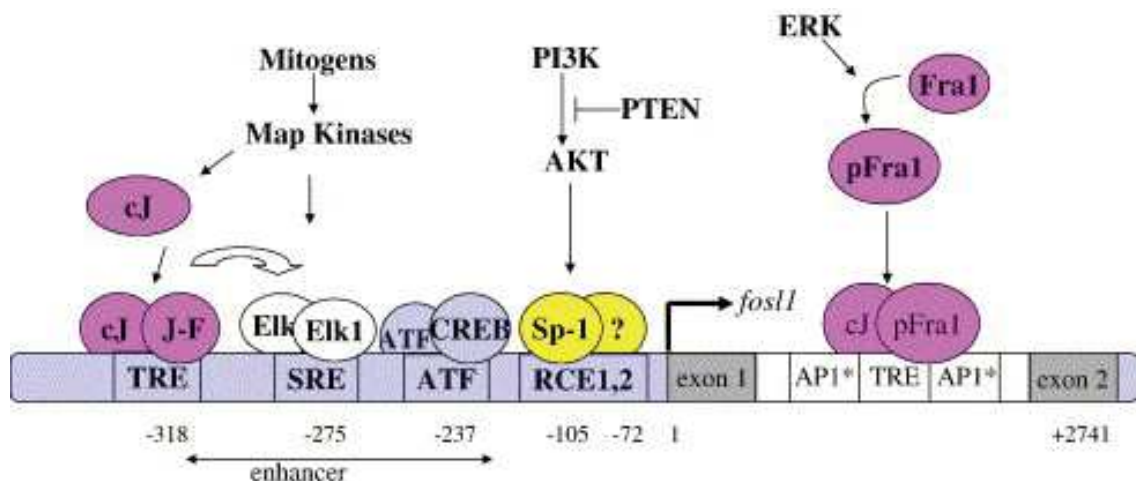


Figure 10 : Protéines impliquées dans la régulation transcriptionnelle de Fra-1. L'expression du gène FOSL1, codant pour la protéine Fra-1, est activée par les voies MAPK et PI3K/AKT, ainsi que par la liaison de Fra-1 à son propre promoteur. La voie MAPK active les facteurs de transcription se liant aux sites TRE (TPA response element), SRE (serum response element) et ATF (activating transcription factor) dans le promoteur de FOSL1 et la voie PI3K/AKT régule l'expression de FOSL1 via le site RCE (retinoblastoma control element). Le dimère Fra-1/c-Jun régule l'expression FOSL1 en se liant au site TRE situé dans le premier intron de FOSL1 (région +105 bp à +2741 bp), qui contient un site AP-1 consensus ainsi que deux sites AP-1-like. (cJ) cJun ; (J-F) JunD ou Fra-2 ; (AP-1*) site AP-1-like. *Source : Young and Colburn (2006).*

** Stabilité et traduction des ARNm AP-1.* Les ARNm codant les protéines AP-1 sont généralement très instables, en particulier l'ARNm de c-Fos. Cependant les demi-vies des ARNm peuvent être plus ou moins longues suivant les conditions d'expression (Piechaczyk and Blanchard, 1994). Leur traduction a également été rapportée comme étant régulée par différents mécanismes selon les ARNm. Ainsi des travaux récents rapportent que la traduction de c-Jun est régulée par des mécanismes à la fois CAP-dépendants et indépendants, alors que celle de c-Fos est négativement régulée par le microRNA mir7B, tandis que celle de JunD et JunB font intervenir d'autres mécanismes (pour revue voir (Vesely et al., 2009)).

** Stabilité des protéines AP-1.* Un autre mécanisme déterminant pour contrôler l'abondance des protéines AP-1 est la régulation de leur stabilité. Des travaux antérieurs du laboratoire et d'autres groupes ont montré que les protéines c-Fos et Fra-1 sont intrinsèquement instables et sont dégradées par le protéasome indépendamment de la voie de l'ubiquitine (Basbous et al., 2008; Bossis et al., 2003; Ferrara et al., 2003; Vial and Marshall, 2003). Cependant suite à leur phosphorylation par la voie ERK1/2 sur deux sérines C-terminales (S362 et S374 pour c-Fos et S252 et S265 pour Fra-1) ces deux protéines sont fortement stabilisées (Basbous et al., 2008; Bossis et al., 2003; Okazaki and Sagata, 1995; Vial and Marshall, 2003). La voie ERK5 a également été décrite pour phosphoryler (sur des sites non complètement caractérisés) et stabiliser c-Fos et Fra-1 (Terasawa et al., 2003). Plus récemment, les travaux de Belguise *et al* ont montré que dans des cellules TNBC n'ayant pas la voie ERK activée, la protéine Fra-1 est stabilisée par phosphorylation sur les sites S265, T223 et T230 par PKC θ (Belguise et al., 2012). Ces résultats montrent que différentes voies de signalisation contribuent à la stabilisation de Fra-1, contrôlant ainsi son abondance intracellulaire. Les mécanismes de stabilisation de FosB et Fra-2 n'ont pas, ou peu, été étudiés.

Il est intéressant de noter, que c-Fos et Fra-1 ont été décrits pour être des senseurs de l'activité ERK et entraîner des réponses cellulaires différentes selon l'intensité de l'activation de la voie ERK en réponse aux stimuli. Ainsi, lorsque l'activité des ERK est faible et transitoire, ces kinases sont capables d'induire l'expression des gènes c-Fos et Fra-1 mais ne sont plus actives au moment où les protéines c-Fos et Fra-1 sont synthétisées. Ces dernières ne sont alors pas phosphorylées et stabilisées, et de ce fait elles sont peu abondantes. En revanche lorsque le stimulus induit une activité ERK forte et durable dans le temps, ces dernières sont encore actives au moment où les protéines c-Fos et Fra-1 sont synthétisées et peuvent les phosphoryler. Sous forme stabilisée c-Fos et Fra-1 s'accumulent et peuvent enclencher les programmes transcriptionnels appropriés en réponse aux stimuli (Chalmers et al., 2007; Murphy and Blenis, 2006; Murphy et al., 2002, 2004).

En ce qui concerne les membres JUN, il a été montré que la phosphorylation de la S73 par la voie des JNK protège c-Jun de la dégradation ubiquitine-dépendante par le protéasome et de ce fait augmente sa demi-vie (Fuchs et al., 1996). La phosphorylation des JUN par les JNK

nécessite la présence d'un site de fixation des JNK (site de docking) et d'un accepteur de phosphorylation (S63 et S73 pour c-Jun). Contrairement à c-Jun qui possède ces deux éléments, JunD ne possède pas le site de docking et est décrite pour être phosphorylée via sa dimérisation avec les deux autres membres JUN. Cependant, JunD est peu ubiquitinée et a une demi-vie longue (Musti et al., 1996). Par ailleurs, la phosphorylation de JunB sur d'autres sites (T150 et S186) par CDC2/CycB, a été montrée nécessaire à sa dégradation efficace par la voie du protéasome au cours de la transition G2/M (Bakiri et al., 2000; Farràs et al., 2008). De façon plus récente, il a été mis en évidence que la dimérisation pouvait également contribuer à contrôler la demi-vie des protéines AP-1, du moins celle de c-Jun. Ainsi dans des cellules de cancer de la thyroïde où la voie Ras est activée, la dimérisation avec Fra-1 phosphorylée par la voie ERK stabilise la protéine c-Jun et favorise son accumulation (Talotta et al., 2010).

Ainsi, le panel des différents dimères AP-1 présents dans une cellule à un moment donné sera fonction du contexte cellulaire, de l'étape du cycle cellulaire et des voies de signalisation activées.

II.2.b. Dimérisation des protéines AP-1

Les protéines FOS et JUN doivent homodimériser (JUN/JUN) ou hétérodimériser (FOS/JUN) pour pouvoir se lier à l'ADN. Des travaux anciens ont abouti à la conclusion que la durée de l'interaction entre les membres FOS et JUN est courte, de l'ordre de la dizaine de secondes (Patel et al., 1994), ce qui signifie que la composition des dimères AP-1 est extrêmement dynamique. L'idée généralement admise est que la nature des dimères formés dans la cellule dépend essentiellement de l'abondance des différents constituants à un moment donné (Cook et al., 1999; Kovary and Bravo, 1991). Cependant, les travaux de Masson *et al* réalisés *in vitro*, en utilisant des formes tronquées des protéines FOS et JUN (contenant le b-zip et quelques dizaines d'acides aminés de part et d'autre), ont montré que certains dimères présentent une stabilité d'interaction sensiblement plus grande que d'autres. Ainsi, les hétérodimères c-Fos/c-Jun, Fra-1/c-Jun et Fra-1/JunD sont favorisés et les homodimères entre membres JUN sont peu stables, à l'exception de c-Jun/c-Jun (Mason et al., 2006). Si l'interaction est dynamique en solution, la liaison des dimères AP-1 au niveau de leurs séquences cibles ADN est un facteur de stabilisation (Patel et al., 1994).

II.2.c. Régulation de l'activité des protéines AP-1 par modifications post-traductionnelles

Comme nous l'avons vu, les voies de signalisation par les MAPK contribuent à l'expression et à la stabilisation des protéines FOS et JUN, mais leur implication dans l'activité AP-1 ne s'arrête pas là. En effet, ces kinases phosphorylent aussi les protéines AP-1 sur des résidus impliqués dans l'activité transcriptionnelle à proprement parler (**Figure 11**).

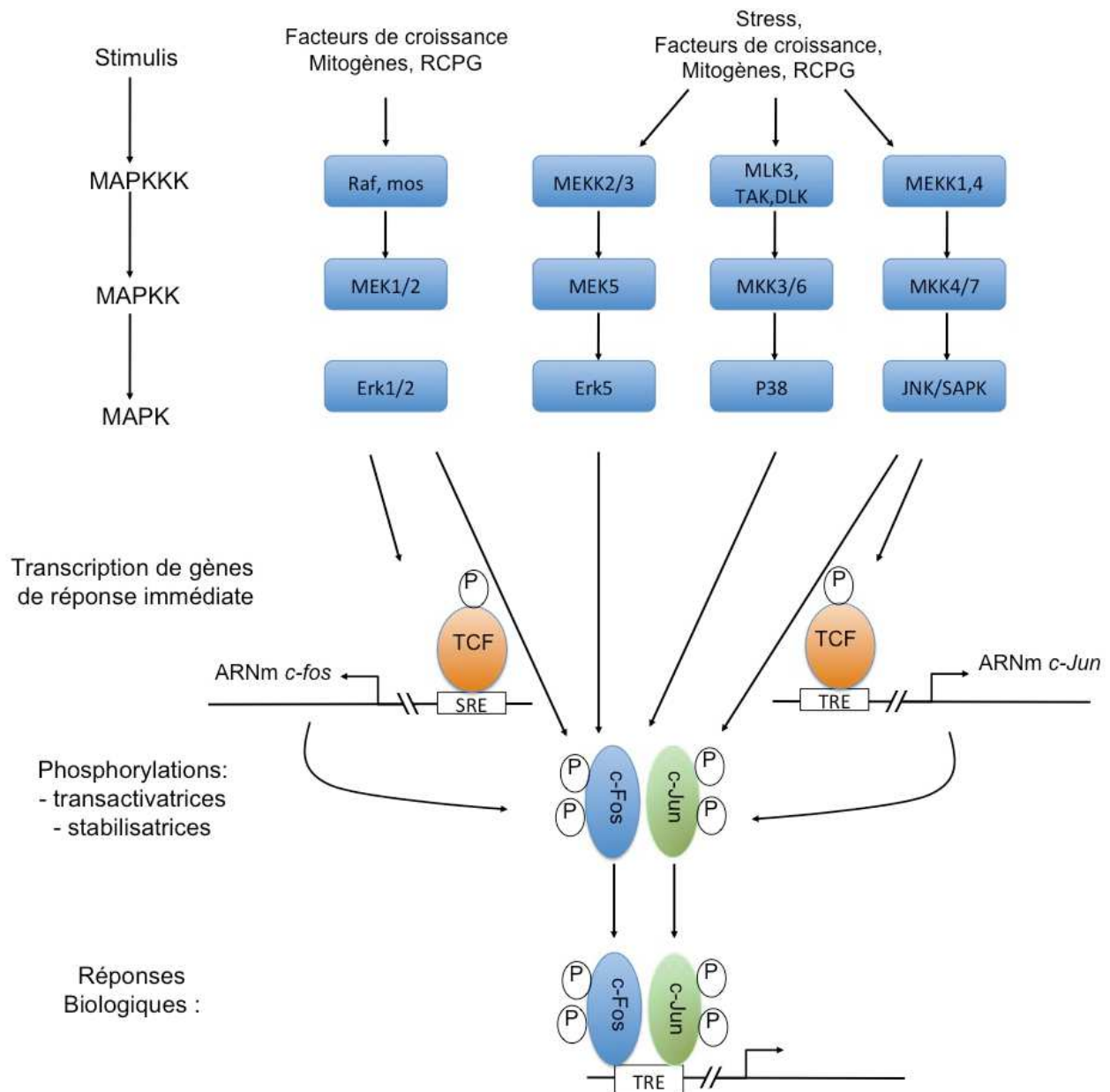


Figure 11 : Différents niveaux de régulation de l'activité AP-1 par les voies MAPK dans la cellule. Suite à une stimulation extracellulaire, l'activation des voies MAPK conduit dans un premier temps à l'expression des gènes précoces et immédiats c-Fos et c-Jun. Puis, dans un second temps, à la phosphorylation de leurs domaines transactivateurs et stabilisateurs. Ainsi, une fois activées les protéines c-Fos et c-Jun se fixent au niveau des motifs AP-1 TRE ou CRE de leurs gènes cibles et stimulent leurs transcriptions. Les membres AP-1 permettent donc la traduction de stimuli extracellulaires en réponse biologique.

L'activité transcriptionnelle des protéines JUN est principalement induite suite à leur phosphorylation par la voie des JNK/SAPK, activée en réponse à des facteurs mitogènes, des hormones, des cytokines et à différents types de stress physiques et chimiques. De manière non exhaustive, on peut mentionner que la phosphorylation de c-Jun par les JNK sur deux sérines (S63 et S73) et deux thréonines (T91 et T93) active son potentiel transcriptionnel (pour revue voir (Shaulian, 2010)). Un des mécanismes sous-jacents étant la dissociation de c-Jun, suite à sa phosphorylation, de l'histone déacétylase HDAC3 (Shaulian, 2010; Weiss et al., 2003). A

l'inverse la phosphorylation sur T239 par la kinase GSK3 et celle sur S243 par une kinase non identifiée ont un effet inhibiteur sur l'activité transcriptionnelle de c-Jun (Morton et al., 2003). Par ailleurs, dans un modèle hépatique, il a été montré que l'expression JunD-dépendante du gène TMP1-1 est contrôlée via la phosphorylation de JunD, sur S100, par les JNK (Smart et al., 2006).

L'activité transcriptionnelle des protéines FOS est principalement induite suite à leur phosphorylation par la voie ERK1/2. La protéine c-Fos, la plus étudiée parmi les membres FOS, est connue pour avoir un domaine transactivateur C-terminal dont la capacité transactivatrice est augmentée suite à sa phosphorylation par ERK1/2. Deux sites accepteurs ont été caractérisés, T325 et T331 (Murphy et al., 2002), mais il est très vraisemblable que d'autres sites de phosphorylation restent à caractériser. L'activation transcriptionnelle de c-Fos par ERK5 et p38 est controversée dans la littérature. Des premiers travaux avaient montré que (i) l'activation de la voie ERK5 phosphoryle c-Fos sur des sites différents de ceux phosphorylés par ERK1/2 (Terasawa et al., 2003) et que (ii) les irradiations UV induisent la phosphorylation de c-Fos par p38, augmentant ainsi son pouvoir transactivateur (Tanos et al., 2005). A l'inverse, des travaux plus récents ont conclu que les voies ERK5 et p38 ne pouvaient se substituer à la voie ERK1/2 pour le contrôle de l'activité transcriptionnelle de c-Fos (Gilley et al., 2009). Concernant FosB, la phosphorylation C-terminale pour son activation transcriptionnelle est admise, mais les sites accepteurs de modification n'ont pas été identifiés (Skinner et al., 1997). Les protéines Fra-1 et Fra-2 sont dépourvues d'une partie du domaine transactivateur C-terminal présent dans c-Fos et FosB. Dans le même sens, il a été observé que ces deux protéines, contrairement à c-Fos et FosB, ne stimulent pas l'expression de gènes rapporteurs sous contrôle de promoteurs artificiels contenant des séquences AP-1 et n'ont pas d'activité transformante dans des cellules de carcinomes embryonnaires murins (Suzuki et al., 1991). Il a donc été considéré pendant longtemps qu'elles n'avaient qu'un faible rôle transactivateur, voire même qu'elles étaient des répresseurs transcriptionnels (Wisdon and Verma, 1993). Cependant, les travaux de Young *et al* ont établi que Fra-1 possède en fait un domaine transactivateur latent, qui n'est révélé qu'après sa phosphorylation par la voie ERK1/2 : la phosphorylation de Fra-1 sur T231 a été identifiée comme nécessaire pour son activation transcriptionnelle (Young et al., 2002). Ainsi dans les contextes cellulaires où la voie ERK1/2 est peu ou pas active, le potentiel transactivateur de Fra-1 ne peut être révélé. Cependant, comme pour c-Fos, il est très vraisemblable que d'autres sites de phosphorylation non encore caractérisés soient impliqués dans le pouvoir transactivateur de Fra-1. De la même manière, il est connu que Fra-2 est phosphorylée par la voie ERK1/2, mais ni les sites de phosphorylation, ni les rôles physiologiques de ces phosphorylations n'ont été établis.

A l'inverse de la phosphorylation, connue pour activer les protéines AP-1, des travaux dont ceux du laboratoire ont montré que la sumoylation diminue leur potentiel transcriptionnel (Bossis et al., 2005; Muller et al., 2000). La sumoylation est une modification post-

traductionnelle qui conduit à l'attachement covalent d'une ou plusieurs molécules de SUMO (ubiquitin-like) sur les chaînes latérales de lysines spécifiques, modifiant ainsi les surfaces d'interaction des protéines et donc leur devenir ou leur activité. Il a été montré que la sumoylation de c-Fos sur la lysine 265 et de c-Jun sur les lysines 229 et 257 diminue l'activité transcriptionnelle du dimère (Bossis et al., 2005; Muller et al., 2000). Dans le cas de c-Fos, le laboratoire a montré que la sumoylation de la lysine 265 est en compétition avec la phosphorylation de la thréonine 232 décrite pour augmenter l'activité transactivatrice. De plus, cette modification entraîne une rétention de c-Fos dans le noyau par des mécanismes encore non caractérisés (Bossis et al., 2005). Les mécanismes sous-jacents liés à la répression transcriptionnelle par la sumoylation ne sont pas encore clairement identifiés mais les données de la littérature suggèrent qu'ils mettraient en jeu le recrutement de co-répresseurs transcriptionnels (Garcia-Dominguez and Reyes, 2009). A l'heure actuelle, le dogme « sumoylation est essentiellement répresseur de la transcription » est remis en question par des études de cartographie à grande échelle des protéines sumoylées sur la chromatine. En effet, des régions hautement transcrites du génome sont très enrichies en SUMO (Chymkowitch et al., 2015).

Il a également été reporté que les protéines AP-1 sont la cible d'une autre modification post-traductionnelle : l'acétylation. Il a été montré que c-Jun est acétylé *in vivo* sur la lysine 271 et que l'interaction directe de SIRT1 (NAD-dependent deacetylase sirtuin-1) avec c-Jun réprime son activité transcriptionnelle et entraîne la diminution de l'expression de MMP9 (protéase de matrice extracellulaire) (Gao and Ye, 2008; Mishra et al., 2016). Il a également été montré que SIRT1 diminue l'acétylation de c-Fos/c-Jun induite par P300, inhibant ainsi l'activité transcriptionnelle du dimère, ce qui a pour conséquence de diminuer l'expression de COX-2 dans les macrophages (Zhang et al., 2010). Ainsi, la désacétylation des protéines AP-1 par SIRT1 inhibe leur activité transcriptionnelle (Xie et al., 2013) et à l'inverse, leur acétylation a pour effet de les activer et de favoriser leur fixation à l'ADN (Vries et al., 2001; Wang et al., 2011). Cependant, l'effet direct de l'acétylation sur l'activité transcriptionnelle n'a pas encore été rapporté.

Comme cela apparaît dans ce paragraphe, l'essentiel des travaux de la littérature concernant la régulation de l'activité du complexe AP-1 s'est concentré sur les protéines c-Fos et c-Jun. Quelques données sont connues sur Fra-1 et pratiquement rien n'a été décrit pour la régulation de Fra-2.

II.2.d. Interaction avec d'autres facteurs de transcription

En plus de leur capacité à dimériser entre elles, les protéines JUN et FOS peuvent également interagir physiquement avec d'autres facteurs de transcription contenant un domaine leucine zipper, tels que les protéines ATF, MAF ou C/EBP, déjà mentionnées au paragraphe II.1. Les

protéines AP-1 sont également capables d'interagir avec d'autres facteurs de transcription sans leucine zipper tels que les récepteurs aux oestrogènes (Cascio et al., 2007; Pedraza-Alva et al., 2009), le récepteur aux glucocorticoïdes, NF-KB (Rahmani et al., 2001; Xiao et al., 2004) ou la protéine de liaison à CREB (CREBP-binding protein). Ces interactions, en plus d'augmenter la complexité du répertoire de protéines capables d'interagir avec AP-1, entraînent des changements d'affinité de AP-1 pour ses gènes cibles. Les différents complexes peuvent avoir des effets synergiques ou au contraire antagonistes, ce qui permet une régulation plus fine des gènes contrôlés par AP-1 (pour revue voir (Chinenov and Kerppola, 2001)).

Une série de travaux récents a montré l'existence d'une coopération entre AP-1 et TEAD pour la régulation de gènes impliqués dans l'invasion (Verfaillie et al., 2015), la motilité (Liu et al., 2016) et la prolifération (Zanconato et al., 2015). De plus, l'interaction avec AP-1 est décrite pour recruter le co-activateur SRCI-3 (Liu et al., 2016) et pour faire intervenir des boucles de chromatine entre les enhancers distaux et les promoteurs (Verfaillie et al., 2015; Zanconato et al., 2015).

Par ailleurs, dans les chondrocytes préhypertrophiques c-Jun et Fra-2 collaborent positivement avec le facteur de transcription SOX9 pour l'établissement du programme hypertrophique et le développement du squelette. Cette collaboration transcriptionnelle passe par l'interaction de la région N-terminale de SOX9 avec la région basique et le domaine adjacent riche en pro/glu de c-Jun (He et al., 2016). c-Jun a également été montré pour interagir avec le domaine C-terminal de p73 et cette interaction augmente l'activation des gènes cibles d'AP-1 et la croissance cellulaire (Subramanian et al., 2015).

II.3. AP-1 et mécanismes de régulation de la transcription des gènes cibles

Comme déjà mentionné, le complexe AP-1 joue un rôle d'intégrateur des signaux extracellulaires en réponse à une large gamme de stimuli et il stimule l'expression de différents programmes génétiques, tels que prolifération, différenciation, apoptose, réponse au stress, etc. Les données de la littérature, décrites brièvement ci-dessus, indiquent que la nature des dimères formés, les cascades de signalisation activées, le contexte cellulaire et l'interaction avec d'autres facteurs de transcription contribuent à réguler positivement ou négativement différentes catégories de gènes (Eferl and Wagner, 2003; Hess et al., 2004b).

Pour illustrer cette complexité, on peut citer les travaux de Bakiri *et al* qui ont mis en évidence, dans des fibroblastes de souris NIH3T3 et des cellules humaines HEK293, que les dimères c-Jun/c-Fos, c-Jun/Fra-1 et c-Jun/Fra-2 ont des affinités quasi identiques pour le promoteur de la cycline D1 alors qu'ils présentent des affinités différentes pour le promoteur de MMP1, ce qui se traduit par des potentiels différents d'activation transcriptionnelle (Bakiri et al., 2002). Les travaux de Toft *et al* ont montré que le gène de la proliférine (impliquée dans

l'angiogenèse et la migration des cellules endothéliales) est régulé positivement par le dimère c-Jun/JunB et négativement par le dimère c-Jun/c-Fos, dans un modèle de fibrosarcome (Toft et al., 2001). Par ailleurs, il est intéressant de noter que la fixation de c-Fos et Fra-1 à différents temps sur la chromatine contrôle l'expression de l'interleukine 8 (IL-8) et de la cycline D1. En effet, la synthèse dans un premier temps de c-Fos conduit à l'activation de l'expression de IL-8 puis la synthèse dans un deuxième temps de Fra-1 conduit à la répression de sa transcription (Hoffmann et al., 2005). A l'inverse, dans un autre système, la fixation de c-Fos sur le promoteur de la cycline D1 réprime son expression et l'échange de c-Fos par Fra-1 est nécessaire à l'initiation de l'expression de cycline D1 (Burch et al., 2004).

Cependant, malgré une littérature très abondante sur AP-1 (plus de 10 000 références) les mécanismes par lesquels les membres AP-1 gouvernent l'expression transcriptionnelle de leurs gènes cibles sont pour l'essentiel encore inconnus. Une difficulté pour ces études réside dans le fait que, suivant le contexte cellulaire et le stimulus, différents dimères AP-1 sont produits et régulent différentes catégories de gènes, ce qui signifie que les données obtenues dans un contexte ne sont pas toujours transposables à un autre. Il est également vraisemblable que les mécanismes de régulation seront, au moins partiellement, différents d'un gène à l'autre ou d'une catégorie de gènes à une autre. Néanmoins, des études *in vitro* et *in cellulo* ont montré que la régulation de la transcription des gènes par AP-1 peut se faire à différents niveaux. Quelques exemples qui soulignent la complexité des mécanismes mis en jeu sont brièvement décrits ci-dessous.

II.3.a. Modifications de la structure chromatinienne par AP-1

AP-1 a été décrit pour être capable d'induire différents réarrangements chromatinien. Des premiers travaux ont montré que la liaison des dimères FOS/JUN induit une pliure de l'ADN et il a été proposé que ce changement de conformation contribuait à la réorganisation architecturale de la région promotrice nécessaire à l'activation transcriptionnelle (Kerppola and Curran, 1997). Par ailleurs, il a été décrit que l'hétérodimère c-Jun/c-Fos est capable de se fixer sur un site AP-1 inséré dans un nucléosome acétylé et que cette fixation détruit la structure nucléosomale, sans entraîner de déplacement des histones, ce qui permet ensuite le recrutement d'autres facteurs de transcription (Ng et al., 1997). Plus récemment, il a été montré que l'expression d'AP-1 et sa liaison avec NFAT est nécessaire à la formation d'une région dépourvue de nucléosome sur le promoteur de l'IL-2 et que la phosphorylation de c-Jun par les JNK joue un rôle dans le recrutement plus tardif d'autres co-facteurs (Ishihara and Schwartz, 2011). La forme phosphorylée de c-Jun est également capable de recruter la Topoisomérase-I sur la chromatine et ainsi de permettre l'ouverture de la double hélice d'ADN lors du passage de la machinerie de transcription en régulant les super-tours de l'ADN (Mialon et al., 2005).

Malgré ces observations, AP-1 n'est pas décrit comme un « facteur pionnier », c'est à dire un facteur capable de se lier sur la chromatine condensée et de modifier sa structure pour permettre la fixation d'autres facteurs et la transcription (**Figure 12**).

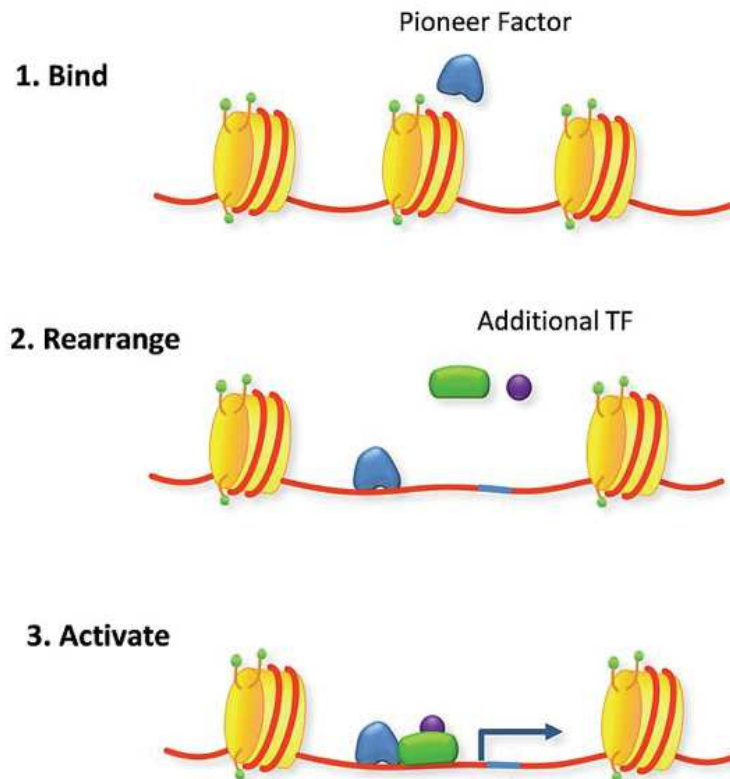


Figure 12 : Ouverture de la chromatine condensée par un facteur pionnier pour initier la transcription. Un facteur pionnier est capable de se lier à des régions chromatiniennes inaccessibles à la plupart des autres facteurs de transcription et d'établir un environnement plus favorable, soit par le biais de remodeleurs chromatinien, soit en ayant lui-même la capacité de changer la conformation de la chromatine afin de permettre la liaison d'autres facteurs.

De façon plus récente, certains membres AP-1 ont été impliqués dans la formation de boucles chromatiniennes nécessaires à l'expression de gènes cibles. Il a été montré, dans des cellules leucémiques, que l'expression de MMP9 est sous le contrôle de p65 (sous-unité de NF- κ B) phosphorylé par la voie AKT et de c-Jun phosphorylé par la voie JNK. L'expression de MMP9 fait intervenir une boucle entre les sites de liaison de p65 et de c-Jun via l'interaction des protéines (Chen and Chang, 2015). Il a également été montré, dans des cellules de cancer du poumon, que la régulation de l'expression du gène de la nitric-oxide-synthetase (h*i*NOS) fait intervenir un enhancer AP-1 localisé à 5.3kb en amont du TSS (Transcription Start Site). L'interaction fonctionnelle du dimère JunD/Fra-2 et de P300 au niveau de l'enhancer est nécessaire à l'expression d'h*i*NOS et fait intervenir une boucle chromatinienne entre l'enhancer et la région promotrice du gène (Guo et al., 2016).

II.3.b. Recrutement de modificateurs des histones par AP-1

Les complexes AP-1 peuvent également recruter des activités HATs (Histones Acétyl Transférases : CBP/P300, PCAF) ou HDACs (Histone De-Acétylases) et ainsi passer du mode activateur à celui de répresseur des gènes cibles. Cette régulation nécessite, dans certains cas, la phosphorylation des protéines AP-1. Dans le cas de c-Jun, il a été montré que l'inhibition de sa phosphorylation empêche le recrutement de PCAF au promoteur du gène MMP9 (Tsai et al., 2008). Il en est de même pour JunB qui nécessite la phosphorylation de sa sérine 79 pour interagir avec P300 et réguler la transcription du gène DMP1, lors de la différenciation ostéoblastique (Narayanan et al., 2004). De la même façon, il a été montré que c-Fos coopère avec P300 pour stimuler la transcription de l'IL-8 dans des macrophages murins (Sun et al., 2005) ou réguler l'apoptose dans des cellules de hamster (Preston et al., 2000). Fra-1 a été décrit pour coopérer, de façon indirecte, avec P300 et PCAF pour réguler la transcription du gène de l'involucrine dans des kératinocytes humains (Crish and Eckert, 2008). A l'inverse, Fra-1 stimule le recrutement de HDAC-1 au promoteur de l'IL-8 pour inhiber sa transcription induite par c-Fos, dans des cellules KB (Hoffmann et al., 2005). Des travaux récents du laboratoire auxquels j'ai été associée à titre secondaire ont montré, entre autres, que Fra-1 est nécessaire au recrutement de P300 sur les enhancers du gène uPA/PLAU, ce recrutement étant requis pour l'expression optimale d'uPA/PLAU (Moquet-Torcy, Tolza et al., 2014).

II.3.c. Recrutement de co-facteurs transcriptionnels par AP-1

Plusieurs études ont montré que les membres AP-1 peuvent directement interagir avec certains composants de la machinerie de transcription ou certains co-activateurs. Ainsi, c-Fos et c-Jun ont été décrits pour interagir avec le domaine C-terminal de TBP (TATA Binding Protein) (Ransone et al., 1993). L'hétérodimère c-Fos/c-Jun établit aussi des contacts directs avec trois protéines de la machinerie basale de transcription : TFIIE- β (TFIIE-34), RAP30 (TFIIF-30) et RAP74 (TFIIF-74) ainsi que les complexes TFIIE et TFIIF, impliqués respectivement dans la stabilisation de la machinerie transcriptionnelle et dans l'ouverture du promoteur (Martin et al., 1996).

Les membres AP-1 ont aussi été décrits pour interagir avec certaines sous-unités du complexe TFIID impliqué dans l'assemblage du complexe transcriptionnel. Ainsi, c-Jun interagit physiquement avec TAF7 (TAFII55) via deux domaines d'interaction indépendants, situés dans ses parties N- et C-terminales. L'interaction se fait de façon préférentielle quand c-Jun est phosphorylée et liée à l'ADN (Munz et al., 2003). Par ailleurs, il est décrit que TAF1 (TAFII250) est recruté par le dimère c-Fos/c-Jun grâce à l'interaction du domaine b-zip de c-Jun avec TAF1, c-Fos n'interagissant pas directement avec TAF1 (Lively et al., 2004). Dans un modèle de cancer de colon agressif, c-Jun a également été rapporté pour interagir avec la sous-

unité TAF4b (TAFII135) alors que JunB interagit avec la sous-unité TAF10 (TAFII28). La liaison de c-Jun sur le promoteur de l'intégrine $\alpha 6$ (impliquée dans l'EMT) est nécessaire au recrutement de TAF4b dans les cellules Caco-H (Kalogeropoulou et al., 2010).

Certains membres AP-1 ont également été décrits pour interagir avec le complexe de remodelage de la chromatine SWI-SNF : l'hétérodimère c-Fos/c-Jun, mais pas l'hétérodimère Fra-2/JunD, interagit avec la sous-unité BAF60 et cette interaction est déterminante pour son pouvoir transactivateur au niveau des gènes tels que la collagénase ou c-met (Ito et al., 2001).

De façon plus récente, nTRIP6, une protéine à domaine LIM, a été identifiée comme un nouvel interacteur d'AP-1 nécessaire au recrutement du Mediator via son interaction avec la sous-unité TRAP3 (Med150). De manière intéressante, le recrutement du Mediator est inhibé par l'interaction de nTRIP6 avec GR (récepteur aux glucocorticoïdes), GR étant connu pour réprimer l'activité AP-1 (Diefenbacher et al., 2014).

II.4. AP-1 et tumorigénèse

Certains composants AP-1 présentent des propriétés oncogéniques et/ou oncosuppressives suivant le contexte cellulaire. Ceci a été mis en évidence dans un premier temps par transfection des différents membres AP-1 dans des fibroblastes embryonnaires murins. Ainsi, il a été montré que c-Fos et c-Jun sont capables de transformer des cellules en culture, leur conférant une croissance indépendante d'ancrage et à haute densité. Des potentiels oncogéniques ont également été rapportés pour FosB et Fra-1 (pour revues voir (Ozanne et al., 2007; Shaulian and Karin, 2002)).

Plus tard, des études utilisant des souris transgéniques surexprimant les différents membres AP-1 ou des souris invalidées génétiquement pour chacun d'entre eux ont montré des fonctions spécifiques pour chacun, soit au cours du développement, soit au cours de l'organogénèse, soit dans la tumorigénèse (**Tableau 5**) (Eferl and Wagner, 2003; Jochum et al., 2001). La surexpression dans la souris WT aborde la question du pouvoir oncogénique intrinsèque des membres AP-1, indépendamment de l'activation de cascades de signalisation. Il est à noter que dans ce système expérimental seules deux protéines de la famille FOS induisent des tumeurs : la surexpression de Fra-2 induit des tumeurs dans le pancréas, le thymus et les poumons tandis que celle de c-Fos induit des ostéosarcomes de faibles pénétrance et incidence. Du fait de la létalité de plusieurs invalidations phénotypiques (c-Jun, JunB, Fra-1 et Fra-2), il est difficile d'évaluer leur impact sur le processus tumorigénique ou anti-tumorigénique. Pour pallier la létalité, quelques études d'invalidation tissu-spécifique ont été menées (principalement par le groupe de EF Wagner). Il a été montré, dans un modèle de souris où le gène JunB a été invalidé spécifiquement dans l'épithélium de la prostate, que JunB coopère avec la voie de signalisation PTEN pour prévenir la cancérisation de la prostate, mettant en évidence le rôle suppresseur de

tumeur de JunB dans ce contexte cellulaire (Thomsen et al., 2015). JunB a également été décrit pour être un suppresseur de tumeur pour le lignage lymphoïde B (Szremska et al., 2003) et pour protéger les cellules souches hématopoïétiques de désordre myéloprolifératif (Passequé et al., 2004). L'invalidation de c-Jun dans les hépatocytes a mis en évidence que c-Jun promeut la tumorigenèse du foie induite par HBV (Trierweiler et al., 2016). Il est à noter également que, le background génétique peut influencer l'activité et l'oncogénicité des membres AP-1. En effet, alors que la délétion de c-Fos entraîne une ostéopétrose dans la souris WT, la délétion de c-Fos dans un background p53^{-/-} entraîne la formation de rhabdomyosarcomes de haute pénétrance, indiquant que c-Fos peut avoir des propriétés anti-oncogéniques dans ces conditions (Eferl and Wagner, 2003).

	Délétion		Surexpression	
	Phénotype	Organes et types cellulaires affectés	Phénotype	Organes et types cellulaires affectés
c-Fos	Ostéopétrose	Os, Ostéoclastes	Ostéosarcomes	Os, ostéoblastes
FosB	Défaut nutritionnel	Cerveau hypothalamus	RAS	RAS
Fra-1	Léthale au stade embryonnaire J9.5	Tissus extra embryonnaires, placenta	Osteosclérose	Os, ostéoclastes
Fra-2	Léthale à la naissance	Os, cœur, intestins	Malformations oculaires Formation de tumeurs	Structure antérieure de l'œil Pancréas, thymus, poumons
c-Jun	Léthale au stade embryonnaire J12.5	Foie, cœur, hépatoblastes, crête neurale	RAS	RAS
JunB	Léthale au stade embryonnaire J10	Tissus extra embryonnaires, placenta	RAS	RAS
JunD	Stérilité chez les mâles	Testicules, spermatides	Cellules T et B réduites	thymus, Thymocytes CD4

Tableau 5 : AP-1 dans le développement et la tumorigenèse chez la souris. Ce tableau récapitule les différents phénotypes observés chez la souris, suite à la délétion ou à la surexpression des différents membres AP-1. (RAS) aucun phénotype observé. *D'après Eferl & Wagner (2003) et Jochum et al. (2001).*

Il est important de noter que les gènes codant pour les protéines AP-1 n'ont jamais été retrouvés ni mutés ni amplifiés dans les tumeurs humaines, sauf dans le cas de c-Jun qui a été trouvé amplifié dans des liposarcomes humains hautement agressifs, le mécanisme tumoral sous-jacent étant le blocage de la différenciation adipocytaire induit par la surexpression de c-Jun (Mariani et al., 2007).

Dans la plupart des cas, les protéines AP-1 sont impliquées dans le cancer en tant qu'effecteurs indispensables à l'activité d'oncogènes activés en amont tels que Ras ou Raf (Ozanne et al., 2007). Cette activation constitutive par mutation ou amplification de voies de signalisation aboutit à la surexpression, à la stabilisation et à l'hyper-activation des membres AP-1, comme illustré au paragraphe II.2. Il en résulte une transcription exacerbée et/ou hors des fenêtres temporelles physiologiques de leurs gènes cibles, ce qui contribue à orienter le devenir cellulaire vers un phénotype tumoral. En effet, l'inhibition de l'activité AP-1 dans des lignées cancéreuses (voie Ras-MAPK activée) par des oligonucléotides anti-sens, des mutants dominants négatifs ou des siRNA dirigés contre certains membres AP-1 (c-Fos, c-Jun et Fra-1) peut réverter le phénotype transformé induit par Ras, établissant ainsi le rôle d'AP-1 dans la cancérisation (pour revue voir : (Ozanne et al., 2007)).

Le processus tumoral fait intervenir la dérégulation de nombreux processus cellulaires qui confèrent à la cellule tumorale de nouvelles propriétés lui permettant de s'affranchir d'un grand nombre de limitations (Hanahan and Weinberg, 2011). De nombreux travaux ont montré que les dimères AP-1 sont impliqués dans de nombreuses fonctions cellulaires qui sont dérégulées au cours de la tumorigenèse : prolifération, différenciation, apoptose, réponse au stress, invasion, angiogenèse, inflammation, transition épithélio-mésenchymateuse et qu'ils participent au processus tumoral. A titre d'exemple, c-Fos et Fra-1 régulent l'expression de métalloprotéinase matricielle (MMP1, MMP13) ainsi que celle de certains composants du système urokinase (uPA, uPAR et PAI-1) (Andrew et al., 2001; Belguise et al., 2005; Das et al., 2004; Moquet-Torcy et al., 2014; Park et al., 2011) qui sont impliqués dans la dégradation de la matrice extracellulaire et dans l'invasion tumorale (Milde-Langosch et al., 2004). De plus, c-Fos régule également l'expression de VEGF α dans des tumeurs murines ainsi que dans des lignées cellulaires mammaires humaines (Qin et al., 2015), alors que Fra-1 est impliqué dans l'EMT en modulant l'expression de Zeb1/2 dans des cellules de cancer du sein (Bakiri et al., 2015).

II.5. Fra-1, Fra-2 et cancer du sein

Parmi les protéines AP-1, il est maintenant bien établi que la surexpression de la protéine Fra-1 est étroitement associée à l'agressivité de divers cancers épithéliaux (ovaire, vessie, thyroïde, colon, poumon, etc.), dont les cancers du sein de type triple négatifs et basal-like (Dhillon and Tulchinsky, 2015; Milde-Langosch et al., 2004; Young and Colburn, 2006).

Concernant le cancer du sein, parmi les observations les plus marquantes, on peut noter que :

- (i) Les lignées cellulaires issues de cancers ER- sont caractérisées par une plus forte activité AP-1 largement liée à la surexpression de Fra-1, en comparaison aux lignées issues de cancers ER+ (Philips et al., 1998; Tam et al., 2013).

- (ii) La surexpression de Fra-1 dans des lignées de carcinomes mammaires peu agressifs entraîne une augmentation de la prolifération, de la motilité et de l'invasion (Belguise et al., 2005).
- (iii) À l'inverse, l'inhibition de l'expression de Fra-1 dans des lignées de cancers du sein agressifs diminue l'invasion et reverse le phénotype mésenchymateux (Belguise et al., 2005).
- (iv) L'expression ectopique de Fra-1 dans des cellules épithéliales murine non tumorigéniques est suffisante pour induire un programme mésenchymateux (EMT) invasif et tumorigénique sur des cellules en culture. L'expression de Fra-1 dans ces mêmes cellules xénogreffées chez la souris nude induit des tumeurs et des métastases (Bakiri et al., 2015).
- (v) L'expression de Fra-1 est nécessaire pour l'établissement du processus métastatique. La déplétion phénotypique (shRNA) de Fra-1 dans des cellules de cancer du sein xénogreffées chez la souris nude supprime leur potentiel métastatique (Desmet et al., 2013). La même observation a été faite dans un modèle de zebrafish xénogreffé avec des cellules de cancer du sein exprimant ou non un siRNA dirigé contre Fra-1 (Zhao et al., 2014).
- (vi) Une protection contre la progression tumorale et l'établissement des métastases a été montrée dans un modèle de cancer du sein chez la souris, suite à l'administration orale d'un vaccin ADN codant pour Fra-1, délivré par une souche atténuée de *S.Typhimurium* (Luo et al., 2003).
- (vii) Fra-1 semble être un élément central dans la régulation de l'EMT et dans la plasticité des cellules tumorales en contrôlant, entre autres, l'expression de facteurs de transcription et de microRNA régulant l'EMT (Dhillon and Tulchinsky, 2015; Tam et al., 2013).
- (viii) L'immunohistochimie montre une forte accumulation de Fra-1 dans les noyaux de cellules tumorales de coupes de carcinomes lobulaires et ductaux, qui contraste avec un marquage plus faible, à la fois cytoplasmique et nucléaire et n'affectant pas toutes les cellules, dans les adénomes mammaires et une absence totale dans les tissus sains, suggérant que Fra-1 pourrait être un marqueur intéressant pour les hyperplasies et néoplasies (Chiappetta et al., 2007).
- (ix) Allant dans le même sens, l'analyse des niveaux de transcrits des différents membres AP-1 dans des tumeurs du sein et leurs tissus adjacents a montré que l'expression de Fra-1, Fra-2, JunB et JunD est significativement augmentée dans les tumeurs, avec Fra-1 largement surexprimé dans les tumeurs ER- (Kharman-Biz et al., 2013).
- (x) L'analyse de « tissue microarrays » a montré qu'une forte expression de Fra-1 corrèle avec les cancers *invasifs* mais est sans réelle corrélation avec les cancers *in situ*. De plus, après

avoir classé les tumeurs par sous-types tumoraux, les auteurs ont montré que l'expression de Fra-1 est associée aux sous-types de plus mauvais pronostic : TNBC et basal-like (Logullo et al., 2011).

- (xi) Enfin, des analyses de banques de données cliniques ont fait apparaître qu'une forte expression de Fra-1 dans les tumeurs corrèle avec les patientes ayant un temps de survie sans métastase plus court, par rapport à celui des patientes dont les tumeurs expriment peu de Fra-1 (Bakiri et al., 2015; Tam et al., 2013).

Fra-2 a été beaucoup moins étudié que Fra-1. Cependant les travaux du groupe de Milde-Langosch ont montré que la surexpression de Fra-2, si elle n'affecte pas, ou peu, la prolifération cellulaire, augmente le pouvoir invasif des cellules cancéreuses (Milde-Langosch et al., 2008). La progression tumorale induite par Fra-2 passerait par un changement dans l'expression des protéines d'adhésion impliquées dans les contacts cellule-cellule et cellule-matrice extracellulaire (Schröder et al., 2010). D'autre part, une étude conduite sur 167 tumeurs primaires du sein a montré une corrélation positive entre l'expression de Fra-2, l'absence d'expression de ER, l'invasion nodale, le stade tumoral et la diminution de la survie sans rechute (Schröder et al., 2010).

L'ensemble de ces données suggèrent que Fra-1, et peut-être même Fra-2, pourraient être des cibles attractives pour une thérapie ciblée des cancers du sein métastatiques. Cependant, viser Fra-1 et Fra-2 eux-mêmes est exclu car il est pour le moment impossible d'inhiber directement ces facteurs de transcription par des approches thérapeutiques simples. Une alternative pourrait être de viser des cibles situées en aval des voies Fra-1 et Fra-2, c'est à dire de cibler les produits des gènes régulés par Fra-1 et/ou Fra-2.

Au démarrage de ma thèse, la caractérisation des transcriptomes régulés par Fra-1 et Fra-2 n'avait pas été abordée dans la littérature. De même, très peu d'éléments concernant les mécanismes transcriptionnels mis en jeu avaient été rapportés.

III. RÉGULATION DE LA TRANSCRIPTION : mécanismes généraux et nouveaux concepts

III.1. Structure tridimensionnelle de la chromatine

Depuis la fin du 20^{ème} siècle, le développement des techniques de microscopie et de biologie moléculaire pour l'étude de l'organisation spatiale du génome dans le noyau a permis de mettre en évidence l'implication de la topologie du génome dans la modulation de l'expression des gènes. La compréhension du fonctionnement des génomes nécessite de connaître leur séquence, mais également leur organisation structurale dans le noyau des cellules.

III.1.a. Organisation hiérarchique de la chromatine

La chromatine est une structure complexe qui présente plusieurs niveaux de compaction, permettant à environ 2 mètres d'ADN d'être contenus dans un noyau cellulaire ayant un diamètre de l'ordre d'une dizaine de microns chez les mammifères. La chromatine est très organisée dans le noyau, et cette organisation résulte de différents niveaux hiérarchiques de compaction qui sont présentés dans la **Figure 13** (pour revue voir (Hübner et al., 2013)).

Le premier niveau de compaction correspond à l'enroulement de l'ADN dans une structure nucléosomique pour former le nucléofilament. Le nucléosome est constitué d'environ 146 paires de bases (pb) d'ADN enroulé autour d'un octamère d'histones, composé de deux exemplaires des histones H2A, H2B, H3 et H4. Cette structure d'un diamètre de 10 nm et d'une hauteur de 5,5 nm constitue l'unité fondamentale de la chromatine. Les nucléosomes sont séparés les uns des autres par un segment d'ADN de longueur variable (10 à 100 pb). La répétition de ces structures forme le nucléofilament de 10 nm de diamètre dont la structure sous forme de collier de perles est visible en microscopie électronique.

Le second niveau de compaction correspond au passage de la fibre de 10 nm en une fibre de 30 nm. Les histones H1 sont capables de lier les ADN de liaison entrant et sortant des nucléosomes sur une trentaine de paires de bases et induisent le rapprochement des nucléosomes. Plusieurs modèles ont été proposés concernant la structure de la fibre de 30 nm. Deux grandes familles de modèles existent : l'une basée sur un axe virtuel autour duquel s'enroulent les nucléosomes successifs (modèle solénoïde) et l'autre dans lequel l'ADN de liaison traverse l'axe d'enroulement conduisant à un agencement des nucléosomes en zigzag (modèle zigzag).

La fibre de 30 nm se replie pour former des boucles de chromatine de tailles variables (150 à 200 kb) d'environ 300 nm, visualisées pendant l'interphase. Puis celles-ci s'enroulent en structures d'ordre supérieur successives pour conduire au chromosome métaphasique, qui est la forme la plus condensée de l'ADN.

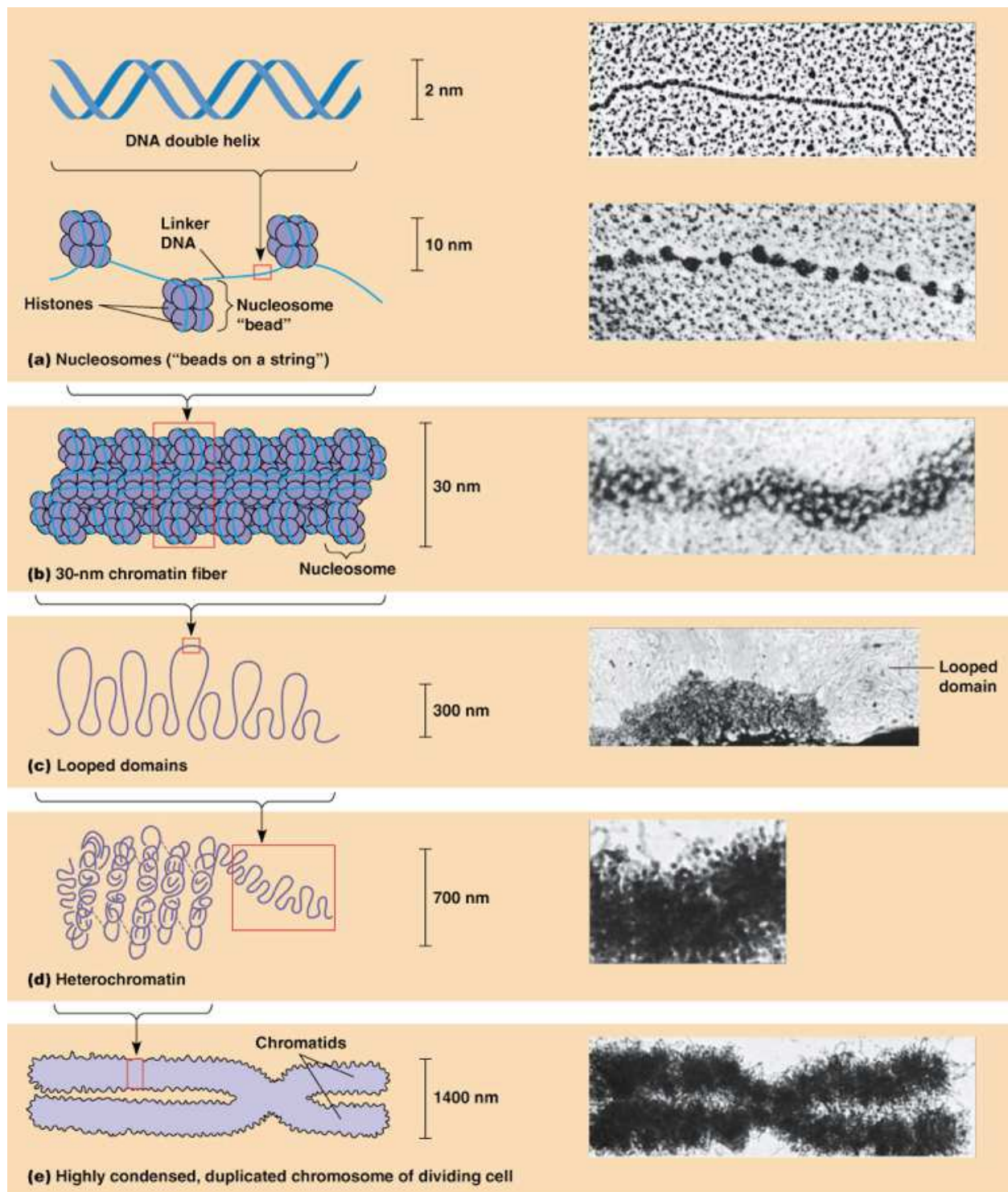


Figure 13 : Les différents niveaux hiérarchiques d'organisation de la chromatine dans un noyau de mammifère. Le premier niveau de compaction de l'ADN correspond à l'enroulement d'environ 200 paires de bases autour d'un octamère d'histones pour former le nucléosome. La succession de nucléosomes forme le nucléofilament aussi appelé « collier de perles ». Le nucléofilament s'enroule ensuite en structures d'ordre supérieur successives. La compaction maximale est présente dans les chromosomes métaphasiques. *Source :* <http://www.mun.ca/biology/desmid/brian/BIOL2060/BIOL2060-18/CB18.html>.

Parallèlement à ces niveaux hiérarchiques de compaction, différents niveaux d'organisation de la chromatine dans le noyau en interphase peuvent être observés, comme illustré dans la **Figure 14** (Phillips-Cremins, 2014). Dans la suite de cette partie je présenterai brièvement ces diverses structures.

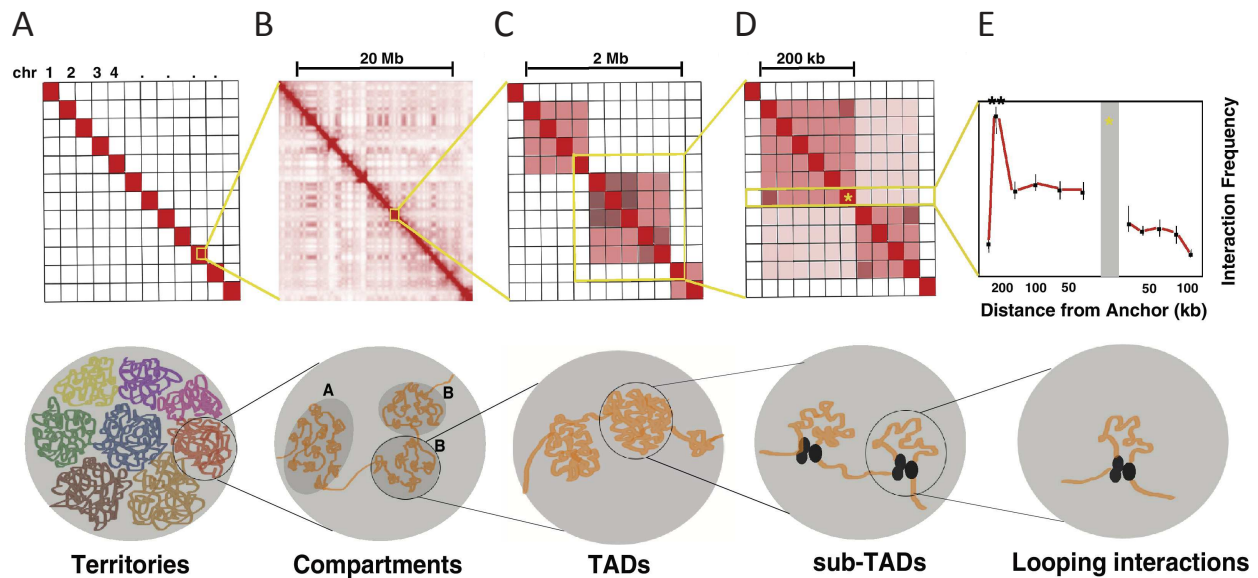


Figure 14 : Organisation hiérarchique du génome mammifère. Les matrices représentent les fréquences d'interactions, des plus faibles (carrés blancs) aux plus élevées (carrés rouges). **A)** Territoires chromosomiques. **B)** Compartiments « A » et « B » à l'intérieur d'un territoire chromosomique. **C)** Groupe de TAD à l'intérieur des compartiments « A » et « B ». **D)** Sous-domaines à l'intérieur des TAD. **E)** Boucle d'interaction qui peut se produire à l'intérieur d'un sous-domaine ou d'un TAD. *Source : Phillips-Cremins (2014).*

III.1.b. Territoires chromosomiques

Le noyau des cellules eucaryotes en interphase est une structure hautement compartimentalisée dans laquelle les chromosomes ne sont pas répartis au hasard. Des expériences de microscopie optique telles que le FISH (Fluorescence In Situ Hybridization), permettant de visualiser simultanément plusieurs chromosomes et d'analyser leur agencement les uns par rapport aux autres, ont révélé que la chromatine de chaque chromosome occupe un territoire nucléaire relativement circonscrit, appelé territoire chromosomique (**Figure 14.A**) (Bolzer et al., 2005; Cheung et al., 1977). Les territoires chromosomiques, bien que définis dans l'espace nucléaire, ne possèdent pas de bordures complètement franches et une certaine imbrication entre les chromosomes est possible. Des expériences de FISH montrent que des boucles provenant de différents territoires chromosomiques peuvent se rencontrer (Branco and Pombo, 2006).

Ces territoires chromosomiques ne sont pas disposés aléatoirement dans le volume nucléaire. Une corrélation a été observée entre la taille et la densité en gènes d'un chromosome et sa position radiale dans le noyau. Ainsi, les chromosomes denses en gènes sont préférentiellement observés au centre du noyau, tandis que ceux qui sont plus longs, souvent

moins denses en gènes, sont situés plus à la périphérie (Lieberman-Aiden et al., 2009). Ces observations ont abouti à une description dichotomique des formes de chromatine observables : l'euchromatine et l'hétérochromatine. L'euchromatine, d'aspect clair au microscope, est retrouvée préférentiellement au centre du noyau. Elle représente une chromatine ouverte, dense en gènes et accessible à la transcription. A l'inverse l'hétérochromatine, plus sombre, se positionne plutôt à la périphérie du noyau. Elle correspond à une chromatine très compactée et caractérisée par une plus faible densité en gènes. L'hétérochromatine peut être constitutive ou facultative. L'hétérochromatine constitutive correspond à des régions génomiques « muettes » qui ne s'expriment pas, telles que les régions répétées et les régions centromériques et télomériques. En revanche, l'hétérochromatine facultative correspond à des régions génomiques codantes qui peuvent s'exprimer dans certains types cellulaires ou selon l'état de différenciation d'une cellule donnée.

Depuis 2002, les approches de Capture de Conformation de Chromosome (3C) ont révolutionné l'étude de la topologie de la chromatine, en permettant de caractériser l'existence de contacts entre deux points du génome (une paire de loci) à l'intérieur du noyau (Dekker et al., 2002). Le principe repose sur une forme de « couper/coller ». La chromatine est d'abord figée par pontage au formaldéhyde et après digestion par une enzyme de restriction deux fragments de restriction, non consécutifs sur la séquence d'ADN mais physiquement proches dans l'espace nucléaire, sont susceptibles d'être « raboutés » grâce à l'action d'une ADN ligase, formant ainsi une molécule chimérique. Ceci permet de convertir une information tridimensionnelle (structure dans l'espace nucléaire) en une donnée unidimensionnelle (la séquence nucléotidique) accessible aux outils de la biologie moléculaire. A travers l'ensemble des chimères obtenues, la banque de produits de 3C représente l'ensemble des contacts qui existaient au moment où le pontage a été réalisé. Concrètement, il existe un nombre fini de fragments susceptibles de former l'une ou l'autre des extrémités des chimères. On parle donc de « librairie de 3C » ou, d'un point de vue mathématique, de « matrice de 3C ». Une décennie plus tard, de nombreuses techniques dérivées du 3C ont vu le jour (4C, 5C, Hi-C, ChIA-PET, etc), en particulier la technique de l'High throughput 3C (Hi-C) qui permet de générer des matrices d'interactions à l'échelle du génome, avec une résolution allant de la mégabase (Mb) à la kilobase (Kb) (Baù and Marti-Renom, 2012; Marti-Renom and Mirny, 2011; Rao et al., 2014).

Les expériences de Hi-C ont révélé une surreprésentation des contacts à l'intérieur d'un même chromosome (contacts intra-chromosomiques) par rapport aux contacts entre des chromosomes différents (contacts inter-chromosomiques) (Lieberman-Aiden et al., 2009) comme le laissaient déjà entrevoir les résultats de 4C (Simonis et al., 2006) et de 5C (Dostie et al., 2006). De plus, l'approche en cellule unique confirme que les interactions sont moins fréquentes quand la distance génomique augmente, et sont encore plus faibles pour les interactions inter-chromosomiques, ce qui est là aussi cohérent avec une organisation territoriale des chromosomes (Nagano et al., 2013).

III.1.c. Domaines chromosomiques (compartiments A et B)

Les expériences de Hi-C ont permis d'identifier un autre niveau d'organisation à l'intérieur de territoires chromosomiques, appelés domaines chromosomiques ou encore compartiments A et compartiments B, ayant une taille médiane d'environ 3 mégabases (**Figure 14.B**) (Lieberman-Aiden et al., 2009; Phillips-Cremins, 2014; Zhang et al., 2012).

Globalement, il est possible de séparer les régions actives (compartiments A) des régions inactives (compartiments B), le long d'un chromosome. Cette répartition en compartiments A et B s'avère très précise mais également très pertinente d'un point de vue fonctionnel. En effet, la séparation en régions actives/inactives intègre plusieurs types de données : transcriptomes, cartographies de modifications d'histones et de sites de fixation de protéines, et bien sûr cartographies d'interactions chromatinienne. Les compartiments A regroupent ainsi des régions riches en gènes, transcriptionnellement actives et hypersensibles à la DNase I, tandis que les compartiments B sont des régions plus pauvres en gènes, présentant moins de transcription et une sensibilité réduite à la DNase I. De plus, les contacts homologues (A-A ou B-B) sont largement favorisés à longue distance, alors qu'à l'inverse très peu de contacts A-B sont observés.

III.1.d. Domaines topologiquement associés (TAD)

Par la suite, l'utilisation de 5C et de Hi-C plus résolutive a mis en évidence le fractionnement des compartiments chromatinien en domaines topologiquement associés ou TAD (Topologically Associating Domain) (**Figure 14.C**). Leur définition repose sur un principe simple : les régions situées à l'intérieur d'un domaine ont tendance à interagir plus fréquemment entre elles et, à l'inverse, les interactions entre régions situées dans deux domaines différents sont moins fréquemment retrouvées. À la jonction entre deux TAD consécutifs il existe des régions dites barrières, où une chute brutale des interactions est observée. La délimitation en TAD s'accorde très bien avec les mécanismes de régulation et de co-régulation des gènes (Shen et al., 2012). Le repliement tridimensionnel du génome en domaines topologiques semble donc fonctionnel. Les génomes humains et murins sont composés de plus de 2000 TAD, de taille variable entre 1 et 2 Mb (Dixon et al., 2012). Par ailleurs, Phillips-Cremins *et al* ont également identifié des sous-domaines nichés à l'intérieur des TAD appelés sub-TAD et ayant une taille de l'ordre de 200 kb (**Figure 14.D**) (Phillips-Cremins, 2014; Phillips-Cremins et al., 2013).

L'existence des TAD a également été confirmée par des données de FISH à haute résolution. En effet, pour une distance génomique comparable, deux loci situés dans un même TAD sont plus proches dans l'espace que deux loci séparés par une bordure et donc situés dans deux TAD différents (Dixon et al., 2012; Nora et al., 2012).

Les frontières des TAD restent encore très mal caractérisées. Le premier isolateur qui a été identifié comme étant spécifiquement recruté sur ces frontières est CTCF (CCCTC-binding Factor), déjà amplement connu pour son rôle dans l'organisation en boucles de chromatine du génome. Cependant, CTCF se trouve aussi à l'intérieur des TAD eux-mêmes et donc la présence d'un site de liaison pour CTCF ne semble pas suffisante pour définir ce type de frontière. Plusieurs auteurs ont proposé que l'action de CTCF serait complétée par d'autres types de complexes protéiques, tels que les cohésines (Smc1 et Smc3) et le Mediator, afin d'établir de véritables frontières (**Figure 15**) (Kagey et al., 2010; Phillips-Cremins, 2014; Phillips-Cremins et al., 2013).

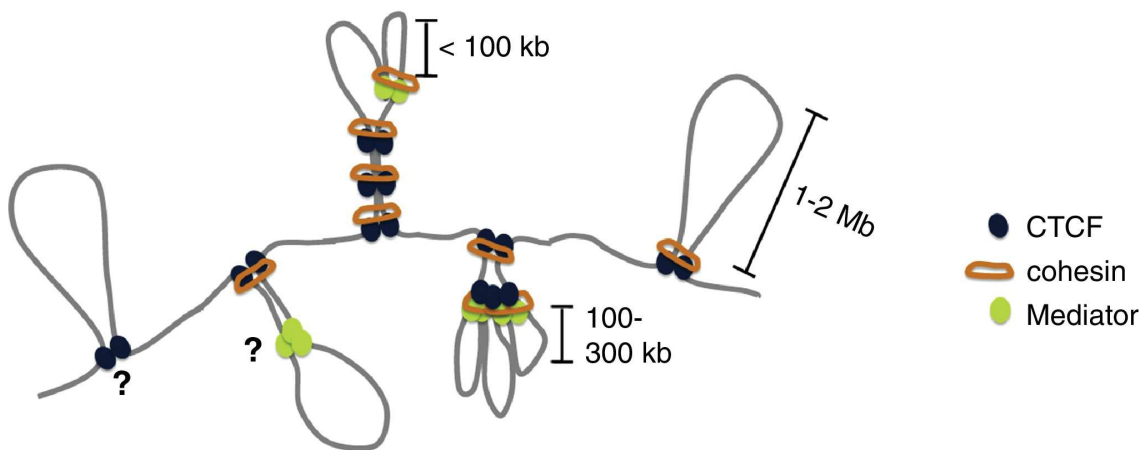


Figure 15 : Implications de CTCF, cohésine et Mediator à l'échelle des TAD. Diverses combinaisons de ces protéines seraient impliquées dans l'architecture des TAD : le duo cohésine/Mediator ancrerait des interactions spécifiques du type cellulaire à courte distance (< 100 kb) ; l'ajout de sites de fixation pour CTCF organiserait les sous-domaines (< 1 Mb) ; CTCF et le complexe cohésine réguleraient les interactions à plus longue distance (1 à 2 Mb). *Source : Phillips-Cremins (2014).*

III.1.e. Boucles de chromatine (DNA looping)

La fibre de chromatine est capable de former différents types de boucles qui ont chacune un rôle distinct selon l'échelle spatiale considérée. En effet, les boucles de chromatine constituent des structures importantes pour les interactions et les régulations existantes entre les activateurs transcriptionnels (enhancers) et les promoteurs. Si les TAD ont une taille variant entre 1 et 2 Mb, l'échelle classique pour les interactions enhancer-promoteur est de l'ordre de 10 à 200 kb (Chung et al., 2016; Phillips-Cremins et al., 2014). Ainsi, ces boucles de régulation enhancers-promoteurs auraient lieu au sein des sub-TAD (**Figure 14.E**).

Un enhancer est capable d'interagir avec plusieurs promoteurs et inversement, multipliant ainsi les possibilités de régulation mais aussi les possibilités d'organisation tridimensionnelle du génome. Les interactions entre enhancers et promoteurs et leur implication dans la régulation de l'expression des gènes seront développées plus loin dans la suite de ce chapitre (voir paragraphe III.5).

III.2. Dynamique de la structure chromatinienne et régulation de la transcription

Comme mentionné précédemment, l'enroulement de l'ADN autour des histones pour former le nucléosome constitue un premier niveau de compaction. Il instaure par la même occasion une contrainte dès lors que l'accessibilité de l'ADN à la machinerie de transcription est requise. Le contrôle de la structure chromatinienne est donc essentiel au contrôle de l'expression génique.

Bien que les nucléosomes soient des complexes protéines-ADN stables, ils ne sont pas statiques et subissent des réarrangements structuraux aboutissant à leur déplacement le long de la fibre chromatinienne (glissement) ou leur disparition transitoire, pour former des régions dépourvues de nucléosomes. Ce changement d'état des nucléosomes est dépendant de facteurs protéiques qui peuvent modifier post-traductionnellement les histones (pour revue voir (Bell et al., 2011)).

III.2.a. Régions dépourvues de nucléosomes ou NFR

Pour pouvoir s'enrouler autour des histones en structure nucléosomique, la séquence d'ADN doit avoir une certaine flexibilité de courbure. Les séquences rigides telles que les poly dA/dT (expl : TATAAA, TATATA et AAAAAA) (Ozsolak et al., 2007) sont des motifs appelés « nucleosome exclusion sequences » (Radwan et al., 2008), car l'enroulement de ces séquences d'ADN autour des histones demanderait beaucoup d'énergie par rapport à des séquences relativement aléatoires. Elles forment donc des régions dépourvues de nucléosomes ou NFR (Nucleosome Free Region) de manière constitutive. Cependant, les NFR constitutifs sont rares dans la cellule.

La formation de NFR est généralement un processus actif qui fait intervenir des enzymes de remodelage de la chromatine (ou remodelers) suite à leur recrutement par les facteurs de transcription (Becker and Hörz, 2002; Clapier and Cairns, 2009). Deux types d'enzymes de remodelage de la chromatine ont été identifiés : celles qui organisent la chromatine et celles qui la désorganisent. Ainsi, la plupart des protéines de la famille SWI/SNF interviennent dans le désassemblage des nucléosomes (elles agissent donc comme activateur de la transcription) tandis que celles de la famille ISWI interviennent plutôt dans l'assemblage des nucléosomes (rôle répresseur de la transcription) (Clapier and Cairns, 2009).

Différentes expériences de sensibilité à la DNase I (Dnase I hypersensitivity ou DHS) ont montré qu'en général les NFR contiennent des sites de liaison pour les facteurs de transcription. De plus, les phases ouvertes de lecture sont généralement caractérisées par des régions de moins grande densité en nucléosomes (Heintzman et al., 2007; Nishida et al., 2006; Ozsolak et al., 2007).

III.2.b. Modifications post-traductionnelles d'histones

Les domaines N-terminaux des histones formant le nucléosome sont flexibles et dirigées vers l'extérieur du nucléosome. Ces régions N-terminales sont sujettes à de nombreuses modifications post-traductionnelles qui jouent un rôle très important dans la régulation de la transcription. Les conséquences fonctionnelles de ces modifications dépendent du type de groupement chimique apposé (méthyle, acétyle, phosphate, ubiquitine, glycosyle, sumo, etc.) ainsi que du type de résidu ciblé (lysine K, arginine R, sérine S, etc.) et de sa position le long de la région N-terminale. La diversité des groupements et des résidus qui peuvent être ciblés donne une combinatoire riche. Ces modifications influencent, d'une part, la compaction et l'accessibilité de l'ADN et, d'autre part, créent des motifs de liaison pour des effecteurs transcriptionnels.

La caractérisation systématique des modifications d'histone a conduit à proposer l'existence d'un code histone qui associerait chaque modification à un état transcriptionnel donné (Jenuwein and Allis, 2001). Cependant, il semble que l'effet dû à la combinaison de plusieurs marques soit plus important que celui de marques isolées, ce qui complexifie l'étude des conséquences de ces modifications individuelles. En effet, ces modifications sont généralement dépendantes les unes des autres, soit positivement soit négativement, et peuvent générer des affinités d'interactions synergiques ou antagonistes pour des protéines qui vont participer à la transition dynamique entre un état transcriptionnellement inactif et un état actif.

Les fonctions connues des différentes modifications d'histones ont été résumées par T. Kouzarides (Kouzarides, 2007). Les deux modifications post-traductionnelles des histones les plus étudiées, l'acétylation et la méthylation, sont détaillées ci-dessous.

* Acétylation des histones :

L'acétylation des histones est généralement associée à une transcription active des gènes (Heintzman et al., 2007, 2009). L'acétylation des résidus lysine des régions N-terminales des histones est une modification très dynamique, qui dépend de l'équilibre entre deux activités antagonistes : l'activité histone acétyltransférase (HAT) et l'activité histone déacétylase (HDAC). Il est généralement admis que ces enzymes sont recrutées, respectivement, sur les gènes à activer ou à éteindre par les différents facteurs de régulation de la transcription.

En fonction des signaux reçus, un même facteur de transcription peut recruter différentes HAT ou HDAC. Par exemple, dans le cas de Fra-1, il a été montré que l'activation de la transcription du gène de l'involucrine dans les kératinocytes passait par le recrutement de P300 et PCAF (Crish and Eckert, 2008), tandis que l'inhibition de la transcription de l'IL-8 passe par le recrutement de HDAC1 (Hoffmann et al., 2005). De plus, certaines HAT peuvent à la fois avoir des rôles redondants pour un même résidu lysine et être spécifiques pour d'autres, c'est le cas par exemple pour P300 et PCAF. Si P300 acétyle spécifiquement le résidu lysine 5 de l'histone H4 (H4K5) et PCAF le résidu lysine 9 de l'histone H3 (H3K9), elles peuvent toutes les

deux acétyler les lysines 14 et 18 de l'histone H3 (H3K14/18).

A l'heure actuelle, chez l'Homme, plus de 20 résidus lysine ont été caractérisés pour être acétylables. Parmi les différentes acétylations d'histones, l'acétylation de l'histone H3 sur la lysine 27 (H3K27ac) est particulièrement utilisée pour discriminer si une région régulatrice (promoteur ou enhancer) est active ou inactive transcriptionnellement (**Figure 16**) (Pundhir et al., 2015).

** Méthylation des histones :*

La méthylation affecte des résidus lysine et arginine des régions N-terminales des histones. Cette réaction est catalysée par des histones méthyltransférases, tandis que le retrait des groupements méthyl est assuré par les histones déméthylases. Contrairement à l'acétylation, la méthylation des histones est beaucoup plus stable et ne modifie pas la charge des acides aminés concernés. Elle n'entraîne donc pas une diminution de l'interaction électrostatique entre région N-terminale et ADN. Par contre, elle permet la reconnaissance par des protéines possédant un chromodomaine qui peuvent, par elles-mêmes ou par recrutement de partenaires, modifier l'état de compaction de la chromatine (He and Lehming, 2003; Zhou et al., 2011).

Contrairement à l'acétylation des histones qui corrèle presque exclusivement avec l'activation transcriptionnelle, le rôle de la méthylation est beaucoup plus complexe (Barski et al., 2007). En effet, un même résidu peut être mono-, di- ou tri-méthylé et, suivant son degré de méthylation, l'effet sur la transcription peut être différent. De plus, la méthylation peut avoir lieu à différents endroits du gène (promoteur, enhancer ou séquence codante) (Rando and Chang, 2009; Zhou et al., 2011).

Parmi les différentes méthylations d'histones, trois sont particulièrement utilisées pour caractériser l'anatomie et l'activité des gènes : la mono-méthylation de l'histone H3 sur la lysine 4 (H3K4me1), la tri-méthylation de l'histone H3 sur la lysine 4 (H3K4me3) et la tri-méthylation de l'histone H3 sur la lysine 36 (H3K36me3) (Heintzman et al., 2007, 2009; Ng et al., 1997; Pundhir et al., 2015). Les profils de H3K4me1 et H3K4me3 permettent de différencier les séquences de régulation, à savoir enhancer ou promoteur. En effet, aujourd'hui il est admis que H3K4me3 est fortement présente aux niveaux des régions promotrices des gènes et est très significativement associée à l'activation de la transcription, tandis que H3K4me1 est présente aux niveaux des régions enhancers (**Figure 16**) (Pundhir et al., 2015). Le profil H3K36me3 quant à lui est associé à une transcription active au niveau de la région codante du gène. En effet, cette modification est déposée sur les histones présentes dans la séquence codante du gène par une enzyme associée à l'ARN Pol II lors de l'étape d'élongation de la transcription (Carrozza et al., 2005; Keogh et al., 2005). Elle est généralement peu abondante dans la région promotrice et très abondante sur la séquence codante des gènes transcrits. La superposition de ces différentes modifications renseigne à la fois sur l'anatomie du gène et son activité.

