



HAL
open science

Rôle et pertinence clinique du facteur de transcription FOXS1 dans le carcinome hépatocellulaire

Kévin Bévant

► **To cite this version:**

Kévin Bévant. Rôle et pertinence clinique du facteur de transcription FOXS1 dans le carcinome hépatocellulaire. Médecine humaine et pathologie. Université de Rennes, 2021. Français. NNT : 2021REN1B011 . tel-03413106

HAL Id: tel-03413106

<https://theses.hal.science/tel-03413106>

Submitted on 3 Nov 2021

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.

THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1

ECOLE DOCTORALE N° 605

Biologie Santé

Spécialité : Cancérologie

Par Kevin Bévant

Rôle et pertinence clinique du facteur de transcription FOXS1 dans le carcinome hépatocellulaire

Thèse présentée et soutenue à Rennes, le 01 avril 2021

Unité de recherche : Inserm UMR 1242 - Oncogenesis, Stress & Signaling - COSS

Rapporteurs avant soutenance : Composition du Jury :

Jean-Charles Nault
MCU-PH
Centre de recherche des Cordeliers - Paris
Julie Pannequin
DR CNRS
Institut de Génomique Fonctionnelle - Montpellier

Marie-Thérèse Dimanche-Boitrel
DR Inserm
U1085 IRSET - Rennes
Laura Fouassier
CR Inserm
Centre de Recherche Saint-Antoine (CRSA) - Paris
Jean-Charles Nault
MCU-PH
Centre de recherche des Cordeliers – Paris
Président : Eric Ogier-Denis
DR Inserm
U1242 COSS – Rennes
Julie Pannequin
DR CNRS
Institut de Génomique Fonctionnelle - Montpellier
Directeur de thèse : Cédric Coulouarn
CR Inserm
U1242 COSS - Rennes

Remerciements

Je tiens à remercier en premier lieu Cédric Coulouarn pour m'avoir encadré au cours de ces trois années de thèse et pour m'avoir appris beaucoup sur la science et la recherche. Ta rigueur, ton encadrement et ta bienveillance m'ont permis d'effectuer ma thèse dans les meilleures conditions.

Je remercie Jean-Charles Nault et Julie Pannequin pour avoir m'avoir fait l'honneur d'être les rapporteurs de ma thèse ainsi que Laura Fouassier, Hélène Beloeil et Eric Ogier-Denis pour en être les examinateurs.

Je remercie particulièrement Gaëlle Angenard pour m'avoir formé depuis le master 1, pour avoir été d'une aide considérable tout au long de ma thèse mais également pour ta bonne humeur et tous les bons moments que nous avons passé au laboratoire.

Je remercie Corentin et Matthis, les camarades de crimes sur qui j'ai pu compter dans les bonnes comme les mauvaises périodes. Ça a été un vrai plaisir de travailler avec vous !

Je remercie Raphaël pour toute son aide, son expertise et sa bonne humeur.

Je remercie également les autres membres de l'équipe : David, Delphine et Panagiotis, avec qui j'ai eu plaisir à échanger sur mon sujet comme sur le leur.

Je remercie tous les doctorants de l'Institut Numecan avec qui j'ai pu travailler ou partager pendant ma thèse. Je remercie tout particulièrement Yoann, Julien, Yann, Kilian, Clémence, Luis, Elise Lelou, Elise Vene, Thomas Gautier et Thomas Guerbette. Les bons moments que nous avons passé au labo comme ailleurs ont rendu ma thèse très agréable. Je n'oublierai pas les nombreuses soirées passées à refaire le monde et la science !

Je remercie tout autant les membres passés de Numecan : Simon, Dounia, Moussa et Tatiana avec qui j'ai passé de très bons moments et qui m'ont beaucoup appris.

Je remercie du fond du cœur Sophie pour son soutien indéfectible et les bons moments que nous avons partagés.

Je remercie aussi tout le personnel de Numecan avec qui j'ai eu le plaisir de travailler. Remerciement spéciale pour Michèle qui a été d'une aide incommensurable. Merci également

à Claudine, Grégory, Adina, Karen, Thomas Poussou, Patricia Jouas et Patricia Leroyer pour leur aide et leur bonne humeur.

Je remercie également tous les doctorants et personnels de l'Institut COSS pour leur accueil et leur bienveillance. Un grand merci à Jérôme pour avoir fait le lien entre Eugène Marquis et le bâtiment 15, ainsi qu'à Vesna, Simon et Manon. Je regrette que le contexte sanitaire ne nous ait pas permis de nous voir davantage en dehors du laboratoire.

Je remercie l'association des doctorants LUCA et tous ses membres. J'ai eu plaisir à organiser des événements avec vous et partager ensemble nos expériences de recherche. J'espère que cette association perdura encore pendant des années ! Merci également à Laurent et Romain et leur bar L'amaryllis pour nous avoir servi de quartier général.

Je remercie bien évidemment tous mes amis qui m'ont supportés et encouragés tout au long de ma thèse

Je remercie la région Bretagne ainsi que les comités 22, 35 et 44 de la ligue contre le cancer sans qui ma thèse n'aurait pas pu être réalisée.

Je remercie Gilles Salbert et Raphaël Métivier pour m'avoir formé aux expériences de CHIP.

Je remercie tous les membres de mon CSI : Nathalie Théret, David Gilot, Massimo Levrero et Marie-Dominique Galibert pour avoir suivi mes travaux au cours de ma thèse et pour les précieux conseils qu'ils m'ont fournis.

Je remercie Badia Mssassi pour toute son aide concernant les démarches administratives de ma soutenance.

Je remercie Alexandra Elbakyan pour avoir facilité considérablement ma veille bibliographique.

Enfin je remercie ma famille et notamment mes parents Yann et Marie-Christine et mes frères Alan et Killian pour leur soutien sans faille, leurs encouragements permanents et toute l'aide qu'ils m'ont apportées.

Liste des publications & des communications

PUBLICATIONS

- Bévant, K., Desoteux, M., Angenard, G., Caruso, S., Louis, C., Papoutsoglou, P., Sulpice, L., Gilot, D., Zucman-Rossi, J., and Coulouarn, C. TGF β -induced TFOX controls epithelial-mesenchymal transition and predicts a poor prognosis in liver cancer. **Under revision in Hepatology, HEP-20-1574.**
- Bévant, K. & Coulouarn, C. A novel immunosuppression-based classification of liver tumors opens new perspectives for adapted therapeutic strategies. *EBioMedicine* 54, (2020).
- Bévant, K. & Coulouarn, C. Landscape of genomic alterations in hepatocellular carcinoma: current knowledge and perspectives for targeted therapies. *Hepatobiliary Surg Nutr* 6, 404–407 (2017).

COMMUNICATIONS

Oral

- 8ème colloque de Génomique Fonctionnelle du Foie à Rennes en 2020
- 2^{ème} Journées scientifiques de l'Ecole Doctorale de l'UBL 2019 à Angers

Poster

- 8ème colloque de Génomique Fonctionnelle du Foie à Rennes en 2020 (Prix poster)
- 12th ILCA (International Liver Cancer Association) à Londres en 2018
- 12ème Journées du CGO (Cancéropôle Grand Ouest) à Vannes en 2018
- 5ème Journée SCAN (Signalisation et Cancer) à Rennes en 2018 (Prix poster)
- GIS "Nutrition-Alimentation-Métabolisme-Santé" (NAMS) à Rennes en 2018

Table des matières

Table des matières	5
Légendes	8
Abréviations	10
Chapitre 1 : Introduction.....	15
I) Le carcinome hépatocellulaire.....	15
1) Epidémiologie	15
2) Facteurs de risques	17
3) Histoire naturelle du carcinome hépatocellulaire.....	22
4) Classification des tumeurs hépatiques.....	29
5) Traitement.....	35
II) Le Transforming Growth Factor β et la transition épithélio-mésenchymateuse.....	44
1) Fonctionnement de la signalisation canonique	44
2) Signalisation non canonique	47
3) Réponse cellulaire au TGF β	47
4) Rôle du TGF β	49
a) Développement.....	49
b) Immunologie.....	50
c) Fibrose.....	50
5) La transition épithélio-mésenchymateuse.....	51
a) EMT de type 1 :	53
b) EMT de type 2 :	54
c) EMT de type 3 :	55
d) Modifications phénotypiques	58
e) Régulation de l'EMT	60
6) Induction de l'EMT par le TGF β	64
7) TGF β et carcinome hépatocellulaire	66
III) Les facteurs de transcription de la famille des forkhead box	71
1) Généralités	71
2) Les FOXA.....	72

3) Les FOXO	76
4) Les FOXP	76
5) FOXG1/K1/M1/Q1	77
6) Les FOXC	78
7) FOXS1	79
Chapitre 2 : Rationnel du projet de thèse.....	85
Objectifs du projet	86
Chapitre 3 : Résultats des travaux de thèse.....	87
I) Expression, localisation et régulation de FOXS1	87
1) Identification de FOXS1 par analyse transcriptomique	87
2) Expression de FOXS1 dans les lignées de CHC.....	87
3) Localisation de FOXS1	88
4) Régulation de <i>FOXS1</i> par les SMAD.....	89
5) Régulation de <i>FOXS1</i> par méthylation.....	95
6) Autres mécanismes de régulation.....	99
II) Rôle de FOXS1	101
1) Caractérisation des lignées FOXS1 positives	101
2) Analyse transcriptomique de l'extinction de FOXS1.....	104
3) Analyse moléculaire et fonctionnelle du gain et perte de fonction de <i>FOXS1</i>	107
4) Analyse des gènes cibles de FOXS1 par CHIP-seq.....	110
III) Pertinence clinique de FOXS1	117
1) Expression de FOXS1 dans les tumeurs.....	117
2) Association avec la survie des patients.....	118
Chapitre 4 : Discussion et conclusion	122
I) Discussion	122
1) Régulation de <i>FOXS1</i>	123
2) Fonction de FOXS1.....	127
3) Pertinence clinique de FOXS1	131
II) Conclusion	136
Bibliographie.....	138
Annexes	154
Figures supplémentaires.....	155

Tables	173
Articles.....	192

Légendes

Figures

Figure 1 : Incidence et facteurs de risques du carcinome hépatocellulaire.	18
Figure 2 : Mutations communément retrouvées dans le CHC.	24
Figure 3 : Signalisation Wnt dans le CHC.	27
Figure 4 : Impact du microenvironnement dans le carcinome hépatocellulaire.	30
Figure 5 : Classification intégrative et pathologique des carcinomes hépatocellulaires. .	32
Figure 6 : Stratégie thérapeutique contre le carcinome hépatocellulaire d’après le BCLC et essais cliniques sur la thérapie combinée.	40
Figure 7 : Classification immunitaire des carcinomes hépatocellulaires.	43
Figure 8 : Biosynthèse et signalisation du TGFβ.	46
Figure 9 : Mécanismes et marqueurs de la transition épithélio-mésenchymateuse dans les hépatocytes.	62
Figure 10 : Induction de l’EMT par le TGFβ.	65
Figure 11 : Dualité fonctionnelle du TGFβ dans le carcinome hépatocellulaire.	70
Figure 12 : Phylogénie et implications de différents membres de la famille des forkhead box dans les “hallmarks” du cancer.	73
Figure 13 : Expression de FOXS1 chez l’humain.	81
Figure 14 : Historique des études du rôle de FOXS1 en cancérologie.	84
Figure 15 : Expression et localisation de FOXS1	90
Figure 16 : FOXS1 est régulé par la signalisation canonique du TGFβ.	94
Figure 17 : L’expression de FOXS1 est impactée par la méthylation de son locus.	97
Figure 18 : FOXS1 est exprimé dans les lignées mésenchymateuses.	102
Figure 19 : L’extinction de FOXS1 induit des modifications associées à l’EMT et l’invasion tumorale.	105
Figure 20 : Analyse moléculaire et fonctionnelle du gain et perte de fonction de FOXS1.	112
Figure 21 : Identification de gènes cibles de FOXS1.	116
Figure 22 : Pertinence clinique de l’expression de FOXS1.	121
Figure 23 : Rôle et pertinence clinique de FOXS1 dans le CHC.	137

Figures supplémentaires

Figure supplémentaire 1 : <i>FOXS1</i> est une cible précoce du TGFβ.	155
Figure supplémentaire 2 : Analyse de l’expression génique dans des coupes de tissus humains.	156
Figure supplémentaire 3 : Organisation génomique et régulation transcriptionnelle de <i>FOXS1</i>	157

Figure supplémentaire 4 : Validation des off-targets prédits des siRNA contre FOXS1.	159
Figure supplémentaire 5 : Validation du gain et perte de fonction de FOXS1.	161
Figure supplémentaire 6 : Surexpression endogène de <i>FOXS1</i> par le système CRISPR/cas9 Synergistic Activation Mediator (SAM).	162
Figure supplémentaire 7 : Extinction de <i>FOXS1</i> dans les Hep3B.	163
Figure supplémentaire 8 : Analyse fonctionnelle de <i>FOXS1</i>.	164
Figure supplémentaire 9 : Expression de <i>FOXS1</i> dans différents modèles d'induction de l'EMT.	166
Figure supplémentaire 10 : L'expression de <i>FOXS1</i> est augmentée dans le CHC et les métastases par rapport au tissu non tumoral.	168
Figure supplémentaire 11 : Expression de <i>FOXS1</i> dans 37 cancers du TCGA.	169
Figure supplémentaire 12 : Expression de <i>FOXS1</i> dans les CHC humain (n=193).	170
Figure supplémentaire 13 : L'expression de <i>FOXS1</i> est associée à une survie sans récurrence plus faible.	172

Tables

Table 1 : Liste des gènes différentiellement exprimé dans les SNU-449 après 72h de traitement.	173
Table 2 : Caractéristiques des lignées cellulaires humaines de CHC.	179
Table 3 : Liste des gènes inclus dans la signature siFOXS1.	180
Table 4 : Signalisations et voies enrichies dans le chIP-seq contre FOXS1-HA après association de chaque pic au site d'initiation de la transcription le plus proche.	185
Table 5 : Signalisations et voies enrichies dans le chIP-seq contre FOXS1-HA après sélection des pics à moins de 10kb du site d'initiation de la transcription le plus proche.	189

Abréviations

AAT : Alpha 1-antitrypsin
ADAM : Disintegrin and metalloproteinase domain-containing protein
ADN : Acide desoxyribonucléique
AFP : Alpha-fetoprotein
APC : Adenomatous polyposis coli
AR : Androgen receptor
ARN : Acide ribonucléique
ASO : Oligonucléotide antisens
ATF : Activating transcription factor
ATM : *ataxia telangiectasia mutated*
AXL : Tyrosine-protein kinase receptor UFO
BAMBI : BMP and activating membrane-bound inhibitor homolog
BCLC : Barcelona Clinic Liver Cancer
BLBC : basal-like breast cancer
BMP : Bone morphogenetic protein
BRCA1/2 : breast cancer 1/2
CAF : Cancer-associated fibroblast
CCL : Chemokine (C-C motif) ligand
CDKIs : cyclin-dependent kinase inhibitors
CHC : Carcinome hépatocellulaire
CHOL : Cholangiocarcinoma
ChIP : Chromatin immunoprecipitation
CSCs : Cellules souches cancéreuses
CTCF ; CCCTC-binding factor
CTCs : circulating tumor cells
CXCR1 : C-X-C motif chemokine receptor 1

DBE : Daf-16 family member-binding element
DDR2 : Discoidin domain-containing receptor 2
DLG : Discs large
DNMT : DNA methyl transferases
ECM : Extracellular matrix
EGF : Epidermal Growth Factor
EGR1 : Early growth response protein 1
EndMT : Endothelial-mesenchymal transition
EMT : Epithelial-to-mesenchymal transition
EMT-TF : Facteur de transcription contrôlant l'EMT
ER : Estrogen receptor
ERK : Extracellular signal-Regulated Kinases
FGF : Fibroblast growth factor
FKH : Forkhead box
FOX : Forkhead box protein
GEF : Guanine nucleotide exchange factors
GLI1 : Zinc finger protein GLI1
GIUC / Gaussia luciferase
GSK-3 β : Glycogen synthase kinase-3 beta
HA : Hémagglutinine
HAT : Histones acetyl transferases
HDAC : Histone deacetylase
hESC : human embryonic stem cell
HGF : Hepatocyte Growth Factor
HIF-1 α : Hypoxia-inducible factor-1-alpha
HMGA2 : High-mobility group AT-hook 2
HNF : Hepatic Nuclear Factor
HSC : Hepatic stellate cells

ICI : Immune checkpoint inhibitor
ID1 : Inhibitor of DNA binding 1
INF- γ : Interferon gamma
IRES : Internal ribosome entry site
JAK : Janus kinases
JNK : c-Jun N-terminal kinases
KICH : Kidney chromophobe
KIRC : Kidney renal clear cell carcinoma
KIRP : Kidney renal papillary cell carcinoma
LAP : Latency-associated protein
LGL : Lethal giant larvae
LIHC : Liver hepatocellular carcinoma
LLC : Large latent complex
LTBP : Latent TGF β binding protein
MAD : Mothers against decapentaplegic
MAPK : Mitogen-activated protein kinases
MELD : Model for End-stage Liver Disease
MET : Mesenchymal-to-epithelial transition
miR : microARN
MMP : matrix metalloproteinase
mTOR : mechanistic target of rapamycin
NAFLD : Non-alcoholic fatty liver disease
NASH : Non alcoholic steatohepatitis
NSCLC : Non small cell lung cancer
NF- κ B : nuclear factor kappa-light-chain-enhancer of activated B cells
NGS : Séquencage de nouvelle generation
NK : Natural killer
NOX4 : NAPDH oxidase 4

PARP6 : Protein mono-ADP-ribosyltransferase PARP6

PD-1 : Programmed cell death 1

PDGF : Platelet-derived growth factor

PD-L1 : Programmed cell death ligand 1

PI3K : Phosphoinositide 3-kinase

PRC2 : Polycomb repressive complex 2

RNF6 : Ring Finger Protein 6

ROCK : Rho-associated protein kinase

ROS : Reactive oxygen species

RT-qPCR : Reverse transcription quantitative polymerase chain reaction

SAM : Synergistic activation mediator

SARA : SMAD anchor for receptor activation

SBE : SMAD binding element

SERPINA1 : Serine protease inhibitor A1

SKIL : Ski-like protein

SLC : Small latent complex

SMA : Small worm phenotype

SMAD : Mothers against decapentaplegic homolog

SMURF : E3 ubiquitin-protein ligase SMURF

SOX : SRY-related HMG-box genes

STAD : Stomach adenocarcinoma

STAT : signal transducer and activator of transcription proteins

SWI/SNF : SWItch/Sucrose Non-Fermentable

TACE : Chimioembolisation par voie artérielle

TAK1 (NR2C2) : Nuclear receptor subfamily 2 group C member 2

TAM : Tumor-associated macrophages

TBET : T-box transcription factor TBX21

TCGA : The Cancer Genome Atlas

TGF β : Transforming Growth Factor beta
TGFBR : Transforming Growth Factor beta Receptor
TRAF6 : TNF receptor-associated factor 6
TRAIL : TNF-related apoptosis-inducing ligand
TSS : Transcription starting site
UCSC : University of California, Santa Cruz
UCSF : University of California, San Francisco
USA : Etats-Unis d'Amérique
VEGF : Vascular Endothelial Growth Factor
VEGFR : Vascular Endothelial Growth Factor Receptor
VHB/HBV : Virus de l'hépatite B
VHC/HCV : Virus de l'hépatie C
XBP1 : X-box binding protein 1
ZEB2-AS1 : ZEB2 antisens 1
ZO-1 : Zonula occludens 1/Tight junction protein 1

Chapitre 1 : Introduction

I) Le carcinome hépatocellulaire

Le cancer du foie est la quatrième cause de mortalité liée au cancer dans le monde et le sixième cancer en termes d'incidence. Ainsi en 2018, 782 000 personnes sont mortes des suites de ce cancer et 841 000 nouveaux cas ont été diagnostiqués. Ce cancer a une incidence et une mortalité deux à trois fois plus élevée chez l'homme que chez la femme, ce qui en fait le 2^{ème} plus meurtrier chez l'homme. Les cancers du foie sont divisés en deux types majeurs auxquels s'ajoutent des formes rares. Le premier est le cholangiocarcinome intrahépatique qui représente entre 10 et 15% des cas et le second, beaucoup plus important, est le carcinome hépatocellulaire (CHC) représentant 75 à 85% des cancers du foie.¹

1) Epidémiologie

Le cancer du foie est présent dans tous les pays du monde mais il existe une forte disparité géographique quant à son incidence et sa mortalité, certaines régions du monde étant ainsi considérées comme à haut risques (Figure 1A). Le cancer du foie est le cancer le plus répandu dans 13 pays incluant l'Égypte, la Gambie, la Guinée, la Mongolie, le Cambodge et le Vietnam.¹ Les zones les plus à risques sont ainsi l'Afrique du Nord et de l'Ouest ainsi que l'Est et le Sud-est de l'Asie.¹ Cette dernière regroupe autour de 75% des cancers du foie mondiaux avec la Chine portant près de 50% des cas totaux. C'est toutefois la Mongolie qui possède la plus forte incidence avec des taux quatre fois supérieurs à ceux de la Chine et de la Corée du Sud en 2018.^{1,2} Dans ces régions à haut risques, les principales causes de CHC sont la consommation de nourriture contaminée par l'aflatoxine B1 et l'infection par le virus de l'hépatite B (VHB) bien que l'hépatite C soit également en cause dans certains pays, notamment en Afrique du Nord (Égypte) et en Mongolie.^{1,3,4} L'incidence est toutefois en diminution dans certaines zones

à risques, notamment en Chine grâce aux programmes visant à réduire l'exposition à l'aflatoxine contaminant la nourriture et aux campagnes de vaccination contre le VHB.³

A l'inverse, les zones les moins touchées par le cancer du foie sont l'Amérique, l'Europe, l'Océanie et le Moyen-Orient. L'incidence a toutefois augmenté dans ces pays jusqu'à la fin des années 2010. Les causes d'apparition y sont différentes de celles des zones à haut risques et principalement liées aux infections par le virus de l'hépatites C (VHC), la consommation d'alcool et la maladie dite du foie gras d'origine non alcoolique ou *non-alcoholic fatty liver disease* (NAFLD).^{1,3,4} Ces différences s'expliquent en partie par des mesures sanitaires plus efficaces et la mise en place de normes dans les pays développés sur la consommation maximale d'aflatoxine ainsi que par les campagnes de vaccination contre l'hépatite B réduisant fortement l'apparition de CHC lié à ces facteurs. L'hépatite C a joué un rôle important dans l'incidence du CHC dans les pays développés notamment pour la génération née entre 1945 et 1965 à l'origine d'un pic d'hépatite C aux USA au début des années 1980 et en Australie dans les années 2000.^{3,5} La progression du CHC dans ces pays est également fortement liée à la progression de l'obésité et de la NAFLD qui est maintenant la maladie du foie la plus répandue dans les pays développés et est responsable de 10 à 20% des CHC aux USA.³

Les différences de facteurs de risques entre les zones à haut et faible risques se traduisent également par une différence sur l'âge d'apparition du CHC. Le CHC apparaît en moyenne après 60 ans dans les pays développés (Japon, USA, Europe) tandis que la maladie apparaît entre 30 et 60 ans dans les zones à risques (59 ans en Corée du Sud, 52 en Chine et 45 en Afrique). Des études ont également mis en évidence le lien entre le CHC VHB-induit et l'apparition précoce du CHC dans des populations africaines.^{3,4,6}

Il existe de plus une disparité de genre sur l'incidence du CHC. Le taux de CHC est en moyenne deux à trois fois plus élevé chez l'homme que chez la femme, sans que cette disparité varie selon les zones à haut ou faible risques (bien que la différence homme/femme soit encore plus marquée en Europe de l'Est).² Des variations épigénétiques, immunitaires ou dans les hormones stéroïdes sexuelles ont été proposées pour expliquer cette disparité mais les mécanismes sous-jacents sont encore mal connus. En revanche, quels que soient le genre ou la localisation géographique, les causes génétiques et environnementales d'apparition du CHC sont relativement bien documentées.

2) Facteurs de risques

Le CHC résulte le plus souvent de la transformation maligne des hépatocytes, les cellules parenchymateuses du foie. Cette transformation se produit comme pour tous les cancers suite à l'accumulation de mutations et d'altérations génétiques, de modifications épigénétiques et de dérégulations de voies de signalisation conduisant la cellule à proliférer de manière incontrôlée et à résister aux signaux apoptotiques. Comme pour d'autres cancers, la présence de cellules souches cancéreuses (CSCs) a été identifiée dans le CHC. Ces CSCs ont toutefois des origines différentes suivant les tumeurs. Elles peuvent émerger directement de populations de cellules souches telle que les cellules souches progénitrices hépatiques (HPCs), des cellules multipotentes, localisées dans le foie au niveau des canaux de Hering, ou bien des hépatoblastes, des cellules bipotentes capables de se différencier en hépatocytes ou cholangiocytes. Ces cellules souches sont particulièrement vulnérables à une transformation tumorale dans des contextes d'inflammations chroniques du foie entraînant un renouvellement rapide des hépatocytes et donc une accumulation d'altérations génétiques et épigénétiques. A l'inverse, les CSCs peuvent émerger directement de cellules cancéreuses ayant effectuées une retro-différenciation. Ces CSCs ne représentent qu'une fraction des cellules cancéreuses et plusieurs origines peuvent coexister au sein d'une même tumeur, les CHC étant des tumeurs extrêmement hétérogènes.⁷

La transformation tumorale des hépatocytes et l'apparition du CHC résulte comme souvent en oncologie d'une interaction complexe entre des facteurs génétiques et l'environnement. Si l'environnement joue un rôle prépondérant dans la mise en place de ce cancer (80% des CHC se développent sur foie cirrhotique), des prédispositions génétiques au CHC ont tout de même été mises en évidence. Celles-ci restent toutefois extrêmement rares au regard d'autres cancers comme les cancers ovariens, du sein et colorectaux. La plupart des syndromes monogéniques retrouvés pour ces cancers (comme le syndrome de Lynch, les mutations de *BRCA1* et *BRCA2*, etc) ne sont pas associés au développement de tumeurs primaires hépatiques hormis les mutations d'*APC* qui peuvent exceptionnellement entraîner l'apparition du CHC.⁸

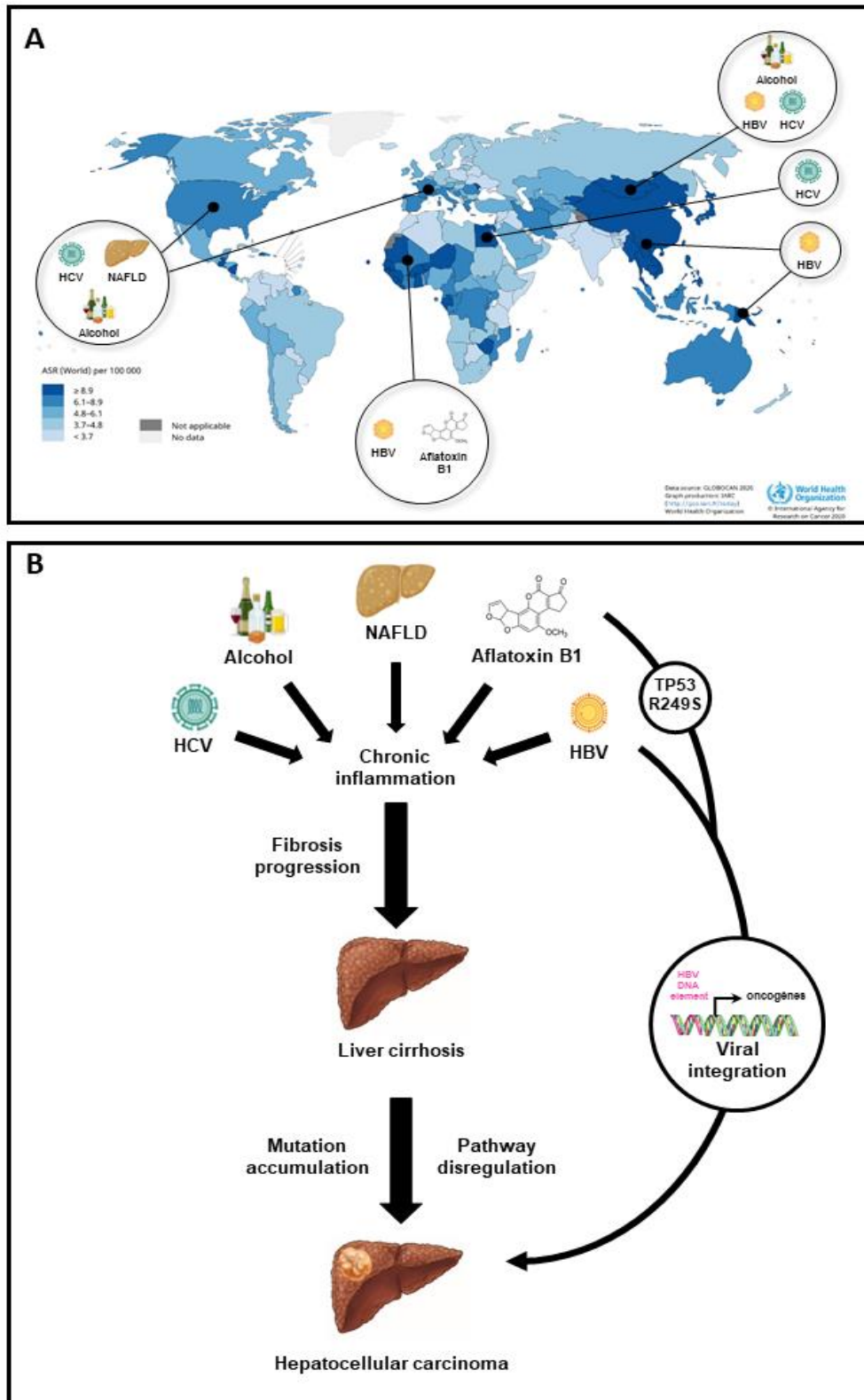


Figure 1 : Incidence et facteurs de risques du carcinome hépatocellulaire.

A : Incidence standardisée du cancer du foie en 2020 à travers le monde avec les étiologies associées. B : Facteurs de risques du CHC et mécanismes de cancérogenèse associés.

Légende : HCV : Virus de l'hépatite C ; NAFLD : *Non alcoholic fatty liver disease* ; HBV : Virus de l'hépatite B

La plupart des altérations génétiques favorisant la survenue du CHC sont associées à des maladies métaboliques entraînant l'apparition de cirrhose. La déficience en Alpha1-antitrypsin (AAT) est le syndrome monogénique « le plus courant » dans les tumeurs hépatiques. Ce syndrome est causé par des mutations de *SERPINA1*, codant pour un inhibiteur de serine protéase, qui va alors s'accumuler dans le réticulum endoplasmique des hépatocytes qui l'expriment fortement. L'accumulation de cette protéase va alors favoriser le développement de maladies hépatiques par rupture de la balance protéase/antiprotéase. Cette déficience est toutefois extrêmement rare, le phénotype muté étant récessif, et, de plus, seulement 12-15% des individus avec le génotype muté développent des maladies hépatiques.⁹ Les autres maladies monogéniques incluent l'hémochromatose, la porphyrie et la tyrosinémie de type I. Des altérations génétiques impactant le métabolisme du glucose et le stockage du glycogène telle que la maladie de Von Gierke ou des sous-types de diabète de type 3 peuvent favoriser l'apparition d'adénomatose hépatique pouvant se transformer en CHC en l'absence de cirrhose.

Il existe de plus des polymorphismes génétiques associés au risque de CHC. De nombreux polymorphismes d'un seul nucléotide impactant des signalisations cellulaires telle que l'inflammation via *TNFA*, *IL1B* ou *TGFB*, le stress oxydatif, le métabolisme du fer, le cycle cellulaire et la signalisation EGFR sont susceptibles de favoriser la cirrhose et le CHC à différentes étapes de leur apparition et de leur progression.¹⁰⁻¹⁵

La cirrhose est donc un terreau fertile pour le développement du CHC, l'inflammation chronique et la forte prolifération des hépatocytes dans ce contexte favorisant l'apparition d'altérations susceptibles d'entraîner une transformation maligne des hépatocytes. Ceci fait de toutes les maladies inflammatoires chroniques du foie, des facteurs de risques pour l'apparition du CHC. Parmi elles, on trouve entre autres les infections par les virus de l'hépatite B et C, la contamination par l'aflatoxine B1, la consommation excessive d'alcool, le diabète, l'obésité et la NAFLD en découlant (Figure 1B).

Les hépatites chroniques B et C comptent parmi les principales causes de CHC et sont responsables de l'apparition de près de 80% des CHC, l'hépatite B étant responsable de plus de la moitié à elle seule.¹⁶ L'incidence du CHC est de 2-5% chaque année pour les individus avec une cirrhose induite par le VHB ou le VHC. Les mécanismes entraînant l'apparition de CHC VHB-induit sont complexes et de nombreuses voies de cancérogenèse ont été identifiées. La

principale étant la cirrhose induite par l'infection chronique du VHB mais l'activation d'oncogènes due à l'intégration d'élément viral dans le génome des cellules a également été mis en évidence et de nombreux cas de CHC sur foie non cirrhotique sont liés à une infection par le VHB (Figure 1B), essentiellement dans les régions où il est endémique (30-50% des CHC VHB-induit sont non cirrhotiques en Asie de l'Est et en Afrique contre moins de 10% aux USA où le VHB n'est pas endémique).^{3,17,18} Le VHC quant à lui est davantage responsable de l'augmentation de l'incidence du CHC dans les pays développés (en parallèle de l'obésité et de la consommation d'alcool). De la même manière que pour le VHB, les infections chroniques par le VHC vont favoriser l'apparition du CHC en provoquant l'apparition d'une fibrose hépatique via l'induction de l'inflammation et de stress oxydatif dans le foie. Toutefois, contrairement au VHB, il s'agit d'un virus à ARN qui ne va donc pas s'intégrer dans le génome des hépatocytes. Des études récentes ont toutefois montré que le risque d'apparition du CHC chez des personnes guéries du VHC était supérieur à celui chez des personnes n'ayant jamais été infectées et ce quelle que soit la thérapie utilisée contre le VHC. Cette augmentation des risques serait liée à des modifications épigénétiques induites par le VHC, notamment des modifications sur l'ensemble du génome de l'acétylation de la lysine 27 de l'histone H3 (H3K27ac, marqueur d'enhancer actif) entraînant des modifications d'expression de certains gènes associées à l'apparition du CHC comme *SOX9*, *SPHK1* et *ETV4*.¹⁹ L'infection par le VHC agit de plus en synergie avec d'autres maladies chroniques du foie comme l'hépatite B ou la consommation d'alcool augmentant significativement les risques de CHC.¹⁷

Parmi les autres facteurs de risques, la NAFLD est à l'origine de 10 à 20% des CHC aux USA.²⁰ Cette maladie, caractérisée par une accumulation de triglycérides dans le foie, est en augmentation dans les pays développés et est maintenant la maladie du foie la plus répandue et un des principaux facteurs de risque pour le CHC dans ces pays. Il s'agit d'une manifestation du syndrome métabolique, de l'obésité et du diabète de type 2. La stéatose hépatique non alcoolique (NASH) est une forme avancée de NAFLD qui peut facilement évoluer vers une fibrose puis une cirrhose et un CHC (Figure 1B). La physiopathologie de la cancérogenèse induite par la NASH est vraisemblablement due à une augmentation de l'oxydation des acides gras suite à la surcharge en triglycérides conduisant à une augmentation des espèces réactives de l'oxygène (ROS) et ainsi à un stress oxydatif dans les hépatocytes.^{17,21} La consommation

chronique d'alcool est le second facteur de risque pour le CHC dans les pays développés.⁴ L'alcool favorise également l'apparition du CHC en induisant la cirrhose (Figure 1B) bien que les risques d'apparition de CHC soient plus faibles sur cirrhose induite par l'alcool que sur cirrhose induite par infection virale.²²

L'aflatoxine B1, quant à elle, est une mycotoxine produite par *Aspergillus flavus* et *Aspergillus parasiticus*. Les pays d'Afrique de l'Ouest sont fortement exposés à cette toxine. Elle induit le CHC essentiellement en causant une mutation R249S dans le gène *TP53*. L'aflatoxine B1 agit de plus en synergie avec le VHB, celui-ci augmentant le risque de mutations de *TP53* (Figure 1B) en favorisant la métabolisation de la forme inactive de l'aflatoxine en sa forme mutagène et en inhibant le système de réparation de l'ADN par excision de base.²³ Cette inhibition est due à l'intégration dans l'ADN de la cellule du gène viral HBVx, qui va réduire l'interaction de p53 avec XBP1,²⁴ interaction nécessaire à la correction de la mutation induite par l'aflatoxine.²⁵

Ces facteurs de risques sont généralement à l'origine de la transformation maligne des hépatocytes. L'apparition et la progression du CHC est toutefois un processus lent (plusieurs années) découlant de nombreuses altérations successives des processus biologiques normaux opérant dans le foie.

3) Histoire naturelle du carcinome hépatocellulaire

Le CHC émerge dans plus de 80% des cas d'un foie cirrhotique. Exception faite des rares origines purement génétique et des CHC non cirrhotiques provenant d'insertion virale du VHB, le CHC émerge d'un long processus de transformation maligne d'hépatocytes très proliférants entourés de tissus fibreux dans le foie cirrhotique. Ces nodules de régénération résultent d'un cycle agression (généralement par les maladies chroniques responsable de la cirrhose)/régénération (prolifération des hépatocytes pour combler la mortalité induite par l'agression). Sur le long terme, ces nodules de régénération peuvent évoluer vers des nodules dysplasiques de bas et de haut grade qui seront à l'origine d'une tumeur précoce de CHC progressant vers des stades avancés.²⁶ Cette évolution se fait essentiellement via l'accumulation de mutations dans les hépatocytes. En moyenne, l'accumulation de 80 mutations est observée

au cours de la progression des CHC, dont une soixantaine sont des mutations non-silencieuses.²⁷ Parmi ces mutations, 2 à 6 sont des mutations dites *drivers* déclenchant la transformation tumorale des hépatocytes. Le type d'altérations génétiques et les gènes impactés varient énormément d'une tumeur à l'autre mais il existe un certain nombre de gènes et voies de signalisation associées qui sont retrouvés entre différents CHC (figure 2).

L'événement le plus courant et également le plus précoce parmi les CHC (et même une partie importante des lésions pré-néoplasiques) est la réactivation de la télomérase. Cette enzyme sert à la synthèse des télomères, qui raccourcissent au cours des cycles de division cellulaire, et permet ainsi d'éviter la sénescence et la mort par apoptose. Elle est réactivée dans plus de 90% des CHC avancés et les mécanismes moléculaires de maintenance des télomères employés par les 10% restant sont pour le moment inconnus. L'étape de réactivation de la télomérase (codé par le gène *TERT*) est variable selon les tumeurs mais toujours précoce et peut apparaître dès les premières dysplasies de bas grade (6% possèdent des mutations du promoteur de *TERT*).¹⁰ Le mécanisme principal de réactivation est une mutation du promoteur de *TERT* sur un *hotspot* localisé à 124 paires de bases en amont du codon start. Cette mutation représente 94% des mutations réactivatrices de *TERT* et est également retrouvée dans d'autres cancers tels que le mélanome où elle est induite par les UV.²⁸ La mutation du promoteur est le processus de réactivation le plus courant (30-60% des cancers) mais d'autres mécanismes moléculaires ont été identifiés.²⁹

Parmi les CHC VHB-induit, les mutations du promoteur de *TERT* sont beaucoup plus rares du fait d'un mécanisme alternatif de réactivation de la télomérase. Le génome du VHB est capable de s'intégrer dans le génome des hépatocytes par un mécanisme de mutagenèse insertionnelle. L'insertion des éléments viraux se fait généralement au niveau d'oncogènes tels que *MLL4*, *CCNE1* et *TERT*. Cette capacité du VHB à activer des oncogènes explique la forte proportion de CHC induit sur foie non cirrhotique. L'insertion d'éléments viraux dans le promoteur de *TERT* est retrouvé dans 26% des CHC VHB-induit et est mutuellement exclusive avec les mutations du promoteur, expliquant la plus faible proportion de mutation dans les CHC VHB-induit.²⁹

Enfin, la télomérase peut être réactivée par des altérations chromosomiques. L'amplification du gène *TERT* a été identifiée par des approches de *whole genome sequencing* (WGS) et *whole*

Pathway(s)	Gene(s)	Alteration	Frequency in HCC
Telomere maintenance	<i>TERT</i>	Promoter mutation	54–60%
		<i>Amplification</i>	5–6%
Cell cycle control	<i>TP53</i>	Mutation or deletion	12–48%
	<i>RB1</i>	Mutation or deletion	3–8%
	<i>CCND1</i>	Amplification	7%
	<i>CDKN2A</i>	Mutation or deletion	2–12%
WNT– β -catenin signalling	<i>CTNNB1</i>	Mutation	11–37%
	<i>AXIN1</i>	Mutation or deletion	5–15%
Oxidative stress	<i>NFE2L2</i>	Mutation	3–6%
	<i>KEAP1</i>	Mutation	2–8%
Epigenetic and chromatin remodelling	<i>ARID1A</i>	Mutation or deletion	4–7%
	<i>ARID2</i>	Mutation	3–18%
	<i>KMT2A (MLL1), KMT2B (MLL4), KMT2C (MLL3) and KMT2D (MLL2)</i>	Mutation	2–6%
AKT–mTOR–MAPK signalling	<i>RPS6KA3</i>	Mutation	2–9%
	<i>TSC1</i> and <i>TSC2</i>	Mutation or deletion	3–8%
	<i>PTEN</i>	Mutation or deletion	1–3%
	<i>FGF3, FGF4</i> and <i>FGF19</i>	Amplification	4–6%
	<i>PI3KCA</i>	Mutation	0–2%
Angiogenesis	<i>VEGFA</i>	Amplification	3–7%

Figure 2 : Mutations communément retrouvées dans le CHC.

Données provenant de Llovet et al, Nature Reviews. Disease Primers, 2016

exome sequencing (WES). Ce mécanisme n'a pas encore été caractérisé mais est retrouvé dans 5% des CHC. Des fusions de gènes, plaçant *TERT* sous le contrôle d'un promoteur fort ont également été identifiées, notamment avec les gènes *SLC12A7* et *SLC7A2* dans 3,6% des CHC.²⁹

La réactivation de la télomérase est connue pour son rôle oncogène permettant la prolifération sans limite des cellules cancéreuses. Une étude a montré toutefois que l'immortalisation des cellules avait lieu en deux phases. La réactivation de la télomérase par mutation du promoteur dans les premiers événements de cancérogenèse n'entraîne pas un arrêt du raccourcissement des télomères longs mais stabilisent seulement les télomères extrêmement courts afin d'éviter la sénescence. Au fur et à mesure des divisions, l'ensemble des télomères raccourcissent jusqu'à entraîner une instabilité génomique, insuffisante pour induire la sénescence comme dans des cellules sans télomérase, mais suffisante pour permettre des altérations génétiques favorisant la cancérogenèse. La réactivation de la télomérase agit ainsi doublement par immortalisation des cellules et instabilité génomique.³⁰

Le second moteur de la cancérogenèse est l'altération du contrôle du cycle cellulaire, notamment via un de ses gènes clés, le plus muté retrouvé dans le CHC ainsi que dans la plupart des cancers : *TP53*. Le gardien du génome est inactivé dans la moitié des tumeurs hépatiques, essentiellement par mutation. Entre 18 et 50% des tumeurs sont mutées suivant les causes du cancer, l'aflatoxine B1 et l'hépatite B favorisant grandement les mutations de *TP53*, notamment la réversion G:C vers T:A au codon 249.^{24,26} D'autres mécanismes d'inhibition de p53 existent dans les tumeurs hépatiques dans lesquelles *TP53* n'est pas muté telle que l'augmentation du nombre de copies du gène *MDM4*, un inhibiteur de p53^{31,32} ou des mutations de *IRF2*, un régulateur de p53 dont les mutations sont mutuellement exclusives à celle de *TP53*.³³ La signalisation peut également être altérée en aval de p53. *CDKN2A* et *RB1*, deux protéines clés contrôlant la progression de la phase G1 à la phase S du cycle cellulaire subissent une délétion homozygote dans 2 à 12% (*CDKN2A*) et 3 à 8% (*RB1*) des tumeurs. Le promoteur de *CDKN2A* est de plus fréquemment hyperméthylé dans le CHC.³⁴ Certaines cyclines sont également fréquemment altérées : L'infection par le VHB est liée à une inactivation de *CCNE1* (5%) et le locus de *CCND1/FGF19* est fréquemment amplifié (5-14%) favorisant ainsi la progression du cycle G1 vers S.^{10,35,36}

L'autre voie de signalisation fortement altérée dans certains sous-types de CHC est la voie Wnt. L'effecteur principal de cette voie est la β -caténine, codée par le gène *CTNNB1*. Cette protéine assure deux fonctions distinctes : un rôle dans la structure des jonctions adhérentes où elle sert de pont entre l'E-cadhérine et l' α -caténine et un second rôle comme facteur de transcription lorsqu'elle est libre dans le cytoplasme et sous l'action de la voie Wnt. L'activation de cette dernière par fixation de Wnt sur Frizzled va déstabiliser le complexe APC/AXIN1/GSK-3 β qui phosphoryle la β -caténine libre et entraîne sa dégradation par le protéasome. La β -caténine va alors pouvoir relocaliser au noyau où elle va agir comme facteur de transcription pour activer les gènes cibles de la voie Wnt (figure 3).³⁷ La dérégulation de la voie Wnt a été établie comme moteur de la cancérogenèse dans les hépatoblastomes et bien qu'elle ne soit pas suffisante pour induire seule un CHC, il est admis que cette dérégulation est un événement important de la cancérogenèse de certains CHC qui va coopérer avec d'autres altérations et dérégulations pour promouvoir le cancer. Les mutations de *CTNNB1* sont, avec *TP53* et le promoteur de *TERT*, les plus fréquentes dans les CHC. *CTNNB1* est muté dans 11 à 37% d'entre eux.³⁸ Ces mutations sont activatrices de la β -caténine en empêchant sa dégradation par le protéasome via le masquage des sites de phosphorylation normalement ciblés par GSK-3 β (figure 3).³⁸ En plus de *CTNNB1*, de nombreux gènes régulateurs de la voie Wnt sont mutés dans le CHC telle que *AXIN1* (5-15%) et *APC* (1-2%).³⁹ Des études récentes ont de plus montré que la β -caténine exerce des actions pro-tumorales indépendamment de la signalisation Wnt. La β -caténine ne migre fortement au noyau que dans les stades tardifs du CHC en association avec le processus métastatique, la β -caténine restant essentiellement localisée à la membrane dans les stades précoces.³⁷ Elle semble être un élément important de la signalisation EGFR interagissant avec les jonctions adhérentes (et notamment l'E-cadhérine) pour favoriser la survie des cellules tumorales.

On trouve ensuite, parmi les autres mutations *drivers*, des altérations de modificateurs épigénétiques, notamment des membres du complexe de remodelage SWI/SNF impliqués dans le positionnement des nucléosomes, comme *ARID1A* et *ARID2* (mutés respectivement dans 4-17% et 3-18% des CHC).^{33,40} Les membres de la famille des histones méthyltransférases MLL sont également régulièrement altérés dans le CHC avec des mutations dans 2-6% des tumeurs. *MLL4* est par ailleurs fortement sujet aux insertions du VHB (10%).^{10,27}

Les mutations activatrices de *NRF2* et inactivatrices de *KEAP1*, deux gènes impliqués dans le stress oxydatif, sont également souvent retrouvées dans les tumeurs hépatiques (5-15%).³³

NRF2

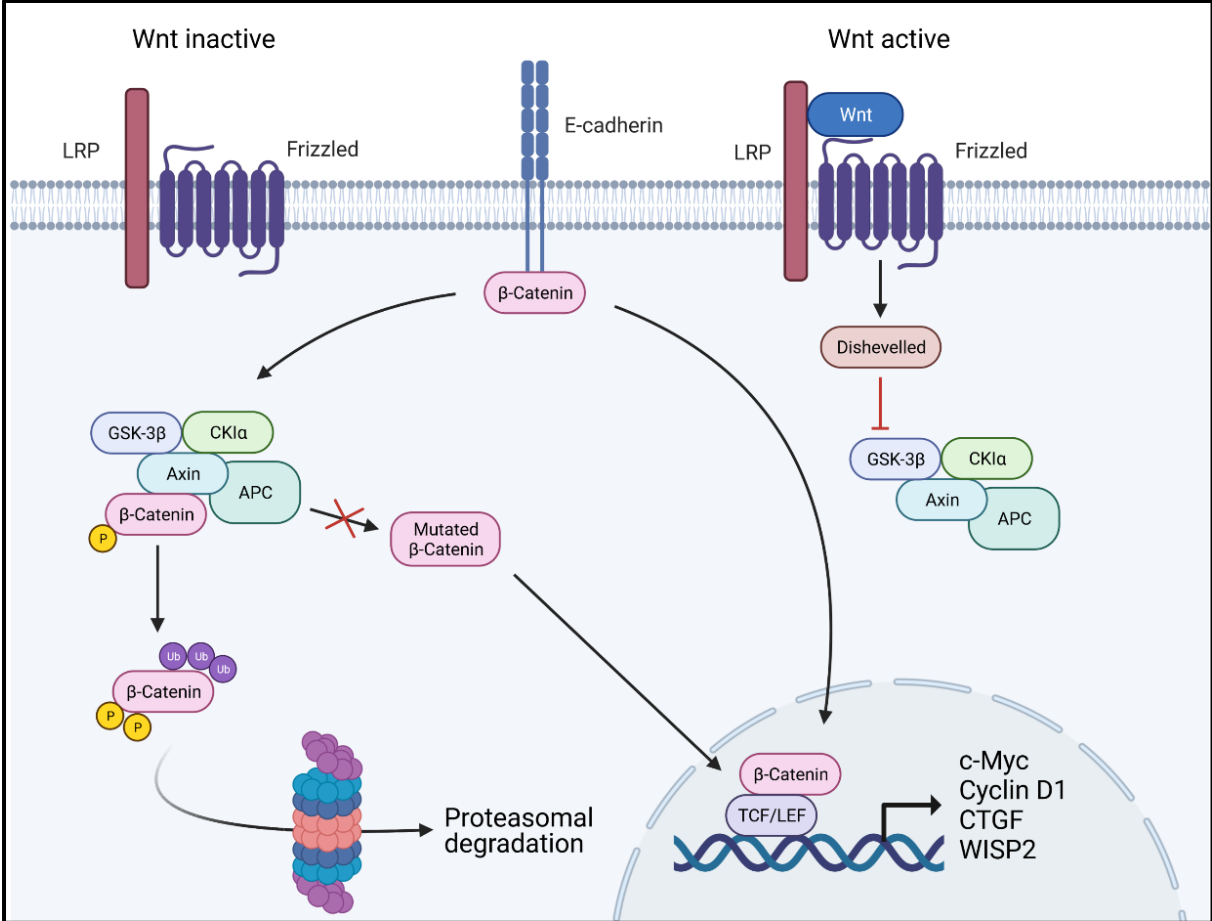


Figure 3 : Signalisation Wnt dans le CHC.

a été identifiée comme oncogène de par sa capacité à protéger les cellules cancéreuses du stress oxydatif. NRF2 est normalement dégradé par le protéasome après ubiquitinylation par le complexe KEAP1/CUL3 et les mutations empêchant sa dégradation sont donc des moteurs de la progression tumorale.¹⁰

Enfin, les voies des signalisations PI3K/AKT/mTOR et RAS/RAF/MAPK fortement impliquées dans le cancer, peuvent également subir de nombreuses mutations au cours de la cancérogenèse. La voie MAPK est activée dans 5-10% des CHC par amplification de *FGF19/CCDN1* tandis que la voie PI3K/AKT est activée par mutations activatrices de *PIK3CA*, inactivation de *TSC1* et *TSC2* ou délétion de *PTEN*, l'inhibiteur de la voie PI3K.

En plus de toutes ces altérations prenant place dans les cellules cancéreuses, vient s'ajouter l'impact du microenvironnement sur l'histoire naturelle du CHC. Le microenvironnement joue un rôle important dans l'apparition du cancer comme moteur de la mutagenèse, mais il est également fortement impliqué dans la progression tumorale. Il va notamment favoriser la prolifération et la survie des cellules cancéreuses mais aussi faciliter l'invasion tumorale et les métastases par la sécrétion de cytokines, chimiokines, facteurs de croissance et facteurs angiogéniques. De plus, la fibrose sous-jacente, caractéristique du microenvironnement des CHC, va également participer à la dérégulation de voies de signalisation dans les cellules cancéreuses par piégeage et relargage de cytokines comme le TGF β et le VEGF et activation de signalisation via des protéines de la matrice. Différents types de collagènes sont par exemple capable d'induire des réponses cellulaires via leurs interactions avec les intégrines ou avec le *Discoidin domain-containing receptor 2* (DDR2).⁴¹ Ces effets protumoraux du microenvironnement sont contrôlés par des types cellulaires spécifiques tels que les *cancer-associated fibroblast* (CAF), les *tumor-associated macrophages* (TAM) et les cellules endothéliales. Ces différentes populations vont également coopérer pour promouvoir la progression tumorale. Dans le foie, les CAF correspondent généralement aux cellules étoilées activées du foie ou *hepatic stellate cells* (HSC). Ces cellules, quiescente dans le foie sain où elles servent au stockage de la vitamine A, s'activent lors de dommages hépatiques sous l'action de cytokines inflammatoires. Le TGF β est particulièrement impliqué dans l'activation de ces cellules et leur différenciation en myofibroblastes. Ces cellules sont les moteurs de la fibrose

lors de maladie chronique du foie mais participent également fortement au microenvironnement tumoral. Elles vont sécréter des facteurs proangiogéniques comme le VEGF et l'angiopoïétine pour promouvoir la formation de néovasculature par les cellules endothéliales pour alimenter la tumeur en nutriments. Elles vont remodeler la matrice extracellulaire et agir sur la tumeur par la sécrétion de cytokines comme le TGF β , le HGF, l'EGF et le FGF. La tumeur va également participer au maintien de l'activation de ces cellules par une communication croisée via la sécrétion de TGF β . Les cellules étoilées vont aussi agir avec les TAM pour supprimer la réponse immunitaire antitumorale en promouvant un microenvironnement immunosuppresseur via le recrutement de lymphocytes T régulateurs, la sécrétion de PD-L1 qui inactive les lymphocytes T et la sécrétion de TGF β , qui favorise en surplus l'invasion et la métastase de certaines cellules cancéreuses (figure 4).^{42,43}

L'histoire naturelle est différente pour chaque tumeur, d'autant plus dans le CHC qui est extrêmement hétérogène. Les CHC diffèrent entre eux par des variations génétiques (mutations, recombinaisons) mais aussi épigénétiques, entraînant des différences importantes dans les signalisations activées au sein de cellules cancéreuses et de leur phénotype. Le microenvironnement tumoral varie également fortement d'une tumeur à l'autre avec des différences de composition de la matrice extracellulaire, d'angiogenèse et de cellules stromales infiltrantes (notamment immunitaires). La présence de cellules souches cancéreuses est également une variable permettant de différencier les tumeurs.⁴⁴ Tous ces paramètres rendent chaque tumeur unique et ces différences peuvent même se retrouver au sein d'un même individu voir d'une même tumeur. Toutefois, malgré leurs différences, la récurrence des événements clés de la cancérogenèse et les signalisations et phénotypes cellulaires associées permettent de regrouper ces tumeurs en différentes catégories.

4) Classification des tumeurs hépatiques

En clinique, les tumeurs hépatiques sont classiquement diagnostiquées par imagerie. Ce diagnostic est généralement confirmé par les anatomo-pathologistes sur biopsie ou après résection de la tumeur si celle-ci est opérable. Une fois la tumeur diagnostiquée, il lui est attribué un stade selon une des classifications existantes. La classification des patients selon le stade tumoral permet de les regrouper en différents sous-groupes associés au pronostic et ainsi

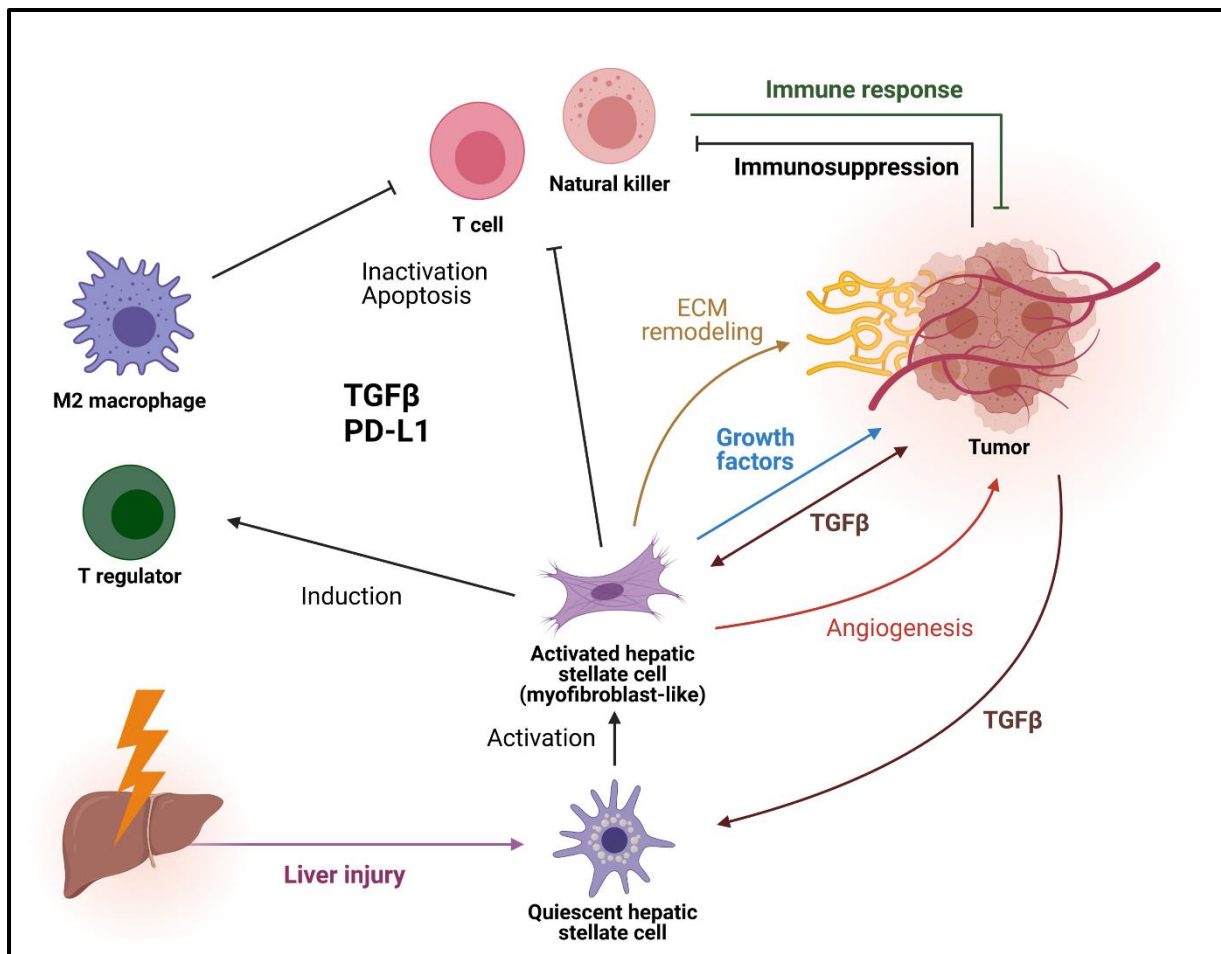


Figure 4 : Impact du microenvironnement dans le carcinome hépatocellulaire.

Légende : ECM : Matrice extracellulaire ; PD-L1 : Programmed cell death ligand 1 ; TGFβ : Transforming Growth Factor beta

d'adapter les traitements. La classification histologique repose sur des variables cliniques associées au pronostic et reflétant l'avancement du cancer telle que la taille et le nombre de tumeurs, l'invasion vasculaire et la présence de métastases distantes, ainsi que des paramètres estimant les dommages hépatiques (synthèse protéiques, détoxification, etc).⁴⁵

De nombreux systèmes de classification existent concernant le CHC (Okuda stage, French score, CLIP score, etc) et parmi eux, deux systèmes incluent des recommandations de traitement associé aux stades tumoraux. Le premier de ces systèmes est *le Barcelona Clinic Liver Cancer* (BCLC), mis au point en 1999 qui classe les patients en 4 stades : Early stage (A), Intermediate stage (B), Advanced Stage (C) et End-Stage (D).⁴⁶ Il a depuis été révisé à plusieurs reprises (ajout d'un very early stage (0) pour les patients asymptomatique avec une tumeur inférieur à 2cm en 2003, incorporation du sorafenib en 1^{ère} ligne de traitement en 2008). Ce système a été intégré aux recommandations de traitement des patients de l'*American Association for Study of Liver Diseases* (AASLD) et de l'*European Association for the Study of the Liver* (EASL) et est donc ainsi extrêmement utilisé en occident. Le second est le *Hong Kong Liver Cancer* (HKLC), développé plus récemment en 2014 par Yau et al. et qui a montré une meilleure capacité de pronostic que le BCLC lors de son essai⁴⁷ mais semble moins performant sur les cohortes occidentales à l'étiologie différente des cohortes asiatiques⁴⁸. Il n'y a pour le moment pas de consensus établi sur le système de classification optimal des tumeurs hépatiques et cela est essentiellement dû à la forte hétérogénéité du CHC, aux maladies sous-jacentes du foie et la variabilité des modalités de traitement dans le monde.⁴⁹

En parallèle de ces systèmes de classifications se sont développées de nombreuses classifications moléculaires des tumeurs hépatiques au cours des dix dernières années. Le problème rencontré classiquement dans les classifications histologiques est que les tumeurs classées sous un même stade tumoral peuvent être extrêmement variables d'un point de vue moléculaire et les recommandations de traitement ne sont alors pas optimisées pour chacune de ces tumeurs. Le BCLC B regroupant les CHC intermédiaires inclut de nombreuses tumeurs hétérogènes mais ne recommande que la chimioembolisation par voie artérielle (TACE) comme thérapie quand bien même certains patients peuvent bénéficier de traitements autres que la TACE.⁵⁰ Les classifications moléculaires visent ainsi à fournir un portrait détaillé des altérations moléculaires conduisant à chaque sous-type de CHC et ainsi à définir des thérapies

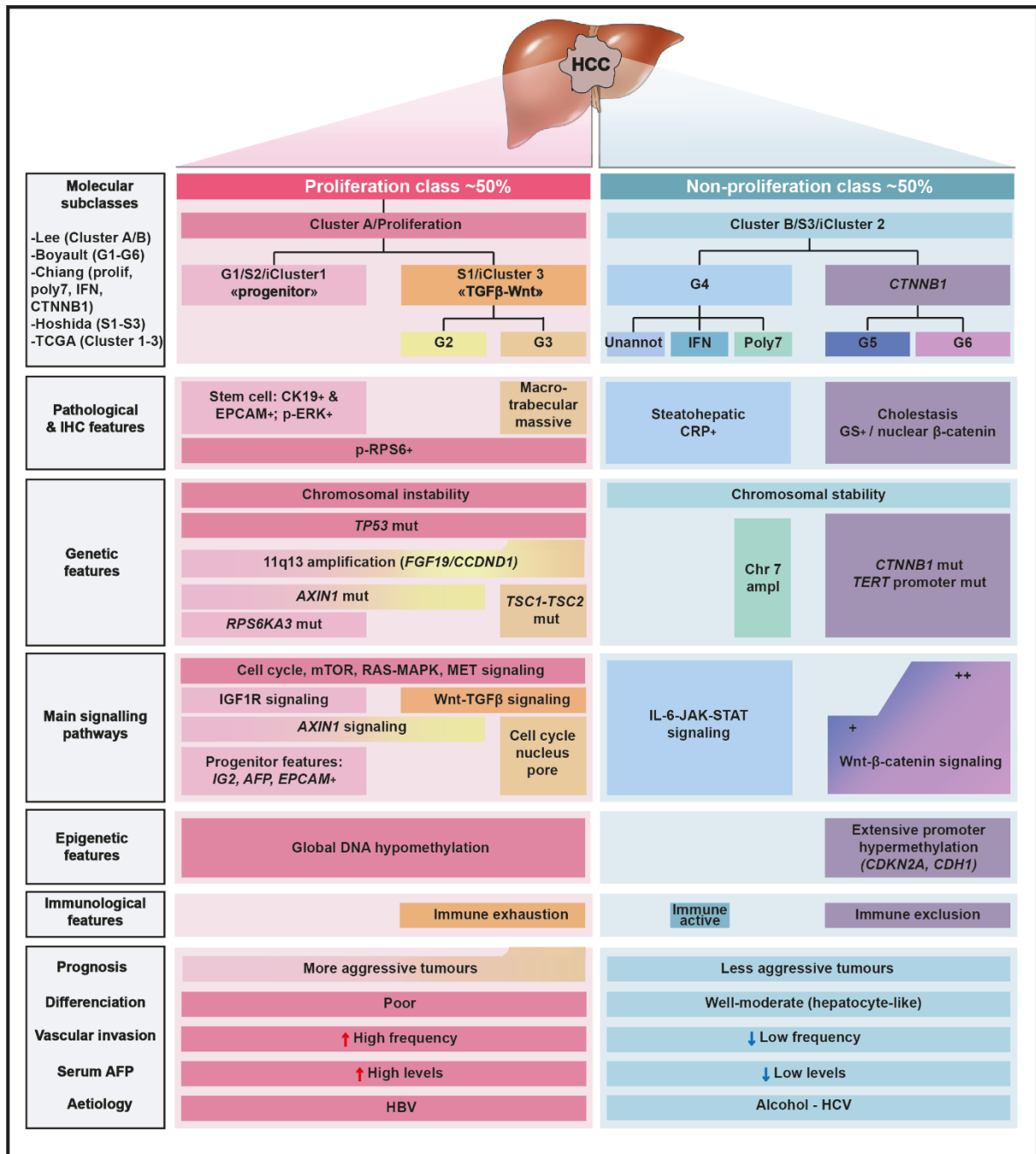


Figure 5 : Classification intégrative et pathologique des carcinomes hépatocellulaires.

Issue de Rebouissou, S., and Nault, JC., J Hepatol, 2020

ciblées ou à identifier des biomarqueurs spécifiques de chacun de ces sous-types (Figure 5). L'essor des technologies de séquençage de nouvelle génération (NGS) a permis d'obtenir rapidement et à relativement faible coût de précieuses informations sur les tumeurs hépatiques telles que le profil mutationnel, le transcriptome et le méthylome. Ces informations permettent non seulement d'identifier des marqueurs pronostics mais également d'améliorer la compréhension de la pathogenèse du CHC.¹⁰

La première classification moléculaire des CHC a été établie par Lee et al, en 2004.⁵¹ Elle regroupe les CHC en deux clusters, associés à une bonne (cluster B) et une mauvaise survie (cluster A). Les tumeurs du cluster A sont associées à une forte prolifération et une expression plus élevée de nombreux gènes anti-apoptotiques. Ces tumeurs expriment aussi fortement des gènes impliqués dans les modifications post-traductionnelles d'histones et dans l'ubiquitination. Le cluster B à l'inverse est associé aux tumeurs différenciées exprimant davantage les gènes impliqués dans les fonctions hépatiques.

Lee et al ont également identifié une signature génique de tumeurs proche de celle des hépatoblastes fœtaux.⁵² Cette signature regroupe des marqueurs des *hepatic oval cell* (des cellules souches bipotentes du foie) qui suggère que les tumeurs présentant cette signature émergent de cellules progénitrices. Cette signature « hépatoblastes » (en opposition avec la signature « hépatocytes ») s'avère de plus être de mauvais pronostic.

Parmi les classifications qui suivirent celle de Lee, celle établie par Boyault en 2007 est incontournable (Figure 5). Elle classe les CHC en 6 sous-types, de G1 à G6 à partir d'analyses transcriptomiques non supervisées de 57 CHC. Les classes G1 à G3 sont associées à une dérégulation du cycle cellulaire et une instabilité chromosomique. G1 regroupe des tumeurs qui présentent peu de copies de VHB et surexpriment des gènes typiques du développement hépatique contrôlés par l'empreinte parentale. G2 à l'inverse est enrichi en tumeurs possédant un grand nombre de copies de VHB et des mutations de *PIK3CA* et *TP53*. Ces deux sous-groupes sont également associés à une activation de la voie PI3K/AKT. G3 est enrichi en CHC présentant des mutations de *TP53*, ainsi que *CDKN2A* méthylé et surexprimant des gènes impliqués dans le cycle cellulaire. G4 est un groupe hétérogène incluant les tumeurs avec mutations de *TCF1*. G5 et G6 sont associés à l'activation de la voie Wnt, généralement par

mutation de *CTNNB1*. G5 est associée aux tumeurs immunoactives quand G6 est caractérisé par des nodules satellites et une sous-expression de l'E-cadhérine.⁵³

Une classification des tumeurs basée sur la signalisation TGF β a été élaborée par notre équipe en 2008. Cette classification sépare les tumeurs répondant au TGF β de celles n'y répondant pas. Elle discrimine deux clusters parmi les tumeurs répondant au TGF β , l'un associé à la signature précoce du TGF β et correspondant à son action cytotatique et l'autre à la signature tardive et reflétant l'action pro-métastatique du TGF β .⁵⁴

Une autre des classifications moléculaires majeures du CHC basée sur l'expression génique a été établie par Hoshida et al. en 2009 en intégrant les profils d'expression de 8 cohortes de patients. Cette analyse a permis d'identifier 3 sous-groupes de CHC, nommé S1, S2 et S3 (Figure 5). La classe S1 est caractérisée par une activation de la voie Wnt et TGF β (signature pro-métastatique établie par notre équipe)⁵⁴ et la classe S2 par l'activation de MYC et AKT. Ces deux sous-groupes sont associés à la prolifération, des tumeurs peu différenciées et plus grandes et une survie plus faible. La classe S3 est quant à elle associée aux mutations de *CTNNB1* et à des tumeurs plus petites et plus différenciées avec des cellules cancéreuses conservant le phénotype hépatocytaire et une meilleure survie.⁵⁵

Les autres classifications moléculaires importantes incluent de manière non exhaustive les classifications de Chiang, et du TCGA. La classification de Chiang se base sur l'analyse de 91 CHC associés au VHC. Elle identifie 5 catégories de CHC : *CTNNB1* muté, prolifératif (avec phosphorylation de IGF1R et RPS6), interféron-associé, polysomie du chromosome 7 et surexpression de VEGFA. Enfin la classification proposée par le TCGA est basée sur une analyse multiplateforme incluant les données de séquençage de 363 patients (nombre de copies et mutation somatique) ainsi que les données d'expression génique (ARN et protéine), d'expression de miRNA et de méthylation de 196 patients. Le clustering non supervisé de ces données a permis d'identifier trois sous-types majeurs. Le cluster 1 est associé aux femmes, d'ethnicité asiatique et généralement jeune, sans surpoids. Ce sont des tumeurs de haut grade, avec invasion macrovasculaire, des cellules peu différenciées et peu de mutations de *CTNNB1*, de *CDKN2A* et du promoteur de *TERT*. A l'inverse, les clusters 2 et 3 sont associés aux mutations de *CTNNB1*, de *HNFI1A* et du promoteur de *TERT*, ainsi qu'à la répression de *CDKN2A* par hyperméthylation. Le cluster 2 comprend les tumeurs de bas grade avec peu

d'invasion vasculaire tandis que le cluster 3 est associé à une forte instabilité chromosomique avec notamment la perte du 17p, à une forte fréquence des mutations de *CTNNB1* et une hypométhylation globale des îlots CpG.³¹

L'ensemble de ces classifications élaborées au cours de la dernière décennie a permis de mettre en évidence deux grands sous-types de CHC : les CHC dits « prolifératifs » et « non prolifératifs », ces deux grandes classes englobant les sous-types provenant des différentes classifications moléculaires (Figure 5). La classe proliférative inclue ainsi le sous-groupe G1 de Boyault, S2 d'Hoshida et le cluster 1 du TCGA dans un sous-groupe dit progéniteur (expression de marqueur du foie fœtale et activation de AKT et IGF1R) et les sous-groupes S1 (Hoshida), G2 et G3 (Boyault) et le cluster 3 (TCGA) dans un second sous-groupe « TGFβ-Wnt », caractérisé, comme son nom l'indique, par la coactivation des signalisations TGFβ et Wnt, mais aussi par une réponse immunitaire atténuée et regroupant les tumeurs les plus agressives. La classe non proliférative quant elle, est plus hétérogène et regroupe des tumeurs moins agressives avec un plus grande stabilité chromosomique en deux sous-groupes : les *CTNNB1* muté (G5-G6) et le G4, groupe hétérogène incluant les tumeurs avec polysomie du chromosome 7 et celles surexprimant les gènes des interférons, associées à une forte réponse immunitaire.⁵⁶

Ces classifications moléculaires, malgré leur capacité à discriminer les patients bien plus efficacement que la classification histologique, n'ont pas encore été intégrées dans les protocoles cliniques. Cela est essentiellement dû à l'absence de thérapie efficace pour cibler spécifiquement chacun de ces sous-types, les options thérapeutiques étant extrêmement limitées dans le CHC.

5) Traitement

Le type de traitement des CHC dépend essentiellement du stade de la tumeur et des moyens dont dispose les médecins. Les approches chirurgicales sont de loin les plus efficaces (pour une approche à visée curative) avec la transplantation hépatique au sommet et la résection dans une moindre mesure. Ces approches ne sont toutefois accessibles qu'aux patients présentant des stades précoces de la maladie (Figure 6). L'accès à la transplantation est limité de par le nombre

de greffons inférieur aux nombres de demandeurs, obligeant à sélectionner les patients qui ont le plus de chance de survie pour la transplantation. La résection, bien que complexe dans les cas de tumeurs multinodulaires ou infiltrantes n'a pas ce souci de la disponibilité des donneurs. Elle est néanmoins associée, malgré son potentiel curatif, à un risque de rechute beaucoup plus important que la transplantation (70% de rechute parmi les réséqués à 5 ans).⁵⁷ De plus, l'accès à la résection est limité par l'état du foie. Les patients présentant un foie cirrhotique sous-jacent à la tumeur et/ou complètement envahi par le cancer ne survivent généralement pas à l'opération. Les patients présentant des stades de tumeurs plus avancés se voient donc refuser la chirurgie au profit de la radiothérapie, de l'embolisation transartérielle et de la chimiothérapie dans les stades les plus avancés ou dans les pays aux ressources plus limitées.³

La transplantation hépatique est une allogreffe permettant le remplacement d'un foie malade (généralement cirrhotique ou atteint de CHC) par un foie provenant d'un donneur sain. Le patient subit généralement une hépatectomie complète mais il existe de plus en plus de cas où seul le lobe atteint par la tumeur est réséqué, ce qui permet d'accélérer la récupération des patients. La majorité des greffons sont des foies entiers provenant de donneurs décédés mais des approches de transplantations à partir de donneur vivant ont émergé à partir des années 1990.⁵⁸ Cette approche a l'avantage de limiter les décès sur liste d'attente, en permettant des greffes beaucoup plus rapides, bien qu'elle soit plus compliquée techniquement et porteuse également de problèmes éthiques (mise en danger du donneur).⁵⁹ Toutefois, que ce soit venant de patients décédés ou vivants, le nombre de transplants disponible est toujours limité et impose ainsi un certain nombre de critères aux patients pour être éligibles à la transplantation. Cela permet de maximiser la survie à long terme en fournissant des transplants aux patients qui en bénéficieront le plus.

L'accès à la transplantation hépatique pour les patients atteints de CHC dans les pays développés est ainsi défini par les critères de Milan depuis la fin des années 1990. Les patients éligibles doivent n'avoir qu'une seule tumeur inférieure à 5cm ou trois tumeurs chacune inférieure à 3cm.⁶⁰ A partir de 2002 a également émergé un modèle permettant l'allocation des organes pour la transplantation hépatique : le MELD (*Model for End-stage Liver Disease*). Ce modèle, initialement créé pour prédire la survie des patients ayant des complications d'hypertension portale, est basé sur la mesure de trois paramètres : les niveaux sanguins de

créatinine, de bilirubine et de prothrombine. L'adoption de ce modèle pour la sélection des patients atteints de CHC éligibles à la transplantation et inclus dans les critères de Milan a permis d'augmenter la survie à 5 ans des patients transplantés au-delà de 80% et a multiplié par cinq les patients atteints de CHC bénéficiant d'une transplantation hépatique.^{58,61} Afin de permettre à de nombreux patients de pouvoir être éligibles à la transplantation, ces derniers peuvent subir des traitements localisés afin de réduire la taille tumorale. La réponse aux traitements locaux est de plus de bon pronostic pour la survie post-transplantation.⁶²

Des critiques se sont toutefois élevées contre les critères de Milan, les trouvant trop restrictifs et empêchant l'accès à la transplantation à des patients qui pourraient tout de même en profiter. Plusieurs propositions visant à accroître modérément la taille limite de la tumeur au-delà des critères de Milan ont ainsi émergé, avec notamment les critères de l'Université de Californie (UCSF)⁶³ et les « *up-to-seven* »⁶⁴ qui ont tous les deux montré une survie globale à cinq ans légèrement inférieure aux critères de Milan.

L'autre approche thérapeutique potentiellement curative est la résection chirurgicale, à savoir l'excision de la tumeur et d'une partie du tissu l'environnant. Cette option est recommandée en présence d'une seule tumeur, d'un score Pugh-Child A avec un niveau normal de bilirubine dans le sérum et l'absence d'hypertension portale significative. Les patients ne doivent également pas présenter d'invasion des veines portales ou hépatiques ainsi qu'aucune maladie extrahépatique. Ces critères ont, comme pour la transplantation, été critiqués car ils n'ont pour vocation qu'à sélectionner les patients « idéaux » pour la résection hépatique quand bien même plusieurs études ont montré qu'étendre ces critères aux patients avec plusieurs tumeurs, de l'hypertension portale ou un score Pugh-Child B n'aurait pas d'effets néfastes sur le pronostic de ces patients et permettrait une meilleure survie pour certains d'entre eux comparée aux traitements classiques.⁶⁵ La résection reste toutefois limitée pour les patients atteints de cirrhose et souffrant d'insuffisance hépatique trop importante pour être opérables. De plus, les patients bénéficiant de résection hépatique restent sujets à d'importants risques de récurrences (50-70% à 5 ans) entraînant la mort chez 30-50% des patients à 5 ans. Les rechutes précoces (inférieures à 2 ans) sont essentiellement dû à l'invasion locale des cellules cancéreuses et aux métastases intrahépatiques qui restent présentes dans le foie après résection tandis que les récurrences plus tardives sont généralement dues à la formation de tumeurs *de novo* avec en cause la cirrhose

sous-jacente.⁶⁶ De manière à améliorer la prédiction de la survie ou de la récurrence des patients, l'introduction de biomarqueurs tumoraux telle que le niveau d'AFP, le niveau de différenciation tumoral et la présence d'invasion microvasculaire ont été proposés. Une signature de 5 gènes (*HNF1*, *RAN*, *RAMP3*, *KRT19* et *TAF9*), validée dans plusieurs cohortes, et permettant de prédire la survie des patients après résection a également été identifiée.⁶⁷ L'utilisation de ces biomarqueurs dans le suivi des patients permet le traitement plus efficace des rechutes.

Chez les patients atteints de CHC à un stade précoce mais non éligibles à la chirurgie, cette dernière peut être substituée par l'ablation de la tumeur par radiofréquence ou par l'ablation percutanée via de l'éthanol absolu.⁶⁸ Ces approches visent à détruire directement la tumeur en insérant à l'intérieur une sonde produisant de la chaleur ou en y injectant de l'éthanol. Elle est recommandée pour le traitement de petites tumeurs inférieures à 2 cm et montre une survie à 5 ans proche de celle obtenue par résection bien que le risque de rechute soit plus élevé.⁶⁹

Pour les patients avec un CHC de stade plus avancés, la chirurgie n'est plus possible et la classification de Barcelone (BCLC) recommande alors les traitements par chimioembolisation et chimiothérapie (Figure 6). La chimioembolisation est pratiquée essentiellement par voie artérielle, chez les patients atteints de stades intermédiaires de CHC. Cette technique consiste à délivrer des traitements anticancéreux encapsulés tels que la doxorubicine, la cisplatine ou la mitomycine C directement au niveau de la tumeur en les couplant à l'obturation des artères irriguant la tumeur via l'encapsulation.⁷⁰ La sélection précise de l'artère à bloquer permet de limiter l'hépatotoxicité du traitement. La chimioembolisation peut alternativement être remplacée par la radioembolisation utilisant l'yttrium-90, un isotope radioactif émettant des rayonnements gamma. Cette approche n'a pas démontré d'efficacité supérieure à la chimioembolisation et est généralement proposée aux patients sous-optimaux pour la chimioembolisation.⁷¹ Ces approches ne sont toutefois pas utilisées à but curatif et les patients en bénéficiant ont une survie moyenne de 26 mois.⁶⁹

Enfin, les patients avec un stade avancé de CHC sont traités par chimiothérapie. Cette approche qui n'a pas non plus de vocation curative permet une survie moyenne de 11 mois. Le traitement standard en première ligne est le sorafénib, un inhibiteur de tyrosine kinase, ayant démontré son efficacité lors de l'essai SHARP et obtenu sa validation par la FDA en 2007⁷² puis validé également dans les cohortes asiatiques avec l'essai ASIA-PACIFIC. Il est resté le seul

traitement efficace contre le CHC avancé jusqu'à récemment avec le lenvatinib, montré comme non inférieur au sorafenib dans l'essai REFLECT et utilisé en Europe et aux USA en première ligne contre le CHC depuis 2018.⁷³ Les essais cliniques RESORCE et CELESTIAL ont également récemment montré l'efficacité du regorafenib et du cabozantinib chez les patients ne répondant pas au sorafenib et sont utilisés en seconde ligne de thérapie depuis lors.^{74,75} Enfin, le ramucirumab, un anticorps monoclonal ciblant le VEGFR2 est également utilisé depuis 2019 suite à l'essai REACH-2 en seconde ligne chez les patients avec une AFP supérieure à 400ng/mL (Figure 6).

La plupart des biothérapies ciblées contre le CHC sont donc basées sur des inhibiteurs multi-kinases et principalement de tyrosine kinase. Toutefois, malgré l'efficacité démontrée de ces thérapies, leur bénéfice reste faible pour les patients, avec un gain de quelques mois tout au plus sur la survie globale (e.g. 3 mois de survie supplémentaire avec le Sorafenib). La faible efficacité des traitements anticancéreux dans le CHC peut être expliquée par plusieurs paramètres. Tout d'abord, comme vu précédemment, les altérations génétiques responsables de l'apparition du CHC telles que les mutations de *TP53*, de *CTNNB1* ou du promoteur de *TERT*, ne peuvent pour le moment pas être ciblées par des thérapies, limitant ainsi les possibilités thérapeutiques bien que des approches pour cibler ces gènes soient à l'étude, notamment via l'utilisation de microARN (miR) ou l'oligonucléotides antisens (ASO). Une autre raison de l'échec de nombreuses thérapies dans les essais cliniques de phase III est l'hétérogénéité tumorale inter-patients. Pour résoudre ce problème il a été proposé de tester les traitements sur des populations stratifiées sur la base de biomarqueurs de réponse aux traitements. Cela nécessite toutefois l'identification de biomarqueurs fiables comme l'a montré l'essai clinique du tivantinib (un inhibiteur de MET) qui s'est soldé par un échec car la surexpression de MET ne s'est pas avérée être un bon biomarqueur de réponse au tivantinib⁷⁶.

Une autre approche pouvant s'avérer prometteuse mais qui n'a pas encore donné de résultats est l'immunothérapie et plus particulièrement l'utilisation des « *immune checkpoint inhibitors* » (ICI). Le nivolumab (anti PD-1) fut le premier de ces ICI à obtenir une autorisation de la FDA pour le traitement du CHC. Il a été suivi du pembrolizumab, testé en seconde ligne chez les patients ne répondant pas au sorafenib. Ces deux traitements ont montré des résultats intéressants dans leur essai clinique de phase 2 : l'essai CheckMate040 pour le nivolumab en

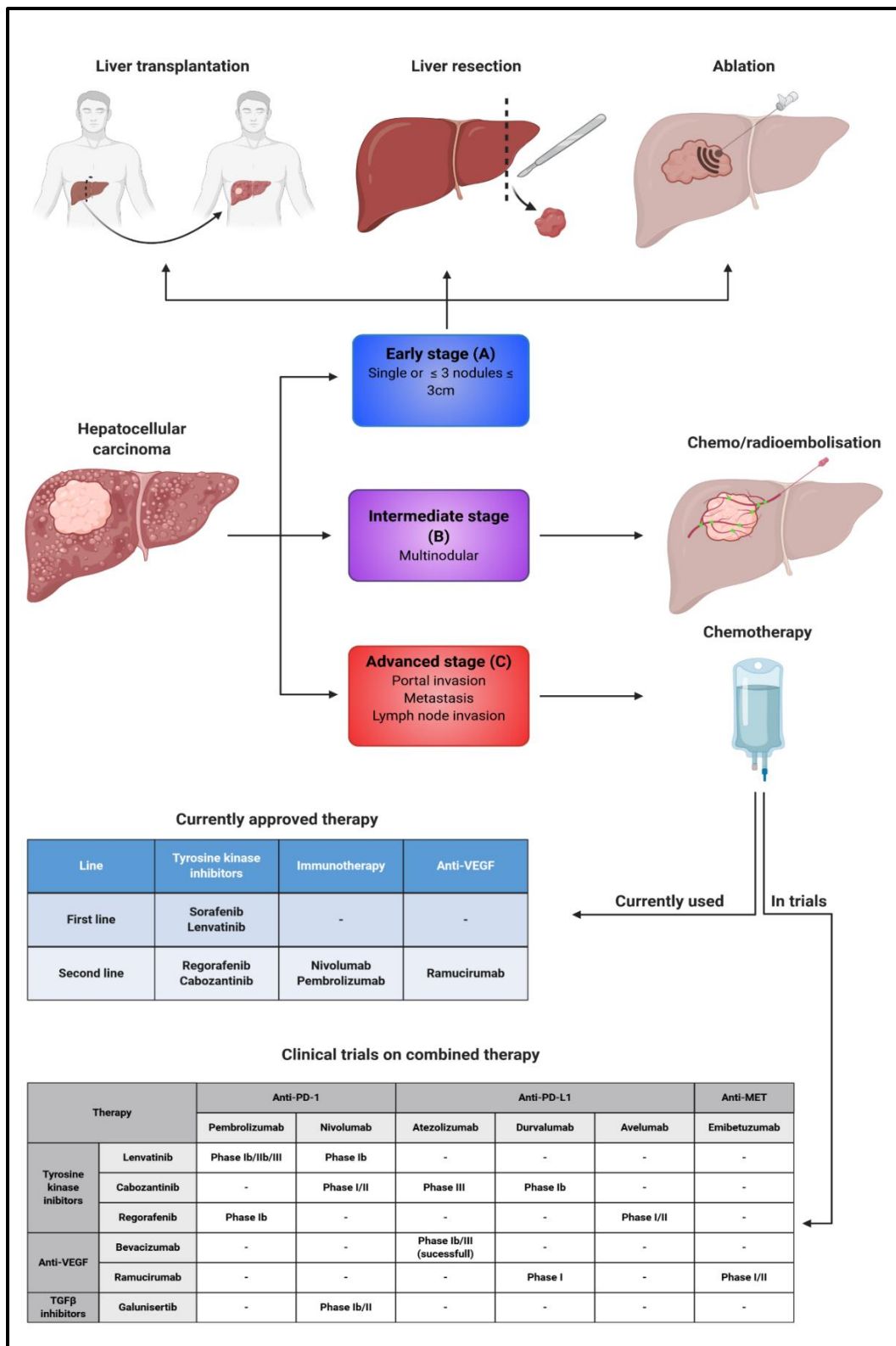


Figure 6 : Stratégie thérapeutique contre le carcinome hépatocellulaire d'après le BCLC et essais cliniques sur la thérapie combinée.

seconde ligne a montré une efficacité chez 20% des patients indépendamment de leur réponse au sorafenib et l'essai KEYNOTE-224 pour le pembrolizumab, a montré une efficacité chez 17% des patients en seconde ligne après le traitement au sorafenib.^{77,78} Toutefois, les essais cliniques de phase III n'ont pas atteint leurs objectifs prédéfinis (pas de différence significative sur la survie globale). L'immunothérapie nécessite ainsi, tout comme la chimiothérapie, l'identification de biomarqueurs pertinents pour identifier les patients les plus à même d'y répondre, l'expression de PD-L1 n'ayant pas été un biomarqueur prédictif dans les essais cliniques de phases II.⁵⁶

L'immunothérapie dans le CHC se heurte de plus au microenvironnement immunosuppresseur de la tumeur. Les mécanismes moléculaires contrôlant l'immunosuppression sont toutefois encore mal connus. Les classifications immunologiques du CHC ont trié les tumeurs en deux grandes catégories : immune ou non immune (tumeurs dites chaude vs froides) (figure 7). La première est caractérisée par la présence de cellules immunitaires (lymphocytes et macrophages) et de réponses inflammatoires. Le système immunitaire pouvant avoir un rôle anti- ou pro-tumoral (sécrétion de facteurs de croissance favorisant la prolifération et l'invasion des cellules tumorales), cette catégorie est divisée en deux sous-catégories (clusters). Le cluster « *Active immune response* » caractérise les tumeurs dans lequel une réponse antitumorale adaptative est en cours tandis que le cluster « *Exhausted immune response* » est fortement associé à l'activation de la voie TGF β (cytokine immunosuppressive) et à l'immunosuppression des lymphocytes T.⁷⁹ Cette classification a été étoffée depuis peu en classifiant les sous-groupes de tumeurs par rapport à leurs mécanismes d'immunosuppression. Cela a permis d'affiner la catégorie des tumeurs non immunes divisée en deux clusters : le cluster TAM, regroupant les tumeurs où l'immunosuppression est orchestrée par les « *tumor associated macrophages* » et le cluster CTNNB1, où l'activation constitutive de la voie Wnt a déjà été associée à l'exclusion des lymphocytes T de la tumeur et à la résistance aux *immune checkpoint inhibitors* bien que les mécanismes n'aient pas été élucidés.⁸⁰

Malgré l'utilité de ces classifications pour orienter la thérapie ciblée du CHC, de nombreux progrès restent encore à faire dans la compréhension des mécanismes d'immunosuppression et plus encore dans l'identification de biomarqueurs de sélection des patients et de réponses aux

traitements. En l'absence de ces biomarqueurs, les essais cliniques récents se sont tournés vers une autre approche : la thérapie combinée^{32,81}.

