

# THÈSE

en vue de l'obtention du grade de

**Docteur de l'Université de Lyon, délivré par l'École Normale Supérieure de Lyon**

**Discipline : Sciences de la Vie**

**Laboratoire : Centre Internationale de la Recherche sur le Cancer – Groupe Epigénétique**

**École Doctorale Biologie Moléculaire, Intégrative et Cellulaire ED 340**

présentée et soutenue publiquement le 6 décembre 2013

par Madame Marion ESSIG épouse MARTIN

## **Analyse de la méthylation de l'ADN des cellules CD133+ dans le cancer du foie et son interaction avec la voie de signalisation TGF- $\beta$**

*Directeur de thèse : Dr Zdenko HERCEG*

*Après l'avis de : Dr Jorg TOST*

*Dr Ilaria MALANCHI*

*Devant la commission d'examen formée de :*

*Dr Pierre-Antoine DEFOSSEZ (examineur)*

*Pr Christine DELPRAT (Présidente du Jury)*

*Dr Zdenko HERCEG (Directeur de thèse)*

*Dr Ilaria MALANCHI (Rapporteuse)*

*Dr Jorg TOST (Rapporteur)*



*A Jérôme, Gaëlle, Bertie  
et les futures petites crapules qui suivront*

*A Papipe*





## Remerciements

Au **Professeur Christine Delprat**, pour avoir immédiatement accepté de faire partie de mon jury, de le présider et pour avoir pris le temps de discuter de mon travail.

To **Dr Tost and Dr Malanchi** thank you for having accepted without any hesitations to be part of my jury, and for taking the time to review and discuss with me about my project.

Au **Dr Pierre-Antoine Defossez** pour avoir accepté non seulement d'examiner mon travail mais également pour avoir pris le temps il y a 6 ans de me faire découvrir l'épigénétique. Je garde un vraiment bon souvenir de ce tout premier stage.

To **Dr Zdenko Herceg**, thank you for having accepted me in the EGE team first for my master and then as a PhD student. I really enjoyed the 4 years that I spend in the epigenetic group not only at a professional level (with all the discussions during lab meetings and seminars that helped me to move my research forward) but also at personal level with all the outings, dinners and coffee breaks. These 4 years in the same team have been a great first professional experience where I did not only learn how to work but where I also had the chance to create warm relationships with my colleagues.

Un grand merci à **Cyrille et Marie-Pierre**, vous faites vraiment partie des piliers de ce groupe et grâce à votre accueil, vos conseils et au temps que vous avez pris pour moi j'ai pu au mieux m'intégrer dans l'équipe. Vous permettez réellement aux gens de travailler dans de meilleures conditions. Mais surtout merci pour votre bonne humeur et votre humour permanents, je vais vraiment regretter de ne plus avoir nos conversations et nos échanges de trouvailles sur internet. Par contre je suis réellement désolée d'avoir été une meilleure élève pour la nomenclature des primers que pour celle des différents tournois de foot...

To **Hector**, I remember the first day when I arrived, you immediately took me in charge and you never let me down since. Beside the fact that you made me enter into the really selective and small world of "CSC believers", your supervision has been a wonderful driving energy for my research. I know that you always say that you hate to teach, but at the end you really took the time to teach me how to conduct a scientific project. Thanks to you, I could use these 4 years to evolve and to improve my way to think and to work. Of course an important part of your supervision was the relaxing time, with beers at flannigan's, the quiz, the bowlings, wii party etc ... it was really nice also to be able to not only talk about science and to have fun outside the lab.

Aux autres thésards du labo. Tout d'abord les "anciennes": **Marie-Pierre, Maria et Sheila**. Au delà du fait que je pense qu'on a formé un bon quatuor faisant parfois perdre un peu la tête à Hector durant nos SS meetings, j'ai beaucoup appris à vos côtés et toujours été bien contente de trouver vos conseils pour survivre à la thèse !

Sheila ton amitié a été un vrai cadeau durant ton séjour en France. Nos discussions plus ou moins sérieuses, nos "enquêtes", nos fous-rires, nos sorties ... mais à tes côtés j'ai aussi découvert une nouvelle culture et même quelques mots dans une nouvelle langue. Même si maintenant on est éloignées géographiquement, je sais qu'on repartagera d'autres super moments, qu'on refera des concours de bières, et qu'on s'échangera de nouveaux potins ;-) Eu gosta de voce muito minha tosquinha!

**Pedro** t'as assuré en temps que co-thésard ;-). Merci de m'avoir fait rire, d'avoir potiné avec moi, d'avoir longuement disserté sur les différents menus macdo mais aussi d'avoir supporté mon (léger)

stress et mes (petites) sautes d'humeur. Je suis désolée de t'avoir transmis ma façon de ranger un bureau et de t'avoir entraîné dans les histoires de TGFbeta (même si je pense que ça t'a fait du bien de changer un peu de sujet parce que sinon en 3 ans t'aurais quand même pas glandé grand chose). Je te souhaite bon courage pour la suite et j'espère vraiment qu'on restera en contact.

A **Davide et Akram** pour votre soutien. Vous avez été comme des grands-frères pendant la fin de cette thèse avec vos encouragements, vos compliments (Akram, ça va me manquer de ne plus me faire accueillir le matin par un "bonjour princesse") et vos conseils. Davide je te remercie en particulier pour le temps que tu as pris pour m'aider à améliorer mon manuscrit de thèse, je pense qu'au final tu seras une des rares personnes à l'avoir lu intégralement.

Je remercie également les autres post-docs **Reetta, Sri, Thai, Nawapol, Nasko, Sunny, Karen** pour la bonne ambiance à laquelle ils ont contribués et les conseils qu'ils ont pu m'apporter.

Merci, **Sid**, outre le fait que tu es une extraordinaire secrétaire toujours prête à se plier en 4 pour nous aider, tu as toujours eu un sourire et un petit mot gentil pour nous aider à bien démarrer la journée.

Aux petites nouvelles **Léa et Nora** quelque soit la durée de votre passage, j'espère que vous apprécierez le temps passé chez EGE autant que moi.

Aux **Drs Philippe Merle et André Verdel**, membres de mon comité de thèse. Vous avez su pendant nos réunions m'aider à faire murir ma discussion.

Aux stagiaires que j'ai encadré complètement ou partiellement, **Laetitia, Jihed et Roméo**. Merci d'avoir essuyé les plâtres de ma supervision. De mon côté cette partie plus en lien avec l'enseignement a été un vrai plaisir durant ma thèse.

A tout ceux que j'ai croisé dans les couloirs et les différents étages du CIRC, en particulier **Thomas, Carla, Nino, Vladimir, Anupam, Clément, Maha, Djamel, Natalia et Stéphanie**

Bien sûr cette thèse n'est pas seulement le fruit d'une bonne relation et de collaboration avec mes collègues, la présence de tous mes amis et de toute ma famille a été plus que nécessaire durant toutes ces années.

A **mes parents** tout d'abord. Soutenir une thèse en biologie, ce n'était pas gagné avec mes 8/20 en terminale, et pourtant vous m'avez toujours fait confiance. Même si mon travail est resté je pense très flou pour vous, vous vous êtes toujours montrés intéressés par ce que je faisais et m'avez soutenu tout le long (bien sûr je ne parle pas que des armoires à gateaux remplies avant les concours, mais quand même ça a beaucoup joué ;-)). Je vous remercie également d'avoir toujours respecté et soutenu mes choix et jusqu'à présent je n'en ai regretté aucun. J'ai la chance de pouvoir continuer à choisir ce que je veux faire après ma thèse, donc Papa ne t'en fais pas je pense que tout ira bien ;- ) reste juste aussi présent que tu l'as été jusqu'à maintenant.

A mes frères et ma soeur, **Agathe, Paul, Antoine, et Nicolas**. Je sais que parfois on vous rabâche un peu trop que j'ai tout fait bien, que j'ai eu un parcours exemplaire. Mon parcours est pourtant semé de petites erreurs, de doutes, mais au final je ne le regrette pas. Ne regardez pas trop en détails ce que j'ai fait, mais concentrez vous plutôt sur vous pour trouver et construire votre propre chemin, celui qui vous ressemblera, qui ne sera surement pas parfait mais qui vous rendra fiers de vous. J'adore vous regarder grandir, faire vos choix et constater que nous avons chacun nos personnalités bien à nous mais

que nous restons malgré tout très unis. Ca compte beaucoup que vous soyez là pour me soutenir et j'attends avec impatience le jour où ce sera moi qui viendrait vous applaudir !

Tout ceux qui me connaissent savent que quand je parle de ma famille, j'inclus beaucoup plus de personne que juste mes parents et mes frères et soeurs. Voilà pourquoi je remercie aussi mes grands-parents **Patch, Mamisa, Vieille-Maman et Papipe** qui a du "jeter un coup d'oeil" depuis là haut. C'est rare et très précieux d'avoir des grands-parents qui vous suivent avec autant d'attention et avec lesquelles on peut partager autant. Il y a également **ma cousine Laetitia, ma marraine France et mon parrain Thierry**. Nous avons développé une superbe complicité, vous me suivez de puis le début de cette aventure et vous guettez toujours le moindre signe de petite fatigue sur facebook pour pouvoir aussitôt me remonter le moral.

J'ai eu la chance d'agrandir ma famille grâce à Jérôme, et je ne pourrais pas donc oublier de remercier mes beaux-parents **René et Christine** qui m'ont également toujours beaucoup soutenu. Je pense notamment aux deux semaines passées aux Frots lors de la rédaction où j'ai rarement été aussi chouchouté. Toutes ces petites attentions m'ont permis d'atteindre mon objectif à temps !

**Valérie**, depuis les cours de danse jusqu'à notre soutien mutuel durant nos longues études (un peu trop longues diront tes parents lol) ça me touche beaucoup que tu sois venue aujourd'hui ! Je remercie également **Fred, Julie, Alexis**, et bien sûr **Raphaël** qui du haut de ces 7 ans a mieux compris que certains adultes ce que je faisais.

Bon bien sûr de temps en temps ça fait du bien de se plaindre à des gens qui savent exactement ce qu'on est en train de vivre. Merci à **Ludo, Charlène, Djo, Marion, Marina et Ludi**. Je ne regrette pas ma transformation provinciale (encore incomplète on est d'accord) puisque mon arrivée à Lyon m'a permis de vous rencontrer, de constituer avec vous un groupe d'amis formidable, mais également de nouer des liens avec chacun individuellement. Je ne vous remercie pas seulement pour les déjeuners ou apéro improvisés histoire d'aider l'un d'entre nous à oublier sa journée pourrie, mais pour tout les autres bon moments partagés (les soirées, les cafés, nos instants "thés/gâteaux" les après-midi – et oui on vieillit). On a vécu plein de choses ensemble depuis les soirées aux foyers quand on avait 20 ans et on continuera même après la thèse.

A **Marco, Alexa, Morgane et Brune**. On n'est jamais assez bien entouré pour partager un café, un verre, des soirées et ensuite se remémorer ensemble les anecdotes ;-)

A **Lisa, Sarah, Maguie et Nico**. Vous êtes mon socle parisien (malgré une petite déroute dans le nord)! J'ai partagé et établi avec chacun d'entre vous des liens particuliers. Quitter Paris n'a pas été si facile, mais je dois dire que l'attention avec laquelle vous avez toujours su garder le contact, prendre de mes nouvelles, prendre le temps de venir me voir et de continuer à partager avec moi les étapes importantes de ma vie, a énormément compté pour moi. J'apprécie d'autant plus votre présence aujourd'hui car vous n'avez pas à vous sentir si éloignés de ce que je fais : tout cela est l'aboutissement d'un parcours qui a commencé quand j'étais encore à Paris, avec vous ;-)

**Jérôme**, mon amour. Tout ceci a commencé il y a 10 ans : un mois avant mon bac tu m'avais promis de m'accompagner jusque là. Puis avec la prépa tu as rempli pour deux ans, ensuite tu m'as laissé partir à Lyon et tu m'y as rejoint pour l'agreg et à peine le temps de souffler un peu, tu as de nouveau signé pour trois ans de thèse. Je ne pense pas que tu savais ce dans quoi je t'embarquais mais depuis le début tu m'as fait confiance, tu ne m'a jamais quitté une seule seconde, tu as toujours cru en moi, tu as partagé avec moi les bonnes nouvelles, et tu m'as réconfortée quand j'avais des doutes. Tu as été une solide

*charpente qui m'a permis pendant mes études et qui nous a permis, de toujours tenir droit. On a beau avoir toutes les connaissances qu'il faut, pour arriver à ce stade il faut également s'engager personnellement, « avoir du mental » comme on dit ;-), mon parcours, cette thèse ce n'est pas le fruit de mon travail, mais de notre union. Maintenant on va pouvoir commencer une nouvelle page dans notre vie ; je ne sais pas encore quel sera son contenu mais je sais qu'on continuera à l'écrire ensemble. Je t'aime*

*A **ma Gallinette**, pour tout tes sourires, tes crapuleries et toute la joie que tu nous apportes chaque jour depuis que tu es arrivée. A **Bertie**, d'avoir été aussi sympa et discret pour me laisser finir ma thèse.*

*Au **Ninkasi**, au **Flannigan's**, et à **l'Etoile**, pour m'avoir fourni tout au long de ma thèse ~~l'alcool~~ l'énergie nécessaire.*

# Table of content

<b>Remerciements .....</b>	<b>5</b>
<b>Index of Figures .....</b>	<b>12</b>
<b>Index of Tables.....</b>	<b>13</b>
<b>ABSTRACT .....</b>	<b>15</b>
<b>RESUME .....</b>	<b>17</b>
<b>INTRODUCTION .....</b>	<b>19</b>
<b>I. The Liver: organisation, function and regeneration.....</b>	<b>21</b>
A. Anatomy and physiology of the liver. ....	21
1. Anatomical divisions and lobulation of the liver.....	21
2. Physiology of the liver .....	22
B. The hepatic cell types .....	22
C. A unique feature of the liver: the regeneration.....	24
1. General description .....	24
2. Role of cytokines and growth factors in liver regeneration:.....	24
3. Hepatic progenitor cells and liver regeneration .....	25
<b>II. Inflammatory liver diseases.....</b>	<b>27</b>
A. Hepatitis .....	27
1. Viral hepatitis.....	27
2. Alcoholic hepatitis.....	28
3. Non-alcoholic hepatosteatosis (NASH).....	28
4. Auto-immune hepatitis.....	29
B. Cirrhosis.....	30
C. Cytokines, growth factors and signaling pathways involved in inflammatory liver diseases .....	31
1. General description of cytokines activated in liver diseases .....	31
2. The IL-6- JAK/STAT signaling pathway. ....	32
3. The TGF- $\beta$ /SMAD signaling pathway.....	34
<b>III. Hepatocellular carcinoma and its links with inflammation .....</b>	<b>40</b>
A. Fundamental concepts on cancer.....	40
1. From hyperplasia to malignant tumor.....	40
2. Tumor classification.....	41
B. Hepatocellular carcinoma .....	42
1. Epidemiology.....	42
2. Risk factors.....	44
3. Molecular alterations in HCC.....	45
C. From chronic inflammation to hepatocellular carcinoma .....	47
1. Inflammatory mechanisms leading to HCC.....	47
2. Inflammation, hepatic progenitor cells and hepatocarcinogenesis.....	49
3. Creation of an inflammatory microenvironment during HCC.....	50
4. Evolution of TGF- $\beta$ functions during HCC development. ....	51
<b>IV. Cancer stem cells in hepatocellular carcinoma.....</b>	<b>54</b>

A. Cancer stem cells concept.....	54
B. Identification of liver cancer stem cells .....	59
C. CD133+ cells as liver CSCs.....	61
1. CD133+ cells as representative population of cancer stem cells.....	62
2. Clinical significance of CD133+ cells in HCC.....	63
3. Molecular characterization and biological functions active in CD133+ cells. ....	64
D. Influence of the microenvironment on CSCs .....	67
1. Cancer niches support and maintain CSC activation .....	67
2. Tumor microenvironment soluble factors influencing CSCs.....	68
3. Influence of the microenvironment on liver progenitor cell transformation. ....	71
<b>V. DNA methylation in Hepatocellular carcinoma .....</b>	<b>73</b>
A. Introduction to epigenetic mechanisms.....	73
B. DNA methylation .....	75
1. CpG sites are methylated by DNMTs.....	75
2. Demethylation processes .....	78
3. Methylation regulates transcription and genome organisation. ....	80
C. Deregulation of DNA methylation and DNMT expression in HCC .....	82
1. Aberrant DNA methylation profiles in HCC .....	82
2. Alteration in DNMT1 DNMT3A, DNMT3B expression .....	84
D. DNA methylation contribution to hepatocarcinogenesis .....	86
1. DNA methylation alterations in precancerous stages.....	86
2. DNA methylation interaction with inflammation .....	88
3. DNA methylation and cancer stem cell phenotype.....	91
<b>HYPOTHESIS AND AIMS OF THE PROJECT .....</b>	<b>93</b>
<b>MATERIALS AND METHODS .....</b>	<b>97</b>
Cell culture .....	99
Cytokines treatment .....	99
Sphere formation assay .....	99
BrDU assay .....	99
Fluorescence Activated cell sorting (FACS).....	100
Magnetic Activated cell sorting (MACS).....	100
Cell sorting .....	104
DNA extraction .....	104
Bisulfite treatment.....	105
Pyrosequencing.....	105
Bead Array Platform .....	108
RNA extraction .....	109
Reverse transcription and quantitative PCR .....	109
Whole genome expression array.....	110
Proteins extraction and Western Blot .....	111
Statistical Analysis.....	111
<b>RESULTS .....</b>	<b>113</b>
I. CD133- and CD133+ liver cancer cells differentially express DNA methylation genes	115
II. A differential DNA methylome defines CD133- and CD133+ liver cancer cells .....	120

III. TGF- $\beta$ , but not IL-6, induces CD133 expression in a stable fashion.....	124
IV. De novo induction of CD133+ cells by TGF- $\beta$ is associated to an increased expression of DNMT3 genes. ....	130
V. Transdifferentiation to CD133+ cells correlates with a methylome reconfiguration .....	135
VI. TGF- $\beta$ -induced methylome matches the basal CD133+ methylome and is reflected on mRNA expression .....	142
<b>DISCUSSION .....</b>	<b>147</b>
I. CD133+ cells represent a distinct sub-population related to cancer stem cells in HCC cell lines. ....	149
II. CD133+ cells phenotype is associated to a specific DNA methylation signature. ....	152
III. CD133+ liver CSCs are triggered by TGF- $\beta$ . ....	155
IV. TGF- $\beta$ treatment induces a global and stable DNA methylation program. ....	158
V. Correlation between TGF- $\beta$ induced DNA methylation signature and gene expression. ....	161
VI. TGF- $\beta$ induced DNA methylation contributes to establish the CD133+ CSCs phenotype in liver cancer. ....	162
VII. Further mechanistic studies. ....	164
<b>CONCLUSIONS .....</b>	<b>167</b>
<b>REFERENCES .....</b>	<b>173</b>
<b>ANNEXE I: Supplementary Tables .....</b>	<b>201</b>
<b>ANNEXE II: Review .....</b>	<b>227</b>
<b>ANNEXE III: Research Paper Manuscript .....</b>	<b>240</b>

## Index of Figures

Figure 1. Functional divisions of the liver by Couinaud. ....	21
Figure 2. Histological organization of the liver. ....	22
Figure 3. Functional anatomy of the liver. ....	23
Figure 4. Multistep model for liver regeneration. ....	25
Figure 5. Model of the hepatic stem cell niche in the canal of Hering. ....	26
Figure 6. Alcoholic liver diseases. ....	28
Figure 7. Changes in hepatic architecture associated with advanced hepatic fibrosis. ....	31
Figure 8. The IL-6/JAK/STAT signaling pathway in hepatocytes. ....	33
Figure 9. The TGF- $\beta$ /Smad signaling pathways. ....	36
Figure 10. Diversity and complexity of TGF- $\beta$ induced biological effect during liver disease progression. ....	37
Figure 11. Representation of phosphorylated sites in SMAD2 and SMAD3. ....	38
Figure 12. Cell type-specific temporal dynamics of R-Smad phosphoisoforms. ....	39
Figure 13 Tumor development. ....	41
Figure 14. Estimated age-standardized incidence rate per 100000 of liver cancer. ....	43
Figure 15. Estimated age-standardized mortality rate per 100000 of liver cancer. ....	43
Figure 16. Sequential gene alterations leading to HCC. ....	47
Figure 17. Molecular mechanisms and cellular processes involved in the road from inflammation to tumor initiation. ....	48
Figure 18. Roles of TGF- $\beta$ during multistep hepatocarcinogenesis. ....	53
Figure 19. The clonal evolution model versus the cancer stem cell model. ....	55
Figure 20. Combination of the CSC and the clonal evolution models. ....	58
Figure 21. Membrane topology of human CD133. ....	61
Figure 22. Biological processes and molecular signaling in CD133+ liver cancer stem cells. ....	67
Figure 23. Signaling between CSCs and tumoral microenvironment. ....	71
Figure 24. The three fundamental epigenetic mechanisms: histone modifications, RNA interference and DNA methylation. ....	74
Figure 25. Chemical reaction of cytosine methylation on the 5' carbone of the base. ....	76
Figure 26. Schematic structure of human DNMTs and DNMT3-like proteins. ....	77
Figure 27. Known and putative pathways of DNA demethylation that involve oxidized methylcytosine intermediates. ....	79
Figure 28. Schematic structures of TET family members. ....	80
Figure 29. Distribution of CpG sites across the genome. ....	81
Figure 30. Aberrant DNA methylation changes during carcinogenesis. ....	83
Figure 31. A hypothetical model depicting cross-talk between activation of inflammatory pathways and epigenome deregulation during liver tumor development. ....	89
Figure 32. Main steps for magnetic activated cell sorting. ....	103
Figure 33 Dot plots of cells fractions enriched in CD133+ Huh7 cells analysed by FACS. ....	104
Figure 34. Chemical steps occurring during bisulfite conversion. ....	105
Figure 35. Pyrosequencing methods (Herceg and Vaissière). ....	106
Figure 36. CD133 expression in liver cancer cell lines. ....	116
Figure 37 <i>CD133</i> gene ( <i>PROM1</i> ) is higher expressed in CD133+ cells. ....	116
Figure 38. CD133+ cells are capable of producing spheres in low attachment conditions. ....	117
Figure 39. Stemness transcription factor expression in CD133+ cells. ....	118
Figure 40. Expression of the genes encoding the key enzymes involved in DNA methylation maintenance in CD133+ cells. ....	119
Figure 41. Experimental design for genome-wide DNA methylation study in CD133+ cells. ....	120
Figure 42. A differential methylome distinguishes CD133+ and CD133- cells. ....	121



Figure 43. Genome-wide DNA methylation array revealed hypomethylation for <i>PROM1</i> in CD133+ cells. ....	122
Figure 44. CD133+ cells are globally hypomethylated compared to their negative counterpart. ....	122
Figure 45. Regional distribution of the differentially methylated CpG loci in CD133+ cells. ....	123
Figure 46. Huh7 and HepG2 cell lines expressed similar levels of TGFBR11. ....	124
Figure 47. Activation of SMAD3 after TGF- $\beta$ exposure. ....	125
Figure 48. IL-6 and TGF- $\beta$ do not alter cell viability of HCC cell lines. ....	125
Figure 49. TGF- $\beta$ and IL-6 signaling pathways target genes expression in CD133+ cells. ....	126
Figure 50. TGF- $\beta$ exposure induces morphological changes in HCC cell lines. ....	127
Figure 51. TGF- $\beta$ induces a persistent increase of CD133+ cells. ....	128
Figure 52. <i>CD133</i> mRNA expression after TGF- $\beta$ exposure. ....	129
Figure 53. Expression of TGF- $\beta$ signaling pathway target genes after TGF- $\beta$ exposure. ....	130
Figure 54. TGF- $\beta$ can induce transdifferentiation of CD133- into CD133+ cells. ....	131
Figure 55. TGF- $\beta$ 's effects on cell cycle. For Huh7 and HepG2 cells. ....	132
Figure 56. Specificity of TGF- $\beta$ 's effect on CD133+ population. ....	133
Figure 57. DNMT and TET expression is modulated by TGF- $\beta$ . ....	134
Figure 58. TGF- $\beta$ induced CD133+ cells are able to grow on low attachment conditions. ....	135
Figure 59. Experimental design for genome-wide DNA methylation study in TGF- $\beta$ exposed cells. ...	136
Figure 60. A differential methylome distinguishes TGF- $\beta$ exposed cells to controls. ....	137
Figure 61. Description of the probe clusters. ....	138
Figure 62. A 580 loci DNA methylation signature can distinguish TGF- $\beta$ exposed cells from their negative counterpart. ....	139
Figure 63. Regional distribution of the differentially methylated CpG sites after TGF- $\beta$ treatment. ....	140
Figure 64. Validation by pyrosequencing of selected differentially methylated loci. ....	141
Figure 65. Correlation between pyrosequencing and Illumina bead array analyses. ....	141
Figure 66. Experimental design for whole genome expression study in TGF- $\beta$ exposed cells. ....	142
Figure 67. Whole genome expression array validation. ....	144
Figure 68. Correlation between whole genome expression (WGX) array and quantitative PCR analyses. ....	144
Figure 69. Overlap between CD133+ and TGF- $\beta$ DNA methylation profiles defines a significant signature of 117 genes. ....	145
Figure 70. DNA methylation changes occur after 2 days of TGF- $\beta$ treatment. ....	166
Figure 71. Model for TGF- $\beta$ 's effect on CD133+ CSCs in HCC and its consequence on the tumor development. ....	170
Figure 72. Model for the DNA methylation role in TGF- $\beta$ induction of liver CSCs. ....	171

## Index of Tables

Table 1. Etiology of hepatic cirrhosis. ....	30
Table 2. Constituent of the different signalling cascade induced by TGF- $\beta$ superfamily ligand ....	35
Table 3. Expression of the TGF- $\beta$ pathway components in HCC. ....	51
Table 4. Cancer stem cells markers in different tumors. ....	56
Table 5. Functional assays to assess cancer stem cells properties. ....	56
Table 6. Cell surface marker for liver CSCs. ....	60
Table 7. List of antibodies used for fluorescent activated cell sorting ....	100
Table 8. List of primers of pyrosequencing assays ....	106
Table 9. List of primers designed for qRT-PCR. ....	109
Table 10. Characteristics of the 3 liver cancer cell lines used for the study. ....	115
Table 11. Correlation between TGF- $\beta$ -induced DNA methylation signature and gene expression. ....	143



## ABSTRACT

Distinct subpopulations of neoplastic cells within tumors, including hepatocellular carcinoma (HCC), display a pronounced ability to initiate new tumors and induce metastasis. Investigations on these cells rapidly described them as essential for tumor growth and based on these observations they have been named “cancer stem cells” (CSCs). Unfortunately, the mechanisms involved in sustaining their programs are only partially known. In HCC, there is an established link between microenvironmental signals from Transforming Growth Factor beta (TGF- $\beta$ ) and survival of certain cell subpopulations which results in a bad prognosis. However, how TGF- $\beta$  establishes and modifies cell behavior in HCC is not fully understood. As DNA methylation is involved in establishing cellular programs, our aim was to characterize the methylome of putative liver CSCs, and its link to the ability of TGF- $\beta$  to induce liver CSCs. We used CD133 expression as a positive marker for liver CSCs. To understand the relevance of DNA methylation programs in liver CSCs, we first defined the methylome signature of CD133+ cells in liver cancer cells using methylation bead arrays. Differentially methylated CpG sites were enriched in known pathways related to CSC survival and to inflammation, including the TGF- $\beta$ /SMAD pathway. Next, we showed that TGF- $\beta$  persistently induces CD133+ cells in opposition to another cytokine related to HCC, interleukin 6. We observed that this increase is associated with genome-wide changes in the methylome induced by TGF- $\beta$  and that are perpetuated through cell divisions<sup>00</sup>. We observed a significant overlap between the CD133+ methylome and the methylome induced by TGF- $\beta$ , indicating that TGF- $\beta$  may induce CSC phenotype through DNA methylation reprogramming. Additionally, we observed genome-wide effects of TGF- $\beta$  that are independent of the induction of CD133. Finally, TGF- $\beta$  methyl-sensitive sites were significantly concentrated in enhancer regions of the genome, and include well-known targets of TGF- $\beta$ , and epigenetic players, such as *de novo* DNA methyl-transferases. In conclusion our results are the first indication of the ability of TGF- $\beta$  to induce genome-wide changes of DNA methylation, leading to a stable switch to a liver cancer stem cell epigenetic program.



## RESUME

Au sein des tumeurs, y compris pour le carcinome hépatocellulaire (CHC), des sous-populations de cellules néoplasiques ont révélé une grande capacité à initier de nouvelles tumeurs et à induire des métastases. Les premières études sur ces cellules ont rapidement montré que la présence de ces cellules était déterminante dans le développement tumoral et elles ont donc été renommées « cellules souches cancéreuses » (CSCs). Malheureusement les mécanismes impliqués dans la maintenance de ces CSCs ne sont que partiellement compris. Par ailleurs dans le CHC un lien a été établi entre les signaux du facteur de croissance de transformation (Transforming Growth Factor, TGF- $\beta$ ) provenant du microenvironnement tumoral et certaines populations de cellules cancéreuses dont la présence est corrélée à un faible pronostic. La façon dont TGF- $\beta$  peut ainsi établir et modifier un phénotype cellulaire dans le CHC reste néanmoins obscure. La méthylation de l'ADN étant un acteur majeur dans la mise en place des programmes cellulaires, notre but a été de caractériser le méthylome de CSCs hépatiques et son lien avec la capacité de TGF- $\beta$  à induire des CSCs. Nous nous sommes appuyés sur l'expression du marqueur CD133 pour définir la population de CSCs hépatiques. Afin comprendre l'importance des marques de méthylation de l'ADN dans les CSCs hépatiques, nous avons dans un premier temps déterminé quelle était la signature des cellules CD133+ au niveau de la méthylation de l'ADN en utilisant des puces de méthylation à grande échelle. Les sites CpG différenciellement méthylés ont montré un enrichissement pour d'une part des voies de signalisation déjà identifiées dans les CSCs et, d'autre part, pour des voies de signalisation associées au processus inflammatoire dont la voie TGF- $\beta$ /SMAD. Par la suite, nous avons montré que TGF- $\beta$  pouvait induire de façon permanente les cellules CD133+ contrairement à une autre cytokine influente dans le cancer du foie, l'interleukine 6. Cette augmentation de cellules CD133+ induite par TGF- $\beta$  est associée à des changements de méthylation de l'ADN sur l'ensemble du génome et qui sont, de plus, maintenus au cours des divisions cellulaires. La comparaison entre les deux méthylomes (liés aux cellules CD133+ et à l'action de TGF- $\beta$ ) a exposé une signature commune significative indiquant que TGF- $\beta$  pourrait promouvoir le phénotype de CSC via le processus de méthylation de l'ADN. Mais nous avons également déterminé qu'une grande partie des effets sur la méthylation induits par TGF- $\beta$  était totalement indépendante de l'induction de cellules CD133+. Enfin, nous avons observé que les sites de méthylation sensibles au signal de TGF- $\beta$  étaient regroupés de façon significative au niveau de régions « enhancer » qui régulent la transcription des gènes. Par ailleurs, ces sites incluaient

également des gènes précédemment identifiés comme cibles de TGF- $\beta$  mais aussi des gènes codant pour des acteurs épigénétiques de premier ordre comme les méthyltransférases de l'ADN. Ces résultats constituent la première description d'une signature de méthylation de l'ADN induite par TGF- $\beta$  permettant une reprogrammation stable vers un profil épigénétique de CSC hépatiques.

## **INTRODUCTION**



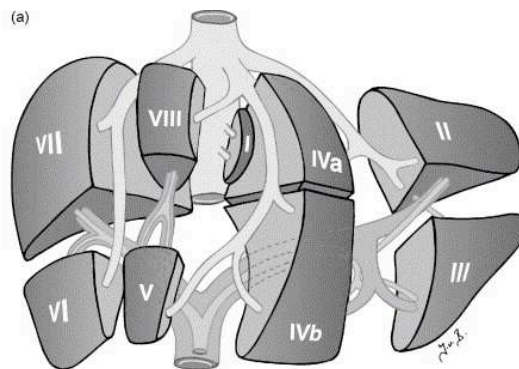


# I. The Liver: organisation, function and regeneration

Residing between the digestive tract and the rest of the body, the liver takes up different functions, including the metabolism of amino acids, carbohydrates, lipids, hormones and vitamins; serum protein' s synthesis; and detoxification of endogenous products and xenobiotics. Thus, it is not surprising that the liver is sensible to a variety of metabolic, toxic, microbial, and circulatory insults that can give rise to different pathologies, including cancer. To improve the comprehension of the context in which inflammation and tumor development may occur in liver, this first chapter will described the general structure of the liver, its function and one of it's unique features: its ability to regenerate after injury.

## A.      Anatomy and physiology of the liver.

### 1.      Anatomical divisions and lobulation of the liver

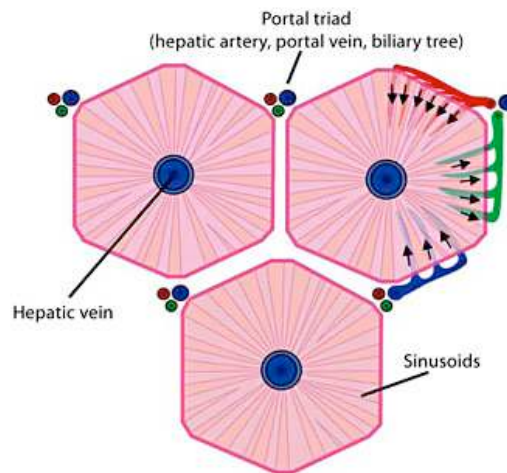


**Figure 1. Functional divisions of the liver by Couinaud.**

Using a functional description, the liver is divided into 8 independent sub segments, so called “Couinaud segments” (Figure 1). As most biochemical exchanges of the liver with body fluids are based on its vascular network, this functional segmentation is based upon the distribution of portal venous branches and the location of the hepatic veins in the parenchyma (Standring, 2008).

The ramification of the vessel system leads into the subdivision of the lobes in lobules, the small functional units of the liver. There is a well-defined hexagonal architecture, with the hepatic vein in the middle, and at the periphery the portal triad, that includes the bile duct, the hepatic artery and the portal vein (Figure 2). Therefore the blood circulation is centripetal

from the periphery to the centre of the lobule, while the bile circulation is centrifuge from the centre to the periphery of the hexagon.



**Figure 2. Histological organization of the liver. (Kline et al., 2011)**

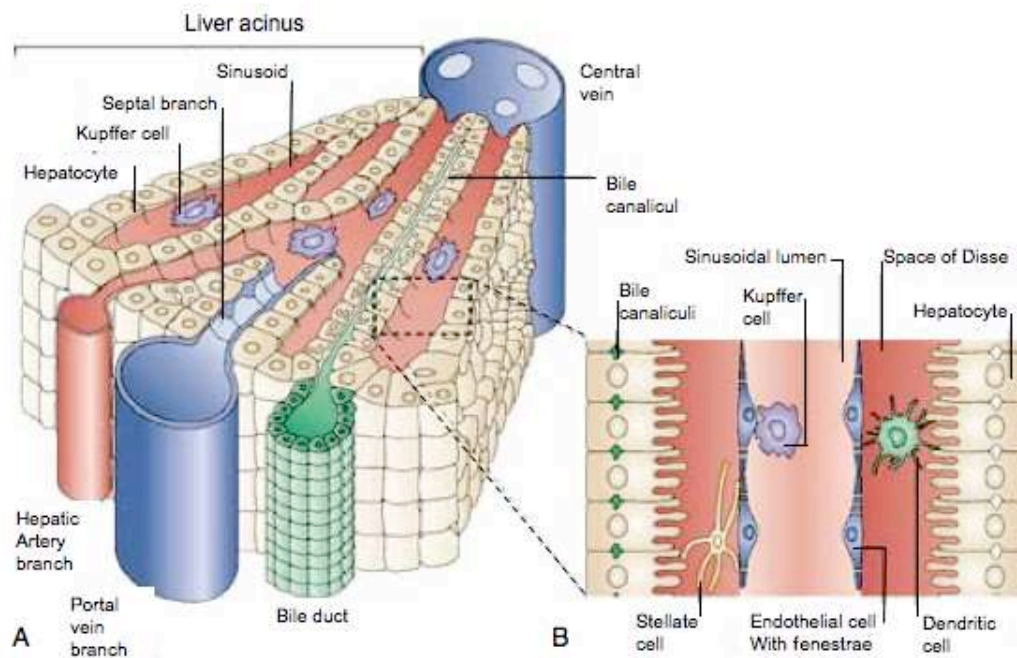
## 2. Physiology of the liver

The localisation of the liver in the circulatory system allows it to receive the portal blood that drains the stomach, small intestine, large intestine, pancreas, and spleen and its principal function is to filter and detoxify this blood. Its main functions are carbohydrate metabolism (glycogen storage), and lipid (e.g. production and storage of cholesterol and triglycerides) and protein management (e.g. production of plasma proteins) (Boron and Boulpaep, 2008). Depending on the metabolic requirements of the body, these products will be stored in the liver, secreted into the blood circulation or excreted into the bile. In addition, due to its large vascularisation and its high number of phagocytes (Kupffer cells), the liver also participates to filtering mechanism for the circulation by extracting foreign particulate matter, including bacteria, endotoxins, parasites, and aging red blood cells.

### B. The hepatic cell types

Five major cell types are essential to hepatic functions: hepatocytes, Kupffer cells, hepatic stellate cells, sinusoidal endothelium, and pit cells (Figure 3).

Hepatocytes represent 80% of the liver parenchymal volume and are the main cellular actors involved in the metabolic functions of the liver (Boron and Boulpaep, 2008). Due to their numerous and various functions and hepatocytes are the principal target in liver's injury. Hepatocytes form an epithelium that constitutes a functional barrier between two fluid compartments: in one hand the bile, in the other hand the blood.(Figure 3B).



**Figure 3. Functional anatomy of the liver.**

A. Scheme of the global organization of a hepatic lobule. B. Sections showing the different cells comprised in the liver (Adams and Eksteen, 2006).

The liver sinusoidal endothelial cells (LSEC) are the cells that compose the sinusoidal blood vessel endothelium. LSECs have a specialized, highly permeable pore system that allows access of circulating molecules to the hepatocytes. These cells also scavenge soluble compounds and can phagocytose small particles.

The Kupffer cells are macrophages localized within the sinusoidal vascular space. They are the first population of cells to be in contact with gut-derived molecules and soluble bacterial products and possess a high capacity for endocytosis and phagocytosis. They may regulate the inflammatory response by acting on numerous cellular and tissular components: T-cell activation, cytotoxicity, stimulation of fibrogenesis, alteration of endothelial cell function and modulation of hepatocyte survival and proliferation (Kmiec, 2001; Sokol, 2002).

Pit cells were firstly described in 1976 (Wisse et al., 1976) and are localized in the liver sinusoids. They possess a high cytotoxic activity and could act as a primary defence barrier to transformed cells and to virus infections (Bouwens and Wisse, 1992).

Finally the hepatic stellate cells exist in the space of Disse and store vitamin A. Upon activation, they become the major source of hepatic extracellular matrix. They can differentiate into myofibroblasts and this process is a critical event in liver fibrosis (Olsen et al., 2011). Upon liver injury, these "activated" cells participate in fibrogenesis through remodelling the extracellular matrix and deposition of type-1 collagen, which can lead to cirrhosis.

All the cells that comprised the liver tissue have specific functions but also work in tight cooperation to allow the liver to respond to the body needs. Due to its anatomical position and physiological function the liver is nevertheless subject to diverse injuries that can result in hepatocytes loss and impair its function. In such a situation the liver has the peculiar capacity to regenerate and repopulate the parenchymal tissue.

### C. A unique feature of the liver: the regeneration

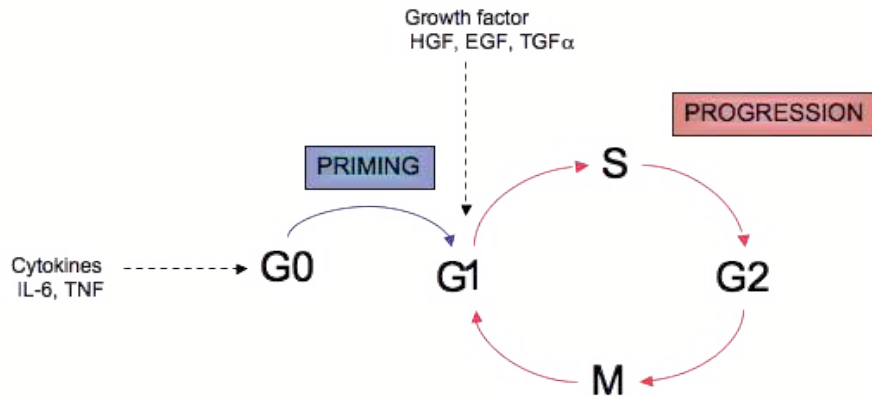
#### 1. General description

As mentioned above, the liver is the only internal human organ capable of regulating its growth and mass. Indeed, after a partial hepatectomy of 70% of the liver, the remaining tissue is able to regenerate, or more precisely, to be repopulated, into a whole liver (Duncan et al., 2009; Michalopoulos and DeFrances, 1997). Liver mass deficit can occur after surgical removal (tumor removal or transplantation from living donor) or after cell loss (functional deficit without mass deficit) caused by toxic or viral agents. When normally the rate of hepatocytes renewal is relatively low (once a year), a rapid regenerative response after loss of two-thirds or more of the liver mass can be observed (Alison et al., 2009). Furthermore in order to not exceed metabolic demands and to maintain an optimal liver mass/ body mass ratio, the liver is also capable of loss of mass by hepatocyte apoptosis. This phenomenon, while less described, can still be observed for drug-induced hyperplasia (Schulte-Hermann et al., 1995) or “large for small” transplant situation (when a large liver is transplanted into a small receiver) (Kam et al., 1987).

#### 2. Role of cytokines and growth factors in liver regeneration:

In case of liver mass or liver function deficit, hepatocytes are the first cells of the liver to enter into the cell cycle and undergo proliferation (Fausto, 2000; Taub, 2004). Genes implicated in this process are activated in sequential order with early genes mainly involved in the transition from quiescence to cell cycle and later genes involved in the progression to the cell cycle, DNA replication and mitosis processes. This multistep process is supported by cytokines and growth factors (Figure 4). The transition from G0 (quiescence) to G1 phase is called “priming” and is mainly triggered by IL-6 and TNF- $\alpha$  signals (Kirillova et al., 1999). The second phase will be supported by HGF (Pediaditakis et al., 2001), TGF- $\alpha$  and EGF signals. Much less is known about how liver regeneration is terminated once the appropriate

liver mass is restored, but it would imply that cytokines such as TGF- $\beta$  will inhibit hepatocyte proliferation (Karkampouna et al., 2012) and cascade signaling negative feedbacks that will turn off the IL-6 pathway (Elliott, 2008).



**Figure 4. Multistep model for liver regeneration.**

Liver regeneration is divided into two phases, priming and cell cycle progression. Priming is a reversible process initiated by cytokines as well as nutritional and hormonal signals. Priming sensitizes the cells to growth factors but is ineffective in their absence. Growth factors are required for cells to move beyond a restriction point in G1 ( adapted from Fausto, 2000).

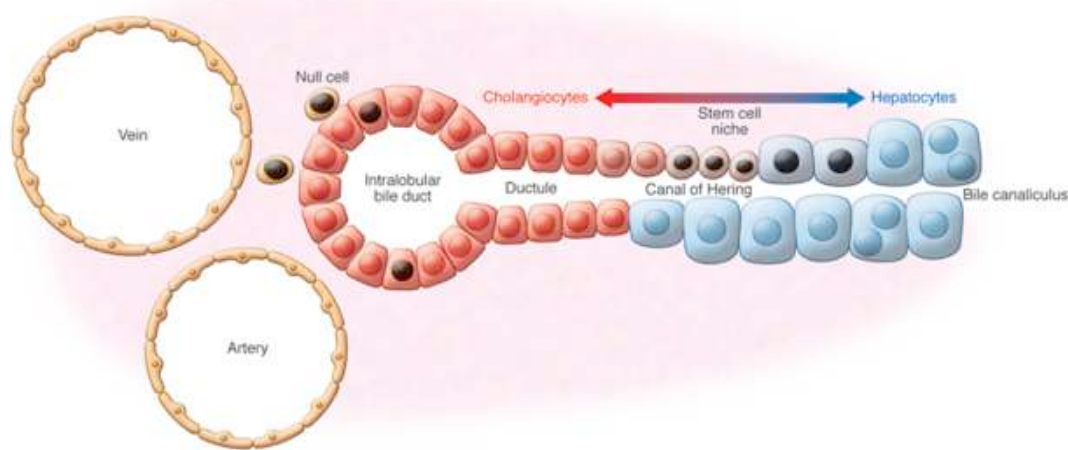
The capacity of mature liver cells to proliferate in response to common forms of injury is remarkable. However, when this response is impaired, the contribution of hepatic progenitors becomes apparent. For example partial hepatectomy is commonly associated with administration of drugs that impair hepatocyte proliferation, triggering the activation of hepatic progenitor cells (HPC) (Alison, 1998).

### 3. Hepatic progenitor cells and liver regeneration

In adult human tissues, HPCs have been localized in the smallest terminal branches of the biliary tree also called “Canals of Hering” (Alison, 2005). HPCs are thus in continuity with hepatocytes at one side and bile duct cells at the other side (Figure 5).

When hepatocytes or cholangiocytes replication are altered, inhibited or slowed down, the HPC population is activated (Roskams et al., 2003a). Then HPCs proliferate and differentiate into hepatocytes and biliary cells. This activation, named “ductular reaction” (POPPER et al., 1957) in humans and “oval cell reaction” in rodents, is observed during liver injuries such as prolonged necrosis, cirrhosis, and chronic inflammatory liver diseases. Moreover, the proportion of HPCs undergoing activation positively correlates with the severity of liver disease (Libbrecht et al., 2000; Lowes et al., 1999). The activation of HPCs and their

differentiation relies not only on the inability of hepatocytes to proliferate, it also depends on microenvironmental factors. Indeed the two models of regeneration are not mutually exclusive, and they have already been observed in some injury models (Rosenberg et al., 2000; Wang et al., 2003). Many cytokines and growth factors have been investigated for oval cells activation (even if some controversies persist between the different models). TNF, TWEAK, IL-6, HGF and EGF are the main actors involved in oval cells proliferation and expansion (Brooling et al., 2005; Knight et al., 2000; Yeoh et al., 2007), while  $LT\alpha$ ,  $LT\beta$ ,  $IFN\alpha$  and  $TGF-\beta$  (Akhurst et al., 2005; Knight and Yeoh, 2005; Nguyen et al., 2007; Preisegger et al., 1999) are responsible for their proliferation arrest.



**Figure 5. Model of the hepatic stem cell niche in the canal of Hering. (Kordes and Häussinger, 2013)**

Liver regeneration, sustained by hepatocyte proliferation and/or HPC activation, is usually accompanied by an inflammatory episode. In humans, HPCs have been observed in samples from patients with liver cancer or chronic diseases (Libbrecht and Roskams, 2002). Moreover these two phenomena are sustained by cytokine actions. Cytokines are small molecules, used for cell signaling, that regulate host responses to infection, immune responses and inflammation.

Therefore, after injuries caused by divers external or internal agents, several types of inflammatory diseases can affect the liver. We will see that during these inflammatory diseases, the entire hepatic structure can be affected and that the microenvironment is highly modified by cytokines.

## II. Inflammatory liver diseases

Inflammation is a beneficial host response to foreign aggressions and necrotic tissue, but it is itself susceptible to generate tissue damages. Inflammation can be classified as either acute or chronic. Acute inflammation constitutes the primary response of the body to injuries and is carry out by the migration of immune cells from the blood into the damaged tissues. Inflammation becomes “chronic” when prolonged and accompanied by a shift in the type of cells present at the site of inflammation. Chronic inflammation is as a process that encompasses simultaneous destruction and healing of the tissue. Several liver conditions can trigger chronic inflammation and they will be described in the next sections.

### A. Hepatitis

Hepatitis is defined by the inflammation of the liver and characterized by the presence of inflammatory cells in the organ tissue. The main risk factors associated with hepatitis are viral infection by hepatitis viruses A (HAV), B (HBV), C (HCV), D (HDV), and E (HEV) (Thomas and Zoulim, 2012), alcohol intake (Mandrekar and Szabo, 2009) and fatty liver disease (Kopeck and Burns, 2011).

#### 1. Viral hepatitis

Viral hepatitis is an inflammatory reaction of the liver caused by hepatotropic viruses (HAV, HBV, HCV, HDV and HEV). The pathophysiology of viral hepatitis covers a broad spectrum from asymptomatic infection to fulminant liver failure. Even if in most cases the infection resolves itself, viral hepatitis infection is one of the primary causes for liver transplantation in the US and other countries (Herzer et al., 2007). In fact, 4% of HBV infected patients and 85% of HCV infected patients will develop chronic hepatitis (Kumar et al., 2012).

In particular for HBV and HCV the host immune response to the virus is the main determinant of the outcome of the infection. The mechanisms of innate immunity protect the host during the initial phases of the infection, and can lead to the resolution of acute infection (Neumann-Haefelin et al., 2005; Thimme et al., 2003). However in HCV infected patients this response often appears not to be sufficient for eradicating the infection. During viral hepatitis, fibrogenesis is also enhanced and may contribute to the development of cirrhosis (Ciurtin and Stoica, 2008; Soussan et al., 2003). Most of the mortality attributed to

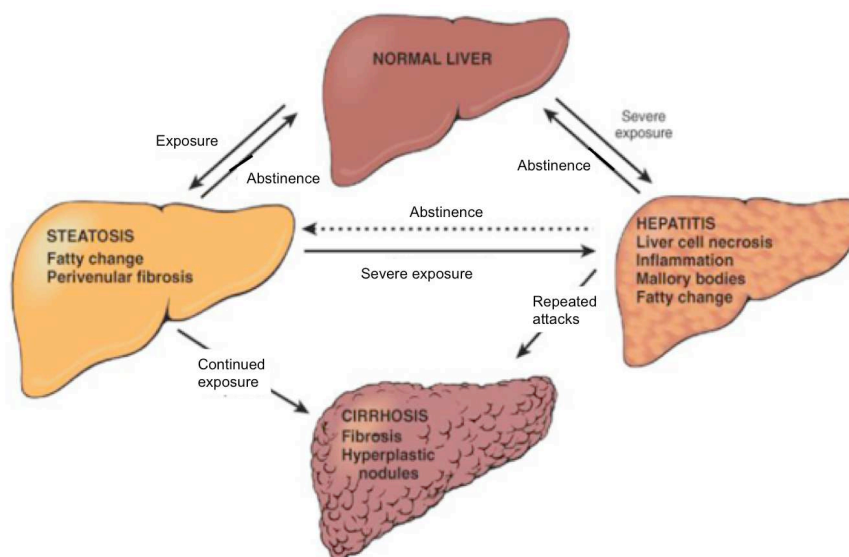


viral hepatitis is the consequences of long-term chronic hepatitis, and its evolution into cirrhosis and/or hepatocellular carcinoma (HCC).

## 2. Alcoholic hepatitis

Chronic alcohol consumption has a variety of adverse effects. However the major forms of alcoholic diseases are: (1) hepatic steatosis (fatty liver), (2) alcoholic hepatitis, and (3) cirrhosis, referred together as alcoholic liver disease (ALD). Ninety to 100% of heavy drinkers develop fatty liver (steatosis), and of those, 10% to 35% develop alcoholic hepatitis (Kumar et al., 2012). Steatosis and alcoholic hepatitis may arise separately, and therefore do not necessarily represent a continuum of changes (Figure 6).

Alcoholic hepatitis is thought to be a precursor to the development of cirrhosis and up to 50% of patients with biopsy-proven alcoholic hepatitis will present cirrhotic-related histological disorders



**Figure 6. Alcoholic liver diseases.**

The interrelationships among hepatic steatosis, hepatitis, and cirrhosis are shown, along with a depiction of key morphologic features at the microscopic level (Kumar, Abbas et al. 2007)

## 3. Non-alcoholic hepatosteatois (NASH)

Free fatty acids (FFAs) from blood circulation can be absorbed by the liver (El-Zayadi, 2008). Any imbalance between the delivery of fat to the liver and its subsequent metabolism and/or secretion will lead to the development of non-alcoholic fatty liver disease (NAFLD). This liver injury associated with an abnormal accumulation of fat encompasses different



forms of diseases from bland fatty infiltration to cirrhosis. Non-alcoholic steatohepatitis (NASH) is an intermediate liver injury state between these two extremes. Biopsies in patients suffering from NASH reveal hepatocyte injuries, apoptosis and infiltration by inflammatory cells (Choi and Diehl, 2005).

As mentioned before, hepatitis may stimulate hepatic cell activation and fibrosis. The progression of fibrosis has been observed in 35% of patients exhibiting NASH. The rate for cirrhosis development over 10 years is between 5 and 20% and the estimated rate for liver-related mortality in patients suffering from NASH reaches 12% (El-Zayadi, 2008).

#### 4. Auto-immune hepatitis

Auto-immune hepatitis (AIH) is an auto-immune liver disorder characterized by an abnormal response of the immune system against a tissue normally present in the body. AIH occurs worldwide, with a reported range of prevalence from 1.9 cases per 100,000 in Norway to 1 per 200,000 in the US general population (Mieli-Vergani and Vergani, 2011).

Due to its functions, the liver is continuously exposed to rich-antigen blood and is highly enriched in phagocytic cells, lymphocytes and antigen-presenting cells (APCs), like LSECs, HSCs, hepatocytes and dendritic cells (DCs). When self-tolerance is lost liver auto-immunity ensues. Two general conditions usually prevail for liver auto-immunity: self-reactive B- and T-lymphocytes must exist in the immunological repertoire and auto-antigens must be presented by APCs (Vergani and Mieli-Vergani, 2008).

The exact aetiology of autoimmune hepatitis is not known. Epidemiological studies indicate that it is most probably a bi-modal disease with genetic susceptibilities (involving one or more genes acting alone or in concert) in combination with environmental factors (Mieli-Vergani and Vergani, 2011).

Chronic liver hepatitis pathologies can exist for extended periods, but are not an end-stage disease. Mechanisms involved in liver regeneration, necrotic hepatocytes clearance and matrix remodelling are constantly solicited and will lead to the deregulation of liver architecture and functions. This stage, when the original organisation of the liver is destructed is referred to as cirrhosis of the liver.

## B. Cirrhosis

Cirrhosis is a long-term consequence of chronic liver disease and can be defined histologically as “a diffuse process characterised by fibrosis and a conversion of normal architecture into structurally abnormal nodules”. This loss of liver architecture is usually associated with a loss of hepatic functions. The main risk factors for cirrhosis are alcoholism, hepatitis B and C, and fatty liver disease, but many other causes are possible and are not mutually exclusive (Table 1).

More precisely the key morphological features of cirrhosis include: diffuse fibrosis, nodules of regenerative parenchyma cells, altered lobular architecture and establishment of intrahepatic shunts between afferent and efferent liver vessels. Subsequent secondary characteristics are: capillarization of the sinusoids (loss of fenestrae by LSEC), vascular thrombosis, obliterative lesions in portal tracts and hepatic veins, and under-perfusion of the parenchyma leading to hepatic tissue hypoxia (Pinzani et al., 2011).

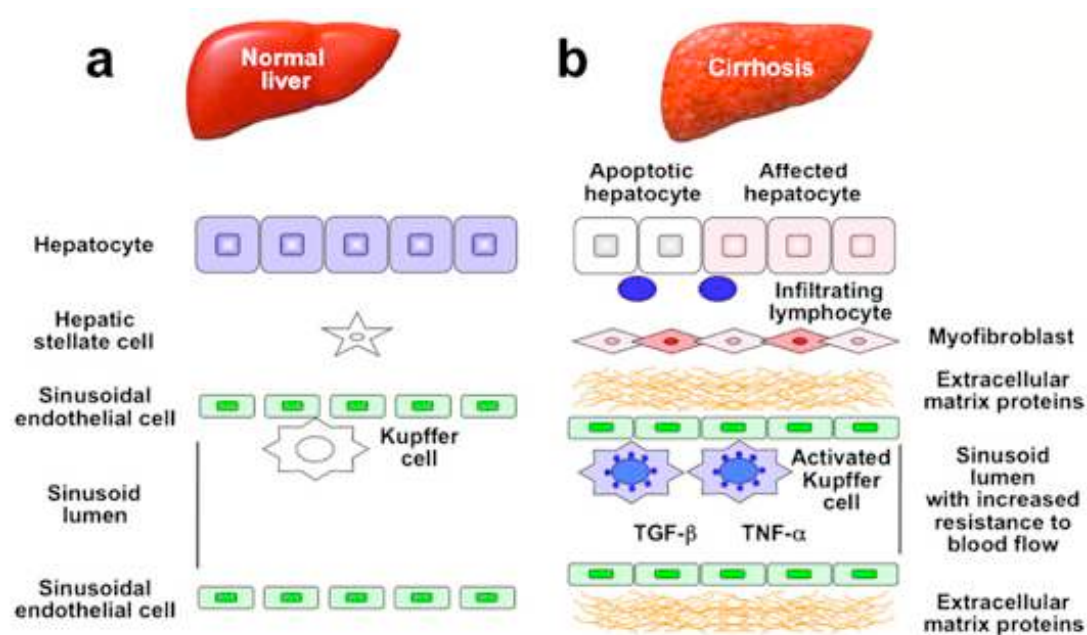
**Table 1. Etiology of hepatic cirrhosis (adapted from Heidelbaugh and Bruderly 2006)**

<b>Etiology of hepatic cirrhosis</b>
<i>Most common causes</i>
Alcohol (60 to 70%)
Biliary obstruction (5 to 10%)
Primary or secondary biliary cirrhosis
Chronic hepatitis B or C (10 %)
Hemochromatosis (5 to 10%)
NAFLD (10%)
<i>Less common causes</i>
Autoimmune chronic hepatitis
Drugs and toxins
Genetic metabolic disease
Infection
Vascular abnormalities
Veno-occlusives disease

Fibrosis is the main mechanism involved in the histological destruction of the liver. In fact, for a long time, cirrhosis was described as the final stage of fibrosis. Fibrosis is excessive production of connective tissue. It is the consequence of a chronic wound healing reaction occurring in response to chronic damage. Figure 7 describes the main changes in hepatic architecture under fibrosis.

The cirrhosis biology (constant stimulus of parenchyma regeneration in an inflammatory microenvironment) will strongly predispose patients for hepatocellular carcinoma (HCC) development. Indeed external stimuli can induce alterations in mature hepatocytes that

under proliferative pressure will create a monoclonal population harbouring dysplastic and further neoplastic hepatocytes (Pinzani et al., 2011).



**Figure 7. Changes in hepatic architecture (a) associated with advanced hepatic fibrosis (b).**

Following liver injury, lymphocytes infiltrate the hepatic parenchyma. Some hepatocytes undergo apoptosis, and Kupffer cells are activated to release fibrogenic mediators such as transforming growth factor- $\beta$  (TGF- $\beta$ ) and tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ). In response to these cytokines, hepatic stellate cells (HSC) transdifferentiate into myofibroblast-like cells and come to secrete large amounts of extracellular material (ECM) proteins. Affected hepatocytes also participate in liver fibrogenesis by stimulating the deposition of ECM proteins. As liver fibrosis progresses, sinusoidal endothelial cells lose their fenestrations, with tonic contraction of HSC increasing resistance to blood flow in hepatic sinusoids (Matsuzaki, 2011).

### C. Cytokines, growth factors and signaling pathways involved in inflammatory liver diseases

#### 1. General description of cytokines activated in liver diseases

As mentioned earlier, all inflammatory actions during chronic liver disease proliferation are mediated through autocrine/paracrine signals involving cytokines. One of the important actions of cytokines is maintaining the balance between proliferation, apoptosis and differentiation (during embryogenesis and organogenesis in particular) and any perturbations to this balance can bring out serious disorders. In chronic liver diseases the balance between protective and damaging signals is fragile, and hepatic failure might arise from excessive apoptosis. Among the various and numerous cytokines involved in liver inflammation, those of most interest to researchers are: TNF- $\alpha$ , IL-6, IL-1 $\alpha$ , IL-1 $\beta$ , TGF- $\beta$  and

IL-10 (Martin and Herceg, 2012). TNF- $\alpha$  is one of the first cytokines released by Kupffer cells, LSECs, HSCs or hepatocytes in all types of liver hepatitis. Its level is elevated in both serum and hepatic tissue in patients with alcoholic liver disease (Hill et al., 1999), with chronic HBV (Falasca et al., 2006), or with steatohepatitis (Fainboim et al., 2007). TNF- $\alpha$  can have both a pro-apoptotic function through the activation of caspases or a survival function through the activation of the nuclear factor kappa B (NF- $\kappa$ B) pathway (Tacke et al., 2009).

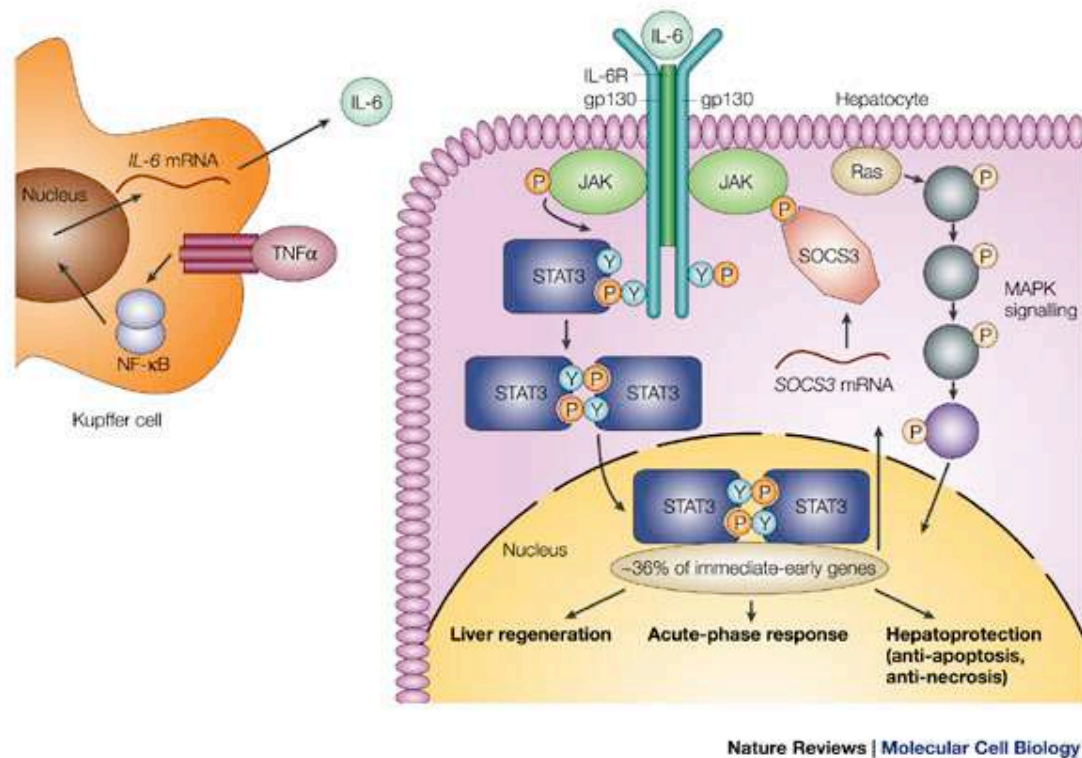
As for the other mentioned cytokine, large-scale studies investigating patient's serum observed that IL-1 $\beta$ , IL-6, TNF- $\alpha$ , TGF- $\beta$ , IL-10 were higher in cirrhosis or chronic hepatitis compared to healthy case (Budhu and Wang, 2006). Moreover comparison between the different forms of liver inflammation revealed a positive correlation between cytokine expression and the disease proression (from hepatitis to cirrhosis) (Kitaoka et al., 2003; Song et al., 2003). These observations suggest that the deregulation of cytokine expression could participate in the evolution of liver disease.

Among the panel of cytokines released in the hepatic environment, two of them fill crucial functions and are always involved in all hepatitis cases, cirrhosis, and fibrosis. On one hand IL-6 is one of the main pro-inflammatory cytokines largely contributing to compensatory hepatocyte proliferation during liver damage (Gao, 2005). On the other hand TGF- $\beta$  is an anti-inflammatory cytokine, involved in arrest of hepatocyte proliferation. However, its fundamental role in sustaining fibrogenesis by activating HSCs makes it a determinant mediator of liver disease progression (Dooley and ten Dijke, 2012). TGF- $\beta$  is involved in all stages of liver diseases (from inflammation to hepatocellular carcinoma) but as it will be described later it can generate multiple biological processes that are sometimes paradoxical. Although much effort has been put into elucidating this signal, TGF- $\beta$  effects are only partially understood. As my work focuses to a large extent on this cytokine, detailed paragraphs will be dedicated to it in this section and the following ones.

## 2. The IL-6- JAK/STAT signaling pathway.

IL-6 belongs to a family including 6 members: IL-6, leukaemia inhibitory factor (LIF), ciliary neutrophic factor (CNTF), oncostatin M (OSM), cardiotrophin-1 and IL-11. The receptors for this family can be composed of a homodimer of the gp130 protein or a heterodimer composed of gp130 with another cytokine specific receptor (Heinrich et al., 1998, 2003). Primary human hepatocytes express IL-6R, gp130, CNTFR, LIFR, OSMR, IL-11R and cardiotrophine-1R (Gao, 2005). The binding of IL-6 to its receptor will activate the

phosphorylation of a Janus Kinase (mostly JAK2) that in turn will phosphorylate STAT3 on the Y705 position. Activated STAT3 forms homodimers and is translocated into the nucleus where it enhances the transcription of several genes belonging mainly to cell survival pathways and implicated in the G1-S phase transition. Besides STAT3, JAK can phosphorylate and activate the protein tyrosine phosphatase SHP2 that will link the cytokine receptor to the mitogen-activated-protein-kinase (MAPK) pathway (fundamental for IL-6 mitogenic function)(Figure 8).



**Figure 8. The IL-6/JAK/STAT signaling pathway in hepatocytes.**

After activation of the IL-6 receptor through the interaction with its ligand, the canonical JAK/STAT pathway is activated. Alternative IL-6 activated pathways include the MAPK pathway. IL-6 signaling includes different regulation systems including a negative feedback triggered by SOCS proteins. (Taub, 2004).

IL-6/JAK/STAT is largely involved in immune regulation, haematopoiesis, inflammation and oncogenesis by regulating cell growth, proliferation and cell survival. In liver injury context, IL-6 is one of the main pro-inflammatory cytokines secreted, among others, by Kupffer cells. IL-6 is mainly involved in acute phase proteins production, liver regeneration (through proliferative effect) and hepatoprotective function (Masubuchi et al., 2003; Ramadori and Armbrust, 2001; Zimmers et al., 2003).

Its protective role has been illustrated in mice studies where IL-6 deficient mice are more sensitive to liver damages (Kovalovich et al., 2000). IL-6 also contributes to fibrogenesis modulation via indirect inhibition of ECM proteases (Shigekawa et al., 2011). However, increasing liver disease severity, from acute hepatitis, to chronic hepatitis, to cirrhosis to HCC has been observed in parallel to increasing IL-6 level (García-Galiano et al., 2007; Kao et al., 2012; Streetz et al., 2003; Zekri et al., 2005). Moreover in HCC, IL-6 is expressed at high levels, and STAT3 is often observed to be activated (He et al., 2010). IL-6 also seems to participate in carcinogenesis, probably through its proliferative effect that supports the expansion of transformed cells. The shift between hepatoprotective and pro-tumorigenic functions were illustrated in a study where an overexpression of IL-6 and IL-6R led to the development of regenerative hyperplasia and adenoma in the liver (Maione et al., 1998).

### 3. The TGF- $\beta$ /SMAD signaling pathway.

The TGF- $\beta$  superfamily ligand includes: bone morphogenetic proteins (BMPs), Growth and differentiation factors (GDFs), Anti-müllerian hormones (AMH), Activin, Nodal and TGF- $\beta$  families. The TGF- $\beta$  family comprises TGF- $\beta$ 1, TGF- $\beta$ 2 and TGF- $\beta$ 3 (Horbelt et al., 2012; Miyazawa et al., 2002). Signaling begins with the binding of a TGF- $\beta$  superfamily ligand to a TGF- $\beta$  type II receptor. The type II receptor is a serine/threonine receptor kinase, which catalyses the phosphorylation of the type I receptor. Each class of ligand binds to a specific type II receptor. In mammals there are seven known type I receptors and five type II receptors (Table 2).

TGF- $\beta$  ligands are initially released in the extracellular milieu in an inactive form, bound to latency associated peptide (LAP) and latent TGF- $\beta$  binding protein (LTBP), which form a complex masking TGF- $\beta$  epitopes preventing any signal activation (Marek et al., 2002). Activation of latent TGF- $\beta$  requires enzymatic proteolysis of this inactive complex.

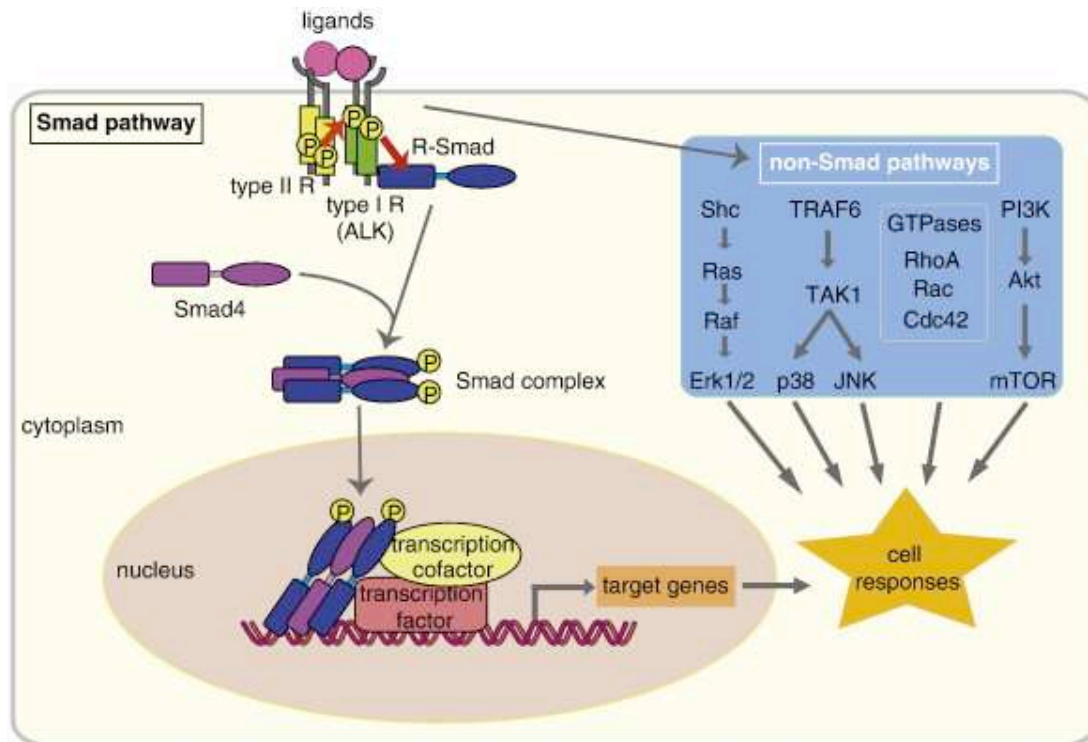
**Table 2. Constituent of the different signalling cascade induced by TGF- $\beta$  superfamily ligand**  
Alternative names are listed in brackets (Akhurst and Hata, 2012)

Molecular category	TGF- $\beta$ pathway *	Activin/Nodal pathway*	BMP pathway*
Ligands	TGF $\beta$ 1, TGF $\beta$ 2, TGF $\beta$ 3	Activin A, activin B, inhibin A, inhibin B, Nodal	BMP2, BMP4, BMP6, BMP7, BMP8A, BMP8B, BMP9, BMP10
Type I receptors	T $\beta$ RI (ALK5), ALK1 (ACVRL or SKR3)	ALK4 (ACVR1B or ACTRIIB), ALK7 (ACVR1C or ACTRIIC)	ALK1 (ACVRL1, SKR3), ALK2 (ACVR1, ACTRI), ALK3 (BMPR1A), ALK6, BMPR1B)
Type II receptors	T $\beta$ RII	ACTRIIA, ABTRIIIB	BMPR2, ACTRIIA, ACTRIIB
R-SMADs	SMAD2, SMAD3	SMAD2, SMAD3	SMAD1, SMAD5, SMAD8
Co-SMAD	SMAD4	SMAD4	SMAD4
I-SMAD	SMAD7	SMAD7	SMAD6, SMAD7

After interaction with a type II receptor and following dimerization with type I receptor, internalisation of the signal continues through the SMAD pathway. Carboxy-terminal phosphorylation of SMAD2 and SMAD3 by activated receptors results in their partnering with the common signaling transducer SMAD4, and translocation to the nucleus. Activated Smads regulate diverse biological effects by partnering with transcription factors resulting in cell-state specific modulation of transcription. Activin and Nodal ligands will transmit signals through the same SMAD2/SMAD3 pathway, while other families of ligands (BMPs, GDFs, AMH) will perpetuate signals through the SMAD1/SMAD5/SMAD9 pathway (Horbelt et al., 2012; Miyazawa et al., 2002). Besides the canonical Smad-mediated TGF- $\beta$  signaling pathway, it has been shown that TGF- $\beta$  superfamily ligands can also regulate cellular or physiological processes through non-canonical pathways by activating other signaling molecules [e.g. Akt, MAPK, mTOR (mammalian target of rapamycin), and Src] (Zhang, 2009) independent of SMAD proteins, which amplifies the complexity of TGF- $\beta$  signaling (Figure 9).

TGF- $\beta$  is mainly known as a cytokine involved in differentiation and anti-inflammatory processes mediated through mechanisms like cell cycle arrest and further apoptosis. During chronic liver disease TGF- $\beta$  is largely secreted by Kupffer cells and LSECs (De Bleser et al., 1997). Hepatic stellate cells are the first targets for TGF- $\beta$ , which will promote their transformation into myofibroblasts, the synthesis of collagen and the production of ECM proteins (Dooley and ten Dijke, 2012).





**Figure 9. The TGF-β /Smad signaling pathways**

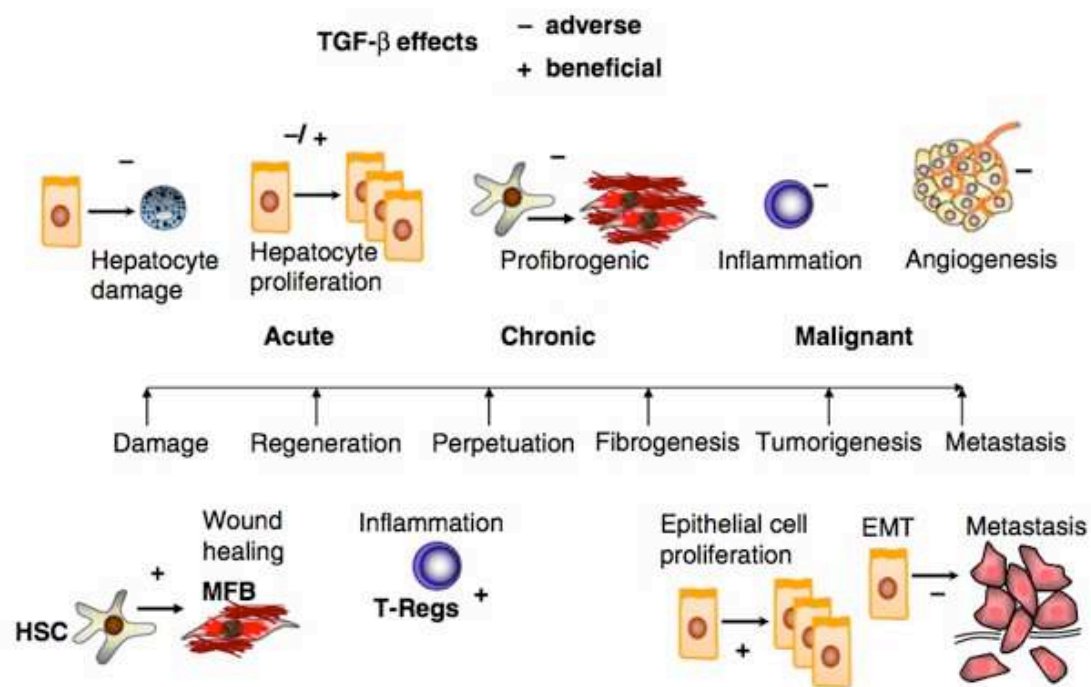
Smad-dependent and Smad-independent TGF-β family signaling. Ligands of TGF-β family members bind to type I and type II receptors. Upon ligand binding, the type II receptors phosphorylate the type I receptors, which then phosphorylate and activate effector Smads. The activated Smads form complexes with Smad4, and translocate into the nucleus. The Smad complex interacts with other transcription factors, co-activators or co-repressors to regulate transcription of target genes. TGF-β also elicits activation of other signaling cascades independent of Smad pathways. TGF-β activates the Ras–Raf–MEK–Erk MAPK pathway through tyrosine phosphorylation of ShcA, and p38 and JNK MAPK signaling through activation of TAK1 by the TRAF6. TGF-β also activates the small GTPases Rho, Rac and Cdc42, and the PI3K–Akt pathway (Sakaki-Yumoto et al., 2013).

TGF-β is thus a major actor in the development of fibrosis. In patients suffering from chronic hepatitis, a positive correlation was observed between the amount of collagen precursor and TGF-β1 expression (Castilla et al., 1991; Dooley et al., 2008). Plasma level of TGF-β also presents a correlation between the cytokine secretion and the extent of liver fibrosis (Tsushima et al., 1999; Xiao et al., 2012).

While TGF-β is thus sustaining growth and differentiation in HSCs (mesenchymal cells) its action totally differs in hepatocytes (epithelial cells). In hepatocytes, TGF-β's action will counteract pro-inflammatory proliferative effect by promoting cell cycle arrest and further apoptosis (Moustakas and Kardassis, 1998; Sheahan et al., 2007; Yoo et al., 2003). In HCV infected patients, TGF-β produced by HCV-specific T cells even appeared to have a protective role and is inversely correlated with inflammation (Li et al., 2012b). Thus, reflecting the complexity of TGF-β intracellular signaling pathways, TGF-β biological effects



are highly diverse and strongly depend on the cellular type and the biological context. Figure 10 recapitulates the various biological actions established for TGF- $\beta$  in liver injury context.

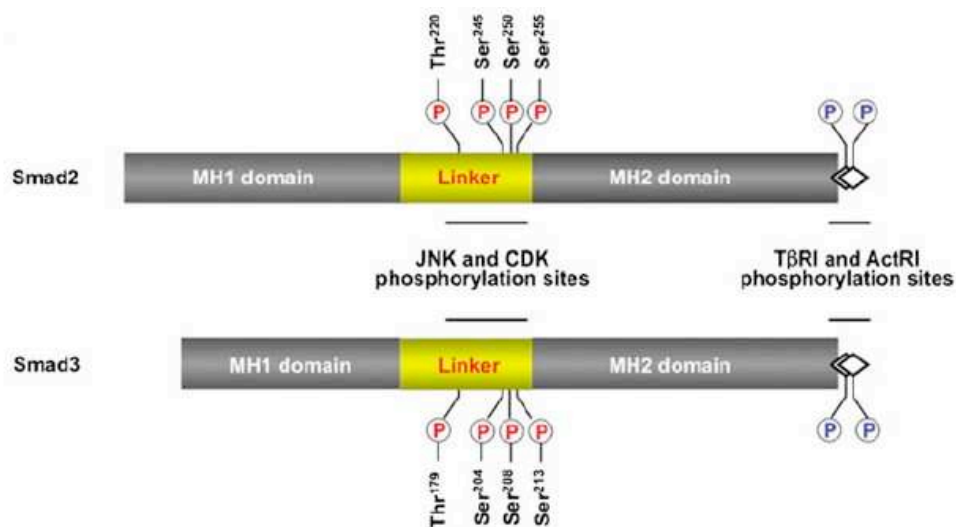


**Figure 10. Diversity and complexity of TGF- $\beta$  induced biological effect during liver disease progression.**

During the life span, the liver undergoes many different phases, as shown along the central time line. Strongly depending on the disease stage, TGF- $\beta$ , and thus its targeting, might have a good (+) or bad (-) outcome in the organ. (Dooley and ten Dijke, 2012).

To complete this elaborate picture, phosphorylated isoforms for pSmad2 and pSmad3 have been described and actively contribute to the diversity of biological actions triggered by TGF- $\beta$ . Smads are modular proteins with conserved Mad-homology-1 and 2 (MH1/2) intermediate linker domains (Figure 11). The phosphorylation sites are traditionally described in the COOH tail domain of R-Smad (pSmadRC isoforms) but can also occur in the linker domain, thus creating a second type of phosphoisoform, the pSmadRL. The linker domain undergoes regulatory phosphorylation by MAPK pathways including extracellular signal regulated kinase (ERK), c-Jun N-terminal kinase (JNK), p38 MAPK, and cyclin-dependent kinase (CDK)-2/4 (Kretzschmar et al., 1999; Mori et al., 2004). These pathways are usually activated by pro-inflammatory cytokines such as TNF- $\alpha$ . Except for pSmad2L, which is cytoplasmic (Kretzschmar et al., 1999; Yamagata et al., 2005), all the phosphoisoforms are localized in the cell nuclei to perpetuate biological signals. The isoforms will activate different sets of genes and thus will differ in their subsequent biological effects. In hepatocytes, TGF- $\beta$ /Activin signaling will involve the pSmad3C and pSmad2C isoforms and

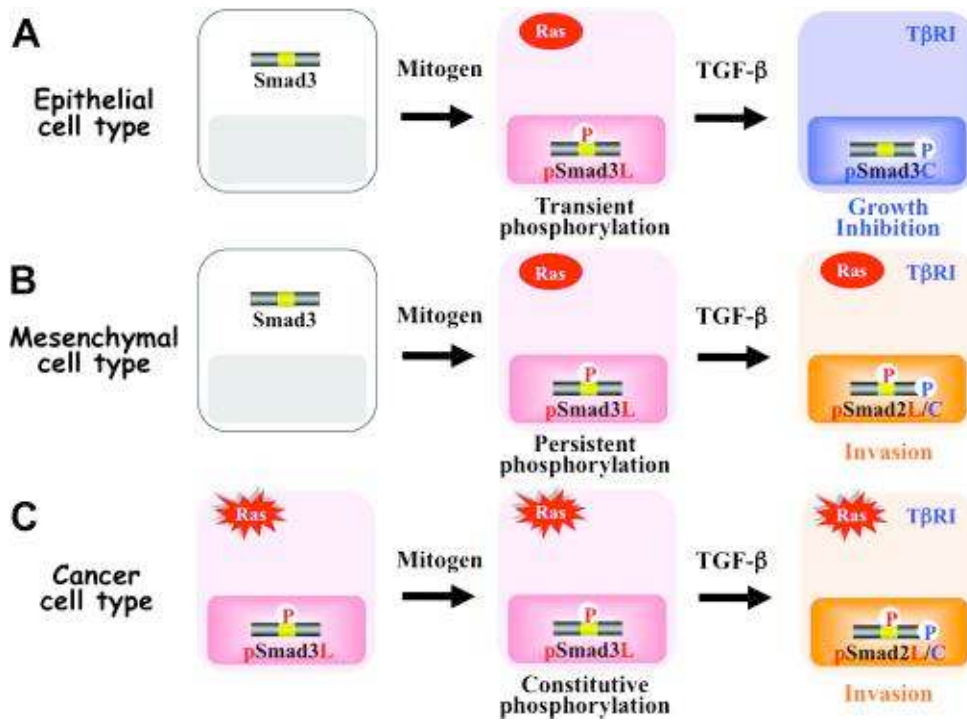
lead to cell cycle arrest (Yang et al., 2006). In contrast JNK signaling will involve pSmad3L and pSmad2L isoforms and trigger a mitogenic signal (Furukawa et al., 2003; Matsuzaki et al., 2007; Mori et al., 2004; Sekimoto et al., 2007). Interestingly pSmad3C and pSmad3L signals oppose each other but the balance between the two can shift. For example in the case of Smad3 mutants lacking linker phosphorylation sites and/or in presence of JNK inhibitors, the growth inhibitory effect can be restored (Murata et al., 2009; Nagata et al., 2009; Sekimoto et al., 2007). Such regulation should be taken into account when one is considering the effectiveness of cytostatic effect of TGF- $\beta$ /Activin on hepatocytes.