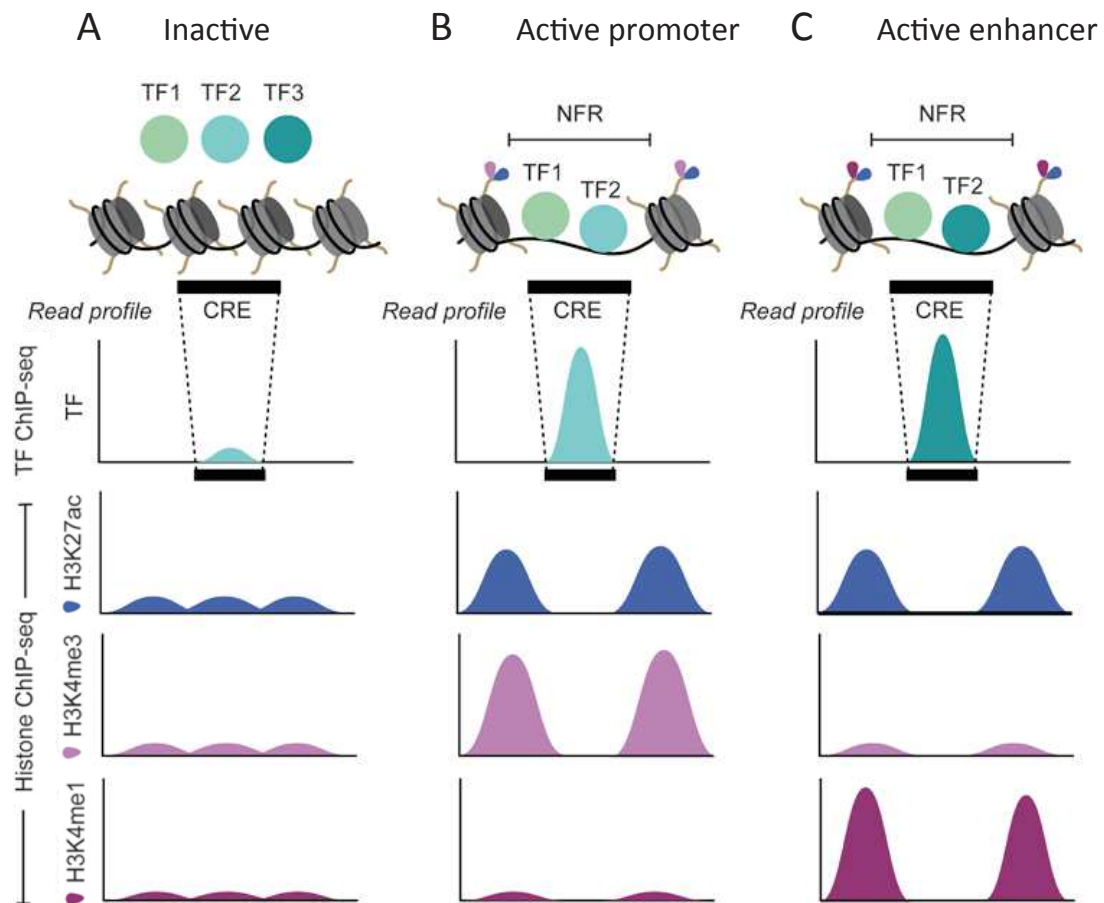


Figure 16 : Modifications d’histones caractéristiques de l’activité des séquences régulatrices enhancer et promoteur. (TF) facteurs de transcription ; (CRE) cis-regulatory elements ou élément de régulation en position cis ; (NFR) nucleosome free region ou région dépourvue de nucléosome. *Source : Pundhir et al. (2015).*

Les modifications d’histones ainsi qu’un certain nombre de protéines interagissant avec la fibre chromatinienne ont pu être caractérisées à l’échelle du génome par immunoprécipitation de chromatine couplé au séquençage à haut débit (ChIP-seq) (Furey, 2012). Ces données sont répertoriées au sein du projet Encode (Encyclopedia of DNA elements, www.encodeproject.org/ENCODE/).

III.3. Les différentes étapes de la transcription

La transcription est un mécanisme très élaboré et complexe dont certains aspects ont été élucidés mais dont les mécanismes intimes ne sont pas encore tous caractérisés.

La transcription comprend plusieurs étapes successives qui permettent de réguler finement la transcription des gènes en réponse aux signaux reçus. Plusieurs étapes limitantes de la transcription ont été définies : l'assemblage du complexe de pré-initiation, l'initiation de la transcription, le départ de l'ARN polymérase du promoteur sur la séquence codante, l'élongation transcriptionnelle et l'étape de terminaison de la transcription (Fuda et al., 2009). Bien que toutes ces étapes puissent être régulées, la régulation se fait principalement au niveau des premières étapes de la transcription pour la plupart des gènes.

Plusieurs complexes protéiques nécessaires à une initiation correcte de la transcription ont été identifiés et regroupés sous le nom de facteurs généraux de transcription (Buratowski et al., 1989). On sait maintenant que ces facteurs interviennent dans la transcription basale et, s'ils sont nécessaires, ils ne sont pas suffisants pour une régulation spatio-temporelle de la transcription des gènes induits. En effet, cette dernière met en jeu des facteurs de transcription synthétisés ou activés de manière spécifique aux tissus ou en réponse à des stimuli particuliers.

III.3.a. Transcription basale et induite

La transcription basale est sous le contrôle du promoteur minimal ou « core promoter », qui se situe en général entre -30 et +30 pb du TSS (Transcription Start Site). Il a été défini comme la plus petite unité nécessaire et suffisante à l'initiation de la transcription par la machinerie de transcription basale, c'est à dire l'ARN polymérase II (Pol II) et les facteurs généraux de transcription (**Figure 17**). Le promoteur minimal a été caractérisé *in vitro*. Il contient des éléments de séquences consensus (boîte TATA, Inr, MTE et DPE) (Butler and Kadonaga, 2002) (pour revue voir (Juven-Gershon and Kadonaga, 2010)). La boîte TATA est une courte région riche en A-T qui se situe généralement entre 25 et 30 pb en amont du site +1 (TSS) (Lee and Young, 2000). Elle permet de positionner la machinerie de transcription par rapport au site d'initiation de la transcription. La boîte TATA est le site de liaison de TBP (TATA-Binding Protein) qui fait partie du facteur général de transcription TFIID. La fixation de TFIID sur la boîte TATA est la première étape dans le recrutement de la machinerie de transcription au promoteur minimal. A l'origine, le promoteur minimal avait été décrit comme un motif quasi-universel pour l'initiation de la transcription. On sait aujourd'hui que les promoteurs minimaux possèdent des structures et des fonctions qui varient en fonction des gènes, de très nombreux promoteurs n'ayant pas de boîtes TATA identifiables. Ces promoteurs possèdent d'autres motifs tels que le motif DPE (Downstream Promoter Element), le motif Inr (Initiator élément) ou le motif BRE (élément reconnu par TFIIB) (Juven-Gershon et al., 2008; Smale, 1997). Ces différents éléments peuvent être reconnus par différents facteurs généraux de transcription afin

de faciliter le recrutement et le positionnement de la machinerie de transcription. Par exemple, certaines sous-unités de TFIID (TAFs ou TFIID Associated Factors) peuvent interagir avec les Inr ou les DPE. Certains promoteurs peuvent posséder l'un et/ou l'autre de ces éléments, mais jamais tous en même temps (Sikorski and Buratowski, 2009). D'autres gènes sont dépourvus de ces éléments mais en général, les promoteurs sans éléments possèdent de multiples sites d'initiation de la transcription (Geng and Johnson, 1993; Lee and Young, 2000).

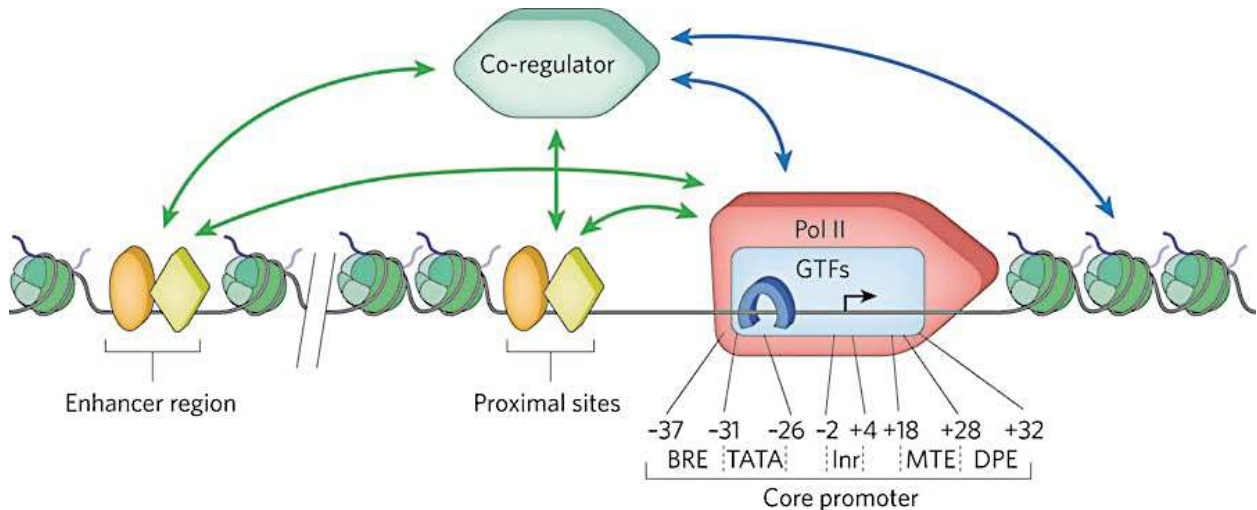


Figure 17 : Interactions régulatrices de la transcription. Les facteurs généraux de transcription (GTFs) se lient à des éléments spécifiques du promoteur. Ces éléments, l'élément de reconnaissance par TFIIB (BRE), la TATA box (TATA), l'élément initiateur (Inr), l'élément Motif Ten (MTE) et l'élément en aval du promoteur (DPE) et leurs positions relatives au site de démarrage de la transcription (transcription start site ou TSS : flèche noire) sont représentés. Les régulateurs transcriptionnels (ovale orange et losange jaune), qui peuvent être soit des activateurs soit des répresseurs, se lient à des séquences d'ADN spécifiques proches du promoteur proximal ou sur des régions distantes appelées enhancers ou silencers. Les régulateurs peuvent interagir avec les GTFs tels que TFIID (rectangle bleu) et TBP (TATA-binding protein, fer à cheval) et le complexe Pol II (missile rouge) pour stimuler ou réprimer la transcription. Ils interagissent aussi avec des co-régulateurs (hexagone vert) qui peuvent interagir avec la machinerie de transcription basale ou des facteurs de modification de la chromatine, tels que les enzymes de modification des histones ou les enzymes de remodelage de la chromatine. Les co-régulateurs peuvent aussi se lier aux nucléosomes (vert) modifiés, stabilisant ainsi la liaison du co-régulateur sur le gène. Les activateurs peuvent recruter, stabiliser ou stimuler ces facteurs tandis que les répresseurs peuvent déstabiliser ou inhiber ces facteurs. *D'après Fuda et al. (2009).*

La transcription, induite ou stimulée, est sous le contrôle du promoteur proximal ainsi que d'éléments de régulation distaux (enhancers et silencers) et est régulée par des facteurs de transcription spécifiques du gène à transcrire (Fuda et al., 2009). Le promoteur dit proximal se situe entre -200 et -30 pb du TSS (**Figure 17**). Outre le promoteur minimal, il contient différents sites de liaison (i) appelés UAS (Upstream Activating Sequences) pour les facteurs de transcription qui, s'ils fixent les facteurs de transcription adéquats, activent la transcription, ou (ii) URS (Upstream Repressing Sequences) qui, après fixation des facteurs ad hoc, répriment la transcription. En fonction du type cellulaire et des signaux reçus, différentes combinaisons de facteurs de transcription capables de lier ces séquences régulatrices sont exprimées. La liaison des différentes combinaisons de facteurs de transcription sur ces séquences proximales en coopération avec le promoteur minimal régule positivement ou négativement la transcription du

gène (Lee and Young, 2000; Maston et al., 2006). Il existe d'autres séquences de régulation distales qui sont généralement composées de sites de liaison plus ou moins nombreux pour les facteurs de transcription : les enhancers ou silencers (**Figure 17**). Ces séquences peuvent réguler de façon positive (enhancers) ou négative (silencers) la transcription des gènes, quelles que soient leurs positions (en cis, en trans, en amont, en aval ou à l'intérieur du gène) ou leurs orientations dans le génome (Ong and Corces, 2011).

III.3.b. Formation du complexe de pré-initiation

Des expériences de transcription *in vitro* sur de l'ADN nu (promoteur tardif de l'adénovirus ou adénovirus Major Late Promoter) ont montré que la formation d'un complexe de pré-initiation (PIC) précède la transcription (Buratowski et al., 1989). Le PIC est nécessaire et suffisant pour une transcription basale des gènes. Le PIC est composé de six complexes protéiques appelés facteurs généraux de transcription : TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH (**Tableau 6**), et est formé par le recrutement successif de ces différents facteurs et de l'ARN Pol II (Buratowski et al., 1989). TFIID est le premier à être recruté au niveau de la TATA box du promoteur. Par la suite, les autres recrutements se font dans l'ordre TFIIA, TFIIB, Pol II/TFIIF, TFIIE et TFIIH.

Le PIC peut aussi être formé en une étape. Des expériences de purification *in vitro* de l'ARN Pol II ont aussi montré qu'elle peut interagir avec tous les facteurs généraux de transcription pour former une holoenzyme, indépendamment de sa liaison à l'ADN. Dans ce cas l'holoenzyme peut être recrutée au promoteur en une seule étape (Koleske and Young, 1994).

Bien que les facteurs généraux de transcription soient nécessaires et suffisants pour une initiation correcte de la transcription *in vitro*, la situation *in vivo* est nettement plus complexe (Robinson and Kadonaga, 1998). Du fait de l'organisation des gènes dans la chromatine mais aussi de leur capacité à répondre à différents stimuli, d'autres facteurs sont nécessaires à la formation du PIC. Maintenant, il est bien établi que les facteurs de transcription peuvent intervenir dans la formation du PIC en interagissant avec ses différents composants. Grâce à leurs domaines de transactivation, les facteurs de transcription liés à l'ADN peuvent stimuler le recrutement d'un ou de plusieurs composants du PIC, ce qui a pour conséquence de faciliter la liaison de la machinerie de transcription et sa stabilisation au promoteur. Les facteurs de transcription peuvent aussi recruter les composants du PIC de façon indirecte grâce à des ponts moléculaires. Le Mediator a été isolé comme l'un de ces co-activateurs permettant de stimuler la transcription des gènes, en facilitant le recrutement de Pol II et en stimulant son départ du promoteur. Plusieurs expériences ont aussi montré qu'en absence de facteur de transcription, le Mediator peut aussi contribuer à la transcription basale (Sikorski and Buratowski, 2009).

Tableau 2: Les facteurs généraux de la transcription chez l'homme.			
facteur	sous-unités	poids moléculaire	Caractéristiques/fonctions
TFIIA	TFIIA α	37 kDa	TFIIA- α et - β résultent du clivage d'un seul peptide. Ils sont impliqués dans la stabilisation de l'interaction TBP-ADN et sont nécessaires à l'activation transcriptionnelle.
	TFIIA β		
	TFIIA γ	13 kDa	Impliqué dans l'activation et l'anti-répression transcriptionnelle.
TFIIB	TFIIB	35 kDa	Impliqué dans la sélection du point +1 et la stabilisation de l'interaction TBP-ADN
TFIID	TBP	38 kDa	Reconnait la boîte TATA
	TAF1 (TAFII250)	250 kDa	Impliqué dans la fixation du promoteur. Possède des activités histone acétyltransférase, ubiquitine-ligase et kinase.
	TAF2 (TAFII150)	150 kDa	Impliqué dans la fixation du promoteur en interagissant avec l'élément DPE.
	TAF3 (TAFII140)	140 kDa	Impliqué dans la fixation du promoteur. Reconnait spécifiquement la lysine 4 de l'histone H3 triméthylée.
	TAF4 (TAFII130/135)	135 kDa	Impliqué dans des interactions avec des coactivateurs
	TAF5 (TAFII100)	100 kDa	Impliqué dans l'assemblage du complexe TFIIF
	TAF6 (TAFII80)	80 kDa	Possède une structure de type histone
	TAF7 (TAFII55)	55 kDa	Interagit avec des activateurs
	TAF8 (TAFII43)	43 kDa	Possède une structure de type histone
	TAF9 (TAFII31/32)	32 kDa	Possède une structure de type histone
	TAF10 (TAFII30)	30 kDa	Interagit avec des activateurs
	TAF11 (TAFII28)	28 kDa	Possède une structure de type histone
	TAF12 (TAFII15/20)	20 kDa	Possède une structure de type histone
	TAF13 (TAFII18)	18 kDa	Possède une structure de type histone
	TAF15 (TAFII68)	68 kDa	Possède un domaine de liaison à l'ARN
TFIIE	TFIIE α	56 kDa	Interagit avec et stimule TFIIF. Impliqué dans l'ouverture du promoteur.
	TFIIE β	34 kDa	Interagit avec l'ADN double-brin
TFIIF	RAP74	58 kDa	Interagit avec l'ADN pendant l'initiation.
	RAP30	26 kDa	Interagit avec la Pol II. Impliqué dans l'élongation
TFIIH	p62	62 kDa	Essentielle à la transcription et à la réparation de l'ADN (NER)
	p52	52 kDa	Essentielle à la transcription et à la réparation de l'ADN (NER)
	MAT1	32 kDa	Impliqué dans la réparation de l'ADN (NER). Régule l'activité du complexe Cdk7/CycH
	p34	34 kDa	Impliqué dans la réparation de l'ADN (NER). Joue un rôle structurale dans le complexe TFIIF
	XPD/ERCC2	80 kDa	Activité 5'-3' hélicase et ATPase. Impliqué dans la réparation de l'ADN (NER).
	p44	44 kDa	Stimule l'activité hélicase de XPD. Impliqué dans la réparation de l'ADN (NER).
	XPB/ERCC3	89 kDa	Activité 3'-5' hélicase et ATPase dépendante de l'ADN. Essentielle à l'ouverture et au dégagement du promoteur.
	Cdk7	40 kDa	Sous-unité kinase du complexe Cdk7/CycH qui phosphoryle le CTD de la Pol II.
	Cycline H	38 kDa	Sous-unité régulatrice du complexe Cdk7/CycH qui phosphoryle le CTD de la Pol II.

Tableau 6 : Compositions et fonctions des facteurs généraux de transcription chez l'Homme. Les facteurs généraux de transcription sont des protéines ou complexes protéiques indispensables à la fixation et à la bonne fonction de Pol II sur le promoteur. Ce tableau détaille leurs compositions et leurs fonctions. *Source : (Thomas and Chiang (2006))*

III.3.c. Initiation de la transcription

La formation du complexe d'initiation est suivie du relargage de la machinerie de transcription du promoteur pour initier la transcription. Cette initiation de la transcription se fait en plusieurs étapes successives. Premièrement, TFIIF qui possède une activité hélicase ouvre l'ADN double brin au niveau des nucléotides -9 à +2 pour former la bulle de transcription. Pendant l'incorporation des 9 premiers nucléotides, le complexe transcriptionnel est assez instable et enclin à produire des transcrits abortifs. A chaque ajout de nucléotide, il y a une compétition entre dissociation et ajout du prochain nucléotide. Ce n'est qu'entre 9 et 11

nucléotides que Pol II est stabilisée et entre dans une phase de synthèse productive. Par la suite, le déplacement de la bulle de transcription en aval permet de quitter le promoteur. Cette étape correspond au passage de l'étape d'initiation à l'étape d'élongation (Conaway et al., 2000).

III.3.d. Transition entre initiation et élongation précoce

La transition de l'initiation de la transcription à l'élongation précoce s'accompagne d'un changement de facteurs finement contrôlé par la phosphorylation du domaine C-terminal (CTD) de sous-unité RPB1 de l'ARN Pol II (voir plus loin, paragraphe III.4). Lors de la formation du PIC, ce dernier est sous forme non-phosphorylée, ce qui favorise son interaction avec des facteurs qui interviennent dans la stabilisation du PIC. Il a été montré que le Mediator interagit avec TFIID et la forme non-phosphorylée du CTD de Pol II (Conaway and Conaway, 2011; Conaway et al., 2005; Mittler et al., 2003; Yang et al., 2004). Cette interaction avec le Mediator stimule l'activité kinase de TFIID, portée par CDK7, qui va alors phosphoryler la sérine-5 du CTD de Pol II. Cette phosphorylation déstabilise l'interaction entre Pol II et les facteurs liés au promoteur (dont le Mediator). La phosphorylation du CTD sur la sérine-5 corrèle donc à son échappée du promoteur. Ainsi, une fois que le Mediator a accompli sa fonction de recrutement de Pol II, il participe aussi à son départ sur la séquence à transcrire (pour revue voir (Nechaev and Adelman, 2011)).

III.3.e. Élongation et notion de pause transcriptionnelle

Suite à l'initiation de la transcription, Pol II a tendance à s'arrêter à environ 20-40 nucléotides en aval du promoteur. Des études à l'échelle du génome entier, dans des cellules de mammifères ou chez la drosophile, ont montré que 20 à 30% des gènes possèdent Pol II sur leur promoteur sans, pour autant, que les gènes soient transcrits. Ce mécanisme est appelé pause transcriptionnelle. De façon intéressante, la pause à l'élongation affecte prioritairement des gènes finement régulés comme ceux impliqués dans le développement ou ceux qui répondent de façon très réactive aux stimuli environnementaux.

Après l'initiation de la transcription, Pol II est bloquée à proximité du promoteur via son interaction avec deux régulateurs négatifs de l'élongation, NELF et DSIF. Le recrutement de P-TEFb (CycT1/CDK9) est nécessaire à la levée du blocage induit par NELF et DSIF. En plus de phosphoryler la sérine-2 du CTD de Pol II, la sous-unité P-TEFb phosphoryle DSIF et NELF. Ces phosphorylations entraînent le relargage de NELF et la levée du blocage à l'élongation. La phosphorylation de DSIF le convertit en activateur transcriptionnel qui voyage avec P-TEFb et Pol II en cours d'élongation. Il a été montré que la régulation du recrutement de P-TEFb peut se faire via son interaction avec des facteurs de transcription tels que NF-KB (Barboric et al., 2001), P53 (Gomes et al., 2006) ou c-Myc (Eberhardy and Farnham, 2002; Rahl et al., 2010).

P-TEFb est régulé par des modifications post-traductionnelles. Par exemple, l'ubiquitylation de CycT1 affecte sa stabilité, et la phosphorylation de la thréonine 186 de CDK9 stimule son activité kinase (Kiernan et al., 2001; Li et al., 2005). Un autre niveau de régulation important de l'activité de P-TEFb passe par son interaction réversible avec un complexe nucléoprotéique inhibiteur formé par la protéine HEXIM1 (Michels et al., 2004) en association avec un petit ARN nucléaire 7SK (Nguyen et al., 2001). L'association de P-TEFb avec ce complexe inhibe son activité kinase. En 2009, les travaux de Cho *et al* (Cho et al., 2009) ont montré que la dissociation du complexe inhibiteur et l'activation subséquente de P-TEFb sont liées à l'acétylation de la CycT1 par l'histone acétyltransférase P300 (**Figure 18**).

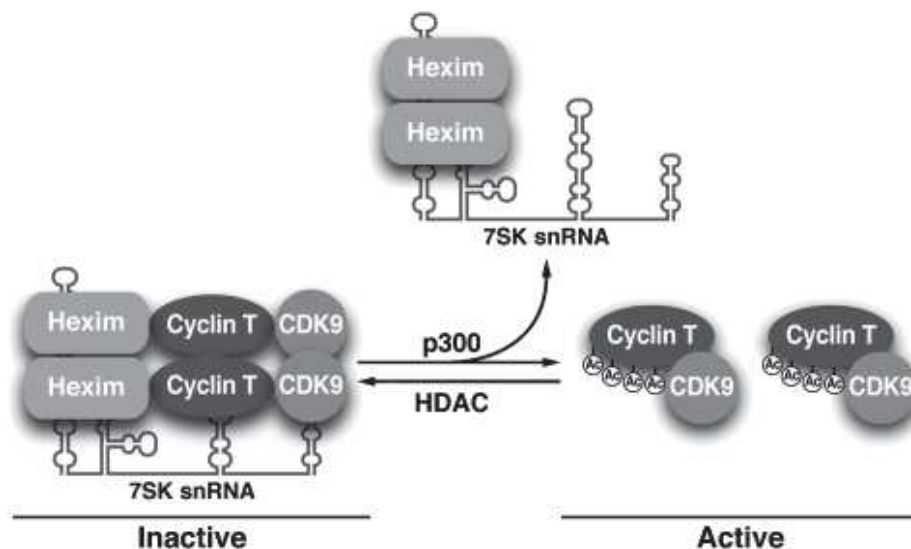


Figure 18 : Régulation de l'activité P-TEFb par P300. L'acétylation réversible de la Cycline T1 par P300 régule la séquestration de P-TEFb dans un complexe inactif (lié à Hexim 1/ARN7sk) ou actif (non lié à Hexim 1/ARN7sk).
Source : Cho et al. (2009).

III.4. Régulation de la transcription par l'ARN Pol II

Chez les eucaryotes, la transcription de l'ADN en ARN est réalisée par trois ARN polymérases selon le type d'ARN transcrit : classiquement, l'ARN polymérase I transcrit les gènes codants pour les ARN ribosomiques, l'ARN polymérase II (Pol II) produit essentiellement les ARN pré-messagers codant les protéines, tandis que l'ARN polymérase III transcrit les gènes codant de petits ARN, tels les ARN de transfert. Nous ne parlerons ici que de la transcription par Pol II.

III.4.a. Domaine C-terminal (CTD) de l'ARN Pol II et phosphorylations

L'ARN Pol II est un complexe multiprotéique d'environ 500 KDa composé de 12 sous-unités numérotées RPB1 à 12. Parmi ces 12 sous-unités, la sous-unité RPB1, qui est la plus grosse, n'est pas nécessaire à la catalyse de la synthèse des ARN *in vitro* mais assure une fonction essentielle lors de la transcription des ARN messagers. Cette sous-unité contient un

domaine C-terminal composé chez l'Homme de 52 répétitions en tandem d'un heptapeptide $Y_1S_2P_3T_4S_5P_6S_7$ sujet à des événements de phosphorylation et déphosphorylation au cours de la transcription.

Suivant son état de phosphorylation, le CTD de RPB1 sert de plateforme de recrutement pour différents facteurs nucléaires qui interviennent dans la maturation des ARN messagers (enzyme de capping, de polyadénylation ou de clivage). Parmi les différents sites de phosphorylation du CTD, les phosphorylations des sérines 2 (P-S2) et 5 (P-S5) sont les mieux étudiées (pour revue voir (Buratowski, 2009; Phatnani and Greenleaf, 2006)).

III.4.b. Phosphorylation de la sérine-5 du CTD de RPB1 (Pol II P-S5)

Le profil de phosphorylation de ces sérines varie durant les différentes étapes de la transcription, parallèlement au déplacement de l'ARN Pol II le long du gène. Le CTD est trouvé sous au moins quatre formes de phosphorylation majeures durant la transcription des gènes codants pour des protéines (**Figure 19**). Dans le complexe de pré-initiation, le CTD n'est pas phosphorylé. A l'étape d'initiation, vers l'extrémité 5' du gène, le CTD est hyperphosphorylé sur la sérine-5 par CDK7 qui fait partie du complexe TFIIF. Les expériences de ChIP ont montré que le niveau de phosphorylation sur la sérine-5 demeure élevé pendant la transcription des 100 premiers nucléotides du gène (Komarnitsky et al., 2000). Par la suite, son niveau décroît lentement dans le sens de déplacement de l'enzyme vers le 3'UTR du gène sous l'effet d'une phosphatase (Cho et al., 1999; Krishnamurthy et al., 2004; Mosley et al., 2009).

Plusieurs événements sont physiquement et fonctionnellement reliés à cette modification. Les enzymes de coiffe des ARNm se lient directement aux résidus sérine-5 phosphorylés du CTD, ce qui permet une addition rapide de la coiffe en 5' sur l'ARN en cours de synthèse (Komarnitsky et al., 2000). Certaines enzymes de modification des histones peuvent aussi s'associer au CTD phosphorylé sur la sérine-5, c'est le cas de Set1 qui méthyle l'histone H3 sur la lysine 4 (chez la levure) (Ng et al., 2003), ce qui a pour but de faciliter l'initiation de la transcription.

Au fur et à mesure que l'ARN Pol II progresse dans l'élongation, on observe une diminution de la phosphorylation sur la sérine-5 inversement corrélée à une augmentation graduelle de la phosphorylation sur la sérine-2 en direction du 3'UTR du gène.

III.4.c. Phosphorylation de la sérine-2 du CTD de RPB1 (Pol II P-S2)

La phosphorylation de la sérine-2 est catalysée par la kinase CDK9 du complexe P-TEFb (CycT1/CDK9) (**Figure 19**). La phosphorylation du CTD sur la sérine-2 donne naissance à une plateforme d'assemblage de complexes qui « voyagent » avec Pol II tout au long du gène. Ceux-ci incluent des facteurs de régulation de l'élongation de la transcription, de la maturation des

ARN messagers et de la terminaison. L'inhibition de l'activité CDK9 par le DRB, empêche la phosphorylation sur la sérine-2 et bloque l'élongation des ARN messagers.

Au cours de l'élongation, l'enzyme qui méthyle la lysine 36 de l'histone H3 (H3K36 méthyltransferase) s'associe au CTD phosphorylé sur la sérine-2 et méthyle les histones H3 du gène transcrit (Hampsey and Reinberg, 2003). Cette modification d'histone permet le recrutement d'autres complexes qui affectent l'acétylation des histones et le remodelage de la chromatine (Carrozza et al., 2005; Keogh et al., 2005). Ainsi, l'information portée par le profil de Pol II P-S2 est propagée et enregistrée sur la chromatine en cours de transcription (Kaplan et al., 2003). Il est supposé que ces modifications ont pour but d'influencer positivement les tours de transcription suivants. De plus, il a été montré chez la levure que les enzymes de polyadénylation et de terminaison de la transcription des ARNm s'associent préférentiellement au CTD phosphorylé sur la sérine-2 (Licatalosi et al., 2002; Meinhart and Cramer, 2004).

III.4.d. Valeur informative du profil Pol II P-S5/P-S2

En conclusion, le profil de phosphorylation des sérines 5 et 2 du CTD de Pol II sur le gène en cours de transcription reflète les différentes étapes de la transcription. Dans le complexe de pré-initiation, le CTD n'est pas phosphorylé. A l'extrémité 5' du gène, on observe une hyperphosphorylation du CTD sur la sérine-5 qui traduit l'étape d'initiation de la transcription. Au cours de l'étape d'élongation, on observe une diminution graduelle de la phosphorylation sur la sérine-5 qui est inversement corrélée à une augmentation de la phosphorylation de la sérine-2. A l'étape de terminaison, la phosphorylation sur la sérine-5 est à son plus bas niveau tandis que celle de la sérine-2 est à son maximum (**Figure 19**) (Phatnani and Greenleaf, 2006).

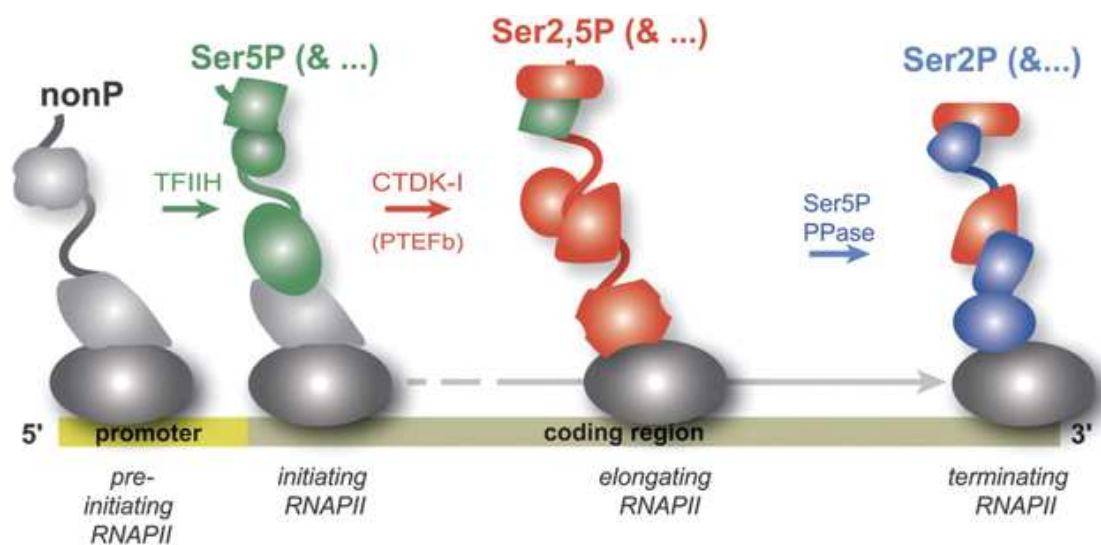


Figure 19 : Phosphorylations du domaine C-terminal de l'ARN Pol II. L'ARN Pol II (ovale gris) est représentée sur quatre positions le long d'un gène. Son CTD (ligne ondulée) est associée à une couleur différente pour indiquer les états de phosphorylation différents: le gris indique le CTD non phosphorylé ; le vert indique la phosphorylation de la sérine 5 (Ser5P) ; le rouge indique la phosphorylation des sérines 5 et 2, (Ser2, 5P) et le bleu indique la phosphorylation de la sérine 2 (Ser2P). *Source : Phatnani & Greenleaf (2006).*

III.4.e. Classement des gènes suivant l'état d'activité de l'ARN Pol II

En 2008 une étude à grande échelle, couplant du GRO-seq (Global Run-On sequencing) à du ChIP-seq Pol II, a permis de généraliser le fait que la régulation de la transcription des gènes par Pol II ne se fait pas qu'au niveau de l'assemblage du complexe de pré-initiation. L'alignement des données de ChIP-seq et de GRO-seq a permis de cartographier la position de Pol II sur tout le génome, tout en visualisant son état d'activité (traduit par la position et l'abondance des transcrits naissants) à une position donnée. Ainsi, Core *et al* montrent que les gènes peuvent être classés en quatre catégories suivant l'activité transcriptionnelle de Pol II (**Figure 20**) (Core et al., 2008).

* Les gènes de la classe I sont composés des gènes où Pol II est active. Ces gènes se caractérisent en ChIP par la détection d'un niveau relativement constant de Pol II tout le long du gène et par la présence d'un gradient régulier de transcrits sens de tailles variables du 5' au 3'UTR.

* Les gènes de classe II sont composés des gènes où une fraction de Pol II est en pause tandis que l'autre fraction est active. Ces gènes se caractérisent par une forte accumulation de Pol II de part et d'autre du TSS ainsi que par la détection de transcrits de tailles variables tout le long de la séquence codante.

* Les gènes de classe III correspondent aux gènes sur lesquels toute la Pol II est en pause. Sur ces gènes, on ne distingue qu'une faible accumulation de transcrits de petite taille centrés sur 50 nucléotides après le TSS.

* Les gènes de classe IV correspondent aux gènes non transcrits sur lesquels il n'y a pas de détection de Pol II, ni de transcrits.

Les classes II et III de gènes traduisent l'existence d'une pause transcriptionnelle à l'élongation. Une des autres observations marquantes de cette étude, qui a été confirmée par d'autres, est que la détection de Pol II au TSS des gènes s'accompagne de la présence de nombreux transcrits de tailles variables de part et d'autre des TSS des gènes de classe I, II et III. Il s'agit de transcrits à la fois sens et anti-sens (Core et al., 2008; Seila et al., 2008). Les transcrits sens ont une taille de 50 nucléotides en moyenne et s'accumulent à de forts niveaux en aval des TSS pour les gènes de classe I et II. Par contre, les transcrits anti-sens, de 250 nucléotides en moyenne, sont détectés sur les trois classes de gènes et s'accumulent à des niveaux plus faibles en amont du TSS (**Figure 20**). L'observation de transcrits sens de courte taille corrobore le fait qu'à l'initiation de la transcription Pol II est encline à faire des transcrits abortifs de courte taille (Marshall and Price, 1992). Par contre la détection de transcrits anti-sens a été beaucoup plus surprenante pour la communauté scientifique et indique que la polymérase peut transcrire de façon bidirectionnelle à partir des promoteurs. A l'heure actuelle, aucun rôle fonctionnel n'a été attribué à ces transcrits anti-sens.

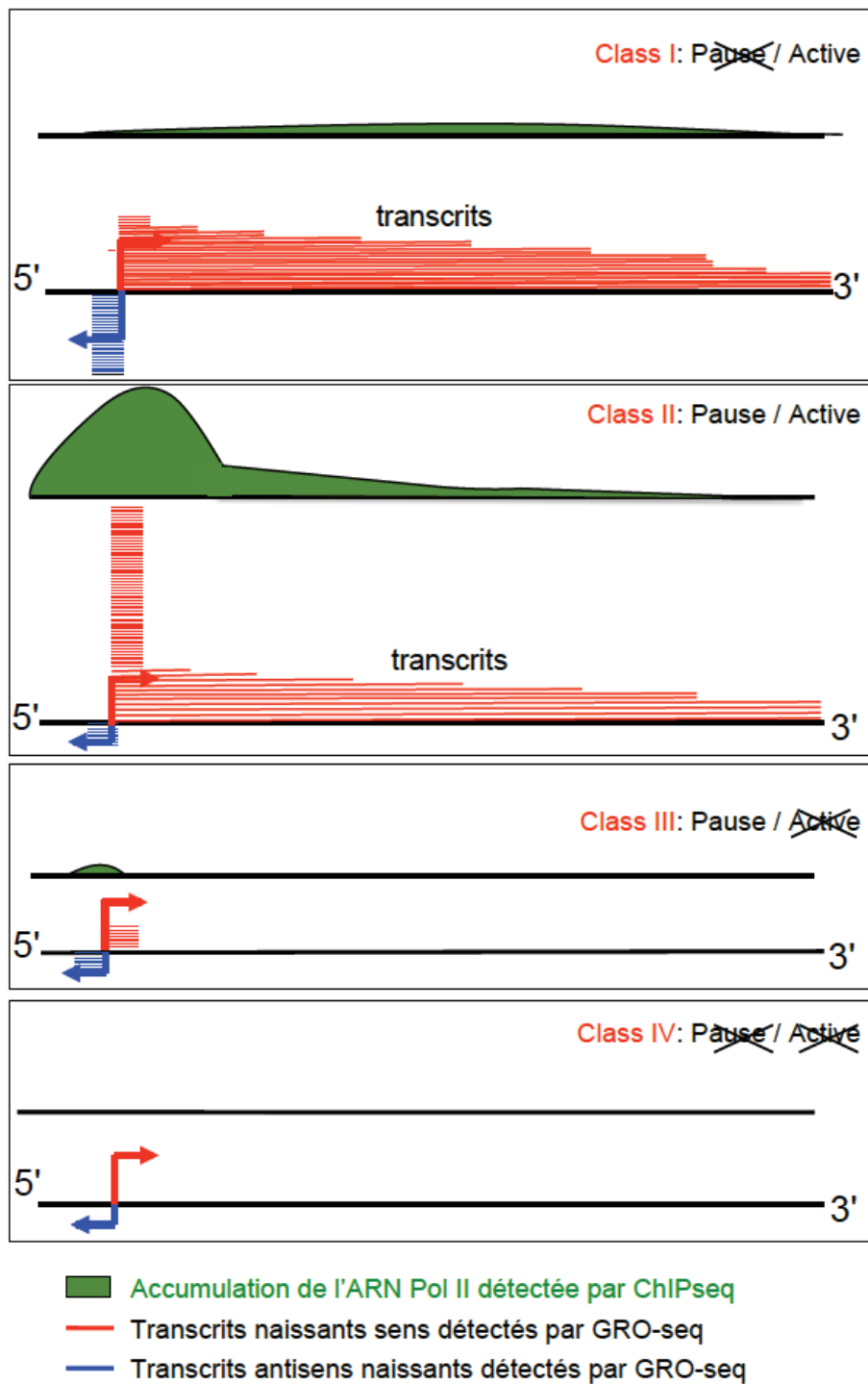


Figure 20 : Schéma représentant la comparaison de la pause transcriptionnelle de l'ARN Pol II à l'activité du gène. Quatre classes de gène peuvent être identifiées sur la base de la comparaison de l'abondance et de la localisation de l'ARN Pol II (ChIP-seq, en vert) et de la transcription du gène associé (GRO-seq, rouge et bleu). Un exemple de chaque classe est représenté. *Illustration adaptée de Core et al. (2008).*

III.5. Concepts émergents

La transcription des gènes est un mécanisme hautement complexe qui nécessite une coordination de l'assemblage d'un grand nombre de facteurs via la reconnaissance de différentes séquences d'ADN régulatrices, ce qui assure une expression spatio-temporelle correcte des gènes.

Avec l'explosion des études genome-wide, les définitions traditionnelles des enhancers, mais aussi des promoteurs, doivent être revisitées. Ces éléments de contrôle peuvent être maintenant caractérisés en terme d'architecture locale et régionale, de présence de modifications d'histones spécifiques, d'association de facteurs de transcription et autres complexes de remodelage de la chromatine, contribuant fonctionnellement à l'initiation de la transcription (Kim and Shiekhattar, 2015).

III.5.a. Enhancers et promoteurs : une classe unique d'éléments fonctionnels

Classiquement, promoteurs et enhancers sont considérés comme des classes distinctes d'éléments de régulation du fait de leur localisation par rapport aux gènes. Cependant, des études récentes ont révélé une grande similarité entre enhancers et promoteurs. Andersson et d'autres ont proposé un modèle dans lequel les promoteurs et les enhancers sont considérés comme une classe unique d'éléments fonctionnels, car possédant une architecture similaire pour l'initiation de la transcription (**Figure 21**). En effet, un grand nombre de données établissent que les enhancers ne sont pas simplement une collection de sites de liaison pour différents co-facteurs transcriptionnels, mais ont aussi la capacité d'actionner la transcription via l'ARN Pol II indépendamment de leur promoteur cible (Andersson, 2015; Andersson et al., 2015; Fuda et al., 2009; Kim and Shiekhattar, 2015; Kim et al., 2015; Li et al., 2016; Nguyen et al., 2016; Weingarten-Gabbay and Segal, 2014).

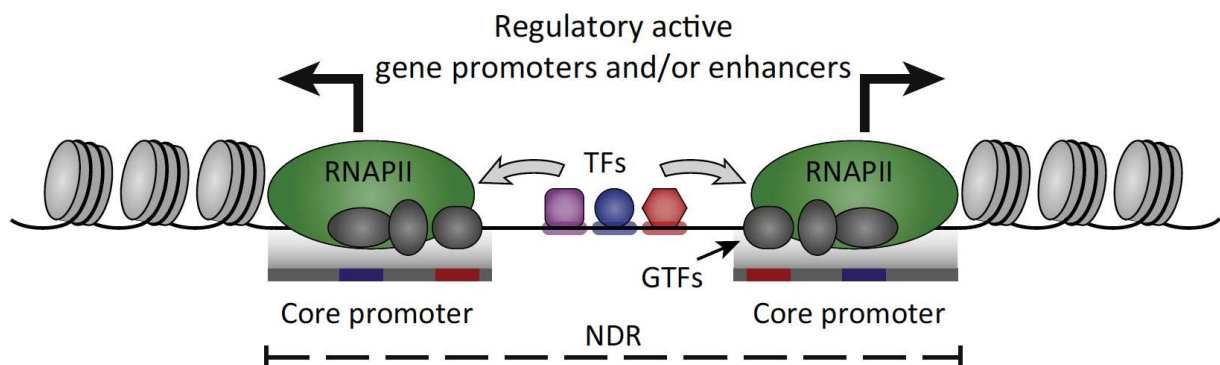


Figure 21 : Les éléments de régulation actifs, enhancer comme promoteur, sont liés par l'ARN Pol II et transcrits. (NDR) nucleosome depleted region ou région dépourvue de nucléosome ; (TFs) facteurs de transcription ; (GTFs) facteurs généraux de transcription ; (RNAPII) ARN polymérase II. Source : Andersson et al. (2015).

Les ARN transcrits par Pol II au niveau des régions enhancers et promoteurs actives sont bidirectionnels et coiffés au niveau de leur extrémité 5'. L'addition de la coiffe a lieu lorsque la taille du transcrit naissant atteint 25 à 50 nucléotides. L'ajout de la coiffe est un processus qui ne requiert pas de séquence ARN spécifique, ainsi les ARN produits par les promoteurs : ARN anti-sens non codant en amont du TSS (uaRNA) et ARN messagers codants (mRNA), tout comme les ARN non codants produits par les enhancers (eRNA) sont coiffés (**Figure 22**). Les ARNm sont généralement longs, multi-exoniques et polyadénylés à leur extrémité 3' ce qui les rend très stables. En revanche, les eRNA et uaRNA sont de petite taille et non-polyadénylés, donc très instables et faiblement représentés (**Figure 22**).

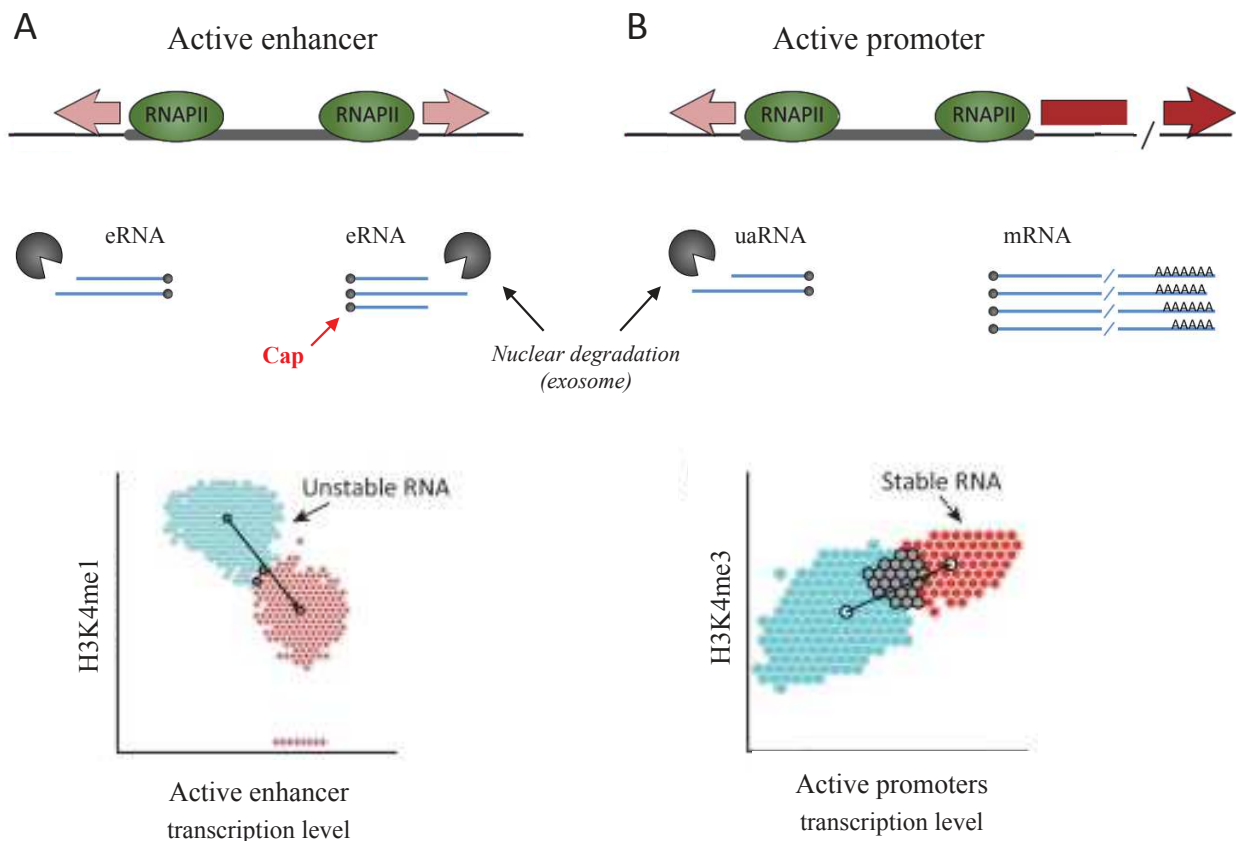


Figure 22 : La stabilité des ARN transcrits distingue les régions régulatrices enhancer et promoteur. L'initiation de transcription au niveau des enhancers (A) et des promoteurs (B) produit différents types d'ARN. Les ARN produits par les enhancers (eRNAs) sont semblables aux ARN anti-sens produits par les promoteurs en amont du TSS (uaRNA) : ils sont faiblement représentés, de petite taille, non-polyadénylés, très instables et retenus dans le noyau où ils sont dégradés par le complexe exosome. Les ARN messagers (mRNA) quant à eux sont généralement plus abondants, longs, multi-exoniques, polyadénylés, stables et exportés vers les polyribosomes pour être traduits. *Source : Andersson et al. (2015).*

La technologie CAGE (Cap Analysis Gene Expression) permet d'identifier les ARN transcrits portant à leur extrémité 5' une coiffe ou CAP (Kodzius et al., 2006; Shiraki et al., 2003). Associée au séquençage à haut débit, la technologie des CAGE permet de visualiser l'expression des ARN capés bidirectionnels et de distinguer si celle-ci est issue de la transcription des régions régulatrices enhancers ou promoteurs. Sur la base du fait que le niveau de

transcription d'un enhancer est corrélé au niveau de transcription du promoteur correspondant, par corrélation d'expression CAGEs entre les eRNA pour les enhancers et uaRNA/mRNA pour les promoteurs, le consortium FANTOM (Functional ANnotation Of the Mammalian genome) a assigné les enhancers à leurs gènes cibles.

Au début des années 2000, le RIKEN Institute situé au Japon a initié le projet FANTOM. Ce projet vise à identifier et cataloguer l'ensemble des gènes et autres éléments fonctionnels dans le génome des mammifères grâce aux nouvelles technologies disponibles. FANTOM est aujourd'hui un consortium international. Plus de 250 experts en biologie cellulaire et en bio-informatique, venant de 114 institutions repartis dans 20 pays ont travaillé au sein de FANTOM5, la 5e édition du projet. L'ensemble des travaux produits au sein de FANTOM5 a donné lieu à 30 articles, dont 2 articles majeurs publiés dans le journal Nature en 2014 : « A promoter level mammalian expression atlas » (Forrest et al., 2014) et « An atlas of active enhancers across human cell types and tissues » (Andersson et al., 2014). L'ensemble des études publiées par FANTOM5 constitue la première cartographie complète des réseaux régulant la transcription des gènes à travers tous les types de cellules. Dans le cadre de mon projet, j'ai eu recours à l'atlas des enhancers actifs définis par FANTOM5 (voir Étude I paragraphe III.2.c).

Au vu de ces nouvelles observations, un enhancer peut être défini comme ayant les caractéristiques suivantes : (i) les enhancers sont des séquences d'ADN qui augmentent l'expression des gènes auxquels ils sont liés, (ii) l'activité des enhancers est indépendante de leur orientation, (iii) les enhancers fonctionnent indépendamment de leur position par rapport à leur gène cible et sur une longue distance, (IV) les enhancers présentent un enrichissement pour la modification d'histone H3K4me1 et une hypersensibilité à la DNase I indiquant une structure chromatinienne peu compacte et (V) les enhancers caractérisés par la présence de l'ARN Pol II et la synthèse de eRNA sont considérés comme étant engagés activement dans la transcription des ARNm correspondants.

Par ailleurs, les études par ChIP-seq ont montré que la présence de facteurs de transcription ne signifie pas systématiquement que l'enhancer est fonctionnel. L'activation fonctionnelle nécessite le recrutement de co-facteurs additionnels ou la combinatoire coordonnée de plusieurs facteurs de transcription. Les histone acétyltransférases P300 et CBP ont été décrites comme marque des enhancers actifs car interagissant avec de nombreux activateurs transcriptionnels et la machinerie générale de la transcription, incluant Pol II. De plus, si les enhancers actifs sont enrichis en H3K27Ac, marque déposée par P300/CBP, les enhancers « poised » sont marqués par H3K27me3 (Kim and Shiekhatar, 2015).

Rajoutant un niveau de complexité supérieur, des études à haut débit des activités fonctionnelles des enhancers et des promoteurs montrent que si les activités promotrices de certains promoteurs sont supérieures à celles des enhancers distaux, leurs activités enhancers sont similaires à celles des enhancers. Les activités différentielles seraient liées à la présence de

motifs de fixation particuliers pour des facteurs de transcription tels que SP1 et EGR1 pour les activités promotrices, AP-1 pour les activités enhancers et RFX, CREB et EGR pour les régions pouvant présenter des activités à la fois enhancer et promoteur (Nguyen et al., 2016; Ruiz-Orera et al., 2015). En d'autres termes, certains promoteurs peuvent se comporter comme des enhancers pour contrôler l'expression de gènes distaux. Ainsi, Skvortsova *et al* ont montré que le promoteur alternatif du gène PIWIL2 fonctionne comme un enhancer pour l'expression du gène PHYHIP localisé 60kb en amont (Skvortsova et al., 2016). Un autre exemple est l'interaction entre les promoteurs des gènes MAFK et INTS1, où le promoteur de INTS1 a une activité enhancer sur le promoteur MAFK (Li et al., 2012).

III.5.b. Communication entre enhancers et promoteurs

La compréhension des mécanismes de communication entre enhancers et promoteurs a grandement évoluée au cours de ces dernières années (Kim and Shiekhhattar, 2015; Kim et al., 2015; Li et al., 2016). Deux mécanismes permettant la communication entre la séquence enhancer et sa séquence promotrice correspondante ont été proposés : il s'agit des mécanismes de tracking et de looping (**Figure 23**).

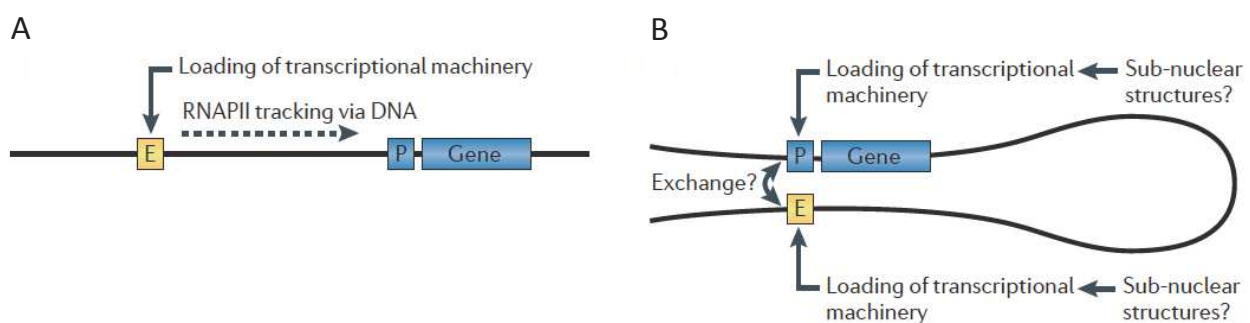


Figure 23 : Représentation schématique de la communication enhancer-promoteur. **A)** Le mécanisme de tracking est un mécanisme de transfert qui fait intervenir le glissement (sliding) de l'ARN Pol II le long de la région comprise entre l'enhancer (E) et le promoteur (P). **B)** Lors du mécanisme de looping l'ARN Pol II est transférée de l'enhancer (E) au promoteur (P) par la formation d'une boucle qui permet d'établir un contact physique entre l'enhancer et le promoteur. *Source : Li et al. (2015).*

Le mécanisme de transfert par tracking propose que l'ARN Pol II recrutée par le facteur de transcription lié à l'enhancer soit transférée de l'enhancer au promoteur par glissement le long de la fibre d'ADN (**Figure 23.A**) (Blackwood and Kadonaga, 1998; Li et al., 2016). Il est supposé que le complexe transcriptionnel scanne la fibre d'ADN jusqu'à ce qu'il atteigne une séquence promotrice compétente pour l'initiation de la transcription. Ce type de transfert n'a été décrit que pour un nombre limité de gènes : β -globine, ϵ -globine, l'antigène spécifique de la prostate (PSA) et uPA (Louie et al., 2003; Moquet-Torcy et al., 2014; Wang et al., 2005; Zhao and Dean, 2004; Zhu et al., 2007). Le peu de données expérimentales qui existent sur le tracking montrent qu'il

est associé à la production de petits ARN non codants, bidirectionnels dans la région comprise entre l'enhancer et le promoteur (Moquet-Torcy et al., 2014; Zhu et al., 2007). Ce mécanisme semblerait plutôt restreint à la communication enhancer-promoteur dans le cas où l'enhancer est situé en amont et à une distance relativement faible.

Le looping est un mécanisme qui permet la formation d'un contact physique grâce à la formation d'une boucle entre, d'une part, la machinerie de transcription recrutée par le facteur de transcription au niveau de l'enhancer et, d'autre part, le promoteur (**Figure 23.B**). Grâce à la technique du 3C (Chromosome Conformation Capture) ce mode de rapprochement physique a été décrit et validé *in vivo* pour quelques gènes (Chung et al., 2016; Phillips-Cremins, 2014; Sexton et al., 2009). Ce type de mécanisme permet d'expliquer le fonctionnement des enhancers quelle que soit leur distance par rapport au promoteur, leur position (en cis ou en trans) et leur orientation dans le génome. Comme mentionné au paragraphe III.1.d (**Figure 15**), la formation de boucle entre enhancer et promoteur ferait intervenir la liaison du Mediator et des cohésines pour stabiliser la structure.

Cependant, il semblerait que les mécanismes de looping et de tracking ne soient pas systématiquement exclusifs et qu'ils puissent avoir lieu de manière coordonnée et complémentaire. Wang *et al* ont montré que pour l'expression du gène de l'antigène spécifique de la prostate (PSA) lors de l'induction par les androgènes, le facteur de transcription recrute l'ARN Pol II au niveau de l'enhancer. Puis, la polymérase se déplace le long de la séquence intergénique tout en synthétisant de courts fragments d'ARN non codants et en restant en contact avec le facteur de transcription et l'enhancer jusqu'à atteindre le promoteur, ce qui conduit à la formation d'une boucle chromatinienne (Wang et al., 2005).

RÉSULTATS

RÉSULTATS

Les protéines Fra-1 et Fra-2 sont co-surexprimées de façon aberrante dans de nombreuses tumeurs triple négatives où elles augmentent l'agressivité tumorale en altérant diverses fonctions cellulaires. Fra-1 et Fra-2 ont été peu étudiées jusqu'à récemment. Ni l'ensemble des mécanismes responsables de leur dérégulation dans les tumeurs, ni les mécanismes de leur action transcriptionnelle, ni, enfin, leurs gènes cibles ne sont encore clairement identifiés. Mes travaux de thèse se sont donc concentrés sur deux aspects :

D'une part, la caractérisation des gènes codants sous contrôle de Fra-1 et/ou Fra-2 de manière à essayer de caractériser des gènes cibles participant à l'agressivité des cancers du sein triple négatifs et qui pourraient être des cibles thérapeutiques potentielles. En effet, malgré leur implication grandissante dans la tumorigenèse et l'agressivité tumorale, le ciblage direct de Fra-1 et/ou Fra-2 dans les tumeurs est pour l'instant exclu car ces deux facteurs de transcription ne sont pas inhibables par des approches utilisables en clinique.

D'autre part, mes efforts ont aussi porté sur la caractérisation des mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour la transcription de leurs gènes cibles. En effet, malgré la très grande abondance de littérature concernant AP-1, seulement quelques données fragmentaires ont été rapportées concernant les mécanismes mis en jeu par AP-1 pour la régulation des gènes. De plus, il est fort vraisemblable que les mécanismes sont différents suivant que les gènes sont activés ou réprimés. On ne peut également pas exclure qu'il puisse y avoir des différences d'un gène activé à un autre en fonction des partenaires transcriptionnels et de la topologie des gènes. Pour cet aspect, je me suis intéressée à élucider le processus transcriptionnel pour un des gènes cibles identifiés dans la première partie de mon travail.

Mon travail de thèse est présenté en trois grandes parties :

Étude I : Identification des transcriptomes gouvernés par Fra-1 et/ou Fra-2 et caractérisation de leurs régions de fixation sur le génome de la lignée de cancer du sein triple négatif MDA-MB231

Dans un premier temps, j'ai cherché à identifier les gènes cibles de Fra-1 et/ou Fra-2 dans une lignée modèle de TNBC, la lignée MDA-MB231 couramment utilisée dans la littérature, dans le but d'identifier des cibles thérapeutiques potentielles. Les lignées TNBC expriment généralement peu ou pas c-Fos et FosB ce qui réduit la combinatoire possible des dimères AP-1 et en fait un modèle de choix pour étudier les rôles des protéines Fra-1 et Fra-2. Mes travaux comprennent des analyses transcriptomiques différentielles (inhibition de Fra-1, de Fra-2 ou des deux protéines de façon concomitante) ainsi que des approches ChIP-seq pour identifier et

caractériser les sites de fixation de Fra-1 et Fra-2 sur la chromatine dans le but d'identifier des cibles directes probables. Cet aspect a également impliqué la validation de l'expression dépendante de Fra-1 et/ou Fra-2 d'un certain nombre de gènes cibles dans différentes lignées TNBC ainsi que des corrélations d'expression dans des banques de tumeurs (CIT et TCGA) de manière à sélectionner des gènes, pour les études ultérieures, qui soient pertinents dans le cadre de la pathologie humaine.

Étude II : Mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour l'expression du gène HMGA1

Je me suis également intéressée aux mécanismes transcriptionnels gouvernés par Fra-1 et/ou Fra-2 pour la régulation de leurs gènes cibles. Pour cette étude, mes efforts se sont concentrés sur un des gènes cibles de Fra-1, le gène HMGA1. En effet, HMGA1 est connu pour son rôle majeur dans l'agressivité des tumeurs épithéliales et son expression a récemment été décrite comme étant corrélée au grade, au stade, à la présence de métastases dans les ganglions et à distance, ainsi qu'au statut triple négatif des tumeurs du sein. Dans cette étude, non complètement encore finalisée, j'ai cherché à déterminer quels sont les co-facteurs transcriptionnels dont le recrutement est dépendant de Fra-1 et/ou Fra-2.

Étude III : Rôle de CD68 dans le phénotype tumoral des TNBC

Lors de mon analyse transcriptomique il est apparu que l'un des gènes fortement régulés positivement par Fra-1 et Fra-2 est CD68. CD68 est considéré dans la littérature comme un marqueur des macrophages. Cependant, quelques études insistent sur le fait que CD68 est exprimé dans d'autres types cellulaires tels que les fibroblastes et les cellules cancéreuses ainsi que dans des tumeurs humaines telles que les gliomes ou les carcinomes fibrolamellaires hépatiques. CD68 est une protéine transmembranaire décrite pour être localisée principalement dans les lysosomes et les endosomes. Sa fonction n'est pas caractérisée et son rôle potentiel dans la tumorigenèse n'a jamais été étudié. J'ai donc cherché à déterminer si l'expression de CD68 est impliquée dans le phénotype tumoral des TNBC au niveau de la prolifération, de la motilité et de l'invasion dans des tests en 2 et 3 dimensions.

RÉSULTATS

Étude I

ÉTUDE I

Identification des transcriptomes gouvernés par Fra-1 et/ou Fra-2 et caractérisation de leurs régions de fixation sur le génome de la lignée de cancer du sein triple négatif MDA-MB231.

Malgré les avancées thérapeutiques, les cancers du sein agressifs et métastatiques triple négatifs (TNBC) ne bénéficient d'aucune thérapie ciblée et restent incurables. Un des objectifs de mon projet est d'identifier des cibles directes de Fra-1 et/ou Fra-2 qui pourraient participer à l'agressivité tumorale et constituer de nouvelles cibles potentielles à visée diagnostique ou thérapeutique pour le traitement de ces cancers.

Dans ce but, j'ai suivi deux axes en parallèle qui sont complémentaires pour mon étude. Le premier est l'identification des gènes cibles de Fra-1 et/ou Fra-2 dans la lignée modèle MDA-MB231, par des approches transcriptomiques en présence et en absence de Fra-1 et/ou Fra-2 réprimés par RNAi. Le second est la cartographie et la caractérisation des sites de liaison de Fra-1 et Fra-2 sur le génome des MDA-MB231 par des approches ChIP-seq.

I. MODÈLE EXPÉRIMENTAL

Le modèle cellulaire triple négatif choisi pour initier mon étude est la lignée de référence MDA-MB231. Elle est la plus utilisée dans la littérature pour l'étude des propriétés des cancers triple négatifs. Bien que mes travaux aient été initiés dans cette lignée modèle, mon projet a impliqué des validations dans d'autres lignées triple négatives et l'analyse des banques de tumeurs humaines pour identifier ce qui a valeur générale de ce qui a valeur spécifique, ce point étant absolument essentiel dans une perspective thérapeutique.

I.1. **Caractérisation du système expérimental**

I.1.a. Mesure des abondances relatives des ARNm AP-1 dans les différentes lignées cellulaires d'intérêt

La lignée MDA-MB231 est issue d'un carcinome mammaire humain triple négatif hautement agressif et métastatique. Cette lignée exprime de hauts niveaux de Fra-1 et Fra-2, sous une forme hyper-phosphorylée due à une forte activité de la voie ERK suite à des mutations activatrices sur K-Ras et B-Raf (Hollestelle et al., 2007). De façon intéressante pour notre étude, les autres protéines FOS, à savoir c-Fos et FosB sont peu ou pas exprimées dans cette lignée, ce qui permet de rechercher spécifiquement les gènes cibles de Fra-1 et/ou Fra-2 et d'entreprendre l'étude des mécanismes transcriptionnels sous-jacents.

L'analyse par RT-qPCR des niveaux relatifs d'ARNm des différents membres AP-1 dans la lignée MDA-MB231 (**Figure 24.A**) confirme que Fra-1 et Fra-2 sont les membres les plus fortement transcrits parmi ceux de la famille FOS. Concernant la famille JUN, c-Jun est le membre le plus abondant suggérant que les complexes AP-1 les plus représentés dans cette lignée sont les hétérodimères c-Jun/Fra-1 et c-Jun/Fra-2.

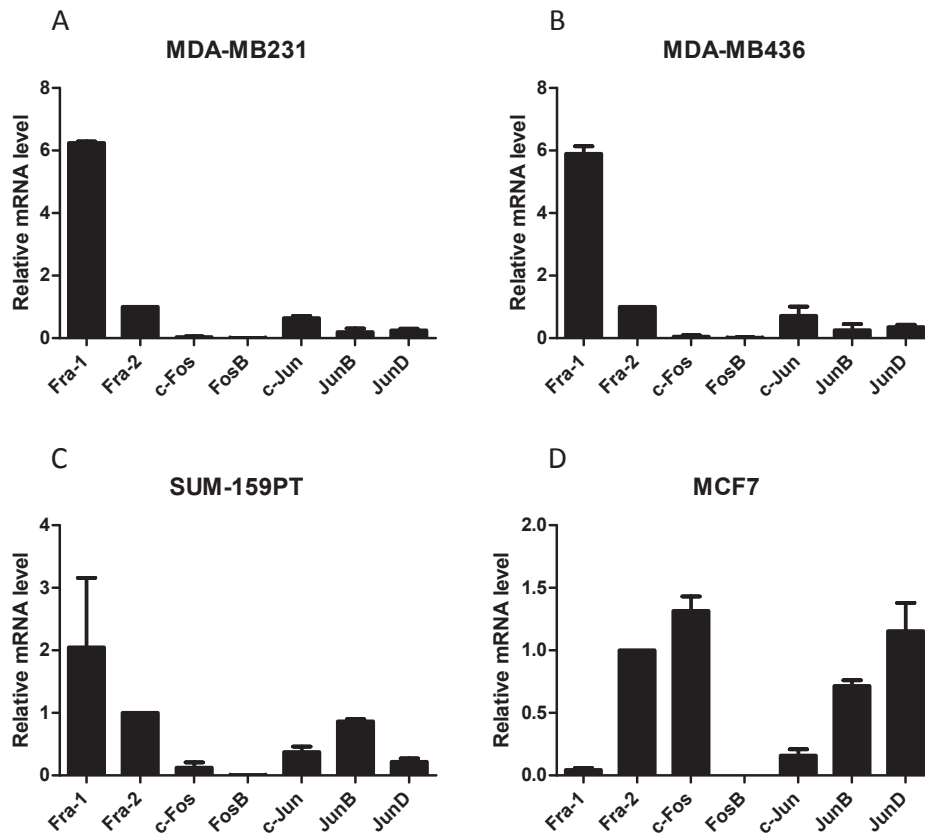


Figure 24 : Mesure des abondances relatives des ARNm des membres AP-1 dans les quatre lignées d'intérêt. Les valeurs de RT-qPCR sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm Fra-2. n=3. Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.

Au cours de mes études, pour déterminer ce qui a une valeur plus générale de ce qui a une valeur spécifique des cellules MDA-MB231, j'ai utilisé deux autres lignées triple négatives humaines : MDA-MB436 et SUM-159PT. Comme les MDA-MB231, ces deux lignées sont décrites comme (i) appartenant aux cancers du sein basaux (basal-like) par Nerve *et al*, (ii) appartenant au sous-type triple négatif (TNBC) par Lehmann *et al*, (iii) plus particulièrement au sous-type mesenchymal stem-like (MSL) par Saha *et al*, ainsi que (iv) représentatives du sous-type claudin-low par Grigoriadis *et al* (Grigoriadis et al., 2012; Lehmann et al., 2011; Neve et al., 2006; Saha and Nanda, 2016). Ces classifications sont détaillées dans l'introduction de la thèse aux paragraphes I.4 et I.5.

Si les abondances relatives des ARNm des membres AP-1 sont similaires pour les MDA-MB231 (**Figure 24.A**) et les MDA-MB436 (**Figure 24.B**) on note des différences dans les SUM-159PT : le différentiel d'expression entre Fra-1 et Fra-2 est moins important et le membre JUN le plus exprimé est JunB (**Figure 24.C**).

De façon intéressante, on observe des profils d'expression des ARNm complètement différents dans la lignée MCF7, ER positive et moins agressive (sous-type luminal A) : les membres FOS les plus exprimés sont c-Fos et Fra-2, et les membres JUN les plus exprimés sont JunD et JunB (**Figure 24.D**).

I.1.b. Comparaison de l'abondance relative des membres AP-1 dans les différentes lignées cellulaires d'intérêt

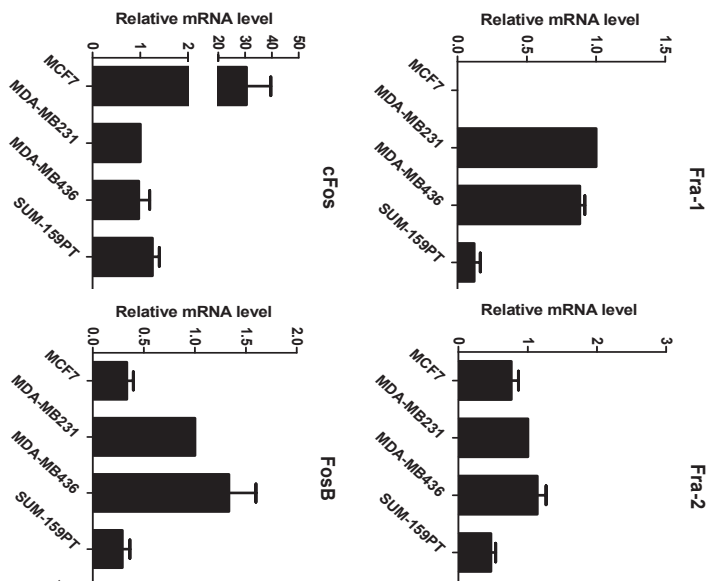
J'ai ensuite comparé l'abondance relative de chaque membre AP-1 dans les quatre lignées mentionnées précédemment, à la fois au niveau ARNm (RT-qPCR) et protéine (immunoblot).

* Membres de la famille FOS : Les niveaux d'ARNm et de protéine pour Fra-1 et Fra-2 sont équivalents entre les lignées MDA-MB231 et MDA-MB436 (**Figure 25.A et B**). En revanche, la lignée SUM-159PT exprime moins de Fra-1 (ARNm et protéine). Dans le cas de Fra-2, on observe une plus grande abondance de la protéine qui est vraisemblablement due à une stabilisation de la protéine et non à une augmentation de la transcription du gène, car son niveau d'ARNm est légèrement plus bas que dans les deux autres lignées. Concernant les membres c-Fos et FosB, ils ont très peu transcrits dans nos lignées d'intérêt et les protéines ne sont pas détectées pas immunoblot (**Figure 25.A et B**).

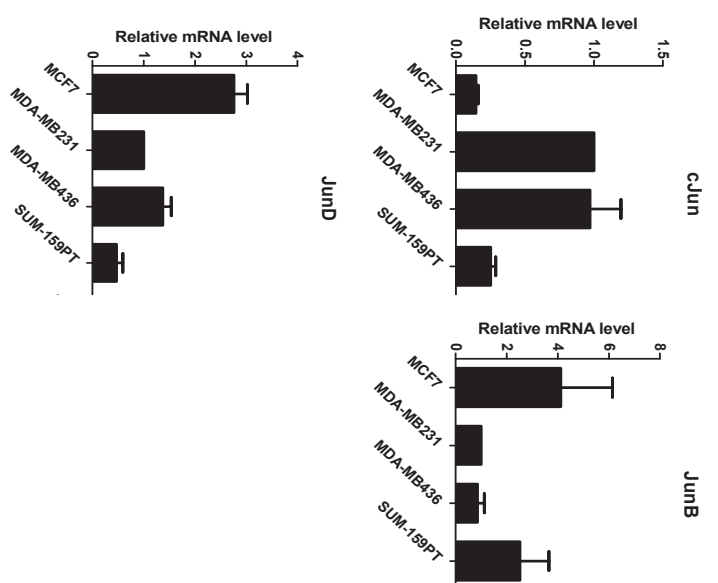
* Membres de la famille JUN : Les niveaux d'ARNm et de protéine pour c-Jun, JunB et JunD sont équivalents entre les lignées MDA-MB231 et MDA-MB436 (**Figure 25.A et C**). Concernant la lignée SUM-159PT, si elle exprime une quantité comparable de JunD (ARNm et protéine) par rapport au MDA-MB231 et MDA-MB436, en revanche elle exprime moins de c-Jun (ARNm et protéine) et beaucoup plus de JunB (ARNm et protéine) (**Figure 25.A et C**).

Il faut donc garder à l'esprit, lors de l'extension des observations faites dans les MDA-MB231 aux deux autres lignées TNBC, que les abondances relatives des différents membres AP-1 ne sont pas exactement les mêmes d'une lignée à l'autre.

A – ARNm des membres FOS



B – ARNm des membres JUN



C – Protéines FOS et JUN

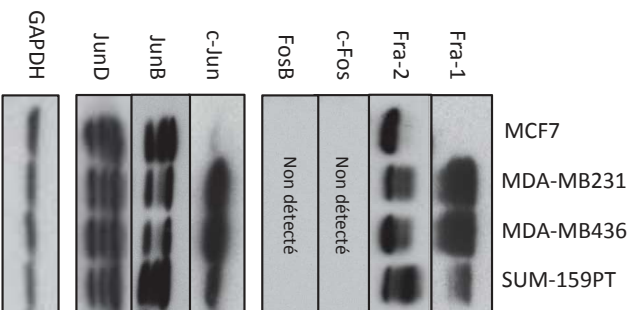


Figure 25 : Analyse de l'expression des membres AP-1 dans différentes lignées de cancer du sein. Mesure de l'abondance d'ARNm des membres FOS (A) et JUN (B) par RT-qPCR. n=3. Les valeurs de RT-qPCR sont normalisées par rapport à l'expression de l'ARNm dans les cellules MDA-MB231. *Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.* (C) Abondance des protéines AP-1 analysée par immunoblot. L'immunoblot GAPDH est présenté en témoin de charge.

I.2. Impact de l'inhibition par RNAi de Fra-1 et/ou Fra-2 sur les autres membres AP-1

Pour identifier les gènes cibles de Fra-1 et/ou Fra-2 dans la lignée modèle MDA-MB231, la stratégie suivie consiste à comparer les ARNm exprimés dans les MDA-MB231 en présence et en absence de Fra-1 et/ou Fra-2, réprimés par transfection de siRNA spécifiques. J'ai donc analysé l'impact de la déplétion phénotypique de l'un sur l'abondance de l'autre, de manière à déterminer s'il pouvait y avoir des biais éventuels (**Figure 26.A**). Il apparaît clairement que la déplétion phénotypique de Fra-1 n'affecte ni l'abondance de l'ARNm ni l'abondance de la protéine Fra-2 et il en est de même dans le cas de la déplétion phénotypique de Fra-2 dans les MDA-MB231 (**Figure 26.A**). Les membres c-Fos et FosB, étant très peu transcrits dans la lignée MDA-MB231 et les protéines n'étant pas détectées pas immunoblot, je n'ai pas analysé l'impact de l'inhibition de Fra-1 et/ou Fra-2 sur leur expression.

En ce qui concerne les deux autres lignées d'intérêt, les résultats ont fait apparaître que la déplétion de Fra-2 n'affecte ni l'abondance de l'ARNm ni l'abondance de la protéine Fra-1. En revanche, la déplétion de Fra-1 entraîne une augmentation de la protéine Fra-2 dans les MDA-MB436 (**Figure 26.B**) ainsi que dans les SUM-159PT (**Figure 26.C**). Cette augmentation ne reflète pas une augmentation de l'ARNm et est donc très vraisemblablement attribuable à une stabilisation de la protéine par un mécanisme que nous n'avons pas investigué.