Cette approche vise à associer plusieurs traitements (généralement chimiothérapie et ICI) dans l'optique d'une part d'agrandir le faisceau de patients éligibles à la thérapie mais également d'obtenir de potentiels effets synergiques entre les traitements (Figure 6). C'est notamment le cas pour la combinaison entre des antiangiogéniques (correspondant à la quasi-totalité des chimiothérapies actuelles contre le CHC qui ciblent, entre autres, l'axe VEGF/VEGFR) et des ICI. L'idée étant d'agir dans un premier temps sur l'hypoxie et les capacités potentiellement immunosuppressives de l'axe VEGF/VEGFR afin de faciliter l'action des immunothérapies. Une quinzaine d'essais cliniques de phase 2 et 3, croisant les différents traitements de chimiothérapies et d'immunothérapies, sont ainsi en cours, sur la base des résultats précliniques de ces combinaisons.⁸² La réussite récente de l'essai IMbrave150 qui a montré la supériorité de la combinaison de l'atezolizumab (anti-PD-L1) avec le bevacizumab (anti-VEGF) sur le sorafenib montre bien l'intérêt de ces thérapies combinées.⁸³

Reste une combinaison extrêmement intéressante basée ici non pas sur l'association d'immunothérapie avec des antiangiogéniques mais avec un inhibiteur de la signalisation canonique du TGF β . Cette approche combinant le nivolumab (ICI) avec le galunisertib (inhibiteur du récepteur TGFBR1) est actuellement testée en essai clinique de phase I/II (NCT02423343). Le galunisertib a été testé en seconde ligne chez les patients non répondants au sorafenib et a montré une bonne réponse pour une sous-catégorie de patients chez qui il a entraîné une diminution de l'AFP et du TGF β 1 circulant.⁸⁴ Il est également depuis testé en combinaison avec le sorafenib. Une autre thérapie basée sur un anti-PD-L1 couplé à deux domaines extracellulaire du TGFBR2 pour piéger le TGF β est également en cours de test.⁸⁵ Cette molécule, le bintrafusp alfa ou M7824 est actuellement en essai clinique dans le cancer du poumon (NCT03631706) mais pourrait également être appliquée au CHC. Cibler la signalisation TGF β peut effectivement s'avérer extrêmement pertinent dans le CHC au regard de l'impact considérable qu'a cette cytokine dans l'apparition et la progression du CHC comme nous allons le voir par la suite.⁸¹

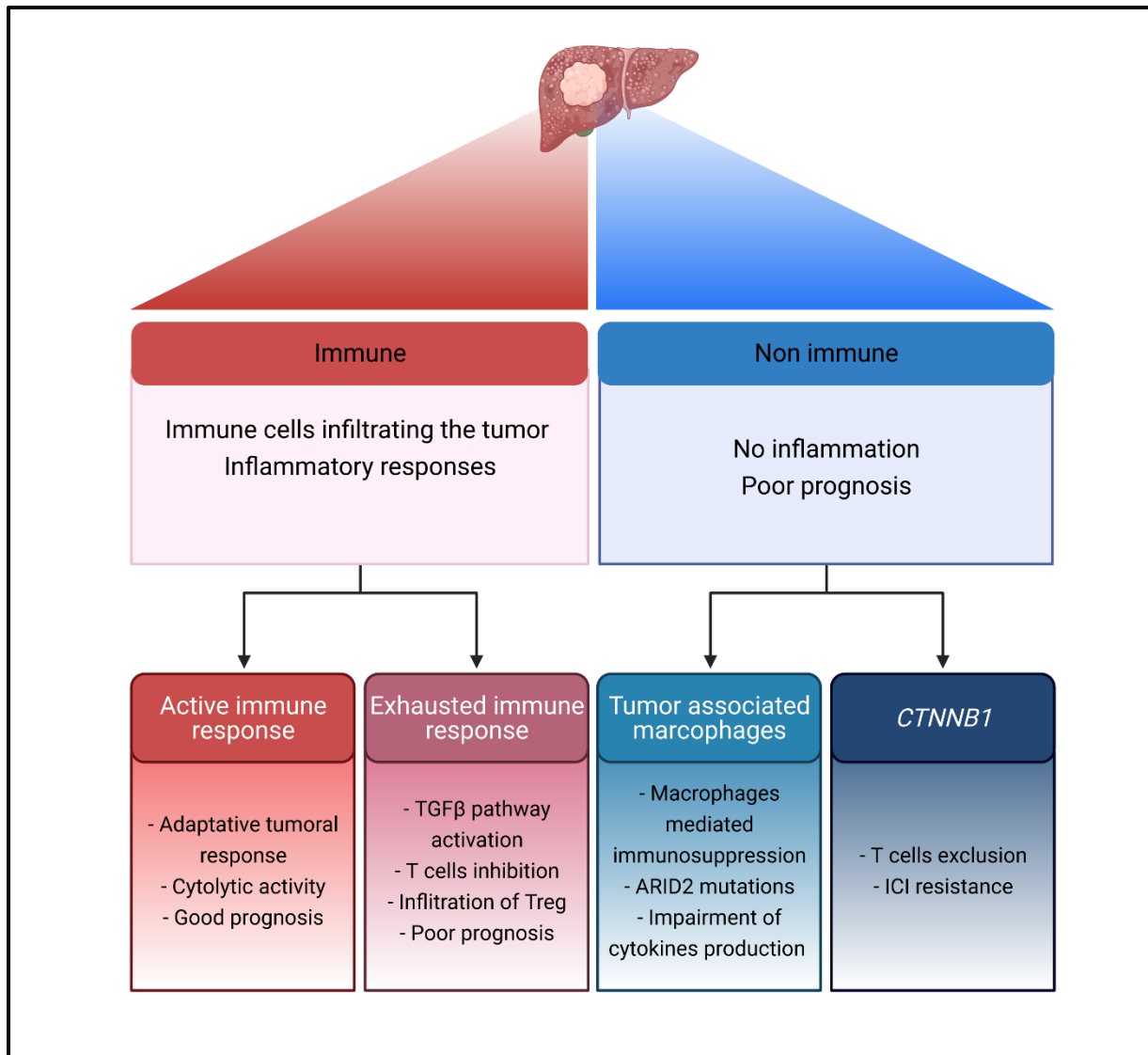


Figure 7 : Classification immunitaire des carcinomes hépatocellulaires.

II) Le Transforming Growth Factor β et la transition épithélio-mésenchymateuse

Le *Transforming Growth Factor beta* ou TGF β est une cytokine pléiotropique impliquée dans de nombreux processus biologiques. Elle est caractérisée par une action versatile extrêmement ambivalente suivant le type et le contexte cellulaire sur lesquels elle agit. Elle est ainsi capable d'inhiber la prolifération, de stimuler la pluripotence des cellules souches, mais également d'induire la différenciation, d'inhiber la croissance tumorale ou d'activer le processus métastatique. Rarement une cytokine n'a porté un nom aussi large ; le TGF β transforme les cellules effectivement mais dans toutes les voies possibles. Cette versatilité du TGF β a rendu perplexe les chercheurs depuis sa découverte dans les années 1980, et si sa signalisation canonique a été identifiée et caractérisée rapidement, les mécanismes par lesquels elle induit l'ensemble de ces effets différents et parfois contradictoires, reste encore sujet à études.

La superfamille du TGF β regroupe chez l'Homme plus de 30 membres répartis en deux sous-familles : TGF β -activin-Nodal et BMP. Tous les membres de cette superfamille agissent sur une combinaison de deux dimères de récepteurs qui vont s'associer en réponse à la fixation du ligand. Ces deux dimères sont des récepteurs de type I (TGFBR1) qui propagent le signal et des récepteurs de type II (TGFBR2) qui activent le signal. Il existe sept récepteurs de type I et cinq de type II chez l'Homme dont les combinaisons sont dépendantes du ligand. La signalisation TGF β n'utilise toutefois que deux récepteurs, le TGFBR1 et le TGFBR2.⁸⁶ Les ligands de la voie TGF β sont au nombre de trois et ainsi judicieusement nommés TGF β 1, TGF β 2 et TGF β 3.

1) Fonctionnement de la signalisation canonique

Le fonctionnement de la voie canonique du TGF β est très bien caractérisé (Figure 8). Le ligand, un dimère sous forme active, va interagir avec un dimère de TGFBR1 et un dimère de TGFBR2. Cette interaction est facilitée par des protéines membranaires dites « ligand trap ». Les récepteurs du TGF β sont des récepteurs transmembranaires de type sérine/thréonine kinase. Le

TGF β a une affinité beaucoup plus forte pour le TGFBR2 que pour le TGBR1 avec lequel il n'interagira pas si celui-ci est isolé. L'activation du récepteur se fait tout d'abord par la fixation du ligand sur le TGFBR2 puis par le recrutement du TGBR1 par ce complexe. L'interaction du TGF β avec les deux récepteurs va entraîner un changement de leur conformation permettant la phosphorylation du TGBR1 par le TGFBR2 au niveau du domaine GS, juste en amont de son domaine kinase conduisant à l'activation de ce dernier.⁸⁷

Interviennent alors les protéines de la famille *Mothers against decapentaplegic homolog* (SMAD), effecteurs intracellulaires de la signalisation TGF β . Leur noms réfèrent à leur homologies avec les *Small worm phenotype* (SMA) de *Caenorhabditis elegans* et les *Mothers against decapentaplegic* (MAD) de *Drosophila melanogaster*. Ces protéines sont au nombre de huit. SMAD1, SMAD5 et SMAD8 sont les effecteurs de la voie BMP dont nous ne discuterons pas plus tandis que SMAD2 et SMAD3 sont ceux de la signalisation TGF β . SMAD4 sert de co-activateur à l'ensemble de ces protéines et SMAD6 et SMAD7 sont quant à eux des inhibiteurs des signalisations TGF β et BMP. Ces protéines sont toutes constituées de deux domaines globulaires appelés MH1 et MH2 permettant respectivement la liaison à l'ADN et l'interaction avec d'autres facteurs de transcription. Ces deux domaines contiennent de multiples sites de phosphorylation responsables de cross-talk avec de nombreuses autres voies de signalisation dites non canoniques.⁸⁷

A l'état basal, les protéines SMAD isolées effectuent un aller-retour constant entre le cytoplasme et le noyau. L'activation du TGBR1 va permettre le recrutement et la phosphorylation de SMAD2 et SMAD3 au niveau C-terminal qui vont alors s'associer pour former un dimère via leur domaine MH2 (ce complexe est généralement appelé R-SMAD pour *receptor-regulated SMAD*). Ce recrutement au niveau du récepteur est facilité par la protéine SARA. Le complexe R-SMAD va ensuite s'associer à SMAD4 pour former un trimère considéré comme étant l'unité fonctionnelle de la signalisation intracellulaire du TGF β . Ce trimère va alors migrer au noyau par l'intermédiaire de facteurs d'imports (pouvant permettre une régulation de la signalisation TGF β par blocage de l'accès au noyau), puis s'associer à différents facteurs de transcription afin de reconnaître et réguler la transcription de gènes.^{86,87}

Afin d'éviter une suractivation de la voie, les SMAD subissent un rétrocontrôle négatif, essentiellement par l'action de SMAD7, elle-même induite par la signalisation TGF β . SMAD7

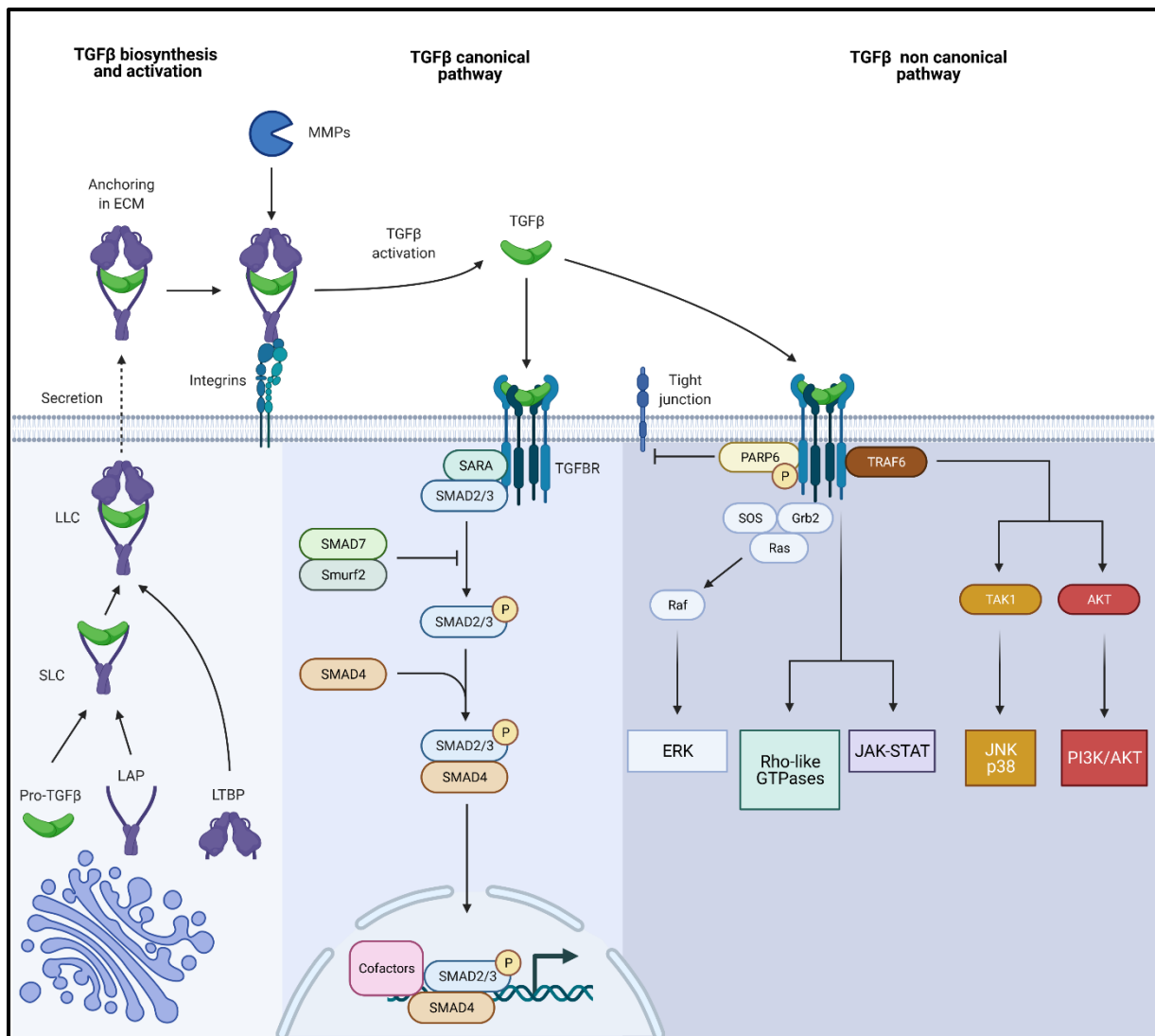


Figure 8 : Biosynthèse et signalisation du TGFβ.

Légende : LAP : Latency-associated protein ; SLC : Small latent complex ; LLC : Large latent complex ; LTBP : Latent TGFβ binding protein ; MMPs : Matrix metalloproteinase ; ECM : Matrice extracellulaire

agit par recrutement de SMURF au niveau du récepteur qui va polyubiquityler et endocyter les récepteurs et conduire à leur dégradation. Les protéines BAMBI et SKIL, elles aussi induites par le TGF β , vont également participer au rétrocontrôle en déstabilisant les complexes SMAD.

2) Signalisation non canonique

Le TGF β est de plus capable d'induire des effets SMAD-indépendants. Le TGFBR2 après fixation du ligand est capable de transduire un signal pro-EMT sans recrutement du TGFBR1 via la phosphorylation de PAR6 qui va aller perturber les jonctions serrées.⁸⁸ La plupart des actions non canoniques passent toutefois par le dialogue et l'activation d'autres voies de signalisation (Figure 8). Les voies MAPK, PI3K sont fortement sujettes à cette communication croisée avec la voie TGF β puisqu'elles sont capables dans certaines circonstances de l'inhiber.⁸⁹ L'association du TGFBR2 avec le TGFBR1 permet par exemple d'activer la voie MAPK par recrutement de Grb2/SOS qui vont activer la cascade de phosphorylation entraînant l'activation de ERK1/2.⁹⁰ De la même manière, les voies JNK et PI3K peuvent être activées par interaction des récepteurs au TGF β avec TRAF-6.^{91,92} Parmi d'autres voies non canoniques activées par le TGF β , on retrouve la voie JAK-STAT et les Rho GTPases, notamment au cours du processus de fibrose dont le TGF β est un médiateur essentiel.^{93,94} La diversité des voies de signalisation impactées par le TGF β et le fait que ces signalisations peuvent aussi bien avoir un effet synergique qu'antagoniste avec la signalisation canonique participent à la complexité de la réponse au TGF β .⁹⁵

3) Réponse cellulaire au TGF β

La capacité des SMAD à interagir avec de nombreux facteurs de transcription explique en partie la diversité des réponses cellulaires au TGF β . En effet, le contexte transcriptionnel de la cellule va être un élément décisif dans la réponse au TGF β puisqu'il va dicter les partenaires avec lesquels les SMAD vont interagir. Ces partenaires vont ainsi déterminer les gènes ciblés par les SMAD aussi bien que l'effet qu'auront les SMAD sur ces gènes cibles (activation ou répression transcriptionnelle) et la durée de la réponse transcriptionnelle.

Les SMAD possèdent une séquence consensus de liaison à l'ADN, appelé *SMAD binding element* (SBE), qui consiste en un motif CAGAC généralement répété, considérant également que la séquence complémentaire reverse permet la fixation des SMAD. Les SMAD ont toutefois une affinité relativement faible pour cette séquence, expliquant la nécessité pour ces protéines d'interagir avec d'autres facteurs de transcription pour assurer leur liaison à l'ADN. De plus, les facteurs de transcription interagissant avec les SMAD peuvent faciliter l'intégration du signal induit par le TGFβ via l'action d'autres voies de signalisation, rajoutant un autre degré de complexité dans la réponse au TGFβ.⁹⁶

Le paysage chromatinien de la cellule est également bien évidemment un paramètre important pour expliquer la diversité des réponses au TGFβ. Les protéines SMAD, bien que n'agissant pas directement sur la chromatine, sont capables d'interagir avec des protéines capables de modifier le paysage chromatinien, telles que des *histones acetyl transferases* (HATs), des *DNA methyl transferases* (comme DNMT3A) et plus généralement le complexe de repositionnement des nucléosomes SWI/SNF.^{97,98} De nombreux gènes régulés par les SMAD sont activés par des modifications épigénétiques contrôlées par ces protéines et non par une activité transcriptionnelle directe des SMAD.⁹⁹ Elles ont toutefois besoin de co-effecteurs et de conditions particulières pour opérer un remodelage complet et profond de la chromatine. Cela se traduit donc forcément par des réponses cellulaires différentes suivant le type cellulaire et le niveau de différenciation de la cellule. Cela étant, les SMAD, notamment par leur interaction avec TRIM33, sont capables d'opérer des changements dans l'état chromatinien de certains de leurs gènes cibles, tout particulièrement ceux dont la chromatine se trouve dans un état intermédiaire dit de « *poised chromatin* » (chromatine regroupant des marques histones activatrices et inhibitrices et associées à des gènes non exprimés mais rapidement activables).^{86,100}

De plus, le TGFβ est soumis à une régulation fine bien avant d'atteindre son récepteur. Certaines cellules expriment ainsi à la surface de leur membrane des protéines qui vont bloquer l'accès du TGFβ ou des BMP au récepteur.⁸⁷ En plus de la régulation que peuvent jouer les « *ligand trap* » sur sa fixation au récepteur, l'accessibilité et la disponibilité du ligand sont contrôlées par la matrice extracellulaire (Figure 8). Le TGFβ a généralement une action autocrine et paracrine, agissant sur les cellules qui le secrètent et les cellules environnantes. Il

est toutefois sécrété de manière inactive avec un propeptide nommé *latency-associated peptide* (LAP) qui nécessite clivage pour pouvoir fixer le récepteur. Ce LAP-TGF β s'associe souvent avec des protéines de liaison, les *latent TGF β binding protein* (LTBP) pour former le complexe large *latent ou large latent complex* (LLC) pouvant s'ancrer dans la matrice extracellulaire (Figure 8).¹⁰¹ Le TGF β peut être libéré de ce complexe par l'intermédiaire de métalloprotéases (notamment MMP2 et MMP9) mais essentiellement par l'interaction de certaines intégrines avec le motif RGD présent dans le LAP. La libération du TGF β a ainsi été démontrée pour les intégrines $\alpha\beta$ 3, $\alpha\beta$ 5, $\alpha\beta$ 6 et $\alpha\beta$ 8. Ceci implique que seuls certains types cellulaires exprimant spécifiquement ces intégrines sont alors capables de libérer le TGF β et de bénéficier de son effet.¹⁰²

La signalisation TGF β , bien que relativement simple en apparence, se révèle donc extrêmement complexe et finement régulée à tous les niveaux. S'ensuit donc une signalisation présente aussi bien lors du développement que pendant la vie de l'organisme, avec une myriade d'effets possibles sur les cellules et un rôle aussi bien physiologique que pathologique.

4) Rôle du TGF β

Le TGF β assure de nombreux rôles et contrôle de nombreux processus biologiques allant du développement embryonnaire à l'homéostasie et la régénération tissulaire. Il impacte également la quasi-totalité des types cellulaires de l'organisme avec une action particulièrement poussée sur les cellules immunitaires et les fibroblastes. Voici quelques exemples de ses nombreux rôles qui sont retrouvés et souvent détournés dans le processus tumoral.

a) Développement

La superfamille du TGF β joue un rôle crucial lors du développement embryonnaire comme régulateurs centraux de la morphogenèse tissulaire (BMP signifiant littéralement *Bone morphogenetic protein*).¹⁰³ La signalisation TGF β est nécessaire, en coopération avec le FGF, à la maintenance de la pluripotence des cellules de l'embryon (hESC), notamment via son interaction avec les régulateurs majeurs de la pluripotence OCT4, SOX2 et NANOG.¹⁰⁴ La perte de la signalisation TGF β induit la différenciation des hESC en neuroectoderme. Le TGF β est toutefois également un moteur de la différenciation cellulaire une fois OCT4-SOX2-NANOG perdue, notamment en ciblant la « *poised chromatin* » établie par OCT4 et en

interagissant avec de nouveaux facteurs de transcription comme FOXH1 pour diriger la différenciation de l'endoderme.¹⁰⁵ Le TGF β joue également un rôle dans le développement fœtal du foie. Un gradient de TGF β /activin est nécessaire à la bonne différenciation des hépatoblastes en hépatocytes et cellules biliaire. Une forte signalisation TGF β /activin est ainsi nécessaire au niveau de la veine porte pour permettre la différenciation en cellules biliaires tandis que la signalisation s'affaiblit dans le parenchyme pour permettre la différenciation en hépatocytes.¹⁰⁶ Chez l'adulte, le TGF β assure le maintien des cellules souches hématopoïétiques, essentiellement en limitant leur prolifération, notamment en empêchant la désactivation de FOXO par AKT.¹⁰⁷

b) Immunologie

En plus de son rôle sur les cellules souches, le TGF β est également une cytokine fortement impliquée dans la réponse inflammatoire.¹⁰⁸ Les souris déficientes en TGF β 1 présentent de graves problèmes inflammatoires et d'auto-immunité montrant l'importance de cette cytokine dans la régulation du système immunitaire, le TGF β étant effectivement une cytokine clé de l'immunosuppression. Il va notamment favoriser l'expansion des lymphocytes T régulateurs (Treg), en induisant l'expression de FOXP3 (facteur de transcription clé des Treg), qui vont eux-mêmes le sécréter pour abolir la réponse antitumorale.¹⁰⁹ Il va également inhiber les cellules présentatrices d'antigènes telle que les cellules dendritiques en réduisant l'expression du complexe majeur d'histocompatibilité de type II. Il agit de plus sur les cellules de l'immunité innée comme les macrophages qu'il va polariser vers un phénotype anti-inflammatoire (M2) et les lymphocytes T et les *natural killers* (NK) dont il va bloquer les fonctions en inhibant l'expression de IFN- γ et TBET, bloquant ainsi leur différenciation en cellules lymphoïdes de type I et leur activité cytotoxique.^{110,111} Cette capacité immunosuppressive du TGF β est bien entendu utilisée par la tumeur pour supprimer la réponse antitumorale et favoriser l'évasion immunitaire.¹⁰⁸

c) Fibrose

De par son action dans l'inflammation, le TGF β est également fortement impliqué dans le processus fibrotique. Au niveau hépatique lors d'inflammation aigue, le TGF β joue un rôle important d'inhibiteur de la prolifération hépatocytaire et de terminaison de la réponse

régénérative. Toutefois, en cas d'inflammation chronique, le TGF β devient alors un puissant inducteur de la fibrose en activant les cellules étoilées du foie et stimulant leur production de matrice extracellulaire (notamment les collagènes I et III) pour remplacer progressivement le tissu hépatique par des cicatrices fibreuses. Ce dépôt de matrice extracellulaire entraîne ainsi une dérégulation de voies de signalisation dû aux interactions intégrines/collagènes et aux cytokines piégées dans la matrice extracellulaire. Le TGF β est de plus souvent sécrété par la tumeur pour maintenir l'activation des cellules étoilées, en plus de son rôle immunosuppresseur.¹¹²

Enfin un des rôles prépondérants du TGF β est sa capacité à induire la transition épithélio-mésenchymateuse ou *epithelial-to-mesenchymal transition* (EMT) que ce soit lors du développement embryonnaire, lors du processus de réparation tissulaire ou dans les cancers. Ce processus, dont le TGF β est le principal inducteur, nécessite d'être détaillé, considérant qu'il sera au cœur du sujet de cette thèse.

5) La transition épithélio-mésenchymateuse

La transition épithélio-mésenchymateuse est un processus cellulaire par lequel des cellules vont perdre leurs propriétés épithéliales et acquérir des propriétés mésenchymateuses, réduisant ainsi leur capacité à adhérer à d'autres cellules ou à la matrice extracellulaire et favorisant leur motilité et leurs capacités migratoires et invasives. Il s'agit d'un processus dynamique et réversible qui implique de nombreuses transformations structurelles et fonctionnelles menant à un changement phénotypique de la cellule (Figure 9). Il ne s'agit cependant pas d'une différenciation cellulaire, l'EMT n'impliquant pas nécessairement un changement de lignage cellulaire. Elle est induite en réponse aux signaux que la cellule reçoit du microenvironnement et particulièrement par l'action du TGF β et contrôlée par plusieurs familles de facteurs de transcription (SNAI/ZEB/TWIST) regroupés sous l'appellation d'EMT-TF.

Le terme EMT apparaît vers la fin des années 1970 et est employé par Elizabeth Hay pour décrire les changements phénotypiques observés dans la première couche cellulaire de l'embryon de poulet. D'abord qualifié de « transformation » épithélio-mésenchymateuse, le terme a changé en « transition » pour accentuer l'aspect transitoire et réversible de ce processus.

¹¹³ L'EMT est un processus physiologique survenant essentiellement au cours du développement embryonnaire. Il devint évident pour les embryologistes à la fin des années 1990 que l'EMT jouait un rôle prépondérant dans les processus de gastrulation à l'origine du mésoderme et de la formation de la crête neurale. Des cytokines comme le TGF β et le FGF1 furent découvertes comme inducteurs de l'EMT, ce qui permit le lien entre l'EMT et le cancer. Toutefois, ce fut chez *Drosophila melanogaster* que furent identifiés les régulateurs transcriptionnels de ce mécanisme avec la découverte de facteurs de transcription clés de l'EMT : Twist et Snail. La recherche sur l'EMT commença à exploser dans les années 1990 lorsque fut identifié Snai2 capable d'induire une EMT dans les cellules cancéreuses, renforçant le lien entre EMT embryonnaire et progression tumorale. D'autres facteurs clés furent identifiés par la suite : ZEB1/2 et E47 ainsi que des miR dont ceux de la famille miR-200. Il fut également découvert que l'EMT participe aux processus de réparation tissulaire ainsi qu'à la pathologie associée (fibrose). La capacité des cellules cancéreuses à effectuer une EMT fut également liée à leurs propriétés souches et métastatiques.¹¹⁴

Le mécanisme inverse par lequel une cellule mésenchymateuse acquiert un phénotype épithélial est qualifié de transition mésenchymato-épithéliale (MET). Elle consiste en la perte des propriétés motiles et invasives de la cellule et en l'apparition de la polarisation cellulaire apico-basale via l'établissement de jonctions avec d'autres cellules. La MET est également un processus physiologique intervenant lors du développement embryonnaire (entre autres au cours du développement cardiaque, de la morphogenèse rénale et de la formation des somites). Ce mécanisme est néanmoins beaucoup moins étudié que l'EMT. La théorie couramment admise est que l'EMT et la MET sont deux versants de la plasticité cellulaire auquel une cellule peut être soumise dépendamment des stimuli qu'elle reçoit. Une balance entre ces deux processus est nécessaire pour réparer les lésions et maintenir l'intégrité tissulaire.¹¹⁵

L'EMT a d'abord été décrite comme un processus binaire qui, une fois engagé, transforme une cellule complètement épithéliale en cellule totalement mésenchymateuse. Cette vision binaire du mécanisme a mené à de nombreux débats sur l'activation ou non du processus dans certaines circonstances. Il est maintenant admis que l'EMT n'est pas un processus binaire mais que des états d'EMT intermédiaires ou partiels existent (Figure 9). Ces états intermédiaires, qualifiés de métastables ou hybrides, ont été associés à tous les sous-types d'EMT. Ces états sont

extrêmement hétérogènes, certains états précoces n'étant associés qu'à une perte de marqueurs épithéliaux sans gain de marqueurs mésenchymateux. De même, il y a de la variabilité quant aux marqueurs mésenchymateux induits. Ils sont fortement dépendants du contexte dans lequel se déroule l'EMT et apportent des propriétés différentes aux cellules. Ces états intermédiaires reflètent la balance complexe des facteurs de transcription et des ARN régulateurs qui orchestrent l'EMT. ¹¹⁶

Selon le contexte physiologique ou pathologique dans lequel elle se déroule, l'EMT a été classée en trois sous-types lors du 3^{ème} symposium international sur l'EMT de 2007 en Pologne et de la réunion de mars 2008 au Cold Spring Harbor Laboratories. Le premier type est associé à l'implantation de l'embryon, l'embryogenèse et le développement des organes, le second à la régénération tissulaire et à la fibrose et le troisième à la progression tumorale et le processus métastatique. ¹¹⁷

a) EMT de type 1 :

Cette catégorie d'EMT n'intervient qu'au cours du développement embryonnaire. Elle n'est pas associée à la fibrose ou à l'invasion cellulaire. Elle fut historiquement la première étudiée et il était autrefois admis que l'EMT n'intervenait que dans ce contexte et que cette plasticité cellulaire était perdue chez l'adulte. ¹¹⁵ Le rôle de l'EMT dans le développement embryonnaire a été principalement étudié chez *Drosophila melanogaster* (où furent identifiés les « *masters genes* » de l'EMT *Snail* et *Twist*), les embryons de poulet et de xénope et *Caenorhabditis elegans*. Dans ce sous-type d'EMT, les cellules mésenchymateuses émanent de cellules épithéliales primitives et vont générer, après migration et dispersion un nouvel épithélium supportant l'organogenèse. ¹¹⁸ L'implantation de l'embryon et l'initiation de la formation du placenta sont deux processus associés à cette EMT. Elle intervient également lors de la séparation du mésendoderme en mésoderme et endoderme ainsi que dans les cellules épithéliales du neuroectoderme pour donner des cellules migrantes qui vont se disperser dans l'embryon et se différencier en plusieurs types cellulaires. L'EMT de type 1 est déclenchée par différentes signalisations cellulaires dont les plus importantes sont la voie canonique Wnt contrôlé par des protéines de la superfamille du TGF- β et les récepteurs aux FGF lors de la

gastrulation ou les signalisations BMPs, MYB et MSX1 interagissant avec Wnt et FGF dans les cellules de la crête neurale. L'EMT de type 1 est impliquée tout au long du développement jusque dans la formation de la valve cardiaque par une EMT spécifique à partir de cellules endothéliales et ainsi appelé justement *endothelial-mesenchymal transition* (EndMT).¹¹⁷

b) EMT de type 2 :

Ce sous-type d'EMT est associé à la réparation tissulaire, et à la fibrose dans les cas pathologiques. Elle intervient dans un contexte inflammatoire sous l'influence de cytokines ou composants de la matrice extracellulaire sécrétés par les cellules immunitaires (majoritairement les macrophages) et les fibroblastes ainsi que sous l'action d'un stress mécanique. Cette EMT a vocation à augmenter le nombre de fibroblastes présents au niveau de la lésion pour la réparer.¹¹⁵ Contrairement à l'EMT de type 1 qui va créer des cellules mésenchymateuses pouvant ensuite se déplacer à travers l'organisme et créer des cellules épithéliales secondaires et de nouveaux épithéliums par MET, l'EMT de type 2 est une EMT plus partielle entraînant une différenciation de cellules épithéliales en fibroblastes (qui ne sont techniquement pas des cellules mésenchymateuse *stricto sensu*) qui vont migrer localement et conserver des marqueurs épithéliaux.¹¹⁸

L'EMT de type 2 peut être sous-catégorisée en 2 types, suivant qu'elle est activée par une inflammation aiguë ou chronique. Dans le cas d'une inflammation aiguë, l'EMT intervient comme une réponse physiologique. Les signaux inflammatoires vont induire l'EMT, permettant de recruter davantage de fibroblastes qui vont déposer de la matrice extracellulaire pour limiter l'impact de la blessure et maintenir la cohésion tissulaire. Une fois l'inflammation terminée, l'EMT s'arrête et le nombre de fibroblastes revient à la normale. Dans le cas d'une inflammation chronique, le pool de fibroblastes continue d'augmenter au détriment des cellules épithéliales, ce qui peut mener à une fibrose avec un remplacement du tissu par de la matrice extracellulaire et peu à peu une destruction de l'architecture de l'organe et de sa fonction. La différence entre ces deux types peut être difficile à cerner, car il n'est pas toujours simple de déterminer quand un signal aigu devient chronique et quand une réponse de réparation devient une fibrose.¹¹⁵ Les facteurs de croissance impliqués dans l'EMT de type 2 (et plus généralement

le processus de réparation tissulaire) sont le FGF, EGF, HGF et bien sur le TGF β . Ces ligands vont induire l'expression des EMT-TF et la répression des protéines de jonction via l'activation des voies MAPK, JNK, PI3K/AKT et TGF β .¹¹⁹

c) EMT de type 3 :

La corrélation entre l'EMT et le cancer apparait dans les années 1980. L'EMT intervient au cours de la progression tumorale pour permettre le détachement de cellules cancéreuses de la masse tumorale épithéliale et leur intravasation dans les vaisseaux sanguins.¹²⁰ A l'instar de l'EMT de type 1 qui permet seulement à certaines cellules d'un tissu épithélial de le quitter, l'EMT de type 3 n'intervient que dans certaines cellules de la tumeur. Celle-ci a généralement lieu au niveau du front d'invasion, mais son exécution est extrêmement hétérogène. Il existe vraisemblablement un gradient d'activation de l'EMT allant d'une activation complète à certains niveaux du front d'invasion à une activation partielle sur le reste du front jusqu'à une absence d'activation dans le centre de la tumeur.¹¹⁶ Ce gradient serait évidemment très variable selon les tumeurs. Le centre des tumeurs peut être sujet à l'hypoxie et cette dernière a également été liée à l'induction de l'EMT, notamment par le facteur de transcription *Hypoxia-inducible factor-1- α* (HIF-1 α) et des modifications épigénétiques. L'induction de l'EMT dans les cellules cancéreuses est aussi extrêmement dépendante du microenvironnement tumoral. Les macrophages (TAM) et les fibroblastes associés à la tumeur (CAF) jouent un rôle prépondérant dans l'activation du processus, principalement via la sécrétion de cytokines comme le TGF β .

Les cellules cancéreuses engageant un processus d'EMT sont capables de dégrader la membrane basale et d'envahir la matrice via l'expression des MMP-2 et 9 et leur capacité motile. Elles maintiennent également une sécrétion autocrine de facteurs de croissance leur permettant de résister à l'apoptose et *l'anoikis* (mort cellulaire induite par la perte des contacts entre cellules et avec la matrice extracellulaire).¹²⁰ Parmi les cellules cancéreuses effectuant une EMT, les plus remarquables sont les cellules cancéreuses circulantes, des cellules de carcinomes ayant pénétré dans la circulation sanguine par intravasation. Ces cellules présentent une myriade de phénotypes EMT, généralement intermédiaires, avec des marqueurs à la fois épithéliaux et mésenchymateux. Ces cellules sont considérées comme les précurseurs des

métastases. Elles peuvent voyager dans la circulation sanguine, seules ou en groupe de plusieurs cellules tumorales maintenues ensemble par la plakoglobin ou par des interactions avec les plaquettes (leur permettant également de se dissimuler des NK). Ces clusters, qui se forment lors de la séparation avec la tumeur primaire et non dans la circulation, sont plus efficaces pour générer des tumeurs secondaires que les cellules circulantes seules.^{116,121} Il est toutefois important de noter que seule une extrême minorité des cellules cancéreuses quittant la tumeur primaire finissent par former des macro-métastases. Les cellules cancéreuses circulantes sont en effet soumises à des conditions extrêmes lors de l'intravasation dans la circulation sanguine (force de poussée, immunité innée, etc) et ne bénéficient plus de l'immunosuppression de la tumeur primaire lorsqu'elles atteignent un autre tissu. La colonisation métastatique reste donc heureusement un processus peu efficace, sans quoi le cancer serait un problème encore bien plus grave qu'il ne l'est déjà.¹²²

La nécessité de l'EMT dans la cascade métastatique a été longuement discutée depuis le début des années 2000. Plusieurs études portant sur le suivi de cellules mésenchymateuses dans les métastases de cancer du poumon chez la souris ou l'impact de la délétion de Snail et Twist sur les métastases du cancer du pancréas chez la souris, ont conclu que l'EMT, bien que favorisant la chimiorésistance, était dispensable à la dissémination métastatique.^{123,124} De nombreux arguments permettent cependant d'appuyer le rôle de l'EMT dans la cascade métastatique. Tout d'abord, la redondance des facteurs de transcription de l'EMT. En effet, la perte de l'un des facteurs, aussi important soit-il, n'est pas toujours suffisante pour supprimer les propriétés invasives et métastatiques des cellules. De plus, il est admis qu'une EMT totale n'arrive que très rarement dans les cellules cancéreuses (presque uniquement dans les carcinosarcome) et n'est pas favorable à la formation de métastases car ces cellules ont perdu leur capacité à initier la formation de tumeur. A l'inverse, les cellules ayant effectué une EMT partielle sont beaucoup plus plastiques et à même de s'adapter aux changements d'environnement et de former des colonies en effectuant une MET.^{125,126} Ces types d'EMT sont de plus dépendants du type cellulaire, ce qui implique que les mécanismes moléculaires induisant l'EMT et les phénotypes qui en découleront seront variables selon les cancers. Enfin, des coopérations étroites entre les cellules mésenchymateuses et épithéliales ont été mises en évidence dans le processus d'invasion et de métastase. La migration collective d'un grand nombre de cellules, toujours

attachées entre elles par des jonctions cellulaires, existe mais les cellules sont généralement accompagnées de quelques cellules mésenchymateuses servant « d'éclaireurs » aux fronts d'invasion tumorales et dégradant la matrice extracellulaire afin de faciliter l'avancée du front de migration. La nature de ces cellules reste toutefois à confirmer, certaines études ayant montré que des fibroblastes pouvaient également remplir cette fonction.^{116,127}

Il semble donc admis que les cellules cancéreuses effectuent une EMT, au moins partielle, pour se détacher de la tumeur primaire. Plusieurs études ont également supposé que le maintien des cellules cancéreuses dans des états intermédiaires d'EMT favorisait leur résistance aux changements d'environnement (le flux sanguin ou lymphatique et le tissu secondaire), mais l'hétérogénéité et la plasticité des états intermédiaires d'EMT rencontrée dans les cellules métastatiques impliquent que le phénotype EMT peut varier pour une même cellule au cours de la cascade métastatique. Il a d'ailleurs été montré que les cellules cancéreuses s'établissant dans un nouveau tissu effectuaient une MET pour s'implanter dans le tissu, regagner les capacités prolifératives du phénotype épithélial et former une tumeur secondaire.^{114,128}

L'EMT a également été associée avec les capacités de cellules souches que possèdent certaines cellules cancéreuses, nommées les *cancer stem cells* (CSCs). Ce phénotype souche (*stemness*) et la capacité d'auto-renouvellement en découlant facilitent les métastases et la formation de tumeurs secondaires. Les CSCs sont également particulièrement résistantes à l'apoptose et aux traitements anticancéreux et peuvent rester en dormance pendant de longue période. Il est fréquent de voir ce phénotype souche chez des cellules effectuant une EMT. Plusieurs EMT-TF (e.g. ZEB1) permettent ainsi le maintien d'un phénotype cellule souche via la répression de certains miR dans les carcinomes pancréatiques peu différenciés. Ces miR ciblent des facteurs de cellules souches comme SOX2 et KLF4, ou SNAI1. Leur répression permet un switch entre une division asymétrique des cellules souches vers une division symétrique permettant ainsi l'augmentation du pool de CSCs.¹²⁹

Néanmoins, il semblerait que l'EMT et la dédifférenciation soit davantage des mécanismes concomitants qu'interdépendants et ainsi régulés par l'intermédiaire des actions spécifiques de chaque EMT-TF plutôt que de leur action conjointe dans l'induction de l'EMT. Certains EMT-TF ont même des actions inhibant le phénotype souche comme PRRX1 dans le cancer du sein.¹³⁰ De plus, les CSCs métastatiques effectuent une MET pour s'implanter au niveau du site

de la tumeur secondaire sans pour autant perdre leurs propriétés souches. Les facteurs contrôlant le processus de MET, telle que ID ou CTGF, sont capables d'induire des gènes impliqués dans la pluripotence. Les deux processus peuvent également être induits indépendamment via des voies de signalisation différentes.¹¹⁶ Plusieurs études ont en revanche montré que l'EMT et le phénotype souche semble davantage lié dans les cas d'EMT intermédiaire. Les cellules en EMT partielle semblent plus à même d'exhiber des propriétés de CSCs.¹²⁶ Cette association entre EMT hybride et *stemness* est également retrouvée dans des cas physiologique avec par exemple les progéniteurs hépatiques, exhibant des marqueurs épithéliaux et mésenchymateux et pouvant donner naissance à des cellules épithéliales et mésenchymateuses. Pour expliquer ce lien, l'existence d'une « fenêtre souche » permettant l'acquisition du caractère souche au cours des stades intermédiaires d'EMT a été proposée. Cette fenêtre dépendrait d'une balance entre les facteurs LIN28/Let-7 permettant une expression modérée de Oct4 contrôlant le phénotype souche. Cette balance serait impactée par des acteurs de l'EMT comme ZEB ou miR-200 et un équilibre de ces facteurs dans les cellules en EMT hybride favoriserait ainsi la balance LIN28/Let-7 et donc le phénotype souche.¹¹⁶

Les CSCs sont de plus les causes principales de récurrence et de résistance aux thérapies, ce qui associe *de facto* la chimiorésistance aux phénotypes EMT intermédiaires. Plusieurs études ont montré que les cellules effectuant une EMT partielle devenaient plus résistantes aux chimiothérapies que les cellules restant épithéliales. Les mécanismes de ces résistances sont encore mal connus mais comprennent de meilleures réparations de dommage à l'ADN, une résistance à l'apoptose et/ou l'élimination des drogues cytotoxiques.¹³¹

d) Modifications phénotypiques

L'initiation de l'EMT commence par la déstabilisation des jonctions cellulaires serrées, adhérentes, communicantes et des desmosomes ainsi que de la machinerie de régulation de la polarité cellulaire (Figure 9). Les protéines de jonctions (claudine, occludine) sont relocalisées ou dégradées, leur expression est réduite tandis que ZO-1 diffuse à travers les contacts cellulaires. Les protéines des jonctions adhérentes, notamment l'E-cadhérine, sont clivées au niveau de la membrane plasmique et dégradées. De plus sous l'action des EMT-TF, l'E-

cadhérine est soumise à une répression transcriptionnelle. Elle est généralement médiée par SNAI1 qui recrute HDAC au niveau du promoteur de l'E-cadhérine, entraînant la déacétylation des histones et le recrutement de PRC2 qui va méthyler les histones et induire une répression transcriptionnelle.¹³² Le promoteur de CDH1 contient également des E-box bipartites servant de domaine de liaison à l'ADN pour ZEB1 et ZEB2 et permettant une répression transcriptionnelle indépendante de PRC2.¹³³ La répression de l'E-cadhérine et son remplacement par la N-cadhérine est un élément majeur de l'EMT souvent mentionné comme le « *cadherin switch* ».

La perte de l'E-cadhérine qui interagit normalement avec la β -caténine pour le maintien des jonctions adhérentes entraîne une relocalisation nucléaire de cette dernière (en cas de mutation ou d'activation de la voie Wnt) ou sa dégradation par le protéasome (figure 3). La β -caténine nucléaire interagit avec des membres de la famille Tcf/Lef et va moduler l'expression d'un grand nombre de gènes impliqués dans la migration et l'invasion cellulaire. De manière assez similaire, la p120-caténine peut relocaliser dans le noyau suite à la déstabilisation des jonctions adhérentes et induire une dé-répression de certains gènes cibles de la β -caténine, (via l'interaction avec les facteurs de transcription Kaiso, Glis2 et le complexe REST/CoREST), ces gènes nécessitant ainsi l'action combinée de p120 et β -caténine pour être induits.¹³⁴ Les desmosomes sont également désorganisés et la diminution de la connexine entraîne la destruction des jonctions communicantes. La plupart des protéines de jonction cellulaire sont ainsi soumises à répression transcriptionnelle lors de l'induction et pour le maintien de l'EMT. Les complexes de polarité cellulaire, intimement liés aux différentes jonctions comme les complexes PAR et Crumbs pour la polarité apicale et Scribble, DLG et LGL pour la polarité basale sont également défaits et soumis à répression transcriptionnelle, ce qui va réduire l'adhérence cellulaire et faciliter la motilité.¹¹³

La seconde étape de l'EMT vise à réorganiser le cytosquelette afin d'augmenter la motilité et d'améliorer la flexibilité nécessaire aux processus d'intravasation et d'extravasation (Figure 9). Cette réorganisation est contrôlée par les RHO GTPases, notamment RHOA, RAC1 et CDC42, également impliqués dans la stabilité des jonctions cellulaires. La N-cadhérine, dont l'expression est induite par TWIST, joue un rôle important dans l'activation de RAC1 qui permet le recrutement de la cortactine au niveau du complexe d'adhésion de la N-cadhérine.¹³²

Le cytosquelette d'actine est alors réorganisé et des saillies membranaires riches en actine comprenant les lamellipodes et les invadopodes émergent de la cellule. Ces protrusions servent à la cellule à se repérer au cours de la migration, à avancer sur la matrice extracellulaire et occasionnellement à la dégrader. Des fibres de stress d'actine sont également formées pour augmenter la contractilité cellulaire. Ce mécanisme est contrôlé par ROCK et DIA1, favorisant la polymérisation de l'actine et la contractilité de l'actomyosine.¹¹³ Une modification du répertoire des intégrines a généralement lieu pour favoriser les capacités migratoires et invasives des cellules, comme les intégrines αV et $\beta 3$ induite par SNAI1 et $\beta 6$ capable d'activer le TGF β latent stocké dans la matrice extracellulaire.¹³²

e) Régulation de l'EMT

L'EMT est un processus extrêmement complexe et étroitement régulé. Au cours du développement embryonnaire, les cellules épithéliales établissent des gardes fous pour maintenir leur phénotype épithélial et résister aux signaux pro-EMT. Cette homéostasie épithéliale explique pourquoi l'EMT n'a lieu que dans une petite proportion des cellules du tissu. De la même manière, de nombreuses cellules cancéreuses ne sont pas sensibles à des signaux pro-EMT médiées par des voies telle que la signalisation TGF β et Wnt due à la présence de modifications épigénétiques ou d'inhibiteurs au sein de la cellule. De nombreuses boucles de rétrocontrôle négatif existent et sont généralement perdues ou perturbées dans les cellules cancéreuses (l'axe SNAI1-miR34 ou ZEB1-miR200) permettant l'activation et le maintien de l'EMT dans ces cellules.¹³¹ Dans les hépatocytes, les facteurs nucléaires hépatiques (HNF) et tout particulièrement HNF4A et FOXA1/2/3 sont les gardiens du phénotype hépatocytaire (Figure 9) et nécessitent d'être réprimés pour que l'EMT puisse être induite.¹³⁵ Le facteur *Tat-interacting protein* (TIP30) joue également un rôle dans le blocage de l'EMT en empêchant la translocation nucléaire de SNAI1.¹³⁶

Trois familles de facteurs de transcription sont particulièrement importantes et ont été intensément étudiées dans le contrôle de l'EMT. Ces facteurs, communément appelés les « *masters genes* », « *Core transcription factor* » de l'EMT ou EMT-TF, sont les membres des familles SNAI, TWIST et ZEB. Ces facteurs assurent le contrôle transcriptionnel de l'EMT

(Figure 9) et sont retrouvés constamment dans les études s'y intéressant (toujours au moins un de ces facteurs de transcription est impliqué dans l'ensemble des EMT de type 1 identifiées jusqu'à présent).¹¹⁴ Snail (l'orthologue de SNAI1 chez *Drosophila melanogaster*) fut le premier facteur de transcription capable d'induire l'EMT identifié, notamment par sa capacité à réprimer l'expression de l'E-cadhérine.¹³⁷ Le mécanisme d'action majeur de ces EMT-TF est la répression transcriptionnelle des protéines de jonction cellulaire incluant l'E-cadhérine mais

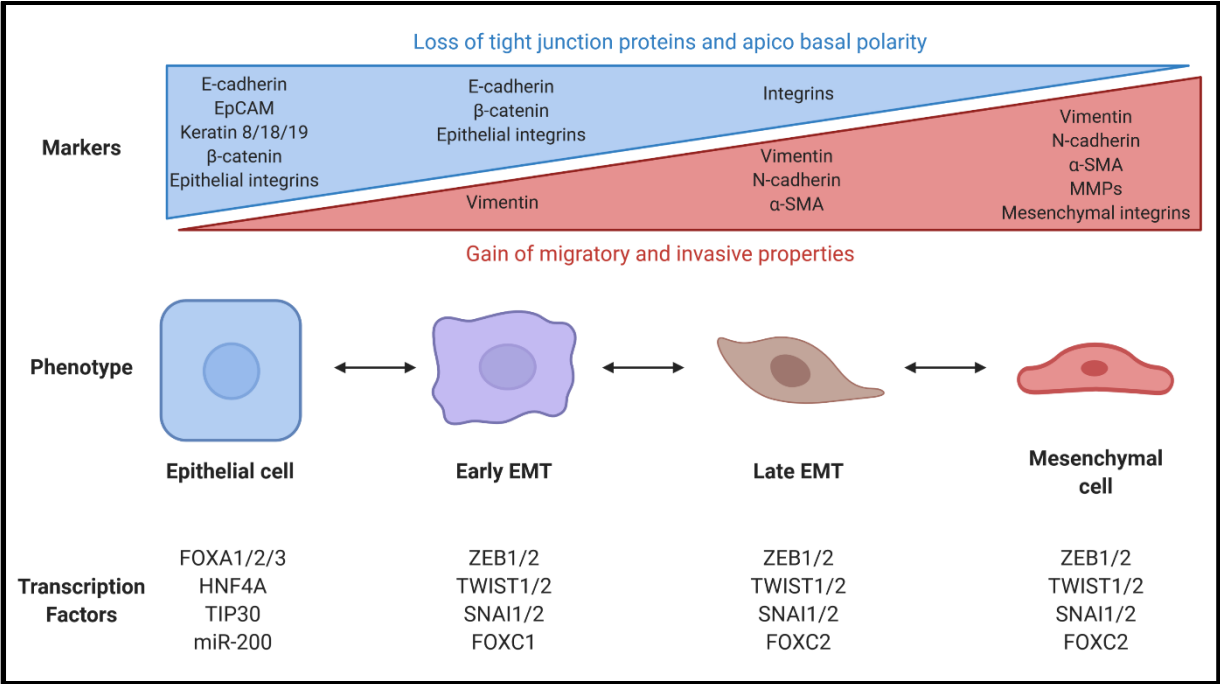


Figure 9 : Mécanismes et marqueurs de la transition épithélio-mésenchymateuse dans les hépatocytes.

également la claudine ou les desmosomes. La plupart des EMT-TF assurent cette fonction en interagissant avec des modificateurs épigénétiques permettant des modifications importantes du paysage chromatinien, comme par exemple SNAI1 qui interagit avec PRC2 pour induire la triméthylation de H3K27 sur les nucléosomes du promoteur de *CDH1*.¹³⁸ Les EMT-TF sont tous capable d'activer un switch vers un phénotype mésenchymateux. Ils agissent généralement en synergie et ciblent les mêmes gènes mais malgré leur fonction redondante dans l'induction de l'EMT, ils possèdent également des actions non-redondantes, indépendantes de l'EMT, conférant des propriétés particulières supplémentaires aux cellules qui les expriment. ZEB1 par exemple favorise la résistance aux dommages à l'ADN subi par irradiation en induisant la réponse ATM tandis que TWIST1 à l'inverse tend à bloquer la réponse aux dommages à l'ADN.¹³⁰ Bien que le contrôle transcriptionnel des EMT-TF soit la partie centrale et la plus étudiée dans la régulation de l'EMT, d'autres réseaux interconnectés y participent également. Les ARN non codants jouent un rôle important dans le contrôle de l'expression des EMT-TF au niveau post-transcriptionnel. Les membres de la famille miR-200 sont fortement associés à la différenciation épithéliale et impliqués dans une boucle de rétrocontrôle avec ZEB1 dont ils ciblent l'ARNm au niveau du 3'UTR, tandis que ZEB1 contrôle leur expression par fixation sur leurs promoteurs. A l'inverse, des miR vont promouvoir l'EMT telle que miR-21 qui favorise l'invasion et les métastases dans le CHC par inhibition du suppresseur de tumeur *Pdcd4*.¹³⁹ D'autres sous-types d'ARN ont également été étudiés pour leur implication dans l'EMT. L'ARN longs non codants ZEB2-AS1 augmente ainsi l'expression de ZEB2. Il permet, en recouvrant la partie 5' d'un intron de ZEB2 contenant un IRES, d'empêcher son épissage et permet ainsi la transcription de ZEB2.¹⁴⁰ Plus récemment, le rôle des ARN circulaires a également été étudié, notamment pour le rôle d'éponge à miRNA pouvant cibler des EMT-TF.¹⁴¹

SNAI, ZEB et TWIST sont donc les acteurs essentiels du processus d'EMT mais leurs actions sont cependant secondées par de nombreux autres facteurs de transcription plus récemment identifiés qui vont coopérer avec eux, faciliter leurs actions et maintenir leur expression. Parmi ces nouveaux facteurs, on trouve entre autres les protéines SOX (*SOX4*, *SOX9*), AP1, PRRX1 et notamment des membres d'une famille de facteurs de transcription qui va nous intéresser particulièrement et dont nous parlerons plus loin : les *forkhead box protein* (FOX).^{116,125}

6) Induction de l'EMT par le TGFβ

Le TGFβ est connu pour être l'un des plus puissants inducteurs de l'EMT. Cette action intervient de manière physiologique au cours du développement embryonnaire et lors du processus de réparation, mais est bien évidemment détournée dans le cancer pour favoriser la migration, l'invasion et la métastase des cellules cancéreuses. Les trois ligands de la voie TGFβ peuvent agir sur l'EMT mais son induction dans les contextes de fibrose et de cancer est essentiellement contrôlée par le TGFβ1. L'EMT induite par le TGFβ repose sur une combinaison d'actions des voies canoniques et non-canoniques (figure 10). Les protéines SMAD sont capables de fixer le promoteur de la plupart des EMT-TF (SNAI1/2, ZEB1/2, FOXC2, HMGA2) et d'induire leur expression.¹¹³ SNAI1 et SNAI2 sont directement induits par SMAD3 et SMAD4 qui viennent se fixer sur leur promoteurs, tandis que ZEB1/2 sont induits indirectement par ETS1 et agissent ensuite en coopération avec les SMAD pour réprimer l'expression des marqueurs épithéliaux. HMGA2 n'est que très peu exprimé chez l'adulte mais est induit dans certains cancers via SMAD3/4, et son expression va induire celle de SNAI1/2 et TWIST1. Les SMAD inhibent également l'expression de *ID1*, en se fixant sur son promoteur après formation d'un complexe avec ATF3.¹⁴² Les SMAD peuvent également induire l'expression de *LEF1* qui peut activer la β-caténine via l'inhibition de GSK-3β ou interagir avec SMAD2 et SMAD4 pour inhiber l'expression de l'E-cadhérine, généralement sous l'action du TGFβ3.¹⁴³

De manière indépendante des SMAD, le TGFβ est également capable d'activer d'autres signalisations telles que ERK de la voie des MAPK. Cette activation se fait par le recrutement et la phosphorylation de ShcA par le TGFβRI permettant ainsi la formation du complexe Shc-GRB2-SOS qui va activer Ras. La cascade de signalisation en découlant va alors potentialiser la réponse pro-EMT des SMAD.¹⁴² Ainsi ERK va favoriser la réponse transcriptionnelle et ERK5 va stabiliser SNAI1 par inhibition de GSK3β. Les voies p38 et JNK sont également activables par le TGFβ via l'activation de TAK1 résultant de l'association de TRAF6 avec les récepteurs du TGFβ. TAK1 et p38 vont permettre l'expression de ATF2, une protéine

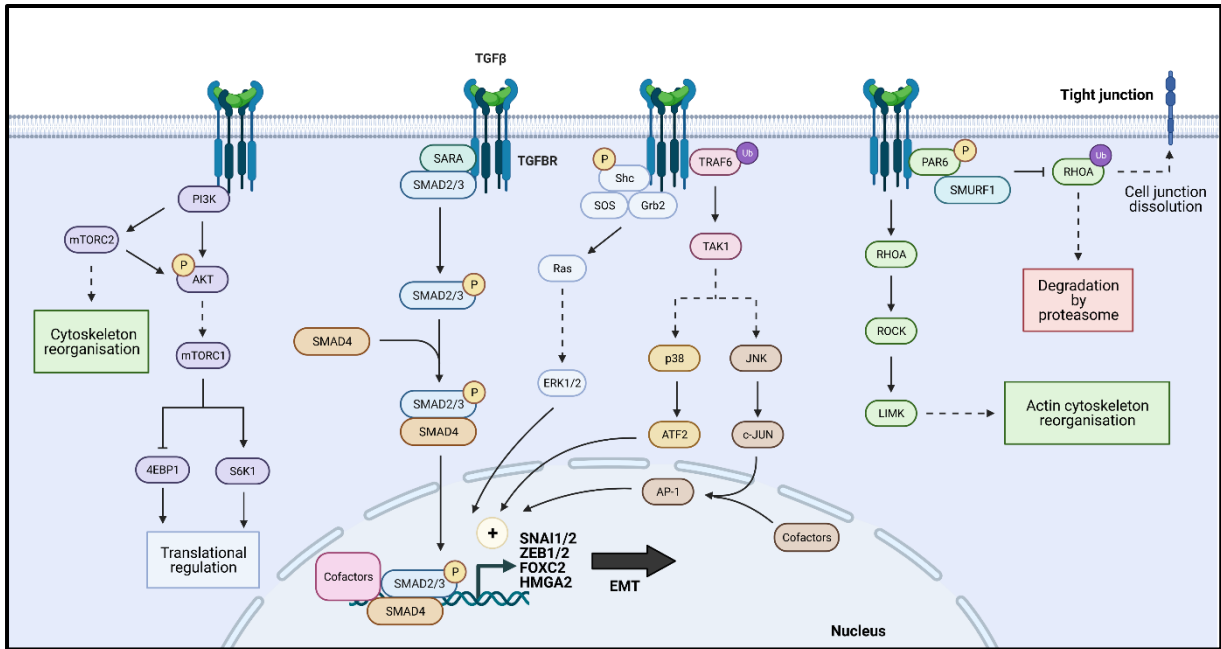


Figure 10 : Induction de l'EMT par le TGFβ.

interagissant avec les SMAD pour activer des gènes de l'EMT. De la même manière, la voie JNK va activer c-JUN, membre du complexe AP-1 qui interagit avec les SMAD pour orienter leur réponse transcriptionnelle.^{113,143}

La voie des Rho-like GTPase est également impliquée dans l'EMT induite par le TGF β , notamment via une régulation fine de RHOA. Le TGF β est capable d'activer RHOA par l'induction des *guanine nucleotide exchange factors* (GEFs) et l'activation de ROCK et LIMK, induisant ainsi un remodelage du cytosquelette d'actine nécessaire à l'EMT. RHOA est toutefois dégradé localement au niveau des jonctions serrées sous l'effet du TGF β grâce à la phosphorylation de PAR6 par le TGFBR2. Cette phosphorylation permet également le recrutement de SMURF1 et entraîne localement l'ubiquitylation et la dégradation de RHOA, favorisant le démantèlement des jonctions serrées.^{113,142}

Enfin, le TGF β peut activer la voie PI3K/AKT, résultant en l'activation de mTORC1 favorisant l'invasion et la mobilité cellulaire par régulation de la traduction et mTORC2 induisant le changement vers un phénotype mésenchymateux.¹¹³

Des coopérations avec d'autres signalisations telles que Wnt et Notch existent également pour induire l'EMT, mais elles ne sont pas directement activées par le TGF β . Dans le CHC, les signalisations TGF β et Wnt semblent toutefois fortement entremêlées puisque les tumeurs avec une co-activation de ces voies forment une sous-catégorie de la classe proliférative.⁵⁶

7) TGF β et carcinome hépatocellulaire

Comme nous l'avons vu, le TGF β est fortement impliqué dans l'EMT aussi bien que dans d'autres processus biologiques, rendant sa signalisation quelque peu complexe. Son action dans le CHC n'est pas une exception. Le TGF β est fortement impliqué dans le processus fibrotique conduisant à l'apparition du CHC, notamment par sa capacité à activer les cellules étoilées du foie et à favoriser le dépôt de matrice extracellulaire. Le TGF β joue également de nombreux rôles lors de l'initiation et la progression du cancer. Si son action sur le microenvironnement est clairement pro-tumorale, de par ses propriétés immunosuppressives et profibrotiques, son impact sur les cellules cancéreuses est beaucoup plus variable avec des effets suppresseurs de tumeurs ou, à l'inverse, pro-métastatique. La signalisation TGF β présente ainsi une dualité

fonctionnelle dans le CHC (ainsi que dans d'autres cancers) caractérisée d'une part par une action cytostatique anticancéreuse et d'autre part par une action métastatique pro-tumorale (Figure 11).

L'action cytostatique du TGF β passe par l'inhibition du cycle cellulaire et plus précisément de son blocage en phase G1 par l'induction de *cyclin-dependent kinase inhibitors* (CDKIs), de p21 et de p15. Le TGF β stoppe également la prolifération par le blocage de l'expression de *MYC* et l'induction de 4EBP1.¹⁴⁴ Cette action anticancéreuse du TGF β intervient essentiellement dans les phases précoces de la cancérogenèse hépatique sur les cellules pré-malignes ou au début de la progression tumorale (Figure 11). La signature moléculaire associée à cette réponse cytostatique est d'ailleurs qualifiée de « *early response* » et associée à un bon pronostic chez les patients atteints de CHC.⁵⁴ L'action cytostatique est également doublée d'une action pro-apoptotique. Les mécanismes d'induction de l'apoptose par le TGF β semblent toutefois varier. Dans le cancer du pancréas par exemple, il induit l'apoptose dans les cellules progénitrices présentant des mutations de *KRAS*.^{108,145} Cet effet pro-apoptotique n'est toutefois que peu retrouvé dans le CHC probablement dû à la faible prévalence des mutations de *KRAS* (<3%). Dans ce cancer, l'action pro-apoptotique du TGF β passe par la production d'espèces réactives de l'oxygène via l'induction de NOX4.¹⁴⁶

Au cours de la progression tumorale cependant, les cellules cancéreuses parviennent d'une manière ou d'une autre à s'affranchir de ces signaux cytostatiques et apoptotiques. Ces tumeurs peuvent alors utiliser le TGF β à leur avantage, notamment pour induire l'EMT et favoriser la migration, l'invasion et les métastases (Figure 11).

La question principale qui se pose alors est par quels mécanismes ces tumeurs parviennent-elles à échapper à l'action cytostatique au bénéfice de l'action pro-métastatique ? Hors, comme observé précédemment, les mécanismes qui orientent et façonnent la réponse au TGF β sont nombreux et complexes.

Ainsi, certaines tumeurs optent simplement pour une inactivation pure et simple de la signalisation TGF β par mutation du récepteur et/ou des protéines SMAD. C'est par exemple le cas dans le cancer du côlon où *TGFBR2* est fréquemment muté ou dans l'adénocarcinome pancréatique avec les mutations de *SMAD4*.¹⁴⁵ Dans le CHC toutefois, ce type de mutation est

rare^{10,31}, surtout au regard du nombre de tumeurs présentant une signature TGFβ tardive (pro-métastatique), impliquant que d'autres mécanismes sont à l'œuvre dans le CHC pour bloquer l'action cytotatique du TGFβ. Cela peut s'opérer par une extinction de la signalisation via une inhibition de ses effecteurs. Des lignées cellulaires de CHC présentent ainsi une surexpression de SMAD7 diminuant l'activité de SMAD3, tandis que d'autres perdent l'expression de *SPTBN1*, un adaptateur associé à SMAD3/4 et favorisant leur action cytotatique.¹⁴⁷

L'impact de la communication croisée avec d'autres voies de signalisation a bien évidemment été exploré. La signalisation EGF et PI3K/AKT serait ainsi capable de contrebalancer l'action apoptotique du TGFβ, notamment en bloquant l'induction de NOX4 par le TGFβ. Ce dernier serait d'ailleurs également capable de transactiver EGFR via l'action de ADAM17 et de la caveoline-1.¹⁴⁸ L'action métastatique du TGFβ pourrait également être liée (en plus de sa capacité à induire les EMT-TF) à l'activation de la signalisation PDGF favorisant l'activation et la translocation de la β-caténine au noyau et protégeant contre l'anoikis.¹⁴⁹ Le processus d'EMT une fois induit semble également exercer une inhibition de l'action cytotatique du TGFβ via SNAIL1, et se renforce en augmentant la sécrétion autocrine de TGFβ, suractivant la voie et transactivant EGFR.^{148,150} Une autre étude a également montré que la phosphorylation aberrante de SMAD3 par l'axe AXL/14-3-3ζ/JNK semblait favoriser l'action pro-métastatique du TGFβ au détriment de l'action cytotatique.¹⁵¹

Ainsi, si certains mécanismes, et notamment des communications croisées, ont été mis en évidence, beaucoup reste à faire dans le CHC pour bien comprendre ce changement d'action du TGFβ. Comme vu précédemment, la réponse cellulaire au TGFβ peut également être orientée par le contexte transcriptionnel de la cellule. Le switch d'action pourrait ainsi dépendre de certains facteurs de transcription avec lesquels les SMAD interagiraient dans certaines cellules cancéreuses pour déclencher une EMT. Ces facteurs de transcription pourraient également n'être induits qu'en cas d'expositions prolongées et chroniques au TGFβ ainsi qu'à d'autres cytokines. Considérant cela, l'identification de nouveaux effecteurs de la signalisation TGFβ induits dans les cellules cancéreuses pourrait permettre d'expliquer ce changement d'action du TGFβ vers un rôle pro-métastatique.

De manière à identifier de nouveaux effecteurs du TGFβ, nous avons opéré un criblage dans une lignée cellulaire de CHC mésenchymateuse, les SNU-449, provenant d'un homme coréen

de 50 ans positif au VHB. Parmi les nombreux gènes dont l'expression est induite par le TGF β dans cette lignée cellulaire, un s'avère particulièrement intéressant : *FOXS1*. Il s'agit d'un facteur de transcription relativement peu étudié (20 résultats sur pubmed, dont 8 publiés pendant ma thèse), fortement induit par le TGF β et appartenant à la famille des *forkhead box protein* (FOX).

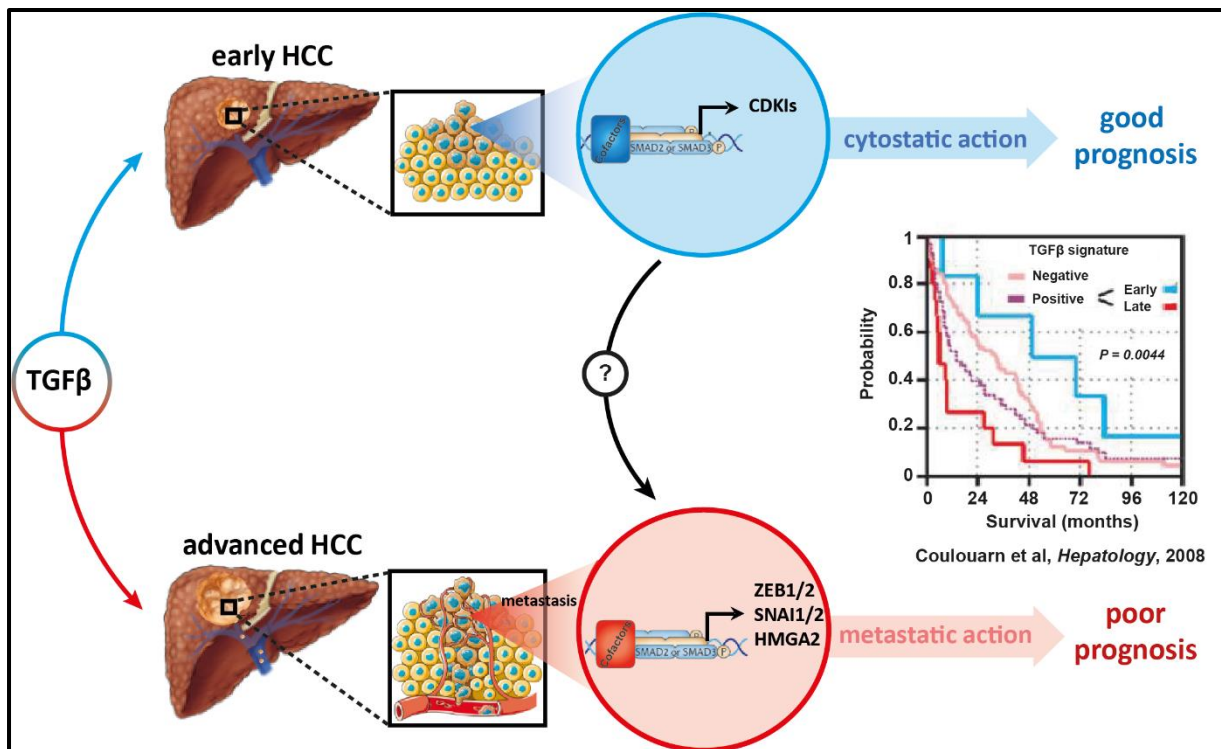


Figure 11 : Dualité fonctionnelle du TGFβ dans le carcinome hépatocellulaire.

III) Les facteurs de transcription de la famille des forkhead box

1) Généralités

Les *forkhead box proteins* (FOX) forment une superfamille d'une cinquantaine de facteurs de transcription conservés au cours de l'évolution. Le premier membre de cette famille a été identifié chez *Drosophila melanogaster* en 1989 par Detlef Weigel et Herbert Jäckle. Ce gène homéotique fut appelé forkhead (fkh) à cause du phénotype que donnait ses mutations chez l'embryon, à savoir le remplacement de la partie ectodermale des intestins par des structures ectopiques de la tête en forme de fourchette, qui donna ensuite son nom à toute la famille.¹⁵² Des centaines de membres de cette famille ont été identifiés chez les eucaryotes, de la levure à l'homme. Le caractère commun de toutes ces protéines est leur domaine de liaison à l'ADN : la *forkhead box* (FKH) ou *winged helix domain* dont la conservation est remarquable. On retrouve ainsi 90% de similarité dans la séquence d'acides aminés entre la protéine Forkhead de *Drosophila melanogaster* et la protéine FOXA1 chez l'homme.¹⁵³ Le domaine FKH canonique est constitué de trois hélices α , trois feuillets β et deux « *wing* » régions entourant le troisième feuillet (donnant son nom alternatif au domaine). La cinquantaine de gènes appartenant à cette famille chez l'homme est divisé en 19 sous-classes (FOXA à FOXS) par analyse phylogénétique de la séquence d'acides aminés de la *forkhead box*.¹⁵⁴ La classification des FOX a ainsi été unifiée sur cette base en 2000 (figure 12B).¹⁵⁵

Les FOX sont impliqués de près ou de loin dans une grande variété de processus biologiques allant de l'homéostasie du glucose à l'apprentissage du langage en passant par le développement embryonnaire, le vieillissement et bien sûr le cancer.¹⁵³ Ces protéines étant des facteurs de transcription, elles agissent en coordonnant l'expression spatio-temporelle de nombreux gènes au cours du processus de développement ou pour permettre le maintien de l'homéostasie tissulaire. Toutefois, ces protéines n'ayant en commun que leur domaine de liaison à l'ADN, leur actions varient fortement d'une protéine à l'autre avec des rôles parfois opposés et dépendant de leur autres domaines d'inhibition, de transactivation ou d'interaction

avec d'autres facteurs de transcription, et entre souvent en compétition pour la fixation sur un même motif d'ADN.¹⁵⁴ Les FOX sont des protéines indispensables aux réponses cellulaires sous l'action de l'environnement. En plus de leur action comme activateurs ou répresseurs transcriptionnels, nombre de ces facteurs de transcription sont également des facteurs dits pionniers. Ces facteurs sont ainsi les premiers à venir se positionner sur des portions d'hétérochromatine condensée et inaccessible à la transcription, où ils vont faciliter le recrutement d'autres facteurs pour débobiner la chromatine et activer la transcription.¹⁵⁶ Cette capacité de certains FOX à interagir avec la chromatine compactée leur permet également de servir de « *bookmarking* » : Ces facteurs peuvent interagir avec la chromatine condensée de la cellule en mitose de manière à restaurer l'expression de certains gènes particuliers à la fin de la division cellulaire et/ou à marquer des gènes de manière à conserver « la mémoire » de son expression génique au cours de la mitose.^{154,156}

Considérant le nombre de FOX et de processus associés, nous ne parlerons ici que des membres de cette famille impliqués dans le cancer et la signalisation TGF β , ce qui représente déjà un nombre non négligeable de ces protéines.

De nombreux membres de la famille des FOX sont impliqués dans l'apparition et le développement du cancer, à la fois comme oncogène et gène suppresseur de tumeur (Figure 12A). Parmi les sous-classes de FOX impliqués dans l'initiation, la progression ou la chimiorésistance tumorale, on trouve notamment les FOXA, FOXC, FOXM, FOXO et FOXP.^{154,157} Ces protéines sont impliquées dans un grand nombre de cancers et dans tous leurs « *hallmarks* » (Figure 12A).

2) Les FOXA

Premièrement, les protéines de la sous-classe FOXA sont au nombre de trois (FOXA1, FOXA2 et FOXA3). Ces protéines sont intimement liées au foie. Autrefois appelé *Hepatic Nuclear Factor 3* (respectivement HNF3A, HNF3B et HNF3G), elles jouent un rôle primordiale dans le développement embryonnaire du foie. Ces protéines sont en effet, par leur capacité de facteurs pionniers, capables d'ouvrir la chromatine pour permettre l'expression de gènes impliqués dans le développement du foie fœtal. Le KO de *Foxa2* est ainsi létal chez la souris tandis que la

A



B

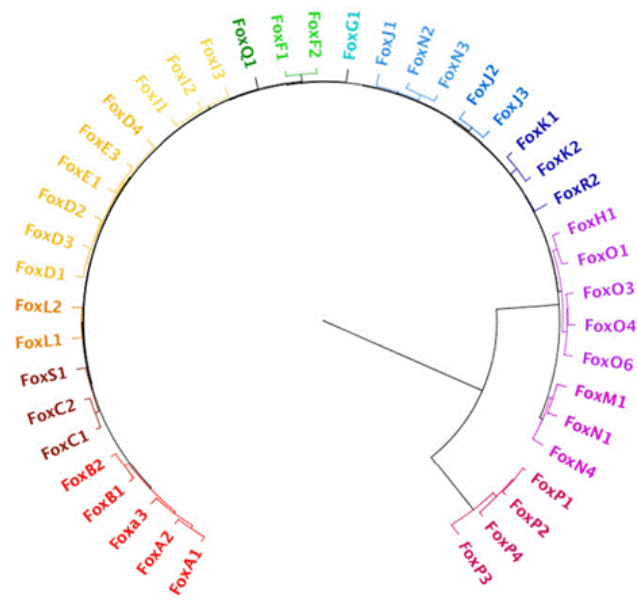


Figure 12 : Phylogénie et implications de différents membres de la famille des forkhead box dans les “hallmarks” du cancer.

A : Exemple d'implications des FOX dans les différents *hallmarks* du cancer, comme définis par Hanahan D. & Weinberg RA en 2011¹⁵⁸. Les FOX soulignés sont des suppresseurs de tumeur. B : Arbre phylogénétique de la famille des forkhead box. Arbre phylogénétique provenant de Kaestner et al, Development, 2016.

délétion conditionnelle des membres de la famille Foxa dans les cellules parenchymateuses du foie adulte entraîne une insuffisance hépatique.¹⁵⁹ Ces protéines sont également importantes chez l'adulte car très liées aux métabolismes hépatiques et aux capacités régénératives du foie. Les FOXA sont ainsi des éléments importants de l'homéostasie du glucose, FOXA2 étant capable d'activer directement la transcription du gène du glucagon ainsi que d'autres gènes impliqués dans le métabolisme et le relargage du glucose par les hépatocytes.¹⁶⁰ FOXA2 est également capable d'induire l'expression de *ALR*, un gène fortement impliqué dans la régénération hépatique après hépatectomie.¹⁶¹ Concernant leur rôle dans le cancer et notamment le CHC, elle est difficile à établir, FOXA1 et FOXA2 pouvant servir aussi bien d'oncogènes que de suppresseurs de tumeurs selon le contexte. Les FOXA sont fortement liés aux signalisations hormonales et sont impliqués dans le dimorphisme sexuel du CHC. FOXA1/2 sont capables d'interagir avec les récepteurs aux androgènes (AR) chez l'homme et le récepteur aux estrogènes (ER α) chez la femme pour permettre leur fixation sur leurs éléments cis-régulateurs. Chez la femme, cela a pour effet de bloquer la progression tumorale, tandis qu'à l'inverse chez l'homme, cela favorise la croissance tumorale.^{162,163} Des mutations au niveau des sites de fixations de ER α et FOXA1 sont d'ailleurs retrouvées dans les CHC de femme.¹⁵⁴ Cependant, leur rôle propre n'est pas clair. FOXA1 peut exercer d'une part une action anti-tumorale par inhibition de la voie PI3K/AKT (par blocage de l'expression de PI3KR1) et une action anti-métastatique par inhibition de l'expression de RNF6, et peut d'autre part favoriser la prolifération des cellules tumorales et augmenter la synthèse d'AFP. FOXA1 est d'ailleurs la cible de plusieurs miR comme miR-212 ou miR-30a-5p connus pour avoir un rôle anti-tumorale. De la même manière, FOXA2 est connu pour avoir des propriétés anti-métastatiques par inhibition de ZEB2 en coordination avec la sirtuine 6, et son inhibition par miR-92a favorise la prolifération et l'invasion tumorale ; mais d'autres études ont montré que son inhibition par miR-141 favorisait la prolifération et l'invasion tumorale et qu'il pouvait contrebalancer les effets anti-tumoraux de miR-200a.¹⁶¹ Les FOXA sont de plus des gardiens du phénotype hépatocytaire et agissent en synergie avec HNF4A pour maintenir les propriétés épithéliales des hépatocytes.¹³⁵

3) Les FOXO

Les FOXO sont un autre sous-groupe dont le rôle dans la cancérogenèse a été fortement étudié. C'est d'ailleurs avec FOXO1 qu'a été établie la première association entre FOX et cancer avec l'implication de FOXO1 dans le rhabdomyosarcome alvéolaire.¹⁵⁷ Ce sous-groupe est composé de quatre membres : FOXO1, FOXO3, FOXO4 et FOXO6. Tous les membres de cette sous-famille sont capables de se fixer sur deux séquences consensus DBE (Daf-16 family member-binding element). Les FOXO sont sujets à régulation par une grande variété de modifications post-traductionnelles incluant phosphorylation, acétylation et ubiquitination. Les membres de la famille FOXO sont exprimés dans tous les tissus, à l'exception de FOXO6 exprimé seulement dans le tissu cérébral. Ces protéines sont généralement des suppresseurs de tumeurs de par leur capacité à contrôler le cycle cellulaire.¹⁶⁴ La perte de FOXO1 par délétion chromosomique est ainsi retrouvée dans de nombreux cancers de la prostate.¹⁶⁵ FOXO1 est de plus capable de bloquer l'EMT induite par ZEB2 dans le CHC, limitant l'invasion et les métastases.¹⁶⁴ De manière intéressante, les FOXO jouent un rôle dans la médiation des effets du TGF β en interagissant avec les protéines SMAD. Dans le glioblastome, FOXO3a est capable d'interagir avec SMAD3 pour induire l'expression de p21. Cette induction peut être inhibée par un autre membre de la famille FOX : FOXG1 qui va bloquer la transcription de p21, ainsi que par l'action de la signalisation PI3K entraînant une rétention de FOXO3a dans le cytoplasme après sa phosphorylation par AKT.¹⁶⁶ Le même type de régulation a lieu dans certaines leucémies où AKT est activée par la fusion de Abl avec Bcr ou Tel, induisant la résistance de ces cellules à l'action cytostatique du TGF β .¹⁶⁷

4) Les FOXP

Le sous-groupe des FOXP est plus ambivalent. Ce sous-groupe est constitué de quatre membres (FOXP1, FOXP2, FOXP3 et FOXP4) qui ont la capacité de former des homo ou hétérodimères. Ces protéines sont normalement impliquées dans le développement embryonnaire, et notamment le développement cérébral relatif à l'apprentissage du langage, mais elles se retrouvent dérégulées dans les cancers où elles peuvent agir comme oncogènes ou suppresseurs de tumeurs.¹⁶⁴ FOXP1 par exemple a été reporté comme oncogène dans les lymphomes B et le

CHC mais est de bon pronostic dans le cancer du sein et le cancer du poumon à petites cellules. Il est supposé que ce sont les capacités de ces FOX à hétérodimériser qui sont responsables de la perte de leur action anti-tumorale, mais cette hypothèse n'a pas été validée expérimentalement.¹⁵⁴ Ces protéines et notamment FOXP3 sont de plus fortement impliqués dans l'évasion immunitaire des tumeurs. FOXP3 est en effet responsable de la différenciation en Treg des lymphocytes T naïfs, cellules immunitaires inhibitrices de la réponse anti-tumorale, notamment via la sécrétion de TGF β .¹⁶⁸

5) FOXG1/K1/M1/Q1

Quelques autres membres de la famille FOX ont été associés à la cancérogenèse hépatique et la progression tumorale du CHC. C'est par exemple le cas de FOXK1, déjà considéré comme un oncogène dans plusieurs cancers, qui a été associé à un mauvais pronostic dans le CHC. FOXK1 semble prendre part à l'acquisition d'un phénotype cellule souche dans le CHC en activant l'expression de gènes impliqués dans la pluripotence et favorisant la formation de sphéroïde *in vitro*.¹⁶⁹ FOXG1, dont nous avons déjà vu le rôle dans le blocage de l'action cytotatique du TGF β dans certains cancers, joue également un rôle dans l'induction de l'EMT dans le CHC. Cette induction passe par l'activation de la voie Wnt où FOXG1 forme un complexe avec la β -caténine pour activer l'expression des gènes cibles de la voie Wnt.¹⁷⁰ Enfin FOXQ1 est impliqué dans le dialogue entre les CAFs et les cellules tumorales hépatiques. Les CAFs sont en effet capables d'induire l'expression de FOXQ1 dans les cellules cancéreuses ce qui favorise la croissance tumorale. FOXQ1 exerce de plus une boucle d'amplification en induisant NDRG1 qui va permettre le recrutement de cellules étoilées par la sécrétion de CCL26.¹⁷¹

A l'inverse des FOXO dont le rôle est pratiquement toujours anti-tumoral, FOXM1, une protéine formant son propre sous-groupe de FOX, agit presque toujours comme un oncogène. Son expression est augmentée dans beaucoup de cancers et associé à un mauvais pronostic. FOXM1 favorise la résistance aux signaux prolifératifs en agissant sur des régulateurs du cycle cellulaire. C'est effet une protéine essentielle de la transition G1-S et G2-M. FOXM1 agit aussi sur l'EMT et les métastases, notamment via des interactions avec la voie PI3K.¹⁶⁴ Son action est liée à celle de FOXO3a car plusieurs mécanismes inhibant FOXO3a (notamment des

phosphorylation) activent également FOXM1 et ces deux protéines sont connues pour coopérer afin de réguler l'expression de ER α .¹⁵⁷

6) Les FOXC

Enfin, le sous-groupe des FOXC est particulièrement intéressant par rapport à FOXS1, étant donné qu'il s'agit du sous-groupe phylogénétiquement le plus proche de ce dernier (Figure 12B). Ce sous-groupe ne contient que deux membres : FOXC1 et FOXC2.¹⁷² Ces protéines sont impliquées dans le développement embryonnaire où elles participent à la spécification des cellules artérielles ainsi qu'à la formation des vaisseaux lymphatiques et sanguins. Leur rôle a toutefois été principalement étudié dans le cancer. Les FOXC sont connus pour avoir un rôle pro-tumoral, notamment en favorisant l'invasion et les métastases par induction de l'EMT et du phénotype cellule souche ainsi que l'angiogenèse.¹⁶⁴ Le rôle central de FOXC2 dans les capacités métastatiques des *basal-like breast cancer* (BLBC) a ainsi été identifié en 2007. Il est induit par la signalisation TGF β en aval des EMT-TF, et va promouvoir la différenciation mésenchymateuse.¹⁷³ Toujours dans le BLBC, FOXC1 favorise la prolifération, la migration et l'invasion, notamment par activation de la voie NF- κ B. Il induit également l'EMT, d'où son niveau d'expression élevé dans les cellules cancéreuses circulantes ou *circulating tumor cells* (CTCs), et il favorise les propriétés souches des cellules cancéreuses par activation de la signalisation *smoothened-independant heghog*. La régulation de *FOXC1* a été étudiée essentiellement dans les cancers du sein. Son expression dans ces cancers semble résulter de dérégulation épigénétique. *FOXC1* est normalement réprimée par le complexe PRC2 mais des mutations de *BRCA1* permettent son expression dans le cancer du sein. Il semble également régulé par méthylation, la déméthylation de son promoteur étant corrélé à une augmentation de son expression. De plus, FOXC1 est contrôlé par la signalisation EGFR via les voies MAPK et PI3K et propage ses effets prolifératifs et pro-migratoires.^{174,175}

Dans le CHC, FOXC2 régule l'expression de nombreux gènes associés à l'EMT comme *ITGB3*, *CTNND1*, *VEGF*, *MMP2* et *PDGFB* tandis que FOXC1 est capable de transactiver *SNAIL* pour réprimer l'E-cadhérine et d'activer *NEDD9* favorisant la motilité cellulaire.^{154,174} De plus, FOXC1 contrôle, dans plusieurs cancers, un grand nombre de métalloprotéases telle que

MMP1, MMP2, MMP7 et MMP9 qui favorisent la dégradation de la matrice extracellulaire et l'invasion tumorale.^{161,174} Il est également capable d'induire l'expression de CXCR1 et CCL2 par fixation direct sur leur promoteur.¹⁷⁵

La proximité phylogénétique du domaine *Forkhead box* de ces protéines avec celui de FOXS1 (77% d'identité de la séquence protéique avec FOXC1 et FOXC2) ainsi que la régulation occasionnelle de ces protéines par le TGFβ pourraient laisser présager une similarité de fonctions de ces différentes protéines. Toutefois, FOXS1 ne possède pas les domaines inhibiteurs et les domaines de transactivation de ces protéines, ce qui signifie que même si FOXS1 cible les mêmes gènes que ces protéines, son action pourrait être très différente. C'est par exemple le cas de FOXL1, l'autre membre des FOX phylogéniquement proche des FOXC et de FOXS1 (74% de similarité du domaine *forkhead box* avec FOXL1). Cette protéine a ainsi à l'inverse des FOXC un rôle suppresseur de tumeur, notamment dans le cancer du pancréas où il induit l'apoptose dans les cellules cancéreuses par induction de TRAIL et inhibe l'expression de ZEB1, réduisant ainsi l'invasion tumorale.¹⁷⁶

7) FOXS1

Les impacts variés des différents membres de la famille des FOX dans les cancers, aussi bien comme oncogènes que suppresseurs de tumeurs, soulignent l'intérêt d'étudier le rôle de FOXS1 dans la cancérogenèse hépatique. De la même manière, l'intrication de ces facteurs de transcription dans de nombreuses signalisations impliquées dans le cancer justifie FOXS1 comme un candidat intéressant pour expliquer, au moins en partie, la dualité fonctionnelle du TGFβ dans le CHC.