**Figure 11. Representation of phosphorylated sites in SMAD2 and SMAD3 (Matsuzaki, 2012)**

In mesenchymal cells (such as HSCs), the pSmadRL isoforms will also inhibit the anti-proliferative effect but here, the phosphorylation on the COOH-tail is necessary to induce phosphorylation on the linker site (Matsuura et al., 2010; Wang et al., 2009). Thus, the third isoform, pSmadRL/C, dually phosphorylated will be present in hepatic stellate cells. These isoforms will promote growth stimulation and fibrogenesis (Furukawa et al., 2003; Li et al., 2009; Matsuzaki, 2009). As shown in Figure 12 the shift between the isoforms is thus continuously used to adapt the transcriptional response of SMAD2/3 proteins to the cell type and the liver histological context.

In summary chronic liver disorder can result in important alterations in liver architecture and biological functions. Chronic liver inflammation affects all hepatic cells, and our comprehension of the disease evolution should take into consideration that all the different cells are continuously interacting with each other, notably through cytokine signals.



**Figure 12. Cell type-specific temporal dynamics of R-Smad phosphoisoforms.**

Although linker phosphorylation is transient after mitogen treatment of normal epithelial cells (A), mitogen-inducible phosphorylation generally persists in various mesenchymal cells (B). Moreover, constitutive linker phosphorylation is found in almost all types of carcinomas and Ras-transformed cells (C). Because mitogenic pSmad3L signaling is followed by the cytostatic pSmad3C signaling in normal epithelial homeostasis, pSmad2L/C and pSmad3L/C rarely exist in normal epithelial cells (A). Resembling observations in mesenchymal cells (B), carcinomas acquire an invasive phenotype via the pSmad2L/C pathway created by a combination of TGF- $\beta$  signal with intracellular Ras signal (C) (Matsuzaki, 2011).

These cytokine signals will create a specific inflammatory environment that will influence cell fate decisions (such as proliferation or differentiation) and the outcome of the disease.

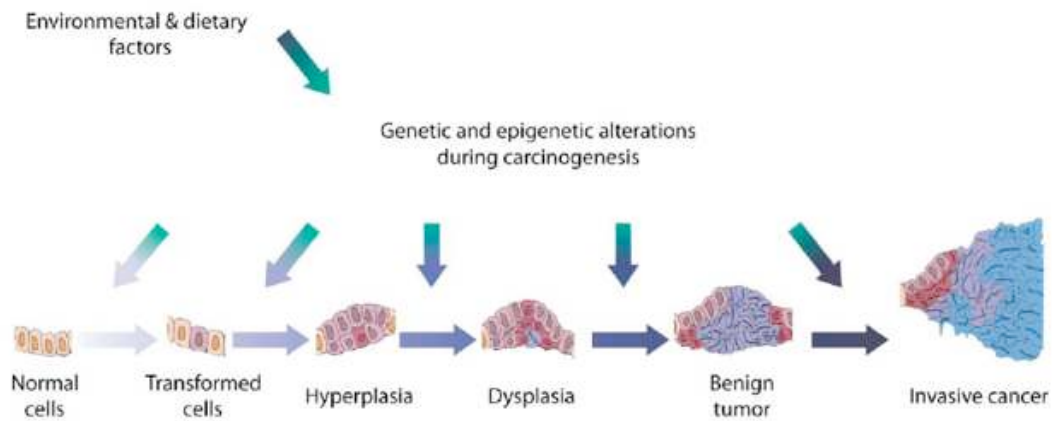
Clinical and epidemiological studies suggest a strong association between chronic infection, inflammation, and cancer (Grivennikov and Karin, 2010; Grivennikov et al., 2010; Lin and Karin, 2007). Indeed, liver cirrhosis represents an ideal predisposing condition for developing hepatocellular carcinoma. The biology of liver cirrhosis is characterized by a constant stimulus for hepatocellular regeneration in a microenvironment characterized by chronic inflammation and altered ECM composition. Abnormal hepatocellular regeneration leading to HCC can be secondary to a step-wise process in which external stimuli induce genetic alterations in mature hepatocytes, thus leading to monoclonal populations that harbour dysplastic and subsequently neoplastic hepatocytes carcinoma (HCC).

### **III. Hepatocellular carcinoma and its links with inflammation**

#### **A. Fundamental concepts on cancer**

##### **1. From hyperplasia to malignant tumor**

Cell growth and differentiation are regular cellular processes, required for the organ development. Alterations in the regulation of these processes can result in loss of control over cell growth, differentiation, and spatial organisation leading to neoplasia or tumor development. Carcinogenesis is a result of stepwise alterations in cellular function (Coleman and Tsongalis, 2009). First, abnormal proliferation after alterations and/or mutations in normal cells that is called hyperplasia. Hyperplasia is considered to be a common and current physiological response to a specific stimulus, and during this process cells remain subject to normal regulatory control mechanisms. On the contrary, in a tumor context, transformed cells proliferate in a non-physiological manner, which is unresponsive to normal stimuli. Then cells are subjected to dedifferentiation, which leads to dysplasia. At the beginning tumor cells retained some of their specialized features and their original morphology are identified as well differentiated (Lodish, 2008). They can thus still be identified as benign since they are well delimited. On the other hand, tumor cells that have lost much of their functions are considered as poorly differentiated. However, although poorly differentiated tumor cells may have underwent an advanced differentiation, their cellular origin may still be recognized through more primitive characteristics. During disease progression, tumor cells can develop the ability to invade surrounding tissues, leading to the appearance of a malignant tumor. The invading ability of the cells can even be extended to other sites within the body ("metastasize") with penetration into the lymphatic vessels ("regional metastasis") and/or the blood vessels ("distant metastasis") (Figure 13). These phenotypic changes confer proliferative, invasive, and metastatic potentials that are the hallmarks of cancer.



**Figure 13 Tumor development.**

Schematic representation of the multistep process of carcinogenesis during which environmental exposure may trigger genetic and epigenetic changes (Herceg, unpublished).

## 2. Tumor classification

Neoplasia encompasses a high number of human diseases with a wide range of characteristics. Therefore, the classification of neoplastic diseases is of great importance for the comprehension, the diagnostic, and the development of appropriate therapies for them.

The broadest classification of tumors uses the embryologic origin of cells. During early embryonic development, three cell lineages are established: ectoderm, endoderm, and mesoderm. All subsequent cells, including adult tumors, can be traced to one of these three cellular origins. As such, cancers can be named as carcinomas if they originate from ectodermal or endodermal tissues and as sarcomas if they originate from mesodermal tissues (Lodish, 2008).

Carcinomas are the most common cancer type and include all the common epithelial tissue cancers such as lung, colon, breast and liver cancers. Sarcomas arise from mesenchymal cell types, which are predominantly connective tissues. Sub-divisions of carcinomas and sarcomas are based on the organ of origin. Progress in gene expression profiling of tumors permitted classification of tumors based on molecular characteristics. Actually, new classification of human tumors based on their gene expression profiles may arise from further research of this area.

In the liver benign and malign tumors can occur. The three common benign tumors are hemangiomas, adenomas and focal nodular hyperplasia. When hemangiomas and focal nodular hyperplasia usually required no treatment, adenomas are typically resected.

Malign tumors comprised cholangiocarcinoma and hepatocellular carcinoma (HCC). Cholangiocarcinoma is relatively rare (annual incidence of 1-2 cases per 100,000 in the Western world) whereas HCC is the most common type of liver cancer. My work has been focused on this pathology. Therefore, more details will be given in the following sections.

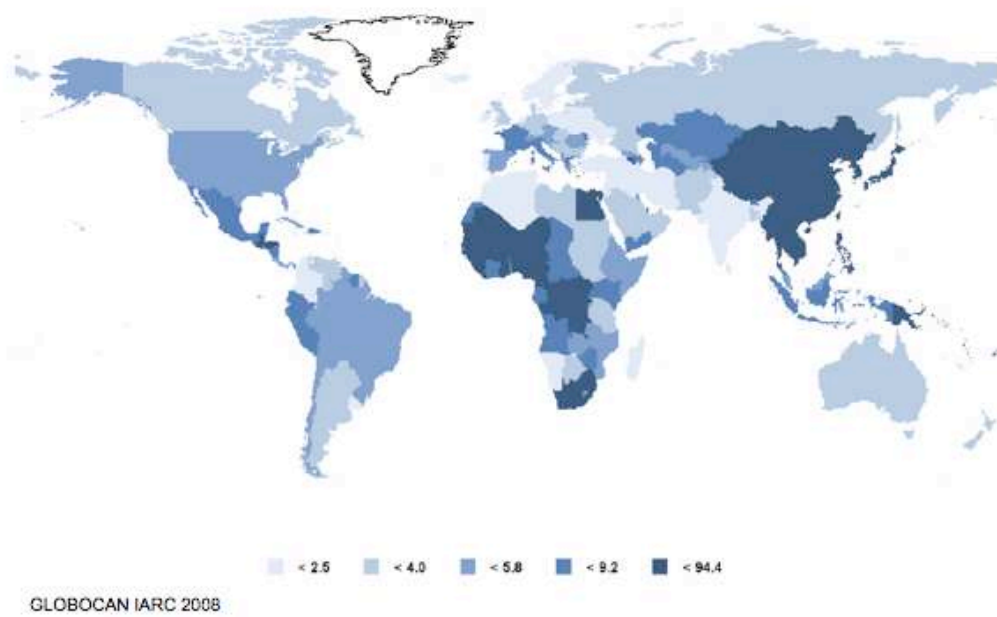
HCC is the most frequent liver tumor, derived from the malignant transformation of hepatocytes. HCC is a major cause of cancer mortality worldwide. Due to late detection, the overall prognosis is generally poor. Understanding the etiology, epidemiology, physiopathology, molecular biology and clinical features of HCC are important to provide appropriate patient care. In addition, understanding the limitations of our current knowledge is crucial to guide future research.

## B. Hepatocellular carcinoma

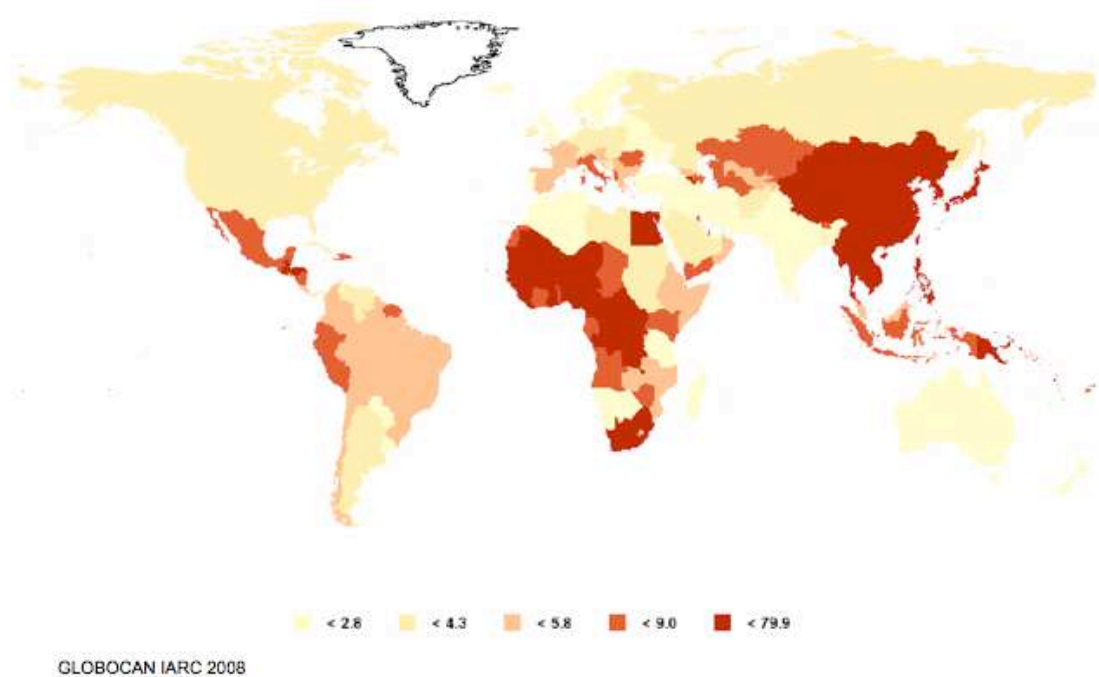
### 1. Epidemiology

Liver cancer is the fifth most common cancer in men (523,000 cases, 7.9% of the total) and the seventh in women (226,000 cases, 6.5% of the total), and most of the burden (85%) occurs in developing countries, and particularly in men: the overall sex ratio male: female is 2.4 (Ferlay et al., 2010).

Regions of higher incidence include Eastern and South-Eastern Asia, Middle and Western Africa, but also Melanesia and Micronesia/Polynesia (particularly in men). Low rates are estimated in developed regions, with the exception of Southern Europe where the incidence in men (10.5 per 100,000) is significantly higher than in other developed regions (Figure 14). There were an estimated 694,000 deaths from liver cancer in 2008 (477,000 in men, 217,000 in women), and because of its high fatality (overall ratio of mortality to incidence of 0.93), liver cancer is the third most common cause of death from cancer worldwide. The geographical distribution of the mortality rates are similar to that observed for incidence (Figure 15) (Ferlay et al., 2010).



**Figure 14. Estimated age-standardized incidence rate per 100000 of liver cancer** (Ferlay et al., 2010)



**Figure 15. Estimated age-standardized mortality rate per 100000 of liver cancer** (Ferlay et al., 2010)

## 2. Risk factors

Most patients with HCC have liver cirrhosis mostly induced by the chronic liver disease's risk factors previously described. 50% of the patients diagnosed for HCC are also infected with hepatitis B virus, with a further 25% infected with hepatitis C virus (Gurtsevitch, 2008). Alcoholic liver disease, non-alcoholic steatohepatitis, intake of aflatoxin-contaminated food, diabetes, and obesity are also known to be major risk factors for HCC development (Fares and Péron, 2013).

HBV infection can stimulate acute and chronic liver disease and is thought to cause HCC via both direct and indirect pathways. Indeed genetic alterations, chromosomal rearrangement and genomic instability can be the direct cause of HBV's DNA integration into the host cell genome (Szabó et al., 2004) or indirect cause associated to the persistent cell's renewal induced by hepatocyte damage and chronic inflammation (But et al., 2008).

HCV infection causes chronic inflammation, cell death, proliferation, and cirrhosis of the liver (But et al., 2008). Thus, HCV-related HCC is found almost exclusively in patients with cirrhosis (But et al., 2008). HCV may cause HCC by various indirect mechanisms including promotion of oxidative stress, upregulation of genes involved in cytokine production and subsequent inflammation, alterations in apoptotic pathways, and tumor formation (Sheikh et al., 2008).

Heavy alcohol intake is the most common cause of liver cirrhosis (Heidelbaugh and Bruderly, 2006) and is a well established risk factor for HCC. The severity of fibrosis and the rate of cirrhosis and HCC development are much higher in patients diagnosed for both HCV infection and alcoholic liver hepatitis than in patients only suffering from HCV infection (Singal and Anand, 2007). The mechanisms by which alcohol acts in synergy with HCV-infection to aggravate liver disease are not fully understood. Nevertheless the dominant mechanism appears to be increased oxidative stress.

HCC in non-cirrhotic livers is rare and mostly occurs as a result of HBV infection, as described earlier (El-Serag and Rudolph, 2007). However, HCC in non-cirrhotic livers can also occur as a result of contamination of foodstuffs with aflatoxin B1 (AFB1) (Wild and Gong, 2010). AFB1 is a mycotoxin produced by the *Aspergillus* fungus that grows readily on food when stored in warm, damp conditions (Abdel-Wahab et al., 2008). When ingested, it is metabolized into the active AFB1-exo-8,9-epoxide, which binds to DNA, to form adducts and cause genomic damage that can promote the tumor formation (Bressac et al., 1991).



### 3. Molecular alterations in HCC

During carcinogenesis, the balance between pathways controlling cell cycle and apoptosis is deregulated; and as a consequence cancer cells gain the capacity to divide indefinitely. These mechanisms are dependent on oncogene expression and/or silencing of tumor suppressor genes (Sulic et al., 2005). Gene expression is regulated either through their DNA sequence (genetic regulation) and/or through the accessibility on the chromatin (epigenetic regulation) (Jones and Baylin, 2002). A complete chapter will be dedicated to epigenetic mechanisms in HCC later in the introduction; the current paragraph will focus mainly on genetic alterations observable in HCC patients.

Hepatocyte transformation occurs with the accumulation of gene alterations related to carcinogenesis. Gene alterations finally cumulate in HCC in order to support, enhance and induce all the cellular processes required for the progression and growth of the tumor. In HCC, several tumor suppressor genes essentially involved in the control of cell cycle have been reported to be mutated, downregulated or inactivated (Shiraha et al., 2013). In this way *TP53*, one of the famous tumor suppressor gene involved in cell cycle arrest, is found mutated in 50% of aflatoxin induced HCC and between 28-42% in non-aflatoxin induced HCC (Bressac et al., 1991; Buendia, 2000; Tannapfel et al., 2001). Two additional recent studies of whole genome or exome sequencing in HCC samples confirmed that mutations are frequently observed in the *TP53* genes (Fujimoto et al., 2012; Guichard et al., 2012). Others cell cycle regulators, like Retinoblastoma protein (Rb) or p16<sup>Ink4A</sup> proteins are inactivated in more than 80% of cases (Azechi et al., 2001). As a last example, the tumor suppressor phosphatase and Tensin homolog (PTEN) protein activity is absent or reduced in 40% cases (Hu et al., 2003). Although the percentage of tumor suppressor gene alterations is lower compared to other solid tumors, it remains a positive contribution for hepatocarcinogenesis. At the opposite end of the spectrum, activation or over-expression of oncogenes appears even less primordial as for example, mutations of the 3 major oncogenes Ras (H-, K- and N-ras) are found in only a few cases (Challen et al., 1992; Stork et al., 1991; Tada et al., 1990).

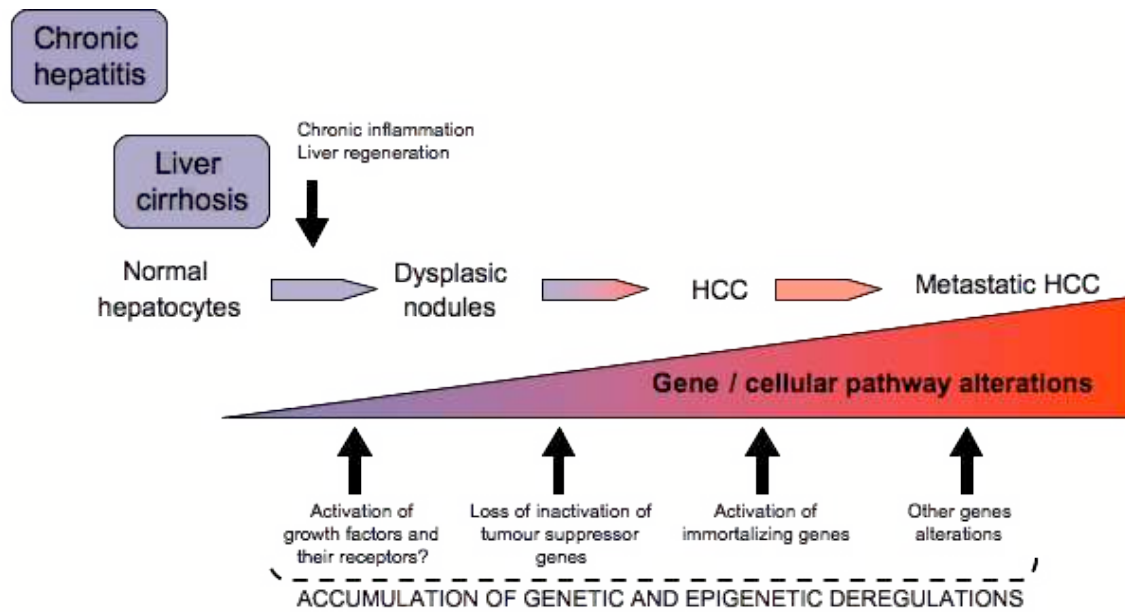
Cell proliferation and tumor growth are also sustained through the reactivation of developmental pathways, notably the Wnt/ $\beta$ -catenin and Hedgehog pathways (Huang et al., 2006; de La Coste et al., 1998; Legoix et al., 1999; Mullor et al., 2002) the activation of these pathways strongly alter the proliferation rate and differentiation of neoplastic cells. This is illustrate by the observation of numerous mutations in genes involved in the Wnt/ $\beta$ -catenin pathway including the gene *CTNNB1* itself (Fujimoto et al., 2012; Guichard et al., 2012). They

act in combination with the expression of several growth factors. For example, TGF- $\alpha$  is expressed in 81% of HCC patient and stimulates hepatocyte proliferation via activation of the EGFR pathway and in a second study mutation in *ERRFI1* (an inhibitor of the EGFR protein) could also contribute to the activation of the pathway (Guichard et al., 2012). Also, alteration of the insulin-like growth factor (IGF)-2 pathway in HCC induces an overexpression of this mitogenic mediator and IGF-2 is even expressed during precancerous lesion stages (De Souza et al., 1995; Yamada et al., 1997). In addition the immortalisation of cancer cells is secure by the maintenance of telomerase activity (found in 90% of HCC cases) (Kojima et al., 1997; Nagao et al., 1999).

Finally, important mutations in the genes coding for proteins that are part of the SWI/SNF complex (such as *ARID1A* and *ARID1B*) were described in HCC samples. In addition others chromatin remodelling complexes harbour mutations in their genes (e.g *SMARCA2*, *SMACB1*) (Fujimoto et al., 2012; Guichard et al., 2012). This type of mutations represents the first hit of a process that will downstream alter the transcription regulation mechanisms for many genes.

All the molecular alterations, including others not detailed in this section, are crucial components of the complex machinery that pilot the initiation and development of hepatocellular carcinoma (Figure 16).

HCC is a complex disease and a better understanding of the underlying mechanisms and the deregulated pathways will bring important information for the development of specific/targeted chemotherapeutic agents that can overcome the mechanisms of drug resistance in the liver. In addition, the delayed prognosis and the lack of appropriate treatment for patients with advanced stages of HCC, highlight the need of to improve patients diagnosis through a better comprehension of the mechanisms implied in the hepatocarcinogenesis.

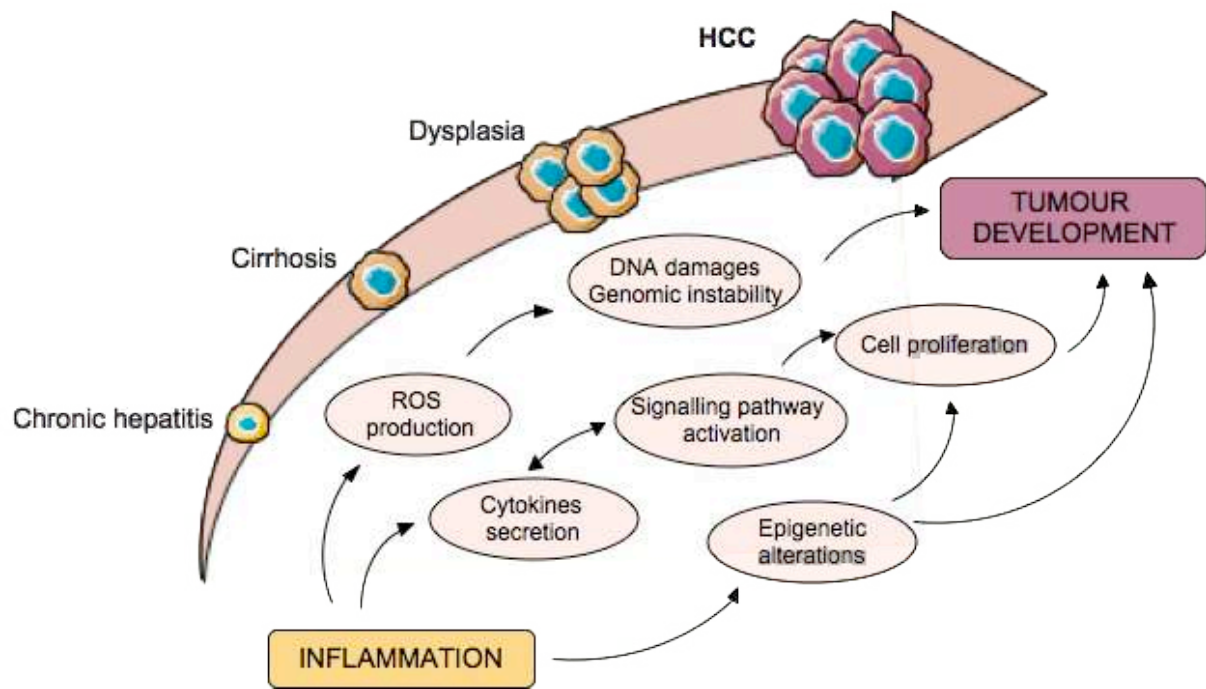


**Figure 16. Sequential gene alterations leading to HCC. (Shiraha et al., 2013)**

### C. From chronic inflammation to hepatocellular carcinoma

#### 1. Inflammatory mechanisms leading to HCC

Epidemiological, pharmacological and genetic evidences provided solid support that inflammation can promote tumor initiation and tumor progression (Grivennikov et al., 2010). Naturally not all types of inflammation lead to cancer: for example, acute inflammation instead contributes to tumor suppression, but as its name indicates, has limited action and evolves rapidly into chronic inflammation. Some mechanisms whereby inflammation promotes tumor initiation have already been mentioned with HCV and HBV risk factor descriptions. All these mechanisms can be grouped into 3 complementary processes (Figure 17): i) induction and/or increase of DNA damage, chromosomal rearrangements and genome instability, ii) perturbation of the proliferation/cell cycle arrest balance iii) epigenetic reprogramming (this section will be developed in future chapters dedicated to epigenetic mechanisms).



**Figure 17. Molecular mechanisms and cellular processes involved in the road from inflammation to tumor initiation.**

In cirrhosis macronodules containing foci of hepatocyte dysplasia are considered to be pre-neoplastic lesions of HCC (Roskams and Kojiro, 2010). In addition, all the cytokines described earlier (i.e.  $\text{TNF}\alpha$ , IL-6, IL-1 $\alpha$  and IL-1 $\beta$ ) and strongly secreted during chronic liver disease are believed to contribute to tumor initiation largely by promoting cell proliferation. Naturally stimulation of cell proliferation alone will not initiate HCC, but associated to carcinogens, inflammatory-induced cell proliferation could make the connection from transformed cells to tumor initiation. Pro-inflammatory cytokines are not the only mediators to be involved in hepatocarcinogenesis, anti-inflammatory cytokines (such as IL-10) are as important to assist tumor initiation via the control of immune surveillance escape (Gonda et al., 2009). Therefore, tumor initiation happened through a delicate deregulation of the pro-inflammatory/anti-inflammatory mechanisms balance.

Concomitantly, inflammation can participate in cancer initiation by promoting DNA damage and genomic instability. These two processes involved in the activation of oncogenes and silencing of tumor suppressor genes are fundamental for cell transformation. Indeed, viral hepatitis, alcohol liver disease and NASH all contribute to the production and accumulation of intracellular ROS (Bartsch and Nair, 2006). ROS can induce DNA damage and genomic instability either directly or indirectly by oxidizing enzymes and proteins involved in mismatched DNA repair. In consequence, hepatocytes harbouring extensive DNA damage

and undergoing prolonged proliferation during chronic inflammation may acquire mutations and growth advantages, thus promoting initiation and progression of hepatocellular carcinoma (Wu et al., 2013; Yan et al., 2009).

Finally inflammation may contribute to cancer development by requisition and activation of hepatic progenitor cells (HPCs). This particular category of cells due to their loss of specialization, and plasticity is indeed more sensitive to transformation.

## 2. Inflammation, hepatic progenitor cells and hepatocarcinogenesis

The observation that HCC cells present specific markers that are common with stem cells and that progression of liver cancer is associated with dedifferentiation (a process by which a specialized, a differentiated cell regresses to a more embryonic and unspecialized form) led to the 'maturation arrest hypothesis', which predicts that liver cancer may arise from stem cells that failed to complete their differentiation (Wu et al., 1996; Yamashita et al., 2008; Yoon et al., 1999). As described before, HPCs are activated when the replication of mature hepatocytes is blocked, in order to take over liver regeneration and repair (Roskams, 2003; Roskams et al., 2003b; Yang et al., 2004). In particular a significant percentage of cirrhotic regenerative nodules are composed of HPC-derived hepatocytes (Lin et al., 2010). Several studies have provided evidence to support the hypothesis of an HPC origin for liver cancer (Knight et al., 2008; Libbrecht, 2006; Tang et al., 2008a). As exposure to different environmental factors can activate inflammation in liver cells, one current model proposes that the inflammatory microenvironment directly promotes HPC activation and transformation. More specifically, IL6, TNF $\alpha$ , IFN $\gamma$  and TWEAK (TNF-like weak inducer of apoptosis, a member of the TNF family), increased the number of rodent HPCs *in vitro* and *in vivo* (Brooling et al., 2005; Knight et al., 2000; Yeoh et al., 2007). In addition, increasing proliferation of HPCs by cytokines is not just a side-effect of inflammation-induced cell proliferation, since the proliferative effects of IFN $\gamma$  and TWEAK on HPCs have been shown to be specific to HPCs (when compared with hepatocytes). Finally in HCC, cells expressing progenitor/ductular markers are more aggressive, chemoresistant and more prone to metastasize (Lee et al., 2006). In this manner, the recruitment of HPCs for hepatocarcinogenesis could be an important feature for the cancer's aggressiveness.

### 3. Creation of an inflammatory microenvironment during HCC

The link between inflammation and HCC is not one-way. If inflammation can promote HCC initiation, the tumor will in turn maintain/create an inflammatory environment to sustain its growth and progression (Grivennikov et al., 2010). To ensure its progression the tumor needs to maintain a cell proliferation rate higher than apoptosis and to hold onto immune surveillance escape. In particular high activation of STAT3 in HCC will not only promote cell proliferation but also induce the secretion of mediators that will impair dendritic cell maturation and lymphocyte T activation (Yu et al., 2007). In the same manner, oncogene activation s not only directly influences cell proliferation but also indirectly contributes to the preservation of a favourable microenvironment by activating the secretions of cytokines involved in inflammation, angiogenesis and metastasis (Mantovani et al., 2008).

mRNA and proteins expression of cytokines in HCC has been demonstrated by immunohistochemistry (IHC), quantitative PCR (qRT-PCR) and ELISA, and compared between tumors versus non tumors samples. Anti-inflammatory (IL-10) and pro-inflammatory (IL-1 $\beta$ , IL-18, TNF- $\alpha$  and IL-6) have all being globally found over expressed in tumors samples compared to healthy tissues, or in plasma of patients (Aroucha et al., 2013; Budhu and Wang, 2006; Jang et al., 2012; Liang et al., 2012). TGF- $\beta$  has been found both lower or higher expressed in tumors in distinct studies (Okumoto et al., 2004; Sasaki et al., 2001; Yuen et al., 2002), underlying its complex function during liver cancer progression (see below). In addition, the levels of cytokines have even been correlated to disease prognosis. For example, IL-6, TNF- $\alpha$  and IL-1 $\beta$  have been linked to the development of metastases (Bortolami et al., 2002; Coskun et al., 2004). High anti-inflammatory levels such as IL-10 and TGF- $\beta$  have been related to shorter free disease, shorter survival period or metastasis (Chau et al., 2000; Hussein et al., 2012; Lee et al., 2012; Li et al., 2012a; Okumoto et al., 2004).

In conclusion, inflammation is not only a path to HCC development it is intimately linked to its evolution. Inflammation and liver cancer disease co-evolving together by continuously regulating each other. As a result the inflammatory landscape is greatly modify between the early and late stages of tumor development, and a cytokine, like TGF- $\beta$ , can display different, even adverse, functions during this development.

#### 4. Evolution of TGF- $\beta$ functions during HCC development.

TGF- $\beta$  is largely overexpressed during hepatic and cirrhotic liver disorders. In HCC cases, either overexpression or downregulation of TGF- $\beta$  itself or of components of the TGF- $\beta$  pathway have been described (Table 3) (Breuhahn et al., 2006). Interestingly, while TGF- $\beta$  receptor type II (TGFBR2) and SMAD4 are commonly found inactivated in several types of carcinoma (Levy and Hill, 2006), in HCC deregulation of the signaling pathway through mutations occurs very rarely (Table 3). TGF- $\beta$  is usually depicted as a suppressive tumor agent (via its cell cycle arrest and apoptotic effects) and inactivation of its signaling pathway could be considered as a strategy of the tumor to bypass its effects. However the fluctuation observed for its regulation in HCC indicates a much more complex role of the TGF- $\beta$  pathway in hepatocarcinogenesis.

**Table 3. Expression of the TGF- $\beta$  pathway components in HCC** (adapted from Breuhahn et al., 2006)

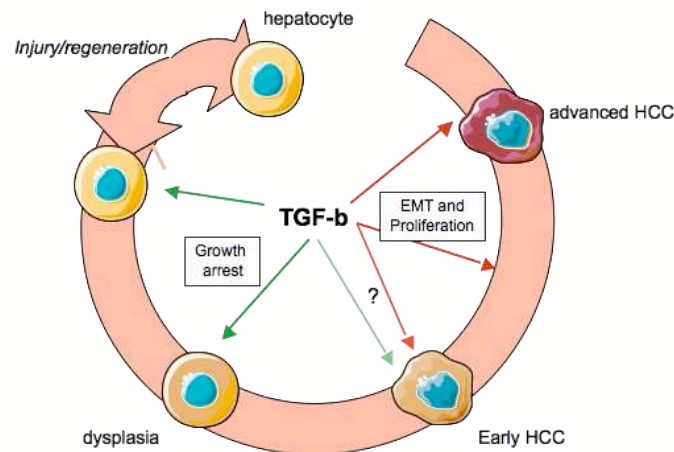
Components	Expression in HCC
<b>TGF-<math>\beta</math></b>	Upregulated in 40%
<b>TBRI</b>	Upregulated in 80% - downregulated in 60%
<b>TB RII</b>	Downregulated in 37-70%
<b>SMAD2</b>	Mutations in 3%
<b>SMAD4</b>	Downregulated in 10%, mutations in 6%
<b>SMAD7</b>	Upregulated in 60% of advanced HCCs

It is true that cell cycle arrest and apoptotic mechanisms triggered by TGF- $\beta$  in hepatocytes would participate in a global anti-tumorigenic effect. In such cases, TGF- $\beta$  operates through the activation of cell cycle inhibitors (e.g. p21 and p15) (Massagué, 2008), the repression of mitogenic agents (e.g. c-myc) (Spender and Inman, 2009) and the stimulation of apoptosis (by interfering with the BIM cell death signaling, the NF- $\kappa$ B anti-apoptotic pathway) (Cavin et al., 2003; Ramesh et al., 2008). This notion is supported in mice models where the decrease of TGFBR2 expression enhances HCC susceptibility (Im et al., 2001). Yet the anti-inflammatory nature of TGF- $\beta$  could also facilitate HCC development by contributing to the immune surveillance escape through modulation of the immune cells' response (Flavell et al., 2010; Yang et al., 2010a).

In the other hand, additional experimental models overproducing TGF- $\beta$  present an increased susceptibility to chemical carcinogens and further HCC development (Factor et al., 1997; Schnur et al., 1999) and persistent high levels of TGF- $\beta$  promote malignancies and

metastases (Padua and Massagué, 2009). In humans, the increase of TGF- $\beta$  even correlates with a decrease in response to effective therapy and TGF- $\beta$  has been proposed as prognostic marker (Ito et al., 1991; Shirai et al., 1994). But above all, TGF- $\beta$ 's pro-tumorigenic effects are mainly grouped under the promotion of epithelial-mesenchymal-transition (EMT) (Gotzmann et al., 2002). EMT designs an orchestrated series of events in which dedifferentiation of epithelial cells occurs by loss of cell-to-cell contacts and the mesenchymal phenotype is acquired by concomitant gain of migratory and invasive abilities (Mikulits, 2009). EMT is essential for numerous developmental processes, wound-healing in fibrotic organs and initiation of metastases in carcinogenesis. EMT allows carcinoma cells to escape the solid tumoral mass and to invade and colonize new sites. TGF- $\beta$  is the main mediator of EMT and processes via the activation of key genes such as *TWIST*, *SNAI-1/2* and *ZEB1/2* and repression of *CDH1* (E-Cadherin) (Inman, 2011). As it is tightly link to metastases, EMT (and by extension, TGF- $\beta$  pro-tumorigenic actions) was traditionally described as advanced/late carcinogenesis stage processes. Opposite roles of TGF- $\beta$  were thus explained by the different stages of carcinogenesis, with early stage associated with a tumor-suppressive function and later stages associated with a tumor-supporting function. This concept is supported by the observation that TGF- $\beta$  does not induces similar intracellular signals in normal or transformed hepatocytes. As an example, the activation of the EGFR signaling pathway and the activation of SNAIL1 are required to inhibit TGF- $\beta$ -induced apoptosis and to enhance EMT (Caja et al., 2007; Franco et al., 2010). But nowadays, deep comprehensive studies on EMT have questioned the idea that this process is associated only with advanced carcinogenesis stages. Cells undergoing morphology changes tightly resembling EMT have been described in early stages of carcinogenesis (Rhim et al., 2012). Thus TGF- $\beta$  could hold at the same time both pro and anti-tumorigenic function (Figure 18). Notably, in HBV infection, one of the initial steps associated with HCC progression is EMT (Cougot et al., 2005). In addition, the observation of some hepatocytes able to respond to TGF- $\beta$  induced EMT during fibrosis, raises the hypothesis that TGF- $\beta$  would induce this phenotype change in hepatocytes in order to escape apoptotic signal (Dooley et al., 2008; Kaimori et al., 2007). This mechanism of apoptotic evasion is naturally fundamental for hepatocarcinogenesis. Finally, TGF- $\beta$  can promote HCC cell proliferation, through modulation of the SMAD3 phosphorylation site. As described earlier, phosphorylation on the linker site will trigger a mitogenic signal. In particular during HBV infection, HBX has shown the ability to shift the phosphorylation on Smad3 linker site and therefore to support growth of HCC cells (Murata et al., 2009).





**Figure 18. Roles of TGF- $\beta$  during multistep hepatocarcinogenesis.**

TGF- $\beta$  inhibits proliferation of pre-malignant hepatocyte. At early stage of HCC, TGF- $\beta$  probably still continues its growth arrest action, but may also initiate in the same time tumor promotion. At advanced stages TGF- $\beta$  clearly support the tumor growth through cell proliferation and EMT (adapted from Yamazaki et al., 2011).

The balance between linker of COOH-tail phosphorylation for SMAD3 activation is one of the proposed mechanisms to explain the switch between tumor-suppressor and tumor-promotor effect of TGF- $\beta$  but in a general manner, this switch between TGF- $\beta$  effects is also the consequence of the multiple genetic and epigenetic changes observed in tumor cell genomes: as examples mutations of the tumor suppressor *TP53* have been described as a trigger for switching TGF- $\beta$  response (Adorno et al., 2009), and epigenetic regulations of *PDGF $\beta$*  and *DAB2* expression are capable in other types of solid tumors to permute TGF- $\beta$  functions (Bruna et al., 2007; Hannigan et al., 2010). But the understanding of TGF- $\beta$  functions during hepatocarcinogenesis remains partial and further studies are required to elucidate its precise roles.

We have reviewed here how inflammation can drive and accompany HCC development. But this specific microenvironment is not the only parameter sustaining tumor growth. Over the past 10 years, the concept of cancer cell hierarchy has greatly evolved and brought out the idea that a small sub-population of cancer cells named “cancer stem cells” harbour unique features that render them indispensable for the tumor development. Such cells have been described in hepatocellular carcinoma, and I will present them in this next section.

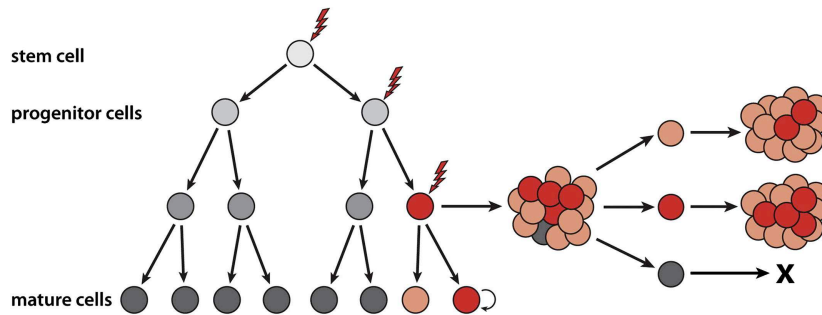
## IV. Cancer stem cells in hepatocellular carcinoma

Traditionally, cancer has been considered as a multistep process defined by the sequential acquisition of key mutations leading to aberrant clonal expansion of a cell. However, recent progress in basic research has transformed this concept at different levels. First, the role of the tumor microenvironment has been well described and is now fully recognized (Lin and Karin 2007; Schafer and Brugge 2007) in contexts such as inflammation. Second, the role of epigenetic deregulation, in combination to genetic aberrations, in most human tumors is more and more striking. Third, a "cancer stem cell" model of tumorigenesis has been strongly supported by experimental evidence. This model suggests that tumors are sustained in their development by a small subpopulation of tumor cells harboring "stem-like" properties.

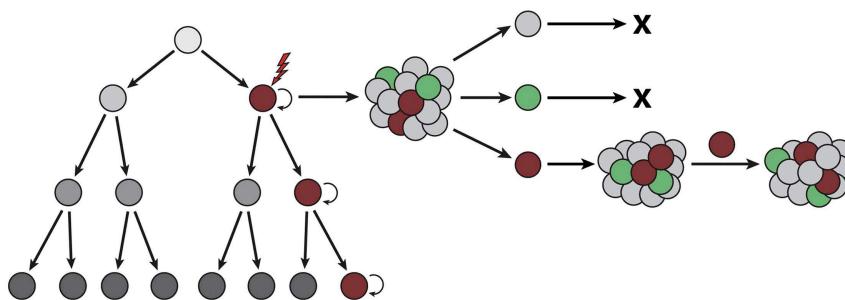
### A. Cancer stem cells concept

Cancer stem cell (CSC) is an operational term to functionally define a distinct subpopulation of tumor cells that present aberrant abilities for self-renewal, proliferation and differentiation (Stingl and Caldas 2007; Visvader and Lindeman 2008). Indeed, classical models of carcinogenesis can be described as "stochastic" or "random," in which any cell in an organ, such as the liver, can be transformed by acquisition of the right combination of mutations (Martinez-Clement, Andreu et al. 2006). As a result, the tumor mass can present some heterogeneity (mainly represented by genetic variations) but cells in the dominant clonal population would possess similar tumorigenic potential and would lead the tumor growth (Figure 19). In consequence, strategies designed to treat and ultimately cure these cancers require killing all these malignant cells. Inversely, the cancer stem cell hypothesis is a fundamentally different model. This model proposes a hierarchical organization, similar to what occurs in healthy tissue with stem cells, where a small subset of cells would be responsible for the tumor development and its cellular heterogeneity. CSCs would thus share with stem cells the ability of self-renewal and differentiation.

### A Clonal evolution model



### B Cancer stem cell model



**Figure 19. The clonal evolution model versus the cancer stem cell model.**

(A) The clonal evolution model is a non-hierarchical model where mutations arising in tumor cells confer a selective growth advantage. Depicted here is a cell (red) that has acquired a series of mutations and produced a dominant clone. Tumor cells (red and orange) arising from this clone have similar tumorigenic capacity. Other derivatives (grey) may lack tumorigenicity due to stochastic events. Tumor heterogeneity results from the diversity of cells present within the tumor. (B) The cancer stem cell model is predicated on a hierarchical organization of cells, where a small subset of cells has the ability to sustain tumorigenesis and generate heterogeneity through differentiation. In the example shown, a mutation(s) in a progenitor cell (depicted as the brown cell) has endowed the tumor cell with stem cell-like properties. These cells have self-renewing capability and give rise to a range of tumor cells (depicted as gray and green cells), thereby accounting for tumor heterogeneity (Visvader and Lindeman, 2012) .

From an experimental point of view, CSCs are usually characterized by a specific combination of one or several extracellular marker(s) (Table 4) and the properties mentioned above are tested via 3 “operational definitions” (Table 5): a specific sub-population within a tumor can be called CSCs if they i) present a superior tumorigenic ability (compared to non cancer CSCs) via *de novo* tumor formation in xenograft model (this assay can be completed or replaced by a clonogenic assay through *in vitro* sphere formation in low attachment conditions), ii) the tumor, if reconstituted should present the same heterogeneity as the original tumor (reflecting the ability to differentiate) iii) CSCs from the new reconstituted tumor should be able to support further transplantation assays (reflecting self-renewal). The xenograft assay is by far the most common assay used to define a sub-population as CSCs. Based on one or several extracellular markers, the subpopulation expressing this (these) marker(s) should present a high capacity to propagate tumor in an immunodeficient

**Table 4. Cancer stem cells markers in different tumors.( adapted from Yi et al., 2013)**

Cancer type	CSC markers
Hematopoietic neoplasms	
Acute myeloid leukemia	CD34 <sup>+</sup> CD38 <sup>-</sup> CD34 <sup>+</sup> CD71 <sup>-</sup> HLA-DR <sup>-</sup> CD44 <sup>+</sup> ALDH <sup>+</sup>
Acute lymphoblastic leukemia	CD34 <sup>+</sup> CD19 <sup>-</sup> or CD34 <sup>+</sup> CD10 <sup>(-)</sup> CD34 <sup>+</sup> CD4 <sup>-</sup> or CD34 <sup>+</sup> CD7 <sup>-</sup>
Solid tumors	
Breast cancer	CD44 <sup>+</sup> CD24 <sup>-/low</sup> ALDH <sup>+</sup>
Brain cancer	CD133 <sup>+</sup>
Prostate cancer	CD44 <sup>+</sup> α <sub>2</sub> β <sub>1</sub> <sup>hi</sup> CD133 <sup>+</sup> CD44 <sup>+</sup> , CD44 <sup>+</sup> α <sub>2</sub> β <sub>1</sub> <sup>+</sup> CD133 <sup>+</sup> CXCR4 <sup>+</sup>
Colon cancer	CD133 <sup>+</sup> EpCAM <sup>+</sup> CD44 <sup>+</sup> CD166 <sup>+</sup> ; ALDH <sup>+</sup>
Melanoma	CD20 <sup>+</sup>
Liver Cancer	CD133 <sup>+</sup> , CD133 <sup>+</sup> ALDH <sup>+</sup> CD90 <sup>+</sup> CD45 <sup>-</sup> CD44 <sup>+</sup>
Pancreatic Cancer	CD44 <sup>+</sup> CD24 <sup>+</sup> EpCAM <sup>+</sup> CD133 <sup>+</sup> CXCR4 <sup>+</sup>
Head and neck cancer	CD44 <sup>+</sup>

**Table 5. Functional assays to assess cancer stem cells properties.**

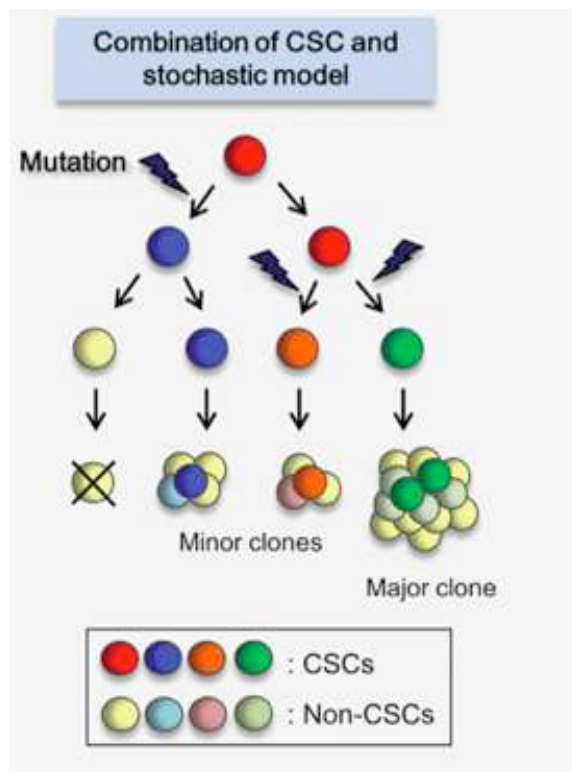
Self-renewal, differentiation capacity and tumor initiation are considered like the 3 fundamental properties of CSCs. Chemoresistance is a supplementary characteristic that has nevertheless been described for many CSCs ( adapted Marquardt et al., 2010)

Property	Definition	Assay
<b>Self-renewal</b>	The ability to undergo symmetric division and thereby indefinitely replenish itself	Re-plating assays. Serial transplantations
<b>Differentiation capacity</b>	The ability to undergo asymmetric division and thereby recapitulate all tumor cell types	Differentiation assays <i>in vitro</i> . Transplantation
<b>Tumor initiation/metastasis</b>	The ability to propagate tumor when transplanted into the proper environment	Sphere formation. Invasion assays. Transplantation
<b>Relapse</b>	The property of resistance to different therapies and the ability to relapse	Chemo/radio-resistance assays

mouse with only a limited number of cells. For example, only 100 CD133+ cancer cells are able to reconstitute a medulloblastoma in NOD/SCID mice whereas 10,000 CD133- were not able to produce any tumor (Calabrese et al., 2007). The first CSCs were described in acute myeloid leukemia (AML) almost 20 years ago (Bonnet and Dick, 1997; Lapidot et al., 1994), since then CSCs have been identified in several others types of solid tumors including, breast (Al-Hajj et al., 2003), liver (Suetsugu et al., 2006), pancreas (Lee et al., 2008), ovarian (Szotek et al., 2006), prostate (Collins et al., 2005), brain (Singh et al., 2003) and colon cancers (O'Brien et al., 2007).

CSCs have been further shown to present additional characteristics such as the expression of ATP-binding cassette transporters (ABC transporters) responsible for drug efflux in the cell (Gatti et al., 2011). In consequence, CSCs present higher resistance to chemotherapy and to irradiation (Grotenhuis et al., 2012). It is then easy to understand that the discovery of this new sub-population generated great enthusiasm because they provided an explanation for chemoresistance and cancer relapse.

In appearance simple, the CSC theory is however complex and source of many controversies. Indeed in the absence of a precise definition (despite an operational characterization that is only rarely fully achieved in every study), a clear classification for CSCs remains impossible. As presented in Table 4, each tissue presents putative CSCs with different extracellular markers and even in one specific tissue several different sub-populations have been described (e.g. ovarian cancer and AML). This heterogeneity within CSC populations can be derived from technical variations used for their study (e.g. cultured vs. fresh sorted cells, extensively passaged vs. early xenograft cells etc.) but also from intra tissue multiple CSC pools (Visvader and Lindeman, 2012). Indeed CSC and clonal evolution models are not mutually exclusive. As presented in Figure 20, within individual cancer patients CSCs can acquire different alterations and became genetically heterogeneous. Finally CSCs heterogeneity can also come from the plasticity of cancer cells that could dedifferentiate and re-acquire a stem cell like phenotype (Figure 20) and generate a second type of CSCs. This dedifferentiation has mainly been described *in vitro*, but several studies presented a stochastic transition between the two states (CSCs and non-CSCs) likely to maintain equilibrium between cell populations (Chaffer et al., 2013; Yang et al., 2012). Such balance between stem and differentiated cells has already been reported in healthy tissue like in mouse testis (Barroca et al., 2009), hence a similar regulation between pluripotency and differentiation could also occur in cancer. The status of CSCs is thus complex and is in constant evolution with the progresses in cancer and stem cell research.



**Figure 20. Combination of the CSC and the clonal evolution models**

A combination of the CSC model and the stochastic (clonal evolution) model has been proposed to account for clonal diversity of CSCs. Each CSC clone is thought to evolve through the acquisition of genetic mutations. Phenotypically and functionally distinct major clones and minor clones may exist in a tumor. Each clone is organized into a hierarchical structure (Sugihara and Saya, 2013).

The last trait subject to discussion is the nomenclature of “cancer stem cells”. CSCs have been named after stem cells because they share with them fundamental properties such as ability to differentiate into heterogeneous lineages and self-renewal. But they also present some differences, mainly that the equilibrium between proliferation, differentiation and apoptosis that characterize regular stem cell is lost in CSCs where the unbalanced cell growth serves exclusively to form of tumor mass (Sampieri and Fodde, 2012). The “cancer stem cell” designation should thus not be confused with a transformed stem cells or an immortalized stem cell. This would imply that CSCs are authentic stem cells, while stem cells and CSCs only shared some properties. And even in the case where CSCs would originate from somatic stem cells, it is very likely that some of the stem cell properties would be lost or altered during the transformation. CSCs have been designed like this to illustrate that they are localized at the base of the pyramidal differentiation process that will construct the tumor (Figure 19). To avoid confusion, CSCs are also called tumor-initiating cells (TICs), this appellation being more in accord with the operational assay used for their definition. However, even this last appellation has been subject to controversy, as the CSCs injected into

to SCID/NOD mice are already initiated and thus will not initiate a new tumor but will rather propagate the tumor they are originated from (Sampieri and Fodde, 2012). Despite this, CSCs and TICs are the most common designations and are often used in a synonymous way. In the present manuscript, the appellation “cancer stem cells” will be used to design this specific sub-population.

Finally it should be underlined that not all cancers develop sustained by CSCs. In melanoma in particular, the high proportion of tumorigenic cells (up to 50%) and the wide spectrum of marker argue against a CSC model for the tumor heterogeneity (Quintana et al., 2010).

## B. Identification of liver cancer stem cells

While liver progenitor cells have been studied for more than 15 years, the observation of cells harboring stem cell properties in hepatocellular carcinoma is much more recent. The first observations date from 2006 by Suetsugu *et al.* describing that CD133+ cells in HCC cell lines have a higher proliferative potential, express a lower level of mature hepatocyte mRNA and most importantly, present a great tumorigenic potential compared to CD133- cells. Since then, numerous investigations have divulged other markers characterizing CSCs in HCC (Tong et al., 2011). Among all the extracellular markers (listed in Table 6), the most common are CD133, CD90, CD44 and the epithelial cell adhesion molecule (EpCAM). More recently CD13 has been identified as a marker for dormant/quiescent CSCs and associated with CD90 and CD133 expression after CSC activation (Haraguchi et al., 2010). Oval cell (OV)-6, delta-like 1 homolog (DLK1) and CD24 have also been identified as potential liver CSCs but have not been deeply exploited (Salnikov et al., 2009; Xu et al., 2012; Yang et al., 2008a). Interestingly, two functional markers have also been used to characterize CSCs in HCC cell lines: the enzymatic activity of aldehyde deshydrogenase (ALDH) involved in detoxification, oxidative stress metabolism and drug resistance (Ma et al., 2008a) and the high expression of ATP binding cassette (ABC) transporters conferring on them a higher ability to efflux xenobiotic substances (Jia et al., 2013; Zhu et al., 2010). This last ability was traditionally visualized after Hoechst 3342 dye staining by fluorescence activated cell sorting (FACS) where CSCs are discerned as a side population (SP) that incorporate the staining less (Chiba et al., 2006). SP was actually one of the first parameters used to characterize CSCs in HCC cell lines (and in other types of cancer) but it was quickly less used in favor of extracellular markers.

**Table 6. Cell surface marker for liver CSCs. ( adapted from Yamashita et al., 2013)**

Cell surface marker	Phenotypes of marker-positive CSCs (source)
CD13	Tumorigenic, cell cycle arrest, chemoresistant (PLC/PRL/5, Huh7, Hep3B)
CD133	Tumorigenic, chemoresistant (PLC8024, Huh7, Hep3B, primary HCC)
CD24	Tumorigenic, chemoresistant, metastatic (PLC/PRL/5, HLE, Huh7, primary HCC)
CD44	Tumorigenic, invasive (PLC, PLC/PRL/5, HLF)
CD90	Tumorigenic, metastatic, circulating (HepG2, Hep3B, PLC, Huh7, MHCC97L, MHCC97H, primary HCC)
DLK1	Tumorigenic, chemoresistant (Hep3B, Huh7)
EpCAM	Tumorigenic, invasive, chemoresistant, circulating (Huh1, Huh7, primary HCC)
OV6	Tumorigenic, chemoresistant, invasive, metastatic (Huh7, SMMC7721, primary HCC)

The relevance of these markers was tested through the classical assays described earlier: proliferation capacity, clonogenic potential and tumorigenic potential (Haraguchi et al., 2010; Kimura et al., 2010; Suetsugu et al., 2006; Yang et al., 2008b; Yin et al., 2007; Zhu et al., 2010). Some studies went further and also investigated the expression of genes related to stem cells (e.g. *NANOG*, *SOX2*, *OCT4*), the chemoresistance, the invasiveness and the metastatic potential (Kohga et al., 2010; Song et al., 2008; Tomuleasa et al., 2010).

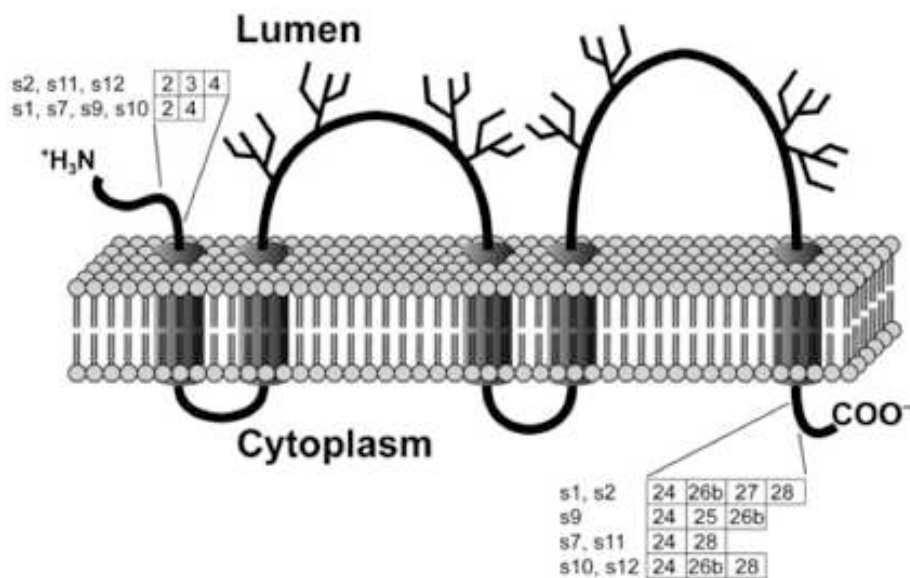
In order to increase the accuracy, some markers were used in combination such as CD44+/CD90+, CD133+/ALDH+, CD133+/EPCAM+ and CD133+/CD44+ (Chen et al., 2012b; Ma et al., 2008a; Yang et al., 2008b; Zhu et al., 2010). This combination of markers demonstrates that CSCs co-expressing two markers are usually more aggressive and more tumorigenic than cells expressing only one marker. But probably due to technical limitations, there is no report investigating the expression of three or more markers and even the combination of two markers seemed to limit the possibilities of biological exploration of CSCs. Therefore, further investigations to clarify the characterization of CSCs are required to refine markers that can be used for their identification.

Among these different extracellular markers, CD133 is by far the most used and CD133+ cells have been subject to numerous investigations to decipher their functions in hepatocellular carcinoma. My work concentrates on this particular CSC sub-population and in the next section I will describe in more detail CD133+ cells in liver cancer.



### C. CD133+ cells as liver CSCs.

CD133 (PROM-1) is a five transmembrane glycoprotein that belongs to the prominin family (Figure 21) (Miraglia et al., 1997). It is encoded by up to 27 exons of the *PROM1* gene located on chromosome 4 which are, like for the murine homologue *Prominin-1*, subject to alternative splicing (Maw et al., 2000). At least seven isoforms (s1, s2, s7, s9, s10, s11, and s12) of 825–865 amino acids in length can be generated in that way in humans (Fargeas et al., 2007; Yu et al., 2002). Its complex gene transcription is controlled in a tissue-specific manner by five alternative promoters, P1–P5, generating at least 16 alternative splicing patterns of the 5'-UTR of CD133 transcripts (Shmelkov et al., 2004). In several tissues including kidney, pancreas, colon, and liver, transcription of the *PROM1* gene initiates from both P1 and P2 (Shmelkov et al., 2004). Although the physiological function of CD133 remains to be elucidated its preferential localization in highly curved plasma membrane protrusions suggests that this protein plays a role as an organizer of the plasma membrane of cellular protrusions (Corbeil et al., 2001; Weigmann et al., 1997).



**Figure 21. Membrane topology of human CD133.**

The N-terminal domain is located outside the cell (lumen), whereas the C-terminal one is within the cytoplasm. Five transmembrane segments are drawn as cylinders, and potential N-glycan structures present in the large extracellular loops ( $\approx 250$  amino acid residues) appear as forks. The presence or absence of a particular exon within the open reading frame is presented with the name of the respective splice variant (named s1-12). Numbering of the exons is such that exon 1 bears the translation start codon. (Grosse-Gehling et al., 2013).

## 1. CD133+ cells as representative population of cancer stem cells

CD133 is primarily known as a marker of adult stem cells for hematopoietic stem cells, endothelial progenitor cells, neuronal and kidney stem cells (Bussolati et al., 2005; Miraglia et al., 1997; Richardson et al., 2004; Uchida et al., 2000; Yin et al., 1997). CD133+ cells were then described as cancer cells presenting specific properties, close to stem cells, distinguishing them from the rest of the cancer cell population (Suetsugu et al., 2006; Yin et al., 2007). CD133 is nowadays used as a CSC marker in several tumors including brain cancer (Singh et al., 2003), prostate cancer (Dalerba et al., 2007), ependymoma (Poppleton and Gilbertson, 2007), colon cancer (Chu et al., 2009), lung cancer (Tirino et al., 2009), laryngeal cancer (Wei et al., 2009), ovarian cancer (Baba et al., 2009) and pancreatic cancer (Olempska et al., 2007).

As described in the previous section, in liver cancer cell lines and in liver cancer samples CD133+ cells were identified as putative liver CSCs through different functional assays. In particular as few as 1000 CD133+ cells from liver cancer were sufficient to induce tumor in NOD/SCID mice, while CD133- do not possess this tumorigenic potential (Yin et al., 2007). In addition the reconstituted tumor presented less than 1% of CD133+ cells, reflecting the original phenotype of the tumor (Ma et al., 2007). In HCC cell lines, CD133+ cell frequency varies from 0% to 95% (Haraguchi et al., 2010; Kohga et al., 2010; Marquardt et al., 2010). In liver cancer specimens, CD133+ cells were detected in all tissues from small studies and in an average 25% of samples from larger studies. CD133+ cell frequency in HCC tissues is usually quite low and does not exceed 5% (Kim et al., 2011; Kohga et al., 2010; Ma et al., 2007; Sasaki et al., 2010; Yin et al., 2007). Interestingly CD133+ cells were also observed in cirrhotic tissues but not in healthy liver patients (Yin et al., 2007).

CD133+ cells display increased capacity for tumorigenesis, self-renewal and sphere formation and the protein CD133 could be not just a marker, but actually contribute to this particular phenotype as suggested in a study by Tong et al (2012). They inactivated CD133 expression through lentiviral based shRNA in PLC8024 HCC cells and observed that inhibition of CD133 expression correlates with a decrease in the ability of sphere formation, self-renewal and tumorigenesis capacities.

CD133+ cell population has been shown to be heterogeneous and can be further sub-divided via co-expression with other CSC markers. In several HCC cell lines, CD44+ cells are all comprised within the CD133+ cell population, but the CD44+/CD133+ cell sub-population is more aggressive and more tumorigenic than the CD44-/CD133+ cell population. CD133+/CD44+ cells also exhibit higher chemoresistance (due to up-regulation of ABC transporters) and higher stemness gene expression (Zhu et al., 2010). ALDH activity can also discriminate the CD133+ cell population (Ma et al., 2008a). ALDH seems to confer

chemoresistance to CD133+ cells and a hierarchical organization for tumorigenicity between the different subpopulations has been established with CD133+/ALDH+ > CD133+/ALDH- > CD133-/ALDH-.

## 2. Clinical significance of CD133+ cells in HCC

In complement to the observation that CD133+ cells can initiate/promote HCC, CD133+ cells seem to be implicated in angiogenesis and metastasis in HCC. CD133+/CD44+ cells from HCC specimens presented a high association with portal vein metastasis (Zhu et al., 2010). Another CD133+/CD24+ cell subpopulation were defined as a metastatic subpopulation (Lee et al., 2011) and finally co-staining of CD133 and ALDH activity in HCC samples were localized in the area adjacent to connective tissue and within invaded vessels, suggesting that these cells could be metastatic (Lingala et al., 2010). These phenotypic differences within the sub-population involved in angiogenesis indicate that CSCs initiating HCC may not be exactly similar to CSCs involved in metastatic progression (this hypothesis is under discussion for other type of CSCs, Visvader and Lindeman, 2012), but in any cases the phenotype of metastatic CSCs seems to always include CD133 expression.

Taken together these observations strongly insinuate that CD133+ cells not only initiate HCC but also participate in its evolution, and thus could be used as a clinical marker for disease evolution and patient prognosis. Song *et al.* (2008) were the first to explore the association between CD133 expression and clinical parameters. They described that the presence of CD133+ cells positively correlates with higher pathological grading and with poor prognosis. Several studies further confirmed these correlations: CD133 expression (assessed by qRT-PCR) was associated with advanced disease stage, higher recurrence and worse overall survival (Ma et al., 2010; Sasaki et al., 2010) and in another study CD133 with other stem cell markers such as Nestin, CD44, ABCG2 was identified as a significant predictor for overall survival and relapse-free survival (Yang et al., 2010b). Lastly CD133 expression was correlated with recurrence rate after surgical therapy (Zen et al., 2011). Although these relations between CD133 expression and HCC evolution provide strong support to use CD133+ cells as predictor/marker for patient outcomes, it should be mentioned that one study conducted by Kim et al, did not observe any correlations between CD133 expression and pathological parameters (Kim et al., 2011).

After the observation that CD133+ liver CSCs are involved in tumorigenesis, self-renewal, chemoresistance, proliferation, metastasis and are linked to the disease evolution, these cells

have become a preferred target for the research of new cancer therapy. But in order to properly and efficiently target CD133+ cells in liver cancer, research studies have tried to determine their molecular characterization and the signaling pathways that are sustaining their biological actions.

### 3. Molecular characterization and biological functions active in CD133+ cells.

Conforming to their analogy to stem cells, CD133+ liver CSCs display signaling pathways and transcriptional pattern involved in pluripotency. Transcription factors involved in the maintenance of pluripotency such as OCT4 (POU5F1), SOX2, NANOG and BMI-1 has been reported to be higher expressed in CD133+ cells (Ma et al., 2010; Machida et al., 2009; Tomuleasa et al., 2010). These observations did not only concerned HCC cell lines but also in human-sample-derived CD133+ spheres. CD133+ cell's phenotype is tightly linked to the expression of these stemness genes as any treatment or stimulus that leads to the decrease of CD133+ cells is followed by a decrease in stemness gene expression (Chiba et al., 2008; Ma et al., 2010). As evidence of their active role in the stemness phenotype observed in CD133+ cells, the inhibition of either NANOG or OCT4 results in reduced tumorigenicity and self-renewal abilities (Lee et al., 2011; Yuan et al., 2010).

Contributing also to the homeostasis of CD133+ cells, the Wnt/ $\beta$ -catenin, Hedgehog and Notch developmental signaling pathways are activated in this population (Ma et al., 2007; Marquardt et al., 2010). Through genome micro arrays that analyzed the expression pattern of CD133+ cells, several downstream components of these pathways have been reported to be up-regulated (Tang et al., 2012). In particular the gene encoding for  $\beta$ -catenin, NOTCH and Smoothed (essential initiator components of respectively, the Wnt, Notch, and Hedgehog signaling pathways) are directly concerned by this transcriptional increase. These pathways are known to be fundamental in embryonic and adult stem cell regulation, and in CSCs could contribute to cell fate decisions (such as EMT initiation), proliferation and apoptosis (Takebe et al., 2011). Interestingly a recent finding reported that the deacetylase HDAC6 can physically interact with CD133 and  $\beta$ -catenin to form a ternary complex that regulates the activation of the Wnt/ $\beta$ -catenin signaling pathway (Mak et al., 2012). The protein CD133 is therefore directly implicated in the activation of this signaling pathway as any downregulation of CD133 leads to the acetylation of  $\beta$ -catenin and its further degradation. In turn, this degradation correlates with decreased proliferation *in vitro* and tumor xenograft growth *in vivo*. This exciting discovery not only supports Wnt/ $\beta$ -catenin

having a fundamental role in CD133+ CSCs induced tumorigenesis but also that CD133 is not only a marker for liver CSCs but could actively contribute to the specific phenotype of liver CSCs. Genomic microarray comparison between CD133+ and CD133- cells in Huh7 and PLC8024 further identified 149 genes differentially expressed, including several genes from the IL-8/CXCL1 signaling pathway (Tang et al., 2012). Increased expression of IL-8 in CD133+ cells activates in turn a feedback loop involving the activation of MAPK pathway. These signals support the proliferation of CD133+ cells and neutralization of IL-8 results in inhibition of CD133+ cell self-renewal, tumorigenesis and angiogenesis. Moreover the inhibition of CD133 protein itself lead to decreased IL-8 production and abolished CSC properties, supporting again the hypothesis that the CD133 protein plays an active role in the liver CSC phenotype. Additional signaling pathways are implicated in CD133+ cell tumorigenesis ability. A correlation between CD133 expression and JNK phosphorylation, for example, can be observed and inhibition of JNK activation highly reduces tumor xenograft assay efficiency (Hagiwara et al., 2012). In an opposite manner, inhibition of the mTOR pathway facilitates the growth of HCC by modulating CD133 homeostasis: mTOR inhibition promotes the conversion of CD133- in CD133+ cells and stemness gene expression (Yang et al., 2011). Reactivation of mTOR signaling is at the opposite followed by CD133 expression decrease.

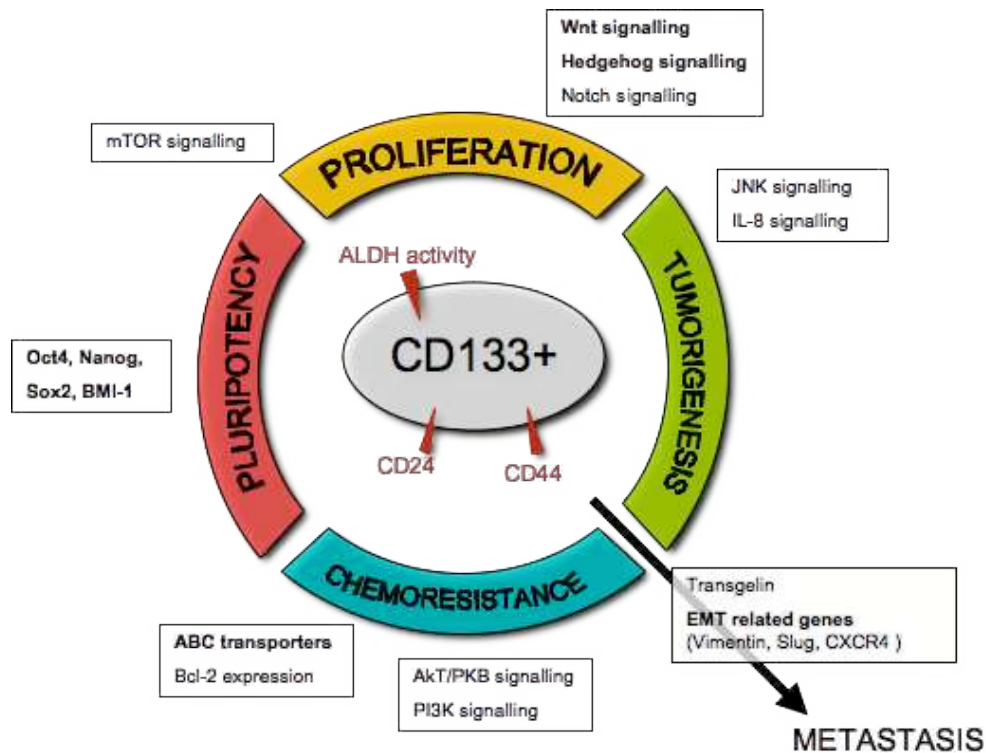
Increased proliferation capacity of CD133+ cells, also derived from their higher chemoresistance and their ability to expulse from the cytoplasm any drugs and xenobiotic substances in opposition to their counterpart CD133- cells. The ABC transporter family members are involved in the transport across external and internal membranes of, among others, metabolites and drugs (Kerr et al., 2011). A higher level of *ABCG2* and *ABCB1* mRNA has been found in CD133+ cells (Ma et al., 2010) and immunostaining revealed a co-expression of *ABCB5* with CD133 and EpCAM (Cheung et al., 2011). The importance of *ABCB5* transporter has been illustrated by the observation that its inhibition further blocked the expression of CD133 and EpCAM proteins and that *ABCB5* expression has been correlated with a higher recurrence rate in patients who had undergone curative partial hepatectomy. CD133+ cells resistance to drugs like doxorubicin and 5-fluoracil has also been demonstrated and the AKT/PKB pathway and BCL-2 signaling pathway would be involved in this chemoresistance process (Ma et al., 2008b). This hypothesis is supported by two observations: first under drug treatment, BCL-2 and phospho-AKT co-localized with CD133 and second, the administration of AKT inhibitor

reduced the expression of survival related proteins. This interaction plays an important role in homeostasis and chemoresistance of CSCs.

In addition to drug chemoresistance, CD133+ cells also exhibit a more efficient resistance to irradiation: after exposition, CD133+ cells display an activation of the MAPK/PI3K signaling pathway, a reduction of ROS production, a greater post-radiation proliferation and a lower radiation induced apoptosis (Piao et al., 2012). These mechanisms could contribute to radioresistance employed during therapy and remaining CD133+ CSCs may be further reactivated and initiate a relapse of the disease. As for the Wnt- $\beta$  catenin pathway, CD133 can directly interact with PI3K regulating subunit (through phosphorylation on its tyrosine 828) and therefore directly modulate the activation of this pathway (Wei et al., 2013).

Finally several studies bring to light molecular mechanisms involved in CSC mediated EMT and metastasis. In Huh7 cells, metalloproteinase MMP-2 and ADAM9 are found up regulated in CD133+ compared to CD133- cells (Kohga et al., 2010). Metalloproteinases facilitate cellular invasion and metastasis and their activation in CD133+ cells was confirmed in PLC/PRF/5 HCC cell lines where the knockdown of CD133 results in a decrease in MMP-2 and ADAM9 expression. A proteomic comparison between CD133+ and CD133- cells revealed one single higher expressed protein in CD133+ cells, the transgelin, a cytoskeleton associated protein involved in TGF- $\beta$ /SMAD3 associated migration (Lee et al., 2010a). SiRNA directed against transgelin results in invasiveness capacity diminution. In addition, expression of proteins involved in EMT process is deregulated in CD133+ Huh7 cells: E-Cadherin is down-regulated while Vimentin, SLUG, SNAIL, TWIST (active contributor to EMT) and CXCR4 (contributor to cell migration) are strongly up-regulated (Lee et al., 2010a; Na et al., 2011; Tsai et al., 2012). This expression pattern has not been completed with functional assays, but it suggests that CD133+ cells will be more sensible to EMT initiation than CD133- cells.

Taken together, these molecular mechanisms (summarized in Figure 22) are important elements to understand the complexity of CD133+ cell biology. They represent promising targets for further CSC-based cancer therapies. CD133+ cell's phenotype is thus represented by a specific panel of gene expression that must be itself supported by specific genomic and epigenomic profile and may be regulated by the tumor microenvironment.



**Figure 22. Biological processes and molecular signaling in CD133+ liver cancer stem cells.** Molecular characteristics can be grouped into four major biological processes (proliferation, pluripotency, tumorigenesis and chemoresistance) and it should be noted that signaling pathways are usually involved in more than one specific processes.

#### D. Influence of the microenvironment on CSCs

##### 1. Cancer niches support and maintain CSC activation

When a tumor develops within a tissue, differentiated cells are not the only component of the tissue to be affected by tumoral transformation. Cancer affect the entire environment and induces structural and functional modifications in the extracellular matrix, the fibroblasts, the vascularization architecture, and cancer-associated inflammation will mobilize immune cells and initiate the liberation in the microenvironment of a panel of various cytokines and growth factors that in turn will influence the structures and functions of all the components, including CSCs. 120 years ago Paget proposed a “seed and soil” hypothesis for metastasis. In a modern context, this hypothesis can be actualized where CSCs represent the seed and the tumor microenvironment the soil, and the interaction between them will promote cancer initiation and development (Korkaya et al., 2011). We previously described how inflammation and carcinogenesis are associated with, for example, oxidative stress generated by ROS that can in turn influence cellular transformation and promote tumorigenesis. In a

same manner ROS can influence the initiation of CSCs, or transform pre-existing CSCs and render them more aggressive (Bao et al., 2013; Pelicci et al., 2013). ROS accumulation can also exert a selective pressure on CSCs that often harbor increased detoxification capacity and thus contribute to maintaining a pool of resistant cells (Diehn et al., 2009). The tumor formation is usually accompanied by a tissue architectural deconstruction, with subsequent tissue anemia and hypoxia. In the bone marrow, hypoxic niches and HIF-1 $\alpha$  play critical roles in the regulation of normal hematopoietic stem cells (Nombela-Arrieta et al., 2013; Takubo et al., 2010), and in cancer activation of HIF-1 $\alpha$  in CSCs niche maintains an undifferentiated phenotype and self-renewal (Bar et al., 2010; Li and Rich, 2010; Wang et al., 2011a; Zhang et al., 2012a). The mechanisms of these processes involve ESC-like programming with the activation of genes such as *NANOG*, *OCT4*, *SOX2* and *KLF4* (Iida et al., 2012; Mathieu et al., 2011). On the other hand, in order to satisfy nutriment and oxygen needs, a second type of niche, perivascular, has been described. The CSCs can be localized in proximity to blood vessels and an angiogenesis process can support the formation and maintenance of CSC populations. For example, Brain CSC expressing nestin and CD133 are found closed to capillaries (Calabrese et al., 2007) and in glioblastoma the perivascular niche promotes glioma cell conversion to a more stem-like state through endothelia-derived nitric oxide-dependent induction of glioma cell Notch signaling. In summary (Charles et al., 2010), there is not one consensus for CSC supporting microenvironment and this is partly due to the fact that each CSCs differs for each tumor type. But it is manifest that tumor microenvironment have an effect (inductive or selective) on CSCs and this will have to be taken into consideration for future development of therapeutic strategies targeted against CSCs.

## 2. Tumor microenvironment soluble factors influencing CSCs.

CSC niches or microenvironments provide a physical anchor and can control stem cell fate through paracrin signals. Soluble factors can thus be secreted by tumor-associated fibroblasts, tumor-associated immune cells and (neo)capillaries (Castaño et al., 2012). I will describe hereafter some selected examples of molecules secreted in the tumor microenvironment and their effect on CSCs.

PDGF can be secreted by endothelial cells or tumor-associated fibroblasts and stimulate various cellular functions, including growth, proliferation, and differentiation (Gialeli et al., 2013). It was notably demonstrated that PDGF is involved in the expansion of breast CSCs (Devarajan et al., 2012). FGF secreted by activated stroma, can induce EMT and is implied in



maintenance of pluripotent cells (Billottet et al., 2008). Preliminary *in vitro* studies suggest that FGF could contribute to self-renewal of lung stem cells and homeostasis of breast CSCs (Fillmore et al., 2010; McQualter et al., 2010). As a last example, tumor-associated macrophages secrete high quantities of EGF which induces the EMT program in several epithelial cell lines *in vitro* and, like FGF, enriches for stem/progenitor cell self-renewal (Condeelis and Pollard, 2006; Ding et al., 2011; Vincent-Salomon and Thiery, 2003).