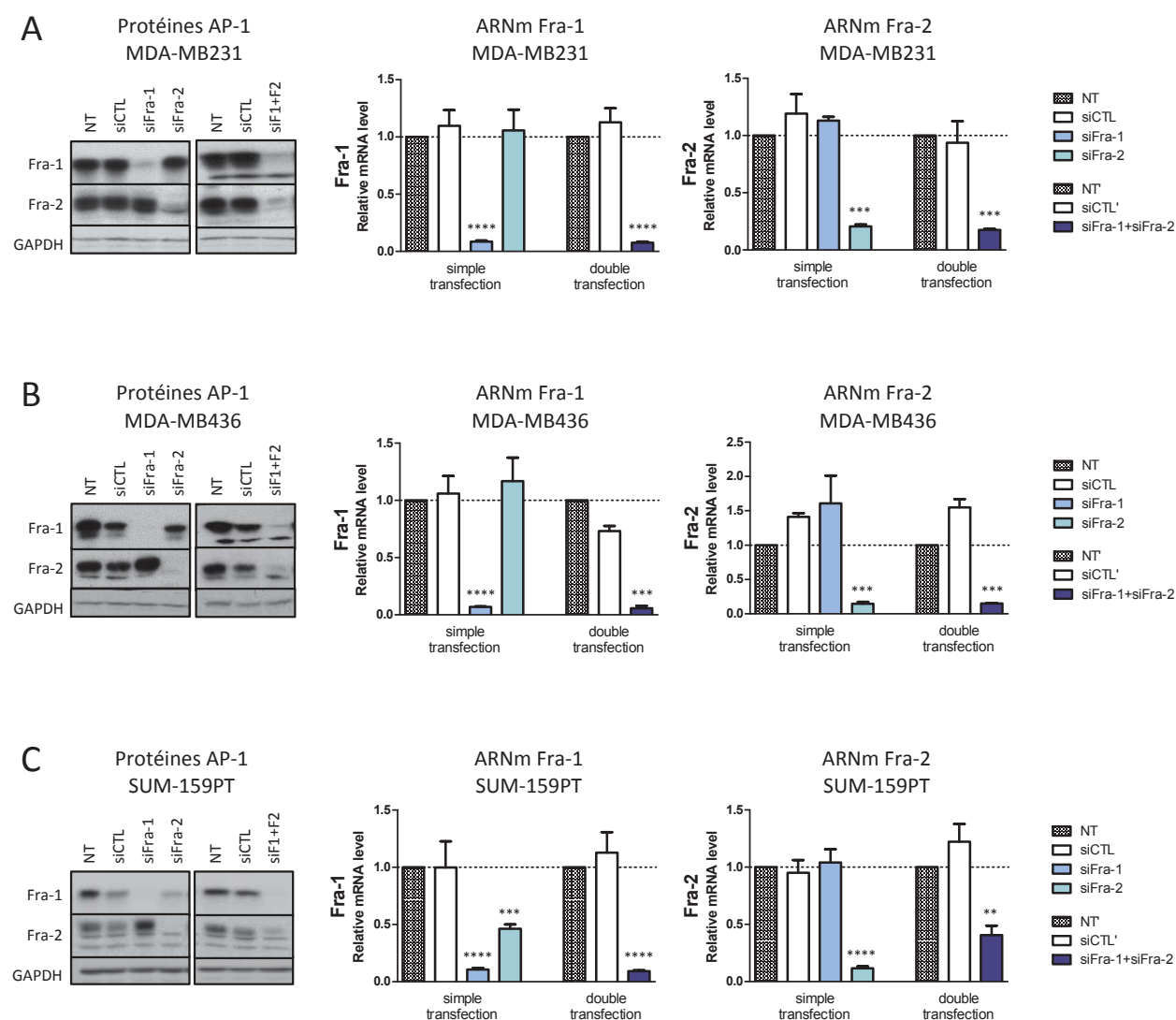


Figure 26 : Impact de l'inhibition de Fra-1 et/ou Fra-2 par RNAi sur les membres FOS, dans les trois lignées triple négatives d'intérêt. Mesure de l'abondance de Fra-1 et Fra-2 au niveau protéine et ARNm après inhibition par siRNA dans les MDA-MB231 (A), MDA-MB436 (B) et SUM-159PT (C). Panel de gauche : Abondance des protéines Fra-1 et Fra-2 analysée par immunoblot, après inhibition de Fra-1 et/ou Fra-2 par siRNA. L'immunoblot GAPDH est présenté en témoin de charge. Panel de droite : Les valeurs de RT-qPCR sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm mesurée pour la condition non-transfectée (NT ou NT'). n=3. Two tailed paired t-test ; (**) p value <0.01 ; (***) p value <0.001 ; (****) p value <0.0001. Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.

II. IDENTIFICATION DES TRANSCRIPTOMES GOUVERNÉS PAR FRA-1 ET/OU FRA-2 DANS LA LIGNÉE MDA-MB231

Mon premier objectif a été d'identifier les cibles géniques de Fra-1 et/ou Fra-2 dans la lignée modèle MDA-MB231. Il est important de souligner que ces gènes peuvent être, non seulement des cibles physiologiques dont l'expression serait stimulée ou réprimée par la surexpression de ces facteurs de transcription, mais aussi des cibles non physiologiques des protéines Fra-1 et Fra-2 exprimées de façon aberrante dans les tumeurs du sein triple négatives.

Les protéines FOS étant potentiellement susceptibles d'avoir des cibles communes, la stratégie suivie consiste à comparer les ARNm exprimés dans les MDA-MB231 en présence et en absence de Fra-1, de Fra-2 et des deux facteurs conjointement, par approche RNAi. Le croisement des transcriptomes obtenus séparément pour Fra-1 et Fra-2 permet de différencier les gènes dont l'expression est régulée à la fois par Fra-1 et Fra-2 de ceux régulés spécifiquement par chacun des deux facteurs de transcription. Et le croisement de ces données avec celles du transcriptome obtenues en absence de Fra-1+Fra-2 permet l'identification des gènes régulés de façon redondante par ces deux facteurs de transcription.

Nous avons choisi d'utiliser trois siRNA dirigés spécifiquement contre Fra-1 et trois siRNA dirigés spécifiquement contre Fra-2, afin d'éviter de sélectionner d'éventuelles cibles off-targets. Par ailleurs, l'utilisation de triplicats technologiques pour chaque condition permet d'augmenter la fiabilité des résultats et d'avoir un recul statistique. Les analyses transcriptomiques ont été conduites sur puce Affymetrix « All genes 2.0ST », en collaboration avec de Dr. L. Vallar (plateforme CRP, Luxembourg). Ce choix se justifie par le fait que jusqu'à ce jour les données transcriptomiques disponibles publiquement sur les tumeurs primaires du sein ont principalement été analysées par technologie Affymetrix.

Cette partie du travail, qui aurait dû être relativement rapide, s'est en fait révélée extrêmement problématique et chronophage en raison (i) de la difficulté à identifier plusieurs siRNA efficaces dirigés contre Fra-1 ou Fra-2, (ii) de la difficulté à trouver des conditions d'extinction fiables pour la suppression conjointe de Fra-1+Fra-2, (iii) des effets off-targets récurrents et non maîtrisables des siRNA Fra-1 et Fra-2 dans des conditions standards de transfection et (iv) des effets off-targets importants du siRNA contrôle commercial (Pool Dharmacon, largement utilisé dans la littérature) que j'ai utilisé dans ma première série d'analyses transcriptomiques (juillet 2013). En effet, la validation par RT-qPCR des résultats obtenus par Affymetrix, a fait apparaître que 50% des gènes retenus étaient des faux positifs résultant directement des effets off-targets du siRNA contrôle (voir plus loin, paragraphe II.1). Outre que cette observation m'obligeait à reprendre mes expériences à leur début, elle attire l'attention sur le fait que les biais liés aux siRNA contrôles commerciaux sont probablement

sous-estimés, voire non réalisés, dans bon nombre d'études transcriptomiques publiées comparant l'expression des gènes entre un siRNA contrôle et des siRNA spécifiques.

J'ai donc dû travailler à redéfinir le design expérimental, ainsi qu'à obtenir des conditions fiables pour répondre à ces problèmes techniques, dans le but d'obtenir des résultats genome-wide robustes, minimisant le plus possible les biais d'analyse liés aux effets off-targets des siRNA (voir plus loin, paragraphe II.2). De ce fait, de nouvelles données exploitables pour identifier les gènes cibles de Fra-1 et/ou de Fra-2 n'ont été disponibles qu'en juillet 2014 alors que j'aurais dû travailler avec ces dernières un an auparavant.

Bien que cet impondérable ait eu des conséquences importantes sur l'avancée de mon projet personnel, j'aimerais cependant souligner que l'optimisation des conditions de suppression des protéines Fra-1 et/ou Fra-2 par RNAi que j'ai dû conduire m'a permis d'être associée, en tant qu'auteur secondaire, à une étude de l'expression aberrante du gène uPA (produisant une protéase essentielle au processus métastatique) sous contrôle de Fra-1 dans les cancers mammaires triple négatifs (*G Moquet-Torcy, C Tolza, M Piechaczyk* and I Jariel-Encontre* (*co-senior authors). Transcriptional complexity and roles of Fra-1/AP-1 at the uPA/Plau locus in aggressive breast cancers. Nucleic Acids Research 42 11011-11024. 2014).*

II.1. Première étude transcriptomique : identification des gènes cibles de Fra-1

Au cours de ma 1^{ère} année de thèse l'approche choisie a été d'utiliser trois siRNA dirigés spécifiquement contre Fra-1 et trois siRNA dirigés spécifiquement contre Fra-2, de façon indépendante, pour éviter la sélection de cibles off-targets et d'identifier les gènes spécifiquement régulés par Fra-1 et/ou par Fra-2.

Le choix des siRNA utilisés est en partie basé sur l'exploitation bioinformatique des données de ChIP-seq pour Fra-1 et Fra-2 réalisés dans la lignée MDA-MB231 (voir plus loin, paragraphe III). Quelques gènes cibles potentiels, contenant des sites de liaison pour Fra-1 et/ou Fra-2 dans leur région promotrice ou intronique, ont été choisis dans le but d'évaluer l'impact des différents siRNA sur leur niveau d'expression. Ceci a été réalisé par RT-qPCR, en utilisant deux couples d'amorces par gène pour valider les résultats. Au préalable, les siRNA avaient été testés pour leur efficacité à antagoniser spécifiquement l'expression de Fra-1 ou Fra-2 par immunoblot, après différents temps de transfection des siRNA ainsi que différentes concentrations. Les siRNA n'entraînant pas une diminution satisfaisante de Fra-1 ou Fra-2, ainsi que ceux entraînant des variations dans les abondances des autres protéines AP-1 ont été écartés. A l'issue de ces expériences, des conditions de transfection ont été choisies : un temps de transfection de 48h et une concentration de 9nM pour chaque siRNA.

Concernant Fra-1, j'ai pu sélectionner trois siRNA satisfaisant les critères ci-dessus : deux qui ciblent la région 3'UTR de Fra-1 et un qui cible sa partie codante. En revanche, pour Fra-2,

le choix des trois siRNA a nécessité des mises au point supplémentaires car les différents siRNA testés présentaient des effets variables sur la diminution de l'expression des gènes cibles analysés. C'est pourquoi la première étude transcriptomique réalisée ne concerne que les gènes cibles de Fra-1.

Les ARN totaux issus de trois expériences indépendantes de transfection des cellules MDA-MB231, soit avec un siRNA contrôle commercial (Pool Dharmacon), soit avec chacun des trois siRNA dirigés contre Fra-1 ont été préparés. Chaque échantillon a été testé pour l'extinction de la protéine Fra-1, la diminution de l'expression des gènes cibles potentiels et la qualité des ARN, sur puce Agilent, avant d'être envoyés à la plateforme d'analyse transcriptomique du CRP au Luxembourg.

Analyse transcriptomique n°1 :

- siRNA CTL Dharmacon (x3)
- siRNA Fra-1 A (x3)
- siRNA Fra-1 B (x3)
- siRNA Fra-1 C (x3)

II.1.a. Résultats et confirmation par RT-qPCR

Les résultats de cette première analyse transcriptomique ont été déconcertants. D'une part, ils ont fait apparaître de façon surprenante que seuls 24 gènes ont une expression dérégulée d'un facteur supérieur à 2 par chacun des trois siRNA dirigés contre Fra-1, soit un recouvrement de 7 à 21% des transcriptomes affectés par chacun des trois siRNA (**Figure 27**). Ces faibles pourcentages sont vraisemblablement dus, au moins en partie, à des effets off-targets liés aux effets indirects microRNA qui sont inhérents et spécifiques à chaque siRNA (Cullen, 2006; Gurtan and Sharp, 2013; Pei and Tuschl, 2006). Dans l'hypothèse inverse, de très larges recouvrements auraient dû être obtenus. Parmi ces 24 gènes, 4 ont une expression plus faible en absence de Fra-1 comparée à la condition siRNA contrôle, ils sont donc up-régulés par Fra-1. Les 20 autres gènes ont une expression plus forte en absence de Fra-1 comparée au siRNA contrôle, ils sont donc down-régulés par Fra-1.

Total des gènes modulés en absence de Fra-1

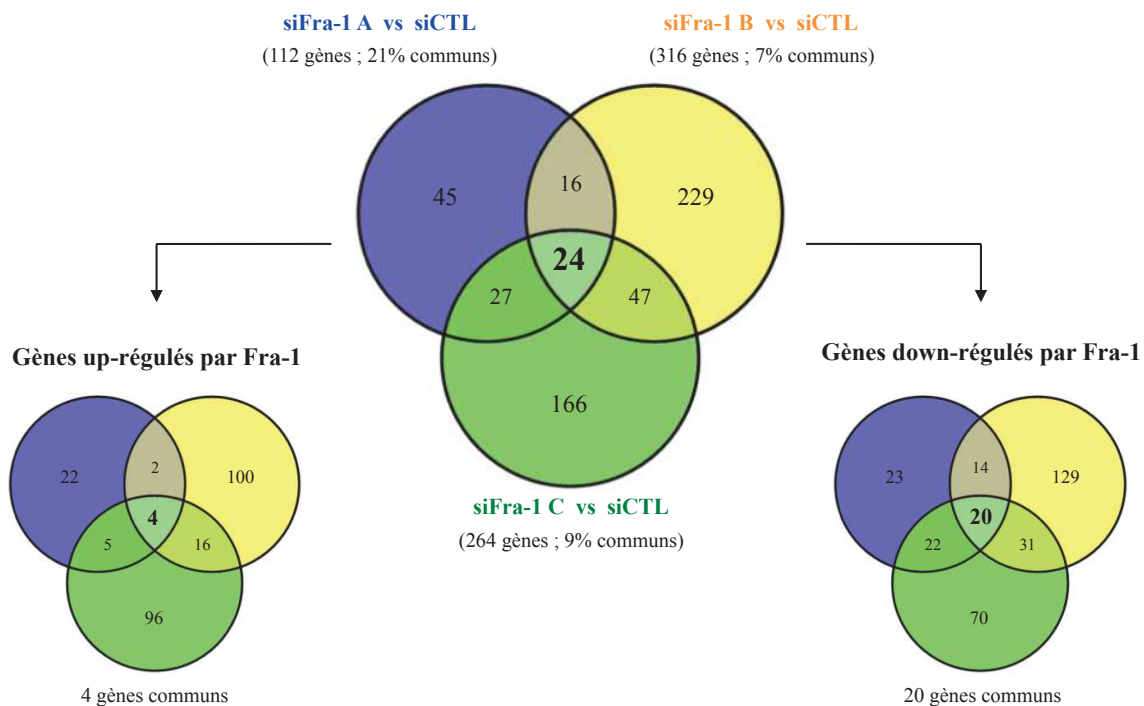


Figure 27 : Gènes différentiellement exprimés en absence de Fra-1, réprimés par RNAi. 24 gènes ont une expression dérégulée d'un facteur supérieur à 2 par chacun des trois siRNA dirigés contre Fra-1, soit un recouvrement des transcriptomes variant entre 7 et 21% selon les siRNA. p value < 0,05.

Pour valider ces résultats par RT-qPCR au laboratoire j'ai choisi parmi les gènes modulés de façon commune par les trois siRNA, 4 gènes up-régulés et 4 gènes down-régulés par Fra-1. Dans cette analyse j'ai inclus la condition non-transfectée qui correspond aux cellules qui n'ont pas été transfectées avec un siRNA, mais qui ont reçu la quantité correspondante d'agent de transfection (Interferin). Les résultats ont fait apparaître que 50% des gènes analysés étaient des faux positifs, liés à un biais généré par le siRNA contrôle.

A titre d'exemple, le gène FBXW8 a été identifié comme étant réprimé par Fra-1 dans l'analyse transcriptomique car son expression est augmentée en absence de Fra-1 (siFra-1) comparée à la condition siRNA contrôle (siCTL). Cependant, quand on regarde les niveaux d'expression par RT-qPCR (**Figure 28.A**) il apparaît que par rapport à la condition non-transfectée (NT), le siCTL entraîne une diminution de moitié de l'ARNm de FBXW8 alors que le siFra-1 ne modifie pas le niveau d'expression. En revanche, si on prend le cas du gène GDF15, l'analyse par RT-qPCR confirme que ce gène est effectivement modulé par l'absence de Fra-1, mais la conclusion n'est plus la même selon la condition contrôle considérée (**Figure 28.B**). Si l'on considère la condition siRNA contrôle (siCTL) comme condition contrôle, comme c'est souvent le cas dans la littérature, alors GDF15 apparaît comme régulé positivement par Fra-1. Par contre, si l'on considère la condition non-transfectée (NT) comme condition contrôle, GDF15 apparaît comme régulé négativement par Fra-1.

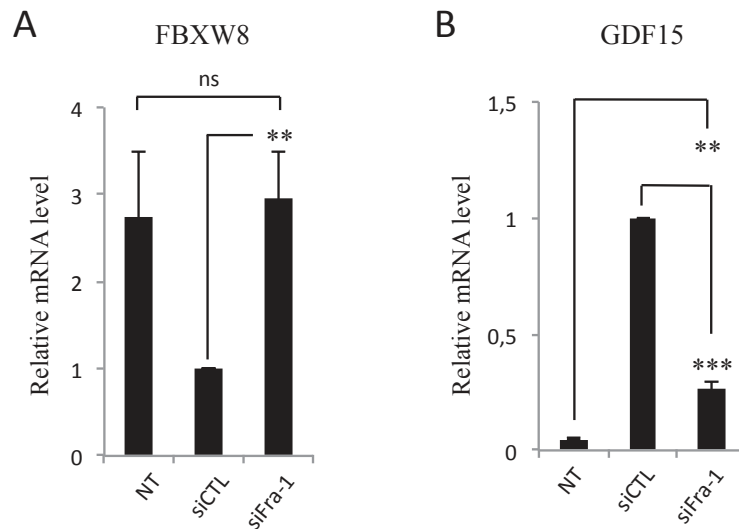


Figure 28 : Exemple de biais générés par l'utilisation du siRNA contrôle. A) FBXW8 B) GDF15. Les valeurs de RT-qPCR sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm mesurée dans la condition siRNA contrôle (siCTL). n=3. Two tailed paired t-test ; (ns) non significatif ; (**) p value <0.01 ; (***) p value <0.001. Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.

Suite à cette observation, il est clair que le siRNA contrôle commercial utilisé (Dharmacon) induit un biais important et conduit à une interprétation erronée des résultats. Nous avons donc testé d'autres siRNA contrôles commerciaux et nous avons choisi le siRNA contrôle Ambion pour poursuivre notre étude (voir plus loin, paragraphe II.2). Celui-ci entraîne de moins grandes variations d'expression, du moins sur les quelques gènes testés (**Figure 29**).

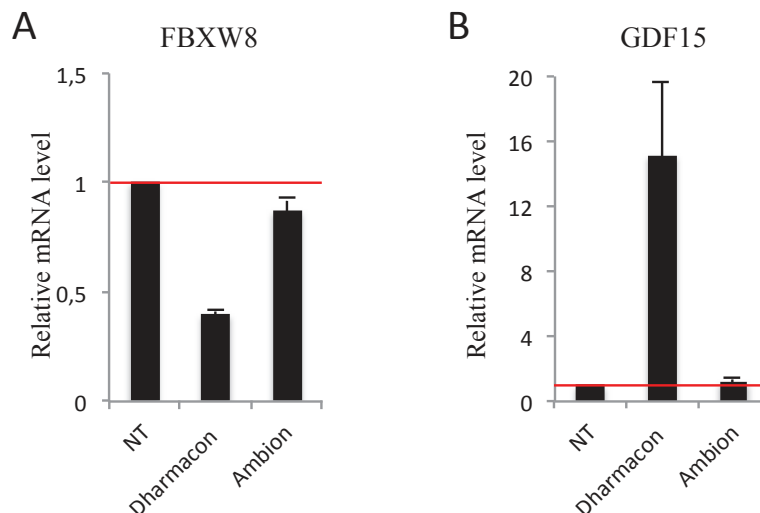


Figure 29 : Comparaison des modulations induites par deux siRNA contrôles commerciaux. A) FBXW8 B) GDF15. Les valeurs de RT-qPCR sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm mesurée pour la condition non-transfectée (NT). n=3. Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.

II.1.b. Discussion sur les biais liés aux siRNA

Ces observations attirent l'attention sur le fait que les biais liés aux siRNA contrôles commerciaux sont probablement sous-estimés, voire non réalisés, dans bon nombre d'études transcriptomiques publiées, basées sur la comparaison de l'expression des gènes entre un siRNA contrôle et des siRNA spécifiques.

Après une lecture approfondie de la littérature concernant l'utilisation de siRNA, il apparaît que même si le design des siRNA est réalisé de telle façon qu'ils aient une homologie parfaite avec seulement une séquence spécifique d'ARNm, ils peuvent avoir une homologie partielle avec d'autres ARNm et agir comme microRNA (**Figure 30**) (Cullen, 2006; Gurtan and Sharp, 2013; Pei and Tuschl, 2006).

De plus, les siRNA et les microRNA partagent la même machinerie cellulaire (Wilson and Doudna, 2013). La transfection de siRNA peut donc entraîner une saturation de la machine RNAi cellulaire et modifier la fonction de microRNA endogène. Par ailleurs, plusieurs études ont décrit que l'introduction de siRNA peut activer le système immunitaire inné en déclenchant notamment une réponse anti-virale du type interféron de type I, induisant une dégradation non spécifique des ARNm (Bridge et al., 2003; Reynolds et al., 2006; Sledz et al., 2003).

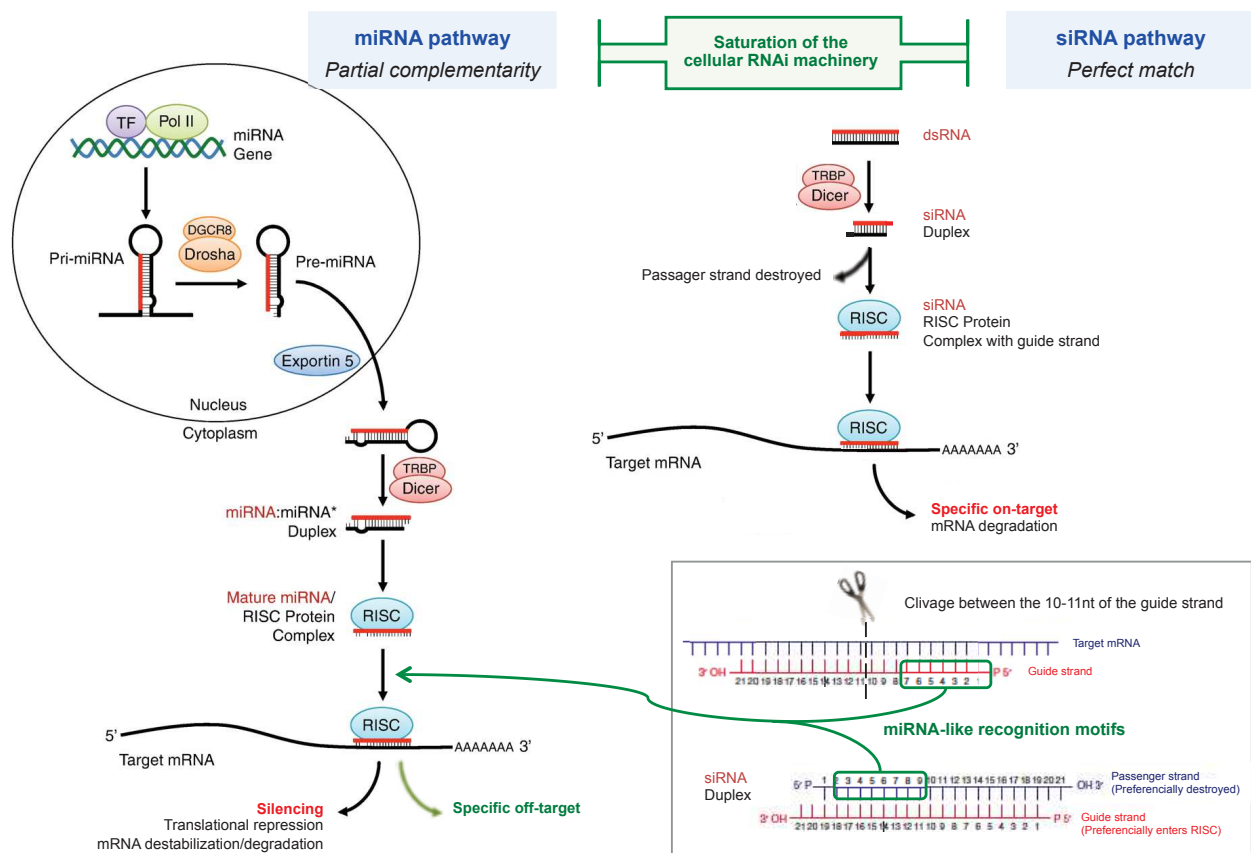


Figure 30 : Représentation schématique de la machinerie des ARN interférents et des biais générés par l'utilisation de siRNA. Adapté d'après Pei & Tuschl (2006) et Gurtan et al. (2013).

Quelques études se sont focalisées sur ces aspects dans le but d'optimiser les protocoles de transfection de siRNA (Cullen, 2006; Pei and Tuschl, 2006). Les suggestions qui ressortent sont les suivantes :

(i) La séquence choisie doit avoir une homologie parfaite avec seulement une séquence spécifique d'ARNm (recherche BLAST), une faible stabilité à l'extrémité 5' du brin guide et une composition faible en G+C pour favoriser l'interaction avec RISC. De plus, l'ARN double brin introduit dans la cellule doit avoir une taille inférieure à 30 nucléotides pour éviter de déclencher la réponse interféron.

(ii) Il est important d'effectuer un titrage des siRNA afin de déterminer la plus faible concentration efficace pour le silencing, de manière à ne pas altérer la machinerie d'interférence. L'utilisation d'un pool de siRNA transfecté en faible quantité est décrite pour diminuer significativement les effets off-targets spécifiques à chaque siRNA, sans altérer leurs effets on-targets.

(iii) Choisir avec soin la condition contrôle négatif. La transfection d'un siRNA contrôle commercial supposé ne cibler aucun transcrite cellulaire est le choix le plus utilisé dans la littérature mais, comme mentionné plus haut, les siRNA contrôles commerciaux ont leurs propres effets off-targets et peuvent amener à identifier des faux positifs. Le fait d'ajouter une condition non-transfectée qui correspond aux cellules qui n'ont pas été transfectées avec un siRNA, mais qui ont reçu la quantité correspondante d'agent de transfection, semble une bonne option pour visualiser ces biais éventuels. Une autre alternative mentionnée dans la littérature consiste à utiliser des siRNA contrôles spécifiques pour chaque séquence utilisée (Buehler et al., 2012). Pour cela, la séquence du siRNA utilisé est mutée au niveau des trois nucléotides centraux, de manière à abroger le clivage par RISC. L'ARNm cible ne sera donc pas éteint, en revanche les effets off-targets spécifiques au siRNA seront conservés. Ainsi, la comparaison entre les siRNA spécifiques et les siRNA contrôles respectifs est supposée révéler les effets spécifiquement liés à l'extinction de l'ARNm d'intérêt. Nous avons essayé cette approche mais elle ne s'est malheureusement pas avérée concluante car les siCTL mutés générés (C911) entraînent également une diminution de la protéine Fra-1 (**Figure 31**).

Mutation des 3 nucléotides centraux du brin guide du siRNA (9-11nt)



Figure 31 : Génération de siCTL spécifiques pour chacun des trois siFra-1, d'après Buehler et al. 2012.

II.2. Deuxième étude transcriptomique : identification des transcriptomes gouvernés par Fra-1 et/ou Fra-2

Parallèlement à la validation par RT-qPCR des résultats de la première analyse transcriptomique, j'ai également pu sélectionner trois siRNA spécifiquement dirigés contre Fra-2 satisfaisant les mêmes critères que ceux appliqués pour la sélection des siRNA spécifiques de Fra-1 (diminution de Fra-2 sans changement notable des autres membres AP-1 et variations similaires de l'expression des gènes cibles potentiels testés). J'ai également mis au point la double extinction phénotypique de Fra-1+Fra-2, sachant que la diminution de la protéine Fra-1, lors de la co-transfection des siRNA dirigés contre Fra-1 et Fra-2, est un peu moins efficace pour des raisons non élucidées (**Figure 26**).

II.2.a. *Modification du protocole expérimental*

Suite à l'exploitation des résultats de l'analyse transcriptomique pour Fra-1 la stratégie utilisée a été modifiée. La nouvelle approche transcriptomique consiste à comparer (i) les transcriptomes des cellules MDA-MB231 non-transfectées et transfectées par le siRNA contrôle (Ambion), de manière à s'affranchir des biais possibles induits par le siRNA contrôle, puis (ii) à comparer les transcriptomes des cellules transfectées avec les siRNA dirigés soit contre Fra-1, soit contre Fra-2, soit contre Fra-1+Fra-2, de manière à caractériser l'ensemble des gènes spécifiquement régulés par Fra-1, spécifiquement régulés par Fra-2 et régulés de façon redondantes par Fra-1 et Fra-2.

De manière à limiter les effets off-targets spécifiques des siRNA sans affecter leur effet on-targets, tout en diminuant le nombre d'échantillons, j'ai choisi (i) de co-transfecter soit les trois siRNA dirigés contre Fra-1, soit les trois siRNA dirigés contre Fra-2, soit les six siRNA pour la double extinction phénotypique et (ii) j'ai paramétré de nouvelles conditions de manière à utiliser la quantité de siRNA la plus faible possible.

Les ARN totaux issus de trois expériences indépendantes de transfection des cellules MDA-MB231 ont été préparés et chaque échantillon a été testé pour l'extinction des protéines Fra-1 et/ou Fra-2, la diminution de l'expression des gènes cibles potentiels et la qualité des ARN, sur puce Agilent, avant d'être envoyés à la plateforme d'analyse transcriptomique du CRP au Luxembourg.

Les conditions expérimentales étant sensiblement différentes pour l'extinction de Fra-1 ou de Fra-2 (1,5nM de chacun des trois siRNA soit 4,5nM final sur 72h), comparées à celle de Fra-1+Fra-2 (1,5nM de chacun des six siRNA soit 9nM final sur 72h), les échantillons devront faire l'objet de deux analyses menées en parallèle, chaque condition nécessitant des contrôles spécifiques.

Analyse transcriptomique n°2 - T1:

- Non-transfectée
- siRNA CTL Ambion (4,5nM)
- Pool siRNA Fra-1 (4,5nM)
- Pool siRNA Fra-2 (4.5nM)

Analyse transcriptomique n°2 - T2:

- Non-transfectée
- siRNA CTL Ambion (9nM)
- Pool siRNA Fra-1 + Pool siRNA Fra-2 (9nM)

II.2.b. Analyse des données brutes

Comme mentionné précédemment, les analyses transcriptomiques ont été conduites sur puce Affymetrix « All genes 2.0ST », en collaboration avec de Dr. L. Vallar (plateforme CRP, Luxembourg). Les analyses de contrôle qualité réalisées par la plateforme sur les données brutes, nous ont fourni deux types d'informations (**Figure 32**) :

(i) Il apparaît clairement que le nouveau siRNA contrôle utilisé (Ambion) possède également des effets off-targets spécifiques. De plus, ces effets off-targets sont différents selon la quantité de siRNA utilisée. En effet, l'analyse en composante principale montre que PC2 sépare les échantillons non-transfectés (T1-NT et T2-NT), des échantillons T1-siCTL ayant reçu une concentration finale en siRNA de 4,5nM, et des échantillons T2-siCTL ayant reçu une concentration finale en siRNA de 9nM.

(ii) Il semblerait que l'absence de Fra-2 ait un impact beaucoup plus faible sur le transcriptome des cellules MDA-MB231, comparé à l'absence de Fra-1 et à l'absence conjointe de Fra-1+Fra-2. En effet, l'analyse en composante principale montre que PC1 ne parvient pas à séparer les échantillons non-transfectés (T1-NT et T2-NT) des échantillons T1-siFra-2, alors qu'elle sépare les échantillons T1-siFra-1 et les échantillons T2-siFra-1+siFra-2.

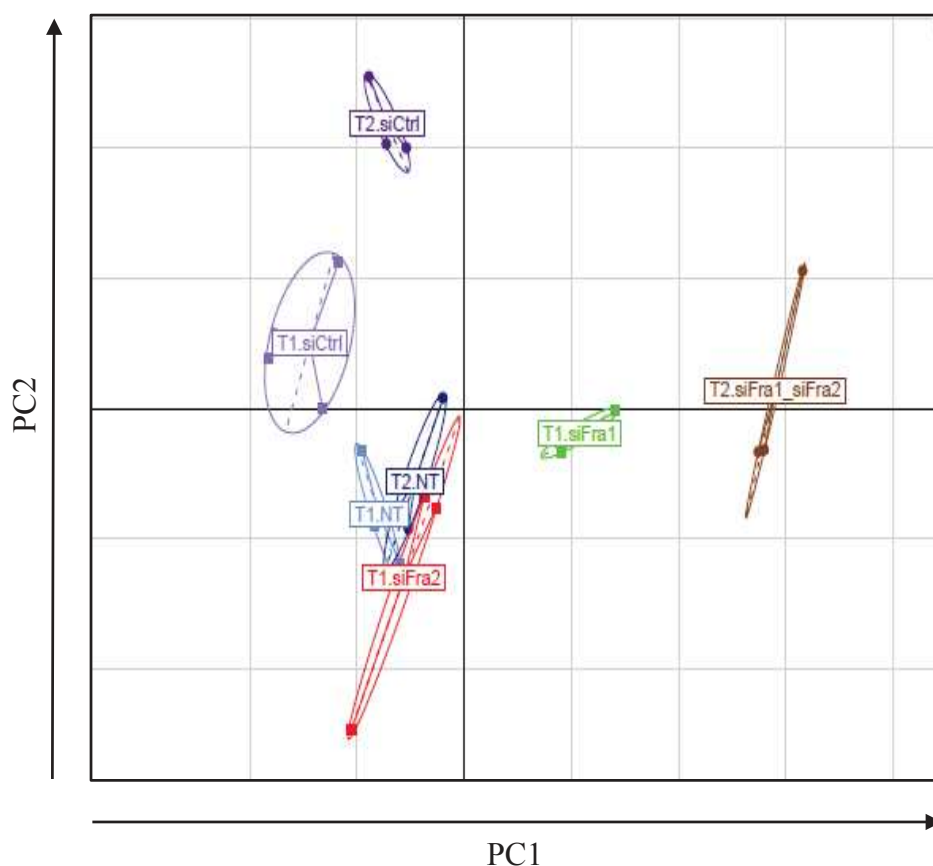


Figure 32 : Analyse en composante principale (PCA). La PCA détermine les gènes qui sont le plus discriminants dans la distribution des données. Ainsi les gènes qui sont liés se regroupent, ce qui réduit le nombre de gènes et donc le nombre de dimensions. Donc plus la variance entre les groupes est grande, plus c'est discriminatif.

II.2.c. Analyse des données transcriptomiques et « nettoyage des listes »

Le nouveau design expérimental, nous permet de nous affranchir des biais liés au siRNA contrôle et d'écartier une partie des faux positifs, ou du moins, de limiter les erreurs d'interprétation des résultats. L'analyse des données transcriptomiques a été réalisée en trois temps.

(i) La première étape de l'analyse consiste à sélectionner les gènes significativement différentiellement exprimés entre chaque condition. L'utilisation des triplicats technologiques pour chaque condition nous permet de calculer le FDR (Fold Discovery Rate) et de sélectionner les gènes significativement modulés entre chaque condition. Un gène est considéré comme significativement différentiellement exprimé lorsque son fold-change (FC) est supérieur ou inférieur à 1 et que sa FDR est inférieure à 0,05 ($FC \geq 1/\leq -1$ et $FDR < 0,05$).

(ii) La seconde étape de l'analyse consiste à croiser chaque transcriptome siRNA spécifique avec ses deux conditions contrôles associées. Ceci permet d'éliminer plusieurs catégories de gènes faux positifs. Il peut s'agir, soit, de faux positifs directement liés aux effets off-targets du siRNA contrôle, soit des faux positifs liés à la transfection en elle-même. Néanmoins, pour ne pas être trop stringant dans l'analyse, nous avons choisi de conserver certains « cas particuliers ».

Il s'agit des gènes pour lesquels le biais lié à l'utilisation du siRNA contrôle (siCTL) est non problématique pour l'interprétation des résultats.

Pour illustrer cette seconde étape je présente le cas du transcriptome siFra-1 : les exemples choisis sont tous des gènes down-régulés par Fra-1, c'est à dire que leur expression est augmentée en absence de Fra-1 (siFra-1) par rapport à la condition non-transfectée (NT), ils ont donc un FC>1 (**Figure 33**).

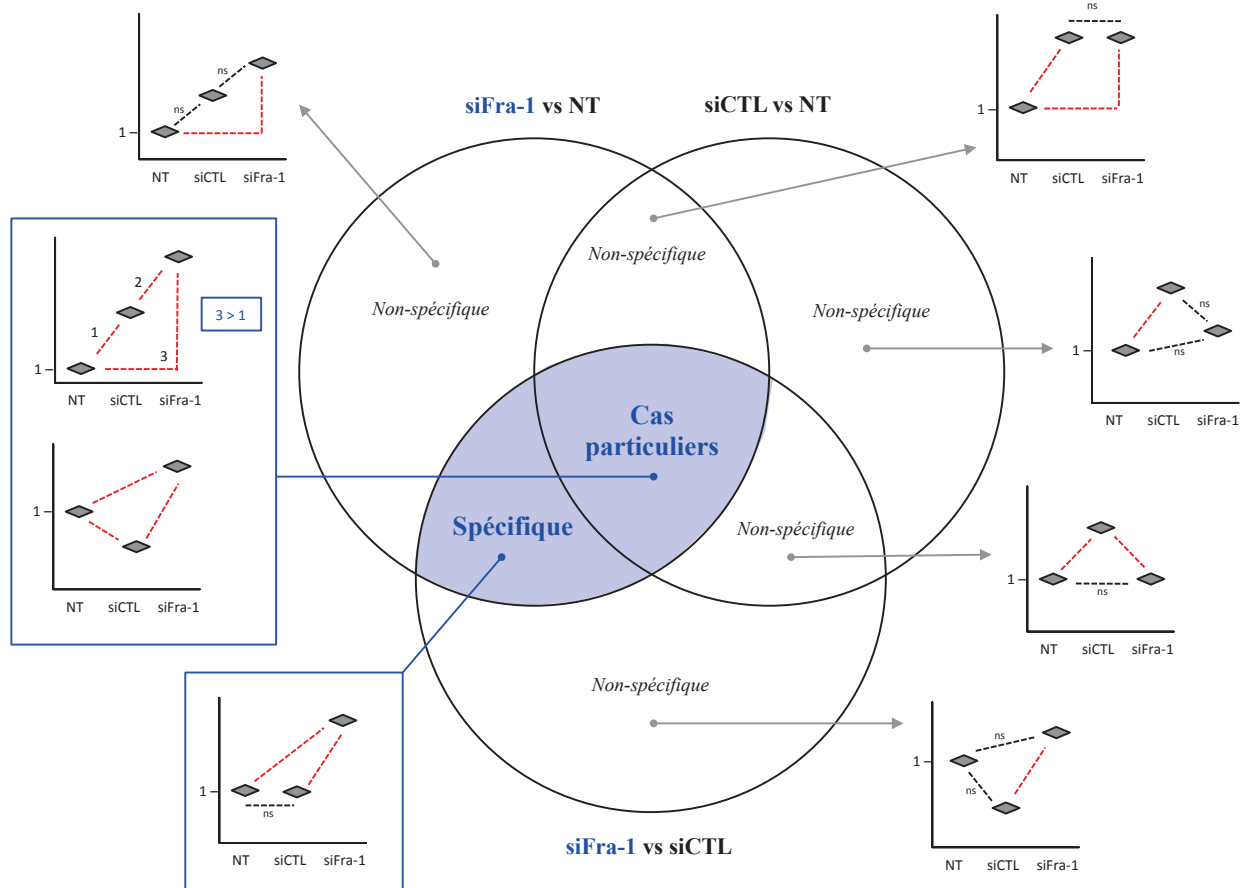
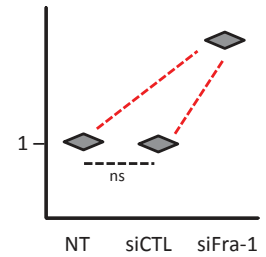


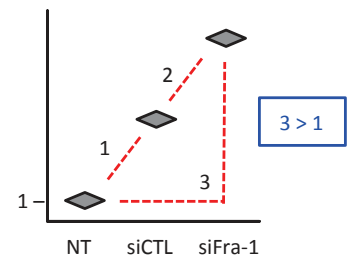
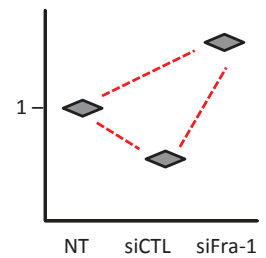
Figure 33 : Croisement du transcriptome siFra-1 avec ses deux conditions contrôles associées pour l'élimination des faux positifs. Rouge : FDR<0,05. Noir : non significatif (ns).

Les gènes affectés de façon spécifique par le pool de siFra-1 sont les gènes identifiés comme significativement modulés dans la condition siFra-1 vs NT, ainsi que dans la condition siFra-1 vs siCTL, mais qui ne sont pas retrouvés comme significativement modulés dans la condition siCTL vs NT.



Les deux « cas particuliers » que nous avons choisi de conserver se trouvent parmi les gènes identifiés comme significativement modulés dans chacune des trois conditions : siFra-1 vs NT, siFra-1 vs siCTL et siCTL vs NT.

- Comme illustré ci-contre, le premier cas que nous avons choisi de conserver est celui où l'effet du siCTL est opposé à celui des siFra-1. C'est à dire qu'en présence du siCTL le $FC < 1$ alors qu'en présence des siFra-1 le $FC > 1$. Dans ce cas, le biais du siRNA contrôle est non problématique pour l'interprétation du résultat siFra-1.
- Le second cas que nous avons choisi de conserver est celui où l'effet de l'absence de Fra-1 est plus importante que l'effet observé dans la condition siCTL. C'est à dire que la valeur absolue du FC observée dans la condition siFra-1 vs NT est supérieure à la valeur absolue du FC observée dans la condition siCTL vs NT. En effet, bien qu'il y ait un biais lié à la transfection en elle-même, l'absence de Fra-1 impacte clairement plus fortement l'expression de ce gène.



À la fin de la seconde étape de filtre on ne considère plus que 1061 gènes sur les 3098 initialement identifiés (soit 34%) pour le transcriptome Fra-1, 226 gènes sur les 1253 initialement identifiés (soit 18%) pour le transcriptome Fra-2 et 2747 gènes sur les 7024 initialement identifiés (soit 40%) pour le transcriptome Fra-1+Fra-2 (**Figure 34**). Ce qui signifie globalement que 64,5% des gènes initialement identifiés sont des faux positifs liés, soit à la transfection elle-même, soit aux effets off-targets du siRNA contrôle.

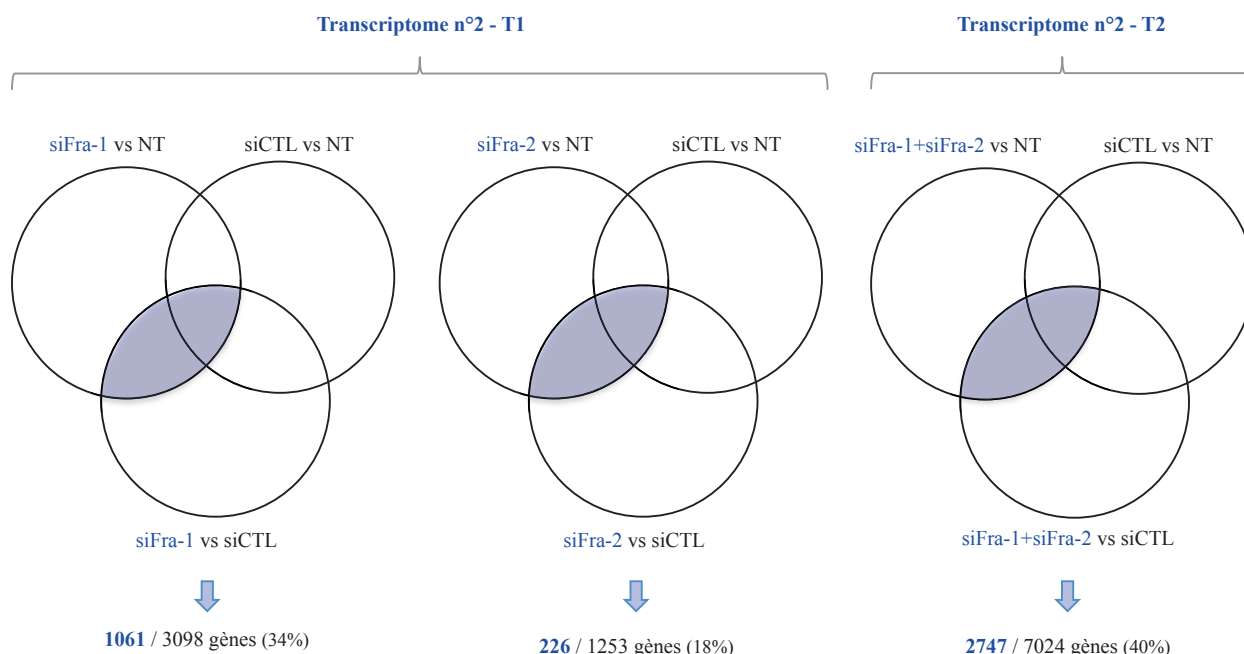


Figure 34 : Croisement de chaque transcriptome siRNA spécifique avec ses deux conditions contrôles associées pour l'élimination des faux positifs. Bleu : catégories de gènes conservées à la fin de la seconde étape de filtre.

(iii) La troisième et dernière étape de l'analyse consiste à croiser les listes obtenues pour siFra-1, pour siFra-2 et pour siFra-1+siFra-2 (**Figure 35**). A cette étape deux groupes de gènes vont être éliminés : les faux positifs directement liés aux effets off-targets spécifiques des siRNA Fra-1 et les faux positifs directement liés aux effets off-targets spécifiques des siRNA Fra-2. En effet, il n'est pas logique qu'un gène qui apparaît modulé par siFra-1 seul ou siFra-2 seul, apparaisse non modulé par siFra-1+siFra-2. À la fin de cette dernière étape, on ne considère plus que 2648 gènes comme significativement différentiellement modulés par Fra-1 et/ou Fra-2, avec un $FC \geq 1/\leq -1$ et une $FDR < 0,05$.

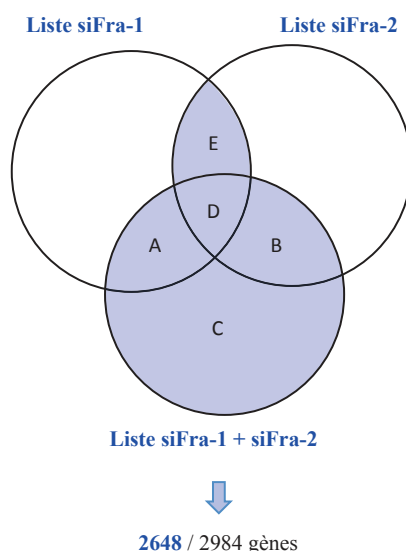


Figure 35 : Croisement des listes siFra-1, siFra-2 et siFra-1+siFra-2. Bleu : catégories de gènes conservées à la fin de la troisième étape de filtre.

II.3. Résultats

Mes analyses ont permis d'identifier 2648 gènes différentiellement exprimés en absence de Fra-1 et/ou Fra-2, parmi lesquels 1103 sont modulés avec un fold-change supérieur ou égal à 1,5 et 281 sont modulés avec un fold-change supérieur ou égal à 2. Les gènes ont une répartition à peu près équivalente entre gènes induits et réprimés par Fra-1 et/ou Fra-2. De façon intéressante, ils se distribuent en cinq groupes d'intérêt (**Tableau 7**).

- A. Les gènes dont l'expression varie uniquement quand Fra-1 et Fra-1+Fra-2 sont réprimés par RNAi, indiquant que ces gènes sont vraisemblablement sous le contrôle préférentiel de Fra-1.
- B. Les gènes dont l'expression varie uniquement quand Fra-2 et Fra-1+Fra-2 sont réprimés par RNAi, indiquant que ces gènes sont vraisemblablement sous le contrôle préférentiel de Fra-2.
- C. Les gènes dont l'expression varie uniquement quand Fra-1+Fra-2 sont réprimés par RNAi, ce qui suggère que ces gènes pourraient être régulés de façon redondante par Fra-1 et par Fra-2.
- D. Les gènes dont l'expression varie quand Fra-1, Fra-2 et Fra-1+Fra-2 sont réprimés par RNAi, se qui suggère que ces gènes pourraient être régulés de façon complémentaire ou synergique par Fra-1 et Fra-2.
- E. Les gènes dont l'expression varie uniquement quand Fra-1 et Fra-2 sont réprimés par RNAi, et pour lesquels Fra-1 et Fra-2 ont des effets opposés (induction vs répression) suggérant que Fra-1 et Fra-2 ont des rôles antagonistes sur ces gènes. Seuls 10 gènes présentent ce profil de régulation et ils ne sont que très faiblement modulés.

	2648 gènes (FC >1/<-1)	1103 gènes (FC ≥1,5/≤-1,5)	281 gènes (FC ≥2/≤-2)
A. Gènes sous le contrôle préférentiel de Fra-1	760 (433 down/ 327 up)	562 (320 down/ 242 up)	199 (123 down/ 76 up)
B. Gènes sous le contrôle préférentiel de Fra-2	61 (30 down/ 31 up)	44 (21 down/ 23 up)	29 (18 down/ 11 up)
C. Gènes régulés de façon redondante par Fra-1 et Fra-2	1771 (864 down/ 907 up)	454 (246 down/ 208 up)	45 (27 down/ 18 up)
D. Gènes régulés de façon complémentaire ou synergique par Fra-1 et Fra-2	46 (30 down/ 16 up)	43 (27 down/ 16 up)	8 (4 down/ 4 up)
E. Gènes régulés de façon antagoniste par Fra-1 et Fra-2	10	0	0

Tableau 7 : Bilan de la nouvelle analyse transcriptomique genome-wide.

Pour plus de détails, la liste des 281 gènes modulés avec un fold-change supérieur ou inférieurs à 2 ($FC \geq 2/\leq -2$ dans au moins une condition et $FDR < 0,05$) est disponible en annexe (**Annexes 1 et 2**). Les valeurs de fold-change indiquées correspondent à la modulation mesurée entre la condition siRNA spécifique versus sa condition contrôle non-transfectée associée (NT ou NT').

L'analyse ontologique des données transcriptomiques en utilisant des programmes comme DAVID (Database for Annotation, Visualisation and Intergrated Discover) et GeneGO (Thompson Reuters) n'a pas révélé de processus affectés préférentiellement par Fra-1 ou par Fra-2, ils s'avèrent être les mêmes. Ces derniers se répartissent dans plusieurs catégories ontologiques qui incluent le contrôle de la mort cellulaire, la prolifération, l'adhésion, la motilité, la migration, l'invasion cellulaire ou encore les relations avec le système immunitaire (**Figure 36**). La dérégulation de chacune de ces fonctions pouvant participer au phénotype et à l'agressivité tumorale.

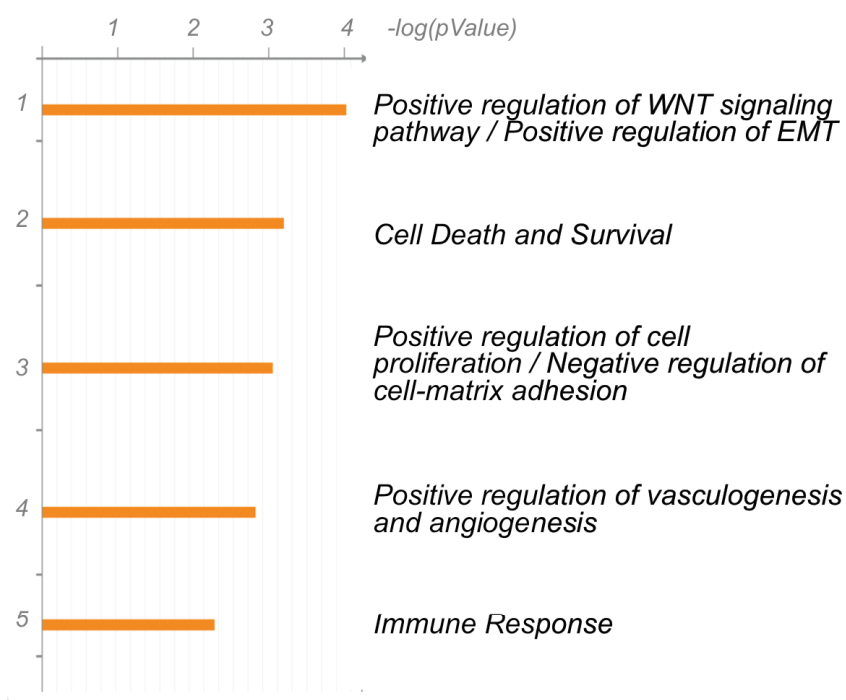


Figure 36 : Bilan de l'analyse ontologique en utilisant le logiciel GeneGo metaCore.

II.4. Conclusion sur les approches transcriptomiques

Les résultats obtenus laissent entrevoir des mécanismes moléculaires sensiblement différents pour la régulation par Fra-1 et Fra-2 des gènes cibles, selon qu'ils sont activés ou réprimés spécifiquement par Fra-1, par Fra-2 ou de façon redondantes par les deux facteurs de transcription. La redondance fonctionnelle de Fra-1 et Fra-2 n'a encore jamais été décrite et les quelques études transcriptomiques publiées après le démarrage de mes travaux de thèse, sont uniquement focalisées sur Fra-1 et décrivent donc un panel de gènes vraisemblablement incomplet.

Quatre études publiées après le démarrage de mes travaux de thèse ont abordé la caractérisation des gènes modulés par Fra-1 dans différents modèles de cancers agressifs. Leurs principales conclusions sont résumées ci-dessous :

* Desmet et al. (2013) PNAS :

En utilisant les lignées RIE (Rat intestinal épithélial) et RK3E (rat kidney epithelial) modifiées pour exprimer le récepteur neurotrophique à tyrosine kinase (TrKB) et son ligand BDNF (brain-derived neurotrophic factor), Desmet *et al* ont montré que l'activation du TrKB conduit à la surexpression de Fra-1 et à la transformation mésenchymateuse subséquente des cellules. De façon intéressante, par approche transcriptomique les auteurs ont décrit que 70% d'un panel de 52 gènes signature de l'EMT, sont régulés par Fra-1. Cependant, il n'a pas été établi si la régulation par Fra-1 est directe ou indirecte. Par ailleurs, via un criblage de drogues par létalité synthétique, ils ont montré que les antagonistes de ADORA2B (Adenosine Receptor A2B) sont préférentiellement toxiques pour les cellules de cancer du sein surexprimant Fra-1 et que l'inhibition de ADORA2B (siRNA ou drogue) conduit à une diminution de la métastatisation. Même si cette inhibition est moins efficace que celle de Fra-1 sur l'agressivité tumorale, le ciblage d'ADORA2B a été proposé par les auteurs comme une nouvelle stratégie thérapeutique potentielle (Desmet et al., 2013).

* Diesch et al. (2014) Plos One :

Diesch *et al* ont montré que Fra-1 régule l'EMT dans la lignée cellulaire BE issue d'un carcinome colorectal hautement invasif et caractérisé par l'activation constitutive de la voie Ras. Pour identifier les gènes cibles, les auteurs ont conduit une approche transcriptomique sur la lignée BE en présence et en l'absence de Fra-1 (siRNA) et dans le but d'identifier les sites de liaison de Fra-1 sur la chromatine, ils ont transfecté les cellules BE avec un transgène Flag-Fra-1, codant pour une forme de Fra-1 étiquetée et immunoprécipitable par son étiquette. Leur analyse transcriptomique a révélé environ 800 gènes régulés positivement et 1400 régulés négativement par Fra-1. L'intersection avec les résultats de ChIP-Seq dans la lignée cellulaire BE a montré que 72% des gènes régulés contiennent des sites de fixation pour Fra-1 situés à ± 5 kb de leur TSS

(Transcriptional Start Site). Ils en ont conclu que ces gènes étaient des cibles directes de Fra-1. L'analyse ontologique a révélé que Fra-1 intervient principalement dans la transition épithélio-mésenchymateuse (EMT) et l'adhésion cellulaire (Diesch et al., 2014).

* Bakiri et al. (2014) Cell death & diff :

Par surexpression de Fra-1 dans la lignée EpH4, lignée épithéliale murine non tumorigénique, Bakiri *et al* ont identifié environ 400 gènes up-régulés et 250 down-régulés par Fra-1. Sur la base d'analyses ontologiques, les voies les plus significativement affectées par ces modulations d'expression sont (i) pour les gènes up-régulés : la signalisation IL-6 et HIF1 α /hypoxie et l'organisation et la dégradation de la matrice extracellulaire et (ii) pour les gènes down-régulés : l'organisation des jonctions cellulaires et des jonctions serrées. L'altération de ces voies est impliquée dans l'EMT (Bakiri et al., 2015).

* Zhao et al. (2014) Cancer Res :

L'étude de Zhao *et al* a été menée sur une lignée de cancer du sein triple négatif humain appartenant au sous groupe basal-like, les BT-549. Les auteurs ont cartographié par ChIP-seq les sites de fixation de Fra-1 et c-Jun sur la chromatine. Ils ont observé que 84% des sites liant Fra-1 lient aussi c-Jun. Les analyses transcriptomiques ($\pm$ siRNA dirigés contre Fra-1 ou c-Jun) ont fait apparaître une expression différentielle de 419 gènes pour Fra-1 et 690 pour c-Jun. Ils ont rapporté que 180 gènes sont induits par à la fois Fra-1 et c-Jun et que 42 gènes sont réprimés par à la fois Fra-1 et c-Jun. Ils sont principalement impliqués dans la réponse immune, l'inflammation et l'adhésion cellulaire. La moitié des gènes régulés par Fra-1 ou c-Jun contiennent des sites de binding à ± 50 kb de leur TSS, ce qui a conduit les auteurs à conclure que ces gènes sont sans doute de cibles directes de Fra-1 et/ou c-Jun. Cependant, la validation expérimentale de leur conclusion n'a pas été apportée (Zhao et al., 2014).

Il ressort de ces études au moins deux notions importantes : (i) Fra-1 est impliqué dans la transition épithélio-mésenchymateuse dans différents modèles de cancers épithéliaux agressifs et (ii) suivant les modèles expérimentaux le nombre de gènes régulés par Fra-1 est variable. Par ailleurs, ces approches transcriptomiques comparent l'expression des gènes entre des siRNA contrôles et des siRNA spécifiques. Or, comme souligné au paragraphe II.1, les biais liés aux siRNA contrôles commerciaux sont importants et ils n'ont pas été mentionnés dans ces études.

Par ailleurs, l'étude de Zhao *et al* ayant été réalisée dans un modèle de cancer du sein triple négatif humain, nous avons comparé nos données transcriptomiques (MDA-MB231 $\pm$ siFra-1) avec les leurs (BT-549 $\pm$ siFra-1) dans le but d'identifier des gènes communs éventuels. En collaboration avec T. Kaoma de la plateforme d'analyse transcriptomique du CRP au Luxembourg, j'ai exploité les données transcriptomiques de Zhao *et al* déposées sur GEO :

numéro d'accès GSE46440. Il s'est avéré que dans les conditions $FC \geq 2/\leq -2$ et $FDR < 0,05$ je ne retrouve pas 419 gènes différentiellement exprimés en absence de Fra-1 comme décrit dans leur papier mais seulement 90 gènes. De plus, dans les conditions $FC \geq 1/\leq -1$ et $FDR < 0,05$, leur liste de gènes différentiellement exprimés révèle 153 gènes, et seuls 12 gènes sont communs avec nos données transcriptomiques (**Tableau 8**). Ceci souligne la variabilité d'un type cellulaire à l'autre.

	Zhao et al. 2014 $FC \geq 1/\leq -1$ et $FDR < 0,05$	TOLZA et al. 2016 $FC \geq 1/\leq -1$ et $FDR < 0,05$		
GnSymbol	siFra-1 vs siCTL	siFra-1 vs siCTL	siFra-1 vs NT	siCTL vs NT
C10orf54	2,82	1,3	1,43	NA
SDPR	2,57	1,97	1,46	NA
KRT7	2,55	1,77	1,85	NA
NFE2L3	2,28	1,75	1,63	NA
IL6	2,06	1,59	1,69	NA
PLSCR4	1,98	1,72	1,72	NA
CD274	1,94	1,20	-1,41	NA
GRAMD3	1,89	1,26	1,33	NA
EDN1	1,84	2,38	1,94	NA
PAPPA	1,82	1,59	1,46	NA
EFHD1	-1,78	-1,10	1,56	NA
JARID2	-1,81	-1,82	-1,45	NA

Tableau 8 : Bilan du croisement entre nos données transcriptomiques et celles publiées par Zhao et al. Notre analyse transcriptomique a été conduite sur puce Affymetrix « All genes 2.0ST » et celle de Zhao *et al* sur puce Affymetrix « Human Gene 1.1ST array ».

Cependant il faut souligner un certain nombre de différences entre nos deux études : (i) les lignées triples négatives utilisées sont différentes, (ii) les pools siFra-1 et siCTL utilisés ne sont pas les mêmes, (iii) la condition NT n'est pas présente dans leur analyse, ils n'ont donc pas pris en compte les biais liés à l'utilisation du siCTL, et enfin (iv) notre analyse transcriptomique a été conduite sur puce Affymetrix « All genes 2.0ST » et celle de Zhao *et al* sur puce Affymetrix « Human Gene 1.1ST array ».

III. IDENTIFICATION ET CARACTÉRISATION DES SITES DE FIXATION DE FRA-1 ET FRA-2 SUR LA CHROMATINE DE LA LIGNÉE MDA-MB231

L'analyse des mécanismes transcriptionnels gouvernés par Fra-1 et Fra-2 nécessite de sélectionner des gènes qui sont des cibles directes et non des cibles de facteurs de transcription eux-mêmes régulés par Fra-1 et/ou Fra-2 ou par des mécanismes post-transcriptionnels dépendants de Fra-1 et/ou Fra-2 (microARN...). Dans ce but, nous avons entrepris de caractériser la nature des éléments fonctionnels liés par Fra-1 et/ou Fra-2 (enhancers ou promoteurs, actifs ou inactifs) puis de procéder à l'attribution des enhancers contenant Fra-1 et/ou Fra-2 aux gènes qu'ils régulent.

III.1. Identification des sites de fixation de Fra-1 et Fra-2

Nous avons cartographié les sites de fixation de Fra-1 et Fra-2 sur le génome des cellules MDA-MB231 par ChIP-seq : immunoprécipitation de chromatine et séquençage à haut débit. Les protéines Fra-1 et Fra-2 doivent hétérodimériser avec les protéines de la famille JUN. c-Jun étant le membre JUN le plus abondant dans la lignée MDA-MB231 (**Figures 24 et 25**) nous avons également effectué des ChIP-seq pour c-Jun dans le but de collecter des informations supplémentaires.

La chromatine associée à Fra-1, Fra-2 et c-Jun a été séquencée par Illumina HighSeq 2500 sur la plateforme MGX, Montpellier. Comme contrôle, l'ADN non soumis à l'immunoprécipitation (Input) a été séquencé en parallèle et l'utilisation de réplicats biologiques pour chaque condition de ChIP-seq nous permet d'augmenter la fiabilité des résultats obtenus. L'analyse bioinformatique des données de ChIP-seq a été réalisée en collaboration avec A.M. Maqbool (IGMM, Montpellier). Brièvement, après les étapes de contrôle qualité, les séquences obtenues (reads) ont été alignées sur le génome de référence (Feb.2009 GRCh37/hg19) et la recherche des pics (peak calling) a été réalisée pour chaque échantillon. Dans le but d'éliminer d'éventuels artéfacts, les réplicats ont été superposés et seuls les pics présents dans les deux réplicats ont été conservés. Enfin, les données d'Input ont été soustraites pour éliminer le bruit de fond.

Ainsi, des distributions de pics uniques et spécifiques ont été obtenues pour Fra-1, Fra-2 et c-Jun : respectivement 9998, 13003 et 8944 pics. De plus, la recherche de motifs *de novo* en utilisant le logiciel HOMER a révélé un enrichissement supérieur à 70% pour les motifs AP-1, indiquant que les immunoprécipitations Fra-1, Fra-2 et c-Jun sont satisfaisantes et que nos données de ChIP-seq sont de bonne qualité.

Comme illustré sur la **Figure 37**, on observe un large recouvrement des pics Fra-1 et Fra-2 dans les MDA-MB231, suggérant que si Fra-1 et Fra-2 peuvent se fixer de façon spécifique sur certains sites génomiques, une large majorité d'entre eux se fixe sur les mêmes sites. En effet, 8070 pics sont communs à Fra-1 et Fra-2, soit 81% des pics Fra-1 et 62% des pics Fra-2. Ces données sont cohérentes avec mes résultats transcriptomiques, suggérant que Fra-1 et Fra-2 interviendraient spécifiquement sur certains gènes, de façon redondante sur d'autres, ou encore agiraient de façon complémentaire ou synergique dans certaines situations.

De façon surprenante, les cellules MDA-MB231 exprimant plus de Fra-1 que de Fra-2 on se serait attendu à obtenir plus de pics Fra-1 que de pics Fra-2. Ceci est sans doute dû à l'efficacité des anticorps utilisés. En effet, l'anticorps dirigé contre Fra-1 est moins efficace que celui dirigé contre Fra-2 en ChIP.

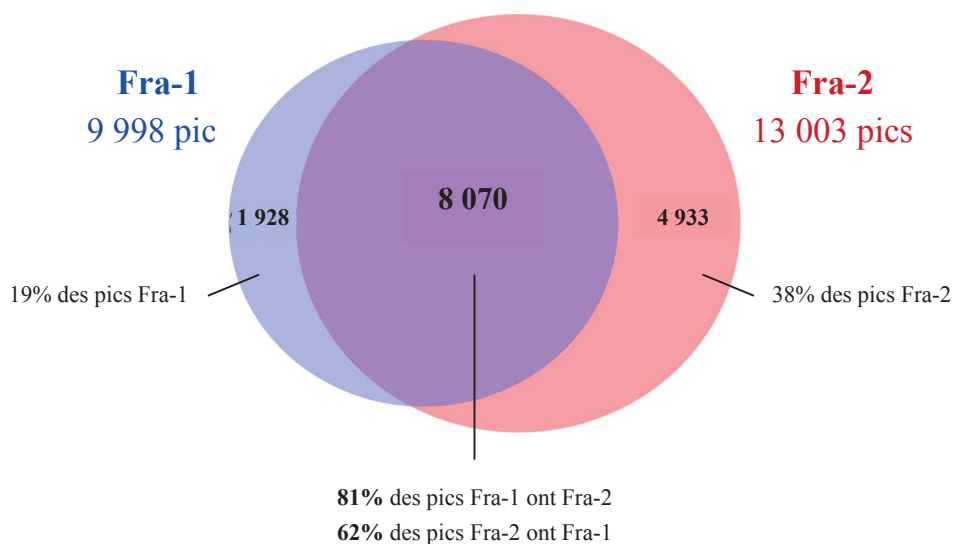


Figure 37 : Intersection entre les pics Fra-1 et Fra-2 obtenus par ChIP-seq dans la lignée MDA-MB231.

Par ailleurs, l'intersection des données de ChIP-seq pour Fra-1, Fra-2 et c-Jun a révélé que 23,2 % des pics Fra-1 et/ou Fra-2 ont également un pic c-Jun et que 38,8% des pics c-Jun sont communs aux pics Fra-1 et/ou Fra-2 (**Figure 38**). Ceci suggère que bien que c-Jun soit le membre JUN le plus abondant dans la lignée MDA-MB231, il n'est pas le partenaire principal de Fra-1 et Fra-2 dans cette lignée.

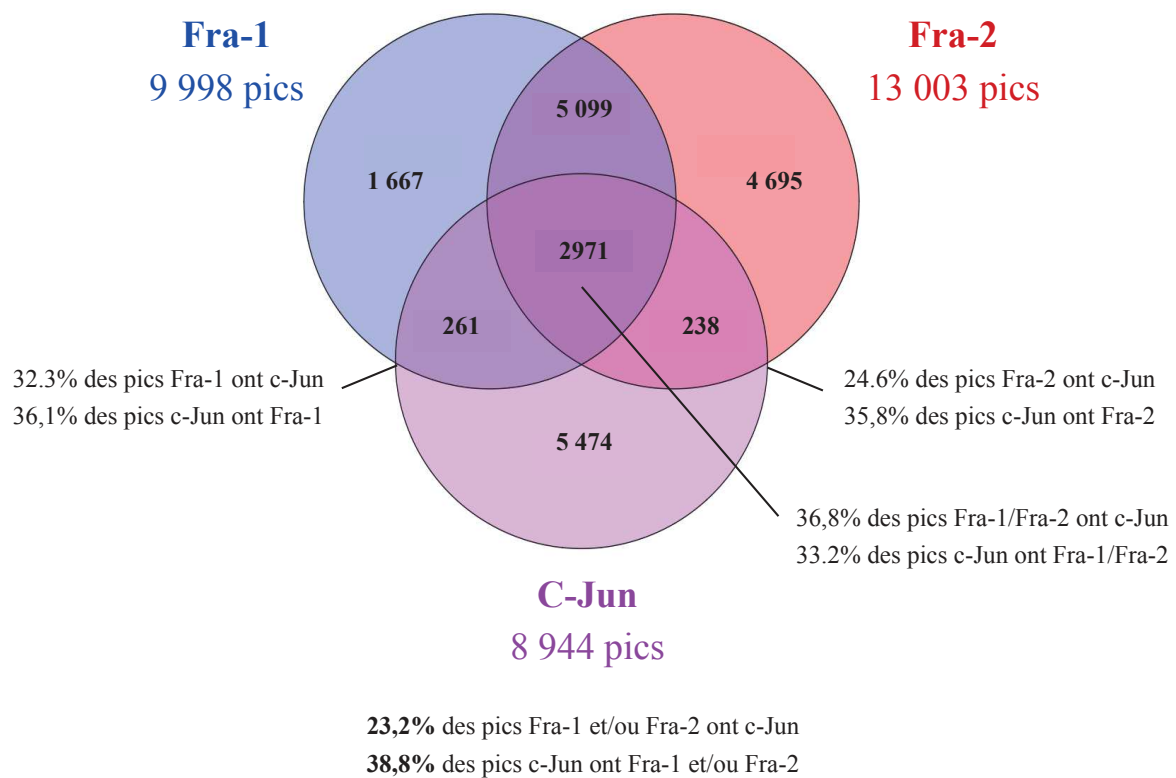


Figure 38 : Intersection entre les pics Fra-1, Fra-2 et c-Jun obtenus par ChIP-seq dans la lignée MDA-MB231

III.2. Caractérisation des sites de fixation de Fra-1 et Fra-2

Il est maintenant établi que les éléments régulateurs possèdent des signatures chromatinienne permettant de les identifier. Ainsi, la mono-méthylation de l'histone H3 sur la lysine 4 (H3K4me1) définit les régions enhancers, la tri-méthylation de l'histone H3 sur la lysine 4 (H3K4me3) définit les régions promoteurs et la présence combinée de ces marques avec l'acétylation de l'histone H3 sur la lysine 27 (H3K27ac) révèle leur activité (voir l'introduction de la thèse, paragraphe III.2.b).

La cartographie des marques d'histones dans la lignée MDA-MB231 n'est pas répertoriée dans la base de données ENCODE. Nous avons donc effectué des ChIP-seq H3K4me1 et H3K4me3 dont le rapport permet de définir les régions enhancers et promoteurs, ainsi que H3K27ac pour identifier si ces éléments sont actifs ou non dans notre lignée. Des ChIP-seq pour cartographier la tri-méthylation de l'histone H3 sur la lysine 36 (H3K36me3) ont également été réalisés afin d'évaluer le niveau de transcription des gènes. Nous avons également effectué des ChIP-seq pour l'histone acétyltransférase P300 et pour Pol II. En effet, P300 a été décrite comme une marque des enhancers actifs et sa présence est utilisée pour leur définition. Par ailleurs, de façon récente, il a été mis en évidence que les enhancers actifs recrutent Pol II et que cette dernière initie la transcription à leur niveau pour produire des eRNA (enhancer RNA) instables et bidirectionnels.

Comme précédemment, des réplicats ont été séquencés sur la plateforme MGX, Montpellier et l'analyse bioinformatique des données de ChIP-seq a été réalisée en collaboration avec A.M. Maqbool (IGMM, Montpellier) de façon identique à celle décrite au paragraphe précédent.

L'exploitation bioinformatique des données de ChIP-seq réalisée en collaboration avec C. Lecellier (IGMM, Montpellier) et F. Bejjani en thèse au laboratoire, est détaillée dans les paragraphes suivants. Cette analyse a révélé que Fra-1 et Fra-2 lient préférentiellement des régions enhancers dans la lignée MDA-MB231.

III.2.a. Identification des régions enhancers dans la lignée MDA-MB231

L'intersection entre les données ChIP-seq pour H3K4me1 et H3K4me3 a permis d'identifier 78297 régions potentiellement enhancers dans la lignée MDA-MB231 wild-type. Ces régions ont été définies en se basant sur le ratio H3K4me1/H3K4me3 : plus précisément une région est définie comme enhancer lorsque le niveau de H3K4me1 est supérieur à celui de H3K4me3 (H3K4me1>me3).

Dans un premier temps, l'intersection entre ces régions enhancers et les données ChIP-seq pour H3K27ac a permis de distinguer deux groupes d'enhancers : les enhancers actifs (35%) et les enhancers inactifs (65%) (**Figure 39**). La présence combinée de H3K27ac, qui révèle l'activité des régions enhancers, a ainsi permis d'identifier environ 27500 enhancers actifs dans la lignée MDA-MB231 wild-type.

Dans un second temps, l'intersection entre les enhancers actifs et les données ChIP-seq pour Pol II et P300 a révélé, de façon surprenante, que seule une minorité des enhancers actifs identifiés ont une présence combinée de Pol II et/ou P300 (**Figure 39**). En effet, seul 6% des enhancers actifs fixent également Pol II et 19% fixent également P300. Ainsi, plus de 75% des enhancers actifs définis par la présence de H3K27ac ne fixent ni Pol II ni P300 (**Figure 39**).

Dans la suite des études bioinformatiques, nous avons donc considéré l'ensemble des enhancers identifiés sans distinction entre ceux fixant ou non Pol II et/ou P300. La définition de ces enhancers est donc exclusivement basée sur les deux critères suivants : ratio H3K4me1>me3 et la présence ou absence de H3K27ac.

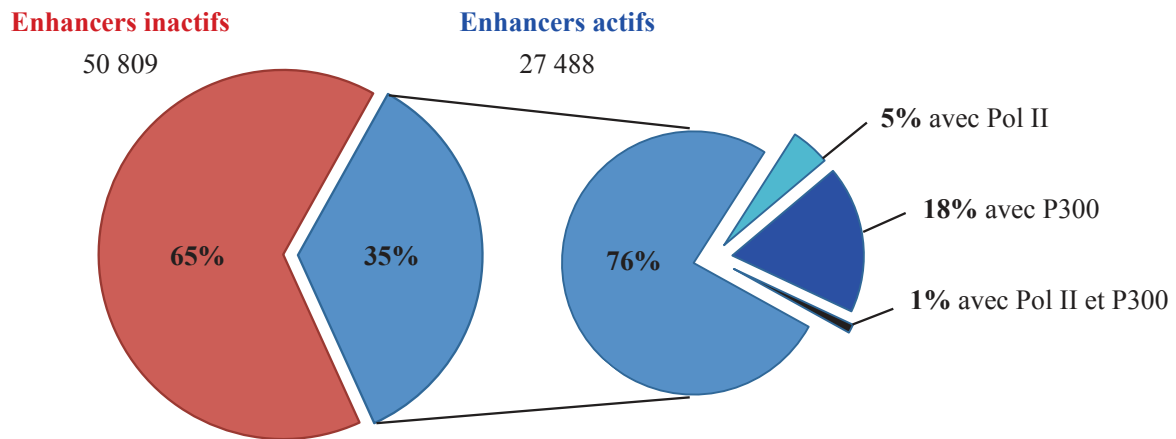


Figure 39 : Descriptif des enhancers identifiés dans la lignée MDA-MB231. Les enhancers actifs et inactifs ont été identifiés en se basant sur le ratio H3K4me1>H3H4me3 et la présence ou l'absence de H3K27ac. Puis la distinction a été faite entre les enhancers actifs liant également Pol II et/ou P300.

III.2.b. Caractérisation des régions régulatrices liées par Fra-1 et/ou Fra-2 dans la lignée MDA-MB231

La distribution des marques H3K4me1 et H3K4me3 autour des pics Fra-1 et/ou Fra-2 a fait apparaître que les facteurs Fra-1 et Fra-2 semblent se lier préférentiellement à des régions présentant les caractéristiques d'enhancers dans la lignée MDA-MB231 wild-type (**Figures 40**). En effet, le signal H3K4me1 moyen autour des pics, caractéristique des enhancers, est trois fois plus fort que celui de H3K4me3, caractéristique des promoteurs.