FOXS1 est un gène à un seul exon, situé sur le chromosome 20 et transcrit en un seul ARNm de 1302 nucléotides codant pour une protéine, FOXS1, de 330 acides aminés. FOXS1 rejoint la grande famille des FOX en 1997. Tout d'abord nommé FREAC-10 et FKHL18, la protéine et le gène sont renommés dans les années 2000 par la nouvelle classification des FOX.^{155,177} FOXS1 continue toutefois d'apparaître jusqu'en 2008 sous le nom de FKHL18.¹⁷⁸

La première analyse de son expression par *dot blot* révèle qu'il n'est exprimé que dans l'aorte.¹⁷⁷ Les bases de données recensant les données d'expression montre qu'il est exprimé dans presque tous les tissus, à l'exception des cellules du sang, bien que son expression soit relativement faible (quelques *transcripts per million*) (figure 13A). L'analyse de son expression par *single cell RNA-seq* montre toutefois que son expression semble restreinte à quelques types cellulaires spécifiques. Elle semble ainsi cantonnée aux fibroblastes et cellules musculaires lisses à travers les différents tissus, à l'exception des testicules où *FOXS1* est exprimé fortement dans les cellules de Sertoli et les cellules péri-tubulaires ainsi que plus faiblement dans les cellules de Leydig et les spermatozoïdes (figure 13B).

L'expression de *FOXS1* dans les testicules a de plus a été observée chez la souris lors du développement embryonnaire, notamment dans les cellules péri-endothéliales. Le *knock-out* de *FOXS1* dans ces souris n'entraîne pas de mortalité ou de stérilité mais induit l'accumulation de sang dans les testicules fœtaux dus à l'apoptose de cellules péri-endothéliales provoquant des fissures de quelques centaines de nanomètres dans les vaisseaux sanguins. Les auteurs de cette découverte supposent que *FOXS1* protégerait ces cellules de l'apoptose en bloquant les sites de fixation reconnu par *FOXO3* et *FOXO4* et ainsi leur induction de *FasL*.¹⁷⁸ Son action dans les fibroblastes a également été partiellement étudiée. *FOXS1* semble en effet un facteur de transcription essentiel dans la différenciation des fibroblastes en myofibroblastes.¹⁷⁹ Ceci peut expliquer son expression relativement faible dans les fibroblastes et plus importante dans les cellules des muscles lisses présentant un phénotype proche des myofibroblastes. En plus de ces articles, l'impact de *FOXS1* a surtout été étudié au cours du développement embryonnaire. Il participe notamment à la formation du système nerveux central et est exprimé au cours du développement embryonnaire de la souris par une sous-population de cellules de la crête neurale qui vont migrer au niveau des somites antéro-dorsaux.^{180,181}

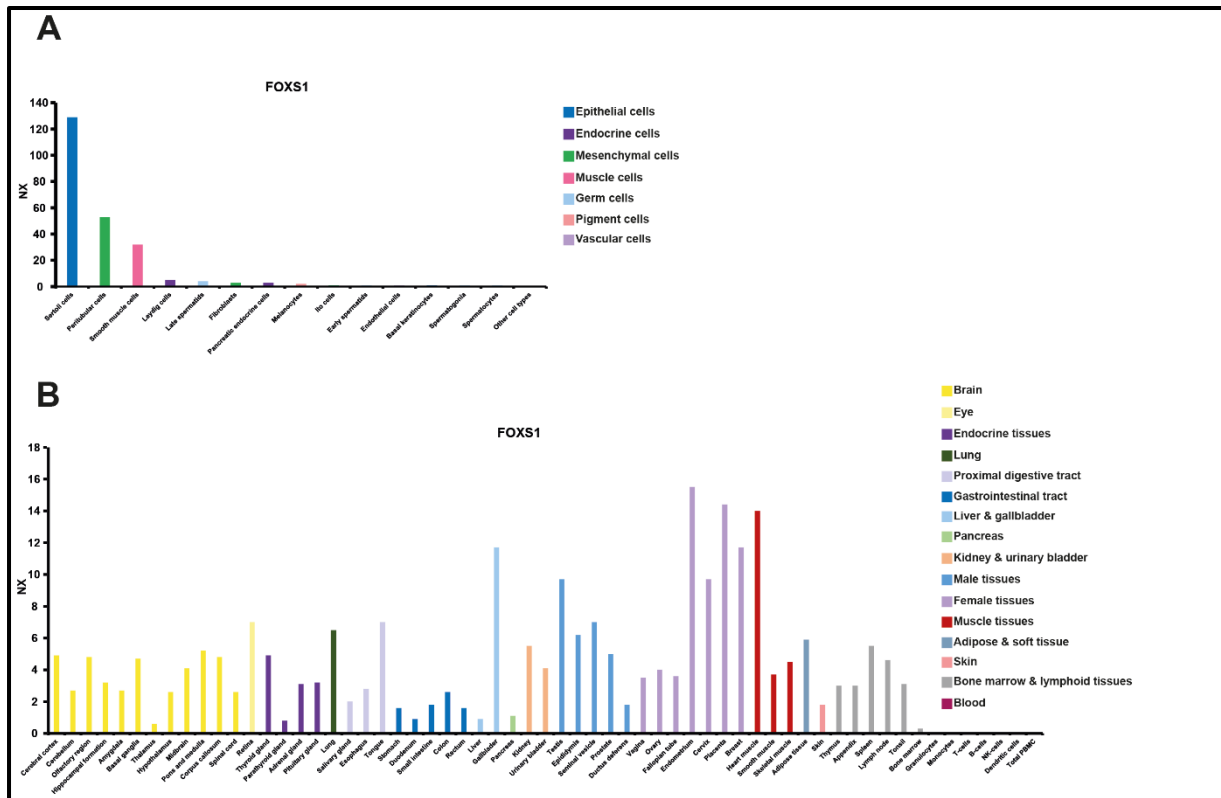


Figure 13 : Expression de FOXS1 chez l'humain.

A : Expression dans les tissus par RNA-seq. B : Expression dans les types cellulaires spécifiques par single-cell RNA-seq. Données provenant de l'Human Protein Atlas.

Au cours de ma thèse, l'impact de FOXS1 en cancérologie et notamment dans le cancer gastrique et le CHC a commencé à être étudié, sans qu'un consensus soit toutefois établi sur son rôle (Figure 14). Le tout premier papier mentionnant FOXS1 dans un contexte cancéreux remonte à 2015 et établit seulement que son promoteur est hypométhylé et son expression augmentée dans les tumeurs adrenocorticales murines induites par gonadectomie.¹⁸² Un an plus tard, un article dans un journal non référencé sur Pubmed établit que l'expression de *FOXS1* est diminuée dans le CHC et est un marqueur de bon pronostic. En 2018, l'auteur de cet article en publie un second se concentrant sur l'action de FOXS1 dans le cancer gastrique et montre une diminution d'expression de *FOXS1* dans les tumeurs de l'estomac. Cet article étudie également l'effet de la surexpression de *FOXS1* et conclue à un effet suppresseur de tumeur, avec un rôle antiprolifératif et inhibiteur de l'EMT, possiblement en agissant sur la signalisation Wnt.¹⁸³ La première conclusion de cet article s'oppose à celle de deux articles précédents dans le cancer gastrique. La première identifie *FOXS1* avec 6 autres gènes comme marqueurs de mauvais pronostic pour la prédiction de la survie des patients atteints de cancer gastrique et la seconde observe une augmentation d'expression de *FOXS1* dans les tumeurs gastriques.^{184,185} De plus, un autre article vient peu après également contredire le rôle anti-tumoral de FOXS1 dans le cancer gastrique (Figure 14). Cette étude plus complète avance que *FOXS1* est davantage exprimé dans les tumeurs gastriques et associé à l'agressivité et à un mauvais pronostic. Elle montre également à l'opposé du premier article que FOXS1 favorise la prolifération et l'EMT. Cette article fait également apparaître une régulation complexe de *FOXS1*, impacté négativement et indirectement par NF-κB1 au niveau transcriptionnel et par miR-125a-5p au niveau post-transcriptionnel.¹⁸⁶ Cette étude valide également la régulation de *FOXS1* par GLI1 et la signalisation Hedgehog dans le cancer gastrique, régulation déjà mise en évidence dans le médulloblastome où FOXS1 est induit par GLI1 et agit comme rétrocontrôle en l'inhibant en retour, réduisant ainsi ses effets pro-oncogéniques..¹⁸⁷

Beaucoup reste à éclaircir sur FOXS1, et si son action physiologique, bien qu'assez peu étudiée, ne soit pas sujette à débat, ce n'est clairement pas le cas de son rôle en cancérologie. A l'instar de nombreuses protéines FOX, FOXS1 semble capable d'agir comme oncogène ou suppresseur de tumeurs. Qui plus est, l'ensemble des études qui se sont intéressées à son action en

cancérologie (8 papiers) et d'autant plus dans le CHC étaient relativement simple et n'ont qu'effleuré la surface de l'action et la régulation de FOXS1. Ainsi, personne n'a encore établi de lien entre cette protéine et la signalisation TGF β , pourtant extrêmement importante en cancérologie. Cette thèse aura donc pour vocation à caractériser le rôle de FOXS1 aussi bien dans la cancérogenèse hépatique et la progression tumorale du CHC que dans la signalisation TGF β et particulièrement sur sa dualité fonctionnelle.

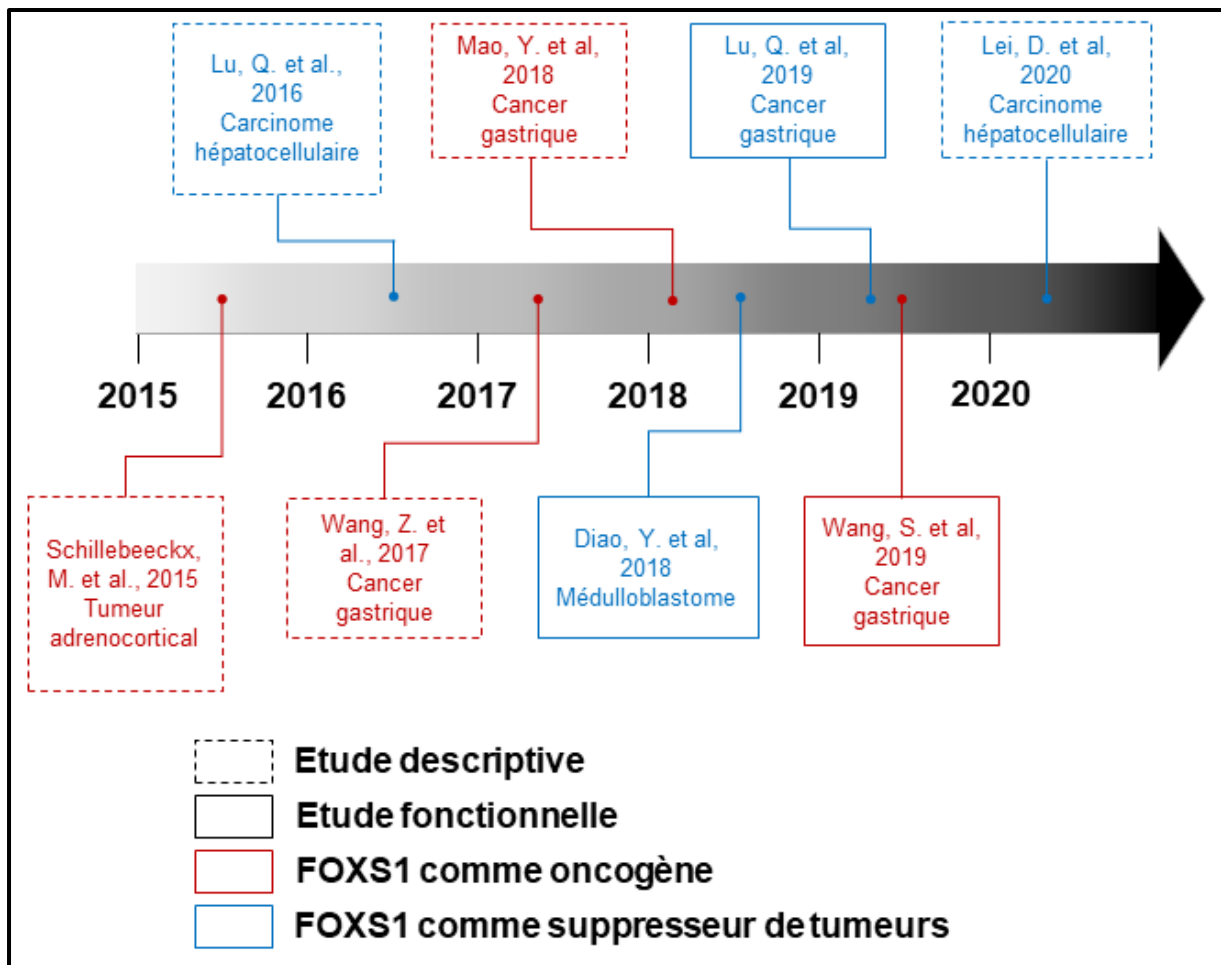


Figure 14 : Historique des études du rôle de FOXS1 en cancérologie.

Chapitre 2 : Rationnel du projet de thèse

Le CHC est la tumeur primitive du foie la plus fréquente et est responsable de la mort de plusieurs centaines de milliers d'individus chaque année. Qui plus est, et contrairement à la majorité des autres cancers, son incidence a continué d'augmenter au cours des dernières décennies, avec plusieurs centaines de milliers de nouveaux cas recensés chaque année. Ce cancer meurtrier manque de plus de stratégies thérapeutiques efficaces pour traiter les patients atteints de formes avancées de la maladie. Malgré les nombreux essais cliniques réalisés, seul le sorafenib et le lenvatinib ont montré une efficacité mais hélas sans espoir curatifs. Les immunothérapies, extrêmement efficaces dans d'autres cancers, n'ont également pas eu les résultats espérés. A l'heure où la stratégie de lutte anti-cancéreuse s'oriente vers la thérapie combinée et la médecine personnalisée, il apparaît indispensable d'identifier des cibles thérapeutiques qui pourront agir en synergie et d'identifier des biomarqueurs de sélection et de réponses aux traitements.

Au regard de ces critères, la signalisation TGF β semble une candidate extrêmement intéressante à cibler. En effet, le TGF β intervient à différents niveaux de l'apparition et de la progression du CHC. Il agit sur le microenvironnement tumoral en activant les cellules étoilées du foie et favorisant la synthèse et le dépôt de composants de la matrice extracellulaire par ces cellules. Ce processus est connu pour favoriser la cancérogenèse et le développement tumoral via la dérégulation de signalisations cellulaires contrôlées par la matrice (piégeage et relargage de cytokines, signalisation médiées par les connexions intégrines/collagènes, etc). Le TGF β est également connu pour son rôle immunosuppresseur, jouant un rôle important dans l'évasion immunitaire des tumeurs. L'environnement immunosuppresseur des tumeurs hépatiques étant suspecté d'être la cause principale de l'échec des immunothérapies contre le CHC, le traitement combiné d'inhibiteurs du TGF β avec les immunothérapies apparaît comme une combinaison évidente et est actuellement testée en essai clinique de phase I/II (NCT02423343, NCT03631706).

La complexité du traitement anti-TGF β réside dans son action beaucoup plus ambivalente sur les cellules cancéreuses. En effet, le TGF β présente deux effets opposés sur les tumeurs

dépendant principalement de la progression tumorale. Sur les CHC précoces, le TGF β induit un effet cytostatique en induisant l'expression de gène de contrôle du cycle cellulaire comme *CDKN1A* et *CDKN1B* et en bloquant des gènes prolifératifs comme *MYC* voir en induisant l'apoptose via l'induction de *NOX4*. Toutefois, au cours de la progression tumorale, les cellules cancéreuses deviennent résistantes à cette action cytostatique et détournent le TGF β vers une action pro-métastatique via l'induction de gènes impliqués dans l'EMT comme *SNAI1/2* et *ZEB1/2*. Ce détournement de la signalisation TGF β existe dans de nombreux cancers, mais les mécanismes qui le contrôlent dans le CHC n'ont pour le moment été que partiellement élucidés. La réponse au TGF β étant extrêmement dépendante du contexte transcriptionnel de la cellule dans laquelle elle a lieu, l'identification de nouveaux effecteurs de la signalisation TGF β dans ce contexte tumorale pourrait permettre de mieux comprendre les mécanismes détournant la signalisation TGF β vers une action pro-métastatique. Cela permettra également d'identifier des biomarqueurs de l'effet pro-métastatique du TGF β permettant de sélectionner les patients pouvant bénéficier de thérapies ciblées avec des inhibiteurs du TGF β .

En cela, FOXS1 s'avère un candidat très intéressant, de par son induction forte par le TGF β qui n'a jamais été rapportée dans la littérature et son action dans le cancer peu étudiée mais également du fait qu'il est membre de la famille des *forkhead box*, des facteurs de transcription connus pour leur capacité à orienter les réponses cellulaires et pour avoir un impact positif comme négatif en cancérologie.

Objectifs du projet

Mon projet de thèse vise donc à caractériser de rôle de FOXS1 dans la cancérogenèse hépatique et la signalisation TGF β . Cela consiste dans un premier temps à analyser l'expression de *FOXS1* dans le CHC, sa localisation cellulaire et sa régulation par le TGF β ainsi que d'autres mécanismes potentiels. Dans un second temps, il s'agit de comprendre le rôle de FOXS1 dans les cellules cancéreuses, ainsi qu'au regard de la voie TGF β (s'il oriente la réponse cellulaire vers une des actions duales du TGF β), ainsi que les mécanismes par lesquels il agit. Enfin dans un dernier temps, il faut étudier la pertinence clinique de l'expression de FOXS1 dans le CHC, et notamment sa capacité à servir de biomarqueur de bon ou mauvais pronostic ainsi que biomarqueur pour identifier les tumeurs où la signalisation TGF β est devenue pro-métastatique.

Chapitre 3 : Résultats des travaux de thèse

I) Expression, localisation et régulation de FOXS1

1) Identification de FOXS1 par analyse transcriptomique

La signature TGF β tardive identifiée en 2008 est, comme nous l'avons vu, associée à un mauvais pronostic dans le CHC et peut-être utilisée pour discriminer les lignées invasives de CHC, notamment les SNU-423, SNU-449 et SNU-475.⁵⁴ Ces lignées ont également été identifiées comme invasives et peu différenciées (sous-groupe CL3) dans la classification transcriptomique des lignées cancéreuses hépatiques établie par le zucmanlab en 2019.¹⁸⁸ J'ai donc utilisé la lignée SNU-449 comme paradigme pour identifier de nouveaux gènes cibles du TGF β pouvant propager son action pro-métastatique et favoriser le phénotype invasif de ces lignées. Cette lignée a été traitée ou non au TGF β 1 à 1ng/mL pendant 72 heures et les gènes différentiellement exprimés ont été étudiés par puce à ARN. 80 gènes ont ainsi été identifiés par des critères stricts (p-value < 0.001 et variation > 2 fois) (figure 15A, table 1). Cette signature contient nombres de gènes déjà bien connus pour être des cibles du TGF β comme *SKIL*, *COL1A1*, *LTBP2*, *BMF*, *TGFBI*, *PDGFA* ainsi que *FOXS1* qui, bien que jamais caractérisé comme cible du TGF β , est l'un des gènes les plus induits de cette analyse (augmentation par dix de l'expression). L'induction de *FOXS1* a été validée par RT-qPCR dans les SNU-449 ainsi que d'autres gènes de la signature (*SKIL*, *COL1A1* et *BMF*, figure 15B).

2) Expression de FOXS1 dans les lignées de CHC

Après avoir validé l'induction de *FOXS1* par le TGF β dans les SNU-449, j'ai étudié le temps nécessaire à cette induction, en réalisant différents temps de traitement de ces cellules au TGF β (figure supplémentaire 1A). Il s'avère que *FOXS1* est induit précocement par le TGF β , avec une augmentation d'expression dès 2 heures de traitement, et un pic d'induction après 8 heures de stimulation. A partir de là, j'ai estimé qu'un traitement de 16 heures était optimal pour observer l'induction de *FOXS1* et j'ai effectué cette stimulation sur un panel de 8 lignées

cellulaires de CHC représentatives de l'hétérogénéité tumorale du CHC (table 2) ainsi que sur des hépatocytes humains en culture primaire pour observer si le TGF β pouvait induire l'expression de *FOXS1* dans un contexte non tumoral (figure 15C). L'analyse de l'expression de *FOXS1* dans ces lignées aussi bien au niveau protéique que transcriptomique révèle que seule une partie des lignées de CHC exprime *FOXS1* et qu'il n'est inductible par le TGF β que dans ces lignées. *FOXS1* n'est ainsi ni exprimé ni induit dans les lignées associées à la signature TGF β précoce (PLC/PRF/5, HepG2-C3A, HuH6 et Hep3B). *FOXS1* n'est de plus ni exprimé ni induit par le TGF β dans les hépatocytes, suggérant que *FOXS1* est possiblement ré-exprimé dans les tumeurs au cours de la cancérogenèse. L'analyse de l'expression de *FOXS1* dans un plus grand panel de lignées cellulaires à partir des données de la *liver cancer cell line database* confirme ces résultats (figure 15D). *FOXS1* est différentiellement exprimé dans les lignées de CHC et davantage retrouvé dans les groupes transcriptomiques CL2 et CL3, eux-mêmes associés aux sous-groupes S1 d'Hoshida et G3 de Boyault, des sous-groupes associés à la signature TGF β tardive.¹⁸⁸ L'expression et l'induction TGF β -dépendante de *FOXS1* a également été étudié dans d'autres lignées cellulaires : de cellules étoilées hépatiques (LX2), de cholangiocarcinomes (HuCCT1, Huh28), de cancer du rein (HK2) et d'adénome pancréatique (PANC-1). Toutes ces lignées expriment *FOXS1* à un niveau basal et induisent son expression sous l'action du TGF β , montrant que le contrôle de *FOXS1* par la signalisation TGF β dépasse le cadre du carcinome hépatocellulaire et est induit dans d'autres tumeurs, ainsi que dans le microenvironnement tumoral du CHC où le TGF β joue un rôle essentiel (figure 15E). Enfin j'ai validé l'expression et l'induction TGF β -dépendante de *FOXS1* directement sur des échantillons de tumeurs de CHC réséqués. Des coupes de tumeurs ou de tissus non tumoraux encerclant la tumeur ont été traités au TGF β ou au LY2157299 (galunisertib) pendant 16 heures et l'expression de *FOXS1* et d'autres gènes cibles du TGF β ont été évalués par RT-qPCR (figure 15F, figure supplémentaire 2). Dans ce modèle, *FOXS1* est également induit par le TGF β à la fois dans le tissu tumoral et non tumoral de la même manière que d'autres gènes cibles du TGF β tel que *SMAD7* ou *SERPINE1*.

3) Localisation de FOXS1

FOXS1 est un facteur de transcription dont la localisation est nucléaire selon la base de données Uniprot. Afin de m'en assurer, j'ai étudié la localisation de *FOXS1* dans les SNU-449 par

immunofluorescence (figure 15G) et dans les SNU-449 et SNU-475 par western blot après fractionnement des protéines nucléaires et cytoplasmiques (figure 15H). *FOXS1* est assez faiblement détecté par immunofluorescence mais on peut toutefois observer un signal diffus dans l'ensemble de la cellule et plus important au niveau du noyau. Cette observation est confirmée par l'analyse en western blot, qui montre que *FOXS1* est présent dans le cytoplasme mais enrichi dans les fractions nucléaires des SNU-449 et SNU-475 à la fois en condition contrôle et après induction par le TGF β . *FOXS1* est donc bien essentiellement un facteur nucléaire comme indiqué dans les bases de données.

4) Régulation de *FOXS1* par les SMAD

FOXS1 est, comme nous l'avons vu, rapidement induit par le TGF β , dès 2h après stimulation (figure supplémentaire 1A). Pour confirmer que *FOXS1* est une cible précoce du TGF β , un traitement au cycloheximide, un inhibiteur de la traduction, a été réalisé sur les SNU-449 en présence ou absence de TGF β . L'expression de *FOXS1* et d'autres gènes cibles du TGF β (*SMAD7*, *COL1A1*, *NANOG* et *ITGB6*) a ensuite été analysée par RT-qPCR (figure supplémentaire 1B). Les résultats montrent que malgré une diminution de son expression, *FOXS1* est toujours induit par le TGF β après blocage de la traduction suggérant que *FOXS1* est induit directement par le TGF β et non en seconde ligne après une première traduction des gènes induits par le TGF β . Étonnamment, *SMAD7*, une cible directe des SMAD n'est plus induit par le TGF β après traitement au cycloheximide. *SMAD7* nécessitant le complexe AP-1 pour son induction par les SMAD, il est possible que la traduction de certains membres de ce complexe soit bloquée par le cycloheximide, empêchant ainsi l'induction de *SMAD7* par le TGF β .

Ensuite afin d'étudier si l'induction de *FOXS1* dépend de la signalisation canonique des SMAD, une expérience d'inhibition de la signalisation par le LY2157299 dans les SNU-449 et SNU-475 a été réalisée. Le LY2157299 inhibe le récepteur TGFBR1, empêchant ainsi la phosphorylation de SMAD2/3 et inhibant de ce fait la signalisation canonique. L'analyse par RT-qPCR de l'expression de *FOXS1* et *SERPINE1* après traitement au LY2157299 avec ou non stimulation par le TGF β montre que le LY2157299 diminue à la fois l'expression basale de *FOXS1* (et de *SERPINE1*) et bloque également son induction par le TGF β , à la fois dans les SNU-449 et SNU-475 (figure 16A). Pour renforcer ce résultat, j'ai également inhibé la signalisation canonique par extinction de SMAD3, SMAD4 ou les deux en même temps via la

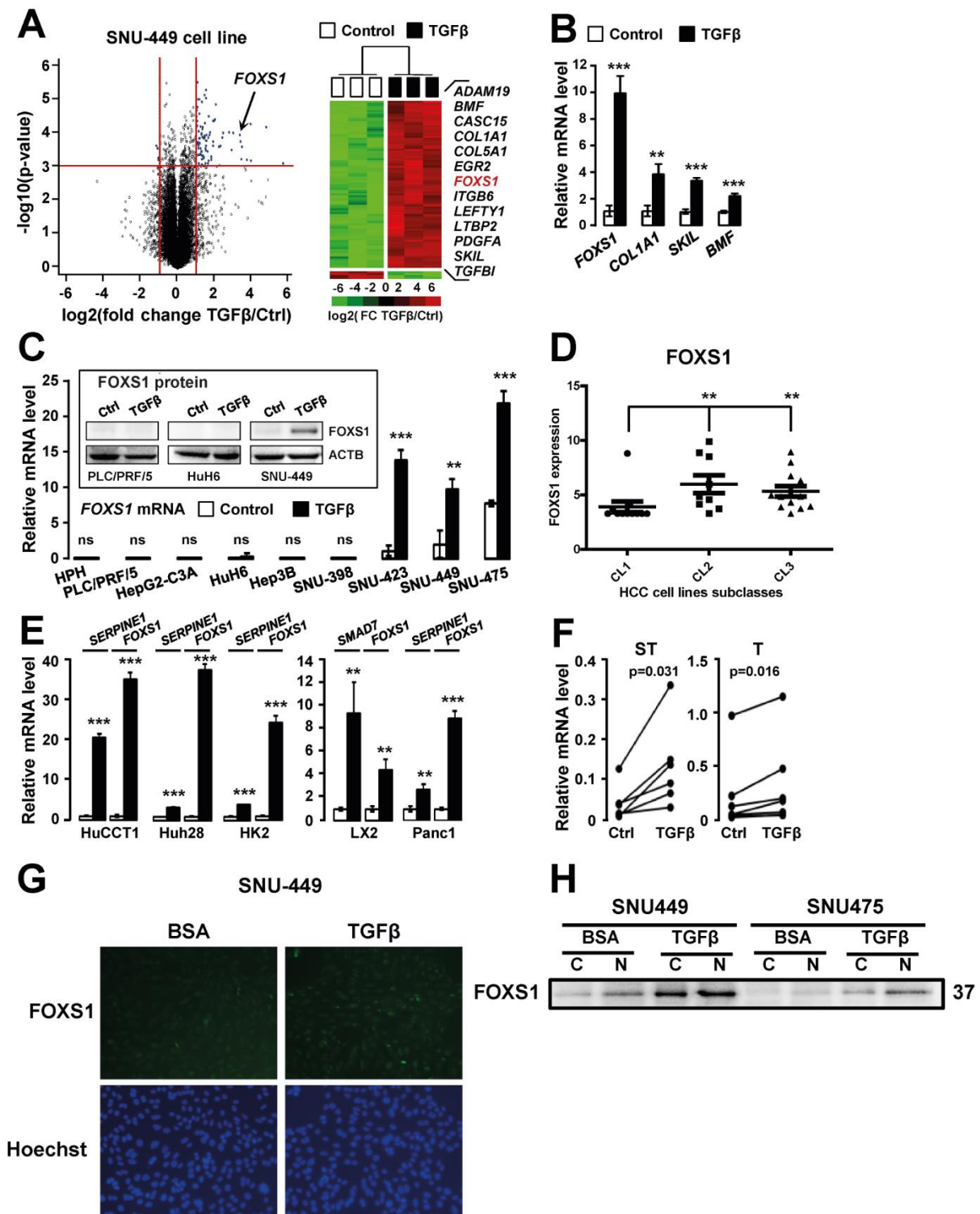


Figure 15 : Expression et localisation de FOXS1

A : Volcano plot et classification hiérarchique des gènes différentiellement exprimés par le TGF β (1ng/mL, 72h) dans la lignée de CHC SNU-449. La ligne horizontale indique une p-

value < 0,001 et la ligne verticale un changement d'expression > 2. La liste complète des gènes est fournie dans la table 1.

B : Niveau d'expression relatif de *FOXS1* et d'autres gènes cibles connus du TGFβ, obtenue par RT-qPCR dans les SNU-449 après 72h de traitement en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL. L'analyse statistique a été réalisée par un t.test (n=3)

C : Analyse par RT-qPCR et western blot de l'expression de *FOXS1* dans des hépatocytes humains en culture primaire ainsi que dans 8 lignées de CHC représentatives des tumeurs humaines après 16h de culture en présence (barre noire) ou absence (barre blanche) de TGFβ à 1ng/mL. L'analyse statistique a été réalisée par un t.test (n=3)

D : Expression de *FOXS1* dans 31 lignées de CHC selon leur groupe transcriptomique établi par Caruso et al. Données provenant- du *Liver Cancer Cell Line database*. L'analyse statistique a été réalisée par une one-way ANOVA.

E : Analyse de l'expression de *SERPINE1* et *FOXS1* dans les lignées HuCCT1 et Huh28 de cholangiocarcinome (1ng/mL de TGFβ pendant 72h), HK2 de cellules tubulaires proximales du rein (GDS3649, 2ng/mL de TGFβ pendant 72h), LX2 de cellules étoilées du foie (1ng/mL de TGFβ pendant 72h) et Panc1 d'adénocarcinome pancréatique (GSE23952, 5ng/mL de TGFβ pendant 72h). L'analyse statistique a été réalisée par un t.test (n=3)

F : Analyse par RT-qPCR de l'expression de *FOXS1* dans le tissu non tumoral entourant la tumeur (ST, panel de gauche ; n=6 patients) ou dans les tumeurs (T, panel de droite, n=6 patients). Echantillons provenant de coupes de tissus de patients après résection chirurgicale. Le tissu a été incubé pendant 48h en présence (barre noire) ou en absence (barre blanche) de TGFβ à 1ng/mL. L'analyse statistique a été réalisée par un test apparié non-paramétrique de Wilcoxon.

G : Analyse par immunofluorescence de la localisation de FOXS1 dans les SNU-449 traitées (panel de droite) ou non (panel de gauche) au TGFβ à 1ng/mL pendant 16h. Après fixation au paraformaldehyde 4%, les cellules ont été incubées pendant 16h avec l'anticorps anti-FOXS1 (Novus Biological) et les noyaux ont été marqués au Hoechst.

H : Analyse de l'expression de *FOXS1* après stimulation ou non par le TGFβ à 1ng/mL pendant 16h par western blot dans les lignées SNU-449 et SNU-475 après fractionnement des protéines nucléaires et cytoplasmiques à l'aide du kit NE-PER™ (ThermoFisher : 78833).

transfection de siRNA contre leur ARNm. L'extinction s'est avérée efficace (figure 16B) et induit une diminution de l'induction par le TGF β de *FOXS1* et *SERPINE1* en cas d'extinction d'un des deux SMAD ou une abolition totale de l'induction en cas d'extinction de SMAD3/4 (figure 16B). Pris ensemble, ces résultats prouvent que, comme *SERPINE1*, *FOXS1* est une cible canonique de la signalisation TGF β .

Je me suis ensuite intéressé à savoir si l'induction de *FOXS1* se faisait par fixation directe des SMAD sur son promoteur proximal. Afin d'étudier cela, j'ai utilisé un promoteur rapporteur constitué du gène de la gaussia luciférase (GLuc) sous le contrôle du promoteur proximal de *FOXS1* (les 1000 paires de bases en amont du site d'initiation de la transcription). Le système contient également un contrôle interne, la *secreted alkaline phosphatase*, exprimé de manière constitutive pour normaliser le signal luciférase. Ce système a été transfecté dans des SNU-449 traitées ou non au TGF β et au LY2157299. Les résultats ne montrent aucun impact des traitements sur l'activité luciférase qui reste à niveau très faible comparé au contrôle positif de l'expérience, mais néanmoins suffisamment élevé par rapport au contrôle négatif sans transfection de la construction pour affirmer qu'il ne s'agit pas de problème de transfection (figure 16C). Ces résultats indiquent donc qu'il ne semble pas y avoir de fixation des SMAD sur le promoteur proximal de *FOXS1*.

Afin de creuser un peu plus la potentielle régulation de *FOXS1* par les SMAD, je suis allé étudier les bases de données de chIP-seq pour identifier les sites de fixation de SMAD2/3 dans différents modèles cellulaires. Les données de chIP-seq dans les SNU-449 ou SNU-475 étant extrêmement limitées, j'ai utilisé les données d'autres modèles cellulaires que j'avais validé comme étant des lignées *FOXS1* positives (PANC-1, LX2, HuCCT1 ainsi que des données sur les *hepatic stellate cells* et le CHC). Il existe davantage de données de chIP-seq sur les lignées hépatiques *FOXS1* négatives et j'ai ainsi pu utiliser HepG2, Huh7, SNU-398 et des hépatocytes comme lignées *FOXS1* négatives. L'analyse des sites de fixation de SMAD2/3 dans les HepG2, LX2 et PANC-1 confirme les résultats du promoteur rapporteur : aucun pic correspondant à la fixation de SMAD au niveau du promoteur proximal de *FOXS1* n'est observé (figure 16D). Cette observation a également été validée par chIP-PCR contre SMAD2/3 dans les SNU-449 où aucun enrichissement n'est retrouvé sur le promoteur (figure 16D-E). Cependant, les résultats de chIP-seq montrent deux sites de fixation de SMAD2 et SMAD3 en aval du locus

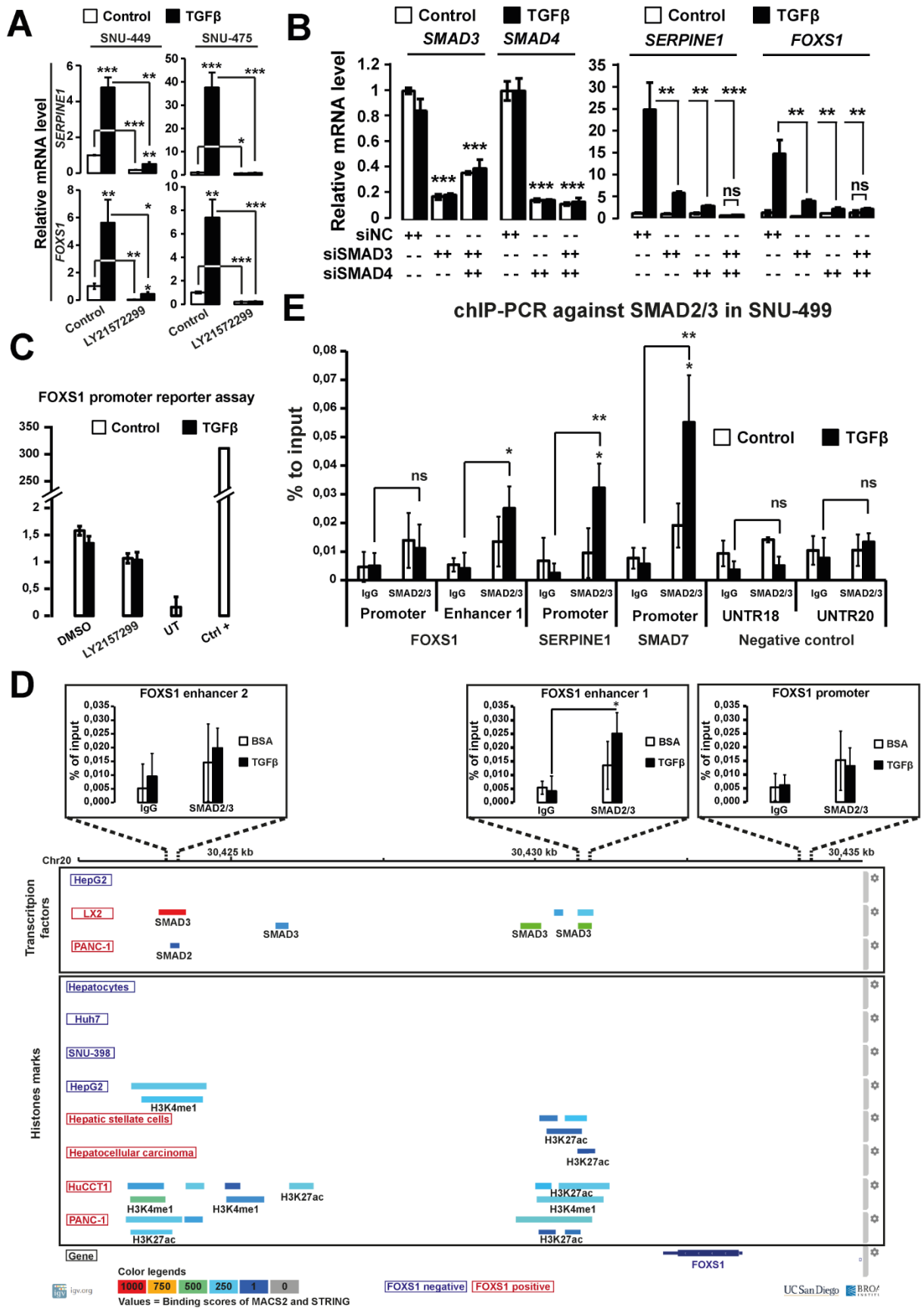


Figure 16 : FOXS1 est régulé par la signalisation canonique du TGFβ.

A : Analyse par RT-qPCR de l'expression de *SERPINE1* et *FOXSI* dans les lignées SNU-449 et SNU-475 après culture en présence (barre noire) ou absence (barre blanche) de TGFβ pendant 16h et en présence ou absence de 10μM de LY2157299. L'analyse statistique a été réalisée par un t.test (n=3)

B : Analyse par RT-qPCR de l'expression de *SERPINE1* et *FOXSI* dans la lignée SNU-449 transfectée par un siSMAD3 et un siSMAD4, seul ou en combinaison (50nM chacun) ou un siRNA contrôle (50nM) après 72h de culture en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL.

C : Analyse luciférase de l'activité du promoteur proximal de *FOXSI*. La lignée SNU-449 a été transfectée par un promoteur rapporteur correspondant aux 1000 paires de bases en amont du TSS de *FOXSI* et contrôlant l'expression de la GLuc. Les cellules ont ensuite été traitées en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL ou de LY2157299 à 10μM pendant 16h. Du milieu EF1A-PG04 est utilisé comme contrôle positif et des cellules non transfectées comme contrôle négatif.

D : Analyse du locus de *FOXSI* à l'aide de l'*integrative genome viewer*. Les données de chIP-seq proviennent de la base de données chIP-atlas. Les lignées cellulaires ont été choisies pour avoir été étudiées en chIP-seq au moins sur SMAD2 ou SMAD3 pour les facteurs de transcription et H3K27ac pour les marques histones.

E : ChIP-PCR contre SMAD2/3 dans les SNU-449. La lignée SNU-449 a été cultivée en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL pendant 16h. La fixation de SMAD2/3 sur le promoteur et les possibles enhancers de *FOXSI* a été étudiée par qPCR. La fixation sur les promoteurs de *SERPINE1* et *SMAD7* a été étudiée comme contrôle positif de la chIP et l'absence de fixation sur deux régions non transcrites du génome comme contrôle négatif. L'analyse statistique a été réalisée par un t.test (n=3)

de *FOXS1* (à 3kb et 10kb du TSS) dans les LX2 et PANC-1 qui ne sont pas retrouvés dans les HepG2. L'analyse des marques histones H3K27ac et H3K4me1 (celle-ci uniquement dans les hépatocytes, Huh7, HepG2, HuCCT1 et PANC-1) dans les lignées *FOXS1* négatives : Hépatocytes, Huh7, SNU-398, HepG2 et les lignées *FOXS1* positives : HSC, HCC, HuCCT1 et PANC-1 montrent que les marques histones d'enhancers actifs (H3K4me1 et H3K27ac) sont retrouvées au niveau des sites de fixation des SMAD dans les lignées *FOXS1* positives mais non dans les lignées *FOXS1* négatives. L'analyse par chIP-PCR contre SMAD2/3 dans les SNU-449 montre un enrichissement des SMAD au niveau de l'enhancer à 3kb du site d'initiation de la transcription (enhancer 1) mais pas au niveau de l'enhancer à 10kb (enhancer 2). Les résultats de cette chIP-PCR ont été validés via des contrôles positifs avec la fixation de SMAD2/3 sur les promoteurs de *SERPINE1* et *SMAD7* et des contrôles négatifs avec l'absence de fixation sur des régions non transcrites du génome (figure 16E). Ces résultats permettent de poser l'hypothèse que l'induction de *FOXS1* par le TGF β passe par la fixation des SMAD sur un enhancer à 3kb du site d'initiation de la transcription de *FOXS1* et que cette fixation est impossible dans les lignées *FOXS1* négatives du fait de l'absence de marques épigénétiques d'enhancers actifs au niveau du site de fixation, expliquant l'absence d'induction de *FOXS1* par le TGF β dans ces lignées. Ces résultats nécessitent toutefois validation pour conclure à une régulation de *FOXS1* par ce mécanisme.

5) Régulation de *FOXS1* par méthylation

S'il est indubitable que les SMAD exercent un contrôle transcriptionnel (direct ou indirect) sur *FOXS1*, d'autres mécanismes peuvent intervenir pour réguler son expression. La méthylation du promoteur reste ainsi un des mécanismes de régulation incontournables dès lors que l'on s'intéresse à la régulation transcriptionnelle. Afin d'étudier si la méthylation du promoteur de *FOXS1* est impliquée dans la régulation de son expression dans le CHC, je suis allé étudier les bases de données contenant des informations sur le méthylome du CHC. A l'aide de Wanderer, un logiciel permettant l'exploration des données de méthylation et d'expression du TCGA, j'ai pu étudier l'état de méthylation du locus de *FOXS1* et son niveau d'expression dans 50 tissus non tumoraux et 256 tumeurs de CHC (figure 17A). Le niveau de méthylation du locus de *FOXS1* est ainsi relativement élevé et homogène dans le tissu non tumoral tandis qu'à l'inverse, il est extrêmement hétérogène dans les tumeurs de CHC. Le niveau moyen de méthylation pour

chacune des 8 sondes sur le locus de *FOXSI* est ainsi significativement inférieur dans les tumeurs par rapport au tissu sain, et à l'inverse, le niveau d'expression de *FOXSI* est significativement supérieur dans les tumeurs par rapport au tissu sain, suggérant que la méthylation du locus pourrait être un mécanisme d'inhibition de l'expression de *FOXSI* dans le foie non tumoral. Il est également important de noter que la variabilité de son expression et de la méthylation de son locus dans les tumeurs est concordante avec le fait qu'il ne semble être exprimé que dans certains sous-types de CHC correspondant aux S1 de Hoshida et aux G3 de Boyault comme nous l'avons vu précédemment.

Ces observations ont également été validées dans une autre cohorte de CHC provenant cette fois de la base de données GEO (*Gene expression omnibus*). Cette cohorte (accession number : GSE77276) comprend les données d'expression et de méthylation de 19 patients, au niveau du foie non-tumoral, de la tumeur primaire et des métastases de la veine porte (figure 17B). Les résultats sont ici similaires à ceux du TCGA avec une diminution significative de la méthylation pour chacune des sondes dans le foie tumoral et les métastases par rapport au foie sain et à l'inverse une augmentation d'expression de *FOXSI* dans la tumeur primaire et les métastases par rapport au foie non tumoral.

Ces analyses descriptives suggèrent donc que la méthylation de son locus est un mécanisme participant à l'inhibition de l'expression de *FOXSI* dans le foie non tumoral et dans certaines tumeurs. Afin d'étudier ce mécanisme plus en détail, j'ai également réalisé des traitements à la décitabine (5-aza-2'-deoxycytidine), un inhibiteur des méthyltransférases de l'ADN entraînant une déméthylation globale de l'ADN, sur plusieurs lignées cellulaires (SNU-449, PLC/PRF/5 et Huh6) avec ou non stimulation par le TGF β afin d'identifier si la déméthylation du locus de *FOXSI* permet d'induire son expression (figure 17C). Les résultats s'avèrent hétérogènes suivant les lignées. La décitabine n'a aucun effet sur l'expression de *FOXSI* dans les PLC/PRF/5 tandis qu'elle induit une légère diminution de l'expression de *FOXSI* dans les SNU-449. A l'inverse, dans les Huh6, la décitabine entraîne une augmentation de l'expression de *FOXSI* sans toutefois entraîner une induction par le TGF β . Ces résultats suggèrent que la méthylation du locus de *FOXSI* ne semble pas impacter sa régulation par le TGF β (qui est vraisemblablement davantage régulée par les marques histones au niveau de l'enhancer) mais joue toutefois un rôle dans l'inhibition de l'expression basale de *FOXSI* dans certaines lignées.

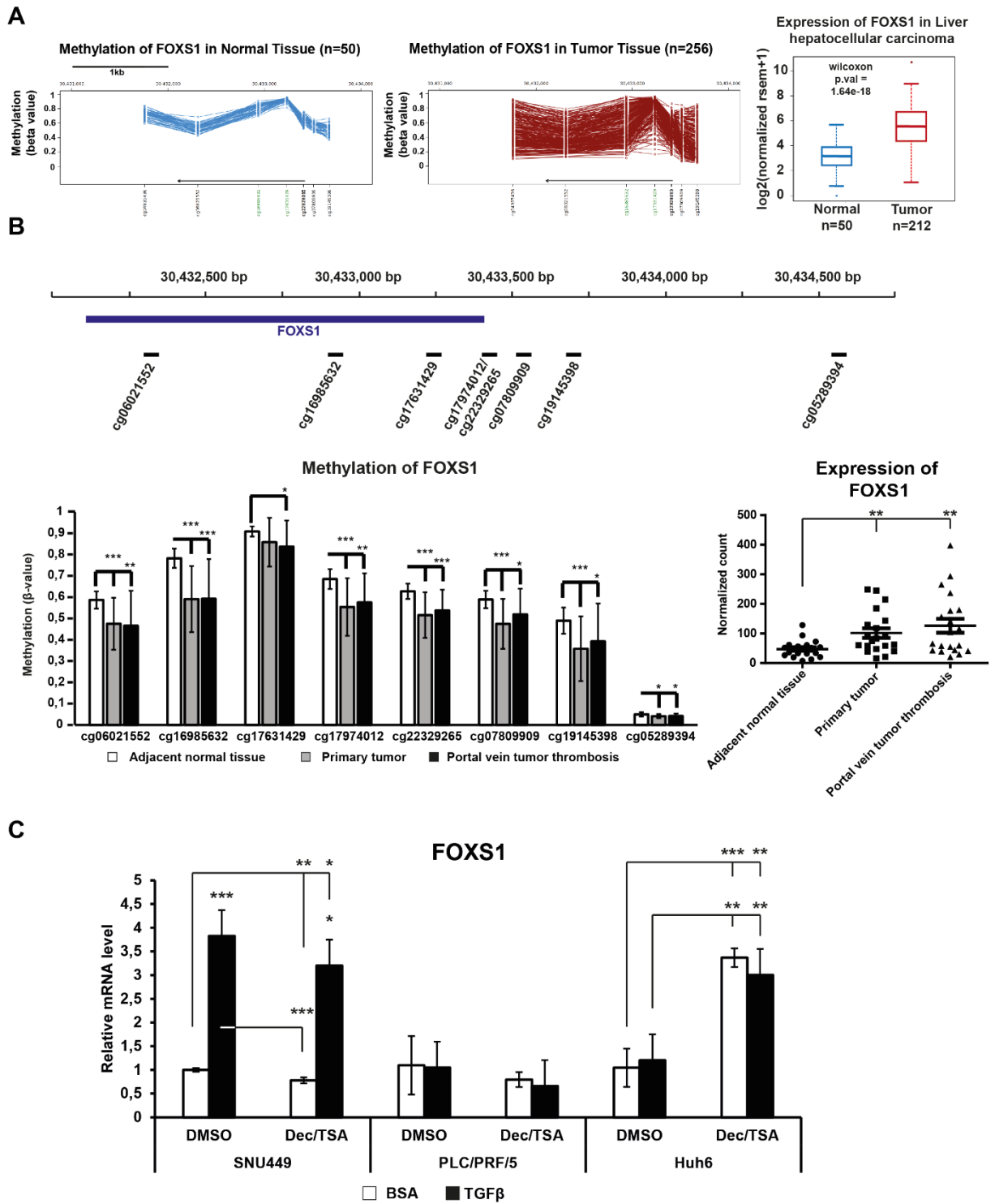


Figure 17 : L'expression de FOXS1 est impactée par la méthylation de son locus.

A : Analyse de la méthylation du locus de *FOXS1* dans le tissu sain (n=50) et les tumeurs (n=256) ainsi que de l'expression de FOXS1 (foie normal : n=50 ; tumeur : n=212) à l'aide de Wanderer et des données du TCGA. Les sondes en vert indiquent un îlot CpG.

B : Analyse de l'expression et de la méthylation du locus de *FOXSI* dans le tissu sain (n=19), les tumeurs primaires (n=19) et les métastases au niveau de la veine porte (n=19). Le panel du haut représente l'emplacement des sondes au niveau du locus de *FOXSI* (GEO : GSE77276).

C : Analyse de la déméthylation sur l'expression de *FOXSI*. Les lignées SNU-449, PLC/PRF/5 et Huh6 ont été cultivées en présence (barre noire) ou absence (barre blanche) de TGF β à 1ng/mL pendant 16h après un traitement de 24h en présence ou absence de décitabine (10 μ M) et de trichostatin A (0,5 μ M) pendant 24h. L'analyse statistique a été réalisée par un t.test (n=3).

Néanmoins, il est important de noter que le niveau d'expression dans les Huh6 après déméthylation du locus, bien que trois fois supérieur au niveau basal, reste bien inférieur au niveau basal de *FOXS1* dans les SNU-449. De même, le mécanisme d'inhibition de *FOXS1* dans les PLC/PRF/5 semble totalement indépendant de la méthylation du locus, ce qui suggère que d'autres mécanismes de régulation sont à l'œuvre pour contrôler l'expression de *FOXS1*.

6) Autres mécanismes de régulation

En plus des mécanismes épigénétiques que nous avons vu, l'expression de *FOXS1* peut être impactée directement par des facteurs de transcription, soit nécessaires à sa transcription, soit bloquant son expression. L'UCSC *Genome Browser* indique, d'après les données de l'*Open Regulation Annotation database* (OREgAnno), trois facteurs de transcription capable de se fixer au niveau du promoteur de *FOXS1* : FOXA1, EGR1 et CTCF (figure supplémentaire 3A). Ces trois facteurs s'avèrent intéressants puisque CTCF est connu pour être un répresseur transcriptionnel, EGR1 régule la réponse aux facteurs de croissance et FOXA1 est, comme nous l'avons vu, un gardien du phénotype hépatocytaire. Afin d'étudier leur rôle potentiel sur l'expression de *FOXS1*, j'ai tout d'abord étudié leur expression dans les lignées SNU-449, Huh6 et PLC/PRF/5 avec stimulation au TGF β (figure supplémentaire 3B). Les niveaux d'expression de ces différents facteurs ne varient pas considérablement entre ces trois lignées cellulaires mais FOXA1 est tout de même régulé de manière différente par le TGF β avec une induction dans les PLC/PRF/5 et une inhibition dans les SNU-449 tandis qu'à l'inverse EGR1 est induit par le TGF β dans les SNU-449 mais réprimé dans les Huh6. Pour étudier l'effet de ces facteurs de transcription sur l'expression de *FOXS1* j'ai ensuite réalisé une extinction de l'expression de chacun de ces facteurs par siRNA dans une lignée spécifique (figure supplémentaire 3C). FOXA1 étant un facteur épithélial et davantage exprimé dans les lignées *FOXS1* négatives, son inhibition a été réalisée dans les PLC/PRF/5. Elle reste toutefois sans effet sur l'expression de *FOXS1* et son induction par le TGF β suggérant soit que FOXA1 ne se fixe pas sur le promoteur de *FOXS1* dans ce modèle cellulaire, soit que son inhibition n'est pas suffisante pour activer l'expression de *FOXS1*. De la même manière, j'ai réprimé CTCF par siRNA dans les PLC/PRF/5 vu son rôle de répresseur, mais son extinction ne modifie pas non plus l'expression de *FOXS1*. Enfin EGR1 étant davantage exprimé dans les SNU-449 et de plus induit par le TGF β dans cette lignée, c'est dans cette dernière qu'a été réalisée son extinction

par siRNA. Cependant, contrairement aux deux autres facteurs, le siRNA contre EGR1 n'est que peu efficace, notamment en condition de stimulation par le TGF β qui compense largement l'extinction induite par le siRNA. Les résultats de l'extinction sont donc difficilement interprétables même s'il semblerait que l'inhibition d'EGR1 augmente l'expression de FOXS1. L'analyse de ces trois facteurs de transcription ne semble donc que peu concluante pour expliquer la régulation transcriptionnel de FOXS1.

Enfin, Wang et al, ont montré un potentiel rôle inhibiteur de NF- κ B sur l'expression de *FOXS1* dans le cancer gastrique.¹⁸⁶ NF- κ B étant généralement activé par la signalisation TNF α , j'ai réalisé un traitement au TNF α à 20ng/mL pendant 16 heures pour étudier son impact sur l'expression de *FOXS1* (figure supplémentaire 3D). Conformément aux prévisions, le TNF α est effectivement capable d'inhiber l'expression de *FOXS1* dans les SNU-449 montrant que d'autres signalisations que celle du TGF β sont capables d'influer sur l'expression de *FOXS1*.

FOXS1 est donc un facteur de transcription, essentiellement nucléaire, et exprimé dans certains sous-types de CHC. Sa régulation est complexe et se fait à différents niveaux. Il est essentiellement contrôlé par l'action de signalisations extracellulaires et notamment celle du TGF β par l'intermédiaire des SMAD, mais dépend également de régulations épigénétiques comme la méthylation de son locus. La question qui se pose ensuite est de comprendre son rôle dans les cellules cancéreuses et au sein de la signalisation du TGF β .

II) Rôle de FOXS1

1) Caractérisation des lignées FOXS1 positives

Avant d'aller creuser directement le rôle de FOXS1, je me suis intéressé aux caractéristiques des lignées FOXS1 positives par rapport aux lignées FOXS1 négatives. Nous avons déjà vu que FOXS1 est enrichi dans les groupes transcriptomiques CL2 et CL3 en relation avec la signature TGF β tardive. La comparaison des lignées FOXS1 négatives et positives montre tout d'abord une différence nette de phénotype. Les lignées PLC/PRF/5, HepG2-C3A et Huh6 exhibent un phénotype épithélial avec des cellules adhérentes entre elles et formant un pseudo-épithélium. A l'inverse, les lignées FOXS1 positives SNU-423, SNU-449 et SNU-475 exhibent un phénotype beaucoup plus mésenchymateux, avec des cellules étalées et formant beaucoup moins de jonctions entre elles (figure 18A). Ces observations ont été confirmées par l'étude de l'expression d'un marqueur épithélial (E-cadhérine) et mésenchymateux (vimentine) en western blot et immunofluorescence (figure 18E-F). Comme attendu, les lignées FOXS1 positives expriment la vimentine et extrêmement peu l'E-cadhérine, confirmant leur phénotype mésenchymateux tandis que les lignées FOXS1 négatives expriment fortement l'E-cadhérine et pas la vimentine et validant ainsi leur phénotype épithélial. Il faut tout de même noter que Hep3B présente un phénotype plus intermédiaire avec une légère expression de la vimentine en western blot suggérant que cette lignée est en EMT précoce.

Pour creuser davantage les différences entre les lignées FOXS1 positives et négatives, j'ai analysé leurs profils d'expression géniques par puce à ARN (figure 18B-C). Les lignées FOXS1 négatives expriment davantage des gènes impliqués dans le métabolisme hépatique tel que *FOXA1*, *APOE*, *AHSG* tandis que les lignées FOXS1 positives expriment davantage des marqueurs mésenchymateux comme *TGFB2*, *VIM*, *MET* ainsi que des marqueurs de dédifférenciation comme *CD44*. Ces observations sont confirmées par *gene set enrichment analysis* (GSEA) montrant un enrichissement de l'expression de gènes impliqués dans la réponse rapide au TGF β , à l'EMT, la migration cellulaire et le processus métastatique tandis que les lignées FOXS1 négatives sont enrichies en gènes exprimés dans les hépatocytes différenciés (métabolisme hépatique, jonctions cellulaires) (figure 18D). Ces résultats montrent

que l'expression de *FOXS1* corrèle avec un phénotype mésenchymateux, ce qui pointe vers une action de *FOXS1* dans l'EMT.

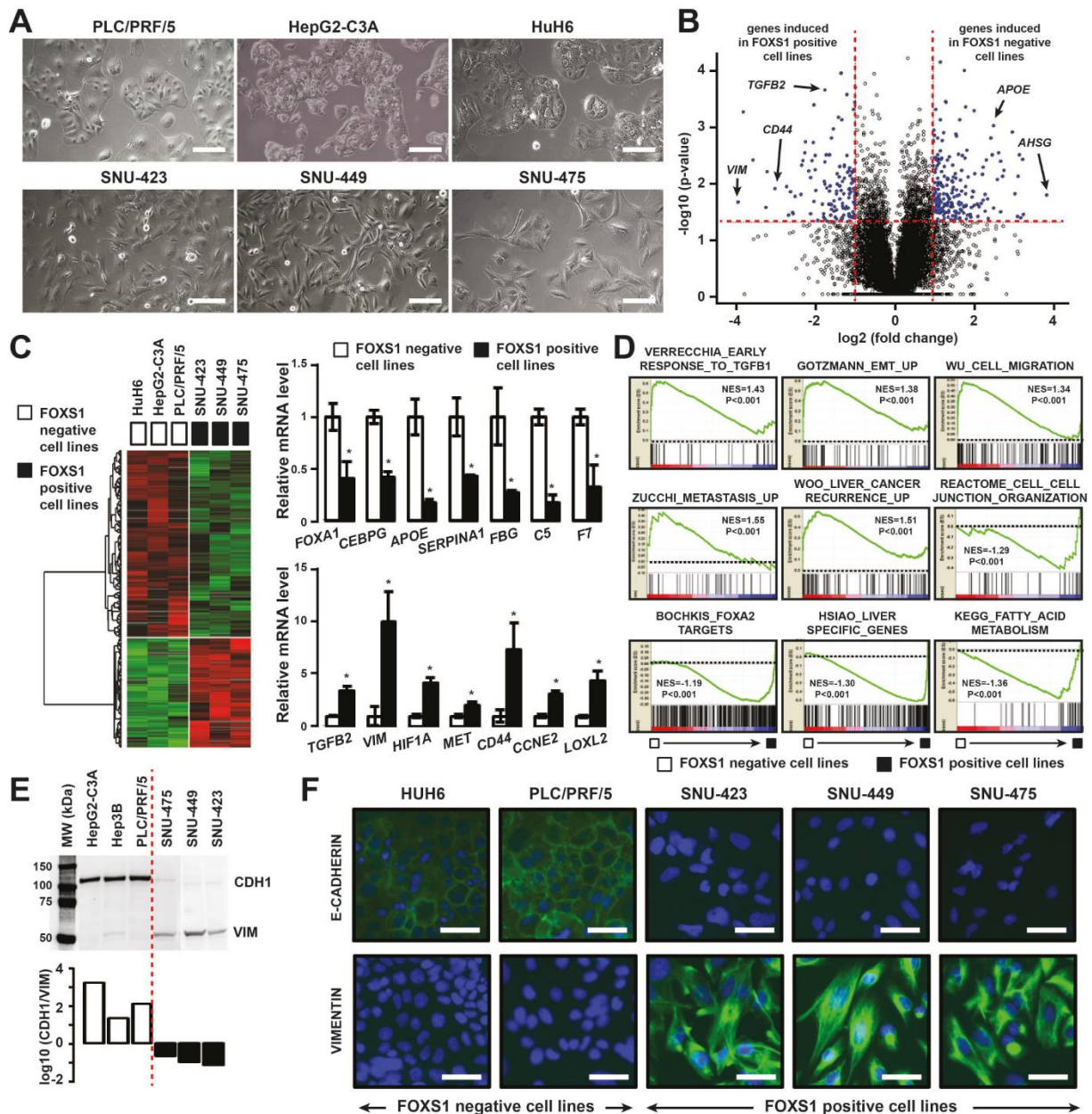


Figure 18 : *FOXS1* est exprimé dans les lignées mésenchymateuses.

A : Microscopie en contraste de phase des lignées *FOXS1* négatives (panel du haut) ou positives (panel du bas). Barre d'échelle : 100 μ m.

B : Volcano plot représentant les gènes différentiellement exprimés entre les lignées *FOXS1* négatives (PLC/PRF/5, HepG2-C3A et Huh6) et positives (SNU-423, SNU-449 et SNU-475).

La ligne hachurée horizontale indique la p-value = 0.05 et les lignes hachurées verticales un changement d'expression supérieur à 2.

C : Classification hiérarchique des gènes différentiellement exprimés et validation en RT-qPCR de certains des gènes dérégulés entre les lignées FOXS1 positive et négatives.

D : *Gene set enrichment analysis* (GSEA) des signatures enrichies dans les profils d'expression des lignées FOXS1 négatives et positives.

E-F : Analyse en western blot (E) et immunofluorescence (F) de l'expression de la vimentine (VIM) et de l'E-cadhérine (CDH1) dans les différentes lignées de CHC.

2) Analyse transcriptomique de l'extinction de FOXS1

Afin d'étudier le rôle de FOXS1 dans les cellules mésenchymateuses, j'ai réalisé une extinction de son expression à l'aide de deux siRNA, le premier ciblant la séquence nucléotidique correspondant à la *forkhead box* et le second ciblant une séquence plus en aval (figure supplémentaire 4A). L'extinction a été réalisée dans la lignée SNU-449 en présence et absence de TGF β et validée par RT-qPCR, et western blot pour le siRNA n°2 (figure 19A). L'effet des deux siRNA sur de potentielles *off-targets* a été vérifié, que ce soit sur les FOX phylogénétiquement proches de FOXS1 (figure supplémentaire 4A) ou sur les *off-targets* prédites à l'aide de BLAST pour chaque siRNA (figure supplémentaire 4B-C) sans qu'aucune ne soit éteinte par les siRNA.

La première chose à noter est que l'extinction de l'expression de *FOXS1* induit une activation transcriptionnelle importante. 177 gènes (non redondants) sont modulés par l'extinction de FOXS1 par les deux siRNA, à la fois dans les conditions contrôle et avec traitement au TGF β (table 3). Parmi ces gènes, 154 (87%) voient leur expression augmentée, ce qui suggère que FOXS1 a un rôle répressif. De plus, l'analyse de cette signature montre un enrichissement en gènes impliqués dans la répression transcriptionnelle, suggérant donc que FOXS1 agit comme répresseur transcriptionnel (figure 19C).

Par ailleurs, l'analyse en GSEA montre un impact de l'extinction de *FOXS1* sur les signalisations Wnt et TGF β , ce qui est cohérent avec son expression dans les lignées cellulaires associées à S1 et G3. Les gènes qu'il régule sont également impliqués dans l'invasion tumorale, l'adhésion focale et la signalisation des intégrines, autant de mécanismes associés à l'EMT (figure 19C-D). Parmi les gènes qu'ils régulent, on retrouve ainsi *CTNNB1* et *TJPI1*, codant pour des protéines clés des jonctions cellulaires, qui sont inhibés par FOXS1 puisque induits par son extinction, ainsi que *AXIN2*, un régulateur négatif de la voie Wnt, lui aussi inhibé par FOXS1. A l'inverse, parmi les gènes induits par FOXS1 se trouvent plusieurs gènes associés à la migration et l'invasion tumorale comme *MMP9*, permettant la dégradation de la matrice extracellulaire et *ITGB6*, une sous-unité des intégrines associée à l'invasion et capable, comme nous l'avons vu, de libérer le TGF β piégé dans la matrice. *SNAI2* est également inhibé par l'extinction de *FOXS1*, renforçant le lien de FOXS1 avec l'EMT puisqu'il agit sur un EMT-TF.

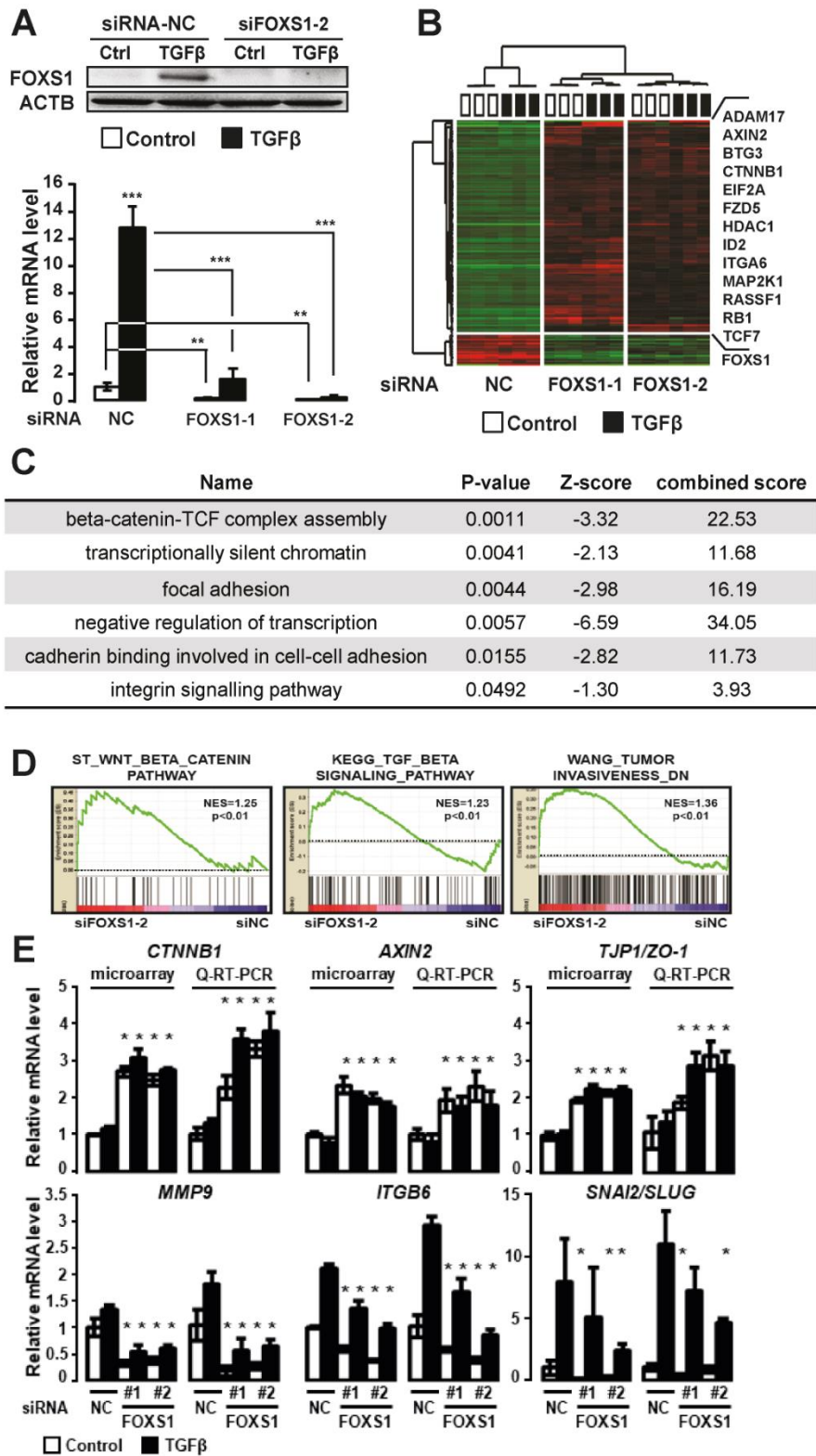


Figure 19 : L'extinction de FOXS1 induit des modifications associées à l'EMT et l'invasion tumorale.

A : Validation de l'extinction de *FOXS1* par western blot (panel du haut) ou RT-qPCR (panel du bas) en présence (barre noire) ou absence (barre blanche) de TGFβ à 1ng/mL pendant 16h.

B : Classification hiérarchique basée sur les profils d'expression de 182 sondes (177 gènes) différentiellement exprimés par l'extinction de FOXS1 (P-value < 0.001 ; Changement d'expression > 2). La liste complète des gènes est fournie en table 3.

C : Signalisations enrichies dans la signature de gènes associée à l'extinction de FOXS1

D : GSEA des signatures enrichies dans les profils d'expression associés à l'extinction de FOXS1.

E : Validation de gènes clés modulés par l'extinction de FOXS1 par RT-qPCR dans les SNU-449 traitées (barre noire) ou non (barre blanche) au TGFβ à 1ng/mL pendant 16h. L'analyse statistique a été réalisée par un t.test (n=3).

3) Analyse moléculaire et fonctionnelle du gain et perte de fonction de *FOXS1*

Afin d'étudier plus précisément l'action de *FOXS1* dans les cellules cancéreuses de CHC, j'ai ensuite réalisé des expériences de gain et perte de fonctions stables dans les lignées PLC/PRF/5 et SNU-449. Le gain de fonction a été réalisé par infection lentivirale suivi de sélection clonale, permettant l'expression exogène de *FOXS1* avec une étiquette hémagglutinine. Cette surexpression a été validée par western blot, RT-qPCR et immunofluorescence dans les PLC/PRF/5, montrant une élévation conséquente de l'expression de *FOXS1* et une localisation nucléaire de la protéine (figure 20A-B-C). Le gain de fonction a été réalisé de la même manière dans les SNU-449 et validé par immunofluorescence et western blot (figures supplémentaires 5A-B). La perte de fonction a été réalisée uniquement dans les SNU-449 à l'aide de l'outil CRISPR/cas9. Un ARN guide ciblant le locus de *FOXS1* en dehors de la séquence nucléotidique de la *forkhead box* (pour limiter les risques de *off-targets* sur d'autres gènes *FOX*) a été utilisé pour induire une coupure double brin. Après sélection clonale, les clones ont été criblés par western blot afin d'identifier les cellules ayant perdu l'expression de la protéine. Les clones positifs ont été séquencés pour valider la mutation du gène (figure supplémentaire 5C).

Ces deux modèles ont été utilisés pour étudier l'impact de *FOXS1* sur l'EMT (figure 20A-B-C). Au niveau moléculaire, le KO de *FOXS1* entraîne une diminution de la vimentine, bien que cela ne soit observable qu'en western blot et une augmentation des marqueurs épithéliaux avec notamment une augmentation de l'E-cadhérine, à la fois au niveau transcriptomique et protéique (bien que difficilement visible en immunofluorescence), et une légère augmentation de ZO-1. Ces résultats sont relativement complémentaires dans le modèle de surexpression, avec une induction d'expression de la vimentine au niveau protéique (en western blot et immunofluorescence), bien que cela ne soit pas visible au niveau transcriptomique, suggérant une régulation post-transcriptionnelle. La surexpression ne semble pas impacter l'E-cadhérine, bien qu'une diminution d'expression de *CDH1* soit visible au niveau transcriptionnel par RT-qPCR (figure 20C) tandis que ZO-1 est réprimé avec une diminution de son expression en RT-qPCR et western blot. Ces résultats sont visibles également dans les SNU-449 surexprimant *FOXS1* (Figure supplémentaire 8D) avec une augmentation de la vimentine et diminution de l'E-cadhérine, bien que moins marquées. Cela est probablement dû au phénotype déjà très

mésenchymateux des SNU-449. Il est également intéressant de noter que l'expression endogène de *FOXS1* semble diminuer dans les SNU-449 surexprimant FOXS1, impliquant donc un possible rétrocontrôle de l'expression de FOXS1 par lui-même.

C'est toutefois au niveau de l'impact sur les EMT-TF que les résultats sont les plus intéressants, les SNU-449 KO pour *FOXS1* perdent l'expression, aussi bien au niveau protéique que transcriptomique, de SNAI1 et SNAI2 suggérant que FOXS1 peut agir en amont de ces deux facteurs de transcription essentiels à l'EMT. La surexpression de FOXS1 induit également l'expression de SNAI1 dans les PLC/PRF/5 mais pas de SNAI2 qui n'est pas du tout exprimé dans ces cellules. Un autre modèle de surexpression endogène de FOXS1 dans les SNU-449 à l'aide du système CRISPR/cas9 SAM (*synergistic activation mediator*) montre également une induction de SNAI1 par FOXS1 (figure supplémentaire 6).

Pour confirmer ce rôle pro-EMT de FOXS1, j'ai observé son expression dans différents modèles d'induction de l'EMT. L'expression de *FOXS1* corrèle avec l'induction de la vimentine et la répression de l'E-cadhérine dans un modèle d'EMT induite par le TGF β dans les cellules épithéliales murines de la glande mammaire NMuMG/E9 (figure supplémentaire 9A). Réciproquement, l'expression de *FOXS1* chute en parallèle de l'expression de *VIM* et de l'induction de *CHDI* après déplétion en TGF β induisant le retour vers le phénotype épithélial.

Chez l'Homme, *FOXS1* est également induit par le TGF β dans la lignée A549 de cancer du poumon non à petites cellules (NSCLC) en parallèle de *VIM* et de l'inhibition de *CDHI* (figure supplémentaire 9B). L'induction de l'EMT par le TGF β dans 13 lignées de NSCLC montre de plus que FOXS1 n'est induit par le TGF β que dans les lignées dans lesquelles ce dernier a induit une EMT et dans aucune lignée dont le phénotype est resté épithélial (figure 9C). L'ensemble de ces données indique donc un rôle important de FOXS1 dans l'EMT induite par le TGF β , soit comme inducteur de l'EMT (à l'image de FOXC1), soit dans le maintien de l'EMT.

Afin d'étudier l'impact de FOXS1 dans l'induction de l'EMT, j'ai réalisé un traitement à 24 et 48 heures au TGF β sur les Hep3B afin d'induire une EMT dans ces cellules. J'ai en parallèle bloqué l'induction de FOXS1 par transfection du siRNA n°2. FOXS1 s'avère ainsi inductible dans les Hep3B par des traitements TGF β plus long et cette induction est bloquée par la transfection du siRNA (figure supplémentaire 7). Les effets observés dans les expériences de

gain et perte de fonction sont plus ou moins retrouvés ici, l'extinction de *FOXS1* induisant une augmentation d'expression de *CTNNB1* et *TJP1* et ce particulièrement en présence de TGF β . De la même manière, l'extinction de *FOXS1* réduit l'expression de *SNAI1* mais uniquement avec stimulation au TGF β tandis qu'à l'inverse, l'extinction bloque l'induction de *SNAI2* par le TGF β . Elle reste toutefois sans effet sur l'expression de l'E-cadhérine qui n'est pas non plus inhibée par le TGF β . L'extinction de *FOXS1* ne bloque pas non plus l'induction de la vimentine par le TGF β mais diminue son expression basale en absence de TGF β , suggérant que le contrôle de l'expression de *VIM* par *FOXS1* pourrait être indépendant du TGF β . L'extinction de *FOXS1* semble donc contrecarrer, au moins partiellement, l'induction de l'EMT par le TGF β en bloquant l'inhibition des protéines de jonctions et l'induction de certains EMT-TF. L'analyse ayant toutefois été réalisée sur l'induction précoce de l'EMT, il serait intéressant d'observer l'effet du KO constitutif de *FOXS1* sur l'induction de l'EMT dans les Hep3B mais je n'ai pas eu le temps de générer cette lignée.

Je me suis ensuite intéressé au rôle fonctionnel de *FOXS1* dans les cellules cancéreuses et dans les processus associés à l'EMT comme la migration, l'invasion, la survie et la résistance aux traitements. L'impact de *FOXS1* sur la migration a été évalué par *gap closure assay* sur les lignées SNU-449 gain et perte de fonction pour *FOXS1*. Les cellules ont également été traitées à la mitomycine C de manière à bloquer la prolifération et n'évaluer que la migration cellulaire. Les résultats de ces expériences montrent clairement un impact positif de *FOXS1* sur la migration cellulaire, les SNU-449 surexprimant *FOXS1* étant capable de recouvrir l'espace libre en 24 heures tandis que les lignées KO en sont incapables en 48 heures (figure 20D). L'impact de *FOXS1* sur l'invasion a été étudié par les capacités des lignées KO à traverser des chambres de Boyden recouvertes de matrigel. Les résultats de cette expérience semblent montrer une diminution des capacités invasives des lignées KO sans toutefois atteindre la significativité statistique (figure supplémentaire 8A). L'impact sur la survie et la prolifération a été évalué par la capacité des cellules KO pour *FOXS1* à former des colonies à partir d'une cellule isolée (*colony formation assay*). Les résultats de l'expérience ne montrent aucune différence entre les capacités des SNU-449 contrôle et KO à former des colonies, indiquant que *FOXS1* ne semble pas impacter les capacités prolifératives et anti-apoptotiques des cellules cancéreuses (figure supplémentaire 8B). Les capacités de résistance aux thérapies ont d'ailleurs

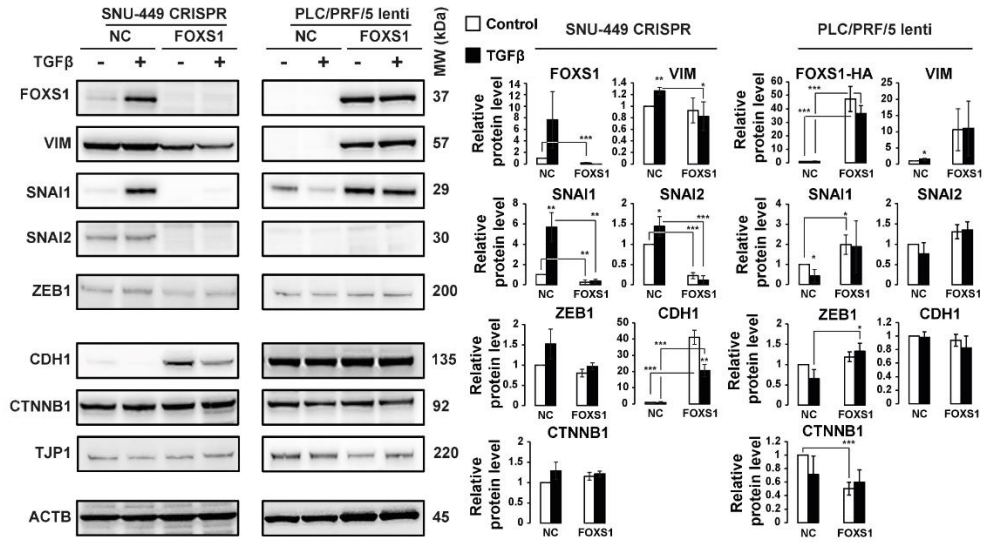
également été étudiées, le phénotype EMT intermédiaire étant, comme nous l'avons vu précédemment, fortement lié à la chimiorésistance. FOXS1 ne semble toutefois pas impliqué dans ce processus, sa surexpression comme son KO ne modifiant pas la sensibilité des cellules au sorafenib (figure supplémentaire 8C). L'impact de FOXS1 sur l'adhésion des cellules a également été évalué par des tests d'adhésion des lignées PLC/PRF/5 contrôle ou surexprimant FOXS1 sur différents éléments de la matrice extracellulaire comme le collagène, la fibronectine ou la laminine. La surexpression de FOXS1 réduit considérablement les propriétés adhérentes des PLC/PRF/5, en accord avec les données moléculaires montrant que FOXS1 inhibe les protéines de jonction et d'adhésion cellulaire (figure supplémentaire 8E). Toutes ces données appuient un rôle pro-EMT de FOXS1. Ce dernier favorise en effet la migration et l'invasion tumorale, probablement par répression des protéines de jonctions cellulaires et induction de certains EMT-TF, notamment les SNAI. FOXS1 étant un facteur de transcription, la question suivante est : Agit-il de manière directe par fixation sur le promoteur de ces gènes ou indirectement par l'activation ou la répression d'autres processus ?

4) Analyse des gènes cibles de FOXS1 par chIP-seq

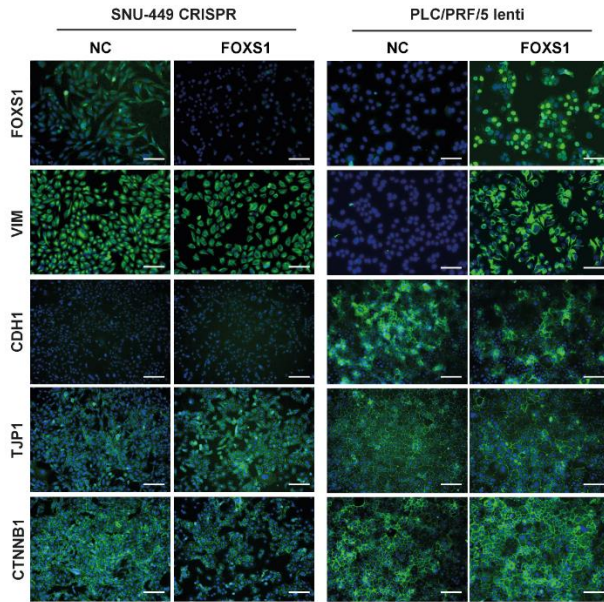
Afin de comprendre plus en détails les mécanismes par lesquelles FOXS1 est capable de contrôler l'EMT, j'ai réalisé une expérience de chIP-seq dans les SNU-449 surexprimant FOXS1 avec une étiquette hémagglutinine avec un anticorps ciblant l'hémagglutinine. Le contrôle négatif a été réalisé avec le même anticorps mais dans des SNU-449 ne surexprimant pas FOXS1. 245 et 3261 pics correspondant à des sites de fixation de FOXS1-HA ont été identifiés respectivement dans le contrôle et les SNU499 FOXS1-HA, validant la spécificité de l'anticorps. L'analyse des gènes associés à ces pics a été réalisée via chIP-ENRICH, soit par association du pic au TSS le plus proche, soit en sélectionnant uniquement les pics à moins de 10kb d'un TSS. La distribution des pics par rapport à leur distance au TSS montre d'ailleurs que 71% des sites de fixation de FOXS1 sont situés à plus de 10kb d'un TSS, suggérant que FOXS1 régule beaucoup de gènes par fixation sur des enhancers distants (figure 21A).