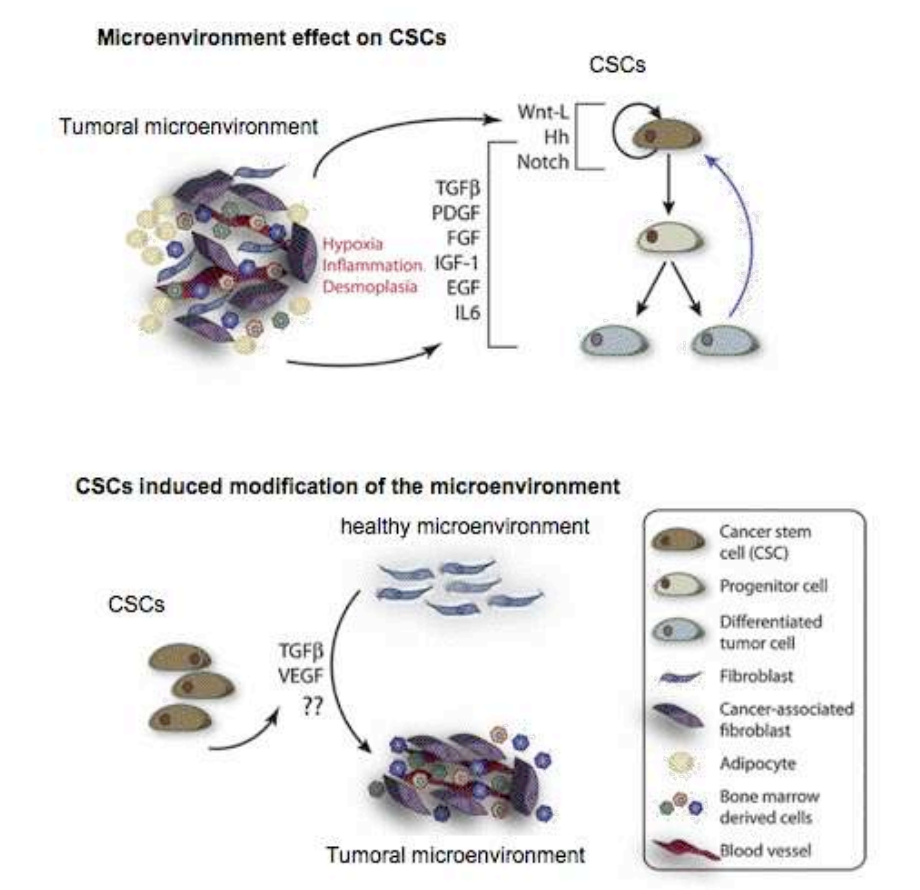
In addition tumor cell stemness is influenced by microenvironmental inflammation. Cytokines secreted in the environment in order to promote tissue repair and regeneration will activate pathways such as Wnt, Hedgehog and Notch, which are important pathways supporting CSCs (Tanno and Matsui, 2011). Thus continuous signaling may lead to aberrant stem cell activation and/or to dysregulation of self-renewal mechanisms and will promote the initiation and maintenance of CSCs. IL-6 in particular has been shown to trigger the conversion of non-CSCs into CSCs in breast cancer via a positive feedback loop involving NF- $\kappa$ B (Iliopoulos et al., 2009). IL-6 can activate the Akt, STAT3 and NF- $\kappa$ B pathways that can lead to transcriptional activation of pluripotency factors such as *OCT4* (Kim et al., 2013; Korkaya et al., 2011). In addition IL-6 can also promote self-renewal, hypoxia resistance and invasiveness, which are classical CSC properties (Dethlefsen et al., 2013; Qiu et al., 2013; Terui et al., 2004; Wang et al., 2012). Nevertheless, research on IL-6 contribution to CSCs have been mainly conducted in breast CSCs, where IL-6 is clearly determinant for the initiation and homeostasis of this population and its functions in other CSC populations remain to be elucidated.

The second important cytokine that has focused research interest for CSC promotion is TGF- $\beta$ . The first data provided by this research indicate that the influence of TGF- $\beta$  on tumorigenesis and CSCs is likely to be complex and to depend on the tissue and carcinogenesis stage. For example, TGF- $\beta$  may regulate chronic myelogenous leukemia (CML) stem cells by regulating the activity of Akt signaling (Miyazono, 2012). In addition to CML, TGF- $\beta$  has been implicated in CSC maintenance/induction for glioblastoma (Peñuelas et al., 2009), breast (Mani et al., 2008a), lung (Pirozzi et al., 2011) and liver cancers (You et al., 2010a). TGF- $\beta$ 's effects are mostly described in glioblastoma initiating cells (GIC) represented by CD133 expression. Although TGF- $\beta$  did not induce any change in the clonogenicity of GIC, inhibition of TGF- $\beta$  leads to the reduction of the number of spheres formed and a decrease in CD133+ population (Ikushima et al., 2009). Preliminary analyses of the mechanisms involved, indicate that TGF- $\beta$  signaling would lead to the transcription of stemness factors like *LIF*, *SOX2* and *SOX4* (Ikushima et al., 2009; Peñuelas et al., 2009). In liver cancer, TGF- $\beta$  treatment can induce the expression of CD133 (through epigenetic

regulation) (You et al., 2010a) and promote specific apoptosis resistance in CD133+ cells through activation of the p38/MAPK pathway (Ding et al., 2009). Moreover TGF- $\beta$  contribution to CSC phenotype is essentially admitted through its role in EMT initiation. Indeed breast cancer cells that underwent EMT acquired stem cell markers (Blick et al., 2010) and it is now recognized that cells undergoing EMT acquire stem cell phenotype (Katsuno et al., 2013; Mani et al., 2008a) and that activation of EMT factors can be associated to stemness factors (Eastham et al., 2007) TGF- $\beta$  could thus participate to the induction of new metastatic CSCs during tumor evolution via EMT initiation (Zhou et al., 2012b).

Finally like inflammation and cancer inter-connections, CSCs can in turn respond to microenvironment stimuli and secrete several factors that will influence its composition and functions (mainly to serve their own survival and support the tumor development) (Figure 23). Results from a recent study demonstrated that secretion of TGF- $\beta$ 2/TGF- $\beta$ 3 from breast cancer cells that disseminated to the lung served to induce stromal fibroblast expression of periostatin (POSTN), a component of the extracellular matrix. In turn, microenvironment-derived POSTN induced recruitment of Wnt ligands, thereby increasing Wnt signaling in CSCs (Malanchi et al., 2012). In another example, it was demonstrated that skin CSCs secreted VEGF, which operated in an autocrine fashion to expand the CSC pool, and in a paracrine manner to promote angiogenesis within the microenvironment (Beck et al., 2011). In addition, VEGF can be also secreted in glioblastoma by CSCs to support the development of local vascularization (Gilbertson and Rich, 2007).

The complexities of these interactions between CSCs and their microenvironment are far from being resolved but preliminary research clearly indicates that these two entities evolve together and influence each other in order to support tumor growth.



**Figure 23. Signaling between CSCs and tumoral microenvironment (adapted from Castaño et al., 2012)**

### 3. Influence of the microenvironment on liver progenitor cell transformation.

Comprehending CSCs involves understanding not only their endogenic properties and their interaction with the tumor microenvironment but also their cellular origin. The presence within a tumor of progenitor cells raises two hypotheses: either the cell of origin is a progenitor cell (maturation arrest theory) or, alternatively, tumor dedifferentiates and acquire progenitor cell features during carcinogenesis (dedifferentiation theory) (Sell, 2010). Animal models have shown that differentiated hepatocytes can be involved in HCC initiation (Roskams, 2006), and the observation of inter-conversion between non-CSCs and CSCs (Chaffer et al., 2013; Yang et al., 2012) suggests that an original transformed cell can further acquired stemness properties (through extracellular signals mentioned earlier, for example). On the other hand, the presence of stem cell markers, activation of notable pathways involved in homeostasis of embryonic and adult stem cells, and the correlation between liver progenitor cells and with liver injury severity and HCC risk, strongly support the “maturation arrest theory”. Stem/progenitors cells are believed to be more flexible to cell

fate decisions, and thus to be more susceptible to any extracellular signals that could interfere with their normal activation and differentiation (Hernandez-Vargas et al., 2009). When hepatic progenitor cells (HPCs) are requisitioned to compensate hepatocyte-driven regeneration failure, they are exposed to the inflammatory microenvironment. They can indeed, like hepatocytes be subject to ROS-induced DNA damage, genetic and epigenetic mutations promoting their transformations (Alison, 2005). They also express extracellular ligands for cytokines and growth factors. The continuous exposition of HPC to these stimuli could deregulate the control of pathways involved in proliferation, self-renewal and cell fate decision like Wnt- $\beta$  catenin/hedgehog and Notch and enhance their transformation in CSCs (Kitisin et al., 2007; Sun and Karin, 2013). This hypothesis is however still under discussion, especially with the description of contrasting observations concerning the effect of extracellular signalings on HPCs. In particular, interactions between TGF- $\beta$  and HPCs seem to be determinant for regulating the balance between their normal activation and their deregulation. TGF- $\beta$  loss of signal results in the expansion of HPCs in mice (Thenappan et al., 2010). Contrastingly, HPCs in regenerative liver harboured the stemness factor Oct4, Nanog, STAT3 together with the receptor TGFBR2, but further examination of stem cells in HCC revealed a loss of TGFBR2 expression together with the activation of the IL-6 pathway (Tang et al., 2008b). These data suggest that impaired TGF- $\beta$  signaling (with additional proliferative signal such as IL-6) can promote the activation and transformation of HPCs into liver CSCs. Joining the controversy for TGF- $\beta$  effects during hepatocarcinogenesis, it is likely that depending on the inflammatory context (viral, alcoholic, cirrhotic), TGF- $\beta$ 's effects on HPCs differ. Moreover the idea that TGF- $\beta$  slows down the activation and transformation of HPCs is not incompatible with the observation that later on, after evolution of the disease and its microenvironment, TGF- $\beta$  could support the growth of liver CSCs.

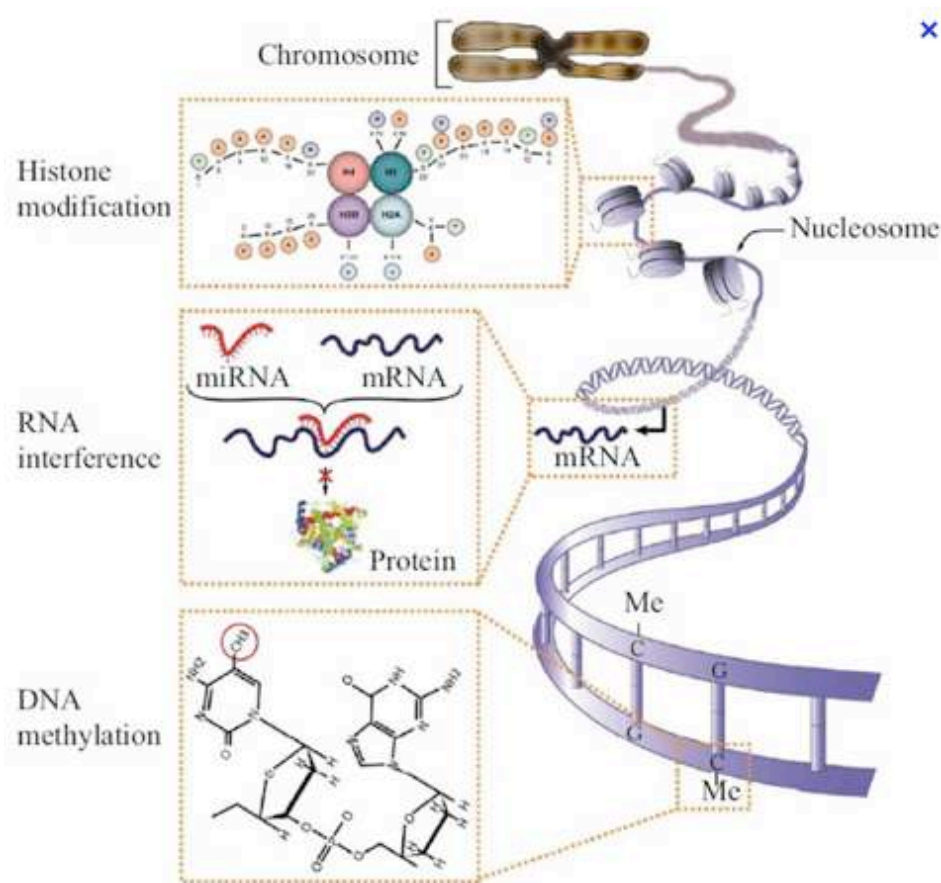
In the previous sections, we described how hepatocellular carcinoma and more precisely liver CSCs are sustained through the activation of specific pathways (like developmental pathways and signaling pathways sustaining secretion of cytokine) promoting tumor growth. Activation of these pathways relies not only on protein phosphorylation but also implies a reprogramming in gene expression. As suggested before, phenotypical changes observed during hepatocarcinogenesis depend on profound genomic and epigenomic modifications (Feo et al., 2009; Herath et al., 2006). Notably, in HCC, genetic alterations are not predominant and alone cannot explain all the alterations observed in cancer cell fate decisions. Epigenetic mechanisms, such as DNA methylation, are thus believed to assume an important role in HCC and cancer stem cell establishment (Sceusi et al., 2011).

## V. DNA methylation in Hepatocellular carcinoma

### A. Introduction to epigenetic mechanisms

The term “epigenetic” refers to all stable and heritable changes of phenotype that occur without generating alterations in the DNA nucleotide sequence (Baylin, 2005; Feinberg et al., 2006; Rountree et al., 2001). This term was first proposed in 1942 by Conrad Waddington to define the causal interactions between genes and their products that explain the phenotypic expression (Waddington 1942).

While every cell in the human body share the same DNA sequence, each acquires specific features allowing the formation of distinct organs and to accomplish the related metabolic functions. This indicates that additional mechanisms independent of the DNA sequence are required. Therefore, different epigenomes may explain differences in cell stages. Epigenetic information relies on three distinct mechanisms: DNA methylation, histone modifications, and non-coding RNA (Figure 24). Changes in these informations allow stable transmission of gene activity states through cell divisions. Alteration of epigenetic mechanisms may therefore contribute to tumor initiation by disrupting gene expression. Indeed, epigenetic mechanisms are now recognized to play a fundamental role in the regulation of important cellular processes and their deregulation contributes to human diseases, most notably cancer (Egger et al., 2004; Herceg and Vaissière, 2011; Sawan et al., 2008). While DNA sequences encode the primary information within the genome, epigenetic modifications offer robust and dynamic possibilities for regulation of the genetic information and for integration of external signals. Human cancer has usually been considered as a genetic disease, but recent evidences have illustrated the important role of epigenetic deregulations in most, if not all, human malignancies; making the concept of tumor development even more complex. The possible interaction between epigenetic mechanisms and environmental signals as part of the cellular adaptation response have raised high interest. (Herceg and Vaissière, 2011) and indeed epigenetic mechanisms appear to play a key role in the interaction between environmental factors and the genome (Herceg, 2007; Jaenisch and Bird, 2003; Shen et al., 2002). Finally, adverse and prolonged exposure to environmental, physical, chemical and infectious agents, as well as lifestyle factors, may induce aberrant epigenetic changes that lead to chronic diseases and neoplastic processes (Herceg et al., 2013).



**Figure 24. The three fundamental epigenetic mechanisms:** histone modifications, RNA interference and DNA methylation (Sawan et al., 2008)

Nucleosomes are the building blocks of chromatin and they represent two turns of genomic DNA (147 base pairs) wrapped around an octamer of two subunits of each of the core histones H2A, H2B, H3, and H4. The amino-terminal portion of the core histone proteins contains a flexible and highly basic tail region, which is conserved across various species and is subject to various post-translational modifications. Histone tails constitute one of the major site for epigenetic regulation of fundamental processes (Herceg and Hainaut, 2007). More than 60 different residues on histones have been described. There are, to date, at least eight different types of histone modification: acetylation, methylation, phosphorylation, ubiquitination, sumoylation, ADP ribosylation, deimination, and proline isomerization (Kouzarides, 2007). Traditionally, two mechanisms are thought to control the function of these modifications. First, these different marks could affect the nucleosome-nucleosome or DNA-nucleosome physical interactions. Second, different marks could represent a binding site for the recruitment of specific proteins involved in gene transcription regulation or in genome spatial organization. Additionally, several reports raise the possibility that all of these modifications are combinatorial and interdependent and therefore may form the

“histone code”, meaning that combination of different modifications may result in distinct and consistent cellular outcomes (Lee et al., 2010b; Rando, 2012).

MicroRNAs (miRNAs) are a class of approximately 22-nt-long non-coding RNAs found in eukaryotes. miRNA processing is mediated by the nuclear Drosha/Pasha complex with RNase III activity and further mediated by the RNase III enzyme Dicer to generate a 22-bp miRNA duplex. miRNA can inhibit gene expression by mRNA degradation or by translational inhibition of target genes. MiRNA genes constitute approximately 1–5% of the predicted genes, with up to 24521 miRNA genes in the human genome (miRBase release 20, June 2013). miRNAs are able to regulate expression of hundreds of target mRNAs simultaneously, thus controlling a variety of cell functions including cell proliferation, stem cell maintenance, and differentiation.

The last important epigenetic mechanisms takes place on the DNA template itself: DNA methylation consists of a chemical modification of the cytosine base. Many fundamental cellular events are the result of epigenetic signals modifying DNA methylation in the genome (Bird, 2002). Changes in DNA methylation have been extensively studied because of their role in major cellular processes, including embryonic development, transcription, chromatin structure, X chromosome inactivation, genomic imprinting and chromosome stability (Baylin et al., 2001; Grønbaek et al., 2007; Jin and Robertson, 2013; Seisenberger et al., 2013) and their frequent association with human diseases (Zardo et al., 2005) As my work focused on this precise epigenetic mechanism, separate sections will be dedicated to it.

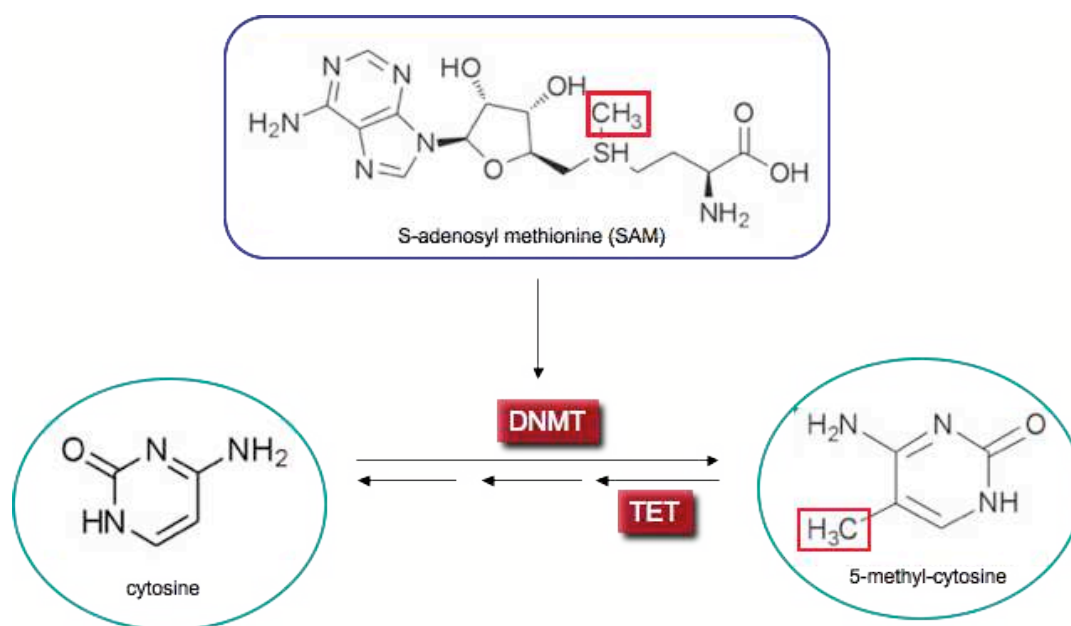
## B. DNA methylation

### 1. CpG sites are methylated by DNMTs

DNA methylation is a chemical modification that results from the transfer of a methyl group from a methyl donor substrate (S-adenosyl-L-methionine, SAM) that affects mainly the 5' position of cytosine bases in CpG conformations (“p” indicates that the cytosine and the guanine are linked by a phosphodiester bond (Doerfler, 1983) (Figure 25).

DNA methylation occurring on non-CpG configuration, such as CpNpG or CpA and CpT sequences, has also been described in the eukaryotic genome (Clark et al., 1995), especially in

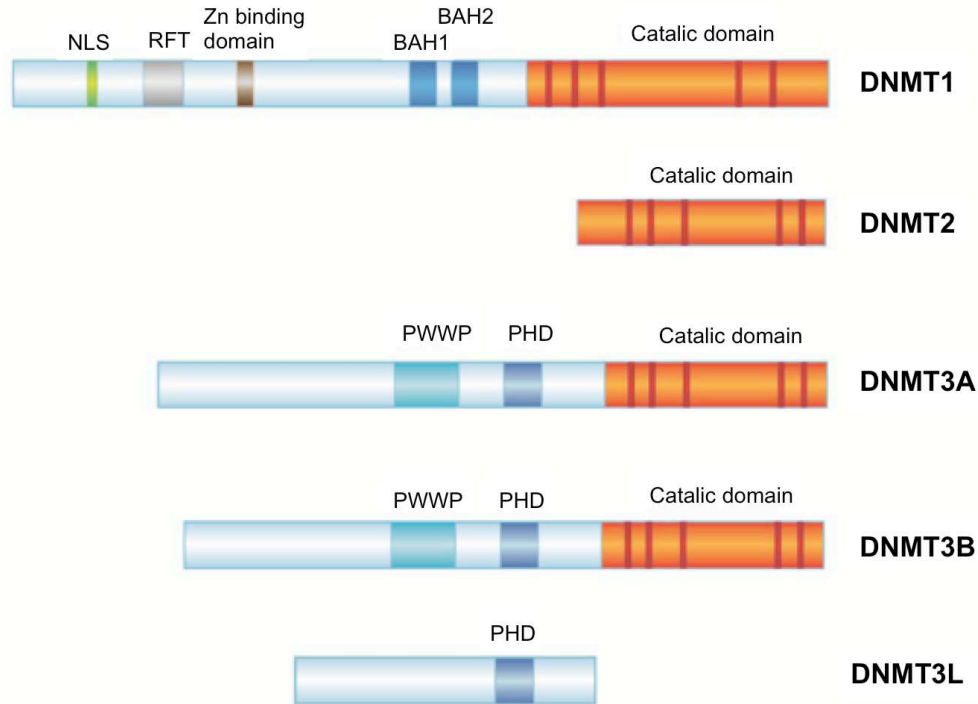
mouse embryonic stem cells (Lister et al., 2009; Ramsahoye et al., 2000), although the role of non CpG methylation is still not clear.



**Figure 25. Chemical reaction of cytosine methylation on the 5' carbone of the base**  
S-adenosyl methionine serves as a methyl group donor. The direct reaction (methylation) is catalyzed by DNMT enzyme while the indirect reaction (demethylation) comprises different intermediaries states and involves TET proteins (adapted from Dricu et al., 2012).

The addition of a methyl group on a cytokine is catalysed by the enzymes belonging to the DNA methyltransferases (DNMTs) family. Five members of the DNMT family have been identified in mammals: DNMT1, DNMT2, DNMT3A, DNMT3B and DNMT3L (Figure 26). However, as far as we know, only DNMT1, DNMT3A and DNMT3B have been implied in the establishment of the global cytosine methylation pattern (Cheng and Blumenthal, 2008). These independently encoded proteins are classified as *de novo* enzymes (DNMT3A and DNMT3B) or as maintenance enzymes (DNMT1), as detailed below. DNMT2 and DNMT3L were not thought to function as cytosine methyltransferases. However, DNMT2 proteins were recently shown by Goll and colleagues to function as RNA methyltransferases (Goll et al., 2006). DNMT3L was shown to stimulate *de novo* DNA methylation by DNMT3A and to mediate transcriptional repression through interaction with histone deacetylase 1 (Chedin et al., 2002; Deplus et al., 2002).





**Figure 26. Schematic structure of human DNMTs and DNMT3-like proteins**

Conserved methyltransferase motifs in the catalytic domain are indicated in red. *NLS*, nuclear localization signal; *RFT*, replication foci-targeting domain; *BAH*, bromo-adjacent homology domain; *PWWP*, a domain containing a conserved proline-tryptophan- tryptophan-proline motif; *PHD*, a cysteine-rich region containing an atypical plant homeodomain; *aa*, amino acids. DNMT3L lacks the critical methyltransferase motifs and is catalytically inactive (adapted from Chen and Riggs, 2011).

DNMT1 appears to be involved in restoring the parental DNA methylation pattern in the newly synthesized DNA daughter strand, thereby ensuring the methylation status of CpG islands through multiple cell generations. DNMT1 exhibits a preference for hemimethylated substrates and it possesses a domain targeting replication foci. It was recently discovered that DNMT1 was guided to replication forks through the protein UHFR1 that would initially recognize the hemimethylated site and further recruit the enzyme (Bostick et al., 2007). Confirming the important role of DNMT1 in proper cell functioning and development, it should be mentioned that the loss of *Dnmt1* function results in embryonic lethality in mice (Li et al., 1992).

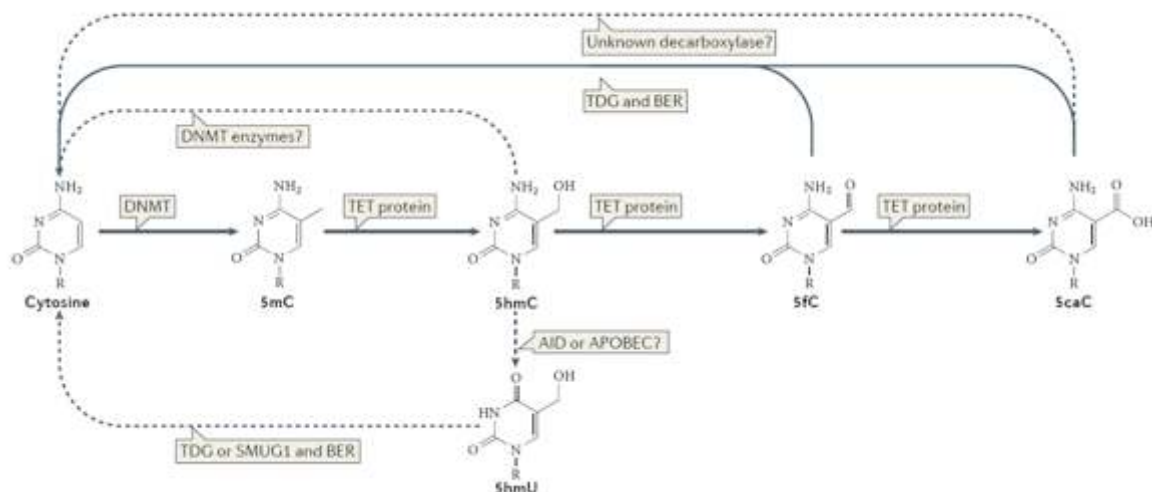
*De novo* DNA methylation during embryogenesis and germ cell development are carried out by the DNMT3 family (DNMT3A and DNMT3B). Inactivation of each of these genes leads to severe phenotypes (Okano et al., 1999). *Dnmt3a* knock-out mice die shortly after birth and embryonic lethality is observed in case of the absence of *Dnmt3b*. Thus, DNMT3A seems to be responsible for the methylation of sequences critical for late developmental stage or those just after birth, whereas DNMT3B may be more important for early developmental stages

(Okano et al., 1999). Besides, DNMT3B appears to be involved in DNA methylation of particular regions of the genome, as it has been shown by the studies of the Immunodeficiency, Centromere instability and Facial abnormalities (ICF) syndrome, a disease caused by genetic mutation in *DNMT3B* (Jin et al., 2008). Finally it should be mentioned that the barrier between *de novo* and maintenance methylation is not impassable and that inter-conversion of activities between DNMT1 and the DNMT3 families has already been described (Egger et al., 2006; Riggs and Xiong, 2004).

## 2. Demethylation processes

Understanding how these patterns of 5-methylcytosine are established and maintained requires the elucidating of mechanisms for both DNA methylation and demethylation. DNA demethylation can be achieved passively, through 3 mechanisms: the limited availability of the donor SAM, the compromised integrity of DNA and the altered expression and/or activity of DNMT1 (Pogribny and Rusyn, 2012). All these mechanisms have for consequence, the non-maintenance of methylation profile through cell divisions and the progressive loss of DNA methylation marks. However, considerable evidences support the existence of genome-wide active demethylation in zygotes (Hajkova et al., 2002; Mayer et al., 2000; Morgan et al., 2005; Oswald et al., 2000) and primary germ cells (Pugs) (Hajkova et al., 2002; Morgan et al., 2005) and locus specific active demethylation in somatic cells, such as neurons (Ma et al., 2009) and T lymphocytes (Bruniquel and Schwartz, 2003). Yet, the mechanism(s) of active demethylation are still currently elucidated. A number of mechanisms for the enzymatic removal of the 5-methyl group of 5mC, the 5mC base, or the 5mC nucleotide have been proposed (shown in Figure 27), The recent discovery of a new modified base, 5-hydroxymethylcytosine (5hmC), now considered as the 6th base of the mammalian DNA (Münzel et al., 2011), is likely to play an important role in active demethylation process and open new area of research.

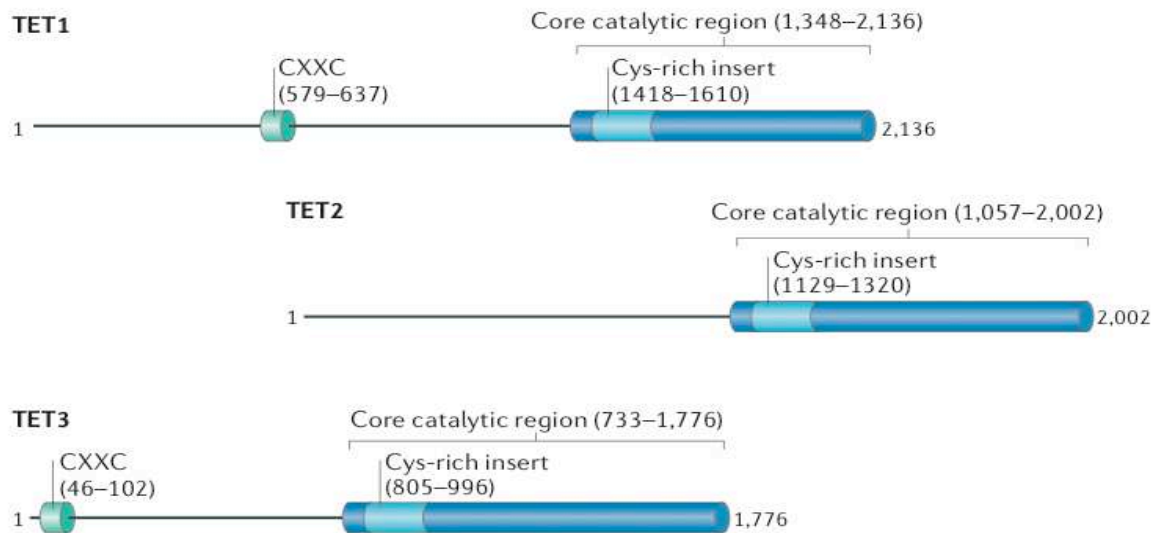
Recently, it has been shown that mouse and human Tet family (Figure 28) proteins can catalyze conversion of 5mC to 5hmC (Ito et al., 2010).



**Figure 27. Known and putative pathways of DNA demethylation that involve oxidized methylcytosine intermediates**

Ten-eleven translocation (TET) proteins sequentially oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC can be removed by thymine DNA glycosylase (TDG) and replaced by cytosine via base excision repair (BER), although the extent to which this mechanism operates in specific cell types during development is unknown. Other proposed mechanisms of demethylation are less well established, including decarboxylation of 5caC, DNA methyltransferase (DNMT)-mediated removal of the hydroxymethyl group of 5hmC and deamination of 5hmC (and 5mC) (see main text) by the cytidine deaminases AID (activation-induced cytidine deaminase) and APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide). AID enzymes deaminate cytosine bases in DNA to yield uracil. AID and the larger family of APOBEC enzymes have been proposed to effect DNA demethylation by deaminating 5mC and 5hmC in DNA to yield thymine and 5hmU, respectively. As these are present in mismatched T:G and 5hmU:G basepairs, they have been proposed to be excised by SMUG1 (single-strand-selective monofunctional uracil DNA glycosylase) or TDG (Pastor et al., 2013).

5hmC might be repaired by a BER process, although, so far, no 5hmC DNA glycosylases have been identified. Interestingly, two new studies identified new intermediates that can be used as substrate for the demethylation process. Indeed, it has been demonstrated that the Tet family of proteins have the capacity to convert 5mC not only to 5hmC, but also to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) *in vitro* and in cultured cells in an enzymatic activity-dependent manner (He et al., 2011). Furthermore, 5hmC can also be oxidized into 5caC, 5fC and 5caC are specifically recognized and excised by TDG, followed by BER (He et al., 2011; Zhang et al., 2012b). Additional processes could include enzymatic activity of DNMT3A/B themselves as an *in vitro* study described that they can present a dehydroxymethylation activity (Chen et al., 2012a). DNMT3B activity in particular would be regulated through the redox balance, with reducing conditions favouring methylation activity and oxidizing conditions favouring dehydroxymethylation activity. Figure 27 summarizes all the possible mechanisms of active demethylation.



**Figure 28. Schematic structures of TET family members.**

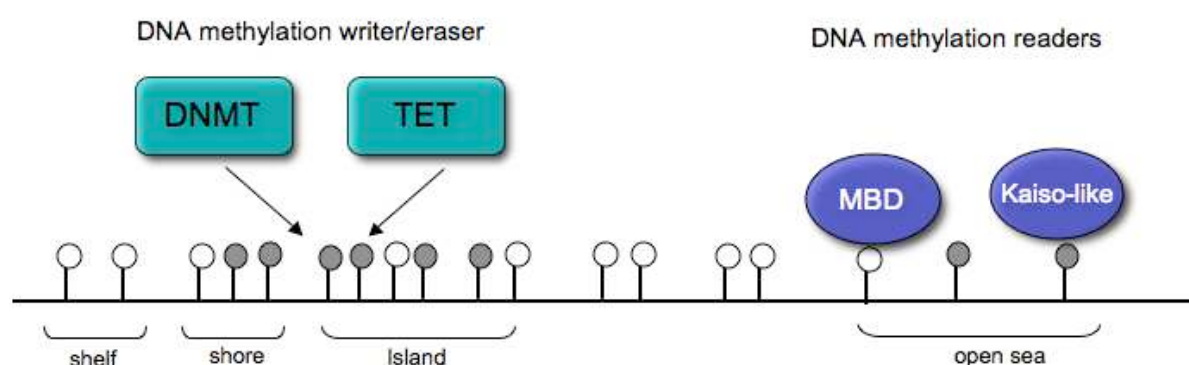
Ten-eleven translocation (TET) proteins contain a DNA-binding CXXC domain towards the amino terminus and a carboxy-terminal catalytic core region that includes a Cys-rich insert and a larger double-stranded  $\beta$ -helix (DSBH) domain. The number of amino acids is indicated, and the numbering corresponds to the human proteins (Pastor et al., 2013).

### 3. Methylation regulates transcription and genome organisation.

A prerequisite for understanding the function of DNA methylation is knowledge of its distribution in the genome. CpG sites are not distributed equally throughout the human genome but are found more frequently within small regions of DNA called CpG islands (Bird, 1986). Regions comprised between CpG islands and CpG “open seas” present a progressive decrease of CpG numbers and are called “shelf” and “shore” regions (Figure 29) (Shen and Laird, 2013). According to calculations of CpG prevalence, nearly 60% of human promoters are characterized by high CpG content (Saxonov et al., 2006). Nevertheless, CpG density itself does not influence gene expression. Almost 28,000 CpG islands are spread within the human genome and among them 20,000 are associated with a gene (Huang and Esteller, 2010), indicating that methylation of those specific regions constitutes a powerful mechanism of gene regulation. Usually, CpG islands are unmethylated in transcriptionally active genes whereas silenced genes are characterized by methylation within promoter region (e.g., tissue-specific or developmental genes). Therefore, the presence of DNA methylation should be tightly controlled in the cell in order to maintain the balance between silencing of repetitive elements and expression of fundamental cellular genes (Lange et al., 2011). It should be specified that DNA methylation works in parallel with other regulatory

mechanism. In consequence an unmethylated sequence within a gene promoter can constitute a permissive state for transcription but this transcription can be blocked through other regulatory mechanisms (including histone modifications and transcription factor availability). The correlation between DNA methylation and gene expression is thus not completely linear (Cooper, 2000).

The genome of higher eukaryotes contains a different types of repetitive sequences (such as Alu, LINEs, and SINEs). A stable inhibition of retrotransposons is necessary to insure the genome stability and integrity (Elgin and Grewal, 2003). Permanent silencing of these DNA sequences is mainly due to DNA methylation, which tightly regulates chromatin. Whereas transposons must be stable and totally silenced to prevent genomic instability, expression of genes involved in development is subject to permissive epigenetic control (Reik, 2007). How DNA methylation contributes to the inhibition of expression still remains unclear and various hypotheses have been proposed. Firstly, for some transcription factors, e.g. AP-2, C-MYC, CREB/ATF, E2F and NF- $\kappa$ B, DNA methylation could create a physical barrier, preventing access to promoter binding sites (Zingg and Jones, 1997). This might be true, but only for a subset of transcription factors. Another model of gene inactivation mediated by DNA methylation is related to DNA methylation “readers” such as methyl-CpG binding domain proteins (MBDs) (Figure 29). In general, DNA methylation is not considered to be sufficient to completely establish the inactive chromatin state. It is more thought to be an initial step, that will be followed by MBD recruitment that, in turn, will interact with histone deacetylases known as epigenetic enzymes linked to repression. The chromatin can thus be compacted and gene silencing is achieved.



**Figure 29. Distribution of CpG sites across the genome.**

CpG site regions have been named according to their density in CpG sites: islands possess a high density of CpG sites, they are surrounded by shelf and shores regions. CpG oceans correspond to regions where CpG sites are spread. Every CpG site can be transformed by DNA methylation writer (DNMT enzymes) or eraser (TET proteins) and can be further bound by DNA methylation readers (MBD or Kaiso like proteins) that will further recruit other chromatin remodelling factors.

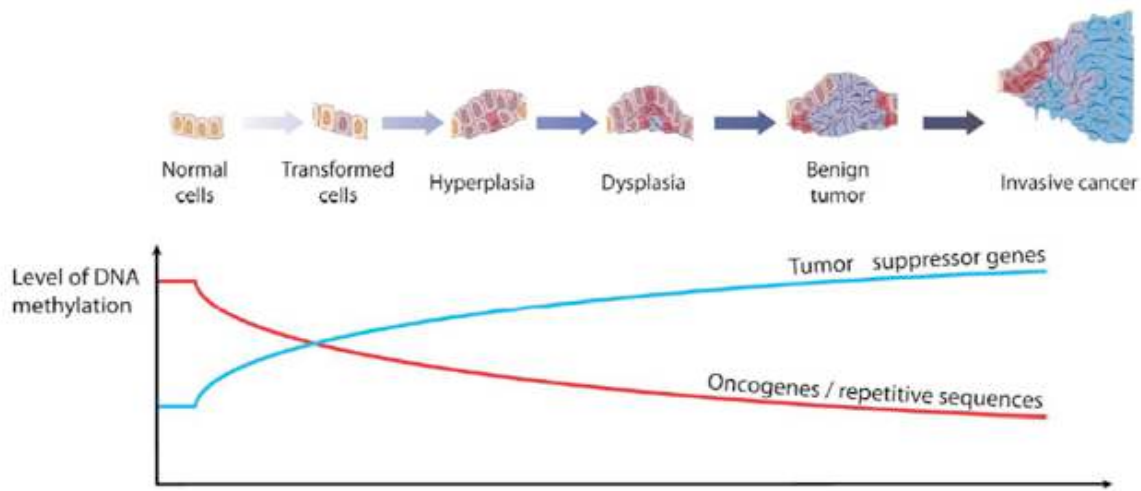
Nevertheless, MBDs are not the only class of proteins capable of acting as HDAC-dependent transcriptional repressors by association with methylated DNA sequences. The Kaiso-like family of proteins also “reads” methylated DNA by zinc finger motifs and has been reported to be involved in gene silencing (Filion et al., 2006; Prokhortchouk et al., 2001). Unlike MBDs, Kaisos also recognize unmethylated sequences. Recently, different studies identified a key role of polycomb group proteins (PcG) in establishing the DNA methylation pattern. It has been suggested that DNMT1 and DNMT3B interact in a specific manner with PcG complexes to establish DNA methylation in combination with histone marks (Hernández-Muñoz et al., 2005; Jin et al., 2009; Viré et al., 2006). For example, it was supposed that target genes are first subjected to H3K27 methylation and then are marked with *de novo* DNA methylation (Ohm et al., 2007; Widschwendter et al., 2007). Moreover, it was also reported that in cancer cells up to 5% of promoters containing CpGs were silenced by H3K27 trimethylation which was independent of DNA methylation (Kondo et al., 2008). As the exact links between PcG regulation and DNA methylation are still unclear, these findings add a novel layer of complexity to epigenetic gene silencing. In summary, the above explanation of DNA methylation-mediated gene silencing clearly illustrates how all epigenetic components interact in a complex manner to regulate gene expression.

### C. Deregulation of DNA methylation and DNMT expression in HCC

#### 1. Aberrant DNA methylation profiles in HCC

As described above, appropriate DNA methylation is essential for development and proper cell functioning, thus any abnormalities in this process may lead to various diseases, including cancer (Jin and Robertson, 2013). The role of DNA methylation in normal cellular processes and the contribution of DNA methylation defects to cancer appearance and progression are now well established. Indeed, tumor cells are characterized by a different methylome from normal cells (Shen and Laird, 2013). Interestingly, both hypo- and hypermethylation events can be observed in cancer (Figure 30). Generally, a global decrease in methylated CpG content is observed. This phenomenon contributes to genomic instability and, less frequently, to activation of silenced oncogenes. On the other hand, CpG island hypermethylation in promoters of specific genes has been shown as a critical hallmark in many cancer cells (Paz et al., 2003). An increasing number of genes has been reported to be inactivated by a DNA methylation mechanism during tumorigenesis that mainly acts as

tumor suppressors in normal tissues. Aberrant DNA hypermethylation of CpG islands is typically associated with inhibition of gene transcription and unscheduled silencing of genes (Baylin, 2005).



**Figure 30. Aberrant DNA methylation changes during carcinogenesis**

Cancer development is mainly characterized by hypermethylation of tumor suppressor genes and hypomethylation of oncogenes. (Herceg, unpublished)

In this manner, several studies have shown that aberrant DNA methylation can promote carcinogenesis, including HCC (Pogribny and Rusyn, 2012; De Zhu, 2005). Comparing tissues from patients with paired- non-cancer liver tissues, the level of genome-wide-5-methylcytosine was significantly reduced in tumorigenic tissues. One of the first epigenetic changes detected in HCC was aberrant genome-wide hypomethylation (Lin et al., 2001). Indeed, *LINE-1* (Long interspersed nuclear element 1) methylation has been shown to be reduced in HCC tumors compared with non cancerous tissues (Lee et al., 2009; Lin et al., 2001). Later, the levels of serum *LINE-1* hypomethylation at initial presentation have been shown to correlate significantly with large tumor sizes, advanced tumor stages as well as HBsAg expression (Tangkijvanich et al., 2007), suggesting that *LINE-1* methylation may be a good prognostic marker. This observation has been confirmed by several other studies (Gao et al., 2013a; Shitani et al., 2012). The development of microarray platforms allowing genome wide analyses for DNA methylation permitted the description of global DNA methylation pattern in HCC. Both hyper- and hypomethylation marks are found (compared to healthy tissue), but hypomethylation marks are always predominant (representing at least 60% of the differentially methylated sites) (Shen et al., 2012; Song et al., 2013; Stefanska et al., 2013a). A recent study performed by Sheng et al. (2013) interrogated more than 450,000 CpG sites within the human genome between HCC and non tumors samples. They found that

10,000 sites presented a difference higher than 30% for DNA methylation. Hypomethylated sites represented 78% of the differentially methylated sites and were comprised mostly in “open sea” regions (60%), whereas hypermethylated sites were mostly comprised within CpG islands (60%). This regional distinction between hypo- and hypermethylation is likely to reflect a difference in DNA methylation function.

In parallel to this genome wide alteration, regional DNA methylation alterations have been reported. Hypermethylation has been detected in particular in CpG islands of tumor suppressor genes (TSGs) (Hamilton, 2010; Huang, 2009; Mao et al., 2012; Nishida et al., 2012a; Wu et al., 2012). These hypermethylated CpG islands result most of the time in gene inactivation. Genes affected are involved in cell proliferation inhibition (*p16INK4A*, *p21*, *p27*, *RASSF1A*, *SOCS1-3*, *RIZ1*, *sFRP1*), cell cycle (*CDKN2A*, *APC*) in apoptosis (*CASP8*, *XAF-1*, *ASPP1*, *ASPP2*), in cell adhesion and cell migration (*CDH1*, *TFPI-2*), gene transcription regulation (*PRDM2*, *RUNX3*) DNA repair (*GSTP1*). All these genes have been found hypermethylated on their promoter in at least 50% of HCC human samples. The status of methylation is often inversely correlated with the gene expression. For example an immunoprecipitation performed on MBD2 on the HepG2 cell line, demonstrate that MBD2 binds to several genes found hypomethylated in HCC and that it colocalizes with the transcription factor CEBPA (Stefanska et al., 2013b). These genes are all related to tumor promoting pathways including inflammation, cell growth, invasion, drug resistance, cell communication etc. In addition the combination of both hypo- and hypermethylated genes can both contribute to the same biological function dysregulation and thus assure the misappropriation of the pathway to the tumor growth. For example, the hypomethylation of *vimentin* and the hypermethylation of *E-Cadherin* involved in EMT transition will both serve the metastatic evolution of the tumor (Kitamura et al., 2011; Zhai et al., 2008).

## 2. Alteration in DNMT1 DNMT3A, DNMT3B expression

In parallel to changes in methylation, alterations in DNMT expression in HCC were investigated in several studies.

*DNMT1*, *DNMT3A* and *DNMT3B* mRNA levels were all higher in HCC samples compared with paired non-HCC samples (Lin et al., 2001). This result was confirmed by further studies that analysed tumors samples, their corresponding non-cancerous tissue, high- and low grade nodule dysplasia (ND) and normal tissue samples (Choi et al., 2003; Oh et al., 2007; Park et al., 2006). They found that DNMT1, DNMT3A and DNMT3B expression was significantly increased in high grade ND, cirrhotic tissues and HCC samples compared to



low grade ND and healthy tissue. Moreover the higher expression of DNMT1 in HCC was correlated to low recurrence-free and overall survival (Saito et al., 2003). More precise mechanisms have been examined for specific genes like for *ASPP2* silencing for which the HBX protein was found to recruit DNMT3A and DNMT3B on its promoter. Further *in vitro* studies with HBX transfected cell lines demonstrates an up-regulation in DNMT3A and DNMT1 expression and a down regulation in DNMT3B expression (Park et al., 2007). Interestingly, here DNMTs specific dysregulations could be related to both DNA aberrant hypo and hypermethylation in liver cancer: indeed DNMT3A has been reported to have more affinity for gene promoters (compared to DNMT3B that would bind preferentially with centromeric regions). Thus in this study, DNMT3A and DNMT1 up-regulation could be responsible for local CpG island hypermethylation, as DNMT3B down-regulation would explain the global hypomethylation observed in non-coding regions. The role of the different splice variants for each DNMT's family member has also been investigated, in particular for the DNMT3b4 isoform. Saito *et al.*, observed that when a global hypomethylation of pericentromeric satellite regions was observed in HCC, no mutation was detectable in DNMT3b whereas the inactive splice variant DNMT3b4 was over-expressed (Saito et al., 2002). DNMT3b4 does not show any catalytic activities but could compete with DNMT3b activity and actually a correlation was found between DNMT3b4 expression and DNA hypomethylation in the pericentromeric satellite regions in HCC patients. DNMT3b splice variants could thus be also implied in the mechanisms of hepatocarcinogenesis.

However, even if DNMTs up-regulation may be observed concomitantly with CpG island hypermethylation in TSG promoters, the role of DNMTs expression in TSG silencing remains uncertain as other analyses in HCC samples concluded that there was no significant correlation between DNMTs expression and DNA methylation (Park et al., 2006). Other studies confirmed this lack of association between DNMTs mRNA's level in tumor samples and DNA hypo- or hypermethylation (Eads et al., 1999; Ehrlich et al., 2006; Oh et al., 2007) . One of these studies actually observed that the detected up-regulation of DNMTs family members was strongly dependent on the housekeeping genes used for the qPCR assay: indeed no upregulation was observed when the normalization of expression was done with proliferation-associated genes (Eads et al., 1999). This could indicate that DNMTs expression is proliferation dependent (which is attested for *DNMT1*) which accounts for all their apparent upregulation in tumors. Therefore, one should be precautious concerning the techniques used to study DNMTs and the ensuing conclusions and hypotheses that can be raised.

In conclusion, even if DNMTs abnormal expression and/or activity have been reported, the exact function of this dysregulation and its link to the aberrant DNA methylation profile observed are still poorly understood.

It should be mentioned that some current hypotheses highlight the role of epigenetic modification in early stages of tumor development and even in cancer predisposition. It has been proposed that epigenetic disruptions are the initiating events leading to the occurrence of “cancer progenitor cells” (Saito et al., 2002). Furthermore, both genetic and epigenetic alterations are known to lead tumor progression. In this context, the existence of DNA methylation abnormalities that appear before mutations and that are involved in tumorigenesis is strong evidence in support of this theory. The next section will describe the evidence and the hint indicating that DNA methylation has a preponderant role in HCC development.

#### D. DNA methylation contribution to hepatocarcinogenesis

##### 1. DNA methylation alterations in precancerous stages

As described above, DNA methylation alterations in HCC affect chromosomal stability, genome integrity, oncogene silencing and TSG expression. Interestingly these events are observed at early stages during liver oncogenesis. Concerning TSG promoter hypermethylation, RASSF1A (link to cell cycle arrest) appears hypermethylated in 50% of fibrosis cases and 75% of cirrhotic tissues (Schagdarsurengin et al., 2003), E-Cadherin (involved in EMT inhibition) methylation is increasing in dysplasia stage 1 and 2 (Kwon et al., 2005), and *in vitro*, and HBV-transfection in cell lines induced hypermethylation of RASSF1A, GSTP1 (involved in DNA repair mechanisms) and CDKN2B (cell cycle effector) (Park et al., 2007). More recently, a subset of 8 TSG (HIC1, GSTP1, SOCS1, RASSF1, CDKN2A, APC, RUNX3 and PRDM2) were analysed for their methylation status on their promoter between tumor, non-tumor matched samples and chronic hepatitis C samples (Nishida et al., 2012b). The promoters of these TSG were hypermethylated in tumors, but interestingly their methylation profile in chronic hepatitis C samples was significantly correlated with shortened time to HCC occurrence. This result insinuates that TSG hypermethylation and silencing are not a consequence of cell transformation in hepatocellular carcinoma, but probably act as tumor initiating events from the early stages of hepatocellular progression. DNMTs have

also been identified to be more expressed since precancerous stages as cirrhosis and dysplasia nodules or even during chronic viral infection (DNMT1 and DNMT3B have in this way been identified as host factors involved in HCV propagation) (Chen et al., 2013a; Choi et al., 2003).

However, these descriptions of early alterations in HCC do not permit to clarify if DNA methylation deregulations are a cause or a (early) consequence of HCC development. The Knudson's hypothesis suggests that cancers arise from a successive accumulation of genetic alterations and further leads to the identification of cancer-related genes. Aberrant promoter hyper- and hypomethylation in cancer (including HCC) are known to occur in well established oncogenes and TSG. These methylation marks should thus be included in the hallmarks characterizing cancer cells. Furthermore epigenetic mechanisms are intimately linked with genetic disorders (Shen and Laird, 2013). Indeed epigenetic marks can directly cause genetic mutations by alteration of the expression of proteins involved in DNA damage repair. The CG base pair is also highly subject to conversion into TA base pair (this mutation link to the methylation status of CG sites has been described in almost 25% of the *TP53* mutations reported in human cancer) (Olivier et al., 2010). In turn genetic defects on epigenetic factors (such as DNMT or TET proteins), will lead to epigenetic alterations (Couronné et al., 2012; Ko et al., 2010; Shen and Laird, 2013). Epigenetic disorders and genetic mutations should thus be included together as genome alterations that can progressively lead to cancer development. Whether epigenetic disorders appear before genetic mutations is still under debate and probably depends on the original tissue and the nature of the environmental risk factors associated. Causative evidences for the implication of DNA methylation processes in HCC initiation include rodent models with nutritional (lipogenic methyl deficient diet) (Christman, 1995; Pogribny et al., 2004) or genetic (*Apc<sup>min/+</sup>;DNMT1<sup>chip/c</sup>*) alterations that result in liver cancer apparition (Yamada et al., 2005). In addition a mouse model of early stage liver fibrosis demonstrated that the hypomethylation of the gene *SPPP1* (involved in inflammation) was correlated to its higher expression and this regulation occurs even before the actual detection of fibrosis (Komatsu et al., 2012). This gene regulation through DNA hypomethylation would be a leading event for liver fibrosis. In order to improve the comprehension of DNA methylation alterations with HCC initiation and development several large scale studies establishing methylation signatures have been conducted. In this manner DNA methylation profiling has been shown to be able to differentiate HCC from preneoplastic lesions (low grade – high grade nodule dysplasia and cirrhosis) (Ammerpohl et al., 2012; Nishida et al., 2008), supporting the idea that DNA methylation profile can serve and thus reflect a particular cellular phenotype and/or

histological context. DNA methylation pattern can fully distinguish HCC samples from adjacent non-tumorigenic tissues and more precisely DNA methylation pattern can discriminate HCC with an etiology associated with HBV, HCV or alcohol intake (Hernandez-Vargas et al., 2010; Lambert et al., 2011). Finally, a successful prediction for HCC (with 95% sensitivity and 100% specificity) was established using quantification of DNA methylation on bacterial artificial chromosomes (Nagashio et al., 2011). All these data suggest that DNA methylation intervenes from precancerous stages to initiate HCC and that the pattern of DNA methylation is specific to each carcinogenic context.

## 2. DNA methylation interaction with inflammation

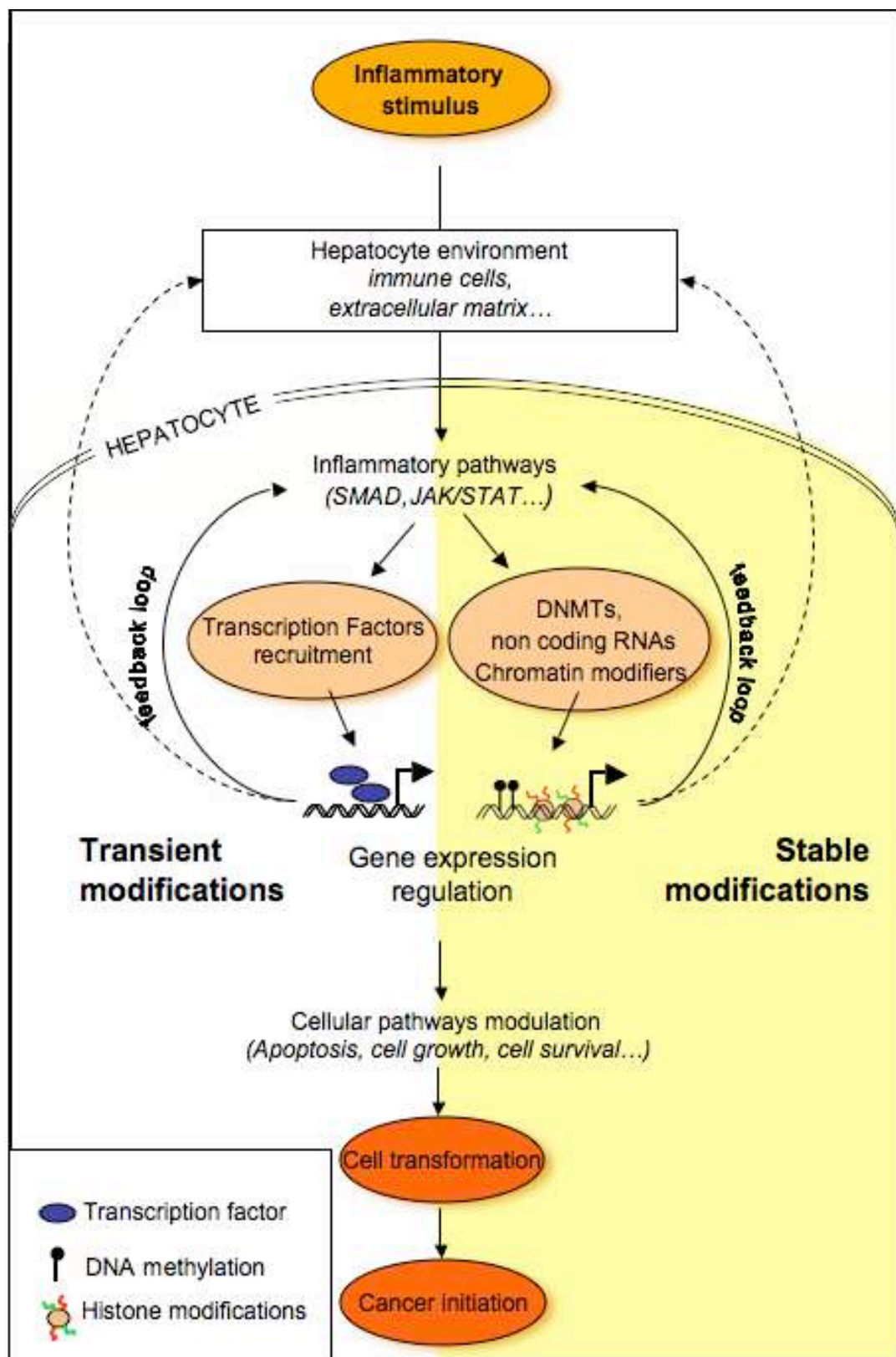
As we have seen before, inflammatory and DNA methylation deregulations are both early events in hepatocarcinogenesis and several observations suggest that they might have a leading role in cancer initiation. Despite this, the question of whether inflammation and DNA methylation act concomitantly to initiate HCC or if one is triggered by the other remains.

As I showed in the previous chapters, cancer may be considered as both a genetic and epigenetic disease. With recent advances exploring epigenetic signatures in tumors, and precancerous and healthy tissues, this definition has been refined and it is believed that epigenetic perturbations act as precursors, before or concomitantly to genetic alterations, to initiate cancer (Shen and Laird, 2013). However, we still don't know what events could be in turn be precursors of epigenetic deregulations. Epigenetic marks observed in tumor samples are not always associated with the etiology of liver cancer, and in the rare cases where a correlation is found, the mechanisms by which an etiological agent can alter the epigenome of hepatocytes remains vague. As a result, two models are drawn for liver cancer initiation (based on epigenetic or inflammatory processes) but these two could be joined into a unique model where liver inflammation could be the precursor event leading to epigenetic alterations and then HCC initiation (Figure 31) (Martin and Herceg, 2012). In this model, inflammation could modify cell activity (leading subsequently to hepatocarcinogenesis) either directly or indirectly through epigenetics. In the direct way, cytokines are able to modulate themselves cellular pathways such as apoptosis, cellular proliferation and cellular survival. In the indirect way, cytokines interfere with cellular pathways through modulation of the gene expressions involved in those pathways by recruiting chromatin modifiers on their promoters and thus activating or silencing their expression. Moreover, these epigenetic modifications can themselves promote the over-expression of inflammatory genes thus

creating a vicious circle. Recent mechanistic and functional studies support this model by demonstrating interconnections between inflammatory pathways and epigenetic modifications. For example, chronic inflammation increases the level of ROS in the cytoplasm and high levels of ROS have been reported to induce *SNAIL* expression that can in turn recruit DNMTs and HDAC to silence several specific genes (Hamilton, 2010; Lim et al., 2008). *In vivo* alcohol intake or *in vitro* HPS treatment (an inflammatory stimulus) can induce H3K9/S10 phosphorylation at cytokine gene promoters (Saccani et al., 2002; Yamamoto et al., 2003) and this specific histone mark happens to be required for NF- $\kappa$ B recruitment to promoters (Anest et al., 2003). IL-6 and TGF- $\beta$  can induce EZH2 (PcG component) (D'Anello et al., 2010) and several studies have shown that TGF- $\beta$  treatment regulates the expression of its target gene through modulation of the promoter DNA methylation (Dong et al., 2012; Eades et al., 2011; Kim and Leonard, 2007; Thillainadesan et al., 2012; You et al., 2010b). Most of the time these epigenetic regulations involve direct recruitment of DNMT or TET on the gene promoters, and are sometimes preceded by histone modifications. In such cases, DNA methylation is thus a more secure system, to ensure the inflammation-induced silencing of genes. Contrary, epigenetic mechanisms can interfere with inflammatory pathways, in particular for the activation of the JAK/STAT3 pathway. HCC sample analyses revealed aberrant silencing of JAK/STAT inhibitors *SOCS-1* and *SOCS-3* by methylation resulting in constitutive activation of the pathway (Calvisi et al., 2006; Niwa et al., 2005). All these examples support the hypothesis that inflammation and epigenetics are not independent events but act in close collaboration to initiate HCC.

**Figure 31. A hypothetical model depicting cross-talk between activation of inflammatory pathways and epigenome deregulation during liver tumor development.**

Different components of the inflammatory response (including transient and stable modifications such as activation of inflammatory pathways nuclear factor (SMAD and JAK/STAT) may induce changes in epigenetic machineries (including DNA methylation, histone modifications and non-coding RNAs), resulting in an 'epigenetic switch' that resets the long-term memory system in hepatocytes. The epigenetic switch in turn may contribute to a persistent inflammatory response through altered gene expression states and a positive feedback loop to exacerbate a chronic state of inflammation. In addition, the deregulated epigenome may maintain an altered transcriptional program that promotes proliferation and oncogenic transformation. This interdependent and self-reinforcing cross-talk between inflammation and the epigenome maintains and amplifies inflammatory signals, resulting in a series of events culminating in the development of liver cancer. The epigenetic switch may also be activated in hepatic or liver progenitor cells whose proliferation is stimulated during liver regeneration and repair. Therefore, an inflammatory microenvironment and an epigenetic switch in response to different environmental factors can directly promote activation of liver progenitor cells and their oncogenic transformation. DNMT, DNA methyl transferase (adapted from Martin and Herceg, 2012).



### 3. DNA methylation and cancer stem cell phenotype

DNA methylation, like other epigenetic marks, is stable, can be passed through cell divisions but remains reversible. The higher dynamism for DNA methylation is observed in embryonic stem cell, at the very early steps of zygote development (Bergman and Cedar, 2013). There is a global demethylation process engaged before implantation of the zygote in order to “erase” the germline programming and to reset totipotency (key master genes such as *NANOG*, *OCT4* and *SOX2* are silenced through hypermethylation in sperm DNA Farthing et al., 2008). After implantation, DNMT3A and DNMT3B are mobilized to establish a new DNA methylation profile. During this wide *de novo* methylation, low CpG content promoters (usually associated with tissue-specific genes) are highly methylated while dense CpG island promoters will remain protected, and thus relatively permissive for the transcription of the genes they belong to (Koh and Rao, 2013). In somatic cells, the DNA methylation pattern is believed to be rather stable and any changes are likely to be rare and to come from “environmental consequences” and/or aging (Bergman and Cedar, 2013). In stem cells, the epigenetic program allows the expression of genes involved in self-renewal and pluripotency but at the same time shall be able to respond to any stimulus to launch differentiation (like liver progenitor cell activation under chronic liver inflammation). In consequence the chromatin state of stem cells is often described as open and flexible and may be subject to epigenetic reprogramming. Deregulation in stem cell differentiation can come from epigenetic alterations where stem cells slowly acquire irreversible silencing of key master regulators required for successful differentiation. In such a model, deregulated stem cells would lose their ability to differentiate while retaining their self-renewal ability. These two conditions are sufficient to favour malignant transformation through additional epigenetic and genetic alterations (Shen and Laird, 2013). Interestingly in this model epigenetic deregulations would be the first hit for stem cell transformation. This transformation could give rise to CSCs and non stem cancer cells. This model has been proposed after several observations: normal mammary gland stem/progenitor cells continuously exposed to estrogen developed a DNA methylation pattern resembling cancer methylome (Cheng et al., 2008) and cancer cells DNA methylation pattern often involves hypermethylation of genes involved in the specific differentiation of their cell of origin (Sproul et al., 2012). Notably genes occupied by PcG (proteins involved in the silencing of genes regulated the differentiation) are more prone to promoter hypermethylation during cell proliferation and malignant transformation (Ohm et al., 2007; Widschwendter et al., 2007). Finally a recent

study observed a distinct methylation signature in Huh7 and PLC/PRF/5 side populations (Zhai et al., 2013). This indicates that DNA methylation remodelling plays an important part in the transformation of CSCs and non stem cancer cells.

Describing of the cellular reprogramming in iPSC (induced pluripotent stem cells) has been beneficial to understanding stem cell dysregulations. iPSC relies initially on key master genes expression (*OCT4*, *SOX2*, *C-MYC* and *KLF4*) followed by an epigenetic remodelling that will permit the secondary transcription of genes involved in pluripotency, self-renewal and proliferation (Li and Laterra, 2012). DNA methylation pattern modification with a global demethylation seems to be necessary (Gao et al., 2013b). Besides, loss of *DNMT3A* has been shown to block hematopoietic stem cell differentiation (Challen et al., 2012) in mice and inhibition of methylation in human HCC cell lines results in increased tumorigenicity in the CSC side population (Marquardt et al., 2010). However the exact role of DNMTs in CSC programming is not obvious as the inhibition of DNMT1 in leukemia was at the opposite correlated with reduction of the tumor growth and impaired CSC functions (Trowbridge et al., 2012). Nevertheless this underlies the important role of DNA methylation pattern acquisition during transformation. In MCF7 (breast cancer cell line) and Huh7 (HCC cell line) cells, the DNA methylation profile for TSG has been compared between CSCs and non stem cancer cells, and DNA methylation level was always found lower in CSCs (Yasuda et al., 2010). This difference can be associated to the less differentiated status of CSCs. In addition, the expression of CD133 protein in CSCs is also regulated through methylation. In liver, ovarian, colorectal and glioma tumors, CD133 promoter is hypomethylated in CD133+ cells compared to CD133- cells (Baba et al., 2009; Yi et al., 2008; You et al., 2010a). Interestingly this type of regulation for CD133 has not been reported in normal cells. As CD133 might be directly involved in the pathways regulating CSCs, this would signify that CD133+ cell's DNA methylation pattern would not only be a signature of CSCs but could contribute to the homeostasis of this subpopulation. These observations strongly support the idea that methylation is not only important in stem cell but also for CSC programming and could serve cancer development through maintenance of this subpopulation.



**HYPOTHESIS  
AND AIMS OF THE PROJECT**



In liver cancer samples and HCC cell lines, CD133+ cells have been reported to represent subpopulations of cells called cancer stem cells. These cells show a higher ability to induce tumors in SCID/ NOD mice and to reproduce the heterogeneity of the tumor. They have been linked to tumor aggressiveness, metastasis and bad prognosis (Ma, 2013). These cells are also believed to support tumor growth, and as they exhibit increased resistance to chemotherapy, they could be responsible for tumor relapses often observed in patients. CSCs thus provoke high interest as they represent a prominent target for future therapy research. Many studies have therefore attempted to characterize these cells. Specific pathway activation such as Wnt/ $\beta$ -catenin and Hedgehog have been described in these cells, but so far a thorough characterization of CD133+ cells in liver cancer is lacking.

Cell fate decisions are governed by non-genetic processes that are maintained through cell divisions. These processes are mediated by epigenetic mechanisms such as DNA methylation and histone modifications. Notably, cancer cells show a loss of their original tissue features and this is associated with the observation that DNA methylation is markedly deregulated in human malignancies. However whether CSCs display a distinct DNA profile (sustaining their distinct phenotype) is not known. DNA methylation can be influenced by both internal cellular and environmental factors. In the case of hepatocellular carcinoma (HCC), the most frequent primary liver cancer form, malignancy development is usually associated with a chronic inflammation (Martin and Herceg, 2012). During chronic hepatitis, hepatocyte proliferation is activated through paracrine signals involving cytokines. Interestingly, the transforming growth factor beta cytokine (TGF- $\beta$ ) has been linked to both tumor suppression at early stages of HCC and tumor progression at later stages. Besides there is recent evidence that TGF- $\beta$  is able to influence the expression of DNMTs, and therefore, potentially affect DNA methylation states (Pan et al., 2013).

In this context, we raise the hypothesis that CD133+ CSCs harbour a specific DNA methylation program supporting their phenotype and that this phenotype might be triggered or influenced by their microenvironment (conditions like inflammation).

The two main questions that we want to answer are the following:

- Do liver CSCs display a specific DNA methylation signature that supports their phenotype ?
- Are liver CSCs and their putative DNA methylation signature sustained by external inflammatory stimulus such as TGF- $\beta$ ?
-

To answer these questions, our main objectives were:

- To select a relevant marker for identification of CSCs in liver cancer cell lines. This would allow us to conveniently perform a genome wide methylation study.
- To establish an assay for magnetic cell separation based on the selected marker which would allow us to perform *in vitro* study and microarray profiling on purified population of CSCs.
- To perform a genome wide methylation assay to compare liver CSCs with non-stem cancer cells in at least two independent HCC cell lines.
- To describe *in vitro* the effect of TGF- $\beta$  exposure on liver CSCs.
- To study the ability of TGF- $\beta$  in inducing DNA methylation changes in liver cancer cells, and to investigate the link between TGF- $\beta$  exposure and liver CSC DNA methylation program.

## **MATERIALS AND METHODS**



#### Cell culture

Huh-7, Hep3B, HepG2 and PLC/PRF/5 (American Type Culture Conditions) were cultured in DMEM medium high glucose with L- Glutamine (Gibco) supplemented with 10% foetal bovine serum (Gibco), 1% Penicillin and Streptomycin (Gibco), 1% Sodium Pyruvate (Gibco) and 1% with Glutamine (Gibco). All cell lines were incubated at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub> and were regularly tested for mycoplasma contamination (MycoAlert detection kit, LONZA).

#### Cytokines treatment

Two cytokines, IL-6 and TGF- $\beta$ 1 (recombinant human, Peprotech) were used to treat HepG2 and Huh7 cell lines. Cells were plated and allowed to get adherent for at least 4h and fresh medium containing 10ng/ml final of IL-6 or TGF- $\beta$ 1 was added. For each condition medium was renewed after 3 days, and cells were collected after 4 days of treatment. For experiments investigating the stability of the effects induced by TGF- $\beta$ 1, after 4 days of treatment, cells were washed once with PBS, fresh medium without cytokines was added, and cells were left in culture for 4 additional days.

#### Sphere formation assay

For sphere formation assay hepatosphere medium (from N.Haraguichi et al, 2010) was used: DMEM F12 (Gibco) completed with L-glutamine 1X (Gibco), sodium Pyruvate 1X (Gibco), non-essential- amino-acid MEM 1X (Gibco), 10mg/L recombinant human insulin (Sigma-Aldrich), 1  $\mu$ M dexamethasone (Sigma), 200  $\mu$ M L-ascorbate-2-phosphate (Sigma-Aldrich), 10mM Nicotinamide (Sigma-Aldrich), penicilline/streptomycine 1X (PAA), 20 ng/mL human EGF and 10ng/mL human FGF. After trypsinization, cells were counted, tested for viability with Trypan blue and washed once with hepatosphere medium. 15,000 cells were plated in low attachment 6 wells plates (Corning) with 2ml of hepatosphere medium. Spheres were counted after 5 or 6 days.

#### BrDU assay

BrDU (Sigma) was added directly in cell's medium at 100uM final and cells were allowed to incorporate it for 1h. After trypsinization, cells were pelleted and 70% ethanol was added drop wise to the pellet. Cells were stored at -20°C for at least 1h.

Pellets were washed once with wash buffer (PBS containing 0.5% BSA) and DNA was

denaturated by adding 0.5ml of 2M HCl for 20min at room temperature (RT). After one washing, to neutralize any residual acid, cells were resuspended in 0,5ml of 0,1M of sodium tetraborate (Na<sub>2</sub>B<sub>4</sub>O<sub>7</sub>) and incubated 2min. Cells were washed twice, resuspended with BrDU antibody (Sigma Aldrich, diluted 1:500 in wash buffer) and incubated 20min at RT. Cells were then incubated with an anti-mouse-FITC antibody (Sigma, diluted 1:100 in wash buffer) for 20min at RT. Labelled cells were finally resuspended in 225ul of wash buffer, and 25ul of propidium iodide was added just before FACS analysis. BrDU staining was analysed using FACS instrument (and FACSCanto II, BD Biosciences) data were collected using BD FACSdiva 6.0 (BD Biosciences software) and analysed with FlowJo or WinMDI software.

#### Fluorescence Activated cell sorting (FACS)

Cells were labelled with anti-human CD44 antibody, anti-human CD133 antibody, anti-human EpCAM antibody, CD90 antibody or anti-human-TGFBRII (see Table 7 for details). Non-conjugated primary antibodies were detected with a secondary antibody conjugated alternatively with FITC, Cy3 or Alexa750 (see Table 7 for details) For each staining, antibodies were diluted in PBS containing 3% foetal bovine serum (see Table 7 for working dilution) and incubated during 30min in the dark at RT. Fluorescence was analysed using FACS instrument (and FACSCanto II, BD Biosciences) and data were collected using BD FACSdiva 6.0 (BD Biosciences software) and analysed with FlowJo or WinMDI software.

**Table 7. List of antibodies used for fluorescent activated cell sorting (FACS).**

	Fluorochrome	Isotype and origin	Company	Working dilution
<i>Primary antibodies</i>				
Anti-human CD44	PE	G44-26 Mouse IgG2b	BD Pharmingen	1:10
Anti-human CD133	none	AC133 mouse IgG1	Miltenyi Biotec	1:10
Anti-human EpCAM	PerCP	clone 1B7 mouse IgG1	Ebiosciences	1:50
Anti-human CD90	FITC	5E10 mouse IgG1	Stem cell technologies	1:10
Anti-human TGFBRII	PE	Goat	RD system	1/10
<i>Secondary antibodies</i>				
Anti-mouse-FITC	-	-	Sigma	1:100
Anti-mouse Cy3	-	-	Sigma	1:200
Anti-mouse-Alexa 750	-	-	Invitrogen	1:100

#### Magnetic Activated cell sorting (MACS).

Huh7 and HepG2 cells were depleted or enriched for CD133+ cells using Miltenyi MACS system (CD133 microbead kit, LS columns and MidiMACS separator, Miltenyi Biotec).



Manufacturer's instructions did not allow us to obtain a satisfactory enrichment for CD133+ cells (at least 2.5 fold enrichment). In order to increase the efficiency of the sorting, the initial protocol was optimized. The procedure described below corresponds to the final optimized protocol. The differences between the manufacturer's instructions and our optimized protocol are summarized in Figure 32 and examples of fractions enriched in CD133+ cells are presented Figure 33.

During the entire procedure, cells were kept in cold PBS completed with 2% FBS and 2mM EDTA. This buffer will be referred hereinafter as MACS buffer.

#### Magnetic Labelling :

After trypsinization, cells were filtered, counted and incubated 30min at 4°C on a wheel with FcR blocking Reagent (diluted 1:3 in MACS buffer, 450ul for 10<sup>8</sup> cells). MicroBeads conjugated to monoclonal anti-human CD133 antibodies were then added to the cell suspension (final dilution 1:4) and incubated 15min at 4°C on a wheel. Reaction was stopped with one wash of 5ml of MACS buffer, cells were centrifuged and resuspended in 2-4-ml of MACS buffer. Cells suspension was then applied onto a pre-rinsed LS column placed in the magnetic field of a MACS separator.

#### Magnetic separation

In order to obtain clear distinct fractions, two different procedures were used for CD133+ cells depletion or CD133+ cells enrichment.



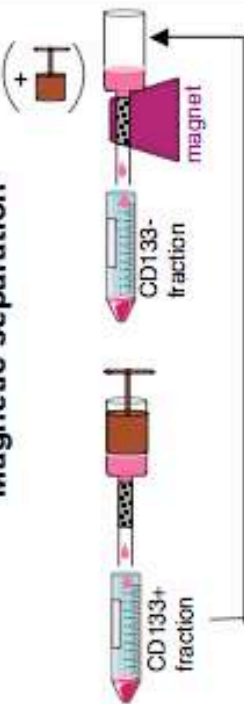

##### CD133+ cells depletion.

Immediately after application of the cell suspension onto the LS column, flow-through containing unlabelled cells was collected. The column was then washed 3 times with 4ml of MACS buffer, and the flow-through fraction was collected and combined with the effluent from the previous step to constitute the CD133 negative fraction. For each experiment aliquots were collected to test by FACS the efficiency of the depletion.

##### CD133+ cells enrichment :

After application of the cell suspension onto the LS column, and flow of the unlabeled cells, the column was washed 3 times with 4 ml of MACS buffer. To improve the efficiency of the washings, a plunger was softly used during all the washing steps. The column was then removed from the separator and placed on a 15ml collection tube. 5ml of MACS buffer was applied onto the column and labelled cells were collected by firmly pushing the plunger in

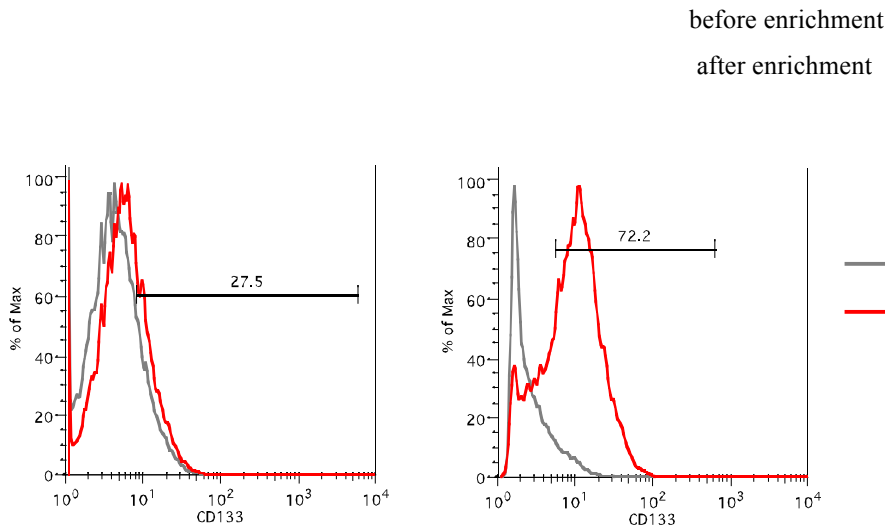
the column. This last step was repeated once in order to avoid any labelled cells remaining in the column. To increase purity of CD133+ cells, the eluted fraction was enriched a second time over a new LS column following the exact same procedure for enrichment. For each experiment, aliquots were kept to test by FACS the efficiency of the enrichment.

		Manufacturer's procedures	Optimized procedures
Cell collection		<ul style="list-style-type: none"> <li>• Trypsinization</li> <li>• Washing</li> </ul>	
Magnetic labeling		<ul style="list-style-type: none"> <li>• Addition of FcR blocking reagent and CD133 MicroBeads</li> <li>• 30 min 4°C</li> </ul>	<ul style="list-style-type: none"> <li>• Addition of FcR blocking reagent</li> <li>• 30 min 4°C</li> <li>• Addition of CD133 MicroBeads</li> <li>• 15 min 4°C</li> </ul> <p><i>objectives : increase the efficiency of the blocking step and the specificity of the staining</i></p>
<b>Deposition of cell suspension onto LS column</b>			
Magnetic separation		<p><u>For depletion and enrichment</u></p> <ul style="list-style-type: none"> <li>• Collection of negative fraction by washing 3X 3ml</li> </ul>	<p><u>For depletion</u></p> <ul style="list-style-type: none"> <li>• Collection of negative fraction by washing 3X 4ml</li> </ul> <p><u>For enrichment</u></p> <ul style="list-style-type: none"> <li>• Collection of negative fraction by washing 3X 4ml with soft use of the plunger</li> </ul> <p><i>Objectives : increase the washings by forcing the flow of unlabelled cells</i></p>
		<ul style="list-style-type: none"> <li>• Collection of positive fraction with 5ml and firm use of plunger</li> <li>• possibility to sort again the positive fraction</li> </ul>	<ul style="list-style-type: none"> <li>• Collection of positive fraction with 2X 5ml and firm use of plunger</li> <li>• possibility to sort again the positive fraction</li> </ul>
FACS analysis	 See Figure		

**Figure 32. Main steps for magnetic activated cell sorting.**

This scheme is representing main steps of the MACS protocol and the differences between manufacturer's and our optimized procedures.

B



**Figure 33** Dot plots of cells fractions enriched in CD133+ Huh7 cells analysed by FACS.

A. CD133+ cells fraction obtained after 2 columns following manufacturer's procedures. B. CD133+ cells fraction obtained after 2 columns following the optimized procedures

#### Cell sorting

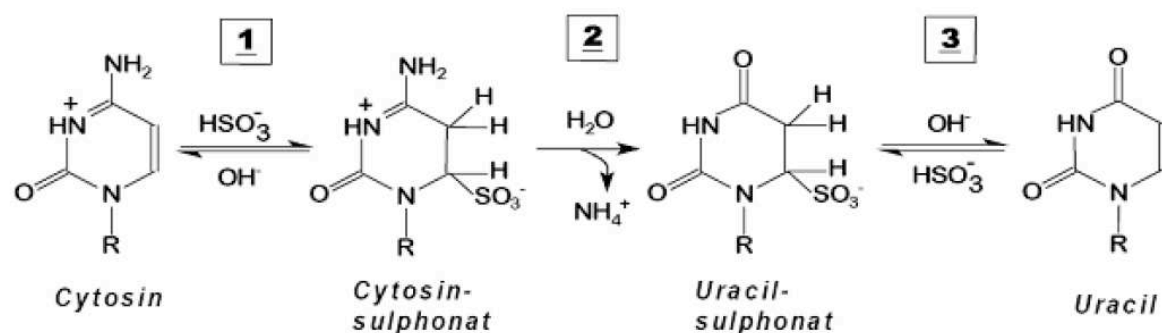
For DNA methylation bead arrays, Huh7 and HepG2 CD133+ cells were sorted using a BD FACS Aria III SORP cell sorter apparatus in the CRCL (Centre de recherche en cancérologie de Lyon) flow cytometry platform. Cells were labelled using anti-CD133 antibody (see Table 7) and a secondary anti-mouse antibody coupled with Cy3 (see Table 7). Cells were first gated with the SSC-A and FSC-A parameters to exclude dead cells and debris, then were filtered for singlet using the FSC-W and FSC-H parameters and finally were sorted according to their fluorescence using the FL2 channel.

#### DNA extraction

After trypsinization, cells were pelleted and resuspended in TAIL buffer (1% SDS, 0.1M NaCl, 0.1M EDTA, 0.05M Tris pH8) with Proteinase K (500ug/ml) and incubated for 2 to 3 hours at 55°C. Saturated NaCl (6M) was then added and after centrifugation (10min, Vmax), the supernatant was transferred into a new tube. DNA was precipitated with isopropanol, and the pellet was cleaned with 70% ethanol. Extracted DNA was finally resuspended in water. Quantity and quality of the extracted DNA were assessed with a ND-8000 spectrophotometer (Nanodrop, Thermo scientific). DNA pellets were stored at -20°C until use..

### Bisulfite treatment

To quantify the percentage of methylated cytosine in individual CpG sites, we performed a bisulfite treatment on the DNA. This technique consists of treating DNA with bisulfite, which causes unmethylated cytosines to be converted into uracil (Figure 34) while methylated cytosines remain unchanged. Then, the methylated and unmethylated cytosines can be easily distinguished. For samples directly analyzed by pyrosequencing, the conversion was performed on 150 to 500 ng of DNA using the the EZ DNA methylation Gold Kit (Zymo Research) and modified DNA was eluted in 15ul of water (samples were stored at -20°C until use). For samples processed on the bead array, the conversion was performed on 600ng of DNA using the EZ DNA methylation Kit (Zymo Research) and modified DNA was eluted in 16ul of water.