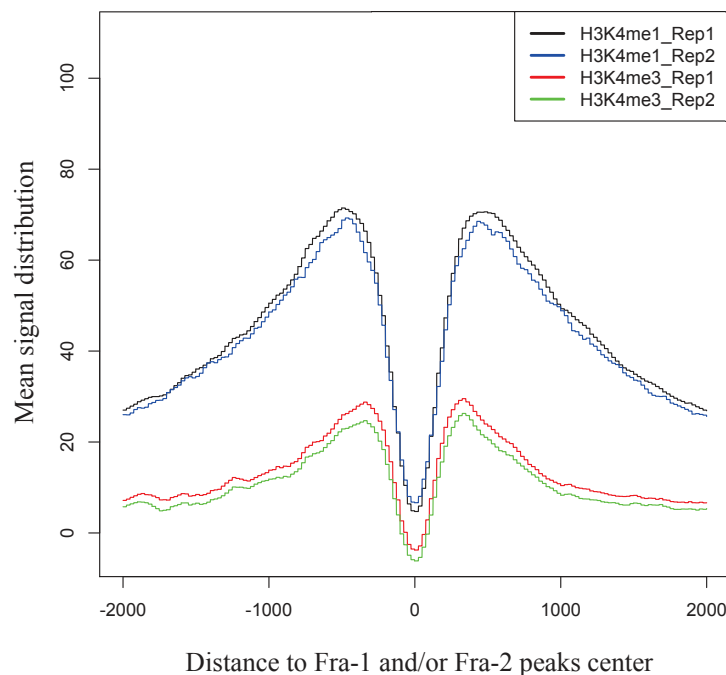
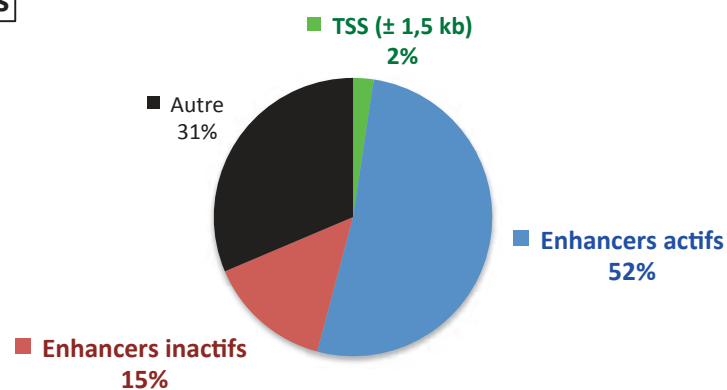


Figure 40 : Distribution des marques H3K4me1 caractéristique des enhancers et H3K4me3 caractéristique des promoteurs autour des pics Fra-1 et/ou Fra-2 (± 2 kb du centre du pic). D'après les données des deux réplicats de ChIP-seq pour les marques H3K4me1 et H3K4me3, les pics Fra-1 et Fra-2 se lient préférentiellement au niveau des régions enhancers dans les MDA-MB231.

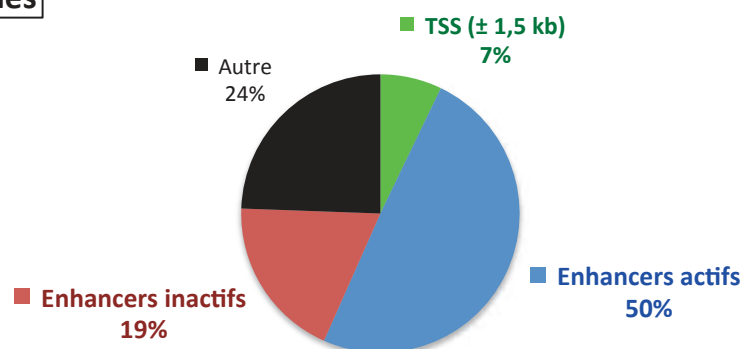
De plus, la distribution des pics Fra-1 et/ou Fra-2 dans le génome de la lignée MDA-MB231 confirme que plus de 70% des pics Fra-1 et/ou Fra-2 se lient au niveau des régions enhancers : 55% dans les enhancers actifs et 15% dans les enhancers inactifs (**Figure 41**).

Il est important de noter que seul 4% des pics Fra-1 et/ou Fra-2 se lient au niveau des régions promoteurs ($\pm 1,5\text{kb}$ des TSS Refseq) dans la lignée MDA-MB231 wild-type. Les 26% restants sont des pics Fra-1 et/ou Fra-2 liés à des régions autres que enhancers et promoteurs (**Figure 41**).

Pics Fra-1 uniques



Pics Fra-2 uniques



Pics Fra-1 et Fra-2

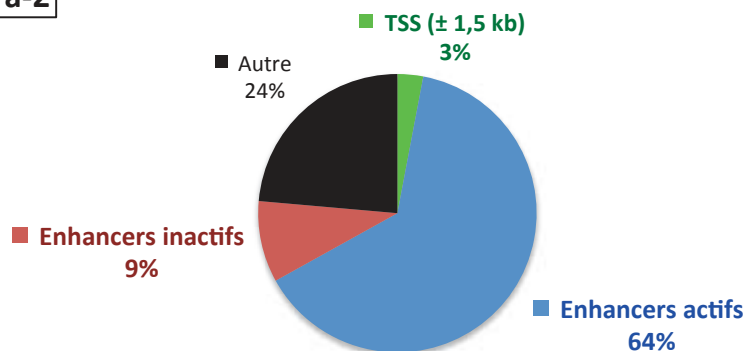


Figure 41 : Distribution des pics Fra-1 et/ou Fra-2 entre les promoteurs ($\pm 1,5\text{kb}$ des TSS) et les enhancers actifs et inactifs. Les pics Fra-1 et/ou Fra-2 se lient préférentiellement au niveau des régions enhancers dans les MDA-MB231, et plus particulièrement dans les régions enhancers actives.

III.2.c. Identification des gènes modulés par Fra-1 et/ou Fra-2 via leurs régions enhancers et/ou promoteurs dans la lignée MDA-MB231

Une des grandes difficultés dans le domaine de la transcription est l'assignation des enhancers aux gènes qu'ils régulent. De manière assez fréquente, les enhancers sont assignés aux gènes les plus proches. Même si cette approximation se vérifie pour un certain nombre de gènes, on sait maintenant que de nombreux enhancers sont situés à plusieurs dizaines ou centaines de kilobases des promoteurs correspondants et qu'ils sont rapprochés tridimensionnellement par des boucles chromatinienne.

Dans le cadre de mon projet, j'ai eu recours à l'atlas des enhancers actifs défini par le consortium FANTOM5 (Andersson et al., 2014) (voir l'introduction de la thèse, paragraphe III.5). FANTOM5 a identifié et répertorié plus de 44000 enhancers actifs à travers la majorité des types cellulaires humains, en utilisant la technologie CAGE (Cap Analysis of Gene Expression) (Kodzius et al., 2006; Shiraki et al., 2003) combinée à la cartographie des marques chromatinienne (modifications d'histones, DNase I, facteurs de transcription, etc.) et ceci dans plus de 400 cellules primaires humaines issues de 150 tissus et environ 250 lignées cancéreuses humaines représentant 150 sous-types distincts de cancers. La lignée MDA-MB231 n'est pas incluse dans les lignées cellulaires analysées par FANTOM5. Sur la base des corrélations d'expression entre les ARN capés bidirectionnels issus de la transcription de régions régulatrices, eRNA pour les enhancers et uaRNA/mRNA pour les promoteurs (**Figure 22** de l'introduction de la thèse, paragraphe III.5), FANTOM5 a assigné les enhancers actifs à leurs gènes cibles. En moyenne, un enhancer a été associé à 2,4 TSS et un TSS avec 4,9 enhancers. De plus, 64 % des enhancers identifiés par FANTOM5 ont été associés à au moins un TSS situé à plus de 500 kilobases de distance.

Dans le but d'identifier les gènes modulés par Fra-1 et/ou Fra-2 via leurs régions enhancers et/ou promoteurs, nos données transcriptomiques et ChIP-seq ont été croisées avec, d'une part, les données enhancers FANTOM5 et, d'autre part, les données TSS Refseq, en collaboration avec C. Lecellier (IGMM, Montpellier).

* Identification des gènes modulés par Fra-1 et/ou Fra-2 via leurs enhancers

Dans un premier temps, l'intersection entre nos données ChIP-seq pour les facteurs Fra-1 et Fra-2 ainsi que les marques histones définissant les enhancers actifs (H3K4me1>me3 et présence de H3K27ac), a révélé que 8600 enhancers identifiés comme actifs dans les MDA-MB231 possèdent des pics de fixation pour Fra-1 et/ou Fra-2 (**Figure 42.A**).

Dans un second temps, l'exploitation bioinformatique des données de FANTOM5 a fait apparaître que 1896 enhancers actifs liés par Fra-1 et/ou Fra-2 dans les MDA-MB231 sont

répertoriés parmi les enhancers FANTOM5. D'après les données de FANTOM5, ces 1896 enhancers sont associés à 2518 TSS (**Figure 42.B**).

Enfin, l'intersection entre ces 2518 TSS et nos données transcriptomiques ($FC \geq 1,5/\leq -1,5$ et $FDR < 0,05$) a permis d'identifier 179 gènes sous le contrôle de Fra-1 et/ou Fra-2 associés à des enhancers FANTOM5 présentant des pics de fixation pour Fra-1 et/ou Fra-2 dans la lignée MDA-MB231 (**Figure 42.C**). La probabilité de trouver par hasard 179 gènes modulés par Fra-1 et/ou Fra-2 via des enhancers fixant Fra-1 et/ou Fra-2 est très faible : la p value du test hypergéométrique est de $8.93E-32$. Ceci suggère que la fixation préférentielle de Fra-1 et Fra-2 dans les régions enhancers a un sens biologique dans la lignée MDA-MB231. Pour plus de détails, la liste des 179 gènes est disponible en annexe (**Annexe 3**).

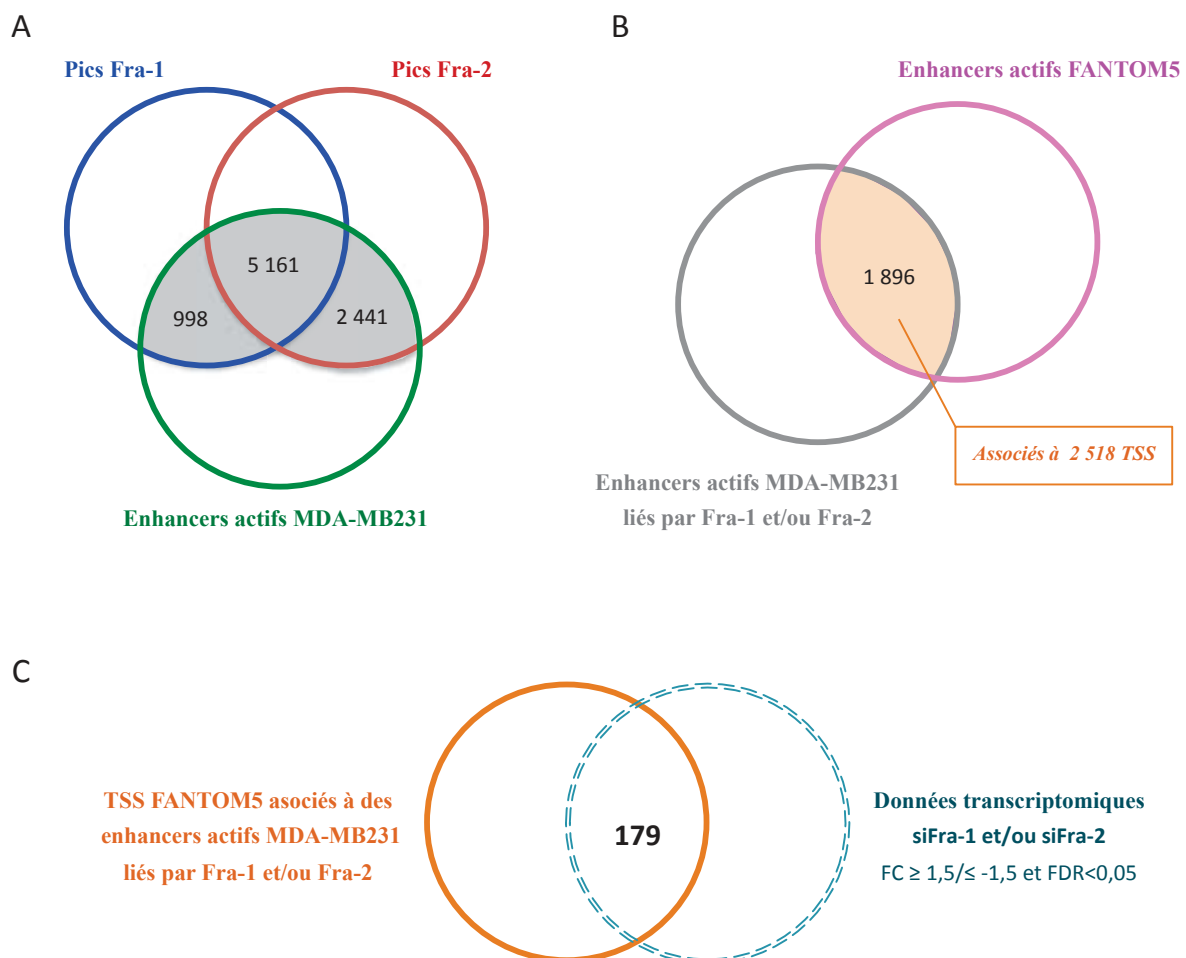


Figure 42 : Identification des gènes sous le contrôle de Fra-1 et/ou Fra-2 associés à des enhancers présentant des pics de fixation pour Fra-1 et/ou Fra-2 dans la lignée MDA-MB231.

** Identification des gènes modulés par Fra-1 et/ou Fra-2 via leurs promoteurs*

L'intersection entre nos données ChIP-seq pour les facteurs Fra-1 et Fra-2, nos données transcriptomiques ($FC \geq 1,5/\leq -1,5$ et $FDR < 0,05$) et les promoteurs ($\pm 1,5$ kb du TSS Refseq) a permis d'identifier 81 gènes sous le contrôle de Fra-1 et/ou Fra-2 et présentant des pics de

fixation pour Fra-1 et/ou Fra-2 dans leur région promotrice dans lignée MDA-MB231. Pour plus de détails, la liste des 81 gènes est disponible en annexe (**Annexe 4**).

III.3. Conclusion sur les approches ChIP-seq et FANTOM5

Les résultats de ChIP-seq Fra-1 et Fra-2 ainsi que l'analyse des marques d'histones ont mis en évidence que les facteurs de transcription Fra-1 et Fra-2 se fixent préférentiellement au niveau de régions chromatiniennes présentant des caractéristiques d'enhancers actifs dans la lignée MDA-MB231. Ces résultats sont en accord avec des données récentes de la littérature qui montrent, par des approches d'analyse de la séquence nucléotidique et de machine learning, que les motifs AP-1 sont enrichis au niveau des séquences enhancers dans différentes lignées (Malik et al., 2014; Nguyen et al., 2016; Yáñez-Cuna et al., 2014).

Une des difficultés dans le domaine de la transcription est l'annotation des enhancers aux gènes qu'ils régulent. Les données récentes de la littérature montrent que plusieurs enhancers peuvent réguler un même promoteur et qu'un même enhancer peut réguler plusieurs promoteurs (Andersson et al., 2014). De plus, tous les enhancers ne sont pas actifs dans tous les types cellulaires. Pour relier mes données de transcriptomes et de ChIP-seq, j'ai utilisé l'atlas du consortium FANTOM5 qui donne une liste d'enhancers robustes associés à leurs promoteurs. Sur cette base, je n'ai pu identifier que 179 gènes dont les enhancers FANTOM5 sont actifs dans les MDA-MB231 et fixent Fra-1 et/ou Fra-2. Ceci suggère, soit qu'un grand nombre des gènes régulés par Fra-1 et Fra-2 correspond à des cibles indirectes, soit que ces gènes sont régulés par des enhancers non annotés comme robustes par FANTOM5.

On observe également un fort recouvrement entre les sites de fixation de Fra-1 et de Fra-2 dans la lignée MDA-MB231. Si certains de ces sites contiennent deux motifs AP-1, susceptibles de permettre la fixation de Fra-1 et Fra-2 en même temps, majoritairement les sites de fixation ne contiennent qu'un motif AP-1. Dans ce dernier cas, on ne peut pas exclure un mécanisme dynamique faisant appel à la fixation, dans un premier temps, de l'un des facteurs, puis dans un deuxième temps, de l'autre. Cependant, il est plus vraisemblable que le recouvrement des pics Fra-1 et Fra-2 reflète la possibilité de fixer indifféremment les deux facteurs sur ces sites. Cette hypothèse va dans le sens de nos données transcriptomiques où une majorité de gènes est régulée de façon redondante par Fra-1 et Fra-2. Il est à souligner que la redondance fonctionnelle de Fra-1 et Fra-2 n'a pas encore été rapportée dans la littérature.

Le ChIP-seq révèle également un nombre de pics de fixation de Fra-2 (13003) supérieur à celui de Fra-1 (9998), bien que la protéine Fra-1 soit très vraisemblablement plus abondante que Fra-2 dans les MDA-MB231 sur la base de l'abondance relative de leur ARNm (**Figure 24.A**). Cette observation pourrait correspondre à la réalité ou être liée à un biais d'efficacité des anticorps utilisés pour les ChIP-seq. Nos données expérimentales de ChIP-qPCR nous amènent à

penser que le nombre moindre de pics Fra-1 en ChIP-seq est une conséquence de la moins grande efficacité des anticorps Fra-1 et ne reflète pas une réalité biologique. En accord avec cette hypothèse, les données transcriptomiques (**Tableau 7**) montrent que le nombre de gènes régulés préférentiellement par Fra-1 est beaucoup plus important que celui des gènes régulés préférentiellement par Fra-2, ce qui souligne l'importance de Fra-1. Néanmoins, nos résultats montrent que les gènes régulés préférentiellement par Fra-2 semblent strictement régulés par Fra-2. En effet, la variation dans leur niveau d'expression est comparable dans les conditions siFra-2 et siFra-1+siFra-2 (**Annexes 1 et 2**). En revanche, une proportion non négligeable des gènes considérés comme préférentiellement régulés par Fra-1 (pas de changement de leur abondance en absence de Fra-2 seul) voit son niveau d'expression varier plus fortement en absence de Fra-1+Fra-2 par rapport à l'absence de Fra-1 seul. Ceci suggère que Fra-2 contribue aussi à l'expression de ces gènes mais que son rôle est masqué par l'abondance de Fra-1.

Par ailleurs, les résultats de l'analyse transcriptomique montrent une répartition à peu près équivalente entre gènes induits et gènes réprimés par Fra-1 et/ou Fra-2, que l'on considère les gènes régulés préférentiellement par Fra-1 ou par Fra-2 ou les gènes régulés de façon redondante ou complémentaire par les deux protéines. Ceci indique que Fra-1 et Fra-2 peuvent se comporter à la fois comme des activateurs ou des répresseurs de la transcription suivant les gènes, ce qui suggère des mécanismes transcriptionnels différents pour les différentes catégories de gènes.

Les données de ChIP-seq font apparaître qu'une fraction seulement de Fra-1 et/ou Fra-2 est associée à c-Jun au niveau des sites de liaison sur la chromatine. Il serait donc intéressant de cartographier les domaines de liaison de JunB et JunD et de déterminer si des corrélations éventuelles avec les différents partenaires JUN peuvent rendre compte des effets activateurs ou répresseurs observés sur la transcription des gènes cibles de Fra-1 et/ou Fra-2.

Un autre mécanisme qui pourrait aussi être à l'origine des activités activatrices ou répressives de Fra-1 et Fra-2 est l'interaction avec différents facteurs de transcription recrutés à proximité. En effet, de façon très récente, il a été montré une coopération entre AP-1 et TEAD4/YAP/TAZ pour la transduction de la voie Hippo, qui régule la croissance des organes et la tumorigenèse, dans des cellules de cancer du sein. Le mécanisme implique la formation d'un complexe qui active de façon synergique les gènes cibles via des enhancers distaux et la formation de boucles chromatinienne entre les enhancers et les promoteurs (Zanconato et al., 2015). Il serait donc intéressant de rechercher par analyse computationnelle (HOMER) la co-occurrence de motifs de liaison pour des facteurs de transcription dans les séquences régulatrices liées par Fra-1 et/ou Fra-2 afin d'évaluer la possibilité de co-régulations spécifiques qui pourraient être synergiques ou antagonistes et différentes suivant que les gènes sont induits ou réprimés.

Enfin, des modifications post-traductionnelles spécifiques de Fra-1 et/ou Fra-2 sur les enhancers pourraient également rendre compte des effets sur la transcription en modifiant leur

surface d'interaction et donc le recrutement de partenaires spécifiques. Le laboratoire dispose de quatre anticorps dirigés contre quatre formes phosphorylées de Fra-1 (sur T217, T230, S252 et S265). Des expériences préliminaires en ChIP-qPCR ont montré que deux d'entre eux sont efficaces pour l'immunoprécipitation de chromatine. Des expériences de ChIP-seq pourraient permettre d'établir si ces modifications sont présentes de manière récurrente dans tous les pics Fra-1 ou si elles ne concernent qu'une fraction et, dans ce cas, elles pourraient être susceptibles de modifier de façon enhancer-spécifique l'activité du facteur de transcription.

IV. SÉLECTION DES GÈNES CANDIDATS POUR POURSUIVRE LES ÉTUDES FONCTIONNELLES ET MÉCANISTIQUES

Le but de mon projet est double : il s'agit d'une part d'identifier des cibles directes de Fra-1 et/ou Fra-2 qui pourraient participer à l'agressivité tumorale et constituer de nouvelles cibles potentielles à visée diagnostique ou thérapeutique, et d'autre part d'étudier les mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour le contrôle de la transcription de leurs gènes.

Pour cela, comme présenté précédemment j'ai suivi en parallèle deux axes complémentaires pour mon étude: (i) l'identification des gènes cibles de Fra-1 et/ou Fra-2 dans la lignée MDA-MB231 par des approches transcriptomiques et (ii) la cartographie et la caractérisation des sites de liaison de Fra-1 et/ou Fra-2 sur le génome des MDA-MB231 par des approches ChIP-seq.

Pour les études fonctionnelles et mécanistiques ultérieures, il était important ensuite de sélectionner des gènes dont l'expression est affectée dans les TNBC et non des gènes dont l'expression est spécifique de la lignée cellulaire utilisée lors des approches transcriptomiques et ChIP-seq. Pour cela j'ai analysé la corrélation de leur expression avec celle de Fra-1 et Fra-2 dans deux banques de données publiques et j'ai validé l'expression d'un certain nombre de ces gènes dans les deux autres lignées TNBC, présentées au paragraphe I.

IV.1. Croisement des données

IV.1.a. Croisement des données transcriptomiques et ChIP-seq

Le but de mon projet étant d'identifier de nouvelles cibles thérapeutiques potentielles, les gènes induits par Fra-1 et/ou Fra-2 ont été considérés en priorité. Dans un premier temps, j'ai concentré mon choix de gènes cibles à étudier sur des gènes dont les variations d'expression sont les plus importantes. Ainsi, 109 gènes sont modulés avec un fold-change compris entre -2 et -11 après suppression de Fra-1 et/ou Fra-2 par RNAi.

Parmi ceux-ci, j'ai sélectionné ceux qui sont susceptibles d'être des cibles directes de Fra-1 et/ou Fra-2, car présentant des signaux de fixation pour Fra-1 et/ou Fra-2 dans leur région promotrice ou à proximité (± 3 kb du TSS) et/ou dans leur(s) enhancer(s) FANTOM5 associé(s).

Puis, j'ai restreint la liste aux gènes ayant un rôle déjà décrit, suspecté ou possible dans la tumorigenèse, soit une première sélection de 14 gènes (**Tableau 9**). Ces derniers se répartissent dans plusieurs catégories ontologiques qui incluent l'adhésion, la motilité, la migration, l'invasion cellulaire ou encore le contrôle de la mort cellulaire. La dérégulation de chacune de ces fonctions pouvant participer au phénotype et à l'agressivité tumorale.

		Fold-change					Pic(s) Fra-1 et/ou Fra-2	
		T1.siCTL vs T1.NT	T1.siFra-1 vs T1.NT	T1.siFra-2 vs T1.NT	T2.siCTL vs T2.NT	T2.si-1+siF2 vs T2.NT	± 3kb du TSS	Enhancer F5
14 gènes induits par Fra-1 et/ou Fra-2 (FC<-2)	CD68	NA	-2,3	NA	-1,1	-5,5	✓	✗
	CD22	1,4	-2,3	NA	1,9	-5,3	✓	✓
	TMEM158	-1,2	-2,5	NA	-1,7	-4,9	✓	✓
	CDCP1	NA	-1,7	NA	NA	-3,0	✓	✓
	SSFA2	-1,1	-1,8	NA	NA	-2,9	✗	✓
	ARNTL2	NA	-1,8	NA	-1,1	-2,8	✓	✗
	PRKCDBP	1,2	-1,6	NA	NA	-2,3	✓	✗
	ABLIM3	NA	-1,6	NA	-1,1	-2,3	✓	✓
	EMP1	NA	NA	NA	NA	-2,2	✓	✓
	HMGA1	-1,1	-1,6	NA	-1,2	-2,2	✓	✗
	MYO10	1,2	-1,4	NA	NA	-2,2	✗	✓
	TNFAIP3	1,1	-1,6	NA	-1,2	-2,2	✓	✗
	RAP2B	NA	-1,3	NA	-1,2	-2,1	✓	✗
	MARCH4	NA	NA	NA	-1,3	-2,0	✓	✗

Tableau 9 : Première sélection de 14 gènes induits par Fra-1 et/ou Fra-2 dans les MDA-MB231.

IV.1.b. Croisement avec deux banques de données de cancers du sein

Dans un second temps, j'ai eu recours à des banques de données publiques pour déterminer si les gènes identifiés comme étant sous contrôle de Fra-1 et/ou Fra-2 dans les MDA-MB231 montrent également une expression corrélée avec celles de Fra-1 et/ou Fra-2 dans des tumeurs humaines primaires. Pour cela, j'ai utilisé les deux banques de données de cancers du sein les plus citées dans la littérature. Il s'agit des données du TCGA (The Cancer Genome Atlas) en collaboration avec T. Gostan (IGMM, Montpellier) et des données du CIT (programme Carte d'Identité des Tumeurs de la Ligue contre le Cancer) en collaboration avec C. Theillet (IRCM, Montpellier).

Ces deux banques comportent quelques différences. Tout d'abord, il s'agit de deux cohortes différentes de tumeurs du sein et la technologie utilisée n'est pas la même. Les données du CIT ont été analysées par technologie Affymetrix et celles du TCGA par technologie Agilent. Par ailleurs, les classifications utilisées ne sont pas exactement les mêmes (**Figure 43**). La principale différence se trouve au niveau du sous-type dit normal breast-like. En effet, le CIT classe ce sous-type comme ER+ (peu agressif) alors que le TCGA le classe comme ER- (très agressif) (**Figure 43**). Comme mentionné dans l'introduction de la thèse (paragraphe I.4), le phénotype des tumeurs normal breast-like est mal caractérisé et leur pronostic est généralement intermédiaire entre les cancers luminaux et basal-like. Cette classification pourrait résulter d'un artefact lié à la contamination des cellules tumorales par le tissu normal environnant.

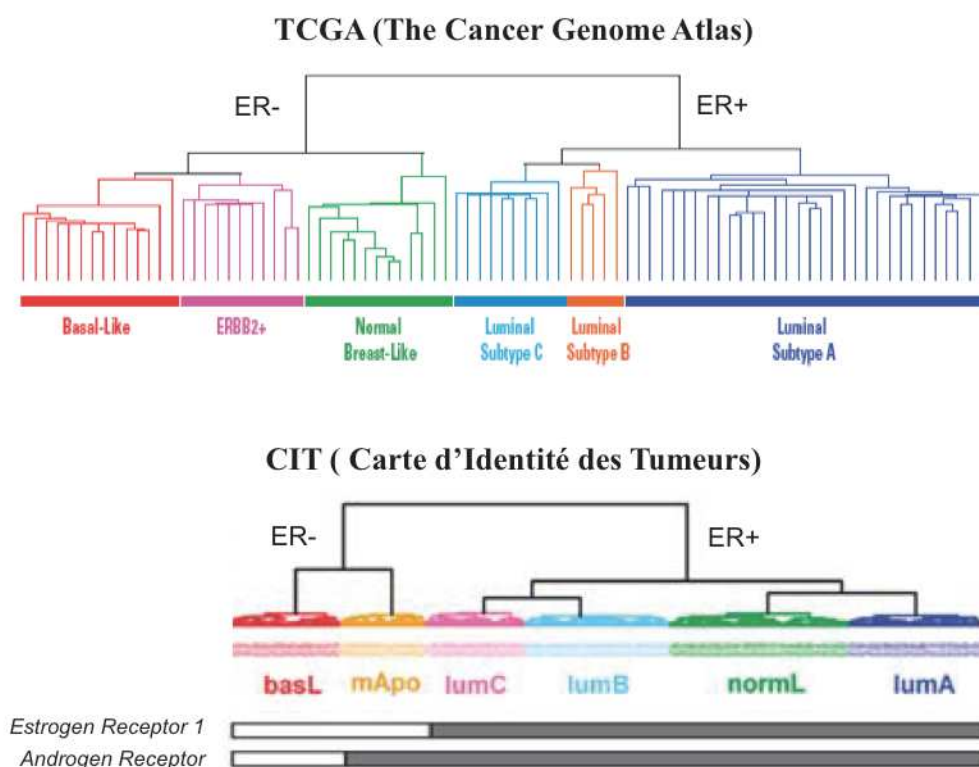


Figure 43 : Classifications des cancers du sein utilisées par les banques de données. Adaptées d'après Sorlie et al. (2001) et Guedji et al. (2012).

Parmi les 14 gènes sélectionnés au paragraphe précédent, seuls 8 montrent une expression corrélée avec celle de Fra-1 et/ou Fra-2 dans les deux banques de tumeurs primaires (**Tableau 10**). Pour illustrer les résultats, je présente (**Figure 44**) les données obtenues pour HMGA1 et CD68 qui font l'objet des études II et III respectivement.

		Fold-change					Pic(s) Fra-1 et/ou Fra-2		Corrélation avec les banques de cancers du sein	
		T1.siCTL vs T1.NT	T1.siFra-1 vs T1.NT	T1.siFra-2 vs T1.NT	T2.siCTL vs T2.NT	T2.si-1+siF2 vs T2.NT	± 3kb du TSS	Enhancer F5	CIT	TCGA
14 gènes induits par Fra-1 et/ou Fra-2 (FC<2)	CD68	NA	-2,3	NA	-1,1	-5,5	✓	✗	✓	✓
	CD22	1,4	-2,3	NA	1,9	-5,3	✓	✓	✗	✗
	TMEM158	-1,2	-2,5	NA	-1,7	-4,9	✓	✓	✓	✓
	CDCP1	NA	-1,7	NA	NA	-3,0	✓	✓	✗	✗
	SSFA2	-1,1	-1,8	NA	NA	-2,9	✗	✓	✗	✗
	ARNTL2	NA	-1,8	NA	-1,1	-2,8	✓	✗	✓	✓
	PRKCDPB	1,2	-1,6	NA	NA	-2,3	✓	✗	✓	✓
	ABLIM3	NA	-1,6	NA	-1,1	-2,3	✓	✓	✗	✗
	EMP1	NA	NA	NA	NA	-2,2	✓	✓	✗	✗
	HMGA1	-1,1	-1,6	NA	-1,2	-2,2	✓	✗	✓	✓
	MYO10	1,2	-1,4	NA	NA	-2,2	✗	✓	✓	✓
	TNFAIP3	1,1	-1,6	NA	-1,2	-2,2	✓	✗	✓	✓
	RAP2B	NA	-1,3	NA	-1,2	-2,1	✓	✗	✓	✓
	MARCH4	NA	NA	NA	-1,3	-2,0	✓	✗	✓	✗

Tableau 10 : L'expression de 8/14 gènes induits par Fra-1 et/ou Fra-2 dans la lignée modèle corrèle avec l'expression de Fra-1 et/ou Fra-2 dans les deux banques de cancers du sein.

HMGA1

L'analyse transcriptomique, par technologie Affymetrix conduite dans les MDA-MB231, a identifié HMGA1 comme un gène up-régulé sous le contrôle préférentiel de Fra-1. En effet, l'expression de HMGA1 varie uniquement quand Fra-1 et Fra-1+Fra-2 sont réprimés par RNAi. L'absence de Fra-2 n'est pas suffisante pour moduler HMGA1 mais pourrait avoir un effet synergique lorsqu'elle est combinée à l'absence de Fra-1, car la modulation est plus importante dans la condition siFra-1+siFra-2 (facteur -2,2) que dans la condition siFra-1 (facteur -1,6) (**Tableaux 9 ou 10**).

L'analyse des expressions de HMGA1, Fra-1 et Fra-2 dans les deux cohortes de tumeurs du sein (**Figure 44.A**), révèle que les expressions de HMGA1 et de Fra-1 corrélaient fortement avec l'agressivité des sous-types de cancer du sein (classification détaillée dans l'introduction de la thèse aux paragraphes I.4 et I.5). La corrélation entre l'expression de Fra-2 et le sous-type tumoral n'apparaît pas de façon claire et ceci dans les deux banques de données analysées.

Par ailleurs, comme on peut le voir sur la **Figure 44.B** HMGA1 montre une expression très fortement corrélée avec celle de Fra-1 dans les tumeurs humaines : avec pour le CIT ($R^2=0,419$; $pV=0,00E0^{***}$) et pour le TCGA ($R^2=0,508$ et $pV=0,00E0^{***}$). L'expression de HMGA1 est plus faiblement corrélée avec celle de Fra-2 dans les tumeurs humaines : il n'y a pas de corrélation pour le CIT ($R^2=0,062$; $pV=0,199$), en revanche pour le TCGA bien que la corrélation soit faible elle est néanmoins significative ($R^2=0,111$; $pV=1,04E-2^*$).

De plus, l'analyse par RT-qPCR de l'abondance des ARNm de HMGA1, Fra-1 et Fra-2 dans les trois lignées TNBC mentionnées précédemment, ainsi que dans la lignée luminal A MCF7 (**Figure 45**) révèle que HMGA1 et Fra-1 ont le même profil d'expression, à savoir fortement exprimés dans les trois lignées ER- et absents dans la lignée ER+. En ce qui concerne Fra-2, l'abondance de l'ARNm est comparable entre les lignées ER- et la lignée ER+.

Le gène HMGA1 (High Mobility Group AT-Hook 1) code pour une protéine liant l'ADN dont la fonction a été largement étudiée. HMGA1 est considéré comme un « master switch » de la tumorigenèse, comme détaillé dans l'introduction de l'étude II. En revanche très peu de données sont connues sur son mode de régulation par AP-1 et plus particulièrement par Fra-1 et Fra-2.

CD68

Comme HMGA1, CD68 est un gène up-régulé sous le contrôle préférentiel de Fra-1. Cependant, un rôle partiellement redondant de Fra-2 ne peut être exclu puisque si l'absence de Fra-2 ne module pas l'expression de CD68, la double extinction Fra-1+Fra-2 a un effet plus fort (facteur -5,5) que la simple extinction Fra-1 (facteur -2,3) (**Tableaux 9 ou 10**).

L'analyse de l'expression de CD68, ne révèle pas de corrélation claire avec le sous-type tumoral (**Figure 44.A**). Il est à noter que ceci pourrait être lié au fait que les macrophages infiltrant (exprimant CD68) peuvent représenter 5 à 40% de la masse d'une tumeur mammaire (Sousa et al., 2015).

Par ailleurs, comme on peut le voir sur la **Figure 44.B** CD68 montre une expression très fortement corrélée avec celle de Fra-1 dans les tumeurs humaines du CIT ($R^2=0,396$; $pV=0,00E0^{***}$) et un peu moins corrélée pour le TCGA ($R^2=0,128$ et $pV=3,1E-3^{**}$). L'expression de CD68 est plus faiblement corrélée avec celle de Fra-2 dans les tumeurs humaines : il n'y a pas de corrélation pour le CIT ($R^2=0,022$; $pV=0,643$), en revanche pour le TCGA bien que la corrélation soit faible elle est néanmoins significative ($R^2=0,253$; $pV=3,31E-3^{**}$).

L'analyse par RT-qPCR de l'abondance des ARNm de CD68, Fra-1 et Fra-2 dans les trois lignées TNBC mentionnées précédemment ainsi que la lignée MCF7 (ER+) (**Figure 45**) révèle que CD68 et Fra-1 ont le même profil d'expression, à savoir fortement exprimés dans les trois lignées ER- et absents dans la lignée ER+.

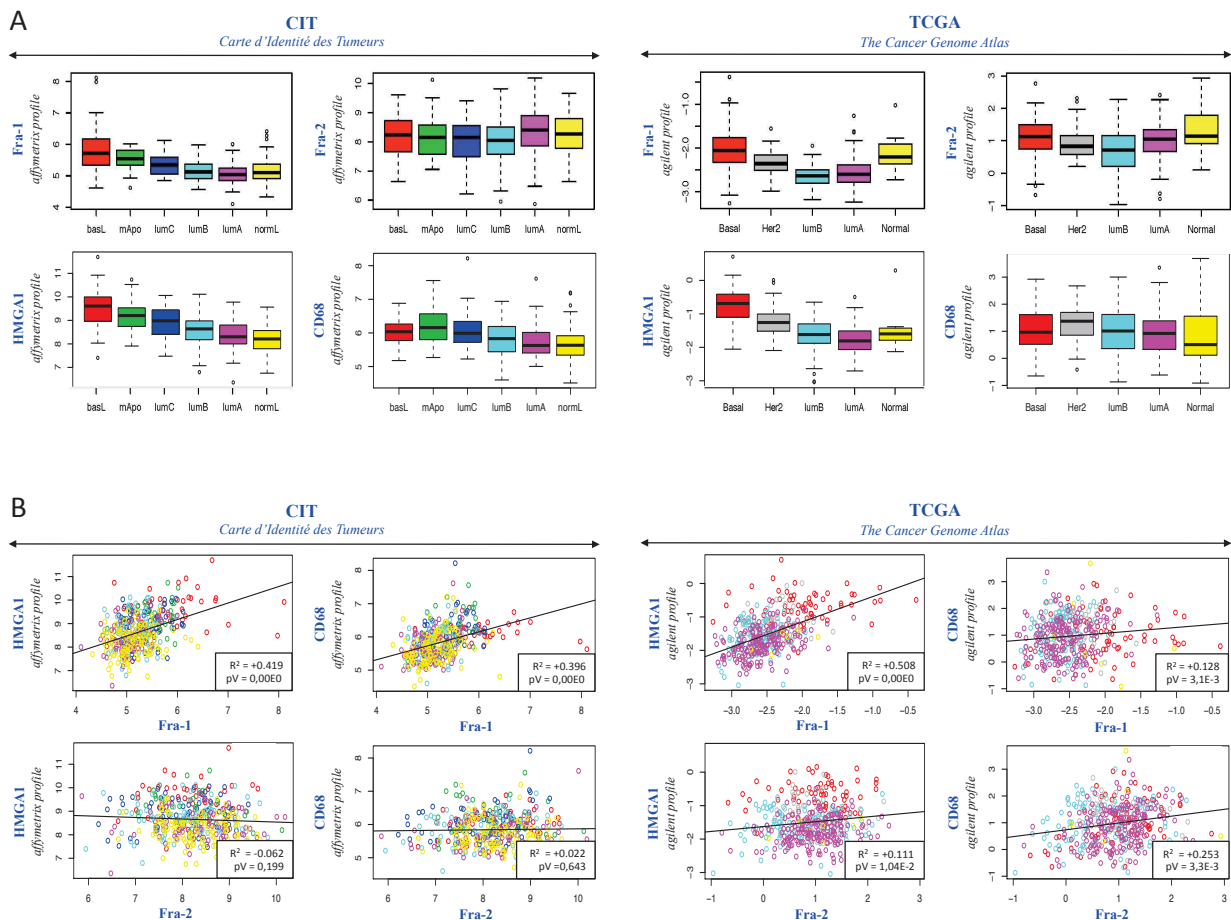


Figure 44 : Données des banques de cancers du sein CIT et TCGA concernant Fra-1, Fra-2, HMGA1 et CD68. A) Expression de Fra1, Fra-2, HMGA1 et CD68 dans les différents sous-types de cancer du sein. **B)** Corrélation entre l'expression de HMGA1/Fra-1, HMGA1/Fra-2, CD68/Fra-1 et CD68/Fra-2. Une version haute définition de cette figure est disponible en annexe (**Annexe 5**).

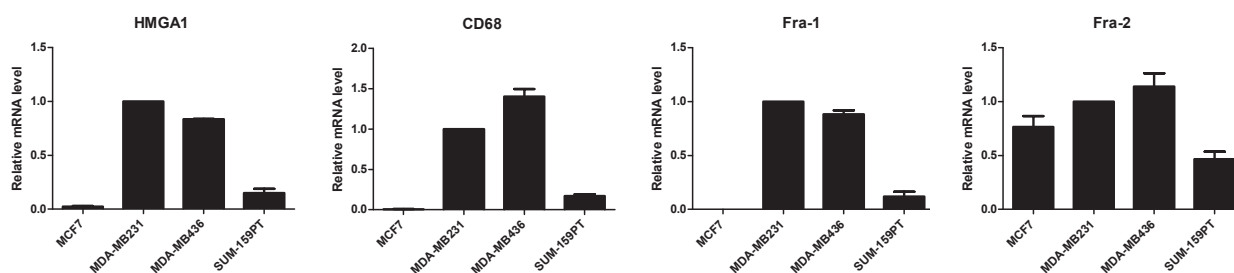


Figure 45 : Analyse de l'expression des ARNm de HMGA1, CD68, Fra-1 et Fra-2 dans différentes lignées de cancers du sein. Mesure de l'abondance d'ARNm par RT-qPCR. n=3. Les valeurs de RT-qPCR sont normalisées par rapport à l'expression de l'ARNm dans les cellules MDA-MB231. Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.

IV.2. Validation par RT-qPCR et extension des observations à deux autres lignées TNBC

Dans un troisième temps, j'ai validé par RT-qPCR que les 8 gènes sélectionnés étaient bien sous le contrôle de Fra-1 et/ou Fra-2 dans la lignée MDA-MB231, mais également dans les MDA-MB436 et les SUM-159PT. Deux couples de primers par gènes ont été utilisés pour augmenter la robustesse des conclusions.

Les résultats de l'analyse par RT-qPCR de la modulation de leur expression après inhibition de Fra-1 et/ou Fra-2 par RNAi dans les trois lignées TNBC sont présentés uniquement pour HMGA1 et CD68 (**Figure 46**). En adéquation avec les données transcriptomiques, les expressions de HMGA1 et CD68 sont sous le contrôle préférentiel de Fra-1 dans la lignée MDA-MB231 (**Figure 46, panel de gauche**).

L'extension des observations aux deux autres lignées TNBC fait apparaître que HMGA1 est également sous le contrôle préférentiel de Fra-1 dans les MDA-MB436 et les SUM-159PT. En revanche, on note une implication de Fra-2 plus ou moins importante dans l'expression de CD68 dans les MDA-MB436 et les SUM-159PT, ce qui pourrait indiquer un rôle complémentaire ou synergique de Fra-1 et Fra-2 pour l'expression de CD68 dans ces deux lignées (**Figure 46**).

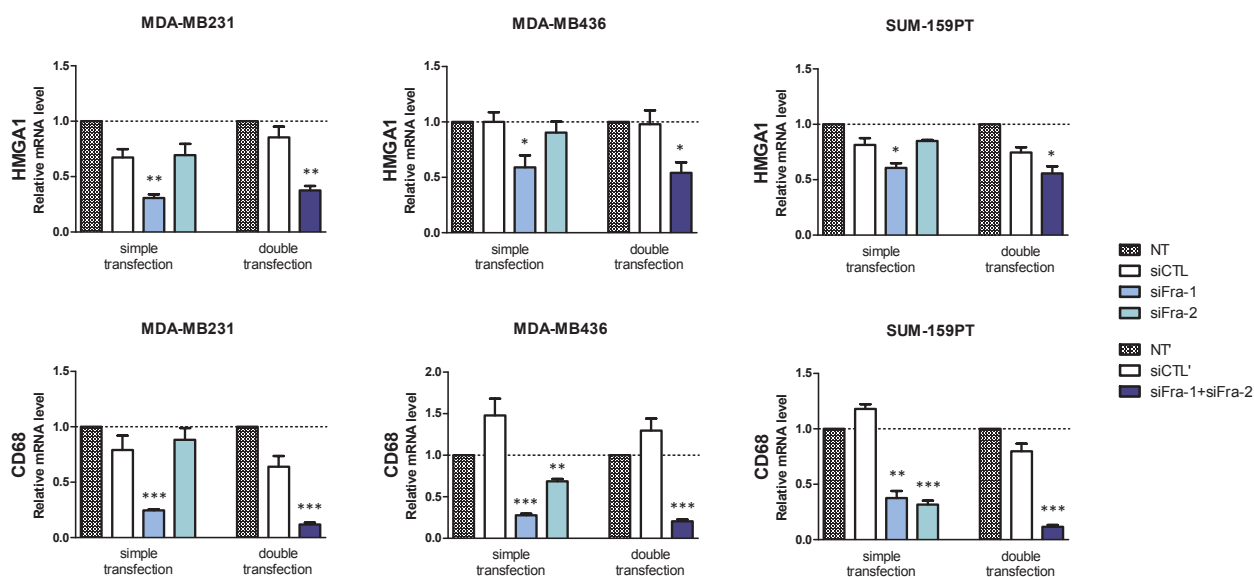


Figure 46 : Modulation du gène HMGA1 et CD68 en absence de Fra-1 et/ou Fra-2 dans les trois lignées TNBC. Les valeurs de RT-qPCR sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm mesurée pour la condition non-transfectée (NT ou NT'). n=3. Two tailed paired t-test ; (*) p value <0.05 (**) p value <0.01 ; (***) p value <0.001. *Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.*

Les résultats des analyses pour les 8 gènes sont présentés dans le **Tableau 11**. Il apparaît que seuls 4 gènes sur les 8 remplissent les critères que nous nous sommes fixés. Il s'agit de HMGA1, CD68, ARNTL2 et PRKCDBP. En effet, ces gènes sont pertinents à étudier du point de vue du cancer mammaire triple négatif, car leur expression est corrélée à celle des facteurs Fra dans les deux banques de données de tumeurs primaires et leur expression a été validée comme étant sous le contrôle de Fra-1 et/ou Fra-2 dans les trois lignées triples négatives de cancer du sein. De plus, ils possèdent des pics de fixation pour Fra-1 et/ou Fra-2 dans leurs régions promotrices. En revanche, ces 4 gènes ne possèdent pas de pics Fra-1 ou Fra-2 dans leurs régions enhancers FANTOM5 associées.

	Fold-change					Pic(s) Fra-1 et/ou Fra-2		Corrélation avec les banques de données		Validation par RTqPCR		
	T1.siCTL vs T1.NT	T1.siFra-1 vs T1.NT	T1.siFra-2 vs T1.NT	T2.siCTL vs T2.NT	T2.si-1+siF2 vs T2.NT	± 3kb du TSS	Enhancer F5	CIT	TCGA	MDA-MB231	MDA-MB436	SUM-159PT
CD68	NA	-2,3	NA	-1,1	-5,5	✓	✗	✓	✓	✓	✓	✓
TMEM158	-1,2	-2,5	NA	-1,7	-4,9	✓	✓	✓	✓	✓	✗	✗
ARNTL2	NA	-1,8	NA	-1,1	-2,8	✓	✗	✓	✓	✓	✓	✓
PRKCDBP	1,2	-1,6	NA	NA	-2,3	✓	✗	✓	✓	✓	✓	✓
HMGA1	-1,1	-1,6	NA	-1,2	-2,2	✓	✗	✓	✓	✓	✓	✓
MYO10	1,2	-1,4	NA	NA	-2,2	✗	✓	✓	✓	✓	✗	✗
TNFAIP3	1,1	-1,6	NA	-1,2	-2,2	✓	✗	✓	✓	✗	✗	✓
RAP2B	NA	-1,3	NA	-1,2	-2,1	✓	✗	✓	✓	✓	✗	✗

Tableau 11 : Bilan des résultats pour les 8 gènes candidats sélectionnés.

L'étude du gène ARNTL2 (ou Bmal2) s'est révélée problématique en raison de la difficulté à obtenir des outils efficaces (anticorps spécifiques utilisables en WB, siRNA spécifiques, ect.). De plus, en 2016 un article montrant l'implication de ARNTL2 dans la promotion du processus métastatique dans un modèle de cancer du sein triple négatif *in vivo*, a été publié (Ha et al., 2016). L'implantation orthotopique *in vivo* des cellules 4T1 après transfection de shRNA anti-ARNTL2 entraîne une réduction significative du nombre de métastases pulmonaires. De la même manière, la surexpression d'ARNTL2 dans ces cellules augmente de façon significative les métastases pulmonaires. Je n'ai donc pas poursuivi mes investigations fonctionnelles sur ARNTL2, néanmoins ceci souligne le fait que les gènes sélectionnés par mon crible sont pertinents à étudier.

Concernant le gène PRKCDBP (ou hSRBC), les données de la littérature sont confuses. En effet, PRKCDBP est considéré comme un probable suppresseur de tumeur car son expression est décrite pour être diminuée dans les tumeurs primaires du poumon, du colon et du sein (Lee et al., 2008, 2011; Wikman et al., 2012; Zöchbauer-Müller et al., 2005). Or les données de corrélation avec les banques de données de cancers du sein associent son augmentation au phénotype agressif. De la même manière, la corrélation entre l'expression de Fra-1 et PRKCDPD suggère un rôle oncogénique. De plus, le facteur de régulation par Fra est plus faible que celui de CD68. PRKCDBP n'a donc pas été considéré en priorité pour les études fonctionnelles ultérieures.

Dans le cadre de ma thèse mes efforts se sont donc essentiellement focalisés sur les gènes HMGA1 et CD68. HMGA1 étant déjà décrit pour son rôle de master switch dans la tumorigenèse, l'étude II est focalisée sur les aspects plus mécanistiques de sa régulation par les facteurs Fra. Le rôle de CD68 quant à lui n'est pas connu, ainsi dans l'étude III j'ai cherché à élucider quel était son rôle dans le processus tumoral.

RÉSULTATS

Étude II

Étude II :

Mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour l'expression du gène HMGA1

Comme présenté à la fin de l'étude I, j'ai choisi de m'intéresser à la caractérisation des mécanismes transcriptionnels mis en jeu par Fra-1 et Fra-2 pour gouverner l'expression de HMGA1 pour trois raisons principales.

D'une part, les résultats acquis au cours de l'étude I ont mis en évidence (i) que l'expression de HMGA1 est régulée de façon positive par Fra-1 dans trois lignées cellulaires issues de cancers du sein triple négatifs, (ii) qu'il existe une forte corrélation entre les expressions de Fra-1 et de HMGA1 avec l'agressivité des tumeurs mammaires et (iii) qu'il existe aussi une corrélation entre les niveaux d'expression de Fra-1 et de HMGA1 dans les cancers du sein. L'ensemble de ces observations suggérerait donc un rôle de Fra-1 dans la régulation de l'expression de HMGA1 dans les tumeurs humaines.

D'autre part, de nombreuses études concourent à montrer que HMGA1 a un rôle clé dans les processus tumorigéniques et métastatiques en affectant, entre autres, la prolifération, la migration et l'invasion cellulaire. Le degré de surexpression de HMGA1 est corrélé à la malignité tumorale et au potentiel métastatique.

Enfin, seulement quelques rares études de la littérature se sont intéressées aux mécanismes de la régulation transcriptionnelle de l'expression de HMGA1. L'implication de AP-1 a été rapportée dans d'autres contextes cellulaires sans que les mécanismes moléculaires sous-jacents ne soient abordés.

Pour introduire cette partie de mon travail, je présenterai une introduction relativement brève sur HMGA1 pour donner une vue d'ensemble de sa structure, de ses fonctions physiopathologiques et des mécanismes de régulation de son expression déjà décrits. Dans la deuxième partie je décrirai les travaux effectués au cours de ma thèse pour essayer de caractériser les mécanismes moléculaires mis en jeu par Fra-1 et/ou Fra-2 pour la régulation de son expression. Cette étude n'est pas encore complètement finalisée et les aspects à développer seront abordés dans la conclusion.

I. HMGA1 : Introduction

I.1. Structure et fonctions moléculaires de HMGA1

I.1.a. Les protéines HMG

HMGA1 appartient à la famille des HMG pour High Mobility Group. C'est une famille hétérogène de protéines non histones associées à la chromatine. Ce sont les protéines chromatinienne les plus abondantes après les histones.

La famille HMG comprend trois groupes : HMGA, HMGB et HMGN. Chaque groupe est caractérisé par un domaine de liaison à l'ADN spécifique (**Figure 47**) : le domaine « AT-hook » pour les protéines HMGA, le domaine « HMG-box » pour les protéines HMGB et le domaine « Nucleosomal Binding Domain » pour les protéines HMGN (Hock et al., 2007).

Ces protéines nucléaires se lient à l'ADN et aux nucléosomes. Elles induisent des changements structuraux qui peuvent être sensiblement différents selon le type de HMG. Parmi les mécanismes on peut noter leur capacité à induire des changements dans la courbure de l'ADN et leur capacité à prévenir ou à faciliter le recrutement de complexes régulateurs ou impliqués dans le remodelage de la chromatine (Reeves and Beckerbauer, 2001).

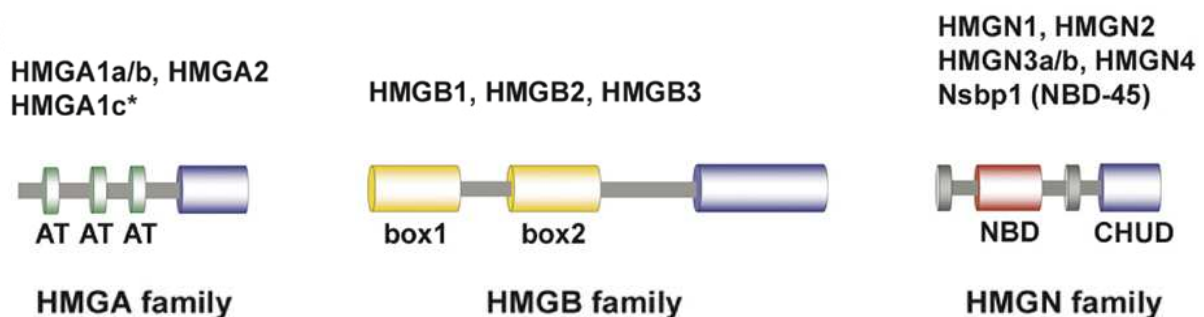


Figure 47 : Structure des protéines HMG. Le groupe HMGA contient quatre membres caractérisés par une extrémité C-terminale acide (bleu) et trois domaines AT-hook (verts), sauf HMGA1c, qui contient seulement deux domaines AT-hook. Chacun des trois membres de la famille HMGB contient deux boîtes HMG (jaunes), ainsi qu'une longue extrémité C-terminale acide (bleu). Les protéines HMGN sont caractérisées par un domaine chargé positivement (rouge) appelé NBD, pour Nucleosomal Binding Domain, et une région C-terminale chargée négativement (bleu) appelée CHUD pour Chromatin Unfolding Domain. *Source : Hock et al (2007).*

Dans la suite de ce chapitre, nous nous intéresserons plus en détail uniquement au groupe des HMGA.

I.1.b. Caractéristiques structurales des HMGA

Chez l'Homme, les gènes codant pour les protéines HMGA sont au nombre de deux : HMGA1 et HMGA2.

Le gène codant pour HMGA1 (High Mobility Group AT-hook 1, anciennement nommé HMGI-Y) se situe sur le locus chromosomique 6p21 et comprend huit exons répartis sur une distance d'environ 10kb. L'épissage alternatif des transcrits donne naissance à trois isoformes protéiques : HMGA1a (107 acides aminés, 11,7 kDa) et HMGA1b (96 acides aminés, 10,6 kDa) et une isoforme mineure HMGA1c (179 acides aminés, 19,7 kDa) très faiblement exprimée (**Figure 48.A**, *isoforme mineure non présentée*) (Cleynen and Van de Ven, 2008; Fusco and Fedele, 2007). Les protéines HMGA1a et HMGA1b diffèrent par une délétion interne de 11 acides aminés pour HMGA1b. Elles contiennent chacune trois domaines basiques d'interaction avec l'ADN appelés domaines AT-Hook, car reconnaissant des séquences riches en adénine et thymidine, et un domaine C-terminal acide (Fusco and Fedele, 2007).

Le gène HMGA2 (High Mobility Group AT-hook 2, anciennement nommé HMGI-C) se situe sur le locus 12q14-15 et comprend cinq exons répartis sur une très longue région chromosomique de plus de 160kb. L'épissage de HMGA2 donne naissance à la protéine HMGA2 de 109 acides aminés (**Figure 48.B**). Elle se distingue des protéines HMGA1 par ses 25 premiers acides aminés, ainsi que par la présence d'un court peptide de 12 acides aminés entre le troisième motif AT-Hook et le domaine C-terminal acide (Fusco and Fedele, 2007).

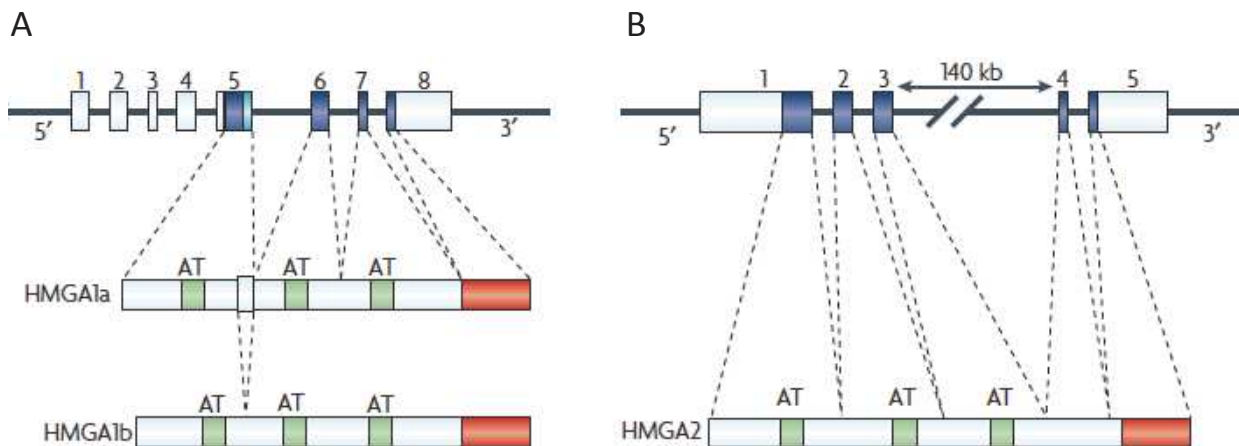


Figure 48 : Représentation schématique des gènes HMGA1 et HMGA2 humains et leurs produits protéiques. A) Le gène HMGA1 humain se compose de 8 exons dont les exons 5 à 8 (boîtes bleues) codent pour les protéines HMGA1. Les protéines HMGA1a et HMGA1b ont une structure semblable et contiennent trois domaines AT-Hook et une séquence C-terminale acide. La protéine HMGA1a possède 11 acides aminés supplémentaires entre les deux premiers domaines AT-Hook comparé à la protéine HMGA1b. B) Le gène HMGA2 comprend 5 exons (boîtes bleues). Les exons 1, 2 et 3 codent pour des domaines AT-Hook (boîte verte), alors que l'exon 5 code pour la séquence acide C-terminale (boîte rouge). L'exon 4 code pour un court peptide de 12 acides aminés qui n'a pas d'homologie avec les protéines HMGA1. Source : Fusco and Fedele (2007).

Les motifs AT-Hook sont des motifs chargés positivement et formés par une séquence de 9 acides aminés contenant une répétition invariante Arg-Gly-Arg-Pro (R-G-R-P) flanquée par d'autres résidus chargés positivement, Arg ou Lys (**Figure 49**) (Cleynen and Van de Ven, 2008). Les motifs AT-Hook reconnaissent et fixent des séquences riches en AT au niveau du petit sillon de l'ADN. C'est plus la structure qu'une séquence bien définie d'ADN qui est reconnue. Chaque motif AT-Hook fonctionne indépendamment et cette indépendance joue un rôle important dans la capacité des protéines HMGA à modifier la conformation de la chromatine. La fixation simultanée de 2 ou 3 des motifs AT-Hook sur des sites ADN distincts augmente la force d'interaction et l'espacement des liaisons sur l'ADN peut avoir un effet sur la conformation de ce dernier. De la même manière, la fixation de plusieurs protéines HMGA agissant de concert pourrait renforcer les changements conformationnels (Cleynen and Van de Ven, 2008). Il est à noter que les motifs AT-Hook peuvent intervenir aussi dans des interactions protéine-protéine.

La fonction de l'extrémité C-terminale acide des protéines HMGA n'est toujours pas complètement caractérisée. Il est considéré que ce domaine serait important pour la modulation des interactions protéine-protéine et qu'il serait impliqué dans l'augmentation de l'activité de facteurs de transcription (Cleynen and Van de Ven, 2008; Noro et al., 2003; Yie et al., 1997).



Figure 49 : Alignement des séquences des protéines HMGA1 et HMGA2. Domaines AT-Hook (gras) ; motif répété RGRP (boîte grise) ; séquence C-terminale acide (boîte rouge) ; délétion interne de 11 acides aminés pour HMGA1b (astérisques) comparativement à HMGA1a ; acides aminés de HMGA1c qui diffèrent de HMGA1a (souligné). *Source : Cleynen & Van De Ven (2008).*

1.1.c. Mécanismes d'action des protéines HMGA1

Comme mentionné ci-dessus, les protéines HMGA1 sont des régulateurs de la structure de la chromatine via des interactions directes entre les motifs AT-Hook et les séquences riches en AT de la fibre chromatinienne. Elles participent ainsi à une grande variété de processus cellulaires, allant du contrôle de la compaction et de l'accessibilité de la chromatine à la régulation de l'expression des gènes (**Figure 50**), qui sont décrits succinctement ci-après.

Les régions riches en AT de l'ADN, reconnues par les protéines HMGA1, ont une forte affinité pour l'interaction avec la matrice nucléaire au niveau de régions nucléaires appelées Matrix/Scaffold Associated regions (MARs/SARs) et de ce fait participent à l'organisation de la topologie de l'ADN génomique en boucles importantes pour la transcription (Cleynen and Van de Ven, 2008).

D'autre part, les protéines HMGA1 rentrent en compétition avec les histones H1 pour la liaison à l'ADN et modifient l'accessibilité de la chromatine. Ceci a été particulièrement mis en évidence sur des cellules précurseurs neuronaux (NPC) au cours du développement. Aux stades précoces du développement, les protéines HMGA1 sont essentielles au maintien d'un état ouvert de la chromatine et du potentiel neurogénique des NPC. L'expression de l'histone H1 antagonise ces effets et inhibe la différenciation neuronale (Kishi et al., 2012).

Grâce à leurs caractéristiques architecturales, les protéines HMGA1 qui n'ont pas d'activité transcriptionnelle intrinsèque, jouent un rôle important dans la transcription. Elles influencent la transcription en orchestrant l'assemblage de facteurs de transcription sur l'ADN (**Figure 50**). De nombreux facteurs de transcription ont été identifiés comme interagissant avec les protéines HMGA (b-caténine, Maf, PDX-1, HIF1, ect.) (Sgarra et al., 2010). Il a également été montré que HMGA1 pouvait moduler la transcription en recrutant sur des promoteurs cellulaires le complexe P-TEFb (CycT1/CDK9) inactif car lié au complexe répresseur CTIP2 (Hexim1 et 7SK RNA) (Eilebrecht et al., 2014). De plus, il est généralement admis que (i) la liaison préférentielle de HMGA1 au niveau des SARs présentes à la base de boucles de chromatine dans un état condensé et réprimé ainsi que (ii) le déplacement des histones H1, conduisent à une ouverture locale de la chromatine et permettent l'initiation de la transcription (Fusco and Fedele, 2007).

Enfin, HMGA1 est décrit pour être un élément important des enhanceosomes qui sont impliqués dans la transcription de gènes spécifiques de tissus. L'activité des enhanceosomes nécessite le recrutement stéréospécifique de facteurs de transcription et la courbure de l'ADN pour ensuite recruter des co-activateurs transcriptionnels. Ainsi, l'activation de la transcription du gène de l'interféron β en réponse à une infection virale nécessite l'assemblage d'un enhanceosome comprenant les facteurs NF-KB, ATF-2/c-Jun, IRF et HMGA1. En premier lieu, la fixation de HMGA1 induit des changements allostériques dans l'ADN qui facilitent le recrutement des facteurs de transcription. Les interactions entre HMGA1 et les facteurs de transcription stabilisent ensuite la structure de l'enhanceosome qui active alors la transcription de façon élevée (Yie et al., 1999). De la même manière HMGA1 est nécessaire à la formation de l'enhanceosome du promoteur précoce de HPV18 (Bouallaga et al., 2003).

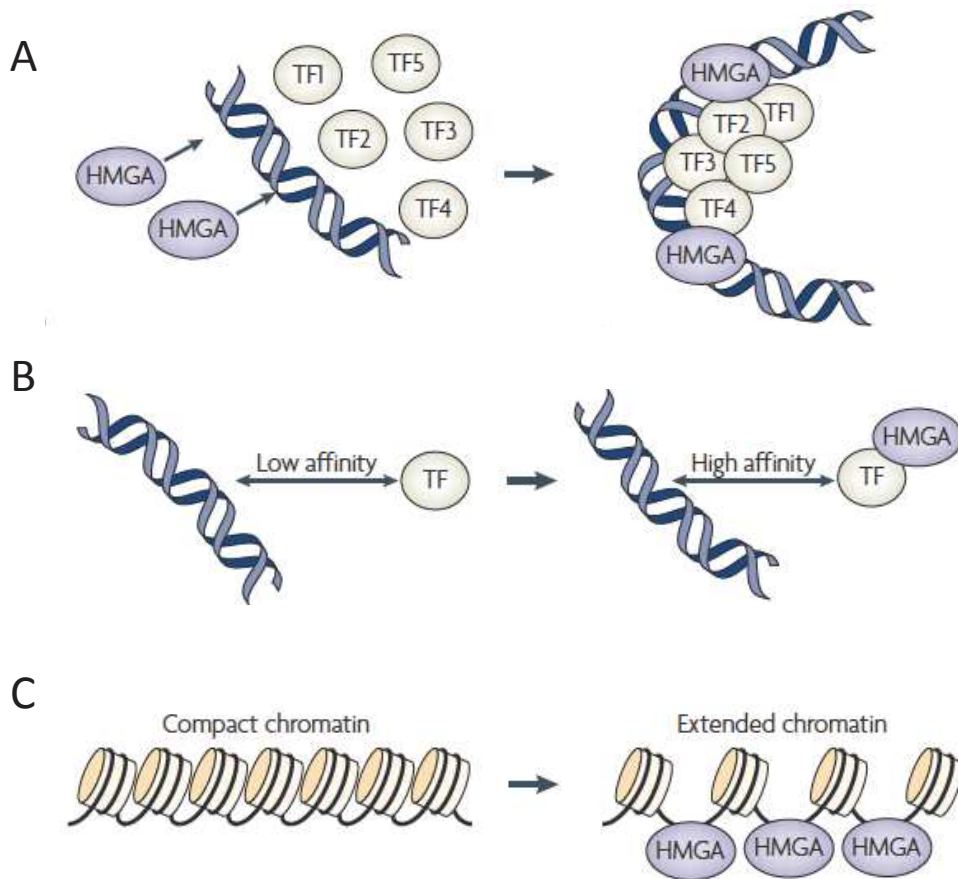


Figure 50 : Mécanismes d'action des protéines HMGA. **A)** En se liant à l'ADN, les protéines HMGA induisent des changements topologiques pour faciliter le recrutement d'autres facteurs de transcription (TF). **B)** En plus de leur capacité de liaison à l'ADN, les protéines HMGA ont également la capacité d'interagir directement avec les facteurs de transcription, de modifier leur structure et ainsi augmenter leur affinité pour leur séquence ADN cible. **C)** Les protéines HMGA ont également la capacité de changer la structure chromatinienne, notamment en déplaçant les histones H1 ce qui conduit à une ouverture locale de la chromatine. *Source : Fusco & Fedele (2007).*

Il est à noter que les protéines HMGA1 sont la cible de modifications post-traductionnelles (phosphorylation, acétylation, méthylation, etc.) susceptibles de modifier leur activité (Young et al., 2010). Par exemple, leur phosphorylation par HIPK2 ou CDC2 diminue leur affinité pour l'ADN (Zhang and Wang, 2007) et la méthylation par PRMT6 des résidus arginines au niveau des AT-Hook est susceptible de modifier les interactions protéine-protéine et protéine-ADN (Sgarra et al., 2006).

I.2. Fonctions physiologiques de HMGA1

Les protéines HMGA1, mais également HMGA2, sont ubiquitaire et abondamment exprimées pendant le développement embryonnaire. En revanche, dans les tissus adultes, le niveau d'expression de HMGA1 est très faible et celui de HMGA2 est quasiment indétectable (Cleynen and Van de Ven, 2008). Si l'expression de HMGA1 est très faible dans les tissus adultes et les cellules différenciées non prolifératives, elle peut être cependant rapidement induite en réponse à divers facteurs de croissance (Reeves et al., 2001).

Le rôle crucial des protéines HMGA dans le développement embryonnaire, la croissance et la détermination de la taille des mammifères a été démontré par la génération de souris knock-out (KO) pour les gènes HMGA1 et/ou HMGA2 (Fedele and Fusco, 2010; Fusco and Fedele, 2007). Les souris KO pour le gène HMGA1 ont une taille discrètement inférieure à la moyenne et développent des hypertrophies cardiaques (Fedele et al., 2006) ainsi qu'un diabète de type 2 (Foti et al., 2005), suggérant un rôle de HMGA1 dans la croissance des cardiomyocytes et dans la modulation de la voie de signalisation de l'insuline. Les souris KO pour le gène HMGA2 ont un phénotype particulier dit « pygmé » correspondant à une forte réduction de la taille des animaux (25% pour les hétérozygotes et 60% pour les homozygotes) associé à une forte diminution du tissu adipeux (Anand and Chada, 2000; Zhou et al., 1995), suggérant un rôle de HMGA2 dans la détermination de la taille et dans la différenciation des adipocytes. Le double KO des gènes HMGA1 et HMGA2 donne naissance à des souris de phénotype dit « super-pygmé », caractérisé par une létalité embryonnaire importante et une réduction de l'ordre de 75% de la taille des animaux. Ces résultats font apparaître une redondance partielle entre les membres de la famille HMGA pour le contrôle de la taille corporelle (Federico et al., 2014).

L'expression ectopique de HMGA1, mais pas celle de HMGA2, dans des cellules épithéliales mammaires induit la transformation cellulaire ainsi que la formation de tumeurs métastatiques lors de xénogreffes chez la souris, mettant ainsi en évidence le caractère oncogénique de l'expression de HMGA1 dans les tissus adultes et des activités tumorigéniques différentes pour HMGA1 et HMGA2 *in vivo* (Reeves et al., 2001). L'activité transformante de HMGA1 a été observée dans différents modèles cellulaires : cellules épithéliales thyroïdiennes de rat (Giancotti et al., 1985, 1987), fibroblastes de rat (Fedele et al., 1998) et lymphoblastes humains (Wood et al., 2000).

En accord avec ces observations, de fortes expressions de HMGA1 sont retrouvées dans de nombreuses tumeurs humaines et à l'heure actuelle HMGA1 est considéré comme un marqueur de tumeur, le degré de son expression étant corrélé à la malignité de la tumeur, au potentiel métastatique, à la chimiorésistance et à un mauvais pronostic (Pallante et al., 2015).

En ce qui concerne le cancer du sein, qui constitue le contexte de mon étude, l'analyse sur « tissu microarray » de l'expression des protéines HMGA1 dans les carcinomes du sein a formellement montré que HMGA1, qui n'est pas détectable dans le tissu mammaire normal, est surexprimé dans la grande majorité des échantillons des carcinomes mammaires analysés (**Figure 51**) (Chiappetta et al., 2004; Evans et al., 2004; Mansueto et al., 2010; Pallante et al., 2015). Par ailleurs, l'expression de HMGA1 a récemment été décrite comme étant corrélée au grade, au stade, à la présence de métastases dans les ganglions et à distance ainsi qu'au statut triple négatif des tumeurs du sein (Huang et al., 2015; Huso and Resar, 2014).

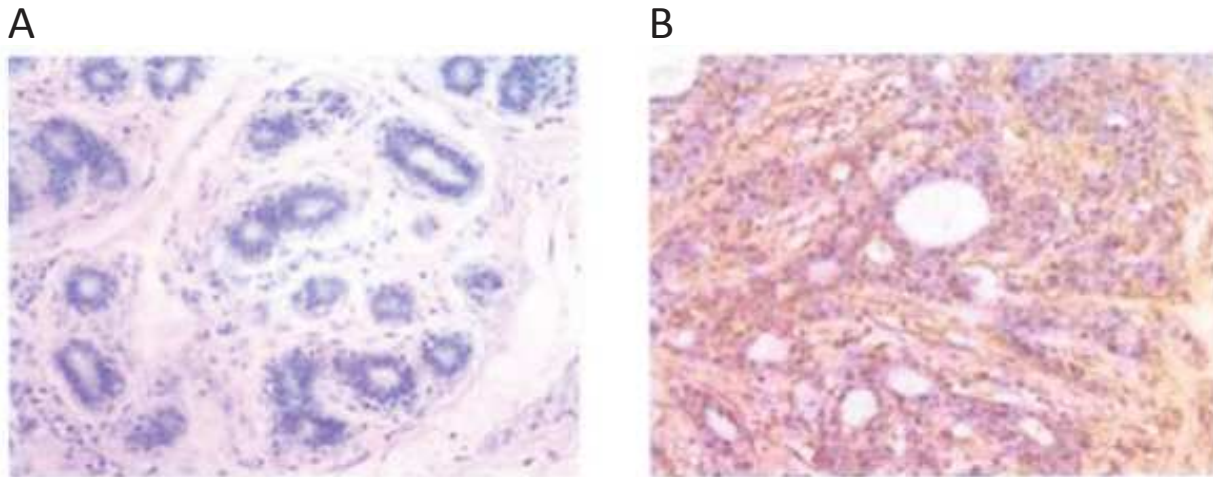


Figure 51 : Expression des protéines HMGA1 dans le tissu mammaire normal et cancéreux. Le marquage immunohistochimique de HMGA1 sur un échantillon de tissu mammaire cancéreux révèle une expression nucléaire très forte (B) alors que sur le tissu mammaire normal le marquage est principalement négatif (A). *Source : Evans et al. (2004)*

De façon intéressante, l'inhibition de HMGA1 dans les cellules MDA-MB231 induit la répression de gènes mésenchymateux (vimentin, snail) et l'induction de gènes épithéliaux (E-cadherin), aboutissant à la transition d'un état hautement prolifératif, invasif et mesenchymal-like des cellules vers un état plus épithélial-like associé à une croissance plus lente (Shah et al., 2013). De plus, l'inhibition de HMGA1 dans les cellules MDA-MB231 prévient leur capacité à former des mammosphères primaires, secondaires et tertiaires dans un modèle en 3 dimensions ainsi que leur capacité à initier des tumeurs après une implantation orthotopique *in vivo* (Di Cello et al., 2013; Pegoraro et al., 2013; Shah et al., 2013).

I.3. Mécanismes connus de la régulation de l'expression de HMGA1

Plusieurs mécanismes transcriptionnels et post-transcriptionnels ont été décrits pour réguler l'expression de HMGA1.

I.3.a. Mécanismes transcriptionnels

Au niveau transcriptionnel, plusieurs facteurs de transcription ont été décrits pour contribuer à l'expression de HMGA1. En particulier, plusieurs études ont montré un lien entre l'expression de HMGA1 et les facteurs de transcription AP-1 (Cleynen et al., 2007; Dhar et al., 2004; Hommura et al., 2004; Ogram and Reeves, 1995; Vallone et al., 1997). Il a notamment été proposé que l'expression élevée de HMGA1 dans plusieurs modèles cellulaires serait principalement contrôlée par la protéine de spécificité 1 (SP1) et les facteurs de transcription AP-1, ces deux facteurs agissant comme des éléments de régulation positive. En effet, en 2004 Hommura *et al* ont montré que le site AP-1 situé à -1091 pb (en amont du TSS) dans la région promotrice de HMGA1 était important pour la régulation de l'expression de HMGA1 dans une lignée de rat (Rat1a) exprimant c-Jun inducible par la doxycycline (Hommura et al., 2004). Puis, en 2007 Cleynen *et al* ont identifié deux régions de contrôle proximal qui sont importantes pour la transcription basale de HMGA1 dans des lignées de cancer du colon. Ils ont rapporté que SP1 et AP-1 se fixaient respectivement à -57 pb et +246 pb dans la région promotrice de HMGA1. De plus, les auteurs ont montré que l'expression de HMGA1 est fortement inducible par l'oncogène Ras, via une région de contrôle distal. Un site AP1 (-1079 pb) et trois sites SP1-like (-1301, -1294 et -1283 pb) semblent impliqués dans cette activation transcriptionnelle (Cleynen et al., 2007).

En 2013, Chiefari *et al* ont identifié un nouveau mécanisme de régulation du gène HMGA1 qui implique le motif d'ADN octamer ATGCAAAT (Oct) (Chiefari et al., 2013). Les auteurs ont montré que bien que les protéines Oct-1 et Oct-2 se fixent sur la même séquence consensus située en amont de la région promotrice de HMGA1, elles exercent des rôles opposés sur sa transcription. En effet, si Oct-2 augmente l'expression de HMGA1, Oct-1 la réprime. De plus, les auteurs ont également montré que la protéine HMGA1 se lie à son propre promoteur et contribue à sa transactivation en coopérant avec Oct-2 (mais pas avec Oct-1).