L'analyse des sites de fixation de FOXS1 ne montre toutefois pas de fixation directe de FOXS1 sur le promoteur de SNAI1 et SNAI2 ainsi que des autres EMT-TF, indiquant donc qu'il n'influe pas sur l'EMT par contrôle direct de l'expression des facteurs de transcription clé de

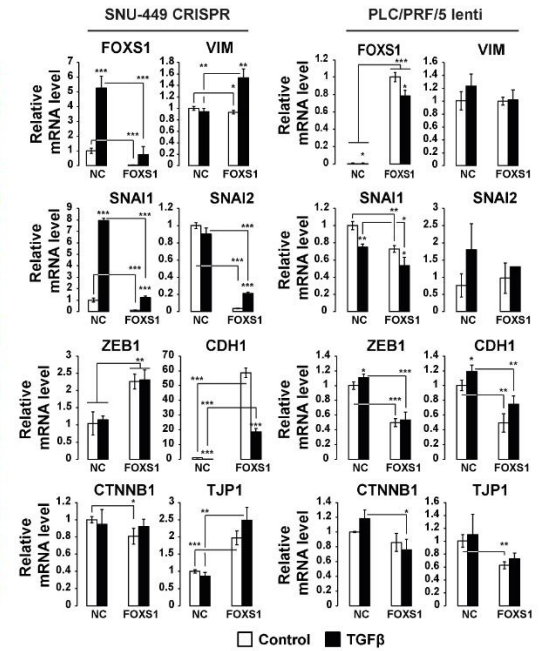
A



B



C



D

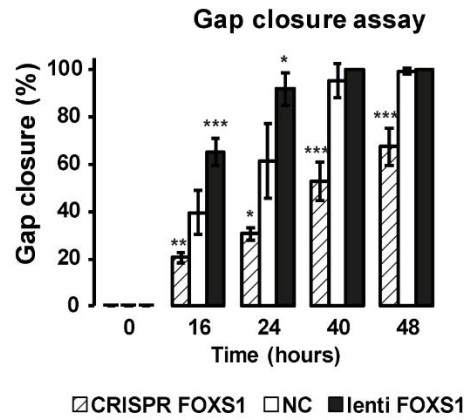
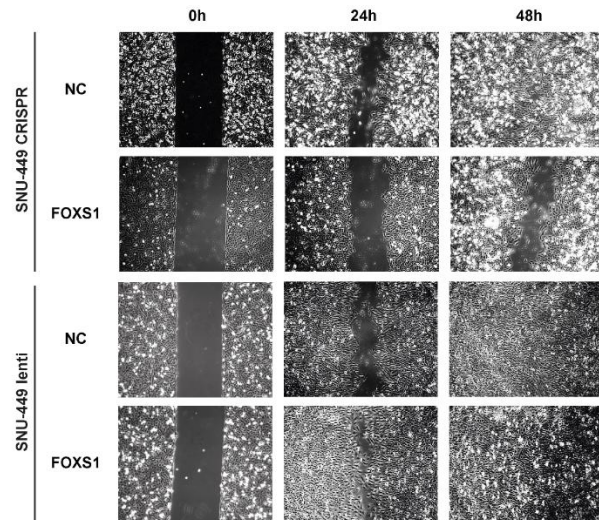


Figure 20 : Analyse moléculaire et fonctionnelle du gain et perte de fonction de FOXS1.

A : Analyse par western blot et quantification des marqueurs EMT après gain et perte de fonction respectivement dans les lignées PLC/PRF/5 et SNU-449. L'analyse a été réalisée en absence (barre blanche) ou présence (barre noire) de TGF β à 1ng/mL pendant 16h.

B : Analyse par immunofluorescence des marqueurs EMT dans les mêmes lignées que ci-dessus. Barre d'échelle : 100 μ m

C : Analyse par RT-qPCR des marqueurs EMT dans les mêmes lignées que précédemment. L'analyse a été réalisée en absence (barre blanche) ou présence (barre noire) de TGF β à 1ng/mL pendant 16h.

D : Test de migration réalisé dans des inserts IBIDI après traitement à la mitomycine C (10 μ g/mL). L'analyse statistique a été réalisé par un t.test (* : P < 0.05 ; ** : P < 0.01 ; *** : P < 0.001 ; n=3)

cette dernière. L'analyse des gènes associés aux pics de fixation de FOXS1-HA montre toutefois un enrichissement dans des signatures géniques associées aux signalisations cellulaires et notamment les signalisations TGF β et EGFR, à la phosphorylation de protéines et à la migration cellulaire (table 4). Ces signatures sont cohérentes avec le rôle pro-EMT de FOXS1. L'analyse des gènes associés à des pics proches de FOXS1 (inférieur à 10kb) montre également un enrichissement dans ces signatures (table 5), bien que moins nombreuses, ainsi qu'un enrichissement dans plusieurs signatures associées à l'activation des lymphocytes T et la réponse Th17 des T helper, indiquant que FOXS1 pourrait également participer à l'action immune du TGF β . L'interprétation de ces résultats est toutefois limitée en l'absence de l'intégration avec les données transcriptomiques des SNU-449 FOXS1-HA qui sera réalisée ultérieurement.

Quelques gènes intéressants sur lesquelles sont retrouvés des sites de fixation de FOXS1-HA au niveau du promoteur ont tout de même été étudiés. C'est notamment le cas de *AXL*, un gène connu pour participer à la transition de la signalisation cytotatique du TGF β vers la signalisation pro-métastatique via phosphorylation du linker des SMAD2/3 ; de *FATI*, une cadhérine dont les mutations sont impliquées dans le maintien du phénotype EMT hybride et de gènes cibles du TGF β comme *SERPINE1* et *SMAD7*, le régulateur négatif de la signalisation TGF β (figure 21B) ainsi que FOXS1 lui-même.^{151,189} De manière extrêmement intéressante, l'expression de *FATI*, *SMAD7* et *SERPINE1* est augmentée dans les SNU-449 KO pour *FOXS1* tandis que leur expression est diminuée dans les PLC/PRF/5 surexprimant FOXS1 (à l'exception de *SERPINE1*) (figure 21C). Ces résultats suggèrent que FOXS1 pourrait inhiber directement l'expression de *FATI* et favoriser ainsi l'EMT mais surtout que FOXS1, qui est induit par le TGF β , est capable d'inhiber l'expression d'autres gènes cibles du TGF β et ainsi d'orienter la réponse cellulaire à cette cytokine. Considérant son rôle pro-EMT, il est donc tout à fait possible que l'expression de FOXS1 permette de diriger la réponse cellulaire au TGF β vers la voie pro-métastatique. Sa capacité à inhiber l'expression de *SMAD7* suppose également qu'il peut participer à la dérégulation de la signalisation TGF β en inhibant le rétrocontrôle négatif de la voie. La fixation de FOXS1 sur son propre locus, et la diminution d'expression de FOXS1 endogène dans les SNU-449 FOXS1-HA (figure supplémentaire 8D) suggère de plus fortement qu'il exerce sur lui-même un rétrocontrôle négatif. Beaucoup reste encore à étudier

pour bien comprendre les mécanismes d'actions de FOXS1, mais ces résultats le montrent clairement comme un médiateur de l'action pro-métastatique de la signalisation TGF β .

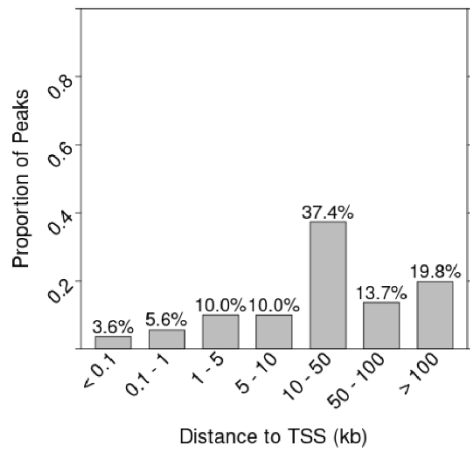
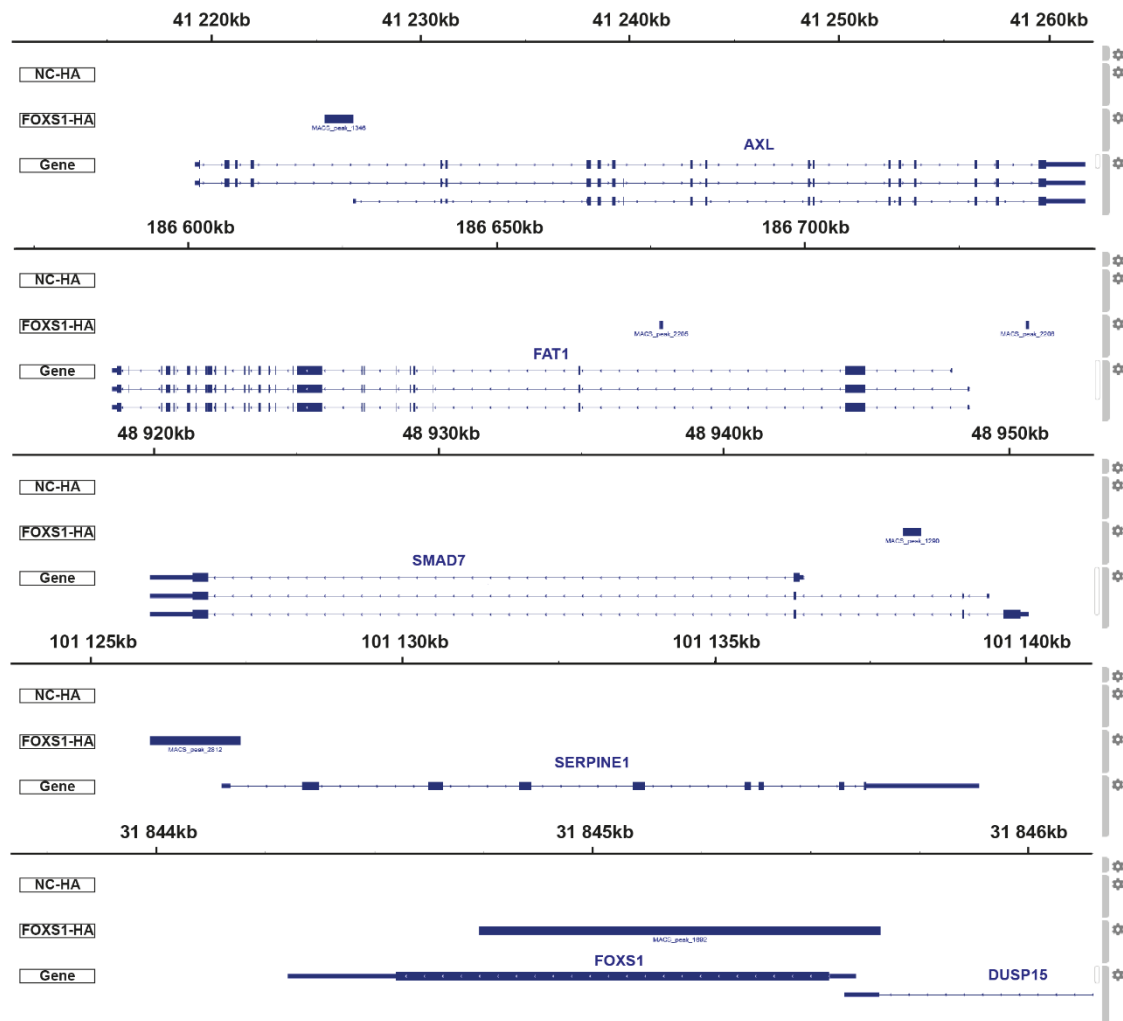
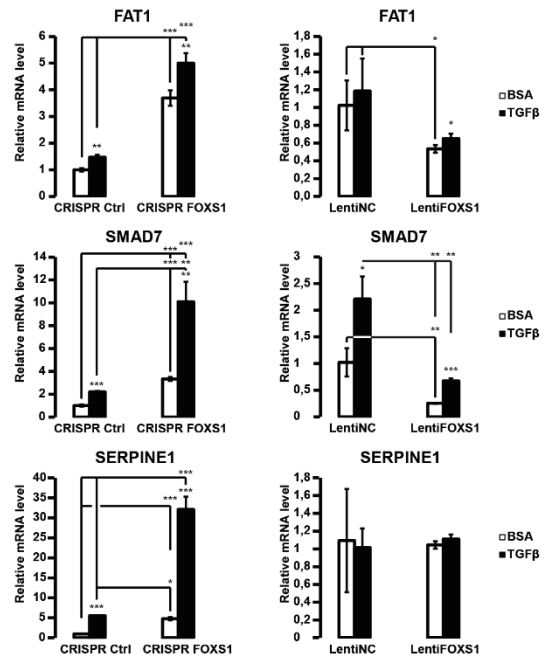
A**Distribution of Distance from Peaks to Nearest TSS****B****C**

Figure 21 : Identification de gènes cibles de FOXS1.

A : Distribution des pics de fixation de FOXS1-HA par rapport à leur distance au site d'initiation de la transcription le plus proche.

B : Sites de fixation de FOXS1-HA au niveau des locus de *AXL*, *FAT1*, *SMAD7*, *SERPINE1* et *FOXS1* visualisés via l'*integrative genome viewer*.

C : Expression de *FAT1*, *SMAD7* et *SERPINE1* dans les lignées gain et perte de fonctions de *FOXS1* en présence (barre noire) ou absence (barre blanche) de TGF β à 1ng/mL pendant 16h.

III) Pertinence clinique de FOXS1

1) Expression de FOXS1 dans les tumeurs

Comme nous l'avons déjà vu dans la partie sur sa régulation (figure 17A-B), l'expression de *FOXS1* augmente dans les tumeurs par rapport au foie non-tumoral, où il n'est normalement exprimé que dans les cellules étoilées du foie activées. Afin de valider cette augmentation, l'expression de FOXS1 a été étudiée dans d'autres cohortes de CHC à partir de la base de données GEO (GSE25097, GSE39791) ou en collaboration avec l'équipe de Jessica Zucman-Rossi aux Cordeliers (cohortes de 193 patients) confirmant une augmentation d'expression de FOXS1 dans les tumeurs par rapport au tissu non tumoral (figures supplémentaires 10A-B et 12A). Cette augmentation d'expression de FOXS1 n'est d'ailleurs pas spécifique au CHC puisqu'elle est également retrouvée dans de nombreux cancers dont le cholangiocarcinome et l'adénocarcinome pancréatique (figures supplémentaires 10C et 11A-B).

Sans surprise, l'expression de FOXS1 corrèle fortement avec l'expression de la vimentine dans les tumeurs hépatiques humaines, à la fois dans la cohorte du TCGA et des Cordeliers montrant encore une fois le lien entre FOXS1 et l'EMT (figures supplémentaires 11D et 12B). L'analyse par hybridation *in situ* de FOXS1 sur une coupe de tumeur montre toutefois une expression homogène de FOXS1 dans la tumeur et pas de surexpression au niveau du front d'invasion tumoral, ce qui aurait été cohérent avec ses propriétés pro-invasives (figure 22A). L'hybridation *in situ* confirme également l'augmentation d'expression de FOXS1 dans la tumeur par rapport au foie non tumoral. FOXS1 n'est d'ailleurs pas du tout exprimé dans les hépatocytes non tumoraux mais est retrouvé dans certaines cellules au niveau des cicatrices fibreuses. Cela est cohérent avec l'expression de FOXS1 dans les cellules étoilées du foie (figure 15E) et le rôle de FOXS1 dans la différenciation des fibroblastes en myofibroblastes¹⁷⁹ ; mais également avec l'augmentation d'expression de *FOXS1* dans le foie cirrhotique et sa corrélation avec le niveau de fibrose (figure supplémentaire 12A).

De la même manière, l'expression de FOXS1 corrèle avec le stade tumoral (figure supplémentaire 11C) mais la variabilité de son expression entre les tumeurs laisse supposer

qu'il est associé à certains sous-types de CHC, comme le montre son enrichissement dans les groupes transcriptomiques CL2 et CL3 des lignées cellulaires de CHC. L'analyse de l'association de FOXS1 avec les sous-groupes G1-G6 définis par Boyault en 2007 dans la cohorte des Cordeliers montre ainsi une expression plus importante de FOXS1 dans le sous-groupe G1 et plus faible dans les sous-groupes G5 et G6 de meilleur pronostic et bien différenciés (figure supplémentaire 12C). L'expression de FOXS1 étant diminuée dans les tumeurs présentant des mutations de CTNNB1, il est logique de ne pas retrouver FOXS1 dans les sous-groupes associés à cette mutation (figure supplémentaire 12D). Il est par contre inattendu de le voir enrichi en G1 et non en G3, ce dernier sous-groupe étant associé à la signature TGF β tardive et aux sous-groupes CL2 et CL3 dans lesquelles FOXS1 est enrichi.¹⁸⁸

2) Association avec la survie des patients

Pour étudier la pertinence clinique de FOXS1, une intégration de la signature génique d'extinction de FOXS1 par siRNA avec les données d'expression de 139 CHC bien caractérisées (la cohorte ayant servi à l'établissement des signatures TGF β précoce et tardive).¹ La classification hiérarchique de cette intégration a permis l'identification de deux clusters, l'un avec les tumeurs associées à la signature génique d'extinction de FOXS1 (cluster 1 : CHC associés au siFOXS1) et l'autre à la présence de FOXS1 (cluster 2 CHC associés au siNC) (figure 22B). Le cluster 1 regroupe majoritairement les tumeurs bien différenciées avec un phénotype hépatocytaire et une signature MET/HGF négative et la quasi-totalité des tumeurs présentant la signature TGF β précoce tandis qu'à l'inverse, le cluster 2 est enrichi en tumeurs présentant un phénotype hépatoblaste et les signatures MET/HGF positives et TGF β tardives. Le cluster 2 regroupe également des tumeurs plus invasives et proliférantes (figure 22C-D). Ces résultats correspondent bien au rôle établi par FOXS1, à savoir un gène pro-EMT favorisant l'action pro-métastatique du TGF β . Enfin, au regard de la survie des patients, l'analyse par Kaplan-Meier plot montre clairement que le cluster 1 est associé à une meilleure survie globale des patients que le cluster 2, indiquant donc que FOXS1 est un marqueur de mauvais pronostic dans le CHC. Ce résultat est d'ailleurs confirmé dans la cohorte du TCGA où l'expression élevée de FOXS1 dans les tumeurs hépatiques (carcinome hépatocellulaire et cholangiocarcinome) est associée à une survie réduite des patients (figure 22E). Ces résultats

globaux montrent donc que FOXS1 est un marqueur de mauvais pronostic dans le carcinome hépatocellulaire.

De plus comme nous l'avons vu précédemment, FOXS1 est induit par le TGF β dans d'autres cancers (figure 15E, figure supplémentaire 9) et les données du TCGA indique qu'il est effectivement exprimé (au moins un transcrit par million) dans la quasi-totalité des tumeurs humaines répertoriées dans le TCGA (à l'exception de la leucémie myéloïde aigue) (figure 22F). L'analyse de son association avec la survie des patients montre qu'il est également associé à une survie globale et/ou une survie sans récurrence réduite dans le cancer de la glande corticosurrénale, le cancer du côlon, du rein, de l'estomac, du sein et de l'œsophage (figure 22G, figure supplémentaire 13). FOXS1 s'avère donc être un marqueur de mauvais pronostic dans les cancers et un potentiel biomarqueur de la signalisation pro-métastatique du TGF β .

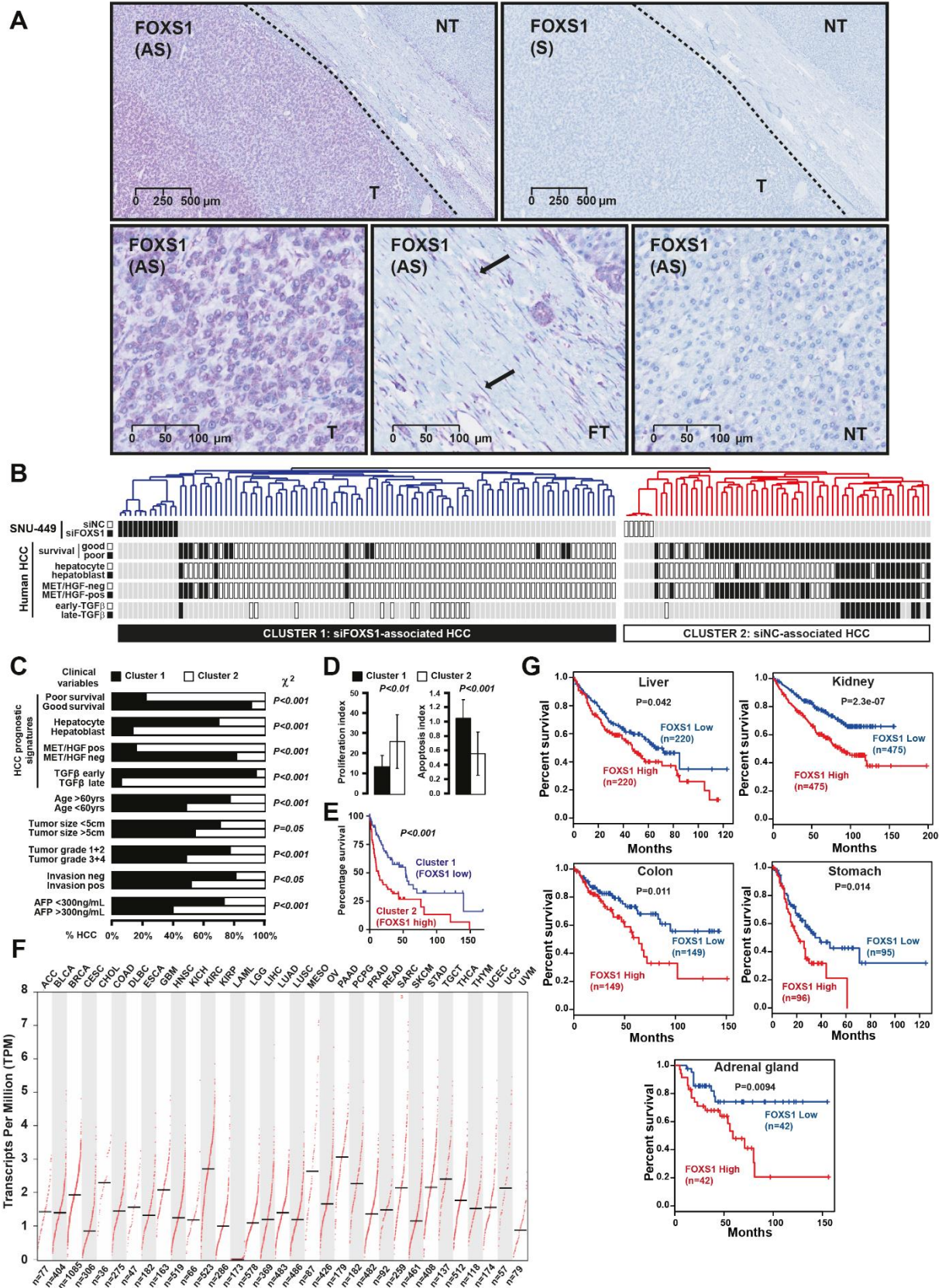


Figure 22 : Pertinence clinique de l'expression de FOXS1.

A : Hybridation *in situ* sur des échantillons cliniques de CHC avec une sonde sens (S) et antisens (AS) montrant une expression augmentée de FOXS1 dans la tumeur (T) comparée au tissu environnant (ST). Flèche noire : Expression de FOXS1 dans des cellules fibroblastiques à l'interface du tissu fibreux.

B : Pertinence clinique de la signature génique d'extinction de *FOXS1*. Dendrogramme représentant les expériences d'extinction dans les SNU-449 transfectées avec le siNC ou le siFOXS1 (comme défini en figure 19B) intégrés avec 139 CHC humains. La classification est basée sur l'expression de 177 gènes différentiellement exprimés après extinction de FOXS1 (table 3). Deux clusters ont été identifiés : le cluster 1 associé à l'extinction de FOXS1 et le cluster 2 associé à son expression. La distribution des CHC humains entre les différents sous-groupes préalablement caractérisés comprenant la survie, l'origine cellulaire, l'activation de MET/HGF et la signature TGF β est indiqué sur la gauche.

C : Analyse statistique de la distribution des CHC entre les clusters 1 et 2 sur la base des signatures géniques et des paramètres cliniques préalablement établis.

D : Index de prolifération et d'apoptose. Les CHC inclus dans le cluster 2 présentent un index de prolifération supérieur et un index apoptotique inférieur.

E : Kaplan-Meier plots et analyse statistique montrant une survie globale réduite chez les patients inclus dans le cluster 2.

F : Expression de FOXS1 dans les différents cancers de la base de données du TCGA d'après GEPIA2.

G : Kaplan-Meier plots et analyses statistiques montrant une diminution de la survie globale des patients présentant une expression élevée de FOXS1. La séparation des tumeurs est basée sur un seuil d'expression de FOXS1 de 45/55 (seuil haut : 45% des tumeurs, seuil bas : 55%) dans les cancers du foie (CHOL + LIHC), du rein (KICH, KIRC, KIRP), du côlon (COAD) et de la glande corticosurrénale (ACC) ; et sur un seuil 75/25 (seuil haut : 3^{ème} quartile, seuil bas : 1^{er} quartile) dans le cancer de l'estomac (STAD). Les données proviennent du TCGA par l'intermédiaire de GEPIA2.

Chapitre 4 : Discussion et conclusion

I) Discussion

L'objectif de mes travaux de thèse était de caractériser le facteur de transcription FOXS1 (expression, régulation, fonction et pertinence clinique) en lien avec la signalisation TGF β et le CHC. Mes travaux ont ainsi montré i) que FOXS1 est un nouveau gène cible du TGF β , induit dans certaines cellules cancéreuses et cellules fibroblastiques du microenvironnement tumoral (les cellules présentant un phénotype mésenchymateux) et ii) que FOXS1 propage les effets pro-métastatiques du TGF β en agissant sur l'EMT. Son expression, généralement augmentée dans les tumeurs humaines, est de plus un signe de mauvais pronostic pour la survie globale et sans récurrence des patients.

Lors des premières expériences d'analyse de l'expression génique des cellules de CHC en réponse au TGF β , FOXS1 est apparu très rapidement comme une cible intéressante. En effet, FOXS1 est un membre de la famille des *forkhead box*, une famille de facteurs de transcription connue pour contrôler de nombreux processus biologiques et orienter les réponses cellulaires dans de nombreux contextes (différenciation cellulaire, survie, apoptose EMT, etc) aussi bien au cours du développement embryonnaire que dans les tissus adultes.¹⁵⁴ Ces protéines sont par conséquent soumises à une régulation fine à laquelle n'échappe pas FOXS1. Les rares études s'intéressant au rôle de cette protéine l'ont été dans un contexte embryonnaire ou cancéreux. De fait, *FOXS1* n'est en effet qu'extrêmement peu exprimé dans le tissu adulte sain à l'exception de quelques types cellulaires bien précis, généralement des cellules fibroblastiques où il est induit lors de la différenciation en myofibroblaste.¹⁷⁹ Dans un contexte embryonnaire cependant, il est retrouvé dans plusieurs types cellulaires comme les cellules péri-endothéliales des testicules où il joue un rôle anti-apoptotique.¹⁷⁸ Il est également exprimé par les cellules de la crête neurale en migration pour former les neurones sensorielles du ganglion spinal.¹⁸¹ Il est intéressant de noter que ces neurones exprimant FOXS1 présentent des capacités prolifératives réduites mais des capacités motiles accrues.

1) Régulation de *FOXS1*

FOXS1 présente ainsi le profil type du gène onco-fœtal, c'est-à-dire du gène exprimé au cours du développement embryonnaire puis éteint dans les tissus adultes et réactivé dans certaines tumeurs. A partir de ce constat, la régulation de l'expression de *FOXS1* par la méthylation de son promoteur semble logique, puisqu'il s'agit d'un des mécanismes d'inhibition les plus communs. De plus, l'analyse de la méthylation du locus de *FOXS1* montre qu'il est relativement fortement méthylé, notamment au niveau de l'îlot CpG présent à l'extrémité 5' du gène (figure 17A, figure supplémentaire 3A), dans le foie sain et que cette méthylation est diminuée (bien que de manière très hétérogène) dans les CHC où l'expression de *FOXS1* est augmentée par rapport au foie non tumoral. Cette observation supporte l'hypothèse d'une régulation par méthylation de l'expression de *FOXS1*. Cette diminution de la méthylation pourrait toutefois n'être également qu'une conséquence de la déméthylation globale du génome dans les cellules cancéreuses sans que cela n'impacte réellement sa régulation.¹⁹⁰ L'analyse de la déméthylation complète du génome dans des lignées de CHC *FOXS1* positives et négatives par traitement à la 5-aza-2'-deoxycytidine (décitabine), un inhibiteur des méthyltransférases de l'ADN, montre de plus un impact variable et relativement modéré sur l'expression de *FOXS1*. Bien sûr, le traitement à la décitabine n'est pas une preuve en soi, la déméthylation totale pouvant entraîner l'expression d'autres facteurs impactant eux-mêmes l'expression de *FOXS1*, mais cette expérience permet de modérer l'impact de la méthylation sur le contrôle de *FOXS1*. Il pourrait être intéressant pour étudier cette question plus en détails de réaliser des séquençages du locus de *FOXS1* après traitement au bisulfite dans les différentes lignées cellulaires *FOXS1* positives et négatives pour vérifier si le niveau de méthylation corrèle efficacement avec l'expression de *FOXS1*.

Les régulations épigénétiques par méthylation vont généralement de pair avec l'accessibilité de la chromatine, contrôlée par les histones. L'analyse de l'état chromatinien du locus de *FOXS1* sur Ensembl (ENSR00000648959) montre effectivement des différences entre les types cellulaires, la chromatine étant considérée comme active dans les fibroblastes et inactive dans les hépatocytes (entre autres). Cela montre bien que les différences d'expression et d'induction de *FOXS1* dépendent de l'état chromatinien de la cellule.

L'expression de *FOXS1* ne dépend toutefois pas que de régulations épigénétiques. Les données que j'ai générées ainsi que celles de la littérature soulignent l'impact de nombreuses signalisations, et toutes particulièrement les signalisations TGF β , TNF α et SHH dans le contrôle transcriptionnel de *FOXS1*. Mes travaux ont clairement montré que *FOXS1* est un gène cible de la signalisation canonique du TGF β et cela dans de nombreux contextes allant du cancer à l'activation des cellules étoilées (et des fibroblastes en général). Il n'est en revanche pas établi si le TGF β ne peut induire *FOXS1* que dans un contexte transcriptionnel et/ou chromatinien particulier ou si le TGF β peut contrôler les changements nécessaires à l'induction de *FOXS1*. L'induction *on* ou *off* de *FOXS1* dans les différentes lignées cellulaires de CHC par des traitements relativement courts montre que l'induction rapide (2h) de *FOXS1* par le TGF β nécessite un contexte transcriptionnel et chromatinien favorable mais la capacité d'induire *FOXS1* par des traitements TGF β plus long dans les Hep3B (figure supplémentaire 7) abonde dans le sens d'une reprogrammation chromatinienne de la cellule lors d'une EMT-TGF β induite permettant l'expression de *FOXS1*. La capacité des SMAD à réguler et interagir avec des protéines de remodelage de la chromatine comme le complexe SWI/SNF n'est plus à démontrer.⁹⁹ Ainsi, les deux enhancers se trouvant en aval du gène *FOXS1* s'avèrent intéressants pour expliquer sa régulation par le TGF β . Bien qu'une validation de l'action de ces enhancers sur l'expression de *FOXS1* s'avère nécessaire, ils présentent un profil au niveau de leurs marques histones H3K4me3 et H3K27ac, présentes uniquement dans les lignées *FOXS1* positives (figure 16D) qui correspond très bien au modèle d'induction de *FOXS1* par le TGF β à court et long terme. L'induction à court terme se ferait par la fixation des SMAD (et de potentiels co-activateurs) sur le ou les enhancers (puisque mes résultats ont montré qu'il ne se fixait pas sur le promoteur), tandis que l'induction à plus long terme nécessiterait tout d'abord un remodelage chromatinien, probablement contrôlé par les SMAD, qui permettrait l'activation du ou des enhancers. L'analyse des marques histones de ces enhancers et de la fixation de SMAD2/3/4 dans la lignée Hep3B à l'état basal et après induction au TGF β pendant 24 ou 48 heures permettrait de valider ce modèle d'induction de *FOXS1*. Cette possible régulation par enhancer est également intéressante quand on sait que l'expression de *FOXS1* est augmentée dans les CHC induits par le VHC. En plus de son impact sur la fibrose, il a récemment été montré que le VHC induisait des modifications épigénétiques au niveau de l'acétylation de

l'histone H3K27 permettant ainsi l'expression de gènes pro-oncogéniques.¹⁹ De là à supposer que *FOXS1* fait partie de ces gènes, il n'y a qu'un pas.

Il est de plus intéressant de noter que l'un de ces enhanceurs (enhancer 1) est prédit par GeneHancer (une base de données des interactions gènes/enhancer) pour réguler *MYLK2*, un gène très proche de *FOXS1* et qui présente le même profil d'expression que lui (corrélation de Pearson $R=0,85$ dans les tumeurs LIHC du TCGA et $R=0,86$ dans les lignées cellulaires de CHC) supposant que ces deux gènes pourraient être régulés de la même manière dans les cellules cancéreuses hépatiques.

L'analyse de la régulation transcriptionnelle de *FOXS1* par d'autres facteurs de transcription (*FOXA1*, *CTCF* et *EGR1*) s'est avérée infructueuse, aucune modulation de ces facteurs ne semblant affecter l'expression de *FOXS1*. *FOXA1* s'avérait pourtant un candidat très intéressant pour expliquer les différences d'expression de *FOXS1* entre les lignées épithéliales et mésenchymateuses. En effet, *FOXA1* est un des gardiens du phénotype épithélial des hépatocytes et il aurait été logique de le voir bloquer l'expression de *FOXS1* pour empêcher une EMT dans les hépatocytes. Qui plus est, son profil d'expression est inverse par rapport à celui de *FOXS1*. Il reste néanmoins intéressant de remarquer que son site de fixation sur le locus de *FOXS1* est le même que celui de *FOXS1*, laissant supposer qu'une séquence conservée de fixation de la *forkhead box* pourrait se trouver sur le locus et que d'autres membres de la famille FOX pourrait participer à la régulation transcriptionnelle de *FOXS1* comme les *FOXC1/2* dont le profil d'expression et la fonction sont similaires.

Au niveau des régulations post-transcriptionnelles, miR-125a-5p a été identifié comme un potentiel régulateur de *FOXS1* par Wang et al, qui ont montré qu'il était capable de fixer l'extrémité 3'UTR de l'ARNm de *FOXS1*.¹⁸⁶ MiR-125a-5p est un suppresseur de tumeurs connus et inhibés dans un certain nombre de cancers incluant le CHC.¹⁹¹ Une régulation par ce microARN est donc cohérente avec le rôle pro-tumoral de *FOXS1*. L'expression de *FOXS1* et celle de miR-125a sont toutefois significativement corrélées ($R=0.38$, $p=0.027$) dans les lignées cellulaires de CHC, ce qui ne soutient pas la théorie de l'inhibition du premier par le second. Des analyses plus complètes devront être réalisées comme l'utilisation d'un miR mimique pour étudier si miR-125a-5p est bien capable d'inhiber *FOXS1* dans le CHC.

Parmi les autres signalisations impactant l'expression de *FOXS1*, la voie du TNF α s'avère avoir un effet opposé à celle du TGF β , réduisant l'expression de *FOXS1* (figure supplémentaire 3D). Le rôle inhibiteur de NF- κ B sur l'expression de *FOXS1* a déjà été relevé dans le cancer gastrique par Wang et al, sans qu'ils ne puissent établir une fixation directe de NF- κ B1 sur le promoteur de *FOXS1*.¹⁸⁶ Il est toutefois relativement surprenant de voir le TNF α et NF- κ B inhiber l'expression de *FOXS1*, cette voie ayant généralement une action pro-EMT notamment par la stabilisation de SNAI1 et l'induction d'autres EMT-TF.¹⁹² L'induction de *FOXS1* semble donc être un mécanisme spécifique des EMT TGF β -induites.

Wang et al, ainsi que Diao et al, ont également mis en évidence une régulation de *FOXS1* par la signalisation *Hedgehog* par l'intermédiaire de la protéine GLI1.^{186,187} Cette régulation n'est toutefois pas clairement établie. Diao et al, ont montré que GLI1 induit l'expression de *FOXS1* dans le médulloblastome tandis que Wang et al, montrent à l'inverse que GLI1 inhibe l'expression de *FOXS1* dans le cancer gastrique. Cette différence pourrait être due à l'existence d'une isoforme courte de GLI1 : tGLI1, connue pour favoriser l'invasion et la migration dans le glioblastome multiforme par rapport à la forme normale.¹⁹³ Des différences d'action de tGLI1 sur le promoteur de *FOXS1* par rapport à ceux de GLI1 pourraient ainsi expliquer ces différences entre cancers. Diao et al, montrent également un rétrocontrôle de *FOXS1* sur l'expression de *GLI1* par interaction des deux protéines entre elles et blocage de l'induction de *GLI1* par lui-même. Cet effet ne semble pas être retrouvé dans le CHC puisque le siRNA contre *FOXS1* n'induit aucune augmentation d'expression de *GLI1*.

Une régulation croisée par la voie Wnt est également envisageable. *FOXS1* est en effet capable de réprimer l'expression de *CTNNB1*, influant d'une part sur les jonctions cellulaires, mais également possiblement sur la voie Wnt dont la β -caténine est un des principaux acteurs. La réciproque semble toutefois également vraie puisque l'expression de *FOXS1* est également réduite dans les tumeurs présentant des mutations activatrices de *CTNNB1* (sous-groupes G5 et G6) suggérant que la β -caténine est potentiellement capable de réprimer l'expression de *FOXS1*.

La régulation transcriptionnelle de *FOXS1* semble donc plutôt complexe puisque de nombreuses voies de signalisation plus ou moins impliquées dans l'EMT semblent y participer. En plus de l'état chromatinien contrôlant l'expression de *FOXS1*, celui-ci serait donc également

fortement dépendant des signaux extracellulaires contrôlés par le microenvironnement tumoral et des voies de signalisation associées au sein de la cellule tumorale pouvant induire ou réprimer son expression. Cette balance entre les différentes cytokines capable d'induire ou d'inhiber *FOXS1* participe probablement à la variabilité d'expression de ce dernier au sein des tumeurs.

2) Fonction de FOXS1

Concernant le rôle de FOXS1, les données montrent clairement un rôle pro-EMT dans le CHC. Son expression dans les lignées cellulaires est associée au phénotype mésenchymateux, il est induit par le TGF β au cours de l'induction de l'EMT et son gain et perte de fonction impactent l'expression des marqueurs et effecteurs de l'EMT. Sa position au sein du processus reste toutefois assez difficile à déterminer. Les analyses moléculaires ont montré que la modulation de son expression impactait *SNAI1* et *SNAI2*, ce qui suggère une action précoce dans l'EMT, ces EMT-TF étant parmi les premiers induits au cours de l'initiation de l'EMT. L'analyse des sites de fixation de FOXS1 par ChIP-seq ne montre toutefois aucune fixation directe sur les promoteurs de *SNAI1* et *SNAI2* ce qui indique que FOXS1 ne régule pas directement l'expression de ces gènes. Il est alors difficile de dire si FOXS1 contrôle ces EMT-TF par la modulation d'autres voies de signalisation, notamment EGFR et TGF β (tables 4-5), et si ces signalisations sont responsables de l'induction ou bien du maintien de l'EMT (par stabilisation des EMT-TF), ce qui placerait FOXS1 à des endroits très différents dans le processus selon le mécanisme en cause.

Considérant le rôle prédominant que semble avoir FOXS1 dans la régulation et l'orientation de la réponse cellulaire au TGF β , il semble légitime de supposer que l'impact de FOXS1 dans l'EMT passe par cette signalisation. Cette dernière est suffisante pour induire une EMT complète, mais les études de Pastushenko et al. sur les états intermédiaires de l'EMT ont montré que l'inhibition de la signalisation TGF β passant par les SMAD bloque l'EMT dans les stades précoces et donc que SMAD2/3 sont impliqués dans les stades hybrides de l'EMT.¹⁹⁴ Etant donné que FOXS1 est contrôlé par les SMAD (figure 16) et est effectivement induit et/ou exprimé dans les cellules ayant déjà entamées le processus d'EMT (figure 18, figure supplémentaire 9), il semble logiquement impliqué dans les stades hybrides de l'EMT. Qui plus

est, FOXS1 module également l'expression de *FAT1*, une cadhérine dont la délétion favorise les états hybrides de l'EMT dans les cancers de la peau et du poumon.¹⁸⁹

Les expériences visant à bloquer l'induction de *FOXS1* au cours de l'EMT TGFβ-induite dans les cellules Hep3B ont montré que l'absence de FOXS1 impacte, au moins partiellement, l'induction de *SNAI1* et *SNAI2* et l'inhibition des protéines de jonctions, validant ainsi l'impact de FOXS1 dans l'EMT TGFβ-induite. L'analyse des gènes cibles de FOXS1 par CHIP-seq semble toutefois montrer que l'action de FOXS1 au sein de la signalisation TGFβ passe par un contrôle des gènes cibles du TGFβ. FOXS1 semble en effet capable d'inhiber certains des gènes cibles du TGFβ comme *SMAD7* ou *SERPINE1*. La répression de *SMAD7* par FOXS1 implique que ce dernier est donc capable de suractiver la signalisation TGFβ en bloquant son rétrocontrôle. *SMAD7* est en effet capable d'inhiber la voie TGFβ en bloquant la phosphorylation de SMAD2 par le TGFBR1. La comparaison des gènes cibles (fixation <10kb du TSS) de SMAD3 dans le foie (d'après les données de CHIP-atlas) avec ceux de FOXS1-HA montre que plus de la moitié des gènes cibles de FOXS1 sont également des cibles de SMAD3 (495/867) et cela atteint 75% (1644/2200) si l'on s'intéresse à l'ensemble des gènes potentiellement régulés par ces facteurs. Son action au sein de la signalisation pourrait donc consister à éteindre ou induire l'expression de certains gènes cibles des SMAD de manière à orienter vers l'action pro-EMT du TGFβ. Il faudra toutefois l'analyse transcriptomique pour valider si FOXS1 est bien capable de réguler ou non certains gènes cibles du TGFβ. Il sera également intéressant de réaliser des expériences de co-immunoprécipitation entre FOXS1 et les SMAD2/3/4 pour voir si FOXS1 n'interagit pas directement avec ces protéines pour réguler les effecteurs de la signalisation TGFβ.

Il est également intéressant de remarquer que FOXS1 inhibe sa propre expression en se fixant sur son locus. Ce mécanisme de rétrocontrôle vise vraisemblablement à empêcher une expression prolongée de FOXS1 après son induction par le TGFβ. Cela suggère qu'il ne doit être normalement exprimé que ponctuellement pour remplir ses fonctions, ce qui est cohérent avec un rôle d'orientation de la réponse au TGFβ, cette dernière s'inhibant également par induction de *SMAD7*. FOXS1 pourrait donc agir comme un boost temporaire de la réponse au TGFβ en bloquant l'induction de *SMAD7* mais également la sienne et en prolongeant ainsi la réponse au TGFβ le temps de sa demi-vie. Ce possible mécanisme de contrôle de la durée de

réponse au TGF β des cellules par FOXS1 serait intéressant à étudier à l'aide d'un système de surexpression de *FOXS1* inductible et d'un rapporteur de l'activité des SMAD. Cela permettrait de suivre l'activation de la voie TGF β et l'action de FOXS1 sur l'activité de la voie.

Un système FOXS1 inductible permettrait également de valider les observations des systèmes de gain et perte de fonction dont le design a empêché une expérience de *rescue*. Le système CRISPR/cas9 et le système d'expression de l'ARN guide contre FOXS1 étant en effet incorporés dans le génome des SNU-449, toutes les tentatives de *rescue* de FOXS1 dans ces cellules modifiées se sont soldées par un échec. Un système inductible d'expression de FOXS1 permettrait d'observer l'effet de FOXS1 sur les marqueurs EMT (VIM, SNAI1/2, etc) puis de valider que ces marqueurs reviennent à leur niveau basal après arrêt de l'induction de FOXS1.

Le rétrocontrôle de *FOXS1* par lui-même est également probablement la cause de l'échec de l'utilisation du système CRISPR SAM à long terme dans les SNU-449 (figure supplémentaire 6). Ce système de surexpression endogène de *FOXS1* a en effet donné des résultats prometteurs et validé l'induction de *SNAI1* par FOXS1 dans un autre modèle de surexpression, mais les cellules ont toutefois perdu l'induction de *FOXS1* après quelques passages. Il est possible que l'inhibition de *FOXS1* par lui-même ait bloqué l'induction par le système CRISPR SAM, résultant après quelques passages en un équilibre de l'expression de *FOXS1* peu différent du niveau basal.

Le rôle de FOXS1 dans l'EMT et le cancer ne fait toutefois pas consensus dans la littérature. Plusieurs auteurs montrent ainsi que FOXS1 réduit la prolifération et l'EMT dans les cellules tumorales.^{183,187,195} L'action de FOXS1 sur la prolifération n'est pas forcément incompatible avec mes résultats dans lesquelles je n'ai pas vu d'impact sur la prolifération. Il est de plus connu que les cellules effectuant une EMT ont généralement tendance à être moins proliférantes au fur et à mesure de la progression vers le phénotype mésenchymateux.¹⁹⁴ L'impact qu'ils relèvent sur l'EMT est en revanche aux antipodes de mes propres résultats, et de ceux de Wang et al, dans le cancer gastrique. Ces différences d'action de FOXS1 pourraient peut-être s'expliquer par des différences de modifications post-traductionnelles sur la protéine, changeant ainsi ses gènes cibles ou ses partenaires selon le contexte cellulaire ou les signalisations en jeu. FOXS1 pourrait donc réguler les mêmes gènes mais de manière diamétralement opposée en recrutant des activateurs transcriptionnels là où dans d'autres contextes, il recruterait des

inhibiteurs. Ce type de mécanismes est courant dans les membres de la famille des *forkhead box*, généralement fortement sujets à modifications post-traductionnelles, et il explique, au moins en partie, le rôle antagoniste qu'ont ces protéines dans des contextes différents (particulièrement étudiées chez les FOXO et FOXM).¹⁵⁴ Je n'ai pour ma part pas observé d'effets opposés de FOXS1 selon les modèles cellulaires, mais cette piste mériterait d'être creusée dans le futur pour obtenir un consensus de l'action de FOXS1 dans le cancer.

L'éclaircissement de ce point nécessite toutefois de bien comprendre le mécanisme d'action de FOXS1 qui reste encore partiellement une boîte noire. L'analyse de son expression génique après diminution de son expression par siRNA a montré qu'il avait un rôle essentiellement inhibiteur, 87% des gènes modulés étant induits par son extinction. Toutefois, contrairement à beaucoup d'autres protéines FOX, aucun domaine d'activation ou d'inhibition n'a été identifié pour le moment dans la protéine FOXS1. Il pourrait donc agir comme un facteur leurre par occupation compétitive des sites de fixation d'autres facteurs de transcription pour empêcher l'induction ou la répression de ses gènes cibles. Ce mécanisme a déjà été proposé par Sato et al, pour expliquer l'apoptose induite dans les cellules péri-endothéliales en cas de KO de *FOXS1*. Ce dernier serait capable de bloquer les sites de fixation de FOXO3 et FOXO4 et leur induction de FASL.¹⁷⁸ Une autre hypothèse, pas forcément exclusive de la première, est qu'il serve de plateforme de recrutement pour d'autres facteurs de transcription. Une interaction entre FOXS1 et BMI1 a déjà été établie par expérience de double hybride.¹⁹⁶ BMI1 est un membre du *polycomb repressive complex 1* (PRC1) impliqué, comme son nom l'indique, dans la répression transcriptionnelle de gènes par remodelage chromatinien. FOXS1 pourrait donc servir de guide à ce complexe pour inhiber l'expression de ses gènes cibles. De la même manière, HDAC2 a été identifié comme une protéine interagissant avec FOXS1.¹⁹⁷ Cette protéine favorise également la répression transcriptionnelle en désacétylant les histones, entraînant généralement une compaction de la chromatine. FOXS1 étant majoritairement retrouvé à distance importante des gènes dans les données de CHIP-seq (figure 21A), il se pourrait que ses mécanismes d'inhibition passent par la désactivation d'enhancers via une interaction avec HDAC2. Des expériences d'ATAC-seq et de CHIP-seq contre les marques histones d'enhancer (H3K27ac et H3K4me3) seraient néanmoins nécessaires pour valider cette hypothèse.

La présence d'un signal de localisation nucléaire (NLS) n'est également pas retrouvée dans la structure protéique de FOXS1 dans les bases de données. Mes expériences ont confirmé que FOXS1, bien que présent dans le cytoplasme, était majoritairement retrouvé dans le noyau où il exerce ses fonctions de facteur de transcription. L'absence de NLS n'est pas toutefois pas spécifique à FOXS1 et est retrouvée chez d'autres membres de la famille comme FOXA1 et FOXA2, eux aussi des facteurs nucléaires. Plusieurs membres de la famille FOX sont en effet dépendants de modifications post-traductionnelles pour permettre leur translocation au noyau. La SUMOylation de FOXA1 favorise par exemple sa translocation au noyau tandis que la phosphorylation des FOXO au niveau de la forkhead box permet leur dissociation de 14-3-3 et leur translocation nucléaire.^{198,199} Toutefois, FOXS1 s'est avéré localisé au noyau dans l'ensemble des modèles cellulaires dans lesquels je l'ai étudié, rendant l'analyse d'un contrôle de son activité par sa localisation quelque peu superflue. Cette potentielle régulation mériterait d'être étudiée dans d'autres cas où FOXS1 semble avoir une action diamétralement différente de celle que j'ai montrée ici. C'est par exemple le cas dans l'étude de Lu et al sur le CHC qui montre FOXS1 comme un marqueur de bon pronostic mais où FOXS1 semble être retrouvé majoritairement dans le cytoplasme.²⁰⁰ La localisation de FOXS1 et les mécanismes la régulant pourrait être un élément de réponse aux différences de rôles de FOXS1 dans la littérature.

3) Pertinence clinique de FOXS1

De plus, comme expliqué précédemment, il n'existe pas plus de consensus actuel dans la littérature sur la pertinence clinique de FOXS1 que sur son rôle dans l'EMT. Des études différentes ont montrés à la fois un rôle oncogénique et suppresseurs de tumeurs dans le carcinome gastrique et le CHC.^{183-186,200} La corrélation de son expression avec la survie des patients tend toutefois à favoriser le rôle oncogénique de FOXS1 puisqu'il est un marqueur de mauvais pronostic dans les cancers gastriques (STAD) et les tumeurs hépatiques (LIHC +CHOL) du TCGA (figure 22G et figure supplémentaire 13). De plus, la diminution d'expression de FOXS1 dans les tumeurs reportées par Lu et al, et Lei et al, a été invalidé dans l'ensemble des cohortes de patients que j'ai étudié (GEO, TCGA et Cordeliers).^{183,195}

Il est important de noter que les deux cohortes de patients proviennent d'hôpitaux chinois chez qui les critères de sélection pour la résection hépatique sont plus flexibles qu'en Occident et il est donc tout à fait possible que les tumeurs proviennent de foies avec des stades de fibroses plus avancés que dans les cohortes occidentales. Hors comme nous l'avons vu, l'expression de FOXS1 est augmentée et corrélée au niveau de cirrhose du fait de son induction dans les cellules étoilées sous l'action du TGF β . L'expression plus faible de FOXS1 dans ces tumeurs est peut-être ainsi davantage due à une augmentation d'expression de FOXS1 dans le foie non tumoral cirrhotique qu'à une diminution dans les cellules cancéreuses.

L'expression induite de FOXS1 dans le microenvironnement tumoral est également un élément de réponse pour justifier les différences d'expression de FOXS1 dans les sous-types de CHC entre les données issues de patients (où FOXS1 est enrichi dans le sous-groupe G1, figure supplémentaire 12C) et les données des lignées cellulaires (où FOXS1 est enrichi dans les groupes CL2 et CL3, figure 15D). Les lignées cellulaires ne reflètent en effet l'expression de FOXS1 que dans les cellules cancéreuses tandis que les données issues de patients reflètent l'expression globale de FOXS1 dans la tumeur mais aussi dans les cellules étoilées activées, celles-ci s'infiltrant dans le stroma tumorale.⁴³ La classification de Boyault n'intègre pas les données du microenvironnement dans ses sous-groupes, il est donc difficile de dire si certains (notamment le G3) peuvent être enrichis en cellules étoilées activées, ce qui « fausserait » ainsi l'expression de FOXS1 dans la tumeur, mais cela reste une hypothèse intéressante pour expliquer les différences entre les données cellulaires et les données issues de patients.

L'impact de FOXS1 dans le microenvironnement n'est pas à écarter pour autant. Le ciblage du microenvironnement tumoral et de la communication croisée qu'il entretient avec la tumeur fait partie des stratégies thérapeutiques émergentes, et des signatures pronostic de patients à partir du stroma ont déjà été identifiées.²⁰¹ Hors, comme nous l'avons vu, une part importante de l'action pro-oncogénique du TGF β dans les tumeurs hépatiques passent par son impact sur le microenvironnement tumoral. Mes résultats montrent l'expression de FOXS1 dans des cellules des cicatrices fibreuses ainsi que l'induction de FOXS1 par le TGF β dans la lignée LX2 de cellules étoilées du foie, supportant un rôle de FOXS1 dans la médiation des effets du TGF β sur le microenvironnement tumoral. FOXS1 a déjà été identifié dans la littérature comme un des facteurs de transcription essentiels de la différenciation des fibroblastes en

myofibroblastes.¹⁷⁹ Il est donc fort probable qu'il soit nécessaire à la différenciation des cellules étoilées en myofibroblastes sous l'action du TGF β et qu'il participe au remodelage matriciel du microenvironnement. Une étude complète est néanmoins nécessaire pour comprendre l'ensemble de son action dans ce processus.

Je n'ai pas étudié le versant immunitaire du TGF β et le rôle que pourrait y jouer FOXS1 au cours de ma thèse, essentiellement par manque de temps et parce qu'il ne s'agit pas de la problématique première de l'axe de recherche de notre équipe. Cela étant dit, creuser le rôle de FOXS1 sur les propriétés immunosuppressives du TGF β dans les cellules immunitaires pourrait s'avérer extrêmement fructueux. La question n'est pas vaine car l'analyse de l'état chromatinien du locus de *FOXS1* sur Ensembl (ENSR00000648959) montre que la chromatine est sous forme « *poised* », c'est à dire inactive mais rapidement activable, dans certaines cellules immunitaires et notamment celles impactées par le TGF β comme les macrophages M2 et les T régulateurs. De plus, le ChIP-seq contre FOXS1-HA montre un enrichissement dans plusieurs signatures associées à l'immunosuppression et à la réponse immunitaire (tables 5 et 6), suggérant que FOXS1 pourrait également être impliqué dans le versant immunitaire du TGF β .

L'association de l'expression de *FOXS1* avec une survie faible dans plusieurs cancers suggère ce dernier comme une potentielle cible thérapeutique. Son ciblage permettrait de limiter son action pro-EMT dans la tumeur et ainsi potentiellement réduire le processus métastatique, mais également de limiter l'activation des cellules étoilées du foie et donc la communication croisée qu'elles entretiennent avec la tumeur et qui favorise la progression tumorale. Néanmoins, avant d'être considéré comme une cible thérapeutique, l'action de FOXS1 dans les tumeurs, notamment sur la croissance tumorale et la dissémination métastatique, doit être validée *in vivo*. Son impact sur le microenvironnement tumoral, l'activation des cellules étoilées par le TGF β et sur la sécrétion de composants de la matrice extracellulaire et de cytokines/facteur de croissances par ces cellules nécessiteront ainsi une étude à part entière.

Si FOXS1 s'avère être un facteur pro-tumoral intéressant à cibler, reste ensuite la question de comment le cibler. L'inhibition d'un facteur de transcription est bien plus complexe que le blocage de kinase ou d'autres protéines à activité enzymatique dont le site actif est plus facilement blocable. Des approches thérapeutiques visant à cibler un facteur de transcription

se sont toutefois développées au cours des dernières années et plusieurs approches peuvent être envisagées.

Une possibilité est de cibler les interactions protéine-protéine entre le facteur de transcription et des activateurs ou inhibiteurs. Des petits inhibiteurs ont déjà été mis au point pour bloquer les interactions entre P53 et MDM2 (un inhibiteur de P53 entraînant sa dégradation par le protéasome) dans les tumeurs solides et les interactions entre CBF β et les protéines RUNX dans les leucémies.^{202,203} Cette approche nécessiterait toutefois d'identifier l'interactome de FOXS1 et de potentielles interactions avec des protéines impliquées dans son action et sa dégradation. Il serait donc davantage pertinent de cibler l'interaction protéine-ADN une fois la séquence consensus de fixation à l'ADN de FOXS1 identifiée, mais ce type d'approche reste encore très limité. Le domaine de liaison à l'ADN des FOX étant de plus bien conservé, de possibles *off-targets* sur d'autres protéines FOX pourrait également entraîner des effets indésirables.

Une autre possibilité serait d'entraîner sa dégradation à l'aide de PROTACs, des molécules bifonctionnelles correspondant à un ligand d'une E3 ligase (permettant l'ubiquitylation et la dégradation de la cible) couplé à un ligand spécifique de la protéine à dégrader. Cette approche a montré son efficacité contre certaines cibles comme les récepteurs aux androgènes ou aux estrogènes (et sont maintenant en essais cliniques : NCT03888612) mais son utilisation contre FOXS1 nécessiterait d'identifier une protéine interagissant fortement et spécifiquement avec FOXS1.

Le ciblage spécifique de FOXS1 s'avère donc complexe et nécessite d'autres études approfondies sur son action *in vivo* et son interactome avant de pouvoir être réellement envisagé. Néanmoins, il reste possible de l'utiliser comme biomarqueur reflétant l'activation du versant pro-métastatique du TGF β . L'essai NCT01246986 testant le galunisertib en seconde ligne a montré une bonne réponse à ce dernier dans une sous-catégorie de patient. Une présélection de patients selon leur expression de *FOXS1* et d'autres potentiels marqueurs du versant pro-métastatique permettrait sûrement d'augmenter la part des patients répondant positivement au galunisertib, mais également d'exclure les patients où le TGF β présente une action anticancéreuse et qui pâtiraient de son inhibition. Afin de valider FOXS1 comme un bon biomarqueur, il faudra toutefois étudier s'il peut être détecté par des approches non invasives

comme l'isolation d'exosomes ou d'ADN tumoral circulant à partir de sang ou d'autres liquides biologiques.

FOXS1 n'est de plus pas le seul facteur impliqué dans le versant pro-métastatique du TGF β et une inhibition complète de la signalisation s'avèrera probablement plus efficace que la perte d'un seul des acteurs, sous couvert donc de bien sélectionner les patients qui pourront y répondre. Un meilleur criblage par l'identification de biomarqueurs pertinents de la dualité fonctionnelle du TGF β pourra permettre d'améliorer les futurs essais cliniques sur des molécules visant à inhiber la signalisation du TGF β en combinaison avec des ICIs comme par exemple le bintrafusp alpha.

II) Conclusion

Ma thèse visait à caractériser le facteur de transcription FOXS1 dans le CHC en étudiant sa régulation, localisation, fonction et pertinence clinique. Mes travaux ont ainsi montré que *FOXS1* est un nouveau gène cible du TGF β , contrôlé par la voie canonique des SMAD probablement via leur fixation sur un enhancer. Ce facteur de transcription est retrouvé essentiellement dans le noyau des cellules cancéreuses où il participe à la signalisation pro-métastatique du TGF β , en favorisant une réponse transcriptionnelle pro-EMT par inhibition spécifique de certains gènes et notamment de gènes cibles du TGF β comme *SMAD7* et par le contrôle indirect de l'expression de facteurs de transcription de l'EMT comme *SNAI1/2* (figure 23).

FOXS1 est augmenté dans les CHC humains par rapport au foie sain. De par sa capacité à induire l'EMT dans les cellules cancéreuses et ainsi à favoriser la migration et l'invasion tumorale, FOXS1 s'avère être un marqueur de mauvais pronostic dans les cancers du foie mais également dans les cancers du rein, de l'estomac et du côlon.

FOXS1 semble de plus impliqué dans les différentes actions du TGF β sur le microenvironnement tumoral. Il est induit par le TGF β au cours de l'activation des cellules étoilées où il participe vraisemblablement à leur différenciation en myofibroblastes. Il est possible qu'il agisse également sur l'action immunosuppressive du TGF β dans les cellules immunitaires.

FOXS1 s'avère donc être un biomarqueur extrêmement pertinent pour identifier les tumeurs dans lesquelles la signalisation pro-métastatique du TGF β est activée et qui peuvent bénéficier de thérapies ciblées à base d'inhibiteur du TGF β comme le galunisertib et de thérapies combinées associant cet inhibiteur à des ICI.

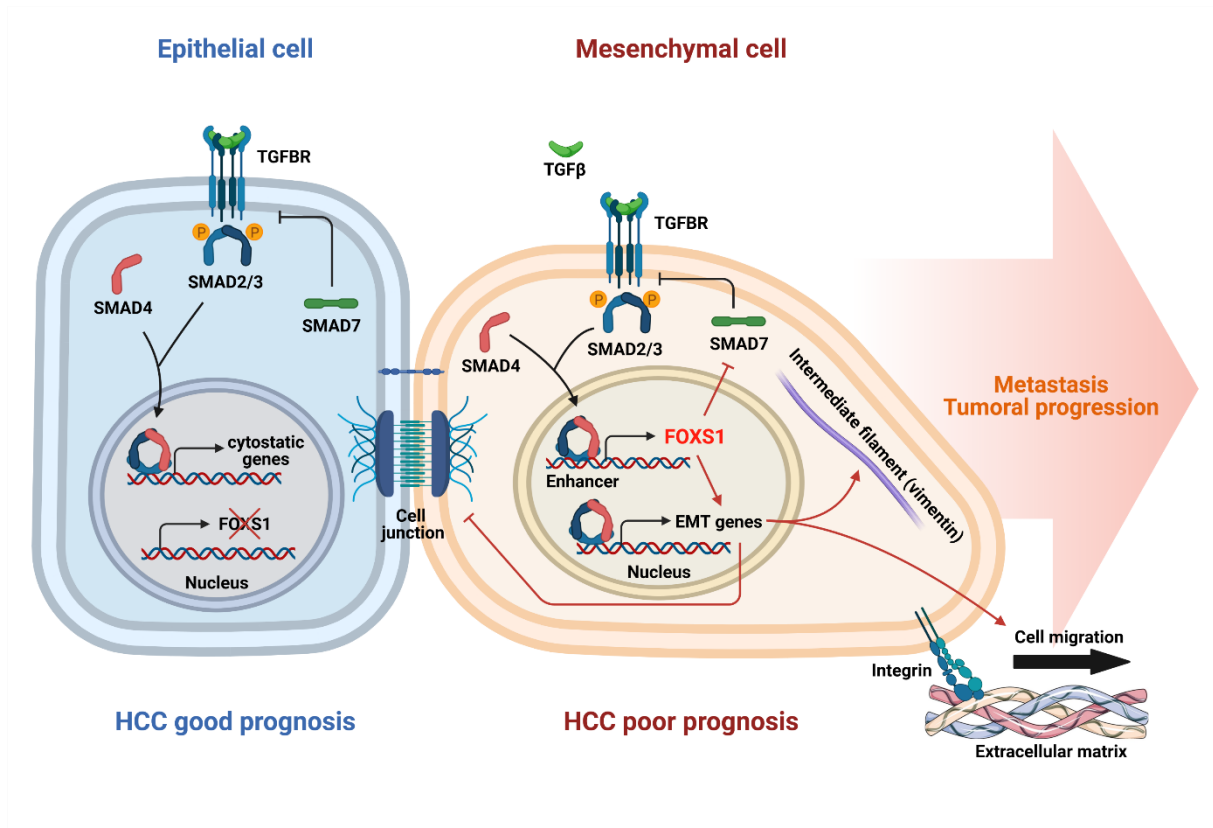


Figure 23 : Rôle et pertinence clinique de FOXS1 dans le CHC.

Bibliographie

1. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394–424 (2018).
2. McGlynn, K. A., Petrick, J. L. & London, W. T. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. *Clin Liver Dis* **19**, 223–238 (2015).
3. Yang, J. D. *et al.* A global view of hepatocellular carcinoma: trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* **16**, 589–604 (2019).
4. Park, J.-W. *et al.* Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int.* **35**, 2155–2166 (2015).
5. Hajarizadeh, B., Grebely, J. & Dore, G. J. Epidemiology and natural history of HCV infection. *Nat Rev Gastroenterol Hepatol* **10**, 553–562 (2013).
6. Yang, J. D. *et al.* Hepatocellular Carcinoma Occurs at an Earlier Age in Africans, Particularly in Association With Chronic Hepatitis B. *Am. J. Gastroenterol.* **110**, 1629–1631 (2015).
7. Yagci, T., Cetin, M. & Ercin, P. B. Cancer Stem Cells in Hepatocellular Carcinoma. *J Gastrointest Canc* **48**, 241–245 (2017).
8. Li, M., Gerber, D. A., Koruda, M. & O’Neil, B. H. Hepatocellular Carcinoma Associated With Attenuated Familial Adenomatous Polyposis: A Case Report and Review of the Literature. *Clinical Colorectal Cancer* **11**, 77–81 (2012).
9. Dragani, T. A. Risk of HCC: Genetic heterogeneity and complex genetics. *Journal of Hepatology* **52**, 252–257 (2010).
10. Zucman-Rossi, J., Villanueva, A., Nault, J.-C. & Llovet, J. M. Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology* **149**, 1226-1239.e4 (2015).
11. Migita, K. *et al.* Cytokine gene polymorphisms in Japanese patients with hepatitis B virus infection--association between TGF-beta1 polymorphisms and hepatocellular carcinoma. *J Hepatol* **42**, 505–510 (2005).

12. Nahon, P. *et al.* Myeloperoxidase and superoxide dismutase 2 polymorphisms comodule the risk of hepatocellular carcinoma and death in alcoholic cirrhosis. *Hepatology* **50**, 1484–1493 (2009).
13. Nahon, P. *et al.* Liver Iron, HFE Gene Mutations, and Hepatocellular Carcinoma Occurrence in Patients With Cirrhosis. *Gastroenterology* **134**, 102–110 (2008).
14. Dharel, N. *et al.* MDM2 promoter SNP309 is associated with the risk of hepatocellular carcinoma in patients with chronic hepatitis C. *Clin Cancer Res* **12**, 4867–4871 (2006).
15. Abu Dayyeh, B. K. *et al.* A Functional Polymorphism in the Epidermal Growth Factor Gene Is Associated With Risk for Hepatocellular Carcinoma. *Gastroenterology* **141**, 141–149 (2011).
16. Parkin, D. M. The global health burden of infection-associated cancers in the year 2002. *International Journal of Cancer* **118**, 3030–3044 (2006).
17. Dhanasekaran, R., Limaye, A. & Cabrera, R. Hepatocellular carcinoma: current trends in worldwide epidemiology, risk factors, diagnosis, and therapeutics. *Hepat Med* **4**, 19–37 (2012).
18. Yang, J. D. *et al.* Characteristics, management, and outcomes of patients with hepatocellular carcinoma in Africa: a multicountry observational study from the Africa Liver Cancer Consortium. *Lancet Gastroenterol Hepatol* **2**, 103–111 (2017).
19. Hamdane, N. *et al.* HCV-Induced Epigenetic Changes Associated With Liver Cancer Risk Persist After Sustained Virologic Response. *Gastroenterology* **156**, 2313-2329.e7 (2019).
20. Younossi, Z. M. *et al.* Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009. *Hepatology* **62**, 1723–1730 (2015).
21. Stickel, F. & Hellerbrand, C. Non-alcoholic fatty liver disease as a risk factor for hepatocellular carcinoma: mechanisms and implications. *Gut* **59**, 1303–1307 (2010).
22. West, J., Card, T. R., Aithal, G. P. & Fleming, K. M. Risk of hepatocellular carcinoma among individuals with different aetiologies of cirrhosis: a population-based cohort study. *Aliment Pharmacol Ther* **45**, 983–990 (2017).

23. Gouas, D., Shi, H. & Hainaut, P. The aflatoxin-induced TP53 mutation at codon 249 (R249S): biomarker of exposure, early detection and target for therapy. *Cancer Lett* **286**, 29–37 (2009).
24. Hussain, S. P., Schwank, J., Staib, F., Wang, X. W. & Harris, C. C. TP53 mutations and hepatocellular carcinoma: insights into the etiology and pathogenesis of liver cancer. *Oncogene* **26**, 2166–2176 (2007).
25. Kew, M. C. Synergistic interaction between aflatoxin B1 and hepatitis B virus in hepatocarcinogenesis. *Liver Int* **23**, 405–409 (2003).
26. Dhanasekaran, R., Bandoh, S. & Roberts, L. R. Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. *F1000Res* **5**, (2016).
27. Schulze, K. *et al.* Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet* **47**, 505–511 (2015).
28. Nault, J. C. *et al.* High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. *Nat Commun* **4**, (2013).
29. Nault, J.-C., Ningarhari, M., Rebouissou, S. & Zucman-Rossi, J. The role of telomeres and telomerase in cirrhosis and liver cancer. *Nat Rev Gastroenterol Hepatol* **16**, 544–558 (2019).
30. Chiba, K. *et al.* Mutations in the promoter of the telomerase gene TERT contribute to tumorigenesis by a two-step mechanism. *Science* **357**, 1416–1420 (2017).
31. Wheeler, D. A. & Roberts, L. R. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* **169**, 1327–1341.e23 (2017).
32. Bévant, K. & Coulouarn, C. Landscape of genomic alterations in hepatocellular carcinoma: current knowledge and perspectives for targeted therapies. *Hepatobiliary Surg Nutr* **6**, 404–407 (2017).
33. Guichard, C. *et al.* Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nature Genetics* **44**, 694–698 (2012).
34. Narimatsu, T. *et al.* p16 promoter hypermethylation in human hepatocellular carcinoma with or without hepatitis virus infection. *Intervirology* **47**, 26–31 (2004).
35. Sung, W.-K. *et al.* Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet* **44**, 765–769 (2012).

36. Sawey, E. T. *et al.* Identification of a Therapeutic Strategy Targeting Amplified FGF19 in Liver Cancer by Oncogenomic Screening. *Cancer Cell* **19**, 347–358 (2011).
37. Kim, E. *et al.* Promotion of growth factor signaling as a critical function of β -catenin during HCC progression. *Nat Commun* **10**, (2019).
38. Coste, A. de L. *et al.* Somatic mutations of the β -catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc Natl Acad Sci U S A* **95**, 8847–8851 (1998).
39. Satoh, S. *et al.* AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1. *Nature Genetics* **24**, 245–250 (2000).
40. Li, M. *et al.* Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. *Nature Genetics* **43**, 828–829 (2011).
41. Walsh, L. A., Nawshad, A. & Medici, D. Discoidin domain receptor 2 is a critical regulator of epithelial-mesenchymal transition. *Matrix Biol* **30**, 243–247 (2011).
42. Affo, S., Yu, L.-X. & Schwabe, R. F. The Role of Cancer-Associated Fibroblasts and Fibrosis in Liver Cancer. *Annu Rev Pathol* **12**, 153–186 (2017).
43. Thompson, A. I., Conroy, K. P. & Henderson, N. C. Hepatic stellate cells: central modulators of hepatic carcinogenesis. *BMC Gastroenterol* **15**, (2015).
44. Jeng, K.-S., Chang, C.-F., Jeng, W.-J., Sheen, I.-S. & Jeng, C.-J. Heterogeneity of hepatocellular carcinoma contributes to cancer progression. *Critical Reviews in Oncology/Hematology* **94**, 337–347 (2015).
45. Hoshida, Y. *et al.* Molecular classification and novel targets in hepatocellular carcinoma: recent advancements. *Semin Liver Dis* **30**, 35–51 (2010).
46. Llovet, J. M., Brú, C. & Bruix, J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin. Liver Dis.* **19**, 329–338 (1999).
47. Yau, T. *et al.* Development of Hong Kong Liver Cancer staging system with treatment stratification for patients with hepatocellular carcinoma. *Gastroenterology* **146**, 1691–1700.e3 (2014).
48. Adhoute, X., Penaranda, G., Bronowicki, J.-P. & Raoul, J.-L. Usefulness of the HKLC vs. the BCLC staging system in a European HCC cohort. *J. Hepatol.* **62**, 492–493 (2015).

49. Liu, P.-H. *et al.* Prognosis of hepatocellular carcinoma: Assessment of eleven staging systems. *Journal of Hepatology* **64**, 601–608 (2016).
50. Raoul, J.-L. *et al.* Evolving strategies for the management of intermediate-stage hepatocellular carcinoma: available evidence and expert opinion on the use of transarterial chemoembolization. *Cancer Treat. Rev.* **37**, 212–220 (2011).
51. Lee, J.-S. *et al.* Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* **40**, 667–676 (2004).
52. Lee, J.-S. *et al.* A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* **12**, 410–416 (2006).
53. Boyault, S. *et al.* Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* **45**, 42–52 (2007).
54. Coulouarn, C., Factor, V. M. & Thorgeirsson, S. S. Transforming Growth Factor- β Gene Expression Signature in Mouse Hepatocytes Predicts Clinical Outcome in Human Cancer. *Hepatology* **47**, 2059–2067 (2008).
55. Hoshida, Y. *et al.* Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res.* **69**, 7385–7392 (2009).
56. Rebouissou, S. & Nault, J.-C. Advances in molecular classification and precision oncology in hepatocellular carcinoma. *Journal of Hepatology* **72**, 215–229 (2020).
57. Tabrizian, P., Jibara, G., Shrager, B., Schwartz, M. & Roayaie, S. Recurrence of hepatocellular cancer after resection: patterns, treatments, and prognosis. *Ann Surg* **261**, 947–955 (2015).
58. Zarrinpar, A. & Busuttil, R. W. Liver transplantation: past, present and future. *Nat Rev Gastroenterol Hepatol* **10**, 434–440 (2013).
59. Goldaracena, N. & Barbas, A. S. Living donor liver transplantation. *Curr Opin Organ Transplant* **24**, 131–137 (2019).
60. Mazzaferro, V. *et al.* Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* **334**, 693–699 (1996).
61. Kamath, P. S., Kim, W. R., & Advanced Liver Disease Study Group. The model for end-stage liver disease (MELD). *Hepatology* **45**, 797–805 (2007).

62. Yao, F. Y. & Fidelman, N. Reassessing the boundaries of liver transplantation for hepatocellular carcinoma: Where do we stand with tumor down-staging? *Hepatology* **63**, 1014–1025 (2016).
63. Yao, F. Y. *et al.* Liver transplantation for hepatocellular carcinoma: expansion of the tumor size limits does not adversely impact survival. *Hepatology* **33**, 1394–1403 (2001).
64. Mazzaferro, V. *et al.* Predicting survival after liver transplantation in patients with hepatocellular carcinoma beyond the Milan criteria: a retrospective, exploratory analysis. *Lancet Oncol* **10**, 35–43 (2009).
65. Roayaie, S. *et al.* The role of hepatic resection in the treatment of hepatocellular cancer. *Hepatology* **62**, 440–451 (2015).
66. Imamura, H. *et al.* Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *J Hepatol* **38**, 200–207 (2003).
67. Nault, J.-C. *et al.* A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. *Gastroenterology* **145**, 176–187 (2013).
68. European Association For The Study Of The Liver & European Organisation For Research And Treatment Of Cancer. EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* **56**, 908–943 (2012).
69. Llovet, J. M. *et al.* Hepatocellular carcinoma. *Nat Rev Dis Primers* **2**, 16018 (2016).
70. Lencioni, R., de Baere, T., Soulen, M. C., Rilling, W. S. & Geschwind, J.-F. H. Lipiodol transarterial chemoembolization for hepatocellular carcinoma: A systematic review of efficacy and safety data. *Hepatology* **64**, 106–116 (2016).
71. Sangro, B., Iñarrairaegui, M. & Bilbao, J. I. Radioembolization for hepatocellular carcinoma. *J Hepatol* **56**, 464–473 (2012).
72. Llovet, J. M. *et al.* Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med* **359**, 378–390 (2008).
73. Kudo, M. *et al.* Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* **391**, 1163–1173 (2018).

74. Bruix, J. *et al.* Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **389**, 56–66 (2017).
75. Abou-Alfa, G. K. *et al.* Cabozantinib in Patients with Advanced and Progressing Hepatocellular Carcinoma. *N Engl J Med* **379**, 54–63 (2018).
76. Dhanasekaran, R., Nault, J.-C., Roberts, L. R. & Zucman-Rossi, J. Genomic Medicine and Implications for Hepatocellular Carcinoma Prevention and Therapy. *Gastroenterology* **156**, 492–509 (2019).
77. El-Khoueiry, A. B. *et al.* Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *The Lancet* **389**, 2492–2502 (2017).
78. Zhu, A. X. *et al.* Pembrolizumab in patients with advanced hepatocellular carcinoma previously treated with sorafenib (KEYNOTE-224): a non-randomised, open-label phase 2 trial. *Lancet Oncol.* **19**, 940–952 (2018).
79. Sia, D. *et al.* Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features. *Gastroenterology* **153**, 812–826 (2017).
80. Fujita, M. *et al.* Classification of primary liver cancer with immunosuppression mechanisms and correlation with genomic alterations. *EBioMedicine* **53**, (2020).
81. Bévant, K. & Coulouarn, C. A novel immunosuppression-based classification of liver tumors opens new perspectives for adapted therapeutic strategies. *EBioMedicine* **54**, (2020).
82. Faivre, S., Rimassa, L. & Finn, R. S. Molecular therapies for HCC: Looking outside the box. *Journal of Hepatology* **72**, 342–352 (2020).
83. Finn, R. S. *et al.* Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma. *N Engl J Med* **382**, 1894–1905 (2020).
84. Faivre, S. *et al.* Novel transforming growth factor beta receptor I kinase inhibitor galunisertib (LY2157299) in advanced hepatocellular carcinoma. *Liver International* **39**, 1468–1477 (2019).
85. Lind, H. *et al.* Dual targeting of TGF- β and PD-L1 via a bifunctional anti-PD-L1/TGF- β RII agent: status of preclinical and clinical advances. *J Immunother Cancer* **8**, e000433 (2020).

86. Massagué, J. TGF β signalling in context. *Nat. Rev. Mol. Cell Biol.* **13**, 616–630 (2012).
87. Shi, Y. & Massagué, J. Mechanisms of TGF- β Signaling from Cell Membrane to the Nucleus. *Cell* **113**, 685–700 (2003).
88. Ozdamar, B. *et al.* Regulation of the Polarity Protein Par6 by TGF β Receptors Controls Epithelial Cell Plasticity. *Science* **307**, 1603–1609 (2005).
89. Zhang, Y. E. Non-Smad Signaling Pathways of the TGF- β Family. *Cold Spring Harb Perspect Biol* **9**, (2017).
90. Lee, M. K. *et al.* TGF-beta activates Erk MAP kinase signalling through direct phosphorylation of ShcA. *EMBO J* **26**, 3957–3967 (2007).
91. Yamashita, M. *et al.* TRAF6 Mediates Smad-Independent Activation of JNK and p38 by TGF- β . *Molecular Cell* **31**, 918–924 (2008).
92. Yang, W.-L. *et al.* The E3 Ligase TRAF6 Regulates Akt Ubiquitination and Activation. *Science* **325**, 1134–1138 (2009).
93. Liu, Y. *et al.* Transforming Growth Factor- β (TGF- β)-mediated Connective Tissue Growth Factor (CTGF) Expression in Hepatic Stellate Cells Requires Stat3 Signaling Activation*. *Journal of Biological Chemistry* **288**, 30708–30719 (2013).
94. Bhowmick, N. A. *et al.* Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. *Mol Biol Cell* **12**, 27–36 (2001).
95. Non-canonical (non-SMAD2/3) TGF- β signaling in fibrosis_ Mechanisms and targets | Elsevier Enhanced Reader.
<https://reader.elsevier.com/reader/sd/pii/S1084952118302659?token=36833EB56A269FF8E6FA183F427E24E21A91A02B71200F739BC246FB97133E6AD7A21A4345F470D7DD9E92709E5CADCA> doi:10.1016/j.semcd.2019.11.013.
96. Schmierer, B. & Hill, C. S. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat Rev Mol Cell Biol* **8**, 970–982 (2007).
97. Xi, Q., He, W., Zhang, X. H.-F., Le, H.-V. & Massagué, J. Genome-wide Impact of the BRG1 SWI/SNF Chromatin Remodeler on the Transforming Growth Factor β Transcriptional Program. *J Biol Chem* **283**, 1146–1155 (2008).

98. Thillainadesan, G. *et al.* TGF- β -dependent active demethylation and expression of the p15ink4b tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol Cell* **46**, 636–649 (2012).
99. Hill, C. S. Transcriptional Control by the SMADs. *Cold Spring Harb Perspect Biol* **8**, (2016).
100. Xi, Q. *et al.* A poised chromatin platform for TGF- β access to master regulators. *Cell* **147**, 1511–1524 (2011).
101. Annes, J. P., Munger, J. S. & Rifkin, D. B. Making sense of latent TGF β activation. *Journal of Cell Science* **116**, 217–224 (2003).
102. Worthington, J. J., Klementowicz, J. E. & Travis, M. A. TGF β : a sleeping giant awoken by integrins. *Trends Biochem Sci* **36**, 47–54 (2011).
103. Mullen, A. C. & Wrana, J. L. TGF- β Family Signaling in Embryonic and Somatic Stem-Cell Renewal and Differentiation. *Cold Spring Harb Perspect Biol* **9**, (2017).
104. Boyer, L. A. *et al.* Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. *Cell* **122**, 947–956 (2005).
105. Brown, S. *et al.* Activin/Nodal Signaling Controls Divergent Transcriptional Networks in Human Embryonic Stem Cells and in Endoderm Progenitors. *STEM CELLS* **29**, 1176–1185 (2011).
106. Clotman, F. *et al.* Control of liver cell fate decision by a gradient of TGF β signaling modulated by Onecut transcription factors. *Genes Dev* **19**, 1849–1854 (2005).
107. Yamazaki, S. *et al.* Cytokine signals modulated via lipid rafts mimic niche signals and induce hibernation in hematopoietic stem cells. *EMBO J* **25**, 3515–3523 (2006).
108. Battle, E. & Massagué, J. Transforming Growth Factor- β Signaling in Immunity and Cancer. *Immunity* **50**, 924–940 (2019).
109. Tone, Y. *et al.* Smad3 and NFAT cooperate to induce Foxp3 expression through its enhancer. *Nat Immunol* **9**, 194–202 (2008).
110. Gorelik, L., Constant, S. & Flavell, R. A. Mechanism of transforming growth factor beta-induced inhibition of T helper type 1 differentiation. *J Exp Med* **195**, 1499–1505 (2002).

111. Laouar, Y., Sutterwala, F. S., Gorelik, L. & Flavell, R. A. Transforming growth factor-beta controls T helper type 1 cell development through regulation of natural killer cell interferon-gamma. *Nat Immunol* **6**, 600–607 (2005).
112. Hernandez-Gea, V. & Friedman, S. L. Pathogenesis of liver fibrosis. *Annu Rev Pathol* **6**, 425–456 (2011).
113. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol* **15**, 178–196 (2014).
114. On behalf of the EMT International Association (TEMTIA) *et al.* Guidelines and definitions for research on epithelial–mesenchymal transition. *Nat Rev Mol Cell Biol* **21**, 341–352 (2020).
115. Cruz-Solbes, A. S. & Youker, K. Epithelial to Mesenchymal Transition (EMT) and Endothelial to Mesenchymal Transition (EndMT): Role and Implications in Kidney Fibrosis. *Results Probl Cell Differ* **60**, 345–372 (2017).
116. Nieto, M. A., Huang, R. Y.-J., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
117. Kalluri, R. & Weinberg, R. A. The basics of epithelial-mesenchymal transition. *J Clin Invest* **119**, 1420–1428 (2009).
118. Zeisberg, M. & Neilson, E. G. Biomarkers for epithelial-mesenchymal transitions. *J Clin Invest* **119**, 1429–1437 (2009).
119. Stone, R. C. *et al.* Epithelial-Mesenchymal Transition in Tissue Repair and Fibrosis. *Cell Tissue Res* **365**, 495–506 (2016).
120. Ribatti, D., Tamma, R. & Annese, T. Epithelial-Mesenchymal Transition in Cancer: A Historical Overview. *Transl Oncol* **13**, 100773 (2020).
121. Aceto, N. *et al.* Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell* **158**, 1110–1122 (2014).
122. Massagué, J. & Obenauf, A. C. Metastatic Colonization. *Nature* **529**, 298–306 (2016).
123. Fischer, K. R. *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472–476 (2015).
124. Zheng, X. *et al.* Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530 (2015).

125. De Craene, B. & Berx, G. Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* **13**, 97–110 (2013).
126. Pastushenko, I. & Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol.* **29**, 212–226 (2019).
127. Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670–691 (2017).
128. Tsai, J. H. & Yang, J. Epithelial–mesenchymal plasticity in carcinoma metastasis. *Genes Dev* **27**, 2192–2206 (2013).
129. Wellner, U. *et al.* The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol* **11**, 1487–1495 (2009).
130. Stemmler, M. P., Eccles, R. L., Brabletz, S. & Brabletz, T. Non-redundant functions of EMT transcription factors. *Nature Cell Biology* **21**, 102 (2019).
131. Chaffer, C. L., San Juan, B. P., Lim, E. & Weinberg, R. A. EMT, cell plasticity and metastasis. *Cancer Metastasis Rev* **35**, 645–654 (2016).
132. Yilmaz, M. & Christofori, G. EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev* **28**, 15–33 (2009).
133. Korpala, M. & Kang, Y. The emerging role of miR-200 family of microRNAs in epithelial-mesenchymal transition and cancer metastasis. *RNA Biol* **5**, 115–119 (2008).
134. Kourtidis, A., Lu, R., Pence, L. & Anastasiadis, P. Z. A central role for cadherin signaling in cancer. *Exp Cell Res* **358**, 78–85 (2017).
135. Cicchini, C. *et al.* Molecular mechanisms controlling the phenotype and the EMT/MET dynamics of hepatocyte. *Liver Int* **35**, 302–310 (2015).
136. Giannelli, G., Koudelkova, P., Dituri, F. & Mikulits, W. Role of epithelial to mesenchymal transition in hepatocellular carcinoma. *Journal of Hepatology* **65**, 798–808 (2016).
137. Skrzypek, K. & Majka, M. Interplay among SNAIL Transcription Factor, MicroRNAs, Long Non-Coding RNAs, and Circular RNAs in the Regulation of Tumor Growth and Metastasis. *Cancers (Basel)* **12**, (2020).
138. Tam, W. L. & Weinberg, R. A. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat Med* **19**, 1438–1449 (2013).

139. Asangani, I. A. *et al.* MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **27**, 2128–2136 (2008).
140. Beltran, M. *et al.* A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial–mesenchymal transition. *Genes Dev* **22**, 756–769 (2008).
141. Shang, B.-Q. *et al.* Functional roles of circular RNAs during epithelial-to-mesenchymal transition. *Molecular Cancer* **18**, 138 (2019).
142. Xu, J., Lamouille, S. & Derynck, R. TGF- β -induced epithelial to mesenchymal transition. *Cell Res* **19**, 156–172 (2009).
143. Gonzalez, D. M. & Medici, D. Signaling mechanisms of the epithelial-mesenchymal transition. *Sci Signal* **7**, re8 (2014).
144. Ikushima, H. & Miyazono, K. TGF β signalling: a complex web in cancer progression. *Nat. Rev. Cancer* **10**, 415–424 (2010).
145. David, C. J. & Massagué, J. Contextual Determinants of TGF- β Action in Development, Immunity and Cancer. *Nat Rev Mol Cell Biol* **19**, 419–435 (2018).
146. Carmona-Cuenca, I. *et al.* Upregulation of the NADPH oxidase NOX4 by TGF-beta in hepatocytes is required for its pro-apoptotic activity. *J Hepatol* **49**, 965–976 (2008).
147. Fabregat, I. *et al.* TGF- β signalling and liver disease. *FEBS J.* **283**, 2219–2232 (2016).
148. Giannelli, G. *et al.* The rationale for targeting TGF- β in chronic liver diseases. *Eur. J. Clin. Invest.* **46**, 349–361 (2016).
149. Fischer, A. N. M. *et al.* PDGF essentially links TGF-beta signaling to nuclear beta-catenin accumulation in hepatocellular carcinoma progression. *Oncogene* **26**, 3395–3405 (2007).
150. Franco, D. L. *et al.* Snail1 suppresses TGF-beta-induced apoptosis and is sufficient to trigger EMT in hepatocytes. *J Cell Sci* **123**, 3467–3477 (2010).
151. Reichl, P. *et al.* Axl activates autocrine transforming growth factor- β signaling in hepatocellular carcinoma. *Hepatology* **61**, 930–941 (2015).
152. Weigel, D., Jürgens, G., Küttner, F., Seifert, E. & Jäckle, H. The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryo. *Cell* **57**, 645–658 (1989).