**Figure 34. Chemical steps occurring during bisulfite conversion.**

### Pyrosequencing

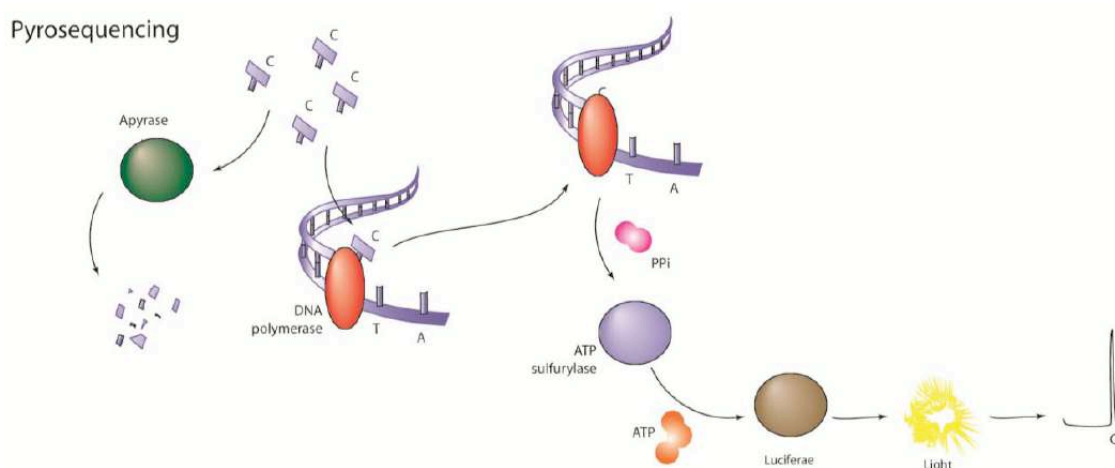
Pyrosequencing is a sequencing-by-synthesis method that quantitatively monitors the real-time incorporation of nucleotides through the enzymatic conversion of released pyrophosphate into a proportional light signal.

Modified DNA (10-25 ng) was amplified in a total volume of 50  $\mu\text{L}$ . 10  $\mu\text{L}$  of PCR reaction and was analyzed on agarose gel, and the remaining 40  $\mu\text{L}$  were used in a pyrosequencing assay. The PCR products were collected and purified from the reaction mixture by binding onto streptavidin-coated sepharose beads (Amersham-GE Healthcare), which recognize biotinylated strands, on the vacuum-based workstation provided with the PSQTM96MA instrument (Qiagen) in a 96-well plate. The biotinylated PCR products were washed in a 70% ethanol bath, denatured with 200nM NaOH solution and then mixed with sequencing primer. The mixture was incubated at 80°C for 2 minutes and allowed to cool down at RT for 20 minutes in order to reach the specific primers annealing temperature.

Pyrosequencing reactions were set up using PyroGold Reagent kit (Qiagen) according to the manufacturer's instructions. The template DNA is immobile, and solutions of A, C, G, and T nucleotides are sequentially added and removed from the reaction. As the nucleotide dATP acts as a natural substrate for luciferase, the modified  $\alpha$ -S-dATP is used as the nucleotide for primer extension as it is equally well incorporated by the polymerase. Light is produced only when the nucleotide solution complements the first unpaired base of the template.

Single-strand DNA template is hybridized to a sequencing primer and incubated with the enzymes DNA polymerase, ATP sulfurylase and apyrase and with the substrates adenosine 5' phosphosulfate (APS) and luciferin (Figure 35).

Pyrosequencing assays (primers for PCR, sequencing primers and regions are described Table 8).



**Figure 35. Pyrosequencing methods (Herceg and Vaissière)**

**Table 8. List of primers of pyrosequencing assays (see next page)**

<b>Genes</b>	<b>Primers for PCR</b>	<b>UCSC localization</b>	<b>Annealing temperature</b>	<b>Sequencing primers</b>	<b>Sequence analyzed</b>	<b>Strand</b>	<b>Infinium450K probe</b>
•DNMT1	For GTATTGGGGATTAAAGAG Rev Btn-TTAATACATCCCTCCTC	Chr19:10,305,553-10,305,597	49	GAAGTGTTATTTTGT	CGCGTTAATTGCTGTGCGCG	+	
•DNMT3a (assay1)	For TTAGTATTGGGTTGGGATAGTAG Rev Btn-CACTCCCTTCAAAACAACATCATCTC	chr2:25,565,470-25,565,648	60	TTATAGGGTTAAGGT	YGGAGYGTAGGTTTTTTTGGT	+	
•DNMT3a (assay2)	For TGAGTTTGGGAGAGGAG Rev Btn-CCCAACCTACTACTACAAA	chr2:25,481,174-25,481,364	59	TTTAGAATTGTAAAG	YGAATTGTGTTTATT	+	cg04058399
•DNMT3b (assay1)	For AGGGTTTGTAGGGGAGGGGA Rev Btn-CCTCACCTCTCTACCCCTT	chr20:31,350,050-31,351,0363	61	ATTTTGTGGGGTGG AA	YGGGGATAGYGGGTGGGAYGYGGG	+	
•DNMT3b (assay2)	For GGTAAGGTAGGGTTTGGGA Rev Btn-GGTAGGTAGGGTTTGGGA	chr20:31,366,350-31,366,545	57	TATTAGATTGAGTTTA GTGTAGTTTGGAGTT	YGGTGTTTTTTGGTGGYGATGTT YGTAAATAAGGTGTGG	+	cg24403338 cg00300969
•TRRAP	For GTGGTTTGGAAATTGTTTAGG Rev Btn-ATCTATACCAACTTCTCCCTCC	chr7:98,520,412-98,520,578	57	TATTTTGTAGAAGA	YGTGTGTTTATGTAGT	+	cg21421984
•TET2	For AGGGTAGGAATGGGTTAG Rev Btn-CTATCTTCCCTCTCTCTTAAC	chr4:106,114,429-106,114,753	53	GAAGTAGGAAAAAAGG	YGATTTTGTATAAA	+	cg22794775
•TWIST	For AGTATATAGTGTGGGGTGG Rev Btn-ACACTCACCTCCTCCTCTC	chr7:19,158,612-19,158,941	52	ATTAATTGAGTAAAG	YGTTTTTTTTGGGTT	+	cg10624122
•CD68	For TAGGGGAGAGAAAAATGGTAGT Rev Btn- TAACCATCCACACCCCAAC	chr17:7,482,223-7,482,541	55	ATTGAGTTTTTAGA	YGTGGAAAAGTTATGTTTTYGGT	+	cg18900669
•PDLM7	For GGAAGAAAGTTAGGGGGTTG Rev Btn- ATTAACCAACCCCTCCTC	chr5:176,921,744-176,921,917	55	AGTTTTTTAGTTTTAGA	YGTATTTTATTTT	+	cg152525325

### Bead Array Platform

Two bead array assays were performed. The first one to compare the methylation profile between CD133+ and CD133- cells, the second to compare the methylation profile between cells treated or not with TGF- $\beta$ 1 (see Results section for further details on experiment design).

Genomic DNA (600 ng) from Huh7 and HepG2 cells was subjected to bisulfite treatment. Quality of modification was checked by PCR using modified and unmodified primers for *GAPDH* gene. Methylation profiles of the different samples were analysed using the 450K Infinium methylation bead arrays (Illumina, San Diego, USA). Briefly, the Infinium Humanmethylation450 beadchip interrogates more than 450,000 methylation sites. 99% of REfSeq genes are covered (including these of low CpG islands density and at high risk for being missed by other commonly used methods). The coverage is targeted across gene regions with sites in the promoter regions, the 5'UTR, the first exon, the gene body and the 3'UTR regions. Beyond genes and CpG islands, multiple additional content categories are also included (CpG sites outside CpG islands, non CpG methylated site identified in human stem cells, DNA hypersensitive sites etc.). In conclusion this methylation bead array provides a high coverage and low bias technique to interrogate DNA methylation profile in different sample types.

The analyses on the bead array was conducted following the recommended protocols for amplification, labelling, hybridisation and scanning. Each methylation analysis was performed in duplicate (for CD133+ versus CD133- samples) or in triplicate (for samples treated with TGF- $\beta$ 1). GenomeStudio Methylation Module software (V2010.3, Illumina) was used to obtain raw data and display beta values. All samples passed data quality controls. Differential methylation data comparing the two phenotypes (CD133+ vs. CD133- or TGF- $\beta$ 1 vs. control) were obtained using the BRB-ArrayTools software with respectively CD133- cells or non treated cells DNA as a reference. Using Infinium annotation data, Infinium sites (cytosines) were classified according to their relation to CpG islands and to the closest annotated gene. Sites unrelated to any annotated gene were classified as intergenic.

To validate the data obtained by Infinium methylation bead arrays in all samples, 8 to 10 CpG sites were selected for their difference of methylation and analyzed a second time by pyrosequencing as described above.



### RNA extraction

Total RNA was isolated using the TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. Briefly 1ml of TRIzol Reagent was added on the cells pellets. Cells were centrifuged and supernatant was collected in a new Eppendorf. 200ul/ml TRIzol was added, cell suspension was vortexed and left at RT for 15min. After centrifugation the aqueous phase was collected in a new tube, and RNA was precipitated with 500ul of isopropanol. The RNA pellet was then washed once with 75% ethanol and finally resuspended in water. RNA quantity and quality were assessed with a ND-8000 spectrophotometer (Nanodrop, Thermo scientific). Pellets were stored at -80°C until use.

### Reverse transcription and quantitative PCR

Reverse transcription reactions were performed using MMLV-RT (Invitrogen) and random hexamers on 500 ng of total RNA per reaction according to the manufacturer's protocol. Quantitative PCR was done in triplicate for each condition using the Mesa Green qPCR MasterMix Plus for SYBR Assay buffer (Eurogentec). The qPCR was performed with a CFX96T touch real time system (BIO-RAD). HPRT1 and GAPDH were used as housekeeping genes and in case of contradictory results two supplementary HCC-specific housekeeping genes (SFRS4 and TBP1) (Waxman and Wurmbach, 2007) were used. The different primers used are listed in Table 9.

**Table 9. List of primers designed for qRT-PCR.**

<i>HPRT1</i>	<i>for 5'-CATTGTAGCCCTCTGTGTGC-3'</i> <i>rev 5'-CACTATTTCTATTCAGTGCTTTGATGT-3'</i>	<i>SOX2</i>	<i>for 5'-AAGACGCTCATGAAGAAGGATAA-3'</i> <i>rev 5'-ACTGTCCATGCGCTGGTT-3'</i>
<i>GAPDH</i>	<i>for 5'-AACGGGAAGCTTGTCATCAA-3'</i> <i>rev 5'-TGGACTCCACGACGTACTCA-3'</i>	<i>BMP1</i>	<i>for 5'-CAAGGCCCACTTCTTCTCAG-3'</i> <i>rev 5'-CATAACTGCCGAACGTGTTG-3'</i>
<i>SFRS4</i>	<i>for 5'-GGCTACGGGAAGATCCTGGA-3'</i> <i>rev 5'-TGCATCACGCAGATCATCAA-3'</i>	<i>ERLIN1</i>	<i>for 5'-GATTGAGGAGGGCCATCTG-3'</i> <i>rev 5'-GGTCCACTGGGGCTAGTTAGT-3'</i>
<i>TBP1</i>	<i>for 5'-TATAATCCCAAGCGGTTTGC-3'</i> <i>rev 5'-CACAGTCCCCACCATATTC-3'</i>	<i>HDAC7</i>	<i>for 5'-GGTGTCTAGACGCACAGAAAT-3'</i> <i>rev 5'-CATGACCGAGTCATAGATCAGC-3'</i>
<i>CD133</i>	<i>for 5'-TCCACAGAAATTTACCTACATTGG-3'</i> <i>rev 5'-CAGCAGAGAGCAGATGACCA-3'</i>	<i>RECE</i>	<i>for 5'-TGAAGAAGTCGGCCAAGAAG-3'</i> <i>rev 5'-CGCTGGCGTTTGTTACTCTT-3'</i>
<i>DNMT3A</i>	<i>for 5'-CCTGAAGCCTCAAGAGCAGT-3'</i> <i>rev 5'-TGGTCTCCTTCTGTTCTTTGC-3'</i>	<i>ZEB1</i>	<i>for 5'-GCTGGGAGGATGACAGAAAG-3'</i> <i>rev 5'-TGCATCTGACTCGCATTCAT-3'</i>
<i>DNMT3B</i>	<i>for 5'-CAATGGCTTCAGATGTTGC-3'</i> <i>rev 5'-TCCTGCCACAAGACAAACAG-3'</i>	<i>COL18A1</i>	<i>for 5'-AGGAAGGACTGGGCAGAAA-3'</i> <i>rev 5'-CTCCCTTGCTCCCCTTATGT-3'</i>
<i>DNMT1</i>	<i>for 5'-GATGTGGCGTCTGTGAGGT-3'</i> <i>rev 5'-CCTTGCAAGCTTTACATTTCC-3'</i>	<i>CALD1</i>	<i>for 5'-CGTCGCAGAGAACTTAGAAGG-3'</i> <i>rev 5'-ATTCCTCTGGTAGGCGATTCT-3'</i>
<i>TET1</i>	<i>for 5'-GCTATACACAGAGCTCACAG-3'</i> <i>rev 5'-GCCAAAAGAGAATGAAGCTCC-3'</i>	<i>CALM2</i>	<i>for 5'-ATGGCTGACCAACTGACTGA-3'</i> <i>rev 5'-CAGTCCCAATTCTTTGTTG-3'</i>
<i>TET2</i>	<i>for 5'-CTTCTCCTCCCTGGAGAACAGCTC-3'</i>	<i>BRD2</i>	<i>for 5'-CCCTAAGAACAGCCACAGAA-3'</i>

	<i>rev</i> 5'-TGCTGGGACTGCTGCATGACT-3'		<i>rev</i> 5'-GGTATCTCAGGTGGAGGAGTAT-3'
<i>TGFB</i>	<i>for</i> 5'-GCACGTGGAGCTGTACCA-3' <i>rev</i> 5'-AAGATAACCACTCTGGCGAGTC-3'	<i>STAT3</i>	<i>for</i> 5'-AACTTCAGACCCGTCAACAAA-3' <i>rev</i> 5'-GGGTCCCCTTTGTAGGAAAC-3'
<i>SNAIL</i>	<i>for</i> 5'-ATCCGAAGCCACACACTG-3' <i>rev</i> 5'-CACTGGTACTTCTTGACATCTG-3'	<i>JAK2</i>	<i>for</i> 5'-GGTGAAAGTCCCATATTCTGGT-3' <i>rev</i> 5'-AGGCCACAGAAAACCTTGCTC-3'
<i>P21</i>	<i>for</i> 5'-GACACCACTGGAGGGTGACT-3' <i>rev</i> 5'-CCACATGGTCTTCCTCTGCT-3'	<i>NANOG</i>	<i>for</i> 5'-CAGCTGTGTGTACTCAATGATAGATT-3' <i>rev</i> 5'-TCTGGAACCAGGTCTTCACC-3'
<i>CCDN1</i>	<i>for</i> 5'-GAAGATCGTCGCCACCTG-3' <i>rev</i> 5'-GACCTCCTCCTCGCACTTCT-3'	<i>OCT4</i>	<i>for</i> 5'-GCTTCGGATTTCGCCTTC-3' <i>rev</i> 5'-CTTAGCCAGGTCCGAGGAT-3'

### Whole genome expression array

Total RNA was isolated using the TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. Briefly 1 ml of TRIzol Reagent was added on the cells pellets. Cells were centrifuged and supernatant was collected in a new eppendorf. 200ul/ml TRIzol was added, and cell suspension was vortexed and let at RT for 15min. After centrifugation the aqueous phase was collected in a new tube, and RNA was precipitated with 500ul of isopropanol. RNA pellet was then washed once with 75% ethanol and finally resuspended in water. RNA quantity and quality were assessed with a ND- 8000 spectrophotometer and bioanalyzer. Pellets were stored at -80°C until use. Using the Illumina TotalPrep RNA Amplification Kit (Lifetechnologies), 500 ng of total RNA from HepG2 and Huh7 cells treated or not with TGF- $\beta$  were reverse-transcribed and biotin-labeled cRNAs were then generated. The distribution of homogeneous cRNAs were checked with the Agilent bioanalyzer instrument and the RNA 6000 Nano kit and 750 ng were hybridized overnight to Human HT-12 Expression BeadChips (Illumina) targeting 25,000 genes with 48,804 probes covering RefSeq (including coding transcript with well-described or provisional annotation and non-coding transcript) and UniGene annotated genes. The hybridized chips were washed and processed to detect biotin- containing transcripts by streptavidin-Cy3 conjugate and scanned on a bead array reader (Illumina). For mRNA expression validation, 10 candidate genes were selected and their expression were re- analyzed by quantitative RT-PCR. To this end, reverse transcription reactions were performed using MMLV-RT (Invitrogen) and random hexamers on 500 ng of total RNA per reaction according to the manufacturer's protocol. Quantitative PCR was done in triplicate for each condition using the Mesa Green qPCR MasterMix Plus for SYBR Assay buffer (Eurogentec). The qPCR was performed with a CFX96T touch real time system (BIO-RAD). Four different housekeeping genes (*HPRT1*, *GAPDH*, *SFRS4* and *TBP1*) were used for internal control. The different primers used are listed in Table 9.

### Proteins extraction and Western Blot

Proteins were extracted in RIPA-like solution (50mM TrisHCl pH8, 150mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 0.1% SDS) complemented with protease inhibitors (complete Mini, Roche). Protein concentration was measured by spectrophotometer (biophotometer, eppendorf). Proteins were separated by SDS-PAGE and transferred on nitrocellulose membranes. Immunostaining was performed with anti SMAD3, anti-phosphorylated SMAD3 and anti-tubulin for loading control. Primary antibodies were detected with secondary antibodies (anti-mouse-HRP and anti-rabbit-HRP, DAKO) and revealed with ECL plus detection kit (on Amersham Hyperfilm ECL films).

### Statistical Analysis

Raw methylation and expression bead array data was exported from Genome Studio (version 2010.3, Illumina) into BRB-ArrayTools software (version 4.3.1, developed by Dr. Richard Simon and the BRB-ArrayTools Development Team). Data was normalized and annotated using the R/Bioconductor package lumi. Unsupervised clustering and class comparisons were performed as previously described. Only those probes with p values <0.001 and FDR<0.05 were considered significant for most analyses, except CD133- vs CD133+ comparison, where only the p value threshold was used. To define a “stable” methylome signature induced by TGF- $\beta$ , we performed a control vs TGF- $\beta$  class comparison blocking by cell line status (Huh7 or HepG2), and including day 4 and day 8 of exposure (day 8 corresponding to 4 days of exposure to TGF- $\beta$  + 4 additional days with control medium). Using Infinium annotation data, Infinium sites (cytosines) were classified according to their relation to CpG islands and to the closest annotated gene. Sites unrelated to any annotated gene were classified as intergenic.

BrB geneset class comparison tool and WebGestalt (WEB-based GENE SeT Analysis Toolkit) and DAVID (Database for Annotation, Visualization and Integrated Discovery) web applications were used for gene set enrichment analyses, including Gene Ontology, pathway, network module, and gene-phenotype associations (Huang et al., 2009; Wang et al., 2013).

Additional R/Bioconductor packages and our own scripts were used for the specific analysis modeling the effect of TGF- $\beta$  and CD133 expression in linear regression. Data loading and preprocessing was performed with the “watermelon” package. This was followed by batch correction using the ComBat function of the “sva” package and linear modeling using “limma”



## **RESULTS**



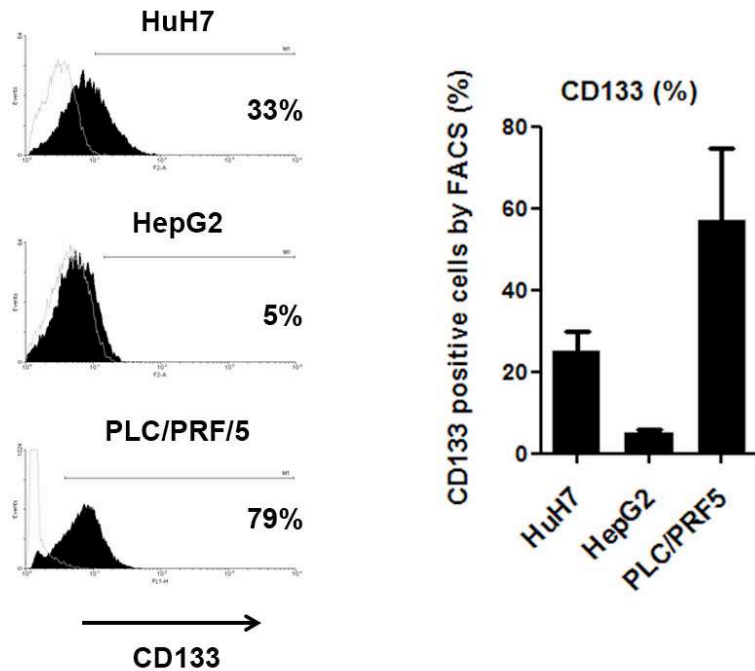
## **I. CD133- and CD133+ liver cancer cells differentially express DNA methylation genes**

CD133 is an established marker of CSCs in different types of human malignancies, including HCC (Grosse-Gehling et al., 2013). In a first step, we estimated the frequency of CD133 expressing cells in three non-related liver cancer cell lines: HuH7, HepG2 and PLC/PRF/5. These 3 cell lines are all originating from liver cancer (HCC or hepatoblastoma), they are all tumorigenic but non metastatic. Their main genetic characteristics are an integrated copy of the HBV genome for PLC/PRF/5 and a *TP53* missense mutation for PLC/PRF/5 and Huh7 (Table 10)

**Table 10. Characteristics of the 3 liver cancer cell lines used for the study.**

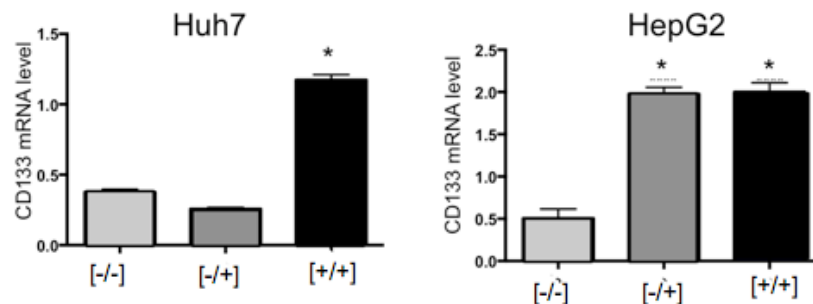
<i><u>Cell lines</u></i>	<i><u>Origin</u></i>	<i><u>HBV infection</u></i>	<i><u>Tumorigenicity</u> (number of cell injected - time for tumor mass emergence)</i>	<i><u>Metastatic potential</u></i>	<i><u>P53 status</u></i>
<b>Huh7</b>	HCC	-	Yes (10 <sup>6</sup> cells – 1 month)	No	Missense mutation (Y220C)
<b>HepG2</b>	Hepatoblastoma	-	Yes (10 <sup>7</sup> cells – 3 months)	No	Wild-type
<b>PLC/PRF/5</b>	HCC	+	Yes (10 <sup>6</sup> cells – 1 month)	No	Missense mutation (R249S)

Non-synchronized, exponentially growing cells were analyzed by FACS after staining with anti-CD133 antibody (AC133), which recognizes all common CD133 isoforms (Grosse-Gehling et al., 2013). The expression of CD133 was evident in all cell lines, ranging from 5% in HepG2, to 79% in PLC/PFR/5 cells and 5-15% standard deviation (Figure 36). Expression of the surface protein correlated well with CD133 expression at the mRNA level as cell population enriched for CD133+ cells with magnetic cell sorting (MACS) displayed a higher expression of CD133 mRNA (Figure 37) compared to cell populations depleted in CD133+ cells.



**Figure 36. CD133 expression in liver cancer cell lines.**

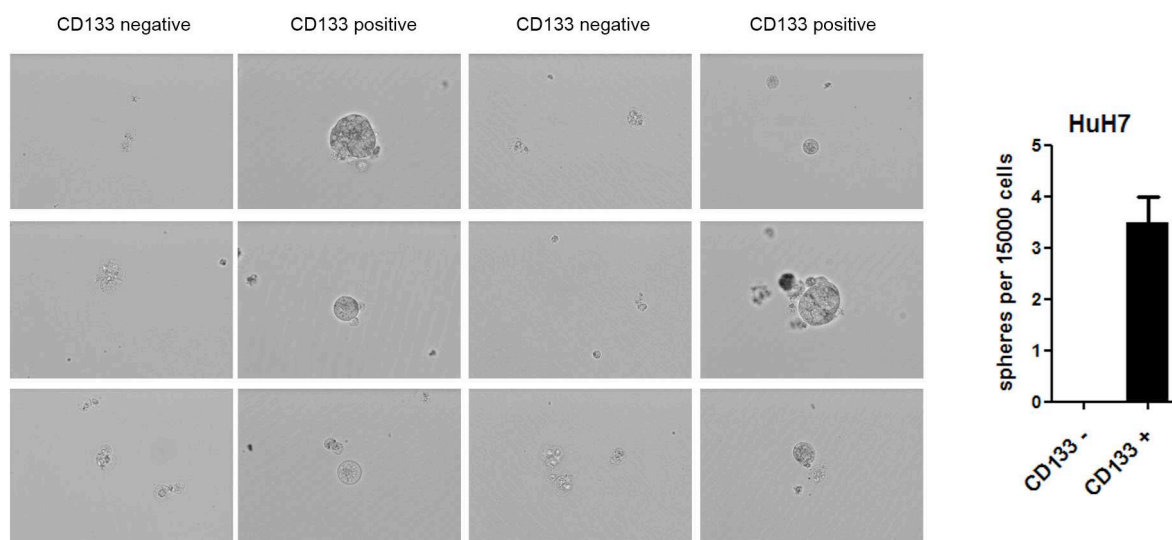
The expression of the stem cell marker CD133 was analysed by fluorescence activated cell sorting (FACS) in 3 independent cell lines, Huh7, HepG2 and PLC/PRF/5. The left panel shows a representative histogram for each of the cell lines (black histograms), with background (secondary antibody) represented by the empty histograms. The average expression  $\pm$  SD, from at least 3 independent assays, is shown on the right panel.



**Figure 37 CD133 gene (*PROM1*) is higher expressed in CD133+ cells.**

Huh7 and HepG2 cells were sorted by MACS and expression of CD133 was investigated by qRT-PCR in subpopulations differentially enriched for CD133+ cells ([-/-]; [-/+]; [+/+]). Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05.





**Figure 38. CD133+ cells are capable of producing spheres in low attachment conditions.**

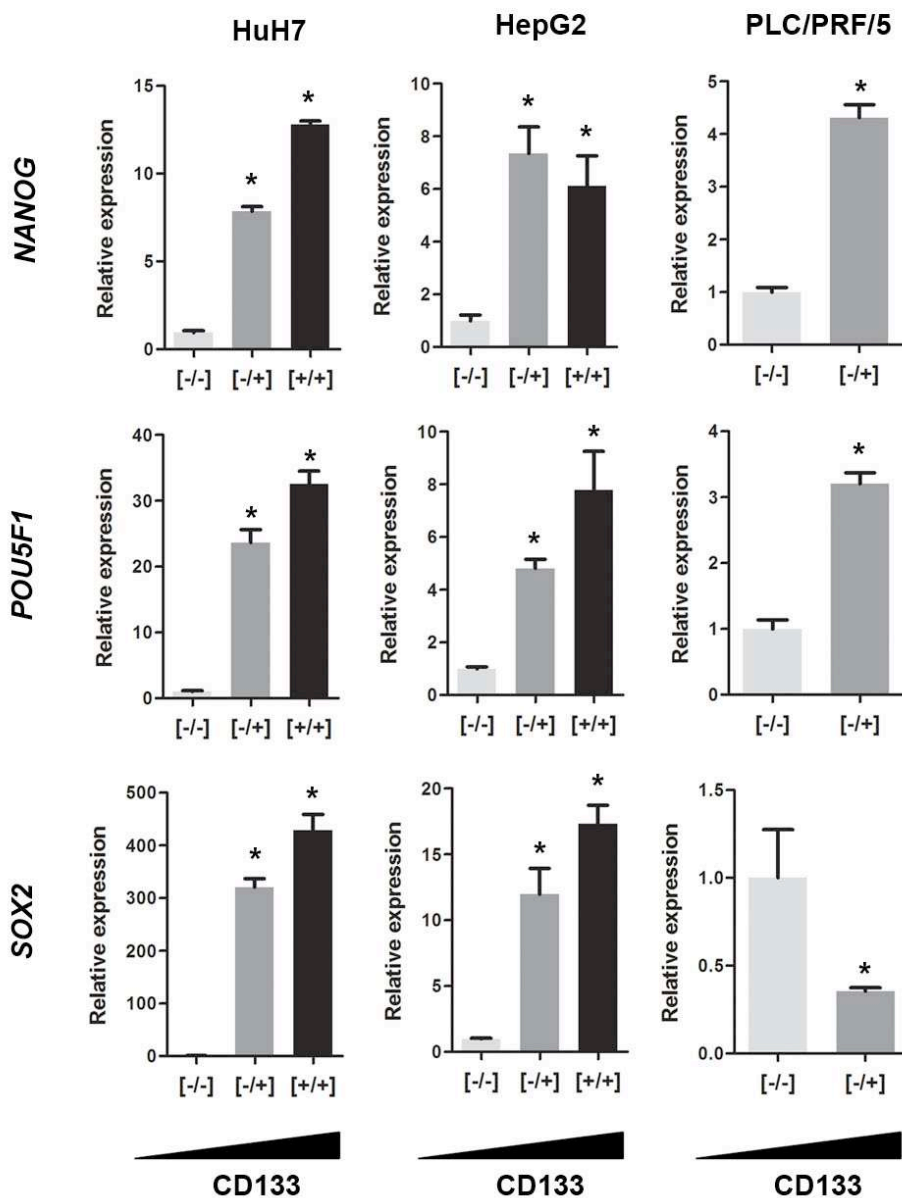
Huh7 CD133+ and CD133- cells were sorted by magnetic activated cell sorting (MACS) and 15,000 cells were plated in ultra-low attachment conditions (see Materials and Methods). The left panel shows representative pictures of spheres formed after 7 days of culture. The right panel shows the total number of spheres counted for each condition (standard deviations of technical triplicates are represented).

Many previous studies have reported that CD133+ cells in these cell lines present important features related to CSCs, such as clonogenicity, tumorigenesis, metastatic potential and chemoresistance. Although the relation of CD133 expression to stemness has already been established, our study demonstrated that Huh7 CD133 positive cells enriched with magnetic cell sorting (MACS) were also able to grow in non-attachment conditions (Figure 38).

In addition, we studied the mRNA expression of well-defined stemness transcription factors that have been reported as differentially expressed in some subpopulations of liver CSCs (Wang et al., 2013). Efficiency of CD133 enrichment by MACS was variable, potentially due to the different starting population for each cell line, and intensity of CD133 protein expression at the cell surface. In spite of this variability, we observed a significant overexpression of *NANOG* and *POU5F1* (*OCT4*) in all three cell lines studied, while *SOX2* was significantly overexpressed in two cell lines (HuH7 and HepG2) (Figure 39). In these two cell lines, mRNA obtained from populations enriched in CD133+ cells at intermediate levels, displayed the expected intermediate levels of mRNA expression (Figure 39).

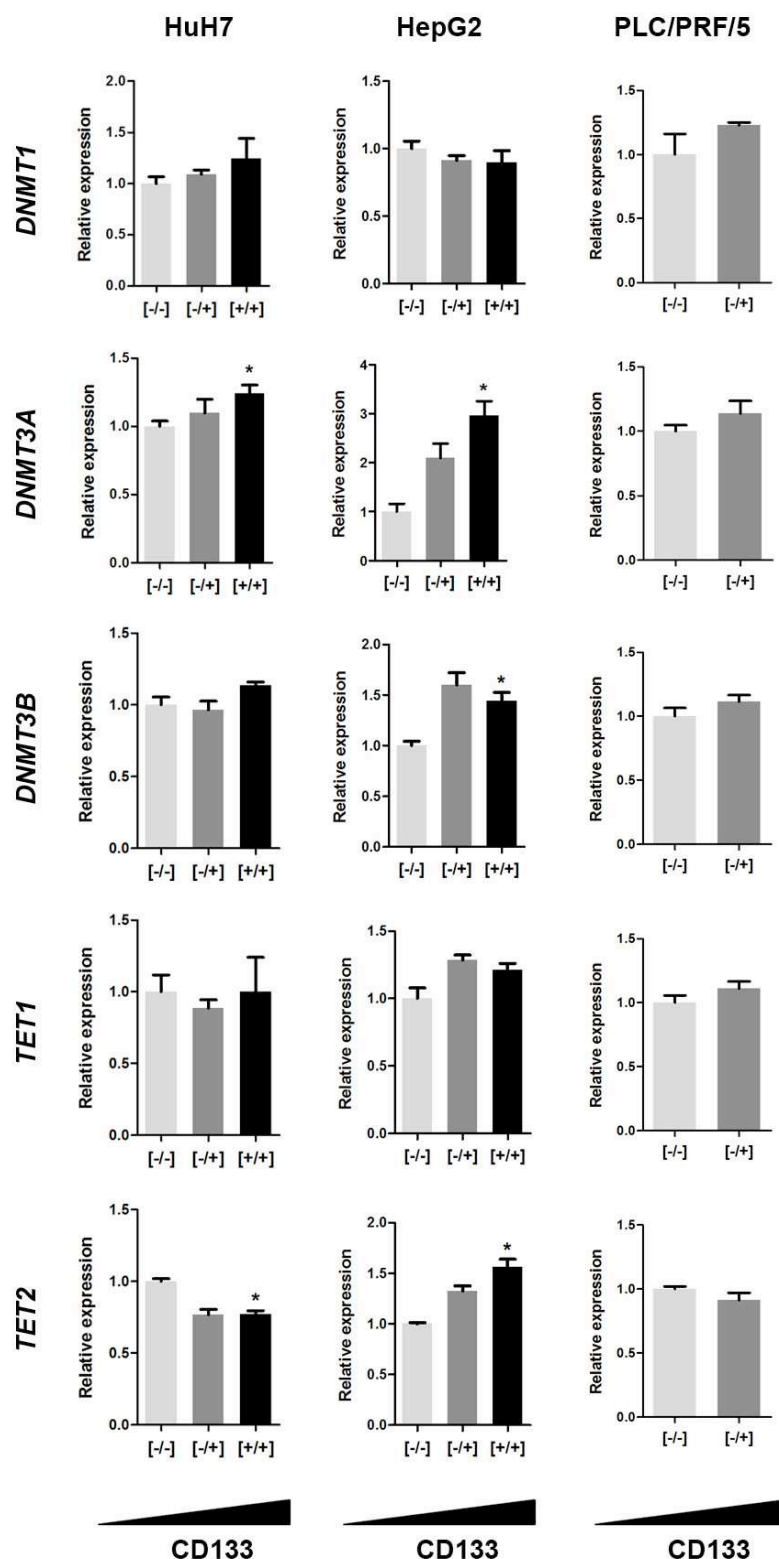
As a an initial step in exploring a potentially different methylation program in CD133+ liver cancer cells, we studied the expression of genes coding for relevant players of the DNA methylation machinery. This included genes involved in maintenance DNA methylation (*DNMT1*), *de novo* DNA methylation (*DNMT3A* and *DNMT3B*) and DNA demethylation (*TET1* and *TET2*). No significant differences were found for any of these genes in PLC/PRF/5 cells. However, *DNMT3A* was consistently overexpressed in both HuH7 and

HepG2 cells progressively enriched for CD133 (Figure 40). In addition, *DNMT3B* was overexpressed in HepG2 CD133-enriched cells, while *TET2* displayed opposite differential expression in HuH7 and HepG2 CD133-enriched cells (Figure 40).



**Figure 39. Stemness transcription factor expression in CD133+ cells.**

HuH7, HepG2 and PLCR/PRF/5 cells were sorted by MACS. RNA was extracted to study the expression of *NANOG*, *POUF5* and *SOX2* by qRT-PCR in subpopulations differentially enriched for CD133+ cells ([-/-] ; [-/+]; [+/+]). Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05.



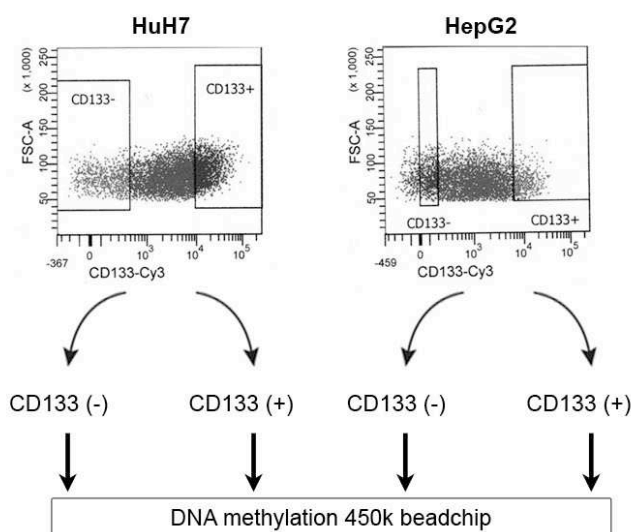
**Figure 40. Expression of the genes encoding the key enzymes involved in DNA methylation maintenance in CD133+ cells.**

Huh7, HepG2 and PLC/PRF/5 cells were sorted by MACS and expression of *DNMT1*, *DNMT3A*, *DNMT3B*, *TET1* and *TET2* were investigated by qRT-PCR in subpopulations differentially enriched for CD133+ cells ([-/-]; [-/+]; [+/+]). Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05.

Together, these data suggest that in at least two independent liver cancer cell lines (HuH7 and HepG2) CD133 marks a specific subpopulation of cells. In addition, the consistent overexpression of *de novo* DNA methylation genes (*DNMT3A* in both cell lines, and *DNMT3B* in HepG2) favors the idea of a potentially unique DNA methylation program. Therefore, we selected HuH7 and HepG2 cell lines for further analysis of DNA methylation in CD133+ cells.

## **II. A differential DNA methylome defines CD133- and CD133+ liver cancer cells**

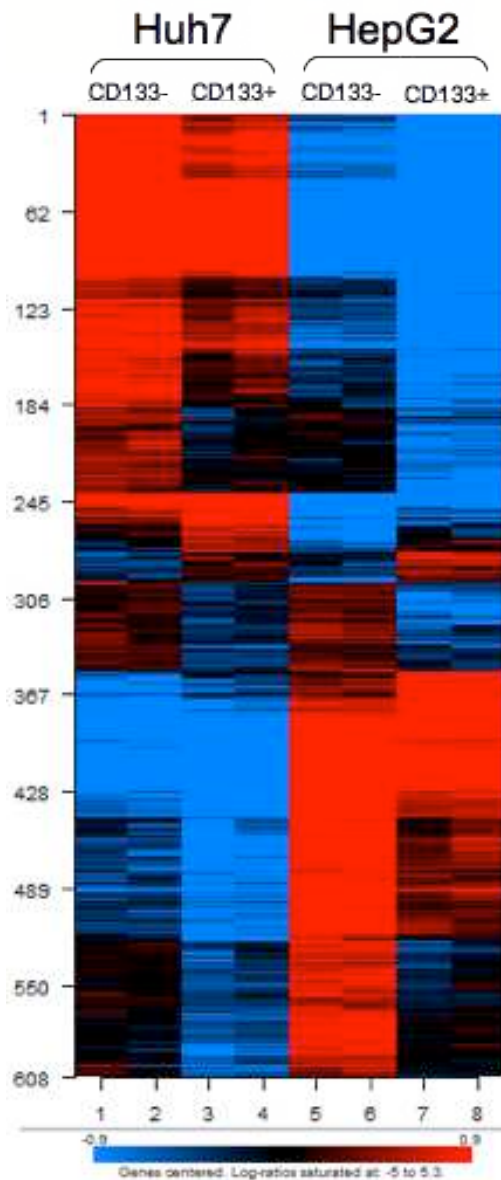
We have shown that CD133+ cells represent a functionally and phenotypically unique fraction of cells. They also display a differential expression of *de novo* DNMTs, and this may be reflected in a differential configuration of their DNA methylome. To study this possibility, we performed a genome-wide DNA methylome analysis in FACS-sorted CD133 negative and positive fractions from Huh7 and HepG2 cells (Figure 41).



**Figure 41. Experimental design for genome-wide DNA methylation study in CD133+ cells.**

Huh7 and HepG2 cells were sorted by FACS using CD133 antibody. DNA from CD133+ and CD133- cells was extracted, converted with bisulfite treatment and processed on the Illumina Infinium 450K bead array. For each cell line, biological duplicates for CD133+ and CD133- subpopulations were processed (see Materials and Methods).

DNA isolated from these fractions was interrogated with the Illumina Infinium 450K bead array, which allows interrogation of more than 450,000 CpG sites, spanning all RefSeq genes, CpG islands, and non-CpG sites (Bibikova et al., 2011). Output data were processed using Illumina GenomeStudio for quality control and data export and the BRB-ArrayTools

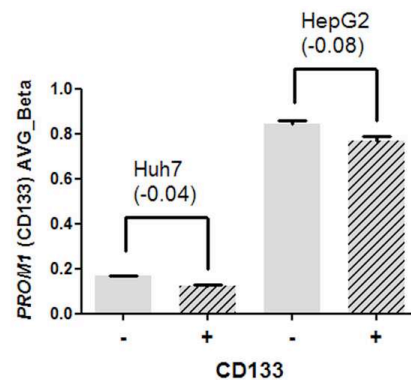


**Figure 42. A differential methylome distinguishes CD133+ and CD133- cells.**

Unsupervised clustering of CD133+ and CD133- cells using the significant CpG ( $n=608$ ) differentially methylated ( $p<0,001$ ; average  $\delta\beta >5\%$ ). Methylation level is expressed in log scale, with higher methylation represented in red and lower methylation shown in blue.

software (see Materials and Methods). In unsupervised analyses, parental cell line was the main factor defining DNA methylation variation (Figure 42). Therefore, our main analysis compared CD133- vs CD133+ fractions accounting for cell of origin (see Materials and Methods). The class comparison analysis resulted in 608 probes differentially methylated at significant  $p$  value ( $p<0.001$ ), although with relatively high FDRs ( $FDR=0.58$ ), probably due to sample variability and cell line differences. Moreover these CpG sites were all selected for an average  $\delta\beta >5\%$ . Supporting the quality of the dataset was the finding of one CpG

site within the *CD133* (*PROM1*) locus, among this list of differentially methylated sites. This CpG site was hypomethylated in CD133+ subpopulations from both cell lines, by 4.4% and 8% in Huh7 and HepG2 cells, respectively (Figure 43).

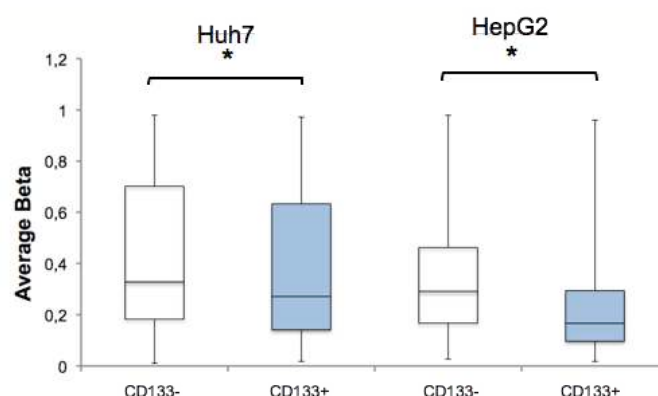


**Figure 43. Genome-wide DNA methylation array revealed hypomethylation for *PROM1* in CD133+ cells.**

Average\_Beta (AVG\_Beta) values obtained from the bead array assay were plotted for on significant CpG site within the CD133 (*PROM1*) promoter. The difference in methylation between CD133- and CD133+ cells (delta\_Beta) is indicated for each cell line.

The 608 differentially methylated genes correspond to 394 RefSeq genes, and represent those CpG sites significantly hypo or hypermethylated in CD133+ cells in both cell lines, relative to their negative counterparts. Most of these probes (n=511, 84%) were hypomethylated in CD133+ cells, while 98 (16%) were hypermethylated (Figure 44).

POU5F1 was displayed. In addition, transcription factors belonging to the STAT and SMAD families were again observed.

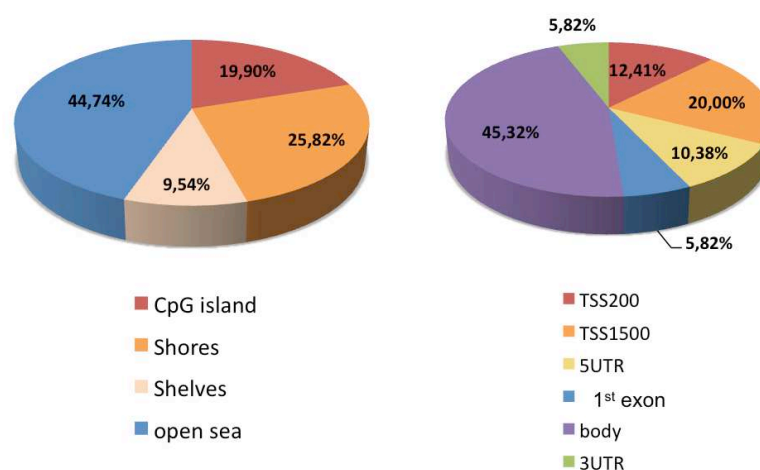


**Figure 44. CD133+ cells are globally hypomethylated compared to their negative counterpart.**

Median methylation (and distribution) for all the 608 differentially methylated loci ( $P < 0.001$ , average delta beta  $> 5\%$ ) distinguishing CD133- vs CD133+ cells in both cell lines. (\*) indicates P value  $< 0.05$  for comparison between CD133+ and CD133- cells in each cell line separately.

Interestingly, an important proportion of differentially methylated loci (44%) were not related to CpG island regions (“open sea” probes in Figure 45). For those probes matching annotated genes, we did not observe any significant overrepresentation of differentially methylated loci in one specific gene-related region (Figure 45). We next carried out pathway analyses and found an enrichment in pathways previously associated with CSC activity, such as Jak-STAT, Wnt and Akt. In addition, there was a significant enrichment for inflammatory pathways, such as NFkB, p38, TNF, and TGF-β signaling pathways (Table S2). Finally transcription factor geneset analysis was realized. As the analysis with the two cell lines combined bring out some transcription factor activated by inflammatory stimulus such as NF-kB, SMAD3 and members of the STAT family (Table S2). Interestingly, when we repeated this analysis for each cell line separately, the stem-cell related transcription factor

In summary, CD133+ liver cancer cells display a distinct DNA methylome compared to their negative counterpart. In spite of the cell line specific profiles, our results revealed a common CD133+ methylome signature, which includes the *PROM1* gene itself. The methylome of CD133+ cells was characterized by a global reduction in DNA methylation, with an overrepresentation of intergenic CpG sites. For those differentially methylated sites related to annotated genes, there was an association with CSC- and inflammation-related pathways. These findings suggest that DNA methylation may have an important contribution to defining the phenotype and functional properties of this cell subpopulation.

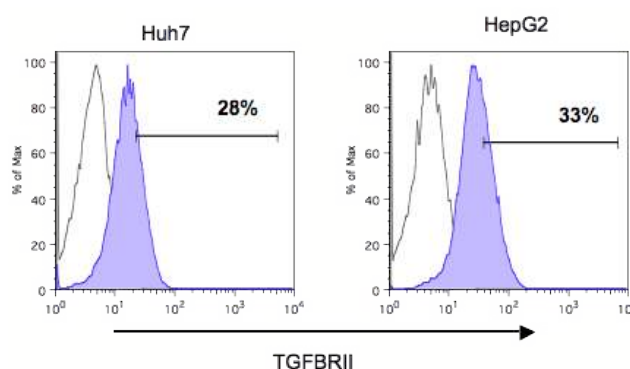


**Figure 45. Regional distribution of the differentially methylated CpG loci in CD133+ cells.**

Significant loci were distributed according to CpG island relationship as Island, shore, shelf, and open sea and are represented in the left pie chart for all the 608 significant loci. The right pie chart represents the distribution of the significant loci in relation to annotated genes (within 200 or 1500 bp from the TSS, 1<sup>st</sup> exon, 3' or 5' UTRs, and gene body).

### III. TGF- $\beta$ , but not IL-6, induces CD133 expression in a stable fashion

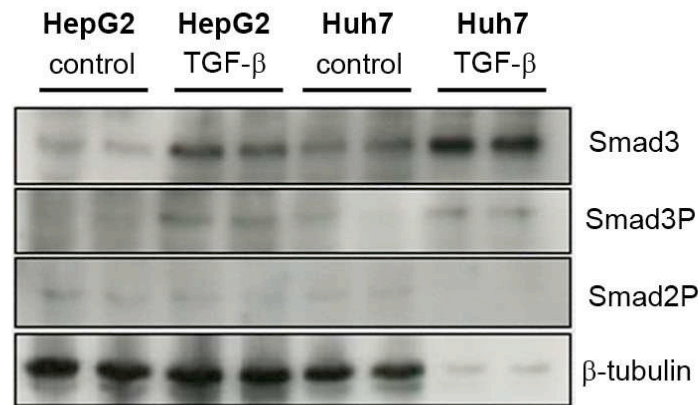
The enrichment in inflammatory pathway observed during our analysis of CD133+ cells DNA methylation signature suggests that these cells might be differentially sensitive to inflammatory stimulus. CD133+ CSCs in other cancers, such as glioblastoma, lung cancer and breast, were described as being maintained by TGF- $\beta$  (Mani et al., 2008a; Peñuelas et al., 2009; Pirozzi et al., 2011). In liver cancer, it has been reported that TGF- $\beta$  exposure increases the percentage of CD133+ cells in the HuH7 cell line (You et al., 2010). In addition, the inflammatory pathways displayed by CD133+ methylome included the TGF- $\beta$  pathway. Therefore we next aimed to validate and extend these observations by investigating the impact of TGF- $\beta$  exposure on CD133+ subpopulation in HepG2 cells. Importantly, both HuH7 and HepG2 cells, express the receptor for TGF- $\beta$  (TGFBRII) at similar levels (Figure 46), and respond to TGF- $\beta$  by phosphorylating the receptor-dependent Smad, SMAD2 (Figure 47).



**Figure 46. Huh7 and HepG2 cell lines expressed similar levels of TGFBR II.**

Huh7 and HepG2 cell lines were immunostained with TGFBR II antibody and analysed by FACS. The dot plots displayed are one representative example of triplicate experiment.

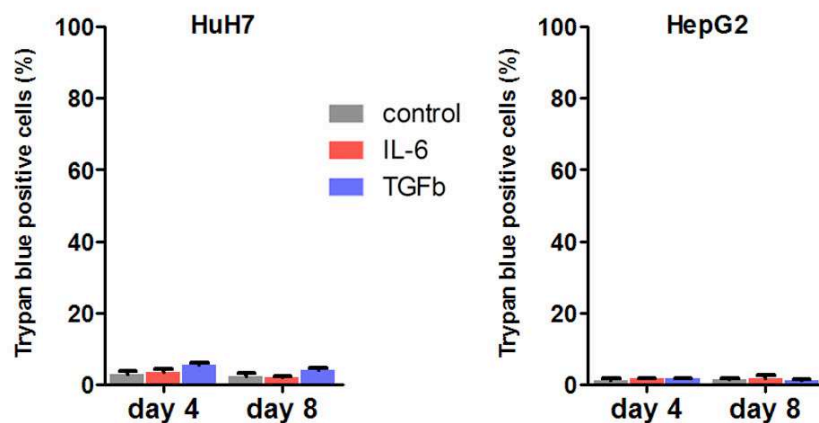




**Figure 47. Activation of SMAD3 after TGF-β exposure.**

Huh7 and HepG2 cells were exposed to TGF-β and proteins were extracted and separated on a SDS-PAGE. Antibodies directed against SMAD3, phosphorylated SMAD3 (Smad3P) and phosphorylated SMAD2 (Smad2P) were used to observe TGF-β pathway activation. β-tubulin was used for loading control.

We reproduced the preliminary findings of You et al, also using an additional cytokine with relevance in HCC, the pro-inflammatory interleukin 6 (IL-6). To this end, we selected concentrations of both cytokines that did not have an effect on cell viability (Figure 48); moreover, these concentrations were already used in previous studies on Huh7 cell line (Matak et al., 2009; You et al., 2010b).

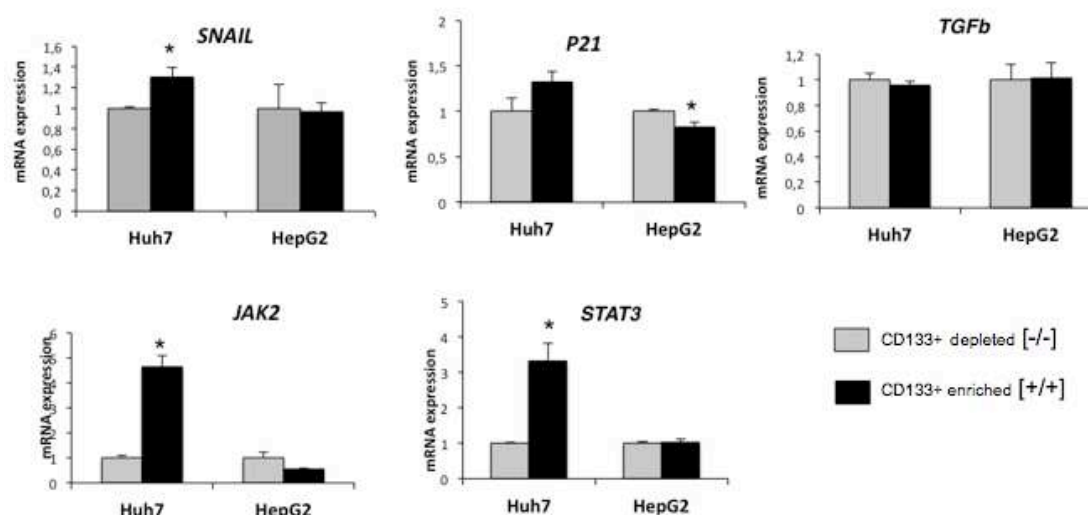


**Figure 48. IL-6 and TGF-β do not alter cell viability of HCC cell lines.**

Huh7 and HepG2 cells were treated with IL-6 or TGF-β (see Materials and Methods). Cell's viability was assessed by trypan blue staining. Mean (+ standard deviation) of three independent experiments are represented.

As genes involved in both SMAD and STAT3 signaling pathways were found to be differentially methylated in CD133+ cells, we first checked if one of these pathways was already activated in CD133+ cells by analyzing the expression of known target genes (Figure 49). Despite the significant increase of *STAT3* and *JAK2* in Huh7, we did not detect any

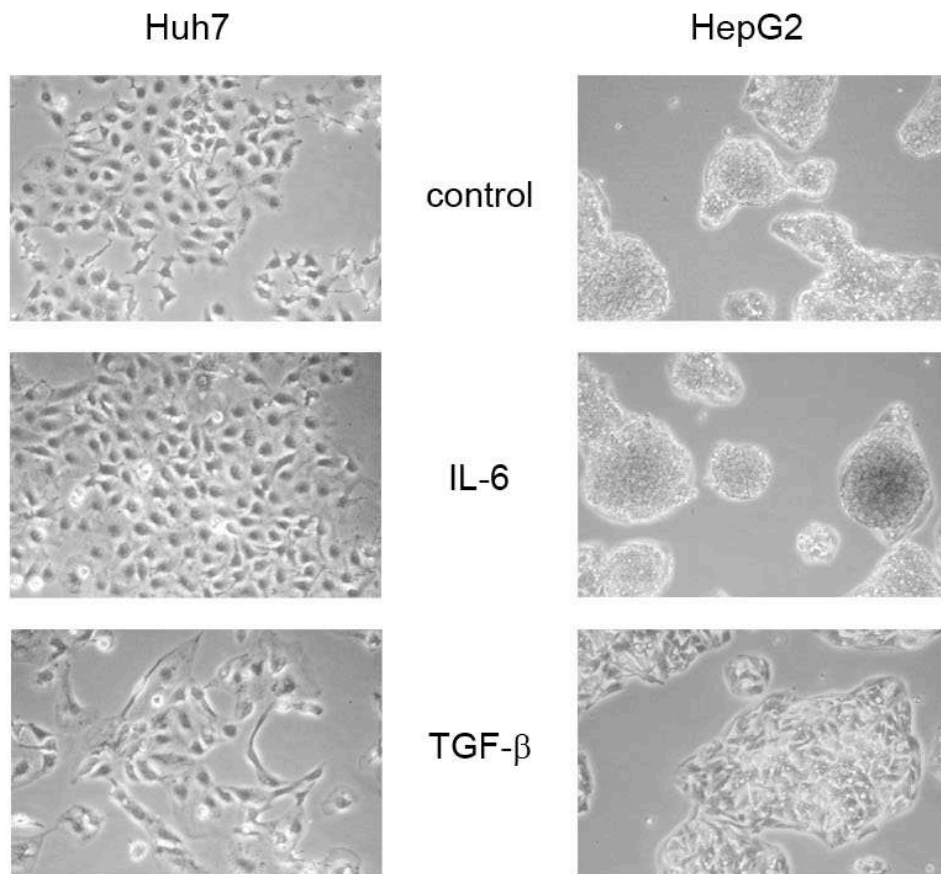
consistent increase of expression for IL-6 target genes in HepG2 or for TGF- $\beta$  target genes or in CD133+ cells (Figure 49). This result indicates, that in the absence of TGF- $\beta$  the medium CD133+ cells do not present any activation of this signaling pathway. As an increased of IL-6 target genes was only observed in CD133+ Huh7 and not in HepG2, it suggests that this basal activation is not a intrinsic property of CD133+ cells but might be rather associated to a specific cell line characteristic.



**Figure 49. TGF- $\beta$  and IL-6 signaling pathways target genes expression in CD133+ cells.**

Huh7 and HepG2 cells were sorted by MACS and expression of *SNAIL*, *P21*, *TGF- $\beta$*  (target genes of the TGF- $\beta$  signalling pathway) *JAK2* and *STAT3* (target genes of the IL-6 signaling pathway) was investigated by qRT-PCR in subpopulations differentially enriched for CD133+ cells ([-/-] ; [+/+]). Expression was normalized to housekeeping gene *GAPDH*. (\*) indicates P value < 0.05.

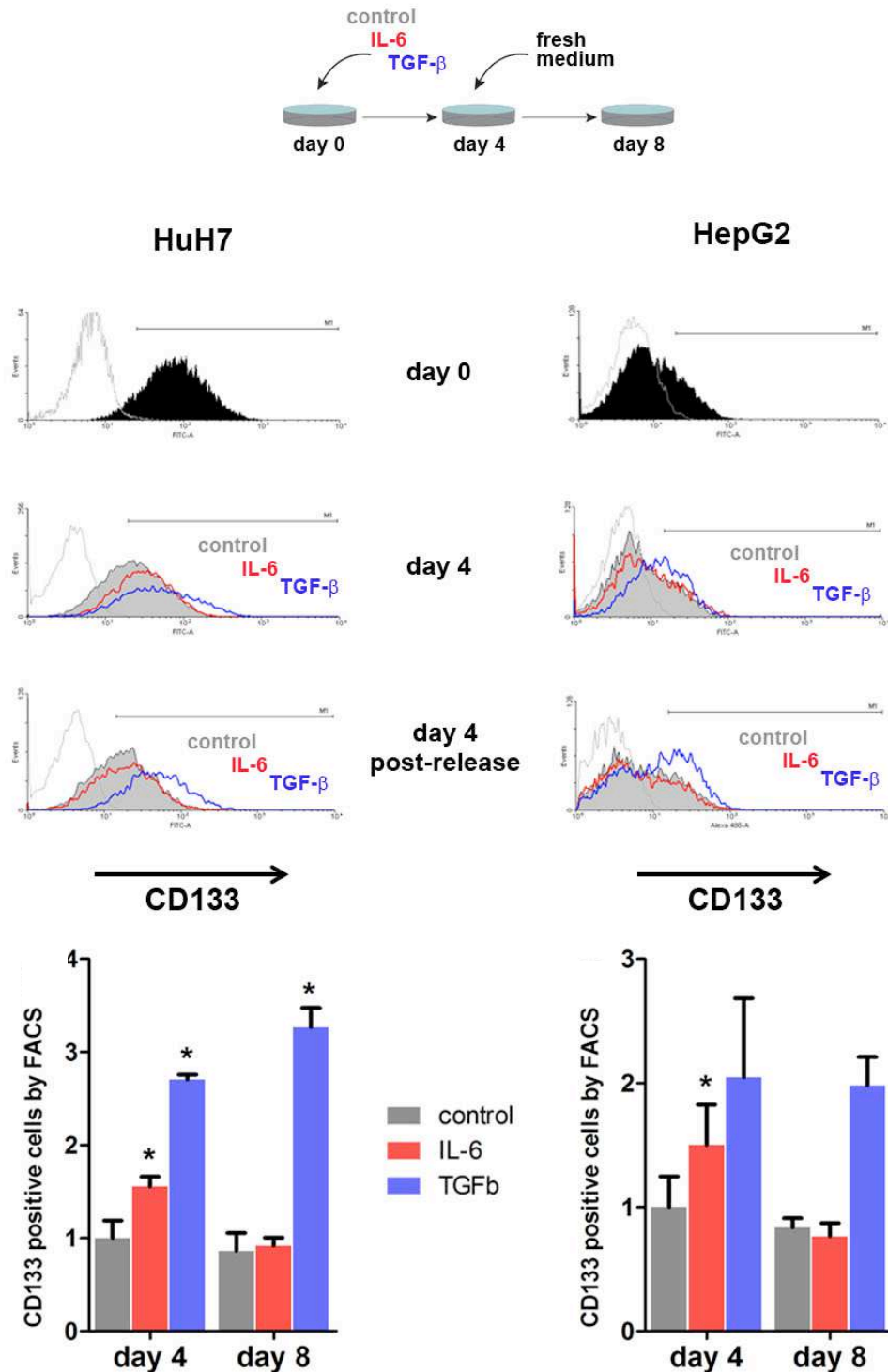
We then exposed cells to IL-6 or TGF- $\beta$  (10 ng/ml) and first check for any morphological changes that could reflect deeper changes in the phenotype. After 4 days of treatment, we didn't observe any morphological change in IL-6 treated cells whereas TGF- $\beta$  treated cells harboured a distinct phenotype compared to non-treated cells. After TGF- $\beta$  exposure, cells became more elongated and were more spread in the culture dishes (Figure 50). These morphological changes strongly remind EMT-associated morphology and suggest that TGF- $\beta$  induced some transformation in our two cell lines.



**Figure 50. TGF- $\beta$  exposure induces morphological changes in HCC cell lines.**

Representative phase contrast images of Huh7 and HepG2 cells left untreated or exposed to IL-6 or TGF- $\beta$  during 4 days.

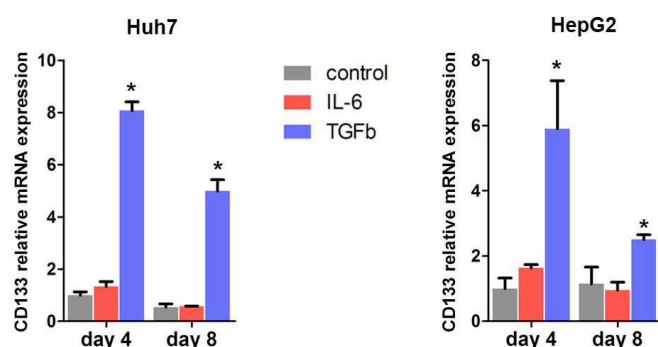
As for CD133 expression, as expected, TGF- $\beta$  exposure during 4 days induced an almost three-fold and two-fold increase in the percentage of CD133+ cells in HuH7 and HepG2 cells, respectively (Figure 51). Interestingly, IL-6 treatment also induced a significant increase in CD133 positivity in both cell lines, although the increase was comparatively mild (about 50%) (Figure 51).



**Figure 51. TGF-β induces a persistent increase of CD133+ cells.**

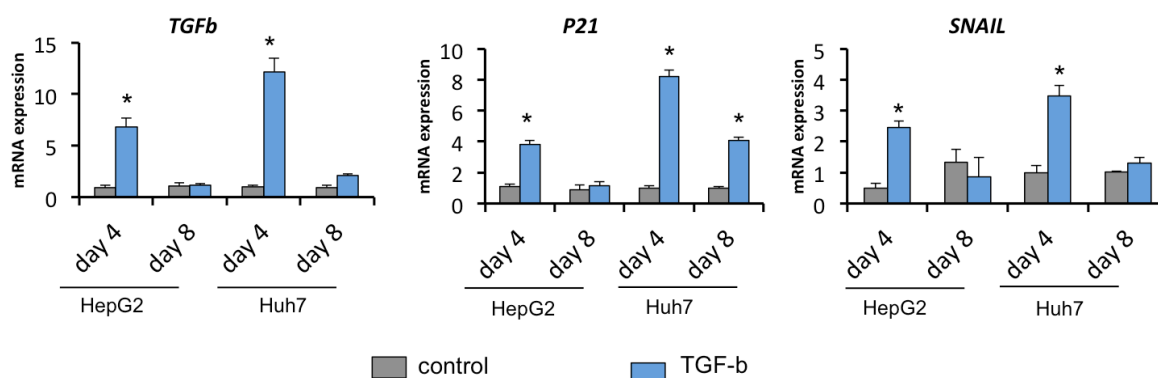
Experimental design is indicated in the upper panel. Huh7 and HepG2 cells were grown in control culture conditions (depicted in gray text and lines), or exposed to 10 ng/ml IL-6 (red) or 10 ng/ml TGF-β (blue) during 4 days. Cells plated in parallel, had their medium replaced by control culture medium and were left in culture for an additional 4 days. FACS expression of surface CD133 protein is shown for day 0, day 4, and day 8 (4 days treatment + 4 days post-release) for all conditions. Histograms are shown for one representative replicate in the middle panel. Fold change compared to the control are shown for three biological replicates in the lower panel barplots. (\*) indicates P value < 0.05.

In order to further analyze the effect of TGF- $\beta$  and IL-6, we next analyzed the persistence of the effect in CD133 expression induced by both cytokines. Indeed the actions of cytokines cover a large number of biological effects ranging from transient proliferation to permanent cell fate conversion. The difference in cytokine action duration relies on adapted mechanisms to regulate gene expression: from transient recruitment of transcription factor to a persistent epigenome reconfiguration. The analysis of the duration of TGF- $\beta$  and IL-6 effect can thus provide a hint about the mechanisms underlying their effect on CD133<sup>+</sup> cells. To this end, we treated both cell lines as previously (TGF- $\beta$  or IL-6 treatment for 4 days). After 4 days, cell culture medium was replaced by standard medium, and cells were left in culture for an additional 4 days. Cells were then collected and screened for CD133 expression using FACS. Of note, only cells treated with TGF- $\beta$  showed a persistent increase in the percentage of CD133<sup>+</sup> cells, of similar magnitude to the increase observed at day 4 (Figure 51). Importantly, only TGF- $\beta$  exposure was able to induce a significant increase in the expression of CD133 in both cell lines at the transcriptional level (8 and 6 fold increase for Huh7 and HepG2, respectively) (Figure 52). Interestingly known target genes of the TGF- $\beta$  signaling pathway such as *TGF $\beta$* , *P21* and *SNAIL* were found upregulated only after 4 days and no persistence in this upregulation was observed after the release of the cytokine (Figure 53). This observation indicates that the persistence of TGF- $\beta$ 's effect on CD133<sup>+</sup> cells is not due to a positive feedback of the pathway that would maintain TGF- $\beta$  intracellular signal activated even in the absence of the cytokine. Altogether these findings suggest that TGF- $\beta$  is able to stably induce CD133 expression (in contrast to the milder and transient effect of IL-6), an observation consistent with epigenetically-induced phenotype persistence.



**Figure 52. CD133 mRNA expression after TGF- $\beta$  exposure.**

Experimental design was the same as described in Figure 50. CD133 mRNA level is shown for day 4, and day 8 (4 days treatment + 4 days post-release). Mean  $\pm$  standard deviation is shown for three biological replicates. Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05.

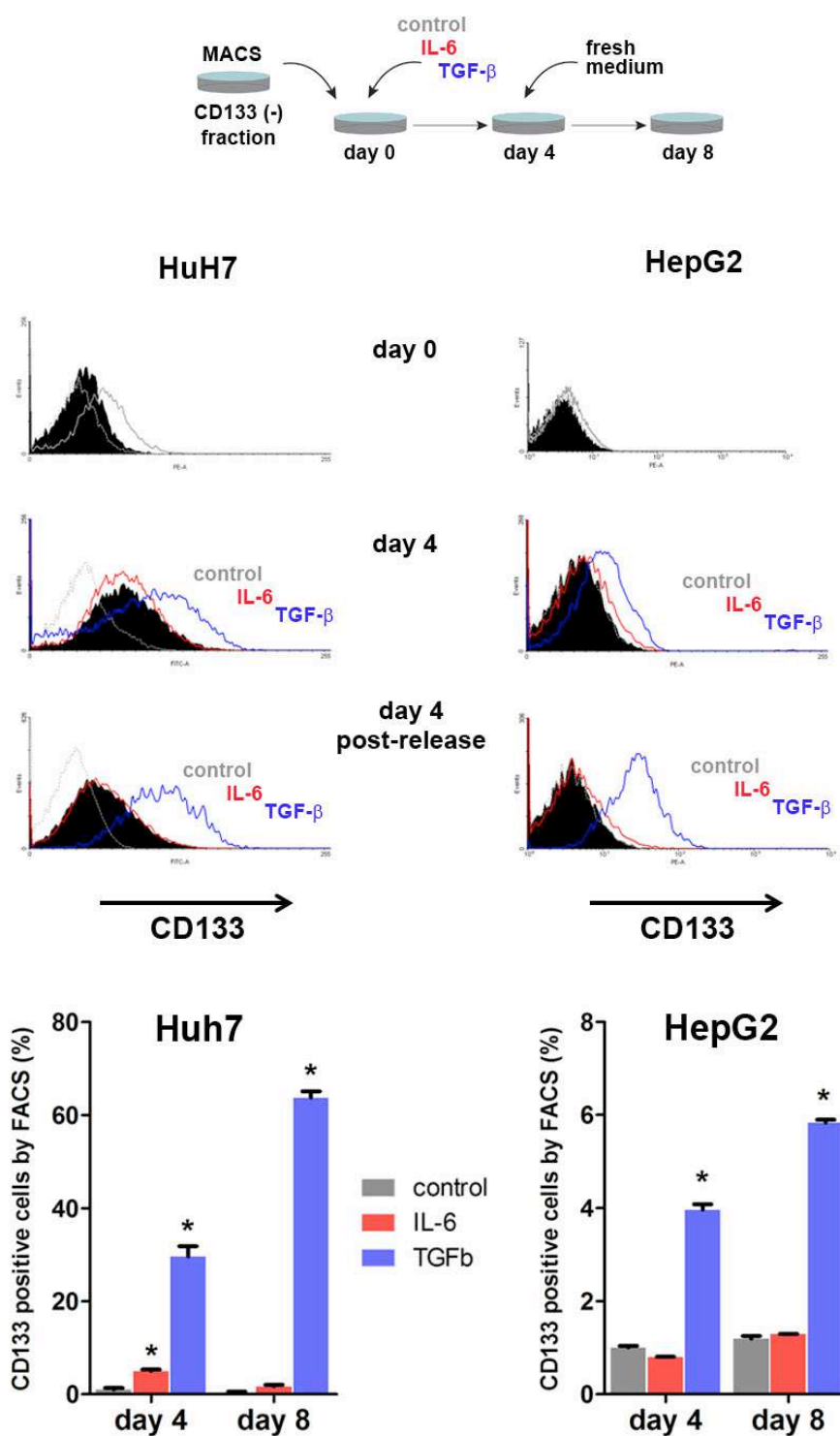


**Figure 53. Expression of TGF- $\beta$  signaling pathway target genes after TGF- $\beta$  exposure.**

Experimental design was the same as described in Figure 50. *TGF- $\beta$* , *P21* and *SNAIL* mRNA levels are shown for day 4, and day 8 (4 days treatment + 4 days post-release). Mean (+standart deviation) is shown for three biological replicates. Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05.

#### **IV. De novo induction of CD133+ cells by TGF- $\beta$ is associated to an increased expression of DNMT3 genes.**

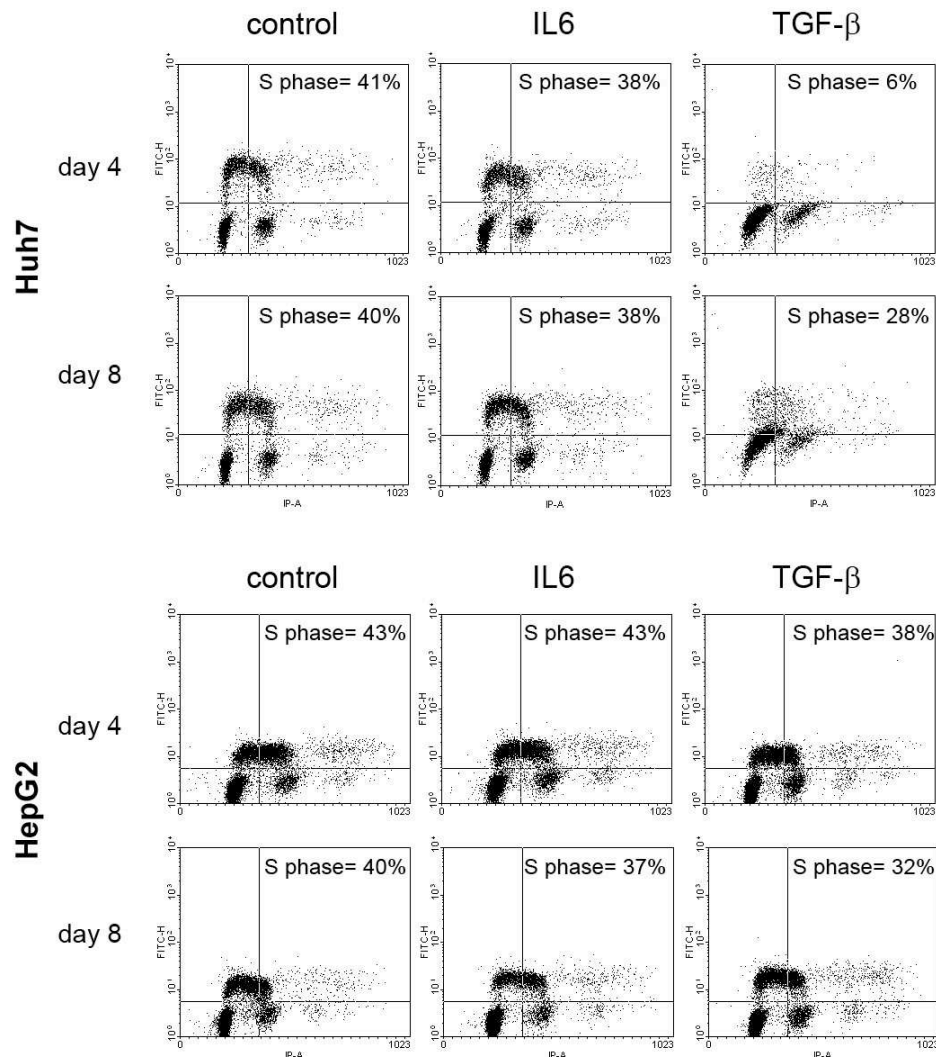
The increase in CD133 positivity can be due to a switch in the expression of CD133, or an increased rate of growth induced by TGF- $\beta$  specifically in the smaller CD133+ fraction of cells. To distinguish between these two possibilities, we repeated the previous experiment in cells negative for CD133 expression, selected by negative enrichment with MACS (see Methods). In both cell lines, TGF- $\beta$  was able to significantly induce a population of CD133+ cells, evident after 4 days of treatment (Figure 54). Also in this case, we replaced the medium after 4 days, and let the cells grow in the absence of cytokines for additional 4 days. After these additional 4 days, the increase in CD133 positive fraction for both cell lines was even higher, relative to the one observed at day 4 (Figure 54). Importantly, although there was a spontaneous induction of a CD133+ fraction in HuH7 cells (from 0 to 20% after 4 days), this percentage did not significantly change at day 8, and is similar to that found in untreated HuH7 cells at basal conditions. This indicates a potential *de novo* balance between the CD133 negative and positive fractions in this cell line. In contrast, the expression of CD133 remained close to zero in HepG2 control cells and only increased after TGF- $\beta$  exposure. This finding indicates that TGF- $\beta$  is able to induce the expression of CD133 surface protein, and not an increased proliferation of CD133+ cells. To actually support this hypothesis, we used BrDU to assess the effect of TGF- $\beta$  on cell cycle. In Huh7 cells we observed an expected lower rate of proliferation of cells treated with TGF- $\beta$  while in HepG2 TGF- $\beta$  has no effect on cell cycle (Figure 55). Thus TGF- $\beta$ 's effect on CD133+ cells seems to not involve any increase of cell



**Figure 54. TGF-β can induce transdifferentiation of CD133- into CD133+ cells.**

Experiment described in Figure 50 was repeated in MACS-sorted CD133- cells, as depicted in the upper panel. Levels of CD133 expression were close to 0%, as shown in the upper histograms for both, Huh7 and HepG2 cells. Mean (+ SD) from three replicates is shown in the lower panels. (\*) indicates P value < 0.05 relative to control conditions.

proliferation. Similar to our previous experiment, the transient effect of IL-6 on CD133+ cells was confirmed, as this cytokine was only able to induce a 50% increase after 4 days in Huh7 but that is not persistent after the replacement with fresh medium (Figure 54). Moreover IL-6 was not able to induce a population of CD133+ cells in HepG2.



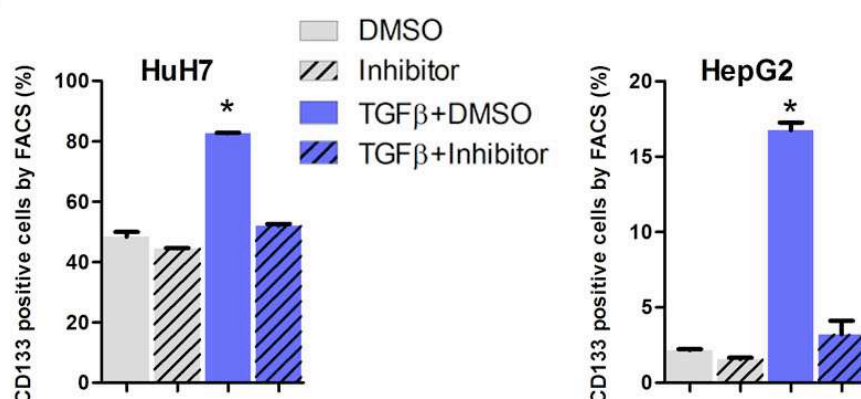
**Figure 55. TGF-β's effects on cell cycle. For Huh7 and HepG2 cells.**

Huh7 and HepG2 cells were grown in control culture conditions (depicted in gray text and lines), or exposed to 10 ng/ml IL-6 (red) or 10 ng/ml TGF-β (blue) during 4 days. Cells plated in parallel, had their medium replaced by control culture medium and left in culture for additional 4 days. TGF-β's effects on cell cycle was assessed by BrDU staining (see Materials and Methods) at day 4, and day 8 (4 days treatment + 4 days post-release) for all conditions. Histograms are shown for one representative replicate.

TGF-β is a member of a large family of pleiotropic cytokines that signal through a receptor complex comprising a diversity of type I and a type II serine/threonine kinases. The recombinant TGF-β1 used in our assays is expected to bind the activin receptor-like kinase



ALK-5 (the TGF- $\beta$  type I receptor) (Callahan et al., 2002). To rule out unspecific effects of this treatment, we used the small molecule inhibitor SB-431542, which targets ALK5 and ALK5-related type I receptors, with no effect on other family members that, for example, recognize bone morphogenetic proteins (BMPs) (Inman et al., 2002). By using this specific inhibitor of TGF- $\beta$  pathway, we were able to completely rescue the effect of TGF- $\beta$  in inducing CD133 expression in both cell lines (Figure 56). Therefore, the ability to induce CD133<sup>+</sup> cells is specific and fully dependent on TGF- $\beta$  type I receptor signaling in both, Huh7 and HepG2 cells (Figure 56).



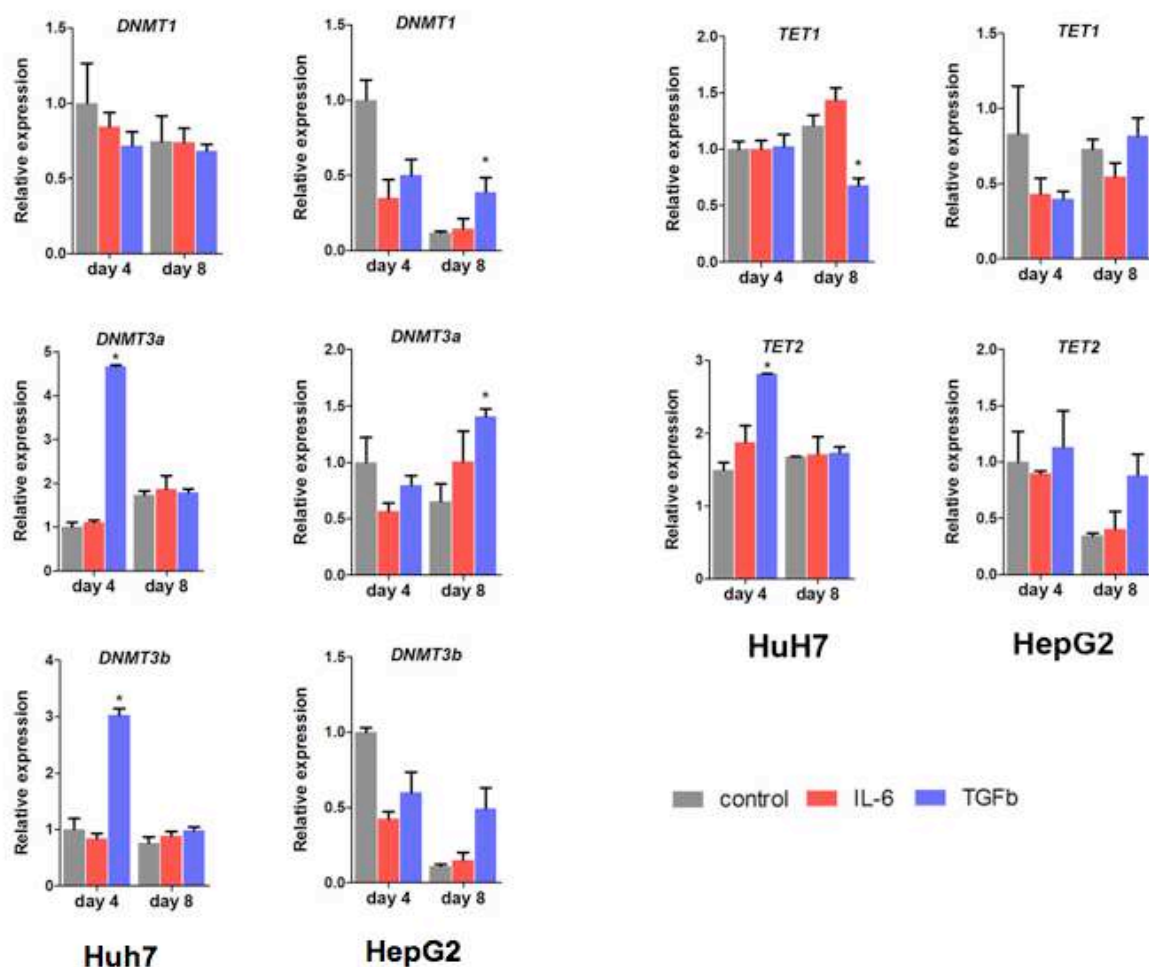
**Figure 56. Specificity of TGF- $\beta$ 's effect on CD133<sup>+</sup> population.**

Huh7 and HepG2 cells were treated for 4 days with TGF- $\beta$  + inhibitor (SB432542, specific inhibitor of the TGFBR1 receptor) or with TGF- $\beta$  + vehicle (DMSO). CD133 expression was observed by FACS. Mean (+ standard deviation) is shown for three biological replicates. (\*) indicates P value < 0.05 relative to cells treated with vehicle. .