Par ailleurs, une autre étude a montré que la voie Wnt/beta-catenin/T-cell factor 4 induit l'expression de HMGA1 dans le cancer du colon (Bush et al., 2013). En effet, la répression de TCF4 par siRNA diminue l'expression de HMGA1 et deux sites de liaison pour TCF4 ont été identifiés par CHIP (-487/-801 pb et -1247/-1599 pb) sur la région promotrice du gène HMGA1.

1.3.b. Mécanismes post-transcriptionnels

En parallèle à la régulation transcriptionnelle de HMGA1, l'abondance de l'ARNm, et donc de la protéine, est contrôlée par des mécanismes post-transcriptionnels faisant intervenir des microARN. En ce qui concerne le cancer du sein, trois microARN, miR26A, Let7a et miR625 ont été décrits pour cibler HMGA1. De façon intéressante, le niveau d'expression de ces microARN est beaucoup plus faible dans les cellules et les tissus cancéreux du sein comparativement aux cellules et tissus adjacents normaux (Liu et al., 2015; Zhao et al., 2015; Zhou et al., 2016).

Ajoutant un niveau de complexité supplémentaire, il existe des pseudogènes de HMGA1 dont deux sont transcrits mais ne sont pas traductibles (HMGA16 et HMGA1P7). Ils se comportent comme des leurres piégeant les microARN et leur surexpression conduit à l'augmentation de l'abondance des protéines HMGA1. Une corrélation entre leur abondance et l'agressivité des tumeurs a été montrée dans le cas des carcinomes thyroïdiens et ovariens (De Martino et al., 2016).

D'autres études sont nécessaires pour élucider l'ensemble des mécanismes moléculaires responsables de la surexpression de HMGA1 dans les cancers. Les protéines AP-1 étant surexprimées dans les cancers du sein agressifs et HMGA1 ayant été identifié comme régulé par Fra-1 et Fra-2 dans la lignée modèle de cancers du sein triple négatifs MDA-MB231, j'ai cherché à approfondir le lien entre AP-1 (Fra-1 et Fra-2) et régulation transcriptionnelle de HMGA1 dans ce contexte tumoral.

II. HMGA1 : Résultats

Comme nous l'avons vu dans l'introduction de cette partie, la régulation de l'expression de HMGA1 fait intervenir des mécanismes transcriptionnels et post-transcriptionnels. Avant de commencer mon étude sur les mécanismes transcriptionnels gouvernés par Fra-1 et/ou Fra-2 pour l'expression de HMGA1, il était crucial de m'assurer que le rôle de Fra-1 et/ou Fra-2 était d'ordre transcriptionnel et non indirect en modulant l'expression de microARN affectant la stabilité et/ou la traductibilité de l'ARNm de HMGA1.

II.1. Fra-1 et Fra-2 régulent la transcription du gène HMGA1

Pour répondre à cette question, j'ai caractérisé l'effet de Fra-1 et/ou Fra-2 sur la synthèse des ARN naissants de HMGA1. Dans ce but, j'ai mis en place au laboratoire une technique de « Nuclear Run-On assay » (Roberts et al., 2015) pour mesurer l'abondance des ARN naissants liés à Pol II en présence et en absence de Fra-1 et/ou Fra-2 par transfection des siRNA spécifiques ou du siRNA contrôle. Les ARN purifiés ont été rétrotranscrits et les cDNA obtenus ont été amplifiés par qPCR en utilisant deux couples de primers situés à +1.22kb et +3.23kb en aval du TSS de HMGA1, pour augmenter la robustesse des conclusions.

Les résultats présentés dans la **Figure 52** montrent que l'abondance des ARN naissants HMGA1 est réduite d'un facteur supérieur ou égal à 2 en absence de Fra-1 et/ou Fra-2, indiquant l'implication de Fra-1 et Fra-2 au niveau de la transcription du gène HMGA1.

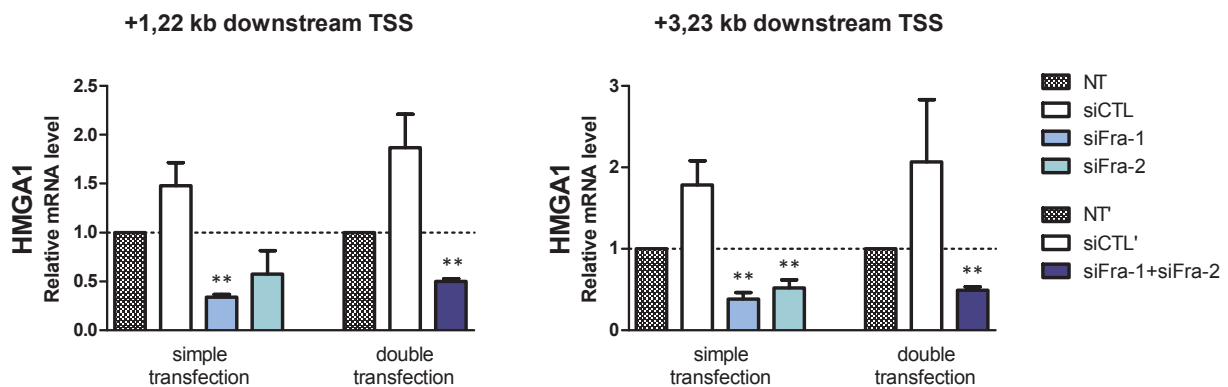


Figure 52 : Abondances relatives des ARN naissants HMGA1 en absence de Fra-1 et/ou Fra-2 dans les MDA-MB231. Abondance relative des ARN naissants HMGA1 analysée par RT-qPCR avec deux couples d'amorces permettant l'amplification des régions situées à +1.22 et +3.23 kb en aval du TSS de HMGA1. Les valeurs sont normalisées par rapport à l'ARN naissant GAPDH et relatives à l'abondance mesurée dans la condition non-transfectée (NT ou NT'). n=3. Two tailed paired t-test ; (**) p value <0.01.

Il est à noter que nos résultats par RT-qPCR ont montré que l'absence de Fra-2 a un effet faible et non significatif sur l'abondance à l'équilibre des l'ARNm HMGA1 (**Figure 46** de l'étude I, paragraphe IV.2). Cette différence dans la contribution de Fra-2 pourrait résider dans

les différences expérimentales liées aux deux approches utilisées, mais aussi à la stabilité de l'ARNm de HMGA1, d'une part, et au temps nécessaire à l'extinction de la protéine Fra-2 suite à la transfection du siRNA, d'autre part. En effet, dans la technique de Run-On, la quantité relative des ARN naissants est mesurée sur une fenêtre de 40 minutes, 72h après la transfection des siRNA et donc à un moment où les protéines Fra-1 et Fra-2 ne sont plus détectables en immunoblot. Par contre, dans la mesure de l'abondance à l'équilibre on évalue les quantités relatives des ARNm HMGA1 sur une période de 72h pendant laquelle les niveaux des protéines Fra-1 et Fra-2 diminuent au cours du temps suite à la transfection des siRNA spécifiques. Une disparition plus lente de la protéine Fra-2 comparativement à Fra-1 associée à une demi-vie relativement importante des ARNm HMGA1 (>12h et 26h dans des cellules épithéliales de poumon et des lymphoblastes infectés par EBV, respectivement (Chen et al., 2016; Foti et al., 2005)) pourrait rendre compte de cette différence.

II.2. Identification des mécanismes moléculaires gouvernés par Fra-1 et/ou Fra-2 pour la transcription du gène HMGA1

Le but de cette étude est de caractériser le ballet des co-activateurs transcriptionnels recrutés par Fra-1 et/ou Fra-2 pour activer l'expression de HMGA1. Pour étudier le rôle de Fra-1 et Fra-2 dans la transcription de HMGA1, j'ai utilisé une approche par ChIP-qPCR pour identifier des changements potentiels dans les marques d'histones ou dans le recrutement de différents co-activateurs transcriptionnels, en présence et en absence de Fra-1 et/ou Fra-2. Dans le but de choisir de façon la plus judicieuse possible les amplicons à utiliser en ChIP-qPCR pour le locus HMGA1, je me suis basée sur les résultats obtenus pour nos différents ChIP-seq présentés dans l'étude I au paragraphe III.

II.2.a. Données de ChIP-seq pour le gène HMGA1

Nos données de ChIP-seq dans les cellules MDA-MB231 montrent 5 sites de fixation pour Fra-1 et Fra-2 au niveau du locus HMGA1 : un pic de fixation à -12kb du TSS, un pic de fixation à -1kb et 3 pics sur la partie 3' du gène, dont un pic situé à +6.5kb dans le 4ième intron et 2 pics à +7.4kb et +7.9kb dans le 5ième et dernier intron (**Figure 53**). Aucun autre pic Fra-1 ou Fra-2 n'est détecté jusqu'à -530kb en amont ni jusqu'à +320kb en aval du TSS de HMGA1.

Les enhancers potentiels répertoriés dans FANTOM5 pour HMGA1 sont au nombre de 3 et se situent à -94kb, -1.7kb et +426.8kb du TSS (**Figure 53**). Cependant aucun d'entre eux ne présente de pic de fixation pour Fra-1 et/ou Fra-2. De plus, ces enhancers potentiels FANTOM5 ne présentent pas les caractéristiques typiques des enhancers actifs dans les MDA-MB231. En effet, dans cette lignée ces régions sont dépourvues des marques histones H3K4me1 et H3K27ac. Il est donc très vraisemblable que ces enhancers potentiels FANTOM5 ne soient pas fonctionnels dans la lignée MDA-MB231. De ce fait, il est très peu probable que la régulation de l'expression de HMGA1 par Fra-1 et Fra-2 soit indirecte et médiée par la fixation de facteurs de transcription,

eux-mêmes induits par Fra-1 et Fra-2, se fixant aux niveaux de ces enhancers FANTOM5 potentiels. Ces observations suggèrent que l'expression de HMGA1 pourrait être régulée par Fra-1 et/ou Fra-2 via le site situé au niveau du promoteur (-1kb) ou peut-être via un ou plusieurs enhancers (+12kb, +6.5kb, +7.4kb et 7.9kb) non répertoriés dans FANTOM5 et spécifiques des cellules MDA-MB231.

Le site de fixation situé à -1kb ne possède pas les caractéristiques d'un enhancer car il ne présente pas d'enrichissement pour H3K4me1. Par contre, il est adjacent à un pic H3K4me3 et H3K27ac caractéristique des promoteurs actifs. Contrairement à la fixation de Fra-1 et Fra-2 sur la partie 3' du gène, le partenaire privilégié de Fra-1 et Fra-2 à la position -1kb ne semble pas être c-Jun. Il en est de même pour le site situé à -12kb qui présente des caractéristiques d'enhancer, bien que P300 soit absente à cette position (**Figure 53**). Il est également à noter que les 5 sites de fixation de Fra-1 et Fra-2 situés à -12kb, -1kb, +6.5kb, +7.4kb et +7.9kb sont caractérisés par des creux dans les profils de modifications d'histones qui traduisent des régions pauvres en nucléosomes, caractéristiques des enhancers et des promoteurs (**Figure 53**).

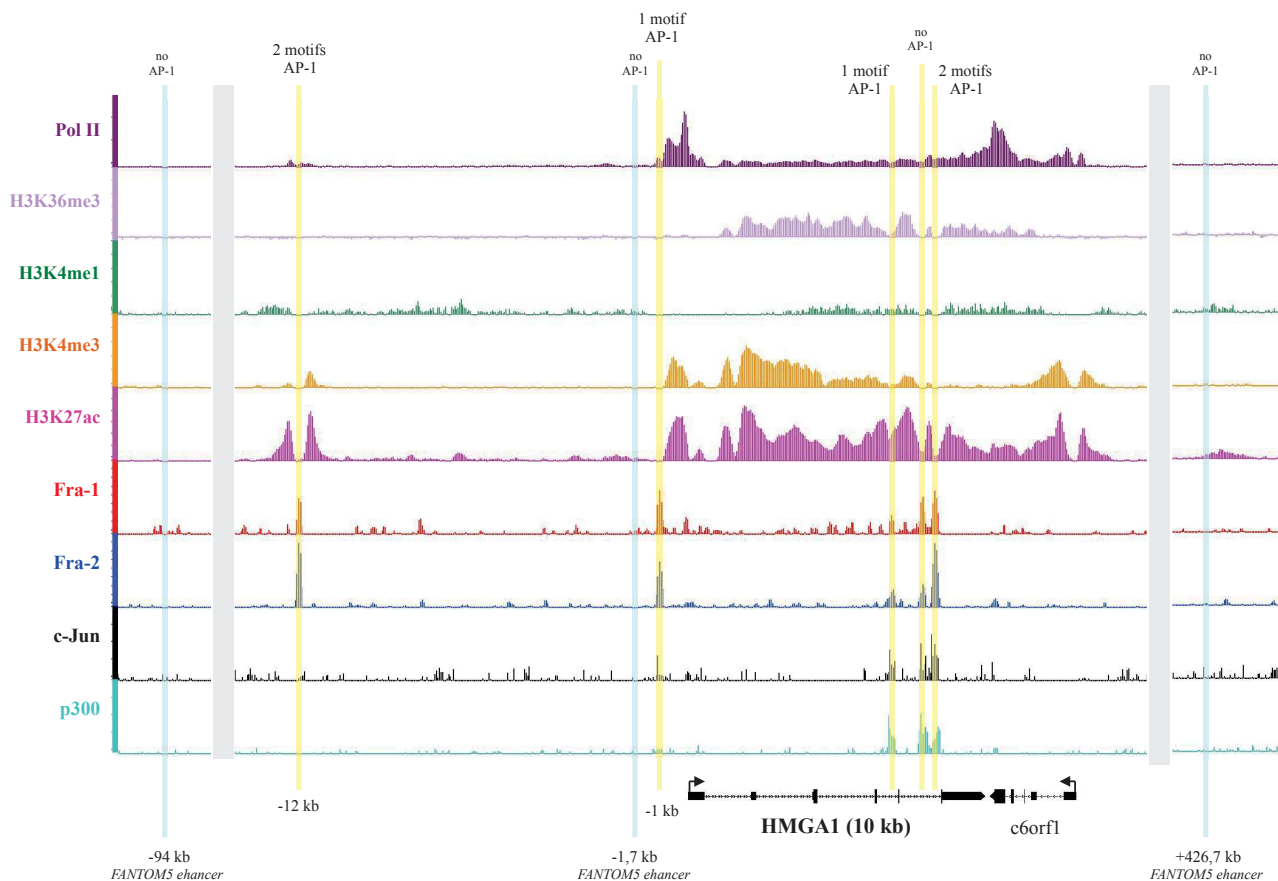


Figure 53 : Cartographie des données de ChIP-seq au niveau du locus de HMGA1 dans la lignée MDA-MB231 wild-type. Cartographie des sites de fixation de Fra-1, Fra-2, c-Jun et des marques caractéristiques des promoteurs et des enhancers actifs : H3K4me3, H3K4me1 H3K27ac, Pol II et P300. La cartographie de H3K36me3 (déposée par la forme élongatrice de Pol II) permet de distinguer le niveau de transcription des gènes. Les positions des enhancers Fantom5 sont indiquées en bleu et les positions des pics Fra-1/Fra-2 en jaune. Les barres grises indiquent des ruptures dans le chromosome pour pouvoir visualiser les enhancers FANTOM5 distaux (-94kb et +426.7kb). Une version haute définition de cette figure est disponible en annexe (**Annexe 6**).

Pour l'étude des mécanismes transcriptionnels j'ai donc choisi des amplicons allant de -12kb à la fin du gène HMGA1 et couvrant, pour certains, le TSS ou les sites de fixation Fra-1/Fra-2 situés à -12kb, -1kb et +7.9kb comme cela est indiqué sur la **Figure 54**.

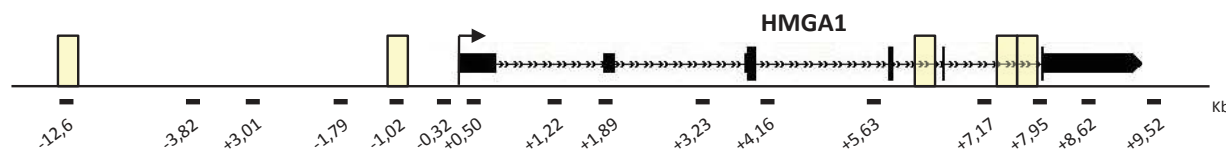


Figure 54 : Structure du locus de HMGA1 et position des fragments amplifiés par ChIP-qPCR. Position des cinq pics Fra-1/Fra-2 (boîtes jaunes) ; Transcription start site ou TSS (flèche). Les chiffres indiquent le milieu des amplicons.

II.2.b. Analyse de certaines modifications d'histones en présence et en absence de Fra-1 et/ou Fra-2 au niveau du gène HMGA1

Comme mentionné dans l'avant propos de ma thèse, les anticorps polyclonaux de lapin anti-Fra-1 et anti-Fra-2 que nous utilisons au laboratoire n'ont plus été fournis par Santa Cruz Biotechnology suite à des problèmes sanitaires et à une fermeture express de leur animalerie (« Thousands of goats and rabbits vanish from major biotech lab » Nature News | 19 February 2016 | 10.1038/nature.2016.19411). Santa Cruz Biotechnology propose depuis des anticorps monoclonaux que nous avons commandés. Cependant aucun des anticorps anti-Fra-1 testés n'a donné satisfaction en ChIP-qPCR, les signaux étant au niveau du bruit fond. Il en est de même pour les anticorps que nous nous sommes procurés auprès d'autres fournisseurs. En l'absence d'anticorps anti-Fra-1 commerciaux efficaces, le laboratoire va sans doute se trouver dans l'obligation de faire produire un anticorps anti-Fra-1 utilisable en ChIP pour la suite de ce travail.

Au cours des expériences présentées ci-dessous, j'ai pu uniquement valider en ChIP-qPCR l'absence de recrutement de Fra-2 en présence du siFra-2 sur le locus HMGA1. Néanmoins, pour chaque expérience l'extinction de l'expression de Fra-1 et Fra-2 a été systématiquement contrôlée par immunoblot (**Figure 55**).

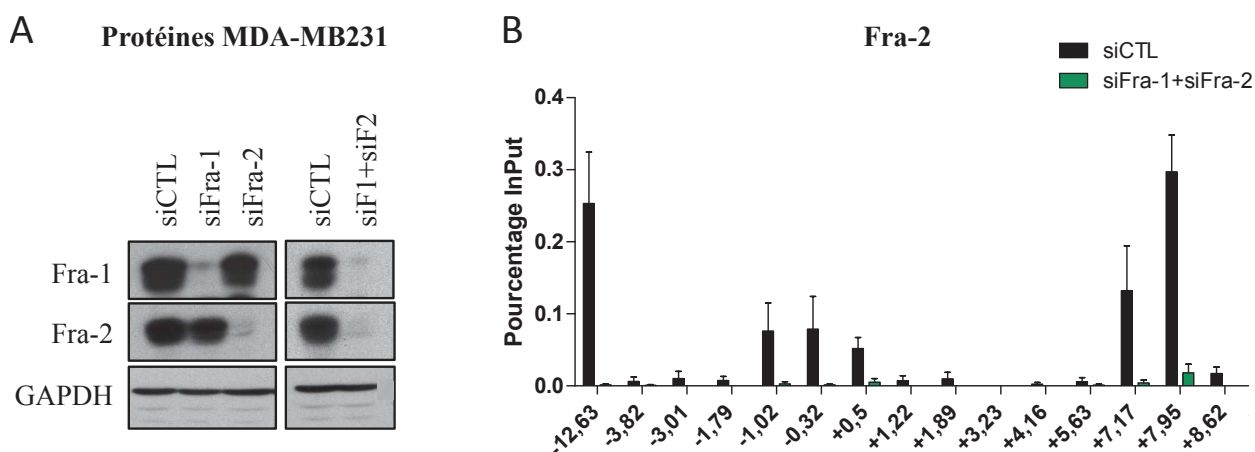


Figure 55 : Expression des protéines Fra-1 et/ou Fra-2 après transfection de siRNA dans la lignée MDA-MB231, pour les expériences de ChIP-qPCR. A) Abondance des protéines Fra-1 et Fra-2 analysée par immunoblot, après inhibition de Fra-1 et/ou Fra-2 par siRNA. L'immunoblot GAPDH est présenté en témoin de charge. B) Données de ChIP-qPCR pour Fra-2 au niveau du locus de HMGA1 après inhibition de Fra-1+Fra-2 par transfection de siRNA dans la lignée MDA-MB231. n=5.

Dans un premier temps, j'ai choisi de déterminer si la suppression de Fra-1 et/ou Fra-2 par RNAi conduit à une altération de l'environnement chromatinien de HMGA1. J'ai testé les marques d'histones caractéristiques des enhancers et des promoteurs actifs à savoir H3K4me1, H3K4me3 et H3K27ac. Toutes les expériences ont été réalisées au moins trois fois de façon indépendante pour avoir un recul statistique. Les résultats sont présentés sur la **Figure 56**.

Les résultats de ChIP-pCR obtenus pour les trois modifications de l'histone H3 analysées en présence du siCTL sont en accord avec les résultats du ChIP-seq effectués sur les MDA-MB231 wild-type. En effet, on observe un enrichissement en H3K4me1 sur le corps du gène à partir de +3kb. On peut noter également une faible mono-méthylation au niveau du site +7.95kb sans doute due à une diminution en histone H3 sur cette position, liée à une absence ou à une faible densité en nucléosome. Un ChIP dirigé contre l'histone H3 non modifiée permettrait d'apporter une réponse formelle. En ce qui concerne H3K4me3, on observe un enrichissement au niveau de la région promotrice avec un creux au niveau du site +0.5kb en accord avec les données de ChIP-seq. De la même manière, H3K27ac suit le profil obtenu en ChIP-seq. Ces observations valident l'approche ChIP-qPCR utilisée pour notre étude.

En présence des siRNA dirigés contre Fra-1 ou contre Fra-2, de même que lorsque les deux protéines sont éteintes conjointement, on n'observe aucun changement majeur quant aux trois modifications d'histones analysées, à l'exception d'une légère diminution de signal en absence de Fra-1 et Fra-1+Fra-2 pour H3K4me3 à partir de +2kb jusqu'à +7kb en aval du TSS de HMGA1 (**Figure 56.B**). Cette position génomique correspondrait au transcrit le plus court de HMGA1. En effet, HMGA1 a été décrit pour produire cinq transcrits codants pour les protéines HMGA1 et l'un d'entre eux est initié à +2kb du TSS de HMGA1.

Globalement, ces résultats indiquent que sur le temps de l'expérience (72h post-transfection des siRNA) l'absence de Fra-1 et/ou Fra-2 n'a pas d'impact majeur sur ces modifications d'histones.

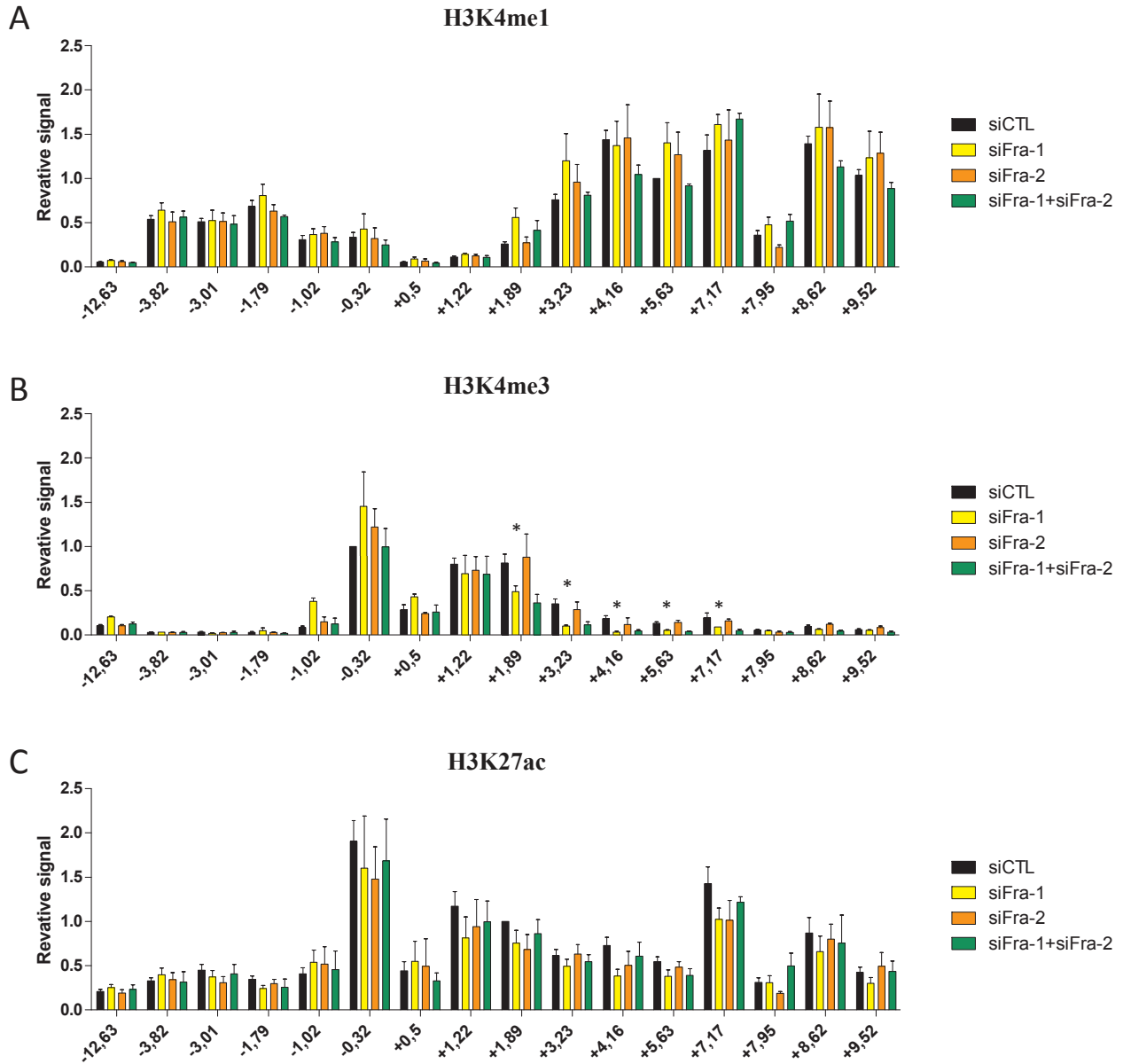


Figure 56 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour H3K4me1, H3K4me3 et H3K27ac, en présence et en absence de Fra-1 et/ou Fra-2 par transfection des siRNA contrôle ou spécifiques dans la lignée MDA-MB231. n=3-5. Lorsqu'une différence significative d'expression (p value<0,05) est observée entre la condition siCTL et l'une des trois conditions siRNA celle-ci est signalée par une étoile (*), two-tailed paired t-test.

II.2.c. Analyse du recrutement de l'ARN polymérase Pol II et de son activité

L'étape suivante a été de déterminer si Fra-1 et/ou Fra-2 étaient impliqués dans le recrutement de la machinerie de transcription ou dans l'activité de l'ARN Pol II. Pour ce faire j'ai utilisé pour les expériences de ChIP : (i) un anticorps anti-Pol II dirigé contre la partie N-terminale de RPB1 (le même anticorps a été utilisé pour effectuer le ChIP-Seq Pol II) qui permet d'immunoprécipiter toutes les formes de Pol II indépendamment de la phosphorylation du CTD, (ii) un anticorps spécifiquement dirigé contre la forme phosphorylée sur la sérine-5 du CTD qui est caractéristique de la forme initiateur de Pol II et (iii) un anticorps spécifiquement dirigé contre la forme phosphorylée sur la sérine-2 du CTD qui est caractéristique de la forme élongatrice et processive de Pol II.

Concernant Pol II total, les résultats obtenus avec le siRNA contrôle sont en accord avec les résultats du ChIP-seq Pol II où l'on observe une accumulation de Pol II de part et d'autre du TSS (**Figure 57.A**). De façon intéressante, nos résultats montrent une nette diminution de l'abondance de Pol II au niveau du TSS en absence de Fra-1. Le niveau de Pol II est, par contre, non significativement réduit en absence de Fra-2. La double extinction Fra-1+Fra-2 conduit à des résultats similaires à ceux obtenus avec la déplétion phénotypique de Fra-1. Ces résultats indiquent donc que, contrairement à Fra-2, Fra-1 est nécessaire pour le recrutement de Pol II sur le gène HMGA1.

L'analyse de la forme initiateur de Pol II, phosphorylée sur la sérine-5 du CTD, révèle également une diminution de cette forme phosphorylée en absence de Fra-1, l'absence de Fra-2 n'ayant pas d'effet notable (**Figure 57.B**). De plus, en présence du siCTL on observe une légère augmentation de la phosphorylation sur la partie 3' du gène au niveau des régions de fixation de Fra-1 et Fra-2. Cette marque est normalement restreinte au TSS et au début des gènes et absente en 3' des gènes.

En ce qui concerne la phosphorylation sur la sérine-2 du CTD on observe également une diminution du niveau de phosphorylation uniquement lorsque Fra-1 est réprimé par siRNA (**Figure 57.C**). L'absence de Fra-2 n'a pas d'effet. Cependant, l'absence conjointe de Fra-1 et Fra-2 entraîne sur tous les amplicons analysés une diminution légèrement accentuée de la forme Phospho-Ser2 de Pol II par rapport à ce qui est observé en absence de Fra-1. Par ailleurs, en présence du siCTL on remarque, de façon surprenante, un pic de phosphorylation au niveau du promoteur. Or cette modification n'est normalement présente que sur le corps des gènes et augmente pour atteindre son maximum sur la fin des gènes.

Ces deux résultats, présence de phospho-Ser5 en fin de gène et phospho-Ser2 en début de gène, suggèrent qu'il pourrait y avoir une boucle entre le promoteur et la région 3' du gène. Cette dernière contient des séquences avec des caractéristiques d'enhancers et qui sont des sites de liaison pour Fra-1 et Fra-2.

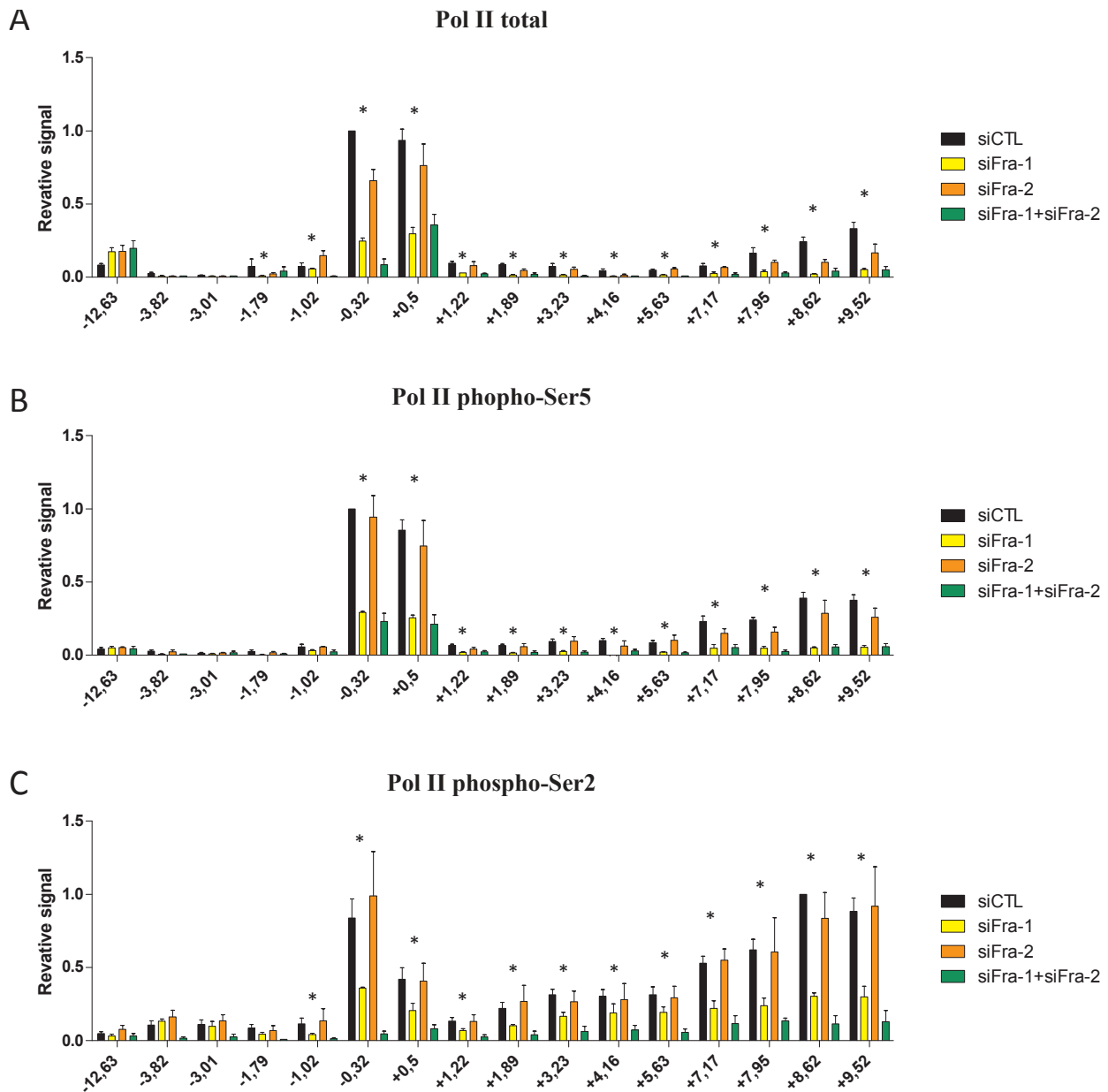


Figure 57 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour Pol II et ses formes phosphorylées sur les ser5 et ser2 du CTD, en présence et en absence de Fra-1 et/ou Fra-2 par transfection des siRNA contrôle ou spécifiques dans la lignée MDA-MB231. n=4. Lorsqu'une différence significative d'expression (p value<0,05) est observée entre la condition siCTL et les conditions siFra-1 et siFra-1+siFra-2, celle-ci est signalée par une étoile (*), two-tailed paired t-test.

En accord avec une diminution du recrutement de Pol II (**Figure 57.A**) et une diminution subséquente de la transcription de HMGA1 en absence de Fra-1 et (**Figure 52**) l'analyse de la modification H3K36me3, qui est une modification déposée sur le corps du gène par une méthyltransférase associée à Pol II sous sa forme élongatrice, montre que l'absence de Fra-1 se traduit également par une diminution de cette modification (**Figure 58**). L'absence de Fra-2 n'a aucun impact sur le niveau de H3K36me3 (**Figure 58**).

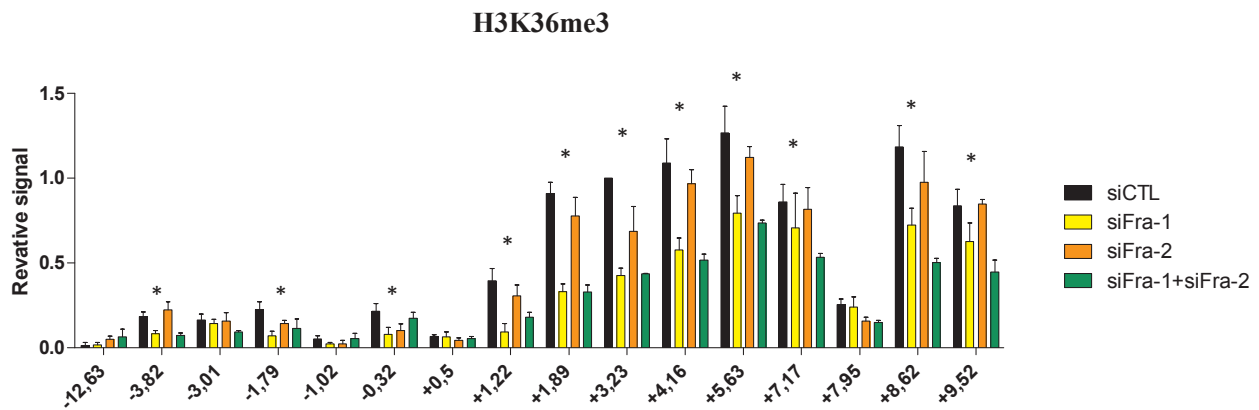


Figure 58 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour H3K36me3, après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231. n=4. Lorsqu'une différence significative d'expression (p value < 0,05) est observée entre la condition siCTL et l'une des trois conditions siRNA celle-ci est signalée par une étoile (*), two-tailed paired t-test.

La diminution de l'abondance des formes phospho-Ser5 et phospho-Ser2 en absence de Fra-1 semble être la conséquence de la diminution du recrutement de Pol II sur le gène et non liée à une diminution de ces modifications. En effet, les rapports Phospho-Ser5/Pol II et phospho-Ser2/Pol II ne varient pas (**Figure 59.A et B**). Par ailleurs, j'ai vérifié que la diminution de Pol II sur le locus HMGA1 n'était pas liée à une diminution globale de Pol II en effectuant des immunoblots pour Pol II et ses formes phosphorylées en présence et en absence de Fra-1 et/ou Fra-2 (**Figure 59.C**).

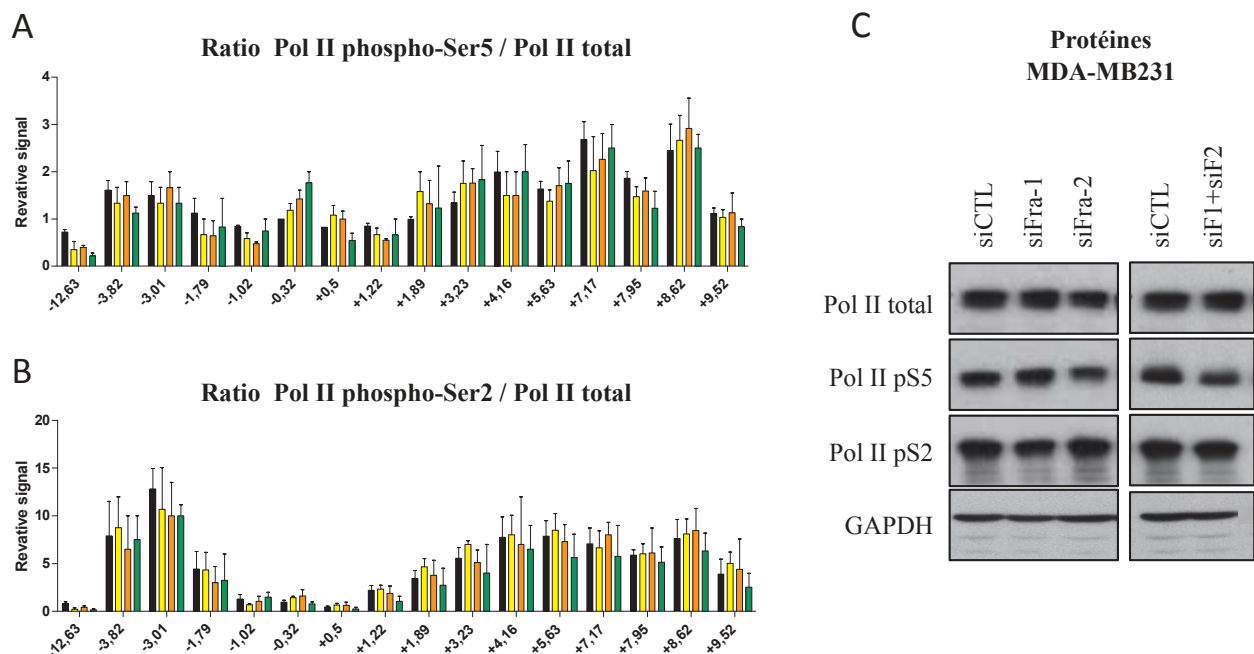


Figure 59 : Expression de Pol II et ses formes phosphorylées après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231. Données de ChIP-qPCR au niveau du locus de HMGA1 pour les rapports Phospho-Ser5/Pol II (A) et phospho-Ser2/Pol II (B) après inhibition de Fra-1 et/ou Fra-2 par siRNA. n=4. C) Abondance de la protéine Pol II et ses formes phosphorylées analysées par immunoblot, après inhibition de Fra-1 et/ou Fra-2 par siRNA.

J'ai également contrôlé la spécificité de la diminution de Pol II et de ses formes phosphorylées sur le locus HMGA1 en absence de Fra-1 en analysant le recrutement de Pol II au promoteur de S26 dont l'expression est indépendante de Fra-1 et Fra-2 (**Figure 60**).

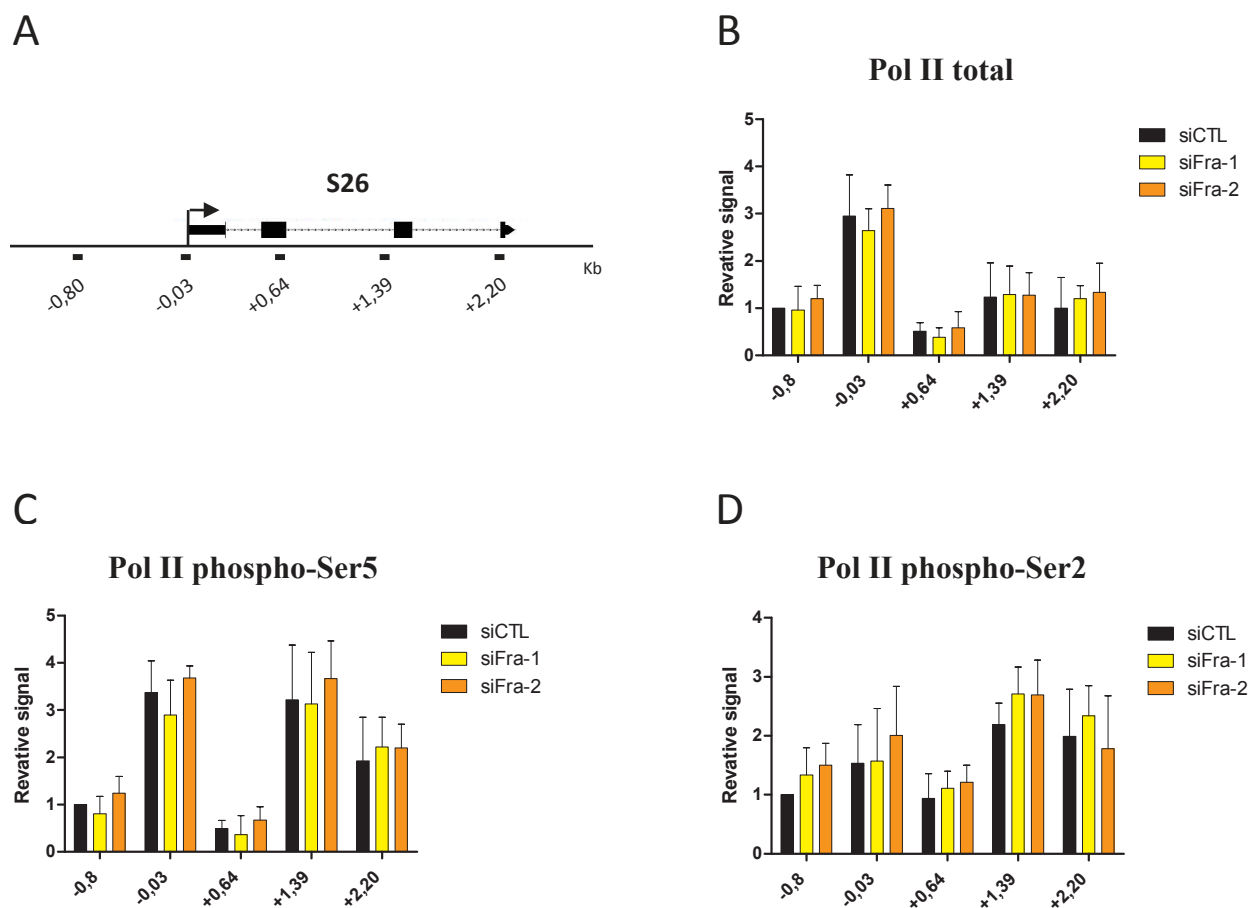


Figure 60 : Données de ChIP-qPCR au niveau du locus de S26 pour la forme initiatrice de Pol II et ses formes phosphorylées sur les Ser-5 et Ser-2 du CTD, après inhibition de Fra-1 ou Fra-2 par siRNA dans la lignée MDA-MB231. n=3.

II.2.d. Analyse du recrutement de P300 au niveau des régions enhancers du gène HMGA1

Fra-1 ayant été montré nécessaire au recrutement de P300 sur les enhancers du gène uPA (Moquet-Torcy et al., 2014), j'ai cherché à déterminer si Fra-1 était également impliqué dans le recrutement de P300 sur le locus HMGA1.

De manière similaire aux données de ChIP-seq on observe, en présence du siCTL, deux pics principaux P300 au niveau des régions qui lient Fra-1 et Fra-2 situées à +7.17kb et +7.95kb du TSS (**Figure 61**). Le recrutement de P300 est fortement diminué en absence de Fra-1, ce qui suggère que, dans le cas de HMGA1 comme dans celui de uPA, le recrutement de P300 nécessite la liaison préalable de Fra-1. Fra-2 n'est pas impliquée dans ce recrutement. On remarque également un léger enrichissement en P300, dans la condition contrôle, au niveau du TSS et jusqu'à -1kb ainsi qu'au niveau de l'enhancer potentiel situé à -12kb du TSS, observation qui pourrait aussi être en faveur d'une boucle sachant que P300 a été décrit comme un marqueur des enhancers et qu'il est impliqué dans la formation de boucles (Guo et al., 2016).

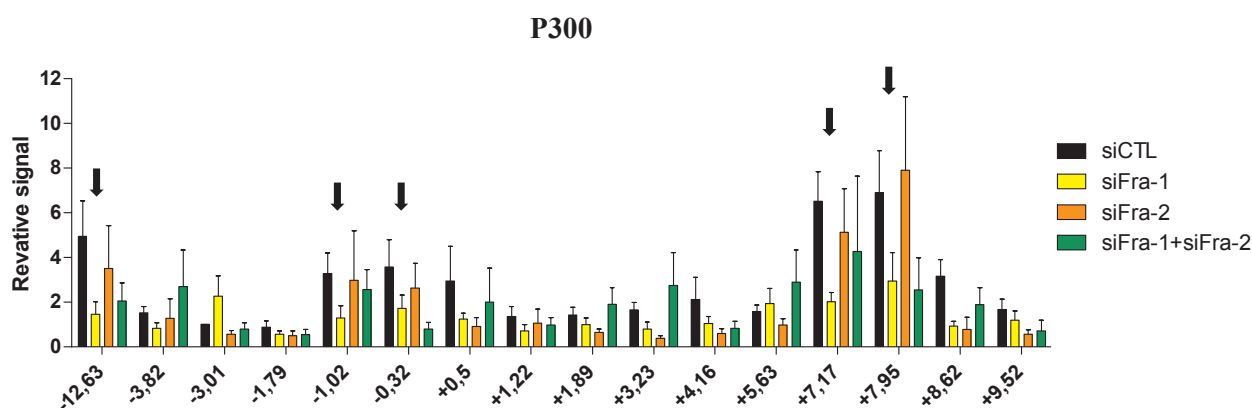


Figure 61 : Données de ChIP-qPCR au niveau du locus de HMGA1 pour P300, après inhibition de Fra-1 et/ou Fra-2 par siRNA dans la lignée MDA-MB231. n=3. Lorsqu'une variation est observée entre la condition siCTL et l'une des trois conditions siRNA celle-ci est signalée par une flèche.

III. HMGA1 : Conclusion

Les travaux réalisés montrent que, contrairement à Fra-2, Fra-1 est nécessaire au recrutement (i) de l'ARN Pol II sur le promoteur de HMGA1 et (ii) de P300 au niveau de deux enhancers potentiels situés dans le dernier intron du gène ainsi qu'au niveau du site à -12kb. Ces résultats suggèrent au moins deux rôles moléculaires de Fra-1 pour la transcription de HMGA1.

La diminution de la forme phosphorylée sur la sérine-2 de Pol II a tendance à être plus importante en l'absence conjointe de Fra-1 et Fra-2 qu'en l'absence de Fra-1 seul. Cette observation pourrait suggérer que Fra-2 a un rôle sur le niveau de cette phosphorylation. Cependant, cette hypothèse n'est pas étayée par les résultats obtenus lors de la déplétion phénotypique de Fra-2 seul qui ne montrent aucun changement dans le niveau de phosphorylation. De même, le ratio phospho-Ser2/Pol II total n'est pas diminué de façon significative en absence de Fra-1+Fra2.

Par ailleurs, les profils obtenus pour Pol II et les formes phospho-Ser5 et phospho-Ser2 ainsi que pour P300 suggèrent la possibilité de l'existence d'une boucle entre le promoteur et la région en aval contenant les deux enhancers potentiels. Il est à noter que le dimère JunD/Fra-2 a été très récemment montré responsable du recrutement de P300 sur l'enhancer du gène hiNOS et que le recrutement de P300 est nécessaire à la formation d'une boucle activatrice entre l'enhancer (AP-1/P300) et le promoteur (Guo et al., 2016). Un mécanisme similaire mais faisant intervenir le recrutement de P300 par Fra-1 pourrait donc être actif pour l'expression de HMGA1 dans les cellules MDA-MB231.

Pour déterminer si une boucle chromatinienne est impliquée dans le contrôle par Fra-1 et/ou Fra-2 de la transcription de HMGA1, plusieurs approches peuvent être envisagées :

(i) Des expériences de ChIP-qPCR dirigées contre les cohésines (SMC1, SMC3) sont envisagées en présence ou en absence de Fra-1 et/ou Fra-2. Un modèle en cours est que les cohésines sont impliquées dans les interactions enhancer-promoteur indépendamment de CTCF (Phillips-Cremins, 2014), CTCF et cohésines étant impliqués dans la formation des TAD (Topologically Associated Domains) qui définissent de larges sous-domaines chromatiniens conservés parmi les différents types cellulaires (Chung et al., 2016). Le complexe multiprotéique Mediator a été montré être un élément important de la liaison entre les facteurs de transcription et l'ARN polymérase et pourrait également participer à la formation d'une boucle potentielle. Nous chercherons donc à déterminer si certaines sous-unités du Mediator (Med1/trap220, Med19, Med25, ect.) sont recrutées au niveau du locus HMGA1 et si leur recrutement est affecté par l'absence de Fra-1 et/ou Fra-2.

(ii) En collaboration avec Fabienne Bejjani, en thèse au laboratoire, nous développerons l'approche capture C qui consiste à visualiser les interactions entre un nombre défini de

promoteurs et toutes les régions du génome. En effet, il n'existe aucune donnée de ChIA-PET ou de Hi-C dans les bases de données concernant les cellules MDA-MB231 et donc la cartographie des interactions chromatinienne n'est pas connue. Cette technique est basée sur une approche 3C (fixation, digestion avec des enzymes de restriction appropriées, ligation des régions en interaction) suivie d'une purification des promoteurs d'intérêt associés aux régions d'interaction par utilisation de sondes nucléotidiques biotinylées et d'un séquençage à haut débit (Davies et al., 2016; Hughes et al., 2014). Cette approche devrait permettre de valider la présence de contact entre le promoteur de HMGA1 et les enhancers potentiels. Le rôle éventuel de Fra-1 et/ou Fra-2 dans des interactions chromatinienne sera établi en effectuant des expériences de capture C en présence et en absence de Fra-1 et/ou Fra-2.

(iii) La détermination de l'implication des différents enhancers potentiels via Fra-1 et/ou Fra-2 sur l'expression de HMGA1 pourrait aussi être apportée par délétion conjointe ou indépendante des sites AP-1 par une approche CRISPR.

(iv) A l'inverse, on peut envisager de forcer la présence de Fra-1 sur un ou plusieurs sites de fixation pour mesurer l'effet sur l'expression de HMGA1 et/ou la formation d'une boucle par une approche originale. Le principe est le suivant : comme rapporté pour c-Fos (Olive et al., 1997), l'idée est de construire un mutant de Fra-1 capable de dimériser avec ses partenaires mais incapable de lier l'ADN. Ceci peut être obtenu en remplaçant le domaine basique de liaison à l'ADN de Fra-1 par un domaine acide en leucine zipper (coil-coiled) qui permet une interaction forte avec le domaine basique de liaison à l'ADN des membres Jun et empêche ainsi la fixation du dimère formé sur les motifs AP-1. Ce mutant A-Fra-1 (pour acide Fra-1) peut être fusionné avec la forme mutée inactive de Cas9 : dCas9. Le site de fixation sur la chromatine devient alors dépendant de l'ARN guide choisi. Ainsi, dans des conditions où l'on réprime Fra-1 par RNAi, on peut mesurer l'effet de la fixation de Fra-1 sur un enhancer spécifique via la liaison dirigée par ARN guide de A-Fra-1-dCas9.

L'ensemble des résultats acquis et des travaux proposés devrait lever le voile sur les mécanismes mis en jeu par Fra-1, et peut être aussi par Fra-2, pour réguler l'expression de HMGA1 dans les cellules MDA-MB231. Les principales observations seront étendues à d'autres lignées triple négatives telles que les MDA-MB436 et les SUM-159PT afin de déterminer si les mécanismes identifiés ont valeur générale dans le contexte du cancer du sein triple négatif. Fra-1 et Fra-2 étant surexprimés dans de nombreux cancers épithéliaux, il serait aussi très intéressant de déterminer par la suite si ces mécanismes peuvent être étendus à d'autres types de cancers.

RÉSULTATS

Étude III

Étude III :

Rôle de CD68 dans le phénotype tumoral des TNBC

Lors de mon analyse transcriptomique il est apparu que l'un des gènes fortement régulés positivement par Fra-1 et/ou Fra-2 dans les cellules MDA-MB231 est CD68. De plus, son expression est également régulée par Fra-1 et/ou Fra-2 dans les deux autres lignées TNBC testées, suggérant que la régulation de CD68 par Fra-1 et/ou Fra-2 pourrait être un mécanisme commun aux cancers du sein triple négatifs.

CD68 est considéré dans la littérature comme un marqueur des macrophages. Cependant, quelques études insistent sur le fait que CD68 est exprimé dans d'autres types cellulaires tels que les fibroblastes et les cellules cancéreuses, ainsi que dans des tumeurs humaines telles que les gliomes ou les carcinomes fibrolamellaires hépatiques (Gottfried et al., 2008a; Travaglione et al., 2002). CD68 est une protéine transmembranaire décrite pour être localisée principalement dans les lysosomes et les endosomes. Sa fonction n'est pas caractérisée et son rôle potentiel dans la tumorigenèse n'a jamais été étudié. J'ai donc cherché à déterminer, dans un premier temps, si l'expression de CD68 est impliquée dans le phénotype tumoral des TNBC au niveau de la prolifération, de la motilité et de l'invasion dans des tests en 2 et 3 dimensions.

Comme nous le verrons ci-après, mes investigations n'ont pas permis de mettre en évidence un rôle de CD68 dans les fonctions testées. J'ai donc amorcé sur la fin de ma thèse des études pour déterminer si CD68 pouvait avoir un rôle dans l'autophagie du fait de sa localisation. Cette étude n'est pas finalisée mais les premiers résultats sont encourageants.

La description des résultats obtenus est précédée d'une introduction sur CD68. Des notions sur l'autophagie seront abordées avant la deuxième partie des résultats ayant trait à l'implication potentielle de CD68 dans le processus autophagique.

I. CD68 : Introduction

I.1. Structure et expression de CD68

Chez l'Homme, le gène codant pour CD68 se situe sur le locus chromosomique 17p13 et comprend six exons répartis sur une distance d'environ 3kb. L'épissage de l'ARNm CD68 conduit à deux isoformes protéiques : CD68a (354 acides aminés ; 37,4 KDa) et CD68b (327 acides aminés ; 17-43 sont manquants comparativement à CD68a). Bien que cela reste à déterminer, il a été suggéré que les différentes isoformes de CD68 puissent avoir des profils d'expression préférentiels : intracellulaire dans les lysosomes ou extracellulaire à la surface des cellules (Holness et al., 1993).

CD68, aussi connu sous le nom de macrosialin chez la souris et LAMP-4 chez l'Homme, est une protéine transmembranaire fortement glycosylée de 95-110 kDa (**Figure 62**). Elle a été décrite pour la première fois comme un antigène reconnu par des anticorps monoclonaux dirigés contre un épitope de macrophage (Micklem et al., 1989; Smith and Koch, 1987). De part son homologie avec les protéines LAMP-1 et LAMP-2 (lysosome-associated membrane glycoprotein), CD68 a été associée à la famille LAMP. LAMP-1 et LAMP-2 sont décrites pour être spécifiques des lysosomes (Holness et al., 1993) et nécessaires pour la fusion des lysosomes avec les phagosomes (Huynh et al., 2007). De façon similaire aux membres de la famille LAMP, les molécules CD68 sont localisées principalement au niveau des membranes lysosomales et des endosomes tardifs (Kreipe et al., 1987; Parwaresch et al., 1986; Pulford et al., 1990; Saito et al., 1991; Tsang and Chan, 1992).

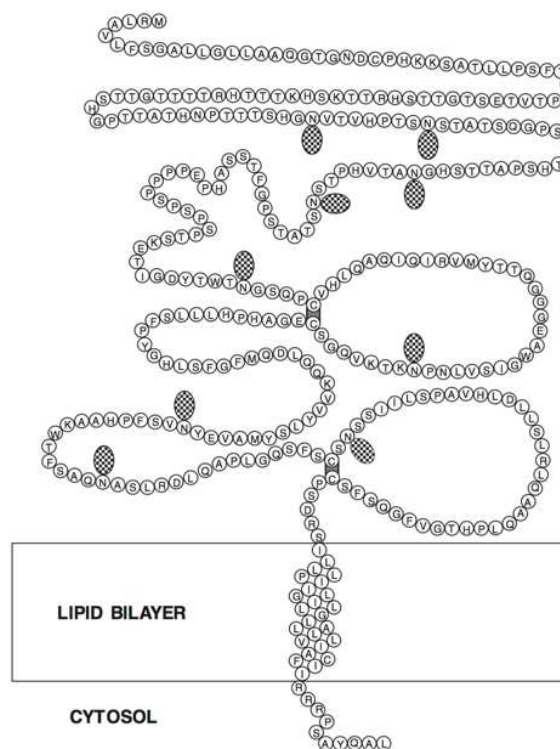


Figure 62 : Structure schématique de la protéine CD68 humaine. La protéine CD68 est caractérisée par une courte région C-terminale situé dans la partie cytoplasmique, un domaine transmembranaire et une longue région N-terminale contenant neuf sites potentiels de glycosylation (rond gris). *Source : (Kostich et al., 2000).*

En 1989, six anticorps monoclonaux dirigés contre la protéine CD68 humaine ont été générés en utilisant des macrophages issus du poumon et de ganglions lymphatiques (Y2/131, EBM11, Ki-M6, Ki-M7, KP1 et Y1/82A) (Holness and Simmons, 1993; Micklem et al., 1989). CD68 a été mis en évidence comme fortement exprimée dans les granules cytoplasmiques, ainsi qu'à la surface des macrophages, des monocytes, des neutrophiles, des basophiles et des cellules NK (Hameed et al., 1994). Du fait de son expression particulièrement élevée dans les macrophages, CD68 a été généralement considérée comme un marqueur des macrophages.

Cependant, quelques articles ont rapporté que CD68 est exprimée à des niveaux proches de celui des macrophages dans d'autres types cellulaires tels que les fibroblastes, les cellules musculaires lisses et les cellules endothéliales. De plus, de nombreuses lignées cancéreuses ainsi que des cancers humains tels que les carcinomes pancréatiques, les gliomes ou les carcinomes fibrolamellaires du foie, expriment de hauts niveaux de CD68 (Beranek, 2005; Gottfried et al., 2008b; Kunisch et al., 2004; Ross et al., 2011; Strojnik et al., 2009).

Malgré ces observations, CD68 est encore utilisé comme marqueur des macrophages dans de très nombreuses publications. En effet, sur plus de 8000 références répondant au mot clé CD68, la très grande majorité des travaux ne concernent pas l'étude de la fonction de CD68 mais l'utilise comme marqueur pour identifier les macrophages. Pourtant, Gottfried *et al* soulignaient déjà en 2008 que l'utilisation de CD68 comme marqueur des macrophages était critiquable en particulier à but diagnostique (Gottfried et al., 2008b). De plus, en 2005, Baranek proposait que CD68 ne soit plus considéré comme un antigène spécifique des macrophages mais plutôt comme un antigène caractéristique des cellules capables de phagocytose du fait de sa localisation endosomique et lysosomale (Beranek, 2005).

I.2. Fonction de CD68

La fonction de la protéine CD68 n'est pour l'heure pas connue. Du fait de sa localisation à la fois intra-cellulaire et à la membrane plasmique, elle a été proposée pour jouer un rôle dans divers processus cellulaires tels que l'interaction avec les lipoprotéines de faible densité oxydées (oxLDL) et le métabolisme lipidique, l'interaction cellule-pathogène, la phagocytose, le métabolisme lysosomal et la présentation des antigènes (De Beer et al., 2003; Holness and Simmons, 1993; Kurushima et al., 2000; Pulford et al., 1990; Ramprasad et al., 1996). Cependant, la délétion de CD68 chez la souris ne prévient pas, ni la capacité des macrophages à capturer le oxLDL, ni leur capacité à phagocyter un pathogène et aucun défaut dans la réponse immunitaire innée suite au contact avec des pathogènes n'a été observé (Song et al., 2011).

Ashley *et al* ont établi que les souris délétées de CD68 ont des ostéoclastes (lignée myéloïde) dont la fonction de résorption osseuse est diminuée. La microarchitecture des os de ces souris montre une augmentation du volume osseux trabéculaire avec une diminution de l'espace intra-trabéculaire et une anomalie de la minéralisation. Les ostéoclastes CD68^{-/-} présentent une accumulation de structures vésiculaires intracellulaires non caractérisées qui sont plus rarement présentes dans les ostéoclastes CD68^{+/+}. Les ostéoclastes CD68^{-/-} sont également plus sensibles aux forces de détachement et moins adhérents à leur substrat. Ces résultats montrent l'importance de CD68 dans un type cellulaire autre que les macrophages pour lesquels le rôle de CD68 n'est toujours pas caractérisé (Ashley et al., 2011).

De façon intéressante, la délétion de c-Fos dans la souris conduit à une ostéopétrose (densification de l'os) liée à une diminution de l'activité de résorption des ostéoclastes (Grigoriadis et al., 1994; Wang et al., 1992). *In vitro*, la transfection d'un transgène codant pour Fra-1 dans des précurseurs ostéoclastiques issus de souris c-Fos^{-/-} est plus efficace que celle des autres membres FOS pour compléter l'absence de c-Fos. Il est à noter que le facteur de différenciation ostéoclastique RANKL, qui induit l'expression de c-Fos dans les ostéoclastes, induit aussi celle de Fra-1 (Matsuo et al., 2000). De plus, chez la souris une restauration complète de la différenciation ostéoclastique est obtenue par knock-in du gène Fra-1 dans le locus c-Fos des souris c-Fos^{-/-} (Fleischmann, 2000), mettant en évidence la capacité de Fra-1 à réguler la différenciation ostéoclastique.

L'ensemble de ces résultats indique que c-Fos, Fra-1 et CD68 sont nécessaires au fonctionnement correct des ostéoclastes. Au vue de nos résultats transcriptomiques dans les cellules MDA-MB231 et des résultats de la littérature présentés ci-dessus, il serait intéressant de déterminer si Fra-1 régule également l'expression de CD68 dans les ostéoclastes.

Comme mentionné dans cette introduction sur CD68, le rôle moléculaire et les fonctions de CD68 ne sont toujours pas clairement caractérisés. Par ailleurs, bien que CD68 soit surexprimé dans certaines tumeurs, son rôle dans la tumorigenèse n'a jamais été abordé. Ainsi, dans l'étude présentée ci-dessous, j'ai principalement cherché à déterminer quelle pourrait être la contribution de CD68 dans le phénotype tumoral des cancers du sein triple négatifs.

II. CD68 : Résultats

II.1. Fra-1 et Fra-2 régulent la transcription du gène CD68

Dans un premier temps, j'ai vérifié que le rôle de Fra-1 et/ou Fra-2 était d'ordre transcriptionnel et non indirect en modulant l'expression de microARN affectant la stabilité et/ou la traductibilité de l'ARNm de CD68. Pour cela, j'ai mesuré l'abondance des ARN naissants CD68, par la technique de « Nuclear Run-On assay » déjà mentionnée dans l'étude II (paragraphe II.1). Les résultats présentés dans la **Figure 63** montrent que l'abondance des ARN naissants CD68 est réduite d'un facteur supérieur ou égal à 2 en absence de Fra-1 et/ou Fra-2, indiquant l'implication de Fra-1 et Fra-2 au niveau de la transcription du gène CD68.

Il est à noter que nos résultats par RT-qPCR ont montré que l'absence de Fra-2 a un effet faible et non significatif sur l'abondance à l'équilibre des l'ARNm CD68 dans la lignée MDA-MB231 (**Figure 46** de l'étude I, paragraphe IV.2). Comme dans le cas de HMGA1, cette différence pourrait être liée aux conditions expérimentales différentes (voir étude II, paragraphe II.1).

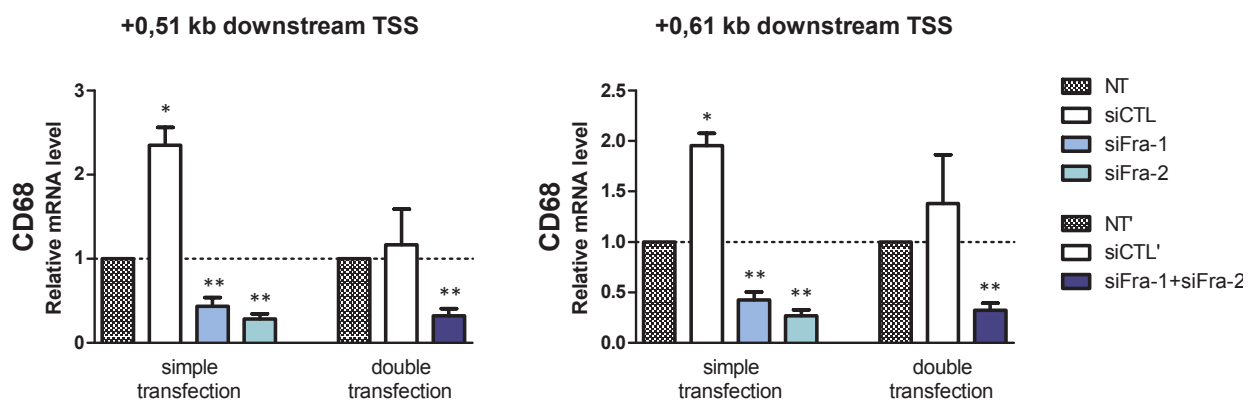


Figure 63 : Abondances relatives des ARN naissants CD68 en absence de Fra-1 et/ou Fra-2 dans les MDA-MB231. Abondance relative des ARN naissants CD68 analysée par RT-qPCR avec deux couples d'amorces permettant l'amplification des régions situées à +0.51 et +0.61 kb en aval du TSS de CD68. Les valeurs sont normalisées par rapport à l'ARN naissant GAPDH et relatives à l'abondance mesurée dans la condition non-transfectée (NT ou NT'). n=3. Two tailed paired t-test ; (**) p value <0.01.

Fra-1 et Fra-2 étant impliqués au niveau de la transcription du gène CD68, j'ai cherché à caractériser le ballet des co-activateurs transcriptionnels recrutés par Fra-1 et/ou Fra-2 pour activer l'expression de CD68. Pour cela j'ai utilisé une approche par ChIP-qPCR similaire à celle présentée dans l'étude II pour le gene HMGA1.

La cartographie des sites de fixation de Fra-1 et Fra-2 sur le génome des cellules MDA-MB231 par ChIP-seq a révélé que le gène CD68 possède des petits pics de fixation pour Fra-1 et Fra-2 sur des motifs AP-1/TRE, au niveau du TSS et en amont de sa région promotrice, ainsi qu'au niveau intragénique. Concernant les régions enhancer potentielles de CD68, nous n'avons

pas observé de pic de fixation pour Fra-1 ou Fra-2 dans les trois régions enhancers décrites comme associées à CD68 par le consortium FANTOM5, à savoir -485kb, -446.6kb et -103.6kb du TSS. Ceci suggère que CD68 pourrait être régulé directement via la fixation de Fra-1 et/ou Fra-2 au niveau de son promoteur. Cependant, on ne peut pas écarter la possibilité que d'autres enhancers Fra-1/Fra-2 dépendants soient non répertoriés dans FANTOM5.

Malheureusement, mes investigations concernant les mécanismes transcriptionnels impliqués dans la régulation de CD68 se sont avérées délicates, en raison de la petite taille du gène CD68 (2.7kb) et de la difficulté à trouver des amplicons efficaces en qPCR qui soient repartis uniformément sur le locus d'intérêt. De plus, l'environnement chromatinien autour de CD68 est riche en gènes transcriptionnellement actifs, avec EIF4A1 situé à 480 pb en amont du TSS de CD68 et MPDU1 situé à 1535 pb en aval de la fin du gène CD68. Ainsi, mis à part une diminution partielle des marques H3K4me3 et H3K27ac sur le corps du gène, en absence conjointe de Fra-1 et Fra-2, les autres résultats obtenus ne permettent pas une interprétation claire (données non montées). Ainsi, dans la suite de l'étude III je me suis focalisée sur l'étude d'un rôle potentiel de CD68 dans la tumorigenèse.

II.2. Modulation de l'expression de CD68 *in vitro*

II.2.a. Analyse de l'abondance de CD68 dans les 3 lignées TNBC versus la lignée MCF7

J'ai évalué l'abondance relative de CD68 dans les trois lignées triple négatives utilisées dans le cadre de ma thèse en comparaison à la lignée ER+ de référence MCF7 (**Figure 64**). Il apparaît que CD68 est fortement exprimé dans les lignées MDA-MB231 et MDA-MB436, aussi bien au niveau ARNm que protéine. Le niveau d'expression est quasi indétectable dans les MCF7 et est relativement faible dans les SUM-159PT. Pour la suite de mes travaux j'ai donc préférentiellement utilisé les lignées MDA-MB231 et MDA-MB436.

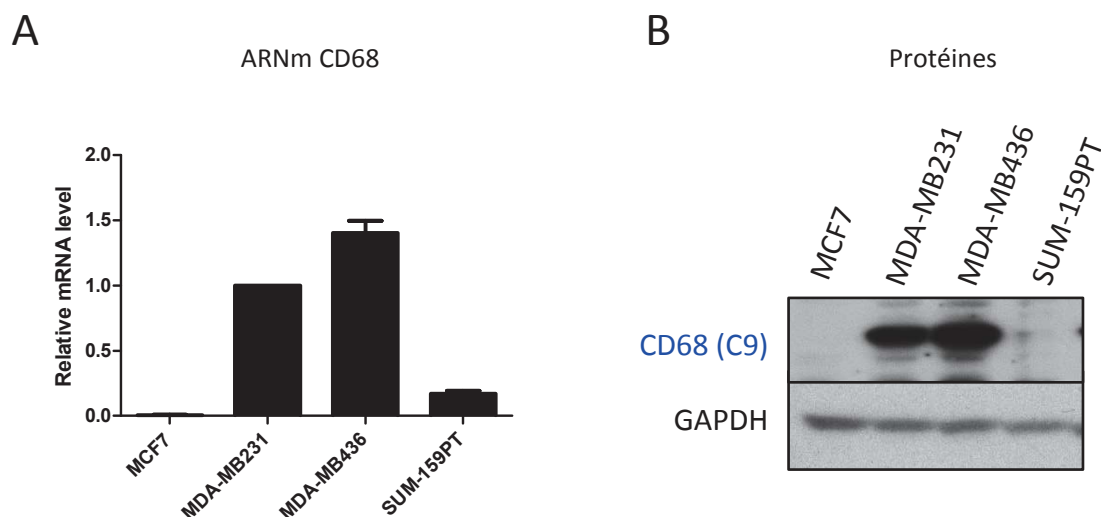


Figure 64 : Abondance de CD68 dans les lignées de TNBC MDA-MB231, MDA-MB436 et SUM-159PT. A) Abondance relative des ARNm analysée par RT-qPCR. Les valeurs sont normalisées par rapport à l'abondance de l'ARNm CD68 dans les MDA-MB231. n=3. *Analyse faite en utilisant deux couples d'amorces par gène pour valider les résultats. Résultats pour le 2nd couple d'amorces non montrés.* **B)** Abondance de la protéine CD68 dans les différentes lignées analysée par immunoblot avec un anticorps anti-CD68 (C9-sc-393951-Santa Cruz Biotechnology), l'immunoblot GAPDH est présenté en témoin de charge.

II.2.b. Mesure de l'inhibition de l'expression CD68 par transfection de siRNA

Les études fonctionnelles sont réalisées en comparant différentes propriétés cellulaires connues pour favoriser l'agressivité tumorale en présence et en absence de CD68 invalidé phénotypiquement par approche siRNA.

Pour éteindre l'expression de CD68, j'ai utilisé un pool commercial de 3 siRNA dans les mêmes conditions de transfection que celles utilisées pour l'extinction de Fra-1 et Fra-2, à savoir 1,5nM de chaque siRNA (soit 4,5nM au total) pendant 72h. Les conditions non-transfected (NT) et siRNA contrôle (siCTL) sont les mêmes que celles présentées aux chapitres précédents.

Les résultats présentés sur la **Figure 65** montrent que le pool siCD68 utilisé est très efficace. En effet, 72h après transfection des siCD68, on observe une diminution des ARNm CD68 supérieure à 90% dans les MDA-MB231 et MDA-MB436. De la même manière on observe une disparition de la protéine par immunoblot dans les deux lignées cellulaires.

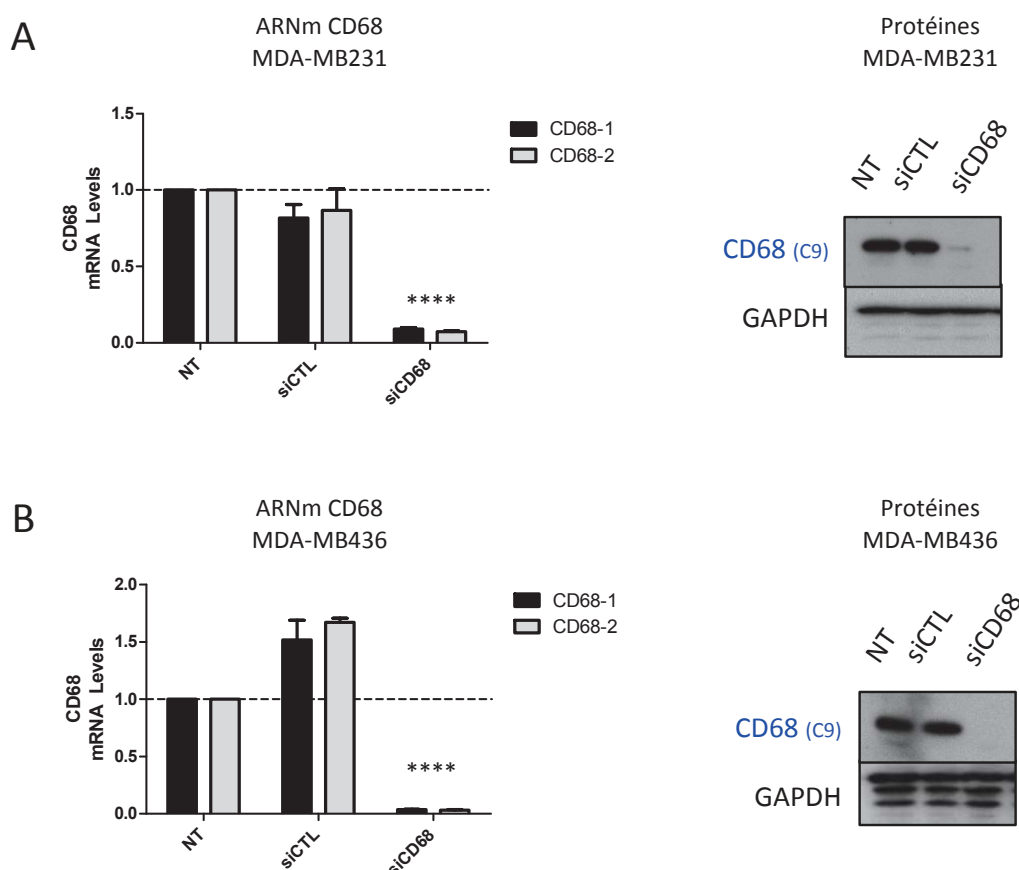


Figure 65 : Inhibition de l'expression de CD68 dans les lignées de TNBC MDA-MB231 et MDA-MB436. Les cellules MDA-MB231 (A) et MDA-MB436 (B) sont transfectées pendant 72h avec le siRNA contrôle ou le pool spécifique siCD68. Panel de gauche : abondance relative des ARNm analysée par RT-qPCR avec deux couples d'amorces (CD68-1 et CD68-2). Les valeurs sont normalisées par rapport à l'ARN S26 et relatives à l'abondance de l'ARNm CD68 dans cellules non-transfectées (NT) mais ayant reçu l'agent de transfection. Panel de droite : le niveau d'expression de la protéine CD68 est analysé par immunoblot, GAPDH est présenté en témoin de charge.

II.3. Impact de l'absence de CD68 sur la prolifération, la viabilité et le cycle cellulaire dans la lignée MDA-MB231

Les processus de prolifération cellulaire, de résistance à l'apoptose et de cycle cellulaire sont fréquemment testés en première approche pour caractériser les effets potentiels d'une protéine sur le phénotype tumoral.

II.3.a. Impact de l'absence de CD68 sur la prolifération

Pour analyser la prolifération cellulaire, j'ai eu recours au test colorimétrique « MTT assay». Les cellules MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 et siFra-1, puis incubées après 2, 4, 6 et 8 jours de transfection avec la solution de MTT avant d'être analysées au spectrophotomètre (lecteur de plaque TECAN). La condition siFra-1 a été rajoutée comme contrôle positif. En effet, de nombreuses études ont montré que Fra-1 favorise la prolifération cellulaire dans différents cancers épithéliaux (Belguise et al., 2005; Casalino et al., 2007; Sayan et al., 2012).