153. Hannehalli, S. & Kaestner, K. H. The evolution of Fox genes and their role in development and disease. *Nat Rev Genet* **10**, 233–240 (2009).
154. Lam, E. W.-F., Brosens, J. J., Gomes, A. R. & Koo, C.-Y. Forkhead box proteins: tuning forks for transcriptional harmony. *Nat Rev Cancer* **13**, 482–495 (2013).
155. Kaestner, K. H., Knochel, W. & Martinez, D. E. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev* **14**, 142–146 (2000).
156. Lalmansingh, A. S., Karmakar, S., Jin, Y. & Nagaich, A. K. Multiple modes of chromatin remodeling by Forkhead box proteins. *Biochim Biophys Acta* **1819**, 707–715 (2012).
157. Myatt, S. S. & Lam, E. W.-F. The emerging roles of forkhead box (Fox) proteins in cancer. *Nat. Rev. Cancer* **7**, 847–859 (2007).
158. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
159. Heslop, J. A. & Duncan, S. A. FoxA factors: the chromatin key and doorstop essential for liver development and function. *Genes Dev* **34**, 1003–1004 (2020).
160. Friedman, J. R. & Kaestner, K. H. The Foxa family of transcription factors in development and metabolism. *Cell Mol Life Sci* **63**, 2317–2328 (2006).
161. Gong, Z., Yu, J., Yang, S., Lai, P. B. S. & Chen, G. G. FOX transcription factor family in hepatocellular carcinoma. *Biochim Biophys Acta Rev Cancer* **1874**, 188376 (2020).
162. Zhao, Y. & Li, Z. Interplay of estrogen receptors and FOXA factors in the liver cancer. *Mol Cell Endocrinol* **418**, 334–339 (2015).
163. Li, Z., Tuteja, G., Schug, J. & Kaestner, K. H. Foxa1 and Foxa2 are Essential for Sexual Dimorphism in Liver Cancer. *Cell* **148**, 72–83 (2012).
164. Bach, D.-H. *et al.* The Dominant Role of Forkhead Box Proteins in Cancer. *Int J Mol Sci* **19**, (2018).
165. Dong, X.-Y. *et al.* FOXO1A is a candidate for the 13q14 tumor suppressor gene inhibiting androgen receptor signaling in prostate cancer. *Cancer Res* **66**, 6998–7006 (2006).
166. Seoane, J., Le, H.-V., Shen, L., Anderson, S. A. & Massagué, J. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117**, 211–223 (2004).

167. Atfi, A., Abécassis, L. & Bourgeade, M.-F. Bcr–Abl activates the AKT/FoxO3 signalling pathway to restrict transforming growth factor- β -mediated cytostatic signals. *EMBO Rep* **6**, 985–991 (2005).
168. Kim, J.-H. *et al.* Molecular networks of FOXP family: dual biologic functions, interplay with other molecules and clinical implications in cancer progression. *Mol Cancer* **18**, (2019).
169. Cao, H. *et al.* High FOXK1 expression correlates with poor outcomes in hepatocellular carcinoma and regulates stemness of hepatocellular carcinoma cells. *Life Sci* **228**, 128–134 (2019).
170. Zheng, X. *et al.* Forkhead box (FOX) G1 promotes hepatocellular carcinoma epithelial-Mesenchymal transition by activating Wnt signal through forming T-cell factor-4/Beta-catenin/FOXG1 complex. *J Exp Clin Cancer Res* **38**, 475 (2019).
171. Luo, Q. *et al.* FOXQ1/NDRG1 axis exacerbates hepatocellular carcinoma initiation via enhancing crosstalk between fibroblasts and tumor cells. *Cancer Lett* **417**, 21–34 (2018).
172. Jackson, B. C., Carpenter, C., Nebert, D. W. & Vasiliou, V. Update of human and mouse forkhead box (FOX) gene families. *Hum Genomics* **4**, 345 (2010).
173. Mani, S. A. *et al.* Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. *PNAS* **104**, 10069–10074 (2007).
174. Gilding, L. N. & Somerville, T. C. P. The Diverse Consequences of FOXC1 Deregulation in Cancer. *Cancers (Basel)* **11**, (2019).
175. Han, B. *et al.* FOXC1: an emerging marker and therapeutic target for cancer. *Oncogene* **36**, 3957–3963 (2017).
176. Zhang, G. *et al.* FOXL1, a novel candidate tumor suppressor, inhibits tumor aggressiveness and predicts outcome in human pancreatic cancer. *Cancer Res* **73**, 5416–5425 (2013).
177. Cederberg, A. *et al.* Chromosome localization, sequence analysis, and expression pattern identify FKHL 18 as a novel human forkhead gene. *Genomics* **44**, 344–346 (1997).
178. Sato, Y. *et al.* Importance of forkhead transcription factor Fkhl18 for development of testicular vasculature. *Mol. Reprod. Dev.* **75**, 1361–1371 (2008).
179. Noizet, M. *et al.* Master regulators in primary skin fibroblast fate reprogramming in a human ex vivo model of chronic wounds. *Wound Repair Regen* **24**, 247–262 (2016).

180. Heglind, M. *et al.* Lack of the central nervous system- and neural crest-expressed forkhead gene Foxs1 affects motor function and body weight. *Mol. Cell. Biol.* **25**, 5616–5625 (2005).
181. Montelius, A. *et al.* Emergence of the sensory nervous system as defined by Foxs1 expression. *Differentiation* **75**, 404–417 (2007).
182. Schillebeeckx, M. *et al.* Novel markers of gonadectomy-induced adrenocortical neoplasia in the mouse and ferret. *Mol. Cell. Endocrinol.* **399**, 122–130 (2015).
183. Lu, Q. *et al.* Overexpression of FOXS1 in gastric cancer cell lines inhibits proliferation, metastasis, and epithelial-mesenchymal transition of tumor through downregulating wnt/ β -catenin pathway. *J. Cell. Biochem.* **120**, 2897–2907 (2019).
184. Wang, Z., Chen, G., Wang, Q., Lu, W. & Xu, M. Identification and validation of a prognostic 9-genes expression signature for gastric cancer. *Oncotarget* **8**, 73826–73836 (2017).
185. Mao, Y., Zhao, Q., Yin, S., Ding, X. & Wang, H. Genome-wide expression profiling and bioinformatics analysis of deregulated genes in human gastric cancer tissue after gastroscopy. *Asia Pac J Clin Oncol* **14**, e29–e36 (2018).
186. Wang, S. *et al.* FOXS1 is regulated by GLI1 and miR-125a-5p and promotes cell proliferation and EMT in gastric cancer. *Sci Rep* **9**, (2019).
187. Diao, Y. *et al.* Identification of novel GLI1 target genes and regulatory circuits in human cancer cells. *Mol Oncol* **12**, 1718–1734 (2018).
188. Caruso, S. *et al.* Analysis of Liver Cancer Cell Lines Identifies Agents With Likely Efficacy Against Hepatocellular Carcinoma and Markers of Response. *Gastroenterology* **157**, 760–776 (2019).
189. Pastushenko, I. *et al.* Fat1 deletion promotes hybrid EMT state, tumour stemness and metastasis. *Nature* **589**, 448–455 (2021).
190. Pfeifer, G. P. Defining Driver DNA Methylation Changes in Human Cancer. *Int J Mol Sci* **19**, (2018).
191. TANG, H., LI, R.-P., LIANG, P., ZHOU, Y.-L. & WANG, G.-W. miR-125a inhibits the migration and invasion of liver cancer cells via suppression of the PI3K/AKT/mTOR signaling pathway. *Oncol Lett* **10**, 681–686 (2015).

192. Jung, H.-Y., Fattet, L. & Yang, J. Molecular Pathways: Linking Tumor Microenvironment to Epithelial–Mesenchymal Transition in Metastasis. *Clin Cancer Res* **21**, 962–968 (2015).
193. Lo, H.-W., Zhu, H., Cao, X., Aldrich, A. & Ali-Osman, F. A novel splice variant of GLI1 that promotes glioblastoma cell migration and invasion. *Cancer Res* **69**, 6790–6798 (2009).
194. Pastushenko, I. *et al.* Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463–468 (2018).
195. Lei, D. *et al.* Forkhead Box S1 Inhibits the Progression of Hepatocellular Carcinoma. *Onco Targets Ther* **13**, 11839–11848 (2020).
196. Ravasi, T. *et al.* An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* **140**, 744–752 (2010).
197. Li, X. *et al.* Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. *Mol Syst Biol* **11**, 775 (2015).
198. Sutinen, P., Rahkama, V., Rytinki, M. & Palvimo, J. J. Nuclear Mobility and Activity of FOXA1 with Androgen Receptor Are Regulated by SUMOylation. *Molecular Endocrinology* **28**, 1719–1728 (2014).
199. Lehtinen, M. K. *et al.* A conserved MST-FOXO signaling pathway mediates oxidative-stress responses and extends life span. *Cell* **125**, 987–1001 (2006).
200. Lu, Q. *et al.* FOXS1 is a prognostic factor for hepatocellular carcinoma. 6.
201. Sulpice, L. *et al.* Molecular profiling of stroma identifies osteopontin as an independent predictor of poor prognosis in intrahepatic cholangiocarcinoma. *Hepatology* **58**, 1992–2000 (2013).
202. Tovar, C. *et al.* MDM2 small-molecule antagonist RG7112 activates p53 signaling and regresses human tumors in preclinical cancer models. *Cancer Res* **73**, 2587–2597 (2013).
203. Illendula, A. *et al.* Small Molecule Inhibitor of CBF β -RUNX Binding for RUNX Transcription Factor Driven Cancers. *EBioMedicine* **8**, 117–131 (2016).

Annexes

Figures supplémentaires

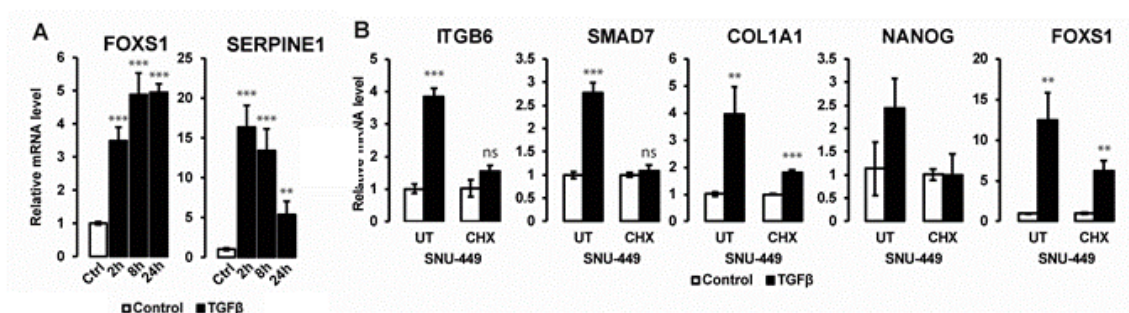


Figure supplémentaire 1 : *FOXS1* est une cible précoce du TGFβ.

A : Analyse par RT-qPCR de l'expression de *FOXS1* dans les SNU-449 après 2h, 8h et 24h de culture en absence (barre blanche) ou présence (barre noire) de 1ng/mL de TGFβ. L'analyse statistique a été réalisée par un t.test (n=3).

B : Analyse par RT-qPCR de l'expression de *ITGB6*, *SMAD7*, *COL1A1*, *NANOG* et *FOXS1* dans les SNU-449 après 16h de culture en absence (barre blanche) ou présence (barre noire) de TGFβ et exposées (CHX) ou non (UT) à 200μg/mL de cycloheximide. L'analyse statistique a été réalisée par un t.test (n=3).

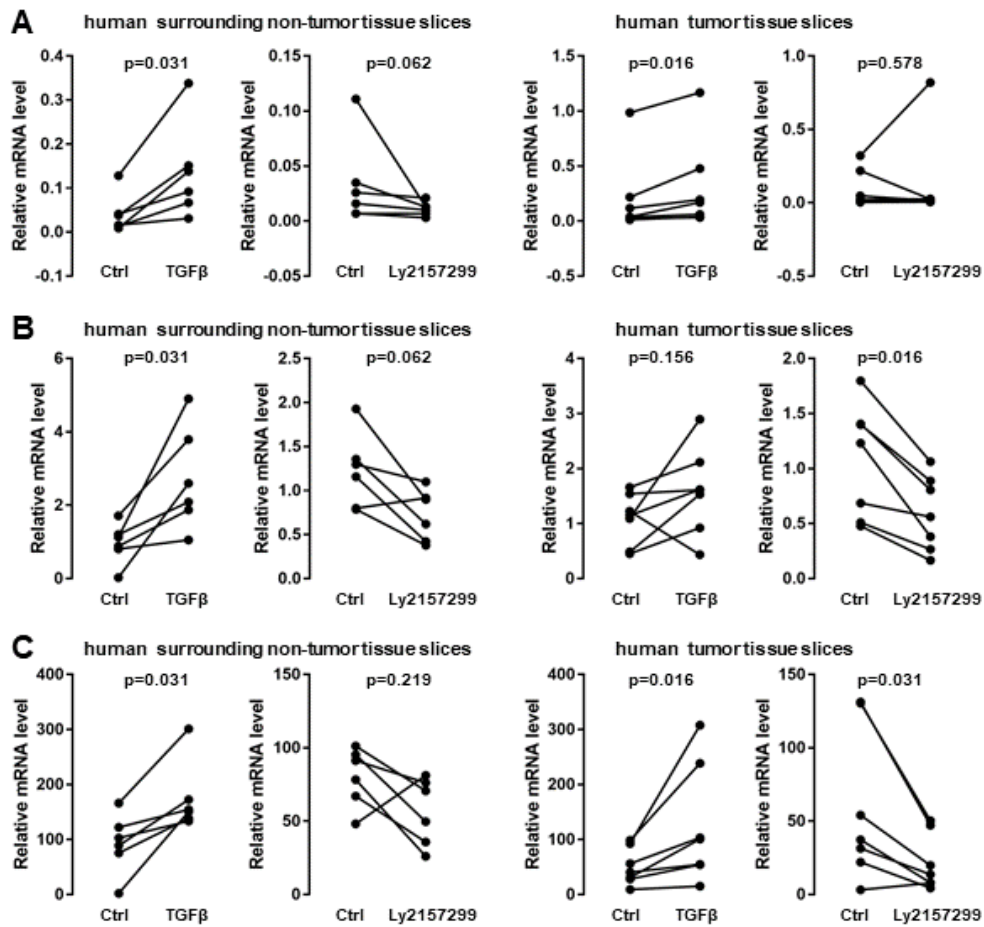


Figure suppl mentaire 2 : Analyse de l'expression g nrique dans des coupes de tissus humains.

Analyse par RT-qPCR de l'expression de (A) *FOXSI*, (B) *SMAD7* et (C) *SERPINE1* dans la tumeur (2 panels de droite, n=7 patients) et dans le tissu environnant (2 panels de gauche, n=6 patients) de patients ayant subi une r section chirurgicale. Les coupes ont  t  incub es 48h en absence ou pr sence de TGF  β   1ng/mL ou 10 M de LY2157299. L'analyse statistique est r alis e par un test de rang appari es non param triques de Wilcoxon.

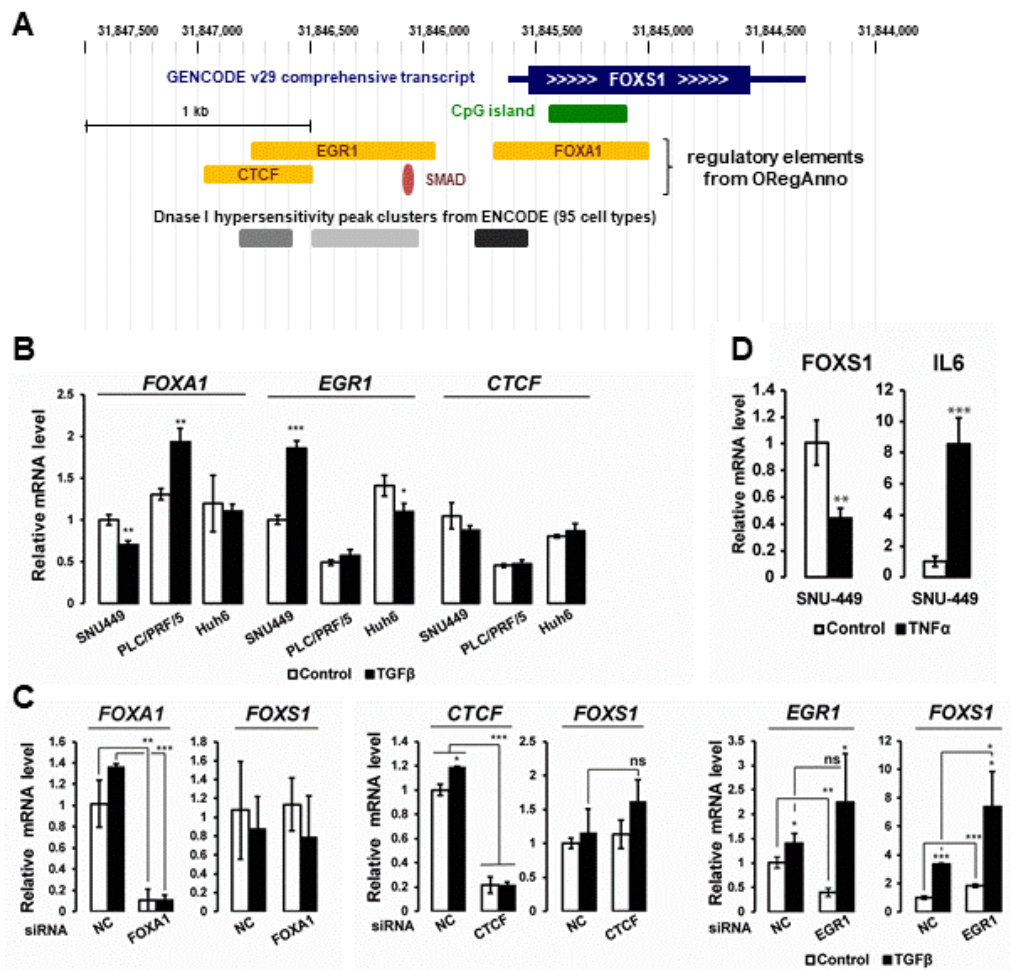


Figure supplémentaire 3 : Organisation génomique et régulation transcriptionnelle de *FOXS1*.

A : Organisation génomique du locus de *FOXS1* au chrs20:31,844,000-31,847,500 (hg38) avec le gène de *FOXS1*, les îlots CpG, les clusters d'hypersensibilité à la DNase I et les éléments de régulation provenant de l'Open Regulatory Annotation Database (ORegAnno) incluant *FOXA1*, *EGR1* et *CTCF*.

B : Analyse par RT-qPCR de l'expression de *FOXA1*, *EGR1* et *CTCF* dans les lignées cellulaires après 16h de culture en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL. L'analyse statistique a été réalisée par un t.test (n=3).

C : Analyse par RT-qPCR de l'expression de *FOXA1* et *CTCF* dans les PLC/PRF/5 et *EGR1* dans les SNU-449 après transfection d'un siRNA ciblant leur ARNm (50nM) ou d'un siRNA contrôle (50nM) pendant 72h suivi de 16h de culture en absence (barre blanche) ou présence (barre noire) de TGFβ à 1ng/mL. L'analyse statistique a été réalisée par un t.test (n=3).

D : Analyse par RT-qPCR de l'expression de *FOXS1* et *IL6* dans les SNU-449 après 16h de culture en absence (barre blanche) ou présence (barre noire) de TNF α à 20ng/mL. L'analyse statistique a été réalisée par un t.test (n=3).

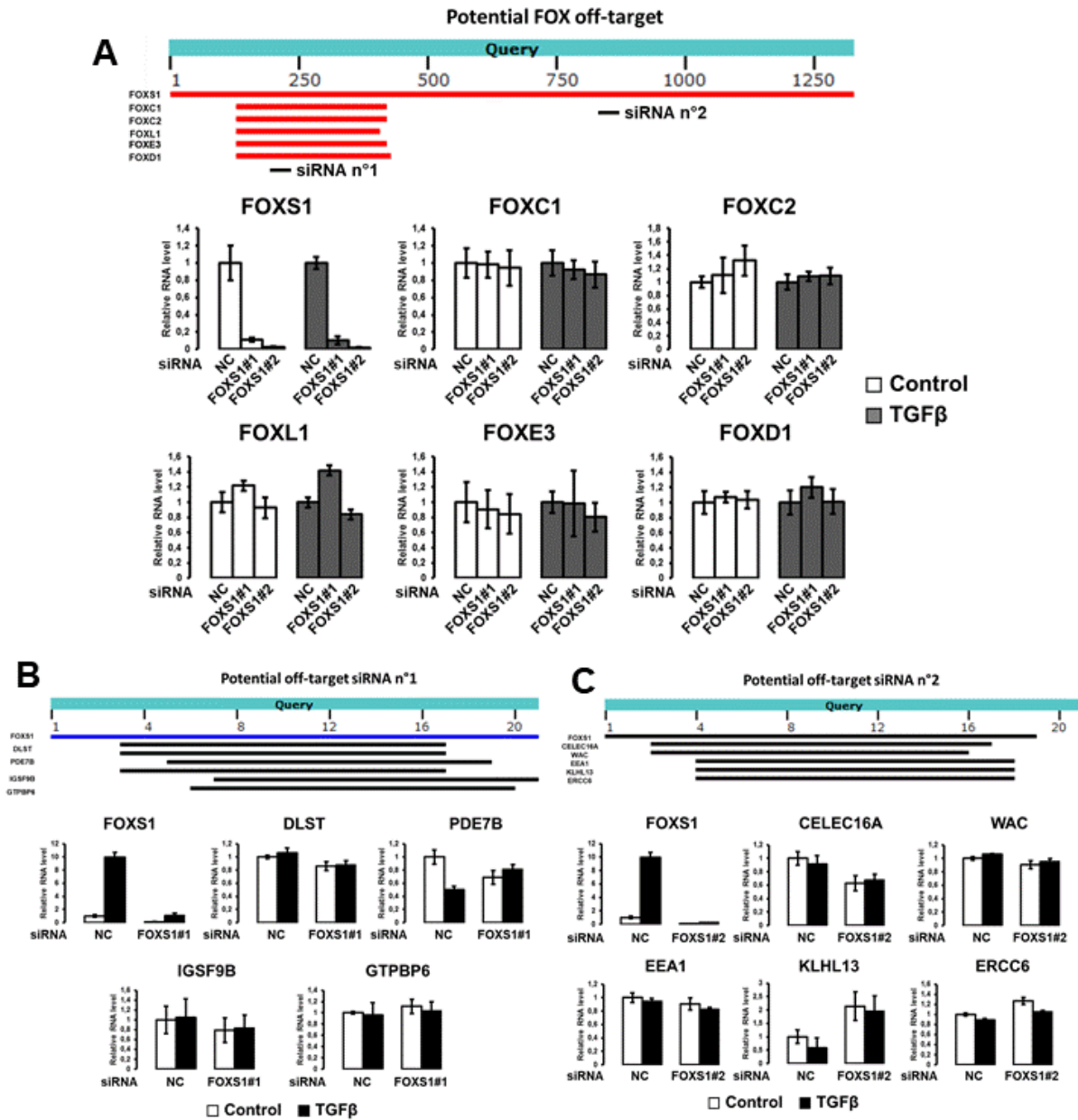


Figure supplémentaire 4 : Validation des off-cibles prédites des siRNA contre FOXS1.

A : Analyse par BLAST de la séquence de l'ARNm de FOXS1 montrant les alignements avec les autres membres de la famille FOX présentant une forte similarité de séquence au niveau de la *forkhead box* contre laquelle est dirigé le siRNA n°1 (panel du haut). Expression des membres de la famille FOX après 72h de transfection des siRNA n°1 et n°2 ciblant FOXS1 (panel du bas)

B-C : Expression des off-targets spécifique du siRNA n°1 (B) ou n°2 (C), prédits par homologie de séquence à l'aide de BLAST, après 72h de transfection des siRNA n°1 ou n°2 ciblant FOXS1.

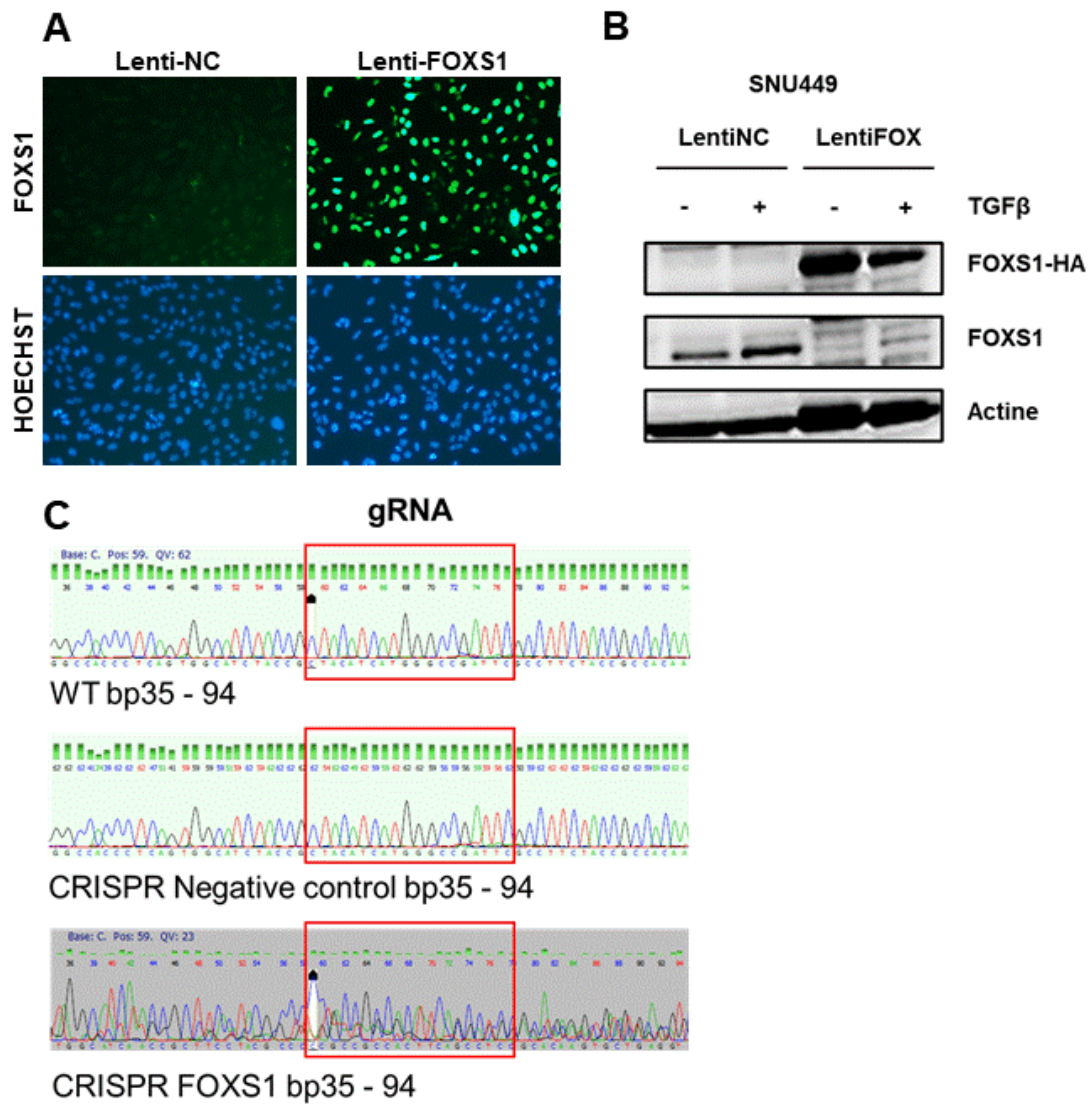


Figure supplémentaire 5 : Validation du gain et perte de fonction de FOXS1.

A-B : Validation de la surexpression de FOXS1 par immunofluorescence (A) et western blot (B) à l'aide de l'anticorps anti-FOXS1. L'expression nucléaire de FOXS1 est validée.

C : Validation des mutations par séquençage par la technique de Sanger du locus de FOXS1 dans des SNU-449 sauvage (WT), infecté par un sgRNA control (CRISPR negative control) ou un sgRNA ciblant *FOXS1* (CRISPR FOXS1).

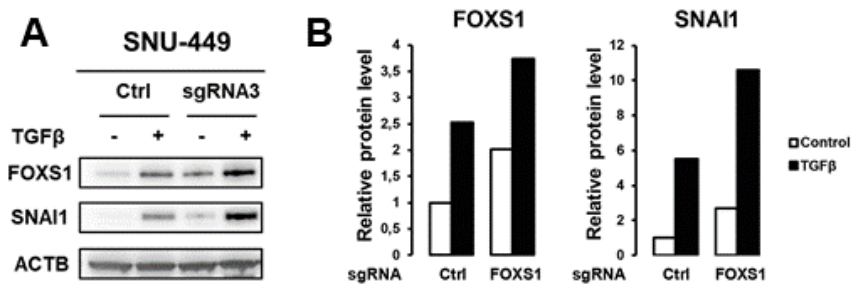


Figure supplémentaire 6 : Surexpression endogène de *FOXS1* par le système CRISPR/cas9 Synergistic Activation Mediator (SAM).

A : Analyse par western blot de l'expression de *FOXS1* et *SNAI1* dans des SNU-499 modifiées pour exprimer le système CRISPR/cas9 SAM et infectées avec un vecteur exprimant un ARN guide contrôle ou ciblant le promoteur de *FOXS1*. Les cellules ont été cultivées en présence ou absence de TGFβ à 1ng/mL pendant 16h.

B : Quantification des protéines par le logiciel ImageJ (n=1).

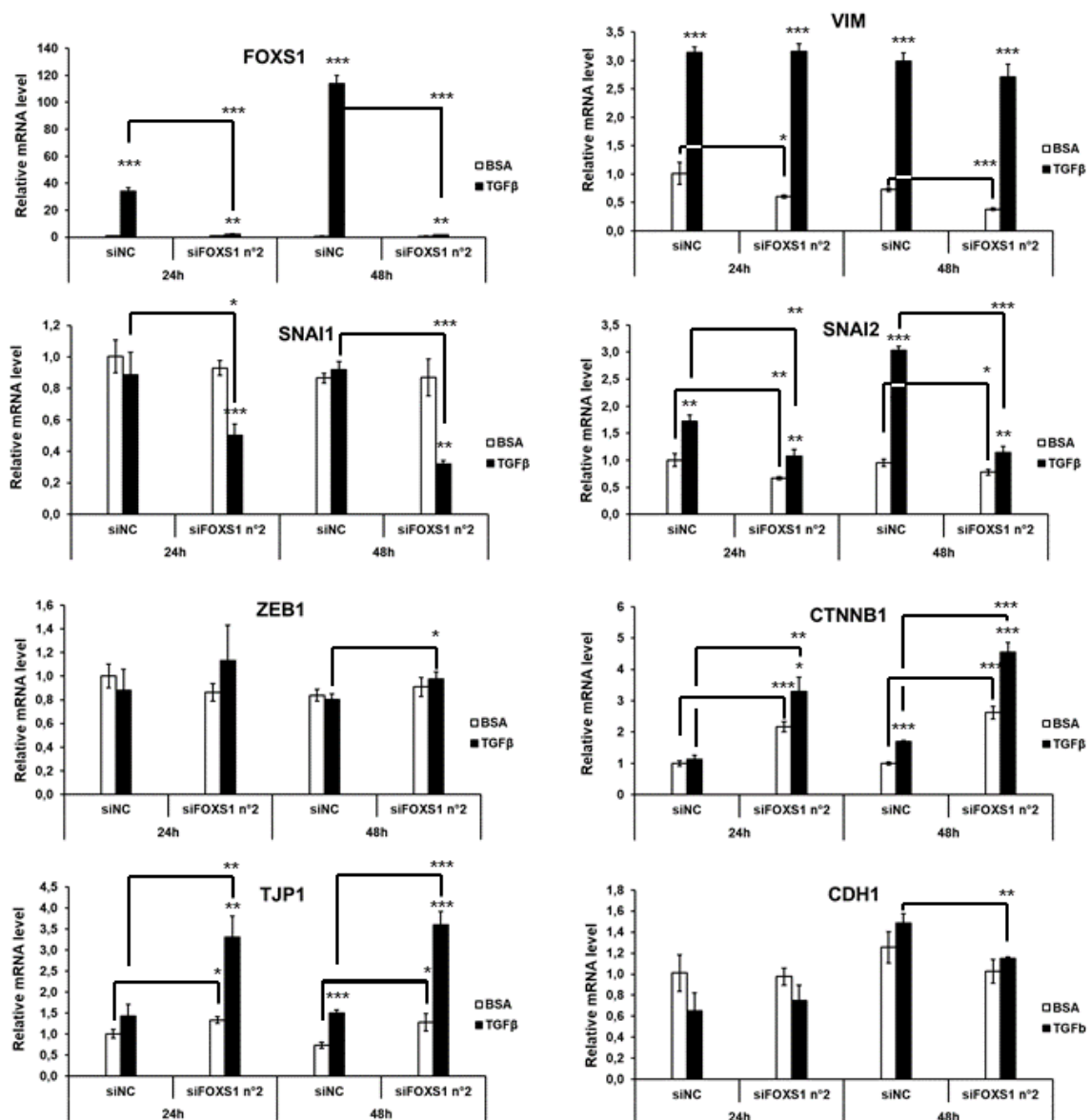


Figure supplémentaire 7 : Extinction de *FOXS1* dans les Hep3B.

Analyse par RT-qPCR de marqueurs de l'EMT dans la lignée Hep3B après transfection d'un siRNA contrôle (siNC) ou ciblant *FOXS1* (siFOXS1 n°2) en présence (barre noire) ou absence (barre blanche) de traitement au TGFβ pendant 24h ou 48h. L'analyse statistique est réalisée par un t.test (* : $p < 0,05$; ** : $p < 0,01$; *** : $p < 0,001$).

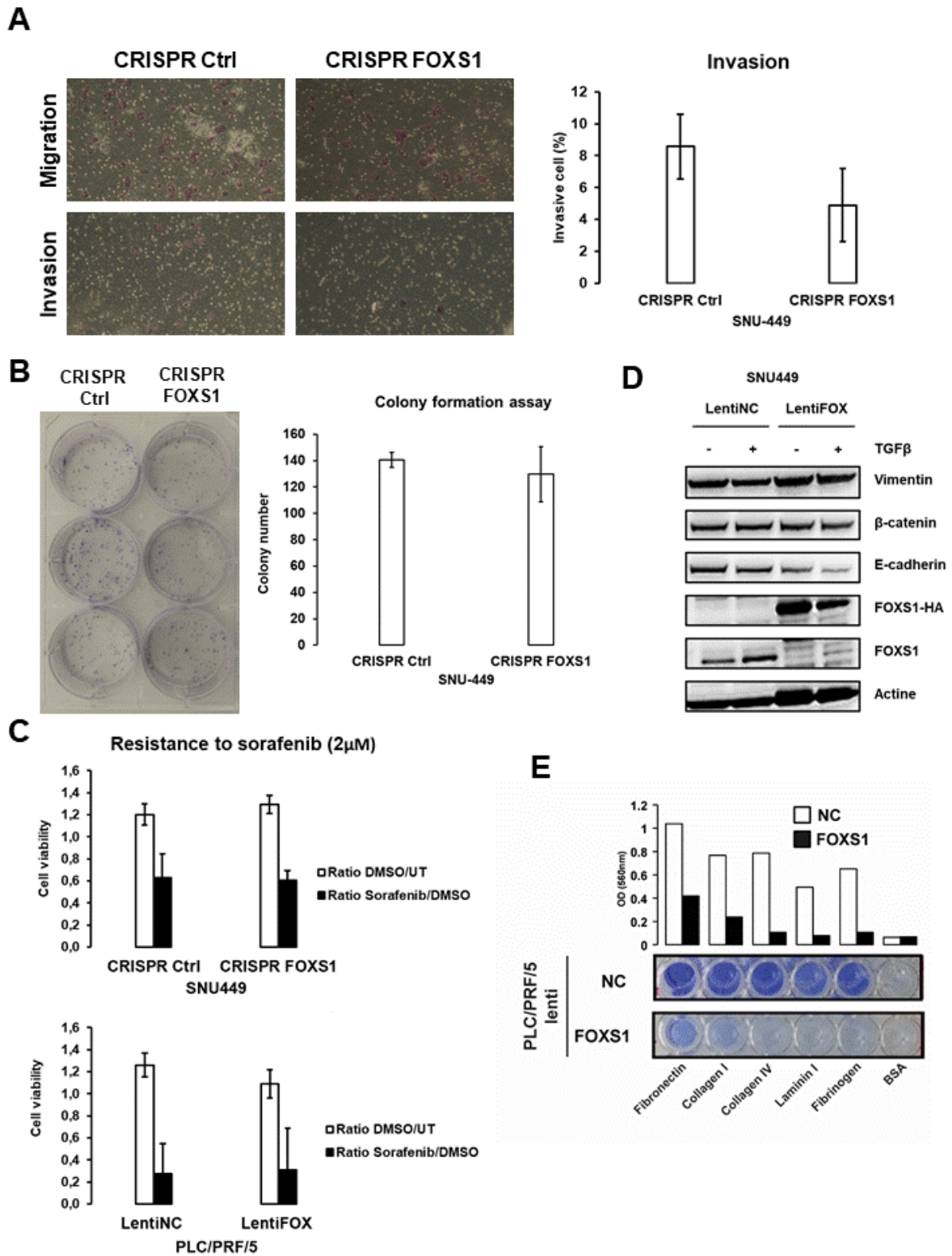


Figure supplémentaire 8 : Analyse fonctionnelle de *FOXS1*.

A : Test d'invasion en chambre de boyden. Le nombre de cellules SNU-449 contrôle ou KO pour FOXS1 capable de traverser des chambres de boyden nues (migration) ou recouvertes de matrigel (invasion) a été évalué après fixation et marquage des cellules.

B : Test de formation de colonie. 250 cellules SNU-449 contrôle ou KO pour FOXS1 ont étéensemencées en plaque 6 puits pendant 10 jours. La capacité des cellules à former des colonies monogéniques a été évaluée après fixation et marquage des cellules.

C : Analyse par prestoblue de la survie des lignées SNU-449 contrôle et KO pour FOXS1 et PLC/PRF/5 contrôle ou surexprimant FOXS1 après traitement (barre noire) ou non (barre blanche) au sorafenib à 2 μ M pendant 24h.

D : Analyse par western blot des marqueurs EMT dans les lignées SNU-449 contrôle ou surexprimant FOXS1.

E : Test d'adhésion des lignées PLC/PRF/5 contrôle et surexprimant FOXS1. Les cellules ont étéensemencées pendant 1h dans des puits recouvert de composants de la matrice extracellulaire (Collagène, fibronectine, laminine, fibrinogène) puis lavées au PBS. Les cellules ont été marquées puis lysées et l'absorbance des solutions de lyse a été mesurée pour évaluer la proportion de cellules adhérentes.

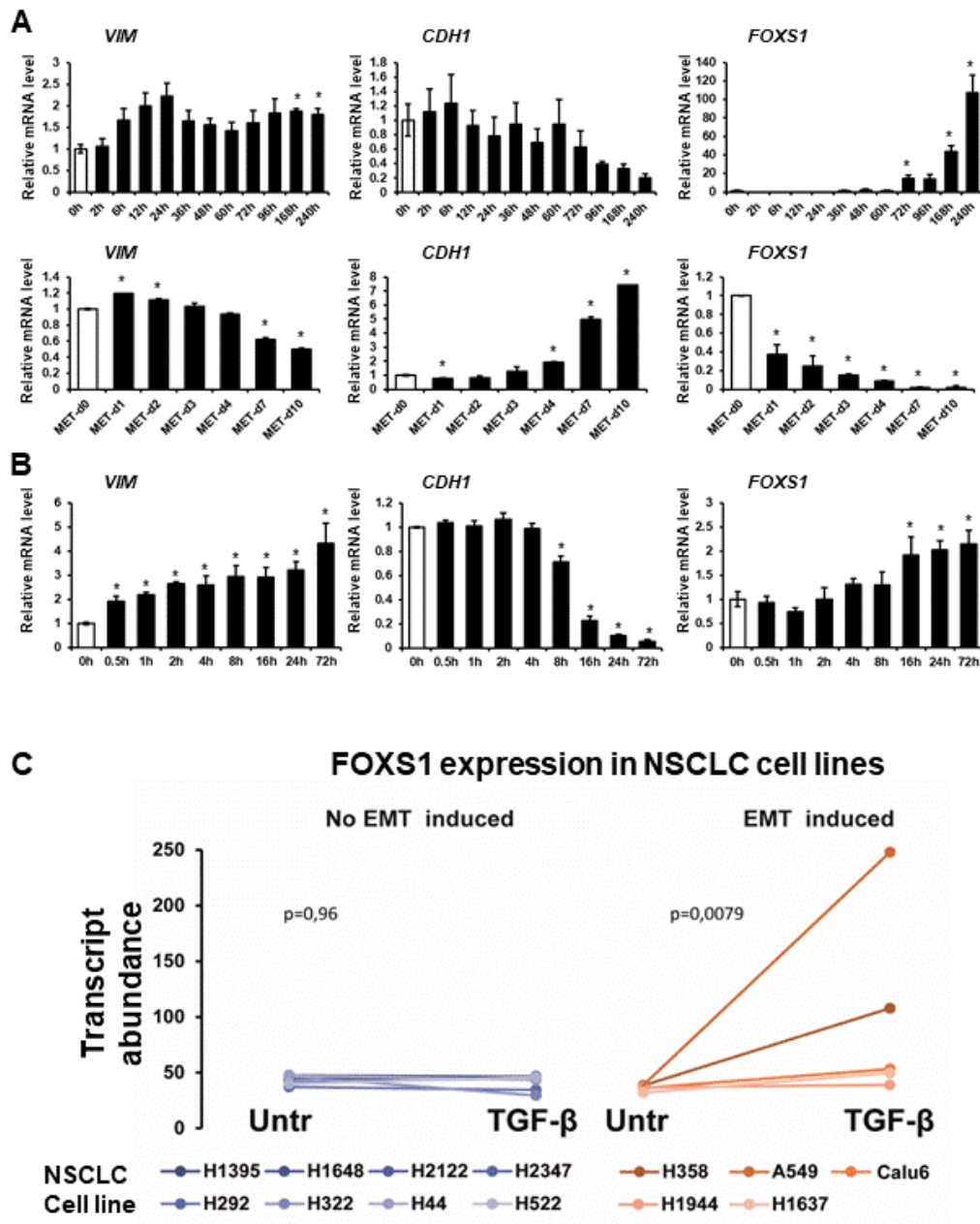


Figure supplémentaire 9 : Expression de *FOXS1* dans différents modèles d'induction de l'EMT.

A : Expression de *VIM*, *CDH1* et *FOXS1* dans une cinétique d'induction de l'EMT par le TGFβ suivi d'une MET dans la lignée murine NMuMG/E9. Panel du haut : Les cellules ont été traitées au TGFβ pour 0, 2h, 6h, 12h, 24h, 36h, 48h, 60h, 72h, 96h, 168h et 240h (RNA-

seq dataset : GSE112797). Panel du bas : Les cellules ont ensuite été privées de TGF β pendant 10 jours (RNA-seq dataset : GSE117474). L'analyse statistique a été réalisée par un t.test.

B : Expression de *VIM*, *CDH1* et *FOXS1* dans une cinétique d'induction de l'EMT par le TGF β dans la lignée A549 de cancer du poumon. Les cellules ont été traitées jusqu'à 72h par le TGF β . L'analyse statistique a été réalisée par un t.test.

C : Expression de *FOXS1* dans 13 lignées de cancers du poumon non à petites cellules traitées à 48 et/ou 120h au TGF β . Les lignées dans lesquelles le TGF β induit une EMT sont représentées en rouge et les lignées dans lesquelles aucune EMT n'est induite sont représentées en bleu (GSE114761). L'analyse statistique a été réalisée par un t.test.

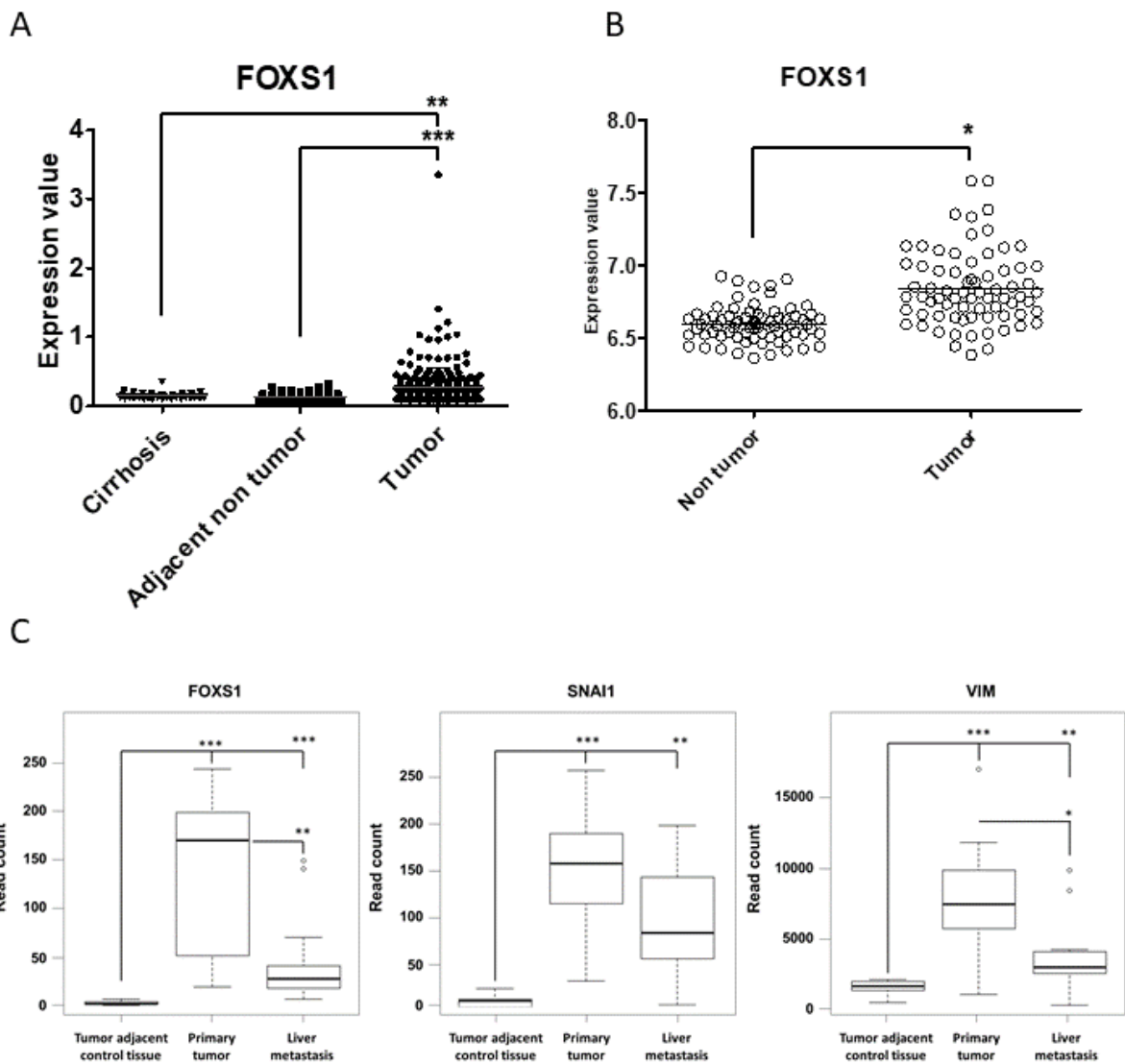


Figure supplémentaire 10 : L'expression de *FOXS1* est augmentée dans le CHC et les métastases par rapport au tissu non tumoral.

A : Expression de *FOXS1* dans le foie cirrhotique (n=40), le foie non tumoral adjacent à la tumeur (n=243) et la tumeur (n=266) à partir des données de GEO (GSE25097).

B : Expression de *FOXS1* dans 72 échantillons de CHC et de foie non tumoral associé (GSE39791)

C : Expression de *FOXS1* dans 14 échantillons d'adénocarcinome pancréatique ainsi que dans les métastases hépatiques et le tissu non tumoral correspondant (GSE151580)

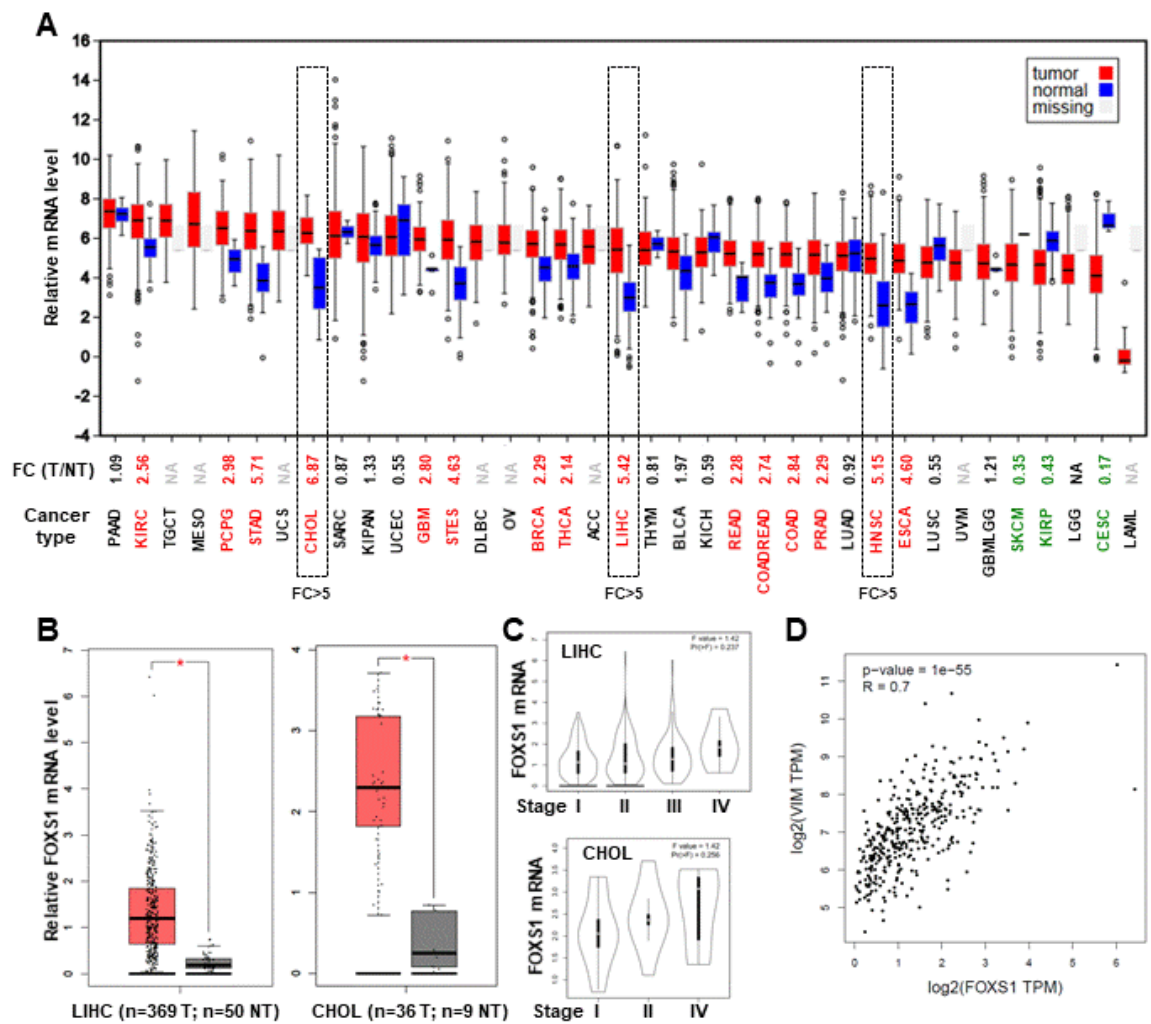


Figure supplémentaire 11 : Expression de *FOXS1* dans 37 cancers du TCGA.

A : Expression de *FOXS1* dans les différents cancers et les tissus sains associés. En rouge sont indiqué les cancers où *FOXS1* est surexprimé et en vert ceux où il est sous-exprimé par rapport au tissu sain.

B : Expression de *FOXS1* dans les tumeurs (rouge) ou le tissu non tumoral (gris) des cohortes LIHC et CHOL.

C : Expression de *FOXS1* selon le stade tumorale (I à IV)

D : Corrélation de l'expression de *FOXS1* et *VIM* dans le CHC (n=369)

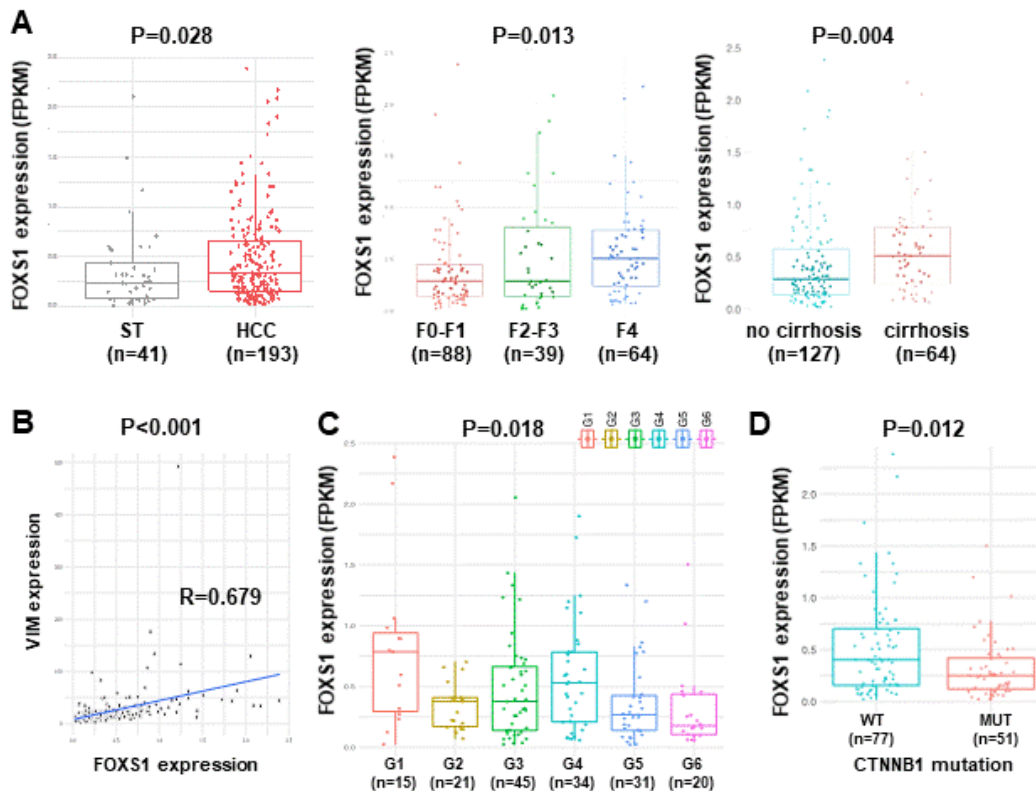


Figure supplémentaire 12 : Expression de *FOXS1* dans les CHC humain (n=193).

A : Expression de *FOXS1* dans 193 CHC comparée à 41 tissus environnants. Expression de *FOXS1* évaluée en fonction de l'histologie du tissu hépatique (score de fibrose et cirrhose)

B : Corrélation entre l'expression de *FOXS1* et *VIM* dans le CHC (n=193)

C : Expression de *FOXS1* selon la classification de Boyault et al, Hepatology, 2007.

D : Expression de *FOXS1* selon les tumeurs *CTNNB1* mutées ou non.

Les données de RNA-seq ont été extraites de Nault JC, et al. Gastroenterology. 2013 and Schullze K, et al. Nat. Genet., 2015.

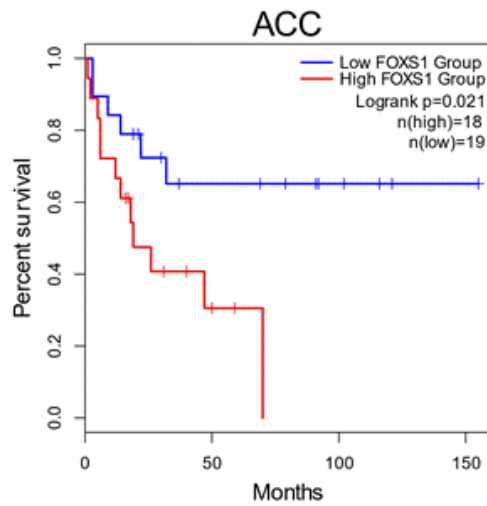
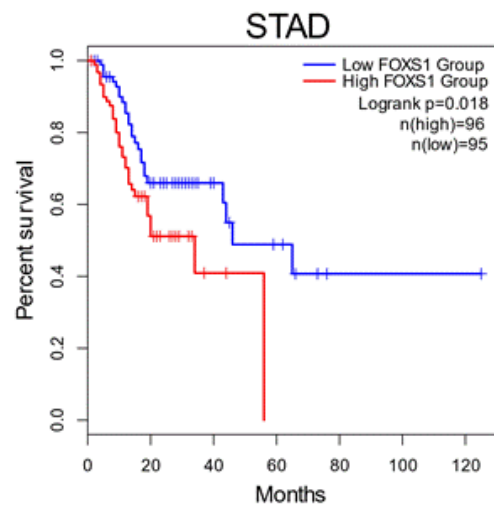
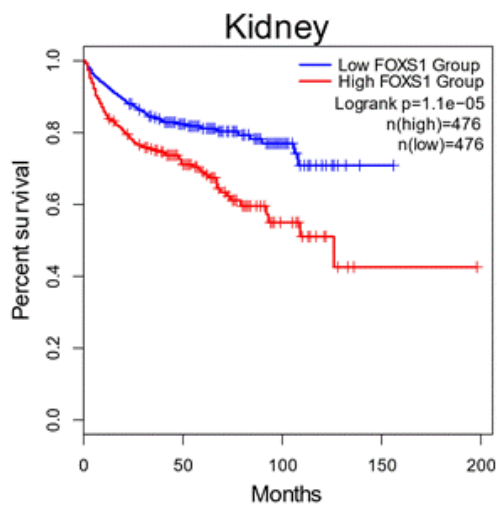
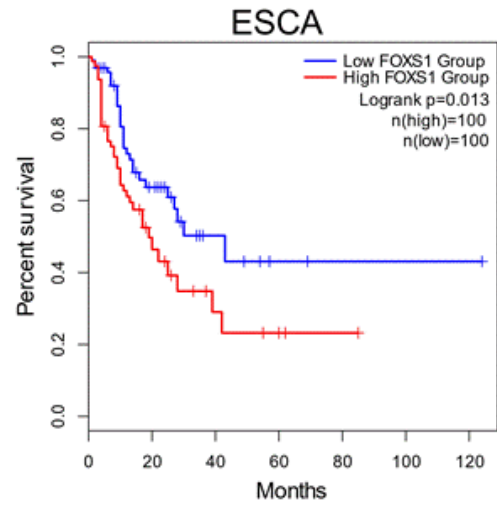
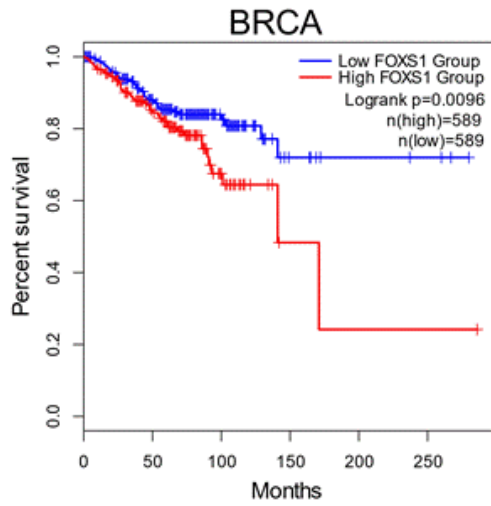


Figure supplémentaire 13 : L'expression de *FOXS1* est associée à une survie sans récurrence plus faible.

Kaplan-Meier plots et analyses statistiques montrant une diminution de la survie sans récurrence des patients présentant une expression élevée de *FOXS1*. La séparation des tumeurs est basée sur un seuil d'expression de *FOXS1* de 45/55 (seuil haut : 45% des tumeurs, seuil bas : 55%) dans les cancers de la vessie (BLCA), du sein (BRCA), du rein (KICH, KIRC, KIRP), et sur un seuil 75/25 (seuil haut : 3^{ème} quartile, seuil bas : 1^{er} quartile) dans les cancers de l'estomac (STAD) et de la glande corticosurrénale (ACC). Les données proviennent du TCGA par l'intermédiaire de GEPIA2.

Tables

Table 1 : Liste des gènes différentiellement exprimé dans les SNU-449 après 72h de traitement.

Gene symbol	Gene description	Accession nb.	P-value	Fold change (TGFβ/Ctrl)
nd	DA292495 BRHIP2 Homo sapiens cDNA clone BRHIP2004426 5', mRNA sequence [DA292495]	DA292495	0,0009	52,63
LDLRAD4	Homo sapiens low density lipoprotein receptor class A domain containing 4 (LDLRAD4), transcript variant a2, mRNA [NM_181482]	NM_181482	0,0001	27,03
LRRC32	Homo sapiens leucine rich repeat containing 32 (LRRC32), transcript variant 1, mRNA [NM_005512]	NM_005512	0,0007	15,38
NKAIN4	Homo sapiens Na ⁺ /K ⁺ transporting ATPase interacting 4 (NKAIN4), mRNA [NM_152864]	NM_152864	0,0001	15,15
LEFTY1	Homo sapiens left-right determination factor 1 (LEFTY1), mRNA [NM_020997]	NM_020997	0,0007	13,16
CRLF1	Homo sapiens cytokine receptor-like factor 1 (CRLF1), mRNA [NM_004750]	NM_004750	0,0003	11,11
MAF	Homo sapiens v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog (MAF), transcript variant 2, mRNA [NM_001031804]	NM_001031804	0,0007	10,75
NKAIN4	Homo sapiens Na ⁺ /K ⁺ transporting ATPase interacting 4 (NKAIN4), mRNA [NM_152864]	NM_152864	0,0002	10,53
FOXS1	Homo sapiens forkhead box S1 (FOXS1), mRNA [NM_004118]	NM_004118	0,0001	10,20
MDFI	Homo sapiens MyoD family inhibitor (MDFI), transcript variant 1, mRNA [NM_001300804]	NM_001300804	0,0001	7,69
P2RY2	Homo sapiens purinergic receptor P2Y, G-protein coupled, 2 (P2RY2), transcript variant 1, mRNA [NM_176072]	NM_176072	0,0004	7,69
COL1A1	Homo sapiens collagen, type I, alpha 1 (COL1A1), mRNA [NM_000088]	NM_000088	0,0001	6,67
CGB	Homo sapiens chorionic gonadotropin, beta polypeptide (CGB), mRNA [NM_000737]	NM_000737	0,0001	5,56
EGR2	Homo sapiens early growth response 2 (EGR2), transcript variant 1, mRNA [NM_000399]	NM_000399	0,0002	5,26
LMO1	Homo sapiens LIM domain only 1 (rhombotin 1) (LMO1), transcript variant 1, mRNA [NM_002315]	NM_002315	0,0003	5,00

SUCNR1	Homo sapiens succinate receptor 1 (SUCNR1), mRNA [NM_033050]	NM_033050	0,0003	4,76
SGCA	Homo sapiens sarcoglycan, alpha (50kDa dystrophin-associated glycoprotein) (SGCA), transcript variant 1, mRNA [NM_000023]	NM_000023	0,0000	4,17
DACT1	Homo sapiens dishevelled-binding antagonist of beta-catenin 1 (DACT1), transcript variant 1, mRNA [NM_016651]	NM_016651	0,0000	4,00
ZNF365	Homo sapiens zinc finger protein 365 (ZNF365), transcript variant A, mRNA [NM_014951]	NM_014951	0,0000	4,00
LOC101927230	Homo sapiens uncharacterized LOC101927230 (LOC101927230), long non-coding RNA [NR_110806]		0,0003	3,85
NXPH4	Homo sapiens neurexophilin 4 (NXPH4), mRNA [NM_007224]	NM_007224	0,0002	3,85
LOC100507431	Homo sapiens uncharacterized LOC100507431 (LOC100507431), long non-coding RNA [NR_125383]	NR_125383	0,0001	3,70
LOC101928173	PREDICTED: Homo sapiens uncharacterized LOC101928173 (LOC101928173), transcript variant X2, ncRNA [XR_245039]	XR_245039	0,0006	3,57
PDGFA	Homo sapiens platelet-derived growth factor alpha polypeptide (PDGFA), transcript variant 2, mRNA [NM_033023]	NM_033023	0,0006	3,57
ZNF365	Homo sapiens zinc finger protein 365 (ZNF365), transcript variant B, mRNA [NM_199450]	NM_199450	0,0007	3,57
HTR1D	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 1D, G protein-coupled (HTR1D), mRNA [NM_000864]	NM_000864	0,0001	3,33
ULK1	Homo sapiens unc-51 like autophagy activating kinase 1 (ULK1), mRNA [NM_003565]	NM_003565	0,0001	3,33
DNM3	Homo sapiens dynamin 3 (DNM3), transcript variant 1, mRNA [NM_015569]	NM_015569	0,0000	3,23
XLOC_I2_007543	BROAD Institute lincRNA (XLOC_I2_007543), lincRNA [TCONS_I2_00013972]		0,0007	3,23
ALOX5AP	Homo sapiens arachidonate 5-lipoxygenase-activating protein (ALOX5AP), transcript variant 1, mRNA [NM_001629]	NM_001629	0,0009	3,13
LTBP2	Homo sapiens latent transforming growth factor beta binding protein 2 (LTBP2), mRNA [NM_000428]	NM_000428	0,0002	3,13
MAFB	Homo sapiens v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB), mRNA [NM_005461]	NM_005461	0,0004	3,03

COL5A1	Homo sapiens collagen, type V, alpha 1 (COL5A1), transcript variant 1, mRNA [NM_000093]	NM_000093	0,0000	2,94
DFNB31	Homo sapiens deafness, autosomal recessive 31 (DFNB31), transcript variant 1, mRNA [NM_015404]	NM_015404	0,0002	2,94
SAMD11	Homo sapiens sterile alpha motif domain containing 11 (SAMD11), mRNA [NM_152486]	NM_152486	0,0002	2,94
SIM2	Homo sapiens single-minded family bHLH transcription factor 2 (SIM2), transcript variant SIM2, mRNA [NM_005069]	NM_005069	0,0001	2,94
ECE1	Homo sapiens endothelin converting enzyme 1 (ECE1), transcript variant 3, mRNA [NM_001113347]	NM_001113347	0,0003	2,86
PMEPA1	Homo sapiens prostate transmembrane protein, androgen induced 1 (PMEPA1), transcript variant 1, mRNA [NM_020182]	NM_020182	0,0004	2,86
PMEPA1	Homo sapiens prostate transmembrane protein, androgen induced 1 (PMEPA1), transcript variant 1, mRNA [NM_020182]	NM_020182	0,0000	2,86
PRSS8	Homo sapiens protease, serine, 8 (PRSS8), mRNA [NM_002773]	NM_002773	0,0003	2,86
ABLIM3	Homo sapiens actin binding LIM protein family, member 3 (ABLIM3), transcript variant 2, mRNA [NM_014945]	NM_014945	0,0000	2,78
ADAM19	Homo sapiens ADAM metallopeptidase domain 19 (ADAM19), mRNA [NM_033274]	NM_033274	0,0002	2,78
PRDM1	Homo sapiens PR domain containing 1, with ZNF domain (PRDM1), transcript variant 1, mRNA [NM_001198]	NM_001198	0,0001	2,78
SIK1	Homo sapiens salt-inducible kinase 1 (SIK1), mRNA [NM_173354]	NM_173354	0,0009	2,78
SKIL	Homo sapiens SKI-like proto-oncogene (SKIL), transcript variant 1, mRNA [NM_005414]	NM_005414	0,0001	2,78
COL5A1	Homo sapiens collagen, type V, alpha 1 (COL5A1), transcript variant 1, mRNA [NM_000093]	NM_000093	0,0002	2,70
C5orf46	Homo sapiens chromosome 5 open reading frame 46 (C5orf46), mRNA [NM_206966]	NM_206966	0,0006	2,63
lnc-C11orf44-1	LNCipedia lincRNA (lnc-C11orf44-1), lincRNA [lnc-C11orf44-1:1]		0,0006	2,63
CASC15	Homo sapiens cancer susceptibility candidate 15 (non-protein coding) (CASC15), long non-coding RNA [NR_015410]	NR_015410	0,0001	2,56

GRPR	Homo sapiens gastrin-releasing peptide receptor (GRPR), mRNA [NM_005314]	NM_005314	0,0005	2,56
CASC15	Homo sapiens cancer susceptibility candidate 15 (non-protein coding) (CASC15), long non-coding RNA [NR_015410]	NR_015410	0,0010	2,50
FSTL3	Homo sapiens follistatin-like 3 (secreted glycoprotein) (FSTL3), mRNA [NM_005860]	NM_005860	0,0007	2,50
KLF12	Kruppel-like factor 12 [Source:HGNC Symbol;Acc:HGNC:6346] [ENST00000472022]	AF113122	0,0001	2,50
SPINK1	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA [NM_003122]	NM_003122	0,0001	2,50
nd	Homo sapiens mRNA; cDNA DKFZp686P18215 (from clone DKFZp686P18215). [BX538202]		0,0005	2,50
LOC101929787	PREDICTED: Homo sapiens uncharacterized LOC101929787 (LOC101929787), transcript variant X1, ncRNA [XR_250866]	XR_250866	0,0001	2,44
LOC79160	Homo sapiens uncharacterized LOC79160 (LOC79160), long non-coding RNA [NR_125359]	NR_125359	0,0001	2,44
PRDM1	Homo sapiens PR domain containing 1, with ZNF domain (PRDM1), transcript variant 1, mRNA [NM_001198]	NM_001198	0,0004	2,44
RTKN2	Homo sapiens rhotekin 2 (RTKN2), transcript variant 1, mRNA [NM_145307]	NM_145307	0,0001	2,44
TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0,0004	2,44
LFNG	Homo sapiens LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase (LFNG), transcript variant 1, mRNA [NM_001040167]	NM_001040167	0,0003	2,33
PLEK2	Homo sapiens pleckstrin 2 (PLEK2), mRNA [NM_016445]	NM_016445	0,0005	2,33
SPOCK1	Homo sapiens sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 1 (SPOCK1), mRNA [NM_004598]	NM_004598	0,0005	2,33
LIMS3L	Homo sapiens LIM and senescent cell antigen-like domains 3-like (LIMS3L), transcript variant 2, non-coding RNA [NR_038099]	NR_038099	0,0002	2,27
SPSB1	Homo sapiens splA/ryanodine receptor domain and SOCS box containing 1 (SPSB1), mRNA [NM_025106]	NM_025106	0,0002	2,27
TGFBI	Homo sapiens transforming growth factor, beta-induced, 68kDa (TGFBI), mRNA [NM_000358]	NM_000358	0,0000	2,27

TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0,0000	2,27
AMIGO2	Homo sapiens adhesion molecule with Ig-like domain 2 (AMIGO2), transcript variant 2, mRNA [NM_181847]	NM_181847	0,0002	2,22
RTKN2	Homo sapiens rhotekin 2 (RTKN2), transcript variant 2, mRNA [NM_001282941]	NM_001282941	0,0002	2,22
SSC5D	Homo sapiens scavenger receptor cysteine rich family, 5 domains (SSC5D), transcript variant 1, mRNA [NM_001144950]	NM_001144950	0,0009	2,22
TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0,0005	2,22
KLF7	Homo sapiens Kruppel-like factor 7 (ubiquitous) (KLF7), transcript variant 1, mRNA [NM_003709]	NM_003709	0,0002	2,17
KPNA7	Homo sapiens karyopherin alpha 7 (importin alpha 8) (KPNA7), mRNA [NM_001145715]	NM_001145715	0,0003	2,17
LRIG3	Homo sapiens leucine-rich repeats and immunoglobulin-like domains 3 (LRIG3), transcript variant 2, mRNA [NM_153377]	NM_153377	0,0003	2,17
nd	long intergenic non-protein coding RNA 313 [Source:HGNC Symbol;Acc:HGNC:16416] [ENST00000427188]	XR_425284	0,0001	2,17
EPHB2	Homo sapiens EPH receptor B2 (EPHB2), transcript variant 2, mRNA [NM_004442]	NM_004442	0,0000	2,13
ITGB6	Homo sapiens integrin, beta 6 (ITGB6), transcript variant 1, mRNA [NM_000888]	NM_000888	0,0000	2,13
MEG3	Homo sapiens maternally expressed 3 (non-protein coding) (MEG3), transcript variant 1, long non-coding RNA [NR_002766]	NR_002766	0,0000	2,13
AADAC	Homo sapiens arylacetamide deacetylase (AADAC), mRNA [NM_001086]	NM_001086	0,0005	2,08
ARSI	Homo sapiens arylsulfatase family, member I (ARSI), mRNA [NM_001012301]	NM_001012301	0,0002	2,08
BMF	Homo sapiens Bcl2 modifying factor (BMF), transcript variant 1, mRNA [NM_001003940]	NM_001003940	0,0001	2,08
MAP3K2	Homo sapiens mitogen-activated protein kinase kinase 2 (MAP3K2), mRNA [NM_006609]	NM_006609	0,0001	2,08
CYTH1	Homo sapiens cytohesin 1 (CYTH1), transcript variant 1, mRNA [NM_004762]	NM_004762	0,0001	2,04
CYTH1	Homo sapiens cytohesin 1 (CYTH1), transcript variant 3, mRNA [NM_001292018]	NM_001292018	0,0002	2,04

EEPD1	Homo sapiens endonuclease/exonuclease/phosphatase family domain containing 1 (EEPD1), mRNA [NM_030636]	NM_030636	0,0002	2,04
EPHB2	Homo sapiens EPH receptor B2 (EPHB2), transcript variant 2, mRNA [NM_004442]	NM_004442	0,0000	2,04
LIMS1	Homo sapiens LIM and senescent cell antigen-like domains 1 (LIMS1), transcript variant 2, mRNA [NM_004987]	NM_004987	0,0000	2,04
PGM2L1	Homo sapiens phosphoglucomutase 2-like 1 (PGM2L1), mRNA [NM_173582]	NM_173582	0,0002	2,04
ABCA1	Homo sapiens ATP-binding cassette, sub-family A (ABC1), member 1 (ABCA1), mRNA [NM_005502]	NM_005502	0,0005	2,00
KLRC1	Homo sapiens killer cell lectin-like receptor subfamily C, member 1 (KLRC1), transcript variant 1, mRNA [NM_002259]	NM_002259	0,0006	0,49
SLC22A1 8AS	Homo sapiens solute carrier family 22 (organic cation transporter), member 18 antisense (SLC22A18AS), transcript variant 1, mRNA [NM_007105]	NM_007105	0,0007	0,49
TRIM55	Homo sapiens tripartite motif containing 55 (TRIM55), transcript variant 2, mRNA [NM_033058]	NM_033058	0,0003	0,48
LOC1019 29641	PREDICTED: Homo sapiens uncharacterized LOC101929641 (LOC101929641), ncRNA [XR_243223]	XR_243223	0,0003	0,45
KCNH5	Homo sapiens potassium channel, voltage gated eag related subfamily H, member 5 (KCNH5), transcript variant 3, mRNA [NM_172375]	NM_172375	0,0002	0,01

Table 2 : Caractéristiques des lignées cellulaires humaines de CHC.

	PLC/PR F/5	HepG2-C3A	HuH6	Hep3B	SNU-398	SNU-423	SNU-449	SNU-475
Gender	M	M	M	M	M	M	M	M
Age (years)	24	15	1	8	42	40	52	43
Histology	hepatocellular carcinoma	hepatoblastoma	hepatoblastoma	hepatocellular carcinoma	hepatocellular carcinoma	hepatocellular carcinoma	hepatocellular carcinoma	hepatocellular carcinoma
Differentiation (Ed.&St.'s grade)	well-differentiated	well-differentiated	na	well-differentiated	anaplastic	grade III/IV	grade II-III/IV	grade II-IV/V
HBV status	HBV+	HBV-	HBV-	HBV+	HBV+	HBV+	HBV+	HBV+
Origin	Mozambique	Argentina	Japan	USA	Korea	Korea	Korea	Korea
E-Cadherin (WB/IF)	High	High	High	High	Absent	Absent	Absent	Low/Absent
Vimentin (WB/IF)	Low/Absent	Low/Absent	Low/Absent	Low	High	High	High	High
Phenotype	Epithelial-like	Epithelial-like	Epithelial-like	Epithelial-like	Mesenchymal-like	Mesenchymal-like	Mesenchymal-like	Mesenchymal-like
TERT promoter mut.	wt	- 124G>A	wt	wt	- 124G>A	- 124G>A	wt	- 124G>A
TP53 mut.	R249S, G199*	wt	A159D, N239D	wt	E224E	slice site at position 7578556	A161T	N239D, G262D
CTNNB1 mut.	wt	wt	G34V	wt	S37C	wt	wt	wt
ARID1A mut.	wt	wt	wt	wt	wt	G623*	1334_1335insQ, E2033fs	wt
ARID2 mut.	wt	wt	wt	wt	wt	-	H353P	wt
AXIN1 mut.	wt	wt	wt	R146*	wt	PYRIR3 45fs	R712*	wt
CDKN2A mut.	R112G	wt	wt	G122V	wt	wt	wt	wt
RBI mut.	wt	wt	wt	S576*	wt	wt	wt	wt
SMAD7 mut.	wt	wt	I317V	wt	wt	wt	wt	wt
TGFBI mut.	G371G (silent)	wt	wt	wt	wt	wt	wt	wt

Table 3 : Liste des gènes inclus dans la signature siFOXSI.

Gene symbol	TGFβ-untreated cells				TGFβ-treated cells			
	siFOXSI-1 vs. siNC		siFOXSI-2 vs. siNC		siFOXSI-1 vs. siNC		siFOXSI-2 vs. siNC	
	FC	P-value	FC	P-value	FC	P-value	FC	P-value
FOXSI	0,11	0,0004	0,02	0,0000	0,09	0,0004	0,02	0,0000
THBD	0,40	0,0001	0,28	0,0000	0,34	0,0000	0,25	0,0001
C19orf33	0,32	0,0000	0,29	0,0000	0,29	0,0001	0,27	0,0007
SPTB	0,26	0,0010	0,28	0,0001	0,33	0,0006	0,27	0,0000
LOC400743	0,27	0,0000	0,35	0,0000	0,29	0,0000	0,36	0,0000
TBC1D2	0,33	0,0001	0,37	0,0001	0,34	0,0000	0,40	0,0002
ALOX12-AS1	0,60	0,0008	0,55	0,0008	0,50	0,0009	0,41	0,0006
NAP1L4	0,36	0,0006	0,43	0,0000	0,34	0,0000	0,45	0,0000
Inc-SCCPDH-1	0,43	0,0000	0,37	0,0000	0,46	0,0003	0,46	0,0002
ZNF192P1	0,31	0,0000	0,37	0,0000	0,29	0,0000	0,46	0,0003
SPHAR	0,59	0,0002	0,50	0,0001	0,61	0,0002	0,49	0,0001
MPP2	0,54	0,0008	0,54	0,0001	0,62	0,0004	0,52	0,0001
ZNF334	0,31	0,0000	0,51	0,0006	0,36	0,0008	0,52	0,0004
AMMECR1L	0,62	0,0000	0,59	0,0001	0,63	0,0003	0,54	0,0002
RPL37	0,58	0,0006	0,62	0,0007	0,60	0,0008	0,55	0,0003
OSBPL5	0,44	0,0008	0,59	0,0002	0,47	0,0001	0,57	0,0004
TBC1D16	0,34	0,0007	0,56	0,0001	0,38	0,0001	0,57	0,0007
THRA	0,44	0,0009	0,58	0,0001	0,42	0,0001	0,57	0,0007
Inc-DYNC1I2-1	0,44	0,0004	0,60	0,0008	0,50	0,0005	0,58	0,0004
KLF13	0,44	0,0001	0,64	0,0006	0,47	0,0001	0,59	0,0007
LINC01481	0,52	0,0001	0,57	0,0001	0,59	0,0001	0,59	0,0005
PTPLA	0,62	0,0000	0,62	0,0002	0,61	0,0007	0,59	0,0007
JDP2	0,53	0,0005	0,59	0,0002	0,59	0,0010	0,60	0,0004
PTH2R	0,58	0,0003	0,59	0,0002	0,56	0,0003	0,60	0,0006
PTGFRN	0,61	0,0008	0,60	0,0001	0,61	0,0003	0,61	0,0007
IFI44	0,48	0,0010	2,14	0,0003	0,39	0,0001	2,22	0,0008
FBXO33	1,78	0,0003	1,57	0,0003	1,81	0,0002	1,55	0,0007
MLX	2,24	0,0000	1,61	0,0001	2,10	0,0000	1,55	0,0006
RAP1B	1,93	0,0005	1,68	0,0005	1,89	0,0001	1,55	0,0007
ZCCHC2	2,67	0,0001	1,63	0,0005	2,62	0,0000	1,55	0,0009
TSPYL1	1,64	0,0001	1,57	0,0003	1,60	0,0003	1,58	0,0007
RB1	2,01	0,0001	1,63	0,0003	1,97	0,0001	1,60	0,0007
C5orf15	1,91	0,0001	1,63	0,0002	1,96	0,0001	1,61	0,0007
CAB39	1,94	0,0000	1,75	0,0001	2,12	0,0000	1,62	0,0008
ADAM17	1,71	0,0001	1,67	0,0002	1,71	0,0002	1,63	0,0004
EIF2A	2,36	0,0004	1,75	0,0003	2,05	0,0001	1,63	0,0005
RASSF1	1,89	0,0007	1,71	0,0002	1,78	0,0003	1,63	0,0009
HNRNPH2	1,68	0,0003	1,64	0,0001	1,75	0,0004	1,64	0,0009
HS2ST1	2,20	0,0001	1,64	0,0001	2,23	0,0001	1,64	0,0007
DERL1	2,17	0,0000	1,72	0,0001	2,16	0,0000	1,65	0,0005
DNAJA3	1,80	0,0000	1,63	0,0001	1,82	0,0001	1,65	0,0005
MRPL22	1,78	0,0001	1,69	0,0002	1,73	0,0002	1,65	0,0010

RAP1GDS1	2,74	0,0001	1,82	0,0002	2,71	0,0000	1,66	0,0006
PDPK1	1,60	0,0003	1,65	0,0002	1,96	0,0000	1,68	0,0005
MCTS1	1,70	0,0000	1,67	0,0002	1,62	0,0006	1,69	0,0009
SGTB	1,73	0,0000	1,84	0,0000	1,79	0,0001	1,69	0,0003
TMEM248	1,82	0,0000	1,81	0,0000	1,87	0,0001	1,69	0,0003
WDR82	1,86	0,0005	1,84	0,0003	1,83	0,0003	1,69	0,0008
YWHAEP1	1,73	0,0004	1,63	0,0004	1,72	0,0001	1,69	0,0002
SLC25A13	2,35	0,0000	1,71	0,0001	2,46	0,0000	1,70	0,0004
CDC23	2,15	0,0000	1,72	0,0001	2,21	0,0000	1,71	0,0004
MMS22L	2,24	0,0003	1,70	0,0006	2,33	0,0000	1,71	0,0009
PPP2R3C	1,84	0,0000	1,76	0,0000	1,85	0,0001	1,71	0,0007
ATP6V1A	2,05	0,0001	1,81	0,0000	1,99	0,0000	1,74	0,0005
IPPK	2,07	0,0000	1,75	0,0004	2,13	0,0000	1,75	0,0007
MAPRE1	2,28	0,0002	1,87	0,0001	2,29	0,0000	1,75	0,0009
TCF7	2,85	0,0002	1,77	0,0002	4,05	0,0000	1,75	0,0002
ANXA7	2,26	0,0001	1,87	0,0001	2,17	0,0000	1,76	0,0005
ARL8B	2,41	0,0001	1,78	0,0001	2,56	0,0000	1,76	0,0003
MBD1	2,19	0,0000	1,83	0,0001	2,10	0,0001	1,76	0,0003
RPAP3	1,78	0,0003	1,73	0,0001	1,84	0,0004	1,76	0,0010
ATP6V1E2	1,94	0,0005	1,83	0,0002	1,81	0,0004	1,77	0,0010
BRCC3	1,99	0,0000	1,77	0,0000	1,86	0,0001	1,77	0,0006
GCNT3	2,56	0,0000	1,64	0,0002	2,22	0,0000	1,77	0,0002
GMPR2	1,93	0,0002	1,79	0,0001	1,91	0,0002	1,77	0,0005
KIAA1715	1,64	0,0002	1,73	0,0001	1,61	0,0004	1,77	0,0002
BRCC3	2,03	0,0000	1,78	0,0001	1,92	0,0002	1,78	0,0003
PPP2R1B	2,11	0,0000	1,74	0,0001	2,04	0,0001	1,78	0,0004
ANKRD13C	2,28	0,0002	1,77	0,0002	2,55	0,0000	1,79	0,0005
UBE2O	1,64	0,0007	1,82	0,0001	1,83	0,0001	1,79	0,0002
PDCD6IP	1,81	0,0001	1,94	0,0002	1,76	0,0003	1,80	0,0007
ETV1	1,92	0,0002	1,67	0,0002	1,77	0,0008	1,82	0,0005
C1GALT1	1,86	0,0001	1,85	0,0001	1,95	0,0002	1,83	0,0006
ENTPD4	2,02	0,0007	1,90	0,0001	1,98	0,0002	1,83	0,0003
HNRNPR	1,97	0,0008	1,95	0,0002	1,98	0,0002	1,84	0,0007
MAP2K1	2,03	0,0000	1,86	0,0000	2,24	0,0000	1,84	0,0001
PVRL2	2,35	0,0001	1,91	0,0001	2,30	0,0000	1,84	0,0002
MSANTD3	1,54	0,0003	1,97	0,0000	1,53	0,0007	1,85	0,0001
TMTC3	1,77	0,0000	1,85	0,0000	1,91	0,0001	1,86	0,0004
ANKRD44	2,52	0,0001	2,13	0,0000	2,37	0,0000	1,87	0,0004
FAM135A	1,84	0,0000	2,09	0,0000	2,04	0,0001	1,87	0,0008
MTMR2	1,96	0,0000	1,95	0,0000	2,01	0,0001	1,87	0,0006
PPAT	1,85	0,0000	1,80	0,0001	1,90	0,0002	1,88	0,0005
ZAK	2,59	0,0001	1,71	0,0004	2,59	0,0000	1,88	0,0002
RBSN	2,15	0,0001	1,92	0,0000	2,14	0,0001	1,89	0,0003
EFCAB14	2,21	0,0001	1,95	0,0000	2,38	0,0001	1,90	0,0007
CEP19	1,76	0,0004	1,74	0,0001	1,79	0,0003	1,91	0,0002
DUS2	1,87	0,0000	1,71	0,0002	1,84	0,0002	1,91	0,0001
ZBTB9	2,09	0,0001	1,99	0,0000	2,18	0,0000	1,91	0,0001
LAP3	1,78	0,0005	1,95	0,0000	1,85	0,0002	1,92	0,0004

MFAP3	1,87	0,0008	1,94	0,0001	1,91	0,0000	1,92	0,0001
LOXL2	2,23	0,0000	2,06	0,0000	2,31	0,0000	1,94	0,0001
SCML1	3,30	0,0001	1,82	0,0001	3,59	0,0000	1,94	0,0003
TEX15	1,71	0,0002	1,54	0,0005	2,22	0,0004	1,94	0,0006
ENTPD4	2,15	0,0010	1,98	0,0001	2,06	0,0004	1,95	0,0002
CRYZ	1,67	0,0003	1,82	0,0001	1,76	0,0002	1,96	0,0001
SUV39H2	1,85	0,0003	2,08	0,0000	1,69	0,0004	1,96	0,0004
TTC27	1,56	0,0001	1,97	0,0000	1,61	0,0003	1,96	0,0003
C1GALT1	1,93	0,0002	1,88	0,0001	2,04	0,0001	1,97	0,0002
PROSC	2,06	0,0004	2,07	0,0001	1,90	0,0007	1,98	0,0005
LOC101059954	3,09	0,0004	2,06	0,0000	2,41	0,0008	1,99	0,0007
NAPG	2,67	0,0000	2,07	0,0000	2,70	0,0000	1,99	0,0001
PRRC2C	2,11	0,0000	2,09	0,0000	2,16	0,0000	1,99	0,0001
ZAK	3,17	0,0002	2,05	0,0000	2,76	0,0001	1,99	0,0004
HAUS6	2,91	0,0001	2,06	0,0001	2,72	0,0000	2,01	0,0004
IL1RAP	3,88	0,0002	2,08	0,0001	4,13	0,0000	2,02	0,0002
SLC25A44	2,08	0,0009	2,04	0,0002	2,18	0,0001	2,02	0,0004
BTG3	1,97	0,0001	2,00	0,0001	2,05	0,0001	2,04	0,0001
CUTC	2,39	0,0001	2,10	0,0001	1,96	0,0001	2,05	0,0002
SMIM13	1,86	0,0003	2,08	0,0000	2,13	0,0000	2,06	0,0001
AIDA	2,02	0,0004	2,19	0,0000	2,05	0,0001	2,07	0,0001
B4GALT4	2,74	0,0000	2,10	0,0000	2,56	0,0000	2,07	0,0002
HDAC1	2,71	0,0001	2,19	0,0001	3,33	0,0000	2,07	0,0005
PHTF2	2,44	0,0000	2,07	0,0001	2,60	0,0001	2,07	0,0002
NRAS	1,86	0,0001	2,19	0,0000	1,86	0,0001	2,08	0,0001
RELL1	3,00	0,0005	2,37	0,0003	2,60	0,0001	2,10	0,0008
KLHL2	2,96	0,0001	2,32	0,0001	2,79	0,0000	2,11	0,0002
TMEM64	2,12	0,0003	2,04	0,0001	2,51	0,0000	2,12	0,0003
PPIF	2,44	0,0000	2,13	0,0000	2,59	0,0000	2,15	0,0005
RNF14	2,59	0,0001	2,58	0,0000	2,48	0,0001	2,19	0,0004
SLC31A1	2,32	0,0001	2,32	0,0000	2,03	0,0000	2,19	0,0001
NUP98	1,93	0,0004	2,27	0,0000	1,84	0,0002	2,20	0,0003
STBD1	3,85	0,0000	2,49	0,0000	3,97	0,0000	2,21	0,0000
UBLCP1	2,36	0,0000	2,19	0,0000	2,38	0,0000	2,21	0,0003
GPD1L	2,37	0,0001	2,04	0,0000	2,23	0,0002	2,24	0,0001
XLOC_I2_010916	2,38	0,0004	2,50	0,0000	2,34	0,0001	2,24	0,0001
RDX	4,05	0,0002	2,51	0,0002	3,91	0,0000	2,26	0,0007
TMTC3	1,80	0,0003	1,86	0,0000	2,21	0,0003	2,27	0,0002
CYHR1	3,24	0,0001	2,60	0,0000	3,22	0,0000	2,28	0,0001
ZNF788	2,23	0,0003	2,13	0,0000	2,23	0,0001	2,28	0,0002
UMPS	2,18	0,0001	2,32	0,0001	2,07	0,0001	2,29	0,0002
KDELR2	2,02	0,0000	2,37	0,0000	2,03	0,0000	2,30	0,0001
CLDN1	2,02	0,0000	1,99	0,0001	1,78	0,0004	2,31	0,0000
SFXN4	2,08	0,0000	2,40	0,0000	1,96	0,0001	2,31	0,0000
AXIN2	2,33	0,0004	1,97	0,0001	2,65	0,0002	2,32	0,0006
MFAP3	2,26	0,0000	2,54	0,0000	2,06	0,0001	2,32	0,0001
SYPL1	2,47	0,0000	2,46	0,0000	2,55	0,0000	2,32	0,0001
CTNNB1	2,71	0,0000	2,47	0,0000	2,61	0,0000	2,34	0,0000

ERLIN1	2,49	0,0001	2,28	0,0000	2,30	0,0001	2,36	0,0002
LOC101926940	2,70	0,0001	2,17	0,0000	2,51	0,0000	2,36	0,0002
SEC24A	3,30	0,0000	2,55	0,0000	3,14	0,0000	2,37	0,0000
CIDCEP	3,89	0,0000	2,33	0,0000	3,44	0,0000	2,39	0,0000
BCAP29	2,41	0,0002	2,63	0,0000	2,71	0,0000	2,42	0,0001
NPTN	3,08	0,0003	2,60	0,0000	2,93	0,0001	2,46	0,0003
CXorf56	2,54	0,0003	2,38	0,0000	2,42	0,0002	2,47	0,0002
ID2	2,01	0,0001	2,58	0,0000	2,30	0,0000	2,48	0,0000
GNA13	2,67	0,0001	2,61	0,0000	2,74	0,0001	2,49	0,0001
CTNNB1	2,37	0,0000	2,42	0,0000	2,49	0,0000	2,50	0,0000
CIDEC	4,07	0,0002	2,58	0,0000	4,09	0,0000	2,52	0,0001
FAM86FP	1,76	0,0001	2,40	0,0000	1,91	0,0001	2,61	0,0001
RNF14	2,92	0,0000	2,92	0,0000	2,90	0,0000	2,65	0,0000
CIDEC	4,24	0,0002	2,75	0,0000	3,59	0,0000	2,68	0,0001
MRPL19	2,46	0,0002	2,65	0,0000	2,34	0,0002	2,71	0,0001
ERLIN2	4,44	0,0000	2,81	0,0000	4,12	0,0000	2,73	0,0001
LOC102724529	3,18	0,0004	2,63	0,0000	3,46	0,0000	2,73	0,0001
CXorf56	2,94	0,0002	2,79	0,0000	2,82	0,0000	2,74	0,0000
C12orf76	3,22	0,0000	2,99	0,0000	2,99	0,0001	2,75	0,0001
IKBIP	2,75	0,0002	3,02	0,0000	2,86	0,0001	2,79	0,0001
EIF2S1	3,06	0,0000	2,81	0,0000	3,10	0,0000	2,81	0,0001
RAB8B	2,45	0,0002	2,79	0,0000	2,43	0,0000	2,85	0,0000
SLC39A14	2,45	0,0000	2,78	0,0000	3,64	0,0000	2,85	0,0000
PGM2L1	3,83	0,0000	3,48	0,0000	4,47	0,0000	2,87	0,0000
GNA13	2,98	0,0002	2,91	0,0000	3,35	0,0000	2,88	0,0000
NUP35	4,47	0,0001	2,98	0,0000	4,47	0,0000	2,88	0,0001
FECH	3,04	0,0003	3,26	0,0000	2,53	0,0003	2,89	0,0002
YWHAG	2,34	0,0001	3,11	0,0000	2,15	0,0000	2,93	0,0000
NARF	3,07	0,0000	3,34	0,0000	2,81	0,0000	3,02	0,0000
RTCA	2,88	0,0001	3,11	0,0000	2,78	0,0000	3,05	0,0000
COMMD10	4,62	0,0001	3,50	0,0000	4,16	0,0000	3,07	0,0001
RDH11	3,90	0,0000	3,58	0,0000	3,92	0,0000	3,09	0,0000
ETF1	3,30	0,0000	3,16	0,0000	3,39	0,0000	3,14	0,0000
ITGA6	3,44	0,0002	2,97	0,0000	3,10	0,0000	3,14	0,0000
NUP35	4,69	0,0000	3,10	0,0000	4,77	0,0000	3,16	0,0000
SWAP70	2,57	0,0005	3,20	0,0000	2,94	0,0001	3,16	0,0000
ITGA6	3,68	0,0001	3,23	0,0000	3,47	0,0000	3,19	0,0000
TAF11	2,81	0,0000	3,24	0,0000	2,56	0,0000	3,21	0,0000
nd	3,41	0,0003	3,50	0,0000	3,45	0,0000	3,33	0,0000
RP2	3,98	0,0001	3,47	0,0000	3,80	0,0000	3,35	0,0000
GGACT	2,90	0,0006	3,80	0,0000	2,24	0,0005	3,45	0,0000
LMBR1	4,53	0,0002	3,78	0,0000	4,39	0,0000	3,47	0,0001
CFL2	3,22	0,0003	3,58	0,0000	3,01	0,0001	3,52	0,0001
FZD5	5,53	0,0000	3,54	0,0000	6,18	0,0000	3,53	0,0000
SLC31A1	3,44	0,0001	3,45	0,0000	3,49	0,0000	3,56	0,0000
HINT3	2,83	0,0004	3,35	0,0000	2,65	0,0000	3,61	0,0000
PMM2	3,54	0,0000	4,06	0,0000	3,46	0,0000	3,63	0,0000
EMC10	4,22	0,0000	3,96	0,0000	3,90	0,0000	3,75	0,0000

UBE2V1	3,81	0,0008	4,10	0,0000	3,99	0,0001	3,82	0,0002
AK2	2,63	0,0003	4,30	0,0000	2,78	0,0000	4,00	0,0000
UBE2V1	4,84	0,0000	4,26	0,0000	4,95	0,0000	4,09	0,0000
SNRPC	4,45	0,0001	4,42	0,0000	4,49	0,0000	4,32	0,0000
DDAH1	1,99	0,0000	4,38	0,0000	2,10	0,0001	4,37	0,0000
ASPH	9,31	0,0000	9,45	< 1e-07	10,18	0,0000	10,90	0,0000

Table 4 : Signalisations et voies enrichies dans le chIP-seq contre FOXS1-HA après association de chaque pic au site d'initiation de la transcription le plus proche.