After having shown that TGF- $\beta$  is able to induce a *de novo* fraction of CD133<sup>+</sup> cells, we asked whether this effect correlated with a differential expression of DNA methylation players, as we have shown that CD133<sup>+</sup> cells overexpress *DNMT3* genes in basal culture conditions (Figure 40). All DNMTs and *TET2* displayed an increase mRNA expression in at least one of the two cell lines, while *TET1* was underexpressed after 4 days of release from TGF- $\beta$  exposure (Figure 57). As shown for the basal CD133-expressing cells, the most consistent finding was the overexpression of *DNMT3A* in both cell lines after TGF- $\beta$  treatment. Of note, in none of the conditions of study IL-6 exposure was able to induce statistically significant changes at the mRNA expression level (Figure 57).

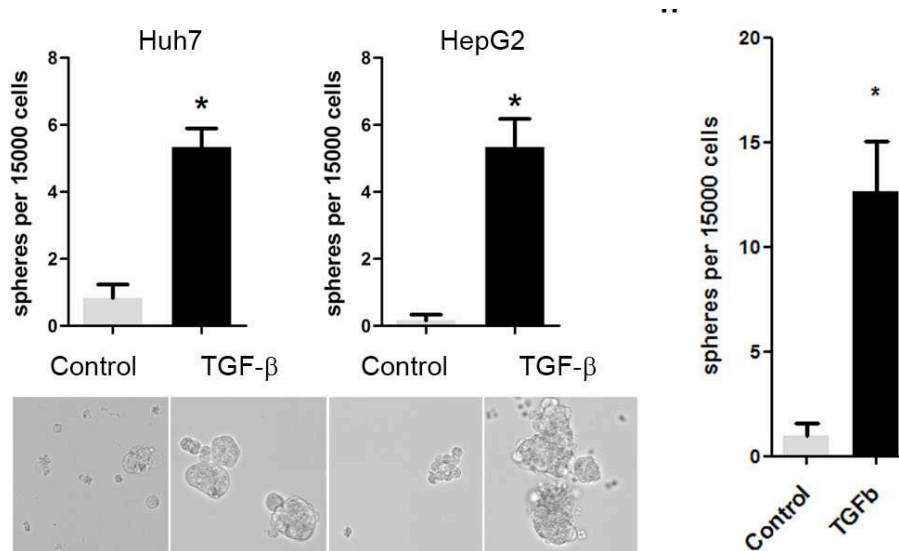
Combined, these data shows the ability of TGF- $\beta$  (in contrast to IL-6) to induce a stable *de novo* fraction of CD133-expressing cells in two independent liver cancer cell lines. This induction correlates with a functional characteristic of basal CD133<sup>+</sup> cells, which is the increased ability to grow under non-attachment cell culture conditions (Figure 58).

Furthermore in Huh7 this functional characteristic was maintained 4 days after the end of treatment (Figure 58). In addition, the differential expression of *de novo* DNMTs and the morphology changes induced by TGF- $\beta$  indicates that the expression of CD133 may be a marker of a more general expression program that defines this cell subpopulation.



**Figure 57. DNMT and TET expression is modulated by TGF- $\beta$ .**

Huh7 (left panels) and HepG2 (right panels) cells were treated as in described in Figure 50, RNA was extracted and qRT-PCR was performed for genes involved in DNA methylation or demethylation. Expression was normalized to housekeeping gene. (\*) indicates P value < 0.05 relative to non-treated cells at the corresponding time point.

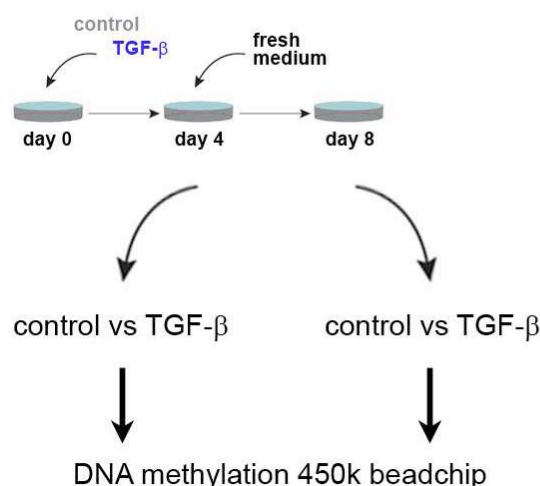


**Figure 58. TGF- $\beta$  induced CD133+ cells are able to grow on low attachment conditions.**

Left panel: Huh7 and HepG2 cells were treated with TGF- $\beta$  for 4 days, cells were then collected and plated (at same density) in low attachment plates (see Materials and Methodes). Spheres were counted after 5 days of culture. Right Panel: Huh7 cells were let 4 additionnal days in culture with fresh medium without TGF- $\beta$ . After those 4 additional days, TGF- $\beta$  induced CD133+ cells were still able to form spheres in low attachment conditions. (\*) indicates P value <0.05.

## **V. Transdifferentiation to CD133+ cells correlates with a methylome reconfiguration**

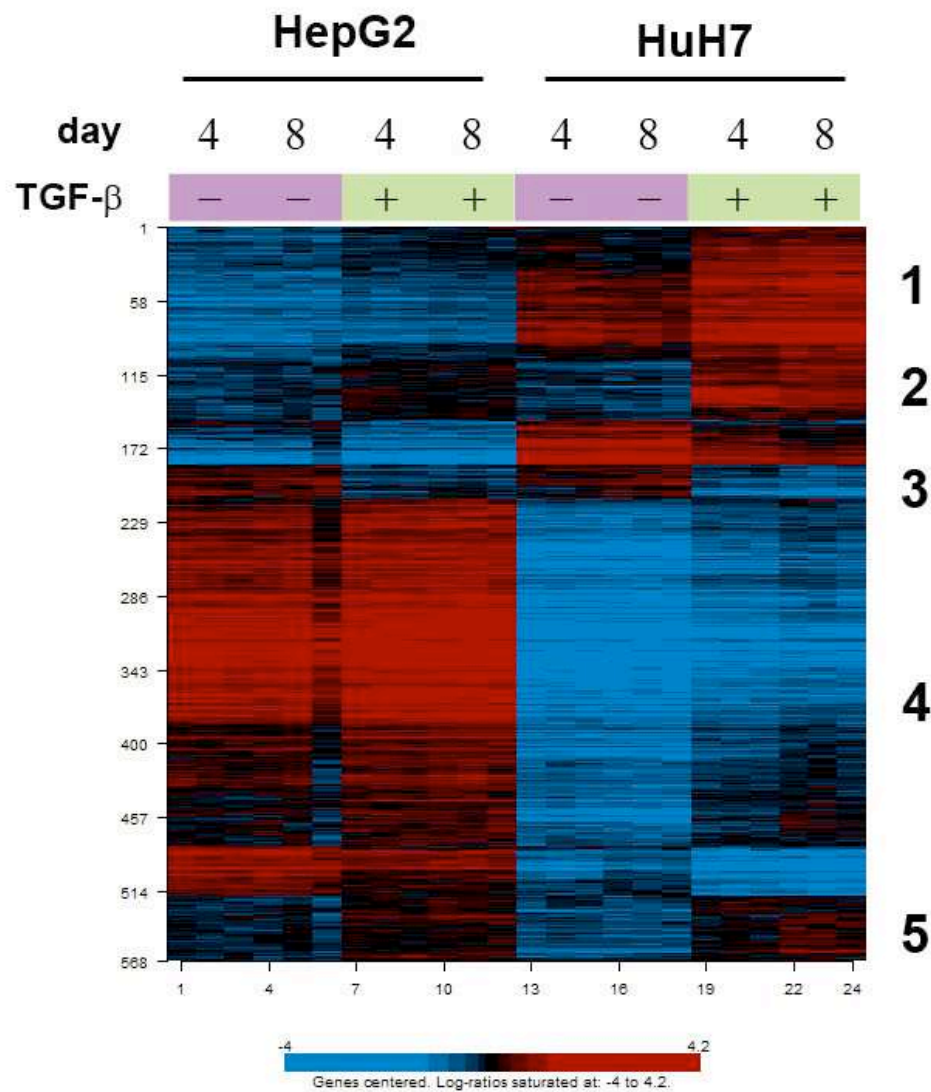
Having shown that CD133+ cells display a unique DNA methylome, and that TGF- $\beta$  is able to induce a *de novo* CD133+ fraction of cells, we further examined the DNA methylome changes induced by TGF- $\beta$  exposure. To this end, we used the same Infinium 450K platform to interrogate DNA methylation changes induced by 4 days of TGF- $\beta$  exposure in both, Huh7 and HepG2 cells (Figure 59). In addition, to define the epigenetic persistence of TGF- $\beta$  effects, we included the DNA from cells released 4 days into normal cell culture medium after the TGF- $\beta$  treatment.



**Figure 59. Experimental design for genome-wide DNA methylation study in TGF- $\beta$  exposed cells.**

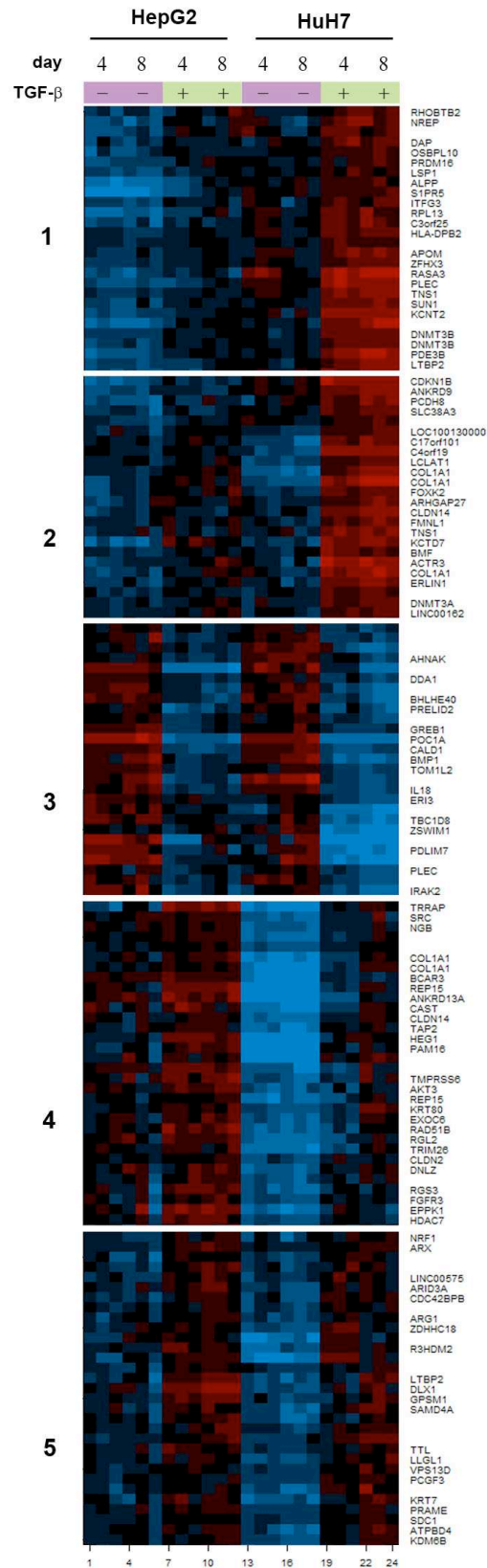
Huh7 and HepG2 cells were treated as described in the upper panel. DNA was extracted, converted with bisulfite treatment, and processed on the Illumina Infinium 450K bead array. For each cell lines, biological triplicates for each conditions were processed (see Materials and Methods).

Our analysis showed that the methylome of HuH7 and HepG2 cells are clearly distinguishable, independently of the experimental conditions (Figure 60), consistent with the CD133 DNA methylation profiling described above (Figure 42). However, in addition to cell type-specific changes we observe striking changes induced by TGF- $\beta$  in a cell type-independent fashion. To define a TGF- $\beta$ -induced DNA methylation signature, we focused on those loci that were significantly hypo or hypermethylated in both cell lines. In addition, we were interested in those changes that were persistent through cell division and stable in the absence of TGF- $\beta$ . Therefore, we selected significant loci (FDR<0.05) that were differentially methylated at both, 4 days of treatment and 4 additional days after release. Finally, we selected those CpG sites that reached an average difference of at least 5% at day 8 (4 days post-release). This results in a 580 CpG sites signature associated to TGF- $\beta$  exposure (Figure 62 and Table S3). In addition, differentially methylated sites were classified into different clusters according to their pattern of expression (Figure 61).



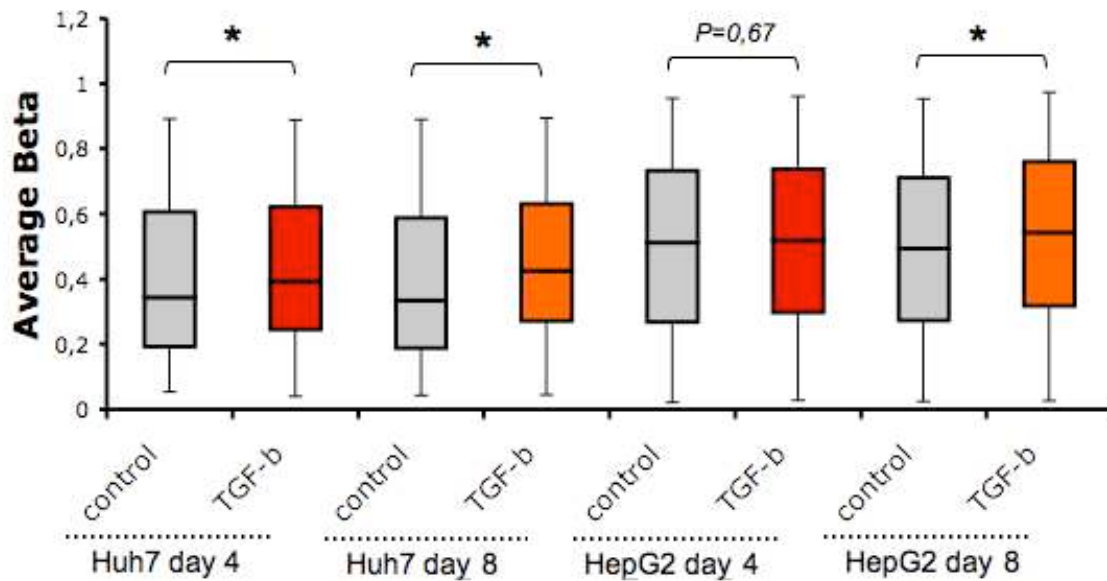
**Figure 60. A differential methylome distinguishes TGF- $\beta$  exposed cells to controls.**

Heatmap represents all probes differentially methylated ( $p < 0.001$ ;  $FDR < 0.05$ ,  $\Delta \beta > 5\%$ ) between control and TGF- $\beta$  treated cells, in both cell lines, and both time points. Methylation level is expressed in log scale, with higher methylation represented in red and lower methylation shown in blue. The numbers on the right point to 5 different probe clusters selected according to their behavior across all samples. A fraction of each cluster is depicted in more details in Figure 61.



**Figure 61. Description of the probe clusters.**

Each cluster presented in Figure 60 are detailed for a fraction in order to illustrate some of the significant genes within each category.

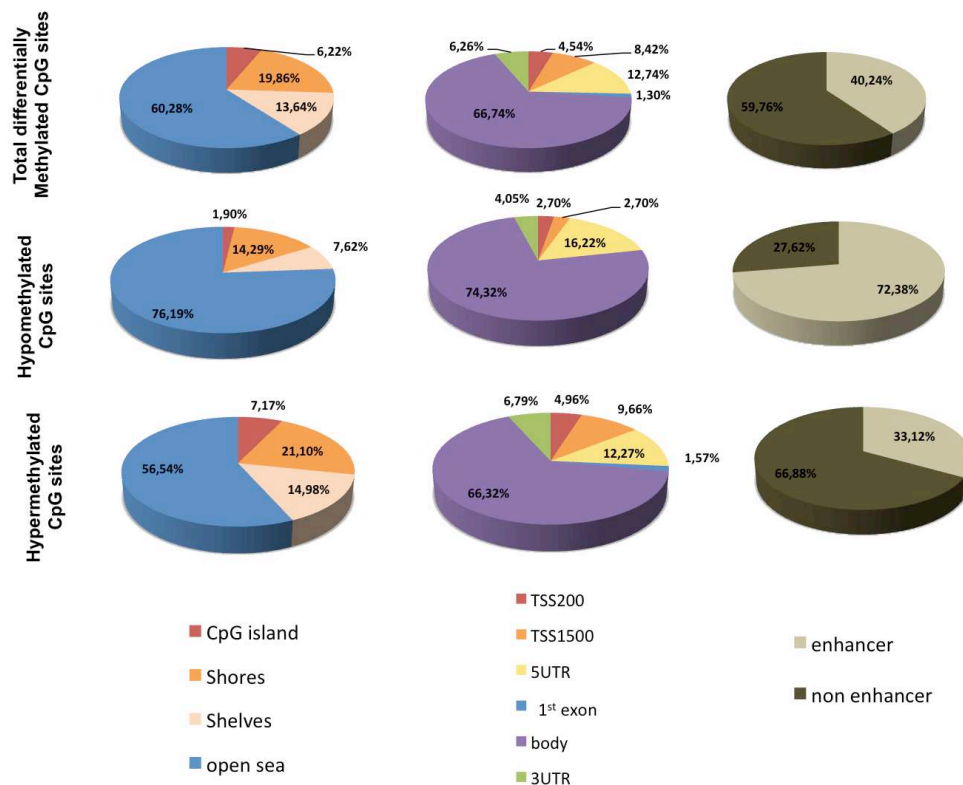


**Figure 62. A 580 loci DNA methylation signature can distinguish TGF-β exposed cells from their negative counterpart.**

Median methylation (and distribution) for all differentially methylated loci distinguishing TGF-β exposed vs control in each cell line. (\*) indicates P value < 0.05 relative to non-treated cells at the corresponding time point for each cell lines separately.

Four out of five clusters represented genomic loci consistently hypermethylated after TGF-β treatment in both cell lines (Figure 61). These loci included both *de novo* DNMTs, *DNMT3A* (one CpG site) and *DNMT3B* (two CpG sites). Differentially methylated sites also included TGF-β-related and chromatin-related genes, such as *CDKN1B*, *COL1A1*, *TRRAP*, *HDAC7*, *ARID3A*, and *KDM6B*. One cluster corresponded to probes significantly hypomethylated after TGF-β exposure, including relevant loci within genes involved in cell migration and inflammation such as *CALD1*, *BMP1*, *IL18*, and *IRAK2*. The majority of differentially probes were hypermethylated after TGF-β treatment (n= 474; 82%). In a similar way to the CD133 methylome, we found an enrichment of differentially methylated probes in “open sea” and gene body regions (60.3% and 66.7%, respectively) (Figure 63). In addition, these differentially methylated sites in open sea regions are related to gene regulation and many of them are localized in enhancer regions (61.8%). This “open sea enrichment” was even higher for hypomethylated loci (76.1%). Finally hypomethylated loci displayed a high enrichment for enhancer regions (2 fold) (Figure 63).





**Figure 63. Regional distribution of the differentially methylated CpG sites after TGF- $\beta$  treatment.**

The 580 significant loci were distributed according to CpG island relationship as Island, north shore, south shore, north shelf, south shelf, and open sea, and are represented in the left pie charts. Middle pie charts represent the distribution of significant loci in relation to annotated genes (within 200 or 1500 bp from the TSS, 1<sup>st</sup> exon, 3' or 5' UTRs, and gene body). Right pie charts represent the fraction of differentially methylated probes annotated to a known UCSC enhancer.

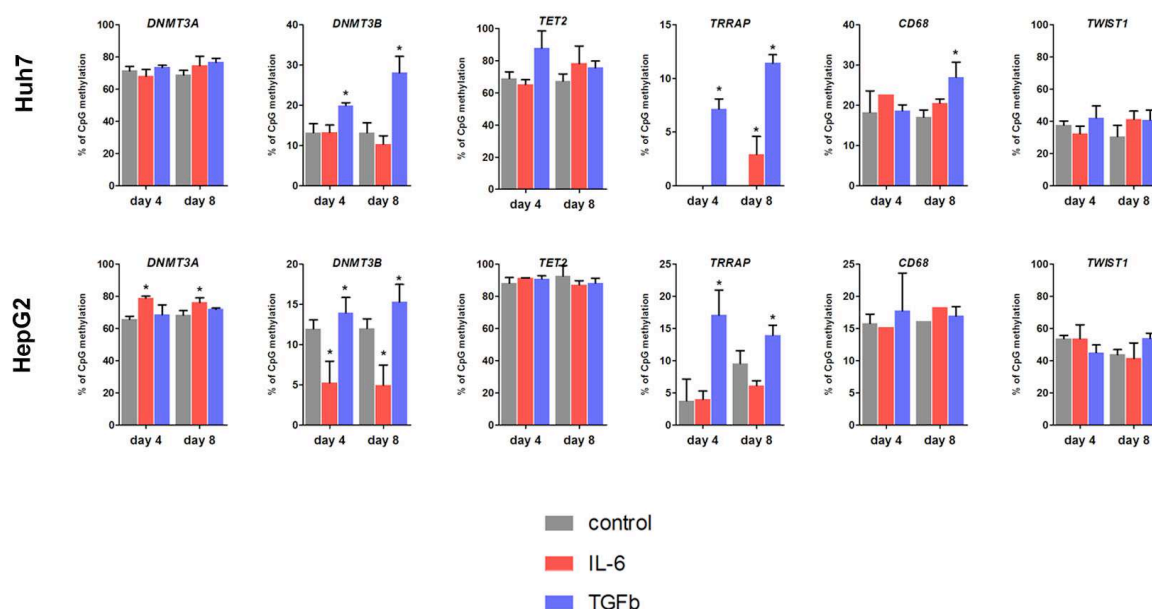
A selection of 6 differentially methylated loci, including DNMTs, were validated using an independent quantitative method, pyrosequencing (Figure 64). Importantly most of the results obtained after pyrosequencing were significantly correlated to the results obtained after the array ( $P < 0.001$ , Figure 65). We observed a weak correlation for only one CpG locus (located in *DNMT3b*), but even so the correlation coefficient remain satisfactory (0.83 when all the 6 CpG sites are included, 0.96 when *DNMT3b* locus is excluded). This correlation indicates that the results obtained after the bead array are fully validated.

We next performed gene ontology analysis and found a notable enrichment in developmental and stemness pathways including Wnt/ $\beta$ catenin, Notch, Shh/Hedgehog, MAPK and JAK/STAT signaling pathways (Table S4).

Our data shows that the effect of TGF- $\beta$  in liver cancer cell lines goes in parallel with a remarkable reconfiguration of the DNA methylome at multiple loci. This reconfiguration is stable and common to two independent cell lines, and affects a significant proportion of

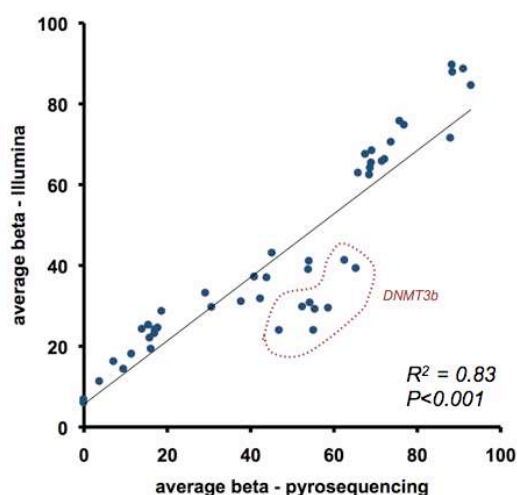


enhancer regions, potentially linked to gene expression changes. The TGF- $\beta$  methyl-sensitive signature described here includes DNA methylation players themselves and a significant enrichment of TGF- $\beta$  pathway loci (Table S3), indicating a potential role for DNA methylation in establishing a TGF- $\beta$ -induced phenotype switch in these cells.



**Figure 64. Validation by pyrosequencing of selected differentially methylated loci.**

A selection of significant loci were validated by pyrosequencing (as described in Methods), in both cell lines. (\*) indicates P value < 0.05 relative to non-treated cells at the corresponding time point.

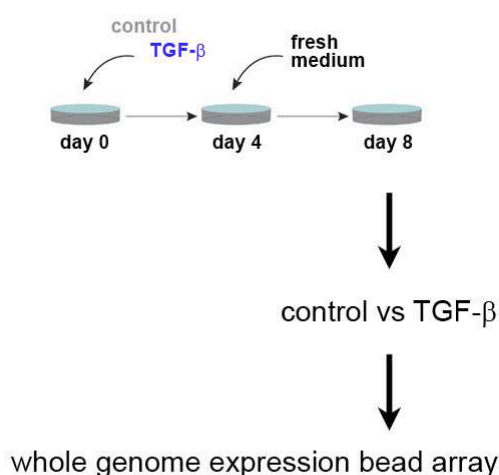


**Figure 65. Correlation between pyrosequencing and Illumina bead array analyses.**

For both Huh7 and HepG2 and for both time point, average beta of 6 CpG loci (located in *TET2*, *TRRAP*, *CD68*, *DNMT3b*, *DNMT3a* and *TWIST*) obtained after pyrosequencing and Illumina bead array were plotted. The dashed line delimits the values for one precise CpG locus (located in *DNMT3b*). Correlation was found statistically significant (Spearman test).

## VI. TGF- $\beta$ -induced methylome matches the basal CD133+ methylome and is reflected on mRNA expression

To gain a better insight into the consequences of TGF- $\beta$ -induced methylome on the phenotype, we performed a whole genome expression analysis in both HuH7 and HepG2 cells. We chose the 8-days time point (4 days of TGF- $\beta$  treatment + 4 days post-release), considered in our model as the one defining long-term, stable changes induced by this cytokine (Figure 66). Bead array transcriptome analysis showed an expected profile of gene expression in both cell lines, including known TGF- $\beta$  targets (Table S5) such as *SMAD6* and *SMAD7* (respective fold-change 0.62 and 0.69). Moreover gene ontology analysis confirmed that the signature we observed is specific to TGF- $\beta$  as this pathway was always displayed in the different analysis (Table S5).



**Figure 66. Experimental design for whole genome expression study in TGF- $\beta$  exposed cells.**

Whole genome expression analysis was performed after 4 days of TGF- $\beta$  exposure (+4 days post-release) in both cell lines, as described in Methods. RNA from control and treated conditions was interrogated with Illumina expression bead arrays.

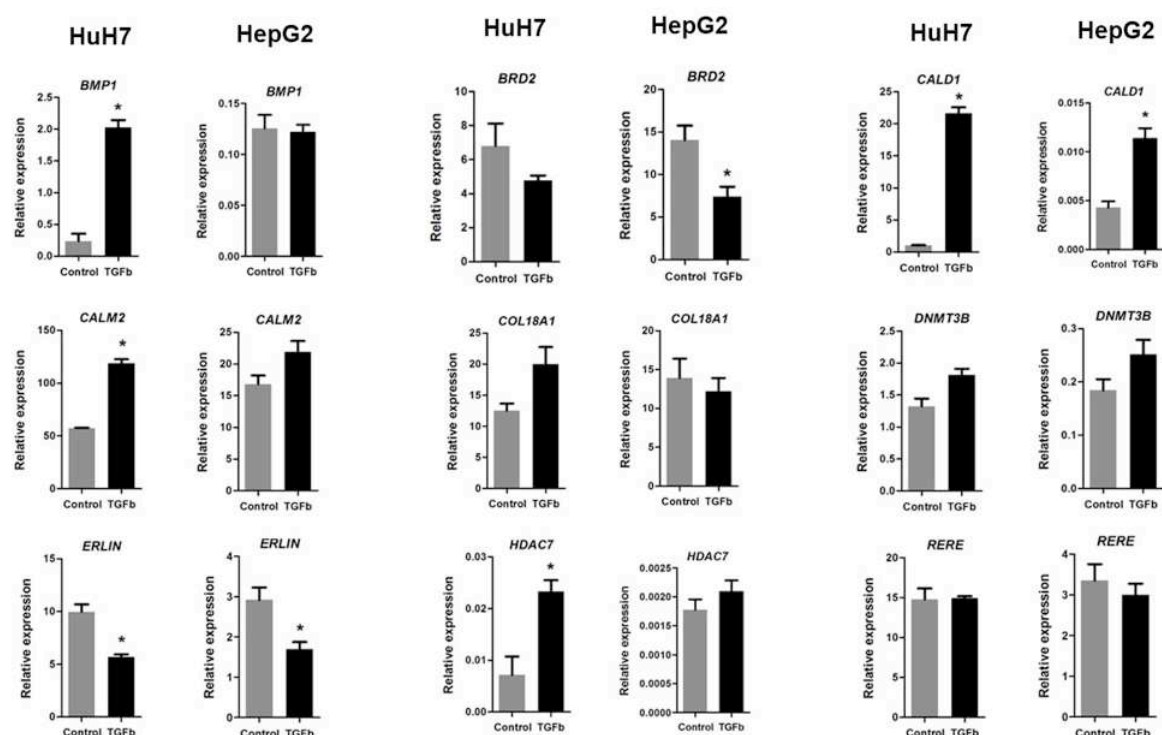
However, when intersecting the expression (n=1032) and methylation (n=242) significant gene lists, there was no significant overlap (26 common genes) (Tables 11 and S6). Interestingly, a majority of overlapping genes (17 out of 26) was positively correlated between mRNA expression and DNA methylation. This was the case for key TGF- $\beta$  pathway targets such as *BMP1*, and *de novo* DNA methylation factor *DNMT3B*.

We validated the data obtained for a set of genes by quantitative PCR (Figure 67). The analysis here were done for each cell line separately and thus the change in gene expression between control and TGF- $\beta$  treated samples are not always significant (in comparison to the statistical analysis for bead array transcriptome that was done using the two cell lines combined together). However data obtained by qPCR were significantly correlated with the data obtained with the bead array transcriptome ( $p < 0.001$ ) (Figure 68). Only two genes in the HepG2 cell line (*COL18A1* and *HDAC7*) presented opposite fold change directions between qPCR and bead array analysis. Nonetheless the overall correlation coefficient was acceptable (0.78 when all the genes were included, 0.88 when *COL18A1* and *HDAC7* in HepG2 were excluded). In conclusion, this good correlation between qPCR and bead array analyses validated the data obtained after the whole-genome expression array after TGF- $\beta$  exposure.

**Table 11. Correlation between TGF- $\beta$ -induced DNA methylation signature and gene expression.**

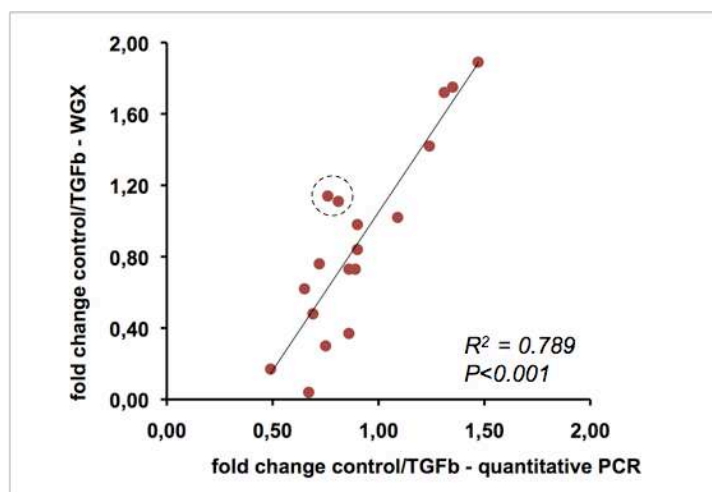
26 overlapping genes are listed below, with red indicating increased expression/methylation, and green indicating reduced expression/methylation after TGF- $\beta$ .

	methylation	expression
<i>ACSL3</i>	1.35	0.76
<i>AHNAK</i>	0.73	0.75
<i>BCR</i>	1.37	1.08
<i>BMP1</i>	0.54	0.89
<i>BRD2</i>	1.52	0.73
<i>C17orf101</i>	1.52	1.19
<i>CALD1</i>	0.57	1.35
<i>CALM2</i>	1.37	1.37
<i>COL18A1</i>	2.17	1.45
<i>DACT2</i>	1.56	1.19
<i>DDA1</i>	0.76	1.47
<i>DDX19B</i>	1.30	1.23
<i>DNMT3B</i>	1.39	1.15
<i>ERLIN1</i>	1.59	0.74
<i>GIPC1</i>	1.59	1.47
<i>HDAC7</i>	1.39	1.28
<i>MAEA</i>	0.63	1.28
<i>NRP2</i>	1.37	1.14
<i>PDLIM1</i>	0.45	1.89
<i>RAP1GAP2</i>	1.30	1.35
<i>RERE</i>	1.39	1.19
<i>SLC22A18</i>	0.63	0.56
<i>SRC</i>	1.33	0.78
<i>STARD13</i>	0.68	1.19
<i>TLE1</i>	1.79	1.45
<i>WDR25</i>	1.39	1.15



**Figure 67. Whole genome expression array validation.**

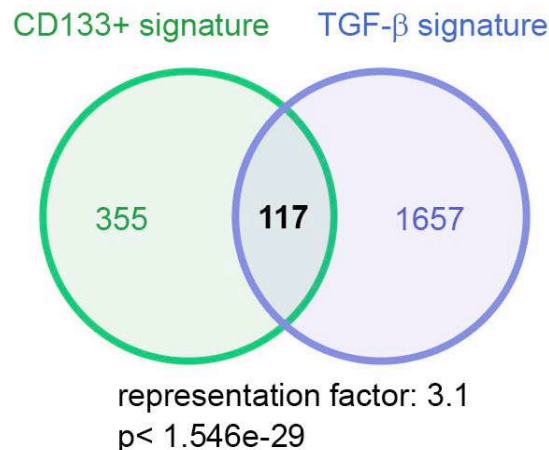
A selection of significant genes was validated by qRT-PCR in both cell lines. (\*) indicates P value < 0.05 relative to non-treated. Expression was normalized to housekeeping gene.



**Figure 68. Correlation between whole genome expression (WGX) array and quantitative PCR analyses.**

Fold change (between control and TGF-b samples) for each cell lines separately for 9 genes (BMP1, BRD2, CALD1, CALM2, COL18A1, DNMT3b, ERLIN1, HDAC7 and RERE) were plotted. The dash line delimits values for which the trend of qPCR and WGx are going in opposite direction (<1 for the quantitative PCR analysis and >1 for the whole genome expression array). Correlation was statistically significant (Pearson test).

Our two independent genome-wide experiments have shown that basal CD133+ cells from two liver cancer cell lines display a common methylome signature, and that TGF- $\beta$  is able in turn to induce a common reconfiguration of the methylome. As TGF- $\beta$  stably induces a *de novo* fraction of CD133+ cells, we asked whether the DNA methylation changes induced by TGF- $\beta$  were similar to the basal CD133+ cells methylation profile, as obtained by FACS sorting from non-treated cell cultures. To answer this question, we studied the overlap between the two signatures (i.e. CD133+ and TGF- $\beta$ ) defined above, common to Huh7 and HepG2 cells. At p values <0.001, the CD133+ signature corresponds to 472 annotated genes, while the TGF- $\beta$  signature represents 1774 genes. We observed a significant overlap of 117 genes when intersecting both signatures (Figure 69 and Table S7). This overlap is highly significant ( $p < 1.5e-29$ ) and represents 3 times more common sites than expected by chance. This result suggests that basal CD133+ cells and TGF- $\beta$ -induced CD133+ cells share a common methylome, potentially involved in sustaining their functional characteristics. Interestingly gene ontology analysis of this common signature highlights the Wnt/ $\beta$ -catenin, mTOR and Notch pathways (Table S7). These pathways were already described in our CD133+ signature and are known to be linked to stem cell phenotype.



**Figure 69. Overlap between CD133+ and TGF- $\beta$  DNA methylation profiles defines a significant signature of 117 genes.**

This analysis suggests that CD133+ cells induced by TGF- $\beta$  are similar, at the methylome level, to CD133+ cells in basal conditions. Based on this assumption, we next asked whether TGF- $\beta$  has an effect on the methylome that is independent on those changes that define the CD133+ subpopulation. To answer this question we used all our bead array data, and modeled the main components of methylome variation in a linear regression (as described in

Materials and Methods). Assuming that two known factors are able to modify the methylome based on our own results (the cell line of origin and the CD133-status), we included these two variables in the model. In addition, we included the potential effect of TGF- $\beta$ , independent of the other two factors. Interestingly, this analysis shows TGF- $\beta$  has an additional effect on the methylome, independent on the induction of CD133+ cells (Table S7).

In summary, TGF- $\beta$ -induced methylome resembles the basal CD133+ methylome and is partially reflected at the transcriptional level. A subset of TGF- $\beta$  methyl-sensitive loci positively correlates with gene expression.

## **DISCUSSION**





## **I. CD133+ cells represent a distinct sub-population related to cancer stem cells in HCC cell lines.**

In the present study, we investigated the characteristics of CD133+ cells in two hepatocellular carcinoma cell lines. Although CD133 is one of the surface markers most consistently used for detection and isolation of putative liver CSCs, it was important to establish in a first step the frequency of expression in our cell lines. CD133 is a well described cell surface marker for stem cells and CSCs from several human tissues (Grosse-Gehling et al., 2013). Even if its own function is not yet fully established, its expression is thus believed to be tightly associated to a specific phenotype (i.e. stem cell phenotype). In this study, we found that CD133 expression on the extracellular surface varies from 2 to 5% (in HepG2) to more than 70 % (in PLC/PRF/5). This variation in the marker expression was not surprising considering that these cell lines, in spite of sharing several properties, are completely independent and harbour some specific features. For Huh7 and HepG2 the percentages we observed by FACS (respectively 15-40% and 2-5%) are similar to what has been reported in the literature (Suetsugu et al., 2006; Yang et al., 2008b; Zhu et al., 2010), with the exception of Kohga et al., (2010) who detected more than 43% of CD133+ cells in the HepG2 cell line. For PLC/PRF/5 the percentage of CD133+ cells detected by FACS varies significantly between different studies, from zero to 95% (Haraguchi et al., 2010). We have detected a rather high percentage, around 70%, of CD133+ cells for PLC/PRF/5. This variability between the studies for the same cell line can be explained partly through the difference in the antibody used but mainly by long-term and independent cell phenotype evolution of different batches of PLC/PRF/5. It is known that immortalized cell lines continue to accumulate genetic alterations through passages (Lin et al., 2003; Noble et al., 2004; O'Driscoll et al., 2006). Thus, similar experiments conducted on different batches of one cell line with a great gap for number of passages can result in different conclusions. We had a similar finding for the CD44 marker in Huh7: while we observed that the majority (80%) of the cell expresses CD44, only a few percent of positive cells was detected by others (Suetsugu et al., 2006; Yang et al., 2008b). For other markers such as CD90 and EpCAM, the percentage of positive cells matches the observations done by the others studies (Kimura et al., 2010; Piao et al., 2012; Yang et al., 2010b).

The expressions of these markers are not mutually exclusive, and the combination of two markers has been investigated in order to refine the identification of liver CSCs. In this way, CD44 and EpCAM have both been used in combination with CD133 in order to better define the CSC population (Chen et al., 2012b; Zhu et al., 2010). CD133+/CD44+ cells and CD133+/EpCAM+ cells have shown higher tumorigenicity and are more aggressive for

metastasis compared to their CD133+/CD44- and CD133+/EpCAM- counterparts. We also examined if the combination of two markers would allow us to define a more precise subpopulation in our cell lines. Unfortunately, we were confronted to the disparity of the marker expression between cell lines and when combining two markers we could not define one subpopulation clearly identifiable in our three cell lines. CD44 expression is close to 80% in Huh7, so in our case the CD44+ population encompasses all the CD133+ population, as opposed to what has been previously reported (Zhu et al., 2010). In contrast, CD44+ cells represent less than 1% of the population in HepG2 and PLC/PRF/5 and thus no CD133+/CD44+ cell was detectable by FACS in those cell lines. We encountered the same difficulty for EpCAM, which is expressed at very high level in Huh7 and HepG2 (60% in average) but that is hardly detectable in PLC/PRF/5 (less than 1%). Finally, the CD90+ population was too small in the three cell lines analyzed (less than 0.5%), and therefore could not be detected in combination with another marker.

In summary, based on the consistent literature supporting its use, the ability to detect a discrete cell subpopulation across our cell lines, and the convenience of using a single surface protein for downstream magnetic cell sorting analyses, we chose CD133 as the surface marker that was best suited to pursue our main objectives. Although the study of CSCs by using only one marker is still under debate, the accuracy of CD133 expression to define liver CSCs has been deeply explored in several reports. CD133+ cells have been identified in HCC tumors several times at low frequency (less than 5%) (Kim et al., 2011; Sasaki et al., 2010; Yin et al., 2007; Zen et al., 2011) and their presence was correlated to the disease stage, poor prognosis, lower overall survival and higher recurrence risk. Moreover, in opposite to CD133- cells, CD133+ cells isolated from HCC tumor samples present highest ability to form spheres and express at high level genes related to stemness phenotype (*NOTCH*, *BMI-1*, *POU5F1*, *NANOG*) (Ma et al., 2008b, 2010). The CD133 expression in HCC delimitates thus a phenotypically distinct subpopulation that appears to not only support the tumor growth and lead the outcome of the disease but also present specific tumorigenic capacities. These phenotypic characteristics are also present in CD133+ cells from HCC cell lines and additional specific features such as clonogenicity (Haraguchi et al., 2010; Yin et al., 2007), metastatic potential (Ma et al., 2008b) and drug or chemo-resistance have been further identified (Hagiwara et al., 2012; Piao et al., 2012). The link between stem cell related phenotype and CD133 expression is thus nowadays abundantly established and our aim was to further continue the epigenetic characterization of CD133+ cells in HCC using those previous studies as a baseline.

As a first step, it was important and crucial to ensure that in our *in vitro* model CD133+ populations did represent a distinct subpopulation related to CSCs. After enrichment using the MACS technology we were able to demonstrate that CD133+ cells express higher levels of the transcripts for *NANOG* and *OCT4* (*POU5F1*) in three independent cell lines and of *SOX2* in two cell lines. In HepG2 and Huh7 cell lines, this expression profile was associated to a higher mRNA expression of the DNA methyltransferases *DNMT3A* and *DNMT3B*. These enzymes are involved in the *de novo* establishment of DNA methylation marks. It was already shown that regulations in DNMTs expression are linked to the stem cell phenotype: DNMT3 family members have been reported to be deregulated in cardiac progenitor cells and induced pluripotent stem cells (iPSCs) and this deregulation was directly linked to the acquisition of a stem cell phenotype (Chen et al., 2013b; Guo et al., 2013). More precisely, the activity of DNMTs has been involved in neural stem cell proliferation (Li et al., 2013), and modulation of this activity induced differentiation of both, somatic and embryonic stem cells (Banerjee and Bacanamwo, 2010; Mahpatra et al., 2010). In our study we did not investigate the activity of DNMTs in CD133+ cells, but their higher expression is likely to be associated with a global increase of activity. Globally this observation, taken together with an increase in stemness genes, suggests that CD133+ cells have a distinct phenotype, probably closer to stem cells. We showed that these gene expression profiles are associated to a difference in functionality, as demonstrated by the higher capacity of Huh7 CD133+ cells to produce spheres. The ability for single cells to proliferate and form spheres in low attachment conditions is thought to be restricted to cells with stem cell properties, and is commonly used as a surrogate to *in vivo* tumor initiating assays (Pastrana et al., 2011). Thus, this observation validates that CD133 expression can be used as criteria to identify cells with stem cell abilities in Huh7. Unfortunately we were unable to reproduce this result in HepG2, probably due to technical limitations. Practically, our optimization of the MACS protocol to sort CD133+ cells allowed us to increase from 2 to 3 fold the initial percentage of CD133+ cells. However, considering the small initial percentage of CD133+ cells in HepG2 (from 2 to 5%) this protocol resulted in a maximum of 30% of CD133+ cells in the final enriched fraction (in contrast, we were able to reach 90% of CD133+ cells for Huh7 final enriched fraction). The gap between depleted and enriched fractions for HepG2 was probably not sufficient to observe a difference in the sphere formation ability. However, HepG2 CD133+ cells were already reported to form more spheres and more colonies compared to CD133- cells (Ma et al., 2007). In that study a successful enrichment to 95% purity for CD133+ cells after sorting was reported. Nevertheless, taken together our preliminary results support in a consistent

manner that CD133+ cells in Huh7 and HepG2 represent a distinct sub-population related to liver CSCs.

## **II. CD133+ cells phenotype is associated to a specific DNA methylation signature.**

We established that CD133+ cells in Huh7 and HepG2 cell lines express *NANOG*, *OCT4*, *SOX2*, *DNMT3A* and *DNMT3B* at significantly higher levels. Our results are consistent with several other studies demonstrating that CD133+ HCC cells have a specific gene expression profile (including *NANOG*, *OCT4* and *SOX2*) (Ma et al., 2008b, 2010). Thus, up-regulation not only of stemness genes such as *BMI-1*, but also of genes involved in stem cell related pathways (i.e Wnt/ $\beta$  catenin pathways, Notch and Hedgehog signalling pathway) and genes involved in drug resistance (i.e. ABC transporter family) has been described in CD133+ cells (Ma et al., 2010). Specific gene expression profile is likely to be dictated by both the availability and activation of a set of transcription factors in accordance with a specific chromatin profile (including histone modifications and DNA methylation marks). Besides, stem/pluripotent cells are known to harbour some particular epigenetic marks such as hyperdynamic chromatin (in order to preserve the possibility to rapidly differentiate depending of the tissue needs), the presence of bivalent domains possessing both active and repressive histone marks and a global DNA hypomethylation – indeed DNA methylation is usually associated with shutdown of genes that are no longer necessary for the differentiation process (Hernandez-Vargas et al., 2009). Large-scale epigenetic signatures are important because they provide comprehensive and integrative information about mechanisms involved in the cellular phenotype maintenance. For example, methylome analysis of breast CSCs revealed activation of inflammatory pathways for breast CSCs maintenance (Hernandez-Vargas et al., 2011). DNA methylation signature can also serve as criteria for phenotype classification. In this manner, when no true genetic-based classifications are available for liver cancer, DNA methylome analyses were shown to be able to classify tumors according to their grade and to their etiology (Hernandez-Vargas et al., 2010; Lambert et al., 2011). This last observation underlies the fact that, as epigenetic marks are inheritable through cell divisions, epigenetic signatures can also provide information on the origin of the cell. Thus, epigenetic signature characterization provides a strong and powerful tool to define one cellular or tissue phenotype. Our observation, never reported before, that the enzymes involved in *de novo* DNA methylation, *DNMT3A* and *DNMT3B* are over-expressed in CD133+ cells strongly suggests that this subpopulation harbours a specific DNA methylation program.

Except for a global *LINE-1* demethylation (Zhang et al., 2011), no epigenetic signature has been yet investigated for CD133+ liver CSCs. Only one study investigated the methylation in the side-population (SP) for Huh7 and PLC/PRF/5 cell lines (Zhai et al., 2013). The DNA methylation array used in this study restricted the analysis to CpG islands within gene promoters. A common signature of 72 hypermethylated loci and 181 hypomethylated loci was described for Huh7 and PLCR/PRF/5. Here, we performed an Illumina Infinium 450K beads arrays assay, on CD133+ HCC derived cells, that interrogates more than 480,000 CpG sites for DNA methylation status among the entire genome (including intergenic regions) and for CpG sites comprised not only in CpG islands but also in shores, shelves or open seas. Our analysis using this array revealed a unique DNA methylation signature in CD133+ HCC-derived cells of 608 differentially methylated sites (with an averaged delta beta > 5%). Although these probes were associated with a good p-values (<0.001), the FDR were higher (0.58). These values reflect probably the minimal number of samples (2 cell lines and 2 biological duplicates per cell line), and therefore, the variability between replicates. However, the analysis of the localisation of the differentially methylated probes was consistent with the CSCs phenotype. First, *CD133* itself (*PROM1*) was found hypomethylated in CD133+ cells (7.9% in HepG2 and 4.4% in Huh7). Although the difference of methylation is modest, the concerned locus is localized in a CpG island within the promoter and therefore is likely to be involved in the increase of CD133 expression that we detected in Huh7 and HepG2 CD133+ cells. This promoter hypomethylation was already described in CD133+ Huh7 cells (You et al., 2010b), and in a more extensive way in other CD133+ CSCs such as glioblastoma (Tabu et al., 2008) and neuronal CSCs (Schiapparelli et al., 2010) and has also been correlated to CD133 expression. Our results, with these previous observations, indicate thus that CD133 expression is regulated by methylation on its promoter. Besides, our pathway analyses revealed a significant enrichment in genes involved in Akt, Wnt, Hedgehog and mTOR signaling pathways. As described in the introduction, previous mechanistic studies and gene expression arrays already reported and validated the deregulation of these pathways in liver cancer CD133+ cells (Ma et al., 2007, 2008b; Yang et al., 2011). Our results could be further consolidated by a precise comparison between CD133+ transcriptome and our DNA methylation signature. Other pathways involved in DNA repair, telomerase function, immortality and cell aging were also displayed and are known to be required for establishing cancer and stem cell phenotypes. The presence in CD133+ signature of genes involved in these pathways and these processes is thus not surprising and supports the idea that this methylome signature is linked to a CSC

phenotype. Moreover, analyses for enrichment in genes regulated by common transcription factors revealed that several genes differentially methylated in CD133+ cells are regulated by OCT4 (POU5F1) (Table S2). This transcription factor is known to be essential for establishment and maintenance of the stem cell phenotype. Taken together, these results sustain the reliability of our analyses and link the CD133+ methylation signature to the recognized properties of CSCs or stem cells. Moreover, it confirms the importance of these mechanisms (activation of developmental pathways, increased DNA repair efficiency and stem cell transcription factor expression) in CSCs and suggests that they are regulated through DNA methylation.

We observed that globally CD133+ cells are hypomethylated (84%) compared to CD133- cells. While we did not compare hypomethylation to gene expression, this general hypomethylation may be correlated to a more “open chromatin” state, that may coincide with the transcriptionally permissive state previously described in stem cells (Hernandez-Vargas et al., 2009).

Comparing our signature to the DNA methylation profile obtained by Zhai *et al.* (2013) in Huh7 and PLC/PRF/5 SP cells, we have 6 genes in common and only 3 of that are differentially methylated in the same direction in CSCs (CD133+ or SP). This can be explained by two major differences in the two experiments. First we did not use the same marker to target CSCs in HCC cell lines (CD133 expression vs. side population) and second the DNA methylation arrays used do not cover the same panel of CpG sites. The array used by Zhai *et al.*, is limited to CpG loci comprised in CpG islands and gene promoters when the Infinium bead array covers a much larger part of the genome with not only gene promoters but also in gene regions, intergenic regions and interrogates CpG loci not only comprised in CpG islands but also in shore, shelves and open sea regions. This can explain first why we obtained a bigger signature (608 versus 253 loci) and that even for the few genes that we found in common, the probe were not necessarily localized in the same region which can explain that the changes in DNA methylation are not similar (for example, a gene can be both subjected to promoter hypermethylation and gene body hypomethylation).

Beyond pathways and molecular characteristics linked to stem/pluripotent phenotype, differentially methylated sites are related to many genes involved in inflammation. Inflammation is known to be involved in carcinogenesis, liver progenitor activation and several types of CSCs including breast, brain and blood (Iliopoulos et al., 2009; Naka et al.,

2010; Peñuelas et al., 2009). But as for liver cancer, interaction between inflammation and CSCs has not been well explored. One study (by You et al., 2010) reported that in liver cancer TGF- $\beta$  supports the expression of CD133 through demethylation of its promoter. The methylation signature characterized in our study provides a strong connection between inflammatory pathways and CSCs. We observed differentially methylated sites within genes involved in JAK/STAT, p38 MAPK, NF- $\kappa$ B, TGF- $\beta$  and several interleukins signaling pathways. More precisely, the genes involved in IL-6 and TGF- $\beta$  signaling pathways were differentially methylated between CD133+ and CD133- cells and included *SMAD3*, *SMAD4*, *STAT3*, *JAK2*, *IL-6*, *TGF- $\beta$ 1*, *TGF $\beta$ RI* and *TGF $\beta$ RII*. Naturally, this methylation profile does not necessary correlate with a constitutive activation of these pathways and indeed for TGF- $\beta$  and IL-6 signalling pathways we did not detect a consistent higher expression of their respective target genes in Huh7 and HepG2 CD133+ cells . If this DNA methylation profile for IL-6 and TGF- $\beta$  related genes is not correlated with gene expression it can still constitute a permissive/restrictive chromatin state that will condition the cellular response to any future exposition to these cytokines. Together our results provide the evidence for the existence of a link between CSCs and inflammation, although further studies are needed to elucidate the exact nature of this interaction.

### **III. CD133+ liver CSCs are triggered by TGF- $\beta$ .**

After the observation that in CD133+ cells the DNA methylation signature highlights inflammatory pathways, we explored the link between inflammation and CD133+ cells in HCC cell lines. Because TGF- $\beta$  displays multiple roles in HCC progression and also was associated to CSC phenotype in other tissues (Cao et al., 2012; Wang et al., 2011b), we focused our study on this cytokine. Although the TGF- $\beta$  signaling pathway was not the most affected by DNA methylation deregulation in CD133+ cells, analyses of transcription factors binding sites revealed that several differentially methylated genes possess binding sites for SMAD3, SMAD2 or SMAD4 (Table S2). As suggested earlier, this methylation profile could mean that CD133+ cells harbour marks on TGF- $\beta$  target genes and thus would respond differently to TGF- $\beta$  stimulus different from their CD133- counterparts. Indeed, it has been suggested that DNA methylation marks do not always have an immediate and direct impact on gene expression, but would instead anchor the intention for gene expression by settling an appropriate chromatin structure. Consistent with this notion, hypomethylation may provide a permissive structure for gene expression and hypermethylation may irrevocably

lock the gene silencing. In this manner, the DNA methylation profile in CD133+ cells could condition their response to TGF- $\beta$  exposure. As an example, Kabashima et al., (2009) compared the effectiveness of TGF- $\beta$  induced EMT between side population (SP) and middle population (MP) in pancreatic cancer and demonstrated that the SP was much more sensitive to TGF- $\beta$ -induced phenotypical switch. This link between TGF- $\beta$  and stem cell phenotype via EMT regulation has also been shown to support CD133+ CSC phenotype in lung cancer (Pirozzi et al., 2011). We were able to demonstrate that TGF- $\beta$  exposure is able to increase the number of CD133+ cells in Huh7 and HepG2. Although the induction of CD133+ cells in Huh7 was already described by You *et al.* (2010) our study demonstrated the stability of the effect and the generality of this phenomenon in an independent cell line. We designed an *in vitro* model, where after the treatment of TGF- $\beta$  we included several days of “rest” in fresh new media. We further utilized this model on CD133- cells and demonstrated that TGF- $\beta$  was not only able to increase but also to induce CD133+ cells that persisted for long time, consistent with TGF- $\beta$ -mediated setting up of long-term memory system..

Although in Huh7, CD133- cells presented the ability of transdifferentiation into CD133+ cells in the absence of any treatment, in HepG2 it was not the case. Nevertheless, homeostasis between CSCs and non-stem cancer cells was already investigated in breast and colon cancer. In both studies non-stem cells presented natural ability to dedifferentiate in CSCs, and inversely CSCs were able to produce non-stem cancer stem (Chaffer et al., 2013; Yang et al., 2012). The interconnection between the two populations was stabilized to equilibrium. Interestingly these two studies demonstrated that downregulation of TGF- $\beta$  itself or ZEB1 (an EMT mediator) were able to disturb this process. Inversely, the addition of TGF- $\beta$  was only able to accelerate the rate of transdifferentiation but not to change the final homeostatic state. Comparing these results with our own observations, it appears that Huh7 cells possess this homeostasis state between CD133+ and CD133- cells (4 days after depletion of CD133+ cells, the percentage of CD133+ cells reached the initial level and did not change after 8 days). However here TGF- $\beta$  seems to do more than just accelerate the recovery of the initial ratio between CD133+ and CD133- cells as after 4 days the number of CD133+ cells was higher than in the untreated population and that this number continued to increase at day 8 even after the end of the treatment. In our model, TGF- $\beta$  seems thus to disturb the homeostasis between CSCs and non-stem cancer cells in favour of CSCs. This deregulation is likely to involve a complex transdifferentiation mechanism.



The level of both CD133 protein and mRNA remained elevated 4 days after the end of TGF- $\beta$  treatment, demonstrating that in this case TGF- $\beta$  did not induce a short and limited response, but truly supported a cellular transdifferentiation that is maintained after the end of the stimulus. Cellular morphology after TGF- $\beta$  treatment was also altered: both Huh7 and HepG2 cells became more elongated mimicking EMT morphology. Interestingly, this morphological switch was maintained after the end of the treatment (data not shown), meaning that beyond the CD133+ transdifferentiation TGF- $\beta$  is inducing a global switch in cellular phenotype and this switch appears to be stable. Furthermore, for comparison we also performed these experiments with IL-6. This pro-inflammatory cytokine is well known for being involved in liver cells proliferation and for breast CSCs initiation (Iliopoulos et al., 2009). It was thus interesting to compare the effects of this cytokine on CD133+ cells to these of TGF- $\beta$ . Interestingly, while IL-6 displayed similar ability as TGF- $\beta$  to increase CD133+ cells, its effect was not stable after the end of the treatment. This indicates that IL-6 has a transient effect and cannot induce a permanent transdifferentiation (unlike for breast CSCs), likely reflecting the differences between tissues and the fact that the cellular context can influence cytokine effects. Furthermore, the comparison between IL-6 and TGF- $\beta$  brings forward several hypotheses regarding their respective mechanisms to induce CD133+ cells. First, IL-6 effect disappears shortly after the release from the cytokine exposure. Second, in HepG2 cells where CD133- cells were not capable of spontaneously transdifferentiating into CD133+ cells, IL-6 cannot induce CD133+ cells. This argues that in Huh7, IL-6 does not really induce CD133+ cells but may just increase the proliferation of naturally induced CD133+ cells, whereas in HepG2 where there is no *de novo* formation of CD133+ IL-6 has no effect. In our model, IL-6 seems to have a transient proliferative effect on CD133+ cells and on the CD133+/CD133- cell homeostasis. In contrast, in Huh7 TGF- $\beta$  significantly decreased the cell proliferation rate during the treatment and cell proliferation re-accelerated after the release of the cytokine. This indicates that in Huh7 the induction of CD133+ cells does not rely on a proliferative effect on *de novo* formed CD133+ cells. For HepG2 cells, no effect on the cell proliferation rate was observed. Therefore, besides the difference observed in HepG2 and Huh7 for *de novo* CD133+ cells induction, we can conclude that for both cell lines the effect of TGF- $\beta$  on CD133+ cells does not involve changes in cellular proliferation and that the stability of the effect is not due to an inhibition of cell divisions. This last observation is critically important as it means that the transdifferentiation from CD133- to CD133+ phenotype can be transmitted through cell divisions. Moreover we showed as You *et al.* that the Huh7 CD133+ cells induced after TGF- $\beta$  treatment are able to produce spheres. We confirmed these results in HepG2 and further demonstrated that this capacity is maintained

after the end of the treatment. In contrast, a study on breast cancer cells showed that TGF- $\beta$  is able to stably induce EMT phenotype but that the effect on mammosphere formation was only transient and was lost at the end of the treatment (Dunphy et al., 2013). This again highlights the importance of the cellular context and the importance of functional assays to define CSCs. Together our results demonstrate that TGF- $\beta$  does not only induce the expression of CD133 marker, but induces a comprehensive and stable cellular reprogramming tightly linked to CD133+ liver CSC phenotype.

#### **IV. TGF- $\beta$ treatment induces a global and stable DNA methylation program.**

We were able to demonstrate that TGF- $\beta$  treatment can induce a stable reprogramming resulting in an increase of CD133+ CSCs. The stability of this reprogramming (including the up-regulation of *DNMT3A* and *DNMT3B*) is consistent with an epigenetically-induced change of phenotype. In particular, it indicates that TGF- $\beta$  could act through DNA methylation mechanisms. TGF- $\beta$  has already been described to act on gene expression through epigenetic mechanisms such as microRNAs regulation and histone modifications (McDonald et al., 2011). DNA methylation changes after TGF- $\beta$  treatment have also been investigated and some specific genes have been described as targets for DNA methylation changes (Dong et al., 2012; Eades et al., 2011; Kim and Leonard, 2007; Thillainadesan et al., 2012; Yeh et al., 2011; You et al., 2010b). Some studies however reported no methylation change after TGF- $\beta$  treatment, suggesting that this mechanism is not indispensable for TGF- $\beta$ -induced changes in gene expression (Acun et al., 2011; Akool et al., 2005; Dumont et al., 2008; McDonald et al., 2011; Pen et al., 2008; Wakabayashi et al., 2011). In those cases where TGF- $\beta$ 's effect is correlated with DNA methylation changes, TGF- $\beta$  seems to be able to induce both hyper- and hypo-methylation and these changes correlate with respectively down- and up-regulation of gene expression. Overall, these TGF- $\beta$ -related DNA methylation changes were only reported for few genes in independent studies and a complete landscape of DNA methylation changes after TGF- $\beta$  is still lacking. Here for the first time, using the Infinium Illumina 450K technology we revealed a global DNA methylation signature induced by TGF- $\beta$  in HCC cell lines. Interestingly, with our *in vitro* model we were able to define a DNA methylation signature that is stably propagated even after the end of the treatment. Using selective criteria ( $p < 0.001$ ,  $FDR < 0.05$  and averaged delta beta  $> 5\%$  4 days after the end of the treatment) we were able to define a signature of 580 loci (representing more than 400 genes) differentially methylated after TGF- $\beta$  treatment. If we compare our results with previous studies it is interesting to note that studies that did not report any

DNA methylation change after TGF- $\beta$  used experimental models where TGF- $\beta$ 's effect was transient. In particular Dumont *et al.* (2008), analyzed different models (reversible or not) of EMT and they reported that *E-Cadherin* silencing was associated to DNA methylation only when the EMT phenotype had been induced in a permanent manner. In addition, McDonald *et al.* (2011), described a TGF- $\beta$ -induced EMT model using non-transformed mouse hepatocytes and they investigated global methylation changes but did not found any differences between cells that underwent EMT and parental cells. Therefore, the implication of DNA methylation for TGF- $\beta$ 's effect strongly depends on the duration of its effect and the cellular context (transformed/non-transformed cells). Here, we investigated DNA methylation signatures associated to TGF- $\beta$  in a precise and constant *in vitro* model where the stability of TGF- $\beta$ 's effect has been clearly established in transformed cells. Admittedly, as suggested McDonald *et al.* (2011), cells that already underwent transformation can harbour epigenetic/genetic instabilities that render them more sensitive to further epigenomic changes. Nevertheless TGF- $\beta$  is known to be involved not only in cell transformation and cancer initiation, but also in cancer progression. In transformed cells, TGF- $\beta$  could set up a new DNA methylation profile that may reprogram the cell to promote tumor progression.

As described above, some target genes of TGF- $\beta$ , such as *TWISTNB*, *BMP1* and *SKI* were found to be a part of our signature, and the TGF- $\beta$  pathway itself was observed in the gene ontology analyses. Thus we confirmed that the known targets of TGF- $\beta$  can be regulated through epigenetic mechanisms involving DNA methylation. Interestingly, the gene ontology analyses revealed notable pathways such as Wnt/ $\beta$ catenin, Notch, Shh/Hedgehog, MAPK and JAK/STAT signaling pathways. Several of these pathways have already been described to interact with the TGF- $\beta$  signaling pathway (Cai *et al.*, 2013; Hussey *et al.*, 2012; Kurpinski *et al.*, 2010; Maitah *et al.*, 2011; Matsuno *et al.*, 2012; Zhou *et al.*, 2012a). DNA methylation profile could thus be one of the mechanisms involved in the cross-talk between these pathways.

TGF- $\beta$  is known to be able to fully reprogram cell and to set up a new gene expression profile (notably during EMT) but the mechanisms involved in this reprogramming are not understood. As mentioned before, several studies demonstrated that for specific genes TGF- $\beta$  regulates gene expression through remodelling the chromatin (i.e. by influencing histone modification marks) (McDonald *et al.*, 2011). Analysis of the biological function of our differentially methylated genes after TGF- $\beta$  treatment highlight categories related to gene

expression regulation, cell differentiation regulation and regulation of transcription. All these functions are consistent with the notion that TGF- $\beta$  could induce a reprogramming of the cellular transcription programme and that this reprogramming may rely on the establishment of a specific DNA methylation signature. This is supported by the observation that many genes in our list are involved in chromatin remodelling (*TRRAP*, *HDAC7*, *PRDMs* and *KDM6B*). Moreover, *DNMT3B* gene itself was found to be differentially methylated in TGF- $\beta$ -treated cells. Therefore, the presence of genes involved in chromatin remodelling and epigenetic mark deposition suggest that TGF- $\beta$  may directly modulate gene expression through DNA methylation but also could act indirectly by regulating expression of chromatin remodelers (through DNA methylation) and that this may further expand TGF- $\beta$ -induced transcription reprogramming through modifying the chromatin context and the gene expression of many other genes.

Detailed analyses revealed that the majority of the CpG were hypermethylated (82%) after TGF- $\beta$  treatment. Globally, the differentially methylated sites were found in open sea regions (1.6 fold enrichment) and this enrichment was even higher in hypomethylated regions (2.1 fold enrichment). In addition, these differentially methylated sites in open seas may be related to gene regulation because many of them are localized in enhancer regions (61%). As for the hypomethylated sites, an 2.1 fold enrichment for enhancer regions was observed and included genes such as *BMP1* and *IRAK2* (involved in inflammatory response). Interestingly these hypomethylated enhancer loci were often localized in the gene body region and the same was observed for hypermethylated sites. For example, hypermethylation within *KDM6B*, *PRDM*, *HDAC*, *TRRAP* and *SKI* genes predominantly takes place in enhancers localized in gene body regions. These observations suggest that many TGF- $\beta$  methyl-sensitive sites are localized in regulatory regions that are not necessarily within CpG islands and promoter regions. This analysis should be further expanded with a genome expression investigation in order to assess the regulatory potential of these CpG sites. Nevertheless, these findings underscore the importance of not restricting DNA methylation analyses to CpG islands and promoter regions.

## **V. Correlation between TGF- $\beta$ induced DNA methylation signature and gene expression.**

We demonstrated that TGF- $\beta$  treatment can induce a stable DNA methylation signature. To further determine if this signature is linked to a specific gene expression profile, we performed a whole genome expression array on Huh7 and HepG2 cell lines. Based on our previous analysis of DNA methylation signature after the release of TGF- $\beta$ , we selected only the last time point (at day8) to monitor gene expression changes linked to permanent epigenetically-induced modifications. We extracted a signature of 1032 significantly deregulated genes that included some known TGF- $\beta$  target genes such as *BMP1*, *BMP4*, *BMP2*, *SMAD6* and *SMAD7*. Interestingly *SMAD6* and *SMAD7*, both of which inhibit the SMAD3/SMAD2 intracellular signal, were both up-regulated and none of the main actors of the TGF- $\beta$  signaling (*TGF- $\beta$ 1*, *TGF- $\beta$ 2*, *TGF- $\beta$ 3*, *SMAD2*, *SMAD3*, *SMAD4*, *TGF- $\beta$ 1*, *TGF- $\beta$ 2* and *TGF- $\beta$ 3*) were found in the signature. It suggests that the TGF- $\beta$  signaling may be shut down after the release of the cytokine and this inactivation could involve SMAD6/SMAD7. This hypothesis is supported by qPCR results that showed up-regulation of TGF- $\beta$  target genes just after the treatment (day 4 time point) but no longer after the release of the cytokine (day 8 time point). Thus TGF- $\beta$  signalling pathway appears to be activated only during the treatment and the gene expression profile obtained would truly reflect the new transcription program stably established after TGF- $\beta$  exposure. In addition, this gene expression signature displays an up-regulation for both *CD133* (*PROM1*) and *DNMT3B*. This validated our previous observations and strengthen the notion that these proteins play an important role in TGF- $\beta$  induced cell reprogramming.

To test whether the DNA methylation changes affect gene expression, we analysed the correlation between the two profiles. We observed that only a small fraction of differentially expressed genes can be attributed to changes in DNA methylation at their loci (26 genes). Several possibilities can explain this observation. First, this gene expression profile is sustained by other epigenetic mechanisms such as histone marks and microRNAs. As previously mentioned, TGF- $\beta$  has already been shown to regulate gene expression through these processes. Second, the gene expression profile we observed may not be directly linked to the DNA methylation signature, but instead may be the outcome of secondary effect of these genes directly targeted by DNA methylation. Indeed, as mentioned earlier, we observed that several DNA methylation changes occur in genes related to chromatin

remodelling and histone modifications and this could in turn modulate the chromatin context for gene transcription.

In addition, the changes in gene expression might be too subtle to be detected in our analysis. Finally, as discussed before, DNA methylation profile, in particular DNA hypomethylation, is not directly linked with gene expression but can constitute an imprint that will condition gene expression regulation in response to future stimulus. Nevertheless, the possibility that those DNA methylation marks are just irrelevant and/or random and do not contribute to the transcriptional profile induced by TGF- $\beta$  should also be considered.

For these genes present in both expression and methylation signatures it is interesting to note that the correlation between DNA methylation and expression is more often positive (17 genes on 26) whereas DNA methylation state is typically inversely correlated with gene expression. The development of new array technologies that interrogate CpG sites across the entire genome and not only on gene promoters rendered the relation between DNA methylation and gene expression more complex (Varley et al., 2013). Whereas DNA methylation on gene promoters is often, yet not always, inversely correlated with gene expression, DNA methylation within gene body is usually positively correlated with gene expression (Maunakea et al., 2010). Concerning the positively correlated genes in our data DNA methylation occurs in gene body (82%), whereas for the negatively correlated genes only 30% display DNA methylation changes in promoter or associated regulatory regions (like 5'UTR). However, for 3 out of 6 inversely correlated genes where the DNA methylation changes occurred in gene body, methyl-sensitive loci are localized in enhancer regions that are also considered as regulatory regions. This highlights the need of not restricting DNA methylation analysis to promoter. In conclusion, the association between CpG site localization, DNA methylation status and gene expression is complex and would require further investigations to fully elucidate the link between DNA methylation and gene expression.

## **VI. TGF- $\beta$ induced DNA methylation contributes to establish the CD133+ CSCs phenotype in liver cancer.**

We demonstrated that CD133+ cells harbour a specific DNA methylation signature, that TGF- $\beta$  can induce transdifferentiation of CD133+ cells, and that this transdifferentiation is accompanied by a methylome reconfiguration. To determine if this TGF- $\beta$  methylome signature is related to the CD133+ phenotype induction, we analyzed the overlap between

the two signatures (CD133+ and TGF- $\beta$ ). Among the 494 annotated genes differentially methylated in CD133+ cells and the 1774 differentially methylated genes after TGF- $\beta$  treatment, a highly significant overlap of 117 genes was observed ( $p < 1.5 \times 10^{-29}$ ). This indicates that CD133+ cells and TGF- $\beta$ -induced CD133+ cells share an important part of their methylome and thus are likely to be phenotypically and functionally similar (as discussed before, the ability of TGF- $\beta$ -induced CD133+ cells to form spheres sustain this hypothesis). Interestingly gene ontology analysis of this common signature highlights the Wnt/ $\beta$ -catenin, mTOR and Notch pathways. These pathways were already revealed independently in the two DNA methylation signatures. The relation between the genes of the common signature to these pathways supports the hypothesis that they are tightly linked to the CD133+ CSC phenotype but also suggests that they could represent the active molecular mechanisms by which TGF- $\beta$  induces CD133+ liver CSCs. More precisely *NOTCH4* (the gene encoding a receptor of the Notch signaling pathway) was present in the common signature, and could represent a good candidate to link liver CSCs and TGF- $\beta$ . In HCC, *NOTCH4* was found to be deregulated (Gao et al., 2008) and has been proposed as a marker for poor prognosis (Ahn et al., 2013). Interestingly, in breast cancer, *NOTCH4* expression appeared to be essential for CSCs maintenance (Harrison et al., 2010; Yu et al., 2012a). In addition *NOTCH4* expression in breast CSCs have also been correlated to EMT marker expression (Yu et al., 2012b) and finally several studies reported possible interaction between the Notch and the TGF- $\beta$  signaling pathways (Sun et al., 2005; Tang et al., 2010). *NOTCH4* expression could thus represent one of the molecular connections between TGF- $\beta$  and CD133+ cells. The exact role of genes found in this common DNA methylation signature requires further investigations.

In spite of this overlap, there is a significant number of differentially methylated sites after TGF- $\beta$  treatment that were not observed in the CD133+ methylome profile. This was confirmed by the linear regression of all our arrays that shows that besides methylation changes imputable to the cell line origin or the CD133 expression, many differentially methylated loci were only related to TGF- $\beta$  treatment (see Materials and Methods for the analysis). In addition to CSC phenotype, TGF- $\beta$  is involved in several other biological processes including differentiation, cell cycle arrest and EMT. These processes are linked to important changes in cell fate decision and thus are sustained by a specific transcriptional program. But as discussed before, the mechanisms underlying these transcriptional programs are not fully elucidated. In addition to histone marks and microRNAs, our analysis

indicates that DNA methylation can play a role in the establishment of TGF- $\beta$ -induced phenotypes.

As we already mentioned, TGF- $\beta$ 's effect strongly depends on the cellular context. Further experiments on isolated CD133+ and CD133- cells should advance our understanding of TGF- $\beta$  mechanisms in different cell subpopulations (to compare precisely the effect on their respective methylome, for example). Furthermore, although TGF- $\beta$ -induced CD133+ cells shared many properties with natural CD133+ cells (*DNMTs* expression, sphere formation ability, differentially methylated genes), these two populations might be not identical. Although several studies have shown that CSCs can be involved in metastatic processes, some observations have raised the possibility that "metastatic CSCs" may differ from the CSCs involved in tumor initiation (Beck and Blanpain, 2013; Visvader and Lindeman, 2012). As TGF- $\beta$  is one of the main factors involved in EMT induction, one might wonder if TGF- $\beta$ -induced CD133+ cells are identical to the initial liver CD133+ cells. The cellular morphology changes observed after TGF- $\beta$  treatment are clearly linked to an EMT phenotype, and could indicate that in our model TGF- $\beta$  does not only induce a CSC phenotype but also promotes EMT. Indeed EMT process is often described as a mechanism of dedifferentiation with re-acquisition of stem cell markers and stem cell phenotype (Eastham et al., 2007; Katsuno et al., 2013; Mani et al., 2008b). This notion may also explain why the two methylation signatures overlap only partially, as TGF- $\beta$ -induced CD133+ cells signature may encompass other specific marks linked to EMT processes. Further experiments such as immunostaining for EMT markers (Vimentin, SNAIL, ZEB1 and TWIST) are necessary to answer this question. In any case it will be interesting to use this two DNA methylation signatures to characterise CD133+ cells in human liver tumor samples and to observe if these cells are more related to one of these two profiles.

## **VII. Further mechanistic studies.**

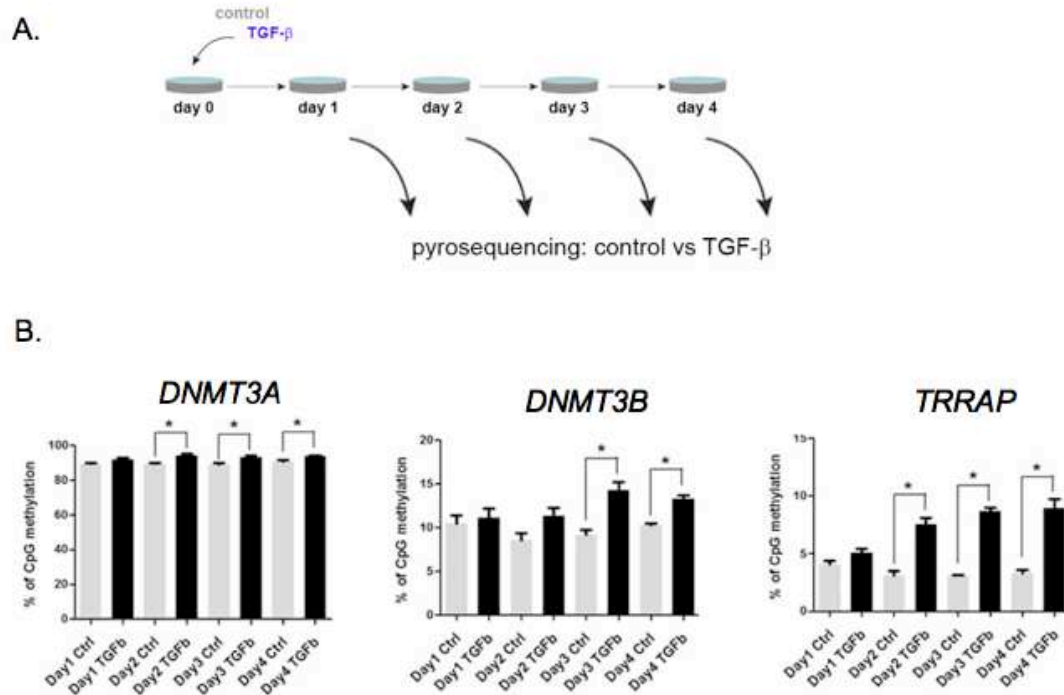
Through two genome wide methylome analyses and a subsequent series of mechanistic studies we provided a strong evidence that CD133+ CSCs in HCC cell lines can be triggered by TGF- $\beta$  and that this relies on a global DNA methylation reprogramming. On the other hand, CD133+ cells signature contains genes that encompass binding sites for members of the SMAD family. Curiously our expression analysis did not reveal any association between DNA methylation and expression of TGF- $\beta$  target genes in CD133+ cells. In addition treatment of Huh7 and HepG2 cells with TGF- $\beta$  inhibitor did not induce any change in



CD133+ population, supporting the hypothesis that the pathway is not activated in natural CD133+ cells. Moreover during CD133+ cells induction by TGF- $\beta$ , we did observe an increase of TGF- $\beta$  target genes during the treatment but not beyond the replacement with fresh medium, indicating that the TGF- $\beta$  pathway is activated only during the period of treatment. This observation argues that TGF- $\beta$  is only necessary for the initiation of the CD133+ phenotype but that once this reprogramming is established, it is transmitted through cell divisions. The imprints related to TGF- $\beta$  observed in CD133+ cells DNA methylation signature may thus represent a past exposure to this cytokine during their initiation. But it can also signify that CD133+ cells possess a favourable epigenetic landscape to efficiently respond to any new exposure to TGF- $\beta$  that would in turn serve to the tumor growth. This suggestion is linked to the global observation that TGF- $\beta$  effect depends on the cell type and cellular context, and that in the tumor mass, CSCs may represent a more sensitive population that will act in synergy with TGF- $\beta$ . However although we demonstrated that DNA methylation plays a role in this CD133+ phenotype induction, we did not investigate to what extent this mechanism is essential for reprogramming. To this end, a treatment with DNMT inhibitor during TGF- $\beta$  induction of CD133+ cells would allow to establish if this transdifferentiation is fully or partially dependant on DNA methylation.

To better understand the mechanisms by how TGF- $\beta$  induces a DNA methylation reprogramming, the links between SMAD binding on the DNA (that represent the terminal step of the TGF- $\beta$  signal) and the DNA methylation machinery should be elucidated. In previous reports describing that TGF- $\beta$  target can be regulated by DNA methylation, the binding of DNA methylation writers/erasers (i.e DNMT3A, DNMT3B, DNA glycosylase) on target gene promoters was observed upon TGF- $\beta$  treatment. However none of these studies investigated how DNA methylation writers were recruited. It is thus essential to determine if these TGF- $\beta$  methyl sensitive regions could be directly recognized by SMAD proteins that will in turn recruit DNA methylation factors.

We performed preliminary experiments (employing pyrosequencing) on Huh7 to determine when the DNA methylation signature was established during the 4 days of TGF- $\beta$  treatment. Our first results on selected genes (Figure 70) suggested that after one day of treatment no change occurred, however after 2 days the DNA methylation profile seems to be almost fully established. These results suggest that DNA methylation is established during a time window of 24h. We can further use this window to perform ChIP for SMAD/ DNMT and TET proteins on selected sites in order to determine the sequence of events that takes place from TGF- $\beta$  signaling activation to DNA methylation.



**Figure 70. DNA methylation changes occur after 2 days of TGF- $\beta$  treatment.**

Huh7 cells were treated as described in A. *DNMT3A*, *DNMT3B* and *TRRAP* methylation profiles were investigated by pyrosequencing after 1, 2, 3 or 4 days of TGF- $\beta$  treatment (10ng/ml) (B).

Mean (+ standart deviation) is shown for three biological replicates. (\*) indicates P value < 0.05.

## CONCLUSIONS



In HCC, CD133+ cells have been intimately linked to cancer stem cells (CSCs). CD133+ cells appear to be involved in the tumor initiation and the tumor growth but so far there has been a lack of studies to understand their molecular characteristics and the mechanisms involved in their maintenance. Because it is known that DNA methylation profile is involved in cell phenotype, this project aimed to characterize the DNA methylation profile of CD133+ cells in liver cancer cell lines and to demonstrate that inflammatory microenvironment-related cytokines can be associated to the mechanisms involved in their initiation/maintenance.

In order to conduct mechanistic studies, we choose to work with *in vitro* models. We demonstrated that CD133 is a marker of distinct subpopulation in two independent HCC cell lines and we established a link between CD133 expression and stemness properties by showing that CD133+ cells express stemness markers and are able to grow in low-attachment conditions.

Thereafter, we explored the epigenetic characteristics of CD133+ cells, focusing our investigations on DNA methylation. We observed that CD133+ cells differentially express genes involved in the DNA methylation machinery (*DNMT* and *TET* proteins) and that CD133+ cells display a distinct DNA methylome linked to specific cellular pathways.