Le principe du test MTT est basé sur la réduction du Thiazolyl Blue Tetrazolium Bromide (MTT) par les déshydrogénases mitochondriales dépendantes du NADH ou du NADPH, dans les cellules vivantes actives. Le formazan issu de cette réaction, forme un précipité dans la mitochondrie de couleur violette. Les cellules mortes ne produisent pas de cristaux de formazan. L'intensité de la coloration violette est donc fonction de la proportion de cellules vivantes. Les résultats présentés dans la **Figure 66**, indiquent que l'absence de CD68 n'a pas d'impact notable sur la prolifération. En effet, la courbe de prolifération dans la condition siCD68 se situe entre les courbes de prolifération des conditions contrôles (NT et siCTL). En revanche, en accord avec les données de la littérature, les cellules transfectées avec le siFra-1 ont une prolifération fortement ralentie. Il est à noter également que la transfection du siCTL commercial entraîne, pour une raison indéterminée, une diminution de la prolifération comparativement à la condition NT.

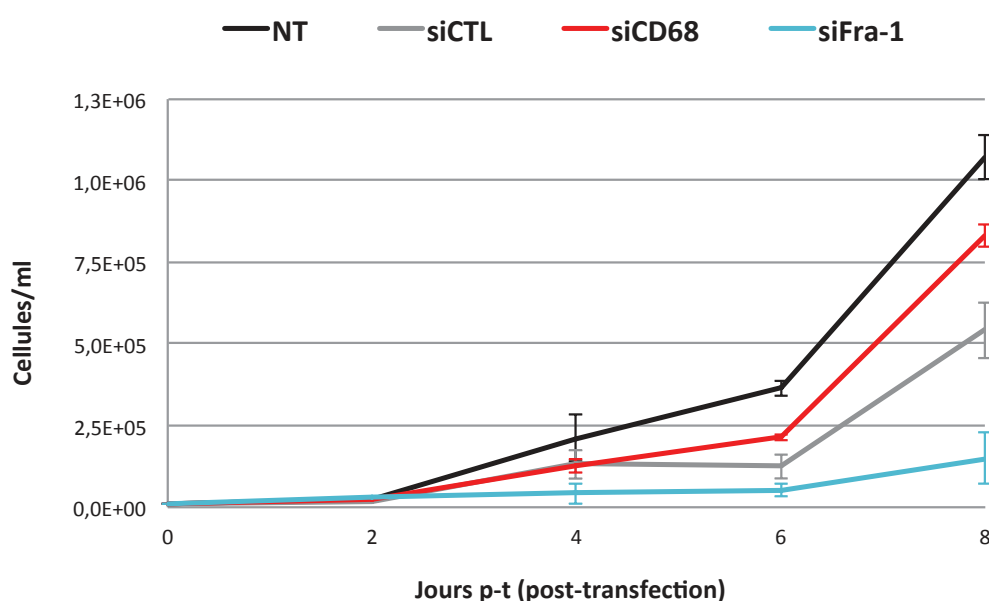


Figure 66 : Mesure de la prolifération des cellules MDA-MB231 après inhibition de CD68 ou Fra-1. n=3. Les cellules MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 et siFra-1, puis incubées après 2, 4, 6 et 8 jours de transfection avec la solution de MTT avant d'être analysées au spectrophotomètre (lecteur de plaque TECAN).

Par ailleurs, lors de l'utilisation du MTT assay, la quantité de précipité formé est proportionnelle à la quantité de cellules vivantes mais également à l'activité métabolique de chaque cellule. Pour m'affranchir d'une éventuelle différence de métabolisme, j'ai également mesuré la prolifération cellulaire en comptant directement les cellules à l'aide du compteur automatique de cellules EVE (NanoEnTek) après 2, 4, 6 et 8 jours de transfection. Cette technique permet également d'écarter les cellules mortes. Les courbes de prolifération sont similaires (résultats non montrés) à celles obtenues par la méthode MTT.

II.3.b. Impact de l'absence de CD68 sur la viabilité

Dans le but de détecter les cellules apoptotiques et nécrotiques, les MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 et siFra-1, puis marquées après 4 et 6 jours de transfection avec la combinaison DiOC6(3)/Pi avant d'être analysées par cytométrie de flux.

L'iodure de propidium (Pi) est un agent intercalant des acides nucléiques qui marque le noyau des cellules ayant perdu leur intégrité membranaire, caractéristique de la nécrose cellulaire. Le 3,3'-Dihexyloxacarbocyanine Iodide (DiOC6(3)) est un marqueur qui permet de mettre en évidence les variations du potentiel de membrane des mitochondries. L'intensité de fluorescence du DiOC6(3) est fonction de la proportion de mitochondries dont la membrane est polarisée. La dépolarisation des membranes mitochondriales, qui survient précocement au cours du processus apoptotique, entraîne un décalage de la fluorescence vers les petites intensités. Ainsi la combinaison DiOC6(3)/Pi permet de différencier les cellules vivantes (DiOC6(3)+/Pi-), des cellules nécrotiques (DiOC6(3)+/Pi+) et apoptotiques (apoptose précoce : DiOC6(3)-/Pi- ; apoptose tardive : DiOC6(3)-/Pi+) (**Figure 67**).

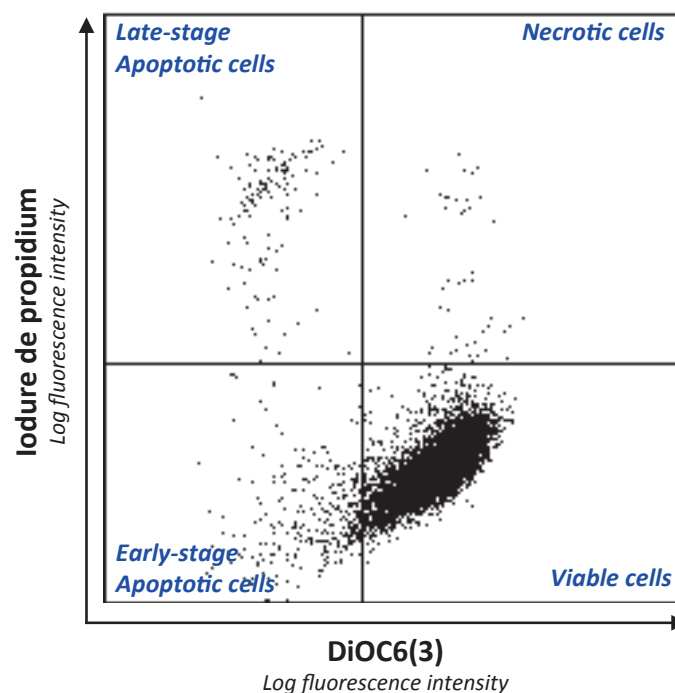


Figure 67 : Mesure de la viabilité cellulaire des cellules MDA-MB231 non-transfectées par cytométrie de flux au moyen du marquage DiOC6(3)/Pi. La combinaison DiOC6(3)/Pi permet de différencier les cellules vivantes (DiOC6(3)+/Pi-), des cellules nécrotiques (DiOC6(3)+/Pi+) et apoptotiques (apoptose précoce : DiOC6(3)-/Pi- ; apoptose tardive : DiOC6(3)-/Pi+).

Les résultats présentés dans la **Figure 68**, indiquent que l'absence de CD68 n'induit pas significativement d'apoptose, ni de nécrose dans les cellules MDA-MB231. En effet, 80-90% des cellules sont viables dans la condition siCD68 aux temps analysés, ce qui est comparable aux conditions contrôles (NT et siCTL). En revanche, les cellules transfectées avec le siFra-1 ont une

viabilité beaucoup plus faible au jour 6 post-transfection clairement attribuable à une augmentation de l'apoptose : 45% de cellules viables, 50% de mort par apoptose et 5% de mort par nécrose. Les données concernant Fra-1 indiquent donc que, dans les MDA-MB231, Fra-1 a une fonction anti-apoptotique. Un effet anti-apoptotique de Fra-1 a également été rapporté récemment dans les cellules H460 de cancer du poumon (Zhong et al., 2016).

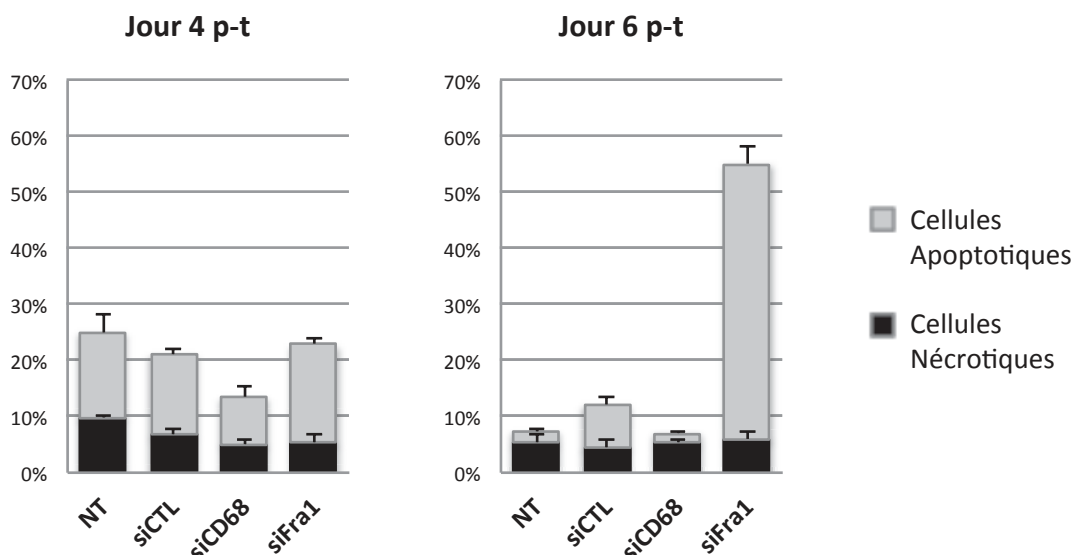


Figure 68 : Pourcentage de cellules MDA-MB231 apoptotiques et nécrotiques mesuré par cytométrie de flux au moyen du marquage DiOC6(3)/Pi, après inhibition de CD68 ou Fra-1 par siRNA. Les données présentées correspondent à la moyenne (et écart type) de 3 expériences indépendantes 4 jours et 6 jours post-transfection.

II.3.c. Impact de l'absence de CD68 sur le cycle cellulaire

Pour quantifier la distribution des cellules dans le cycle cellulaire, les MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 ou siFra-1, puis fixées à l'éthanol 70% et marquées à l'iodure de propidium (Pi) après 4 et 6 jours de transfection, avant d'être analysées par cytométrie de flux. L'éthanol a pour effet de fixer les noyaux et de détruire partiellement les membranes plasmiques, ce qui permet au Pi de s'insérer dans les acides nucléiques. Ainsi, la quantité de Pi fixé est proportionnelle à la quantité d'ADN cellulaire (**Figure 69**).

La majorité des cellules sont en G1, car c'est l'étape la plus longue du cycle cellulaire. En G1 les cellules présentent une seule copie de leur patrimoine génétique et forment un pic bien distinct. Les cellules en phase S de réplication ont une quantité variable d'ADN, ce qui explique leur distribution étalée. En phase G2/M, les cellules ont doublé leurs chromosomes, ce qui se traduit par la formation d'un second pic au double de la valeur de G1. En général, les cellules donnent une distribution G1 / S / G2/M autour de 60% / 20% / 20%. La phase SubG1 quant à elle est représentative du degré de mort cellulaire et correspond à la dégradation/fragmentation de l'ADN des cellules en fin d'apoptose ou en nécrose.

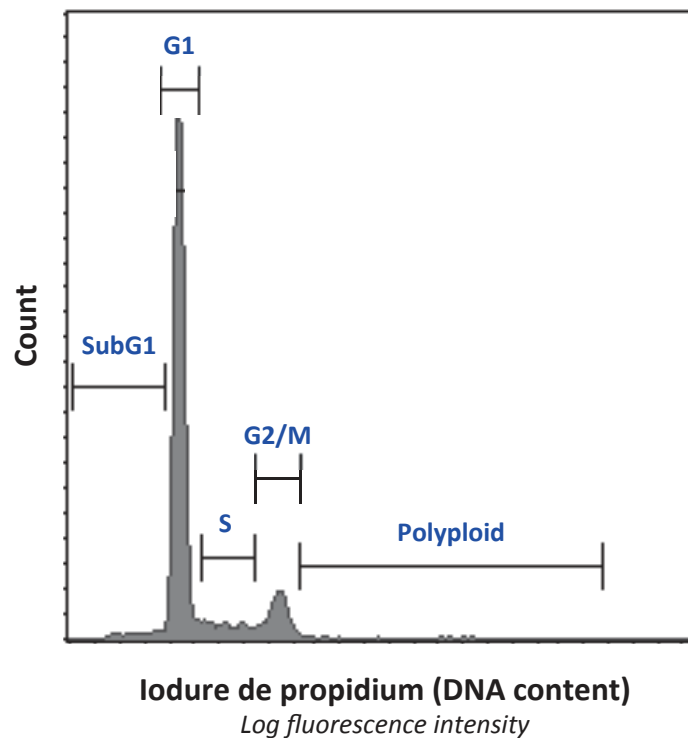


Figure 69 : Mesure de la distribution dans le cycle cellulaire des cellules MDA-MB231 non-transfectées par cytométrie de flux au moyen du marquage Pi.

Les résultats présentés dans la **Figure 70**, indiquent que l'absence de CD68 ne modifie pas de façon significative la progression des cellules MDA-MB231 dans le cycle cellulaire. En effet, il n'y a aucune différence entre la condition siCD68 et les conditions contrôles (NT et siCTL) au jour 4 post-transfection. Au jour 6 post-transfection on observe cependant une légère augmentation (10%) des cellules en phase G1 dans la condition siCD68.

En ce qui concerne la condition siFra-1, l'absence de Fra-1 ne modifie pas non plus de façon drastique la distribution des cellules dans le cycle cellulaire. Néanmoins, on observe une plus grande proportion de cellules en phase SubG1. Cette observation est en accord avec les résultats présentés précédemment, à savoir une viabilité plus faible des cellules MDA-MB231 en absence de Fra-1 (**Figure 68**).

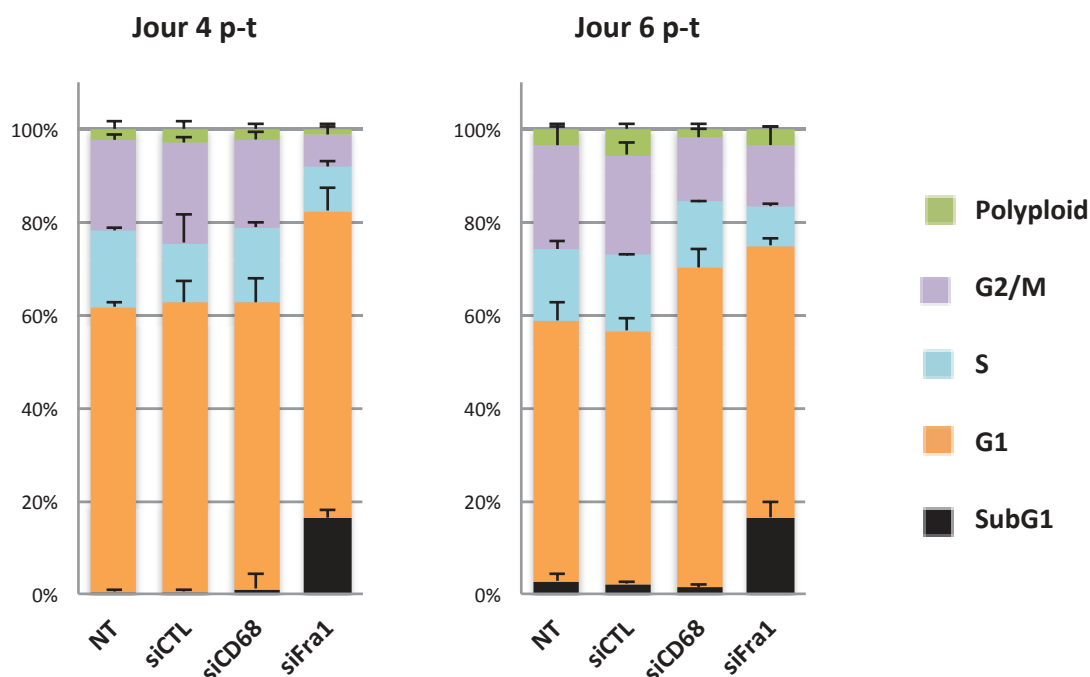


Figure 70 : Distribution dans le cycle cellulaire des cellules MDA-MB231 par cytométrie de flux au moyen du marquage Pi, après inhibition de CD68 ou Fra-1 par siRNA. Les données présentées correspondent à la moyenne (et écart type) de 3 expériences indépendantes 4 jours et 6 jours post-transfection.

II.4. Impact de l'absence de CD68 sur la motilité, l'invasion et la capacité à former des mammosphères

Les cancers du sein triple négatifs sont caractérisés par une forte propension à former des métastases et CD68 est une protéine transmembranaire généralement considérée comme un marqueur des macrophages, qui sont des cellules particulièrement mobiles. Du fait de sa forte glycosylation, quand elle est présente à la surface cellulaire CD68 peut interagir avec un certain nombre de lectines et sélectines qui permettent des phénomènes d'adhérences faibles mais de très haute spécificité. Ceci pourrait permettre aux cellules cancéreuses d'interagir avec les leucocytes, les plaquettes ou encore les cellules endothéliales des capillaires sanguins favorisant ainsi leur dissémination.

II.4.a. Impact de l'absence de CD68 sur la motilité en 2D

Pour analyser la motilité cellulaire, j'ai eu recours à une variante de la méthode dite de « wound-healing assay » associée à la vidéomicroscopie. Cette technique consiste à utiliser des inserts en silicone dans le but de créer une brèche calibrée et constante au milieu du tapis cellulaire.

Les cellules MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 et siFra-1 pendant 72h, puis un nombre identique de cellules a été mis en culture dans les inserts en silicone pour chaque condition de manière à obtenir un tapis cellulaire régulier avec une

confluence de 100% une fois que les cellules ont re-adhéré, soit 24h plus tard. Dès que les inserts sont retirés, l'analyse de la motilité est alors suivie sur 24h par vidéomicroscopie (**Figure 71**).

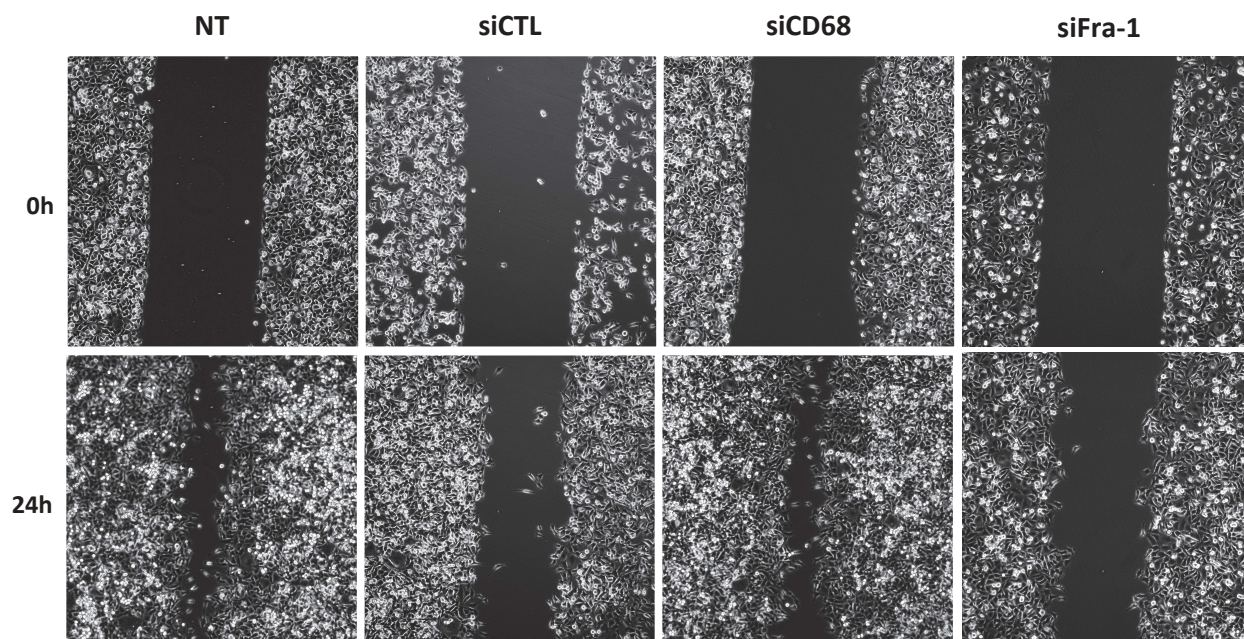


Figure 71 : Analyse de la motilité des cellules MDA-MB231 par la méthode de wound-healing associée à la vidéomicroscopie, après inhibition de CD68 ou Fra-1 par siRNA. Microscope Olympus objectif 10X. Deux expériences indépendantes avec minimum 2 inserts par condition.

Les résultats présentés dans la **Figure 71**, indiquent que l'absence de CD68 ne modifie pas de façon significative la motilité cellulaire des cellules MDA-MB231. En effet, il n'y a aucune différence entre la condition siCD68 et la condition contrôle NT : les cellules sont fortement mobiles et referment la brèche laissée par l'insert en silicone en 24h. De plus, les cellules MDA-MB231 transfectées avec le siCD68 gardent un profil mésenchymateux associé à une migration aléatoire. Ces résultats montrent que CD68 n'est pas impliqué dans la motilité des cellules MDA-MB231.

Cependant, les résultats de la condition contrôle siCTL posent questionnement (**Figure 71**). En effet, bien que les cellules aient été mises en culture dans les inserts à la même concentration que les autres conditions, les cellules transfectées avec le siCTL sont moins confluentes. De ce fait, les cellules comblent les espaces interstitiels avant de combler la brèche centrale. Le suivi de refermeture de la brèche sur 12h supplémentaires a montré que les cellules transfectées par le siCTL sont capables de fermer la bèche. Il se pourrait donc que la différence observée sur 24h soit liée uniquement à la différence de confluence au démarrage de l'expérience. Une nouvelle expérience avec un tapis cellulaire confluent au départ permettrait d'établir si le siCTL affecte ou non la motilité des cellules MDA-MB231.

En ce qui concerne la condition siFra-1, les résultats indiquent clairement que l'absence de Fra-1 diminue significativement la motilité cellulaire des cellules MDA-MB231 (**Figure 71**).

comme cela a déjà été rapporté dans la littérature pour différents types de cellules épithéliales (Adiseshaiah et al., 2007; Belguise et al., 2005; Hanson et al., 2016; Milde-Langosch et al., 2004; Ramos-Nino et al., 2007; Zhang et al., 2016a). En effet, les cellules transfectées avec le siFra-1 ne parviennent pas à refermer la brèche en 24h, ni même en 36h (non montré). De plus, on peut noter un changement de morphologie des cellules MDA-MB231 transfectées avec le siFra-1, qui sont passées d'un profil mésenchymateux à un profil plus épithélial associé à une migration plus cohésive.

II.4.b. Impact de l'absence de CD68 sur l'invasion et la capacité à former des mammosphères en 3 dimensions

Pour analyser l'invasion, j'ai eu recours à la technique des « mammosphères » suivi d'un test d'invasion tridimensionnelle en matrigel. Le matrigel permet de maintenir les cellules en culture en conditions non adhérentes.

Les MDA-MB231 ont été transfectées ou non avec les siCTL, siCD68 et siFra-1 pendant 72h, puis des quantités équivalentes de cellules ont été remises en culture dans une solution à base de matrigel en utilisant des plaques 96 puits de très faible adhérence. Après 24h de culture, les sphéroïdes formés sont prélevés et remis en culture individuellement dans une matrice à base de collagène I, qui constitue une matrice biologiquement active semblable à la matrice basale dans les tissus mammifères.

Les résultats présentés dans la **Figure 72**, indiquent que l'absence de CD68 ne modifie pas de façon significative la capacité des cellules MDA-MB231 à former des mammosphères ni leur capacité à envahir en 3 dimensions.

En ce qui concerne la condition siFra-1, les résultats indiquent clairement que l'absence de Fra-1 diminue significativement la capacité des cellules MDA-MB231 à envahir en 3 dimensions. De plus, on peut noter que la taille des sphéroïdes formés est légèrement plus faible en absence de Fra-1 (**Figure 72**). Ces résultats sont en accord avec les données de la littérature montrant un rôle de Fra-1 dans l'invasion et la métastatisation (Belguise et al., 2005; Desmet et al., 2013; Wu et al., 2015; Zhang et al., 2016b).

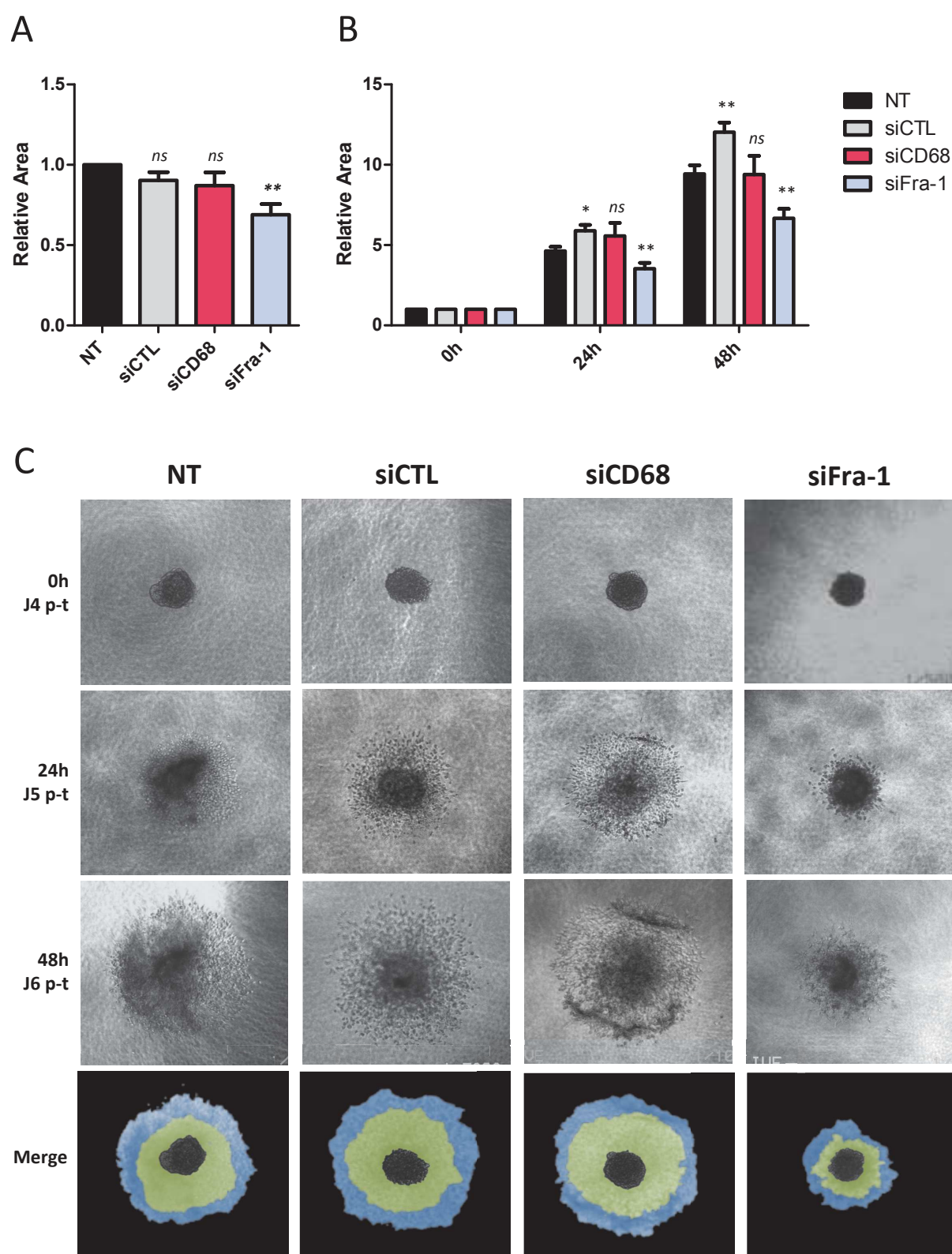


Figure 72 : Évolution de la morphologie en 3D des mammosphères au cours du temps, après inhibition de CD68 ou Fra-1 par siRNA dans les MDA-MB231. **A)** Quantification de l'aire de surface des sphéroïdes au jour 4 post-transfection. Les valeurs sont normalisées sur la condition non-transfectée (NT). **B)** Quantification de l'aire de surface des mammosphères au cours du temps. Les valeurs sont normalisées sur la taille des sphéroïdes au jour 4 post-transfection (0h) pour chaque condition. Two-tailed paired t-test ; (*) p value <0.05 (**) p value <0.01 ; (ns) non significatif. Les données présentées correspondent à deux expériences indépendantes avec minimum 5 mammosphères par condition. **C)** Visualisation de l'évolution de la morphologie en 3D des mammosphères au cours du temps par microscopie, après inhibition de CD68 ou Fra-1 par siRNA dans les MDA-MB231.

Cependant, il est important de prendre en considération le fait que dans les cellules MDA-MB231, l'extinction des protéines CD68 et Fra-1 est optimale après 72h de transfection (J3 p-t) avec des pools de trois siRNA (4,5nM final) et cela jusqu'à environ 120h (J5 p-t) (**Figure 73**). En effet, les protéines commencent à être à nouveau détectables, faiblement, au jour 5 post-transfection. Les abondances protéiques sont cependant encore relativement faibles au jour 6 et augmentent au jour 7, surtout pour CD68, mais sont encore en deçà des niveaux initiaux.

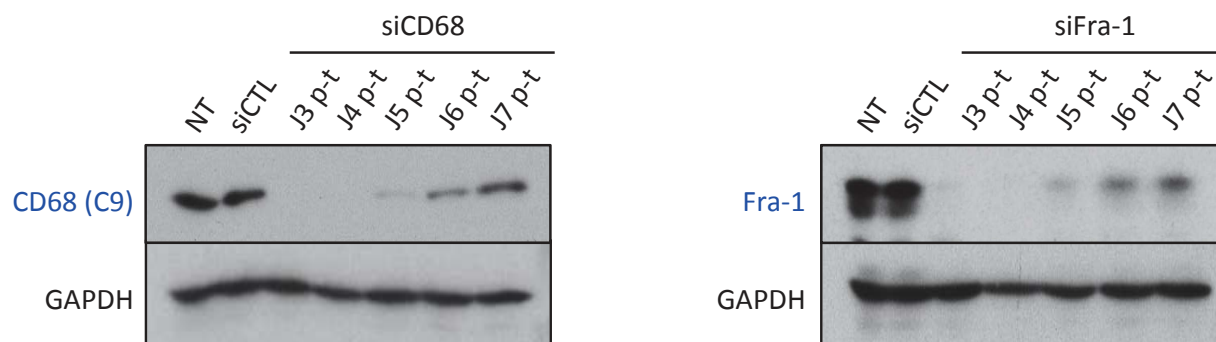


Figure 73 : Suivi de l'abondance des protéines CD68 et Fra-1 jusqu'à 7 jours post-transfection. L'extinction des protéines CD68 et Fra-1 est optimale entre J3 et J5 post-transfection.

Le protocole expérimental pour les études d'invasion en matrigel et mammosphères présentées ci-dessus nécessite une échelle de temps d'au moins 6 jours après la transfection. Cependant, la capacité des cellules à s'autorenouveler (capacité stem cell), en étant capables de constituer des mammosphères de seconde génération par passages successifs, n'a pas pu être investiguée dans ces conditions car elle nécessite une échelle de temps plus grande.

Dans le but de caractériser l'implication éventuelle de CD68 dans la capacité d'auto renouvellement j'ai opté pour une extinction stable de CD68 par transfection de shRNA. Pour choisir le ou les shRNA les plus adéquats, j'ai transfecté individuellement les 3 siRNA constitutifs du pool siCD68 utilisé pour les expériences décrites ci-dessus. Après avoir sélectionné le siRNA le plus efficace pour l'extinction complète de CD68 sur 72h, j'ai transfecté les cellules MDA-MD231 avec le shRNA correspondant et j'ai procédé à la sélection de clones.

De façon inattendue et très décevante, aucun des 12 clones générés, ni même la population de clones, ne présentaient une diminution de CD68 au niveau protéine et ARN. Le niveau de CD68 était équivalent à celui présent dans les clones exprimant un shRNA contrôle. Dans les mêmes conditions de sélection, les cellules non-transfectées étaient mortes en quelques jours. Une nouvelle tentative en utilisant des concentrations de drogue (puromycine) plus élevées pour la sélection des clones n'a pas été plus fructueuse. Une hypothèse pouvant rendre compte de cet échec pourrait être une contre sélection des cellules dans lesquelles CD68 est inactivé sur le long terme.

Les résultats acquis au cours de cette étude ont néanmoins montré que CD68, dont l'expression est fortement stimulée par Fra-1, n'est impliqué ni dans la motilité ni dans l'invasion des cellules MDA-MB231.

II.5. Rôle potentiel de CD68 dans l'autophagie ?

CD68 étant décrit pour avoir une forte homologie avec la famille LAMP (lysosome-associated membrane glycoprotein), il était intéressant de déterminer si CD68 pouvait avoir un rôle dans la phagocytose, l'endocytose ou l'autophagie. En fin de thèse, je me suis intéressée au rôle potentiel de CD68 dans l'autophagie.

II.5.a. Localisation de CD68 dans la lignée MDA-MB231

Pour analyser la localisation cellulaire de CD68, j'ai eu recours à l'immunofluorescence sur cellules fixées. Lors de cette étude, il est apparu que l'anticorps C9 (sc-393951 Santa Cruz Biotechnology) utilisé en immunoblot dans le cadre du travail présenté ci-dessus ne permet pas de distinguer un signal suffisamment spécifique en immunofluorescence. J'ai donc eu recours à l'utilisation de deux anticorps monoclonaux largement utilisés dans la littérature pour le marquage immunohistochimique de CD68 (Holness and Simmons, 1993; Kelly et al., 1988; Kunisch et al., 2004; Micklem et al., 1989). Il s'agit des clones EBM11 (M0718-DAKO) et KP1 (M0814-DAKO).

Si les deux anticorps révèlent un signal vésiculaire intra-cytoplasmique dans les conditions non-transfectée (NT) et transfectée avec le siCTL, ils donnent des résultats différents dans la condition transfectée avec le siCD68 (**Figure 74.A**). En effet, alors que le signal a complètement disparu avec l'anticorps EBM11 comme attendu, du signal est encore présent lors de la détection avec l'anticorps KP1. Ceci suggère que l'anticorps KP1 n'est pas complètement spécifique de CD68. Au vu de nos résultats, il est vraisemblable que l'anticorps KP1 soit susceptible de reconnaître d'autres protéines, sans doute les protéines de la famille LAMP en raison de la localisation du signal. Cette observation pose la question de la spécificité de l'anticorps KP1 anti-CD68 utilisé en immunohistochimie dans de nombreuses études de la littérature.

J'ai également confirmé ces résultats par cytométrie de flux (**Figure 74.B**). En effet, comparé aux conditions contrôles (NT et siCTL) la condition transfectée avec le siCD68 entraîne une diminution presque complète de l'intensité du signal émis par l'anticorps EBM11. La détection avec l'anticorps KP1 ne montre pas de diminution d'intensité du signal, ce qui traduit un manque de spécificité pour CD68. Concernant l'anticorps C9, comparé aux conditions contrôles la condition transfectée avec le siCD68 entraîne une diminution de l'intensité du signal, en accord avec la détection par immunoblot, cependant le signal positif est faible et proche du bruit de fond (marquage des cellules avec l'anticorps secondaire seul).

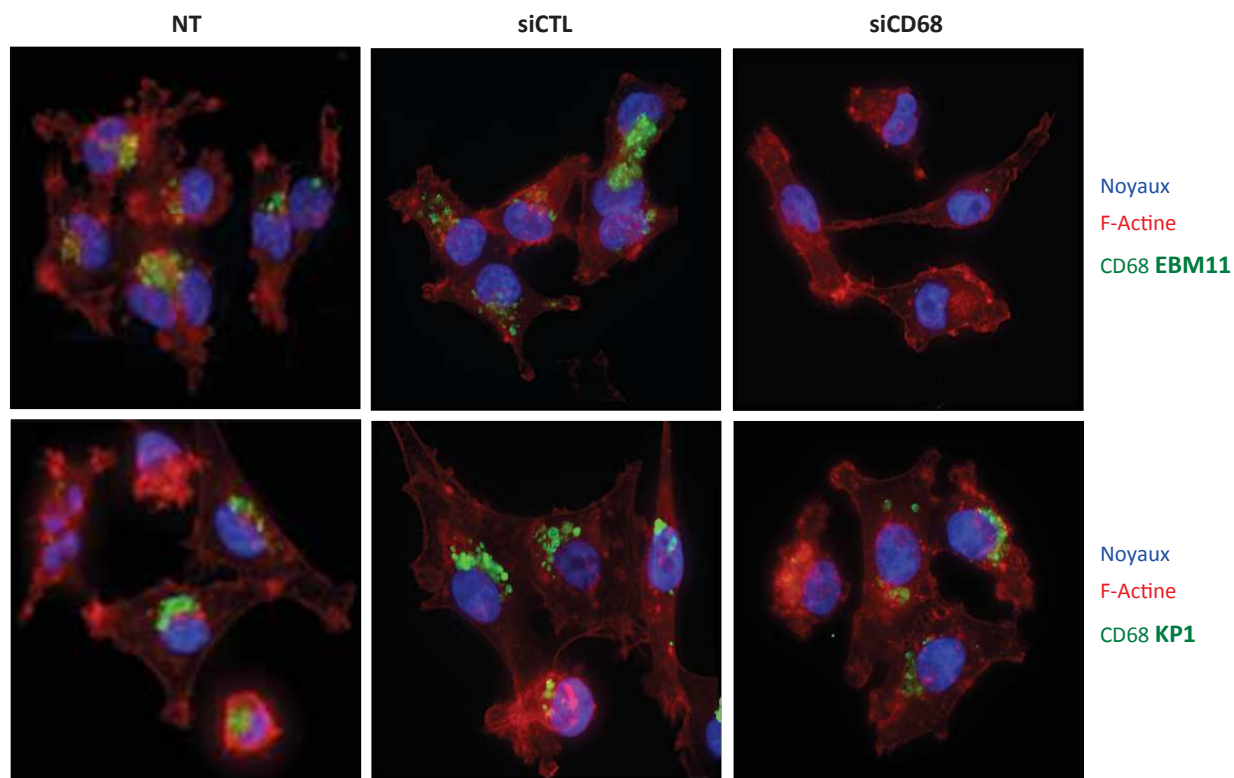
Les anticorps EBM11 et KP1 ont également été testés en immunoblot et aucun d'eux ne donne un signal spécifique. Les anticorps monoclonaux EBM11 et KP1 ont été produits à partir de la protéine entière contrairement à l'anticorps C9 qui est dirigé contre un peptide N-terminal. Ils pourraient donc reconnaître des épitopes conformationnels qui sont déstructurés après migration sur PAGE-SDS.

Les résultats obtenus par immunofluorescence et cytométrie de flux établissent que, dans les cellules MDA-MB231, la protéine CD68 est uniquement détectée au niveau intracellulaire dans des granules/vésicules cytoplasmiques (**Figure 74**).

Pour caractériser plus avant la nature des granules cytoplasmiques marqués par CD68, j'ai effectué des expériences de co-marquage par immunofluorescence sur cellules fixées visant à établir si l'on pouvait détecter une co-localisation avec les lysosomes et/ou les autophagosomes. En témoin négatif j'ai utilisé le co-marquage avec les filaments d'actine et les mitochondries.

Les résultats présentés sur la **Figure 75** indiquent une co-localisation partielle de CD68 avec les lysosomes ou les endosomes tardifs (lysotracker, marqueur des compartiments acides de la cellule) et les autophagosomes (marqueur LC3), suggérant une localisation probable de CD68 dans les autophagolysosomes. Ces résultats suggèrent que CD68 pourrait jouer un rôle dans le trafic vésiculaire lié au processus autophagique.

A



B

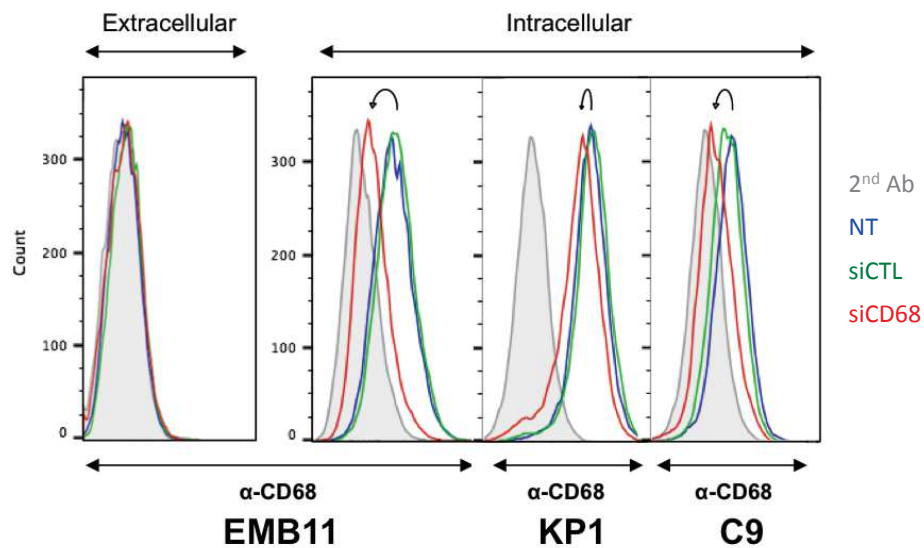


Figure 74 : Localisation de CD68 dans la lignée MDA-MB231. A) Immunofluorescence sur cellules fixées. **B)** Cytométrie en flux sur cellules vivantes (marquage extracellulaire) et fixées (marquage intracellulaire).

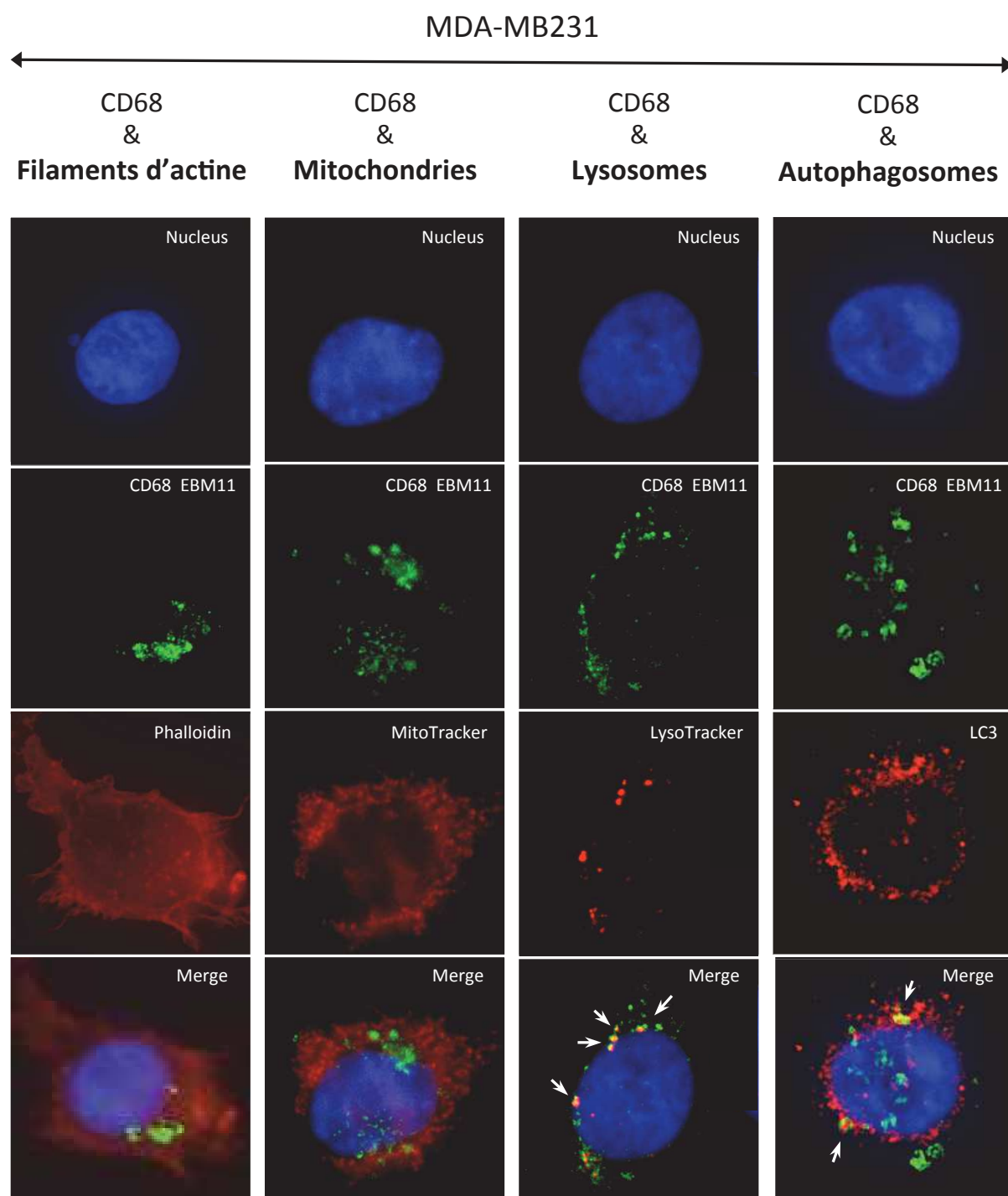


Figure 75 : Expériences de co-localisation dans la lignée MDA-MB231, par immunofluorescence sur cellules fixées.

II.5.b. Le processus autophagique

L'autophagie est un processus clé dans le maintien de l'homéostasie cellulaire en épurant la cellule des organites endommagés, des métabolites toxiques ou des pathogènes. Cette voie catabolique peut aussi être activée par des stress tel que la carence nutritionnelle. Les voies de signalisation conduisant au déclenchement de l'autophagie sont nombreuses et souvent interconnectées (Klionsky and Codogno, 2013). La protéine mTOR (mammalian target of rapamycin), senseur de l'état métabolique cellulaire, joue un rôle clé dans la régulation de l'autophagie (Jung et al., 2009). La voie autophagique conduit à la dégradation par le lysosome de composants cytoplasmiques (organelles endommagées, protéines avec une durée de vie longue, lipides, ect.) dans le but de procurer une source d'énergie essentielle à la survie de la cellule lors de carence nutritionnelle ou de stress. L'autophagie peut également induire la mort cellulaire en cas d'auto-digestion excessive ou de dégradation de constituants essentiels. L'autophagie est un processus biologique complexe finement régulé qui permet à la cellule de s'adapter aux conditions environnementales.

De nombreux travaux ont fait apparaître un lien entre la dérégulation de l'autophagie et différentes pathologies, telles que les maladies neurodégénératives, le diabète, l'infection par des pathogènes et le cancer (Chen and Debnath, 2010; Jing and Lim, 2012; Levine and Kroemer, 2008; Liang and Jung, 2010). Dans le cas du cancer, l'autophagie est décrite pour favoriser ou réprimer la tumorigenèse suivant le contexte cellulaire. Par exemple, l'expression constitutive de H-Ras ou K-Ras dans les cellules cancéreuses conduit à une augmentation de l'autophagie qui est nécessaire à la survie des cellules tumorales et à la formation des tumeurs (Guo et al., 2011a).

L'autophagie peut être divisée en quatre étapes principales (**Figure 76**) :

(i) L'étape d'initiation avec la formation d'une structure membranaire appelée phagophore constituée d'une bicouche phospholipidique. L'origine des membranes du phagophore est toujours sujette à débat, mais différentes études ont montré qu'elles pouvaient émerger du réticulum endoplasmique, de la mitochondrie, de l'appareil de Golgi ou encore des membranes nucléaire et plasmique (Hayashi-Nishino et al., 2009; Itoh et al., 2008; Lamb et al., 2013; Ravikumar et al., 2010).

(ii) Au cours de l'étape d'élongation, le phagophore s'allonge et se referme autour du contenu cytoplasmique à séquestrer pour former une vésicule à double membrane appelée autophagosome.

(iii) L'autophagosome fusionne ensuite avec le lysosome au cours de l'étape de maturation pour former un autophagolysosome.

(iv) La dernière étape du processus autophagique est la dégradation du contenu de l'autophagolysosome par les hydrolases du lysosome. Les produits de dégradation sont relargués dans le cytoplasme et réutilisés par la cellule.

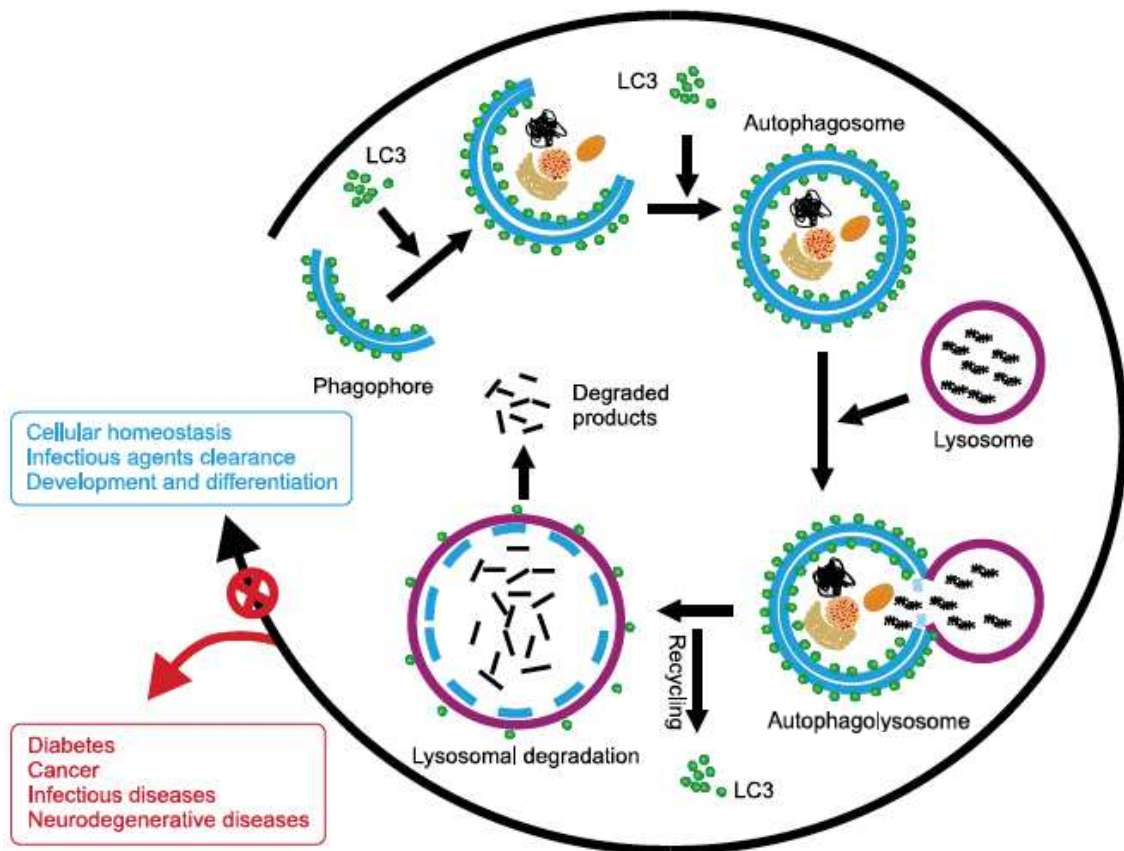


Figure 76 : Représentation schématique du processus autophagique. L'autophagie contribue au maintien de l'homéostasie cellulaire et sa dérégulation est impliquée dans plusieurs maladies humaines. *Source : Jing & Lim. (2012).*

La formation de l'autophagosome repose essentiellement sur l'action concertée de deux systèmes de conjugaison de type ubiquitination : le système de conjugaison Atg12-Atg5/Atg16L1 et le système de conjugaison LC3. Classiquement, la conjugaison de l'ubiquitine et des molécules dites « ubiquitin-like » sur leurs substrats nécessite une enzyme d'activation (E1), une enzyme de conjugaison (E2) et une ligase (E3).

Dans le premier système de conjugaison, la protéine Atg12 est activée par la protéine Atg7. Agissant comme une enzyme d'activation E1, Atg7 transfère Atg12 sur la protéine Atg10. Atg10 qui a une activité E2 de conjugaison lie de manière covalente Atg12 à Atg5. Aucune enzyme E3 ligase ne semble intervenir à cette étape. Le conjugué Atg12-Atg5 interagit alors avec Atg16L1 pour former le complexe Atg5-Atg12/Atg16L1 qui joue le rôle d'enzyme E3 pour le second système de conjugaison, la conjugaison de LC3 (**Figure 77**). Le complexe Atg5-Atg12/Atg16L1

est associé à la membrane externe de l'autophagosome en formation et est ensuite dissocié une fois l'autophagosome formé.

Le deuxième système de conjugaison concerne la protéine « ubiquitin-like » LC3 : c'est un système unique car c'est le seul connu permettant la conjugaison d'une protéine à un lipide, le phosphatidylethanolamine (PE). LC3 est tout d'abord synthétisé sous la forme d'un précurseur pro-LC3 qui est immédiatement clivé par la protéase Atg4 au niveau d'un résidu glycine situé en C-terminal, pour produire la protéine LC3-I cytosolique. Par l'intermédiaire de sa glycine C-terminale, LC3-I est liée de manière covalente au PE, on parle alors de la forme LC3-II. Cette conjugaison nécessite l'activité de Atg7 (E1), de Atg3 (E2) et du complexe Atg12-Atg5/Atg16L1 (E3) (**Figure 77**). Grâce à sa liaison au PE, la protéine LC3-II s'ancree à la fois dans la membrane externe et interne des autophagosomes. Après fermeture de celui-ci, les protéines LC3-II fixées sur la membrane externe de l'autophagosome sont clivées par Atg4 et recyclées. Celles présentes sur la membrane interne de l'autophagosome restent intactes jusqu'à l'étape de fusion avec le lysosome où elles sont dégradées avec le contenu de l'autophagosome. La forme LC3-II est donc présente tout au long du processus autophagique. C'est un marqueur spécifique utilisé pour suivre l'autophagie et son flux dégradatif (Klionsky et al., 2016).

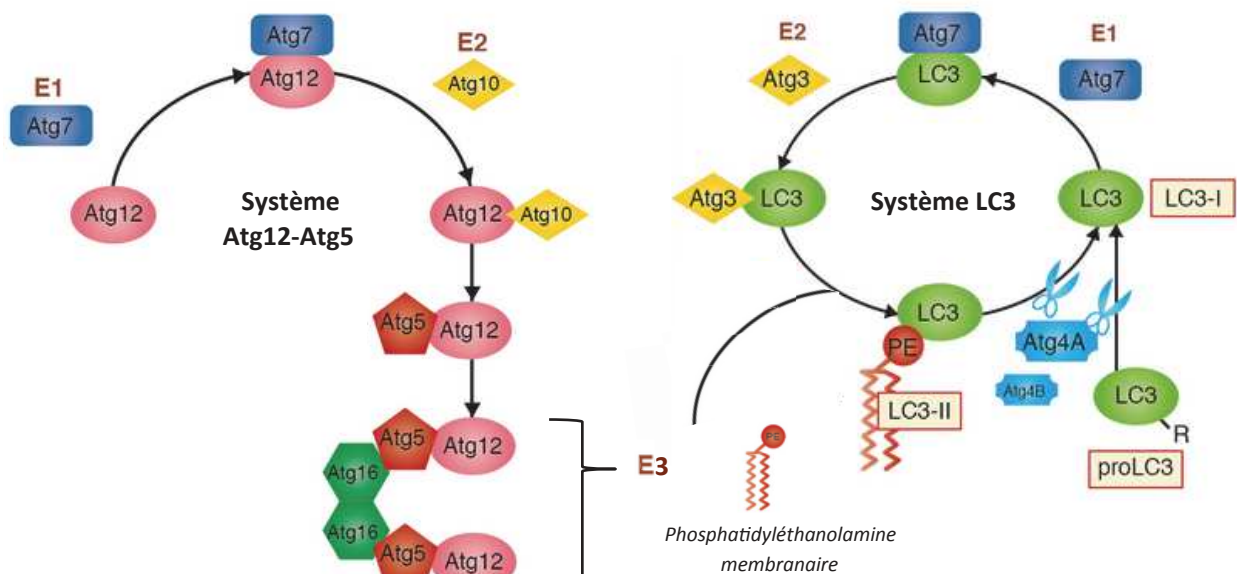


Figure 77 : Systèmes de conjugaison Atg12-Atg5 et LC3. Adapté de (Randall-Demllo et al., 2013) et <http://mbl.co.jp/bio/e/product/autophagy/autophagy.html>.

L'utilisation du marqueur LC3-II pour suivre l'autophagie doit être complétée par l'étude du flux autophagique. En effet, quand on observe une induction de l'autophagie, caractérisée par une augmentation du marquage LC3-II en immunoblot ou en microscopie, on ne peut pas savoir si elle est la conséquence d'une induction de la formation des autophagosomes ou d'une

inhibition de la maturation, due à un blocage de l'étape de fusion (**Figure 78**). De plus, si le flux autophagique est important, l'induction de l'autophagie peut être en partie masquée. Il faut donc être particulièrement attentif à la dynamique du flux autophagique (Klionsky et al., 2016).

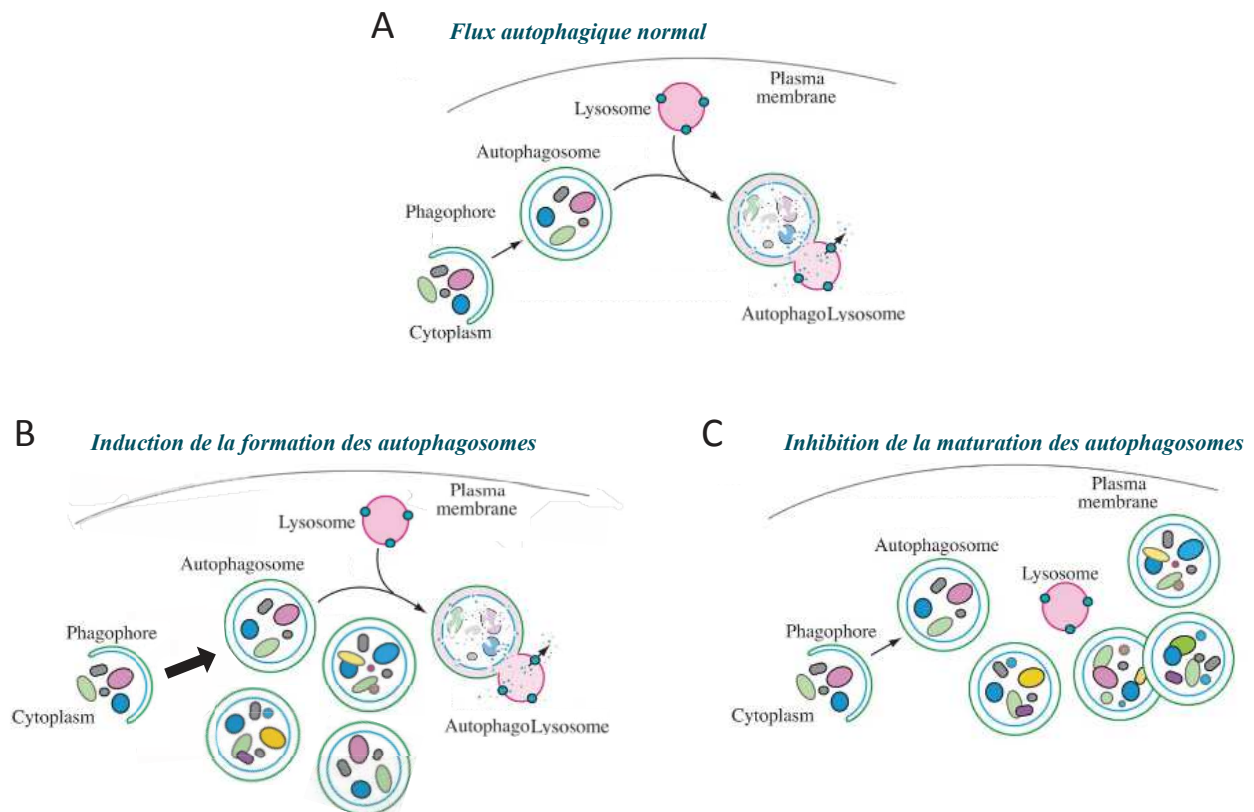


Figure 78 : Représentation schématique de la dérégulation du flux autophagique. A) Le flux autophagique est un terme utilisé pour définir l'activité autophagique, allant de la séquestration du matériel cellulaire à sa dégradation par les lysosomes. B) Augmentation de l'autophagie due à une augmentation de la formation des autophagosomes. C) Augmentation de l'autophagie due à l'inhibition de la maturation, c'est à dire le blocage de l'étape de fusion avec le lysosome *Illustration adaptée de Klionsky et al. (2016).*

En pratique, dans la littérature quatre conditions sont couramment utilisées pour permettre l'interprétation des résultats : condition basale, induction de l'autophagie, blocage du flux et induction de l'autophagie couplée au blocage du flux. Pour cela, dans mes études j'ai eu recours à l'utilisation de la torin-1 un inhibiteur de mTOR (sérine/thréonine kinase jouant le rôle de senseur des conditions environnementales) pour induire l'autophagie, ainsi qu'à la chloroquine un inhibiteur de l'acidification du lysosome qui bloque l'activité protéolytique lysosomale et empêche la fusion lysosome/autophagosome, pour bloquer le flux.

II.5.c. Modulation de la voie autophagique et impact sur la localisation de CD68 dans la lignée MDA-MB231

Dans ce dernier paragraphe sont présentés des résultats préliminaires qui pourraient être en faveur d'un lien entre CD68 et l'autophagie dans les cellules MDA-MB231.

En collaboration avec L. Esperte et M. Biard (CPBS, Montpellier), j'ai utilisé deux techniques classiques et complémentaires pour suivre l'autophagie en marquant la protéine LC3 : la microscopie à fluorescence et l'immunoblot (**Figure 79**) (Klionsky et al., 2016). En immunofluorescence, LC3 sous sa forme LC3-I n'est pas détectable car diffuse dans le cytoplasme alors que LC3-II est visualisée dans des structures ponctuelles correspondant aux autophagosomes (**Figure 79.A**). En immunoblot, la lipidation de LC3 modifie ses propriétés électrophorétiques : LC3-I migre avec un poids moléculaire apparent de 18 kDa tandis que LC3-II migre avec un poids moléculaire apparent de 16 kDa (**Figure 79.B**). On peut facilement distinguer les deux formes en immunoblot pour suivre l'induction de l'autophagie.

* Condition basale versus blocage du flux par la chloroquine

La comparaison de l'abondance de LC3-II en condition basale et lors du blocage du flux par la chloroquine montre que le flux autophagique basal est important dans les cellules MDA-MB231 (**Figure 79**). En effet, le seul fait de bloquer le flux autophagique par la chloroquine est suffisant pour observer une accumulation de la protéine LC3-II comme cela est visible sur l'immunoblot et en immunofluorescence. De plus, par immunofluorescence, on observe un changement dans la distribution de LC3-II lors du blocage du flux. La protéine est présente sur des structures larges qui correspondent vraisemblablement à des autophagosomes de grande taille.

De façon intéressante, on observe que l'abondance de la protéine CD68 augmente aussi lors du blocage du flux et se distribue sur ces mêmes structures, dans lesquelles CD68 et LC3-II semblent enchassées (**Figure 79.A**). Étant donné que CD68 présente une localisation partielle avec les lysosomes et/ou endosomes tardifs (**Figure 75**), il se pourrait que ces vésicules de grande taille correspondent à la fusion entre les autophagosomes (marquage LC3-II) et les lysosomes et/ou les endosomes (marquage CD68). Ceci signifierait que dans nos conditions, la chloroquine n'a pas inhibé la fusion mais a uniquement bloqué l'acidification des lysosomes et donc l'activité des enzymes lysosomales. Le marquage d'une protéine lysosomale telle que LAMP-1 ou LAMP-2, d'une part, et d'une protéine caractéristique des endosomes précoces (Rab5) ou tardifs (Rab7), d'autre part, permettrait de répondre à cette question.

* Induction de l'autophagie par la torin-1 combinée ou non au blocage du flux

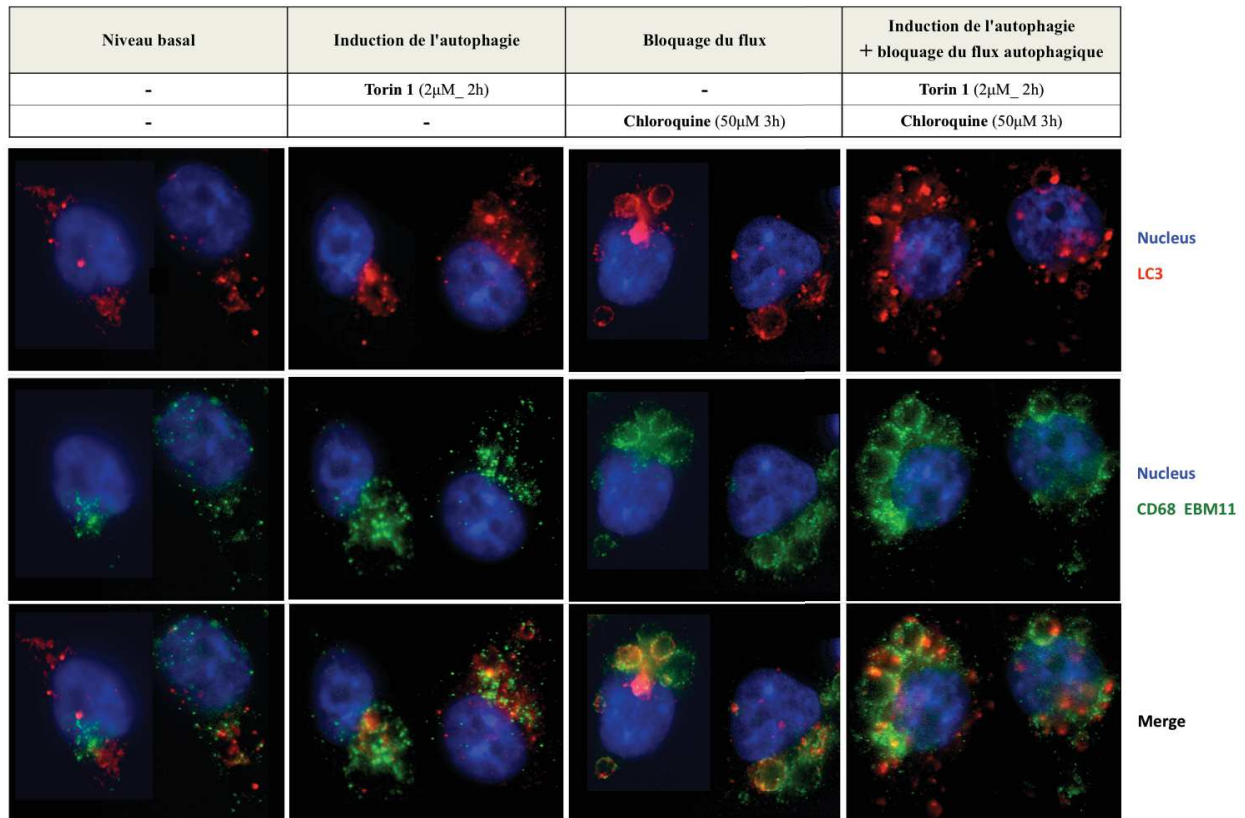
L'induction de l'autophagie par l'ajout de torin-1 seule, entraîne une légère augmentation du signal LC3-II en immunoblot et en immunofluorescence comparée à la condition basale, ce qui

est en accord avec le flux autophagique important des nos cellules (**Figure 79**). Lorsque l'induction de l'autophagie est couplée au blocage du flux on détecte un incrément dans l'accumulation de LC3-II en immunoblot et en immunofluorescence (**Figure 79**). De plus, en immunofluorescence on remarque l'apparition de points LC3-II de taille plus importante sur les structures vésiculaires.

En ce qui concerne CD68, on observe de façon curieuse une diminution de sa détection en immunoblot lors de l'induction de l'autophagie par la torin-1 seule comparée à la condition basale (**Figure 79.B**), alors qu'en immunofluorescence le signal ne décroît pas, voire augmente sensiblement (**Figure 79.A**). Cette discordance, encore non expliquée, est beaucoup plus marquée lorsque l'induction de l'autophagie est couplée au blocage du flux, où il apparaît de façon nette que les abondances visualisées en immunoblot et en immunofluorescence ne sont pas cohérentes (**Figure 79**). Ces différences pourraient être liées au fait que les anticorps utilisés pour les révélations de CD68 en immunoblot et en immunofluorescence sont différents. En effet, comme mentionné précédemment, l'anticorps C9 est dirigé contre un peptide N-terminal (27-60) alors que l'anticorps EBM11 reconnaît vraisemblablement un épitope conformationnel (voir paragraphe II.5.a). L'utilisation d'autres anticorps CD68 dirigés contre d'autres épitopes devrait répondre à cette question. Une explication plausible à ces résultats préliminaires pourrait être que l'induction de l'autophagie par la torin-1 induit une modification post-traductionnelle de CD68 qui masque l'épitope N-terminal reconnu par l'anticorps C9. Pour déterminer si cet effet est spécifique de la torin-1 ou est lié à l'induction de l'autophagie, on peut envisager l'utilisation d'autres drogues (rapamycine, perifosine, etc.) ou encore l'induction d'un stress métabolique tel que la carence en nutriments (culture en milieu EBSS).

Néanmoins, ces résultats préliminaires qui mettent en évidence une co-distribution de CD68 et de LC3-II, vraisemblablement dans des autophagolysosomes, dans des conditions où le flux autophagique est inhibé, suggèrent une implication de CD68 dans le processus autophagique des cellules MDA-MB231.

A



B

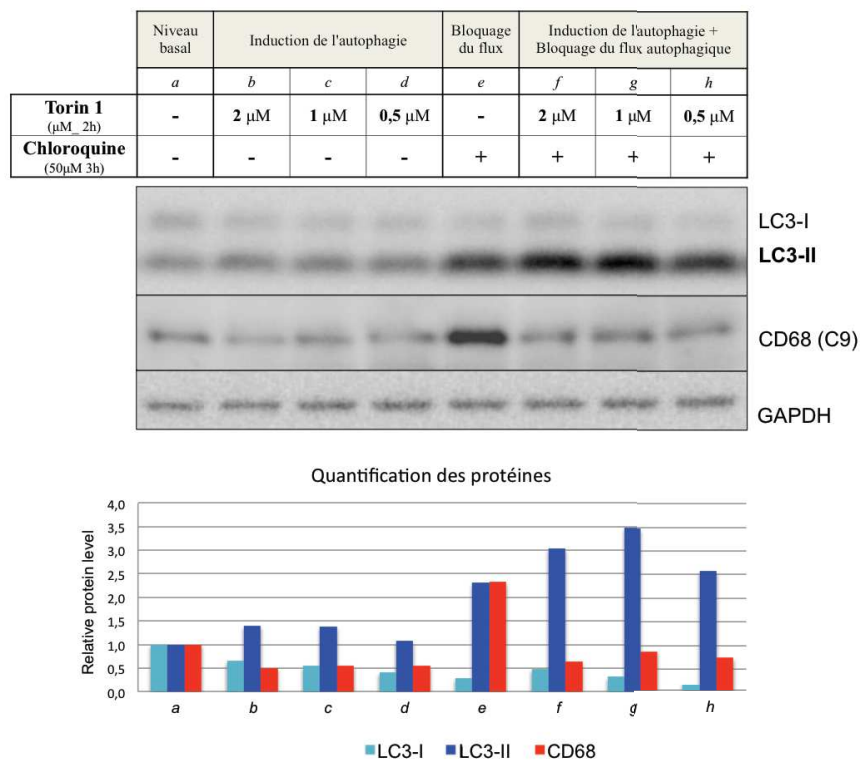


Figure 79 : Détection des protéines CD68 et LC3, dans les cellules MDA-MB231. A) Immunofluorescence sur cellules fixées. Une version haute définition de cette figure est disponible en annexe (**Annexe 7**). **B)** Immunoblot et quantification des protéines suite à l'utilisation de drogues modulant la voie autophagique.

III. CD68 : Conclusion

Mes études préliminaires n'ont pas mis en évidence de rôle de CD68 dans la prolifération, la migration ou l'invasion des cellules MDA-MB231. Il est à noter cependant que les études ont été menées sur des fenêtres de temps restreintes du fait de l'inactivation phénotypique transitoire liée à l'utilisation des siRNA. L'obtention d'une inactivation sur le plus long terme par transfection stable de shRNA sera donc nécessaire pour poursuivre mes investigations. Pour la suite de l'étude du rôle de CD68 dans le phénotype tumoral des MDA-MB231 et des TNBC de manière plus générale, il sera également important de déterminer précisément la localisation de CD68. En effet, la détection en immunofluorescence de CD68, des autophagosomes (par l'intermédiaire de LC3) et des compartiments acides (lysosomes et endosomes tardifs par l'intermédiaire du lysotracker) indique une co-localisation partielle. Il sera donc important d'effectuer des marquages avec des protéines spécifiques de chacun des compartiments. Par exemple Rab5 pour les endosomes précoces (dont la lumière n'est pas acide), Rab7 pour les endosomes tardifs et cathepsine-D pour les lysosomes, afin de déterminer plus précisément dans quel(s) type(s) de compartiment cellulaire est localisé CD68 de manière à orienter nos recherches. Néanmoins, il est important de noter que les endosomes tardifs peuvent fusionner avec les autophagosomes formant des structures appelées amphisomes capables de fusionner à leur tour avec les lysosomes pour former un autolysosome (**Figure 80**) (Klionsky, 2007). Ainsi, quelle que soit la localisation précise de CD68, son implication dans le flux autophagique nécessite investigation.

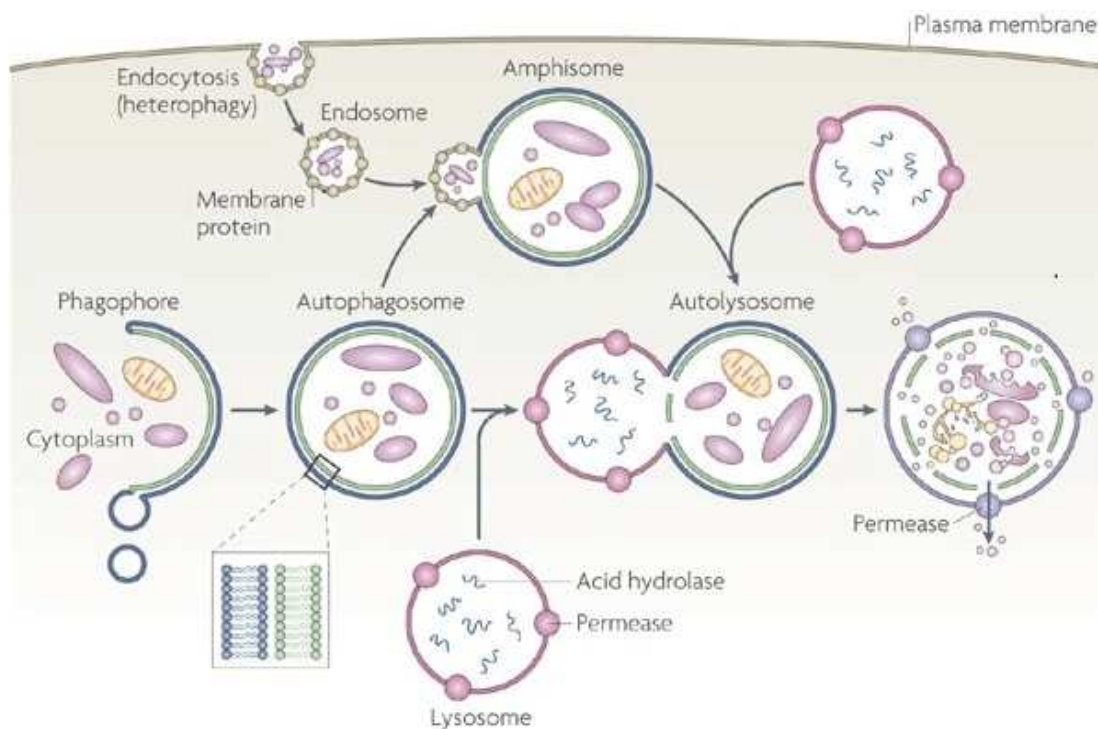


Figure 80 : Fusion entre les différentes vésicules intracellulaires. Les amphisomes, structure hybride obtenue par fusion entre les endosomes tardifs et les autophagosomes représentent le point d'intersection entre la voie endosomale et la voie autophagique. *Source : (Klionsky, 2007).*

Il est maintenant établi que l'autophagie est impliquée dans de nombreuses pathologies dont le cancer. Le rôle de l'autophagie dans le cancer s'avère complexe et dépend probablement en grande partie du contexte cellulaire. En effet, l'autophagie serait un mécanisme suppresseur de tumeur au moment de l'initiation de la tumeur en permettant, entre autres, une élimination des mutagènes tels que des mitochondries altérées (mitophagie), source de ROS, ce qui permettrait de prévenir des altérations de l'ADN, ou encore entraînerait la mort cellulaire suite une forte augmentation de son activité dans certaines conditions (Gozuacik and Kimchi, 2004; Jiang et al., 2015). Cependant, il est clair que l'autophagie promeut aussi la progression maligne suite à l'initiation tumorale. En effet, l'augmentation de la taille des tumeurs solides induit des stress, en particulier dans leur région centrale, en raison de l'hypoxie et du manque en nutriments. Dans ces conditions, l'autophagie en apportant une source alternative d'énergie permet aux cellules tumorales de survivre. Cette notion est supportée par le fait que des défauts dans la machinerie autophagique entraîne une diminution de la prolifération, de la dissémination et du potentiel métastatique des cellules cancéreuses (Jiang et al., 2015). De plus, les tumeurs humaines à un stade avancé montrent généralement une augmentation du flux autophagique qui corrèle avec le caractère invasif et métastatique et avec un mauvais pronostic (Galluzzi et al., 2015; Guo et al., 2013; Lazova et al., 2012).