Geneset.ID	Description	P.value	FDR	Effect	Nb of genes in dataset	Nb of peaks
GO:190253 1	regulation of intracellular signal transduction	1,16E-07	0,001	0,41	1714	262
M2634	EGFR_UP.V1_UP	1,40E-07	0	0,92	191	55
GO:000646 8	protein phosphorylation	2,67E-06	0,007	0,36	1837	270
GO:000193 2	regulation of protein phosphorylation	6,79E-06	0,009	0,4	1284	195
GO:004232 5	regulation of phosphorylation	8,27E-06	0,009	0,38	1375	204
GO:003033 4	regulation of cell migration	8,88E-06	0,009	0,49	671	125
M2790	EIF4E_DN	1,02E-05	0,001	1,02	99	33
M2839	TGFB_UP.V1_UP	2,95E-05	0,002	0,76	188	47
M2855	STK33_NOMO_UP	3,54E-05	0,002	0,65	286	62
GO:003023 4	enzyme regulator activity	4,42E-05	0,041	0,41	979	145
GO:003033 5	positive regulation of cell migration	4,66E-05	0,039	0,57	394	75
GO:200014 5	regulation of cell motility	5,63E-05	0,039	0,43	723	129
GO:000268 3	negative regulation of immune system process	6,03E-05	0,039	0,62	380	60
M2871	CORDENONSI_YAP_CONSERVED_SIGNATURE	6,49E-05	0,002	1,2	57	20
GO:000561 5	extracellular space	7,05E-05	0,044	0,37	1387	173
GO:004858 5	negative regulation of response to stimulus	7,95E-05	0,041	0,34	1372	204
GO:001922 0	regulation of phosphate metabolic process	9,31E-05	0,042	0,32	1608	230
GO:007223 4	metanephric nephron tubule development	9,82E-05	0,042	1,89	20	9
GO:000467 5	transmembrane receptor protein serine/threonine kinase activity	0,00011	0,052	2,14	16	9
GO:005117 4	regulation of phosphorus metabolic process	0,00012	0,043	0,31	1621	231
M2858	STK33_UP	0,00012	0,004	0,61	280	59
GO:006104 5	negative regulation of wound healing	0,00013	0,043	1,18	64	18

GO:0040017	positive regulation of locomotion	0,00014	0,043	0,52	437	79
GO:0051270	regulation of cellular component movement	0,00014	0,043	0,4	791	138
GO:2000147	positive regulation of cell motility	0,00014	0,043	0,53	409	75
GO:0072243	metanephric nephron epithelium development	0,00015	0,043	1,76	23	9
GO:0008106	alcohol dehydrogenase (NADP+) activity	0,00021	0,065	2,06	16	6
GO:0072170	metanephric tubule development	0,00022	0,054	1,72	22	9
GO:0071548	response to dexamethasone	0,00022	0,054	1,38	41	12
M2698	P53_DN.V1_UP	0,00022	0,006	0,69	191	45
M2857	STK33_SKM_UP	0,00028	0,007	0,62	270	51
GO:0033002	muscle cell proliferation	0,00028	0,065	0,73	166	39
GO:0005125	cytokine activity	0,00029	0,067	0,77	218	31
GO:0060548	negative regulation of cell death	0,0003	0,066	0,38	890	127
GO:0051272	positive regulation of cellular component movement	0,00032	0,068	0,5	419	75
M2728	RAF_UP.V1_UP	0,00032	0,007	0,64	194	48
GO:0071364	cellular response to epidermal growth factor stimulus	0,00036	0,071	1,41	39	11
GO:0000165	MAPK cascade	0,00037	0,071	0,37	856	133
GO:0040012	regulation of locomotion	0,00037	0,071	0,38	789	133
M6778	BIOCARTA_IL10_PATHWAY	0,00039	0,059	1,98	17	6
GO:1903035	negative regulation of response to wounding	0,0004	0,073	1,03	75	20
GO:0023014	signal transduction by protein phosphorylation	0,00044	0,073	0,36	892	139
GO:0048661	positive regulation of smooth muscle cell proliferation	0,00044	0,073	1,03	72	19
GO:0050687	negative regulation of defense response to virus	0,00045	0,073	1,8	22	6
GO:0051336	regulation of hydrolase activity	0,00048	0,075	0,3	1342	195

GO:000152 2	pseudouridine synthesis	0,00053	0,079	2,12	18	5
GO:005084 0	extracellular matrix binding	0,00057	0,075	1,18	50	14
GO:004232 7	positive regulation of phosphorylation	0,00059	0,079	0,35	922	138
GO:001056 2	positive regulation of phosphorus metabolic process	0,00062	0,079	0,33	1045	155
GO:004593 7	positive regulation of phosphate metabolic process	0,00062	0,079	0,33	1045	155
GO:009877 2	molecular function regulator	0,00062	0,075	0,3	1392	198
GO:000223 7	response to molecule of bacterial origin	0,00062	0,079	0,58	316	48
GO:003355 4	cellular response to stress	0,00063	0,079	0,28	1796	223
GO:005092 1	positive regulation of chemotaxis	0,00064	0,079	0,82	126	26
GO:007220 7	metanephric epithelium development	0,00065	0,079	1,51	26	9
GO:001063 2	regulation of epithelial cell migration	0,00066	0,079	0,66	176	41
GO:005170 7	response to other organism	0,00068	0,079	0,41	790	92
GO:004320 7	response to external biotic stimulus	0,00072	0,083	0,41	791	92
GO:003399 3	response to lipid	0,00076	0,084	0,36	871	126
GO:000512 6	cytokine receptor binding	0,00079	0,075	0,64	263	38
GO:190253 3	positive regulation of intracellular signal transduction	0,00083	0,086	0,35	897	130
GO:000193 4	positive regulation of protein phosphorylation	0,00085	0,086	0,35	883	133
GO:003249 6	response to lipopolysaccharide	0,00087	0,086	0,58	302	46
GO:007084 9	response to epidermal growth factor	0,00087	0,086	1,29	43	11
GO:004340 8	regulation of MAPK cascade	0,00087	0,086	0,39	659	107
GO:000687 9	cellular iron ion homeostasis	0,00091	0,089	1,23	47	11
M2768	ESC_J1_UP_LATE.V1_UP	0,00095	0,017	0,63	189	41

GO:004691 6	cellular transition metal ion homeostasis	0,001	0,094	1	83	16
----------------	--	-------	-------	---	----	----

Table 5 : Signalisations et voies enrichies dans le chIP-seq contre FOXS1-HA après sélection des pics à moins de 10kb du site d'initiation de la transcription le plus proche.

Genes et.ID	Description	P.valu e	FDR	Effect	Nb of genes in dataset	Nb of peaks
GO:0071222	cellular response to lipopolysaccharide	6,9208E-06	0,02041183	1,22016997	149	17
GO:0071219	cellular response to molecule of bacterial origin	1,0135E-05	0,02041183	1,19452658	154	17
GO:0072358	cardiovascular system development	2,3074E-05	0,02041183	0,65489477	609	52
GO:0006949	syncytium formation	2,5469E-05	0,02041183	1,61419144	52	9
GO:0001932	regulation of protein phosphorylation	2,6154E-05	0,02041183	0,51192199	1284	87
GO:0001944	vasculature development	2,8E-05	0,02041183	0,65368339	600	51
GO:0052695	cellular glucuronidation	3,222E-05	0,02041183	2,51170741	17	4
GO:1902531	regulation of intracellular signal transduction	3,2534E-05	0,02041183	0,44983503	1714	116
GO:1901798	positive regulation of signal transduction by p53 class mediator	3,6021E-05	0,02041183	2,32252129	19	5
GO:0046631	alpha-beta T cell activation	3,9535E-05	0,02041183	1,22555312	116	14
GO:0005615	extracellular space	4,0322E-05	0,02512031	0,53012173	1387	78
GO:0001568	blood vessel development	5,461E-05	0,02452294	0,6414171	577	49
GO:0043408	regulation of MAPK cascade	5,6997E-05	0,02452294	0,62988239	659	50
GO:0090342	regulation of cell aging	7,0516E-05	0,02800589	1,74740642	35	7
GO:0042325	regulation of phosphorylation	8,3429E-05	0,03076744	0,47153134	1375	90
GO:0060142	regulation of syncytium formation by plasma membrane fusion	9,2602E-05	0,03187345	2,02955586	26	5
GO:0006508	proteolysis	0,00012212	0,03670069	0,43444895	1662	102
GO:0042517	positive regulation of tyrosine phosphorylation of Stat3 protein	0,00012789	0,03670069	1,78892716	36	6
GO:2000772	regulation of cellular senescence	0,00012795	0,03670069	1,82987105	28	6
GO:0000768	syncytium formation by plasma membrane fusion	0,00013829	0,03757753	1,54209297	48	8
GO:0071216	cellular response to biotic stimulus	0,00014887	0,03843088	1,01375356	176	17

M2790	EIF4E_DN	0,000 16771	0,031 52965	1,052 82757	99	17
GO:00 04675	transmembrane receptor protein serine/threonine kinase activity	0,000 1807	0,102 82336	2,098 06765	16	5
GO:00 72359	circulatory system development	0,000 19922	0,047 17011	0,497 95353	924	71
GO:00 61045	negative regulation of wound healing	0,000 20793	0,047 17011	1,337 63241	64	10
GO:00 44257	cellular protein catabolic process	0,000 21013	0,047 17011	0,584 07476	663	49
GO:00 50431	transforming growth factor beta binding	0,000 22232	0,102 82336	2,073 60915	17	5
GO:00 07413	axonal fasciculation	0,000 23727	0,051 04331	1,882 15093	21	6
GO:00 50687	negative regulation of defense response to virus	0,000 27792	0,055 72949	1,932 7153	22	5
GO:00 06468	protein phosphorylation	0,000 28064	0,055 72949	0,391 48039	1837	116
GO:00 32102	negative regulation of response to external stimulus	0,000 30933	0,056 94305	0,827 03657	253	23
GO:00 61041	regulation of wound healing	0,000 31902	0,056 94305	1,033 00937	128	15
GO:00 35710	CD4-positive, alpha-beta T cell activation	0,000 35242	0,056 94305	1,334 75495	70	9
GO:00 46632	alpha-beta T cell differentiation	0,000 35424	0,056 94305	1,197 78097	90	11
GO:00 19220	regulation of phosphate metabolic process	0,000 36114	0,056 94305	0,408 13029	1608	100
GO:00 46633	alpha-beta T cell proliferation	0,000 36143	0,056 94305	1,845 29307	27	5
GO:00 32496	response to lipopolysaccharide	0,000 36396	0,056 94305	0,811 63066	302	23
GO:00 72539	T-helper 17 cell differentiation	0,000 41102	0,062 41492	2,066 66415	19	4
GO:00 03158	endothelium development	0,000 43744	0,063 34513	1,063 73453	111	14
GO:00 46634	regulation of alpha-beta T cell activation	0,000 45757	0,063 34513	1,285 30636	74	9
GO:00 30334	regulation of cell migration	0,000 46318	0,063 34513	0,533 31935	671	54
GO:00 51174	regulation of phosphorus metabolic process	0,000 46622	0,063 34513	0,400 2433	1621	100
GO:00 30163	protein catabolic process	0,000 54681	0,072 38859	0,510 75509	805	56
GO:00 06511	ubiquitin-dependent protein catabolic process	0,000 61359	0,074 7532	0,580 9465	555	42

GO:00 02832	negative regulation of response to biotic stimulus	0,000 61935	0,074 7532	1,584 14374	37	6
GO:00 98868	bone growth	0,000 6411	0,074 7532	1,798 8315	23	5
GO:00 72538	T-helper 17 type immune response	0,000 64549	0,074 7532	1,966 99082	21	4
GO:00 06063	uronic acid metabolic process	0,000 65154	0,074 7532	1,986 46138	22	4
GO:00 19585	glucuronate metabolic process	0,000 65154	0,074 7532	1,986 46138	22	4
GO:00 46640	regulation of alpha-beta T cell proliferation	0,000 67799	0,074 95549	1,907 17745	23	4
GO:00 02237	response to molecule of bacterial origin	0,000 68234	0,074 95549	0,770 98702	316	23
M2871	CORDENONSI_YAP_CONSERVED_SIG NATURE	0,000 69659	0,046 0779	1,235 8255	57	10
GO:00 04857	enzyme inhibitor activity	0,000 69672	0,154 37276	0,698 09347	379	28
M2828	PDGF_ERK_DN.V1_DN	0,000 73529	0,046 0779	1,014 87517	147	13
GO:00 19941	modification-dependent protein catabolic process	0,000 73695	0,079 26854	0,572 04768	562	42
GO:00 42516	regulation of tyrosine phosphorylation of Stat3 protein	0,000 76647	0,080 76128	1,536 29777	43	6
GO:00 01933	negative regulation of protein phosphorylation	0,000 81297	0,083 94685	0,651 64965	396	32
GO:00 15020	glucuronosyltransferase activity	0,000 81697	0,154 37276	1,709 96743	34	5
GO:00 05126	cytokine receptor binding	0,000 83445	0,154 37276	0,849 17596	263	18
GO:00 30198	extracellular matrix organization	0,000 90425	0,089 90826	0,678 59511	314	29
GO:00 43409	negative regulation of MAPK cascade	0,000 94201	0,089 90826	0,939 64479	147	15
GO:00 43062	extracellular structure organization	0,000 95291	0,089 90826	0,675 4073	315	29
GO:00 23014	signal transduction by protein phosphorylation	0,000 95731	0,089 90826	0,472 61778	892	60
GO:00 51603	proteolysis involved in cellular protein catabolic process	0,000 95777	0,089 90826	0,540 63143	627	45
GO:00 44092	negative regulation of molecular function	0,000 98867	0,091 15182	0,450 45995	1082	66

Articles

Landscape of genomic alterations in hepatocellular carcinoma: current knowledge and perspectives for targeted therapies

Kévin Bévant, Cédric Coulouarn

Inserm, Inra, Univ Rennes, UMR 1241, Nutrition Metabolisms and Cancer, Rennes, France

Correspondence to: Cédric Coulouarn. Inserm, Inra, Univ Rennes, UMR 1241, CHU Pontchaillou, 2 rue Henri Le Guilloux, F-35033 Rennes, France.

Email: cedric.coulouarn@inserm.fr.

Provenance: This is an invited Editorial commissioned by Editor-in-Chief Yilei Mao (Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China).

Comment on: Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu; Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 2017;169:1327-41.e23.

Submitted Oct 19, 2017. Accepted for publication Oct 27, 2017.

doi: 10.21037/hbsn.2017.10.02

View this article at: <http://dx.doi.org/10.21037/hbsn.2017.10.02>

Hepatocellular carcinoma (HCC) is a deadly cancer worldwide characterized by a rising incidence and limited therapeutic options (1). HCC is associated with multiple risk factors, including chronic hepatitis B and C (HBV/HCV), alcohol intake, aflatoxin B1 exposure, diabetes and obesity. In addition, liver carcinogenesis is a long and multistep process associated with the accumulation of multiple genetic and epigenetic alterations resulting in tumor heterogeneity (1). Thus, each HCC tumor is characterized by a unique molecular fingerprint made of a specific combination of somatic alterations (2). Deciphering this molecular fingerprint is essential to develop effective personalized treatments. Over the last decades, genome-wide unsupervised strategies including next-generation sequencing (NGS) and gene expression profiling allowed a deep characterization of genomic alterations in HCC, as well as the definition of clinically relevant HCC subtypes (3-6). In direct continuity with this experienced strategy, The Cancer Genome Atlas (TCGA) network recently reported a comprehensive and integrative genomic characterization of a large cohort of 363 HCC (7). A multiplex molecular analysis was performed, including exome sequencing, DNA copy number and methylation analysis, gene expression profiling, and proteomics (Figure 1). The study validates known and identifies novel driver gene candidates in HCC, defines putative key therapeutic targets, and data integration from multiple genomic platforms provides mechanistic molecular insights for the observed alterations (7).

Whole-genome and exome sequencing provided an accurate landscape of recurrent somatic mutations and driver genes in HCC (3,4,7). The most frequently observed mutational events include *TERT* promoter activating mutations (~40% HCC), and the mutually exclusive mutations of *TP53* and *CTNNB1* genes (~30% HCC). Taken together, these studies highlighted a complex network of signaling pathways that constitute the core functional hallmarks of HCC (5,6,8).

Telomere maintenance

Reactivation of telomerase is an early and recurrent event in liver carcinogenesis as it is detected in preneoplastic lesions and in >90% HCC. These observations open new opportunities for early HCC detection, specifically if *TERT* alterations could be detected in circulating tumor cells. Sequencing analysis demonstrated that *TERT* reactivation mainly results from activating mutations or viral insertion at the promoter site, and focal DNA amplification (8). Interestingly, HCV-associated HCC are more likely to be mutated in *TERT* promoter, as compared to HBV-associated HCC in which *TERT* reactivation is frequently linked to HBV integration. *TERT* promoter mutations are also frequently associated with *CTNNB1* mutations, suggesting a cooperation of the two pathways in liver tumorigenesis. Thus, interfering with telomerase and the Wnt/ β -catenin pathway may represent a promising strategy.

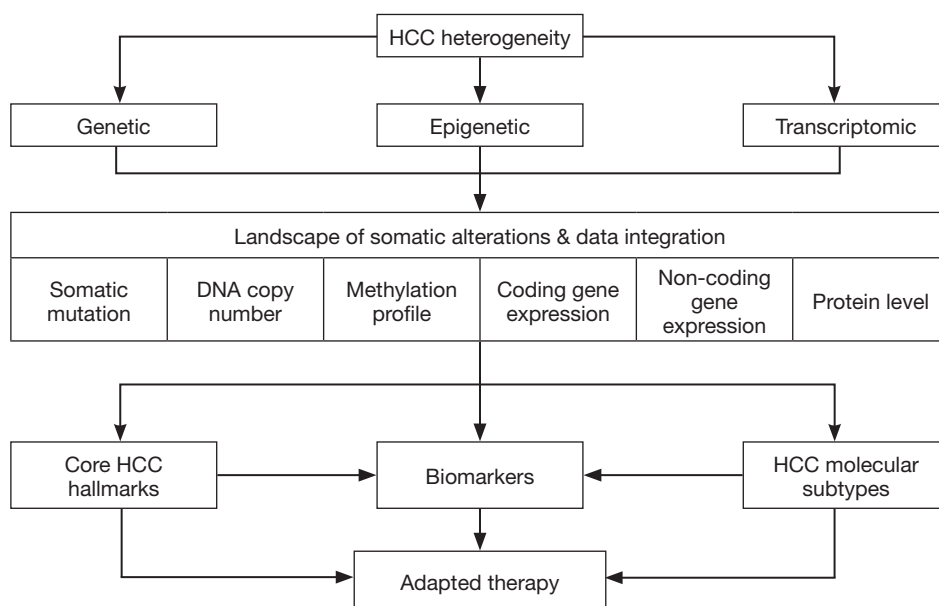


Figure 1 Integrative genomics to decrypt HCC tumor heterogeneity and to identify new therapeutic strategies. HCC is a heterogeneous disease characterized by a unique combination of genetic, epigenetic and transcriptomic alterations. The TCGA study illustrates the potential of large-scale analysis using multiple genomic platforms (e.g., exome sequencing, DNA copy number and methylation analysis, gene expression and protein profiling) to characterize the molecular landscape of genomic alterations in HCC. Data integration is used to define the core hallmarks altered in most HCC, as well as to stratify HCC into clinically relevant subtypes. Such integrative approach not only provides mechanistic insights into the molecular mechanisms involved in HCC carcinogenesis but also identifies key oncogenic pathways to be targeted for combined personalized therapies.

Wnt/ β -catenin pathway

The Wnt/ β -catenin pathway plays a key role in liver physiology and is frequently altered in HCC. Although not addressed in the TCGA study, a specific spectrum of *CTNNB1* activating mutations is closely associated with the activity of β -catenin and its oncogenic potential in HCC (9). Negative regulators of the Wnt/ β -catenin pathway, including AXIN1/2 and APC, are also inactivated by mutations in HCC. Thus, various strategies have been developed to interfere with the Wnt/ β -catenin pathway, e.g., by using small intracellular inhibitors or soluble decoy receptors for Wnt ligands. However, such pharmacological inhibitors have not been translated into the clinic yet, due to the difficulty of designing highly selective and non-toxic molecules (8). A better characterization of canonical and non-canonical effects of the Wnt/ β -catenin pathway and its crosstalk with the tumor microenvironment may help to model potential side effects of such inhibitors and to design effective combined therapies. Notably, an association of Wnt and TGF β inhibitors may be relevant

in Hoshida's S1 HCC subtype in which a crosstalk between Wnt/ β -catenin and TGF β pathways has been previously identified (6).

Cell cycle control

The P53 cell cycle pathway is frequently altered in HCC (8). Interestingly, by using a signature of P53 transcriptional targets, the TCGA study identifies a group of P53 inactive tumors but independently to *TP53* mutations. The amplification of P53 inhibitor *MDM4* is notably identified as an alternate mechanism of P53 inactivation (7). P53 inactive HCC subtype is associated with a poor prognosis, stemness features and activation of the sonic hedgehog signaling (7). DNA hypermethylation was also found to be the major mechanism of silencing of cyclin dependent kinase inhibitor *CDKN2A* (>50% HCC) whereas mutations of *CDKN2A* and other cell cycle regulators (e.g., *RB1*, *CCNE1*) occur in less than 4% HCC (7). Frequent alterations of numerous targets linked to the activity of

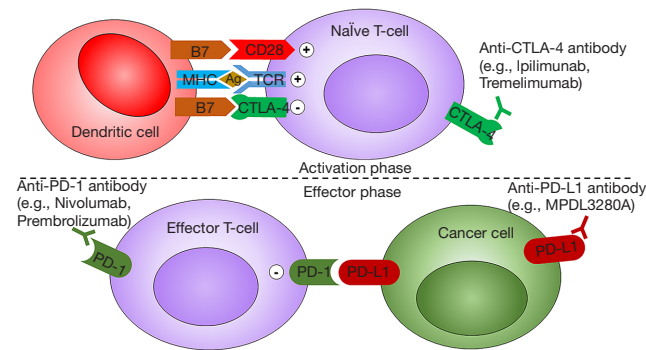


Figure 2 Targeted therapy using specific immune checkpoint inhibitors. T-cell activation involves co-stimulatory signals initiated by the interaction of B7 with CD28 while T-cell inactivation involves interaction of B7 with *CTLA-4*. T-cells are also negatively regulated by the interaction of PD-1 with its ligand PD-L1 commonly over-expressed in tumor cells. Targeted therapies using immune checkpoint inhibitors represent a promising approach to boost the activation of immune cell effectors and to avoid their inactivation. The identification of a specific HCC subtype expressing immune checkpoint genes (e.g., *CTLA4*, *PD1*, *PDL1*) at high levels opens promising therapeutic opportunities using immune checkpoint inhibitors.

proliferation-associated receptor tyrosine kinases (RTK) (e.g., *RAS*, *PI3K*, *PTEN*) are in agreement with the sensitivity of HCC to RTK inhibitors (10).

Chromatin remodeling

As previously reported, recurrent somatic mutations of chromatin modifier genes, including *ARID1/2*, *BAP1* or histone methyl transferases *KMT2C* and *KMT2D* were identified (7). Although this point was not addressed in the TCGA study, integration of long non-coding RNA profiles would also provide mechanistic insights into HCC tumorigenesis. Indeed, long non-coding RNA play a key role in epigenetics, notably by acting as scaffolds for chromatin remodeling factors (11). Changes in chromatin remodeling or DNA methylation as a result of somatic mutations or gene expression alterations suggest that therapeutic targeting of epigenetic pathways is relevant in HCC. Accordingly, by integrating multiple transcriptomic profiles, we recently reported that histone deacetylase inhibitors may target multiple core hallmarks of HCC (5).

Molecular HCC subtyping

Molecular stratification of HCC has been extensively described based on genetic, epigenetic and transcriptomic profiles. The originality of the TCGA study comes from the integration of multiple genomic data types using a joint latent variable model (12). The so-called iCluster algorithm identified 3 HCC subtypes (iC1-3) which partially overlap with S1-3 subtypes, as defined by Hoshida (6). However, the absence of significant differences in term of survival between the 3 iClusters (except for the poor prognosis iC1 subtype) challenges the approach of integrating all genomic data types simultaneously. It will be interesting to determine whether some specific data types (e.g., quantitative mRNA levels) are more efficient than others for patient stratification.

Novel HCC driver genes

The TCGA study identified 8 novel HCC driver candidates (*LZTR1*, *AZIN1*, *RP1L1*, *EEF1A1*, *GPATCH4*, *CREB3L3*, *AHCTF1*, *HIST1H1C*) mutated in 1-4% HCC (7). Functional studies and modelling of the impact of the identified mutations will be required to validate these genes as true drivers in HCC. Interestingly, cholangiocarcinoma-associated *IDH1/2* mutations were identified in 4 tumors. Although exhibiting histological features of HCC, integrating gene expression profiles of *IDH1/2* mutated tumors with those of well-defined HCC, intrahepatic cholangiocarcinoma (iCCA) and mixed iCCA/HCC tumors may help to clarify their cellular origin. Indeed, these HCC were associated with features previously observed in iCCA and mixed iCCA/HCC, including a very poor prognosis and the expression of stemness markers (13,14). Thus, the existence of *IDH1/2* mutated HCC supports the hypothesis of tumors possibly arising from cancer stem cells.

Immune-based therapeutic strategies

Striking differences in the relative composition of immune cell populations were observed between non-tumor and tumor tissues highlighting a switch from activating effector cells to resting suppressive immune cells. Based on the expression of 66 curated immune cell markers, a HCC subset with high expression of immune checkpoint genes (e.g., *CTLA4*, *PDCD1/PD1*, *CD274/PDL1*) was identified, opening interesting opportunities for efficient targeted therapies using immune checkpoint inhibitors, e.g., anti-CTLA4 and anti-PD1/PDL1 antibodies (Figure 2).

In conclusion, integrative genomic studies, including the TCGA study, clearly point out the heterogeneity of HCC alterations and suggest that combined therapies targeting several cancer hallmarks simultaneously may represent a clinically relevant strategy (*Figure 1*). Once a consensus in HCC subtyping will be established, another important step to be reached for a better clinical translation will be the definition of specific biomarkers for each HCC subtype, ideally detectable with poorly invasive methods. In this context, further investigations will be required to determine whether specific alterations reflecting each HCC subtype could be detected in liquid biopsies.

Acknowledgements

Funding: Author's laboratory is supported by funds from Inserm, University of Rennes, Institut National du Cancer, Ligue Contre le Cancer (cd22, cd35, cd85), and Association Française pour l'Etude du Foie (AFEF). KB is supported by a doctoral fellowship from Ligue Contre le Cancer (cd22) and Conseil Régional de Bretagne.

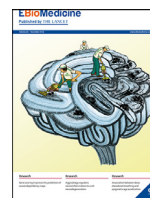
Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References

1. Thorgeirsson SS, Grisham JW. Molecular pathogenesis of human hepatocellular carcinoma. *Nat Genet* 2002;31:339-46.
2. Schulze K, Nault JC, Villanueva A. Genetic profiling of hepatocellular carcinoma using next-generation sequencing. *J Hepatol* 2016;65:1031-42.
3. Schulze K, Imbeaud S, Letouze E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet* 2015;47:505-11.
4. Totoki Y, Tatsuno K, Covington KR, et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat Genet* 2014;46:1267-73.
5. Allain C, Angenard G, Clement B, et al. Integrative Genomic Analysis Identifies the Core Transcriptional Hallmarks of Human Hepatocellular Carcinoma. *Cancer Res* 2016;76:6374-81.
6. Hoshida Y, Nijman SM, Kobayashi M, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 2009;69:7385-92.
7. Cancer Genome Atlas Research Network. Electronic address: wheeler@bcm.edu; Cancer Genome Atlas Research Network. Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. *Cell* 2017;169:1327-41.e23.
8. Zucman-Rossi J, Villanueva A, Nault JC, et al. Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology* 2015;149:1226-39.e4.
9. Rebouissou S, Franconi A, Calderaro J, et al. Genotype-phenotype correlation of CTNNB1 mutations reveals different ss-catenin activity associated with liver tumor progression. *Hepatology* 2016;64:2047-61.
10. Llovet JM, Villanueva A, Lachenmayer A, et al. Advances in targeted therapies for hepatocellular carcinoma in the genomic era. *Nat Rev Clin Oncol* 2015;12:408-24.
11. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell* 2016;29:452-63.
12. Shen R, Olshen AB, Ladanyi M. Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis. *Bioinformatics* 2009;25:2906-12.
13. Lee JS, Heo J, Libbrecht L, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 2006;12:410-6.
14. Coulouarn C, Cavard C, Rubbia-Brandt L, et al. Combined hepatocellular-cholangiocarcinomas exhibit progenitor features and activation of Wnt and TGFbeta signaling pathways. *Carcinogenesis* 2012;33:1791-6.

Cite this article as: Bévant K, Coulouarn C. Landscape of genomic alterations in hepatocellular carcinoma: current knowledge and perspectives for targeted therapies. *HepatoBiliary Surg Nutr* 2017;6(6):404-407. doi: 10.21037/hbsn.2017.10.02



Commentary

A novel immunosuppression-based classification of liver tumors opens new perspectives for adapted therapeutic strategies

Kevin Bévant, Cédric Coulouarn*

Inserm, CHU Pontchaillou, 2 rue Henri Le Guilloux, F-35033 Rennes, France



ARTICLE INFO

Article History:

Received 28 February 2020

Accepted 28 February 2020

Available online xxx

Liver cancer, including hepatocellular carcinoma (HCC), is the 4th leading cause of cancer-related death worldwide, notably as a result of tumor heterogeneity which hampers the development of efficient treatments. Integrative genomics helped with characterizing heterogeneity in cancer, leading to molecular stratification of tumors and adapted therapeutic strategies for specific tumor subtypes.

In this article of *EBioMedicine*, Fujita and colleagues report the results of an elegant meta-analysis of RNA and whole genome sequencing data from 234 liver primary tumors [1]. Based on immune signatures and genomic alterations, four tumor subtypes are highlighted. Three subtypes exhibit mutually exclusive immunosuppressive features associated with infiltration of regulatory T cells or tumor-associated macrophages (TAM) and *CTNNB1* mutations. One subtype associated with a better prognosis shows cytolytic activity. The study brings important information on the immunosuppressive properties of HCC microenvironment. Notably, the non-inflamed TAM subtype is enriched in *ARID2* mutations resulting in impairment of chemokine production and a weaker inflammatory and interferon- γ response. By taking into account immunosuppressive mechanisms, the study improves the current binary immune classification into “inflamed” or “non-inflamed” tumors. From a clinical perspective, understanding HCC immunosuppressive properties is critical to improve the efficacy of immune-based therapeutic strategies.

Over the last decade, immune checkpoint inhibitors and CAR-T cells emerged as innovative therapeutic options. Both have shown promising results in several cancers but not in liver cancer, particularly as a result of frequent intrinsic tumor immunosuppression [2,3]. Notably, CAR-T cell therapy is hampered in HCC by the limited migration of T cell into the tumor, the immunosuppressive tumor microenvironment and the lack of well-defined tumor antigens [4].

Concerning immune checkpoint inhibitors, the phase III randomized controlled trials comparing nivolumab versus sorafenib in first line or pembrolizumab versus placebo did not reach their pre-specified endpoints. The lack of efficacy for these two PD-1 inhibitors suggests that a better stratification of HCC patients based on the immune microenvironment is required to identify the patients who may benefit from immune-based therapies [5]. By improving the current immune classification of HCC, the study of Fujita et al. may help to identify the patients who most likely could respond to immunotherapy. Importantly, this proposed immunological classification is validated in Asian and Western datasets and therefore could be applied to a large cohort of patients independently from HCC etiology.

Immunotherapy in HCC not only lacks relevant biomarkers to identify the best responder patients, but also a deep characterization of molecular mechanisms generating an immunosuppressive microenvironment. Fujita et al. identifies tumor-specific immunosuppressive mechanisms which could be targeted to improve efficacy of treatments, including innovative strategies combining immune-modulatory agents and anti-tumor compounds. Several studies have already demonstrated the potential benefit of such combined therapies. Thus, co-treatments with PD-1 inhibitors and kinase inhibitors (e.g. lenvatinib, sunitinib) improve anti-tumor responses and tumor regression. Combination of PD-1 and CTLA-4 inhibitors that have shown great results in melanoma are also currently tested in HCC. Similarly, the efficacy of PD-L1 inhibitors (e.g. avelumab) are being evaluated with sorafenib or axitinib in patients with advanced HCC. Sorafenib, a multi-kinase inhibitor, exhibits immunomodulatory although the underlying molecular mechanisms are unclear [6]. In immunodeficient mice, sorafenib enhances the efficacy of human CAR-T cells against HCC [7]. These results suggest that combining immunotherapy and drugs presenting immunomodulatory effects represents a promising effective therapeutic strategy in HCC. Thus, efficient immunotherapy in HCC will need both a clear comprehension of immunosuppressive mechanisms and the development of effective drug combination. In this context, molecules or pathways involved in immunosuppression and cancer progression represent ideal targets. Transforming Growth Factor beta ($TGF\beta$) is such a relevant target, inducing tumor microenvironment remodeling and exhibiting potent immunosuppressive features. Interestingly, Fujita et al. report enrichment of a $TGF\beta$ signature in the TAM HCC subtype, in agreement with our recent study demonstrating that $TGF\beta$ and AXL induce CXCL5 and neutrophil recruitment in poor prognosis HCC [8]. Accordingly, galunisertib (LY2157299), an inhibitor of $TGF\beta$

DOI of original article: <http://dx.doi.org/10.1016/j.ebiom.2020.102659>.

* Corresponding author.

E-mail address: cedric.coulouarn@inserm.fr (C. Coulouarn).<https://doi.org/10.1016/j.ebiom.2020.102737>2352-3964/© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

pathway, is currently being investigated in several clinical trials. Combined therapeutic strategies associating galunisertib with PD-1/PD-L1 inhibitors, or bifunctional trap fusion proteins targeting both TGF β and PD-L1, such as the M7824 compound, are also attractive [9]. Supporting this strategy, it was recently reported that TGF β attenuates the efficacy of immune checkpoint inhibitors by modulating the tumor microenvironment and restricting T cell infiltration [10].

In conclusion, the study by Fujita and colleagues contributes to improve our knowledge on immunosuppressive features of specific HCC subtypes and opens new avenues for the development of effective targeted immunotherapy.

Declaration of Competing Interest

All authors have no conflicts of interest to disclose.

Funding sources

This research was supported by Inserm, Université de Rennes 1 and Ligue Contre le Cancer (CD22, CD35, CD85). KB is supported by a PhD fellowship from Ligue Contre le Cancer (CD22) and Région Bretagne.

References

- [1] Fujita M, Yamaguchi R, Hasegawa T, Shimada S, Arihiro K, Hayashi S, et al. Classification of primary liver cancer with immunosuppression mechanisms and correlation with genomic alterations. *EBioMedicine* 2020. doi: 10.1016/j.ebiom.2020.102659.
- [2] Heinhuis KM, Ros W, Kok M, Steeghs N, Beijnen JH, Schellens JHM. Enhancing antitumor response by combining immune checkpoint inhibitors with chemotherapy in solid tumors. *Ann. Oncol.* 2019;30(2):219–35.
- [3] Hilmi M, Vienot A, Rousseau B, Neuzillet C. Immune therapy for liver cancers. *Cancers* 2019;12(1).
- [4] Hendrickson PG, Olson M, Luetkens T, Weston S, Han T, Atanackovic D, et al. The promise of adoptive cellular immunotherapies in hepatocellular carcinoma. *Oncoimmunology* 2020;9(1):1673129.
- [5] Rebouissou S, Nault JC. Advances in molecular classification and precision oncology in hepatocellular carcinoma. *J. Hepatol.* 2020;72(2):215–29.
- [6] Zongyi Y, Xiaowu L. Immunotherapy for hepatocellular carcinoma. *Cancer Lett.* 2020;470:8–17.
- [7] Wu X, Luo H, Shi B, Di S, Sun R, Su J, et al. Combined antitumor effects of Sorafenib and GPC3-CAR T cells in mouse models of hepatocellular carcinoma. *Mol. Ther.* 2019;27(8):1483–94.
- [8] Haider C, Hnat J, Wagner R, Huber H, Timelthaler G, Grubinger M, et al. Transforming growth factor- β and Axl Induce CXCL5 and neutrophil recruitment in hepatocellular carcinoma. *Hepatology* 2019;69(1):222–36.
- [9] Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: looking outside the box. *J. Hepatol.* 2020;72(2):342–52.
- [10] Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554(7693):544–8.



TGF β -induced FOXS1 controls epithelial-mesenchymal transition and predicts a poor prognosis in liver cancer

Journal:	<i>Hepatology</i>
Manuscript ID	Draft
Wiley - Manuscript type:	Original
Date Submitted by the Author:	n/a
Complete List of Authors:	Bévant, Kevin; INSERM Desoteux, Matthis; INSERM Angenard, Gaëlle; INSERM, U1241 Caruso, Stefano Louis, Corentin; INSERM Papoutsoglou, Panagiotis; Inserm U1241, U1241 Sulpice, Laurent; Hopital Ponchaillou, Service de Chirurgie Gilot, david; Zucman-Rossi, Jessica; INSERM, FUNGEST Coulouarn, Cedric; INSERM, U991
Keywords:	Liver cancer, hepatocellular carcinoma, TGF-beta, EMT, FOXS1

SCHOLARONE™
Manuscripts

1
2
3 **TGF β -induced FOXS1 controls epithelial-mesenchymal transition and predicts a**
4
5 **poor prognosis in liver cancer**
6
7
8
9

10 Kevin Bévant¹, Matthis Desoteux¹, Gaëlle Angenard¹, Stefano Caruso²,
11
12 Corentin Louis¹, Panagiotis Papoutsoglou¹, Laurent Sulpice¹, David Gilot³,
13
14 Jessica Zucman-Rossi^{2,4} and Cédric Coulouarn^{1*}
15
16
17
18
19

20
21 ¹ Inserm, Univ Rennes, UMR 1241, CHU Rennes, Service de Chirurgie Hépatobiliaire et
22 Digestive, F-35000, Rennes, France;
23
24

25
26 ² Centre de Recherche des Cordeliers, Inserm, Sorbonne Université, Université de Paris,
27 Université Paris 13, Functional Genomics of Solid Tumors laboratory, F-75006 Paris, France;
28
29

30
31 ³ Univ Rennes, Cnrs, IGDR [(Institut de génétique et développement de Rennes)]-UMR 6290,
32 ARC Labelled team, F-35000, Rennes, France;
33
34

35
36 ⁴ European Hospital Georges Pompidou, AP-HP, F-75015, Paris, France.
37
38
39
40

41
42 **Correspondence:** Cédric COULOUARN, corresponding author *
43
44

45 Inserm UMR 1241, CHU Pontchaillou, 2 rue Henri Le Guilloux, F-35033 Rennes, France.
46
47

48 Tel : +33 2 2323 3881, Fax +33 2 9954 0137 ; E-mail: cedric.coulouarn@inserm.fr
49
50
51
52

53 **Running title:** TGF-beta-induced FOXS1 in human HCC
54
55
56
57
58
59
60

1
2
3 **Keywords:** Liver cancer; hepatocellular carcinoma; TGF-beta; EMT; FOXS1
4
5

6 **Electronic word count:** 5.737 words
7

8
9 **Number of figures and tables:** 6 figures and no table
10

11 **Disclosures (for all authors):** The authors disclose no conflict of interest.
12
13

14
15 **Grant support:** CC team was supported by Inserm, Université de Rennes 1, Ligue Contre le
16 Cancer (CD22, CD35, CD85), INCa, and ITMO Cancer AVIESAN (Alliance Nationale pour les
17 Sciences de la Vie et de la Santé) dans le cadre du Plan cancer (Non-coding RNA in
18 cancerology: fundamental to translational : [Grant C18007NS]). KB and MD are supported by a
19 PhD fellowship from Ligue Contre le Cancer (CD22) and Région Bretagne. CL is supported by a
20 PhD fellowship from Université de Rennes 1. JZR team was supported by Inserm, Ligue
21 Nationale contre le Cancer (Equipe Labellisée), Labex OncoImmunology (investissement
22 d'avenir), grant IREB, Coup d'Elan de la Fondation Bettencourt-Schueller, the SIRIC CARPEM,
23 Raymond Rosen Award from the Fondation pour le Recherche Médicale, Prix René and Andrée
24 Duquesne - Comité de Paris Ligue Contre le Cancer and Fondation Mérieux. SC is supported by
25 a funding from Labex OncoImmunology and CARPEM.
26
27
28
29
30
31
32
33
34
35
36
37
38

39 **Transcript Profiling:** Microarray datasets have been deposited into the public Gene Expression
40 Omnibus (GEO) database under the accession number GSE148795.
41
42
43

44 **Author contributions:** Study concept and design: CC, KB; Acquisition of data: KB, MD, GA,
45 SC, CC; Analysis and interpretation of data: ALL; Drafting of the manuscript: KB, MD, CC; Study
46 supervision: CC; Revision of the manuscript and approval of the final version of the manuscript:
47 ALL.
48
49
50
51

52
53 **Writing Assistance:** none
54
55
56
57
58
59
60

1
2
3 **Abbreviations:** AFP, alpha-fetoprotein; CDH1, E-cadherin; EMT, epithelial-mesenchymal
4 transition; FC, fold change; FOX, forkhead box; FOXS1 , forkhead box S1; GEO, gene
5 expression omnibus; GSEA, gene set enrichment analysis; HCC, hepatocellular carcinoma;
6
7
8
9 HPH, human primary hepatocytes; IVT, in vitro transcription; MET, mesenchymal-epithelial
10 transition; siRNA, small interfering RNA; ST, surrounding non-tumor tissue; TCGA, the cancer
11 genome atlas; TGF β , transforming growth factor beta; VIM, vimentin.
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Abstract

Background and Aims: Transforming Growth Factor beta (TGF β) is a pleiotropic cytokine which controls fundamental cellular processes driving tumor onset and progression. In hepatocellular carcinoma (HCC), we previously highlighted TGF β signatures recapitulating its functional duality (i.e. cytostatic versus pro-metastatic features) to discriminate good and poor prognosis patients. Here, we aimed at characterizing Forkhead Box S1 (FOXS1), a predicted transcription factor that we hypothesize to mediate pro-metastatic properties of TGF β .

Approach and Results: FOXS1 expression was profiled in well-characterized HCC cell lines. Gain and loss of functions (CRISPR-Cas9) were combined with functional assays and analysis of epithelial-mesenchymal transition (EMT) hallmarks. FOXS1 clinical relevance was determined by integrative genomics. Our data identifies *FOXS1* as a novel canonical SMAD-dependent TGF β target gene in cells with a mesenchymal phenotype. FOXS1 drives the expression of genes required for initial steps of EMT (e.g. *SNAI1*) but is also necessary to maintain a mesenchymal phenotype. FOXS1 induces cancer cell migration. In human HCC, FOXS1 correlates with TGF β phenotype switching, and is highly expressed in poorly differentiated and aggressive molecular subtypes. Importantly, FOXS1 expression level and activity predict a poor prognosis (e.g. reduced patient survival) not only in HCC, but also in colon, stomach and kidney cancers.

Conclusions: FOXS1 is a novel TGF β target gene controlling EMT. Its induction in cancer is associated with a poor prognosis. Determining the expression of FOXS1 in advanced tumors, in which the pro-metastatic arm of TGF β is activated, could help to select patients who may benefit from targeted therapies such as galunisertib.

Introduction

Liver cancer is the sixth most common cancer worldwide in terms of incidence and the fourth leading cause of cancer-related death with more than 780,000 deaths annually.(1) Hepatocellular carcinoma (HCC) accounts for 75-85% cases of malignant liver primary tumors.(1) HCC is associated with a rising incidence and limited curative treatments due to a late diagnosis.(2) For advanced HCC, only multikinase inhibitors demonstrated moderate survival benefits, either in first-line (e.g. sorafenib) or in second-line (e.g. regorafenib).(2) Advances in functional genomics highlighted HCC heterogeneity and identified clinically relevant HCC subgroups.(3, 4) Notably, we contributed to define a poor prognosis molecular subtype characterized by the activation of transforming growth factor beta (TGF β). (3-6)

TGF β is a pleiotropic cytokine involved in processes frequently deregulated in cancer, including cell proliferation, differentiation and migration. However, the actions of TGF β in cancer are complex since TGF β exhibits either onco-suppressive or pro-tumorigenic properties, depending on the tumor stage.(7, 8) At an early stage TGF β , exerts anti-proliferative properties in pre-malignant epithelial cells, but at an advanced stage, TGF β promotes tumor growth and metastatic progression, notably as a potent inducer of epithelial-mesenchymal transition (EMT).(7, 8) EMT is a dynamic and reversible process associated with disruption of cell-cell junctions and reorganization of cytoskeleton, favoring cell motility, metastasis and chemoresistance. At the molecular level, EMT is tightly executed by transcription factors, including SNAI, TWIST and ZEB families, regulating the expression of epithelial (e.g. E-cadherin, CDH1) and mesenchymal (e.g. Vimentin, VIM) markers.(9-11) Understanding the molecular mechanisms switching TGF β actions from tumor suppressive toward EMT and tumor promotion is clinically relevant to select patients who may benefit from interference with the TGF β pathway.

1
2
3 Forkhead box (FOX) proteins are transcriptional regulators characterized by a conserved
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Forkhead box (FOX) proteins are transcriptional regulators characterized by a conserved
forkhead DNA-binding domain.(12) So far, more than forty FOX proteins classified into 19
subfamilies (FOXA to FOXS) have been discovered in human.(13) They exhibit functional
diversity acting in cell proliferation, apoptosis, metabolism, differentiation and migration.(12) By
gene expression profiling, we identified novel TGF β -regulated genes in HCC cell lines, including
FOXS1, highly expressed in cells with a mesenchymal-like phenotype. Our goal was to
determine the role of *FOXS1* in the TGF β gene regulatory network and the clinical relevance of
its expression in cancer. *FOXS1* shares a high degree of similarity with EMT-related factor
FOXC2, which promotes HCC cell invasion.(14) Thus, we hypothesized that *FOXS1* could
mediate pro-metastatic actions of TGF β in poor prognosis HCC.

Materials and methods

Cell culture

Cell lines were purchased from ATCC (www.lgcstandards-atcc.org) and cultured as previously
described.(15) ATCC performed cell lines authentication by STR DNA profiling. Freshly isolated
human primary hepatocytes (HPH) were from Biopredic International (St Grégoire, France) and
cultured in William's E medium. Cells were treated with 1ng/mL recombinant human TGF β 1
(R&D system, Minneapolis, USA) and 10 μ M LY2157299 (Sigma-Aldrich, St. Louis, USA) after
overnight serum starvation.

Gain and loss of function

FOXS1 siRNA were from Ambion: si*FOXS1*#1 (5'-AAUCGGCCCAUGAUGUAGCgg-3'),
si*FOXS1*#2 (5'-CAUGCAAAAUUCAGUGCCgg-3') and negative control siRNA (siNC,
AM4620). Silencing of SMAD3 (Dharmacon, #4088), SMAD4 (Dharmacon, #4089) and *FOXA1*
(Ambion, 4392420) was performed according to manufacturer's instruction. Cells were
transfected at 70% confluency with 50nM siRNA using lipofectamine RNAiMAX and Opti-MEM.

1
2
3 Gain of function was performed by infecting cells with a HA-tagged-FOXS1-expressing lentivirus
4 (Lenti-pReceiver-Lv120, GeneCopoeia). FOXS1 knock-out was performed using custom
5 CRISPR-Cas9 (Sigma-Aldrich). Stable cell lines were generated after puromycin selection for 2
6 weeks and validated by RT-qPCR, Western Blot and sequencing of *FOXS1* locus.
7
8
9
10

11 **Gene expression profiling**

12 Gene expression profiling was performed using a low-input QuickAmp labeling kit and human
13 SurePrint G3 8x60K microarrays (Agilent Technologies, Santa Clara, USA). Datasets were
14 deposited into Gene Expression Omnibus (GSE148795). Integration of transcriptomic data was
15 conducted using publicly available datasets (GSE1898 and GSE4024) as previously
16 described.(16) The Cancer Genome Atlas (TCGA) RNA-seq data were used as validating
17 datasets (<https://portal.gdc.cancer.gov>). In total, 3 independent cohorts of HCC patients were
18 analyzed (697 cases). RT-qPCR was performed using a SYBR Green (Applied Biosystems,
19 Carlsbad, CA) and analyzed as previously described.(17) The list of primers is provided in Table
20 S1A.
21
22
23
24
25
26
27
28
29
30
31
32
33

34 **In situ hybridization (ISH)**

35 FOXS1 riboprobes were generated by *in vitro* transcription (IVT) from PCR products
36 incorporating the promoter of T7 RNA polymerase. IVT was performed with 40 units T7 in
37 presence of 0.35mM digoxigenin-11-UTP (Roche Diagnostics, Meylan, France). ISH was
38 performed on the Discovery Automated stainer (Ventana Medical Systems, Tucson, USA).(18)
39
40
41
42
43
44

45 **Protein expression analysis**

46 Protein analysis by immunoblotting was as previously described.(19) EMT antibody sampler kit
47 (#9782) was from Cell Signaling. For immunofluorescence studies, cells were fixed with 4% PFA
48 in PBS for 15min at room temperature. The list of antibodies is provided in Table S1B.
49
50
51
52
53
54

55 **Functional tests**

1
2
3 Cell viability was evaluated using a PrestoBlue® reagent (Invitrogen). Cell migration was
4 evaluated using silicone inserts with a defined cell-free gap (Ibidi) and ImageJ software.
5
6

7 8 **Human HCC**

9
10 Human HCC and surrounding non-tumor (ST) tissues were obtained from freshly resected
11 specimen (Biological Resource Centre, BB-0033-00056). The research protocol fulfilled national
12 laws and regulations and was approved by the local ethics committee and the institutional review
13 board of Inserm (IRB00003888). Precision-cut tissue slices (300µm) were performed with a
14 Vibratome Leica VT1200S and maintained in culture in William's E medium.
15
16
17
18
19

20 21 **Statistical analysis**

22
23 Statistical analyses were performed using R-3.5.1 and GraphPad 7.0. Differentially expressed
24 genes were identified by a 2-sample univariate *t* test and a random variance model.(16) For
25 group comparison of quantitative variables, *t* testing was applied. Categorical data were
26 analyzed by chi-squared testing. Survival was estimated by the Kaplan-Meier method and log-
27 rank test.
28
29
30
31
32
33

34 35 **Results**

36 37 38 **Novel TGFβ-regulated genes in SNU-449 liver cancer cells**

39
40 We previously highlighted early and late TGFβ signatures respectively associated with onco-
41 suppressive and pro-oncogenic properties of TGFβ in cancer.(5) Notably, the late TGFβ
42 signature predicts a poor prognosis in HCC patients and invasiveness of HCC cell lines,
43 including SNU-423, SNU-449 and SNU-475.(5) Here, SNU-449 was used as a paradigm to get
44 insights into the molecular mechanisms associated with this phenotype. Gene expression
45 profiling identified 80 genes differentially expressed after 72hrs TGFβ exposure (Fig. 1A and
46 Table S2). This signature established by using stringent statistical criteria ($P < 0.001$ and $FC > 2$)
47 included well-known TGFβ target genes such as *BMF*, *COL1A1*, *PDGFA*, *SKIL* and *TGFB1* (Fig.
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 1A) but also novel TGF β target gene candidates such as *FOXS1*, a transcriptional regulator of
4 the FOX family.(12) Upregulation of *FOXS1*, *COL1A1*, *SKIL* and *BMF* by TGF β was further
5 validated by RT-qPCR in SNU-449 (Fig. 1B).
6
7
8
9

10 **FOXS1 is a canonical SMAD-dependent TGF β target**

11
12 *FOXS1* expression was evaluated at basal level and after 16hrs TGF β treatment in a panel of 8
13 cell lines representative of HCC heterogeneity (Table S3), as well as in freshly isolated HPH.
14 Neither a basal expression nor a significant induction by TGF β was observed in HPH or in
15 PLC/PRF/5, HepG2-C3A, HuH6, and Hep3B HCC cell lines (Fig. 1C). However, *FOXS1* was
16 expressed and strongly induced by TGF β in SNU-423, SNU-449 and SNU-475, all associated
17 with the late TGF β signature (Fig. 1C). Accordingly, a reduced expression of *FOXS1* and
18 *SERPINE1*, a well-known TGF β transcriptional target, was observed in SNU-449 and SNU-475
19 cultured in presence of LY2157299 (galunisertib), a selective TGF β receptor type I kinase
20 inhibitor, both at basal level and upon TGF β exposure (Fig. 1D). Time course analysis and co-
21 treatments using cycloheximide translation inhibitor indicated that *FOXS1* is an early and direct
22 TGF β target (Fig. S1A-B). Similar to *SERPINE1*, a SMAD-dependent induction of *FOXS1* by
23 TGF β was demonstrated. Induction was abolished in presence of siRNAs targeting SMAD3
24 and/or SMAD4, in SNU-449 (Fig. 1E) and HMEC-TR mammary cells (Fig. 1F). Actually,
25 SMAD2/3 was shown to bind *FOXS1* promoter in SNU-449 (Fig. S1C). *FOXS1* induction by
26 TGF β was also observed in cholangiocarcinoma and pancreatic cancer cell lines, and in LX2
27 hepatic stellate cell line (Fig. 1G). *FOXS1* was induced or repressed in human tissues slices
28 (HCC and ST from patients undergoing surgery) exposed to TGF β or LY2157299, respectively
29 (Fig. 1H and Fig. S2). Altogether, these data demonstrate that *FOXS1* is a novel canonical
30 SMAD-dependent TGF β target.
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53 **FOXS1 expression is associated with a mesenchymal phenotype**

1
2
3 Next, we characterized *FOXS1* negative *versus* positive (i.e. basal expression and induction by
4 TGF β evidenced) HCC cell lines (Fig. 1C). *FOXS1* negative cell lines exhibited an epithelial-like
5 morphology with a cobblestone-like growth pattern while *FOXS1* positive cell lines exhibited a
6 mesenchymal-like phenotype with a spreading growth pattern (Fig. 2A). Gene expression
7 profiling supported the observed phenotypes. *FOXS1* positive cell lines exhibited an increased
8 expression of mesenchymal markers, including *VIM* and *TGFB2*, together with a decreased
9 expression of hepatocyte differentiation markers, including *FOXA1*, *FGG* and *APOE* (Fig. 2B-C
10 and Table S4). GSEA highlighted a positive enrichment of signatures associated with metastasis
11 and poor prognosis (e.g. TGF β , EMT, migration, cancer recurrence) in the gene expression
12 profiles of *FOXS1* positive cell lines together with a negative enrichment of signatures
13 associated with a well-differentiated hepatocyte phenotype (e.g. liver specific genes and
14 metabolisms) (Fig. 2D). The phenotype of *FOXS1* negative and positive cell lines was further
15 confirmed by the expression of *CDH1* and *VIM*, as evidenced by western-blotting (Fig. 2E) and
16 immunofluorescence (Fig. 2F). Analysis of *FOXS1* locus suggested that the absence of *FOXS1*
17 expression in well-differentiated HCC cell lines could result from an epigenetic silencing and/or
18 from a transcriptional repression by liver enriched *FOXA1* (Fig. S3A-B). Supporting an
19 epigenetic regulation, a globally lower methylation of *FOXS1* locus was observed in human HCC
20 as compared to the surrounding non-tumor tissue (Fig. S3C). Although *FOXA1* was induced by
21 TGF β in *FOXS1* negative cells, its silencing was not able to reactivate *FOXS1* expression (Fig.
22 S3D-E). Altogether, these data demonstrate that the basal expression of *FOXS1* and its
23 induction by TGF β in HCC cell lines correlate with a mesenchymal phenotype.

47 ***FOXS1* induces transcriptional changes associated with EMT and tumor progression**

48
49
50 Transient loss of function was achieved in SNU-449 by two siRNAs, which significantly
51 suppressed endogenous *FOXS1* expression and its induction by TGF β without affecting the
52 expression of other FOX family members (Fig. 3A and Fig. S4). Gene expression profiling
53
54
55
56
57
58
59
60

1
2
3 identified a robust subset of 177 non-redundant genes differentially expressed by the two
4 siRNAs in both untreated and TGF β treated cells (Fig. 3B and Table S5). Interestingly, 87%
5 genes were induced after FOXS1 silencing, suggesting that FOXS1 may act as a transcriptional
6 repressor (Fig. 3B-C). GSEA demonstrated that FOXS1 silencing impacted genes involved in β -
7 catenin and TGF β pathways, cell adhesion and associated with tumor invasion (Fig. 3C-D). The
8 expression of several key markers was further validated by RT-qPCR. Cell adhesion associated
9 genes (e.g. *CTNNB1*, *TJP1/ZO1*) were induced after FOXS1 silencing, while invasion and EMT-
10 associated genes (e.g. *MMP2*, *SNAI2*) were suppressed (Fig. 3E). Altogether, these data
11 demonstrate that FOXS1 induces specific transcriptional changes associated with EMT and
12 tumor progression.
13
14
15
16
17
18
19
20
21
22
23

24 **FOXS1 controls EMT and promotes cell migration**

25
26 To get deeper insight into the molecular mechanisms involved, we generated clonal and stable
27 cell lines with FOXS1 gain and loss of function in PLC/PRF/5 and SNU-449, respectively (Fig.
28 4A and Fig. S5). In SNU-449, FOXS1 knock-out by CRISPR-Cas9 resulted in a decreased basal
29 expression of mesenchymal markers (e.g. *VIM*, *SNAI1*, *SNAI2*) and a re-expression of epithelial
30 markers *CDH1* evocative of a mesenchymal-epithelial transition (MET) (Fig. 4A). Interestingly,
31 induction of *SNAI1* and *SNAI2* by TGF β was abolished, suggesting that FOXS1 acts upstream
32 to these master EMT transcription factors (Fig. 4A). An opposite regulation was observed in
33 epithelial-like PLC/PRF/5 after FOXS1 gain of function, including the expression of
34 mesenchymal markers *VIM*, *SNAI1* and *ZEB1*, as well as repression of *CTNNB1* (Fig. 4A).
35 These observations were validated by immunofluorescence (Fig. 4B) and by RT-qPCR (Fig. 4C).
36 Activation of endogenous *FOXS1* by CRISPR-Cas9 Synergistic Activation Mediator in SNU-449
37 cells also resulted in *SNAI1* induction (Fig. S6). Further supporting a role in EMT, FOXS1
38 expression positively correlates with *VIM* expression and negatively correlates with *CDH1*
39 expression in several *in vitro* models of EMT and MET (Fig. S7). The analysis of 34 HCC cell
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

lines also demonstrated that the gene expression profiles of highly FOXS1 expressing cells are enriched in signatures associated with EMT, cell migration and liver recurrence in patients with HCC(20) (Fig. S8). Accordingly, FOXS1 gain and loss of function in SNU-449 cells resulted in a significant increased and decreased of cell migration, respectively (Fig. 4D). Altogether, these data demonstrate that FOXS1 acts as an upstream regulator of EMT and is required to maintain a mesenchymal phenotype.

FOXS1 expression predicts a poor prognosis in human cancer

The clinical relevance of *FOXS1* expression was first determined by meta-analyzing previously published cancer datasets. Increased expression of *FOXS1* was observed in numerous cancers, including HCC and cholangiocarcinoma, where it correlates with tumor stage and *VIM* expression (Fig. S9). In an independent cohort of 193 HCC, we validated a specific increased of *FOXS1* expression in HCC as compared to surrounding non-tumor tissues and a correlation with *VIM* expression. Interestingly, *FOXS1* expression correlated with the degree of fibrosis, *FOXS1* being significantly more expressed in tumors developed on a cirrhotic background (Fig. S10A-B). Specific increased of *FOXS1* expression in the tumor tissue was further demonstrated by *in situ* hybridization on human resected HCC (Fig. 5A). Interestingly, FOXS1 positive cells were detected in fibrotic areas (Fig. 5A), in agreement with our previous observation of FOXS1 induction in hepatic stellate cells upon TGF β treatment (Fig. 1G). Heterogeneous expression of *FOXS1* in HCC tissues suggested a correlation with clinically relevant HCC subtypes (Fig. S9B and S10A). Supporting this hypothesis, *FOXS1* expression was significantly associated with G1-G6 molecular classification(21) with a higher expression in poor prognosis G1 HCC subtypes and a lower expression in well-differentiated G5-G6 HCC subtypes with no *CTNNB1* mutations (Fig. S10C-D). To address the clinical relevance of FOXS1 more precisely, we integrated the endogenous FOXS1 loss of function gene expression signature generated in SNU-449 (Fig. 3B) with the gene expression profiles of 139 cases of human HCC extensively characterized.(22-24)

1
2
3 Hierarchical clustering analysis of the integrated dataset identified 2 major clusters (Fig. 5B).
4
5 Cluster 1 and cluster 2 included human HCC associated with the siFOXS1 (low FOXS1 activity)
6
7 and the siNC (high FOXS1 activity) SNU-449 samples, respectively. Interestingly, the biological
8
9 and clinical parameters of HCC were not randomly distributed between clusters 1 and 2 (Fig. 5B-
10
11 D). Strikingly, cluster 1 (low FOXS1) included significantly more tumors, which were previously
12
13 assigned to a better prognosis group, than cluster 2 (Fig. 5B-C). Conversely, cluster 2 (high
14
15 FOXS1), was significantly enriched in HCC previously defined by poor prognosis signatures,
16
17 including poor survival-associated genes(23), hepatoblast/progenitor traits(24) and activation of
18
19 a pro-oncogenic MET/HGF pathway(22) (Fig. 5B-C). More interestingly, cluster 2 was
20
21 significantly enriched in HCC harboring the pro-metastatic late TGF β signature while cluster 1
22
23 was enriched in HCC with the early TGF β signature associated with the cytostatic properties of
24
25 TGF β (5) (Fig. 5B-C). In addition, cluster 2 was significantly enriched in bigger (>5cm) and less
26
27 differentiated (Edmondson grade 3+4) HCC with signs of vascular invasion and occurring in
28
29 younger patients (Age <60 years). Serum alpha-fetoprotein (AFP) was significantly higher in
30
31 these FOXS1-high aggressive HCC (Fig. 5B-C). Higher proliferation and lower apoptosis
32
33 indexes were significantly associated with FOXS1-high HCC (Fig. 5D). Accordingly, high FOXS1
34
35 expression was found to predict a reduced survival (Fig. 5E). Finally, *FOXS1* expression was
36
37 directly evaluated in 17 types of human cancer from RNA-seq data generated by TCGA. *FOXS1*
38
39 was found to be detected across all 17 types of cancer tissues (Fig. 5F). More importantly, high
40
41 expression of *FOXS1* significantly ($P<0.05$) predicted a poor survival in liver, colon, stomach and
42
43 kidney cancers (Fig. 5G). Altogether, these data demonstrate that *FOXS1* expression is a
44
45 predictor of poor prognosis in human cancer.
46
47
48
49

50 Discussion

51
52
53
54 In this study, we characterized the role of FOXS1 and the clinical relevance of its expression in
55
56 cancer. We report that *FOXS1* is a novel TGF β target which expression and induction occur in
57
58
59
60

1
2
3 cells with a mesenchymal-like phenotype. We further report that *FOXS1* controls EMT, induces
4 cell migration and predicts a poor prognosis in cancer.
5
6
7

8 FOX proteins are tightly regulated during embryonic development and are required for the
9 normal specification, differentiation and function of numerous tissues, including the liver.(12)
10 *FOXS1* has been poorly investigated so far. Previous studies reported a critical role in
11 embryogenesis, especially for normal brain and testis development.(25-27) Interestingly, *FOXS1*
12 was reported to be silenced in most normal tissues, its expression being restricted to aorta and
13 to a lesser extent to kidney.(28) Our data suggest that *FOXS1* could be epigenetically silenced
14 in non-tumor tissues through methylation of CpG islands. *FOXS1* re-expression in cancer is in
15 agreement with observations reporting hypomethylation-associated upregulation of *Foxs1* in a
16 mouse model of gonadectomy-induced adrenocortical neoplasia.(29) This specific gene
17 expression pattern highlights *FOXS1* as a novel oncofetal gene possibly inactivated by
18 methylation at birth and reactivated in cancer, similar to AFP in liver cancer.(30)
19
20
21
22
23
24
25
26
27
28
29
30
31

32 At the transcriptional level, we demonstrated that *FOXS1* is a direct target of SMAD2/3/4-
33 dependent canonical TGF β pathway in mesenchymal-like cells. In addition to epigenetic
34 silencing, we hypothesized that a transcriptional repression by liver-enriched FOXA1 could occur
35 in epithelial cells. FOXS1 repression by FOXA1 was supported by the differential expression of
36 FOXA1 in FOXS1 negative versus positive cell lines both at basal level and upon TGF β
37 treatment. However, FOXA1 silencing was not able to reactivate FOXS1 expression in negative
38 cell lines, suggesting that transcriptional repression is probably not the main mechanism of
39 *FOXS1* silencing in well-differentiated cells. A modulation by the Sonic Hedgehog pathway has
40 been highlighted, although the reports are conflicting. *FOXS1* was induced by GLI1 in
41 rhabdomyosarcoma and medulloblastoma(31) but repressed by GLI1 in gastric cancer(32)
42 suggesting that a more complex interplay between cancer specific transcriptional or post-
43 transcriptional regulators. Indeed, miR-125a-5p was identified as a potent inhibitor of FOXS1
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 expression in gastric cancer. Interestingly, miR-125a-5p is a well-known tumor suppressor
4 microRNA in HCC(33),acting as a potent inhibitor of EMT and cell invasion(34). Altogether, the
5 data demonstrate a complex FOXS1 regulation occurring at epigenetic, transcriptional and post-
6 transcriptional level in a specific cellular context among tissue and cancer types.
7
8
9

10
11
12 Regarding FOXS1 role in cancer, there is no consensus in the literature yet. FOXS1 has been
13 reported to exhibit anti-proliferative effects in medulloblastoma.(31) In gastric cancer, conflicting
14 data have been reported.(32, 35) FOXS1 overexpression inhibits EMT in SNU-216 cells but
15 promotes cell proliferation, migration, invasion and EMT in SGC7901 and BGC823 cells.(32)
16 These opposite results suggest that FOXS1 function is highly dependent of cell specific
17 transcriptional contexts associated to each cancer subtypes. It was previously hypothesized that
18 FOXS1 prevent apoptosis of endothelial cells by blocking FOXO3 and FOXO4 action through
19 promoter binding competition.(27) Our gene expression profiling suggests that FOXS1 may
20 indeed act as a transcriptional repressor. If FOXS1 acts as a decoy factor, it could be blocking
21 activators or inhibitors at the same locus thus changing its action depending on the
22 transcriptional context. Two-hybrid experiments have highlighted a potential interaction between
23 FOXS1 and BMI1.(36) BMI1 is a core component of the polycomb repressive complex 1.(37)
24 This interaction could explain the transcriptional repressive activity of FOXS1.
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40
41 Regardless of previously conflicting reported studies in gastric cancer, our data in HCC show
42 clearly a regulatory role of FOXS1 in EMT, as evidenced by the modulation of molecular EMT
43 markers (e.g. VIM and CDH1) and effectors (e.g. SNAI1 and SNAI2) as well as the impact of
44 FOXS1 on cell migration. Those results are in agreement with the reported role of FOXS1 during
45 embryonic development where FOXS1 is expressed in migrating neural crest cells.(26) Our data
46 place FOXS1 in a classical model of EMT where an inducible signal (TGF β) activate EMT-
47 transcription factors (FOXS1, SNAI1, SNAI2) which regulate the expression of EMT markers
48 (VIM, CDH1) (Fig. 6). EMT occurs in a gradual manner characterized by several states referred
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 to as epithelial, early hybrid, late hybrid and full EMT states.(9, 38) Thus, even if an over-
4
5 activation of TGF β signaling can promote a full EMT in different cancers such as HCC(39), a
6
7 recent study showed that an inhibition of TGF β or Smad2 signaling blocks EMT progression at
8
9 the early stages of EMT.(9) Total absence of *FOXS1* expression in epithelial cells and
10
11 reappearance of epithelial features in *FOXS1* knock-out SNU-449 cells suggest that *FOXS1*
12
13 could not only act as an inducer of EMT but also as a maintainer of the mesenchymal state.
14
15

16
17 As for its role in cancer, the clinical relevance of *FOXS1* expression has been controversial. High
18
19 expression of *FOXS1* is associated with better relapse free survival in patient with breast
20
21 cancer.(31) Conversely, *FOXS1* is enriched in a 9-genes expression signature of poor prognosis
22
23 in gastric cancer(40), in agreement with our analysis identifying *FOXS1* as a marker of poor
24
25 prognosis in liver, kidney, colon and stomach cancer. We also validated the clinical relevance of
26
27 *FOXS1* by using an integrative transcriptomic approach based on a set of genes that reflect
28
29 *FOXS1* activity (Fig. 6). In addition, our data highlight *FOXS1* as a promising biomarker to
30
31 identify tumor where the pro-metastatic arm of TGF β arm is active (Fig. 6) and thus, patients
32
33 who may benefit from targeted therapies using TGF β inhibitors. This is particularly relevant in
34
35 the context of treatment personalization in HCC with a phase II clinical trial (NCT02178358)
36
37 currently evaluating LY2157299 (galunisertib) in patients with advanced HCC.(41)
38
39

40
41 In conclusion, our study identified *FOXS1* as a novel canonical TGF β target gene. *FOXS1*
42
43 controls EMT and cell migration through induction of *SNAI1/2*, *VIM* and decrease of cell junction
44
45 proteins. *FOXS1* is overexpressed in HCC and predicts a poor prognosis in liver, stomach, colon
46
47 and kidney cancer.
48

49
50 **Acknowledgments:** The authors thank C. Allain, J. Chivot, A. Gautron, B. Maillot, R. Métivier
51
52 and N. Tournadre for their technical help. The authors thank S. Dréano for DNA sequencing and
53
54 the core facilities of Biosit and Biogenouest (CRB Santé BB-0033-00056, GEH, H2P2,
55
56
57
58
59
60

1
2
3 ImpPACcell). The authors thank L. Gómez-Quiroz, M. Levrero and G. Salbert for helpful
4
5 discussions.
6
7

8 **Figure legends**

9

10
11 **Fig. 1. FOXS1 is a novel SMAD-dependent TGF β target.** (A) Volcano plot and hierarchical
12 clustering of genes differentially expressed by TGF β in SNU-449. (B) RT-qPCR of *FOXS1* and
13 well-known targets in SNU-449 (1ng/mL TGF β , 72hrs). (C) RT-qPCR and western blot (inset)
14 analysis of *FOXS1* expression in human primary hepatocytes (HPH) and 8 HCC cell lines
15 (1ng/mL TGF β , 16hrs). (D) RT-qPCR of *SERPINE1* and *FOXS1* in SNU-449 and SNU-475 cells
16 (1ng/mL TGF β , 10 μ M LY2157299, 16hrs). (E) RT-qPCR analysis of *SMAD3*, *SMAD4*,
17 *SERPINE1* and *FOXS1* in SNU-449 cells transfected with siSMAD3 and siSMAD4, alone or in
18 combination (50nM each), or a negative control siRNA (siNC, 50nM) (1ng/mL TGF β , 72hrs). (F)
19 *SERPINE1* and *FOXS1* expression in HMEC-TR mammary epithelial cells transfected with
20 siSMAD4 or siNC (GSE28448: 5ng/mL TGF β , 24hrs). (G) *SERPINE1*, *SMAD7* and *FOXS1*
21 expression in HuCCT1 and Huh28 cholangiocarcinoma cells (1ng/mL TGF β , 72hrs), HK2
22 proximal tubular cells (GDS3649: 2ng/mL TGF β , 72hrs), LX2 hepatic stellate cell line (1ng/mL
23 TGF β , 16hrs) and Panc1 pancreatic adenocarcinoma cells (GSE23952: 5ng/mL TGF β , 42hrs).
24 (H) RT-qPCR of *FOXS1* in human surrounding non-tumor (ST, left panel, n=6 patients) or tumor
25 (T, right panel, n=7 patients) tissue slices derived from patients after surgical resection (1ng/mL
26 TGF β , 48hrs). Statistical analysis was performed by a Wilcoxon matched-pairs signed rank test.
27 In B-G panels, statistical analysis was performed by a t-test (** P<0.01; *** P<0.001; ns P>0.05;
28 n \geq 3 replicates)
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51

52 **Fig. 2. FOXS1 positive HCC cell lines exhibit a mesenchymal phenotype.** (A) Phase
53 contrast micrographs of representative *FOXS1* negative (upper panels) and positive (lower
54 panels) HCC cell lines (scale bar: 100 μ m). (B) Volcano plot of genes differentially expressed
55
56
57
58
59
60

1
2
3 between FOXS1 negative (PLC/PRF/5, HepG2-C3A, HuH6) and positive (SNU-423, SNU-449,
4 SNU-475) cell lines. (C) Hierarchical clustering analysis and examples of genes down- and up-
5 regulated in FOXS1 positive cell lines (* $P < 0.05$). (D) GSEA of signatures significantly enriched
6 (P < 0.01) in the gene expression profiles of FOXS1 positive (black box) and negative (white box)
7 HCC cell lines. (E-F) Expression of epithelial marker E-cadherin (CDH1) and mesenchymal
8 marker Vimentin (VIM) at protein level in HCC cell lines, as determined by western-blot (E) and
9 immunofluorescence (F) in FOXS1 negative and positive HCC cell lines. Scale bar: 50 μ m.
10
11
12
13
14
15
16
17
18
19

20 **Fig. 3. Specific transcriptional changes evocative of EMT and tumor progression after**
21 **FOXS1 silencing.** (A) Validation of endogenous FOXS1 silencing in SNU-449 cells by western
22 blot (upper part) and RT-qPCR (lower part) (1ng/mL TGF β , 72hrs). (B) Venn diagram analysis of
23 genes differentially expressed by FOXS1 silencing, in presence or absence of TGF β (P < 0.001;
24 FC > 1.5). A stringent signature made of 182 probes (177 non-redundant genes) reflecting
25 endogenous FOXS1 silencing was defined. Hierarchical clustering analysis based on the
26 expression of 182 probes highlighted a global gene induction (87% genes). (C) Top enriched
27 pathways in FOXS1 silencing signature. (D) GSEA using gene signatures involved in Wnt/ β -
28 catenin and TGF β pathways, and tumor invasion. (E) Validation by RT-qPCR of key genes
29 modulated after FOXS1 silencing from the gene expression profiling by microarray in SNU-449
30 (1ng/mL TGF β , 72hrs). Statistical analysis was performed by a t-test (* $P < 0.05$; ** $P < 0.01$; ***
31 $P < 0.001$; n \geq 3 replicates).
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

47 **Fig. 4. FOXS1 induces EMT markers and promotes cell migration.** (A) Western blot analysis
48 of EMT markers after stable FOXS1 loss and gain of function in clonal SNU-449 and
49 PLC/PRF/5, respectively (1ng/mL TGF β , 16hrs). (B) Evaluation of EMT markers by
50 immunofluorescence (scale bar: 100 μ m). (C) RT-qPCR of EMT markers (1ng/mL TGF β , 16hrs).
51
52
53
54
55
56
57
58
59
60

1
2
3 (D) Migration assay in cells after mitomycin C treatment (10µg/mL). Statistical analysis was
4 performed by a t-test (* P<0.05; ** P<0.01; *** P<0.001; n≥3 replicates).
5
6
7
8

9 **Fig. 5. FOXS1 predicts a poor prognosis in human cancer.** (A) *In situ* hybridization on human
10 HCC sample using FOXS1 sense (S) and antisense (AS) riboprobes. Black arrows: FOXS1
11 expression in fibroblast-like cells in the interface fibrotic tissue (FT). (B) Clinical relevance of
12 FOXS1 silencing signature in HCC. Dendrogram overview of *in vitro* experiments using SNU-
13 449 cells transfected with siNC or siFOXS1 integrated with 139 cases of human HCC based on
14 the expression of 177 genes differentially expressed after FOXS1 silencing. Two major clusters
15 were identified: cluster 1 associated with FOXS1 silencing and cluster 2 associated with high
16 FOXS1 expression. Distribution of HCC between previously described subgroups with respect to
17 survival (23) (good vs poor prognosis), cell origin (24) (hepatoblast vs hepatocyte), activation of
18 MET/HGF (22) (Negative vs Positive) and TGFβ signaling pathway (5) (early vs late) is indicated
19 on the left side. (C) Statistical analysis of HCC distribution between clusters 1 and 2 based on
20 previous gene signatures and clinical parameters. (D) HCC included in cluster 2 exhibited a
21 higher proliferation index and a lower apoptotic index, as previously defined (24). (E) Kaplan-
22 Meier plots and log-rank statistics analysis revealed a significant decreased in overall survival for
23 patients included in cluster 2. (F) Expression of FOXS1 in 17 cancer types from the TCGA
24 consortium. (G) Kaplan-Meier plots and log-rank statistics analysis revealed a significant
25 decreased in overall survival for patients with a high expression of FOXS1 in liver, colon,
26 stomach and kidney cancer.
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 **Fig. 6. Functional role and clinical relevance of FOXS1 in HCC.**
50
51
52
53

54 REFERENCES

55
56
57
58
59
60

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394-424.
2. Villanueva A. Hepatocellular Carcinoma. *N Engl J Med* 2019;380:1450-1462.
3. Hoshida Y, Nijman SM, Kobayashi M, Chan JA, Brunet JP, Chiang DY, Villanueva A, et al. Integrative transcriptome analysis reveals common molecular subclasses of human hepatocellular carcinoma. *Cancer Res* 2009;69:7385-7392.
4. Zucman-Rossi J, Villanueva A, Nault JC, Llovet JM. Genetic Landscape and Biomarkers of Hepatocellular Carcinoma. *Gastroenterology* 2015;149:1226-1239 e1224.
5. Coulouarn C, Factor VM, Thorgeirsson SS. Transforming growth factor-beta gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer. *Hepatology* 2008;47:2059-2067.
6. Giannelli G, Villa E, Lahn M. Transforming growth factor-beta as a therapeutic target in hepatocellular carcinoma. *Cancer Res* 2014;74:1890-1894.
7. Ikushima H, Miyazono K. TGFbeta signalling: a complex web in cancer progression. *Nat Rev Cancer* 2010;10:415-424.
8. Seoane J, Gomis RR. TGF-beta Family Signaling in Tumor Suppression and Cancer Progression. *Cold Spring Harb Perspect Biol* 2017;9.
9. Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, Van Keymeulen A, et al. Identification of the tumour transition states occurring during EMT. *Nature* 2018;556:463-468.
10. Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol* 2017;14:611-629.
11. Stemmler MP, Eccles RL, Brabletz S, Brabletz T. Non-redundant functions of EMT transcription factors. *Nat Cell Biol* 2019;21:102-112.
12. Lam EW, Brosens JJ, Gomes AR, Koo CY. Forkhead box proteins: tuning forks for transcriptional harmony. *Nat Rev Cancer* 2013;13:482-495.
13. Wang J, Li W, Zhao Y, Kang, Fu W, Zheng X, Pang X, et al. Members of FOX family could be drug targets of cancers. *Pharmacol Ther* 2017.
14. Yang F, Lv L, Zhang K, Cai Q, Liu J, Jiang Y. Elevated FOXC2 Expression Promotes Invasion of HCC Cell Lines and is Associated with Poor Prognosis in Hepatocellular Carcinoma. *Cell Physiol Biochem* 2017;44:99-109.
15. Allain C, Angenard G, Clement B, Coulouarn C. Integrative Genomic Analysis Identifies the Core Transcriptional Hallmarks of Human Hepatocellular Carcinoma. *Cancer Res* 2016;76:6374-6381.
16. Coulouarn C, Corlu A, Glaise D, Guenon I, Thorgeirsson SS, Clement B. Hepatocyte-stellate cell cross-talk in the liver engenders a permissive inflammatory microenvironment that drives progression in hepatocellular carcinoma. *Cancer Res* 2012;72:2533-2542.
17. Coulouarn C, Factor VM, Andersen JB, Durkin ME, Thorgeirsson SS. Loss of miR-122 expression in liver cancer correlates with suppression of the hepatic phenotype and gain of metastatic properties. *Oncogene* 2009;28:3526-3536.
18. Merdrignac A, Angenard G, Allain C, Petitjean K, Bergeat D, Bellaud P, Fautrel A, et al. A novel transforming growth factor beta-induced long noncoding RNA promotes an inflammatory microenvironment in human intrahepatic cholangiocarcinoma. *Hepatology Commun* 2018;2:254-269.
19. Dubois-Pot-Schneider H, Fekir K, Coulouarn C, Glaise D, Aninat C, Jarnouen K, Le Guevel R, et al. Inflammatory cytokines promote the retrodifferentiation of tumor-derived hepatocyte-like cells to progenitor cells. *Hepatology* 2014;60:2077-2090.
20. Caruso S, Calatayud AL, Pilet J, La Bella T, Rekik S, Imbeaud S, Letouze E, et al. Analysis of Liver Cancer Cell Lines Identifies Agents With Likely Efficacy Against Hepatocellular Carcinoma and Markers of Response. *Gastroenterology* 2019;157:760-776.