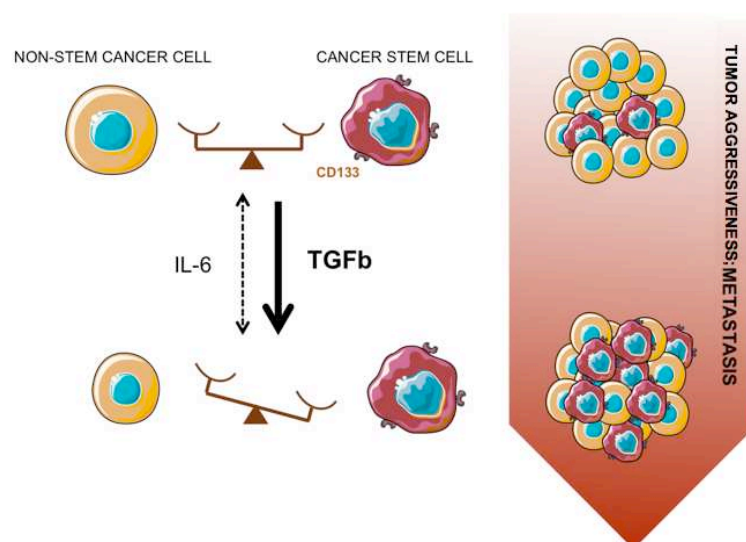
Cellular pathways revealed by CD133+ methylome analysis were notably enriched in inflammatory pathways including the TGF- $\beta$ /SMAD signaling pathway. Subsequently we analyzed the effect of TGF- $\beta$  exposure on CD133+ populations and demonstrated that TGF- $\beta$  was able to induce CD133 expression (at both mRNA and surface protein levels) in HCC cell lines. This induction was stable over cell divisions (in contrast to IL-6) and associated to functional stemness properties (growth in low attachment conditions) as well as dependent on the TGFBR1 receptor signal transmission.

TGF- $\beta$  exposure was also accompanied with an increase in expression for genes involved in DNA methylation machinery. In consequence we explored global DNA methylation changes stably induced by TGF- $\beta$ . We described a unique methylation profile induced by TGF- $\beta$  in HCC cell lines and our analyses revealed that this profile is closely linked to TGF- $\beta$  function and partially explains TGF- $\beta$ -induced gene expression.

Finally comparison between the two DNA methylation signatures (CD133+ and TGF- $\beta$ ) revealed a significant overlap of 117 genes that display links with pathways related to stem cell.

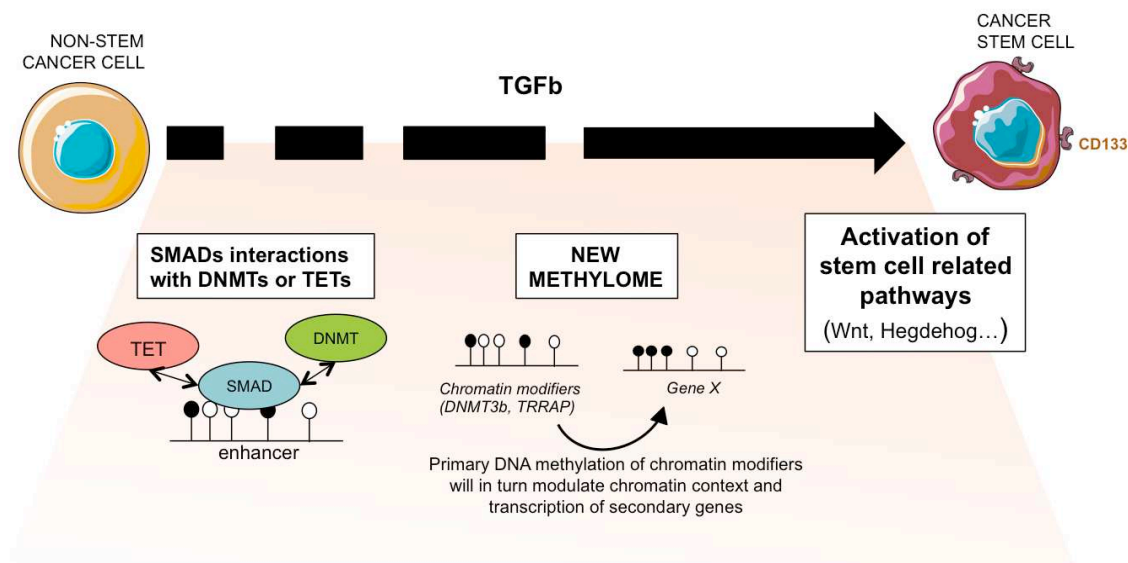
We can propose a model in which a natural balance between non-stem cancer cells (such as CD133-) and cancer stem cells (CD133+) cells exists and where exposition to TGF- $\beta$  would alter this balance in favor of CD133+ cells (Figure 71). Loss of balance results in an increase in CSCs population within the tumor mass which would in turn serve the tumor growth by increasing its aggressiveness and accelerating metastasis.

The second part of the model proposes how TGF- $\beta$  induces cancer stem cells through DNA methylation mechanisms (Figure 72). As DNMT and TET proteins are already known to be recruited on genes after TGF- $\beta$  treatment, we proposed that SMAD binding on regulatory regions (such as enhancer) would participate in DNMT or TET recruitment to establish a new DNA methylation program. This new methylome would further support the establishment of a CSC phenotype through expression of stem-cell related pathways.



**Figure 71. Model for TGF- $\beta$ 's effect on CD133+ CSCs in HCC and its consequence on the tumor development.**

In HCC, a natural homeostasis state is likely to exist within the tumor between CD133- and CD133+ cells. External stimulus, such as cytokines that are released in the microenvironment during inflammation, may alter this balance between non-stem cancer cells and CSCs. For example, our results show that TGF- $\beta$  can stably alter this balance in favor of CSCs (CD133+ cells) in a permanent fashion. Inversely, IL-6 effect's is less strong that TGF- $\beta$  and mostly is not stable (this effect is represented by a dotted double arrow). This switch induced by TGF- $\beta$  in the CSC population could serve the tumor growth by increasing its aggressiveness and favor metastasis development.



**Figure 72. Model for the DNA methylation role in TGF- $\beta$  induction of liver CSCs.**

Our results show that in HCC, TGF- $\beta$  can induce transdifferentiation of non-stem cancer cells (CD133-) into CSCs (CD133+) and can induce a new DNA methylation profile.

This model proposes that DNA methylation could be directly involved in the TGF- $\beta$ -induced initiation of CSCs. Activation of the TGF- $\beta$  signaling pathway would lead to the binding of SMAD proteins on regulating regions (such as enhancers) and could then recruit DNA methylation machinery complexes (including DNMT and TET proteins) to establish a new DNA methylation programming. This DNA methylation profile would sustain a specific genome expression program involving, among others, stem cell related signaling pathways (such as Wnt, Notch and Hedgehog signalling pathways).

In details, DNA methylation and genome expression signature could be set up in two steps: first, SMAD and DNMT/TET interactions would target a primary panel of genes (including chromatin modifiers) for DNA methylation changes; second, transcription deregulations of these chromatin modifiers will in turn modulate the epigenetic profile and the transcription of secondary target genes.

To verify this last part of the model, further mechanistic studies investigating the exact relationship and potential interactions between SMAD and proteins involved in DNA methylation profile establishment are required.

Regarding CD133+ cells epigenetic characterization of the signature provides a reliable database to investigate the exact role of DNA methylation in CSC phenotype establishment and to further identify key genes or pathways involved in CSC maintenance. Finally it will be interesting to compare our results obtained in *in vitro* models with DNA methylation signature from CD133+ cells from liver biopsies. It will allow us to reinforce our conclusions about CD133+ cells molecular characterization and to adapt further research to improve our understanding of CSCs.





## REFERENCES



- Abdel-Wahab, M., Mostafa, M., Sabry, M., el-Farrash, M., and Yousef, T. (2008). Aflatoxins as a risk factor for hepatocellular carcinoma in Egypt, Mansoura Gastroenterology Center study. *Hepatogastroenterology*. 55, 1754–1759.
- Acun, T., Oztas, E., Yagci, T., and Yakicier, M.C. (2011). SIP1 is downregulated in hepatocellular carcinoma by promoter hypermethylation. *BMC Cancer* 11, 223.
- Adams, D.H., and Eksteen, B. (2006). Aberrant homing of mucosal T cells and extra-intestinal manifestations of inflammatory bowel disease. *Nat. Rev. Immunol.* 6, 244–251.
- Adorno, M., Cordenonsi, M., Montagner, M., Dupont, S., Wong, C., Hann, B., Solari, A., Bobisse, S., Rondina, M.B., Guzzardo, V., et al. (2009). A Mutant-p53/Smad complex opposes p63 to empower TGFβ-induced metastasis. *Cell* 137, 87–98.
- Ahn, S., Hyeon, J., and Park, C.-K. (2013). Notch1 and Notch4 are markers for poor prognosis of hepatocellular carcinoma. *Hepatobiliary Pancreat. Dis. Int. HBPD INT* 12, 286–294.
- Akhurst, R.J., and Hata, A. (2012). Targeting the TGFβ signalling pathway in disease. *Nat. Rev. Drug Discov.* 11, 790–811.
- Akhurst, B., Matthews, V., Husk, K., Smyth, M.J., Abraham, L.J., and Yeoh, G.C. (2005). Differential lymphotoxin-beta and interferon gamma signaling during mouse liver regeneration induced by chronic and acute injury. *Hepatology*. 41, 327–335.
- Akool, E.-S., Doller, A., Müller, R., Gutwein, P., Xin, C., Huwiler, A., Pfeilschifter, J., and Eberhardt, W. (2005). Nitric oxide induces TIMP-1 expression by activating the transforming growth factor beta-Smad signaling pathway. *J. Biol. Chem.* 280, 39403–39416.
- Alison, M. (1998). Liver stem cells: a two compartment system. *Curr. Opin. Cell Biol.* 10, 710–715.
- Alison, M.R. (2005). Liver stem cells: implications for hepatocarcinogenesis. *Stem Cell Rev.* 1, 253–260.
- Alison, M.R., Islam, S., and Lim, S. (2009). Stem cells in liver regeneration, fibrosis and cancer: the good, the bad and the ugly. *J. Pathol.* 217, 282–298.
- Ammerpohl, O., Pratschke, J., Schafmayer, C., Haake, A., Faber, W., von Kampen, O., Brosch, M., Sipos, B., von Schönfels, W., Balschun, K., et al. (2012). Distinct DNA methylation patterns in cirrhotic liver and hepatocellular carcinoma. *Int. J. Cancer* 130, 1319–1328.
- Anest, V., Hanson, J.L., Cogswell, P.C., Steinbrecher, K.A., Strahl, B.D., and Baldwin, A.S. (2003). A nucleosomal function for IκB kinase-α in NF-κB-dependent gene expression. *Nature* 423, 659–663.
- Aroucha, D.C.B.L., do Carmo, R.F., Moura, P., Silva, J.L.A., Vasconcelos, L.R.S., Cavalcanti, M.S.M., Muniz, M.T.C., Aroucha, M.L., Siqueira, E.R.F., Cahú, G.G.O.M., et al. (2013). High tumor necrosis factor-α/interleukin-10 ratio is associated with hepatocellular carcinoma in patients with chronic hepatitis C. *Cytokine* 62, 421–425.
- Azechi, H., Nishida, N., Fukuda, Y., Nishimura, T., Minata, M., Katsuma, H., Kuno, M., Ito, T., Komeda, T., Kita, R., et al. (2001). Disruption of the p16/cyclin D1/retinoblastoma protein pathway in the majority of human hepatocellular carcinomas. *Oncology* 60, 346–354.
- Baba, T., Convery, P.A., Matsumura, N., Whitaker, R.S., Kondoh, E., Perry, T., Huang, Z., Bentley, R.C., Mori, S., Fujii, S., et al. (2009). Epigenetic regulation of CD133 and tumorigenicity of CD133+ ovarian cancer cells. *Oncogene* 28, 209–218.
- Banerjee, S., and Bacanamwo, M. (2010). DNA methyltransferase inhibition induces mouse embryonic stem cell differentiation into endothelial cells. *Exp. Cell Res.* 316, 172–180.
- Bao, B., Azmi, A.S., Li, Y., Ahmad, A., Ali, S., Banerjee, S., Kong, D., and Sarkar, F.H. (2013). Targeting CSCs in Tumor Microenvironment: The Potential Role of ROS-Associated miRNAs in Tumor Aggressiveness. *Curr. Stem Cell Res. Ther.*
- Bar, E.E., Lin, A., Mahairaki, V., Matsui, W., and Eberhart, C.G. (2010). Hypoxia increases the expression of stem-cell markers and promotes clonogenicity in glioblastoma neurospheres. *Am. J. Pathol.* 177, 1491–1502.
- Barroca, V., Lassalle, B., Coureuil, M., Louis, J.P., Le Page, F., Testart, J., Allemand, I., Riou, L., and Fouchet, P. (2009). Mouse differentiating spermatogonia can generate germinal stem cells in vivo. *Nat. Cell Biol.* 11, 190–196.

- Bartsch, H., and Nair, J. (2006). Chronic inflammation and oxidative stress in the genesis and perpetuation of cancer: role of lipid peroxidation, DNA damage, and repair. *Langenbecks Arch. Surg. Dtsch. Ges. Für Chir.* 391, 499–510.
- Baylin, S.B. (2005). DNA methylation and gene silencing in cancer. *Nat. Clin. Pract. Oncol.* 2 Suppl 1, S4–11.
- Baylin, S.B., Esteller, M., Rountree, M.R., Bachman, K.E., Schuebel, K., and Herman, J.G. (2001). Aberrant patterns of DNA methylation, chromatin formation and gene expression in cancer. *Hum. Mol. Genet.* 10, 687–692.
- Beck, B., and Blanpain, C. (2013). Unravelling cancer stem cell potential. *Nat. Rev. Cancer* 13, 727–738.
- Beck, B., Driessens, G., Goossens, S., Youssef, K.K., Kuchnio, A., Caauwe, A., Sotiropoulou, P.A., Loges, S., Lapouge, G., Candi, A., et al. (2011). A vascular niche and a VEGF-Nrp1 loop regulate the initiation and stemness of skin tumours. *Nature* 478, 399–403.
- Bergman, Y., and Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nat. Struct. Mol. Biol.* 20, 274–281.
- Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., et al. (2011). High density DNA methylation array with single CpG site resolution. *Genomics* 98, 288–295.
- Billottet, C., Tuefferd, M., Gentien, D., Rapinat, A., Thiery, J.-P., Broët, P., and Jouanneau, J. (2008). Modulation of several waves of gene expression during FGF-1 induced epithelial-mesenchymal transition of carcinoma cells. *J. Cell. Biochem.* 104, 826–839.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.
- Bird, A.P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* 321, 209–213.
- De Bleser, P.J., Niki, T., Rogiers, V., and Geerts, A. (1997). Transforming growth factor-beta gene expression in normal and fibrotic rat liver. *J. Hepatol.* 26, 886–893.
- Blick, T., Hugo, H., Widodo, E., Waltham, M., Pinto, C., Mani, S.A., Weinberg, R.A., Neve, R.M., Lenburg, M.E., and Thompson, E.W. (2010). Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi)/CD24 (lo/-) stem cell phenotype in human breast cancer. *J. Mammary Gland Biol. Neoplasia* 15, 235–252.
- Bonnet, D., and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* 3, 730–737.
- Boron, W.F., and Boulpaep, E.L. (2008). *Medical Physiology* (Elsevier Health Sciences).
- Bortolami, M., Venturi, C., Giacomelli, L., Scalerta, R., Bacchetti, S., Marino, F., Floreani, A., Lise, M., Naccarato, R., and Farinati, F. (2002). Cytokine, infiltrating macrophage and T cell-mediated response to development of primary and secondary human liver cancer. *Dig. Liver Dis. Off. J. Ital. Soc. Gastroenterol. Ital. Assoc. Study Liver* 34, 794–801.
- Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760–1764.
- Bouwens, L., and Wisse, E. (1992). Pit cells in the liver. *Liver* 12, 3–9.
- Bressac, B., Kew, M., Wands, J., and Ozturk, M. (1991). Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa. *Nature* 350, 429–431.
- Breuhahn, K., Longerich, T., and Schirmacher, P. (2006). Dysregulation of growth factor signaling in human hepatocellular carcinoma. *Oncogene* 25, 3787–3800.
- Brooling, J.T., Campbell, J.S., Mitchell, C., Yeoh, G.C., and Fausto, N. (2005). Differential regulation of rodent hepatocyte and oval cell proliferation by interferon gamma. *Hepatol. Baltim. Md* 41, 906–915.
- Bruna, A., Darken, R.S., Rojo, F., Ocaña, A., Peñuelas, S., Arias, A., Paris, R., Tortosa, A., Mora, J., Baselga, J., et al. (2007). High TGFbeta-Smad activity confers poor prognosis in glioma patients and promotes cell proliferation depending on the methylation of the PDGF-B gene. *Cancer Cell* 11, 147–160.
- Bruniquel, D., and Schwartz, R.H. (2003). Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process. *Nat. Immunol.* 4, 235–240.
- Budhu, A., and Wang, X.W. (2006). The role of cytokines in hepatocellular carcinoma. *J. Leukoc. Biol.* 80, 1197–

- Buendia, M.A. (2000). Genetics of hepatocellular carcinoma. *Semin. Cancer Biol.* 10, 185–200.
- Bussolati, B., Bruno, S., Grange, C., Buttiglieri, S., Deregis, M.C., Cantino, D., and Camussi, G. (2005). Isolation of renal progenitor cells from adult human kidney. *Am. J. Pathol.* 166, 545–555.
- But, D.-Y.-K., Lai, C.-L., and Yuen, M.-F. (2008). Natural history of hepatitis-related hepatocellular carcinoma. *World J. Gastroenterol.* WJG 14, 1652–1656.
- Cai, J., Schleidt, S., Pelta-Heller, J., Hutchings, D., Cannarsa, G., and Iacovitti, L. (2013). BMP and TGF- $\beta$  pathway mediators are critical upstream regulators of Wnt signaling during midbrain dopamine differentiation in human pluripotent stem cells. *Dev. Biol.* 376, 62–73.
- Caja, L., Ortiz, C., Bertran, E., Murillo, M.M., Miró-Obradors, M.J., Palacios, E., and Fabregat, I. (2007). Differential intracellular signalling induced by TGF- $\beta$  in rat adult hepatocytes and hepatoma cells: implications in liver carcinogenesis. *Cell. Signal.* 19, 683–694.
- Calabrese, C., Poppleton, H., Kocak, M., Hogg, T.L., Fuller, C., Hamner, B., Oh, E.Y., Gaber, M.W., Finklestein, D., Allen, M., et al. (2007). A perivascular niche for brain tumor stem cells. *Cancer Cell* 11, 69–82.
- Callahan, J.F., Burgess, J.L., Fornwald, J.A., Gaster, L.M., Harling, J.D., Harrington, F.P., Heer, J., Kwon, C., Lehr, R., Mathur, A., et al. (2002). Identification of novel inhibitors of the transforming growth factor  $\beta$ 1 (TGF- $\beta$ 1) type 1 receptor (ALK5). *J. Med. Chem.* 45, 999–1001.
- Calvisi, D.F., Ladu, S., Gorden, A., Farina, M., Conner, E.A., Lee, J.-S., Factor, V.M., and Thorgeirsson, S.S. (2006). Ubiquitous activation of Ras and Jak/Stat pathways in human HCC. *Gastroenterology* 130, 1117–1128.
- Cao, L., Shao, M., Schilder, J., Guise, T., Mohammad, K.S., and Matei, D. (2012). Tissue transglutaminase links TGF- $\beta$ , epithelial to mesenchymal transition and a stem cell phenotype in ovarian cancer. *Oncogene* 31, 2521–2534.
- Castaño, Z., Fillmore, C.M., Kim, C.F., and McAllister, S.S. (2012). The bed and the bugs: interactions between the tumor microenvironment and cancer stem cells. *Semin. Cancer Biol.* 22, 462–470.
- Castilla, A., Prieto, J., and Fausto, N. (1991). Transforming growth factors  $\beta$  1 and  $\alpha$  in chronic liver disease. Effects of interferon  $\alpha$  therapy. *N. Engl. J. Med.* 324, 933–940.
- Cavin, L.G., Romieu-Mourez, R., Panta, G.R., Sun, J., Factor, V.M., Thorgeirsson, S.S., Sonenshein, G.E., and Arsura, M. (2003). Inhibition of CK2 activity by TGF- $\beta$ 1 promotes I $\kappa$ B- $\alpha$  protein stabilization and apoptosis of immortalized hepatocytes. *Hepatology* 38, 1540–1551.
- Chaffer, C.L., Marjanovic, N.D., Lee, T., Bell, G., Kleer, C.G., Reinhardt, F., D'Alessio, A.C., Young, R.A., and Weinberg, R.A. (2013). Poised Chromatin at the ZEB1 Promoter Enables Breast Cancer Cell Plasticity and Enhances Tumorigenicity. *Cell* 154, 61–74.
- Challen, C., Guo, K., Collier, J.D., Cavanagh, D., and Bassendine, M.F. (1992). Infrequent point mutations in codons 12 and 61 of ras oncogenes in human hepatocellular carcinomas. *J. Hepatol.* 14, 342–346.
- Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* 44, 23–31.
- Charles, N., Ozawa, T., Squatrito, M., Bleau, A.-M., Brennan, C.W., Hambardzumyan, D., and Holland, E.C. (2010). Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* 6, 141–152.
- Chau, G.Y., Wu, C.W., Lui, W.Y., Chang, T.J., Kao, H.L., Wu, L.H., King, K.L., Loong, C.C., Hsia, C.Y., and Chi, C.W. (2000). Serum interleukin-10 but not interleukin-6 is related to clinical outcome in patients with resectable hepatocellular carcinoma. *Ann. Surg.* 231, 552–558.
- Chedin, F., Lieber, M.R., and Hsieh, C.-L. (2002). The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc. Natl. Acad. Sci. U. S. A.* 99, 16916–16921.
- Chen, Z., and Riggs, A.D. (2011). DNA methylation and demethylation in mammals. *J. Biol. Chem.* 286, 18347–18353.
- Chen, C., Pan, D., Deng, A.-M., Huang, F., Sun, B.-L., and Yang, R.-G. (2013a). DNA methyltransferases 1 and 3B are required for hepatitis C virus infection in cell culture. *Virology* 441, 57–65.
- Chen, C.-C., Wang, K.-Y., and Shen, C.-K.J. (2012a). The mammalian de novo DNA methyltransferases DNMT3A

- and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J. Biol. Chem.* 287, 33116–33121.
- Chen, Y., Yu, D., Zhang, H., He, H., Zhang, C., Zhao, W., and Shao, R.-G. (2012b). CD133(+)EpCAM(+) phenotype possesses more characteristics of tumor initiating cells in hepatocellular carcinoma Huh7 cells. *Int. J. Biol. Sci.* 8, 992–1004.
- Chen, Z., Pan, X., Yao, Y., Yan, F., Chen, L., Huang, R., and Ma, G. (2013b). Epigenetic Regulation of Cardiac Progenitor Cells Marker c-kit by Stromal Cell Derived Factor-1 $\alpha$ . *PloS One* 8, e69134.
- Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA methyltransferases: a structural perspective. *Struct. Lond. Engl.* 1993 16, 341–350.
- Cheng, A.S.L., Culhane, A.C., Chan, M.W.Y., Venkataramu, C.R., Ehrich, M., Nasir, A., Rodriguez, B.A.T., Liu, J., Yan, P.S., Quackenbush, J., et al. (2008). Epithelial progeny of estrogen-exposed breast progenitor cells display a cancer-like methylome. *Cancer Res.* 68, 1786–1796.
- Cheung, S.T., Cheung, P.F.Y., Cheng, C.K.C., Wong, N.C.L., and Fan, S.T. (2011). Granulin-epithelin precursor and ATP-dependent binding cassette (ABC)B5 regulate liver cancer cell chemoresistance. *Gastroenterology* 140, 344–355.
- Chiba, T., Kita, K., Zheng, Y.-W., Yokosuka, O., Saisho, H., Iwama, A., Nakauchi, H., and Taniguchi, H. (2006). Side population purified from hepatocellular carcinoma cells harbors cancer stem cell-like properties. *Hepatology* 44, 240–251.
- Chiba, T., Miyagi, S., Saraya, A., Aoki, R., Seki, A., Morita, Y., Yonemitsu, Y., Yokosuka, O., Taniguchi, H., Nakauchi, H., et al. (2008). The polycomb gene product BMI1 contributes to the maintenance of tumor-initiating side population cells in hepatocellular carcinoma. *Cancer Res.* 68, 7742–7749.
- Choi, S., and Diehl, A.M. (2005). Role of inflammation in nonalcoholic steatohepatitis. *Curr. Opin. Gastroenterol.* 21, 702–707.
- Choi, M.S., Shim, Y.-H., Hwa, J.Y., Lee, S.K., Ro, J.Y., Kim, J.-S., and Yu, E. (2003). Expression of DNA methyltransferases in multistep hepatocarcinogenesis. *Hum. Pathol.* 34, 11–17.
- Christman, J.K. (1995). Dietary effects on DNA methylation: do they account for the hepatocarcinogenic properties of lipotrope deficient diets? *Adv. Exp. Med. Biol.* 369, 141–154.
- Chu, P., Clanton, D.J., Snipas, T.S., Lee, J., Mitchell, E., Nguyen, M.-L., Hare, E., and Peach, R.J. (2009). Characterization of a subpopulation of colon cancer cells with stem cell-like properties. *Int. J. Cancer* 124, 1312–1321.
- Ciurtin, C., and Stoica, V. (2008). Hepatitis virus C infection, adipokines and hepatic steato-fibrosis. *J. Med. Life* 1, 49–54.
- Clark, S.J., Harrison, J., and Frommer, M. (1995). CpNpG methylation in mammalian cells. *Nat. Genet.* 10, 20–27.
- Coleman, W.B., and Tsongalis, G.J. (2009). *Molecular Pathology: The Molecular Basis of Human Disease* (Academic Press).
- Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J., and Maitland, N.J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res.* 65, 10946–10951.
- Condeelis, J., and Pollard, J.W. (2006). Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell* 124, 263–266.
- Cooper, G.M. (2000). *Regulation of Transcription in Eukaryotes*.
- Corbeil, D., Röper, K., Fargeas, C.A., Joester, A., and Huttner, W.B. (2001). Prominin: a story of cholesterol, plasma membrane protrusions and human pathology. *Traffic Cph. Den.* 2, 82–91.
- Coskun, U., Bukan, N., Sancak, B., Günel, N., Ozenirler, S., Unal, A., and Yucel, A. (2004). Serum hepatocyte growth factor and interleukin-6 levels can distinguish patients with primary or metastatic liver tumors from those with benign liver lesions. *Neoplasma* 51, 209–213.
- Cougot, D., Neuveut, C., and Buendia, M.A. (2005). HBV induced carcinogenesis. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 34 Suppl 1, S75–78.
- Couronné, L., Bastard, C., and Bernard, O.A. (2012). TET2 and DNMT3A mutations in human T-cell lymphoma. *N. Engl. J. Med.* 366, 95–96.

- D'Anello, L., Sansone, P., Storci, G., Mitrugno, V., D'Uva, G., Chieco, P., and Bonafé, M. (2010). Epigenetic control of the basal-like gene expression profile via Interleukin-6 in breast cancer cells. *Mol. Cancer* 9, 300.
- Dalerba, P., Dylla, S.J., Park, I.-K., Liu, R., Wang, X., Cho, R.W., Hoey, T., Gurney, A., Huang, E.H., Simeone, D.M., et al. (2007). Phenotypic characterization of human colorectal cancer stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 104, 10158–10163.
- Deplus, R., Brenner, C., Burgers, W.A., Putmans, P., Kouzarides, T., de Launoit, Y., and Fuks, F. (2002). Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res.* 30, 3831–3838.
- Dethlefsen, C., Højfeldt, G., and Hojman, P. (2013). The role of intratumoral and systemic IL-6 in breast cancer. *Breast Cancer Res. Treat.* 138, 657–664.
- Devarajan, E., Song, Y.-H., Krishnappa, S., and Alt, E. (2012). Epithelial-mesenchymal transition in breast cancer lines is mediated through PDGF-D released by tissue-resident stem cells. *Int. J. Cancer J. Int. Cancer* 131, 1023–1031.
- Diehn, M., Cho, R.W., Lobo, N.A., Kalisky, T., Dorie, M.J., Kulp, A.N., Qian, D., Lam, J.S., Ailles, L.E., Wong, M., et al. (2009). Association of Reactive Oxygen Species Levels and Radioresistance in Cancer Stem Cells. *Nature* 458, 780–783.
- Ding, B.-S., Nolan, D.J., Guo, P., Babazadeh, A.O., Cao, Z., Rosenwaks, Z., Crystal, R.G., Simons, M., Sato, T.N., Worgall, S., et al. (2011). Endothelial-derived angiocrine signals induce and sustain regenerative lung alveolarization. *Cell* 147, 539–553.
- Ding, W., Mouzaki, M., You, H., Laird, J.C., Mato, J., Lu, S.C., and Rountree, C.B. (2009). CD133+ liver cancer stem cells from methionine adenosyl transferase 1A-deficient mice demonstrate resistance to transforming growth factor (TGF)-beta-induced apoptosis. *Hepatology* 49, 1277–1286.
- Doerfler, W. (1983). DNA methylation and gene activity. *Annu. Rev. Biochem.* 52, 93–124.
- Dong, C., Wu, Y., Yao, J., Wang, Y., Yu, Y., Rychahou, P.G., Evers, B.M., and Zhou, B.P. (2012). G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J. Clin. Invest.* 122, 1469–1486.
- Dooley, S., and ten Dijke, P. (2012). TGF- $\beta$  in progression of liver disease. *Cell Tissue Res.* 347, 245–256.
- Dooley, S., Hamzavi, J., Ciucan, L., Godoy, P., Ilkavets, I., Ehnert, S., Ueberham, E., Gebhardt, R., Kanzler, S., Geier, A., et al. (2008). Hepatocyte-specific Smad7 expression attenuates TGF-beta-mediated fibrogenesis and protects against liver damage. *Gastroenterology* 135, 642–659.
- Dricu, A., Oana, S., Sandra, A., Elise, D., Daianu, O., Stoleru, B., Mihaela, A., Daianu, T., and Gabriela, L. (2012). DNA Methylation, Stem Cells and Cancer. In *Methylation - From DNA, RNA and Histones to Diseases and Treatment*, A. Dricu, ed. (InTech),.
- Dumont, N., Wilson, M.B., Crawford, Y.G., Reynolds, P.A., Sigaroudinia, M., and Tlsty, T.D. (2008). Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. *Proc. Natl. Acad. Sci. U. S. A.* 105, 14867–14872.
- Duncan, A.W., Dorrell, C., and Grompe, M. (2009). Stem cells and liver regeneration. *Gastroenterology* 137, 466–481.
- Dunphy, K.A., Seo, J.-H., Kim, D.J., Roberts, A.L., Tao, L., Drenzo, J., Balboni, A.L., Crisi, G.M., Hagen, M.J., Chandrasekaran, T., et al. (2013). Oncogenic transformation of mammary epithelial cells by transforming growth factor beta independent of mammary stem cell regulation. *Cancer Cell Int.* 13, 74.
- Eades, G., Yao, Y., Yang, M., Zhang, Y., Chumsri, S., and Zhou, Q. (2011). miR-200a regulates SIRT1 expression and epithelial to mesenchymal transition (EMT)-like transformation in mammary epithelial cells. *J. Biol. Chem.* 286, 25992–26002.
- Eads, C.A., Danenberg, K.D., Kawakami, K., Saltz, L.B., Danenberg, P.V., and Laird, P.W. (1999). CpG Island Hypermethylation in Human Colorectal Tumors Is Not Associated with DNA Methyltransferase Overexpression. *Cancer Res.* 59, 2302–2306.
- Eastham, A.M., Spencer, H., Soncin, F., Ritson, S., Merry, C.L.R., Stern, P.L., and Ward, C.M. (2007). Epithelial-mesenchymal transition events during human embryonic stem cell differentiation. *Cancer Res.* 67, 11254–11262.
- Egger, G., Liang, G., Aparicio, A., and Jones, P.A. (2004). Epigenetics in human disease and prospects for

- epigenetic therapy. *Nature* 429, 457–463.
- Egger, G., Jeong, S., Escobar, S.G., Cortez, C.C., Li, T.W.H., Saito, Y., Yoo, C.B., Jones, P.A., and Liang, G. (2006). Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl. Acad. Sci. U. S. A.* 103, 14080–14085.
- Ehrlich, M., Woods, C.B., Yu, M.C., Dubeau, L., Yang, F., Campan, M., Weisenberger, D.J., Long, T., Youn, B., Fiala, E.S., et al. (2006). Quantitative analysis of associations between DNA hypermethylation, hypomethylation, and DNMT RNA levels in ovarian tumors. *Oncogene* 25, 2636–2645.
- El-Serag, H.B., and Rudolph, K.L. (2007). Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 132, 2557–2576.
- El-Zayadi, A.-R. (2008). Hepatic steatosis: a benign disease or a silent killer. *World J. Gastroenterol. WJG* 14, 4120–4126.
- Elgin, S.C.R., and Grewal, S.I.S. (2003). Heterochromatin: silence is golden. *Curr. Biol. CB* 13, R895–898.
- Elliott, J. (2008). SOCS3 in liver regeneration and hepatocarcinoma. *Mol. Interv.* 8, 19–21, 2.
- Factor, V.M., Kao, C.Y., Santoni-Rugiu, E., Voitach, J.T., Jensen, M.R., and Thorgeirsson, S.S. (1997). Constitutive expression of mature transforming growth factor beta1 in the liver accelerates hepatocarcinogenesis in transgenic mice. *Cancer Res.* 57, 2089–2095.
- Fainboim, L., Chervashvsky, A., Paladino, N., Flores, A.C., and Arruvito, L. (2007). Cytokines and chronic liver disease. *Cytokine Growth Factor Rev.* 18, 143–157.
- Falasca, K., Ucciferri, C., Dalessandro, M., Zingariello, P., Mancino, P., Petrarca, C., Pizzigallo, E., Conti, P., and Vecchiet, J. (2006). Cytokine patterns correlate with liver damage in patients with chronic hepatitis B and C. *Ann. Clin. Lab. Sci.* 36, 144–150.
- Fares, N., and Péron, J.-M. (2013). [Epidemiology, natural history, and risk factors of hepatocellular carcinoma]. *Rev. Prat.* 63, 216–217, 220–222.
- Fargeas, C.A., Huttner, W.B., and Corbeil, D. (2007). Nomenclature of prominin-1 (CD133) splice variants - an update. *Tissue Antigens* 69, 602–606.
- Farthing, C.R., Ficz, G., Ng, R.K., Chan, C.-F., Andrews, S., Dean, W., Hemberger, M., and Reik, W. (2008). Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. *PLoS Genet.* 4, e1000116.
- Fausto, N. (2000). Liver regeneration. *J. Hepatol.* 32, 19–31.
- Feinberg, A.P., Ohlsson, R., and Henikoff, S. (2006). The epigenetic progenitor origin of human cancer. *Nat. Rev. Genet.* 7, 21–33.
- Feo, F., Frau, M., Tomasi, M.L., Brozzetti, S., and Pascale, R.M. (2009). Genetic and epigenetic control of molecular alterations in hepatocellular carcinoma. *Exp. Biol. Med.* Maywood NJ 234, 726–736.
- Ferlay, J., Shin, H.-R., Bray, F., Forman, D., Mathers, C., and Parkin, D.M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer J. Int. Cancer* 127, 2893–2917.
- Filion, G.J.P., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E., and Defossez, P.-A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol. Cell. Biol.* 26, 169–181.
- Fillmore, C.M., Gupta, P.B., Rudnick, J.A., Caballero, S., Keller, P.J., Lander, E.S., and Kuperwasser, C. (2010). Estrogen expands breast cancer stem-like cells through paracrine FGF/Tbx3 signaling. *Proc. Natl. Acad. Sci. U. S. A.* 107, 21737–21742.
- Flavell, R.A., Sanjabi, S., Wrzesinski, S.H., and Licona-Limón, P. (2010). The polarization of immune cells in the tumour environment by TGFbeta. *Nat. Rev. Immunol.* 10, 554–567.
- Franco, D.L., Mainez, J., Vega, S., Sancho, P., Murillo, M.M., de Frutos, C.A., Del Castillo, G., López-Blau, C., Fabregat, I., and Nieto, M.A. (2010). Snail1 suppresses TGF-beta-induced apoptosis and is sufficient to trigger EMT in hepatocytes. *J. Cell Sci.* 123, 3467–3477.
- Fujimoto, A., Totoki, Y., Abe, T., Borojevich, K.A., Hosoda, F., Nguyen, H.H., Aoki, M., Hosono, N., Kubo, M., Miya, F., et al. (2012). Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* 44, 760–764.



- Furukawa, F., Matsuzaki, K., Mori, S., Tahashi, Y., Yoshida, K., Sugano, Y., Yamagata, H., Matsushita, M., Seki, T., Inagaki, Y., et al. (2003). p38 MAPK mediates fibrogenic signal through Smad3 phosphorylation in rat myofibroblasts. *Hepatology*. Baltimore, Md 38, 879–889.
- Gao, B. (2005). Cytokines, STATs and liver disease. *Cell. Mol. Immunol.* 2, 92–100.
- Gao, J., Song, Z., Chen, Y., Xia, L., Wang, J., Fan, R., Du, R., Zhang, F., Hong, L., Song, J., et al. (2008). Deregulated expression of Notch receptors in human hepatocellular carcinoma. *Dig. Liver Dis. Off. J. Ital. Soc. Gastroenterol. Ital. Assoc. Study Liver* 40, 114–121.
- Gao, X.-D., Qu, J.-H., Chang, X.-J., Lu, Y.-Y., Bai, W.-L., Wang, H., Xu, Z.-X., An, L.-J., Wang, C.-P., Zeng, Z., et al. (2013a). Hypomethylation of long interspersed nuclear element-1 promoter is associated with poor outcomes for curative resected hepatocellular carcinoma. *Liver Int. Off. J. Int. Assoc. Study Liver*.
- Gao, Y., Chen, J., Li, K., Wu, T., Huang, B., Liu, W., Kou, X., Zhang, Y., Huang, H., Jiang, Y., et al. (2013b). Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* 12, 453–469.
- García-Galiano, D., Sánchez-Garrido, M.A., Espejo, I., Montero, J.L., Costán, G., Marchal, T., Membrives, A., Gallardo-Valverde, J.M., Muñoz-Castañeda, J.R., Arévalo, E., et al. (2007). IL-6 and IGF-1 are independent prognostic factors of liver steatosis and non-alcoholic steatohepatitis in morbidly obese patients. *Obes. Surg.* 17, 493–503.
- Gatti, L., Cossa, G., Beretta, G.L., Zaffaroni, N., and Perego, P. (2011). Novel insights into targeting ATP-binding cassette transporters for antitumor therapy. *Curr. Med. Chem.* 18, 4237–4249.
- Gialeli, C., Nikitovic, D., Kletsas, D., Theocharis, A.D., Tzanakakis, G.N., and Karamanos, N.K. (2013). PDGF/PDGFR Signaling and Targeting in Cancer Growth and Progression: Focus on Tumor Microenvironment and Cancer-Associated Fibroblasts. *Curr. Pharm. Des.*
- Gilbertson, R.J., and Rich, J.N. (2007). Making a tumour's bed: glioblastoma stem cells and the vascular niche. *Nat. Rev. Cancer* 7, 733–736.
- Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.-L., Zhang, X., Golic, K.G., Jacobsen, S.E., and Bestor, T.H. (2006). Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science* 311, 395–398.
- Gonda, T.A., Tu, S., and Wang, T.C. (2009). Chronic inflammation, the tumor microenvironment and carcinogenesis. *Cell Cycle Georgetown, Tex* 8, 2005–2013.
- Gotzmann, J., Huber, H., Thallinger, C., Wolschek, M., Jansen, B., Schulte-Hermann, R., Beug, H., and Mikulits, W. (2002). Hepatocytes convert to a fibroblastoid phenotype through the cooperation of TGF-beta1 and Ha-Ras: steps towards invasiveness. *J. Cell Sci.* 115, 1189–1202.
- Grivennikov, S.I., and Karin, M. (2010). Inflammation and oncogenesis: a vicious connection. *Curr. Opin. Genet. Dev.* 20, 65–71.
- Grivennikov, S.I., Greten, F.R., and Karin, M. (2010). Immunity, inflammation, and cancer. *Cell* 140, 883–899.
- Grønbaek, K., Hother, C., and Jones, P.A. (2007). Epigenetic changes in cancer. *APMIS Acta Pathol. Microbiol. Immunol. Scand.* 115, 1039–1059.
- Grosse-Gehling, P., Fargeas, C.A., Dittfeld, C., Garbe, Y., Alison, M.R., Corbeil, D., and Kunz-Schughart, L.A. (2013). CD133 as a biomarker for putative cancer stem cells in solid tumours: limitations, problems and challenges. *J. Pathol.* 229, 355–378.
- Grotenhuis, B.A., Wijnhoven, B.P.L., and van Lanschot, J.J.B. (2012). Cancer stem cells and their potential implications for the treatment of solid tumors. *J. Surg. Oncol.* 106, 209–215.
- Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I.B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., et al. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat. Genet.* 44, 694–698.
- Guo, X., Liu, Q., Wang, G., Zhu, S., Gao, L., Hong, W., Chen, Y., Wu, M., Liu, H., Jiang, C., et al. (2013). microRNA-29b is a novel mediator of Sox2 function in the regulation of somatic cell reprogramming. *Cell Res.* 23, 142–156.
- Gurtsevitch, V.E. (2008). Human oncogenic viruses: hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis. *Biochem. Biokhimiia* 73, 504–513.

- Hagiwara, S., Kudo, M., Nagai, T., Inoue, T., Ueshima, K., Nishida, N., Watanabe, T., and Sakurai, T. (2012). Activation of JNK and high expression level of CD133 predict a poor response to sorafenib in hepatocellular carcinoma. *Br. J. Cancer* 106, 1997–2003.
- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* 100, 3983–3988.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* 117, 15–23.
- Hamilton, J.P. (2010). Epigenetic mechanisms involved in the pathogenesis of hepatobiliary malignancies. *Epigenomics* 2, 233–243.
- Hannigan, A., Smith, P., Kalna, G., Lo Nigro, C., Orange, C., O'Brien, D.I., Shah, R., Syed, N., Spender, L.C., Herrera, B., et al. (2010). Epigenetic downregulation of human disabled homolog 2 switches TGF-beta from a tumor suppressor to a tumor promoter. *J. Clin. Invest.* 120, 2842–2857.
- Haraguchi, N., Ishii, H., Mimori, K., Tanaka, F., Ohkuma, M., Kim, H.M., Akita, H., Takiuchi, D., Hatano, H., Nagano, H., et al. (2010). CD13 is a therapeutic target in human liver cancer stem cells. *J. Clin. Invest.* 120, 3326–3339.
- Harrison, H., Farnie, G., Howell, S.J., Rock, R.E., Stylianou, S., Brennan, K.R., Bundred, N.J., and Clarke, R.B. (2010). Regulation of breast cancer stem cell activity by signaling through the Notch4 receptor. *Cancer Res.* 70, 709–718.
- He, G., Yu, G.-Y., Temkin, V., Ogata, H., Kuntzen, C., Sakurai, T., Sieghart, W., Peck-Radosavljevic, M., Leffert, H.L., and Karin, M. (2010). Hepatocyte IKKbeta/NF-kappaB inhibits tumor promotion and progression by preventing oxidative stress-driven STAT3 activation. *Cancer Cell* 17, 286–297.
- He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307.
- Heidelbaugh, J.J., and Bruderly, M. (2006). Cirrhosis and chronic liver failure: part I. Diagnosis and evaluation. *Am. Fam. Physician* 74, 756–762.
- Heinrich, P.C., Behrmann, I., Müller-Newen, G., Schaper, F., and Graeve, L. (1998). Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem. J.* 334 ( Pt 2), 297–314.
- Heinrich, P.C., Behrmann, I., Haan, S., Hermanns, H.M., Müller-Newen, G., and Schaper, F. (2003). Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochem. J.* 374, 1–20.
- Herath, N.I., Leggett, B.A., and MacDonald, G.A. (2006). Review of genetic and epigenetic alterations in hepatocarcinogenesis. *J. Gastroenterol. Hepatol.* 21, 15–21.
- Herceg, Z. (2007). Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. *Mutagenesis* 22, 91–103.
- Herceg, Z., and Hainaut, P. (2007). Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. *Mol. Oncol.* 1, 26–41.
- Herceg, Z., and Vaissière, T. (2011). Epigenetic mechanisms and cancer: an interface between the environment and the genome. *Epigenetics Off. J. DNA Methylation Soc.* 6, 804–819.
- Herceg, Z., Lambert, M.-P., van Veldhoven, K., Demetriou, C., Vineis, P., Smith, M.T., Straif, K., and Wild, C.P. (2013). Towards incorporating epigenetic mechanisms into carcinogen identification and evaluation. *Carcinogenesis* 34, 1955–1967.
- Hernández-Muñoz, I., Taghavi, P., Kuijl, C., Neefjes, J., and van Lohuizen, M. (2005). Association of BMI1 with polycomb bodies is dynamic and requires PRC2/EZH2 and the maintenance DNA methyltransferase DNMT1. *Mol. Cell. Biol.* 25, 11047–11058.
- Hernandez-Vargas, H., Sincic, N., Ouzounova, M., and Herceg, Z. (2009). Epigenetic signatures in stem cells and cancer stem cells. *Epigenomics* 1, 261–280.
- Hernandez-Vargas, H., Lambert, M.-P., Le Calvez-Kelm, F., Gouysse, G., McKay-Chopin, S., Tavtigian, S.V., Scoazec, J.-Y., and Herceg, Z. (2010). Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. *PloS One* 5, e9749.
- Hernandez-Vargas, H., Ouzounova, M., Le Calvez-Kelm, F., Lambert, M.-P., McKay-Chopin, S., Tavtigian, S.V.,

- Puisieux, A., Matar, C., and Herceg, Z. (2011). Methylome analysis reveals Jak-STAT pathway deregulation in putative breast cancer stem cells. *Epigenetics Off. J. DNA Methylation Soc.* 6, 428–439.
- Herzer, K., Sprinzl, M.F., and Galle, P.R. (2007). Hepatitis viruses: live and let die. *Liver Int. Off. J. Int. Assoc. Study Liver* 27, 293–301.
- Hill, D.B., Devalaraja, R., Joshi-Barve, S., Barve, S., and McClain, C.J. (1999). Antioxidants attenuate nuclear factor-kappa B activation and tumor necrosis factor-alpha production in alcoholic hepatitis patient monocytes and rat Kupffer cells, in vitro. *Clin. Biochem.* 32, 563–570.
- Horbelt, D., Denkis, A., and Knaus, P. (2012). A portrait of Transforming Growth Factor  $\beta$  superfamily signalling: Background matters. *Int. J. Biochem. Cell Biol.* 44, 469–474.
- Hu, T.-H., Huang, C.-C., Lin, P.-R., Chang, H.-W., Ger, L.-P., Lin, Y.-W., Changchien, C.-S., Lee, C.-M., and Tai, M.-H. (2003). Expression and prognostic role of tumor suppressor gene PTEN/MMAC1/TEP1 in hepatocellular carcinoma. *Cancer* 97, 1929–1940.
- Huang, J. (2009). Current progress in epigenetic research for hepatocarcinogenesis. *Sci. China Ser. C Life Sci. Chin. Acad. Sci.* 52, 31–42.
- Huang, T.H.-M., and Esteller, M. (2010). Chromatin remodeling in mammary gland differentiation and breast tumorigenesis. *Cold Spring Harb. Perspect. Biol.* 2, a004515.
- Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Huang, S., He, J., Zhang, X., Bian, Y., Yang, L., Xie, G., Zhang, K., Tang, W., Stelter, A.A., Wang, Q., et al. (2006). Activation of the hedgehog pathway in human hepatocellular carcinomas. *Carcinogenesis* 27, 1334–1340.
- Hussein, Y.M., Morad, F.E., Gameel, M.A., Emam, W.A., El Sawy, W.H., El Tarhouny, S.A., Bayomy, E.S., and Raafat, N. (2012). MAGE-4 gene m-RNA and TGF in blood as potential biochemical markers for HCC in HCV-infected patients. *Med. Oncol. Northwood Lond. Engl.* 29, 3055–3062.
- Hussey, G.S., Link, L.A., Brown, A.S., Howley, B.V., Chaudhury, A., and Howe, P.H. (2012). Establishment of a TGF $\beta$ -induced post-transcriptional EMT gene signature. *PloS One* 7, e52624.
- Iida, H., Suzuki, M., Goitsuka, R., and Ueno, H. (2012). Hypoxia induces CD133 expression in human lung cancer cells by up-regulation of OCT3/4 and SOX2. *Int. J. Oncol.* 40, 71–79.
- Ikushima, H., Todo, T., Ino, Y., Takahashi, M., Miyazawa, K., and Miyazono, K. (2009). Autocrine TGF- $\beta$  Signaling Maintains Tumorigenicity of Glioma-Initiating Cells through Sry-Related HMG-Box Factors. *Cell Stem Cell* 5, 504–514.
- Iliopoulos, D., Hirsch, H.A., and Struhl, K. (2009). An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. *Cell* 139, 693–706.
- Im, Y.H., Kim, H.T., Kim, I.Y., Factor, V.M., Hahm, K.B., Anzano, M., Jang, J.J., Flanders, K., Haines, D.C., Thorgeirsson, S.S., et al. (2001). Heterozygous mice for the transforming growth factor-beta type II receptor gene have increased susceptibility to hepatocellular carcinogenesis. *Cancer Res.* 61, 6665–6668.
- Inman, G.J. (2011). Switching TGF $\beta$  from a tumor suppressor to a tumor promoter. *Curr. Opin. Genet. Dev.* 21, 93–99.
- Inman, G.J., Nicolás, F.J., Callahan, J.F., Harling, J.D., Gaster, L.M., Reith, A.D., Laping, N.J., and Hill, C.S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* 62, 65–74.
- Ito, N., Kawata, S., Tamura, S., Takaishi, K., Shirai, Y., Kiso, S., Yabuuchi, I., Matsuda, Y., Nishioka, M., and Tarui, S. (1991). Elevated levels of transforming growth factor beta messenger RNA and its polypeptide in human hepatocellular carcinoma. *Cancer Res.* 51, 4080–4083.
- Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466, 1129–1133.
- Jaenisch, R., and Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* 33, 245–254.
- Jang, J.W., Oh, B.S., Kwon, J.H., You, C.R., Chung, K.W., Kay, C.S., and Jung, H.S. (2012). Serum interleukin-6 and C-reactive protein as a prognostic indicator in hepatocellular carcinoma. *Cytokine* 60, 686–693.

- Jia, Q., Zhang, X., Deng, T., and Gao, J. (2013). Positive correlation of Oct4 and ABCG2 to chemotherapeutic resistance in CD90(+)CD133(+) liver cancer stem cells. *Cell. Reprogramming* 15, 143–150.
- Jin, B., and Robertson, K.D. (2013). DNA methyltransferases, DNA damage repair, and cancer. *Adv. Exp. Med. Biol.* 754, 3–29.
- Jin, B., Tao, Q., Peng, J., Soo, H.M., Wu, W., Ying, J., Fields, C.R., Delmas, A.L., Liu, X., Qiu, J., et al. (2008). DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. *Hum. Mol. Genet.* 17, 690–709.
- Jin, B., Yao, B., Li, J.-L., Fields, C.R., Delmas, A.L., Liu, C., and Robertson, K.D. (2009). DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. *Cancer Res.* 69, 7412–7421.
- Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3, 415–428.
- Kaimori, A., Potter, J., Kaimori, J.-Y., Wang, C., Mezey, E., and Koteish, A. (2007). Transforming growth factor-beta1 induces an epithelial-to-mesenchymal transition state in mouse hepatocytes in vitro. *J. Biol. Chem.* 282, 22089–22101.
- Kam, I., Lynch, S., Svanas, G., Todo, S., Polimeno, L., Francavilla, A., Penkrot, R.J., Takaya, S., Ericzon, B.G., and Starzl, T.E. (1987). Evidence that host size determines liver size: studies in dogs receiving orthotopic liver transplants. *Hepatology* 7, 362–366.
- Kao, J.-T., Lai, H.-C., Tsai, S.-M., Lin, P.-C., Chuang, P.-H., Yu, C.-J., Cheng, K.-S., Su, W.-P., Hsu, P.-N., Peng, C.-Y., et al. (2012). Rather than interleukin-27, interleukin-6 expresses positive correlation with liver severity in naïve hepatitis B infection patients. *Liver Int. Off. J. Int. Assoc. Study Liver* 32, 928–936.
- Karkampouna, S., Ten Dijke, P., Dooley, S., and Julio, M.K. (2012). TGF $\beta$  signaling in liver regeneration. *Curr. Pharm. Des.* 18, 4103–4113.
- Katsuno, Y., Lamouille, S., and Derynck, R. (2013). TGF- $\beta$  signaling and epithelial-mesenchymal transition in cancer progression. *Curr. Opin. Oncol.* 25, 76–84.
- Kerr, I.D., Haider, A.J., and Gelissen, I.C. (2011). The ABCG family of membrane-associated transporters: you don't have to be big to be mighty. *Br. J. Pharmacol.* 164, 1767–1779.
- Kim, H.-P., and Leonard, W.J. (2007). CREB/ATF-dependent T cell receptor-induced FoxP3 gene expression: a role for DNA methylation. *J. Exp. Med.* 204, 1543–1551.
- Kim, H., Choi, G.H., Na, D.C., Ahn, E.Y., Kim, G.I., Lee, J.E., Cho, J.Y., Yoo, J.E., Choi, J.S., and Park, Y.N. (2011). Human hepatocellular carcinomas with “Stemness”-related marker expression: keratin 19 expression and a poor prognosis. *Hepatology* 54, 1707–1717.
- Kim, S.-Y., Kang, J.W., Song, X., Kim, B.K., Yoo, Y.D., Kwon, Y.T., and Lee, Y.J. (2013). Role of the IL-6-JAK1-STAT3-Oct-4 pathway in the conversion of non-stem cancer cells into cancer stem-like cells. *Cell. Signal.* 25, 961–969.
- Kimura, O., Takahashi, T., Ishii, N., Inoue, Y., Ueno, Y., Kogure, T., Fukushima, K., Shiina, M., Yamagiwa, Y., Kondo, Y., et al. (2010). Characterization of the epithelial cell adhesion molecule (EpCAM)+ cell population in hepatocellular carcinoma cell lines. *Cancer Sci.* 101, 2145–2155.
- Kirillova, I., Chaisson, M., and Fausto, N. (1999). Tumor necrosis factor induces DNA replication in hepatic cells through nuclear factor kappaB activation. *Cell Growth Differ. Mol. Biol. J. Am. Assoc. Cancer Res.* 10, 819–828.
- Kitamura, Y., Shirahata, A., Sakuraba, K., Goto, T., Mizukami, H., Saito, M., Ishibashi, K., Kigawa, G., Nemoto, H., Sanada, Y., et al. (2011). Aberrant methylation of the Vimentin gene in hepatocellular carcinoma. *Anticancer Res.* 31, 1289–1291.
- Kitaoka, S., Shiota, G., and Kawasaki, H. (2003). Serum levels of interleukin-10, interleukin-12 and soluble interleukin-2 receptor in chronic liver disease type C. *Hepatogastroenterology.* 50, 1569–1574.
- Kitisin, K., Shetty, K., Mishra, L., and Johnson, L.B. (2007). Hepatocellular stem cells. *Cancer Biomark. Sect. Dis. Markers* 3, 251–262.
- Kline, T.L., Zamir, M., and Ritman, E.L. (2011). Relating function to branching geometry: a micro-CT study of the hepatic artery, portal vein, and biliary tree. *Cells Tissues Organs* 194, 431–442.

- Kmiec, Z. (2001). Cooperation of liver cells in health and disease. *Adv. Anat. Embryol. Cell Biol.* 161, III–XIII, 1–151.
- Knight, B., and Yeoh, G.C. (2005). TNF/LT $\alpha$  double knockout mice display abnormal inflammatory and regenerative responses to acute and chronic liver injury. *Cell Tissue Res.* 319, 61–70.
- Knight, B., Yeoh, G.C., Husk, K.L., Ly, T., Abraham, L.J., Yu, C., Rhim, J.A., and Fausto, N. (2000). Impaired preneoplastic changes and liver tumor formation in tumor necrosis factor receptor type 1 knockout mice. *J. Exp. Med.* 192, 1809–1818.
- Knight, B., Tirnitz-Parker, J.E.E., and Olynyk, J.K. (2008). C-kit inhibition by imatinib mesylate attenuates progenitor cell expansion and inhibits liver tumor formation in mice. *Gastroenterology* 135, 969–979, 979.e1.
- Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., et al. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 468, 839–843.
- Koh, K.P., and Rao, A. (2013). DNA methylation and methylcytosine oxidation in cell fate decisions. *Curr. Opin. Cell Biol.* 25, 152–161.
- Kohga, K., Tatsumi, T., Takehara, T., Tsunematsu, H., Shimizu, S., Yamamoto, M., Sasakawa, A., Miyagi, T., and Hayashi, N. (2010). Expression of CD133 confers malignant potential by regulating metalloproteinases in human hepatocellular carcinoma. *J. Hepatol.* 52, 872–879.
- Kojima, H., Yokosuka, O., Imazeki, F., Saisho, H., and Omata, M. (1997). Telomerase activity and telomere length in hepatocellular carcinoma and chronic liver disease. *Gastroenterology* 112, 493–500.
- Komatsu, Y., Waku, T., Iwasaki, N., Ono, W., Yamaguchi, C., and Yanagisawa, J. (2012). Global analysis of DNA methylation in early-stage liver fibrosis. *BMC Med. Genomics* 5, 5.
- Kondo, Y., Shen, L., Cheng, A.S., Ahmed, S., Bumber, Y., Charo, C., Yamochi, T., Urano, T., Furukawa, K., Kwabi-Addo, B., et al. (2008). Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat. Genet.* 40, 741–750.
- Kopec, K.L., and Burns, D. (2011). Nonalcoholic fatty liver disease: a review of the spectrum of disease, diagnosis, and therapy. *Nutr. Clin. Pract. Off. Publ. Am. Soc. Parenter. Enter. Nutr.* 26, 565–576.
- Kordes, C., and Häussinger, D. (2013). Hepatic stem cell niches. *J. Clin. Invest.* 123, 1874–1880.
- Korkaya, H., Liu, S., and Wicha, M.S. (2011). Regulation of cancer stem cells by cytokine networks: attacking cancer's inflammatory roots. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 17, 6125–6129.
- Kouzarides, T. (2007). SnapShot: Histone-modifying enzymes. *Cell* 131, 822.
- Kovalovich, K., DeAngelis, R.A., Li, W., Furth, E.E., Ciliberto, G., and Taub, R. (2000). Increased toxin-induced liver injury and fibrosis in interleukin-6-deficient mice. *Hepatol. Baltim. Md* 31, 149–159.
- Kretschmar, M., Doody, J., Timokhina, I., and Massagué, J. (1999). A mechanism of repression of TGF $\beta$ /Smad signaling by oncogenic Ras. *Genes Dev.* 13, 804–816.
- Kumar, V., Abbas, A.K., Fausto, N., and Mitchell, R. (2012). *Robbins Basic Pathology* (Elsevier Health Sciences).
- Kurpinski, K., Lam, H., Chu, J., Wang, A., Kim, A., Tsay, E., Agrawal, S., Schaffer, D.V., and Li, S. (2010). Transforming growth factor-beta and notch signaling mediate stem cell differentiation into smooth muscle cells. *Stem Cells Dayt. Ohio* 28, 734–742.
- Kwon, G.Y., Yoo, B.C., Koh, K.C., Cho, J.W., Park, W.S., and Park, C.K. (2005). Promoter methylation of E-cadherin in hepatocellular carcinomas and dysplastic nodules. *J. Korean Med. Sci.* 20, 242–247.
- De La Coste, A., Romagnolo, B., Billuart, P., Renard, C.A., Buendia, M.A., Soubrane, O., Fabre, M., Chelly, J., Beldjord, C., Kahn, A., et al. (1998). Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8847–8851.
- Lambert, M.-P., Paliwal, A., Vaissière, T., Chemin, I., Zoulim, F., Tommasino, M., Hainaut, P., Sylla, B., Scoazec, J.-Y., Tost, J., et al. (2011). Aberrant DNA methylation distinguishes hepatocellular carcinoma associated with HBV and HCV infection and alcohol intake. *J. Hepatol.* 54, 705–715.
- Lange, M., Demajo, S., Jain, P., and Di Croce, L. (2011). Combinatorial assembly and function of chromatin regulatory complexes. *Epigenomics* 3, 567–580.

- Lapidot, T., Sirard, C., Vormoor, J., Murdoch, B., Hoang, T., Caceres-Cortes, J., Minden, M., Paterson, B., Caligiuri, M.A., and Dick, J.E. (1994). A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367, 645–648.
- Lee, C.J., Dosch, J., and Simeone, D.M. (2008). Pancreatic cancer stem cells. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 26, 2806–2812.
- Lee, D., Chung, Y.-H., Kim, J.A., Lee, Y.S., Lee, D., Jang, M.K., Kim, K.M., Lim, Y.-S., Lee, H.C., and Lee, Y.S. (2012). Transforming growth factor beta 1 overexpression is closely related to invasiveness of hepatocellular carcinoma. *Oncology* 82, 11–18.
- Lee, E.-K., Han, G.-Y., Park, H.W., Song, Y.-J., and Kim, C.-W. (2010a). Transgelin promotes migration and invasion of cancer stem cells. *J. Proteome Res.* 9, 5108–5117.
- Lee, H.S., Kim, B.-H., Cho, N.-Y., Yoo, E.J., Choi, M., Shin, S.-H., Jang, J.-J., Suh, K.-S., Kim, Y.S., and Kang, G.H. (2009). Prognostic implications of and relationship between CpG island hypermethylation and repetitive DNA hypomethylation in hepatocellular carcinoma. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* 15, 812–820.
- Lee, J.-S., Heo, J., Libbrecht, L., Chu, I.-S., Kaposi-Novak, P., Calvisi, D.F., Mikaelyan, A., Roberts, L.R., Demetris, A.J., Sun, Z., et al. (2006). A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat. Med.* 12, 410–416.
- Lee, J.-S., Smith, E., and Shilatifard, A. (2010b). The language of histone crosstalk. *Cell* 142, 682–685.
- Lee, T.K.W., Castilho, A., Cheung, V.C.H., Tang, K.H., Ma, S., and Ng, I.O.L. (2011). CD24(+) liver tumor-initiating cells drive self-renewal and tumor initiation through STAT3-mediated NANOG regulation. *Cell Stem Cell* 9, 50–63.
- Legoix, P., Bluteau, O., Bayer, J., Perret, C., Balabaud, C., Belghiti, J., Franco, D., Thomas, G., Laurent-Puig, P., and Zucman-Rossi, J. (1999). Beta-catenin mutations in hepatocellular carcinoma correlate with a low rate of loss of heterozygosity. *Oncogene* 18, 4044–4046.
- Levy, L., and Hill, C.S. (2006). Alterations in components of the TGF-beta superfamily signaling pathways in human cancer. *Cytokine Growth Factor Rev.* 17, 41–58.
- Li, Y., and Laterra, J. (2012). Cancer stem cells: distinct entities or dynamically regulated phenotypes? *Cancer Res.* 72, 576–580.
- Li, Z., and Rich, J.N. (2010). Hypoxia and hypoxia inducible factors in cancer stem cell maintenance. *Curr. Top. Microbiol. Immunol.* 345, 21–30.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.
- Li, F., Zeng, B., Chai, Y., Cai, P., Fan, C., and Cheng, T. (2009). The linker region of Smad2 mediates TGF-beta-dependent ERK2-induced collagen synthesis. *Biochem. Biophys. Res. Commun.* 386, 289–293.
- Li, G.-C., Ye, Q.-H., Dong, Q.-Z., Ren, N., Jia, H.-L., and Qin, L.-X. (2012a). TGF beta1 and related-Smads contribute to pulmonary metastasis of hepatocellular carcinoma in mice model. *J. Exp. Clin. Cancer Res. CR* 31, 93.
- Li, S., Vriend, L.E.M., Nasser, I.A., Popov, Y., Afdhal, N.H., Koziel, M.J., Schuppan, D., Exley, M.A., and Alatrakchi, N. (2012b). Hepatitis C virus-specific T-cell-derived transforming growth factor beta is associated with slow hepatic fibrogenesis. *Hepatology* 56, 2094–2105.
- Li, W., Yu, M., Luo, S., Liu, H., Gao, Y., Wilson, J.X., and Huang, G. (2013). DNA methyltransferase mediates dose-dependent stimulation of neural stem cell proliferation by folate. *J. Nutr. Biochem.* 24, 1295–1301.
- Liang, H., Block, T.M., Wang, M., Nefsky, B., Long, R., Hafner, J., Mehta, A.S., Marrero, J., Gish, R., and Norton, P.A. (2012). Interleukin-6 and oncostatin M are elevated in liver disease in conjunction with candidate hepatocellular carcinoma biomarker GP73. *Cancer Biomark. Sect. Dis. Markers* 11, 161–171.
- Libbrecht, L. (2006). Hepatic progenitor cells in human liver tumor development. *World J. Gastroenterol. WJG* 12, 6261–6265.
- Libbrecht, L., and Roskams, T. (2002). Hepatic progenitor cells in human liver diseases. *Semin. Cell Dev. Biol.* 13, 389–396.

- Libbrecht, L., Desmet, V., Van Damme, B., and Roskams, T. (2000). Deep intralobular extension of human hepatic "progenitor cells" correlates with parenchymal inflammation in chronic viral hepatitis: can "progenitor cells" migrate? *J. Pathol.* 192, 373–378.
- Lim, S.-O., Gu, J.-M., Kim, M.S., Kim, H.-S., Park, Y.N., Park, C.K., Cho, J.W., Park, Y.M., and Jung, G. (2008). Epigenetic changes induced by reactive oxygen species in hepatocellular carcinoma: methylation of the E-cadherin promoter. *Gastroenterology* 135, 2128–2140, 2140.e1–8.
- Lin, W.-W., and Karin, M. (2007). A cytokine-mediated link between innate immunity, inflammation, and cancer. *J. Clin. Invest.* 117, 1175–1183.
- Lin, C.H., Hsieh, S.Y., Sheen, I.S., Lee, W.C., Chen, T.C., Shyu, W.C., and Liaw, Y.F. (2001). Genome-wide hypomethylation in hepatocellular carcinogenesis. *Cancer Res.* 61, 4238–4243.
- Lin, H.-K., Hu, Y.-C., Yang, L., Altuwaijri, S., Chen, Y.-T., Kang, H.-Y., and Chang, C. (2003). Suppression versus induction of androgen receptor functions by the phosphatidylinositol 3-kinase/Akt pathway in prostate cancer LNCaP cells with different passage numbers. *J. Biol. Chem.* 278, 50902–50907.
- Lin, W.-R., Lim, S.-N., McDonald, S.A.C., Graham, T., Wright, V.L., Peplow, C.L., Humphries, A., Kocher, H.M., Wright, N.A., Dhillon, A.P., et al. (2010). The histogenesis of regenerative nodules in human liver cirrhosis. *Hepatology* 51, 1017–1026.
- Lingala, S., Cui, Y.-Y., Chen, X., Ruebner, B.H., Qian, X.-F., Zern, M.A., and Wu, J. (2010). Immunohistochemical staining of cancer stem cell markers in hepatocellular carcinoma. *Exp. Mol. Pathol.* 89, 27–35.
- Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315–322.
- Lodish, H. (2008). *Molecular Cell Biology* (W. H. Freeman).
- Lowes, K.N., Brennan, B.A., Yeoh, G.C., and Olynyk, J.K. (1999). Oval cell numbers in human chronic liver diseases are directly related to disease severity. *Am. J. Pathol.* 154, 537–541.
- Ma, S. (2013). Biology and clinical implications of CD133(+) liver cancer stem cells. *Exp. Cell Res.* 319, 126–132.
- Ma, D.K., Jang, M.-H., Guo, J.U., Kitabatake, Y., Chang, M.-L., Pow-Anpongkul, N., Flavell, R.A., Lu, B., Ming, G.-L., and Song, H. (2009). Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* 323, 1074–1077.
- Ma, S., Chan, K.-W., Hu, L., Lee, T.K.-W., Wo, J.Y.-H., Ng, I.O.-L., Zheng, B.-J., and Guan, X.-Y. (2007). Identification and characterization of tumorigenic liver cancer stem/progenitor cells. *Gastroenterology* 132, 2542–2556.
- Ma, S., Chan, K.W., Lee, T.K.-W., Tang, K.H., Wo, J.Y.-H., Zheng, B.-J., and Guan, X.-Y. (2008a). Aldehyde dehydrogenase discriminates the CD133 liver cancer stem cell populations. *Mol. Cancer Res.* 6, 1146–1153.
- Ma, S., Lee, T.K., Zheng, B.-J., Chan, K.W., and Guan, X.-Y. (2008b). CD133+ HCC cancer stem cells confer chemoresistance by preferential expression of the Akt/PKB survival pathway. *Oncogene* 27, 1749–1758.
- Ma, S., Tang, K.H., Chan, Y.P., Lee, T.K., Kwan, P.S., Castilho, A., Ng, I., Man, K., Wong, N., To, K.-F., et al. (2010). miR-130b Promotes CD133(+) liver tumor-initiating cell growth and self-renewal via tumor protein 53-induced nuclear protein 1. *Cell Stem Cell* 7, 694–707.
- Machida, K., Tsukamoto, H., Mkrtchyan, H., Duan, L., Dynnyk, A., Liu, H.M., Asahina, K., Govindarajan, S., Ray, R., Ou, J.-H.J., et al. (2009). Toll-like receptor 4 mediates synergism between alcohol and HCV in hepatic oncogenesis involving stem cell marker Nanog. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1548–1553.
- Mahapatra, S., Firpo, M.T., and Bacanamwo, M. (2010). Inhibition of DNA methyltransferases and histone deacetylases induces bone marrow-derived multipotent adult progenitor cells to differentiate into endothelial cells. *Ethn. Dis.* 20, S1–60–4.
- Maione, D., Di Carlo, E., Li, W., Musiani, P., Modesti, A., Peters, M., Rose-John, S., Della Rocca, C., Tripodi, M., Lazzaro, D., et al. (1998). Coexpression of IL-6 and soluble IL-6R causes nodular regenerative hyperplasia and adenomas of the liver. *EMBO J.* 17, 5588–5597.
- Maitah, M.Y., Ali, S., Ahmad, A., Gadgeel, S., and Sarkar, F.H. (2011). Up-regulation of sonic hedgehog contributes to TGF- $\beta$ 1-induced epithelial to mesenchymal transition in NSCLC cells. *PloS One* 6, e16068.

- Mak, A.B., Nixon, A.M.L., Kittanakom, S., Stewart, J.M., Chen, G.I., Curak, J., Gingras, A.-C., Mazitschek, R., Neel, B.G., Stagljar, I., et al. (2012). Regulation of CD133 by HDAC6 promotes b-catenin signaling to suppress cancer cell differentiation. *Cell Rep.* 2, 951–963.
- Malanchi, I., Santamaria-Martínez, A., Susanto, E., Peng, H., Lehr, H.-A., Delaloye, J.-F., and Huelsken, J. (2012). Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* 481, 85–89.
- Mandrekar, P., and Szabo, G. (2009). Signalling pathways in alcohol-induced liver inflammation. *J. Hepatol.* 50, 1258–1266.
- Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008a). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.
- Mani, S.A., Guo, W., Liao, M.-J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., et al. (2008b). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.
- Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. *Nature* 454, 436–444.
- Mao, J., Yu, H., Wang, C., Sun, L., Jiang, W., Zhang, P., Xiao, Q., Han, D., Han, D., Saiyin, H., et al. (2012). Metallothionein MT1M is a tumor suppressor of human hepatocellular carcinomas. *Carcinogenesis* 33, 2568–2577.
- Marek, A., Brodzicki, J., Liberek, A., and Korzon, M. (2002). TGF-beta (transforming growth factor-beta) in chronic inflammatory conditions - a new diagnostic and prognostic marker? *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* 8, RA145–151.
- Marquardt, J.U., Factor, V.M., and Thorgeirsson, S.S. (2010). Epigenetic regulation of cancer stem cells in liver cancer: current concepts and clinical implications. *J. Hepatol.* 53, 568–577.
- Martin, M., and Herceg, Z. (2012). From hepatitis to hepatocellular carcinoma: a proposed model for cross-talk between inflammation and epigenetic mechanisms. *Genome Med.* 4, 8.
- Massagué, J. (2008). TGFbeta in Cancer. *Cell* 134, 215–230.
- Masubuchi, Y., Bourdi, M., Reilly, T.P., Graf, M.L.M., George, J.W., and Pohl, L.R. (2003). Role of interleukin-6 in hepatic heat shock protein expression and protection against acetaminophen-induced liver disease. *Biochem. Biophys. Res. Commun.* 304, 207–212.
- Matak, P., Chaston, T.B., Chung, B., Srai, S.K., McKie, A.T., and Sharp, P.A. (2009). Activated macrophages induce hepcidin expression in HuH7 hepatoma cells. *Haematologica* 94, 773–780.
- Mathieu, J., Zhang, Z., Zhou, W., Wang, A.J., Heddleston, J.M., Pinna, C.M.A., Hubaud, A., Stadler, B., Choi, M., Bar, M., et al. (2011). HIF induces human embryonic stem cell markers in cancer cells. *Cancer Res.* 71, 4640–4652.
- Matsuno, Y., Coelho, A.L., Jarai, G., Westwick, J., and Hogaboam, C.M. (2012). Notch signaling mediates TGF-b1-induced epithelial-mesenchymal transition through the induction of Snai1. *Int. J. Biochem. Cell Biol.* 44, 776–789.
- Matsuura, I., Chiang, K.-N., Lai, C.-Y., He, D., Wang, G., Ramkumar, R., Uchida, T., Ryo, A., Lu, K., and Liu, F. (2010). Pin1 promotes transforming growth factor-beta-induced migration and invasion. *J. Biol. Chem.* 285, 1754–1764.
- Matsuzaki, K. (2009). Modulation of TGF-beta signaling during progression of chronic liver diseases. *Front. Biosci. Landmark Ed.* 14, 2923–2934.
- Matsuzaki, K. (2011). Smad phosphoisoform signaling specificity: the right place at the right time. *Carcinogenesis* 32, 1578–1588.
- Matsuzaki, K. (2012). Smad phosphoisoform signals in acute and chronic liver injury: similarities and differences between epithelial and mesenchymal cells. *Cell Tissue Res.* 347, 225–243.
- Matsuzaki, K., Murata, M., Yoshida, K., Sekimoto, G., Uemura, Y., Sakaida, N., Kaibori, M., Kamiyama, Y., Nishizawa, M., Fujisawa, J., et al. (2007). Chronic inflammation associated with hepatitis C virus infection perturbs hepatic transforming growth factor beta signaling, promoting cirrhosis and hepatocellular carcinoma. *Hepatol. Baltim. Md* 46, 48–57.



- Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466, 253–257.
- Maw, M.A., Corbeil, D., Koch, J., Hellwig, A., Wilson-Wheeler, J.C., Bridges, R.J., Kumaramanickavel, G., John, S., Nancarrow, D., Röper, K., et al. (2000). A frameshift mutation in prominin (mouse)-like 1 causes human retinal degeneration. *Hum. Mol. Genet.* 9, 27–34.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* 403, 501–502.
- McDonald, O.G., Wu, H., Timp, W., Doi, A., and Feinberg, A.P. (2011). Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat. Struct. Mol. Biol.* 18, 867–874.
- McQualter, J.L., Yuen, K., Williams, B., and Bertoncello, I. (2010). Evidence of an epithelial stem/progenitor cell hierarchy in the adult mouse lung. *Proc. Natl. Acad. Sci. U. S. A.* 107, 1414–1419.
- Michalopoulos, G.K., and DeFrances, M.C. (1997). Liver regeneration. *Science* 276, 60–66.
- Mieli-Vergani, G., and Vergani, D. (2011). Autoimmune hepatitis. *Nat. Rev. Gastroenterol. Hepatol.* 8, 320–329.
- Mikulits, W. (2009). Epithelial to mesenchymal transition in hepatocellular carcinoma. *Future Oncol. Lond. Engl.* 5, 1169–1179.
- Miraglia, S., Godfrey, W., Yin, A.H., Atkins, K., Warnke, R., Holden, J.T., Bray, R.A., Waller, E.K., and Buck, D.W. (1997). A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood* 90, 5013–5021.
- Miyazawa, K., Shinozaki, M., Hara, T., Furuya, T., and Miyazono, K. (2002). Two major Smad pathways in TGF-beta superfamily signalling. *Genes Cells Devoted Mol. Cell. Mech.* 7, 1191–1204.
- Miyazono, K. (2012). Tumour promoting functions of TGF-b in CML-initiating cells. *J. Biochem. (Tokyo)* 152, 383–385.
- Morgan, H.D., Santos, F., Green, K., Dean, W., and Reik, W. (2005). Epigenetic reprogramming in mammals. *Hum. Mol. Genet.* 14 Spec No 1, R47–58.
- Mori, S., Matsuzaki, K., Yoshida, K., Furukawa, F., Tahashi, Y., Yamagata, H., Sekimoto, G., Seki, T., Matsui, H., Nishizawa, M., et al. (2004). TGF-beta and HGF transmit the signals through JNK-dependent Smad2/3 phosphorylation at the linker regions. *Oncogene* 23, 7416–7429.
- Moustakas, A., and Kardassis, D. (1998). Regulation of the human p21/WAF1/Cip1 promoter in hepatic cells by functional interactions between Sp1 and Smad family members. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6733–6738.
- Mullor, J.L., Sánchez, P., and Ruiz i Altaba, A. (2002). Pathways and consequences: Hedgehog signaling in human disease. *Trends Cell Biol.* 12, 562–569.
- Münzel, M., Globisch, D., and Carell, T. (2011). 5-Hydroxymethylcytosine, the sixth base of the genome. *Angew. Chem. Int. Ed Engl.* 50, 6460–6468.
- Murata, M., Matsuzaki, K., Yoshida, K., Sekimoto, G., Tahashi, Y., Mori, S., Uemura, Y., Sakaida, N., Fujisawa, J., Seki, T., et al. (2009). Hepatitis B virus X protein shifts human hepatic transforming growth factor (TGF)-beta signaling from tumor suppression to oncogenesis in early chronic hepatitis B. *Hepatol. Baltim. Md* 49, 1203–1217.
- Na, D.C., Lee, J.E., Yoo, J.E., Oh, B.-K., Choi, G.H., and Park, Y.N. (2011). Invasion and EMT-associated genes are up-regulated in B viral hepatocellular carcinoma with high expression of CD133-human and cell culture study. *Exp. Mol. Pathol.* 90, 66–73.
- Nagao, K., Tomimatsu, M., Endo, H., Hisatomi, H., and Hikiji, K. (1999). Telomerase reverse transcriptase mRNA expression and telomerase activity in hepatocellular carcinoma. *J. Gastroenterol.* 34, 83–87.
- Nagashio, R., Arai, E., Ojima, H., Kosuge, T., Kondo, Y., and Kanai, Y. (2011). Carcinogenetic risk estimation based on quantification of DNA methylation levels in liver tissue at the precancerous stage. *Int. J. Cancer J. Int. Cancer* 129, 1170–1179.
- Nagata, H., Hatano, E., Tada, M., Murata, M., Kitamura, K., Asechi, H., Narita, M., Yanagida, A., Tamaki, N., Yagi, S., et al. (2009). Inhibition of c-Jun NH2-terminal kinase switches Smad3 signaling from oncogenesis to tumor-suppression in rat hepatocellular carcinoma. *Hepatol. Baltim. Md* 49, 1944–1953.

- Naka, K., Hoshii, T., Muraguchi, T., Tadokoro, Y., Ooshio, T., Kondo, Y., Nakao, S., Motoyama, N., and Hirao, A. (2010). TGF-beta-FOXO signalling maintains leukaemia-initiating cells in chronic myeloid leukaemia. *Nature* 463, 676–680.
- Neumann-Haefelin, C., Blum, H.E., Chisari, F.V., and Thimme, R. (2005). T cell response in hepatitis C virus infection. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 32, 75–85.
- Nguyen, L.N., Furuya, M.H., Wolfrain, L.A., Nguyen, A.P., Holdren, M.S., Campbell, J.S., Knight, B., Yeoh, G.C.T., Fausto, N., and Parks, W.T. (2007). Transforming growth factor-beta differentially regulates oval cell and hepatocyte proliferation. *Hepatology* 45, 31–41.
- Nishida, N., Nagasaka, T., Nishimura, T., Ikai, I., Boland, C.R., and Goel, A. (2008). Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. *Hepatology* 47, 908–918.
- Nishida, N., Kudo, M., Nagasaka, T., Ikai, I., and Goel, A. (2012a). Characteristic patterns of altered DNA methylation predict emergence of human hepatocellular carcinoma. *Hepatology* 56, 994–1003.
- Nishida, N., Kudo, M., Nagasaka, T., Ikai, I., and Goel, A. (2012b). Characteristic patterns of altered DNA methylation predict emergence of human hepatocellular carcinoma. *Hepatology* 56, 994–1003.
- Niwa, Y., Kanda, H., Shikauchi, Y., Saiura, A., Matsubara, K., Kitagawa, T., Yamamoto, J., Kubo, T., and Yoshikawa, H. (2005). Methylation silencing of SOCS-3 promotes cell growth and migration by enhancing JAK/STAT and FAK signalings in human hepatocellular carcinoma. *Oncogene* 24, 6406–6417.
- Noble, J.R., Zhong, Z.-H., Neumann, A.A., Melki, J.R., Clark, S.J., and Reddel, R.R. (2004). Alterations in the p16(INK4a) and p53 tumor suppressor genes of hTERT-immortalized human fibroblasts. *Oncogene* 23, 3116–3121.
- Nombela-Arrieta, C., Pivarnik, G., Winkel, B., Canty, K.J., Harley, B., Mahoney, J.E., Park, S.-Y., Lu, J., Protopopov, A., and Silberstein, L.E. (2013). Quantitative imaging of haematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. *Nat. Cell Biol.* 15, 533–543.
- O'Brien, C.A., Pollett, A., Gallinger, S., and Dick, J.E. (2007). A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445, 106–110.
- O'Driscoll, L., Gammell, P., McKiernan, E., Ryan, E., Jeppesen, P.B., Rani, S., and Clynes, M. (2006). Phenotypic and global gene expression profile changes between low passage and high passage MIN-6 cells. *J. Endocrinol.* 191, 665–676.
- Oh, B.-K., Kim, H., Park, H.-J., Shim, Y.-H., Choi, J., Park, C., and Park, Y.N. (2007). DNA methyltransferase expression and DNA methylation in human hepatocellular carcinoma and their clinicopathological correlation. *Int. J. Mol. Med.* 20, 65–73.
- Ohm, J.E., McGarvey, K.M., Yu, X., Cheng, L., Schuebel, K.E., Cope, L., Mohammad, H.P., Chen, W., Daniel, V.C., Yu, W., et al. (2007). A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat. Genet.* 39, 237–242.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.
- Okumoto, K., Hattori, E., Tamura, K., Kiso, S., Watanabe, H., Saito, K., Saito, T., Togashi, H., and Kawata, S. (2004). Possible contribution of circulating transforming growth factor-beta1 to immunity and prognosis in unresectable hepatocellular carcinoma. *Liver Int. Off. J. Int. Assoc. Study Liver* 24, 21–28.
- Olempska, M., Eisenach, P.A., Ammerpohl, O., Ungefroren, H., Fandrich, F., and Kalthoff, H. (2007). Detection of tumor stem cell markers in pancreatic carcinoma cell lines. *Hepatobiliary Pancreat. Dis. Int. HBPDI* 6, 92–97.
- Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb. Perspect. Biol.* 2, a001008.
- Olsen, A.L., Bloomer, S.A., Chan, E.P., Gaça, M.D.A., Georges, P.C., Sackey, B., Uemura, M., Janmey, P.A., and Wells, R.G. (2011). Hepatic stellate cells require a stiff environment for myofibroblastic differentiation. *Am. J. Physiol. Gastrointest. Liver Physiol.* 301, G110–118.
- Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol. CB* 10, 475–478.