Il a également été montré que le traitement par tamoxifène des cancers du sein ER+ induit de l'autophagie dans les cellules cancéreuses et que cette induction d'autophagie est impliquée dans les mécanismes de résistance au tamoxifène. Le mécanisme sous-jacent fait intervenir l'induction de l'expression de LAMP-3, une protéine lysosomale. La délétion phénotypique de LAMP-3, dans des cellules MCF7 rendues résistantes au tamoxifène, diminue le flux autophagique et re-sensibilise les cellules au tamoxifène. Les analyses menées sur des patientes traitées en première ligne au tamoxifène montrent qu'une expression élevée de l'ARNm LAMP-3 corrèle avec une survie plus courte sans progression tumorale et une survie plus courte après la rechute (Nagelkerke et al., 2014). Une augmentation de l'expression de LAMP-3 a également été rapportée dans les carcinomes squameux de l'œsophage et corrèle avec un mauvais pronostic (Liao et al., 2015). Par ailleurs, quelques études ont rapporté que les cellules cancéreuses présentant des mutations activatrices H-Ras ou K-Ras sont dépendantes de l'autophagie pour leur survie et leur croissance (Guo et al., 2011b; Lock et al., 2011, 2014). Cependant, cette conclusion est remise en question par les travaux publiés récemment par Eng *et al* qui montrent que l'invalidation d'ATG7 dans des cellules de pancréas (Panc10.50) ou de poumon (A549) xénogreffées chez la souris n'affecte pas la croissance tumorale (Eng et al., 2016).

Le complexe AP-1 est un facteur de réponse précoce immédiat aux stress de différentes natures. Il n'est donc pas improbable que les stress induisant l'autophagie dans les cellules saines et régulant le processus autophagique dans les cellules tumorales impliquent aussi les, ou certains, membres AP-1. Une seule étude publiée en 2010 a montré que l'induction de

l'autophagie par carence en nutriments conduit à une diminution de l'expression de JunB par des mécanismes transcriptionnels et post-transcriptionnels, dans des fibroblastes embryonnaires de souris et des fibroblastes transformés (Yogev et al., 2010). Les auteurs ont également montré que la surexpression de JunB ou c-Jun, contrairement à celle de JunD, c-Fos ou Fra-1, inhibe l'autophagie induite par la carence en nutriments et que cet effet dépend de l'activité transcriptionnelle de ces deux protéines AP-1.

Aucune étude n'a recherché si Fra-1, en plus de ses rôles bien établis dans la l'EMT, la migration, l'invasion et la métastatisation, pouvait être impliquée dans le processus autophagique dans les cellules cancéreuses, via la régulation de l'expression de protéines associées à ce mécanisme. Mes études transcriptomiques menées sur la lignée de cancers du sein triple négatif MDA-MB231 ont fait apparaître que Fra-1 et/ou Fra-2 régulent positivement plusieurs gènes impliqués dans le processus autophagique tels que ATG4A, ATG7 ou AGT12, ce qui renforce la notion que Fra-1 et Fra-2 pourraient contribuer au flux autophagique dans les cellules tumorales.

L'étude d'un rôle potentiel de CD68, mais aussi de Fra-1 et/ou Fra-2 dans l'autophagie au sein des cellules tumorales nécessitera de soumettre les cellules à des stress tels que la carence en nutriments, l'hypoxie ou l'acidose, qui sont présents dans les tumeurs solides. Ces études permettront de déterminer si CD68, Fra-1 ou Fra-2 favorisent via l'autophagie la survie ou encore la prolifération des cellules dans des conditions défavorables. D'autres sources de stress également intéressantes à tester sont les drogues chimiothérapeutiques (doxorubicine, méthotrexate, ect.) utilisées en clinique pour traiter les patientes atteintes de cancers du sein triple négatifs. En effet, comme mentionné plus haut, l'augmentation de l'autophagie dans certaines cellules cancéreuses les rend plus résistantes aux traitements anticancéreux. Une meilleure compréhension des mécanismes sous-jacents permettrait d'améliorer la connaissance dans les processus de résistances aux thérapies et pourrait conduire à envisager de nouvelles approches thérapeutiques.

DISCUSSION GÉNÉRALE

DISCUSSION GÉNÉRALE

La transcription régulée des gènes est un mécanisme crucial pour que chaque type cellulaire puisse avoir son devenir propre et pour que chaque cellule puisse répondre aux stimuli extracellulaires qu'elle reçoit et s'adapter à son milieu environnant. Ainsi chaque type cellulaire aura un transcriptome particulier qui variera en fonction des signaux reçus. Il est clair que le transcriptome des cellules est grandement dérégulé et modifié lors de la cancérisation ce qui permet aux cellules cancéreuses d'acquérir de nouvelles caractéristiques leur conférant des avantages sélectifs, des capacités d'invasion et d'échappement au système immunitaire ainsi qu'une résistance aux drogues anti-tumorales.

L'identification des modifications des transcriptomes dans les différents cancers est apparue, aux yeux de la communauté scientifique, comme un moyen de choix pour identifier de nouvelles cibles thérapeutiques dans le but de développer de nouvelles approches cliniques, en particulier pour les cancers à pronostic encore sombres tels que les cancers du sein triple négatifs. Les nombreuses études menées ont permis de classer les tumeurs du sein en différents sous-types moléculaires (luminal A et B, HER2+, triple négatif, normal breast-like, claudin-low, ect.) mais ont également mis en évidence une grande hétérogénéité entre les tumeurs des patientes d'un même sous-type, ce qui a conduit à affiner les classements des TNBC en sous 6 sous-groupes avec des devenir plus ou moins sombres. Malheureusement, malgré ces études, aucune nouvelle thérapie efficace et ciblée n'a encore vu le jour pour le traitement de ces cancers et la recherche de cibles moléculaires potentiellement thérapeutiques est toujours d'actualité. C'est dans ce contexte que s'est inscrit mon projet de thèse visant à mieux comprendre les rôles transcriptionnels de Fra-1 et Fra-2, deux facteurs de transcription de la famille AP-1 fortement exprimés dans les TNBC et impliqués dans la progression cancéreuse, avec comme perspective l'identification d'une ou plusieurs cibles intéressantes dans le cadre de la lutte anti-tumorale.

Mon travail de thèse a fait appel à des études transcriptomiques pour caractériser les gènes (dé)régulés par Fra-1 et Fra-2 dans les TNBC. Les résultats ont montré que les gènes cibles se répartissent en 4 catégories principales : (i) les gènes préférentiellement régulés par Fra-1, (ii) ou préférentiellement par Fra-2 et (iii) les gènes régulés de façon complémentaire ou (iii) de façon redondante par les deux facteurs. Ces résultats mettant en évidence des redondances fonctionnelles pour ces protéines et soulignent l'importance d'identifier les transcriptomes des facteurs de transcription indépendamment et conjointement dès lors que l'on est en présence de facteurs appartenant à des familles multigéniques, de manière à avoir la vue la plus synthétique possible des gènes affectés par l'expression des différents membres de la famille. Dans le cas de la famille AP-1 il est établi que suivant les types cellulaires et les stimuli extracellulaires reçus différents membres de la famille sont exprimés de façon plus ou moins transitoire. Il en ressort

que les transcriptomes obtenus dans un contexte cellulaire ne seront pas systématiquement valables dans un autre.

Les analyses transcriptomiques ne permettent pas de déterminer si les gènes modulés sont des cibles directes ou indirectes. Pour répondre à cette question, j'ai donc mené en parallèle des approches ChIP-seq pour cartographier les sites de liaison de Fra-1 et Fra-2 sur le modèle cellulaire étudié (MDA-MB231). Les données obtenues ont été complétées par des ChIP-seq pour différentes marques d'histones dans le but de caractériser la nature des sites de fixation. En effet, si un grand nombre de données de ChIP-seq sont présentes dans la base ENCODE, les cellules que j'ai utilisées comme lignée modèle de TNBC ne sont pas répertoriées dans cette base de données. Il est à noter qu'aucune lignée TNBC n'y est encore présente, le dépôt de nos données devrait donc être utile pour la communauté scientifique. Nos analyses ont mis en évidence que Fra-1 et Fra-2 se fixent essentiellement sur des enhancers ce qui est en accord avec les données récentes de la littérature qui, sur la base d'études bioinformatiques, ont identifié les motifs AP-1 comme étant des caractéristiques des enhancers.

Suite à cette étude, nous nous sommes heurtés à l'une des grosses difficultés du domaine de la transcription qui est l'attribution des enhancers aux gènes qu'ils régulent. Les très nombreuses études menées à grande échelle depuis quelques années ont permis de mettre en évidence un certain nombre de caractéristiques des enhancers et des promoteurs, en particulier des modifications d'histones différentielles. De plus, l'utilisation conjointe des techniques de GRO-seq et de ChIP-seq a montré que les enhancers recrutent l'ARN Pol II et initient la transcription bidirectionnelle de petits ARN instables. Ces observations, entre autres, ont conduit à reconsidérer la définition des enhancers et des promoteurs qui sont maintenant proposés pour constituer une classe unique d'éléments fonctionnels avec la même architecture pour l'initiation de la transcription et avec une activité différentielle enhancer/promoteur liée essentiellement à leur capacité de fixer différents facteurs de transcription. D'autre part, le niveau de transcription des enhancers a été montré corrélé à celui des promoteurs correspondants. Sur cette base, le consortium international FANTOM (Functional ANnotation Of the Mammalian genome) en utilisant la technique des CAGEs (Cap Analysis of Gene Expression) a identifié et caractérisé l'activité des promoteurs et enhancers sur plus de 400 types de cellules primaires humaines prélevées à partir de plus de 150 tissus. Les données ont été complétées avec environ 250 lignées cancéreuses humaines (mais ne contenant pas les MDA-MB231) représentant 150 sous-types distincts de cancers. L'analyse des corrélations d'expression a permis au consortium FANTOM de proposer un atlas des enhancers actifs et de les corréler aux promoteurs correspondants. Si certaines des interactions enhancer/promoteur sont soutenues par les données issues d'approches de ChIA-PET (Chromatin Interaction Analysis by Paired-End Tags sequencing) ou de capture de conformation de la chromatine (approche 3C et dérivés), la majorité d'entre elles n'a pas encore été validée expérimentalement. Malgré la puissance de ces données à grande échelle permettant d'obtenir des renseignements globaux au niveau du génome

et de conduire à de nouveaux concepts, leur intégration dans l'interprétation des données expérimentales sur un type cellulaire particulier reste délicate. Dans le cas de mon étude, l'utilisation de l'annotation FANTOM5 a permis d'identifier seulement 179 gènes régulés par Fra-1 et/ou Fra-2 dont les enhancers robustes FANTOM5 contiennent Fra-1 et/ou Fra-2, ce qui paraît peu par rapport au nombre de gènes identifiés par l'approche transcriptomique. Il se pourrait donc soit qu'une majorité des gènes identifiés le soit de façon indirecte par Fra-1 et/ou Fra-2 via la régulation de l'expression de facteurs de transcription ou de microARN, soit qu'un certain nombre d'enhancers fonctionnels dans les cellules modèles MDA-MB231 ne sont pas annotés comme robustes par le consortium FANTOM5. Il apparaît donc que l'étude des mécanismes de la transcription, si elle bénéficie de grandes avancées dans le domaine grâce aux études à grande échelle et aux progrès technologiques, nécessite aussi des études plus ciblées sur des gènes particuliers pour appréhender les mécanismes moléculaires fins de la régulation transcriptionnelle. En effet, dans le cas de Fra-1 et Fra-2, les résultats obtenus montrent que suivant les gènes ces deux facteurs de transcription ont une action positive ou négative sur la transcription, indiquant que les mécanismes sous-jacents sont très vraisemblablement différents. Il est à noter également que suivant la topologie des gènes (positionnement de l'enhancer par rapport au promoteur et nature des sites de fixation pour différents facteurs de transcription sur ces éléments du génome) les mécanismes mis en jeu risquent d'être sensiblement différents, même pour deux gènes régulés dans le même sens. Cette notion est illustrée par nos travaux qui ont montré que si Fra-1 n'est pas impliqué dans le recrutement de l'ARN Pol II sur le promoteur du gène uPA/PLAU, sa présence est, par contre, nécessaire au recrutement de Pol II sur le promoteur du gène HMGA1. Il apparaît donc que les mécanismes moléculaires fins devront être caractérisés pour de nombreux gènes avant de pouvoir déterminer ce qui peut avoir valeur plus générale en lien avec la topologie des gènes.

Mon étude transcriptomique a, par ailleurs, fait apparaître que les niveaux de régulation des gènes par Fra-1 et/ou Fra-2, dans les conditions étudiées, sont relativement faibles avec seulement 281 gènes dont le niveau d'expression varie d'un facteur supérieur ou égal à 2. Ceci pourrait indiquer que, pour la majorité des gènes régulés, la contribution de Fra-1 et/ou Fra-2 est faible. Cependant, on ne peut pas exclure que cette observation est liée à nos conditions expérimentales qui visaient (i) à limiter au maximum les effets off-targets des siRNA en transfectant des quantités minimales de siRNA et (ii) à ne pas observer des effets d'effets liés à l'analyse des transcrits à des temps trop longs après l'extinction des facteurs d'intérêt. Néanmoins, dans ces conditions, quelques gènes montrent une forte diminution de leur expression en absence de Fra-1 et/ou Fra-2, ce qui traduit un contrôle positif de leur expression par ces facteurs de transcription. Parmi ces gènes, le gène CD68 présente l'un des facteurs d'induction les plus importants et la régulation de son expression par les protéines Fra a été validée dans deux autres lignées triple négatives. L'expression de CD68 est également corrélée avec l'agressivité des cancers du sein et le niveau d'expression de Fra-1 et/ou Fra-2 suivant les

banques de données analysées. Bien que généralement considéré comme un marqueur des macrophages, CD68 est aussi exprimé par des cellules tumorales et sa participation potentielle au phénotype tumoral n'a encore jamais été rapportée. Mes résultats ont montré que CD68 ne joue pas de rôle majeur dans la prolifération et l'invasion cellulaire dans des conditions optimales de croissance des cellules en culture. Cependant, mes données préliminaires suggèrent que CD68 pourrait être impliqué dans l'autophagie ou le flux autophagique. Ces aspects devront être approfondis et pourraient conduire à la mise en évidence, pour la première fois, d'un rôle de Fra-1 et/ou Fra-2 dans la régulation de l'autophagie dans les cellules cancéreuses via la régulation de gènes spécifiques. Cet aspect est d'autant plus intéressant que l'autophagie est rapportée pour promouvoir la progression maligne dans des conditions de croissance défavorables des cellules, telles qu'elles peuvent être rencontrées dans les tumeurs solides, et pour augmenter la résistance des cellules cancéreuses aux drogues.

Il est à noter que de nombreux autres cancers épithéliaux agressifs (ovaire, thyroïde, vessie, colon, ect.) surexpriment également Fra-1 et Fra-2. Mes observations pourraient donc, à plus long terme, être étendues au delà du cancer du sein.

ANNEXES

Gènes up-régulés par Fra-1 et/ou Fra-2 (FC ≤-2 ; FDR<0,05)				Gènes up-régulés par Fra-1 et/ou Fra-2 (FC ≤-2 ; FDR<0,05)			
GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT	GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT
LOC100507127	-3,64	-1,83	-11,06	FAM138B	NA	NA	-2,04
MMP1	-4,32	-2,00	-7,86	ZXDB	NA	NA	-2,03
MPP4	-3,00	NA	-6,52	CD274	-1,41	NA	-2,02
FOSL1	-2,69	NA	-6,39	ACSL5	-1,67	NA	-2,02
LOC100216001	-2,42	NA	-5,62	SNORA2B	NA	NA	-2,02
CD68	-2,27	NA	-5,54	MARCH4	NA	NA	-2,01
NAV3	-2,15	NA	-5,52	WASF3	-1,50	NA	-2,01
FAM171A1	-1,94	NA	-5,32	APBB1P	-1,36	NA	-2,00
CD22	-2,31	NA	-5,31	SMTN	-1,42	NA	-2,00
TMEM158	-2,52	NA	-4,91	SLC9A7	-1,64	NA	-2,00
LOC730755	NA	-1,83	-4,53	PTPRE	-1,43	NA	-2,00
NLRP10	-1,61	-1,48	-4,15				
BCL2A1	-2,13	NA	-3,63				
MAP2K1	-1,66	NA	-3,63				
ZNF114	-2,09	NA	-3,61				
RBBP9	-1,83	NA	-3,48				
HLA-DRA	-1,80	NA	-3,35				
EVI2A	-1,70	NA	-3,25				
FAM40B	-1,71	NA	-3,23				
CPED1	NA	NA	-3,23				
VEPH1	-1,80	NA	-3,05				
CDCP1	-1,68	NA	-2,99				
MIR222	NA	NA	-2,96				
WASF2	-1,52	-1,29	-2,94				
SSFA2	-1,84	NA	-2,92				
PDE2A	-1,68	-1,34	-2,89				
PTPN22	-1,75	-1,87	-2,88				
FBP1	-1,41	NA	-2,83				
ARNTL2	-1,77	NA	-2,80				
PITPNC1	-1,87	NA	-2,79				
KHDRBS3	NA	-1,50	-2,79				
MXD1	-1,64	NA	-2,76				
PEAR1	-1,54	NA	-2,75				
ZNF521	-1,63	NA	-2,73				
STEAP1	NA	NA	-2,69				
LOC100129940	NA	NA	-2,69				
RPSAP52	-1,87	NA	-2,69				
AIM1	NA	NA	-2,68				
AOX1	NA	NA	-2,65				
ZNF185	-1,59	NA	-2,64				
DOCK4	-1,68	NA	-2,60				
RN5S237	-1,91	NA	-2,57				
ADAMTS6	-2,24	-1,55	-2,57				
ADAMTSL1	-1,32	NA	-2,54				
MRPL9	-1,56	NA	-2,52				
SLCO1B7	-1,64	NA	-2,48				
FAM46C	-1,70	NA	-2,47				
TMEM194B	-1,42	-1,26	-2,47				
SLC22A3	-1,66	NA	-2,47				
FBXO28	-1,42	-1,43	-2,46				
CHRN81	-1,43	NA	-2,45				
HK2	-1,41	NA	-2,42				
ABT1	-1,55	NA	-2,41				
MIR221	NA	NA	-2,39				
RAB8B	NA	-1,45	-2,39				
SNORD99	NA	NA	-2,39				
LOC338588	NA	NA	-2,38				
MT1X	-1,58	NA	-2,37				
SMAGP	-1,36	NA	-2,37				
SASH1	-1,83	NA	-2,34				
PRKCDBP	-1,57	NA	-2,33				
ABLIM3	-1,64	NA	-2,32				
NOL6	-1,61	NA	-2,32				
GOS2	-1,54	NA	-2,31				
PTX3	-1,26	NA	-2,31				
LPXN	-1,71	NA	-2,31				
PLBD2	-1,70	NA	-2,30				
SLCO1B3	-1,73	NA	-2,27				
LOC646329	NA	NA	-2,25				
SLC26A2	-1,37	NA	-2,24				
EMP1	NA	NA	-2,23				
MLEC	-1,36	NA	-2,22				
RAD21	-1,66	NA	-2,22				
HMGA1	-1,62	NA	-2,21				
TFAP4	-1,42	NA	-2,20				
MYO10	-1,40	NA	-2,20				
ADAMTS16	-1,44	NA	-2,20				
SNORD104	-1,47	NA	-2,18				
TNFAIP3	-1,64	NA	-2,18				
C4orf22	-1,44	NA	-2,17				
LOC152225	NA	-1,51	-2,16				
SNORA76	-1,42	NA	-2,16				
POLR2A	-1,49	NA	-2,15				
LOC100127888	-1,66	NA	-2,14				
LOC286009	NA	NA	-2,13				
JUN	-1,47	NA	-2,12				
ACSL4	-1,49	NA	-2,12				
EIF2D	-1,62	NA	-2,09				
IL13RA2	-1,38	NA	-2,09				
DDX58	-1,71	NA	-2,08				
DMBT1	-1,91	-1,84	-2,08				
BEND7	NA	NA	-2,07				
CNOT6	-1,68	NA	-2,07				
SOX9	-1,79	NA	-2,07				
HIST1H4B	NA	NA	-2,06				
SNORD14E	-1,53	NA	-2,06				
RAP2B	-1,34	NA	-2,05				
IL24	-1,32	-1,75	-2,05				

Annexe 1 : Liste des gènes identifiés comme étant induits par Fra-1 et/ou Fra-2 (FC ≤-2 dans au moins une condition et FDR<0,05). Les valeurs de fold-change indiquées correspondent à la modulation mesurée entre la condition siRNA spécifique versus sa condition contrôle non-transfectée (NT) associée. Les gènes sont classés par ordre croissant pour siFra-1+siFra-2 vs NT.

Gènes down-régulés par Fra-1 et/ou Fra-2 (FC ≥2 ; FDR<0,05)				Gènes down-régulés par Fra-1 et/ou Fra-2 (FC ≥2 ; FDR<0,05)			
GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT	GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT
SCEL	3,71	NA	12,45	TES	NA	NA	2,26
TNFSF18	6,17	NA	11,86	SORBS2	1,69	1,34	2,25
LIPH	3,01	NA	7,09	SGCB	1,52	NA	2,24
KLRC3	2,59	NA	7,05	IFIH1	NA	NA	2,23
F2RL2	2,15	2,53	6,46	TMEM140	NA	NA	2,23
TGFB2	3,43	NA	5,23	PCDH89	NA	NA	2,23
FLJ22447	2,16	NA	4,96	MAP3K1	1,70	NA	2,23
KLRC4-KLRC1	2,28	NA	4,94	PLXDC2	1,77	NA	2,23
PLD1	2,65	NA	4,86	ARHGAP23	1,79	NA	2,23
IL8	2,17	NA	4,38	TMEM139	1,79	NA	2,22
NSAP11	NA	NA	4,24	SLC46A3	1,75	NA	2,22
PLP	3,55	NA	4,11	SOX4	1,48	1,39	2,22
ID3	2,07	NA	4,00	LOC100131096	NA	NA	2,22
GABRE	2,29	NA	3,97	DNAJB4	1,55	NA	2,21
ID1	2,30	NA	3,91	KIAA1432	1,57	NA	2,21
ANKFN1	2,72	NA	3,86	WWC3	1,45	NA	2,21
CFH	2,52	NA	3,86	F13A1	2,35	NA	2,21
LOC100653350	2,26	NA	3,84	C6orf165	1,80	NA	2,20
ABCC3	1,79	1,97	3,84	CDC42BPA	1,78	NA	2,20
TSPAN2	2,64	NA	3,64	PLXNA1	1,66	NA	2,19
HES1	2,22	NA	3,57	LINC00452	1,96	NA	2,18
MAP2K6	1,58	2,61	3,56	FAM26E	1,35	NA	2,18
LOC100507539	1,86	NA	3,46	ALDH1L2	NA	NA	2,17
LOC100506670	2,16	NA	3,45	MIR328	NA	NA	2,17
KRT80	1,74	1,62	3,41	FLI46906	2,04	NA	2,16
CYP1A1	NA	NA	3,40	CC2D2A	1,45	NA	2,16
PHGDH	2,44	NA	3,36	C19orf21	1,49	NA	2,15
C12orf39	2,27	NA	3,33	CREB3L3	NA	2,91	2,14
LOC100128054	2,75	NA	3,31	CPE	1,45	NA	2,14
SLAMF7	2,21	1,75	3,19	SLC29A3	1,61	NA	2,13
LRP5	2,32	NA	3,19	SLCO4C1	2,05	NA	2,12
GXYLT2	2,24	NA	3,09	STXBP6	1,80	NA	2,12
SLC44A5	1,60	NA	3,08	MIR1299	NA	NA	2,12
INHBB	1,83	1,56	3,02	PLCB1	1,57	NA	2,12
LINC00357	2,02	NA	3,01	ZNF197-AS1	NA	NA	2,11
TMEM156	1,75	NA	2,91	MIR181B1	1,74	NA	2,11
HPSE	1,81	NA	2,91	SDC2	1,56	1,38	2,10
TIMP4	1,58	1,66	2,91	PDIA3P	NA	NA	2,10
SCIN	1,57	NA	2,90	ZAK	1,47	NA	2,10
PLXND1	1,48	NA	2,88	FOLR1	1,55	NA	2,10
CCDC148	2,16	NA	2,87	SDC4	1,67	NA	2,09
ARHGAP40	1,96	NA	2,82	CLIC2	1,83	1,47	2,09
HEG1	1,91	1,30	2,81	GPR65	NA	NA	2,09
BFSF1	1,70	NA	2,77	PTPN13	1,99	NA	2,08
KCNH1	1,63	NA	2,74	SLC41A2	1,48	NA	2,08
KRT7	1,85	NA	2,73	CNN2	1,73	NA	2,08
LOC100506178	2,07	NA	2,73	NFE2L3	1,63	NA	2,08
KRT81	NA	NA	2,73	ARHGAP24	1,44	NA	2,08
SYT11	NA	2,02	2,72	TMC6	1,39	NA	2,08
SOGA3	1,29	NA	2,71	AHCYL2	1,51	NA	2,07
LPCAT2	1,96	NA	2,70	FAT4	NA	NA	2,07
SYNPO	2,29	NA	2,70	SCAND3	1,68	NA	2,07
ENPP5	2,00	NA	2,68	GULP1	1,64	NA	2,06
CLU	1,90	1,42	2,68	ATP7A	1,46	NA	2,05
ME1	1,75	NA	2,67	LOC728463	NA	NA	2,05
HSPG2	1,78	NA	2,66	LOC729974	1,68	NA	2,04
LOC100652886	NA	NA	2,66	CABLES1	NA	NA	2,04
TSPAN13	1,64	1,64	2,65	GALNT10	1,75	NA	2,04
MAP2	NA	NA	2,65	SFTA3	1,50	NA	2,03
SSBP2	1,68	NA	2,64	SGPL1	1,77	NA	2,03
PIK3C2B	1,55	NA	2,63	PPFIA2	1,33	1,58	2,03
PTPRJ	1,73	NA	2,61	BANK1	1,46	NA	2,02
GDPD1	2,01	NA	2,61	SEMA3C	NA	NA	2,02
CHRNA1	1,75	NA	2,60	AGR2	NA	1,47	2,01
CA11	NA	NA	2,57	JUP	1,42	NA	2,01
KLHL4	NA	2,03	2,57	IL13RA1	1,54	NA	2,00
KLRC2	1,90	NA	2,55	TBC1D8B	NA	NA	2,00
STK31	1,85	1,50	2,54	SPTSSB	NA	NA	2,00
MEGF9	1,63	NA	2,53	PCDH814	1,33	NA	2,00
SMAD6	NA	NA	2,52	ADAM17	1,64	NA	2,00
DTNA	1,71	NA	2,51	GPC6	2,05	NA	1,97
GALNT5	1,45	1,87	2,49	NNMT	2,15	NA	1,77
EDN1	1,94	NA	2,49	C4BPB	2,06	NA	1,69
VGLL3	1,94	NA	2,48				
NTN4	1,68	NA	2,46				
C10orf91	1,91	NA	2,45				
CTSO	1,90	NA	2,43				
ICAM1	2,22	NA	2,42				
SLC7A5	1,82	NA	2,42				
TBC1D2B	1,60	NA	2,41				
SCN9A	NA	NA	2,41				
NEBL	1,43	NA	2,41				
NAALAD2	1,69	NA	2,40				
PLCE1	1,30	NA	2,36				
TMOD2	1,41	1,32	2,36				
FZD4	1,72	NA	2,36				
PIR-FIGF	NA	NA	2,34				
PLSCR4	1,72	NA	2,33				
NQO1	NA	NA	2,33				
GPR137C	1,42	NA	2,32				
SERPINF1	2,01	NA	2,32				
CDS1	1,82	NA	2,32				
LOC643650	1,37	NA	2,30				
TPTE2P1	1,75	NA	2,29				
CAMK2D	1,55	NA	2,28				
ZNF608	2,17	NA	2,28				
NCAM2	1,31	NA	2,28				
SLC1A3	1,77	NA	2,27				
THBS1	1,43	NA	2,27				

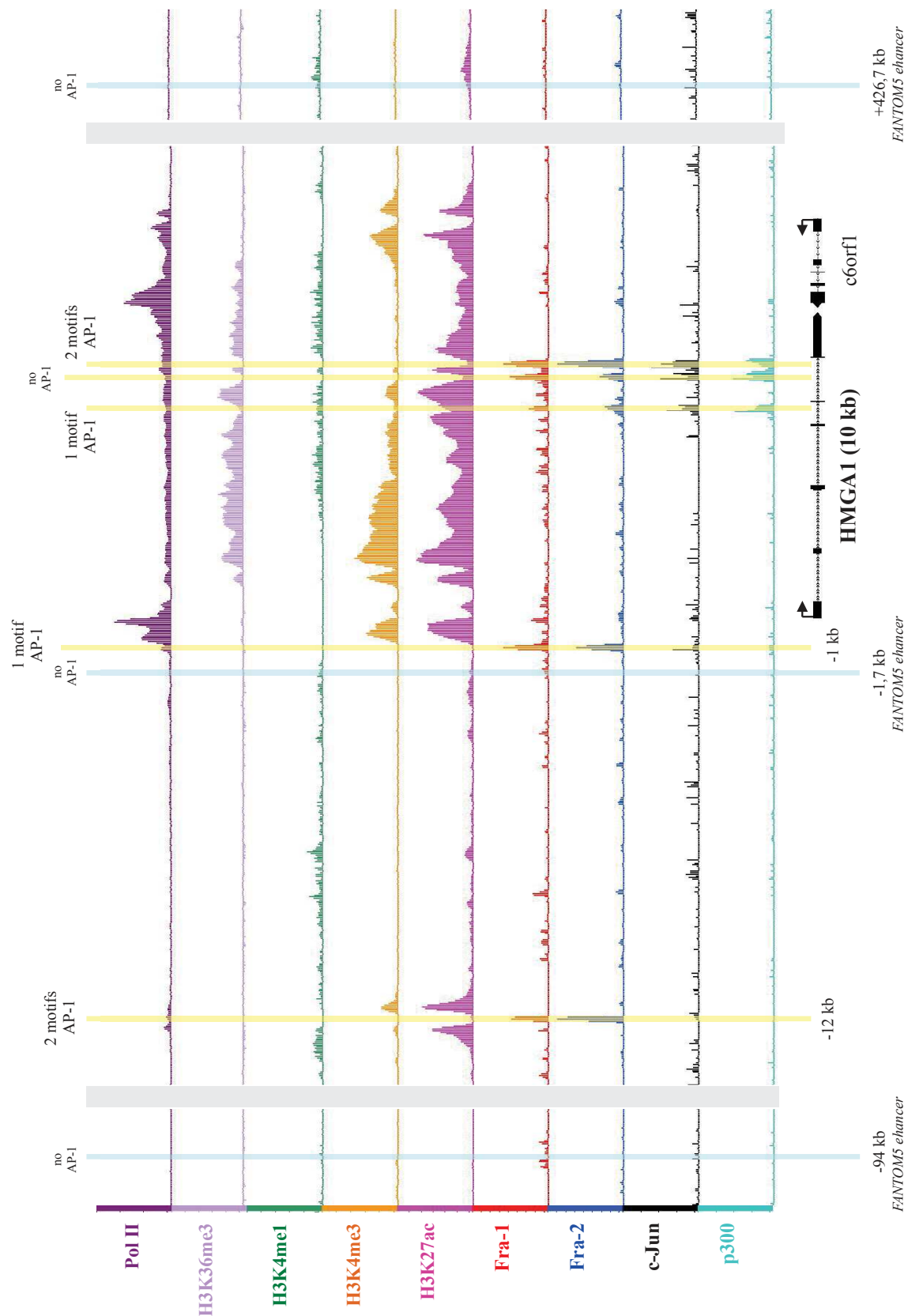
Annexe 2 : Liste des gènes identifiés comme étant réprimés par Fra-1 et/ou Fra-2 (FC ≥2 dans au moins une condition et FDR<0,05). Les valeurs de fold-change indiquées correspondent à la modulation mesurée entre la condition siRNA spécifique versus sa condition contrôle non-transfectée (NT) associée. Les gènes sont classés par ordre décroissant pour siFra-1+siFra-2 vs NT.

Gènes modulés par Fra-1 et/ou Fra-2 avec des pics Fra-1 et/ou Fra-2 dans leur(s) région(s) enhancers(s) (FC ≥1,5/≤-1,5 ; FDR<0,05)				Gènes modulés par Fra-1 et/ou Fra-2 avec des pics Fra-1 et/ou Fra-2 dans leur(s) région(s) enhancers(s) (FC ≥1,5/≤-1,5 ; FDR<0,05)			
GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT	GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT
ABHD6	1.3	1.5	1.53	MAP3K7	-1.3	NA	-1.63
ABLIM1	NA	NA	1.82	MAPKAPK3	NA	NA	-1.72
ABLIM3	-1.64	NA	-2.32	MATN3	1.47	NA	1.97
ABTB2	1.38	NA	1.98	MCL1	-1.44	NA	-1.71
ACSL4	-1.49	NA	-2.12	MGLL	-1.28	NA	-1.61
ADAM12	NA	1.36	1.56	MIR222	NA	NA	-2.96
ADAM17	1.64	NA	2	MMP14	-1.37	NA	-1.55
ADAM19	1.51	1.56	1.51	MSN	-1.19	NA	-1.53
ADAMTS6	-2.24	-1.55	-2.57	MXD1	-1.64	NA	-2.76
ADAMTSL1	-1.32	NA	-2.54	MXI1	1.58	NA	1.59
ANKRD46	1.33	NA	1.59	MYADM	-1.37	NA	-1.64
APBB1IP	-1.36	NA	-2	MYO10	-1.4	NA	-2.2
ARHGEF17	1.43	NA	1.53	NAV2	-1.69	NA	-1.82
ARRDC4	NA	NA	1.5	NCEH1	-1.33	NA	-1.67
ASAP1	-1.35	NA	-1.97	NCOA3	1.36	NA	1.75
ASB1	-1.42	NA	-1.93	NF2	NA	NA	-1.7
ASL	-1.34	NA	-1.72	NFATC2	NA	NA	-1.56
ATG7	-1.25	NA	-1.55	NFKBIA	NA	NA	1.56
ATL3	NA	-1.37	-1.58	NQO2	NA	NA	1.69
BACH1	NA	NA	-1.67	NR2F6	NA	NA	1.67
BCAS1	NA	NA	1.68	NRG1	-1.32	NA	-1.68
BZW1	-1.23	NA	-1.58	P2RY2	-1.34	NA	-1.77
C17orf53	NA	NA	-1.52	PALLD	1.29	NA	1.63
CAMK1D	NA	1.39	1.62	PDE2A	-1.68	-1.34	-2.89
CD22	-2.31	NA	-5.31	PIK3C2B	1.55	NA	2.63
CD44	-1.44	NA	-1.91	PITPNC1	-1.87	NA	-2.79
CD47	NA	NA	1.62	PLBD2	-1.7	NA	-2.3
CD83	-1.36	NA	-1.8	PP1F	NA	NA	-1.55
CDC27	NA	NA	1.52	PRDM8	-1.58	NA	-1.77
CDCP1	-1.68	NA	-2.99	PTPRE	-1.43	NA	-2
CDKN1B	1.46	NA	1.98	PYGB	NA	1.46	1.69
CEP112	1.46	NA	1.76	RAB8B	NA	-1.45	-2.39
CLYBL	1.44	NA	1.65	RHOB	NA	NA	1.89
CNNM2	NA	NA	1.51	RIN2	-1.67	NA	-1.8
COG1	-1.3	NA	-1.59	RPL29	NA	NA	-1.65
COL1A1	1.48	NA	1.69	RPLP1	NA	NA	-1.51
CREB3L1	NA	NA	1.7	RPSAP52	-1.87	NA	-2.69
DAGLA	1.29	1.33	1.55	SAMD12	-1.6	NA	-1.64
DBF4B	NA	NA	-1.52	SASH1	-1.83	NA	-2.34
DHX40	1.27	NA	1.57	SDC2	1.56	1.38	2.1
DOCK10	NA	NA	-1.69	SEC61A2	1.36	NA	1.68
DOCK3	1.48	NA	1.85	SEMA3C	NA	NA	2.02
DOCK4	-1.68	NA	-2.6	SKIL	1.29	NA	1.62
DRG2	-1.28	-1.45	-1.82	SLC1A3	1.77	NA	2.27
DTNA	1.71	NA	2.51	SLC22A3	-1.66	NA	-2.47
DUSP4	1.45	NA	1.58	SLC22A5	1.66	NA	1.76
DUSP5	-1.26	NA	-1.66	SLC25A19	NA	NA	-1.54
DUSP7	-1.44	NA	-1.9	SLC41A2	1.48	NA	2.08
EDN1	1.94	NA	2.49	SMAD6	NA	NA	2.52
EGFR	-1.58	NA	-1.79	SORBS2	1.69	1.34	2.25
EHF	1.38	NA	1.85	SRGAP1	-1.49	NA	-1.95
EIF5A2	NA	NA	1.74	SSFA2	-1.84	NA	-2.92
EMP1	NA	NA	-2.23	SSH1	NA	NA	1.55
ERCC1	-1.39	NA	-1.74	STRAP	-1.34	NA	-1.61
ERO1LB	1.34	NA	1.78	TBC1D8	NA	NA	1.66
ETS1	-1.32	NA	-1.71	TCF7L2	1.3	NA	1.69
EVI2B	-1.27	NA	-1.55	TES	NA	NA	2.26
FAM89B	NA	NA	1.58	TGFA	NA	NA	1.55
FAM8A1	1.67	NA	1.83	TGFB2	3.43	NA	5.23
FANCE	-1.29	NA	-1.65	TGFB1	NA	NA	1.99
FBP1	-1.41	NA	-2.83	TGFB3	NA	1.52	1.55
FBXO32	NA	NA	-1.67	TMC6	1.39	NA	2.08
FCF1	NA	NA	-1.55	TMCO3	NA	NA	1.62
FCHO2	NA	-1.65	-1.57	TMEM158	-2.52	NA	-4.91
FEZ2	NA	NA	-1.58	TMEM63B	NA	NA	1.51
FNDC3B	1.59	NA	1.9	TMEM87B	1.54	NA	1.69
FOSL1	-2.69	NA	-6.39	TMEM92	1.29	NA	1.55
FOSL2	NA	-1.76	-1.68	TMOD2	1.41	1.32	2.36
FPGS	NA	NA	-1.54	TNFAIP3	-1.64	NA	-2.18
FRMD6	-1.37	NA	-1.67	TNS3	1.42	NA	1.9
FXYS	NA	NA	-1.56	TPCN1	NA	NA	1.51
GADD45B	-1.27	NA	-1.81	TRIP6	NA	NA	-1.5
GALNT2	1.57	NA	1.98	TSPAN2	2.64	NA	3.64
GALNT5	1.45	1.87	2.49	TUBD1	NA	NA	1.63
GSTO2	NA	NA	1.79	VAMP3	NA	-1.37	-1.99
HECW2	1.27	NA	1.71	VEP1	-1.8	NA	-3.05
HEG1	1.91	1.3	2.81	ZBTB38	-1.5	NA	-1.45
HES1	2.22	NA	3.57	ZFP36L1	-1.3	NA	-1.78
HGS	NA	NA	-1.85	ZNF253	NA	NA	1.66
HK2	-1.41	NA	-2.42	ZNF554	NA	NA	1.55
HMG2	NA	NA	-1.85	ZNF608	2.17	NA	2.28
HYI	-1.43	NA	-1.51				
IER5	-1.33	NA	-1.59				
IL31RA	-1.44	NA	-1.62				
IL8	2.17	NA	4.38				
IRF2BPL	1.55	NA	1.52				
KRT34	NA	NA	-1.55				
KRT7	1.85	NA	2.73				
KRT80	1.74	1.62	3.41				
LAMC1	1.7	NA	1.94				
LEPREL1	1.56	NA	1.68				
LLGL2	1.7	NA	1.4				
LOC646329	NA	NA	-2.25				
LOC81691	-1.39	NA	-1.53				
LOX	1.45	1.31	1.91				
LRP10	1.28	NA	1.8				
LRP12	1.43	NA	1.89				
LRR1	NA	NA	-1.58				

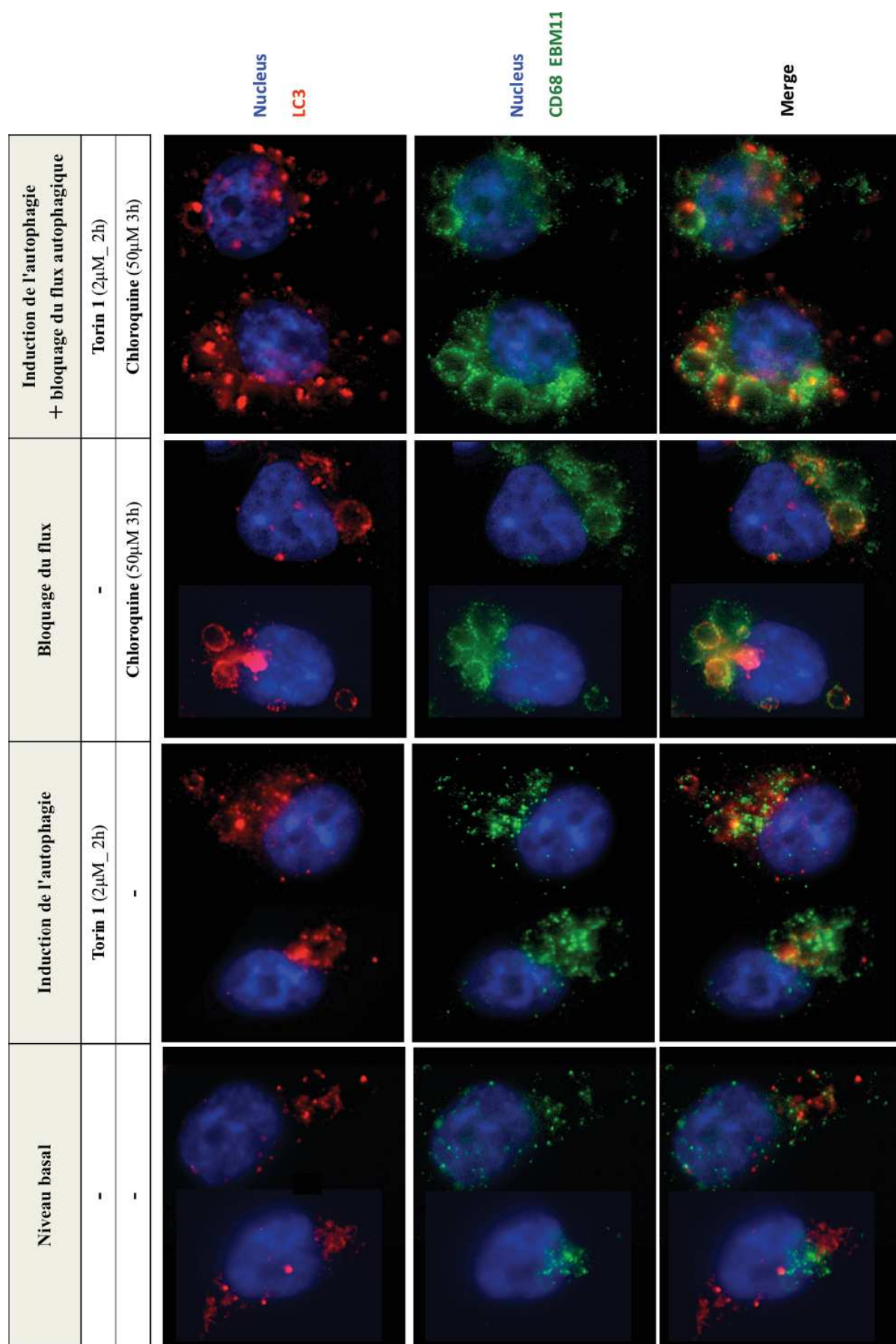
Annexe 3 : Liste des gènes identifiés comme étant modulés par Fra-1 et/ou Fra-2 (FC ≥1,5/≤-1,5 dans au moins une condition et FDR<0,05) et ayant des pics pour Fra-1 et/ou Fra-2 au niveau de leur(s) région(s) enhancer(s). Les valeurs de fold-change indiquées correspondent à la modulation mesurée entre la condition siRNA spécifique versus sa condition contrôle non-transfectée (NT) associée. Les gènes sont classés par ordre alphabétique.

Gènes modulés par Fra-1 et/ou Fra-2 avec des pics Fra-1 et/ou Fra-2 dans leur région promoteur (FC ≥1,5/≤-1,5 ; FDR<0,05)			
GnSymbol	siFra-1 vs NT	siFra-2 vs .NT	siF1+siF2 vs NT
ABCC3	1.79	1.97	3.84
ACADVL	-1.54	NA	-1.59
ACTR1B	NA	NA	-1.59
AGTPBP1	1.3	NA	1.78
ARNTL2	-1.77	NA	-2.8
BCL2A1	-2.13	NA	-3.63
BEND7	NA	NA	-2.07
BFSF1	1.7	NA	2.77
C10orf54	1.43	NA	1.98
C19orf33	-1.55	NA	-1.87
CCDC121	NA	NA	1.71
CD3EAP	NA	NA	-1.64
CDC27	NA	NA	1.52
CLCF1	-1.91	NA	-1.82
CLU	1.9	1.42	2.68
CMIP	1.57	NA	1.69
EFCAB5	1.27	NA	1.6
EIF5A2	NA	NA	1.74
EMP1	NA	NA	-2.23
ERCC1	-1.39	NA	-1.74
ETS1	-1.32	NA	-1.71
F13A1	2.35	NA	2.21
FGF1	-1.34	NA	-1.86
FOSL1	-2.69	NA	-6.39
FOSL2	NA	-1.76	-1.68
GAB2	NA	NA	-1.54
GADD45B	-1.27	NA	-1.81
GSPT1	NA	NA	-1.52
HEG1	1.91	1.3	2.81
HGS	NA	NA	-1.85
HIST1H4B	NA	NA	-2.06
HK2	-1.41	NA	-2.42
HMGAI	-1.62	NA	-2.21
HRH1	NA	NA	-1.67
IL15RA	NA	-1.53	-1.73
IL24	-1.32	-1.75	-2.05
IL31RA	-1.44	NA	-1.62
KIF20B	-1.26	NA	-1.58
KRT80	1.74	1.62	3.41
LAMB3	NA	NA	-1.85
LAMC1	1.7	NA	1.94
LAPTM5	NA	NA	-1.5
LIMK1	NA	NA	1.89
LPXN	-1.71	NA	-2.31
LY6K	-1.34	NA	-1.93
MAP1LC3C	1.37	NA	1.72
MARCH4	NA	NA	-2.01
MCL1	-1.44	NA	-1.71
NCEH1	-1.33	NA	-1.67
NFATC2	NA	NA	-1.56
NME1-NME2	NA	NA	-1.69
NPTN	1.52	NA	1.71
NUAK1	1.98	NA	1.78
PDCD1LG2	NA	NA	-1.67
PDE2A	-1.68	-1.34	-2.89
PPP1R15A	NA	NA	-1.55
PSMCS	NA	NA	-1.52
RAB3D	NA	NA	1.82
RNF168	-1.41	NA	-1.9
SLC16A5	1.59	NA	1.84
SLC1A3	1.77	NA	2.27
SLMAP	NA	-1.53	-1.57
SPNS2	-1.49	NA	-1.51
SSH1	NA	NA	1.55
SSH2	-1.34	NA	-1.56
SYNPO	2.29	NA	2.7
TAGLN2	1.37	NA	1.53
TES	NA	NA	2.26
THBS1	1.43	NA	2.27
THOC6	-1.44	NA	-1.53
TMEM156	1.75	NA	2.91
TMEM87B	1.54	NA	1.69
TNFAIP3	-1.64	NA	-2.18
TOMM34	-1.59	NA	-1.85
UBE2D1	-1.52	NA	-1.87
UHRF1	-1.26	NA	-1.84
ULBP2	NA	NA	1.51
WASF2	-1.52	-1.29	-2.94
XPO6	NA	-1.36	-1.95
ZNF227	1.42	NA	1.59
ZNF235	NA	NA	1.68

Annexe 4 : Liste des gènes identifiés comme étant modulés par Fra-1 et/ou Fra-2 (FC ≥1,5/≤-1,5 dans au moins une condition et FDR<0,05) et ayant des pics pour Fra-1 et/ou Fra-2 au niveau de leur région promoteur. Les valeurs de fold-change indiquées correspondent à la modulation mesurée entre la condition siRNA spécifique versus sa condition contrôle non-transfectée (NT) associée. Les gènes sont classés par ordre alphabétique.



Annexe 6 : Cartographie des données de ChIP-seq au niveau du locus de HMGA1 dans la lignée MDA-MB231 wild-type. Cartographie des sites de fixation de Fra-1, Fra-2, c-Jun et des marques caractéristiques des promoteurs et des enhancers actifs : H3K4me3, H3K4me1 H3K27ac, Pol II et P300. La cartographie de H3K36me3 (déposée par la forme élongatrice de Pol II) permet de distinguer le niveau de transcription des gènes. Les positions des enhancers Fantom5 sont indiquées en bleu et les positions des pics Fra-1/Fra-2 en jaune. Les barres grises indiquent des ruptures dans le chromosome pour pouvoir visualiser les enhancers FANTOM5 distaux (-94kb et +426.7kb).



Annexe 7 : Détection des protéines CD68 et LC3, dans les cellules MDA-MB231. Immunofluorescence sur cellules fixées.

MATÉRIELS ET MÉTHODES

MATÉRIELS ET MÉTHODES

Culture cellulaire :

Les lignées cellulaires de cancer du sein MDA-MB231, MDA-MB436, SUM-159PT et MCF7 proviennent de l'ATCC (American Type Culture Collection, Rockville, MD, USA). Les cellules MDA-MB231, MDA-MB436 et MCF7 sont cultivées dans du milieu DMEM (Dulbecco's modified Eagles) supplémenté avec 10% de sérum de veau foetal et un cocktail streptomycine/pénicilline, dans une chambre humide à 37°C et 5% CO₂. Les cellules SUM-159PT sont cultivées dans du milieu F-12 HAM en présence de 5% de sérum fœtal de veau, 10µg/ml insuline, 1µg/lm hydrocortisone et streptomycine/pénicilline. Toutes les cellules sont testées en routine pour la contamination aux mycoplasmes.

Transfection transitoire de siRNA :

Les cellules sont transfectées avec des pools de siRNA, soit à une concentration finale de 4,5nM pour chacun des siRNA (siCTL, siFra-1, siFra-2 et siCD68), soit à une concentration finale de 9nM pour le siCTL' correspondant à la condition contrôle des pools siFra-1+siFra-2 pendant 72h. Les cellules non-transfectées ont reçu la quantité correspondante d'agent de transfection. L'agent de transfection Interferin (Polyplus) est utilisé suivant le protocole du fabricant.

Pool siRNA	siRNA	Brin	Sequence (5' - 3')
Pool siCTL Darmacon	siCTL-A	S	AUGAACGUGAAUUGCUCUAA-dTdT
	siCTL-A	AS	UUGAGCAAUUCACGUUCAU-dTdT
	siCTL-B	S	AGGCUAUGAAGAGAGAUAC-dTdT
	siCTL-B	AS	GUAUCUCUCUUCAUAGCCU-dTdT
	siCTL-C	S	AUGAACGUGAAUUGCUCUAA-dTdT
	siCTL-C	AS	CUAAUACAGGCCAAUACAU-dTdT
	siCTL-D	S	UAGCGACUAAACACAUCAA-dtdT
	siCTL-D	AS	UUGAUGUGUUUAGUCGCUA-dTdT
Pool siCTL Ambion	Silencer® Control siRNAs. Aucune donnée fournie par le fabricant.		
Pool siFra-1	siFra-1-A	S	CCAAAGCUGCCCACUGUUU-dTdT
	siFra-1-A	AS	AAACAGUGGGCAGCUUUGG-dTdT
	siFra-1-B	S	GACAGUAUCCCACAUCCAA-dTdT
	siFra-1-B	AS	UUGGAUGUGGGAUACUGUC-dTdT
	siFra-1-C	S	CUGACUGCCACUCAUGGUG-dTdT
	siFra-1-C	AS	CACCAUGAGUGGCAGUCAG-dTdT
Pool siFra-2	Fra-2-A	S	CACGCAUGCUCAGUGCCUU-dTdT
	Fra-2-A	AS	AAGGCACUGAGCAUGCGUG-dTDT
	Fra-2-B	S	CCUCCAUGUCCAACCAUACC-dTdT
	Fra-2-B	AS	GGUAUGGGUUGGACAUGGAGG-dTdT
	Fra-2-C	S	GGAUUAUCCCGGAACUUU-dTdT
	Fra-2-C	AS	AAAGUUCCCGGGAUAAUCC-dTdT
Pool siCD68	siCD68-A siCD68-B siCD68-C	siRNA (h): sc-35019 SANTA CRUZ BIOTECHNOLOGY	

Immunoblot :

Les cellules sont directement lysées dans le tampon de lyse Laemmli (125 mM Tris (pH 6.8), 200 mM DTT, 20% Glycerol, 4% SDS et bleu de bromophénol). Les lysats sont ensuite dénaturés pendant 15 min à 90°C. Les protéines sont fractionnées sur un gel de polyacrylamide en gradient (7-20% PAGE-SDS). Elles sont ensuite électro-transférées en milieu liquide sur une membrane de polyvinylidene fluoride (PVDF). Pour la détection en immunoblot, la membrane est bloquée sur la nuit à 4°C dans une solution de PBS contenant 0,1% de Tween-20 et 10% de lait. La membrane est incubée avec l'anticorps primaire dans une solution de PBS contenant 0,1% de Tween-20, 5% de lait pendant 2h à température ambiante. La membrane est ensuite incubée à température ambiante pendant 2h avec l'anticorps secondaire ad hoc couplé à la peroxidase (HRP). La révélation est réalisée en utilisant le kit Luminata Forte (Milipore_WBLUF0500).

Anticorps Immunoblot	Reference	Dilution
GAPDH	FL335 sc.25778	1/1000
Fra-1	N17 sc.183	1/1000
Fra-2	Q20 sc.604	1/500
c-Fos	(4)-G sc.52.G	1/1000
FosB	H75 sc.7203	1/500
c-Jun	H79 sc.1694	1/1000
JunD	329 sc.74	1/500
JunB	C37F9 06/2013	1/1000
CD68	C9 sc.393951	1/500

Extraction ARN :

Pour les microarrays : Les ARNs totaux sont extraits avec le miRNeasy Mini Kit (Qiagen : cat. no. 217004), selon le protocole fourni par le fabricant. L'ADN contaminant est éliminé par une étape de digestion à la DNase à l'aide du RNase-Free DNase Set (Qiagen : cat. no. 79254), selon le protocole fourni par le fabricant. La quantité d'ARN extrait est quantifiée en mesurant l'absorbance à 260nm et la qualité des ARN est vérifiée par le ratio de l'absorbance à 260nm/280nm ainsi que sur puce Agilent. Les échantillons ont ensuite été envoyés à la plateforme d'analyse transcriptomique du CRP au Luxembourg.

Pour la validation par RT-pPCR : Les ARNm sont extraits avec le GenElute Mammalian Total RNA kit (Sigma ; RTN35), selon le protocole fourni par le fabricant. Après une étape de digestion à la DNase I réalisée selon le protocole du fabricant (BioLabs), la quantité d'ARN extrait est quantifiée en mesurant l'absorbance à 260nm et la qualité des ARN a été vérifiée par le ratio de l'absorbance à 260nm/280nm.

RT-qPCR :

La transcription inverse est réalisée sur 1 µg d'ARN, en utilisant le Maxima First Strand cDNA Synthesis Kit (Thermo Scientific), selon les instructions du fabricant. L'ADNc obtenu est dilué au 1/10 et 5µl d'échantillon sont utilisés pour analyse par PCR semi-quantitative sur l'appareil LightCycler 480 Real-Time qPCR (Roche) en utilisant la taq platinum (Invitrogen) et un mix de sybergreen préparé à l'IGMM. Les couples de primers utilisés sont conçus pour permettre une hybridation à une température de 70°. La quantification relative de l'expression des gènes est réalisée à l'aide la méthode "advanced relative quantification" Roche qui effectue les calculs au plus juste en tenant compte des paramètres de la gamme standart pour chaque couple de primers donné. Les quantités relatives des ARN sont ensuite normalisées par rapport à celles obtenues pour un gène de référence ne variant pas dans les différentes conditions analysées, ici la sous-unité ribosomique S26. Les valeurs sont ensuite normalisées par rapport à l'une des conditions arbitrairement fixée à 1.

RNA qPCR	Primers	Brin	Sequence (5' - 3')
S26	S26-1	S	CTGCACTAACTGTGCCCGATGCGTG
	S26-1	AS	GACGCTCGCTTCAGAAATGTCCCTG
Fra-1	Fra1-1	S	TCAGCTCATCGCAAGAGTAGCAGCAGC
	Fra1-1	AS	GGAGGAGACATTGGCTAGGGTGGCATC
	Fra1-2	S	GGCCTTGTGAACAGATCAGCC
	Fra1-2	AS	TTCTCCGGTTCCTGCACTTG
Fra-2	Fra2-1	S	GTCGGCCACGGGAAGTTTATT
	Fra2-1	AS	CCGATTTTCAGACAGGGCCAAA
	Fra2-2	S	GGCCAGCAGAAATTCGGGTA
	Fra2-2	AS	GAGGGTATGGGTGGACATGG
c-Fos	cFos-1	S	TGAGAAGCCAAGACTGAGCCG
	cFos-1	AS	TCATCGTGGCGTTAGGCAAA
	cFos-2	S	TTCATTCCCACGGTCACTGCC
	cFos-2	AS	GCCTCCTGTCATGGTCTTCAC
FosB	FosB-1	S	TTTATTACTCCCCTCCCCCG
	FosB-1	AS	GCCGCTGTACCTCTTGAAAGT
	FosB-2	S	CAGTACTCCACACCAGGCAT
	FosB-2	AS	GTCTCCTCTCGGGGTCTCCTA
JunB	JunB-1	S	CCCCTACCGGAGTCTCAAAGC
	JunB-1	AS	GACAATCAGGCGTTCCAGCTC
	JunB-2	S	GTCAAAGCCCTGGACGATCTG
	JunB-2	AS	GGTTGGTGTAACGGGAGGTG
c-Jun	cJun-1	S	GAGGACCGGAGACAAGTGGCA
	cJun-1	AS	CCTTCTCTCTTGCGTGGCTC
	cJun-2	S	CCACGCAAGAGAAGAAGGACG
	cJun-2	AS	TTGTCCCCTCCTCCGTGCTA
JunD	JunD-1	S	CGCCTCATCATCCAGTCCAAC
	JunD-1	AS	TGGGGTAGAGGAAGTGTGAGC
	JunD-2	S	TCAAGGACGAGCCACAGACGG
	JunD-2	AS	CTCCTGCGTGTCCATGTCGAT
HMGA1	HMGA1-1	S	CCCAAGCCCCATCTCATCCTG
	HMGA1-1	AS	GCTGTCCAACCTAGTACGGGG
	HMGA1-2	S	TGTCCAGGTGAGGCCAAGAG
	HMGA1-2	AS	GGGCAGGGTGAGCTCCTTGTT
CD68	CD68-1	S	TGCCACCAGCCCAGGATTCAC
	CD68-1	AS	TCAGTACAGAGATGCCCCAGG
	CD68-2	S	ACCTCGACCTGCTCTCCCTGA
	CD68-2	AS	GAGAGGCAGCAAGATGGACCG

Immunoprecipitation de chromatine (ChIP) et anticorps :

Le ChIP (Chromatin ImmunoPrecipitation) est réalisé comme décrit par Gomard *et al* (2010). Brièvement, 2.10^6 cellules MDA-MB231 sont transfectées avec les pools de siRNA pendant 72h. Les cellules sont fixées au Paraformaldehyde 1% (Euromedex) pendant 5 minutes à température ambiante (TA) suivi par l'ajout de 2.5 mM glycine pendant 10 minutes à TA pour neutraliser la réaction. Les cellules sont ensuite lysées pendant 10 minutes sur glace dans une solution de lyse (5mM PIPES, 85mM KCl, 0,5% NP40, 10 mM NaButyrate, protease inhibitors) contenant 0,125% de SDS et les noyaux sont récupérés par centrifugation à 5000 rpm pendant 10 minutes à 4°C. Les noyaux sont lysés pendant 2-3h à 4°C dans le tampon NLB (5mM TrisHCl pH 7,5, 1% SDS, 10 mM EDTA, 10 mM NaButyrate, protease inhibitors) avant une sonication au Bioruptor Pico (Diagenode) dans des conditions standardisées. Après 10 cycles de sonication (30 secondes on/off), la quantité d'ADN est quantifiée en mesurant l'absorbance à 280nm et la chromatine des différentes conditions siRNA est normalisée à une valeur $A_{280}=0,13$. Pour l'étape d'immunoprécipitation de chromatine, 3µg d'anticorps sont utilisés pour 100µl de chromatine en présence de billes magnétiques Dynabeads proteinG dans un volume final de 1ml de tampon d'IP (50mM Tris-HCL pH7,5, 1,1% TritonX100, 167mM NaCl, 1mM EDTA, 0,01% SDS, 1 mM NaButyrate, protease inhibitor). Pour l'immunoprécipitation des formes phosphorylées de Pol II (p-Ser5 et p-Ser2) les billes magnétiques sont incubées au préalable pendant 20 minutes avec 3µg d'anticorps anti-mouse IgM (04-6800 InVitroGen), car les anticorps anti-pSer5 et anti-pSer2 utilisés sont des IgM. L'immunoprécipitation se fait sur la nuit à 4°C sur roue. Après lavage des billes, avec les tampons low salt, high salt, LiCL salt et TE, la chromatine immunoprécipitée est éluée par 250µl une solution 1% SDS, 100mM NaHCO₃. Après réversion du pontage à 65°C sur la nuit, l'ADN est purifié sur colonne Nucleospin gel and PCR cleanup (Macherey-Nagel). L'analyse de l'immunoprécipitation est effectuée par PCR semi-quantitative.

Tampons pour le lavage des billes :

* Low salt buffer : 50mM Tris-HCl pH7,5, 150mM NaCl, 1% TritonX100, 0,1% SDS, 1mM EDTA, 1mM NaButyrate, protease inhibitor

* High salt buffer : 50mM Tris-HCl pH7,5, 500mM NaCl, 1% TritonX100, 0,1% SDS, 1mM EDTA, 1mM NaButyrate, protease inhibitor

* LiCL buffer : 20mM Tris-HCl pH7,5, 250mM LiCl, 1% NP40, 1% Deoxycholic acid, 1mM EDTA, 1mM Na-Butyrate

* TE buffer : 10 mM Tris-HCl pH7,5, 1 mM EDTA, 0,02% Tween 20

ChIP qPCR	Primers	Brin	Sequence (5' - 3')
HMGA1	-12,63	S AS	CGGCCTGAGTCACACGCATGA ACACTTCTCTTTCCCTGCGA
	-3,82	S AS	AGGTGGGTCATTAGCTCACAG CTCTGCTGCTTCCAACCTTCG
	-3,01	S AS	TGCATGCTGGTGACACAAAGG GGGAATGGGGAGTTTGAGTCA
	-1,79	S AS	CTTGAGAGCGCGAAGTGAGAG GATTCTCAGCCTCCAGCACC
	-1,02	S AS	GGGAGGAAGAAGGGGAGCAGA CAGGGAATTGAAATGCGCTCC
	-0,32	S AS	TTTATAGTTAGGGGGCGGCGTG TCACCTCTAGGGGCTTCCAGG
	+0,5	S AS	TTAAATCCCCGGGCTCATTTG CCAATTAGAGGAGCGCAGCAC
	+1,22	S AS	CCCCACAACTGGTTTCGCTG CCCGCAAAATAAAGCCCCAGA
	+1,89	S AS	AAGTGAGGGATTGGGGTTAGG CCCTTCCAGTGAGGCCAAGA
	+3,23	S AS	CAGTTGGGCAGAGGAGGGGTA AGGGAGGAGATGGTTGAAGCT
	+4,16	S AS	TGGTAGGGAGTCAGGTGGGTG AATCATTCTGCACCCCTCCATG
	+5,63	S AS	CCCACCCAAGATAAGTGTCAG GACAAGAGCCCAGAAGACCCC
	+7,17	S AS	GCAGAAGAGAAACAGCGAAGG AGCTTCTTGACTCAGCCCACC
	+7,56	S AS	CATTCCCGCCAGTGAGTGAG TGCAGGGGAACCAAGTCCTTG
	+7,95	S AS	GTGGAGGTTGCTGAGTCACCC GGCAGTTGTTGGTGTGAGCTC
	+8,31	S AS	GCCCTCACCACCACACTACAC AGGGTGCGGCTAAGTGGGATG
	+8,62	S AS	CCCAAGCCCCATCTCATCCTG GCTGTCCAACCTAGTACGGGG
	+8,9	S AS	TGTCCAGGTGAGGCCCAAGAG GGGCAGGGTGAGCTCCTTGTT
	+9,52	S AS	CAGGTAGGATGGGTGACACTT GCTTAAATGACCTTCCCCGG
S26	-0,8	S AS	CGTGTTGTGTGTGTACTGGG TGGCTCCGCTGTACTATGCTG
	-0,03	S AS	CTGCACTAACTGTGCCCGATGCGTG GACGCTCGCTTCAGAAATGTCCCTG
	+0,64	S AS	CTGCACTAACTGTGCCCGATGCGTG GACGCTCGCTTCAGAAATGTCCCTG
	+1,39	S AS	TGGGCTGAACAGGTGCTTTGG GGGAAGCACATAGGCTAAGGA
	+2,20	S AS	GTTCTTTGGGGGAAGGGAGTC TTACATGGGCTTTGGTGGGGG

Anticoprs ChIP	Reference
GAPDH	Santa Cruz - FL335 sc25778
Fra-2	Santa Cruz - H103 sc13017
H3K4me1	Abcam - 8895
H3K4me3	Abcam - 8580
H3K27ac	Abcam - 4729
H3K36me3	Active Motif - 61101
Pol II totale	Santa Cruz - F12 sc55492
Pol II p-S5	Anti-mouse IgM (Heavy Chain Secondary Antibody) Covance - H14 MMS-134R
Pol II p-S2	Anti-mouse IgM (Heavy Chain Secondary Antibody) Covance - H5 MMS-129R
P300	Abcam - 14984 KAT3B/p300 (3G250/NM-11)

Mesure des ARN naissants (Run-On) :

Le Run-On a été réalisé comme décrit par Roberts *et al* (Nature protocols. 2015). Brièvement, 2.10^6 cellules MDA-MB231 sont transfectées avec les pools de siRNA pendant 72h. Les noyaux sont isolés à 4°C dans une solution contenant du NP-40 pour figer l'état transcriptionnel puis incubés à 37°C pendant 40 minutes en présence de ribonucléotides supplémentés par du BrdU, pour permettre la reprise de la transcription et l'élongation des transcrits en cours. Le BrdU s'incorpore dans les ARNs nouvellement transcrits par l'ARN Pol II. Les ARNs sont extraits avec le miRNeasy Mini Kit (Qiagen : cat. no. 217004), selon le protocole fourni par le fabricant. Une étape de digestion à la DNase est réalisée à l'aide du RNase-Free DNase Set (Qiagen : cat. no. 79254), selon le protocole fourni par le fabricant. La quantité d'ARN extrait est ensuite quantifiée en mesurant l'absorbance à 260nm au NanoDrop. Pour l'étape d'immunoprécipitation des ARNs lié à Pol II, 2µg d'anticorps anti-BrdU (Santa Cruz sc-32323) et 8µg d'ARN sont utilisés en présence de billes Dynabeads ProteinG. Après 30 min d'incubation sur roue à TA, les billes sont lavées et les ARNs naissants sont élués au trizol, purifiés par phénol/chloroforme et précipités à l'éthanol. Les ARNs sont resuspendus dans 30µl d'eau nuclease free. 10µl sont utilisés pour la transcription reverse avec le kit Maxima First Strand cDNA Synthesis Kit (Thermo Scientific). L'ADNc obtenu est dilué au 1/10 et 5µl sont utilisés pour analyse sur LightCycler 480 Real-Time qPCR (Roche). La quantification relative est réalisée à l'aide la méthode "advanced relative quantification" Roche. Les valeurs obtenues sont normalisées par rapport au niveau mesuré pour GAPDH. Les couples de primers utilisés sont conçus pour permettre une amplification à une température de 70°C sur le LightCycler 480 (Roche).

Run-On qPCR	Primers	Brin	Sequence (5' - 3')
GAPDH	+2,40	S	AATCCCATCACCATCTTCCAG
		AS	GAGCCACACCATCCTAGTTG
HMGA1	+1,22	S	CCCCACAACTGGTTTCGCTG
		AS	CCCGCAAAATAAAGCCCCAGA
	+3,23	S	CAGTTGGGCAGAGGAGGGGTA
		AS	AGGGAGGAGATGGTTGAAGCT
CD68	+0,53	S	ACTCACAGGACAACCACCACA
		AS	GTGAGTGGCAGTTGAGGGTCC
	+0,61	S	GCCATGGAAACGTACAGTTC
		AS	GAAGTGGTGAATCCTGGGCTG

Mesure de la viabilité cellulaire :

Les changements de polarisation de la membrane mitochondriale sont évalués en utilisant la combinaison DiOC6(3)/Pi. Les cellules sont collectées par trypsinisation et lavées dans 5 ml de PBS. 1.10^6 cellules sont ensuite incubées en présence de DiOC6(3)/Pi (80 nM DiOC6(3) et 1 µg/mL de Pi) pendant 30 minutes à 37°C dans le noir. Les échantillons sont ensuite lavés dans 5 ml de PBS avant d'être analysés par cytométrie de flux sur le FACSCalibur (BD Biosciences). La quantification est réalisée à l'aide du logiciel CellQuest Pro (BD Biosciences).

Mesure de la distribution dans le cycle cellulaire :

Suite à la mesure de la viabilité cellulaires, les échantillons sont centrifugés et les culots sont resuspendus et fixés avec une solution d'éthanol à 70% à 4°C ajoutée goutte à goutte en agitant délicatement. Les cellules sont ensuite immédiatement placées à -20°C pour au minimum une nuit. Les culots sont ensuite lavés dans 5 ml de PBS et les cellules sont incubées en présence de Pi : 50 µg/ml de Pi dilué extemporanément dans du PBS supplémenté avec 0.1% de D-glucose et 1 µg/mL de RNase A. Les cellules sont marquées pendant 30 minutes à 37°C dans le noir, puis immédiatement stockées à -20°C pour au minimum une nuit, avant d'être analysées par cytométrie de flux sur le FACSCalibur (BD Biosciences). La quantification est réalisée à l'aide du logiciel CellQuest Pro (BD Biosciences).

Immunofluorescence sur cellules fixées :

Les cellules MDA-MB231 étant adhérentes, celles-ci sont directement cultivées sur lamelles de verre après transfection ou non de siRNA. Après lavage au PBS, les cellules sont fixées soit au paraformaldehyde 4% pendant 30 minutes à TA pour le co-marquage de CD68 avec la phalloïdine, le mitotracker et le lysotracker, soit par une solution de méthanol/1% paraformaldehyde à -20°C pendant 5 minutes pour le co-marquage de CD68 et LC3. Les cellules fixées sont incubées pendant 30 minutes dans un tampon de blocage (Saponine 0.1%, Sérum de veau foetal 1%, PBS) puis avec les anticorps primaires dilués dans le tampon de blocage pendant 1h30 à TA et à l'abri de la lumière. Après lavage avec le tampon de blocage, les cellules sont incubées avec les anticorps secondaires pendant 1h à l'abri de la lumière. Après lavage au PBS, les lamelles sont incubées 1 min avec une solution de Hoechst 33342 à 0,2µg/ml pour marquer les noyaux. Après un lavage final au PBS, les lamelles sont montées sur les lames de microscope à l'aide de permafluor.

Immunofluorescence	Reference	Dilution
CD68 KP1	Monoclonal Mouse Anti-Human CD68. Clone KP1. Isotype: IgG1, kappa DAKO - M 0814	1/100
CD68 EBM11	Mouse IgG concentration: 460 ug/ml. Isotype: IgG1, kappa DAKO- M 0718	1/100
Ac secondaire CD68	Donkey anti-mouse Alexa-488	1/4000
LC3	Anti-LC3B antibody produced in rabbit Sigma - L7543	1/1000
Ac secondaire LC3	Chicken anti-rabbit Alexa-647	1/4000
Phalloïdine	Phalloïdine-Tetramethylrhodamine B isothiocyanate Sigma - P1951	1/1000
MitoTracker	MitoTracker® Orange CMTMRos Molecular Probes – M7510	1/10000
LysoTracke	LysoTracker® Red DND-99 Molecular Probes – L7528	1/20000

BIBLIOGRAPHIE

BIBLIOGRAPHIE

- Adiseshaiah, P., Lindner, D.J., Kalvakolanu, D.V., and Reddy, S.P. (2007). FRA-1 proto-oncogene induces lung epithelial cell invasion and anchorage-independent growth in vitro, but is insufficient to promote tumor growth in vivo. *Cancer Res.* 67, 6204–6211.
- Adiseshaiah, P., Li, J., Vaz, M., Kalvakolanu, D.V., and Reddy, S.P. (2008). ERK signaling regulates tumor promoter induced c-Jun recruitment at the Fra-1 promoter. *Biochem. Biophys. Res. Commun.* 371, 304–308.
- Anand, A., and Chada, K. (2000). In vivo modulation of Hmgic reduces obesity. *Nat. Genet.* 24, 377–380.
- Andersson, R. (2015). Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 37, 314–323.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* 507, 455–461.
- Andersson, R., Sandelin, A., and Danko, C.G. (2015). A unified architecture of transcriptional regulatory elements. *Trends Genet. TIG* 31, 426–433.
- Andrew, A.S., Klei, L.R., and Barchowsky, A. (2001). AP-1-dependent induction of plasminogen activator inhibitor-1 by nickel does not require reactive oxygen. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 281, L616–623.
- Ashley, J.W., Shi, Z., Zhao, H., Li, X., Kesterson, R.A., and Feng, X. (2011). Genetic ablation of CD68 results in mice with increased bone and dysfunctional osteoclasts. *PloS One* 6, e25838.
- Bakiri, L., Lallemand, D., Bossy-Wetzel, E., and Yaniv, M. (2000). Cell cycle-dependent variations in c-Jun and JunB phosphorylation: a role in the control of cyclin D1 expression. *EMBO J.* 19, 2056–2068.
- Bakiri, L., Matsuo, K., Wisniewska, M., Wagner, E.F., and Yaniv, M. (2002). Promoter specificity and biological activity of tethered AP-1 dimers. *Mol Cell Biol* 22, 4952–4964.
- Bakiri, L., Macho-Maschler, S., Custic, I., Niemiec, J., Guío-Carrión, A., Hasenfuss, S.C., Eger, A., Müller, M., Beug, H., and Wagner, E.F. (2015). Fra-1/AP-1 induces EMT in mammary epithelial cells by modulating Zeb1/2 and TGFβ expression. *Cell Death Differ.* 22, 336–350.
- Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N., and Peterlin, B.M. (2001). NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* 8, 327–337.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837.
- Basbous, J., Jariel-Encontre, I., Gomard, T., Bossis, G., and Piechaczyk, M. (2008). Ubiquitin-independent- versus ubiquitin-dependent proteasomal degradation of the c-Fos and Fra-1 transcription factors: is there a unique answer? *Biochimie* 90, 296–305.

- Baù, D., and Marti-Renom, M.A. (2012). Genome structure determination via 3C-based data integration by the Integrative Modeling Platform. *Methods* 58, 300–306.
- Becker, P.B., and Hörz, W. (2002). ATP-dependent nucleosome remodeling. *Annu. Rev. Biochem.* 71, 247–273.
- de Beer, M.C., Zhao, Z., Webb, N.R., van der Westhuyzen, D.R., and de Villiers, W.J.S. (2003). Lack of a direct role for macrosialin in oxidized LDL metabolism. *J. Lipid Res.* 44, 674–685.
- Belguise, K., Kersual, N., Galtier, F., and Chalbos, D. (2005). FRA-1 expression level regulates proliferation and invasiveness of breast cancer cells. *Oncogene* 24, 1434–1444.
- Belguise, K., Milord, S., Galtier, F., Moquet-Torcy, G., Piechaczyk, M., and Chalbos, D. (2012). The PKC θ pathway participates in the aberrant accumulation of Fra-1 protein in invasive ER-negative breast cancer cells. *Oncogene* 31, 4889–4897.
- Bell, O., Tiwari, V.K., Thomä, N.H., and Schübeler, D. (2011). Determinants and dynamics of genome accessibility. *Nat. Rev. Genet.* 12, 554–564.
- Beranek, J.T. (2005). CD68 is not a macrophage-specific antigen. *Ann. Rheum. Dis.* 64, 342–343–344.
- Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. *Science* 281, 60–63.
- Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R., et al. (2005). Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol.* 3, e157.
- Borsig, L., Wolf, M.J., Roblek, M., Lorentzen, A., and Heikenwalder, M. (2014). Inflammatory chemokines and metastasis--tracing the accessory. *Oncogene* 33, 3217–3224.
- Bossis, G., Ferrara, P., Acquaviva, C., Jariel-Encontre, I., and Piechaczyk, M. (2003). c-Fos proto-oncoprotein is degraded by the proteasome independently of its own ubiquitinylation in vivo. *Mol. Cell. Biol.* 23, 7425–7436.
- Bossis, G., Malnou, C.E., Farras, R., Andermarcher, E., Hipskind, R., Rodriguez, M., Schmidt, D., Muller, S., Jariel-Encontre, I., and Piechaczyk, M. (2005). Down-regulation of c-Fos/c-Jun AP-1 dimer activity by sumoylation. *Mol. Cell. Biol.* 25, 6964–6979.
- Bouallaga, I., Teissier, S., Yaniv, M., and Thierry, F. (2003). HMG-I(Y) and the CBP/p300 coactivator are essential for human papillomavirus type 18 enhanceosome transcriptional activity. *Mol. Cell. Biol.* 23, 2329–2340.
- Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4, e138.
- Bridge, A.J., Pebernard, S., Ducraux, A., Nicoulaz, A.-L., and Iggo, R. (2003). Induction of an interferon response by RNAi vectors in mammalian cells. *Nat. Genet.* 34, 263–264.
- Brouckaert, O., Wildiers, H., Floris, G., and Neven, P. (2012). Update on triple-negative breast cancer: prognosis and management strategies. *Int. J. Womens Health* 4, 511–520.
- Buehler, E., Chen, Y.-C., and Martin, S. (2012). C911: A bench-level control for sequence specific siRNA off-target effects. *PloS One* 7, e51942.

- Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. *Mol. Cell* 36, 541–546.
- Buratowski, S., Hahn, S., Guarente, L., and Sharp, P.A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56, 549–561.
- Burch, P.M., Yuan, Z., Loonen, A., and Heintz, N.H. (2004). An extracellular signal-regulated kinase 1- and 2-dependent program of chromatin trafficking of c-Fos and Fra-1 is required for cyclin D1 expression during cell cycle reentry. *Mol. Cell. Biol.* 24, 4696–4709.
- Bush, B.M., Brock, A.T., Deng, J.A., Nelson, R.A., and Sumter, T.F. (2013). The Wnt/ β -catenin/T-cell factor 4 pathway up-regulates high-mobility group A1 expression in colon cancer. *Cell Biochem. Funct.* 31, 228–236.
- Butler, J.E.F., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* 16, 2583–2592.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M.P., et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581–592.
- Casalino, L., Bakiri, L., Talotta, F., Weitzman, J.B., Fusco, A., Yaniv, M., and Verde, P. (2007). Fra-1 promotes growth and survival in RAS-transformed thyroid cells by controlling cyclin A transcription. *EMBO J.* 26, 1878–1890.
- Cascio, S., Bartella, V., Garofalo, C., Russo, A., Giordano, A., and Surmacz, E. (2007). Insulin-like growth factor 1 differentially regulates estrogen receptor-dependent transcription at estrogen response element and AP-1 sites in breast cancer cells. *J. Biol. Chem.* 282, 3498–3506.
- Chalmers, C.J., Gilley, R., March, H.N., Balmanno, K., and Cook, S.J. (2007). The duration of ERK1/2 activity determines the activation of c-Fos and Fra-1 and the composition and quantitative transcriptional output of AP-1. *Cell. Signal.* 19, 695–704.
- Chen, N., and Debnath, J. (2010). Autophagy and tumorigenesis. *FEBS Lett.* 584, 1427–1435.
- Chen, Y.-J., and Chang, L.-S. (2015). NF κ B- and AP-1-mediated DNA looping regulates matrix metalloproteinase-9 transcription in TNF- α -treated human leukemia U937 cells. *Biochim. Biophys. Acta* 1849, 1248–1259.
- Chen, C.-Y.A., Chang, J.T., Ho, Y.-F., and Shyu, A.-B. (2016). MiR-26 down-regulates TNF- α /NF- κ B signalling and IL-6 expression by silencing HMGA1 and MALT1. *Nucleic Acids Res.* 44, 3772–3787.
- Cheung, S.W., Tishler, P.V., Atkins, L., Sengupta, S.K., Modest, E.J., and Forget, B.G. (1977). Gene mapping by fluorescent in situ hybridization. *Cell Biol. Int. Rep.* 1, 255–262.
- Chiappetta, G., Botti, G., Monaco, M., Pasquinelli, R., Pentimalli, F., Di Bonito, M., D’Aiuto, G., Fedele, M., Iuliano, R., Palmieri, E.A., et al. (2004). HMGA1 protein overexpression in human breast carcinomas: correlation with ErbB2 expression. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 10, 7637–7644.
- Chiappetta, G., Ferraro, A., Botti, G., Monaco, M., Pasquinelli, R., Vuttariello, E., Arnaldi, L., Di Bonito, M., D’Aiuto, G., Pierantoni, G.M., et al. (2007). FRA-1 protein overexpression is a feature of hyperplastic and neoplastic breast disorders. *BMC Cancer* 7, 17.

- Chiefari, E., Arcidiacono, B., Possidente, K., Iiritano, S., Ventura, V., Pandolfo, R., Brunetti, F.S., Greco, M., Foti, D., and Brunetti, A. (2013). Transcriptional regulation of the HMGA1 gene by octamer-binding proteins Oct-1 and Oct-2. *PloS One* 8, e83969.
- Chinenov, Y., and Kerppola, T.K. (2001). Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene* 20, 2438–2452.
- Cho, H., Kim, T.K., Mancebo, H., Lane, W.S., Flores, O., and Reinberg, D. (1999). A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev.* 13, 1540–1552.
- Cho, S., Schroeder, S., Kaehlcke, K., Kwon, H.-S., Pedal, A., Herker, E., Schnoelzer, M., and Ott, M. (2009). Acetylation of cyclin T1 regulates the equilibrium between active and inactive P-TEFb in cells. *EMBO J.* 28, 1407–1417.
- Chung, I.-M., Ketharnathan, S., Kim, S.-H., Thiruvengadam, M., Rani, M.K., and Rajakumar, G. (2016). Making Sense of the Tangle: Insights into Chromatin Folding and Gene Regulation. *Genes* 7.
- Chymkowitz, P., Nguéa P, A., and Enserink, J.M. (2015). SUMO-regulated transcription: challenging the dogma. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 37, 1095–1105.
- Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78, 273–304.
- Cleynen, I., and Van de Ven, W.J.M. (2008). The HMGA proteins: a myriad of functions (Review). *Int. J. Oncol.* 32, 289–305.
- Cleynen, I., Huysmans, C., Sasazuki, T., Shirasawa, S., Van de Ven, W., and Peeters, K. (2007). Transcriptional control of the human high mobility group A1 gene: basal and oncogenic Ras-regulated expression. *Cancer Res.* 67, 4620–4629.
- Collignon, J., Lousberg, L., Schroeder, H., and Jerusalem, G. (2016). Triple-negative breast cancer: treatment challenges and solutions. *Breast Cancer Dove Med. Press* 8, 93–107.
- Conaway, R.C., and Conaway, J.W. (2011). Function and regulation of the Mediator complex. *Curr. Opin. Genet. Dev.* 21, 225–230.
- Conaway, J.W., Shilatifard, A., Dvir, A., and Conaway, R.C. (2000). Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* 25, 375–380.
- Conaway, R.C., Sato, S., Tomomori-Sato, C., Yao, T., and Conaway, J.W. (2005). The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* 30, 250–255.
- Cook, S.J., Aziz, N., and McMahon, M. (1999). The repertoire of fos and jun proteins expressed during the G1 phase of the cell cycle is determined by the duration of mitogen-activated protein kinase activation. *Mol. Cell. Biol.* 19, 330–341.
- Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848.
- Crish, J.F., and Eckert, R.L. (2008). Synergistic activation of human involucrin gene expression by Fra-1 and p300--evidence for the presence of a multiprotein complex. *J. Invest. Dermatol.* 128, 530–541.

- Cullen, B.R. (2006). Enhancing and confirming the specificity of RNAi experiments. *Nat. Methods* 3, 677–681.
- Das, R., Mahabeleshwar, G.H., and Kundu, G.C. (2004). Osteopontin induces AP-1-mediated secretion of urokinase-type plasminogen activator through c-Src-dependent epidermal growth factor receptor transactivation in breast cancer cells. *J. Biol. Chem.* 279, 11051–11064.
- Davies, J.O.J., Telenius, J.M., McGowan, S.J., Roberts, N.A., Taylor, S., Higgs, D.R., and Hughes, J.R. (2016). Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat. Methods* 13, 74–80.
- De Martino, M., Forzati, F., Arra, C., Fusco, A., and Esposito, F. (2016). HMGA1-pseudogenes and cancer. *Oncotarget* 7, 28724–28735.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* 295, 1306–1311.
- Desmet, C.J., Gallenne, T., Prieur, A., Rey, F., Visser, N.L., Wittner, B.S., Smit, M.A., Geiger, T.R., Laoukili, J., Iskit, S., et al. (2013). Identification of a pharmacologically tractable FRA-1/ADORA2B axis promoting breast cancer metastasis. *Proc. Natl. Acad. Sci.* 110, 5139–5144.
- Dhar, A., Hu, J., Reeves, R., Resar, L.M.S., and Colburn, N.H. (2004). Dominant-negative c-Jun (TAM67) target genes: HMGA1 is required for tumor promoter-induced transformation. *Oncogene* 23, 4466–4476.
- Dhillon, A.S., and Tulchinsky, E. (2015). FRA-1 as a driver of tumour heterogeneity: a nexus between oncogenes and embryonic signalling pathways in cancer. *Oncogene* 34, 4421–4428.
- Di Cello, F., Shin, J., Harbom, K., and Brayton, C. (2013). Knockdown of HMGA1 inhibits human breast cancer cell growth and metastasis in immunodeficient mice. *Biochem. Biophys. Res. Commun.* 434, 70–74.
- Diefenbacher, M.E., Reich, D., Dahley, O., Kemler, D., Litfin, M., Herrlich, P., and Kassel, O. (2014). The LIM domain protein nTRIP6 recruits the mediator complex to AP-1-regulated promoters. *PloS One* 9, e97549.
- Diesch, J., Sanij, E., Gilan, O., Love, C., Tran, H., Fleming, N.I., Ellul, J., Amalia, M., Haviv, I., Pearson, R.B., et al. (2014). Widespread FRA1-dependent control of mesenchymal transdifferentiation programs in colorectal cancer cells. *PloS One* 9, e88950.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C., et al. (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 16, 1299–1309.
- Eberhardy, S.R., and Farnham, P.J. (2002). Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J. Biol. Chem.* 277, 40156–40162.
- Eferl, R., and Wagner, E.F. (2003). AP-1: a double-edged sword in tumorigenesis. *Nat. Rev. Cancer* 3, 859–868.

- Eilebrecht, S., Le Douce, V., Riclet, R., Targat, B., Hallay, H., Van Driessche, B., Schwartz, C., Robette, G., Van Lint, C., Rohr, O., et al. (2014). HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res.* *42*, 4962–4971.
- Eng, C.H., Wang, Z., Tkach, D., Toral-Barza, L., Ugwonal, S., Liu, S., Fitzgerald, S.L., George, E., Frias, E., Cochran, N., et al. (2016). Macroautophagy is dispensable for growth of KRAS mutant tumors and chloroquine efficacy. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 182–187.
- Evans, A., Lennard, T.W.J., and Davies, B.R. (2004). High-mobility group protein 1(Y): metastasis-associated or metastasis-inducing? *J. Surg. Oncol.* *88*, 86–99.
- Farmer, P., Bonnefoi, H., Becette, V., Tubiana-Hulin, M., Fumoleau, P., Larsimont, D., Macgrogan, G., Bergh, J., Cameron, D., Goldstein, D., et al. (2005). Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene* *24*, 4660–4671.
- Farràs, R., Baldin, V., Gallach, S., Acquaviva, C., Bossis, G., Jariel-Encontre, I., and Piechaczyk, M. (2008). JunB breakdown in mid-/late G2 is required for down-regulation of cyclin A2 levels and proper mitosis. *Mol. Cell. Biol.* *28*, 4173–4187.
- Fedele, M., and Fusco, A. (2010). HMGA and cancer. *Biochim. Biophys. Acta* *1799*, 48–54.
- Fedele, M., Berlingieri, M.T., Scala, S., Chiariotti, L., Viglietto, G., Rippel, V., Bullerdiek, J., Santoro, M., and Fusco, A. (1998). Truncated and chimeric HMGI-C genes induce neoplastic transformation of NIH3T3 murine fibroblasts. *Oncogene* *17*, 413–418.
- Fedele, M., Fidanza, V., Battista, S., Pentimalli, F., Klein-Szanto, A.J.P., Visone, R., De Martino, I., Curcio, A., Morisco, C., Del Vecchio, L., et al. (2006). Haploinsufficiency of the Hmgal gene causes cardiac hypertrophy and myelo-lymphoproliferative disorders in mice. *Cancer Res.* *66*, 2536–2543.
- Federico, A., Forzati, F., Esposito, F., Arra, C., Palma, G., Barbieri, A., Palmieri, D., Fedele, M., Pierantoni, G.M., De Martino, I., et al. (2014). Hmgal/Hmga2 double knock-out mice display a “superpygmy” phenotype. *Biol. Open* *3*, 372–378.
- Ferrara, P., Andermarcher, E., Bossis, G., Acquaviva, C., Brockly, F., Jariel-Encontre, I., and Piechaczyk, M. (2003). The structural determinants responsible for c-Fos protein proteasomal degradation differ according to the conditions of expression. *Oncogene* *22*, 1461–1474.
- Fleischmann, A. (2000). Fra-1 replaces c-Fos-dependent functions in mice. *Genes Dev* *14*, 2695–2700.
- Forrest, A.R.R., Rehli, M., Baillie, J.K., de Hoon, M.J.L., Haberle, V., Lassmann, T., Kulakovskiy, I.V., Lizio, M., Itoh, M., Andersson, R., et al. (2014). A promoter-level mammalian expression atlas. *Nature* *507*, 462–470.
- Foti, D., Chiefari, E., Fedele, M., Iuliano, R., Brunetti, L., Paonessa, F., Manfioletti, G., Barbetti, F., Brunetti, A., Croce, C.M., et al. (2005). Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. *Nat. Med.* *11*, 765–773.
- Foulkes, W.D., Smith, I.E., and Reis-Filho, J.S. (2010). Triple-negative breast cancer. *N. Engl. J. Med.* *363*, 1938–1948.
- Fuchs, S.Y., Dolan, L., Davis, R.J., and Ronai, Z. (1996). Phosphorylation-dependent targeting of c-Jun ubiquitination by Jun N-kinase. *Oncogene* *13*, 1531–1535.

- Fuda, N.J., Ardehali, M.B., and Lis, J.T. (2009). Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* 461, 186–192.
- Furey, T.S. (2012). ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions. *Nat. Rev. Genet.* 13, 840–852.
- Fusco, A., and Fedele, M. (2007). Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* 7, 899–910.
- Galluzzi, L., Pietrocola, F., Bravo-San Pedro, J.M., Amaravadi, R.K., Baehrecke, E.H., Cecconi, F., Codogno, P., Debnath, J., Gewirtz, D.A., Karantza, V., et al. (2015). Autophagy in malignant transformation and cancer progression. *EMBO J.* 34, 856–880.
- Gao, Z., and Ye, J. (2008). Inhibition of transcriptional activity of c-JUN by SIRT1. *Biochem. Biophys. Res. Commun.* 376, 793–796.
- Garcia-Dominguez, M., and Reyes, J.C. (2009). SUMO association with repressor complexes, emerging routes for transcriptional control. *Biochim. Biophys. Acta* 1789, 451–459.
- Geng, Y., and Johnson, L.F. (1993). Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA-less mouse thymidylate synthase promoter. *Mol. Cell. Biol.* 13, 4894–4903.
- Giancotti, V., Berlingieri, M.T., DiFiore, P.P., Fusco, A., Vecchio, G., and Crane-Robinson, C. (1985). Changes in nuclear proteins on transformation of rat epithelial thyroid cells by a murine sarcoma retrovirus. *Cancer Res.* 45, 6051–6057.
- Giancotti, V., Pani, B., D’Andrea, P., Berlingieri, M.T., Di Fiore, P.P., Fusco, A., Vecchio, G., Crane-Robinson, C., and Goodwin, G.H. (1987). Histone and nonhistone proteins from normal and virus-transformed rat thyroid epithelial cells. *Basic Appl. Histochem.* 31, 229–238.
- Gilley, R., March, H.N., and Cook, S.J. (2009). ERK1/2, but not ERK5, is necessary and sufficient for phosphorylation and activation of c-Fos. *Cell. Signal.* 21, 969–977.
- Gomes, N.P., Bjerke, G., Llorente, B., Szostek, S.A., Emerson, B.M., and Espinosa, J.M. (2006). Gene-specific requirement for P-TEFb activity and RNA polymerase II phosphorylation within the p53 transcriptional program. *Genes Dev.* 20, 601–612.
- Gottfried, E., Kunz-Schughart, L.A., Weber, A., Rehli, M., Peuker, A., Müller, A., Kastenberger, M., Brockhoff, G., Andreesen, R., and Kreutz, M. (2008a). Expression of CD68 in non-myeloid cell types. *Scand. J. Immunol.* 67, 453–463.
- Gozuacik, D., and Kimchi, A. (2004). Autophagy as a cell death and tumor suppressor mechanism. *Oncogene* 23, 2891–2906.
- Grigoriadis, A., Mackay, A., Noel, E., Wu, P.J., Natrajan, R., Frankum, J., Reis-Filho, J.S., and Tutt, A. (2012). Molecular characterisation of cell line models for triple-negative breast cancers. *BMC Genomics* 13, 619.
- Grigoriadis, A.E., Wang, Z.Q., Cecchini, M.G., Hofstetter, W., Felix, R., Fleisch, H.A., and Wagner, E.F. (1994). c-Fos: a key regulator of osteoclast-macrophage lineage determination and bone remodeling. *Science* 266, 443–448.
- Gruver, A.M., Portier, B.P., and Tubbs, R.R. (2011). Molecular pathology of breast cancer: the journey from traditional practice toward embracing the complexity of a molecular classification. *Arch. Pathol. Lab. Med.* 135, 544–557.

- Guedj, M., Marisa, L., de Reynies, A., Orsetti, B., Schiappa, R., Bibeau, F., MacGrogan, G., Lerebours, F., Finetti, P., Longy, M., et al. (2012). A refined molecular taxonomy of breast cancer. *Oncogene* 31, 1196–1206.
- Guo, J.Y., Chen, H.-Y., Mathew, R., Fan, J., Strohecker, A.M., Karsli-Uzunbas, G., Kamphorst, J.J., Chen, G., Lemons, J.M.S., Karantza, V., et al. (2011a). Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. *Genes Dev.* 25, 460–470.
- Guo, Z., Zheng, L., Liao, X., and Geller, D. (2016). Up-Regulation of Human Inducible Nitric Oxide Synthase by p300 Transcriptional Complex. *PloS One* 11, e0146640.
- Gurtan, A.M., and Sharp, P.A. (2013). The role of miRNAs in regulating gene expression networks. *J. Mol. Biol.* 425, 3582–3600.
- Ha, N.-H., Long, J., Cai, Q., Shu, X.O., and Hunter, K.W. (2016). The Circadian Rhythm Gene *Arntl2* Is a Metastasis Susceptibility Gene for Estrogen Receptor-Negative Breast Cancer. *PLoS Genet.* 12, e1006267.
- Hameed, A., Hruban, R.H., Gage, W., Pettis, G., and Fox, W.M. (1994). Immunohistochemical expression of CD68 antigen in human peripheral blood T cells. *Hum. Pathol.* 25, 872–876.
- Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. *Cell* 113, 429–432.
- Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.
- Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674.
- Hanson, R.L., Brown, R.B., Steele, M.M., Grandgenett, P.M., Grunkemeyer, J.A., and Hollingsworth, M.A. (2016). Identification of FRA-1 as a novel player in pancreatic cancer in cooperation with a MUC1: ERK signaling axis. *Oncotarget*.
- Hayashi-Nishino, M., Fujita, N., Noda, T., Yamaguchi, A., Yoshimori, T., and Yamamoto, A. (2009). A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. *Nat. Cell Biol.* 11, 1433–1437.
- He, H., and Lehming, N. (2003). Global effects of histone modifications. *Brief. Funct. Genomic. Proteomic.* 2, 234–243.
- He, X., Ohba, S., Hojo, H., and McMahon, A.P. (2016). AP-1 family members act with Sox9 to promote chondrocyte hypertrophy. *Dev. Camb. Engl.* 143, 3012–3023.
- Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39, 311–318.
- Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112.
- Hess, J., Angel, P., and Schorpp-Kistner, M. (2004a). AP-1 subunits: quarrel and harmony among siblings. *J. Cell Sci.* 117, 5965–5973.
- Hock, R., Furusawa, T., Ueda, T., and Bustin, M. (2007). HMG chromosomal proteins in development and disease. *Trends Cell Biol.* 17, 72–79.

- Hoffmann, E., Thiefes, A., Buhrow, D., Dittrich-Breiholz, O., Schneider, H., Resch, K., and Kracht, M. (2005). MEK1-dependent delayed expression of Fos-related antigen-1 counteracts c-Fos and p65 NF-kappaB-mediated interleukin-8 transcription in response to cytokines or growth factors. *J. Biol. Chem.* *280*, 9706–9718.
- Holbro, T., Civenni, G., and Hynes, N.E. (2003). The ErbB receptors and their role in cancer progression. *Exp. Cell Res.* *284*, 99–110.
- Hollestelle, A., Elstrodt, F., Nagel, J.H.A., Kallemeijn, W.W., and Schutte, M. (2007). Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines. *Mol. Cancer Res. MCR* *5*, 195–201.
- Holness, C.L., and Simmons, D.L. (1993). Molecular cloning of CD68, a human macrophage marker related to lysosomal glycoproteins. *Blood* *81*, 1607–1613.
- Holness, C.L., da Silva, R.P., Fawcett, J., Gordon, S., and Simmons, D.L. (1993). Macrosialin, a mouse macrophage-restricted glycoprotein, is a member of the lamp/lgp family. *J. Biol. Chem.* *268*, 9661–9666.
- Hommura, F., Katabami, M., Leaner, V.D., Donniger, H., Sumter, T.F., Resar, L.M.S., and Birrer, M.J. (2004). HMG-I/Y is a c-Jun/activator protein-1 target gene and is necessary for c-Jun-induced anchorage-independent growth in Rat1a cells. *Mol. Cancer Res. MCR* *2*, 305–314.
- Hopp, T.A., Weiss, H.L., Hilsenbeck, S.G., Cui, Y., Allred, D.C., Horwitz, K.B., and Fuqua, S.A.W. (2004). Breast cancer patients with progesterone receptor PR-A-rich tumors have poorer disease-free survival rates. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *10*, 2751–2760.
- Huang, R., Huang, D., Dai, W., and Yang, F. (2015). Overexpression of HMGA1 correlates with the malignant status and prognosis of breast cancer. *Mol. Cell. Biochem.* *404*, 251–257.
- Hübner, M.R., Eckersley-Maslin, M.A., and Spector, D.L. (2013). Chromatin organization and transcriptional regulation. *Curr. Opin. Genet. Dev.* *23*, 89–95.
- Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulitou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* *46*, 205–212.
- Huso, T.H., and Resar, L.M.S. (2014). The high mobility group A1 molecular switch: turning on cancer - can we turn it off? *Expert Opin. Ther. Targets* *18*, 541–553.
- Huynh, K.K., Eskelinen, E.-L., Scott, C.C., Malevanets, A., Saftig, P., and Grinstein, S. (2007). LAMP proteins are required for fusion of lysosomes with phagosomes. *EMBO J.* *26*, 313–324.
- Ishihara, S., and Schwartz, R.H. (2011). Two-step binding of transcription factors causes sequential chromatin structural changes at the activated IL-2 promoter. *J. Immunol. Baltim. Md 1950* *187*, 3292–3299.
- Ito, T., Yamauchi, M., Nishina, M., Yamamichi, N., Mizutani, T., Ui, M., Murakami, M., and Iba, H. (2001). Identification of SWI.SNF complex subunit BAF60a as a determinant of the transactivation potential of Fos/Jun dimers. *J. Biol. Chem.* *276*, 2852–2857.
- Itoh, T., Fujita, N., Kanno, E., Yamamoto, A., Yoshimori, T., and Fukuda, M. (2008). Golgi-resident small GTPase Rab33B interacts with Atg16L and modulates autophagosome formation. *Mol. Biol. Cell* *19*, 2916–2925.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* *293*, 1074–1080.

- Jiang, X., Overholtzer, M., and Thompson, C.B. (2015). Autophagy in cellular metabolism and cancer. *J. Clin. Invest.* *125*, 47–54.
- Jing, K., and Lim, K. (2012). Why is autophagy important in human diseases? *Exp. Mol. Med.* *44*, 69–72.
- Jochum, W., Passegue, E., and Wagner, E.F. (2001). AP-1 in mouse development and tumorigenesis. *Oncogene* *20*, 2401–2412.
- Jung, C.H., Jun, C.B., Ro, S.-H., Kim, Y.-M., Otto, N.M., Cao, J., Kundu, M., and Kim, D.-H. (2009). ULK-Atg13-FIP200 complexes mediate mTOR signaling to the autophagy machinery. *Mol. Biol. Cell* *20*, 1992–2003.
- Juven-Gershon, T., and Kadonaga, J.T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* *339*, 225–229.
- Juven-Gershon, T., Hsu, J.-Y., Theisen, J.W., and Kadonaga, J.T. (2008). The RNA polymerase II core promoter - the gateway to transcription. *Curr. Opin. Cell Biol.* *20*, 253–259.
- Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* *467*, 430–435.
- Kalogeropoulou, M., Voulgari, A., Kostourou, V., Sandaltzopoulos, R., Dikstein, R., Davidson, I., Tora, L., and Pintzas, A. (2010). TAF4b and Jun/activating protein-1 collaborate to regulate the expression of integrin $\alpha 6$ and cancer cell migration properties. *Mol. Cancer Res. MCR* *8*, 554–568.
- Kaplan, C.D., Laprade, L., and Winston, F. (2003). Transcription elongation factors repress transcription initiation from cryptic sites. *Science* *301*, 1096–1099.
- Karin, M., Liu, Z. g, and Zandi, E. (1997). AP-1 function and regulation. *Curr. Opin. Cell Biol.* *9*, 240–246.
- Kelly, P.M., McGovern, M., Gatter, K.C., Theaker, J.M., and McGee, J.O. (1988). EBM/11 reactivity in malignant histiocytosis. *J. Clin. Pathol.* *41*, 1305–1309.
- Keogh, M.-C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* *123*, 593–605.
- Kerppola, T.K., and Curran, T. (1997). The transcription activation domains of Fos and Jun induce DNA bending through electrostatic interactions. *EMBO J.* *16*, 2907–2916.
- Kharman-Biz, A., Gao, H., Ghiasvand, R., Zhao, C., Zendejdel, K., and Dahlman-Wright, K. (2013). Expression of activator protein-1 (AP-1) family members in breast cancer. *BMC Cancer* *13*, 441.
- Kiernan, R.E., Emiliani, S., Nakayama, K., Castro, A., Labbé, J.C., Lorca, T., Nakayama Ki, K., and Benkirane, M. (2001). Interaction between cyclin T1 and SCF(SKP2) targets CDK9 for ubiquitination and degradation by the proteasome. *Mol. Cell. Biol.* *21*, 7956–7970.
- Kim, T.-K., and Shiekhatar, R. (2015). Architectural and Functional Commonalities between Enhancers and Promoters. *Cell* *162*, 948–959.

- Kim, T.-K., Hemberg, M., and Gray, J.M. (2015). Enhancer RNAs: a class of long noncoding RNAs synthesized at enhancers. *Cold Spring Harb. Perspect. Biol.* 7, a018622.
- Kishi, Y., Fujii, Y., Hirabayashi, Y., and Gotoh, Y. (2012). HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. *Nat. Neurosci.* 15, 1127–1133.
- Klionsky, D.J. (2007). Autophagy: from phenomenology to molecular understanding in less than a decade. *Nat. Rev. Mol. Cell Biol.* 8, 931–937.
- Klionsky, D.J., and Codogno, P. (2013). The mechanism and physiological function of macroautophagy. *J. Innate Immun.* 5, 427–433.
- Klionsky, D.J., Abdelmohsen, K., Abe, A., Abedin, M.J., Abeliovich, H., Acevedo Arozena, A., Adachi, H., Adams, C.M., Adams, P.D., Adeli, K., et al. (2016). Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). *Autophagy* 12, 1–222.
- Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., Sasaki, D., Imamura, K., Kai, C., Harbers, M., et al. (2006). CAGE: cap analysis of gene expression. *Nat. Methods* 3, 211–222.
- Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. *Nature* 368, 466–469.
- Komarnitsky, P., Cho, E.J., and Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* 14, 2452–2460.
- Kostich, M., Fire, A., and Fambrough, D.M. (2000). Identification and molecular-genetic characterization of a LAMP/CD68-like protein from *Caenorhabditis elegans*. *J. Cell Sci.* 113 (Pt 14), 2595–2606.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.
- Kovary, K., and Bravo, R. (1991). The jun and fos protein families are both required for cell cycle progression in fibroblasts. *Mol. Cell. Biol.* 11, 4466–4472.
- Kreike, B., van Kouwenhove, M., Horlings, H., Weigelt, B., Peterse, H., Bartelink, H., and van de Vijver, M.J. (2007). Gene expression profiling and histopathological characterization of triple-negative/basal-like breast carcinomas. *Breast Cancer Res. BCR* 9, R65.
- Kreipe, H., Radzun, H.J., Parwaresch, M.R., Haislip, A., and Hansmann, M.L. (1987). Ki-M7 monoclonal antibody specific for myelomonocytic cell lineage and macrophages in human. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 35, 1117–1126.
- Krishnamurthy, S., He, X., Reyes-Reyes, M., Moore, C., and Hampsey, M. (2004). Ssu72 Is an RNA polymerase II CTD phosphatase. *Mol. Cell* 14, 387–394.
- Kunisch, E., Fuhrmann, R., Roth, A., Winter, R., Lungershausen, W., and Kinne, R.W. (2004). Macrophage specificity of three anti-CD68 monoclonal antibodies (KP1, EBM11, and PGM1) widely used for immunohistochemistry and flow cytometry. *Ann. Rheum. Dis.* 63, 774–784.
- Kurushima, H., Ramprasad, M., Kondratenko, N., Foster, D.M., Quehenberger, O., and Steinberg, D. (2000). Surface expression and rapid internalization of macrosialin (mouse CD68) on elicited mouse peritoneal macrophages. *J. Leukoc. Biol.* 67, 104–108.

- Lamb, C.A., Yoshimori, T., and Tooze, S.A. (2013). The autophagosome: origins unknown, biogenesis complex. *Nat. Rev. Mol. Cell Biol.* *14*, 759–774.
- Lazova, R., Camp, R.L., Klump, V., Siddiqui, S.F., Amaravadi, R.K., and Pawelek, J.M. (2012). Punctate LC3B expression is a common feature of solid tumors and associated with proliferation, metastasis, and poor outcome. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *18*, 370–379.
- Lee, T.I., and Young, R.A. (2000). Transcription of eukaryotic protein-coding genes. *Annu. Rev. Genet.* *34*, 77–137.
- Lee, J.-H., Byun, D.-S., Lee, M.-G., Ryu, B.-K., Kang, M.-J., Chae, K.-S., Lee, K.Y., Kim, H.-J., Park, H., and Chi, S.-G. (2008). Frequent epigenetic inactivation of hSRBC in gastric cancer and its implication in attenuated p53 response to stresses. *Int. J. Cancer* *122*, 1573–1584.
- Lee, J.-H., Kang, M.-J., Han, H.-Y., Lee, M.-G., Jeong, S.-I., Ryu, B.-K., Ha, T.-K., Her, N.-G., Han, J., Park, S.J., et al. (2011). Epigenetic alteration of PRKCDBP in colorectal cancers and its implication in tumor cell resistance to TNF α -induced apoptosis. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* *17*, 7551–7562.
- Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* *121*, 2750–2767.
- Levine, B., and Kroemer, G. (2008). Autophagy in the pathogenesis of disease. *Cell* *132*, 27–42.
- Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* *148*, 84–98.
- Li, Q., Price, J.P., Byers, S.A., Cheng, D., Peng, J., and Price, D.H. (2005). Analysis of the large inactive P-TEFb complex indicates that it contains one 7SK molecule, a dimer of HEXIM1 or HEXIM2, and two P-TEFb molecules containing Cdk9 phosphorylated at threonine 186. *J. Biol. Chem.* *280*, 28819–28826.
- Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* *17*, 207–223.
- Liang, C., and Jung, J.U. (2010). Autophagy genes as tumor suppressors. *Curr. Opin. Cell Biol.* *22*, 226–233.
- Liao, X., Chen, Y., Liu, D., Li, F., Li, X., and Jia, W. (2015). High Expression of LAMP3 Is a Novel Biomarker of Poor Prognosis in Patients with Esophageal Squamous Cell Carcinoma. *Int. J. Mol. Sci.* *16*, 17655–17667.
- Licatalosi, D.D., Geiger, G., Minet, M., Schroeder, S., Cilli, K., McNeil, J.B., and Bentley, D.L. (2002). Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. *Mol. Cell* *9*, 1101–1111.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* *326*, 289–293.
- Liu, K., Zhang, C., Li, T., Ding, Y., Tu, T., Zhou, F., Qi, W., Chen, H., and Sun, X. (2015). Let-7a inhibits growth and migration of breast cancer cells by targeting HMGA1. *Int. J. Oncol.* *46*, 2526–2534.

- Liu, X., Li, H., Rajurkar, M., Li, Q., Cotton, J.L., Ou, J., Zhu, L.J., Goel, H.L., Mercurio, A.M., Park, J.-S., et al. (2016). Tead and AP1 Coordinate Transcription and Motility. *Cell Rep.* *14*, 1169–1180.
- Lively, T.N., Nguyen, T.N., Galasinski, S.K., and Goodrich, J.A. (2004). The basic leucine zipper domain of c-Jun functions in transcriptional activation through interaction with the N terminus of human TATA-binding protein-associated factor-1 (human TAF(II)250). *J. Biol. Chem.* *279*, 26257–26265.
- Lock, R., Roy, S., Kenific, C.M., Su, J.S., Salas, E., Ronen, S.M., and Debnath, J. (2011). Autophagy facilitates glycolysis during Ras-mediated oncogenic transformation. *Mol. Biol. Cell* *22*, 165–178.
- Lock, R., Kenific, C.M., Leidal, A.M., Salas, E., and Debnath, J. (2014). Autophagy-dependent production of secreted factors facilitates oncogenic RAS-driven invasion. *Cancer Discov.* *4*, 466–479.
- Logullo, A.F., Stiepcich, M.M.Á., de Toledo Osório, C.A.B., Nonogaki, S., Pasini, F.S., Rocha, R.M., Soares, F.A., and Brentani, M.M. (2011). Role of Fos-related antigen 1 in the progression and prognosis of ductal breast carcinoma. *Histopathology* *58*, 617–625.
- Louie, M.C., Yang, H.Q., Ma, A.-H., Xu, W., Zou, J.X., Kung, H.-J., and Chen, H.-W. (2003). Androgen-induced recruitment of RNA polymerase II to a nuclear receptor-p160 coactivator complex. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 2226–2230.
- Luo, Y., Zhou, H., Mizutani, M., Mizutani, N., Reisfeld, R.A., and Xiang, R. (2003). Transcription factor Fos-related antigen 1 is an effective target for a breast cancer vaccine. *Proc. Natl. Acad. Sci.* *100*, 8850–8855.
- Malik, A.N., Vierbuchen, T., Hemberg, M., Rubin, A.A., Ling, E., Couch, C.H., Stroud, H., Spiegel, I., Farh, K.K.-H., Harmin, D.A., et al. (2014). Genome-wide identification and characterization of functional neuronal activity-dependent enhancers. *Nat. Neurosci.* *17*, 1330–1339.
- Mansueto, G., Forzati, F., Ferraro, A., Pallante, P., Bianco, M., Esposito, F., Iaccarino, A., Troncone, G., and Fusco, A. (2010). Identification of a New Pathway for Tumor Progression: MicroRNA-181b Up-Regulation and CBX7 Down-Regulation by HMGA1 Protein. *Genes Cancer* *1*, 210–224.
- Mariani, O., Brennetot, C., Coindre, J.-M., Gruel, N., Ganem, C., Delattre, O., Stern, M.-H., and Aurias, A. (2007). JUN oncogene amplification and overexpression block adipocytic differentiation in highly aggressive sarcomas. *Cancer Cell* *11*, 361–374.
- Marshall, N.F., and Price, D.H. (1992). Control of formation of two distinct classes of RNA polymerase II elongation complexes. *Mol. Cell. Biol.* *12*, 2078–2090.
- Martin, M.L., Lieberman, P.M., and Curran, T. (1996). Fos-Jun dimerization promotes interaction of the basic region with TFIIE-34 and TFIIF. *Mol. Cell. Biol.* *16*, 2110–2118.
- Marti-Renom, M.A., and Mirny, L.A. (2011). Bridging the resolution gap in structural modeling of 3D genome organization. *PLoS Comput. Biol.* *7*, e1002125.
- Mason, J.M., Schmitz, M.A., Müller, K.M., and Arndt, K.M. (2006). Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 8989–8994.

- Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29–59.
- Masuda, H., Baggerly, K.A., Wang, Y., Zhang, Y., Gonzalez-Angulo, A.M., Meric-Bernstam, F., Valero, V., Lehmann, B.D., Pietenpol, J.A., Hortobagyi, G.N., et al. (2013). Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 19, 5533–5540.
- Matsuo, K., Owens, J.M., Tonko, M., Elliott, C., Chambers, T.J., and Wagner, E.F. (2000). Fos11 is a transcriptional target of c-Fos during osteoclast differentiation. *Nat. Genet.* 24, 184–187.
- Mayer, I.A., Abramson, V.G., Lehmann, B.D., and Pietenpol, J.A. (2014). New strategies for triple-negative breast cancer--deciphering the heterogeneity. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 20, 782–790.
- Meinhart, A., and Cramer, P. (2004). Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors. *Nature* 430, 223–226.
- Mialon, A., Sankinen, M., Söderström, H., Junttila, T.T., Holmström, T., Koivusalo, R., Papageorgiou, A.C., Johnson, R.S., Hietanen, S., Elenius, K., et al. (2005). DNA topoisomerase I is a cofactor for c-Jun in the regulation of epidermal growth factor receptor expression and cancer cell proliferation. *Mol. Cell. Biol.* 25, 5040–5051.
- Michels, A.A., Fraldi, A., Li, Q., Adamson, T.E., Bonnet, F., Nguyen, V.T., Sedore, S.C., Price, J.P., Price, D.H., Lania, L., et al. (2004). Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.* 23, 2608–2619.
- Micklem, K., Rigney, E., Cordell, J., Simmons, D., Stross, P., Turley, H., Seed, B., and Mason, D. (1989). A human macrophage-associated antigen (CD68) detected by six different monoclonal antibodies. *Br. J. Haematol.* 73, 6–11.
- Milde-Langosch, K., Röder, H., Andritzky, B., Aslan, B., Hemminger, G., Brinkmann, A., Bamberger, C.M., Löning, T., and Bamberger, A.-M. (2004). The role of the AP-1 transcription factors c-Fos, FosB, Fra-1 and Fra-2 in the invasion process of mammary carcinomas. *Breast Cancer Res. Treat.* 86, 139–152.
- Milde-Langosch, K., Janke, S., Wagner, I., Schröder, C., Streichert, T., Bamberger, A.-M., Jänicke, F., and Löning, T. (2008). Role of Fra-2 in breast cancer: influence on tumor cell invasion and motility. *Breast Cancer Res. Treat.* 107, 337–347.
- Mishra, M., Flaga, J., and Kowluru, R.A. (2016). Molecular Mechanism of Transcriptional Regulation of Matrix Metalloproteinase-9 in Diabetic Retinopathy: MMP-9 AND DIABETIC RETINOPATHY. *J. Cell. Physiol.* 231, 1709–1718.
- Mittler, G., Stühler, T., Santolin, L., Uhlmann, T., Kremmer, E., Lottspeich, F., Berti, L., and Meisterernst, M. (2003). A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* 22, 6494–6504.
- Moquet-Torcy, G., Tolza, C., Piechaczyk, M., and Jariel-Encontre, I. (2014). Transcriptional complexity and roles of Fra-1/AP-1 at the uPA/Plau locus in aggressive breast cancer. *Nucleic Acids Res.* 42, 11011–11024.
- Morton, S., Davis, R.J., McLaren, A., and Cohen, P. (2003). A reinvestigation of the multisite phosphorylation of the transcription factor c-Jun. *EMBO J.* 22, 3876–3886.

- Mosley, A.L., Pattenden, S.G., Carey, M., Venkatesh, S., Gilmore, J.M., Florens, L., Workman, J.L., and Washburn, M.P. (2009). Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation. *Mol. Cell* 34, 168–178.
- Müller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M.E., McClanahan, T., Murphy, E., Yuan, W., Wagner, S.N., et al. (2001). Involvement of chemokine receptors in breast cancer metastasis. *Nature* 410, 50–56.
- Muller, S., Berger, M., Lehenbre, F., Seeler, J.S., Haupt, Y., and Dejean, A. (2000). c-Jun and p53 activity is modulated by SUMO-1 modification. *J. Biol. Chem.* 275, 13321–13329.
- Munz, C., Psichari, E., Mandilis, D., Lavigne, A.-C., Spiliotaki, M., Oehler, T., Davidson, I., Tora, L., Angel, P., and Pintzas, A. (2003). TAF7 (TAFII55) plays a role in the transcription activation by c-Jun. *J. Biol. Chem.* 278, 21510–21516.
- Murakami, M., Ui, M., and Iba, H. (1999). Fra-2-positive autoregulatory loop triggered by mitogen-activated protein kinase (MAPK) and Fra-2 phosphorylation sites by MAPK. *Cell Growth Differ. Mol. Biol. J. Am. Assoc. Cancer Res.* 10, 333–342.
- Murphy, L.O., and Blenis, J. (2006). MAPK signal specificity: the right place at the right time. *Trends Biochem. Sci.* 31, 268–275.
- Murphy, L.O., Smith, S., Chen, R.-H., Fingar, D.C., and Blenis, J. (2002). Molecular interpretation of ERK signal duration by immediate early gene products. *Nat. Cell Biol.* 4, 556–564.
- Murphy, L.O., MacKeigan, J.P., and Blenis, J. (2004). A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. *Mol. Cell. Biol.* 24, 144–153.
- Musti, A.M., Treier, M., Peverali, F.A., and Bohmann, D. (1996). Differential regulation of c-Jun and JunD by ubiquitin-dependent protein degradation. *Biol. Chem.* 377, 619–624.
- Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502, 59–64.
- Nagelkerke, A., Sieuwerts, A.M., Bussink, J., Sweep, F.C.G.J., Look, M.P., Foekens, J.A., Martens, J.W.M., and Span, P.N. (2014). LAMP3 is involved in tamoxifen resistance in breast cancer cells through the modulation of autophagy. *Endocr. Relat. Cancer* 21, 101–112.
- Narayanan, K., Srinivas, R., Peterson, M.C., Ramachandran, A., Hao, J., Thimmapaya, B., Scherer, P.E., and George, A. (2004). Transcriptional regulation of dentin matrix protein 1 by JunB and p300 during osteoblast differentiation. *J. Biol. Chem.* 279, 44294–44302.
- Nechaev, S., and Adelman, K. (2011). Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation. *Biochim. Biophys. Acta* 1809, 34–45.
- Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.-P., Tong, F., et al. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* 10, 515–527.
- Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719.

- Ng, K.W., Ridgway, P., Cohen, D.R., and Tremethick, D.J. (1997). The binding of a Fos/Jun heterodimer can completely disrupt the structure of a nucleosome. *EMBO J.* *16*, 2072–2085.
- Nguyen, T.A., Jones, R.D., Snaveley, A.R., Pfenning, A.R., Kirchner, R., Hemberg, M., and Gray, J.M. (2016). High-throughput functional comparison of promoter and enhancer activities. *Genome Res.* *26*, 1023–1033.
- Nguyen, V.T., Kiss, T., Michels, A.A., and Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* *414*, 322–325.
- Nishida, H., Suzuki, T., Kondo, S., Miura, H., Fujimura, Y., and Hayashizaki, Y. (2006). Histone H3 acetylated at lysine 9 in promoter is associated with low nucleosome density in the vicinity of transcription start site in human cell. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* *14*, 203–211.
- Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* *485*, 381–385.
- Noro, B., Licheri, B., Sgarra, R., Rustighi, A., Tessari, M.A., Chau, K.-Y., Ono, S.J., Giancotti, V., and Manfioletti, G. (2003). Molecular dissection of the architectural transcription factor HMGA2. *Biochemistry (Mosc.)* *42*, 4569–4577.
- Oakman, C., Viale, G., and Di Leo, A. (2010). Management of triple negative breast cancer. *Breast Edinb. Scotl.* *19*, 312–321.
- Obiorah, I., and Jordan, V.C. (2011). Progress in endocrine approaches to the treatment and prevention of breast cancer. *Maturitas* *70*, 315–321.
- Ogram, S.A., and Reeves, R. (1995). Differential regulation of a multipromoter gene. Selective 12-O-tetradecanoylphorbol-13-acetate induction of a single transcription start site in the HMG-I/Y gene. *J. Biol. Chem.* *270*, 14235–14242.
- Okazaki, K., and Sagata, N. (1995). The Mos/MAP kinase pathway stabilizes c-Fos by phosphorylation and augments its transforming activity in NIH 3T3 cells. *EMBO J.* *14*, 5048–5059.
- Olive, M., Krylov, D., Echlin, D.R., Gardner, K., Taparowsky, E., and Vinson, C. (1997). A dominant negative to activation protein-1 (AP1) that abolishes DNA binding and inhibits oncogenesis. *J. Biol. Chem.* *272*, 18586–18594.
- Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* *12*, 283–293.
- Ozanne, B.W., Spence, H.J., McGarry, L.C., and Hennigan, R.F. (2007). Transcription factors control invasion: AP-1 the first among equals. *Oncogene* *26*, 1–10.
- Ozsolak, F., Song, J.S., Liu, X.S., and Fisher, D.E. (2007). High-throughput mapping of the chromatin structure of human promoters. *Nat. Biotechnol.* *25*, 244–248.
- Pallante, P., Sepe, R., Puca, F., and Fusco, A. (2015). High mobility group a proteins as tumor markers. *Front. Med.* *2*, 15.
- Park, J.S., Park, J.H., Khoi, P.N., Joo, Y.E., and Jung, Y.D. (2011). MSP-induced RON activation upregulates uPAR expression and cell invasiveness via MAPK, AP-1 and NF- κ B signals in gastric cancer cells. *Carcinogenesis* *32*, 175–181.

- Parwaresch, M.R., Radzun, H.J., Kreipe, H., Hansmann, M.L., and Barth, J. (1986). Monocyte/macrophage-reactive monoclonal antibody Ki-M6 recognizes an intracytoplasmic antigen. *Am. J. Pathol.* *125*, 141–151.
- Passegué, E., Wagner, E.F., and Weissman, I.L. (2004). JunB deficiency leads to a myeloproliferative disorder arising from hematopoietic stem cells. *Cell* *119*, 431–443.
- Patel, L.R., Curran, T., and Kerppola, T.K. (1994). Energy transfer analysis of Fos-Jun dimerization and DNA binding. *Proc. Natl. Acad. Sci. U. S. A.* *91*, 7360–7364.
- Pedraza-Alva, G., Zingg, J.M., Donda, A., and Pérez-Martínez, L. (2009). Estrogen receptor regulates MyoD gene expression by preventing AP-1-mediated repression. *Biochem. Biophys. Res. Commun.* *389*, 360–365.
- Pegoraro, S., Ros, G., Piazza, S., Sommaggio, R., Ciani, Y., Rosato, A., Sgarra, R., Del Sal, G., and Manfioletti, G. (2013). HMGA1 promotes metastatic processes in basal-like breast cancer regulating EMT and stemness. *Oncotarget* *4*, 1293–1308.
- Pei, Y., and Tuschl, T. (2006). On the art of identifying effective and specific siRNAs. *Nat. Methods* *3*, 670–676.
- Perou, C.M. (2011). Molecular stratification of triple-negative breast cancers. *The Oncologist* *16 Suppl 1*, 61–70.
- Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. *Nature* *406*, 747–752.
- Phatnani, H.P., and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* *20*, 2922–2936.
- Philips, A., Teyssier, C., Galtier, F., Rivier-Covas, C., Rey, J.M., Rochefort, H., and Chalbos, D. (1998). FRA-1 expression level modulates regulation of activator protein-1 activity by estradiol in breast cancer cells. *Mol. Endocrinol. Baltim. Md* *12*, 973–985.
- Phillips-Cremins, J.E. (2014). Unraveling architecture of the pluripotent genome. *Curr. Opin. Cell Biol.* *28*, 96–104.
- Phillips-Cremins, J.E., Sauria, M.E.G., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S.K., Ong, C.-T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* *153*, 1281–1295.
- Piechaczyk, M., and Blanchard, J.M. (1994). c-fos proto-oncogene regulation and function. *Crit. Rev. Oncol. Hematol.* *17*, 93–131.
- Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X., and Perou, C.M. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res. BCR* *12*, R68.
- Preston, G.A., Srinivasan, D., and Barrett, J.C. (2000). Apoptotic response to growth factor deprivation involves cooperative interactions between c-Fos and p300. *Cell Death Differ.* *7*, 215–226.
- Pulford, K.A., Sipos, A., Cordell, J.L., Stross, W.P., and Mason, D.Y. (1990). Distribution of the CD68 macrophage/myeloid associated antigen. *Int. Immunol.* *2*, 973–980.

- Pundhir, S., Poirazi, P., and Gorodkin, J. (2015). Emerging applications of read profiles towards the functional annotation of the genome. *Front. Genet.* 6, 188.
- Qin, L., Xu, Y., Xu, Y., Ma, G., Liao, L., Wu, Y., Li, Y., Wang, X., Wang, X., Jiang, J., et al. (2015). NCOA1 promotes angiogenesis in breast tumors by simultaneously enhancing both HIF1 α - and AP-1-mediated VEGFa transcription. *Oncotarget* 6, 23890–23904.
- Radwan, A., Younis, A., Luykx, P., and Khuri, S. (2008). Prediction and analysis of nucleosome exclusion regions in the human genome. *BMC Genomics* 9, 186.
- Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. *Cell* 141, 432–445.
- Rahmani, M., Péron, P., Weitzman, J., Bakiri, L., Lardeux, B., and Bernuau, D. (2001). Functional cooperation between JunD and NF-kappaB in rat hepatocytes. *Oncogene* 20, 5132–5142.
- Ramos-Nino, M.E., Blumen, S.R., Pass, H., and Mossman, B.T. (2007). Fra-1 governs cell migration via modulation of CD44 expression in human mesotheliomas. *Mol. Cancer* 6, 81.
- Ramprasad, M.P., Terpstra, V., Kondratenko, N., Quehenberger, O., and Steinberg, D. (1996). Cell surface expression of mouse macrosialin and human CD68 and their role as macrophage receptors for oxidized low density lipoprotein. *Proc. Natl. Acad. Sci. U. S. A.* 93, 14833–14838.
- Randall-Demllo, S., Chieppa, M., and Eri, R. (2013). Intestinal epithelium and autophagy: partners in gut homeostasis. *Front. Immunol.* 4, 301.
- Rando, O.J., and Chang, H.Y. (2009). Genome-wide views of chromatin structure. *Annu. Rev. Biochem.* 78, 245–271.
- Ransone, L.J., Kerr, L.D., Schmitt, M.J., Wamsley, P., and Verma, I.M. (1993). The bZIP domains of Fos and Jun mediate a physical association with the TATA box-binding protein. *Gene Expr.* 3, 37–48.
- Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.
- Rapoport, B.L., Specialist Physician and Medical Oncologist, The Medical Oncology Center of Rosebank, Johannesburg, South Africa, Nayler, S., Specialist Histopathologist, Gritzman and Thatcher Inc. and Wits Donald Gordon Medical Center, Johannesburg, South Africa, Demetriou, G.S., Specialist Physician and Medical Oncologist, Department of Internal Medicine Division of Medical Oncology, University of the Witwatersrand and Wits Donald Gordon Medical Center, Johannesburg, South Africa, Moodley, S.D., Specialist Physician and Medical Oncologist, University of the Witwatersrand and Wits Donald Gordon Medical Center, Johannesburg, South Africa, Benn, C.A., and Specialist Surgeon, Netcare Milpark Hospital Breast Centre and Helen Joseph Hospital, University of the Witwatersrand, Johannesburg, South Africa (2014). Triple Negative Breast Cancer Pathologic Diagnosis and Current Chemotherapy Treatment Options. *Oncol. Hematol. Rev. US* 10, 25.
- Ravikumar, B., Moreau, K., Jahreiss, L., Puri, C., and Rubinsztein, D.C. (2010). Plasma membrane contributes to the formation of pre-autophagosomal structures. *Nat. Cell Biol.* 12, 747–757.

- Reeves, R., and Beckerbauer, L. (2001). HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys. Acta* 1519, 13–29.
- Reeves, R., Edberg, D.D., and Li, Y. (2001). Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol. Cell. Biol.* 21, 575–594.
- Reynolds, A., Anderson, E.M., Vermeulen, A., Fedorov, Y., Robinson, K., Leake, D., Karpilow, J., Marshall, W.S., and Khvorova, A. (2006). Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA N. Y. N* 12, 988–993.
- Roberts, T.C., Hart, J.R., Kaikkonen, M.U., Weinberg, M.S., Vogt, P.K., and Morris, K.V. (2015). Quantification of nascent transcription by bromouridine immunocapture nuclear run-on RT-qPCR. *Nat. Protoc.* 10, 1198–1211.
- Robinson, K.M., and Kadonaga, J.T. (1998). The use of chromatin templates to recreate transcriptional regulatory phenomena in vitro. *Biochim. Biophys. Acta* 1378, M1-6.
- Ross, H.M., Daniel, H.D.J., Vivekanandan, P., Kannangai, R., Yeh, M.M., Wu, T.-T., Makhoul, H.R., and Torbenson, M. (2011). Fibrolamellar carcinomas are positive for CD68. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* 24, 390–395.
- Ruiz-Orera, J., Hernandez-Rodriguez, J., Chiva, C., Sabidó, E., Kondova, I., Bontrop, R., Marqués-Bonet, T., and Albà, M.M. (2015). Origins of De Novo Genes in Human and Chimpanzee. *PLOS Genet* 11, e1005721.
- Saha, P., and Nanda, R. (2016). Concepts and targets in triple-negative breast cancer: recent results and clinical implications. *Ther. Adv. Med. Oncol.* 8, 351–359.
- Saito, N., Pulford, K.A., Breton-Gorius, J., Massé, J.M., Mason, D.Y., and Cramer, E.M. (1991). Ultrastructural localization of the CD68 macrophage-associated antigen in human blood neutrophils and monocytes. *Am. J. Pathol.* 139, 1053–1059.
- Sayan, A.E., Stanford, R., Vickery, R., Grigorenko, E., Diesch, J., Kulbicki, K., Edwards, R., Pal, R., Greaves, P., Jariel-Encontre, I., et al. (2012). Fra-1 controls motility of bladder cancer cells via transcriptional upregulation of the receptor tyrosine kinase AXL. *Oncogene* 31, 1493–1503.
- Schröder, C., Schumacher, U., Müller, V., Wirtz, R.M., Streichert, T., Richter, U., Wicklein, D., and Milde-Langosch, K. (2010). The transcription factor Fra-2 promotes mammary tumour progression by changing the adhesive properties of breast cancer cells. *Eur. J. Cancer Oxf. Engl.* 1990 46, 1650–1660.
- Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. *Science* 322, 1849–1851.
- Sexton, T., Bantignies, F., and Cavalli, G. (2009). Genomic interactions: chromatin loops and gene meeting points in transcriptional regulation. *Semin. Cell Dev. Biol.* 20, 849–855.
- Sgarra, R., Lee, J., Tessari, M.A., Altamura, S., Spolaore, B., Giancotti, V., Bedford, M.T., and Manfioletti, G. (2006). The AT-hook of the chromatin architectural transcription factor high mobility group A1a is arginine-methylated by protein arginine methyltransferase 6. *J. Biol. Chem.* 281, 3764–3772.

- Sgarra, R., Zammiti, S., Lo Sardo, A., Maurizio, E., Arnoldo, L., Pegoraro, S., Giancotti, V., and Manfioletti, G. (2010). HMGA molecular network: From transcriptional regulation to chromatin remodeling. *Biochim. Biophys. Acta* 1799, 37–47.
- Shah, S.N., Cope, L., Poh, W., Belton, A., Roy, S., Talbot, C.C., Sukumar, S., Huso, D.L., and Resar, L.M.S. (2013). HMGA1: a master regulator of tumor progression in triple-negative breast cancer cells. *PloS One* 8, e63419.
- Shaulian, E. (2010). AP-1--The Jun proteins: Oncogenes or tumor suppressors in disguise? *Cell. Signal.* 22, 894–899.
- Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. *Nat. Cell Biol* 4, E131–E136.
- Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenko, V.V., et al. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature* 488, 116–120.
- Shiraki, T., Kondo, S., Katayama, S., Waki, K., Kasukawa, T., Kawaji, H., Kodzius, R., Watahiki, A., Nakamura, M., Arakawa, T., et al. (2003). Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15776–15781.
- Sikorski, T.W., and Buratowski, S. (2009). The basal initiation machinery: beyond the general transcription factors. *Curr. Opin. Cell Biol.* 21, 344–351.
- Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38, 1348–1354.
- Skinner, M., Qu, S., Moore, C., and Wisdom, R. (1997). Transcriptional activation and transformation by FosB protein require phosphorylation of the carboxyl-terminal activation domain. *Mol. Cell. Biol.* 17, 2372–2380.
- Skvortsova, Y.V., Kondratieva, S.A., Zinovyeva, M.V., Nikolaev, L.G., Azhikina, T.L., and Gainetdinov, I.V. (2016). Intragenic Locus in Human PIWIL2 Gene Shares Promoter and Enhancer Functions. *PloS One* 11, e0156454.
- Sledz, C.A., Holko, M., de Veer, M.J., Silverman, R.H., and Williams, B.R.G. (2003). Activation of the interferon system by short-interfering RNAs. *Nat. Cell Biol.* 5, 834–839.
- Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochim. Biophys. Acta* 1351, 73–88.
- Smart, D.E., Green, K., Oakley, F., Weitzman, J.B., Yaniv, M., Reynolds, G., Mann, J., Millward-Sadler, H., and Mann, D.A. (2006). JunD is a profibrogenic transcription factor regulated by Jun N-terminal kinase-independent phosphorylation. *Hepatol. Baltim. Md* 44, 1432–1440.
- Smith, M.J., and Koch, G.L. (1987). Differential expression of murine macrophage surface glycoprotein antigens in intracellular membranes. *J. Cell Sci.* 87 (Pt 1), 113–119.
- Song, L., Lee, C., and Schindler, C. (2011). Deletion of the murine scavenger receptor CD68. *J. Lipid Res.* 52, 1542–1550.

- Sonobe, M.H., Yoshida, T., Murakami, M., Kameda, T., and Iba, H. (1995). fra-2 promoter can respond to serum-stimulation through AP-1 complexes. *Oncogene* 10, 689–696.
- Sorlie, T., Perou, C.M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* 98, 10869–10874.
- Sousa, S., Brion, R., Lintunen, M., Kronqvist, P., Sandholm, J., Mönkkönen, J., Kellokumpu-Lehtinen, P.-L., Lanttia, S., Tynnenen, O., Joensuu, H., et al. (2015). Human breast cancer cells educate macrophages toward the M2 activation status. *Breast Cancer Res. BCR* 17.
- Stover, D.G., Bell, C.F., and Tolaney, S.M. (2016). Neoadjuvant and Adjuvant Chemotherapy Considerations for Triple-Negative Breast Cancer. *Am. J. Hematol. Oncol.* 12.
- Strojanik, T., Kavalari, R., Zajc, I., Diamandis, E.P., Oikonomopoulou, K., and Lah, T.T. (2009). Prognostic impact of CD68 and kallikrein 6 in human glioma. *Anticancer Res.* 29, 3269–3279.
- Subramanian, D., Bunjobpol, W., and Sabapathy, K. (2015). Interplay between TAp73 Protein and Selected Activator Protein-1 (AP-1) Family Members Promotes AP-1 Target Gene Activation and Cellular Growth. *J. Biol. Chem.* 290, 18636–18649.
- Sun, H., Lu, J., Xu, X., Jin, S., Wang, X., Wei, L., Dong, M., and Huang, B. (2005). Histone acetyltransferase activity of p300 enhances the activation of IL-18 promoter. *J. Cell. Biochem.* 94, 566–572.
- Suzuki, T., Okuno, H., Yoshida, T., Endo, T., Nishina, H., and Iba, H. (1991). Difference in transcriptional regulatory function between c-Fos and Fra-2. *Nucleic Acids Res.* 19, 5537–5542.
- Szremska, A.P., Kenner, L., Weisz, E., Ott, R.G., Passegué, E., Artwohl, M., Freissmuth, M., Stoxreiter, R., Theussl, H.-C., Parzer, S.B., et al. (2003). JunB inhibits proliferation and transformation in B-lymphoid cells. *Blood* 102, 4159–4165.
- Talotta, F., Mega, T., Bossis, G., Casalino, L., Basbous, J., Jariel-Encontre, I., Piechaczyk, M., and Verde, P. (2010). Heterodimerization with Fra-1 cooperates with the ERK pathway to stabilize c-Jun in response to the RAS oncoprotein. *Oncogene* 29, 4732–4740.
- Tam, W.L., Lu, H., Buikhuisen, J., Soh, B.S., Lim, E., Reinhardt, F., Wu, Z.J., Krall, J.A., Bieri, B., Guo, W., et al. (2013). Protein kinase C α is a central signaling node and therapeutic target for breast cancer stem cells. *Cancer Cell* 24, 347–364.
- Tanos, T., Marinissen, M.J., Leskow, F.C., Hochbaum, D., Martinetto, H., Gutkind, J.S., and Coso, O.A. (2005). Phosphorylation of c-Fos by members of the p38 MAPK family. Role in the AP-1 response to UV light. *J. Biol. Chem.* 280, 18842–18852.
- Terasawa, K., Okazaki, K., and Nishida, E. (2003). Regulation of c-Fos and Fra-1 by the MEK5-ERK5 pathway. *Genes Cells Devoted Mol. Cell. Mech.* 8, 263–273.
- Thomas, M.C., and Chiang, C.-M. (2006). The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* 41, 105–178.
- Thomsen, M.K., Bakiri, L., Hasenfuss, S.C., Wu, H., Morente, M., and Wagner, E.F. (2015). Loss of JUNB/AP-1 promotes invasive prostate cancer. *Cell Death Differ.* 22, 574–582.

- Toft, D.J., Rosenberg, S.B., Bergers, G., Volpert, O., and Linzer, D.I. (2001). Reactivation of proliferin gene expression is associated with increased angiogenesis in a cell culture model of fibrosarcoma tumor progression. *Proc Natl Acad Sci USA* 98, 13055–13059.
- Travaglione, S., Falzano, L., Fabbri, A., Stringaro, A., Fais, S., and Fiorentini, C. (2002). Epithelial cells and expression of the phagocytic marker CD68: scavenging of apoptotic bodies following Rho activation. *Toxicol. Vitro Int. J. Publ. Assoc. BIBRA* 16, 405–411.
- Trierweiler, C., Hockenjos, B., Zatloukal, K., Thimme, R., Blum, H.E., Wagner, E.F., and Hasselblatt, P. (2016). The transcription factor c-JUN/AP-1 promotes HBV-related liver tumorigenesis in mice. *Cell Death Differ.* 23, 576–582.
- Tsai, L.N., Ku, T.K.S., Salib, N.K., and Crowe, D.L. (2008). Extracellular signals regulate rapid coactivator recruitment at AP-1 sites by altered phosphorylation of both CREB binding protein and c-jun. *Mol. Cell. Biol.* 28, 4240–4250.
- Tsang, W.Y., and Chan, J.K. (1992). KP1 (CD68) staining of granular cell neoplasms: is KP1 a marker for lysosomes rather than the histiocytic lineage? *Histopathology* 21, 84–86.
- Vallone, D., Battista, S., Pierantoni, G.M., Fedele, M., Casalino, L., Santoro, M., Viglietto, G., Fusco, A., and Verde, P. (1997). Neoplastic transformation of rat thyroid cells requires the junB and fra-1 gene induction which is dependent on the HMGI-C gene product. *EMBO J.* 16, 5310–5321.
- Verfaillie, A., Imrichova, H., Atak, Z.K., Dewaele, M., Rambow, F., Hulselmans, G., Christiaens, V., Svetlichnyy, D., Luciani, F., Van den Mooter, L., et al. (2015). Decoding the regulatory landscape of melanoma reveals TEADS as regulators of the invasive cell state. *Nat. Commun.* 6, 6683.
- Vesely, P.W., Staber, P.B., Hoefler, G., and Kenner, L. (2009). Translational regulation mechanisms of AP-1 proteins. *Mutat. Res.* 682, 7–12.
- Vial, E., and Marshall, C.J. (2003). Elevated ERK-MAP kinase activity protects the FOS family member FRA-1 against proteasomal degradation in colon carcinoma cells. *J. Cell Sci.* 116, 4957–4963.
- Vinson, C., Myakishev, M., Acharya, A., Mir, A.A., Moll, J.R., and Bonovich, M. (2002). Classification of human B-ZIP proteins based on dimerization properties. *Mol. Cell. Biol.* 22, 6321–6335.
- Vries, R.G., Prudenziati, M., Zwartjes, C., Verlaan, M., Kalkhoven, E., and Zantema, A. (2001). A specific lysine in c-Jun is required for transcriptional repression by E1A and is acetylated by p300. *EMBO J.* 20, 6095–6103.
- Wang, Q., Carroll, J.S., and Brown, M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Mol. Cell* 19, 631–642.
- Wang, W.-M., Wu, S.-Y., Lee, A.-Y., and Chiang, C.-M. (2011). Binding site specificity and factor redundancy in activator protein-1-driven human papillomavirus chromatin-dependent transcription. *J. Biol. Chem.* 286, 40974–40986.
- Wang, Z.Q., Ovitt, C., Grigoriadis, A.E., Möhle-Steinlein, U., Rüther, U., and Wagner, E.F. (1992). Bone and haematopoietic defects in mice lacking c-fos. *Nature* 360, 741–745.

- Weingarten-Gabbay, S., and Segal, E. (2014). A shared architecture for promoters and enhancers. *Nat. Genet.* *46*, 1253–1254.
- Weiss, C., Schneider, S., Wagner, E.F., Zhang, X., Seto, E., and Bohmann, D. (2003). JNK phosphorylation relieves HDAC3-dependent suppression of the transcriptional activity of c-Jun. *EMBO J.* *22*, 3686–3695.
- Wikman, H., Sielaff-Frimpong, B., Kropidlowski, J., Witzel, I., Milde-Langosch, K., Sauter, G., Westphal, M., Lamszus, K., and Pantel, K. (2012). Clinical relevance of loss of 11p15 in primary and metastatic breast cancer: association with loss of PRKCDBP expression in brain metastases. *PLoS One* *7*, e47537.
- Wilson, R.C., and Doudna, J.A. (2013). Molecular mechanisms of RNA interference. *Annu. Rev. Biophys.* *42*, 217–239.
- Wisdon, R., and Verma, I.M. (1993). Transformation by Fos proteins requires a C-terminal transactivation domain. *Mol. Cell. Biol.* *13*, 7429–7438.
- Wong, C., Rougier-Chapman, E.M., Frederick, J.P., Datto, M.B., Liberati, N.T., Li, J.M., and Wang, X.F. (1999). Smad3-Smad4 and AP-1 complexes synergize in transcriptional activation of the c-Jun promoter by transforming growth factor beta. *Mol. Cell. Biol.* *19*, 1821–1830.
- Wood, L.J., Maher, J.F., Bunton, T.E., and Resar, L.M. (2000). The oncogenic properties of the HMG-I gene family. *Cancer Res.* *60*, 4256–4261.
- Wu, J., Ji, A., Wang, X., Zhu, Y., Yu, Y., Lin, Y., Liu, Y., Li, S., Liang, Z., Xu, X., et al. (2015). MicroRNA-195-5p, a new regulator of Fra-1, suppresses the migration and invasion of prostate cancer cells. *J. Transl. Med.* *13*, 289.
- Xiao, W., Hodge, D.R., Wang, L., Yang, X., Zhang, X., and Farrar, W.L. (2004). NF-kappaB activates IL-6 expression through cooperation with c-Jun and IL6-AP1 site, but is independent of its IL6-NFkappaB regulatory site in autocrine human multiple myeloma cells. *Cancer Biol. Ther.* *3*, 1007–1017.
- Xie, J., Zhang, X., and Zhang, L. (2013). Negative regulation of inflammation by SIRT1. *Pharmacol. Res.* *67*, 60–67.
- Yáñez-Cuna, J.O., Arnold, C.D., Stampfel, G., Boryn, L.M., Gerlach, D., Rath, M., and Stark, A. (2014). Dissection of thousands of cell type-specific enhancers identifies dinucleotide repeat motifs as general enhancer features. *Genome Res.* *24*, 1147–1156.
- Yang, F., DeBeaumont, R., Zhou, S., and Näär, A.M. (2004). The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 2339–2344.
- Yersal, O., and Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J. Clin. Oncol.* *5*, 412–424.
- Yie, J., Liang, S., Merika, M., and Thanos, D. (1997). Intra- and intermolecular cooperative binding of high-mobility-group protein I(Y) to the beta-interferon promoter. *Mol. Cell. Biol.* *17*, 3649–3662.
- Yie, J., Merika, M., Munshi, N., Chen, G., and Thanos, D. (1999). The role of HMG I(Y) in the assembly and function of the IFN-beta enhanceosome. *EMBO J.* *18*, 3074–3089.

- Yogev, O., Goldberg, R., Anzi, S., Yogev, O., and Shaulian, E. (2010). Jun proteins are starvation-regulated inhibitors of autophagy. *Cancer Res.* 70, 2318–2327.
- Young, M.R., and Colburn, N.H. (2006). Fra-1 a target for cancer prevention or intervention. *Gene* 379, 1–11.
- Young, M.R., Nair, R., Bucheimer, N., Tulsian, P., Brown, N., Chapp, C., Hsu, T.-C., and Colburn, N.H. (2002). Transactivation of Fra-1 and consequent activation of AP-1 occur extracellular signal-regulated kinase dependently. *Mol. Cell. Biol.* 22, 587–598.
- Young, N.L., Plazas-Mayorca, M.D., DiMaggio, P.A., Flaniken, I.Z., Beltran, A.J., Mishra, N., LeRoy, G., Floudas, C.A., and Garcia, B.A. (2010). Collective mass spectrometry approaches reveal broad and combinatorial modification of high mobility group protein A1a. *J. Am. Soc. Mass Spectrom.* 21, 960–970.
- Zanconato, F., Forcato, M., Battilana, G., Azzolin, L., Quaranta, E., Bodega, B., Rosato, A., Bicciato, S., Cordenonsi, M., and Piccolo, S. (2015). Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* 17, 1218–1227.
- Zhang, Q., and Wang, Y. (2007). Homeodomain-interacting protein kinase-2 (HIPK2) phosphorylates HMGA1a at Ser-35, Thr-52, and Thr-77 and modulates its DNA binding affinity. *J. Proteome Res.* 6, 4711–4719.
- Zhang, R., Chen, H.-Z., Liu, J.-J., Jia, Y.-Y., Zhang, Z.-Q., Yang, R.-F., Zhang, Y., Xu, J., Wei, Y.-S., Liu, D.-P., et al. (2010). SIRT1 suppresses activator protein-1 transcriptional activity and cyclooxygenase-2 expression in macrophages. *J. Biol. Chem.* 285, 7097–7110.
- Zhang, X., Wu, J., Luo, S., Lechler, T., and Zhang, J.Y. (2016a). FRA1 promotes squamous cell carcinoma growth and metastasis through distinct AKT and c-Jun dependent mechanisms. *Oncotarget* 7, 34371–34383.
- Zhang, Y., McCord, R.P., Ho, Y.-J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., and Dekker, J. (2012). Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148, 908–921.
- Zhao, H., and Dean, A. (2004). An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res.* 32, 4903–4919.
- Zhao, C., Qiao, Y., Jonsson, P., Wang, J., Xu, L., Rouhi, P., Sinha, I., Cao, Y., Williams, C., and Dahlman-Wright, K. (2014). Genome-wide Profiling of AP-1-Regulated Transcription Provides Insights into the Invasiveness of Triple-Negative Breast Cancer. *Cancer Res.* 74, 3983–3994.
- Zhao, X.-X., Yuan, Q.-Z., Mu, D.-P., Sun, D.-W., Bo, Q.-A., Pan, G.-Z., Li, G.-Q., Cui, T., Ding, P.-P., You, F.-P., et al. (2015). MicroRNA-26a inhibits proliferation by targeting high mobility group AT-hook 1 in breast cancer. *Int. J. Clin. Exp. Pathol.* 8, 368–373.
- Zhong, G., Chen, X., Fang, X., Wang, D., Xie, M., and Chen, Q. (2016). Fra-1 is upregulated in lung cancer tissues and inhibits the apoptosis of lung cancer cells by the P53 signaling pathway. *Oncol. Rep.* 35, 447–453.
- Zhou, V.W., Goren, A., and Bernstein, B.E. (2011). Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* 12, 7–18.

- Zhou, W., Zhong, C., Luo, X., Zhang, Y., Zhang, G., Zhou, D., and Liu, L. (2016). miR-625 suppresses cell proliferation and migration by targeting HMGA1 in breast cancer. *Biochem. Biophys. Res. Commun.* 470, 838–844.
- Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. (1995). Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* 376, 771–774.
- Zhu, X., Ling, J., Zhang, L., Pi, W., Wu, M., and Tuan, D. (2007). A facilitated tracking and transcription mechanism of long-range enhancer function. *Nucleic Acids Res.* 35, 5532–5544.
- Zöchbauer-Müller, S., Fong, K.M., Geradts, J., Xu, X., Seidl, S., End-Pfützenreuter, A., Lang, G., Heller, G., Zielinski, C.C., Gazdar, A.F., et al. (2005). Expression of the candidate tumor suppressor gene hSRBC is frequently lost in primary lung cancers with and without DNA methylation. *Oncogene* 24, 6249–6255.

Titre / Résumé

Fra-1 et Fra-2 dans les cancers du sein triple négatifs : mécanismes transcriptionnels et identification de cibles thérapeutiques potentielles. Les cancers du sein triple négatifs (TNBC) ne bénéficient d'aucune thérapie ciblée et restent incurables. L'identification de nouvelles cibles moléculaires diagnostiques et surtout, thérapeutiques constitue donc un enjeu majeur pour le traitement de ces cancers. Fra-1 et Fra-2, deux constituants du complexe transcriptionnel AP-1, sont surexprimés dans les TNBC où ils participent à l'agressivité tumorale et/ou la métastatisation. Les gènes cibles de Fra-1 et Fra-2, ainsi que les mécanismes moléculaires via lesquels ils gouvernent la transcription de leurs gènes cibles sont très peu caractérisés. Dans ce contexte, l'objectif général de ma thèse a été d'identifier les gènes codants sous contrôle de Fra-1 et/ou Fra-2 et de caractériser les sites de fixation de Fra-1 et Fra-2 sur la chromatine pour identifier des cibles directes, en combinant des approches transcriptomiques et génomiques combinées à l'interférence à l'ARN dans la lignée modèle MDA-MB231. Les résultats obtenus, associés à l'analyse de banques de données humaines, ont permis la sélection de gènes cibles pour les études mécanistiques et fonctionnelles. Le gène HMGA1, choisi en raison de son rôle maintenant bien établi dans l'agressivité des tumeurs épithéliales, a fait l'objet d'une étude portant sur les mécanismes transcriptionnels gouvernés par Fra-1 et/ou Fra-2 pour sa régulation. L'étude fonctionnelle a été focalisée sur CD68 dont l'expression est fortement induite par Fra-1 et Fra-2. CD68 code pour une protéine transmembranaire dont la fonction n'est pas clairement établie et dont l'implication dans le phénotype tumoral n'a jamais été étudiée.

Title / Abstract

Fra-1 and Fra-2 in triple negative breast cancers: transcriptional mechanisms and identification of potential therapeutic targets. Triple negative breast cancers (TNBC) are characterized by a poor prognosis and no targeted therapy is currently available. The identification of new diagnostic and therapeutic targets is crucial for the treatment of these cancers. Fra-1 and Fra-2, two members of the AP-1 transcriptional complex, are frequently overexpressed in TNBC, where they contribute to the tumorigenic phenotype. The panel of genes under the control of Fra-1 and/or Fra-2 in TNBC, as well as the molecular mechanisms by which they control their target gene expression are mostly unknown. The aim of my thesis was to identify the panel of genes controlled by Fra-1 and/or Fra-2 in TNBC and to characterize the binding sites of Fra-1 and Fra-2 on chromatin to select direct targets for further studies, by using transcriptomic and ChIP-seq approaches combined to RNAi in the model cell line MDA-MB231. The results allowed us to select target genes for transcriptional and functional studies. The study of the transcriptional mechanisms governed by Fra-1 and/or Fra-2 was carried out on the HMGA1 gene, already known for its crucial role in the aggressiveness of epithelial tumours. The functional study was focused on CD68, as its expression is highly induced by Fra-1 and Fra-2. CD68 encodes a transmembrane protein which cellular function is still not known and its potential role in tumorigenesis has not been studied yet.

Mots Clés / Key words

Cancer du sein / Breast Cancer / AP-1 / Fra-1 / Fra-2 / HMGA1 / CD68

Laboratoire

Institut de Génétique Moléculaire de Montpellier, équipe « Oncogenèse et Immunothérapies » CNRS UMR 5535, 1919 Route de Mende, 34293 Montpellier Cedex 5, France.