21. Boyault S, Rickman DS, de Reynies A, Balabaud C, Rebouissou S, Jeannot E, Herault A, et al. Transcriptome classification of HCC is related to gene alterations and to new therapeutic targets. *Hepatology* 2007;45:42-52.
22. Kaposi-Novak P, Lee JS, Gomez-Quiroz L, Coulouarn C, Factor VM, Thorgeirsson SS. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest* 2006;116:1582-1595.
23. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, Durnez A, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* 2004;40:667-676.
24. Lee JS, Heo J, Libbrecht L, Chu IS, Kaposi-Novak P, Calvisi DF, Mikaelyan A, et al. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 2006;12:410-416.
25. Heglind M, Cederberg A, Aquino J, Lucas G, Ernfors P, Enerback S. Lack of the central nervous system- and neural crest-expressed forkhead gene *Foxs1* affects motor function and body weight. *Mol Cell Biol* 2005;25:5616-5625.
26. Montelius A, Marmigere F, Baudet C, Aquino JB, Enerback S, Ernfors P. Emergence of the sensory nervous system as defined by *Foxs1* expression. *Differentiation* 2007;75:404-417.
27. Sato Y, Baba T, Zubair M, Miyabayashi K, Toyama Y, Maekawa M, Owaki A, et al. Importance of forkhead transcription factor *Fkh18* for development of testicular vasculature. *Mol Reprod Dev* 2008;75:1361-1371.
28. Cederberg A, Betz R, Lagercrantz S, Larsson C, Hulander M, Carlsson P, Enerback S. Chromosome localization, sequence analysis, and expression pattern identify *FKHL 18* as a novel human forkhead gene. *Genomics* 1997;44:344-346.
29. Schillebeeckx M, Pihlajoki M, Gretzinger E, Yang W, Thol F, Hiller T, Lobs AK, et al. Novel markers of gonadectomy-induced adrenocortical neoplasia in the mouse and ferret. *Mol Cell Endocrinol* 2015;399:122-130.
30. Coulouarn C, Derambure C, Lefebvre G, Daveau R, Hiron M, Scotte M, Francois A, et al. Global gene repression in hepatocellular carcinoma and fetal liver, and suppression of *dudulin-2* mRNA as a possible marker for the cirrhosis-to-tumor transition. *J Hepatol* 2005;42:860-869.
31. Diao Y, Rahman MF, Vyatkin Y, Azatyan A, St Laurent G, Kapranov P, Zaphiropoulos PG. Identification of novel *GLI1* target genes and regulatory circuits in human cancer cells. *Mol Oncol* 2018;12:1718-1734.
32. Wang S, Ran L, Zhang W, Leng X, Wang K, Liu G, Song J, et al. *FOXS1* is regulated by *GLI1* and *miR-125a-5p* and promotes cell proliferation and EMT in gastric cancer. *Sci Rep* 2019;9:5281.
33. Ming M, Ying M, Ling M. *miRNA-125a-5p* inhibits hepatocellular carcinoma cell proliferation and induces apoptosis by targeting *TP53* regulated inhibitor of apoptosis 1 and *Bcl-2-like-2* protein. *Exp Ther Med* 2019;18:1196-1202.
34. Tang L, Zhou L, Wu S, Shi X, Jiang G, Niu S, Ding D. *miR-125a-5p* inhibits colorectal cancer cell epithelial-mesenchymal transition, invasion and migration by targeting *TAZ*. *Oncotargets Ther* 2019;12:3481-3489.
35. Lu Q, Ma X, Li Y, Song W, Zhang L, Shu Y, Wan B. Overexpression of *FOXS1* in gastric cancer cell lines inhibits proliferation, metastasis, and epithelial-mesenchymal transition of tumor through downregulating *wnt/beta-catenin* pathway. *J Cell Biochem* 2019;120:2897-2907.
36. Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, Tan K, Akalin A, et al. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* 2010;140:744-752.
37. Cao R, Tsukada Y, Zhang Y. Role of *Bmi-1* and *Ring1A* in H2A ubiquitylation and *Hox* gene silencing. *Mol Cell* 2005;20:845-854.
38. Zhang J, Tian XJ, Zhang H, Teng Y, Li R, Bai F, Elankumaran S, et al. TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops. *Sci Signal* 2014;7:ra91.

- 1
2
3 39. Bertran E, Crosas-Molist E, Sancho P, Caja L, Lopez-Luque J, Navarro E, Egea G, et al.
4 Overactivation of the TGF-beta pathway confers a mesenchymal-like phenotype and CXCR4-
5 dependent migratory properties to liver tumor cells. *Hepatology* 2013;58:2032-2044.
6 40. Wang Z, Chen G, Wang Q, Lu W, Xu M. Identification and validation of a prognostic 9-
7 genes expression signature for gastric cancer. *Oncotarget* 2017;8:73826-73836.
8 41. Faivre S, Santoro A, Kelley RK, Gane E, Costentin CE, Gueorguieva I, Smith C, et al.
9 Novel transforming growth factor beta receptor I kinase inhibitor galunisertib (LY2157299) in
10 advanced hepatocellular carcinoma. *Liver Int* 2019;39:1468-1477.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

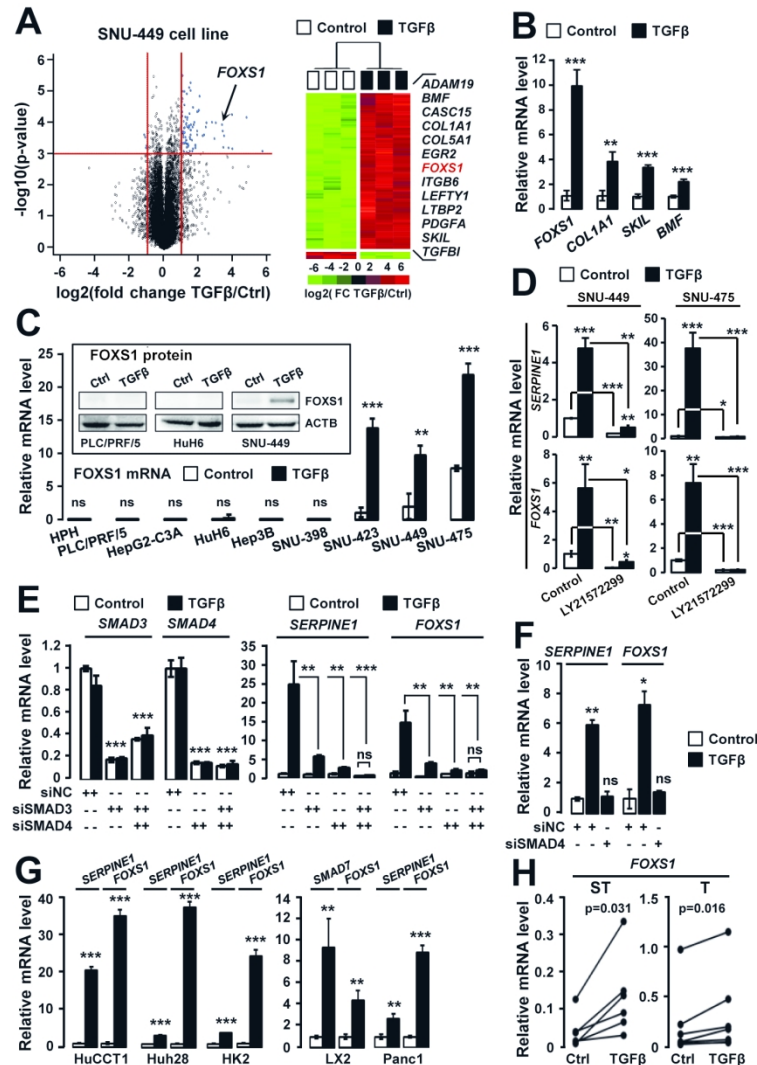


Fig. 1. FOXS1 is a novel SMAD-dependent TGF β target. (A) Volcano plot and hierarchical clustering of genes differentially expressed by TGF β in SNU-449. (B) RT-qPCR of FOXS1 and well-known targets in SNU-449 (1ng/mL TGF β , 72hrs). (C) RT-qPCR and western blot (inset) analysis of FOXS1 expression in human primary hepatocytes (HPH) and 8 HCC cell lines (1ng/mL TGF β , 16hrs). (D) RT-qPCR of SERPINE1 and FOXS1 in SNU-449 and SNU-475 cells (1ng/mL TGF β , 10 μ M LY21572299, 16hrs). (E) RT-qPCR analysis of SMAD3, SMAD4, SERPINE1 and FOXS1 in SNU-449 cells transfected with siSMAD3 and siSMAD4, alone or in combination (50nM each), or a negative control siRNA (siNC, 50nM) (1ng/mL TGF β , 72hrs). (F) SERPINE1 and FOXS1 expression in HMEC-TR mammary epithelial cells transfected with siSMAD4 or siNC (GSE28448: 5ng/mL TGF β , 24hrs). (G) SERPINE1, SMAD7 and FOXS1 expression in HuCCT1 and Huh28 cholangiocarcinoma cells (1ng/mL TGF β , 72hrs), HK2 proximal tubular cells (GDS3649: 2ng/mL TGF β , 72hrs), LX2 hepatic stellate cell line (1ng/mL TGF β , 16hrs) and Panc1 pancreatic adenocarcinoma cells (GSE23952: 5ng/mL TGF β , 42hrs). (H) RT-qPCR of FOXS1 in human surrounding non-tumor (ST, left panel, n=6 patients) or tumor (T, right panel, n=7 patients) tissue slices derived from patients after surgical resection (1ng/mL TGF β , 48hrs). Statistical analysis was performed by a Wilcoxon matched-pairs signed rank test. In B-G panels, statistical analysis was performed by a t-test (** P<0.01; *** P<0.001; ns P>0.05; n \geq 3 replicates)

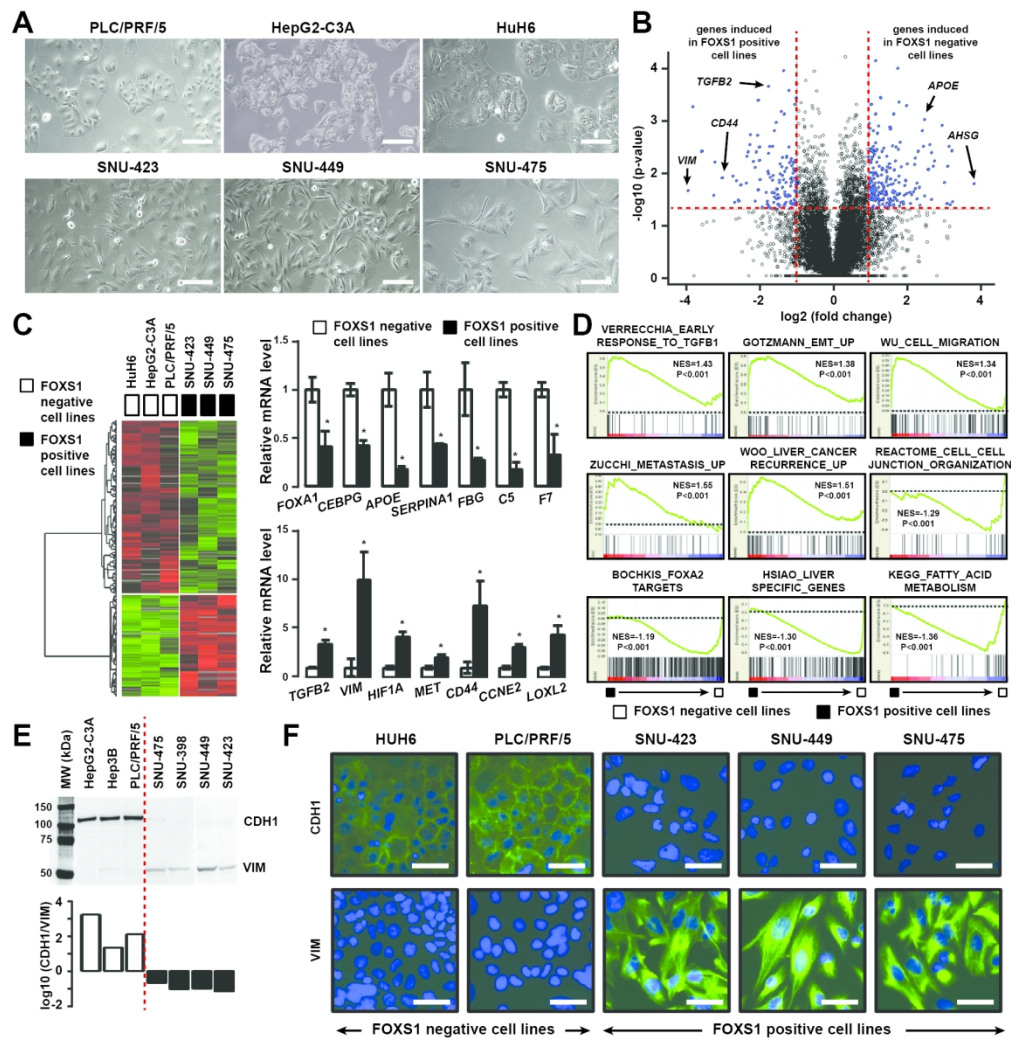


Fig. 2. FOXS1 positive HCC cell lines exhibit a mesenchymal phenotype. (A) Phase contrast micrographs of representative FOXS1 negative (upper panels) and positive (lower panels) HCC cell lines (scale bar: 100 μ m). (B) Volcano plot of genes differentially expressed between FOXS1 negative (PLC/PRF/5, HepG2-C3A, HuH6) and positive (SNU-423, SNU-449, SNU-475) cell lines. (C) Hierarchical clustering analysis and examples of genes down- and up-regulated in FOXS1 positive cell lines (* P<0.05). (D) GSEA of signatures significantly enriched (P<0.01) in the gene expression profiles of FOXS1 positive (black box) and negative (white box) HCC cell lines. (E-F) Expression of epithelial marker E-cadherin (CDH1) and mesenchymal marker Vimentin (VIM) at protein level in HCC cell lines, as determined by western-blot (E) and immunofluorescence (F) in FOXS1 negative and positive HCC cell lines. Scale bar: 50 μ m.

163x168mm (300 x 300 DPI)

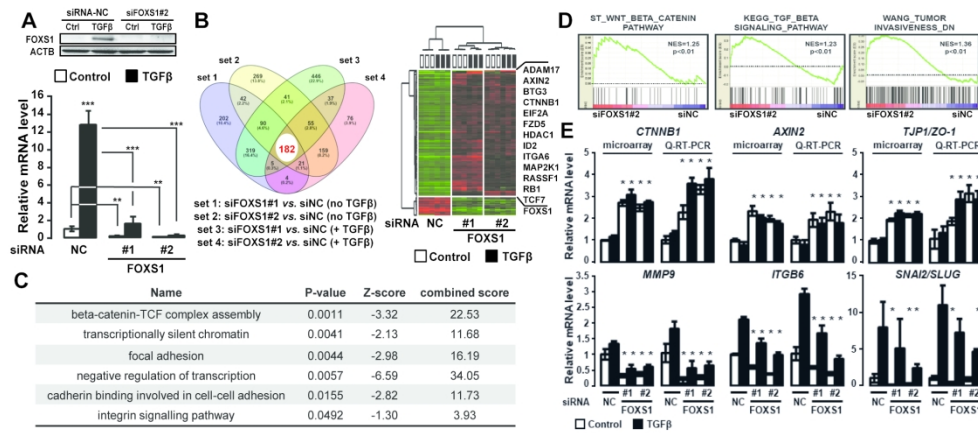


Fig. 3. Specific transcriptional changes evocative of EMT and tumor progression after FOXS1 silencing. (A) Validation of endogenous FOXS1 silencing in SNU-449 cells by western blot (upper part) and RT-qPCR (lower part) (1ng/mL TGFβ, 72hrs). (B) Venn diagram analysis of genes differentially expressed by FOXS1 silencing, in presence or absence of TGFβ ($P < 0.001$; $FC > 1.5$). A stringent signature made of 182 probes (177 non-redundant genes) reflecting endogenous FOXS1 silencing was defined. Hierarchical clustering analysis based on the expression of 182 probes highlighted a global gene induction (87% genes). (C) Top enriched pathways in FOXS1 silencing signature. (D) GSEA using gene signatures involved in Wnt/ β -catenin and TGFβ pathways, and tumor invasion. (E) Validation by RT-qPCR of key genes modulated after FOXS1 silencing from the gene expression profiling by microarray in SNU-449 (1ng/mL TGFβ, 72hrs). Statistical analysis was performed by a t-test (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; $n \geq 3$ replicates).

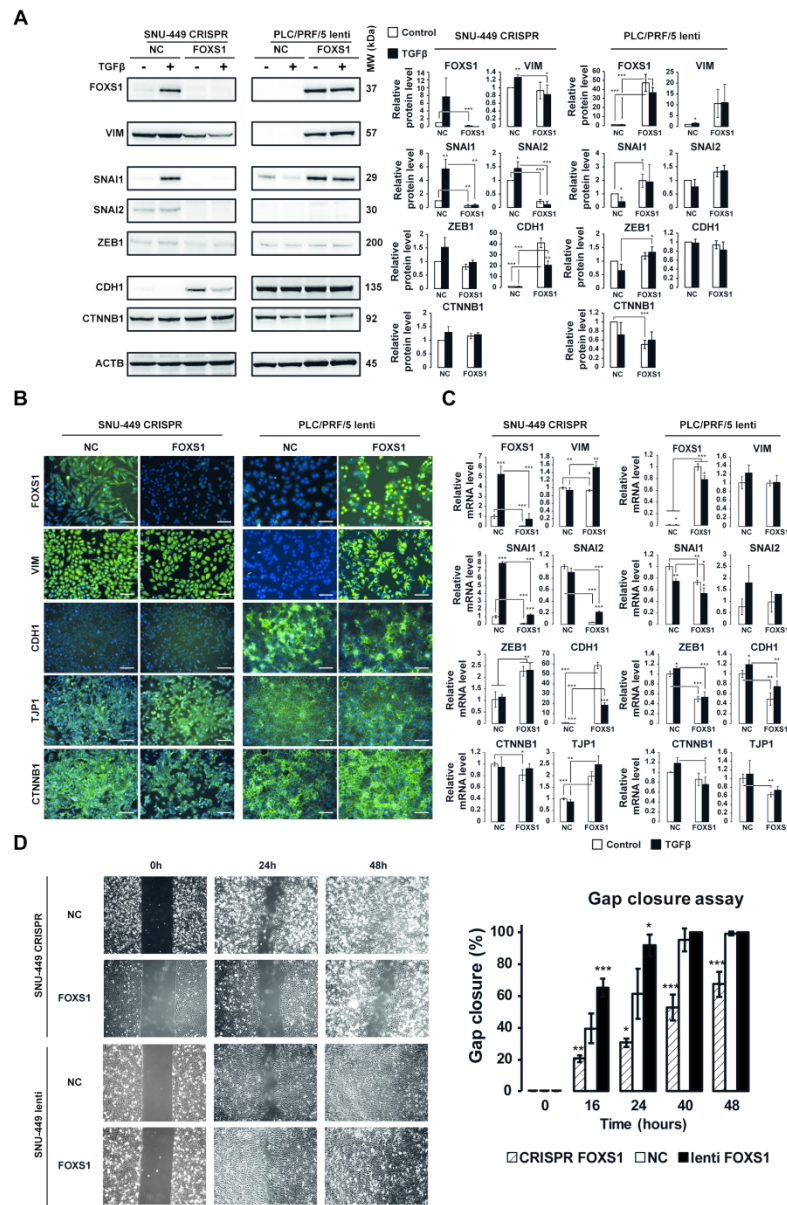


Fig. 4. FOXS1 induces EMT markers and promotes cell migration. (A) Western blot analysis of EMT markers after stable FOXS1 loss and gain of function in clonal SNU-449 and PLC/PRF/5, respectively (1ng/mL TGFβ, 16hrs). (B) Evaluation of EMT markers by immunofluorescence (scale bar: 100μm). (C) RT-qPCR of EMT markers (1ng/mL TGFβ, 16hrs). (D) Migration assay in cells after mitomycin C treatment (10μg/mL). Statistical analysis was performed by a t-test (* P<0.05; ** P<0.01; *** P<0.001; n≥3 replicates).

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

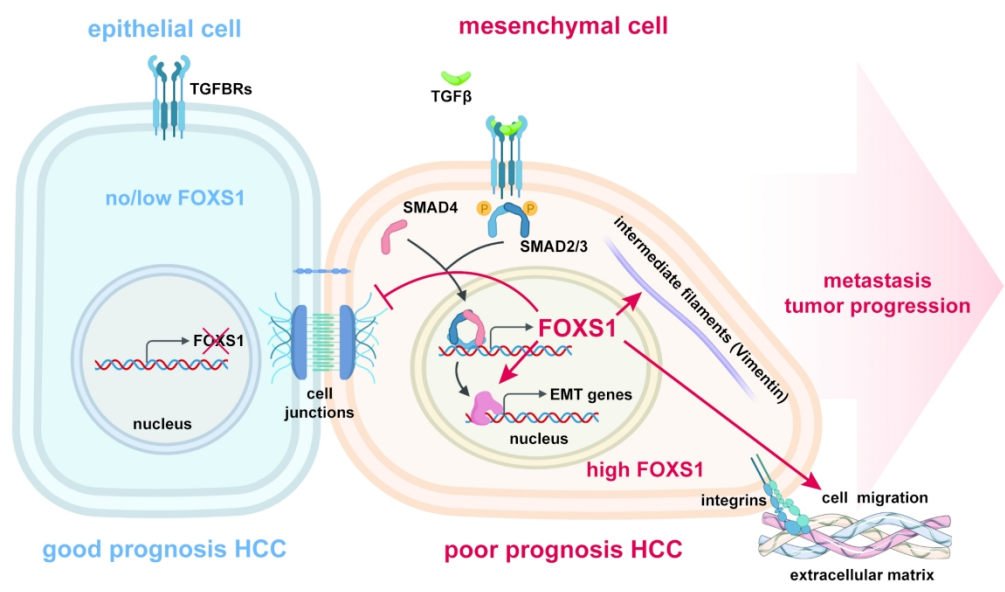


Fig. 6. Functional role and clinical relevance of FOXS1 in HCC.

1
2
3 **TGF β -induced FOXS1 controls epithelial-mesenchymal transition and predicts a**
4
5 **poor prognosis in liver cancer**
6
7
8
9

10 Kevin Bévant¹, Matthis Desoteux¹, Gaëlle Angenard¹, Stefano Caruso²,
11
12 Corentin Louis¹, Panagiotis Papoutsoglou¹, Laurent Sulpice¹, David Gilot³,
13
14
15 Jessica Zucman-Rossi^{2,4} and Cédric Coulouarn^{1*}
16
17
18
19
20

21 **Supplementary data**
22
23
24
25
26
27

28 **SUPPLEMENTARY FIGURE LEGENDS**
29
30

31 **Supplementary Fig. S1. FOXS1 is an early and direct TGF β target.** (A) Q-RT-PCR analysis
32 of *FOXS1* expression in SNU-449 cell line after 2hrs, 8hrs and 24hrs of culture in absence (white
33 bars) or presence (black bars) of 1ng/mL TGF β (** P<0.01; n=3 replicates). (B) Q-RT-PCR
34 analysis of *ITGB6*, *SMAD7*, *COL1A1*, *NANOG* and *FOXS1* expression in SNU-449 cell line after
35 16hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGF β and exposed
36 (CHX) or not (UT) to 200 μ g/mL cycloheximide translation inhibitor (** P<0.01; *** P<0.001; ns
37 P>0.05 ; n=3 replicates). (C) Chromatin immunoprecipitation PCR (chIP-PCR) experiment using
38 primers targeting *SMAD7* and *FOXS1* promoters in SNU-449 cell line after 16hrs of culture in
39 absence (white bars) or presence (black bars) of 1ng/mL TGF β . UNTR5 is an untranscribed
40 region of the genome used as a negative control.
41

42 **Supplementary Fig. S2. Gene expression analysis in human tissue slices.** Expression of
43 *FOXS1* (A), *SMAD7* (B) and *SERPINE1* (C) by quantitative RT-PCR in human surrounding non-
44 tumor (left 2 panels ; n=6 patients) or tumor (right 2 panels ; n=7 patients) tissue slices from
45 patients undergoing surgical resection. Tissue slices were incubated 48hrs in absence or
46 presence of 1ng/mL TGF β or 10 μ M LY2157299 TGF β inhibitor. Statistical analysis was
47 performed by a non-parametric Wilcoxon matched-pairs signed rank test.
48

49 **Supplementary Fig. S3. FOXS1 genomic organization and putative regulation.** (A) Genomic
50 organization of *FOXS1* locus at chrs20:31,844,000-31,847,500. Highlighted are *FOXS1*
51 transcript, CpG island, regulatory elements from Open Regulatory Annotation (OREgAnno),
52 including *FOXA1*, Dnase I hypersensitivity peak clusters from ENCODE, and sequence
53 conservation between 100 vertebrate species. A *SMAD3* binding site was predicted by Ensembl
54 at position 31,846,071-31,846,080; no binding site for other EMT-associated transcription factors
55 was predicted. (B) CpG analysis in *FOXS1* genomic sequence starting 1kb upstream of ATG
56
57
58
59
60

(putative proximal promoter) and ending 300bp downstream of TAA stop codon. A CpG island in FOXS1 coding sequence was predicted by MethPrimer (www.urogene.org/cgi-bin/methprimer/methprimer.cgi). (C) Methylation status of FOXS1 locus in surrounding non-tumor tissue (n=30, upper part) or HCC (n=30, lower part) using Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer (TCGA dataset). (D) Q-RT-PCR analysis of FOXA1 expression in SNU-449 and PLC/PRF/5 HCC cell lines after 16hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ (** P<0.01, n=3 replicates). (E) Q-RT-PCR analysis of FOXA1 and FOXS1 expression in PLC/PRF/5 HCC cell line transfected with siFOXA1 siRNA (FOXA1, 50nM each) or a negative control siRNA (NC, 50nM) after 72hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ.

Supplementary Fig. S4. Validation of predicted FOXS1 siRNA off-targets. (A) BLAST analysis of FOXS1 mRNA sequence (Query) and alignments with other FOX family members, highlighting a high degree of similarity in the forkhead box region where is located FOXS1 siRNA#1 (upper part). Expression of FOX family members after 72hrs transfection of SNU-449 cells with siRNA#1 and #2 targeting FOXS1 (lower part). (B) Expression of putative siRNA#1 off-targets predicted by sequence homology using BLAST, after siRNA transfection. (C) Expression of putative siRNA#2 off-targets predicted by sequence homology using BLAST, after siRNA transfection.

Supplementary Fig. S5. Validation of FOXS1 stable gain and loss of function. (A) Validation of FOXS1 overexpression (lentiviral construct) and knock-out (CRISPR-Cas9) by immunofluorescence using an anti-FOXS1 antibody. Nuclear location of FOXS1 was highlighted. (B) Mutation of FOXS1 locus was validated by sanger sequencing in wild-type (WT) SNU-449 cells, as well as in cells infected with a control sgRNA or with a sgRNA targeting FOXS1 locus.

Supplementary Fig. S6. Overexpression of endogenous FOXS1 by CRISPR-Cas9 Synergistic Activation Mediator. (A) Western Blot analysis of FOXS1 and SNAIL in SNU-449 cells specifically engineered to express the CRISPR-Cas9 Synergistic Activation Mediator system and infected with a vector expressing a control single guide RNA (sgRNA) or a sgRNA targeting FOXS1 promoter. Cells were cultured in presence or absence of 1ng/mL TGFβ. (B) Western Blot quantification was performed using ImageJ software (n=1).

Supplementary Fig. S7. FOXS1 expression in independent EMT-MET *in vitro* models. (A) Expression of VIM, CDH1 and FOXS1 in independent gene expression datasets. *Upper panels:* time-course analysis of TGFβ-induced EMT in NMuMG/E9 murine mammary gland epithelial cells. Cells were treated with TGFβ for 0, 2h, 6h, 12h, 24h, 36h, 48h, 60h, 72h, 96h, 168h and 240h (RNA-seq dataset: GSE112797). *Lower panels:* time-course analysis of MET over a period of 10 days on withdrawal of TGFβ in NMuMG/E9 cells (RNA-seq dataset: GSE117474). Statistical analysis was performed by a t-test (* P<0.05). (B) Expression of VIM, CDH1 and FOXS1 in an independent gene expression datasets of EMT (GSE17708). A549 epithelial cells treated for up to 72 hours with TGFβ to induce EMT. Statistical analysis was performed by a t-test (* P<0.05).

Supplementary Fig. S8. FOXS1 expression in a panel of 34 liver cancer cell lines. (A) Expression of FOXS1 in 34 liver cancer cell lines (LCCL). Cell lines are regrouped in three groups according to FOXS1 expression (low, medium and high). CL1, CL2 and CL3 were as defined in Caruso et al., *Gastroenterology*. 2019. Notably, cell lines from CL3 group exhibit a mesenchymal-like phenotype. (B) Expression of FOXS1, FOXA1, CDH1, TGFβ1, VIM and SERPINE1 in cell lines expressing FOXS1 at low, medium and high level. (C) GSEA of specific signatures enriched in the gene expression profiles of high FOXS1 subgroup (left side in red) or

1
2
3 in the low FOXS1 subgroup (right side in blue). Data were extracted from Caruso et al.,
4 *Gastroenterology*. 2019 Sep;157(3):760-776.
5

6 **Supplementary Fig. S9. FOXS1 expression in 37 human cancers from the TCGA**
7 **consortium.** (A) This figure shows mRNASeq expression profiles for each TCGA disease chart.
8 ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive
9 carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma;
10 **CHOL, Cholangiocarcinoma;** COAD, Colon adenocarcinoma; COADREAD, Colorectal
11 adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA,
12 Esophageal carcinoma; FPPP, FFPE Pilot Phase II; GBM, Glioblastoma multiforme; GBMLGG,
13 Glioma; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIPAN,
14 Pan-kidney (KICH+KIRC+KIRP); KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal
15 papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain Lower Grade Glioma;
16 **LIHC, Liver hepatocellular carcinoma;** LUAD, Lung adenocarcinoma; LUSC, Lung squamous
17 cell carcinoma; MESO, Mesothelioma; OV, Ovarian serous cystadenocarcinoma; PAAD,
18 Pancreatic adenocarcinoma; PCPG, Pheochromocytoma and Paraganglioma; PRAD, Prostate
19 adenocarcinoma; READ, Rectum adenocarcinoma; SARC, Sarcoma; SKCM, Skin Cutaneous
20 Melanoma; STAD, Stomach adenocarcinoma; STES, Stomach and Esophageal carcinoma;
21 TGCT, Testicular Germ Cell Tumors; THCA, Thyroid carcinoma; THYM, Thymoma; UCEC,
22 Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma; UVM, Uveal Melanoma.
23 (B) Expression of *FOXS1* in tumor (red box) and non-tumor (grey box) tissues from LIHC and
24 CHOL cohorts. (C) Expression of *FOXS1* according to tumor stage (I to IV). (D) Correlation
25 between *FOXS1* and *VIM* expression in HCC (n=369).
26
27

28 **Supplementary Fig. S10. FOXS1 expression in human HCC (n=193).** (A) Expression of
29 *FOXS1* in 193 HCC tissues as compared to 41 surrounding non-tumor tissues. *FOXS1*
30 expression in HCC was evaluated as a function of the histology of the surrounding non-tumor
31 liver (Fibrosis score F0-F4, and cirrhosis). (B) Correlation between *FOXS1* and *VIM* expression
32 in HCC (n=193). (C) Expression of *FOXS1* according to the G1-G6 molecular classification, as
33 defined in Boyault et al., *Hepatology*, 2007. (D) Expression of *FOXS1* according to mutations of
34 *CTNNB1*, encoding β -catenin, a gene frequently mutated in HCC. RNA-seq data were extracted
35 from Nault JC, et al. *Gastroenterology*. 2013 and Schullze K, et al. *Nat. Genet.*, 2015.
36
37
38
39
40

41 SUPPLEMENTARY TABLES

42
43
44 **Supplementary Table S1.** List of primers (A) and antibodies (B) used in the study.
45

46 **Supplementary Table S2.** List of genes differentially expressed in SNU-449 after 72hrs of
47 TGF β treatment.
48

49 **Supplementary Table S3.** Characteristics of human HCC cell lines.
50

51 **Supplementary Table S4.** List of genes differentially expressed between FOXS1-positive and -
52 negative HCC cell lines.
53

54 **Supplementary Table S5.** Gene expression signature of endogenous FOXS1 silencing in SNU-
55 449 cells.
56
57
58
59
60

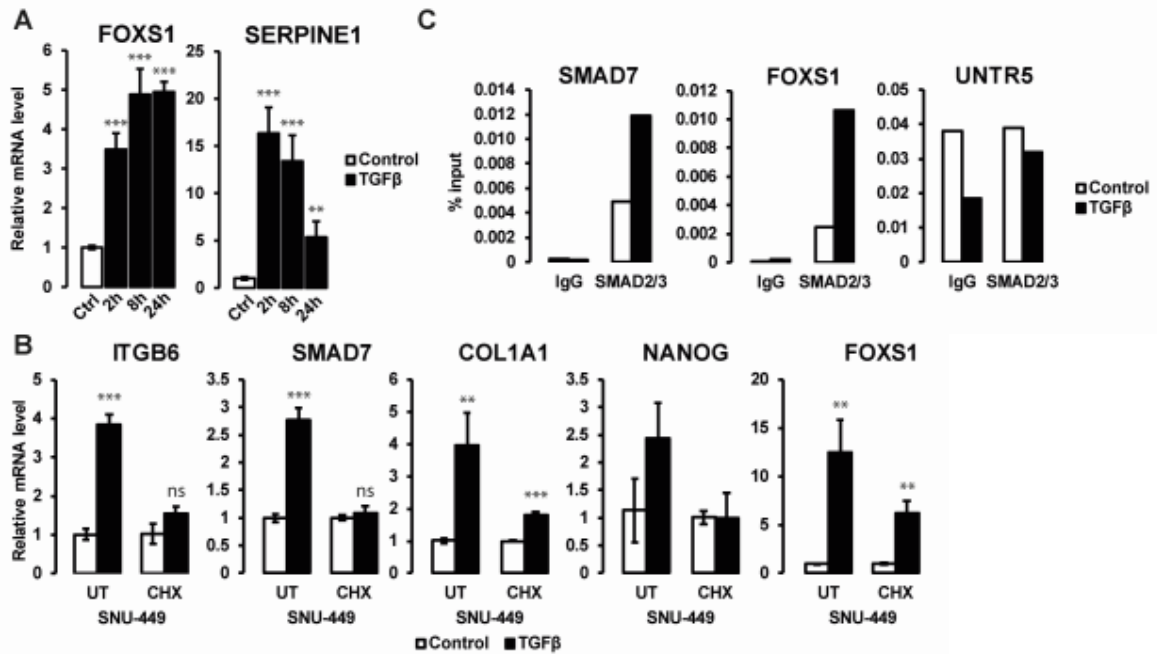


Fig. S1. FOXS1 is an early and direct TGFβ target.

(A) Q-RT-PCR analysis of *FOXS1* expression in SNU-449 cell line after 2hrs, 8hrs and 24hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ (** P<0.01; n=3 replicates).

(B) Q-RT-PCR analysis of *ITGB6*, *SMAD7*, *COL1A1*, *NANOG* and *FOXS1* expression in SNU-449 cell line after 16hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ and exposed (CHX) or not (UT) to 200μg/mL cycloheximide translation inhibitor (** P<0.01; *** P<0.001; ns P>0.05 ; n=3 replicates).

(C) Chromatin immunoprecipitation PCR (ChIP-PCR) experiment using primers targeting *SMAD7* and *FOXS1* promoters in SNU-449 cell line after 16hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ . UNTR5 is an untranscribed region of the genome used as a negative control.

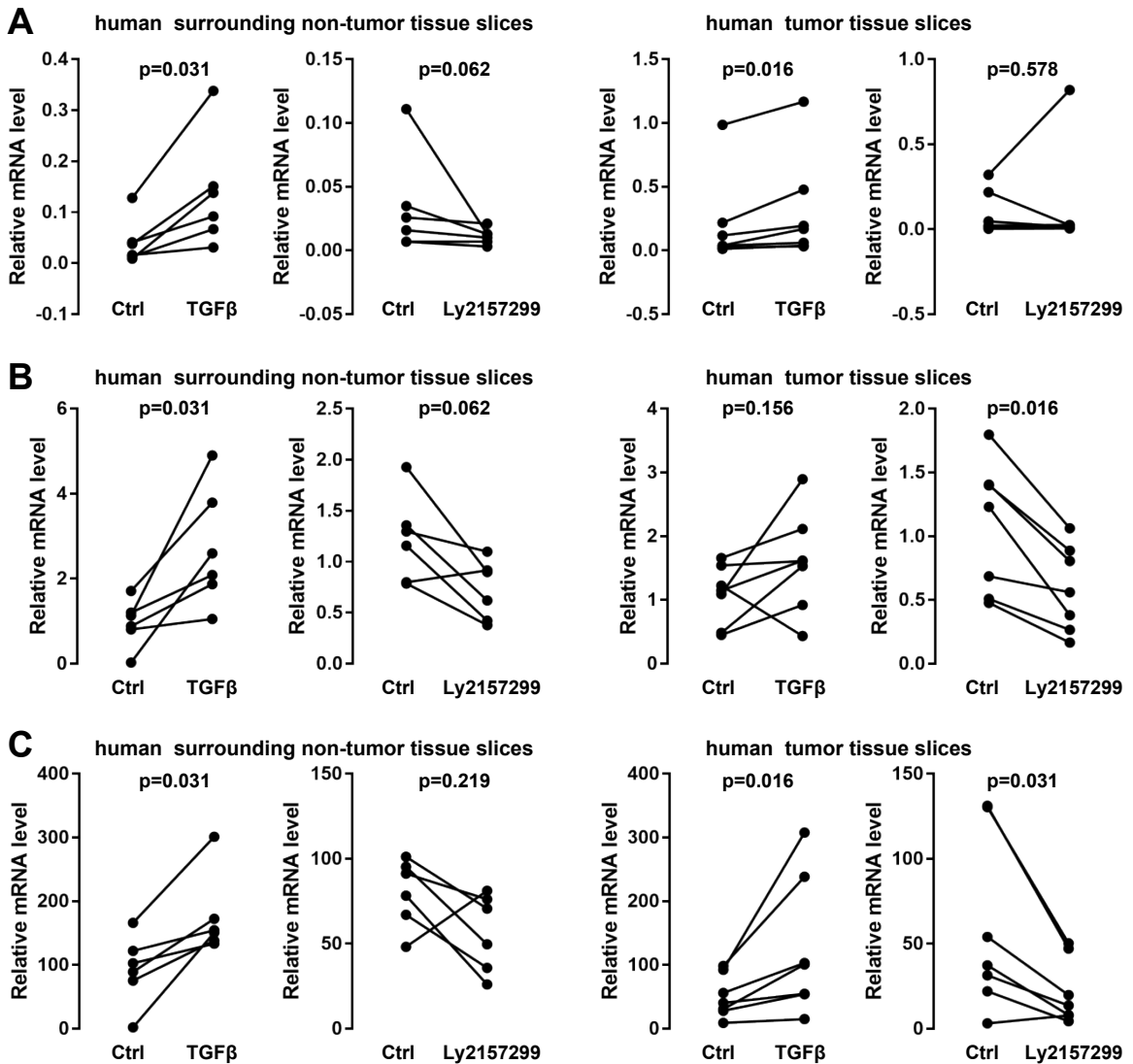


Fig. S2. Gene expression analysis in human tissue slices.

Expression of *FOXS1* (A), *SMAD7* (B) and *SERPINE1* (C) by quantitative RT-PCR in human surrounding non-tumor (left 2 panels ; n=6 patients) or tumor (right 2 panels ; n=7 patients) tissue slices from patients undergoing surgical resection. Tissue slices were incubated 48hrs in absence or presence of 1ng/mL TGF β or 10 μ M LY2157299 TGF β inhibitor. Statistical analysis was performed by a non-parametric Wilcoxon matched-pairs signed rank test.

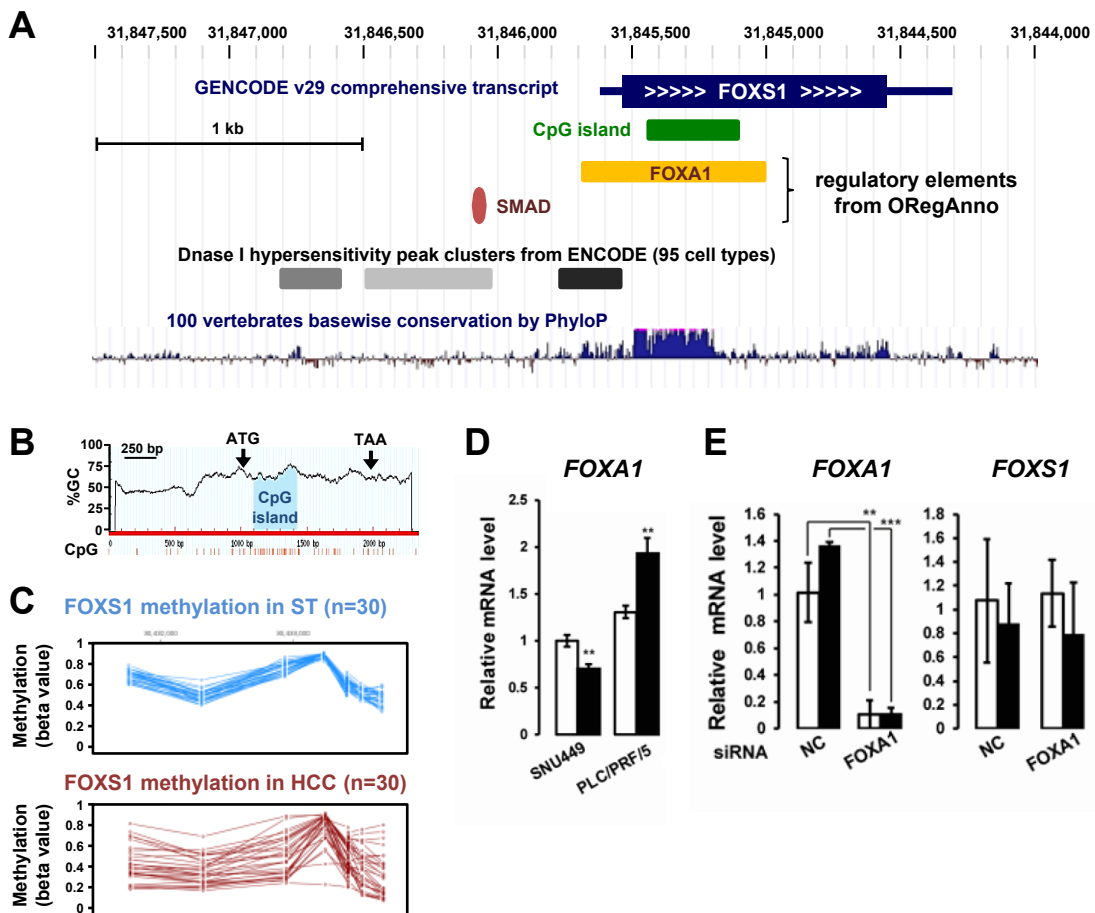


Fig. S3. FOXS1 genomic organization and putative regulation.

(A) Genomic organization of *FOXS1* locus at chr20:31,844,000-31,847,500. Highlighted are *FOXS1* transcript, CpG island, regulatory elements from Open Regulatory Annotation (ORegAnno), including *FOXA1*, Dnase I hypersensitivity peak clusters from ENCODE, and sequence conservation between 100 vertebrate species. A SMAD3 binding site was predicted by Ensembl at position 31,846,071-31,846,080; no binding site for other EMT-associated transcription factors was predicted.

(B) CpG analysis in *FOXS1* genomic sequence starting 1kb upstream of ATG (putative proximal promoter) and ending 300bp downstream of TAA stop codon. A CpG island in *FOXS1* coding sequence was predicted by MethPrimer (www.urogene.org/cgi-bin/methprimer/methprimer.cgi).

(C) Methylation status of *FOXS1* locus in surrounding non-tumor tissue (n=30, upper part) or HCC (n=30, lower part) using Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer (TCGA dataset).

(D) Q-RT-PCR analysis of *FOXA1* expression in SNU-449 and PLC/PRF/5 HCC cell lines after 16hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ (** P<0.01, n=3 replicates).

(E) Q-RT-PCR analysis of *FOXA1* and *FOXS1* expression in PLC/PRF/5 HCC cell line transfected with siFOXA1 siRNA (FOXA1, 50nM each) or a negative control siRNA (NC, 50nM) after 72hrs of culture in absence (white bars) or presence (black bars) of 1ng/mL TGFβ.

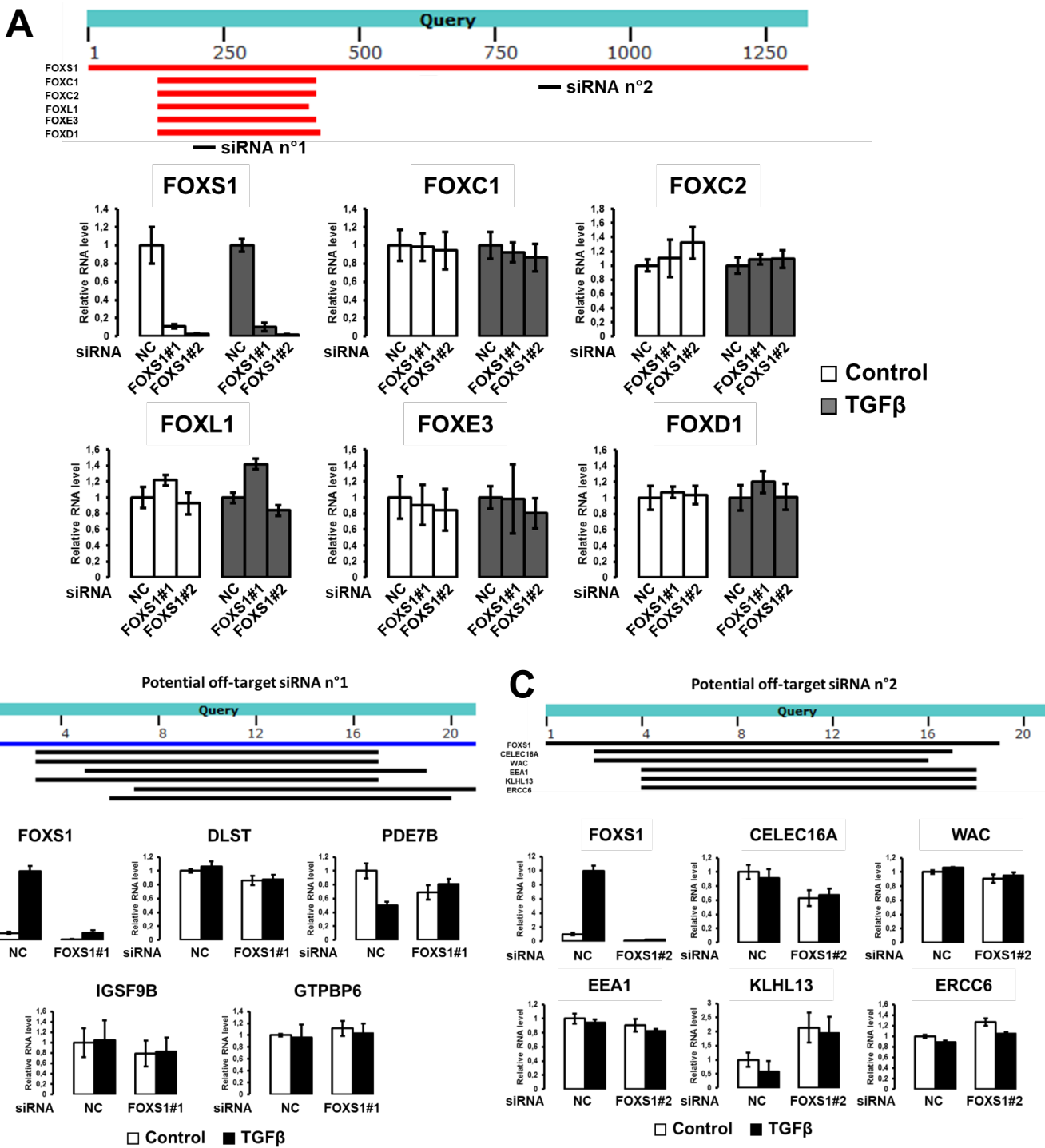


Fig. S4. Validation of predicted FOXS1 siRNA off-targets.

(A) BLAST analysis of FOXS1 mRNA sequence (Query) and alignments with other FOX family members, highlighting a high degree of similarity in the forkhead box region where is located FOXS1 siRNA#1 (upper part). Expression of FOX family members after 72hrs transfection of SNU-449 cells with siRNA#1 and #2 targeting FOXS1 (lower part).

(B) Expression of putative siRNA#1 off-targets predicted by sequence homology using BLAST, after siRNA transfection.

(C) Expression of putative siRNA#2 off-targets predicted by sequence homology using BLAST, after siRNA transfection.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

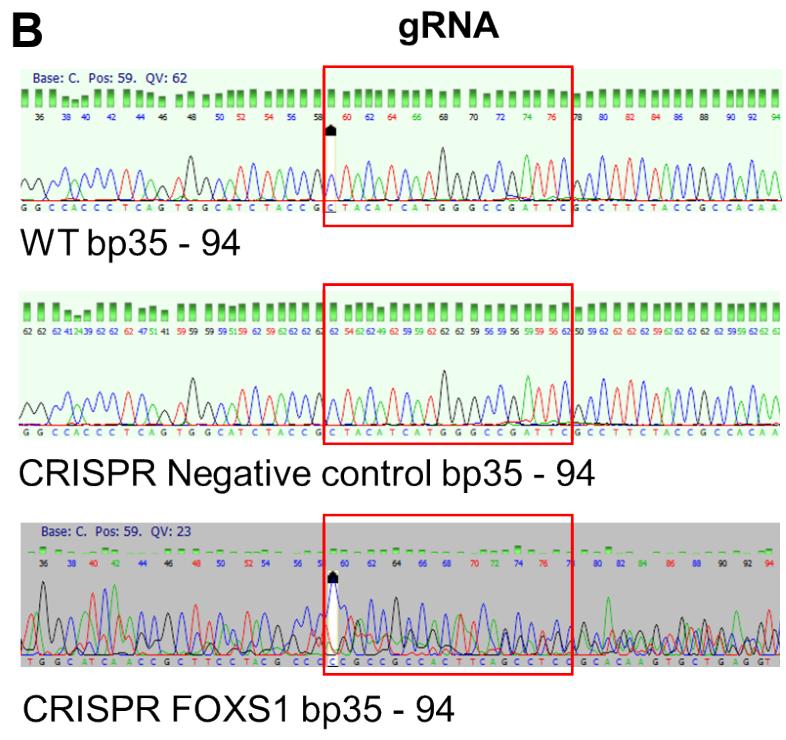
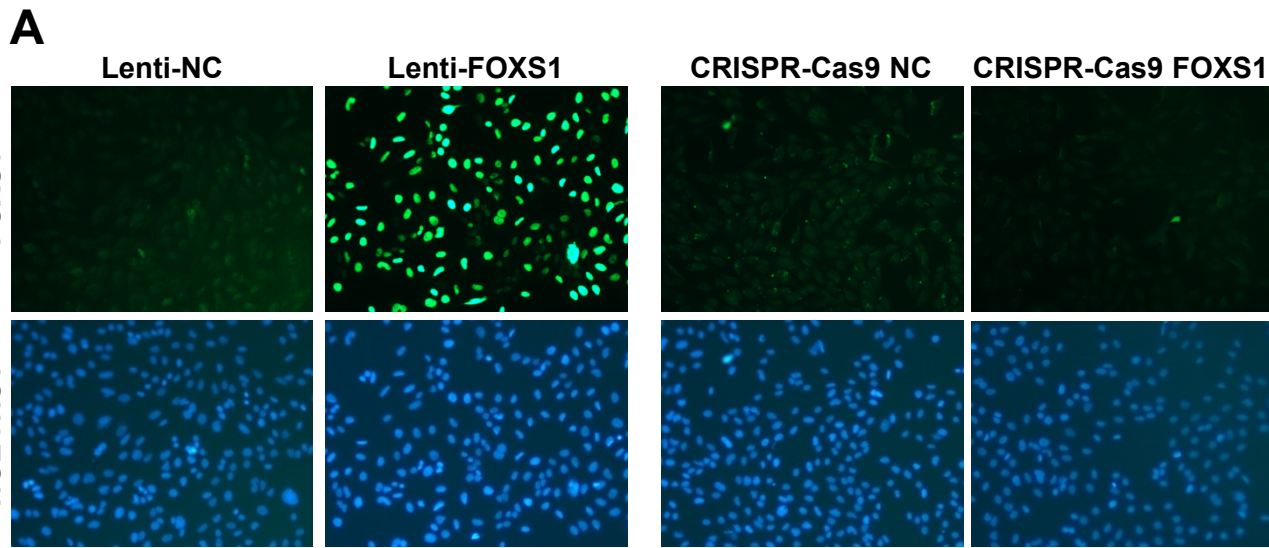


Fig. S5. Validation of FOXS1 stable gain and loss of function.

(A) Validation of FOXS1 overexpression (lentiviral construct) and knock-out (CRISPR-Cas9) by immunofluorescence using an anti-FOXS1 antibody. Nuclear location of FOXS1 was highlighted.

(B) Mutation of *FOXS1* locus was validated by sanger sequencing in wild-type (WT) SNU-449 cells, as well as in cells infected with a control sgRNA or with a sgRNA targeting *FOXS1* locus.

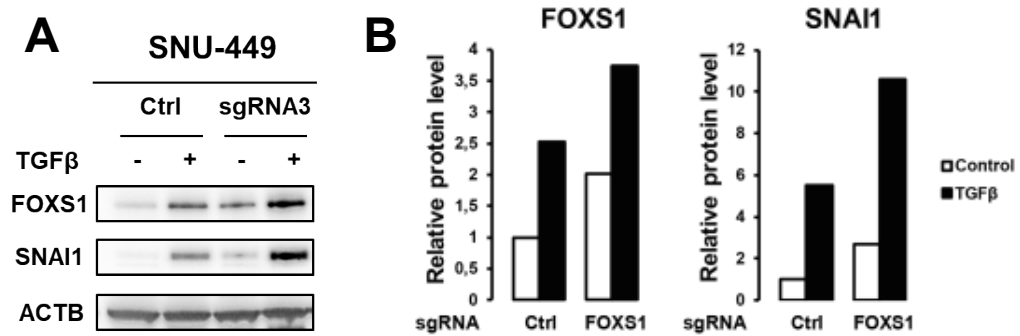


Fig. S6. Overexpression of endogenous *FOXS1* by CRISPR-Cas9 Synergistic Activation Mediator.

(A) Western Blot analysis of FOXS1 and SNAI1 in SNU-449 cells specifically engineered to express the CRISPR-Cas9 Synergistic Activation Mediator system and infected with a vector expressing a control single guide RNA (sgRNA) or a sgRNA targeting *FOXS1* promoter. Cells were cultured in presence or absence of 1ng/mL TGFβ.

(B) Western Blot quantification was performed using ImageJ software (n=1).

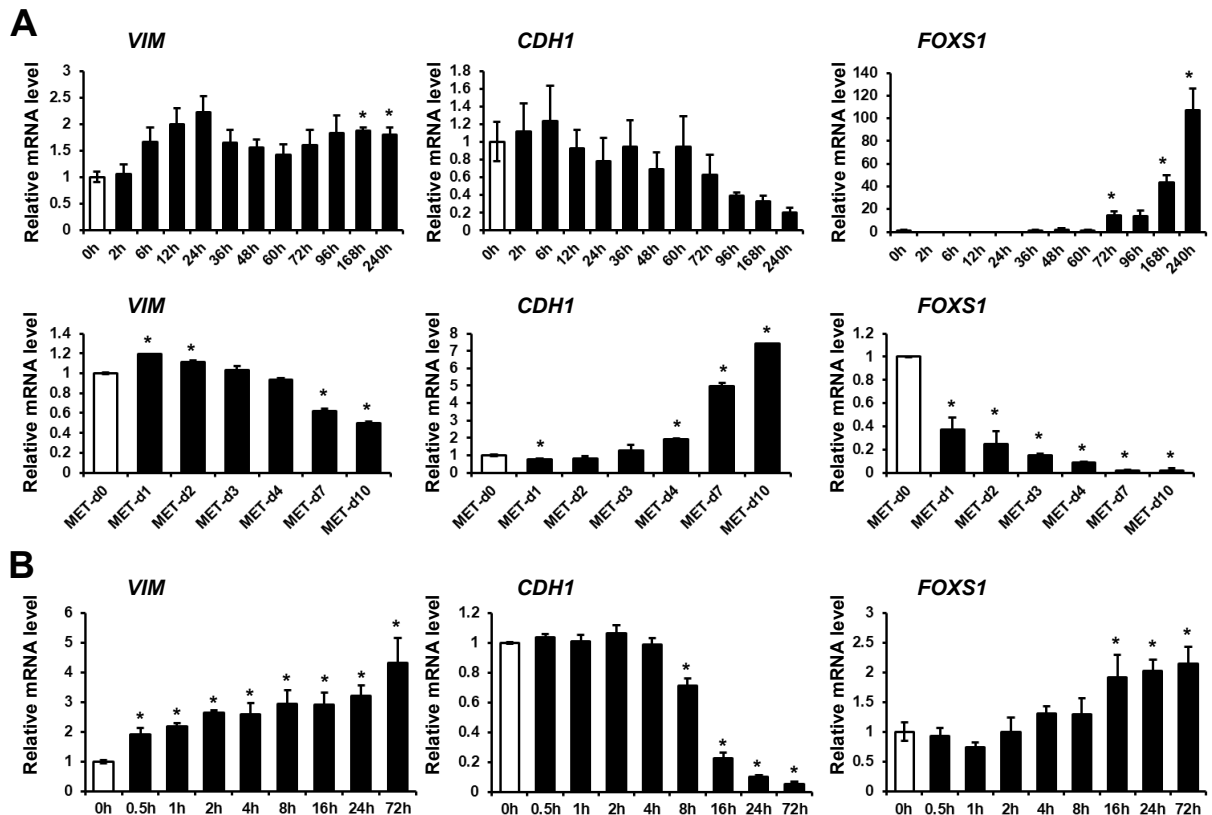


Fig. S7. FOXS1 expression in independent EMT-MET *in vitro* models.

(A) Expression of *VIM*, *CDH1* and *FOXS1* in independent gene expression datasets. *Upper panels*: time-course analysis of TGFβ-induced EMT in NMuMG/E9 murine mammary gland epithelial cells. Cells were treated with TGFβ for 0, 2h, 6h, 12h, 24h, 36h, 48h, 60h, 72h, 96h, 168h and 240h (RNA-seq dataset: GSE112797). *Lower panels*: time-course analysis of MET over a period of 10 days on withdrawal of TGFβ in NMuMG/E9 cells (RNA-seq dataset: GSE117474). Statistical analysis was performed by a t-test (* $P < 0.05$).

(B) Expression of *VIM*, *CDH1* and *FOXS1* in an independent gene expression datasets of EMT (GSE17708). A549 epithelial cells treated for up to 72 hours with TGFβ to induce EMT. Statistical analysis was performed by a t-test (* $P < 0.05$).

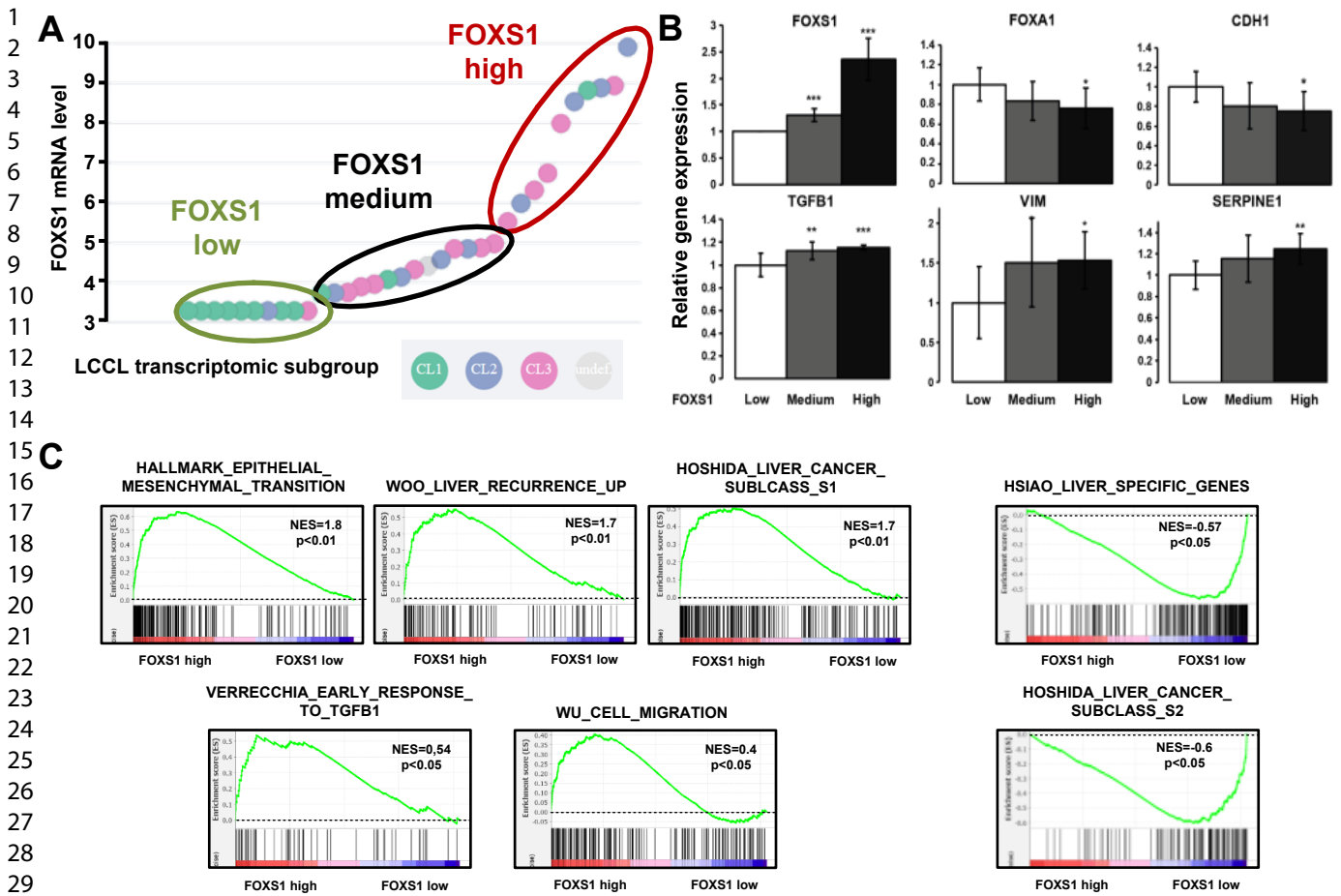


Fig. S8. FOXS1 expression in a panel of 34 liver cancer cell lines.

(A) Expression of *FOXS1* in 34 liver cancer cell lines (LCCL). Cell lines are regrouped in three groups according to *FOXS1* expression (low, medium and high). CL1, CL2 and CL3 were as defined in Caruso et al., *Gastroenterology*. 2019. Notably, cell lines from CL3 group exhibit a mesenchymal-like phenotype.

(B) Expression of *FOXS1*, *FOXA1*, *CDH1*, *TGFB1*, *VIM* and *SERPINE1* in cell lines expressing *FOXS1* at low, medium and high level.

(C) GSEA of specific signatures enriched in the gene expression profiles of high *FOXS1* subgroup (left side in red) or in the low *FOXS1* subgroup (right side in blue).

Data were extracted from Caruso et al., *Gastroenterology*. 2019 Sep;157(3):760-776.

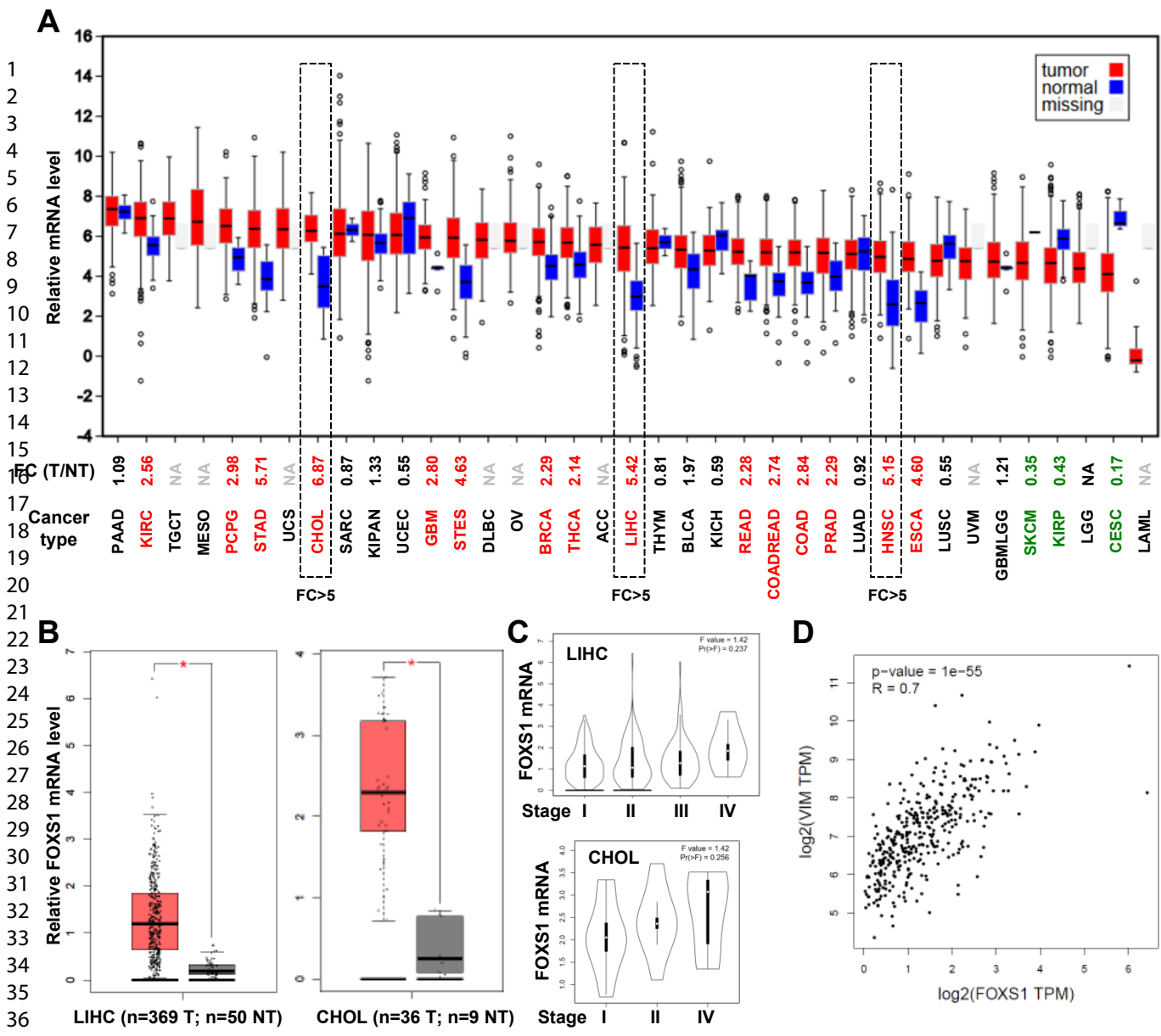


Fig. S9. FOXS1 expression in 37 human cancers from the TCGA consortium.

(A) This figure shows mRNASeq expression profiles for each TCGA disease chart. ACC, Adrenocortical carcinoma; BLCA, Bladder Urothelial Carcinoma; BRCA, Breast invasive carcinoma; CESC, Cervical squamous cell carcinoma and endocervical adenocarcinoma; **CHOL, Cholangiocarcinoma**; COAD, Colon adenocarcinoma; COADREAD, Colorectal adenocarcinoma; DLBC, Lymphoid Neoplasm Diffuse Large B-cell Lymphoma; ESCA, Esophageal carcinoma; FPPP, FFPE Pilot Phase II; GBM, Glioblastoma multiforme; GBMLGG, Glioma; HNSC, Head and Neck squamous cell carcinoma; KICH, Kidney Chromophobe; KIPAN, Pan-kidney (KICH+KIRC+KIRP); KIRC, Kidney renal clear cell carcinoma; KIRP, Kidney renal papillary cell carcinoma; LAML, Acute Myeloid Leukemia; LGG, Brain Lower Grade Glioma; **LIHC, Liver hepatocellular carcinoma**; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; MESO, Mesothelioma; OV, Ovarian serous cystadenocarcinoma; PAAD, Pancreatic adenocarcinoma; PCPG, Pheochromocytoma and Paraganglioma; PRAD, Prostate adenocarcinoma; READ, Rectum adenocarcinoma; SARC, Sarcoma; SKCM, Skin Cutaneous Melanoma; STAD, Stomach adenocarcinoma; STES, Stomach and Esophageal carcinoma; TGCT, Testicular Germ Cell Tumors; THCA, Thyroid carcinoma; THYM, Thymoma; UCEC, Uterine Corpus Endometrial Carcinoma; UCS, Uterine Carcinosarcoma; UVM, Uveal Melanoma.

(B) Expression of *FOXS1* in tumor (red box) and non-tumor (grey box) tissues from LIHC and CHOL cohorts.

(C) Expression of *FOXS1* according to tumor stage (I to IV).

(D) Correlation between *FOXS1* and *VIM* expression in HCC (n=369).

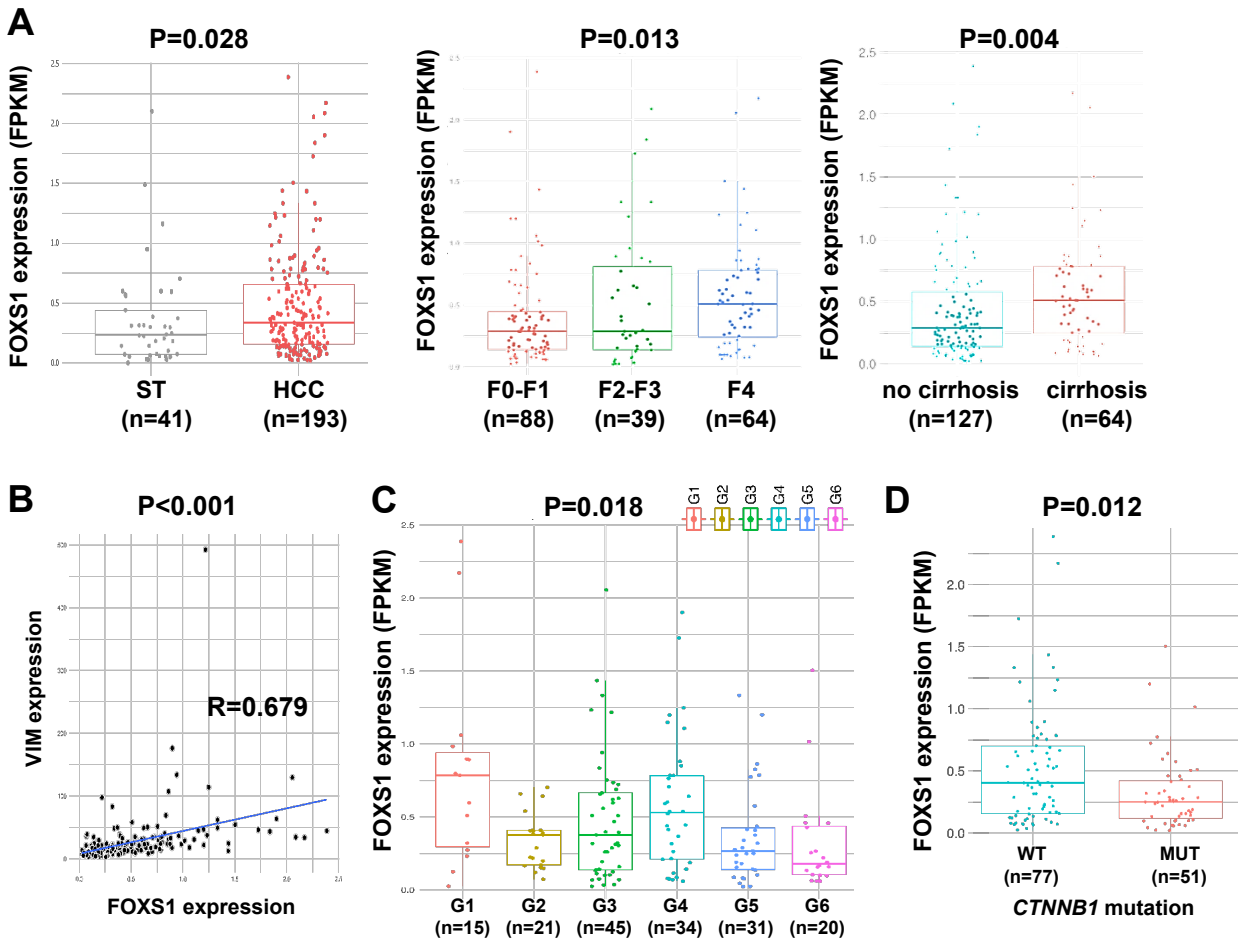


Fig. S10. FOXS1 expression in human HCC (n=193).

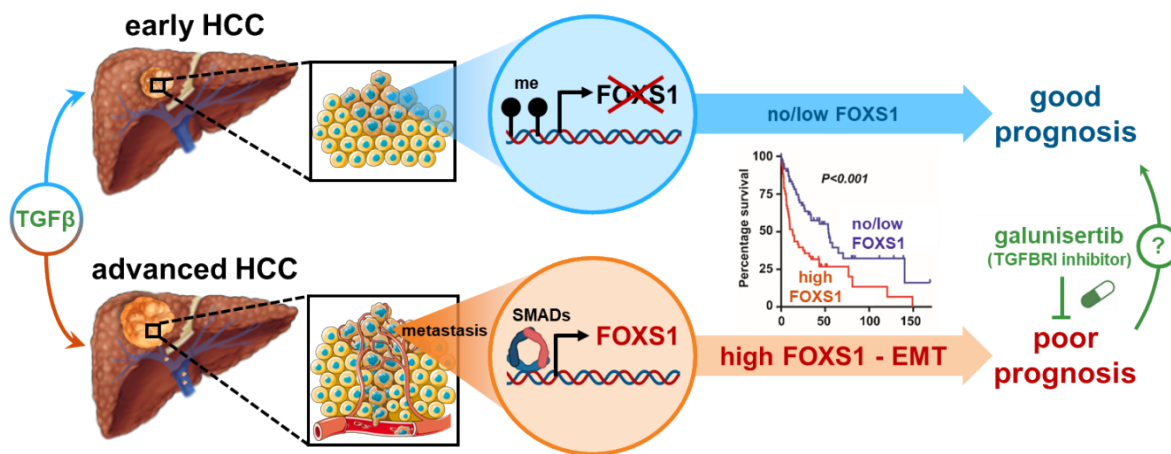
(A) Expression of *FOXS1* in 193 HCC tissues as compared to 41 surrounding non-tumor tissues. *FOXS1* expression in HCC was evaluated as a function of the histology of the surrounding non-tumor liver (Fibrosis score F0-F4, and cirrhosis).

(B) Correlation between *FOXS1* and *VIM* expression in HCC (n=193).

(C) Expression of *FOXS1* according to the G1-G6 molecular classification, as defined in Boyault et al., *Hepatology*, 2007.

(D) Expression of *FOXS1* according to mutations of *CTNNB1*, encoding β -catenin, a gene frequently mutated in HCC.

RNA-seq data were extracted from Nault JC, et al. *Gastroenterology*. 2013 and Schullze K, et al. *Nat. Genet.*, 2015.



Graphical abstract

The study identifies FOXS1 as a novel canonical TGFβ target gene controlling epithelial-mesenchymal transition (EMT). FOXS1 expression level and activity predict a poor prognosis in several cancers, including HCC. FOXS1 may constitute a clinically relevant biomarker of tumors in which the pro-metastatic arm of TGFβ is active and thus, of patients who may benefit from personalized targeted therapies using TGFβ inhibitors.

Supplementary Table S1. List of primers (A) and antibodies (B) used in the study.

(A) List of primers used in the study.

Gene name	Forward	Reverse
AXIN2	CTCCTTATCGTGTGGGCAGT	CCAACTCCAGCTTCAGCTTT
CDH1	CGAGAGCTACACGTTACCGG	GGGTGTCGAGGGAAAAATAGG
CLDN1	GCCCTGCCCCAGTGGAGGAT	CGGGTTGCTTGCAATGTGCTGCT
CTCF	ATGTGCGATTACGCCAGTGTA	TGAAACGGACGCTCTCCAGTA
CTNNB1	AAAATGGCAGTGCCTTTAG	TTTGAAGGCAGTCTGTGCGTA
EGR1	CCACGCCGAACACTGACATT	GAGGGGTTAGCGAAGGCTG
FOXA1	CCCCACTACTCCTTCAACCA	GTGTTTAGGACGGGTCTGGA
FOXS1	CCGTCTTGTCCCCGGAATC	CCAGGGTTCCTTGTGGTCAGT
KLF12	CCTTTCCATAGCCAGAGCAG	TTGCATCCCTCAAATCACA
MMP2	TTCCGCTTCCAGGGCACA	CACCTTCTGAGTTCCACCAA
MMP9	CCTGGAGACCTGAGAACCAATCT	CCACCCGAGTGTAAACCATAGC
SERPINE1	AAGCCTAATCAGCCCACCAT	TAGCATTTGACACCACCCT
SMAD2	AGACACCAGTTTTGCCTCCA	GAGGTGGCGTTTTCTGGAATA
SMAD3	CTAGGGCTGCTCTCCAATGT	AAGACCTCCCCTCCGATGTA
SMAD4	GCTGCTGGAATTGGTGTGATG	AGGTGTTTCTTTGATGCTCTGTCT
SMAD7	CCAACTGCAGACTGTCCAGA	GGGCTCCTGGACACAGTAGA
SNAI1	CACTATGCCGCGCTCTTTC	GGCTGCTGGAAGGTAAACTC
SNAI2	TCGGACCCACACATTACCTT	TTGGAGCAGTTTTTGCCTG
TBP	GAGCTGTGATGTGAAGTTTCC	TCTGGGTTTGATCATTCTGTAG
TJP1	GGAGAGGTGTTCCGTGTTGT	GAGCGGACAAATCCTCTCTG
TP53	CGCTTCGAGATGTTCCGAGA	CAGTGCAGGCCAACTTGTTT
TWIST1	CCTCCAAGTCTGCAGCTCTC	AAAAAGAAAGCGCCCAACGG
TWIST2	CGGGTCATGGCTAACGTG	CAGCTTGCCATCCTGGAGTC
VIMENTIN	GACGCCATCAACCCGAGTT	CTTTGTGCTTGGTTAGCTGGT
ZEB1	CCCACACGACCACAGATACG	TGAGGAGAAGTGGTTGCCTG
ZEB2	CCCAACTCCGATGAACTGCT	GGAGGGTTACTGTTGGGAGC

(B) List of antibodies used in the study.

Antibodies	Host	Company
Actin (D68A)	rabbit	Cell Signaling
Anti-HA Tag	mouse	Genecopoeia
FOXS1	rabbit	Novus Biological
β -catenin	rabbit	Cell signaling
E-cadherin (24E10)	rabbit	Cell signaling
Slug (C19G7)	rabbit	Cell signaling
SNAI1 (C15D3)	rabbit	Cell signaling
TWIST1	rabbit	Cell signaling
Vimentine (D21H3)	rabbit	Cell signaling
ZO-1 (D7D12)	rabbit	Cell signaling
Anti-mouse IgG/HRP	goat	DAKO
Anti-rabbit IgG/HRP	goat	Cell Signaling
Dylight 488 anti-rabbit IgG (H+L)	goat	Eurobio

Supplementary Table S2. List of genes differentially expressed in SNU-449 after 72hrs of TGF β treatment.