- Padua, D., and Massagué, J. (2009). Roles of TGFbeta in metastasis. *Cell Res.* 19, 89–102.
- Park, H.-J., Yu, E., and Shim, Y.-H. (2006). DNA methyltransferase expression and DNA hypermethylation in human hepatocellular carcinoma. *Cancer Lett.* 233, 271–278.
- Park, I.Y., Sohn, B.H., Yu, E., Suh, D.J., Chung, Y.-H., Lee, J.-H., Surzycki, S.J., and Lee, Y.I. (2007). Aberrant epigenetic modifications in hepatocarcinogenesis induced by hepatitis B virus X protein. *Gastroenterology* 132, 1476–1494.
- Pastor, W.A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* 14, 341–356.
- Pastrana, E., Silva-Vargas, V., and Doetsch, F. (2011). Eyes wide open: a critical review of sphere-formation as an assay for stem cells. *Cell Stem Cell* 8, 486–498.
- Paz, M.F., Fraga, M.F., Avila, S., Guo, M., Pollan, M., Herman, J.G., and Esteller, M. (2003). A systematic profile of DNA methylation in human cancer cell lines. *Cancer Res.* 63, 1114–1121.
- Pediaditakis, P., Lopez-Talavera, J.C., Petersen, B., Monga, S.P., and Michalopoulos, G.K. (2001). The processing and utilization of hepatocyte growth factor/scatter factor following partial hepatectomy in the rat. *Hepatology* 34, 688–693.
- Pelicci, P.G., Dalton, P., and Giorgio, M. (2013). The other face of ROS: a driver of stem cell expansion in colorectal cancer. *Cell Stem Cell* 12, 635–636.
- Pen, A., Moreno, M.J., Durocher, Y., Deb-Rinker, P., and Stanimirovic, D.B. (2008). Glioblastoma-secreted factors induce IGFBP7 and angiogenesis by modulating Smad-2-dependent TGF-beta signaling. *Oncogene* 27, 6834–6844.
- Peñuelas, S., Anido, J., Prieto-Sánchez, R.M., Folch, G., Barba, I., Cuartas, I., García-Dorado, D., Poca, M.A., Sahuquillo, J., Baselga, J., et al. (2009). TGF-beta increases glioma-initiating cell self-renewal through the induction of LIF in human glioblastoma. *Cancer Cell* 15, 315–327.
- Piao, L.S., Hur, W., Kim, T.-K., Hong, S.W., Kim, S.W., Choi, J.E., Sung, P.S., Song, M.J., Lee, B.-C., Hwang, D., et al. (2012). CD133+ liver cancer stem cells modulate radioresistance in human hepatocellular carcinoma. *Cancer Lett.* 315, 129–137.
- Pinzani, M., Rosselli, M., and Zuckermann, M. (2011). Liver cirrhosis. *Best Pract. Res. Clin. Gastroenterol.* 25, 281–290.
- Pirozzi, G., Tirino, V., Camerlingo, R., Franco, R., La Rocca, A., Liguori, E., Martucci, N., Paino, F., Normanno, N., and Rocco, G. (2011). Epithelial to mesenchymal transition by TGFb-1 induction increases stemness characteristics in primary non small cell lung cancer cell line. *PloS One* 6, e21548.
- Pogribny, I.P., and Rusyn, I. (2012). Role of epigenetic aberrations in the development and progression of human hepatocellular carcinoma. *Cancer Lett.*
- Pogribny, I.P., James, S.J., Jernigan, S., and Pogribna, M. (2004). Genomic hypomethylation is specific for preneoplastic liver in folate/methyl deficient rats and does not occur in non-target tissues. *Mutat. Res.* 548, 53–59.
- POPPER, H., KENT, G., and STEIN, R. (1957). Ductular cell reaction in the liver in hepatic injury. *J. Mt. Sinai Hosp. N. Y.* 24, 551–556.
- Poppleton, H., and Gilbertson, R.J. (2007). Stem cells of ependymoma. *Br. J. Cancer* 96, 6–10.
- Preisegger, K.H., Factor, V.M., Fuchsbichler, A., Stumptner, C., Denk, H., and Thorgeirsson, S.S. (1999). Atypical ductular proliferation and its inhibition by transforming growth factor beta1 in the 3,5-diethoxycarbonyl-1,4-dihydrocollidine mouse model for chronic alcoholic liver disease. *Lab. Investig. J. Tech. Methods Pathol.* 79, 103–109.
- Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* 15, 1613–1618.
- Qiu, B., Zhang, D., Wang, Y., Ou, S., Wang, J., Tao, J., and Wang, Y. (2013). Interleukin-6 is overexpressed and augments invasiveness of human glioma stem cells in vitro. *Clin. Exp. Metastasis.*
- Quintana, E., Shackleton, M., Foster, H.R., Fullen, D.R., Sabel, M.S., Johnson, T.M., and Morrison, S.J. (2010).

- Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* 18, 510–523.
- Ramadori, G., and Armbrust, T. (2001). Cytokines in the liver. *Eur. J. Gastroenterol. Hepatol.* 13, 777–784.
- Ramesh, S., Qi, X.-J., Wildey, G.M., Robinson, J., Molkentin, J., Letterio, J., and Howe, P.H. (2008). TGF beta-mediated BIM expression and apoptosis are regulated through SMAD3-dependent expression of the MAPK phosphatase MKP2. *EMBO Rep.* 9, 990–997.
- Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. U. S. A.* 97, 5237–5242.
- Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr. Opin. Genet. Dev.* 22, 148–155.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425–432.
- Rhim, A.D., Mirek, E.T., Aiello, N.M., Maitra, A., Bailey, J.M., McAllister, F., Reichert, M., Beatty, G.L., Rustgi, A.K., Vonderheide, R.H., et al. (2012). EMT and dissemination precede pancreatic tumor formation. *Cell* 148, 349–361.
- Richardson, G.D., Robson, C.N., Lang, S.H., Neal, D.E., Maitland, N.J., and Collins, A.T. (2004). CD133, a novel marker for human prostatic epithelial stem cells. *J. Cell Sci.* 117, 3539–3545.
- Riggs, A.D., and Xiong, Z. (2004). Methylation and epigenetic fidelity. *Proc. Natl. Acad. Sci. U. S. A.* 101, 4–5.
- Rosenberg, D., Ilic, Z., Yin, L., and Sell, S. (2000). Proliferation of hepatic lineage cells of normal C57BL and interleukin-6 knockout mice after cocaine-induced periportal injury. *Hepatol. Baltim. Md* 31, 948–955.
- Roskams, T. (2003). Progenitor cell involvement in cirrhotic human liver diseases: from controversy to consensus. *J. Hepatol.* 39, 431–434.
- Roskams, T. (2006). Liver stem cells and their implication in hepatocellular and cholangiocarcinoma. *Oncogene* 25, 3818–3822.
- Roskams, T., and Kojiro, M. (2010). Pathology of early hepatocellular carcinoma: conventional and molecular diagnosis. *Semin. Liver Dis.* 30, 17–25.
- Roskams, T., Yang, S.Q., Koteish, A., Durnez, A., DeVos, R., Huang, X., Achten, R., Verslype, C., and Diehl, A.M. (2003a). Oxidative stress and oval cell accumulation in mice and humans with alcoholic and nonalcoholic fatty liver disease. *Am. J. Pathol.* 163, 1301–1311.
- Roskams, T.A., Libbrecht, L., and Desmet, V.J. (2003b). Progenitor cells in diseased human liver. *Semin. Liver Dis.* 23, 385–396.
- Rountree, M.R., Bachman, K.E., Herman, J.G., and Baylin, S.B. (2001). DNA methylation, chromatin inheritance, and cancer. *Oncogene* 20, 3156–3165.
- Saccani, S., Pantano, S., and Natoli, G. (2002). p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment. *Nat. Immunol.* 3, 69–75.
- Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2002). Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 99, 10060–10065.
- Saito, Y., Kanai, Y., Nakagawa, T., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2003). Increased protein expression of DNA methyltransferase (DNMT) 1 is significantly correlated with the malignant potential and poor prognosis of human hepatocellular carcinomas. *Int. J. Cancer J. Int. Cancer* 105, 527–532.
- Sakaki-Yumoto, M., Katsuno, Y., and Derynck, R. (2013). TGF- $\beta$  family signaling in stem cells. *Biochim. Biophys. Acta* 1830, 2280–2296.
- Salnikov, A.V., Kusumawidjaja, G., Rausch, V., Bruns, H., Gross, W., Khamidjanov, A., Ryschich, E., Gebhard, M.-M., Moldenhauer, G., Büchler, M.W., et al. (2009). Cancer stem cell marker expression in hepatocellular carcinoma and liver metastases is not sufficient as single prognostic parameter. *Cancer Lett.* 275, 185–193.
- Sampieri, K., and Fodde, R. (2012). Cancer stem cells and metastasis. *Semin. Cancer Biol.* 22, 187–193.

- Sasaki, A., Kamiyama, T., Yokoo, H., Nakanishi, K., Kubota, K., Haga, H., Matsushita, M., Ozaki, M., Matsuno, Y., and Todo, S. (2010). Cytoplasmic expression of CD133 is an important risk factor for overall survival in hepatocellular carcinoma. *Oncol. Rep.* 24, 537–546.
- Sasaki, Y., Tsujiuchi, T., Murata, N., Tsutsumi, M., and Konishi, Y. (2001). Alterations of the transforming growth factor-beta signaling pathway in hepatocellular carcinomas induced endogenously and exogenously in rats. *Jpn. J. Cancer Res. Gann* 92, 16–22.
- Sawan, C., Vaissière, T., Murr, R., and Herceg, Z. (2008). Epigenetic drivers and genetic passengers on the road to cancer. *Mutat. Res.* 642, 1–13.
- Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U. S. A.* 103, 1412–1417.
- Sceusi, E.L., Loose, D.S., and Wray, C.J. (2011). Clinical implications of DNA methylation in hepatocellular carcinoma. *HPB* 13, 369–376.
- Schagdarsurengin, U., Wilkens, L., Steinemann, D., Flemming, P., Kreipe, H.H., Pfeifer, G.P., Schlegelberger, B., and Dammann, R. (2003). Frequent epigenetic inactivation of the RASSF1A gene in hepatocellular carcinoma. *Oncogene* 22, 1866–1871.
- Schiapparelli, P., Enguita-Germán, M., Balbuena, J., Rey, J.A., Lázcoz, P., and Castresana, J.S. (2010). Analysis of stemness gene expression and CD133 abnormal methylation in neuroblastoma cell lines. *Oncol. Rep.* 24, 1355–1362.
- Schnur, J., Nagy, P., Sebestyén, A., Schaff, Z., and Thorgeirsson, S.S. (1999). Chemical hepatocarcinogenesis in transgenic mice overexpressing mature TGF beta-1 in liver. *Eur. J. Cancer Oxf. Engl.* 1990 35, 1842–1845.
- Schulte-Hermann, R., Bursch, W., and Grasl-Kraupp, B. (1995). Active cell death (apoptosis) in liver biology and disease. *Prog. Liver Dis.* 13, 1–35.
- Seisenberger, S., Peat, J.R., and Reik, W. (2013). Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr. Opin. Cell Biol.* 25, 281–288.
- Sekimoto, G., Matsuzaki, K., Yoshida, K., Mori, S., Murata, M., Seki, T., Matsui, H., Fujisawa, J., and Okazaki, K. (2007). Reversible Smad-dependent signaling between tumor suppression and oncogenesis. *Cancer Res.* 67, 5090–5096.
- Sell, S. (2010). On the stem cell origin of cancer. *Am. J. Pathol.* 176, 2584–2494.
- Sheahan, S., Bellamy, C.O., Dunbar, D.R., Harrison, D.J., and Prost, S. (2007). Deficiency of G1 regulators P53, P21Cip1 and/or pRb decreases hepatocyte sensitivity to TGFbeta cell cycle arrest. *BMC Cancer* 7, 215.
- Sheikh, M.Y., Choi, J., Qadri, I., Friedman, J.E., and Sanyal, A.J. (2008). Hepatitis C virus infection: molecular pathways to metabolic syndrome. *Hepatol. Baltim. Md* 47, 2127–2133.
- Shen, H., and Laird, P.W. (2013). Interplay between the cancer genome and epigenome. *Cell* 153, 38–55.
- Shen, J., Wang, S., Zhang, Y.-J., Kappil, M., Wu, H.-C., Kibriya, M.G., Wang, Q., Jasmine, F., Ahsan, H., Lee, P.-H., et al. (2012). Genome-wide DNA methylation profiles in hepatocellular carcinoma. *Hepatol. Baltim. Md* 55, 1799–1808.
- Shen, L., Ahuja, N., Shen, Y., Habib, N.A., Toyota, M., Rashid, A., and Issa, J.-P.J. (2002). DNA methylation and environmental exposures in human hepatocellular carcinoma. *J. Natl. Cancer Inst.* 94, 755–761.
- Shigekawa, M., Takehara, T., Kodama, T., Hikita, H., Shimizu, S., Li, W., Miyagi, T., Hosui, A., Tatsumi, T., Ishida, H., et al. (2011). Involvement of STAT3-regulated hepatic soluble factors in attenuation of stellate cell activity and liver fibrogenesis in mice. *Biochem. Biophys. Res. Commun.* 406, 614–620.
- Shiraha, H., Yamamoto, K., and Namba, M. (2013). Human hepatocyte carcinogenesis (review). *Int. J. Oncol.* 42, 1133–1138.
- Shirai, Y., Kawata, S., Tamura, S., Ito, N., Tsushima, H., Takaishi, K., Kiso, S., and Matsuzawa, Y. (1994). Plasma transforming growth factor-beta 1 in patients with hepatocellular carcinoma. Comparison with chronic liver diseases. *Cancer* 73, 2275–2279.
- Shitani, M., Sasaki, S., Akutsu, N., Takagi, H., Suzuki, H., Nojima, M., Yamamoto, H., Tokino, T., Hirata, K., Imai, K., et al. (2012). Genome-wide analysis of DNA methylation identifies novel cancer-related genes in hepatocellular carcinoma. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 33, 1307–1317.

- Shmelkov, S.V., Jun, L., St Clair, R., McGarrigle, D., Derderian, C.A., Usenko, J.K., Costa, C., Zhang, F., Guo, X., and Rafii, S. (2004). Alternative promoters regulate transcription of the gene that encodes stem cell surface protein AC133. *Blood* 103, 2055–2061.
- Singal, A.K., and Anand, B.S. (2007). Mechanisms of synergy between alcohol and hepatitis C virus. *J. Clin. Gastroenterol.* 41, 761–772.
- Singh, S.K., Clarke, I.D., Terasaki, M., Bonn, V.E., Hawkins, C., Squire, J., and Dirks, P.B. (2003). Identification of a cancer stem cell in human brain tumors. *Cancer Res.* 63, 5821–5828.
- Sokol, R.J. (2002). Liver cell injury and fibrosis. *J. Pediatr. Gastroenterol. Nutr.* 35 Suppl 1, S7–10.
- Song, L.H., Binh, V.Q., Duy, D.N., Kun, J.F.J., Bock, T.C., Kremsner, P.G., and Luty, A.J.F. (2003). Serum cytokine profiles associated with clinical presentation in Vietnamese infected with hepatitis B virus. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 28, 93–103.
- Song, M.-A., Tiirikainen, M., Kwee, S., Okimoto, G., Yu, H., and Wong, L.L. (2013). Elucidating the landscape of aberrant DNA methylation in hepatocellular carcinoma. *PLoS One* 8, e55761.
- Song, W., Li, H., Tao, K., Li, R., Song, Z., Zhao, Q., Zhang, F., and Dou, K. (2008). Expression and clinical significance of the stem cell marker CD133 in hepatocellular carcinoma. *Int. J. Clin. Pract.* 62, 1212–1218.
- Soussan, P., Tuveri, R., Nalpas, B., Garreau, F., Zavala, F., Masson, A., Pol, S., Brechot, C., and Kremsdorf, D. (2003). The expression of hepatitis B spliced protein (HBSP) encoded by a spliced hepatitis B virus RNA is associated with viral replication and liver fibrosis. *J. Hepatol.* 38, 343–348.
- De Souza, A.T., Hankins, G.R., Washington, M.K., Orton, T.C., and Jirtle, R.L. (1995). M6P/IGF2R gene is mutated in human hepatocellular carcinomas with loss of heterozygosity. *Nat. Genet.* 11, 447–449.
- Spender, L.C., and Inman, G.J. (2009). TGF-beta induces growth arrest in Burkitt lymphoma cells via transcriptional repression of E2F-1. *J. Biol. Chem.* 284, 1435–1442.
- Sproul, D., Kitchen, R.R., Nestor, C.E., Dixon, J.M., Sims, A.H., Harrison, D.J., Ramsahoye, B.H., and Meehan, R.R. (2012). Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol.* 13, R84.
- Standring, S. (2008). *Gray's Anatomy: The Anatomical Basis of Clinical Practice* (Churchill Livingstone/Elsevier).
- Stefanska, B., Bouzelmat, A., Huang, J., Suderman, M., Hallett, M., Han, Z.-G., Al-Mahtab, M., Akbar, S.M.F., Khan, W.A., Raqib, R., et al. (2013a). Discovery and Validation of DNA Hypomethylation Biomarkers for Liver Cancer Using HRM-Specific Probes. *PLoS One* 8, e68439.
- Stefanska, B., Suderman, M., Machnes, Z., Bhattacharyya, B., Hallett, M., and Szyf, M. (2013b). Transcription onset of genes critical in liver carcinogenesis is epigenetically regulated by methylated DNA binding protein MBD2. *Carcinogenesis*.
- Stork, P., Loda, M., Bosari, S., Wiley, B., Poppenhusen, K., and Wolfe, H. (1991). Detection of K-ras mutations in pancreatic and hepatic neoplasms by non-isotopic mismatched polymerase chain reaction. *Oncogene* 6, 857–862.
- Streetz, K.L., Tacke, F., Leifeld, L., Wüstefeld, T., Graw, A., Klein, C., Kamino, K., Spengler, U., Kreipe, H., Kubicka, S., et al. (2003). Interleukin 6/gp130-dependent pathways are protective during chronic liver diseases. *Hepatol. Baltim. Md* 38, 218–229.
- Suetsugu, A., Nagaki, M., Aoki, H., Motohashi, T., Kunisada, T., and Moriwaki, H. (2006). Characterization of CD133+ hepatocellular carcinoma cells as cancer stem/progenitor cells. *Biochem. Biophys. Res. Commun.* 351, 820–824.
- Sugihara, E., and Saya, H. (2013). Complexity of cancer stem cells. *Int. J. Cancer J. Int. Cancer* 132, 1249–1259.
- Sulic, S., Panic, L., Dikic, I., and Volarevic, S. (2005). Deregulation of cell growth and malignant transformation. *Croat. Med. J.* 46, 622–638.
- Sun, B., and Karin, M. (2013). Inflammation and liver tumorigenesis. *Front. Med.* 7, 242–254.
- Sun, Y., Lowther, W., Kato, K., Bianco, C., Kenney, N., Strizzi, L., Raafat, D., Hirota, M., Khan, N.I., Bargo, S., et al. (2005). Notch4 intracellular domain binding to Smad3 and inhibition of the TGF-beta signaling. *Oncogene* 24, 5365–5374.
- Szabó, E., Páska, C., Kaposi Novák, P., Schaff, Z., and Kiss, A. (2004). Similarities and differences in hepatitis B

- and C virus induced hepatocarcinogenesis. *Pathol. Oncol. Res.* 10, 5–11.
- Szotek, P.P., Pieretti-Vanmarcke, R., Masiakos, P.T., Dinulescu, D.M., Connolly, D., Foster, R., Dombkowski, D., Preffer, F., Maclaughlin, D.T., and Donahoe, P.K. (2006). Ovarian cancer side population defines cells with stem cell-like characteristics and Mullerian Inhibiting Substance responsiveness. *Proc. Natl. Acad. Sci. U. S. A.* 103, 11154–11159.
- Tabu, K., Sasai, K., Kimura, T., Wang, L., Aoyanagi, E., Kohsaka, S., Tanino, M., Nishihara, H., and Tanaka, S. (2008). Promoter hypomethylation regulates CD133 expression in human gliomas. *Cell Res.* 18, 1037–1046.
- Tacke, F., Luedde, T., and Trautwein, C. (2009). Inflammatory pathways in liver homeostasis and liver injury. *Clin. Rev. Allergy Immunol.* 36, 4–12.
- Tada, M., Omata, M., and Ohto, M. (1990). Analysis of ras gene mutations in human hepatic malignant tumors by polymerase chain reaction and direct sequencing. *Cancer Res.* 50, 1121–1124.
- Takebe, N., Harris, P.J., Warren, R.Q., and Ivy, S.P. (2011). Targeting cancer stem cells by inhibiting Wnt, Notch, and Hedgehog pathways. *Nat. Rev. Clin. Oncol.* 8, 97–106.
- Takubo, K., Goda, N., Yamada, W., Iriuchishima, H., Ikeda, E., Kubota, Y., Shima, H., Johnson, R.S., Hirao, A., Suematsu, M., et al. (2010). Regulation of the HIF-1 $\alpha$  level is essential for hematopoietic stem cells. *Cell Stem Cell* 7, 391–402.
- Tang, K.H., Ma, S., Lee, T.K., Chan, Y.P., Kwan, P.S., Tong, C.M., Ng, I.O., Man, K., To, K.-F., Lai, P.B., et al. (2012). CD133(+) liver tumor-initiating cells promote tumor angiogenesis, growth, and self-renewal through neurotensin/interleukin-8/CXCL1 signaling. *Hepatology*. Baltimore, Md 55, 807–820.
- Tang, Y., Kitisin, K., Jogunoori, W., Li, C., Deng, C.-X., Mueller, S.C., Ransom, H.W., Rashid, A., He, A.R., Mendelson, J.S., et al. (2008a). Progenitor/stem cells give rise to liver cancer due to aberrant TGF- $\beta$  and IL-6 signaling. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2445–2450.
- Tang, Y., Kitisin, K., Jogunoori, W., Li, C., Deng, C.-X., Mueller, S.C., Ransom, H.W., Rashid, A., He, A.R., Mendelson, J.S., et al. (2008b). Progenitor/stem cells give rise to liver cancer due to aberrant TGF- $\beta$  and IL-6 signaling. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2445–2450.
- Tang, Y., Urs, S., Boucher, J., Bernaiche, T., Venkatesh, D., Spicer, D.B., Vary, C.P.H., and Liaw, L. (2010). Notch and transforming growth factor- $\beta$  (TGF $\beta$ ) signaling pathways cooperatively regulate vascular smooth muscle cell differentiation. *J. Biol. Chem.* 285, 17556–17563.
- Tangkijvanich, P., Hourpai, N., Rattanatanyong, P., Wisedopas, N., Mahachai, V., and Mutirangura, A. (2007). Serum LINE-1 hypomethylation as a potential prognostic marker for hepatocellular carcinoma. *Clin. Chim. Acta Int. J. Clin. Chem.* 379, 127–133.
- Tannapfel, A., Busse, C., Weinans, L., Benicke, M., Katalinic, A., Geissler, F., Hauss, J., and Wittekind, C. (2001). INK4a-ARF alterations and p53 mutations in hepatocellular carcinomas. *Oncogene* 20, 7104–7109.
- Tanno, T., and Matsui, W. (2011). Development and maintenance of cancer stem cells under chronic inflammation. *J. Nippon Med. Sch. Nippon Ika Daigaku Zasshi* 78, 138–145.
- Taub, R. (2004). Liver regeneration: from myth to mechanism. *Nat. Rev. Mol. Cell Biol.* 5, 836–847.
- Terui, K., Enosawa, S., Haga, S., Zhang, H.Q., Kuroda, H., Kouchi, K., Matsunaga, T., Yoshida, H., Engelhardt, J.F., Irani, K., et al. (2004). Stat3 confers resistance against hypoxia/reoxygenation-induced oxidative injury in hepatocytes through upregulation of Mn-SOD. *J. Hepatology*. 41, 957–965.
- Thenappan, A., Li, Y., Kitisin, K., Rashid, A., Shetty, K., Johnson, L., and Mishra, L. (2010). Role of transforming growth factor  $\beta$  signaling and expansion of progenitor cells in regenerating liver. *Hepatology*. Baltimore, Md 51, 1373–1382.
- Thillainadesan, G., Chitilian, J.M., Iovic, M., Ablack, J.N.G., Mymryk, J.S., Tini, M., and Torchia, J. (2012). TGF- $\beta$ -dependent active demethylation and expression of the p15<sup>ink4b</sup> tumor suppressor are impaired by the ZNF217/CoREST complex. *Mol. Cell* 46, 636–649.
- Thimme, R., Wieland, S., Steiger, C., Ghayeb, J., Reimann, K.A., Purcell, R.H., and Chisari, F.V. (2003). CD8(+) T cells mediate viral clearance and disease pathogenesis during acute hepatitis B virus infection. *J. Virol.* 77, 68–76.
- Thomas, D., and Zoulim, F. (2012). New challenges in viral hepatitis. *Gut* 61 Suppl 1, i1–5.

- Tirino, V., Camerlingo, R., Franco, R., Malanga, D., La Rocca, A., Viglietto, G., Rocco, G., and Pirozzi, G. (2009). The role of CD133 in the identification and characterisation of tumour-initiating cells in non-small-cell lung cancer. *Eur. J. Cardio-Thorac. Surg. Off. J. Eur. Assoc. Cardio-Thorac. Surg.* 36, 446–453.
- Tomuleasa, C., Soritau, O., Rus-Ciucă, D., Pop, T., Todea, D., Mosteanu, O., Pintea, B., Foris, V., Susman, S., Kacsó, G., et al. (2010). Isolation and characterization of hepatic cancer cells with stem-like properties from hepatocellular carcinoma. *J. Gastrointest. Liver Dis. JGLD* 19, 61–67.
- Tong, C.M., Ma, S., and Guan, X.-Y. (2011). Biology of hepatic cancer stem cells. *J. Gastroenterol. Hepatol.* 26, 1229–1237.
- Trowbridge, J.J., Sinha, A.U., Zhu, N., Li, M., Armstrong, S.A., and Orkin, S.H. (2012). Haploinsufficiency of *Dnmt1* impairs leukemia stem cell function through derepression of bivalent chromatin domains. *Genes Dev.* 26, 344–349.
- Tsai, S.-T., Tsou, C.-C., Mao, W.-Y., Chang, W.-C., Han, H.-Y., Hsu, W.-L., Li, C.-L., Shen, C.-N., and Chen, C.-H. (2012). Label-free quantitative proteomics of CD133-positive liver cancer stem cells. *Proteome Sci.* 10, 69.
- Tsushima, H., Kawata, S., Tamura, S., Ito, N., Shirai, Y., Kiso, S., Doi, Y., Yamada, A., Oshikawa, O., and Matsuzawa, Y. (1999). Reduced plasma transforming growth factor-beta1 levels in patients with chronic hepatitis C after interferon-alpha therapy: association with regression of hepatic fibrosis. *J. Hepatol.* 30, 1–7.
- Uchida, N., Buck, D.W., He, D., Reitsma, M.J., Masek, M., Phan, T.V., Tsukamoto, A.S., Gage, F.H., and Weissman, I.L. (2000). Direct isolation of human central nervous system stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 97, 14720–14725.
- Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., et al. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res.* 23, 555–567.
- Vergani, D., and Mieli-Vergani, G. (2008). Aetiopathogenesis of autoimmune hepatitis. *World J. Gastroenterol.* WJG 14, 3306–3312.
- Vincent-Salomon, A., and Thiery, J.P. (2003). Host microenvironment in breast cancer development: epithelial-mesenchymal transition in breast cancer development. *Breast Cancer Res. BCR* 5, 101–106.
- Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.-M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439, 871–874.
- Visvader, J.E., and Lindeman, G.J. (2012). Cancer stem cells: current status and evolving complexities. *Cell Stem Cell* 10, 717–728.
- Wakabayashi, Y., Tamiya, T., Takada, I., Fukaya, T., Sugiyama, Y., Inoue, N., Kimura, A., Morita, R., Kashiwagi, I., Takimoto, T., et al. (2011). Histone 3 lysine 9 (H3K9) methyltransferase recruitment to the interleukin-2 (IL-2) promoter is a mechanism of suppression of IL-2 transcription by the transforming growth factor- $\beta$ -Smad pathway. *J. Biol. Chem.* 286, 35456–35465.
- Wang, G., Matsuura, I., He, D., and Liu, F. (2009). Transforming growth factor- $\beta$ -inducible phosphorylation of Smad3. *J. Biol. Chem.* 284, 9663–9673.
- Wang, J., Duncan, D., Shi, Z., and Zhang, B. (2013). WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res.* 41, W77–83.
- Wang, X., Foster, M., Al-Dhalimy, M., Lagasse, E., Finegold, M., and Grompe, M. (2003). The origin and liver repopulating capacity of murine oval cells. *Proc. Natl. Acad. Sci. U. S. A.* 100 Suppl 1, 11881–11888.
- Wang, Y., Liu, Y., Malek, S.N., Zheng, P., and Liu, Y. (2011a). Targeting HIF1 $\alpha$  eliminates cancer stem cells in hematological malignancies. *Cell Stem Cell* 8, 399–411.
- Wang, Y., Yu, Y., Tsuyada, A., Ren, X., Wu, X., Stubblefield, K., Rankin-Gee, E.K., and Wang, S.E. (2011b). Transforming growth factor- $\beta$  regulates the sphere-initiating stem cell-like feature in breast cancer through miRNA-181 and ATM. *Oncogene* 30, 1470–1480.
- Wang, Y., Li, L., Guo, X., Jin, X., Sun, W., Zhang, X., and Xu, R.C. (2012). Interleukin-6 signaling regulates anchorage-independent growth, proliferation, adhesion and invasion in human ovarian cancer cells. *Cytokine* 59, 228–236.
- Waxman, S., and Wurmbach, E. (2007). De-regulation of common housekeeping genes in hepatocellular



- carcinoma. *BMC Genomics* 8, 243.
- Wei, X.D., Zhou, L., Cheng, L., Tian, J., Jiang, J.J., and Maccallum, J. (2009). In vivo investigation of CD133 as a putative marker of cancer stem cells in Hep-2 cell line. *Head Neck* 31, 94–101.
- Wei, Y., Jiang, Y., Zou, F., Liu, Y., Wang, S., Xu, N., Xu, W., Cui, C., Xing, Y., Liu, Y., et al. (2013). Activation of PI3K/Akt pathway by CD133-p85 interaction promotes tumorigenic capacity of glioma stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6829–6834.
- Weigmann, A., Corbeil, D., Hellwig, A., and Huttner, W.B. (1997). Prominin, a novel microvilli-specific polytopic membrane protein of the apical surface of epithelial cells, is targeted to plasmalemmal protrusions of non-epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12425–12430.
- Widschwendter, M., Fiegl, H., Egle, D., Mueller-Holzner, E., Spizzo, G., Marth, C., Weisenberger, D.J., Campan, M., Young, J., Jacobs, I., et al. (2007). Epigenetic stem cell signature in cancer. *Nat. Genet.* 39, 157–158.
- Wild, C.P., and Gong, Y.Y. (2010). Mycotoxins and human disease: a largely ignored global health issue. *Carcinogenesis* 31, 71–82.
- Wisse, E., van't Noordende, J.M., van der Meulen, J., and Daems, W.T. (1976). The pit cell: description of a new type of cell occurring in rat liver sinusoids and peripheral blood. *Cell Tissue Res.* 173, 423–435.
- Wu, P.C., Fang, J.W., Lau, V.K., Lai, C.L., Lo, C.K., and Lau, J.Y. (1996). Classification of hepatocellular carcinoma according to hepatocellular and biliary differentiation markers. Clinical and biological implications. *Am. J. Pathol.* 149, 1167–1175.
- Wu, Y., Li, J., Sun, C.Y., Zhou, Y., Zhao, Y.F., and Zhang, S.J. (2012). Epigenetic inactivation of the canonical Wnt antagonist secreted frizzled-related protein 1 in hepatocellular carcinoma cells. *Neoplasia* 59, 326–332.
- Wu, Y., Antony, S., Meitzler, J.L., and Doroshow, J.H. (2013). Molecular mechanisms underlying chronic inflammation-associated cancers. *Cancer Lett.*
- Xiao, X., Gang, Y., Gu, Y., Zhao, L., Chu, J., Zhou, J., Cai, X., Zhang, H., Xu, L., Nie, Y., et al. (2012). Osteopontin contributes to TGF- $\beta$ 1 mediated hepatic stellate cell activation. *Dig. Dis. Sci.* 57, 2883–2891.
- Xu, X., Liu, R.-F., Zhang, X., Huang, L.-Y., Chen, F., Fei, Q.-L., and Han, Z.-G. (2012). DLK1 as a potential target against cancer stem/progenitor cells of hepatocellular carcinoma. *Mol. Cancer Ther.* 11, 629–638.
- Yamada, T., De Souza, A.T., Finkelstein, S., and Jirtle, R.L. (1997). Loss of the gene encoding mannose 6-phosphate/insulin-like growth factor II receptor is an early event in liver carcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 94, 10351–10355.
- Yamada, Y., Jackson-Grusby, L., Linhart, H., Meissner, A., Eden, A., Lin, H., and Jaenisch, R. (2005). Opposing effects of DNA hypomethylation on intestinal and liver carcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 102, 13580–13585.
- Yamagata, H., Matsuzaki, K., Mori, S., Yoshida, K., Tahashi, Y., Furukawa, F., Sekimoto, G., Watanabe, T., Uemura, Y., Sakaida, N., et al. (2005). Acceleration of Smad2 and Smad3 phosphorylation via c-Jun NH(2)-terminal kinase during human colorectal carcinogenesis. *Cancer Res.* 65, 157–165.
- Yamamoto, Y., Verma, U.N., Prajapati, S., Kwak, Y.-T., and Gaynor, R.B. (2003). Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression. *Nature* 423, 655–659.
- Yamashita, T., Forgues, M., Wang, W., Kim, J.W., Ye, Q., Jia, H., Budhu, A., Zanetti, K.A., Chen, Y., Qin, L.-X., et al. (2008). EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma. *Cancer Res.* 68, 1451–1461.
- Yamashita, T., Honda, M., Nakamoto, Y., Baba, M., Nio, K., Hara, Y., Zeng, S.S., Hayashi, T., Kondo, M., Takatori, H., et al. (2013). Discrete nature of EpCAM<sup>+</sup> and CD90<sup>+</sup> cancer stem cells in human hepatocellular carcinoma. *Hepatology* 57, 1484–1497.
- Yamazaki, K., Masugi, Y., and Sakamoto, M. (2011). Molecular pathogenesis of hepatocellular carcinoma: altering transforming growth factor- $\beta$  signaling in hepatocarcinogenesis. *Dig. Dis. Basel Switz.* 29, 284–288.
- Yan, B., Peng, Y., and Li, C.-Y. (2009). Molecular analysis of genetic instability caused by chronic inflammation. *Methods Mol. Biol.* Clifton NJ 512, 15–28.
- Yang, G., Quan, Y., Wang, W., Fu, Q., Wu, J., Mei, T., Li, J., Tang, Y., Luo, C., Ouyang, Q., et al. (2012). Dynamic equilibrium between cancer stem cells and non-stem cancer cells in human SW620 and MCF-7 cancer cell

- populations. *Br. J. Cancer* 106, 1512–1519.
- Yang, L., Pang, Y., and Moses, H.L. (2010a). TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol.* 31, 220–227.
- Yang, S., Koteish, A., Lin, H., Huang, J., Roskams, T., Dawson, V., and Diehl, A.M. (2004). Oval cells compensate for damage and replicative senescence of mature hepatocytes in mice with fatty liver disease. *Hepatol. Baltim. Md* 39, 403–411.
- Yang, W., Yan, H.-X., Chen, L., Liu, Q., He, Y.-Q., Yu, L.-X., Zhang, S.-H., Huang, D.-D., Tang, L., Kong, X.-N., et al. (2008a). Wnt/beta-catenin signaling contributes to activation of normal and tumorigenic liver progenitor cells. *Cancer Res.* 68, 4287–4295.
- Yang, X.-R., Xu, Y., Yu, B., Zhou, J., Qiu, S.-J., Shi, G.-M., Zhang, B.-H., Wu, W.-Z., Shi, Y.-H., Wu, B., et al. (2010b). High expression levels of putative hepatic stem/progenitor cell biomarkers related to tumour angiogenesis and poor prognosis of hepatocellular carcinoma. *Gut* 59, 953–962.
- Yang, Y.-A., Zhang, G.-M., Feigenbaum, L., and Zhang, Y.E. (2006). Smad3 reduces susceptibility to hepatocarcinoma by sensitizing hepatocytes to apoptosis through downregulation of Bcl-2. *Cancer Cell* 9, 445–457.
- Yang, Z., Zhang, L., Ma, A., Liu, L., Li, J., Gu, J., and Liu, Y. (2011). Transient mTOR inhibition facilitates continuous growth of liver tumors by modulating the maintenance of CD133+ cell populations. *PloS One* 6, e28405.
- Yang, Z.F., Ho, D.W., Ng, M.N., Lau, C.K., Yu, W.C., Ngai, P., Chu, P.W.K., Lam, C.T., Poon, R.T.P., and Fan, S.T. (2008b). Significance of CD90+ cancer stem cells in human liver cancer. *Cancer Cell* 13, 153–166.
- Yasuda, H., Soejima, K., Watanabe, H., Kawada, I., Nakachi, I., Yoda, S., Nakayama, S., Satomi, R., Ikemura, S., Terai, H., et al. (2010). Distinct epigenetic regulation of tumor suppressor genes in putative cancer stem cells of solid tumors. *Int. J. Oncol.* 37, 1537–1546.
- Yeh, K.-T., Chen, T.-H., Yang, H.-W., Chou, J.-L., Chen, L.-Y., Yeh, C.-M., Chen, Y.-H., Lin, R.-I., Su, H.-Y., Chen, G.C.W., et al. (2011). Aberrant TGF $\beta$ /SMAD4 signaling contributes to epigenetic silencing of a putative tumor suppressor, RunX1T1 in ovarian cancer. *Epigenetics Off. J. DNA Methylation Soc.* 6, 727–739.
- Yeoh, G.C.T., Ernst, M., Rose-John, S., Akhurst, B., Payne, C., Long, S., Alexander, W., Croker, B., Grail, D., and Matthews, V.B. (2007). Opposing roles of gp130-mediated STAT-3 and ERK-1/ 2 signaling in liver progenitor cell migration and proliferation. *Hepatol. Baltim. Md* 45, 486–494.
- Yi, J.M., Tsai, H.-C., Glöckner, S.C., Lin, S., Ohm, J.E., Easwaran, H., James, C.D., Costello, J.F., Riggins, G., Eberhart, C.G., et al. (2008). Abnormal DNA methylation of CD133 in colorectal and glioblastoma tumors. *Cancer Res.* 68, 8094–8103.
- Yi, S.-Y., Hao, Y.-B., Nan, K.-J., and Fan, T.-L. (2013). Cancer stem cells niche: a target for novel cancer therapeutics. *Cancer Treat. Rev.* 39, 290–296.
- Yin, A.H., Miraglia, S., Zanjani, E.D., Almeida-Porada, G., Ogawa, M., Leary, A.G., Olweus, J., Kearney, J., and Buck, D.W. (1997). AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 90, 5002–5012.
- Yin, S., Li, J., Hu, C., Chen, X., Yao, M., Yan, M., Jiang, G., Ge, C., Xie, H., Wan, D., et al. (2007). CD133 positive hepatocellular carcinoma cells possess high capacity for tumorigenicity. *Int. J. Cancer J. Int. Cancer* 120, 1444–1450.
- Yoo, J., Ghiassi, M., Jirmanova, L., Balliet, A.G., Hoffman, B., Fornace, A.J., Jr, Liebermann, D.A., Bottinger, E.P., and Roberts, A.B. (2003). Transforming growth factor-beta-induced apoptosis is mediated by Smad-dependent expression of GADD45b through p38 activation. *J. Biol. Chem.* 278, 43001–43007.
- Yoon, D.S., Jeong, J., Park, Y.N., Kim, K.S., Kwon, S.W., Chi, H.S., Park, C., and Kim, B.R. (1999). Expression of biliary antigen and its clinical significance in hepatocellular carcinoma. *Yonsei Med. J.* 40, 472–477.
- You, H., Ding, W., and Rountree, C.B. (2010a). Epigenetic regulation of cancer stem cell marker CD133 by transforming growth factor-beta. *Hepatol. Baltim. Md* 51, 1635–1644.
- You, H., Ding, W., and Rountree, C.B. (2010b). Epigenetic regulation of cancer stem cell marker CD133 by transforming growth factor-beta. *Hepatol. Baltim. Md* 51, 1635–1644.
- Yu, F., Jiao, Y., Zhu, Y., Wang, Y., Zhu, J., Cui, X., Liu, Y., He, Y., Park, E.-Y., Zhang, H., et al. (2012a). MicroRNA

- 34c Gene Down-regulation via DNA Methylation Promotes Self-renewal and Epithelial-Mesenchymal Transition in Breast Tumor-initiating Cells. *J. Biol. Chem.* 287, 465–473.
- Yu, F., Jiao, Y., Zhu, Y., Wang, Y., Zhu, J., Cui, X., Liu, Y., He, Y., Park, E.-Y., Zhang, H., et al. (2012b). MicroRNA 34c gene down-regulation via DNA methylation promotes self-renewal and epithelial-mesenchymal transition in breast tumor-initiating cells. *J. Biol. Chem.* 287, 465–473.
- Yu, H., Kortylewski, M., and Pardoll, D. (2007). Crosstalk between cancer and immune cells: role of STAT3 in the tumour microenvironment. *Nat. Rev. Immunol.* 7, 41–51.
- Yu, Y., Flint, A., Dvorin, E.L., and Bischoff, J. (2002). AC133-2, a novel isoform of human AC133 stem cell antigen. *J. Biol. Chem.* 277, 20711–20716.
- Yuan, F., Zhou, W., Zou, C., Zhang, Z., Hu, H., Dai, Z., and Zhang, Y. (2010). Expression of Oct4 in HCC and modulation to wnt/ $\beta$ -catenin and TGF- $\beta$  signal pathways. *Mol. Cell. Biochem.* 343, 155–162.
- Yuen, M.F., Norris, S., Evans, L.W., Langley, P.G., and Hughes, R.D. (2002). Transforming growth factor-beta 1, activin and follistatin in patients with hepatocellular carcinoma and patients with alcoholic cirrhosis. *Scand. J. Gastroenterol.* 37, 233–238.
- Zardo, G., Fazi, F., Travaglini, L., and Nervi, C. (2005). Dynamic and reversibility of heterochromatic gene silencing in human disease. *Cell Res.* 15, 679–690.
- Zekri, A.-R.N., Ashour, M.S.E.-D., Hassan, A., Alam El-Din, H.M., El-Shehaby, A.M.R., and Abu-Shady, M.A. (2005). Cytokine profile in Egyptian hepatitis C virus genotype-4 in relation to liver disease progression. *World J. Gastroenterol.* WJG 11, 6624–6630.
- Zen, C., Zen, Y., Mitry, R.R., Corbeil, D., Karbanová, J., O'Grady, J., Karani, J., Kane, P., Heaton, N., Portmann, B.C., et al. (2011). Mixed phenotype hepatocellular carcinoma after transarterial chemoembolization and liver transplantation. *Liver Transplant. Off. Publ. Am. Assoc. Study Liver Dis. Int. Liver Transplant. Soc.* 17, 943–954.
- Zhai, B., Yan, H.-X., Liu, S.-Q., Chen, L., Wu, M.-C., and Wang, H.-Y. (2008). Reduced expression of E-cadherin/catenin complex in hepatocellular carcinomas. *World J. Gastroenterol.* WJG 14, 5665–5673.
- Zhai, J.-M., Yin, X.-Y., Hou, X., Hao, X.-Y., Cai, J.-P., Liang, L.-J., and Zhang, L.-J. (2013). Analysis of the genome-wide DNA methylation profile of side population cells in hepatocellular carcinoma. *Dig. Dis. Sci.* 58, 1934–1947.
- Zhang, Y.E. (2009). Non-Smad pathways in TGF-beta signaling. *Cell Res.* 19, 128–139.
- Zhang, C., Xu, Y., Zhao, J., Fan, L., Jiang, G., Li, R., Ling, Y., Wu, M., and Wei, L. (2011). Elevated expression of the stem cell marker CD133 associated with Line-1 demethylation in hepatocellular carcinoma. *Ann. Surg. Oncol.* 18, 2373–2380.
- Zhang, H., Li, H., Xi, H.S., and Li, S. (2012a). HIF1a is required for survival maintenance of chronic myeloid leukemia stem cells. *Blood* 119, 2595–2607.
- Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G.-L., Luo, C., Jiang, H., and He, C. (2012b). Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat. Chem. Biol.* 8, 328–330.
- Zhou, B., Liu, Y., Kahn, M., Ann, D.K., Han, A., Wang, H., Nguyen, C., Flodby, P., Zhong, Q., Krishnaveni, M.S., et al. (2012a). Interactions between  $\beta$ -catenin and transforming growth factor- $\beta$  signaling pathways mediate epithelial-mesenchymal transition and are dependent on the transcriptional co-activator cAMP-response element-binding protein (CREB)-binding protein (CBP). *J. Biol. Chem.* 287, 7026–7038.
- Zhou, C., Liu, J., Tang, Y., and Liang, X. (2012b). Inflammation linking EMT and cancer stem cells. *Oral Oncol.* 48, 1068–1075.
- Zhu, Z., Hao, X., Yan, M., Yao, M., Ge, C., Gu, J., and Li, J. (2010). Cancer stem/progenitor cells are highly enriched in CD133+CD44+ population in hepatocellular carcinoma. *Int. J. Cancer J. Int. Cancer* 126, 2067–2078.
- De Zhu, J. (2005). The altered DNA methylation pattern and its implications in liver cancer. *Cell Res.* 15, 272–280.
- Zimmers, T.A., Pierce, R.H., McKillop, I.H., and Koniaris, L.G. (2003). Resolving the role of IL-6 in liver regeneration. *Hepatology* 38, 1590–1591; author reply 1591.
- Zingg, J.M., and Jones, P.A. (1997). Genetic and epigenetic aspects of DNA methylation on genome expression,

evolution, mutation and carcinogenesis. *Carcinogenesis* 18, 869–882.

## **ANNEXE I: Supplementary Tables**



**Supplementary Table 1: List of the 395 annotated CpG loci differentially methylated between CD133+ and CD133- cells.**

The list of significant CpG loci between CD133+ and CD133- cells was obtained using the BrB class comparison tool, blocked by cell lines and filtered for p-value <0.001 and averaged delta beta (between the two cell lines) >5%. Among the 608 significant loci, here are presented the ones that are associated to annotated genes (395). For each locus, information are given as presented below:

NAME OF THE GENE (probe's ID; relation to CpG island; gene region; true enhancer; averaged deltabeta between CD133+ and CD133- cells). Legend:

*Relation to CpG island:* 1= Island / N2 = N Shore / S2 = S Shore / N3 = N Shelf / S3 = S Shelf / 4 = Open sea

*Gene regions:* a= TSS1500 / b = TSS200 / c= 5'UTR / d = 1st exon / e = Body / f = 3'UTR

*Enhancer:* T = true enhancer

A2BP1(cg09243507;4;c;T;-0.078)	-	ABLM3(cg02306139;4;f;-0.074)	-	ACCN1(cg02254574;N2;d;T;-0.07)	-
ACER3(cg17838626;N2;a;0.103)	-	ACSL6(cg19096799;1;e;0.05)	-	ADAMTS17(cg10135717;S2;a;-0.06)	-
AGAP2(cg13879455;S2;a;-0.05)	-	AK5(cg11661204;4;e;T;-0.068)	-	AMBN(cg13523386;4;b;-0.121)	-
AMN1(cg01048346;4;f;-0.082)	-	AMPD3(cg11854154;1;b;-0.099)	-	AMPH(cg23092449;N2;e;-0.071)	-
AP3B2(cg20194856;4;e;-0.123)	-	APBA1(cg13580827;4;e;T;0.066)	-	ARHGAP4(cg19271175;N3;e;-0.117)	-
ARHGEF10(cg23223533;4;e;T;-0.107)	-	ARNTL(cg09527192;N2;a;0.107)	-	ASCL1(cg27420520;1;d;0.084)	-
ATP10B(cg08743199;4;e;-0.088)	-	ATP12A(cg16602799;1;c;-0.06)	-	ATP8A2(cg26153234;N2;a;-0.105)	-
B3GNT4(cg11970289;1;c;-0.097)	-	B4GALNT3(cg15580052;4;e;T;-0.085)	-	BCL11A(cg07469838;4;e;T;0.081)	-
BCL11A(cg23678058;4;e;0.067)	-	BHLHA9(cg25681339;1;d;T;0.06)	-	BICC1(cg17548395;4;e;T;-0.074)	-
BMP3(cg01285706;S2;e;T;-0.096)	-	BRSK2(cg15411034;1;e;-0.057)	-	C11orf36(cg08093323;S3;e;-0.057)	-
C14orf180(cg24102702;S3;a;-0.07)	-	C17orf85(cg03933290;4;e;0.059)	-	C18orf20(cg01722450;4;a;-0.079)	-
C19orf57(cg23925190;1;c;0.068)	-	C22orf34(cg09906324;4;b;-0.086)	-	C5orf20(cg13349425;4;b;-0.051)	-
C6orf195(cg00483640;4;f;-0.097)	-	C8orf55(cg10918419;1;a;-0.07)	-	C9orf135(cg14359824;4;b;-0.103)	-
CA8(cg00994693;1;c;T;-0.076)	-	CARD14(cg01255509;S3;a;T;-0.082)	-	CASD1(cg16194437;4;f;-0.094)	-
CAV1(cg24987440;4;e;T;-0.105)	-	CCDC102B(cg2227582;4;e;-0.089)	-	CCDC42B(cg16120422;1;e;0.063)	-
CCNH(cg27584762;S2;a;-0.06)	-	CD300LD(cg20871097;4;a;-0.06)	-	CD8A(cg13946520;1;e;T;-0.081)	-
CDGAP(cg10243989;N2;e;-0.088)	-	CDH23(cg18645316;4;e;T;-0.12)	-	CDKN1C(cg21741284;S2;a;0.056)	-
CHRNA7(cg05168573;4;e;-0.184)	-	CLCN7(cg08499756;S2;a;-0.109)	-	CLDN7(cg14034852;1;e;0.083)	-
CNNM4(cg14228484;S2;d;-0.161)	-	CNTN2(cg01537455;N2;e;-0.125)	-	COLEC11(cg19183742;4;e;-0.076)	-
CPEB1(cg17453840;1;a;-0.077)	-	CPNE7(cg09436290;S2;f;-0.07)	-	CRHR2(cg23185751;N2;e;-0.087)	-
CRTAC1(cg08649440;4;e;T;-0.134)	-	CRTAC1(cg26298855;4;e;T;-0.054)	-	CSGALNACT2(cg07178550;1;a;-0.082)	-
CSMD1(cg22619018;1;a;-0.084)	-	CSRNP3(cg02461406;4;e;T;-0.098)	-	CTBP2(cg10491546;4;c;T;-0.103)	-
CXorf27(cg02835735;4;a;0.199)	-	CYP2C8(cg14717122;4;f;-0.165)	-	DAB1(cg01074356;N3;c;-0.097)	-
DACT1(cg10796078;1;e;0.062)	-	DBX1(cg13696942;N2;e;-0.084)	-	DCLK2(cg02587405;N2;e;-0.127)	-
DCP2(cg21539223;N2;a;-0.051)	-	DGKB(cg00308440;4;e;T;-0.111)	-	DIO2(cg09005221;4;e;-0.119)	-
DLGAP2(cg16520712;N2;e;-0.087)	-	DNAH10(cg22472488;4;e;-0.092)	-	DNAH10(cg27508046;S3;e;-0.058)	-
DNAH14(cg17482224;N2;a;-0.096)	-	DNAJB13(cg23327896;4;e;-0.104)	-	DNER(cg15603812;4;e;-0.078)	-
DOCK9(cg21712331;4;e;T;-0.067)	-	DOCK9(cg27518324;4;e;T;-0.065)	-	DSCAM(cg12089511;4;e;-0.085)	-
DTX1(cg22685245;1;a;-0.063)	-	EBF2(cg18239431;N2;e;-0.096)	-	EBF3(cg12129080;N2;e;0.104)	-
EBF3(cg12509733;4;e;-0.11)	-	ENTPD4(cg26837477;N2;c;-0.082)	-	EPHX4(cg07197585;N2;a;-0.05)	-
ESRRB(cg10526223;N2;e;-0.08)	-	EXD3(cg03466780;N2;e;-0.087)	-	FAM101B(cg05738687;N2;f;0.085)	-
FAM110A(cg26275986;N3;c;-0.052)	-	FAM122B(cg12054337;S2;a;-0.1)	-	FAM181A(cg24013741;4;e;-0.089)	-
FAM184B(cg15471073;S3;e;-0.099)	-	FAM60A(cg03184588;S2;a;-0.067)	-	FBN1(cg04125371;1;c;0.083)	-
FBN2(cg23213887;1;e;0.101)	-	FBN3(cg25809434;1;e;-0.075)	-	FGF11(cg04293888;1;e;-0.091)	-
FGF12(cg09725157;4;a;-0.145)	-	FGFBP2(cg12073319;4;b;-0.102)	-	FGFR2(cg12835048;1;c;T;0.109)	-
FKBP8(cg10011232;S2;a;-0.069)	-	FLJ30058(cg20201673;N2;c;-0.146)	-	FMN1(cg21278889;4;e;T;-0.12)	-
FOXF1(cg03366439;1;a;-0.085)	-	FOXJ3(cg10309580;4;c;T;0.057)	-	FUZ(cg23348158;S2;b;0.082)	-
FXYD4(cg02788264;4;a;-0.074)	-	FXYD5(cg12055183;S3;e;-0.081)	-	FXYD5(cg26824126;S2;e;-0.073)	-
GABRP(cg18483611;4;f;-0.097)	-	GALNTL4(cg05197062;N2;e;0.09)	-	GIMAP7(cg17643598;4;b;-0.058)	-

GIPC1(cg20742389;1;e;0.09) - GJB5(cg13170235;N3;e;-0.087) - GLB1L2(cg15391574;1;d;T;-0.083) -  
 GLB1L3(cg25999867;1;a;T;-0.073) - GLRA1(cg23158483;S2;a;-0.107) - GLYATL1(cg23804666;4;a;-0.093) -  
 GNA13(cg00904578;4;f;T;-0.053) - GPR101(cg06469252;1;b;-0.063) - GPR101(cg25075069;1;b;-0.185) -  
 GPR137B(cg04077303;S2;e;-0.11) - GPR174(cg19388557;4;d;0.182) - GPR6(cg01261007;1;b;T;-0.08) -  
 GPR68(cg05903992;1;e;-0.118) - GPX5(cg19743215;4;e;-0.096) - GREM2(cg01176047;N2;f;T;-0.121) -  
 GRK1(cg20009378;4;e;0.12) - GSTM2(cg12647497;N2;a;T;-0.074) - GUCY1A2(cg12800047;4;e;T;-0.104) -  
 H2AFY(cg23137088;S2;a;-0.055) - HCCA2(cg07523553;1;e;-0.086) - HDGF(cg14155849;S2;b;-0.051) -  
 HHIPL1(cg23584871;N3;e;-0.097) - HLA-DQA1(cg17421046;4;e;-0.082) - HOXA13(cg03107888;S2;a;T;-0.089) -  
 HR(cg15137760;1;b;T;-0.128) - HRNBP3(cg02743895;4;c;-0.068) - HS3ST1(cg20075156;1;b;-0.06) -  
 HS3ST2(cg03532483;4;e;T;-0.066) - HS3ST3A1(cg02568806;S2;e;-0.073) - HTR7(cg10650018;N2;e;0.07) -  
 HTRA1(cg10588377;1;e;-0.081) - IER2(cg15382580;1;c;-0.055) - IGFBP3(cg16447589;S2;a;-0.081) -  
 IGSF5(cg12170845;4;b;-0.067) - IL17D(cg12475590;4;e;T;-0.08) - IL17RD(cg10882522;S2;a;-0.106) -  
 IL2RB(cg26757673;4;c;T;-0.105) - INCA1(cg14928057;S2;f;0.073) - INSC(cg20523653;4;f;-0.078) -  
 IRF4(cg06392169;1;c;-0.055) - IRF6(cg11570233;N2;c;-0.089) - IRS4(cg10293967;1;d;-0.127) -  
 JAKMIP1(cg05382097;1;c;T;-0.059) - JAKMIP3(cg00459068;1;e;-0.083) - JAKMIP3(cg05741225;4;a;-0.083) -  
 JAKMIP3(cg14932408;4;f;0.106) - JRKL(cg03075605;S2;c;-0.106) - KCNG2(cg14759931;1;e;-0.09) -  
 KCNIP1(cg00224471;4;e;T;-0.096) - KCNQ1(cg08066631;4;e;-0.067) - KCNQ2(cg13379325;1;e;-0.127) -  
 KCP(cg12582654;1;b;0.09) - KIAA0922(cg03464224;4;e;T;-0.066) - KIAA1026(cg19598142;1;e;-0.077) -  
 KIAA1257(cg07468062;4;e;T;-0.074) - KIAA1257(cg27640833;1;b;T;-0.079) - KIF6(cg19240233;4;e;T;-0.091) -  
 KLF12(cg15891850;4;e;T;-0.088) - KLHL31(cg07406888;4;a;0.099) - KLK1(cg26415633;N3;a;-0.06) -  
 KLRK1(cg21584184;4;e;-0.064) - KRTAP10-12(cg00163674;4;a;-0.093) - LAMA2(cg20640433;4;b;T;-0.078) -  
 LDOC1(cg25870731;1;d;-0.069) - LEPREL1(cg18670076;4;f;0.073) - LMBRD1(cg03001484;S2;b;-0.09) -  
 LMF1(cg04538473;S2;e;0.054) - LMF1(cg08074182;S2;e;-0.114) - LOC100130331(cg17078190;4;e;T;-0.072) -  
 LOC153328(cg00584238;1;e;-0.067) - LOC284276(cg17726692;4;b;T;-0.055) - LOC285796(cg10036368;4;e;-0.059) -  
 LOC349114(cg05363335;1;b;T;-0.073) - LRP11(cg12232274;S2;a;-0.051) - LRRC32(cg01439670;N2;c;-0.082) -  
 LTBP1(cg16572410;4;e;T;0.161) - LY6K(cg22146357;N2;b;T;-0.108) - LYPD1(cg17112958;1;c;-0.069) -  
 MAGEC2(cg10739728;4;b;-0.068) - MAML3(cg18127159;4;e;T;-0.094) - MAPT(cg00480298;4;e;-0.086) -  
 MED12L(cg21401219;1;a;-0.113) - MGMT(cg02172216;4;e;-0.095) - MIR1207(cg01940297;4;a;T;-0.058) -  
 MIR125B1(cg20475322;4;a;T;-0.073) - MIR487B(cg19560831;4;b;-0.086) - MIR518C(cg22725901;4;a;-0.135) -  
 MIR548F5(cg21884062;N2;e;-0.098) - MPPED2(cg20871721;4;e;T;-0.087) - MRV1(cg08298091;4;e;T;-0.113) -  
 MS4A3(cg14328641;4;a;-0.104) - MST1R(cg06521550;S2;b;-0.084) - MXRA5(cg07842130;N2;e;0.129) -  
 MYLK4(cg01681032;N2;e;-0.076) - MYO1F(cg26269802;4;b;0.106) - MYO7A(cg00497905;4;e;0.106) -  
 MYOM2(cg08575875;4;a;-0.089) - MYOM2(cg22740895;4;e;-0.106) - NANOS1(cg23089913;4;a;-0.06) -  
 NBL1(cg19136075;N2;a;-0.1) - NCRNA00094(cg13789015;N2;a;-0.094) - NEUROG1(cg02604503;S2;a;-0.12) -  
 NFASC(cg03854265;4;e;0.088) - NFIC(cg01033360;N2;e;-0.073) - NGFRAP1(cg13486082;N2;a;-0.117) -  
 NKAIN3(cg07592254;4;e;-0.092) - NKX2-3(cg06854084;1;b;-0.136) - NLGN1(cg02910194;4;e;T;-0.119) -  
 NLRP12(cg04695373;4;a;-0.113) - NOS1(cg10914558;N2;c;-0.12) - NOX4(cg26893231;4;d;-0.093) -  
 NPHS2(cg10711209;1;b;T;-0.052) - NTN3(cg04085822;1;d;-0.094) - NTRK3(cg18772882;4;e;T;-0.102) -  
 NUBP2(cg03226752;N2;e;-0.09) - NXF3(cg12584889;4;d;-0.15) - ODZ2(cg00192966;N2;e;-0.106) -  
 OLFM3(cg23631062;4;b;-0.064) - OPCML(cg08945802;4;e;-0.053) - OR11H4(cg24137472;4;b;-0.064) -  
 OR2T8(cg20528165;4;b;-0.107) - OR5V1(cg19323832;4;d;-0.105) - OR5W2(cg16734913;4;d;-0.086) -  
 OR8U8(cg12746908;4;e;T;-0.114) - OSBPL5(cg16507827;N3;e;-0.094) - P2RX5(cg03552992;S2;b;-0.057) -  
 PACRG(cg08555556;4;e;-0.06) - PAK6(cg12423123;1;c;-0.079) - PALM3(cg18720973;S2;d;0.053) -  
 PAX7(cg11704005;4;e;T;-0.16) - PCDH10(cg14410319;N2;d;-0.081) - PCDH8(cg19712603;1;b;T;0.071) -  
 PCDHGA2(cg18781988;1;e;-0.076) - PDE10A(cg26723355;1;a;-0.058) - PDE3A(cg11416338;S2;d;-0.063) -  
 PDE4A(cg08291069;1;d;0.095) - PDE4B(cg26963271;S2;c;T;-0.106) - PDE4D(cg11258089;1;e;T;-0.15) -  
 PDGFD(cg05246098;S2;e;-0.19) - PDZD4(cg09329826;1;e;T;-0.132) - PDZRN4(cg15235614;1;e;-0.058) -  
 PHLDB1(cg04142864;1;c;-0.086) - PIK3R5(cg13453139;4;c;-0.094) - PIP5KL1(cg14062643;S2;e;-0.083) -  
 PKHD1(cg10775844;4;e;-0.076) - PKLR(cg25651783;1;e;-0.097) - PLAC1L(cg08648317;4;d;-0.091) -



PLCH2(cg04847649;N3;a;0.061) - POU2F3(cg20663364;1;a;0.078) - PPP1R16B(cg22128431;1;d;-0.109) -  
 PPP2R5C(cg10764933;4;e;-0.075) - PRKAG2(cg00452039;N2;e;-0.119) - PROM1(cg04203238;1;b;-0.061) -  
 PRSS16(cg24760023;N3;a;-0.076) - PSMB9(cg03735531;1;e;0.063) - PTCHD1(cg18005219;1;e;-0.119) -  
 PTGS2(cg24887140;1;b;-0.084) - PTPRN2(cg02030008;1;e;0.103) - PTPRN2(cg04106894;4;e;-0.088) -  
 PTPRN2(cg08597719;S2;e;-0.11) - PTPRN2(cg09496385;N2;e;-0.081) - PTPRN2(cg19746375;4;e;-0.073) -  
 PYGO1(cg22510134;S2;a;-0.089) - RAB27A(cg05982017;4;a;0.072) - RAB34(cg19982230;N2;a;-0.073) -  
 RAB7A(cg27297137;N2;a;0.104) - RAB9B(cg26699341;S2;a;-0.071) - RADIL(cg11526630;N3;e;-0.08) -  
 RASA3(cg23264059;N2;e;-0.083) - RASGRP4(cg11876705;4;a;-0.052) - RGMA(cg04697454;S2;a;-0.062) -  
 RGS12(cg06789048;S2;e;0.126) - RND3(cg12632313;4;e;T;-0.074) - RNF220(cg04023150;1;c;0.052) -  
 RPS4Y2(cg01943289;4;a;0.073) - RPTOR(cg02386420;4;e;T;-0.074) - RSF1(cg06695566;N3;e;0.052) -  
 RSPO1(cg03654735;1;a;-0.068) - RYR3(cg15428578;1;e;0.05) - SARS(cg16257375;4;e;T;0.053) -  
 SATL1(cg10423328;4;b;0.193) - SCARA3(cg05626079;4;e;T;-0.081) - SCIN(cg08912317;1;b;T;-0.105) -  
 SCN3B(cg20662169;4;f;-0.094) - SDK1(cg03457472;N2;a;-0.082) - SDK1(cg27381557;N2;e;-0.094) -  
 SEMA3A(cg19762801;4;f;-0.115) - SFMBT2(cg10451314;4;e;T;0.085) - SGCD(cg01474424;4;d;-0.059) -  
 SH3BGR1(cg17405145;4;c;-0.068) - SHISA3(cg11065575;S3;f;0.116) - SHROOM3(cg24175289;4;e;T;-0.072) -  
 SIX3(cg08696165;S2;f;-0.09) - SLC10A3(cg23493704;S2;a;-0.112) - SLC16A2(cg03424927;S3;e;-0.154) -  
 SLC20A2(cg17999393;4;e;-0.057) - SLC22A12(cg03999137;4;e;-0.072) - SLC27A2(cg25150243;N2;a;T;-0.105) -  
 SLC30A7(cg16925880;S3;e;0.067) - SLC38A2(cg27491190;S2;a;-0.063) - SLC45A1(cg06322323;N2;e;-0.075) -  
 SLC46A3(cg20752818;1;b;-0.06) - SLC5A10(cg02758593;4;f;0.117) - SLC6A13(cg20958098;4;e;-0.1) -  
 SLC6A7(cg18725599;4;f;T;0.08) - SLC6A1(cg26101428;S2;a;-0.062) - SLITRK2(cg13663706;1;c;-0.146) -  
 SLITRK4(cg14340500;S2;b;-0.145) - SMARCA1(cg02827112;S2;c;-0.103) - SMARCD3(cg21192063;1;e;-0.093) -  
 SMOC2(cg16588799;4;e;-0.069) - SNAP91(cg20631014;1;b;0.062) - SNCAIP(cg07289841;4;c;T;0.101) -  
 SNTG1(cg08954417;S2;a;T;-0.086) - SORCS2(cg23910743;1;e;T;-0.068) - SOX17(cg15377283;1;e;T;-0.085) -  
 SOX2OT(cg08085357;N2;e;-0.082) - SPEG(cg03906033;S2;e;-0.081) - SPG20(cg25921358;S2;a;0.104) -  
 SPINK5L3(cg10598168;4;a;-0.064) - SRPK3(cg02279124;S2;e;T;-0.12) - SRRM3(cg15198101;1;e;-0.081) -  
 SSBP4(cg01532080;1;a;-0.058) - SSPO(cg19857140;4;e;T;-0.072) - ST18(cg01923473;4;a;-0.123) -  
 ST18(cg14965368;4;c;T;-0.138) - ST18(cg18061579;4;c;-0.11) - STARD8(cg24253627;4;f;-0.106) -  
 STAU2(cg17811994;4;f;-0.111) - SUPT7L(cg06062945;N3;e;0.104) - SV2C(cg15118835;N2;e;-0.057) -  
 SYCN(cg22290648;1;d;-0.068) - T-SP1(cg10657228;4;e;-0.076) - TACR3(cg07824172;S2;a;-0.053) -  
 TAGLN(cg10130564;4;b;-0.09) - TBX5(cg16249035;1;d;-0.069) - TCF7L1(cg07847030;1;e;-0.066) -  
 TEK(cg09827833;4;c;-0.092) - TGIF2LX(cg00378950;4;e;0.091) - TMEM132C(cg01055386;S2;e;-0.094) -  
 TMEM14E(cg15012282;4;a;-0.092) - TMEM171(cg25177452;1;c;-0.11) - TMEM220(cg00549475;N2;e;-0.102) -  
 TMEM233(cg25898092;S2;e;-0.06) - TMEM90A(cg13222752;1;c;-0.056) - TMSB4X(cg03120461;1;b;-0.056) -  
 TNIK(cg23970740;4;e;T;-0.156) - TNPO2(cg07655025;S2;b;-0.055) - TPM3(cg14835484;S3;e;T;-0.109) -  
 TPPP3(cg08074621;1;b;-0.093) - TRAF3IP3(cg17518842;4;e;-0.115) - TRIM68(cg01719157;S2;b;-0.119) -  
 TRIM68(cg19859515;S2;b;-0.104) - TRIOBP(cg12325455;4;b;T;0.059) - TSHR(cg16108059;S2;c;T;0.057) -  
 TTBK1(cg24901317;N3;e;-0.088) - TTLL8(cg09131332;N2;e;-0.097) - UVRAG(cg14839892;4;e;T;-0.079) -  
 VAV2(cg14499274;4;e;T;-0.054) - VCAM1(cg25762679;4;d;-0.075) - VCAN(cg17771652;N2;c;-0.057) -  
 VIPR2(cg03642066;N2;e;0.129) - WDR72(cg20667778;4;e;-0.077) - WHAMM(cg15558558;N2;b;-0.087) -  
 WNT2(cg18496858;4;e;T;-0.085) - WNT6(cg24813176;1;e;-0.069) - WNT9A(cg06641299;1;a;-0.056) -  
 XRCC5(cg09977847;4;f;-0.062) - ZEB2(cg15120754;1;e;-0.067) - ZFP92(cg21500538;N3;b;-0.138) -  
 ZFYVE28(cg00554437;N2;e;-0.089) - ZIC4(cg18930354;1;e;-0.105) - ZNF236(cg15059932;1;a;-0.065) -  
 ZNF469(cg05603784;S2;e;-0.052) - ZNF662(cg24384244;1;c;-0.098) - ZNF667(cg11314748;N3;c;-0.075) -  
 ZNF862(cg01844514;N2;e;-0.067)

### Supplementary table 2: Gene ontology analysis of the CD133+ DNA methylation signature

Pathways analyses were performed using the BrB geneset class comparison tool for KEGG and Biocarta and the WebGesalt and DAVID web applications. For analyses using WebGesalt and DAVID web applications pathways and genesets were filter for p-value <0.05.

Transcription factor genesets enrichment analyses were performed using BrB Array tool (independently for each cell line) and Webgesalt web application (two cell lines combined).

The tables presented here show the results obtained using the BrB geneset class comparison tool (top 50 for KEEG and top 100 for Biocarta).