Gene symbol	Gene description [GenBank Accession number]	Accession nb.	P-value	Fold change (TGF β /Ctrl)
DA292495	DA292495 BRHIP2 Homo sapiens cDNA clone BRHIP2004426 5', mRNA sequence [DA292495]	DA292495	0.000873	52.63
LDLRAD4	Homo sapiens low density lipoprotein receptor class A domain containing 4 (LDLRAD4), transcript variant a2, mRNA [NM_181482]	NM_181482	0.0000769	27.03
LRRC32	Homo sapiens leucine rich repeat containing 32 (LRRC32), transcript variant 1, mRNA [NM_005512]	NM_005512	0.0007413	15.38
NKAIN4	Homo sapiens Na ⁺ /K ⁺ transporting ATPase interacting 4 (NKAIN4), mRNA [NM_152864]	NM_152864	0.0000632	15.15
LEFTY1	Homo sapiens left-right determination factor 1 (LEFTY1), mRNA [NM_020997]	NM_020997	0.000702	13.16
CRLF1	Homo sapiens cytokine receptor-like factor 1 (CRLF1), mRNA [NM_004750]	NM_004750	0.0002942	11.11
MAF	Homo sapiens v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog (MAF), transcript variant 2, mRNA [NM_001031804]	NM_001031804	0.0006723	10.75
NKAIN4	Homo sapiens Na ⁺ /K ⁺ transporting ATPase interacting 4 (NKAIN4), mRNA [NM_152864]	NM_152864	0.0002019	10.53
FOXS1	Homo sapiens forkhead box S1 (FOXS1), mRNA [NM_004118]	NM_004118	0.0001324	10.20
MDFI	Homo sapiens MyoD family inhibitor (MDFI), transcript variant 1, mRNA [NM_001300804]	NM_001300804	0.0001139	7.69
P2RY2	Homo sapiens purinergic receptor P2Y, G-protein coupled, 2 (P2RY2), transcript variant 1, mRNA [NM_176072]	NM_176072	0.0003603	7.69
COL1A1	Homo sapiens collagen, type I, alpha 1 (COL1A1), mRNA [NM_000088]	NM_000088	0.0001113	6.67
CGB	Homo sapiens chorionic gonadotropin, beta polypeptide (CGB), mRNA [NM_000737]	NM_000737	0.0001154	5.56
EGR2	Homo sapiens early growth response 2 (EGR2), transcript variant 1, mRNA [NM_000399]	NM_000399	0.0001571	5.26
LMO1	Homo sapiens LIM domain only 1 (rhombotin 1) (LMO1), transcript variant 1, mRNA [NM_002315]	NM_002315	0.0002682	5.00
SUCNR1	Homo sapiens succinate receptor 1 (SUCNR1), mRNA [NM_033050]	NM_033050	0.0002878	4.76
SGCA	Homo sapiens sarcoglycan, alpha (50kDa dystrophin-associated glycoprotein) (SGCA), transcript variant 1, mRNA [NM_000023]	NM_000023	0.0000168	4.17
DACT1	Homo sapiens dishevelled-binding antagonist of beta-catenin 1 (DACT1), transcript variant 1, mRNA [NM_016651]	NM_016651	0.0000459	4.00
ZNF365	Homo sapiens zinc finger protein 365 (ZNF365), transcript variant A, mRNA [NM_014951]	NM_014951	0.0000489	4.00

LOC101927230	Homo sapiens uncharacterized LOC101927230 (LOC101927230), long non-coding RNA [NR_110806]		0.0002677	3.85
NXPH4	Homo sapiens neurexophilin 4 (NXPH4), mRNA [NM_007224]	NM_007224	0.0002031	3.85
LOC100507431	Homo sapiens uncharacterized LOC100507431 (LOC100507431), long non-coding RNA [NR_125383]	NR_125383	0.0000514	3.70
LOC101928173	PREDICTED: Homo sapiens uncharacterized LOC101928173 (LOC101928173), transcript variant X2, ncRNA [XR_245039]	XR_245039	0.0005653	3.57
PDGFA	Homo sapiens platelet-derived growth factor alpha polypeptide (PDGFA), transcript variant 2, mRNA [NM_033023]	NM_033023	0.0005938	3.57
ZNF365	Homo sapiens zinc finger protein 365 (ZNF365), transcript variant B, mRNA [NM_199450]	NM_199450	0.0007015	3.57
HTR1D	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 1D, G protein-coupled (HTR1D), mRNA [NM_000864]	NM_000864	0.0001306	3.33
ULK1	Homo sapiens unc-51 like autophagy activating kinase 1 (ULK1), mRNA [NM_003565]	NM_003565	0.0001096	3.33
DNM3	Homo sapiens dynamin 3 (DNM3), transcript variant 1, mRNA [NM_015569]	NM_015569	0.0000381	3.23
XLOC_I2_007543	BROAD Institute lincRNA (XLOC_I2_007543), lincRNA [TCONS_I2_00013972]		0.0007302	3.23
ALOX5AP	Homo sapiens arachidonate 5-lipoxygenase-activating protein (ALOX5AP), transcript variant 1, mRNA [NM_001629]	NM_001629	0.0008857	3.13
LTBP2	Homo sapiens latent transforming growth factor beta binding protein 2 (LTBP2), mRNA [NM_000428]	NM_000428	0.0001534	3.13
MAFB	Homo sapiens v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog B (MAFB), mRNA [NM_005461]	NM_005461	0.0004018	3.03
COL5A1	Homo sapiens collagen, type V, alpha 1 (COL5A1), transcript variant 1, mRNA [NM_000093]	NM_000093	0.0000113	2.94
DFNB31	Homo sapiens deafness, autosomal recessive 31 (DFNB31), transcript variant 1, mRNA [NM_015404]	NM_015404	0.0001901	2.94
SAMD11	Homo sapiens sterile alpha motif domain containing 11 (SAMD11), mRNA [NM_152486]	NM_152486	0.0001914	2.94
SIM2	Homo sapiens single-minded family bHLH transcription factor 2 (SIM2), transcript variant SIM2, mRNA [NM_005069]	NM_005069	0.0001315	2.94
ECE1	Homo sapiens endothelin converting enzyme 1 (ECE1), transcript variant 3, mRNA [NM_001113347]	NM_001113347	0.0003256	2.86
PMEPA1	Homo sapiens prostate transmembrane protein, androgen induced 1 (PMEPA1), transcript variant 1, mRNA [NM_020182]	NM_020182	0.000366	2.86
PMEPA1	Homo sapiens prostate transmembrane protein, androgen induced 1 (PMEPA1), transcript variant 1, mRNA [NM_020182]	NM_020182	0.0000063	2.86
PRSS8	Homo sapiens protease, serine, 8 (PRSS8), mRNA [NM_002773]	NM_002773	0.0002701	2.86

1					
2					
3					
4	ABLIM3	Homo sapiens actin binding LIM protein family, member 3 (ABLIM3), transcript variant 2, mRNA [NM_014945]	NM_014945	0.0000091	2.78
5	ADAM19	Homo sapiens ADAM metalloproteinase domain 19 (ADAM19), mRNA [NM_033274]	NM_033274	0.0001578	2.78
6					
7	PRDM1	Homo sapiens PR domain containing 1, with ZNF domain (PRDM1), transcript variant 1, mRNA [NM_001198]	NM_001198	0.0001414	2.78
8					
9	SIK1	Homo sapiens salt-inducible kinase 1 (SIK1), mRNA [NM_173354]	NM_173354	0.0008676	2.78
10					
11	SKIL	Homo sapiens SKI-like proto-oncogene (SKIL), transcript variant 1, mRNA [NM_005414]	NM_005414	0.0000581	2.78
12					
13					
14	COL5A1	Homo sapiens collagen, type V, alpha 1 (COL5A1), transcript variant 1, mRNA [NM_000093]	NM_000093	0.0001682	2.70
15					
16					
17	C5orf46	Homo sapiens chromosome 5 open reading frame 46 (C5orf46), mRNA [NM_206966]	NM_206966	0.0005732	2.63
18					
19	lnc-C11orf44-1	LNCipedia lincRNA (lnc-C11orf44-1), lincRNA [lnc-C11orf44-1:1]		0.0006299	2.63
20					
21					
22					
23	CASC15	Homo sapiens cancer susceptibility candidate 15 (non-protein coding) (CASC15), long non-coding RNA [NR_015410]	NR_015410	0.0000786	2.56
24					
25					
26	GRPR	Homo sapiens gastrin-releasing peptide receptor (GRPR), mRNA [NM_005314]	NM_005314	0.0005194	2.56
27					
28					
29	CASC15	Homo sapiens cancer susceptibility candidate 15 (non-protein coding) (CASC15), long non-coding RNA [NR_015410]	NR_015410	0.0009539	2.50
30					
31					
32	FSTL3	Homo sapiens follistatin-like 3 (secreted glycoprotein) (FSTL3), mRNA [NM_005860]	NM_005860	0.0006748	2.50
33					
34					
35	KLF12	Kruppel-like factor 12 [Source:HGNC Symbol;Acc:HGNC:6346] [ENST00000472022]	AF113122	0.0001479	2.50
36					
37	SPINK1	Homo sapiens serine peptidase inhibitor, Kazal type 1 (SPINK1), mRNA [NM_003122]	NM_003122	0.0000885	2.50
38					
39					
40	nd	Homo sapiens mRNA; cDNA DKFZp686P18215 (from clone DKFZp686P18215). [BX538202]		0.0005274	2.50
41					
42					
43	LOC101929787	PREDICTED: Homo sapiens uncharacterized LOC101929787 (LOC101929787), transcript variant X1, ncRNA [XR_250866]	XR_250866	0.0001169	2.44
44					
45					
46					
47	LOC79160	Homo sapiens uncharacterized LOC79160 (LOC79160), long non-coding RNA [NR_125359]	NR_125359	0.0001254	2.44
48					
49					
50	PRDM1	Homo sapiens PR domain containing 1, with ZNF domain (PRDM1), transcript variant 1, mRNA [NM_001198]	NM_001198	0.0004009	2.44
51					
52					
53	RTKN2	Homo sapiens rhotekin 2 (RTKN2), transcript variant 1, mRNA [NM_145307]	NM_145307	0.0001249	2.44
54					
55	TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0.0004021	2.44
56					
57					
58	LFNG	Homo sapiens LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase (LFNG), transcript variant 1, mRNA [NM_001040167]	NM_001040167	0.0003232	2.33
59					
60					

1	PLEK2	Homo sapiens pleckstrin 2 (PLEK2), mRNA [NM_016445]	NM_016445	0.0004734	2.33
2	SPOCK1	Homo sapiens sparc/osteonectin, cwcx and kazal-like domains proteoglycan (testican) 1 (SPOCK1), mRNA [NM_004598]	NM_004598	0.0004942	2.33
3	LIMS3L	Homo sapiens LIM and senescent cell antigen-like domains 3-like (LIMS3L), transcript variant 2, non-coding RNA [NR_038099]	NR_038099	0.0002299	2.27
4	SPSB1	Homo sapiens splA/ryanodine receptor domain and SOCS box containing 1 (SPSB1), mRNA [NM_025106]	NM_025106	0.0001597	2.27
5	TGFBI	Homo sapiens transforming growth factor, beta-induced, 68kDa (TGFBI), mRNA [NM_000358]	NM_000358	0.0000312	2.27
6	TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0.0000119	2.27
7	AMIGO2	Homo sapiens adhesion molecule with Ig-like domain 2 (AMIGO2), transcript variant 2, mRNA [NM_181847]	NM_181847	0.0001694	2.22
8	RTKN2	Homo sapiens rhotekin 2 (RTKN2), transcript variant 2, mRNA [NM_001282941]	NM_001282941	0.0002168	2.22
9	SSC5D	Homo sapiens scavenger receptor cysteine rich family, 5 domains (SSC5D), transcript variant 1, mRNA [NM_001144950]	NM_001144950	0.0008854	2.22
10	TUBA4A	Homo sapiens tubulin, alpha 4a (TUBA4A), transcript variant 1, mRNA [NM_006000]	NM_006000	0.0004931	2.22
11	KLF7	Homo sapiens Kruppel-like factor 7 (ubiquitous) (KLF7), transcript variant 1, mRNA [NM_003709]	NM_003709	0.0002433	2.17
12	KPNA7	Homo sapiens karyopherin alpha 7 (importin alpha 8) (KPNA7), mRNA [NM_001145715]	NM_001145715	0.0002695	2.17
13	LRIG3	Homo sapiens leucine-rich repeats and immunoglobulin-like domains 3 (LRIG3), transcript variant 2, mRNA [NM_153377]	NM_153377	0.0003266	2.17
14	nd	long intergenic non-protein coding RNA 313 [Source:HGNC Symbol;Acc:HGNC:16416] [ENST00000427188]	XR_425284	0.0001039	2.17
15	EPHB2	Homo sapiens EPH receptor B2 (EPHB2), transcript variant 2, mRNA [NM_004442]	NM_004442	0.0000119	2.13
16	ITGB6	Homo sapiens integrin, beta 6 (ITGB6), transcript variant 1, mRNA [NM_000888]	NM_000888	0.00002	2.13
17	MEG3	Homo sapiens maternally expressed 3 (non-protein coding) (MEG3), transcript variant 1, long non-coding RNA [NR_002766]	NR_002766	0.0000038	2.13
18	AADAC	Homo sapiens arylacetamide deacetylase (AADAC), mRNA [NM_001086]	NM_001086	0.0005116	2.08
19	ARSI	Homo sapiens arylsulfatase family, member I (ARSI), mRNA [NM_001012301]	NM_001012301	0.0002143	2.08
20	BMF	Homo sapiens Bcl2 modifying factor (BMF), transcript variant 1, mRNA [NM_001003940]	NM_001003940	0.0000505	2.08
21	MAP3K2	Homo sapiens mitogen-activated protein kinase kinase kinase 2 (MAP3K2), mRNA [NM_006609]	NM_006609	0.0000691	2.08

1					
2	CYTH1	Homo sapiens cytohesin 1 (CYTH1), transcript variant 1, mRNA [NM_004762]	NM_004762	0.0000692	2.04
3					
4	CYTH1	Homo sapiens cytohesin 1 (CYTH1), transcript variant 3, mRNA [NM_001292018]	NM_001292018	0.000232	2.04
5					
6					
7	EEPD1	Homo sapiens endonuclease/exonuclease/phosphatase family domain containing 1 (EEPD1), mRNA [NM_030636]	NM_030636	0.0002349	2.04
8					
9					
10					
11	EPHB2	Homo sapiens EPH receptor B2 (EPHB2), transcript variant 2, mRNA [NM_004442]	NM_004442	0.0000251	2.04
12					
13	LIMS1	Homo sapiens LIM and senescent cell antigen-like domains 1 (LIMS1), transcript variant 2, mRNA [NM_004987]	NM_004987	0.000038	2.04
14					
15	PGM2L1	Homo sapiens phosphoglucomutase 2-like 1 (PGM2L1), mRNA [NM_173582]	NM_173582	0.000236	2.04
16					
17	ABCA1	Homo sapiens ATP-binding cassette, sub-family A (ABC1), member 1 (ABCA1), mRNA [NM_005502]	NM_005502	0.0005224	2.00
18					
19					
20	KLRC1	Homo sapiens killer cell lectin-like receptor subfamily C, member 1 (KLRC1), transcript variant 1, mRNA [NM_002259]	NM_002259	0.0006104	0.49
21					
22					
23	SLC22A18AS	Homo sapiens solute carrier family 22 (organic cation transporter), member 18 antisense (SLC22A18AS), transcript variant 1, mRNA [NM_007105]	NM_007105	0.0006927	0.49
24					
25					
26					
27	TRIM55	Homo sapiens tripartite motif containing 55 (TRIM55), transcript variant 2, mRNA [NM_033058]	NM_033058	0.0003219	0.48
28					
29					
30	LOC101929641	PREDICTED: Homo sapiens uncharacterized LOC101929641 (LOC101929641), ncRNA [XR_243223]	XR_243223	0.0002746	0.45
31					
32					
33	KCNH5	Homo sapiens potassium channel, voltage gated eag related subfamily H, member 5 (KCNH5), transcript variant 3, mRNA [NM_172375]	NM_172375	0.0001595	0.01
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					

Supplementary Table S4. List of genes differentially expressed between FOXS1-positive and -negative HCC cell lines.

Gene symbol	Gene description	Accession nb.	P-value	Fold change (FOXS1 positive/negative)
VIM	vimentin	NM_003380	0.0218772	16.13
S100A6	S100 calcium binding protein A6	NM_014624	0.000507	14.93
FSTL1	folliculin-like 1	NM_007085	0.0037708	12.50
PLK2	polo-like kinase 2 (Drosophila)	NM_006622	0.0060105	10.00
IGFBP7	insulin-like growth factor binding protein 7	NM_001553	0.0267836	10.00
CD44	CD44 molecule (Indian blood group)	NM_000610	0.0121337	8.33
FZD2	frizzled homolog 2 (Drosophila)	NM_001466	0.011379	6.67
COL6A3	collagen, type VI, alpha 3	NM_004369	0.0370894	6.67
CYR61	cysteine-rich, angiogenic inducer, 61	NM_001554	0.0141182	6.25
GREM1	gremlin 1, cysteine knot superfamily, homolog (Xenopus laevis)	NM_013372	0.0336541	6.25
AXL	AXL receptor tyrosine kinase	NM_001699	0.0168215	5.88
BASP1	brain abundant, membrane attached signal protein 1	NM_006317	0.0348693	5.88
MYOF	myoferlin	NM_013451	0.0046894	5.26
AMIGO2	adhesion molecule with Ig-like domain 2	NM_001143668	0.0158944	5.26
AFAP1	actin filament associated protein 1	NM_001134647	0.0017488	5.00
PFKP	phosphofructokinase, platelet	NM_002627	0.00275	5.00
TRIM16L	tripartite motif-containing 16-like	NM_001037330	0.0038281	5.00
DCBLD2	discoidin, CUB and LCCL domain containing 2	NM_080927	0.0059678	4.55
LOXL2	lysyl oxidase-like 2	NM_002318	0.0018	4.35
RAGE	renal tumor antigen	NM_014226	0.0038905	4.35
LGALS1	lectin, galactoside-binding, soluble, 1	NM_002305	0.0078234	4.35
CNN1	calponin 1, basic, smooth muscle	NM_001299	0.0354466	4.35
HIF1A	hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	NM_001530	0.0003751	4.17
MYLK	myosin light chain kinase	NM_005965	0.0031206	4.17
OGFRL1	opioid growth factor receptor-like 1	NM_024576	0.0056836	4.17
CEP55	centrosomal protein 55kDa	NM_001127182	0.0212508	4.17
ADAM12	ADAM metallopeptidase domain 12	NM_003474	0.0134408	4.00

1	NNMT	nicotinamide N-methyltransferase	NM_006169	0.0466616	4.00
2	ARHGAP29	Rho GTPase activating protein 29	NM_004815	0.0235148	3.57
3	MSN	moesin	NM_002444	0.0350986	3.57
4	UBASH3B	ubiquitin associated and SH3 domain containing, B	NM_032873	0.0407602	3.57
5	TGFB2	transforming growth factor, beta 2	NM_001135599	0.0002047	3.45
6	FHL1	four and a half LIM domains 1	NM_001159699	0.0040312	3.45
7	BCAR3	breast cancer anti-estrogen resistance 3	NM_003567	0.0090764	3.45
8	PLEC	similar to Plectin 1 (PLTN) (PCN) (Hemidesmosomal protein 1) (HD1); plectin 1, intermediate filament binding protein 500kDa	NM_000445	0.0102814	3.45
9	FRMD6	FERM domain containing 6	NM_001042481	0.0433394	3.45
10	ARL3	ADP-ribosylation factor-like 3	NM_004311	0.0029974	3.33
11	PDLIM7	PDZ and LIM domain 7 (enigma)	NM_005451	0.0120542	3.33
12	FAM129A	family with sequence similarity 129, member A	NM_052966	0.0470019	3.33
13	TMEFF1	transmembrane protein with EGF-like and two follistatin-like domains 1; chromosome 9 open reading frame 30; hypothetical LOC729538	NM_003692	0.0137973	3.23
14	TRIM43	Tripartite motif-containing protein 43-like protein 1; tripartite motif-containing 43; similar to Tripartite motif-containing protein 43-like protein 2	NM_138800	0.0233122	3.23
15	ACTN1	actinin, alpha 1	NM_001102	0.0029925	3.13
16	ZFPM2	zinc finger protein, multitype 2	NM_012082	0.0110576	3.13
17	GPR176	G protein-coupled receptor 176	NM_007223	0.0234355	3.13
18	CPA4	carboxypeptidase A4	NM_001163446	0.0243904	3.13
19	FXVD5	FXVD domain containing ion transport regulator 5	NM_001164605	0.0374444	3.13
20	CCNE2	cyclin E2	NM_057749	0.0014749	3.03
21	GNG11	guanine nucleotide binding protein (G protein), gamma 11	NM_004126	0.0291662	3.03
22	TCP11	t-complex 11 homolog (mouse)	NM_001093728	0.001026	2.94
23	SPANXB2	SPANX family, member B1; SPANX family, member F1; SPANX family, member B2	NM_145664	0.0206382	2.94
24	MAP7D3	MAP7 domain containing 3	NM_001173516	0.0314871	2.94
25	CD47	CD47 molecule	NM_001777	0.0495231	2.94
26	LTBP2	latent transforming growth factor beta binding protein 2	NM_000428	0.0263883	2.86

DEGS1	degenerative spermatocyte homolog 1, lipid desaturase (Drosophila)	NM_003676	0.0406881	2.86
PITRM1	pitrilysin metallopeptidase 1	NM_014889	0.0066028	2.78
FLNA	filamin A, alpha (actin binding protein 280)	NM_001456	0.0102853	2.78
ALDH3A1	aldehyde dehydrogenase 3 family, memberA1	NM_000691	0.0137527	2.78
CEP170	centrosomal protein 170kDa	NM_001042404	0.0154674	2.78
UGT1A1	UDP glucuronosyltransferase 1 family, polypeptide A1	NM_019075	0.0304072	2.78
ST6GALNAC2	ST6 (alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 2	NM_006456	0.0438569	2.78
C14orf149	chromosome 14 open reading frame 149	NM_144581	0.000626	2.70
TFE3	transcription factor binding to IGHM enhancer 3	NM_006521	0.0075072	2.70
LMNA	lamin A/C	NM_005572	0.0000997	2.63
C1S	complement component 1, s subcomponent	NM_001734	0.0048418	2.63
ULBP1	UL16 binding protein 1	NM_025218	0.0109839	2.63
PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)	NM_000300	0.0191111	2.63
MAP7D1	MAP7 domain containing 1	NM_018067	0.0006946	2.56
NFIB	nuclear factor I/B	NM_001190737	0.0133247	2.56
TUBB2A	tubulin, beta 2A	NM_001069	0.0255762	2.56
CDCA2	cell division cycle associated 2	NM_152562	0.0417427	2.56
NEK7	NIMA (never in mitosis gene a)-related kinase 7	NM_133494	0.0444494	2.56
SPIN1	spindlin 1	NM_006717	0.0091466	2.50
EMP3	epithelial membrane protein 3	NM_001425	0.0227472	2.50
ERBB2IP	erbb2 interacting protein	NM_001006600	0.02749	2.50
IFI27L2	interferon, alpha-inducible protein 27-like 2	NM_032036	0.0372615	2.50
CITED2	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2	NM_001168388	0.0431242	2.50
MOSPD1	motile sperm domain containing 1	NM_019556	0.0047051	2.44
CAPN2	calpain 2, (m/II) large subunit	NM_001146068	0.0048799	2.44
PICALM	phosphatidylinositol binding clathrin assembly protein	NM_001008660	0.0079566	2.44
CEACAM8	carcinoembryonic antigen-related cell adhesion molecule 8	NM_001816	0.0154476	2.44
NKX2-8	NK2 homeobox 8	NM_014360	0.0258898	2.38

1	TMEM158	transmembrane protein 158	NM_015444	0.028142	2.38
2					
3	VOPP1	similar to EGFR-coamplified and overexpressed protein; EGFR-coamplified and overexpressed protein	NM_030796	0.0002467	2.33
4					
5	KIAA1279	KIAA1279	NM_015634	0.0030593	2.33
6					
7	ASAP1	ArfGAP with SH3 domain, ankyrin repeat and PH domain 1	NM_018482	0.0063268	2.33
8					
9	WBP5	WW domain binding protein 5	NM_001006612	0.0486981	2.33
10					
11	LASP1	LIM and SH3 protein 1	NM_006148	0.0058774	2.27
12	ZNF93	zinc finger protein 93	NM_031218	0.0127539	2.27
13	ARL6IP5	ADP-ribosylation-like factor 6 interacting protein 5	NM_006407	0.0276335	2.27
14	RAB18	RAB18, member RAS oncogene family	NM_021252	0.0352943	2.27
15					
16	CDA	cytidine deaminase	NM_001785	0.038189	2.27
17	TMEM14C	transmembrane protein 14C	NM_001165258	0.0052204	2.22
18	ARHGAP23	Rho GTPase activating protein 23	NM_020876	0.0071682	2.22
19	LOC221442	adenylate cyclase 10 pseudogene	NM_001010871	0.0123894	2.22
20	UHMK1	U2AF homology motif (UHM) kinase 1	NM_001184763	0.0152913	2.22
21	G6PD	glucose-6-phosphate dehydrogenase	NM_000402	0.0177551	2.22
22	CFHR2	complement factor H-related 2	NM_005666	0.038774	2.22
23					
24	MYL1	myosin, light chain 1, alkali; skeletal, fast	NM_079420	0.039576	2.22
25	PYGB	phosphorylase, glycogen; brain	NM_002862	0.0470875	2.22
26	NRIP3	nuclear receptor interacting protein 3	NM_020645	0.0142177	2.17
27					
28	SLC9A6	solute carrier family 9 (sodium/hydrogen exchanger), member 6	NM_001042537	0.006104	2.13
29	OPTN	optineurin	NM_001008211	0.0062544	2.13
30	TOX2	TOX high mobility group box family member 2	NM_001098796	0.0081463	2.13
31					
32	PSRC1	proline/serine-rich coiled-coil 1	NM_001005290	0.0098832	2.13
33	PPME1	protein phosphatase methylesterase 1	NM_016147	0.0326571	2.13
34	PFN2	profilin 2	NM_002628	0.0403955	2.13
35	PTPN1	protein tyrosine phosphatase, non-receptor type 1	NM_002827	0.000457	2.08
36	RRAS	related RAS viral (r-ras) oncogene homolog	NM_006270	0.0056715	2.08
37	GDI1	GDP dissociation inhibitor 1	NM_001493	0.0094438	2.08
38					
39	COL4A3BP	collagen, type IV, alpha 3 (Goodpasture antigen) binding protein	NM_001130105	0.0123175	2.08
40					
41					
42					
43					
44					
45					
46					

1	TSEN15	tRNA splicing endonuclease 15 homolog (<i>S. cerevisiae</i>)	NM_001127394	0.0215841	2.08
2	AIDA	hypothetical LOC653631; hypothetical LOC646050; hypothetical LOC646890; axin interactor, dorsalization associated	NM_022831	0.0412513	2.08
3	PSMD10	proteasome (prosome, macropain) 26S subunit, non- ATPase, 10	NM_002814	0.0003344	2.04
4	MET	met proto-oncogene (hepatocyte growth factor receptor)	NM_000245	0.008475	2.04
5	TMEM126B	transmembrane protein 126B	NM_001193537	0.0141663	2.04
6	NF2	neurofibromin 2 (merlin)	NM_000268	0.0165951	2.04
7	AKR1B1	aldo-keto reductase family 1, member B1 (aldose reductase)	NM_001628	0.0182534	2.04
8	CCR2	chemokine (C-C motif) receptor 2	NM_000647	0.0237576	2.04
9	RRAS2	related RAS viral (r-ras) oncogene homolog 2; similar to related RAS viral (r-ras) oncogene homolog 2	NM_012250	0.0363583	2.04
10	FZD6	frizzled homolog 6 (<i>Drosophila</i>)	NM_001164615	0.0371627	2.04
11	WDR26	WD repeat domain 26	NM_001115113	0.0119659	2.00
12	ZNF626	zinc finger protein 626	NM_001076675	0.0154565	2.00
13	CECR1	cat eye syndrome chromosome region, candidate 1	NM_017424	0.0170295	2.00
14	NOTCH4	Notch homolog 4 (<i>Drosophila</i>)	NM_004557	0.0047213	0.50
15	P4HB	prolyl 4-hydroxylase, beta polypeptide	NM_000918	0.0010771	0.50
16	VWCE	von Willebrand factor C and EGF domains	NM_152718	0.0165737	0.50
17	ZFP90	zinc finger protein 90 homolog (mouse)	NM_133458	0.0267531	0.50
18	SELS	selenoprotein S	NM_018445	0.029907	0.50
19	MRPS31	mitochondrial ribosomal protein S31	NM_005830	0.0195121	0.49
20	CORO1A	coronin, actin binding protein, 1A	NM_001193333	0.0316224	0.49
21	CYFIP2	cytoplasmic FMR1 interacting protein 2	NM_001037332	0.0013443	0.49
22	AKR1A1	aldo-keto reductase family 1, member A1 (aldehyde reductase)	NM_006066	0.0095779	0.49
23	NFXL1	nuclear transcription factor, X-box binding-like 1	NM_152995	0.0123731	0.49
24	ANKRD57	ankyrin repeat domain 57	NM_023016	0.0200394	0.49
25	PDXDC1	pyridoxal-dependent decarboxylase domain containing 1	NM_015027	0.0149695	0.49
26	MTUS1	mitochondrial tumor suppressor 1	NM_001001927	0.0062605	0.49

1				
2	NDUFAB1	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	NM_005003	0.0033625
3				0.48
4	FADS1	fatty acid desaturase 1	NM_013402	0.0296629
5	AKAP6	A kinase (PRKA) anchor protein 6	NM_004274	0.0004556
6	YIPF3	Yip1 domain family, member 3	NM_015388	0.0055165
7				0.48
8	IGDCC3	immunoglobulin superfamily, DCC subclass, member 3	NM_004884	0.0129571
9	SARDH	sarcosine dehydrogenase	NM_007101	0.0206338
10				0.48
11	TMTC4	transmembrane and tetratricopeptide repeat containing 4	NM_001079669	0.0036395
12				0.48
13	MKNK2	MAP kinase interacting serine/threonine kinase 2	NM_017572	0.0262635
14	TMEM98	similar to transmembrane protein 98; transmembrane protein 98	NM_001033504	0.0148516
15				0.48
16	RPN2	ribophorin II	NM_001135771	0.0185036
17	RAB33B	RAB33B, member RAS oncogene family	NM_031296	0.0215592
18	DGCR8	DiGeorge syndrome critical region gene 8	NM_001190326	0.022739
19	BEX2	brain expressed X-linked 2	NM_001168399	0.0487006
20				0.47
21	NCK2	NCK adaptor protein 2	NM_001004720	0.0294789
22	SLITRK6	SLIT and NTRK-like family, member 6	NM_032229	0.0331201
23				0.47
24	RAB26	RAB26, member RAS oncogene family	NM_014353	0.0047453
25	POLR1D	polymerase (RNA) I polypeptide D, 16kDa	NM_015972	0.0222068
26	FCER2	Fc fragment of IgE, low affinity II, receptor for (CD23)	NM_002002	0.0234349
27				0.47
28	CTDSP1	CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 1	NM_021198	0.0014983
29	MANBA	mannosidase, beta A, lysosomal	NM_005908	0.0030353
30	TP53INP2	tumor protein p53 inducible nuclear protein 2	NM_021202	0.0035635
31				0.46
32	IL12RB1	interleukin 12 receptor, beta 1	NM_005535	0.00383
33	SC5DL	sterol-C5-desaturase (ERG3 delta-5-desaturase homolog, S. cerevisiae)-like	NM_001024956	0.0080301
34				0.46
35	SEPX1	selenoprotein X, 1	NM_016332	0.0108025
36	ACP2	acid phosphatase 2, lysosomal	NM_001131064	0.0202783
37				0.46
38	TARP	TCR gamma alternate reading frame protein; T cell receptor gamma variable 9; T cell receptor gamma constant 1	NM_001003806	0.0096901
39				0.46
40	HSD17B7	hydroxysteroid (17-beta) dehydrogenase 7	NM_016371	0.01259
41	ORMDL3	ORM1-like 3 (S. cerevisiae)	NM_139280	0.0300465
42				0.46
43				
44				
45				
46				

1	TEK	TEK tyrosine kinase, endothelial	NM_000459	0.043744	0.46
2	MGC12982	hypothetical protein MGC12982	BC006113	0.0126916	0.46
3	SALL4	sal-like 4 (Drosophila)	NM_020436	0.0164084	0.46
4	C3orf32	chromosome 3 open reading frame 32	NM_015931	0.0166471	0.46
5	C11orf2	chromosome 11 open reading frame2	NM_013265	0.0210295	0.46
6	PPAP2B	phosphatidic acid phosphatase type 2B	NM_003713	0.0253967	0.45
7	LEPREL4	synaptonemal complex protein SC65	NM_006455	0.0306304	0.45
8	CYP19A1	cytochrome P450, family 19, subfamily A, polypeptide 1	NM_000103	0.0168898	0.45
9	ETNK1	ethanolamine kinase 1	NM_001039481	0.0043041	0.45
10	BAIAP2L2	BAI1-associated protein 2-like 2	NM_025045	0.0256251	0.45
11	KLHDC2	kelch domain containing 2	NM_014315	0.0228383	0.45
12	LITAF	lipopolysaccharide-induced TNF factor	NM_001136472	0.0095804	0.45
13	HPD	4-hydroxyphenylpyruvate dioxygenase	NM_001171993	0.015704	0.45
14	MYO5C	myosin VC	NM_018728	0.028854	0.45
15	VAMP8	vesicle-associated membrane protein 8 (endobrevin)	NM_003761	0.0195398	0.44
16	CYP51A1	cytochrome P450, family 51, subfamily A, polypeptide 1	NM_000786	0.0339774	0.44
17	GCM1	glial cells missing homolog 1 (Drosophila)	NM_003643	0.0000641	0.44
18	PIGF	phosphatidylinositol glycan anchor biosynthesis, class F	NM_002643	0.0007086	0.44
19	AMT	aminomethyltransferase	NM_000481	0.0337982	0.44
20	OAT	ornithine aminotransferase (gyrate atrophy)	NM_000274	0.0025596	0.44
21	LRRC1	leucine rich repeat containing 1	NM_018214	0.0097007	0.44
22	HSP90B1	heat shock protein 90kDa beta (Grp94), member 1	NM_003299	0.0212113	0.44
23	GLYATL1	glycine-N-acyltransferase-like 1	NM_080661	0.0213014	0.44
24	WWP1	WW domain containing E3 ubiquitin protein ligase 1	NM_007013	0.0291177	0.44
25	DHRS3	dehydrogenase/reductase (SDR family) member 3	NM_004753	0.0293894	0.44
26	NRSN1	neurensin 1	NM_080723	0.0218264	0.44
27	SERPINA1	serpin peptidase inhibitor, clade A (alpha-1 antiprotease, antitrypsin), member 1	NM_000295	0.0036913	0.43
28	SPRY1	sprouty homolog 1, antagonist of FGF signaling (Drosophila)	NM_005841	0.0475281	0.43
29	DCI	dodecenoyl-Coenzyme A delta isomerase (3,2 trans-enoyl-Coenzyme A isomerase)	NM_001919	0.0060188	0.43

1	MVD	mevalonate (diphospho) decarboxylase	NM_002461	0.0106013	0.43
2	DAXX	death-domain associated protein	NM_001141969	0.0034765	0.43
3	ADH4	alcohol dehydrogenase 4 (class II), pi polypeptide	NM_000670	0.0312836	0.43
4	ATP1B1	ATPase, Na ⁺ /K ⁺ transporting, beta 1 polypeptide	NM_001677	0.0016851	0.42
5	RHOA	ras homolog gene family, member U	NM_021205	0.011354	0.42
6	DUSP9	dual specificity phosphatase 9	NM_001395	0.0280272	0.42
7	CYP27A1	cytochrome P450, family 27, subfamily A, polypeptide 1	NM_000784	0.030994	0.42
8	CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma	NM_001806	0.0012092	0.42
9	SH3D19	SH3 domain containing 19	NM_001009555	0.037744	0.42
10	SFN	stratifin	NM_006142	0.0457033	0.42
11	CHPT1	choline phosphotransferase 1	NM_020244	0.0191907	0.42
12	PDIA5	protein disulfide isomerase family A, member 5	NM_006810	0.0153554	0.41
13	C2orf72	chromosome 2 open reading frame 72	NM_001144994	0.0048662	0.41
14	TMEM135	transmembrane protein 135	NM_001168724	0.0236466	0.41
15	DENND1B	DENN/MADD domain containing 1B	NM_001195215	0.0427562	0.41
16	RAB17	RAB17, member RAS oncogene family	NM_022449	0.0256368	0.41
17	DHCR7	7-dehydrocholesterol reductase	NM_001163817	0.0074577	0.41
18	UGT2B10	UDP glucuronosyltransferase 2 family, polypeptide B10	NM_001075	0.0323595	0.41
19	CAT	catalase	NM_001752	0.033503	0.41
20	QPRT	quinolinate phosphoribosyltransferase	NM_014298	0.0332848	0.40
21	FGFR4	fibroblast growth factor receptor 4	NM_002011	0.0003279	0.40
22	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	NM_000617	0.0003336	0.40
23	UBC	scavenger receptor class B, member 1	NM_001082959	0.0234863	0.40
24	CXADR	coxsackie virus and adenovirus receptor pseudogene 2; coxsackie virus and adenovirus receptor	NM_001338	0.0107251	0.39
25	CDHR2	protocadherin 24	NM_017675	0.0157996	0.39
26	RAB43	RAB43, member RAS oncogene family; hypothetical LOC100131426	NM_198490	0.0037608	0.39
27	PTPRU	protein tyrosine phosphatase, receptor type, U	NM_001195001	0.008345	0.39
28	TMEM129	transmembrane protein 129	NM_001127266	0.0247158	0.39
29	GCH1	GTP cyclohydrolase 1	NM_000161	0.0079242	0.39
30	AGFG2	ArfGAP with FG repeats 2	NM_006076	0.0082976	0.39

KCNK1	potassium channel, subfamily K, member 1	NM_002245	0.0172193	0.39
SLC25A33	solute carrier family 25, member 33	NM_032315	0.0218303	0.38
CMTM4	CKLF-like MARVEL transmembrane domain containing 4	NM_178818	0.0409192	0.38
FOXA1	forkhead box A1	NM_004496	0.0305435	0.38
UGT2B4	UDP glucuronosyltransferase 2 family, polypeptide B4	NM_021139	0.0330413	0.38
CREB3L3	cAMP responsive element binding protein 3-like 3	NM_032607	0.0180624	0.38
GLDC	glycine dehydrogenase (decarboxylating)	NM_000170	0.0188667	0.38
NDRG2	NDRG family member 2	NM_016250	0.0032483	0.37
NIPSNAP1	nipsnap homolog 1 (<i>C. elegans</i>)	NM_003634	0.018358	0.37
GMCL1	germ cell-less homolog 1 (<i>Drosophila</i>)-like; germ cell-less homolog 1 (<i>Drosophila</i>)	NM_178439	0.0452003	0.37
PAH	phenylalanine hydroxylase	NM_000277	0.0260096	0.37
WDHD1	WD repeat and HMG-box DNA binding protein 1	NM_001008396	0.0221452	0.36
MARVELD2	MARVEL domain containing 2	NM_001038603	0.0089786	0.36
NOMO3	NODAL modulator 3; NODAL modulator 1; NODAL modulator 2	NM_173614	0.0146842	0.36
AADAT	aminoadipate aminotransferase	NM_016228	0.0234341	0.36
PROC	protein C (inactivator of coagulation factors Va and VIIIa)	NM_000312	0.0231937	0.36
VIL1	villin 1	NM_007127	0.0046915	0.36
ITPKC	inositol 1,4,5-trisphosphate 3-kinase C	NM_025194	0.0379615	0.35
ALDH6A1	aldehyde dehydrogenase 6 family, member A1	NM_005589	0.0093598	0.35
ARID3A	AT rich interactive domain 3A (BRIGHT-like)	NM_005224	0.0262267	0.35
GRTP1	growth hormone regulated TBC protein 1	NM_024719	0.0467801	0.35
MBL2	mannose-binding lectin (protein C) 2, soluble (opsonic defect)	NM_000242	0.0075849	0.35
SLC39A4	solute carrier family 39 (zinc transporter), member 4	NM_017767	0.0130256	0.35
DPEP1	dipeptidase 1 (renal)	NM_001128141	0.0115937	0.35
FZD4	frizzled homolog 4 (<i>Drosophila</i>)	NM_012193	0.0172832	0.35
ODAM	odontogenic, ameloblast associated	NM_017855	0.0467975	0.34
DHCR24	24-dehydrocholesterol reductase	NM_014762	0.0491099	0.34
TCTE3	t-complex-associated-testis-expressed 3	NM_174910	0.0318942	0.34

SMARCA1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	NM_003069	0.0066164	0.34
FUT7	fucosyltransferase 7 (alpha (1,3) fucosyltransferase)	NM_004479	0.0434782	0.34
ZDHHC11	zinc finger, DHHC-type containing 11	NM_024786	0.0347613	0.34
XBP1	X-box binding protein 1	NM_001079539	0.0012675	0.34
HCLS1	hematopoietic cell-specific Lyn substrate 1	NM_005335	0.0080815	0.33
SCD	stearoyl-CoA desaturase (delta-9-desaturase)	NM_005063	0.0390941	0.33
NFKBIZ	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, zeta	NM_001005474	0.0316724	0.33
HSPB2	heat shock 27kDa protein 2	NM_001541	0.013069	0.33
MDK	midkine (neurite growth-promoting factor 2)	NM_001012333	0.001366	0.33
ETV4	ets variant 4	NM_001079675	0.0166282	0.33
C16orf72	chromosome 16 open reading frame 72	NM_014117	0.0226884	0.32
GRB7	growth factor receptor-bound protein 7	NM_001030002	0.002904	0.32
S100G	S100 calcium binding protein G	NM_004057	0.03401	0.32
VTN	vitronectin	NM_000638	0.0159195	0.32
GAMT	guanidinoacetate N-methyltransferase	NM_000156	0.0218132	0.32
EFNA1	ephrin-A1	NM_004428	0.0423641	0.32
NAB1	NGFI-A binding protein 1 (EGR1 binding protein 1)	NM_005966	0.0007033	0.31
CLN8	ceroid-lipofuscinosis, neuronal 8 (epilepsy, progressive with mental retardation)	NM_018941	0.0414024	0.31
ASS1	argininosuccinate synthetase 1	NM_000050	0.0433779	0.31
MPST	mercaptopyruvate sulfurtransferase	NM_001013436	0.0243661	0.31
AP1AR	chromosome 4 open reading frame 16	NM_018569	0.0087043	0.31
C5orf33	chromosome 5 open reading frame 33	NM_001085411	0.0290653	0.31
PECR	peroxisomal trans-2-enoyl-CoA reductase	NM_018441	0.0226553	0.30
SEMA4G	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4G	NM_017893	0.0135568	0.30
RTKN	rhotekin	NM_001015055	0.0057105	0.30
ABCC2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	NM_000392	0.030707	0.30
SLC2A2	solute carrier family 2 (facilitated glucose transporter), member 2	NM_000340	0.0031336	0.30

1	FGFR3	fibroblast growth factor receptor 3	NM_000142	0.0255992	0.29
2	ABAT	4-aminobutyrate aminotransferase	NM_000663	0.0262607	0.29
3	CX3CL1	chemokine (C-X3-C motif) ligand 1	NM_002996	0.0339986	0.29
4	TTC39B	tetratricopeptide repeat domain 39B	NM_001168339	0.0000901	0.29
5	TST	thiosulfate sulfurtransferase (rhodanese)	NM_003312	0.0034207	0.28
6	CPB2	carboxypeptidase B2 (plasma)	NM_001872	0.0154714	0.28
7	FAM8A1	family with sequence similarity 8, member A1	NM_016255	0.0004068	0.28
8	TRIM24	tripartite motif-containing 24	NM_003852	0.0113789	0.28
9	LEAP2	liver expressed antimicrobial peptide 2	NM_052971	0.0258057	0.28
10	APOC2	apolipoprotein C-II	NM_000483	0.0437346	0.28
11	PTP4A1	protein tyrosine phosphatase type IVA, member 1	NM_003463	0.0248734	0.28
12	FGB	fibrinogen beta chain	NM_001184741	0.0024887	0.27
13	ESPN	espin	NM_031475	0.0377086	0.27
14	A1CF	APOBEC1 complementation factor	NM_014576	0.0075852	0.27
15	OAF	OAF homolog (Drosophila)	NM_178507	0.0359734	0.26
16	F7	coagulation factor VII (serum prothrombin conversion accelerator)	NM_000131	0.0403171	0.26
17	ASNS	asparagine synthetase	NM_001673	0.0272056	0.25
18	HSD17B11	hydroxysteroid (17-beta) dehydrogenase 11	NM_016245	0.0172677	0.25
19	MAL2	mal, T-cell differentiation protein 2	NM_052886	0.013537	0.25
20	RPS6KA3	ribosomal protein S6 kinase, 90kDa, polypeptide 3	NM_004586	0.0141008	0.24
21	SLC39A14	solute carrier family 39 (zinc transporter), member 14	NM_001128431	0.0251043	0.24
22	STARD10	StAR-related lipid transfer (START) domain containing 10	NM_006645	0.0004816	0.24
23	PROS1	protein S (alpha)	NM_000313	0.0065447	0.24
24	PCSK6	proprotein convertase subtilisin/kexin type 6	NM_002570	0.0235361	0.23
25	REEP6	receptor accessory protein 6	NM_138393	0.0393564	0.23
26	TF	transferrin	NM_001063	0.0265097	0.22
27	FAM60A	similar to hCG2020539; family with sequence similarity 60, member A; similar to family with sequence similarity 60, member A	NM_021238	0.0228732	0.21
28	GJB1	gap junction protein, beta 1, 32kDa	NM_000166	0.006832	0.21
29	SERPINF2	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 2	NM_000934	0.020088	0.21

1	AGT	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	NM_000029	0.0051017	0.21
2	IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	NM_005896	0.015613	0.20
3	HMGCS1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)	NM_002130	0.0315203	0.20
4	FGG	fibrinogen gamma chain	NM_000509	0.0037182	0.19
5	DSP	desmoplakin	NM_001008844	0.0190942	0.19
6	DUSP6	dual specificity phosphatase 6	NM_001946	0.0425929	0.19
7	MMD	monocyte to macrophage differentiation-associated	NM_012329	0.0014744	0.18
8	BCAM	basal cell adhesion molecule (Lutheran blood group)	NM_001013257	0.0030386	0.18
9	APOE	hypothetical LOC100129500; apolipoprotein E	NM_000041	0.0008954	0.17
10	GPAM	glycerol-3-phosphate acyltransferase, mitochondrial	NM_020918	0.0277664	0.17
11	ASGR1	asialoglycoprotein receptor 1	NM_001197216	0.0085157	0.16
12	SERPINA6	serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 6	NM_001756	0.0027892	0.16
13	C5	complement component 5	NM_001735	0.0039878	0.15
14	ANXA4	annexin A4	NM_001153	0.032391	0.15
15	MGST2	microsomal glutathione S-transferase 2	NM_002413	0.0074218	0.14
16	IQGAP2	IQ motif containing GTPase activating protein 2	NM_006633	0.0057675	0.14
17	GPX2	glutathione peroxidase 2 (gastrointestinal)	NM_002083	0.0011704	0.12
18	SERPINF1	serpin peptidase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1	NM_002615	0.015391	0.12
19	CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha	NM_004364	0.0391771	0.11
20	HGD	homogentisate 1,2-dioxygenase (homogentisate oxidase)	NM_000187	0.0073303	0.11
21	FKBP10	FK506 binding protein 10, 65 kDa	NM_021939	0.0029788	0.11
22	SLC2A3	solute carrier family 2 (facilitated glucose transporter), member 3	NM_006931	0.0414818	0.11
23	ASGR2	asialoglycoprotein receptor 2	NM_001181	0.0036893	0.10
24	ITIH2	inter-alpha (globulin) inhibitor H2	NM_002216	0.0360781	0.10
25	AHSG	alpha-2-HS-glycoprotein	NM_001622	0.0160616	0.07
26	GPC3	glypican 3	NM_001164617	0.011857	0.05
27	SSBP1	single-stranded DNA binding protein 1	NM_003143	0.0000083	0.03

Supplementary Table S5. Gene expression signature of endogenous FOXS1 silencing in SNU-449 cells.

Gene symbol	Gene description	TGF β -untreated cells				TGF β -treated cells			
		siFOXS1#1 vs siNC		siFOXS1#2 vs siNC		siFOXS1#1 vs siNC		siFOXS1#2 vs siNC	
		FC	P-value	FC	P-value	FC	P-value	FC	P-value
FOXS1	Homo sapiens forkhead box S1 (FOXS1), mRNA [NM_004118]	0.11	0.0004	0.02	0.0000	0.09	0.0004	0.02	0.0000
SPTB	Homo sapiens spectrin, beta, erythrocytic (SPTB), transcript variant 1, mRNA [NM_001024858]	0.26	0.0010	0.28	0.0001	0.33	0.0006	0.27	0.0000
LOC400743		0.27	0.0000	0.35	0.0000	0.29	0.0000	0.36	0.0000
ZNF192P1	Homo sapiens zinc finger protein 192 pseudogene 1 (ZNF192P1), non-coding RNA [NR_103448]	0.31	0.0000	0.37	0.0000	0.29	0.0000	0.46	0.0003
ZNF334	Homo sapiens zinc finger protein 334 (ZNF334), transcript variant 3, mRNA [NM_001270497]	0.31	0.0000	0.51	0.0006	0.36	0.0008	0.52	0.0004
C19orf33	Homo sapiens chromosome 19 open reading frame 33 (C19orf33), mRNA [NM_033520]	0.32	0.0000	0.29	0.0000	0.29	0.0001	0.27	0.0007
TBC1D2	Homo sapiens TBC1 domain family, member 2 (TBC1D2), transcript variant 1, mRNA [NM_001267571]	0.33	0.0001	0.37	0.0001	0.34	0.0000	0.40	0.0002
TBC1D16	Homo sapiens TBC1 domain family, member 16 (TBC1D16), transcript variant 1, mRNA [NM_019020]	0.34	0.0007	0.56	0.0001	0.38	0.0001	0.57	0.0007
NAP1L4	nucleosome assembly protein 1-like 4 [Source:HGNC Symbol;Acc:HGNC:7640] [ENST00000534372]	0.36	0.0006	0.43	0.0000	0.34	0.0000	0.45	0.0000
THBD	Homo sapiens thrombomodulin (THBD), mRNA [NM_000361]	0.40	0.0001	0.28	0.0000	0.34	0.0000	0.25	0.0001
lnc-SCCPDH-1	LNCipedia lincRNA (lnc-SCCPDH-1), lincRNA [lnc-SCCPDH-1:4]	0.43	0.0000	0.37	0.0000	0.46	0.0003	0.46	0.0002
OSBPL5	Homo sapiens oxysterol binding protein-like 5 (OSBPL5), transcript variant 1, mRNA [NM_020896]	0.44	0.0008	0.59	0.0002	0.47	0.0001	0.57	0.0004
THRA	Homo sapiens thyroid hormone receptor, alpha (THRA), transcript variant 2, mRNA [NM_003250]	0.44	0.0009	0.58	0.0001	0.42	0.0001	0.57	0.0007

1										
2	Inc-DYNC1I2-1	Homo sapiens cDNA FLJ14080 fis, clone HEMBB1002152. [AK024142]	0.44	0.0004	0.60	0.0008	0.50	0.0005	0.58	0.0004
3										
4	KLF13	Homo sapiens Kruppel-like factor 13 (KLF13), transcript variant 1, mRNA [NM_015995]	0.44	0.0001	0.64	0.0006	0.47	0.0001	0.59	0.0007
5										
6	IFI44	Homo sapiens interferon-induced protein 44 (IFI44), mRNA [NM_006417]	0.48	0.0010	2.14	0.0003	0.39	0.0001	2.22	0.0008
7										
8	LINC01481	Homo sapiens long intergenic non-protein coding RNA 1481 (LINC01481), transcript variant 1, long non-coding RNA [NR_120459]	0.52	0.0001	0.57	0.0001	0.59	0.0001	0.59	0.0005
9										
10	JDP2	Homo sapiens Jun dimerization protein 2 (JDP2), transcript variant 1, mRNA [NM_130469]	0.53	0.0005	0.59	0.0002	0.59	0.0010	0.60	0.0004
11										
12	MPP2	Homo sapiens membrane protein, palmitoylated 2 (MAGUK p55 subfamily member 2) (MPP2), transcript variant 9, mRNA [NM_001278381]	0.54	0.0008	0.54	0.0001	0.62	0.0004	0.52	0.0001
13										
14	RPL37	Homo sapiens ribosomal protein L37 (RPL37), mRNA [NM_000997]	0.58	0.0006	0.62	0.0007	0.60	0.0008	0.55	0.0003
15										
16	PTH2R	Homo sapiens parathyroid hormone 2 receptor (PTH2R), transcript variant 1, mRNA [NM_005048]	0.58	0.0003	0.59	0.0002	0.56	0.0003	0.60	0.0006
17										
18	SPHAR	Homo sapiens S-phase response (cyclin related) (SPHAR), mRNA [NM_006542]	0.59	0.0002	0.50	0.0001	0.61	0.0002	0.49	0.0001
19										
20	ALOX12-AS1	Homo sapiens ALOX12 antisense RNA 1 (ALOX12-AS1), long non-coding RNA [NR_040089]	0.60	0.0008	0.55	0.0008	0.50	0.0009	0.41	0.0006
21										
22	PTGFRN	Homo sapiens prostaglandin F2 receptor inhibitor (PTGFRN), mRNA [NM_020440]	0.61	0.0008	0.60	0.0001	0.61	0.0003	0.61	0.0007
23										
24	AMMECR1L	Homo sapiens AMMECR1-like (AMMECR1L), transcript variant 1, mRNA [NM_031445]	0.62	0.0000	0.59	0.0001	0.63	0.0003	0.54	0.0002
25										
26	PTPLA	Homo sapiens protein tyrosine phosphatase-like (proline instead of catalytic arginine), member A, mRNA (cDNA clone IMAGE:4304680), with apparent retained intron. [BC027709]	0.62	0.0000	0.62	0.0002	0.61	0.0007	0.59	0.0007
27										
28										
29										
30										
31										
32										
33										
34										
35										
36										
37										
38										
39										
40										
41										
42										
43										
44										
45										
46										

MSANTD3	Homo sapiens Myb/SANT-like DNA-binding domain containing 3 (MSANTD3), transcript variant 1, mRNA [NM_080655]	1.54	0.0003	1.97	0.0000	1.53	0.0007	1.85	0.0001
TTC27	Homo sapiens tetratricopeptide repeat domain 27 (TTC27), transcript variant 1, mRNA [NM_017735]	1.56	0.0001	1.97	0.0000	1.61	0.0003	1.96	0.0003
PDPK1	Homo sapiens 3-phosphoinositide dependent protein kinase 1 (PDPK1), transcript variant 1, mRNA [NM_002613]	1.60	0.0003	1.65	0.0002	1.96	0.0000	1.68	0.0005
TSPYL1	Homo sapiens TSPY-like 1 (TSPYL1), mRNA [NM_003309]	1.64	0.0001	1.57	0.0003	1.60	0.0003	1.58	0.0007
KIAA1715	Homo sapiens KIAA1715 (KIAA1715), mRNA [NM_030650]	1.64	0.0002	1.73	0.0001	1.61	0.0004	1.77	0.0002
UBE2O	Homo sapiens ubiquitin-conjugating enzyme E2O (UBE2O), mRNA [NM_022066]	1.64	0.0007	1.82	0.0001	1.83	0.0001	1.79	0.0002
CRYZ	Homo sapiens crystallin, zeta (quinone reductase) (CRYZ), transcript variant 1, mRNA [NM_001130042]	1.67	0.0003	1.82	0.0001	1.76	0.0002	1.96	0.0001
HNRNPH2	Homo sapiens heterogeneous nuclear ribonucleoprotein H2 (H') (HNRNPH2), transcript variant 1, mRNA [NM_019597]	1.68	0.0003	1.64	0.0001	1.75	0.0004	1.64	0.0009
MCTS1	Homo sapiens malignant T cell amplified sequence 1 (MCTS1), transcript variant 1, mRNA [NM_014060]	1.70	0.0000	1.67	0.0002	1.62	0.0006	1.69	0.0009
ADAM17	Homo sapiens ADAM metallopeptidase domain 17 (ADAM17), mRNA [NM_003183]	1.71	0.0001	1.67	0.0002	1.71	0.0002	1.63	0.0004
TEX15	Homo sapiens testis expressed 15 (TEX15), mRNA [NM_031271]	1.71	0.0002	1.54	0.0005	2.22	0.0004	1.94	0.0006
SGTB	Homo sapiens small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta (SGTB), mRNA [NM_019072]	1.73	0.0000	1.84	0.0000	1.79	0.0001	1.69	0.0003
YWHAEP1	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon pseudogene 1 (YWHAEP1), non-coding RNA [NR_029404]	1.73	0.0004	1.63	0.0004	1.72	0.0001	1.69	0.0002
CEP19	Homo sapiens centrosomal protein 19kDa (CEP19), mRNA [NM_032898]	1.76	0.0004	1.74	0.0001	1.79	0.0003	1.91	0.0002

FAM86FP	Homo sapiens family with sequence similarity 86, member F, pseudogene (FAM86FP), non-coding RNA [NR_024254]	1.76	0.0001	2.40	0.0000	1.91	0.0001	2.61	0.0001
TMTC3	Homo sapiens transmembrane and tetratricopeptide repeat containing 3 (TMTC3), mRNA [NM_181783]	1.77	0.0000	1.85	0.0000	1.91	0.0001	1.86	0.0004
FBXO33	Homo sapiens F-box protein 33 (FBXO33), mRNA [NM_203301]	1.78	0.0003	1.57	0.0003	1.81	0.0002	1.55	0.0007
MRPL22	Homo sapiens mitochondrial ribosomal protein L22 (MRPL22), transcript variant 2, mRNA [NM_001014990]	1.78	0.0001	1.69	0.0002	1.73	0.0002	1.65	0.0010
RPAP3	Homo sapiens RNA polymerase II associated protein 3 (RPAP3), transcript variant 1, mRNA [NM_024604]	1.78	0.0003	1.73	0.0001	1.84	0.0004	1.76	0.0010
LAP3	Homo sapiens leucine aminopeptidase 3 (LAP3), mRNA [NM_015907]	1.78	0.0005	1.95	0.0000	1.85	0.0002	1.92	0.0004
DNAJA3	Homo sapiens DnaJ (Hsp40) homolog, subfamily A, member 3 (DNAJA3), transcript variant 1, mRNA [NM_005147]	1.80	0.0000	1.63	0.0001	1.82	0.0001	1.65	0.0005
TMTC3	Homo sapiens transmembrane and tetratricopeptide repeat containing 3 (TMTC3), mRNA [NM_181783]	1.80	0.0003	1.86	0.0000	2.21	0.0003	2.27	0.0002
PDCD6IP	Homo sapiens programmed cell death 6 interacting protein (PDCD6IP), transcript variant 4, mRNA [NM_001256192]	1.81	0.0001	1.94	0.0002	1.76	0.0003	1.80	0.0007
TMEM248	Homo sapiens transmembrane protein 248 (TMEM248), mRNA [NM_017994]	1.82	0.0000	1.81	0.0000	1.87	0.0001	1.69	0.0003
PPP2R3C	Homo sapiens protein phosphatase 2, regulatory subunit B", gamma (PPP2R3C), mRNA [NM_017917]	1.84	0.0000	1.76	0.0000	1.85	0.0001	1.71	0.0007
FAM135A	Homo sapiens family with sequence similarity 135, member A (FAM135A), transcript variant 2, mRNA [NM_020819]	1.84	0.0000	2.09	0.0000	2.04	0.0001	1.87	0.0008
PPAT	Homo sapiens phosphoribosyl pyrophosphate amidotransferase (PPAT), mRNA [NM_002703]	1.85	0.0000	1.80	0.0001	1.90	0.0002	1.88	0.0005
SUV39H2	Homo sapiens suppressor of variegation 3-9 homolog 2 (Drosophila) (SUV39H2), transcript variant 3, mRNA [NM_024670]	1.85	0.0003	2.08	0.0000	1.69	0.0004	1.96	0.0004

WDR82	Homo sapiens WD repeat domain 82 (WDR82), mRNA [NM_025222]	1.86	0.0005	1.84	0.0003	1.83	0.0003	1.69	0.0008
C1GALT1	core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:24337] [ENST00000402468]	1.86	0.0001	1.85	0.0001	1.95	0.0002	1.83	0.0006
SMIM13	Homo sapiens small integral membrane protein 13 (SMIM13), mRNA [NM_001135575]	1.86	0.0003	2.08	0.0000	2.13	0.0000	2.06	0.0001
NRAS	Homo sapiens neuroblastoma RAS viral (v-ras) oncogene homolog (NRAS), mRNA [NM_002524]	1.86	0.0001	2.19	0.0000	1.86	0.0001	2.08	0.0001
DUS2	Homo sapiens dihydrouridine synthase 2 (DUS2), transcript variant 1, mRNA [NM_017803]	1.87	0.0000	1.71	0.0002	1.84	0.0002	1.91	0.0001
MFAP3	Homo sapiens microfibrillar-associated protein 3 (MFAP3), transcript variant 3, mRNA [NM_001242336]	1.87	0.0008	1.94	0.0001	1.91	0.0000	1.92	0.0001
RASSF1	Homo sapiens Ras association (RalGDS/AF-6) domain family member 1 (RASSF1), transcript variant C, mRNA [NM_170713]	1.89	0.0007	1.71	0.0002	1.78	0.0003	1.63	0.0009
C5orf15	Homo sapiens chromosome 5 open reading frame 15 (C5orf15), mRNA [NM_020199]	1.91	0.0001	1.63	0.0002	1.96	0.0001	1.61	0.0007
ETV1	Homo sapiens ets variant 1 (ETV1), transcript variant 1, mRNA [NM_004956]	1.92	0.0002	1.67	0.0002	1.77	0.0008	1.82	0.0005
RAP1B	Homo sapiens RAP1B, member of RAS oncogene family (RAP1B), transcript variant 1, mRNA [NM_015646]	1.93	0.0005	1.68	0.0005	1.89	0.0001	1.55	0.0007
GMPR2	Homo sapiens guanosine monophosphate reductase 2 (GMPR2), transcript variant 2, mRNA [NM_001002000]	1.93	0.0002	1.79	0.0001	1.91	0.0002	1.77	0.0005
C1GALT1	Homo sapiens core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 (C1GALT1), mRNA [NM_020156]	1.93	0.0002	1.88	0.0001	2.04	0.0001	1.97	0.0002
NUP98	Homo sapiens nucleoporin 98kDa (NUP98), transcript variant 3, mRNA [NM_005387]	1.93	0.0004	2.27	0.0000	1.84	0.0002	2.20	0.0003

CAB39	Homo sapiens calcium binding protein 39 (CAB39), transcript variant 1, mRNA [NM_016289]	1.94	0.0000	1.75	0.0001	2.12	0.0000	1.62	0.0008
ATP6V1E2	Homo sapiens ATPase, H+ transporting, lysosomal 31kDa, V1 subunit E2 (ATP6V1E2), mRNA [NM_080653]	1.94	0.0005	1.83	0.0002	1.81	0.0004	1.77	0.0010
MTMR2	Homo sapiens myotubularin related protein 2 (MTMR2), transcript variant 2, mRNA [NM_201278]	1.96	0.0000	1.95	0.0000	2.01	0.0001	1.87	0.0006
HNRNPR	Homo sapiens heterogeneous nuclear ribonucleoprotein R (HNRNPR), transcript variant 2, mRNA [NM_005826]	1.97	0.0008	1.95	0.0002	1.98	0.0002	1.84	0.0007
BTG3	Homo sapiens BTG family, member 3 (BTG3), transcript variant 2, mRNA [NM_006806]	1.97	0.0001	2.00	0.0001	2.05	0.0001	2.04	0.0001
BRCC3	Homo sapiens BRCA1/BRCA2-containing complex, subunit 3 (BRCC3), transcript variant 1, mRNA [NM_024332]	1.99	0.0000	1.77	0.0000	1.86	0.0001	1.77	0.0006
DDAH1	Homo sapiens dimethylarginine dimethylaminohydrolase 1 (DDAH1), transcript variant 1, mRNA [NM_012137]	1.99	0.0000	4.38	0.0000	2.10	0.0001	4.37	0.0000
RB1	Homo sapiens retinoblastoma 1 (RB1), mRNA [NM_000321]	2.01	0.0001	1.63	0.0003	1.97	0.0001	1.60	0.0007
ID2	Homo sapiens inhibitor of DNA binding 2, dominant negative helix-loop-helix protein (ID2), mRNA [NM_002166]	2.01	0.0001	2.58	0.0000	2.30	0.0000	2.48	0.0000
ENTPD4	Homo sapiens ectonucleoside triphosphate diphosphohydrolase 4 (ENTPD4), transcript variant 1, mRNA [NM_004901]	2.02	0.0007	1.90	0.0001	1.98	0.0002	1.83	0.0003
AIDA	Homo sapiens axin interactor, dorsalization associated (AIDA), mRNA [NM_022831]	2.02	0.0004	2.19	0.0000	2.05	0.0001	2.07	0.0001
KDELR2	Homo sapiens KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2 (KDELR2), transcript variant 1, mRNA [NM_006854]	2.02	0.0000	2.37	0.0000	2.03	0.0000	2.30	0.0001
CLDN1	Homo sapiens claudin 1 (CLDN1), mRNA [NM_021101]	2.02	0.0000	1.99	0.0001	1.78	0.0004	2.31	0.0000

BRCC3	Homo sapiens BRCA1/BRCA2-containing complex, subunit 3 (BRCC3), transcript variant 2, mRNA [NM_001018055]	2.03	0.0000	1.78	0.0001	1.92	0.0002	1.78	0.0003
MAP2K1	Homo sapiens mitogen-activated protein kinase kinase 1 (MAP2K1), mRNA [NM_002755]	2.03	0.0000	1.86	0.0000	2.24	0.0000	1.84	0.0001
ATP6V1A	Homo sapiens ATPase, H+ transporting, lysosomal 70kDa, V1 subunit A (ATP6V1A), mRNA [NM_001690]	2.05	0.0001	1.81	0.0000	1.99	0.0000	1.74	0.0005
PROSC	Homo sapiens proline synthetase co-transcribed homolog (bacterial) (PROSC), mRNA [NM_007198]	2.06	0.0004	2.07	0.0001	1.90	0.0007	1.98	0.0005
IPPK	Homo sapiens inositol 1,3,4,5,6-pentakisphosphate 2-kinase (IPPK), mRNA [NM_022755]	2.07	0.0000	1.75	0.0004	2.13	0.0000	1.75	0.0007
SLC25A44	Homo sapiens solute carrier family 25, member 44 (SLC25A44), transcript variant 2, mRNA [NM_014655]	2.08	0.0009	2.04	0.0002	2.18	0.0001	2.02	0.0004
SFXN4	Homo sapiens sideroflexin 4 (SFXN4), transcript variant 1, mRNA [NM_213649]	2.08	0.0000	2.40	0.0000	1.96	0.0001	2.31	0.0000
ZBTB9	Homo sapiens zinc finger and BTB domain containing 9 (ZBTB9), mRNA [NM_152735]	2.09	0.0001	1.99	0.0000	2.18	0.0000	1.91	0.0001
PPP2R1B	Homo sapiens protein phosphatase 2, regulatory subunit A, beta (PPP2R1B), transcript variant 2, mRNA [NM_181699]	2.11	0.0000	1.74	0.0001	2.04	0.0001	1.78	0.0004
PRRC2C	Homo sapiens proline-rich coiled-coil 2C (PRRC2C), mRNA [NM_015172]	2.11	0.0000	2.09	0.0000	2.16	0.0000	1.99	0.0001
TMEM64	Homo sapiens transmembrane protein 64 (TMEM64), transcript variant 1, mRNA [NM_001008495]	2.12	0.0003	2.04	0.0001	2.51	0.0000	2.12	0.0003
CDC23	Homo sapiens cell division cycle 23 (CDC23), mRNA [NM_004661]	2.15	0.0000	1.72	0.0001	2.21	0.0000	1.71	0.0004
RBSN	Homo sapiens rabenosyn, RAB effector (RBSN), transcript variant 1, mRNA [NM_022340]	2.15	0.0001	1.92	0.0000	2.14	0.0001	1.89	0.0003
ENTPD4	Homo sapiens ectonucleoside triphosphate diphosphohydrolase 4 (ENTPD4), transcript variant 1, mRNA [NM_004901]	2.15	0.0010	1.98	0.0001	2.06	0.0004	1.95	0.0002

DERL1	Homo sapiens derlin 1 (DERL1), transcript variant 1, mRNA [NM_024295]	2.17	0.0000	1.72	0.0001	2.16	0.0000	1.65	0.0005
UMPS	Homo sapiens uridine monophosphate synthetase (UMPS), transcript variant 1, mRNA [NM_000373]	2.18	0.0001	2.32	0.0001	2.07	0.0001	2.29	0.0002
MBD1	Homo sapiens methyl-CpG binding domain protein 1 (MBD1), transcript variant 11, mRNA [NM_001204142]	2.19	0.0000	1.83	0.0001	2.10	0.0001	1.76	0.0003
HS2ST1	Homo sapiens heparan sulfate 2-O-sulfotransferase 1 (HS2ST1), transcript variant 1, mRNA [NM_012262]	2.20	0.0001	1.64	0.0001	2.23	0.0001	1.64	0.0007
EFCAB14	Homo sapiens EF-hand calcium binding domain 14 (EFCAB14), mRNA [NM_014774]	2.21	0.0001	1.95	0.0000	2.38	0.0001	1.90	0.0007
LOXL2	Homo sapiens lysyl oxidase-like 2 (LOXL2), mRNA [NM_002318]	2.23	0.0000	2.06	0.0000	2.31	0.0000	1.94	0.0001
ZNF788	Homo sapiens zinc finger family member 788 (ZNF788), non-coding RNA [NR_027049]	2.23	0.0003	2.13	0.0000	2.23	0.0001	2.28	0.0002
MLX	Homo sapiens MLX, MAX dimerization protein (MLX), transcript variant 3, mRNA [NM_170607]	2.24	0.0000	1.61	0.0001	2.10	0.0000	1.55	0.0006
MMS22L	Homo sapiens MMS22-like, DNA repair protein (MMS22L), mRNA [NM_198468]	2.24	0.0003	1.70	0.0006	2.33	0.0000	1.71	0.0009
ANXA7	Homo sapiens annexin A7 (ANXA7), transcript variant 2, mRNA [NM_004034]	2.26	0.0001	1.87	0.0001	2.17	0.0000	1.76	0.0005
MFAP3	Homo sapiens microfibrillar-associated protein 3 (MFAP3), transcript variant 3, mRNA [NM_001242336]	2.26	0.0000	2.54	0.0000	2.06	0.0001	2.32	0.0001
MAPRE1	Homo sapiens microtubule-associated protein, RP/EB family, member 1 (MAPRE1), mRNA [NM_012325]	2.28	0.0002	1.87	0.0001	2.29	0.0000	1.75	0.0009
ANKRD13C	Homo sapiens ankyrin repeat domain 13C (ANKRD13C), mRNA [NM_030816]	2.28	0.0002	1.77	0.0002	2.55	0.0000	1.79	0.0005
SLC31A1	Homo sapiens solute carrier family 31 (copper transporter), member 1 (SLC31A1), mRNA [NM_001859]	2.32	0.0001	2.32	0.0000	2.03	0.0000	2.19	0.0001
AXIN2	Homo sapiens axin 2 (AXIN2), mRNA [NM_004655]	2.33	0.0004	1.97	0.0001	2.65	0.0002	2.32	0.0006

1										
2										
3	YWHAG	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma (YWHAG), mRNA [NM_012479]	2.34	0.0001	3.11	0.0000	2.15	0.0000	2.93	0.0000
4										
5										
6	SLC25A13	Homo sapiens solute carrier family 25 (aspartate/glutamate carrier), member 13 (SLC25A13), transcript variant 2, mRNA [NM_014251]	2.35	0.0000	1.71	0.0001	2.46	0.0000	1.70	0.0004
7										
8										
9										
10	PVRL2	Homo sapiens poliovirus receptor-related 2 (herpesvirus entry mediator B) (PVRL2), transcript variant delta, mRNA [NM_001042724]	2.35	0.0001	1.91	0.0001	2.30	0.0000	1.84	0.0002
11										
12										
13										
14	EIF2A	Homo sapiens eukaryotic translation initiation factor 2A, 65kDa (EIF2A), mRNA [NM_032025]	2.36	0.0004	1.75	0.0003	2.05	0.0001	1.63	0.0005
15										
16										
17	UBLCP1	Homo sapiens ubiquitin-like domain containing CTD phosphatase 1 (UBLCP1), mRNA [NM_145049]	2.36	0.0000	2.19	0.0000	2.38	0.0000	2.21	0.0003
18										
19										
20	GPD1L	Homo sapiens glycerol-3-phosphate dehydrogenase 1-like (GPD1L), mRNA [NM_015141]	2.37	0.0001	2.04	0.0000	2.23	0.0002	2.24	0.0001
21										
22										
23	CTNNB1	Homo sapiens catenin (cadherin-associated protein), beta 1, 88kDa (CTNNB1), transcript variant 1, mRNA [NM_001904]	2.37	0.0000	2.42	0.0000	2.49	0.0000	2.50	0.0000
24										
25										
26	LOC_I2_01091	BROAD Institute lincRNA (XLOC_I2_010916), lincRNA [TCONS_I2_00020852]	2.38	0.0004	2.50	0.0000	2.34	0.0001	2.24	0.0001
27										
28										
29	CUTC	Homo sapiens cutC copper transporter (CUTC), mRNA [NM_015960]	2.39	0.0001	2.10	0.0001	1.96	0.0001	2.05	0.0002
30										
31	ARL8B	Homo sapiens ADP-ribosylation factor-like 8B (ARL8B), mRNA [NM_018184]	2.41	0.0001	1.78	0.0001	2.56	0.0000	1.76	0.0003
32										
33										
34	BCAP29	Homo sapiens B-cell receptor-associated protein 29 (BCAP29), transcript variant 2, mRNA [NM_018844]	2.41	0.0002	2.63	0.0000	2.71	0.0000	2.42	0.0001
35										
36	PHTF2	Homo sapiens putative homeodomain transcription factor 2 (PHTF2), transcript variant 2, mRNA [NM_020432]	2.44	0.0000	2.07	0.0001	2.60	0.0001	2.07	0.0002
37										
38										
39	PPIF	Homo sapiens peptidylprolyl isomerase F (PPIF), mRNA [NM_005729]	2.44	0.0000	2.13	0.0000	2.59	0.0000	2.15	0.0005
40										
41										
42										
43										
44										
45										
46										

RAB8B	Homo sapiens RAB8B, member RAS oncogene family (RAB8B), mRNA [NM_016530]	2.45	0.0002	2.79	0.0000	2.43	0.0000	2.85	0.0000
SLC39A14	Homo sapiens solute carrier family 39 (zinc transporter), member 14 (SLC39A14), transcript variant 2, mRNA [NM_015359]	2.45	0.0000	2.78	0.0000	3.64	0.0000	2.85	0.0000
MRPL19	Homo sapiens mitochondrial ribosomal protein L19 (MRPL19), mRNA [NM_014763]	2.46	0.0002	2.65	0.0000	2.34	0.0002	2.71	0.0001
SYPL1	Homo sapiens synaptophysin-like 1 (SYPL1), transcript variant 2, mRNA [NM_182715]	2.47	0.0000	2.46	0.0000	2.55	0.0000	2.32	0.0001
ERLIN1	Homo sapiens ER lipid raft associated 1 (ERLIN1), mRNA [NM_006459]	2.49	0.0001	2.28	0.0000	2.30	0.0001	2.36	0.0002
ANKRD44	Homo sapiens ankyrin repeat domain 44 (ANKRD44), transcript variant B, mRNA [NM_153697]	2.52	0.0001	2.13	0.0000	2.37	0.0000	1.87	0.0004
CXorf56	Homo sapiens chromosome X open reading frame 56 (CXorf56), transcript variant 1, mRNA [NM_022101]	2.54	0.0003	2.38	0.0000	2.42	0.0002	2.47	0.0002
GCNT3	Homo sapiens glucosaminyl (N-acetyl) transferase 3, mucin type (GCNT3), mRNA [NM_004751]	2.56	0.0000	1.64	0.0002	2.22	0.0000	1.77	0.0002
SWAP70	Homo sapiens SWAP switching B-cell complex 70kDa subunit (SWAP70), transcript variant 1, mRNA [NM_015055]	2.57	0.0005	3.20	0.0000	2.94	0.0001	3.16	0.0000
ZAK	Homo sapiens sterile alpha motif and leucine zipper containing kinase AZK (ZAK), transcript variant 1, mRNA [NM_016653]	2.59	0.0001	1.71	0.0004	2.59	0.0000	1.88	0.0002
RNF14	Homo sapiens ring finger protein 14 (RNF14), transcript variant 1, mRNA [NM_004290]	2.59	0.0001	2.58	0.0000	2.48	0.0001	2.19	0.0004
AK2	Homo sapiens adenylate kinase 2 (AK2), transcript variant 2, mRNA [NM_013411]	2.63	0.0003	4.30	0.0000	2.78	0.0000	4.00	0.0000
ZCCHC2	Homo sapiens zinc finger, CCHC domain containing 2 (ZCCHC2), transcript variant 1, mRNA [NM_017742]	2.67	0.0001	1.63	0.0005	2.62	0.0000	1.55	0.0009

NAPG	Homo sapiens N-ethylmaleimide-sensitive factor attachment protein, gamma (NAPG), mRNA [NM_003826]	2.67	0.0000	2.07	0.0000	2.70	0.0000	1.99	0.0001
GNA13	Homo sapiens guanine nucleotide binding protein (G protein), alpha 13 (GNA13), transcript variant 1, mRNA [NM_006572]	2.67	0.0001	2.61	0.0000	2.74	0.0001	2.49	0.0001
LOC101926940	Homo sapiens uncharacterized LOC101926940 (LOC101926940), long non-coding RNA [NR_104632]	2.70	0.0001	2.17	0.0000	2.51	0.0000	2.36	0.0002
HDAC1	Homo sapiens histone deacetylase 1 (HDAC1), mRNA [NM_004964]	2.71	0.0001	2.19	0.0001	3.33	0.0000	2.07	0.0005
CTNNB1	Homo sapiens catenin (cadherin-associated protein), beta 1, 88kDa (CTNNB1), transcript variant 3, mRNA [NM_001098210]	2.71	0.0000	2.47	0.0000	2.61	0.0000	2.34	0.0000
RAP1GDS1	Homo sapiens RAP1, GTP-GDP dissociation stimulator 1 (RAP1GDS1), transcript variant 1, mRNA [NM_001100426]	2.74	0.0001	1.82	0.0002	2.71	0.0000	1.66	0.0006
B4GALT4	Homo sapiens UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 4 (B4GALT4), transcript variant 1, mRNA [NM_212543]	2.74	0.0000	2.10	0.0000	2.56	0.0000	2.07	0.0002
IKBIP	Homo sapiens IKBKB interacting protein (IKBIP), transcript variant 1, mRNA [NM_153687]	2.75	0.0002	3.02	0.0000	2.86	0.0001	2.79	0.0001
TAF11	Homo sapiens TAF11 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 28kDa (TAF11), transcript variant 1, mRNA [NM_005643]	2.81	0.0000	3.24	0.0000	2.56	0.0000	3.21	0.0000
HINT3	Homo sapiens histidine triad nucleotide binding protein 3 (HINT3), mRNA [NM_138571]	2.83	0.0004	3.35	0.0000	2.65	0.0000	3.61	0.0000
TCF7	Homo sapiens transcription factor 7 (T-cell specific, HMG-box) (TCF7), transcript variant 1, mRNA [NM_003202]	2.85	0.0002	1.77	0.0002	4.05	0.0000	1.75	0.0002
RTCA	Homo sapiens RNA 3'-terminal phosphate cyclase (RTCA), transcript variant 2, mRNA [NM_003729]	2.88	0.0001	3.11	0.0000	2.78	0.0000	3.05	0.0000

GGACT	Homo sapiens gamma-glutamylamine cyclotransferase (GGACT), transcript variant 1, mRNA [NM_033110]	2.90	0.0006	3.80	0.0000	2.24	0.0005	3.45	0.0000
HAUS6	Homo sapiens HAUS augmin-like complex, subunit 6 (HAUS6), transcript variant 1, mRNA [NM_017645]	2.91	0.0001	2.06	0.0001	2.72	0.0000	2.01	0.0004
RNF14	Homo sapiens ring finger protein 14 (RNF14), transcript variant 1, mRNA [NM_004290]	2.92	0.0000	2.92	0.0000	2.90	0.0000	2.65	0.0000
CXorf56	Homo sapiens chromosome X open reading frame 56 (CXorf56), transcript variant 1, mRNA [NM_022101]	2.94	0.0002	2.79	0.0000	2.82	0.0000	2.74	0.0000
KLHL2	Homo sapiens kelch-like family member 2 (KLHL2), transcript variant 1, mRNA [NM_007246]	2.96	0.0001	2.32	0.0001	2.79	0.0000	2.11	0.0002
GNA13	Homo sapiens guanine nucleotide binding protein (G protein), alpha 13 (GNA13), transcript variant 1, mRNA [NM_006572]	2.98	0.0002	2.91	0.0000	3.35	0.0000	2.88	0.0000
RELL1	Homo sapiens RELT-like 1 (RELL1), transcript variant 1, mRNA [NM_001085400]	3.00	0.0005	2.37	0.0003	2.60	0.0001	2.10	0.0008
FECH	Homo sapiens ferrochelatase (FECH), transcript variant 1, mRNA [NM_001012515]	3.04	0.0003	3.26	0.0000	2.53	0.0003	2.89	0.0002
EIF2S1	Homo sapiens eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa (EIF2S1), mRNA [NM_004094]	3.06	0.0000	2.81	0.0000	3.10	0.0000	2.81	0.0001
NARF	Homo sapiens nuclear prelamin A recognition factor (NARF), transcript variant 3, mRNA [NM_001038618]	3.07	0.0000	3.34	0.0000	2.81	0.0000	3.02	0.0000
NPTN	Homo sapiens neuroplastin (NPTN), transcript variant b, mRNA [NM_012428]	3.08	0.0003	2.60	0.0000	2.93	0.0001	2.46	0.0003
LOC101059954	PREDICTED: Homo sapiens uncharacterized LOC101059954 (LOC101059954), ncRNA [XR_172435]	3.09	0.0004	2.06	0.0000	2.41	0.0008	1.99	0.0007
ZAK	Homo sapiens sterile alpha motif and leucine zipper containing kinase AZK (ZAK), transcript variant 1, mRNA [NM_016653]	3.17	0.0002	2.05	0.0000	2.76	0.0001	1.99	0.0004

1	LOC102724529	PREDICTED: Homo sapiens uncharacterized LOC102724529 (LOC102724529), ncRNA [XR_425227]	3.18	0.0004	2.63	0.0000	3.46	0.0000	2.73	0.0001
2	C12orf76	Homo sapiens chromosome 12 open reading frame 76 (C12orf76), mRNA [NM_207435]	3.22	0.0000	2.99	0.0000	2.99	0.0001	2.75	0.0001
3	CFL2	Homo sapiens cofilin 2 (muscle) (CFL2), transcript variant 1, mRNA [NM_021914]	3.22	0.0003	3.58	0.0000	3.01	0.0001	3.52	0.0001
4	CYHR1	cysteine/histidine-rich 1 [Source:HGNC Symbol;Acc:HGNC:17806] [ENST00000528663]	3.24	0.0001	2.60	0.0000	3.22	0.0000	2.28	0.0001
5	SCML1	Homo sapiens sex comb on midleg-like 1 (Drosophila) (SCML1), transcript variant 1, mRNA [NM_001037540]	3.30	0.0001	1.82	0.0001	3.59	0.0000	1.94	0.0003
6	SEC24A	Homo sapiens SEC24 family member A (SEC24A), transcript variant 1, mRNA [NM_021982]	3.30	0.0000	2.55	0.0000	3.14	0.0000	2.37	0.0000
7	ETF1	Homo sapiens eukaryotic translation termination factor 1 (ETF1), transcript variant 1, mRNA [NM_004730]	3.30	0.0000	3.16	0.0000	3.39	0.0000	3.14	0.0000
8	nd	nd	3.41	0.0003	3.50	0.0000	3.45	0.0000	3.33	0.0000
9	ITGA6	Homo sapiens integrin, alpha 6 (ITGA6), transcript variant 2, mRNA [NM_000210]	3.44	0.0002	2.97	0.0000	3.10	0.0000	3.14	0.0000
10	SLC31A1	Homo sapiens solute carrier family 31 (copper transporter), member 1 (SLC31A1), mRNA [NM_001859]	3.44	0.0001	3.45	0.0000	3.49	0.0000	3.56	0.0000
11	PMM2	Homo sapiens phosphomannomutase 2 (PMM2), mRNA [NM_000303]	3.54	0.0000	4.06	0.0000	3.46	0.0000	3.63	0.0000
12	ITGA6	Homo sapiens integrin, alpha 6 (ITGA6), transcript variant 2, mRNA [NM_000210]	3.68	0.0001	3.23	0.0000	3.47	0.0000	3.19	0.0000
13	UBE2V1	Homo sapiens ubiquitin-conjugating enzyme E2 variant 1 (UBE2V1), transcript variant 4, mRNA [NM_001032288]	3.81	0.0008	4.10	0.0000	3.99	0.0001	3.82	0.0002
14	PGM2L1	Homo sapiens phosphoglucomutase 2-like 1 (PGM2L1), mRNA [NM_173582]	3.83	0.0000	3.48	0.0000	4.47	0.0000	2.87	0.0000
15	STBD1	Homo sapiens starch binding domain 1 (STBD1), mRNA [NM_003943]	3.85	0.0000	2.49	0.0000	3.97	0.0000	2.21	0.0000
16	IL1RAP	Homo sapiens interleukin 1 receptor accessory protein (IL1RAP), transcript variant 1, mRNA [NM_002182]	3.88	0.0002	2.08	0.0001	4.13	0.0000	2.02	0.0002

CIDECP	Homo sapiens cell death-inducing DFFA-like effector c pseudogene (CIDECP), non-coding RNA [NR_002786]	3.89	0.0000	2.33	0.0000	3.44	0.0000	2.39	0.0000
RDH11	Homo sapiens retinol dehydrogenase 11 (all-trans/9-cis/11-cis) (RDH11), transcript variant 1, mRNA [NM_016026]	3.90	0.0000	3.58	0.0000	3.92	0.0000	3.09	0.0000
RP2	Homo sapiens retinitis pigmentosa 2 (X-linked recessive) (RP2), mRNA [NM_006915]	3.98	0.0001	3.47	0.0000	3.80	0.0000	3.35	0.0000
RDX	Homo sapiens radixin (RDX), transcript variant 3, mRNA [NM_002906]	4.05	0.0002	2.51	0.0002	3.91	0.0000	2.26	0.0007
CIDEC	Homo sapiens cell death-inducing DFFA-like effector c (CIDEC), transcript variant 3, mRNA [NM_022094]	4.07	0.0002	2.58	0.0000	4.09	0.0000	2.52	0.0001
EMC10	Homo sapiens ER membrane protein complex subunit 10 (EMC10), transcript variant 1, mRNA [NM_175063]	4.22	0.0000	3.96	0.0000	3.90	0.0000	3.75	0.0000
CIDEC	Homo sapiens cell death-inducing DFFA-like effector c (CIDEC), transcript variant 3, mRNA [NM_022094]	4.24	0.0002	2.75	0.0000	3.59	0.0000	2.68	0.0001
ERLIN2	Homo sapiens ER lipid raft associated 2 (ERLIN2), transcript variant 1, mRNA [NM_007175]	4.44	0.0000	2.81	0.0000	4.12	0.0000	2.73	0.0001
SNRPC	Homo sapiens small nuclear ribonucleoprotein polypeptide C (SNRPC), transcript variant 1, mRNA [NM_003093]	4.45	0.0001	4.42	0.0000	4.49	0.0000	4.32	0.0000
NUP35	Homo sapiens nucleoporin 35kDa (NUP35), transcript variant 4, mRNA [NM_001287585]	4.47	0.0001	2.98	0.0000	4.47	0.0000	2.88	0.0001
LMBR1	Homo sapiens limb development membrane protein 1 (LMBR1), mRNA [NM_022458]	4.53	0.0002	3.78	0.0000	4.39	0.0000	3.47	0.0001
COMMD10	Homo sapiens COMM domain containing 10 (COMMD10), mRNA [NM_016144]	4.62	0.0001	3.50	0.0000	4.16	0.0000	3.07	0.0001
NUP35	Homo sapiens nucleoporin 35kDa (NUP35), transcript variant 1, mRNA [NM_138285]	4.69	0.0000	3.10	0.0000	4.77	0.0000	3.16	0.0000

1										
2										
3	UBE2V1	Homo sapiens ubiquitin-conjugating enzyme E2 variant 1 (UBE2V1), transcript variant 4, mRNA [NM_001032288]	4.84	0.0000	4.26	0.0000	4.95	0.0000	4.09	0.0000
4										
5										
6	FZD5	Homo sapiens frizzled class receptor 5 (FZD5), mRNA [NM_003468]	5.53	0.0000	3.54	0.0000	6.18	0.0000	3.53	0.0000
7										
8	ASPH	Homo sapiens aspartate beta-hydroxylase (ASPH), transcript variant 3, mRNA [NM_032466]	9.31	0.0000	9.45	< 1e-07	10.18	0.0000	10.90	0.0000
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										
27										
28										
29										
30										
31										
32										
33										
34										
35										
36										
37										
38										
39										
40										
41										
42										
43										
44										
45										
46										

For Peer Review

Titre : Rôle et pertinence clinique du facteur de transcription FOXS1 dans le carcinome hépatocellulaire

Mots clés : Carcinome hépatocellulaire, FOX, TGF β , transition épithélio-mésenchymateuse, métastase

Résumé : Les cancers du foie sont les quatrièmes cancers les plus mortels dans le monde. Parmi eux, le carcinome hépatocellulaire (CHC) est la tumeur primitive la plus fréquente. Il n'existe pas de stratégies thérapeutiques curatives contre ce cancer excepté la chirurgie. Le CHC se développe essentiellement sur des foies cirrhotiques. Dans ce contexte, la cytokine Transforming Growth Factor β (TGF β) joue un rôle important dans le développement du cancer. Toutefois, son rôle est complexe puisque il présente une action pro-cytostatique dans les stades précoces mais une action pro-métastatique dans les stades avancés du cancer. Les mécanismes moléculaires entraînant ce changement d'action ne sont pas encore élucidés. L'identification de nouveaux effecteurs du TGF β pourrait permettre d'expliquer cette dualité fonctionnelle.

Dans cette optique, j'ai identifié *FOXS1*, un nouveau gène cible du TGF β appartenant à la famille des *forkhead box* (FOX) connue pour être impliquée dans différents cancers. Mes données montrent que *FOXS1* est un facteur de transcription nucléaire induit par la signalisation canonique des SMAD. Il induit la transition épithélio-mésenchymateuse par l'inhibition de protéines de jonction cellulaire et induction indirecte de *SNAI1/2*. Il façonne de plus la réponse cellulaire au TGF β en inhibant l'expression de *SMAD7*, le rétrocontrôle de la signalisation TGF β . Dans les tumeurs, l'expression de *FOXS1* est un marqueur de mauvais pronostic, associée à une survie plus faible dans les cancers du foie, du rein, de l'estomac et du côlon.

Title : Role and clinical relevance of the transcription factor FOXS1 in hepatocellular carcinoma

Keywords : Hepatocellular carcinoma, FOX, TGF β , epithelial-to-mesenchymal transition, metastasis

Abstract : Liver cancers is the fourth leading cause of cancer-related death worldwide. Among them, hepatocellular carcinoma (HCC) is the most common liver primary tumor. However, there is no curative therapy except surgery against this cancer. HCC usually occurs in patient with underlying chronic liver disease leading to liver fibrosis. In this context, the cytokine Transforming Growth Factor β (TGF β) plays an important role in HCC development. However TGF β action in HCC is complex because it exhibits anti-tumorigenic properties in early stage of cancer and oncogenic properties in advanced stage. The mechanisms underlining this switch of action of TGF β are yet not fully understand. Identification of new factors involved in the TGF β pathway could explain this functional duality.

In this regard, I identified the transcription factor *FOXS1* as a new target gene of the TGF β pathway. It belongs to the forkhead box (FOX) family of which many members are involved in cancer, acting as oncogenes or tumor suppressors. My data show that *FOXS1* is a transcription factor located in the nucleus and induced through the canonical SMAD pathway. It is a potent inducer of the epithelial-to-mesenchymal transition notably through the inhibition of junction proteins expression and the indirect induction of *SNAI1/2*. Moreover, it impacts the TGF β pathway, especially through the inhibition of *SMAD7* expression, the negative feedback effector of the pathway. In human tumor, *FOXS1* expression is a marker of poor prognosis, associated with a poor survival in patients with liver, kidney, stomach and colon cancer.