Kegg Pathway	Pathway description	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
hsa00190	Oxidative phosphorylation	0.00001	0.58048	0.51 (-)
hsa00240	Pyrimidine metabolism	0.00001	0.07109	0.375 (+)
hsa00510	N-Glycan biosynthesis	0.00001	0.07686	0.43 (+)
hsa00562	Inositol phosphate metabolism	0.00001	0.32165	0.475 (-)
hsa00970	Aminoacyl-tRNA biosynthesis	0.00001	0.22696	0.45 (+)
hsa03010	Ribosome	0.00001	0.306	0.505 (+)
hsa03020	RNA polymerase	0.00001	0.03534	0.465 (+)
hsa03050	Proteasome	0.00001	0.17625	0.385 (+)
hsa04070	Phosphatidylinositol signaling system	0.00001	0.3742	0.49 (-)
hsa04110	Cell cycle	0.00001	0.24012	0.415 (+)
hsa04120	Ubiquitin mediated proteolysis	0.00001	0.07977	0.435 (+)
hsa04150	mTOR signaling pathway	0.00001	0.25134	0.37 (-)
hsa04210	Apoptosis	0.00001	0.51838	0.435 (+)
hsa04620	Toll-like receptor signaling pathway	0.00001	0.41763	0.405 (-)
hsa04910	Insulin signaling pathway	0.00001	0.16509	0.31 (-)
hsa04912	GnRH signaling pathway	0.00001	0.46157	< 0.005 (-)
hsa05040	Huntington@	0.00001	0.11076	0.4 (+)
hsa00632	Benzoate degradation via CoA ligation	0.00001	0.2315	0.405 (+)
hsa03030	DNA polymerase	0.00002	0.41433	0.475 (-)
hsa04664	Fc epsilon RI signaling pathway	0.00004	0.35842	0.345 (-)
hsa04520	Adherens junction	0.00004	0.45498	0.22 (+)
hsa03022	Basal transcription factors	0.00004	0.51979	0.465 (+)
hsa00310	Lysine degradation	0.00005	0.6996	0.275 (-)
hsa01030	Glycan structures - biosynthesis 1	0.00007	0.75985	0.06 (-)
hsa04660	T cell receptor signaling pathway	0.00012	0.3487	0.015 (+)
hsa05050	Dentatorubropallidoluysian atrophy (DRPLA)	0.00016	0.15375	0.185 (+)
hsa00624	1- and 2-Methylnaphthalene degradation	0.00016	0.17452	0.485 (-)
hsa00642	Ethylbenzene degradation	0.00019	0.1618	0.415 (-)
hsa04140	Regulation of autophagy	0.00025	0.0952	0.435 (+)
hsa00271	Methionine metabolism	0.0004	0.0677	0.32 (+)
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	0.00041	0.28364	0.05 (-)
hsa00280	Valine, leucine and isoleucine degradation	0.00043	0.10758	0.49 (+)
hsa04810	Regulation of actin cytoskeleton	0.00081	0.13578	0.455 (-)
hsa00564	Glycerophospholipid metabolism	0.00082	0.03043	0.47 (-)
hsa00290	Valine, leucine and isoleucine biosynthesis	0.00092	0.14385	0.43 (-)
hsa00450	Selenoamino acid metabolism	0.00117	0.06669	0.46 (+)
hsa04630	Jak-STAT signaling pathway	0.00131	0.46693	0.095 (-)
hsa00220	Urea cycle and metabolism of amino groups	0.00134	0.62552	0.265 (+)

hsa00790	Folate biosynthesis	0.00138	0.43195	0.47 (-)
hsa00860	Porphyrin and chlorophyll metabolism	0.00173	0.83511	0.01 (+)
hsa04310	Wnt signaling pathway	0.00176	0.65729	0.07 (-)
hsa04540	Gap junction	0.00182	0.4357	< 0.005 (-)
hsa00061	Fatty acid biosynthesis	0.00258	0.63494	0.44 (-)
hsa04530	Tight junction	0.00269	0.74886	0.22 (+)
hsa04370	VEGF signaling pathway	0.00324	0.33269	0.415 (-)
hsa00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	0.00333	0.05099	0.435 (-)
hsa05210	Colorectal cancer	0.00366	0.26363	0.21 (-)
hsa00062	Fatty acid elongation in mitochondria	0.00413	0.62532	0.495 (-)

Biocarta Pathway	Pathway description	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
h_41bbPathway	The 4-1BB-dependent immune response	0.00001	0.03977	0.02 (-)
h_aktPathway	AKT Signaling Pathway	0.00001	0.65267	0.38 (-)
h_At1rPathway	Angiotensin II mediated activation of JNK Pathway via Pyk2 dependent signaling	0.00001	0.44572	< 0.005 (-)
h_ceramidePathway	Ceramide Signaling Pathway	0.00001	0.17209	0.47 (-)
h_fasPathway	FAS signaling pathway ( CD95 )	0.00001	0.33412	0.45 (+)
h_fbw7Pathway	Cyclin E Destruction Pathway	0.00001	0.51389	0.385 (+)
h_hcmvPathway	Human Cytomegalovirus and Map Kinase Pathways	0.00001	0.20715	0.37 (-)
h_hifPathway	Hypoxia-Inducible Factor in the Cardiovascular System	0.00001	0.61665	0.275 (+)
h_HivnefPathway	HIV-I Nef: negative effector of Fas and TNF	0.00001	0.03053	0.455 (+)
h_hSWI-SNFpathway	Chromatin Remodeling by hSWI/SNF ATP-dependent Complexes	0.00001	0.38452	0.36 (+)
h_insulinPathway	Insulin Signaling Pathway	0.00001	0.05021	0.265 (-)
h_integrinPathway	Integrin Signaling Pathway	0.00001	0.32081	0.425 (-)
h_keratinocytePathway	Keratinocyte Differentiation	0.00001	0.00875	< 0.005 (-)
h_malPathway	Role of MAL in Rho-Mediated Activation of SRF	0.00001	0.09997	0.44 (+)
h_mapkPathway	MAPKinase Signaling Pathway	0.00001	0.00484	0.075 (-)
h_metPathway	Signaling of Hepatocyte Growth Factor Receptor	0.00001	0.74561	0.015 (-)
h_mTORPathway	mTOR Signaling Pathway	0.00001	0.91252	0.435 (-)
h_nfataPathway	NFAT and Hypertrophy of the heart (Transcription in the broken heart)	0.00001	0.51736	0.495 (-)
h_nthiPathway	NFkB activation by Nontypeable Hemophilus influenzae	0.00001	0.05437	0.405 (+)
h_p38mapkPathway	p38 MAPK Signaling Pathway	0.00001	0.01993	0.445 (-)
h_pparaPathway	Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)	0.00001	0.47045	0.115 (+)
h_pyk2Pathway	Links between Pyk2 and Map Kinases	0.00001	0.40695	< 0.005 (-)
h_RacCycDPathway	Influence of Ras and Rho proteins on G1 to S Transition	0.00001	0.51009	0.43 (+)
h_shhPathway	Sonic Hedgehog (Shh) Pathway	0.00001	0.77156	0.25 (-)
h_stressPathway	TNF/Stress Related Signaling	0.00001	0.00257	0.425 (-)
h_tnfr1Pathway	TNFR1 Signaling Pathway	0.00001	0.37629	0.425 (-)
h_vdrPathway	Control of Gene Expression by Vitamin D Receptor	0.00001	0.0114	0.43 (+)
h_wntPathway	WNT Signaling Pathway	0.00001	0.25	0.47 (-)
h_pdgfPathway	PDGF Signaling Pathway	0.00001	0.51784	0.015 (-)
h_ecmPathway	Erk and PI-3 Kinase Are Necessary for Collagen Binding in Corneal Epithelia	0.00001	0.70036	0.33 (-)

h_nfkbPathway	NF-kB Signaling Pathway	0.00001	0.07896	0.45 (+)
h_tnfr2Pathway	TNFR2 Signaling Pathway	0.00001	0.00614	0.41 (+)
h_gleevecPathway	Inhibition of Cellular Proliferation by Gleevec	0.00002	0.61576	0.3 (-)
h_rasPathway	Ras Signaling Pathway	0.00002	0.40621	0.225 (-)
h_cd40Pathway	CD40L Signaling Pathway	0.00002	0.00182	0.485 (+)
h_carm1Pathway	Transcription Regulation by Methyltransferase of CARM1	0.00003	0.78417	0.445 (+)
h_pmlPathway	Regulation of transcriptional activity by PML	0.00003	0.1759	< 0.005 (+)
h_crebPathway	Transcription factor CREB and its extracellular signals	0.00003	0.70412	0.48 (-)
h_akap95Pathway	AKAP95 role in mitosis and chromosome dynamics	0.00004	0.29029	0.21 (-)
h_tollPathway	Toll-Like Receptor Pathway	0.00004	0.42637	0.365 (-)
h_mcmPathway	CDK Regulation of DNA Replication	0.00004	0.68595	0.165 (+)
h_p27Pathway	Regulation of p27 Phosphorylation during Cell Cycle Progression	0.00004	0.50279	0.38 (+)
h_il1rPathway	Signal transduction through IL1R	0.00005	0.0005	0.37 (-)
h_g1Pathway	Cell Cycle: G1/S Check Point	0.00005	0.03305	0.47 (+)
h_eif4Pathway	Regulation of eIF4e and p70 S6 Kinase	0.00006	0.53956	0.465 (-)
h_deathPathway	Induction of apoptosis through DR3 and DR4/5 Death Receptors	0.00006	0.07647	0.44 (+)
h_biopeptidesPathway	Bioactive Peptide Induced Signaling Pathway	0.0001	0.77201	0.105 (-)
h_ptenPathway	PTEN dependent cell cycle arrest and apoptosis	0.00011	0.59771	0.43 (+)
h_Par1Pathway	Thrombin signaling and protease-activated receptors	0.00013	0.81364	0.115 (-)
h_tsp1Pathway	TSP-1 Induced Apoptosis in Microvascular Endothelial Cell	0.00017	0.77398	0.08 (+)
h_skp2e2fPathway	E2F1 Destruction Pathway	0.00018	0.48524	0.14 (+)
h_cdMacPathway	Cadmium induces DNA synthesis and proliferation in macrophages	0.0002	0.17841	0.025 (-)
h_epoPathway	EPO Signaling Pathway	0.0002	0.01263	< 0.005 (-)
h_setPathway	Granzyme A mediated Apoptosis Pathway	0.00022	0.02493	0.3 (+)
h_ctcfPathway	CTCF: First Multivalent Nuclear Factor	0.00022	0.54154	0.445 (+)
h_bcellsurvivalPathway	B Cell Survival Pathway	0.00024	0.80934	0.505 (+)
h_arapPathway	ADP-Ribosylation Factor	0.00026	0.81531	0.5 (-)
h_p53Pathway	p53 Signaling Pathway	0.00027	0.01592	0.415 (+)
h_akap13Pathway	Rho-Selective Guanine Exchange Factor AKAP13 Mediates Stress Fiber Formation	0.00029	0.45844	0.085 (+)
h_pcafpathway	The information-processing pathway at the IFN-beta enhancer	0.00032	0.2853	0.13 (-)
h_TPOPathway	TPO Signaling Pathway	0.00037	0.59633	0.025 (-)
h_sam68Pathway	Regulation of Splicing through Sam68	0.00037	0.132	0.315 (-)
h_rbPathway	RB Tumor Suppressor/Checkpoint Signaling in response to DNA damage	0.00041	0.1595	0.045 (+)
h_tall1Pathway	TACI and BCMA stimulation of B cell immune responses.	0.00043	0.01395	0.375 (-)
h_igf1rPathway	Multiple antiapoptotic pathways from IGF-1R signaling lead to BAD phosphorylation	0.00044	0.6878	0.48 (+)
h_gpcrPathway	Signaling Pathway from G-Protein Families	0.00046	0.60871	0.145 (-)
h_egfPathway	EGF Signaling Pathway	0.00049	0.5386	< 0.005 (-)
h_chemicalPathway	Apoptotic Signaling in Response to DNA Damage	0.00049	0.0597	0.455 (-)
h_dreampathway	Repression of Pain Sensation by the Transcriptional Regulator DREAM	0.0005	0.79856	0.195 (-)
h_carm-erPathway	CARM1 and Regulation of the Estrogen Receptor	0.00056	0.44434	0.13 (+)
h_ck1Pathway	Regulation of ck1/cdk5 by type 1 glutamate receptors	0.00061	0.77694	0.47 (+)

h_ghPathway	Growth Hormone Signaling Pathway	0.00061	0.54468	0.27 (-)
h_atrbcrPathway	Role of BRCA1, BRCA2 and ATR in Cancer Susceptibility	0.00063	0.74089	0.41 (+)
h_RELAPathway	Acetylation and Deacetylation of RelA in The Nucleus	0.00067	0.12386	0.185 (+)
h_ngfPathway	Nerve growth factor pathway (NGF)	0.00068	0.00178	0.095 (-)
h_mitochondriaPathway	Role of Mitochondria in Apoptotic Signaling	0.00076	0.03942	0.485 (+)
h_g2Pathway	Cell Cycle: G2/M Checkpoint	0.00079	0.04833	0.36 (+)
h_myosinPathway	PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase	0.00105	0.37667	0.285 (-)
h_fMLPPathway	fMLP induced chemokine gene expression in HMC-1 cells	0.00105	0.53319	0.425 (-)
h_il6Pathway	IL 6 signaling pathway	0.00131	0.41332	0.28 (-)
h_spryPathway	Sprouty regulation of tyrosine kinase signals	0.00135	0.47504	0.025 (-)
h_eradPathway	ER-associated degradation (ERAD) Pathway	0.00138	0.38552	0.505 (+)
h_cellcyclePathway	Cyclins and Cell Cycle Regulation	0.00163	0.07621	0.32 (+)
h_ctbp1Pathway	SUMOylation as a mechanism to modulate CtBP-dependent gene responses	0.00173	0.02631	0.37 (-)
h_mPRPathway	How Progesterone Initiates the Oocyte Maturation	0.00176	0.74406	0.39 (+)
h_HuntingtonPathway	Inhibition of Huntington@	0.00193	0.71966	0.505 (-)
h_ndkDynaminPathway	Endocytotic role of NDK, Phosphins and Dynamin	0.00197	0.22558	0.525 (+)
h_s1pPathway	SREBP control of lipid synthesis	0.00202	0.50156	0.47 (-)
h_badPathway	Regulation of BAD phosphorylation	0.00218	0.69342	0.44 (+)
h_HBxPathway	Calcium Signaling by HBx of Hepatitis B virus	0.00225	0.16559	0.38 (+)
h_telPathway	Telomeres, Telomerase, Cellular Aging, and Immortality	0.00232	0.02529	0.415 (+)
h_anthraxPathway	Anthrax Toxin Mechanism of Action	0.00278	0.00312	0.015 (-)
h_plk3Pathway	Regulation of cell cycle progression by Plk3	0.003	0.08473	0.235 (+)
h_erk5Pathway	Role of Erk5 in Neuronal Survival	0.003	0.73791	0.085 (-)
h_eifPathway	Eukaryotic protein translation	0.00307	0.05977	0.415 (+)
h_tffPathway	Trefoil Factors Initiate Mucosal Healing	0.00307	0.59771	0.05 (-)
h_prc2Pathway	The PRC2 Complex Sets Long-term Gene Silencing Through Modification of Histone Tails	0.00316	0.57215	0.38 (-)
h_vegfPathway	VEGF, Hypoxia, and Angiogenesis	0.0034	0.57541	0.01 (-)
h_tgfbPathway	TGF beta signaling pathway	0.00358	0.029	0.455 (-)
h_il2rbPathway	IL-2 Receptor Beta Chain in T cell Activation	0.00366	0.50279	0.32 (-)

Transcription Factors Genesets for Huh7	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
AR_T00040	0.12252	0.1803	< 0.005 (+)
BRCA1_T04074	0.22524	0.47933	< 0.005 (+)
CEBPA_T00105	0.61094	0.22754	< 0.005 (+)
CEBPE_T04883	0.25524	0.15789	< 0.005 (-)
CREB1_T00163	0.03534	0.13982	< 0.005 (-)
CREM_T01803	0.19792	0.20882	< 0.005 (-)
E2F-4_T01546	0.18586	0.05574	< 0.005 (+)
EGR1_T00241	0.19381	0.31679	< 0.005 (+)
EGR2_T00242	0.08213	0.21141	< 0.005 (+)
EGR4_T05190	0.93558	0.94853	< 0.005 (+)

ELK1_T00250	0.97053	0.94448	< 0.005 (+)
ERG_T00265	0.012	0.02143	< 0.005 (-)
ESR2_T04651	0.1722	0.1145	< 0.005 (-)
ETS2_T00113	0.0988	0.08036	< 0.005 (-)
ETV4_T00685	0.91247	0.91138	< 0.005 (+)
FOS_T00123	0.16479	0.49582	< 0.005 (-)
HOXA9_T01709	0.17471	0.26822	< 0.005 (+)
HOXB7_T01734	0.02468	0.1957	< 0.005 (+)
JUN_T00029	0.15169	0.16118	< 0.005 (-)
LEF1_T02905	0.01967	0.1124	< 0.005 (+)
MYB_T00137	0.77314	0.91752	< 0.005 (+)
NFIC_T00176	0.03118	0.07819	< 0.005 (-)
NFKB1_T00591	0.29091	0.60122	< 0.005 (+)
PAX5_T00070	0.04622	0.06566	< 0.005 (-)
PAX8_T02898	0.08185	0.17936	< 0.005 (-)
POU5F1_T00652	0.01302	0.02375	< 0.005 (-)
PPARD_T02745	0.0195	0.00814	< 0.005 (-)
REL_T00168	0.05833	0.14667	< 0.005 (-)
SMAD4_T04292	0.06999	0.18164	< 0.005 (-)
SP3_T02338	0.18431	0.12085	< 0.005 (+)
SPI1_T02068	0.80299	0.35605	< 0.005 (+)
STAT1_T01492	0.11198	0.08857	< 0.005 (-)
TAL1_T00790	0.62333	0.83836	< 0.005 (+)
TFAP2C_T02468	0.11435	0.1608	< 0.005 (-)
TP73_T04931	0.50866	0.67087	< 0.005 (+)
USF2_T00878	0.51047	0.06983	< 0.005 (-)
WT1_T00899	0.59505	0.73575	< 0.005 (-)

Transcription Factor GeneSets for HepG2	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
AR_T00040	0.58857	0.47065	< 0.005 (-)
ATF1_T00968	0.03071	0.08507	< 0.005 (-)
ATF4_T00051	0.01936	0.01994	< 0.005 (-)
BRCA1_T04074	0.01183	0.15827	< 0.005 (+)
BRCA2_T06444	0.02685	0.02046	< 0.005 (+)
CREB1_T00163	0.00045	0.0029	< 0.005 (-)
CREB2_T00051	0.01936	0.01994	< 0.005 (-)
EGR1_T00241	0.00028	0.01209	< 0.005 (-)
EGR4_T05190	0.03842	0.24444	< 0.005 (+)
ELK1_T00250	0.49193	0.69234	< 0.005 (-)
EPAS1_T02718	0.00441	0.11929	< 0.005 (+)
ERG_T00265	0.00001	0.0048	< 0.005 (+)
ETV4_T00685	0.00143	0.01906	< 0.005 (-)
GLI_T00330	0.18339	0.36482	< 0.005 (+)
GLI2_T04961	0.03426	0.10361	< 0.005 (+)
HOXA10_T01713	0.07188	0.0222	< 0.005 (-)
HOXA9_T01709	0.01105	0.12195	< 0.005 (+)
HOXB3_T01723	0.01001	0.03591	< 0.005 (-)

HOXB7_T01734	0.00722	0.03153	< 0.005 (-)
MYB_T00137	0.63608	0.66718	< 0.005 (-)
NFKB1_T00591	0.00241	0.00966	< 0.005 (+)
NFKB2_T00394	0.01972	0.02339	< 0.005 (-)
PAX5_T00070	0.00126	0.07603	< 0.005 (+)
PAX8_T02898	0.00017	0.00743	< 0.005 (-)
POU2F1_T00641	0.1886	0.33271	< 0.005 (+)
POU3F2_T00630	0.04999	0.06395	< 0.005 (-)
POU5F1_T00652	0.39964	0.30771	< 0.005 (-)
RARG_T00720	0.01285	0.04103	< 0.005 (-)
RELA_T00594	0.22894	0.35394	< 0.005 (+)
SMAD2_T04095	0.22688	0.20776	< 0.005 (+)
SPI1_T02068	0.00359	0.28796	< 0.005 (-)
STAT5A_T05735	0.00162	0.07653	< 0.005 (+)
TFAP2C_T02468	0.00906	0.06922	< 0.005 (+)

### Supplementary table 3: List of 464 annotated CpG loci differentially methylated after TGF- $\beta$ exposure.

The list of significant CpG loci between TGF- $\beta$ -exposed control cells was obtained using the BrB class comparison tool, blocked by cell lines, filtered for p-value <0.001, FDR<0.05 and averaged delta beta at day 8 (between the two cell lines) >5%. Among the 580 significant loci, here are presented the ones that are associated to annotated genes (464). For each locus, information are given as presented below:

NAME OF THE GENE (probe's ID; relation to CpG island; gene region; true enhancer; averaged deltabeta between CD133+ and CD133- cells). Legend:

*Relation to CpG island:* 1= Island / N2 = N Shore / S2 = S Shore / N3 = N Shelf / S3 = S Shelf / 4 = Open sea

*Gene regions:* a= TSS1500 / b = TSS200 / c= 5'UTR / d = 1st exon / e = Body / f = 3'UTR

*Enhancer:* T = true enhancer

AAK1(cg22059413;4;e;T;-0.075) - ABI1(cg26142490;4;e;T;-0.086) - ACOX3(cg01850230;N2;e;0.071) - ACOX3(cg15137838;4;e;0.069) - ACSF3(cg05145233;S2;e;0.114) - ACSF3(cg01942849;S2;e;0.116) - ACSL3(cg22909085;4;c;T;0.078) - ACTN4(cg01030321;4;e;0.08) - ACTR3(cg07093324;S3;e;0.113) - ADAMTS10(cg22262095;1;e;0.054) - ADAMTSL2(cg21211688;S3;e;0.057) - ADCY3(cg06525750;N2;e;T;0.091) - ADD3(cg18994606;4;c;T;0.073) - AGPAT3(cg01331461;4;b;0.058) - AGRN(cg00525597;N2;e;0.128) - AHNNAK(cg22365167;4;c;T;-0.111) - AKAP2(cg13430173;4;e;T;0.089) - AKT3(cg19503731;4;a;0.105) - ALB(cg02951062;4;b;0.079) - ALB(cg13495204;4;e;T;0.112) - ALPP(cg22007776;N3;a;0.074) - ALPP(cg13605579;N2;a;0.084) - ANKRD11(cg09222577;4;c;T;0.063) - ANKRD13A(cg12870014;4;e;0.22) - ANKRD2(cg07898573;S3;c;0.076) - ANKRD9(cg16727416;1;c;0.05) - ANO4(cg11891256;4;e;T;-0.136) - ANO6(cg10892950;4;e;T;-0.081) - APLP2(cg22186291;4;e;0.057) - APOM(cg00812578;N3;e;0.079) - AQP2(cg05126095;N3;f;0.066) - ARG1(cg02862362;4;e;0.075) - ARHGAP22(cg23901444;1;c;T;0.075) - ARHGAP27(cg22153994;S2;a;0.051) - ARHGEF1(cg05503433;S2;e;0.102) - ARHGEF12(cg05099464;4;e;T;-0.088) - ARHGEF3(cg11229273;4;e;T;-0.1) - ARID3A(cg18084554;N2;c;0.088) - ARNT2(cg01446731;4;e;0.061) - ARSI(cg09001514;S2;e;0.086) - ARX(cg01962233;1;a;0.085) - ATP2C2(cg04342425;4;e;0.076) - ATP6V0D1(cg06120000;4;e;T;0.062) - ATPBD4(cg17014914;4;e;T;0.156) - BCAR3(cg22514229;4;e;T;0.115) - BCR(cg23225050;N3;e;T;0.074) - BEND3(cg01963573;N2;f;0.053) - BGN(cg21635956;4;e;0.077) - BHLHE40(cg16582517;S2;f;-0.153) - BMF(cg09749364;4;e;T;0.068) - BMP1(cg00920938;S2;e;T;-0.099) - BRD2(cg02707277;4;e;0.097) - BRSK2(cg16022876;S2;e;0.065) - C10orf114(cg10973934;N2;a;0.121) - C16orf45(cg02399788;4;e;T;0.079) - C16orf48(cg05807722;S2;f;0.055) - C17orf101(cg06601666;N2;e;0.122) - C17orf28(cg17713161;4;e;0.082) - C17orf28(cg13537156;4;e;0.085) - C19orf22(cg12045715;S3;e;0.075) - C22orf9(cg05345310;N2;e;T;-0.103) - C3orf25(cg14627091;4;e;0.055) - C3orf26(cg24156261;4;e;T;-0.073) - C4orf11(cg16732415;4;a;0.108) - C4orf19(cg10122877;4;a;0.098) - C5orf13(cg06838283;N3;e;0.139) - C9orf30(cg14475840;S2;c;-0.058) - CACNA1H(cg04480708;N2;e;T;0.091) - CACNA1H(cg26744056;S3;e;0.08) - CALD1(cg26915370;4;a;-0.118) - CALD1(cg19764295;4;e;T;0.067) - CALM2(cg21361646;4;f;0.105) - CAMK2B(cg25333258;1;c;0.125) - CARD14(cg15443128;1;e;0.05) - CAST(cg06465076;4;e;T;0.158) - CBFA2T3(cg27248474;N3;c;T;0.079) - CD37(cg19753641;N2;e;0.068) - CDC42BPB(cg05895018;S3;e;0.065) - CDH18(cg00606396;4;e;-0.065) - CDKN1B(cg12751042;N3;f;0.102) - CLDN14(cg11664818;4;c;0.05) - CLDN14(cg08264376;4;e;0.118) - CLDN2(cg17051440;4;c;T;0.055) - COL18A1(cg04617640;N2;a;0.135) - COL1A1(cg16781907;N3;e;T;0.11) - COL1A1(cg18618815;N2;e;0.072) - COL1A1(cg08681473;4;e;T;0.113) - COL1A1(cg18405262;4;e;0.121) - COL1A1(cg10820084;4;e;T;0.126) - COL1A1(cg25026926;4;e;T;0.131) - COL1A1(cg00638021;4;e;T;0.145) - COL1A1(cg21847118;4;e;0.153) - COL1A1(cg27604897;4;e;T;0.165) - COL1A1(cg15435765;4;e;0.182) - COL4A1(cg02658690;4;e;0.056) - CORO2A(cg12800915;N2;c;0.081) - COTL1(cg09985344;N3;e;T;0.063) - CR2(cg10109747;N2;a;0.061) - CRAMP1L(ch.16.97779F;4;e;0.054) - CRIM1(cg00068038;4;e;T;-0.1) - CTNNB1(cg09489743;N2;a;0.08) - CTSH(cg18738367;S2;a;0.061) - CUL9(cg07126235;4;f;-0.075) - DAAM2(cg05609656;4;c;0.064) - DACT2(cg01082907;N2;e;0.132) - DAP(cg18801045;4;e;T;0.069) - DAZAP1(cg11656553;1;e;0.11) - DCK(cg08514408;4;e;T;-0.114) - DDA1(cg17799563;S2;e;T;-0.095) - DDR1(cg16993957;4;f;0.104) - DDX19B(cg10009207;4;e;0.082) - DENND3(cg22899502;4;e;0.058) - DISC1(cg22367981;4;e;T;-0.051) - DLX1(cg01244270;S2;f;0.138) - DNAH17(cg02374982;1;e;0.054) - DNAH17(cg20690714;1;e;0.073)



DNAH17(cg10217661;1;e;0.086) - DNAH17(cg27052709;N3;e;T;0.132) - DNAH17(cg23752152;N2;e;0.058) -  
 DNAH17(cg18098769;N2;e;0.077) - DNAH17(cg09773127;N2;e;0.089) - DNAH17(cg11267683;S3;e;0.07) -  
 DNAH17(cg25875702;S3;e;0.127) - DNAH17(cg03953828;S2;e;0.069) - DNAH17(cg12818883;4;e;0.058) -  
 DNAH17(cg08856333;4;e;0.06) - DNAH17(cg11406388;4;e;0.066) - DNAH17(cg18733967;4;e;0.066) -  
 DNAH17(cg10332979;4;e;0.067) - DNAH17(cg10089081;4;e;0.073) - DNAH17(cg15406425;4;e;T;0.076) -  
 DNAH17(cg14330206;4;e;0.081) - DNAH17(cg09656382;4;e;0.088) - DNAH17(cg00235657;4;e;0.09) -  
 DNAH17(cg06799664;4;e;0.094) - DNAH17(cg05361750;4;e;0.101) - DNAH17(cg11347599;4;e;0.11) -  
 DNAH17(cg05414903;4;b;0.129) - DNAH17(cg06342438;4;e;0.143) - DNAH17(cg10308396;4;e;0.144) -  
 DNAH17(cg23753610;4;e;0.16) - DNAH17(cg07255197;4;b;0.173) - DNAH17(cg09687005;4;e;0.189) -  
 DNAH17(cg09577144;4;b;0.266) - DNAJB6(ch.7.3356624R;N2;e;0.056) - DNLZ(cg00900642;N2;f;T;0.055) -  
 DNMT3A(cg04058399;4;e;T;0.066) - DNMT3B(cg24403338;4;c;0.083) - DNMT3B(cg17475857;4;c;) -  
 DOCK4(cg07996838;4;e;T;0.082) - DOCK5(cg00157012;4;e;0.08) - EFCAB4B(cg17904739;N2;c;0.115) -  
 EFNA2(cg22676470;S2;e;0.093) - ELFN1(cg20718434;4;b;0.064) - ELL2(cg16998950;4;e;T;0.107) - ELMO1(cg17592875;4;e;T;0.057)  
 - ELMO1(cg02902948;4;e;0.066) - ENTPD6(cg07588216;4;e;T;-0.123) - EPPK1(cg21135560;1;d;0.078) - ERI3(cg22709362;4;e;T;-  
 0.075) - ERLIN1(cg26746309;N3;e;0.119) - ETFB(cg11364420;4;a;0.151) - EXOC6(cg12332526;4;e;T;0.079) -  
 FAM183A(cg04900427;1;b;0.122) - FAM38A(cg04602696;S2;e;T;0.052) - FBLIM1(cg06747907;4;c;0.109) -  
 FBXO32(cg06109876;4;a;0.111) - FBXO38(cg23296792;4;e;T;0.162) - FGFR3(cg02350535;1;e;0.053) - FGFR3(cg25301756;1;e;0.084)  
 - FGFR3(cg08145949;N2;e;0.076) - FGFR4(cg11849703;S3;e;0.052) - FGR(cg21115433;S2;b;0.063) - FKBP9(cg17786776;N2;a;0.097)  
 - FLJ32810(cg02008402;4;e;T;0.09) - FLJ39653(cg10051588;N3;e;T;-0.072) - FLT4(cg24096745;N3;e;-0.066) -  
 FMNL1(cg19481029;N3;e;0.093) - FOXP2(cg20412356;4;e;T;0.051) - FOXP2(cg04593859;4;e;T;0.052) -  
 FRK(cg25996663;4;e;T;0.082) - FXN(cg13859886;4;e;T;0.073) - GALNT6(cg21253043;N2;c;T;0.074) -  
 GALNT9(cg12055993;N2;e;0.087) - GAPVD1(ch.9.2043201R;4;e;0.063) - GATA6(cg21684845;N2;e;0.147) -  
 GDNF(cg05330056;N2;e;T;0.103) - GFOD1(cg23336266;4;e;0.081) - GHRH(cg24493068;4;a;0.066) - GIPC1(cg20742389;1;e;0.088) -  
 GIPR(cg13320842;S2;e;0.079) - GMEB2(cg10170269;N2;c;0.085) - GNG7(cg07938763;S3;e;0.071) - GPC6(cg06326425;4;e;T;0.078) -  
 GPR133(cg02153041;4;e;0.084) - GPR179(cg07668558;4;a;0.078) - GPRC5A(cg24765748;4;e;T;0.117) -  
 GPRC5C(cg19212391;S3;e;0.062) - GPSM1(cg14271150;4;e;0.095) - GRB10(cg24977055;1;c;-0.068) - GREB1(cg25649765;4;a;-  
 0.058) - HCCA2(cg08313842;4;e;0.065) - HDAC5(cg24396400;4;f;T;0.078) - HDAC7(cg23522915;4;e;T;0.065) -  
 HEG1(cg19533443;4;e;T;0.205) - HIF3A(cg14088357;1;a;0.066) - HLA-DPB2(cg01184577;4;e;0.053) - HLCS(cg04126652;N2;c;0.058)  
 - HMGA2(cg21822187;4;e;T;-0.174) - HS6ST3(cg11411106;4;e;T;0.063) - HTT(cg12636882;4;e;T;0.123) -  
 ICAM4(cg20036207;N2;a;T;0.072) - IFFO2(cg15675456;4;e;T;0.085) - IFT140(cg06363243;4;e;0.06) -  
 IGDCC3(cg23529231;4;f;T;0.072) - IL18(cg15418499;4;c;T;-0.058) - IL1R1(cg06392753;4;e;0.151) - IL20(cg08479073;4;a;0.098) -  
 INCA1(cg03128860;S3;e;0.082) - IRAK2(cg13419330;4;e;T;-0.11) - ITFG3(cg27316939;4;c;0.051) - ITGB1BP2(cg12391921;4;d;0.104)  
 - ITGB2(cg04217515;S2;e;0.067) - ITPR3(cg14003231;4;e;T;-0.068) - KCNT2(cg09173378;4;e;T;0.065) -  
 KCTD7(cg07522403;4;f;0.167) - KDM6B(cg02308232;N3;a;T;0.098) - KIAA1486(cg10859755;1;e;0.097) -  
 KIAA1949(cg06462220;4;e;0.059) - KIF26B(cg02765564;S3;e;0.087) - KIF26B(cg25143359;4;e;0.056) - KLHL25(cg25843651;4;c;T;-  
 0.084) - KLHL5(cg25123225;4;c;T;0.073) - KRT27(cg04578777;S3;d;0.085) - KRT7(cg02322205;S2;e;0.127) -  
 KRT80(cg02387510;4;a;0.099) - KRT80(cg23243343;4;e;0.101) - LAMC1(cg22809683;4;e;T;-0.055) - LATS2(cg11192895;4;e;T;0.082)  
 - LBH(cg04254242;4;f;T;0.104) - LCLAT1(cg23457506;4;c;T;0.125) - LEPR(cg01933519;4;c;T;-0.153) -  
 LIMD1(cg08062273;4;e;T;0.051) - LIMD1(cg16283183;4;f;T;0.128) - LIPC(cg16391792;4;a;0.116) - LITAF(cg04359558;4;e;0.189) -  
 LLGL1(cg09658183;4;e;T;0.05) - LMNA(cg08881019;4;e;-0.071) - LOC100130000(cg23675441;S2;e;0.077) -  
 LOC100132354(cg02272576;4;e;T;0.062) - LOC100192378(cg10435486;S2;e;0.121) - LOC127841(cg03461777;4;a;0.084) -  
 LOC146880(cg18402166;S2;a;0.061) - LOC283731(cg01568668;1;e;T;-0.094) - LOC284009(cg16002660;4;e;T;-0.118) -  
 LOC340094(cg24763840;4;b;0.075) - LOC646982(cg08559342;4;e;T;0.054) - LOC728264(cg12675571;4;e;0.071) -  
 LPP(cg14177865;4;f;T;0.094) - LPP(cg24454374;4;c;0.224) - LPPR2(cg12587615;N2;e;T;0.091) - LRRC3(cg22139500;S2;f;0.074) -  
 LRRC49(cg11323113;4;f;T;0.106) - LRRFIP1(cg08151442;4;e;T;0.058) - LSP1(cg03541934;S3;e;0.069) -

LTBP2(cg25980489;N2;e;0.139) - LTBP2(cg08189843;S3;e;0.101) - LTBP2(cg16056219;4;e;T;0.132) -  
 LUZP1(cg26530045;4;c;T;0.095) - LYPD5(cg12072164;S3;d;0.069) - MAD1L1(cg27109748;4;e;T;-0.058) - MAEA(cg04277993;N2;e;-  
 0.109) - Magmas(cg01406317;N3;e;0.232) - MAP1LC3B(ch.16.2068605F;N3;f;0.107) - MASP2(cg20817175;4;b;0.063) -  
 MCC(cg21806985;S2;b;0.059) - MCF2L(cg14460195;1;e;0.067) - MCPH1(cg01544777;4;e;0.06) - MICAL2(cg14081744;4;e;T;0.089) -  
 MINK1(cg17901582;4;e;T;0.05) - MIR1208(cg07018107;4;e;-0.071) - MIR143(cg04317047;4;b;0.086) -  
 MIR181D(cg18745782;S2;a;0.1) - MIR192(cg00376448;N3;a;0.056) - MIR196B(cg05250768;1;b;0.084) -  
 MIR19A(cg02297838;S2;a;0.054) - MITF(cg12198198;4;e;0.076) - MYO1E(ch.15.814613R;4;e;0.056) -  
 MYO1F(cg26269802;4;b;0.052) - NAP1L4(cg23951171;4;c;0.082) - NAV1(cg14282634;4;e;T;0.063) -  
 NCRNA00162(cg08636922;N2;e;0.092) - NEDD9(cg25250968;4;c;-0.087) - NFE2L2(cg27507284;4;e;T;0.098) -  
 NFYC(cg24700993;4;e;0.055) - NGB(cg05836974;S2;a;0.086) - NHSL1(cg00409658;N2;f;0.096) - NHSL2(cg02154531;N2;e;0.098) -  
 NISCH(cg16438525;N3;e;T;-0.081) - NLRP1(cg11828470;4;e;T;-0.09) - NRF1(cg26544062;4;f;0.055) -  
 NRP2(cg19731541;S3;e;T;0.065) - NUDT13(cg11930274;4;e;T;-0.054) - ONECUT2(cg11817589;S3;e;0.093) -  
 OSBPL10(cg22539670;4;e;T;0.052) - PALLD(cg16018921;4;e;T;-0.171) - PAQR7(cg01566199;4;d;0.068) -  
 PAQR9(cg08784462;1;a;0.086) - PCDH8(cg17535595;1;b;T;0.084) - PCGF3(cg00169930;N3;e;0.05) - PCID2(cg09075968;4;e;T;-  
 0.125) - PDE3B(cg25393009;4;e;0.076) - PDE6B(cg04774239;4;e;0.073) - PDGFRA(cg12845923;4;e;T;0.05) -  
 PDK2(cg06647382;S3;e;0.056) - PDLIM1(cg10266121;4;e;T;-0.122) - PDLIM7(cg15225325;N2;e;T;-0.183) -  
 PHLDB1(cg20309703;1;c;0.058) - PKIG(cg19554235;N2;a;0.062) - PLCE1(cg23439277;4;e;-0.06) - PLEC1(cg16001422;N2;e;T;-0.073)  
 PLEC1(cg01870834;N2;e;0.097) - PPP1R15A(cg03707168;S3;e;0.066) - PPP3CC(cg20910008;S3;e;-0.151) -  
 PPP4C(cg02077558;4;e;0.065) - PRAME(cg17648213;N3;e;0.089) - PRDM1(cg19064302;S2;e;T;0.061) -  
 PRDM16(cg17421241;N2;e;T;0.08) - PRDM6(cg19328475;N2;e;T;0.081) - PRELID2(cg13019306;4;e;T;-0.064) -  
 PRKCQ(cg04351665;1;b;0.078) - PRR16(cg12453014;1;c;0.058) - PSD4(cg13224420;4;e;0.099) - PSMB9(cg04908668;S2;e;0.088) -  
 PTGER1(cg27524460;S2;a;0.053) - R3HDM2(cg02363202;4;e;T;-0.087) - R3HDM2(cg26650359;4;e;T;-0.084) -  
 R3HDM2(cg26247373;4;e;T;0.077) - RAB38(cg01568784;S2;b;0.13) - RAB39(cg17498773;1;e;T;0.085) -  
 RAB8B(cg14376033;4;f;0.141) - RAD51L1(cg04782982;4;e;T;0.146) - RAP1GAP(cg10038867;N3;c;0.123) -  
 RAP1GAP2(cg22647670;4;e;T;0.059) - RASA3(cg07007754;N2;e;0.07) - RASA3(cg16739503;S2;e;T;0.156) -  
 RB1(cg15108060;4;e;T;0.144) - RBM19(cg12020794;4;e;T;0.073) - RBM26(cg16969872;4;e;0.056) - REP15(cg16706260;4;b;0.057) -  
 REP15(cg07809176;4;c;0.251) - RERE(cg01024458;4;e;0.087) - RFC2(cg17069650;4;e;0.147) - RGL2(cg03789294;N2;e;0.051) -  
 RGL2(cg25361447;N2;e;0.127) - RGL2(cg08727352;N2;e;0.156) - RGS3(cg14327394;4;e;T;0.059) -  
 RHOBTB2(cg20642630;4;c;T;0.096) - RIN3(ch.14.1488981R;4;e;0.118) - RNF121(cg06363692;4;e;T;-0.166) -  
 RPL13(cg01995548;S2;e;0.068) - RPS18(cg09591519;S2;e;0.242) - RRBP1(cg05955436;N2;c;0.119) - RREB1(cg14919455;4;c;T;0.055)  
 RREB1(cg07714276;4;c;T;0.111) - RUNDC2C(cg04194479;4;e;0.083) - S1PR5(cg26918756;1;e;0.141) -  
 SAMD4A(cg09397716;4;e;0.102) - SCN2B(cg04563671;4;f;T;0.072) - SCN4B(cg05269359;4;f;0.102) -  
 SCRNB(cg10093739;N2;e;0.084) - SDC1(cg10329928;4;e;T;0.079) - SDC2(cg15980656;4;e;T;0.144) - SFXN3(cg17858697;S3;e;0.08) -  
 SGCG(cg04678336;4;e;T;0.096) - SGMS1(cg11508429;4;c;T;-0.064) - SIK3(cg08190615;4;e;T;0.156) - SKI(cg01949002;4;e;0.108) -  
 SLC12A5(cg09595245;N2;a;0.066) - SLC22A18(cg24409566;S3;e;T;-0.131) - SLC26A1(cg21743826;4;e;0.062) -  
 SLC2A1(cg09502149;4;e;T;-0.139) - SLC38A3(cg04682699;4;c;T;0.052) - SLC43A2(cg19880947;1;e;0.072) -  
 SLC45A3(cg10581876;N3;c;0.083) - SLC5A10(cg21495715;4;a;0.071) - SLC5A10(cg02758593;4;f;0.084) - SNORD114-  
 25(cg10472263;4;b;0.064) - SNRNP40(cg15084470;N2;e;0.219) - SNX10(cg02389084;4;e;0.132) - SPARC(cg23174201;4;e;0.083) -  
 SPRED2(cg08467103;4;b;T;-0.147) - SPSB1(cg09256832;4;c;-0.077) - SRC(cg01141721;N2;c;0.061) -  
 SREBF1(cg23875758;N2;f;T;0.134) - SRPX2(cg02779592;4;b;-0.092) - SSBP3(cg16096432;4;e;T;0.055) - ST14(cg22110158;4;e;T;-  
 0.109) - ST3GAL1(cg03965649;4;c;T;0.1) - ST3GAL2(cg01389386;N2;c;-0.074) - ST5(cg08726522;4;e;-0.092) -  
 STARD13(cg07499182;4;e;T;-0.077) - STK31(cg14898779;4;a;0.073) - SYNJ2(cg10288437;4;e;T;0.078) -  
 TANC1(cg23401088;4;c;T;0.085) - TAP2(cg25744682;4;e;0.101) - TBC1D1(cg00812557;4;e;T;0.088) - TBC1D8(cg20893936;4;e;T;-  
 0.094) - TBCD(cg21156912;4;e;T;0.142) - TEAD2(cg01468567;S2;f;T;0.085) - TET2(cg22794775;4;c;T;0.057) -  
 TFEB(cg17513832;S3;e;T;0.09) - TFPI2(cg22799321;1;e;0.055) - THRAP3(cg26661718;4;c;T;0.102) -

THRAP3(cg16056044;4;c;T;0.109) - TLE1(cg13895650;4;e;T;0.131) - TMEM35(cg16510657;4;d;0.098) -  
 TMEM49(cg00040016;4;e;T;-0.058) - TMPRSS6(cg19979738;S3;e;0.101) - TMSB15A(cg01737010;S2;a;0.097) -  
 TNIK(cg10180092;4;e;T;-0.139) - TNS1(cg12338137;4;e;T;0.056) - TNS1(cg12004641;4;e;0.057) - TNS1(cg07492051;4;e;T;0.101) -  
 TNS3(cg07488141;4;c;T;-0.07) - TOLLIP(cg19554037;S2;e;0.136) - TOM1L2(cg04324276;4;e;T;-0.077) -  
 TRAK2(cg07226481;4;c;T;0.111) - TRIM26(cg05489957;N3;c;0.137) - TRIM26(cg08850243;4;c;0.083) -  
 TRIM35(cg05755408;4;e;T;0.06) - TRIM40(cg23698950;4;e;0.073) - TRIM71(cg24629438;4;e;T;0.054) -  
 TRPM5(cg26204383;1;e;0.109) - TRRAP(cg21421984;4;e;T;0.109) - TSPAN18(cg20968743;4;c;T;-0.127) -  
 TTC25(cg23017728;4;e;T;0.057) - TTC7A(cg02596427;4;e;T;0.068) - TUBGCP2(ch.10.2988224F;N2;e;0.061) -  
 TWISTNB(cg00726147;S2;a;0.077) - TXNDC11(cg03382501;4;e;T;0.078) - TXNDC12(ch.1.1540554F;N3;e;0.073) -  
 UBE2CBP(ch.6.1693624F;4;e;0.069) - UNC13D(cg16354117;4;e;T;0.085) - UNC84A(cg01947415;N3;f;0.076) -  
 USP40(cg27049539;4;e;T;-0.116) - USP43(cg12130768;S2;e;-0.071) - VAC14(cg00259097;4;e;T;0.055) -  
 VGLL4(cg03370106;4;e;T;0.05) - VPS13D(cg04920869;4;e;0.054) - VT1A(cg09958560;4;e;T;-0.077) -  
 WDR25(cg21550372;4;e;T;0.086) - WDR51A(cg19774788;4;e;-0.304) - XPNPEP1(cg18780288;4;e;T;0.138) -  
 ZBTB38(cg02004979;4;c;0.076) - ZCCHC14(cg27414087;N3;e;T;0.065) - ZDHHC18(cg06511276;S2;e;T;0.053) -  
 ZFH3(cg04667640;4;e;0.084) - ZFPM1(cg07099810;S3;e;0.068) - ZFYVE28(cg08393972;4;e;0.119) -  
 ZNF367(cg13951450;4;e;T;0.078) - ZNF385A(cg04964471;N3;e;0.068) - ZSWIM1(cg13780718;4;f;-0.093) -

**Supplementary table 4: Gene ontology analysis of the TGF- $\beta$  DNA methylation signature.**

Pathways and Transcription factors genesets analyses were performed as described for Supplementary table 2.

The tables presented here show the results obtained using the BrB geneset class comparison tool (top 50 for KEEG and top 100 for Biocarta).

Kegg Pathway	Pathway description	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
hsa00190	Oxidative phosphorylation	0.00001	0.14617	0.405 (-)
hsa00240	Pyrimidine metabolism	0.00001	0.61785	0.405 (-)
hsa00511	N-Glycan degradation	0.00001	0.00016	0.385 (+)
hsa00562	Inositol phosphate metabolism	0.00001	0.05171	0.205 (+)
hsa00563	Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	0.00001	0.00026	0.35 (+)
hsa00564	Glycerophospholipid metabolism	0.00001	0.9684	0.185 (+)
hsa00970	Aminoacyl-tRNA biosynthesis	0.00001	0.27511	0.52 (-)
hsa04070	Phosphatidylinositol signaling system	0.00001	0.01026	0.085 (+)
hsa04110	Cell cycle	0.00001	0.27071	0.45 (+)
hsa04150	mTOR signaling pathway	0.00001	0.85336	0.37 (+)
hsa04210	Apoptosis	0.00001	0.25794	0.47 (+)
hsa04520	Adherens junction	0.00001	0.50218	0.015 (+)
hsa04530	Tight junction	0.00001	0.88315	0.145 (+)
hsa04540	Gap junction	0.00001	0.00452	0.005 (-)
hsa04810	Regulation of actin cytoskeleton	0.00001	0.53567	0.235 (+)
hsa04910	Insulin signaling pathway	0.00001	0.9539	0.075 (+)
hsa04912	GnRH signaling pathway	0.00001	0.21833	0.015 (+)
hsa05120	Epithelial cell signaling in Helicobacter pylori infection	0.00001	0.00023	0.135 (+)
hsa05210	Colorectal cancer	0.00002	0.72525	0.225 (-)
hsa00561	Glycerolipid metabolism	0.00002	0.06381	0.27 (+)
hsa00271	Methionine metabolism	0.00002	0.01987	0.4 (+)
hsa01032	Glycan structures - degradation	0.00003	0.01856	0.3 (+)
hsa04120	Ubiquitin mediated proteolysis	0.00003	0.72438	0.53 (-)
hsa00193	ATP synthesis	0.00003	0.12841	0.33 (-)
hsa00290	Valine, leucine and isoleucine biosynthesis	0.00003	0.08402	0.045 (+)
hsa03030	DNA polymerase	0.00004	0.43373	0.51 (-)
hsa00051	Fructose and mannose metabolism	0.00004	0.82038	< 0.005 (+)
hsa04330	Notch signaling pathway	0.00007	0.10715	0.255 (+)
hsa00510	N-Glycan biosynthesis	0.00008	0.39709	0.5 (+)
hsa04620	Toll-like receptor signaling pathway	0.00013	0.7175	0.49 (-)
hsa03320	PPAR signaling pathway	0.00015	0.03013	< 0.005 (+)
hsa00062	Fatty acid elongation in mitochondria	0.00029	0.02233	0.47 (+)
hsa00310	Lysine degradation	0.0003	0.97273	0.465 (-)
hsa00790	Folate biosynthesis	0.0004	0.39985	0.445 (+)
hsa03010	Ribosome	0.00043	0.13153	0.415 (-)
hsa04660	T cell receptor signaling pathway	0.00057	0.71642	0.075 (+)
hsa00280	Valine, leucine and isoleucine degradation	0.00075	0.8321	0.17 (-)
hsa03050	Proteasome	0.00128	0.00469	0.425 (-)
hsa04370	VEGF signaling pathway	0.00158	0.96511	0.495 (-)
hsa01031	Glycan structures - biosynthesis 2	0.00169	0.69978	0.12 (+)
hsa00230	Purine metabolism	0.0019	0.14026	0.315 (-)

hsa00600	Sphingolipid metabolism	0.00213	0.07846	0.47 (-)
hsa03022	Basal transcription factors	0.00218	0.37548	0.525 (+)
hsa04710	Circadian rhythm	0.00224	0.34024	0.275 (-)
hsa05050	Dentatorubropallidolysian atrophy (DRPLA)	0.00299	0.88051	0.1 (+)
hsa00071	Fatty acid metabolism	0.00342	0.23322	0.015 (+)
hsa04310	Wnt signaling pathway	0.00379	0.92809	0.175 (-)
hsa00520	Nucleotide sugars metabolism	0.00394	0.37041	0.415 (-)
hsa00750	Vitamin B6 metabolism	0.00403	0.24408	0.435 (-)
hsa04920	Adipocytokine signaling pathway	0.00427	0.90064	0.005 (+)

Biocarta Pathway	Pathway description	LS permutation p-value	KS permutation p-value	Efron-Tibshirani's GSA testp-value
h_arapPathway	ADP-Ribosylation Factor	0.00001	0.00112	0.01 (+)
h_carm-erPathway	CARM1 and Regulation of the Estrogen Receptor Transcription Regulation by Methyltransferase of CARM1	0.00001	0.00001	0.04 (+)
h_carm1Pathway	Dicer Pathway	0.00001	0.00001	0.03 (+)
h_dicerPathway	Map Kinase Inactivation of SMRT Corepressor	0.00001	0.00003	< 0.005 (+)
h_egfr_smrtePathway	METS affect on Macrophage Differentiation	0.00001	0.00004	0.035 (+)
h_etsPathway	Cyclin E Destruction Pathway	0.00001	0.00007	0.015 (+)
h_fb7Pathway	Cell Cycle: G1/S Check Point	0.00001	0.1774	0.2 (+)
h_g1Pathway	HIV-I Nef: negative effector of Fas and TNF	0.00001	0.86165	0.245 (-)
h_hivnefPathway	Signal transduction through IL1R	0.00001	0.00067	0.48 (+)
h_il1rPathway	Integrin Signaling Pathway	0.00001	0.00027	< 0.005 (-)
h_integrinPathway	MAPKinase Signaling Pathway	0.00001	0.81883	0.015 (+)
h_mapkPathway	mTOR Signaling Pathway	0.00001	0.48708	0.44 (+)
h_mTORPathway	NfκB activation by Nontypeable Hemophilus influenzae	0.00001	0.13397	0.44 (+)
h_nthiPathway	Regulation of p27 Phosphorylation during Cell Cycle Progression	0.00001	0.00004	0.295 (-)
h_p27Pathway	Mechanism of Gene Regulation by Peroxisome Proliferators via PPARα(α)	0.00001	0.2105	0.345 (+)
h_pparaPathway	Phosphoinositides and their downstream targets.	0.00001	0.4484	0.015 (+)
h_ptdinsPathway	Influence of Ras and Rho proteins on G1 to S Transition	0.00001	0.98657	< 0.005 (+)
h_RacCycDPPathway	Sumoylation by RanBP2 Regulates Transcriptional Repression	0.00001	0.00001	0.475 (+)
h_ranbp2Pathway	Nuclear receptors coordinate the activities of chromatin remodeling complexes and coactivators to facilitate initiation of transcription in carcinoma cells	0.00001	0.00001	0.05 (+)
h_rarrxrPathway	Rho cell motility signaling pathway	0.00001	0.51469	< 0.005 (+)
h_rhoPathway	TNFR1 Signaling Pathway	0.00001	0.49425	0.005 (+)
h_tnfr1Pathway	Control of Gene Expression by Vitamin D Receptor	0.00001	0.14317	0.385 (+)
h_vdrPathway	WNT Signaling Pathway	0.00002	0.00552	0.095 (+)
h_wntPathway	Eukaryotic protein translation	0.00002	0.04796	0.28 (+)
h_eifPathway	Human Cytomegalovirus and Map Kinase Pathways	0.00002	0.00897	0.115 (-)
h_hcmvPathway	Endocytotic role of NDK, Phosphins and Dynamin	0.00002	0.02751	0.52 (-)
h_ndkDynaminPathway	Links between Pyk2 and Map Kinases	0.00003	0.99864	0.4 (-)
h_pyk2Pathway	TGF beta signaling pathway	0.00003	0.19723	0.01 (-)
h_tgfbPathway	Role of BRCA1, BRCA2 and ATR in Cancer Susceptibility	0.00003	0.14163	0.07 (-)
h_atrbcaPathway	AKT Signaling Pathway	0.00003	0.08263	0.405 (-)
h_aktPathway	FAS signaling pathway ( CD95 )	0.00004	0.77015	0.475 (+)
h_fasPathway				0.465 (+)

h_Par1Pathway	Thrombin signaling and protease-activated receptors	0.00004	0.96416	0.24 (-)
h_skp2e2fPathway	E2F1 Destruction Pathway	0.00004	0.00133	0.16 (+)
h_ptenPathway	PTEN dependent cell cycle arrest and apoptosis	0.00004	0.02462	0.34 (+)
h_nfataPathway	NFAT and Hypertrophy of the heart (Transcription in the broken heart)	0.00005	0.35749	0.02 (-)
h_eif4Pathway	Regulation of eIF4e and p70 S6 Kinase	0.00006	0.47107	0.115 (-)
h_ctcfPathway	CTCF: First Multivalent Nuclear Factor	0.00006	0.21047	0.03 (-)
h_srcRPTTPPathway	Activation of Src by Protein-tyrosine phosphatase alpha	0.00008	0.01561	0.085 (-)
h_malPathway	Role of MAL in Rho-Mediated Activation of SRF	0.00008	0.43263	0.14 (-)
h_p38mapkPathway	p38 MAPK Signaling Pathway	0.00008	0.51076	0.4 (-)
h_g2Pathway	Cell Cycle: G2/M Checkpoint	0.00009	0.25382	0.275 (-)
h_pdgfPathway	PDGF Signaling Pathway	0.0001	0.99863	0.005 (-)
h_mCalpainPathway	mCalpain and friends in Cell motility	0.00012	0.05214	0.015 (+)
h_vipPathway	Neuropeptides VIP and PACAP inhibit the apoptosis of activated T cells	0.00014	0.54943	0.345 (+)
h_myosinPathway	PKC-catalyzed phosphorylation of inhibitory phosphoprotein of myosin phosphatase	0.00014	0.10408	0.065 (+)
h_cellcyclePathway	Cyclins and Cell Cycle Regulation	0.00015	0.23479	0.45 (-)
h_crebPathway	Transcription factor CREB and its extracellular signals	0.00016	0.79175	0.215 (+)
h_rasPathway	Ras Signaling Pathway	0.00016	0.81429	0.475 (+)
h_metPathway	Signaling of Hepatocyte Growth Factor Receptor	0.00027	0.99856	0.42 (-)
h_rac1Pathway	Rac 1 cell motility signaling pathway	0.00034	0.9935	0.065 (+)
h_arenrf2Pathway	Oxidative Stress Induced Gene Expression Via Nrf2	0.00035	0.38526	< 0.005 (-)
h_pitx2Pathway	Multi-step Regulation of Transcription by Pitx2	0.00039	0.59898	0.485 (+)
h_ps1Pathway	Presenilin action in Notch and Wnt signaling	0.00045	0.13323	0.41 (+)
h_cd40Pathway	CD40L Signaling Pathway	0.00049	0.00148	0.375 (-)
h_stressPathway	TNF/Stress Related Signaling	0.00061	0.08625	0.54 (-)
h_iresPathway	Internal Ribosome entry pathway	0.00062	0.0911	0.265 (-)
h_shhPathway	Sonic Hedgehog (Shh) Pathway	0.00068	0.52316	0.04 (+)
h_cxcr4Pathway	CXCR4 Signaling Pathway	0.00069	0.87326	0.17 (-)
h_alkPathway	ALK in cardiac myocytes	0.00071	0.19796	0.005 (-)
h_rarPathway	Degradation of the RAR and RXR by the proteasome	0.00073	0.00008	0.01 (+)
h_bard1Pathway	BRCA1-dependent Ub-ligase activity	0.00076	0.01974	< 0.005 (+)
h_tercPathway	Overview of telomerase RNA component gene	0.00086	0.0041	0.45 (+)
h_At1rPathway	hTerc Transcriptional Regulation	0.00087	0.99866	0.025 (-)
h_arfPathway	Angiotensin II mediated activation of JNK	0.00091	0.89704	0.47 (+)
h_telPathway	Tumor Suppressor Arf Inhibits Ribosomal Biogenesis	0.00096	0.72591	0.35 (+)
h_stathminPathway	Telomeres, Telomerase, Cellular Aging, and Immortality	0.0013	0.48951	0.065 (+)
h_pkcPathway	Stathmin and breast cancer resistance to antimicrotubule agents	0.00132	0.01442	< 0.005 (-)
h_gpcrPathway	Activation of PKC through G protein coupled receptor	0.00156	0.64652	0.49 (+)
h_biopeptidesPathway	Signaling Pathway from G-Protein Families	0.00166	0.97672	0.025 (-)
h_ceramidePathway	Bioactive Peptide Induced Signaling Pathway	0.00167	0.77289	0.42 (-)
h_chemicalPathway	Ceramide Signaling Pathway	0.00168	0.54377	< 0.005 (-)
h_edg1Pathway	Apoptotic Signaling in Response to DNA Damage	0.0018	0.58186	0.015 (-)
	Phospholipids as signalling intermediaries			

h_nfkbPathway	NF-kB Signaling Pathway	0.00182	0.29276	0.47 (+)
h_ppargPathway	Role of PPAR-gamma Coactivators in Obesity and Thermogenesis	0.00182	0.10436	< 0.005 (+)
h_fmLPpathway	fMLP induced chemokine gene expression in HMC-1 cells	0.00189	0.9857	0.03 (-)
h_notchpathway	Proteolysis and Signaling Pathway of Notch	0.00192	0.00366	0.335 (+)
h_gsk3Pathway	Inactivation of Gsk3 by AKT causes accumulation of b-catenin in Alveolar Macrophages	0.00192	0.37959	0.34 (+)
h_mcmPathway	CDK Regulation of DNA Replication	0.00198	0.75774	0.01 (-)
h_pepiPathway	Proepithelin Conversion to Epithelin and Wound Repair Control	0.00212	0.00291	< 0.005 (+)
h_gleevecpathway	Inhibition of Cellular Proliferation by Gleevec	0.00238	0.99886	0.44 (-)
h_p35alzheimersPathway	Deregulation of CDK5 in Alzheimers Disease	0.0024	0.17286	0.235 (-)
h_extrinsicPathway	Extrinsic Prothrombin Activation Pathway	0.00242	0.19078	0.045 (+)
h_pmlPathway	Regulation of transcriptional activity by PML	0.00258	0.44092	0.32 (+)
h_dreampathway	Repression of Pain Sensation by the Transcriptional Regulator DREAM	0.00259	0.50941	0.185 (+)
h_mPRPathway	How Progesterone Initiates the Oocyte Maturation	0.00265	0.63271	0.04 (+)
h_tollPathway	Toll-Like Receptor Pathway	0.00288	0.4299	0.485 (-)
h_freePathway	Free Radical Induced Apoptosis	0.00322	0.00228	0.36 (+)
h_ck1Pathway	Regulation of ck1/cdk5 by type 1 glutamate receptors	0.00338	0.56702	0.08 (+)
h_HuntingtonPathway	Inhibition of Huntington@	0.00369	0.03343	0.02 (+)
h_egfPathway	EGF Signaling Pathway	0.00376	0.99867	0.03 (-)
h_dbpbPathway	Transcriptional activation of dbpb from mRNA	0.00377	0.00171	0.045 (+)
h_hesPathway	Segmentation Clock	0.00435	0.04175	0.445 (-)
h_igf1mtorpathway	Skeletal muscle hypertrophy is regulated via AKT/mTOR pathway	0.00459	0.77548	0.485 (+)
h_HBxPathway	Calcium Signaling by HBx of Hepatitis B virus	0.098	0.00137	0.375 (-)
h_plcdPathway	Phospholipase C d1 in phospholipid associated cell signaling	0.11031	0.32886	< 0.005 (-)

### Supplementary table 5: List of genes differentially expressed after TGF- $\beta$ exposure and gene ontology analysis.

List of significant differentially expressed after TGF- $\beta$  exposure was obtained using the BrB class comparison tool, blocked by cell lines, filtered for p-value <0.001, FDR<0.05. Pathway enrichment analysis was performed as described for Supplementary table 2.

The tables presented here show the list of genes with a fold change >0.5 and the gene ontology analysis using the BrB geneset class comparison tool (molecular function).

Symbol	Name	geometric mean control	geometric mean TGFb	fold change (control/ TGFb)
ACLY	ATP citrate lyase	564.29	352.57	<b>1.6</b>
ACSS2	acyl-CoA synthetase short-chain family member 2	2514.54	847.86	<b>2.97</b>
ACSS2	acyl-CoA synthetase short-chain family member 2	2309.46	784.72	<b>2.94</b>
ACTB	actin, beta	2832.13	4747.66	<b>0.6</b>
ADRBK1	adrenergic, beta, receptor kinase 1	1167.95	774.28	<b>1.51</b>
AK4	adenylate kinase 4	333.62	222.76	<b>1.5</b>
AKR1B10	aldo-keto reductase family 1, member B10 (aldose reductase)	453.5	270.11	<b>1.68</b>
ALDOC	aldolase C, fructose-bisphosphate	875.08	524.1	<b>1.67</b>
ALPK2	alpha-kinase 2	101.34	242.15	<b>0.42</b>
AMBP	alpha-1-microglobulin/bikunin precursor	735.9	449.98	<b>1.64</b>
ANXA2	annexin A2	2009.01	3322.02	<b>0.6</b>
ARG2	arginase, type II	237.99	130.18	<b>1.83</b>
ATP6V0E2	ATPase, H+ transporting V0 subunit e2	229.28	467.55	<b>0.49</b>
BGN	biglycan	93.35	155.24	<b>0.6</b>
CADM1	cell adhesion molecule 1	230.86	389.95	<b>0.59</b>
CAPN12	calpain 12	279.64	119.11	<b>2.35</b>
CAPN12	calpain 12	223.68	107.74	<b>2.08</b>
CCBL1	cysteine conjugate-beta lyase, cytoplasmic	454.55	285.39	<b>1.59</b>
CD24	CD24 molecule	139.94	275.79	<b>0.51</b>
CD24	CD24 molecule	625.16	1916.31	<b>0.33</b>
CDCA5	cell division cycle associated 5	274.07	455.69	<b>0.6</b>
CDK10	cyclin-dependent kinase 10	421.28	276.83	<b>1.52</b>
CDK2AP1	cyclin-dependent kinase 2 associated protein 1	619.69	1052.67	<b>0.59</b>
CGN	cingulin	363.79	239.1	<b>1.52</b>
CHST3	carbohydrate (chondroitin 6) sulfotransferase 3	133.94	246.51	<b>0.54</b>
CIDEB	cell death-inducing DFFA-like effector b	315.86	210.97	<b>1.5</b>
CLDN1	claudin 1	701.6	1429.54	<b>0.49</b>
COL4A5	collagen, type IV, alpha 5	233.76	413.54	<b>0.57</b>
CRELD1	cysteine-rich with EGF-like domains 1	442.29	290.34	<b>1.52</b>
CRELD1		198.91	130.92	<b>1.52</b>
CRLS1	cardiolipin synthase 1	743.86	1240.3	<b>0.6</b>
CTSA	cathepsin A	434.41	773.53	<b>0.56</b>
CXCL16	chemokine (C-X-C motif) ligand 16	700.7	441.59	<b>1.59</b>
DHCR7	7-dehydrocholesterol reductase	2007.33	1269.42	<b>1.58</b>
DHCR7	7-dehydrocholesterol reductase	1679.23	1062.13	<b>1.58</b>
DKK1	dickkopf 1 homolog (Xenopus laevis)	187.92	526.11	<b>0.36</b>
DLK1	delta-like 1 homolog (Drosophila)	2537.34	469.88	<b>5.4</b>
DLK1	delta-like 1 homolog (Drosophila)	513.36	184.21	<b>2.79</b>
DUSP5	dual specificity phosphatase 5	358.92	201.14	<b>1.78</b>
DUSP6	dual specificity phosphatase 6	376.55	175.76	<b>2.14</b>
EBP	emopamil binding protein (sterol isomerase)	1442.64	725.37	<b>1.99</b>
ELOVL6	ELOVL fatty acid elongase 6	370.18	198.95	<b>1.86</b>
EPCAM	epithelial cell adhesion molecule	379.85	661.89	<b>0.57</b>
ERP27	endoplasmic reticulum protein 27	80.75	152.14	<b>0.53</b>
FABP5	fatty acid binding protein 5 (psoriasis-associated)	192.15	318.53	<b>0.6</b>



FAM13A	family with sequence similarity 13, member A	206.76	137.32	<b>1.51</b>
FAM162A	family with sequence similarity 162, member A	187.38	105.47	<b>1.78</b>
FDFT1	farnesyl-diphosphate farnesyltransferase 1	3400.49	1978.14	<b>1.72</b>
FDFT1	farnesyl-diphosphate farnesyltransferase 1	724.69	428.51	<b>1.69</b>
FDPS	farnesyl diphosphate synthase	606.88	349.13	<b>1.74</b>
FOXN2	forkhead box N2	323.2	188.66	<b>1.71</b>
FTH1	ferritin, heavy polypeptide 1	542.72	293.32	<b>1.85</b>
FTH1P2	ferritin, heavy polypeptide 1 pseudogene 2	725.45	476.55	<b>1.52</b>
GK	glycerol kinase	420.34	253.78	<b>1.66</b>
GK	glycerol kinase	331.05	213.38	<b>1.55</b>
GLRX	glutaredoxin (thioltransferase)	1262.39	801.12	<b>1.58</b>
GPD1L	glycerol-3-phosphate dehydrogenase 1-like	228.78	422.07	<b>0.54</b>
GPX2	glutathione peroxidase 2 (gastrointestinal)	265.49	453.49	<b>0.59</b>
GPX2		775.35	1339.18	<b>0.58</b>
GSDMB	gasdermin B	506.8	276.73	<b>1.83</b>
GSDMB	gasdermin B	214.58	136.99	<b>1.57</b>
GSDMB		2183.8	1394.19	<b>1.57</b>
H2AFY2	H2A histone family, member Y2	214.06	408.1	<b>0.52</b>
HGD	homogentisate 1,2-dioxygenase	1053.88	691.1	<b>1.52</b>
HMGCS1	3-hydroxy-3-methylglutaryl-CoA synthase 1 (soluble)	3500.41	1742.4	<b>2.01</b>
HSD17B7P 2	hydroxysteroid (17-beta) dehydrogenase 7 pseudogene 2	281.12	154.56	<b>1.82</b>
IGSF1	immunoglobulin superfamily, member 1	624.5	374.47	<b>1.67</b>
INPP1	inositol polyphosphate-1-phosphatase	700.8	432.6	<b>1.62</b>
ISG20	interferon stimulated exonuclease gene 20kDa	350.49	227.61	<b>1.54</b>
KIAA1984	KIAA1984	253.66	147.9	<b>1.72</b>
KIAA1984		294.73	172.85	<b>1.71</b>
KLF9	Kruppel-like factor 9	235.65	143.99	<b>1.64</b>
KLHL5	kelch-like 5 (Drosophila)	251.17	522.35	<b>0.48</b>
KNG1	kininogen 1	575.99	257.98	<b>2.23</b>
KRT19	keratin 19	244.95	699.73	<b>0.35</b>
LPIN1	lipin 1	378.77	234.11	<b>1.62</b>
MAPK13	mitogen-activated protein kinase 13	208.88	425	<b>0.49</b>
MASP1	mannan-binding lectin serine peptidase 1 (C4/C2 activating component of Ra-reactive factor)	283.68	130.38	<b>2.18</b>
MATN3	matrilin 3	128.76	287.45	<b>0.45</b>
MEP1A	meprin A, alpha (PABA peptide hydrolase)	375.81	122.62	<b>3.06</b>
MSMO1	methylsterol monooxygenase 1	1397	695.35	<b>2.01</b>
MTTP	microsomal triglyceride transfer protein	496.97	306.57	<b>1.62</b>
MVD	mevalonate (diphospho) decarboxylase	482.45	269.37	<b>1.79</b>
NCR3LG1	natural killer cell cytotoxicity receptor 3 ligand 1	318.36	207.24	<b>1.54</b>
NMB	neuromedin B	187.66	313.73	<b>0.6</b>
NR2F1	nuclear receptor subfamily 2, group F, member 1	419.98	280.76	<b>1.5</b>
NRP1	neuropilin 1	109.25	203.54	<b>0.54</b>
PCSK9	proprotein convertase subtilisin/kexin type 9	561.17	226.47	<b>2.48</b>
PDLIM1	PDZ and LIM domain 1	523.22	970.6	<b>0.54</b>
PDLIM3	PDZ and LIM domain 3	165.05	92.01	<b>1.79</b>
PFKFB4	6-phosphofructo-2-kinase/fructose-2,6- biphosphatase 4	615.5	366.73	<b>1.68</b>
PHLDA1	pleckstrin homology-like domain, family A, member 1	280.14	699.81	<b>0.4</b>
PIR	pirin (iron-binding nuclear protein)	380.9	247.62	<b>1.54</b>
PIR	pirin (iron-binding nuclear protein)	303.41	200.28	<b>1.51</b>
PLCXD1	phosphatidylinositol-specific phospholipase C, X domain containing 1	1096.4	607.03	<b>1.81</b>
PPIC	peptidylprolyl isomerase C (cyclophilin C)	104.84	185.78	<b>0.56</b>
PPIC	peptidylprolyl isomerase C (cyclophilin C)	136.21	276.95	<b>0.49</b>
PROM1	prominin 1	166.12	405.7	<b>0.41</b>
QPCT	glutaminy-peptide cyclotransferase	92.68	179.44	<b>0.52</b>

RGL1	ral guanine nucleotide dissociation stimulator-like 1	143.27	291.97	<b>0.49</b>
RHOBTB3	Rho-related BTB domain containing 3	234.86	409.59	<b>0.57</b>
RPL14	ribosomal protein L14	342.85	225.07	<b>1.52</b>
SCD	stearoyl-CoA desaturase (delta-9-desaturase)	5170.79	3241.11	<b>1.6</b>
SCN9A	sodium channel, voltage-gated, type IX, alpha subunit	146.2	261.36	<b>0.56</b>
SGSH	N-sulfoglucosamine sulfohydrolase	268.25	482.61	<b>0.56</b>
SLC17A2	solute carrier family 17 (sodium phosphate), member 2	319.48	171.13	<b>1.87</b>
SLC1A7	solute carrier family 1 (glutamate transporter), member 7	154.75	95.69	<b>1.62</b>
SLC22A18	solute carrier family 22, member 18	346.16	200.56	<b>1.73</b>
SLC26A6	solute carrier family 26, member 6	1984.26	1226.58	<b>1.62</b>
SLC38A3	solute carrier family 38, member 3	147.61	98.31	<b>1.5</b>
SLC38A3		280.16	186.55	<b>1.5</b>
SLC7A2	solute carrier family 7 (cationic amino acid transporter, y+ system), member 2	648	345.74	<b>1.87</b>
SNORA12	small nucleolar RNA, H/ACA box 12	265.54	163.35	<b>1.63</b>
SOAT2	sterol O-acyltransferase 2	428.99	170.49	<b>2.52</b>
SOAT2	sterol O-acyltransferase 2	207.1	110.54	<b>1.87</b>
SPP1	secreted phosphoprotein 1	163.7	664.59	<b>0.25</b>
SPP1	secreted phosphoprotein 1	175.06	761.54	<b>0.23</b>
ST6GALNA C6	ST6 (alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 6	295.18	173.55	<b>1.7</b>
STAT1	signal transducer and activator of transcription 1, 91kDa	251.15	427.18	<b>0.59</b>
STC2	stanniocalcin 2	415.34	232.36	<b>1.79</b>
TESC	tescalcin	173.93	339.61	<b>0.51</b>
TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	871.16	490.16	<b>1.78</b>
TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	2995.94	1802.58	<b>1.66</b>
TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	2789.58	1682.29	<b>1.66</b>
TGM2	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase)	99.43	186.24	<b>0.53</b>
THBS1	thrombospondin 1	106.14	216.71	<b>0.49</b>
TK1	thymidine kinase 1, soluble	201.7	342.63	<b>0.59</b>
TKT	transketolase	1170.32	766.64	<b>1.53</b>
TMC6	transmembrane channel-like 6	346.78	203.47	<b>1.7</b>
TMEM150 A	transmembrane protein 150A	569.33	377.32	<b>1.51</b>
TMSB4X	thymosin beta 4, X-linked	447.67	2153.79	<b>0.21</b>
TNFRSF19	tumor necrosis factor receptor superfamily, member 19	147.03	491.99	<b>0.3</b>
TNFRSF21	tumor necrosis factor receptor superfamily, member 21	592.82	1305.55	<b>0.45</b>
TNFSF4	tumor necrosis factor (ligand) superfamily, member 4	165.99	92.03	<b>1.8</b>
TNFSF4	tumor necrosis factor (ligand) superfamily, member 4	129.86	83.75	<b>1.55</b>
TNS3	tensin 3	632.11	1208.55	<b>0.52</b>
TRIB3	tribbles homolog 3 (Drosophila)	1916.69	1019.26	<b>1.88</b>
TSPO	translocator protein (18kDa)	259.65	495.59	<b>0.52</b>
UGDH	UDP-glucose 6-dehydrogenase	546.86	338.2	<b>1.62</b>
UGDH		904.61	557.07	<b>1.62</b>
UGT2B11	UDP glucuronosyltransferase 2 family, polypeptide B11	293.91	176.27	<b>1.67</b>
UGT2B17	UDP glucuronosyltransferase 2 family, polypeptide B17	304.43	136.17	<b>2.24</b>

UGT2B4	UDP glucuronosyltransferase 2 family, polypeptide B4	192.45	96.53	<b>1.99</b>
VCAN	versican	128.95	276.32	<b>0.47</b>
WNK4	WNK lysine deficient protein kinase 4	195.31	118.45	<b>1.65</b>
ZMYM3	zinc finger, MYM-type 3	321.28	177.3	<b>1.81</b>

<u>Gene ontology ID</u>	<u>Gene ontology Term</u>	<u>Observed/Expected</u>
GO:0034713	type I transforming growth factor beta receptor binding	8.1
GO:0016846	carbon-sulfur lyase activity	7.72
GO:0009008	DNA-methyltransferase activity	7.29
GO:0070402	NADPH binding	6.62
GO:0008409	5'-3' exonuclease activity	5.61
GO:0016634	oxidoreductase activity, acting on the CH-CH group of donors, oxygen as acceptor	4.55
GO:0004602	glutathione peroxidase activity	4.55
GO:0016863	intramolecular oxidoreductase activity, transposing C=C bonds	4.05
GO:0005527	macrolide binding	4.05
GO:0016796	exonuclease activity, active with either ribo- or deoxyribonucleic acids and producing 5'-phosphomonoesters	3.89
GO:0005160	transforming growth factor beta receptor binding	3.64
GO:0016878	acid-thiol ligase activity	3.5
GO:0005024	transforming growth factor beta-activated receptor activity	3.5
GO:0004033	aldo-keto reductase (NADP) activity	3.47
GO:0017136	NAD-dependent histone deacetylase activity	3.45
GO:0034979	NAD-dependent protein deacetylase activity	3.28
GO:0070325	lipoprotein particle receptor binding	3.24
GO:0015036	disulfide oxidoreductase activity	3.24
GO:0045309	protein phosphorylated amino acid binding	3.17
GO:0050661	NADP binding	2.98
GO:0005048	signal sequence binding	2.91
GO:0001077	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	2.91
GO:0016877	ligase activity, forming carbon-sulfur bonds	2.85
GO:0001228	RNA polymerase II transcription regulatory region sequence-specific DNA binding transcription factor activity involved in positive regulation of transcription	2.79
GO:0016620	oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor	2.73
GO:0051219	phosphoprotein binding	2.62
GO:0048306	calcium-dependent protein binding	2.62
GO:0004675	transmembrane receptor protein serine/threonine kinase activity	2.59
GO:0005520	insulin-like growth factor binding	2.51
GO:0017048	Rho GTPase binding	2.29
GO:0016769	transferase activity, transferring nitrogenous groups	2.22
GO:0019829	cation-transporting ATPase activity	2.21
GO:0004693	cyclin-dependent protein kinase activity	2.21
GO:0016814	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in cyclic amidines	2.19
GO:0016903	oxidoreductase activity, acting on the aldehyde or oxo group of donors	2.15
GO:0050840	extracellular matrix binding	2.14

GO:0008483	transaminase activity	2.08
GO:0005501	retinoid binding	2.03
GO:0035064	methyated histone residue binding	2.02
GO:0004520	endodeoxyribonuclease activity	2.02

**Supplementary table 6: List of genes overlapping the methylome and transcriptome signatures after TGF- $\beta$  exposure.**

List of common genes between TGF- $\beta$  methylome and TGF- $\beta$  transcriptome. Ratio between control and TGF- $\beta$  treated samples for methylation and expression arrays are displayed. Gene regions (UCSC refgene group) and enhancer annotations are also indicated.

	Methylation array (control/TGF)	whole genome expression array (control/TGFb)	USCS_REFGENE GROUP	ENHANCER
<i>ACSL3</i>	0.74	1.32	5UTR	TRUE
<i>AHNAK</i>	1.37	1.33	5UTR	TRUE
<i>BCR</i>	0.73	0.93	body	TRUE
<i>BMP1</i>	1.85	1.12	body	TRUE
<i>BRD2</i>	0.66	1.37	body	
<i>C17orf101</i>	0.66	0.84	body	
<i>CALD1</i>	1.75	0.74	TSS/body	TRUE
<i>CALM2</i>	0.73	0.73	3UTR	
<i>COL18A1</i>	0.46	0.69	TSS/body	
<i>DACT2</i>	0.64	0.84	body	
<i>DDA1</i>	1.32	0.68	body	TRUE
<i>DDX19B</i>	0.77	0.81	body	
<i>DNMT3B</i>	0.72	0.87	5UTR/TSS	
<i>ERLIN1</i>	0.63	1.36	body	
<i>GIPC1</i>	0.63	0.68	body	
<i>HDAC7</i>	0.72	0.78	body	TRUE
<i>MAEA</i>	1.58	0.78	body	
<i>NRP2</i>	0.73	0.88	body	TRUE
<i>PDLIM1</i>	2.21	0.53	body	TRUE
<i>RAP1GAP2</i>	0.77	0.74	body	TRUE
<i>RERE</i>	0.72	0.84	body	
<i>SLC22A18</i>	1.58	1.79	body/TSS	TRUE
<i>SRC</i>	0.75	1.28	5UTR	
<i>STARD13</i>	1.48	0.84	body	TRUE
<i>TLE1</i>	0.56	0.69	body	TRUE
<i>WDR25</i>	0.72	0.87	body	TRUE

**Supplementary table 7. List of genes overlapping the two methylome signatures (CD133+ and TGF- $\beta$ ).**

List of common genes between CD133+ methylation and TGF- $\beta$  signatures. Pathway enrichment analysis was performed using WebGesalt web application as described for Supplementary table 2.

APBA1; ARHGAP31; ATP1A4; BCL11A; BRSK2; CARD11; CARD14; CBFA2T3; CCDC40; CLDN10; CNNM4; CRTAC1; CSGALNACT1; CSGALNACT2; CSRNP3; CTBP2; DAB1; DNAH10; DNAH17; DOK7; EBF2; EBF3; ELMO1; FBN2; FEZ1; FLT1; FMN1; GALNT9; GFRA2; GIPC1; GLI2; GLIS1; GNG4; GPR133; GPR68; GRK1; HR; IFT140; IGDCC4; IGFBP3; JAKMIP3; KAL1; KAZN; KCNK1; KLF12; LAMC2; LINC00162; LINC00461; LOC154449; LTBP4; MAP7; MAST1; MDGA1; MOB2; MPPED2; MS4A3; MYO1F; MYO7A; NBL1; NBPF3; NFIC; NGF; NKX2-3; NMNAT2; NMU; NOS1; NOTCH4; NRD1; OPCML; OSBPL5; OXGR1; PCDH8; PDE4A; PDE4B; PHLDB1; PLEC; PLEKHA7; POU2F3; PRDM16; PRKAG2; PRSS33; PTPRN2; RADIL; RASA3; RBFOX1; RGMA; RPS6KA2; RPTOR; SLC6A7; SLC05A1; SLITRK4; SMOC2; SOGA2; SPARC; SPEG; SPG20; SRRM3; ST18; ST8SIA2; STK32C; SYNE1; TBX21; TCF7L1; TMEM26; TNFAIP8L3; TSPAN18; VANGL2; VAV2; VAX2; VIPR2; ZFYVE28; ZIC4; ZNF862;

Wikipathways	Number of genes	Adjusted P value
Myometrial Relaxation and Contraction Pathways	4	0.0153
G Protein Signaling Pathways	3	0.0170
Neural Crest Differentiation	3	0.0179
TOR signaling	2	0.0179
Notch Signaling Pathway	2	0.0214
AMPK signaling	2	0.0371

KEEG	Number of genes	Adjusted P Value
Glycosaminoglycan biosynthesis - chondroitin sulfate	2	0.0144
Chemokine signaling pathway	4	0.0144
Wnt signaling pathway	3	0.0276
Basal cell carcinoma	2	0.0276
mTOR signaling pathway	2	0.0276
Notch signaling pathway	2	0.0276
Pathways in cancer	4	0.0278
B cell receptor signaling pathway	2	0.0334
Focal adhesion	3	0.0334
Salivary secretion	2	0.0414

## **ANNEXE II: Review**

“ From Hepatitis to Hepatocellular carcinoma: a proposed model for cross-talk between inflammation and epigenetic mechanisms”.





OPINION

# From hepatitis to hepatocellular carcinoma: a proposed model for cross-talk between inflammation and epigenetic mechanisms

Marion Martin and Zdenko Herceg\*

## Abstract

Inflammation represents the body's natural response to tissue damage; however, chronic inflammation may activate cell proliferation and induce deregulation of cell death in affected tissues. Chronic inflammation is an important factor in the development of hepatocellular carcinoma (HCC), although the precise underlying mechanism remains unknown. Epigenetic events, which are considered key mechanisms in the regulation of gene activity states, are also commonly deregulated in HCC. Here, we review the evidence that chronic inflammation might deregulate epigenetic processes, thus promoting oncogenic transformation, and we propose a working hypothesis that epigenetic deregulation is an underlying mechanism by which inflammation might promote HCC development. In this scenario, different components of the inflammatory response might directly and indirectly induce changes in epigenetic machineries ('epigenetic switch'), including those involved in setting and propagating normal patterns of DNA methylation, histone modifications and non-coding RNAs in hepatocytes. We discuss the possibility that self-reinforcing cross-talk between inflammation and epigenetic mechanisms might amplify inflammatory signals and maintain a chronic state of inflammation culminating in cancer development. The potential role of inflammation-epigenome interactions in the emergence and maintenance of cancer stem cells is also discussed.

**Keywords** Cancer stem cells, epigenetic mechanisms, epigenetic switch, hepatitis, hepatocellular carcinoma, inflammation.

## Hepatocellular carcinoma: the importance of inflammation and epigenetics

Hepatocellular carcinoma (HCC) is the major form of primary liver cancer in the world [1], accounting for 662,000 deaths worldwide per year [2]. HCC is frequently diagnosed at an advanced stage, resulting in rather poor survival rates. HCC typically starts with a pre-existing liver disease caused by infection with hepatitis B virus (HBV) or hepatitis C virus (HCV), chronic aflatoxin exposure or alcohol consumption [3]. Chronic liver damage associated with chronic exposure to these agents results in cirrhosis (scarring of the liver characterized by the formation of fibrous tissue and destruction of normal architecture of the organ), which can eventually progress to liver cancer. People infected with HCV have an 80% chance of developing cirrhosis, and HBV-infected people have a 30-fold higher risk of developing cancer [1,3]. Therefore, there is a need for better understanding of the mechanisms underlying HCC development and progression, as these might improve our ability to detect the disease at earlier stages and design new efficient strategies for detection and treatment.

Deregulation of the epigenome (the totality of epigenetic marks in a cell, including DNA methylation, histone modifications and non-coding RNAs) is thought to play an important role in tumor development and progression. Epigenetic events are considered key mechanisms in the regulation of gene activity, and abnormal expression of a large number of tumor-suppressor genes and cancer-associated genes has been observed in a wide range of human cancers [4-8]. Epigenetic alterations might occur as early events in carcinogenesis and might precede genetic alterations during oncogenic transformation [9]. Moreover, large-scale studies involving epigenomic technologies have identified new genes targeted by aberrant epigenetic changes, and indicated epigenetic patterns that are consistently associated with different cancer types, including lung cancer, colorectal cancer and HCC. These 'epigenetic signatures' can be associated with predisposition factors or clinical outcome [4,10,11], and can be defined as specific epigenetic changes or a

\*Correspondence: [herceg@iarc.fr](mailto:herceg@iarc.fr)

Epigenetics Group, International Agency for Research on Cancer (IARC), 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France

combination thereof that are consistently associated with etiological or clinicopathological features of a tumor.

The inflammatory response is the tissue's natural response to damage. The primary functions of the inflammatory process are to defend the organism against harmful agents and products, remove damaged cells and facilitate the renewal of damaged tissues. However, chronic inflammation might activate cell proliferation and deregulation of cell death in affected tissues [12,13]. Inflammation is an important factor in HCC development, and chronic hepatitis might promote hepatocarcinogenesis through induction of cirrhosis, although inflammation-mediated HCC tumors might also develop in the absence of cirrhotic lesions. Because deregulation of epigenetic mechanisms is one of the hallmarks of cancer, including HCC, it is possible that inflammation might act through epigenetic mechanisms to promote liver cancer. Several comprehensive reviews on the role of either epigenetic deregulation [11] or inflammation [14] individually in liver cancer have been published. Here, we focus on the potential role of epigenetic mechanisms in inflammation-mediated processes during hepatocarcinogenesis and discuss how a cross-talk between inflammation and the epigenome can be exploited in the development of novel and efficient strategies for the treatment and prevention of liver cancer.

### **Deregulation of the inflammatory response during hepatocarcinogenesis**

HCC is one of the well-known examples of inflammation-related cancer that slowly develops on a background of chronic inflammation. The molecular links that connect inflammation and liver tumors are not fully known; however, recent studies are beginning to unravel the underlying mechanisms. In this section, we will discuss the development of chronic inflammation in response to liver damage and viral infection, and then describe the role of cytokine secretion during HCC development. Finally, activation of nuclear factor (NF)- $\kappa$ B and STAT3 (signal transducers and activators of transcription 3) pathways and their consequences for hepatocarcinogenesis will be discussed.

### **Development of chronic inflammation in response to liver damage and viral infection**

One of the main functions of the liver is to detoxify the organism. Consequently, hepatocytes are constantly subjected to diverse infectious or toxic agents that can generate liver damage and initiate an inflammatory response [15]. The purpose of local inflammation is to clear the damage by activating apoptosis of affected hepatocytes, and to promote repair of the tissue by activating cell proliferation [12]. Under normal conditions, when damage is limited and can be rapidly repaired, the

inflammatory state is transient. However, if the tissue damage is severe or if the inflammatory stimulus persists, the inflammatory process is maintained and may progress to chronic inflammation with continuous proliferation of hepatocytes (see below). The cellular pathways that are activated during a prolonged inflammatory response may trigger a wide range of potentially harmful processes, such as induction of DNA damage through reactive oxygen species (ROS) accumulation [16]. Therefore, hepatocytes harboring extensive DNA damage and undergoing prolonged proliferation during chronic inflammation may result in the acquisition of mutations and growth advantages, thus promoting initiation and progression of hepatocellular carcinoma [17].

This scenario is more likely to be associated with excessive alcohol consumption or fat accumulation, or occur during infection with HBV or HCV. HBV or HCV infection can lead to cirrhosis and further development of HCC when the immune system fails to efficiently clear the virus from the liver, resulting in a chronic form of the disease [18-20]. During an infection, hepatitis virus antigens activate immune cells that trigger apoptosis of infected hepatocytes, thus inducing compensatory proliferation [21]. Chronic infection may also lead to the development of HCC through induction of mutations and chromosomal instability [22]. These genetic changes can occur during prolonged cell proliferation and, in the case of HBV infection, can be induced by integration of the viral DNA into human chromosomal DNA [18,20,23,24]. HBV and HCV can also directly initiate cell transformation through the actions of viral proteins that interfere with cellular pathways controlling cell survival, cell proliferation and apoptosis [25]. For example, the HBV protein encoded by the *HBX* gene is able to directly influence the transcription of genes involved in several signaling pathways, including *c-JUN*, *c-FOS*, *c-MYC*, *AP-1* and *P53* [19,26]. In HCV, the core protein (the viral gene product synthesized in the early phase of HCV infection) is known, among others, to inhibit apoptosis through activation of *c-MYC* and inhibition of the *P53* gene [27,28]. Therefore, chronic inflammation in the liver associated with viral infection may contribute to hepatocarcinogenesis through the deregulation of important cellular pathways.

### **Cytokine secretion and HCC development**

During chronic inflammation of the liver, hepatocyte proliferation is activated by local and infiltrated immune cells through paracrine signals involving cytokines [29,30] (Table 1). Among a wide range of cytokines involved in liver inflammation, TNF $\alpha$ , interleukins (IL6, IL1 $\alpha$ , IL1 $\beta$  and IL10), and TGF $\beta$  (transforming growth factor beta) are thought to play major roles [31]. Several large-scale studies investigating serum levels of cytokines revealed

**Table 1. Detected changes in selected cytokines in hepatocellular carcinoma and liver inflammation**

Cytokine	Upregulated or downregulated	Sample/material studied	References
TNF $\alpha$	Up and down	Cirrhotic tissue, HCC patient serum, solid tumors	[31,33,38,41,42]
IL6	Up and down	Cirrhotic tissue, hepatitis C, HCC patient serum	[31-34,40,41,43,137]
IL1 $\alpha$	Up	HCC patient serum	[31,33]
IL1 $\beta$	Up	Cirrhotic tissue, HCC patient serum, solid tumors	[31,33,38,42]
TGF $\beta$	Up and down	Urine, HCC patient serum	[31,36]
IL10	Up	Cirrhotic tissue, hepatitis B, HCC patient serum	[31,32,35,37,39,60]

HCC, hepatocellular carcinoma; IL, interleukin; TGF, transforming growth factor; TNF, tumor necrosis factor.

higher levels of IL1 $\beta$ , IL6, TNF $\alpha$ , TGF $\beta$  and IL10 in individuals with hepatitis in comparison with healthy controls [32-40]. Curiously, serum levels of IL6 and TNF $\alpha$  have been found to be lower in patients with HCC [41], whereas in solid tumors TNF $\alpha$  and IL1 $\beta$  levels are higher in normal tissue than in tumor cells [42]. These studies have provided inconsistent results; therefore, the precise impact of cytokine deregulation associated with chronic inflammation in liver cancer development remains unclear. Nevertheless, changes in cytokine expression are detectable in pre-cancerous stages such as chronic hepatitis and cirrhosis. Comparison between different forms of liver inflammation has revealed higher levels of IL6, TNF $\alpha$ , IL1 $\beta$  and IL10 in patients that have developed cirrhosis compared with those infected with HBV or HCV in the absence of cirrhosis [37,38,43]. Hence, cytokine expression is positively correlated with disease progression, suggesting that the deregulation of cytokine expression may be an early event in hepatocarcinogenesis that could actively participate in cancer development.

#### NF- $\kappa$ B and STAT3 pathway activation and their consequences for hepatocarcinogenesis

Cellular pathways activated by cytokines are involved in cell growth, cell survival, cell proliferation and apoptosis. Here we discuss the importance of the NF- $\kappa$ B and janus kinase (JAK)/STAT3 pathways in key biological functions that are deregulated in HCC [15,44,45].

NF- $\kappa$ B belongs to the REL transcription factor family and exists as a homodimer or heterodimer. In the absence of stimuli, the dimers remain inactive in the cytoplasm [46]. In liver cells, binding of TNF $\alpha$  or IL1 $\alpha$  to the cellular membrane leads to the activation of NF- $\kappa$ B, which can enter the nucleus and initiate transcription of several inflammatory target genes [46]. Even though NF- $\kappa$ B activation has been demonstrated in several solid tumors [47], few studies have investigated NF- $\kappa$ B status in HCC. In a small study involving 15 primary tumors, activation of NF- $\kappa$ B was detected in 87% of peritumoral tissues and in 80% of tumor tissues compared with healthy controls [48]. More recently, investigation of a larger cohort indicated activation of NF- $\kappa$ B in 25% of the tumor samples [49].

In contrast to studies using human cancer samples, there are many studies that have aimed to understand the role of NF- $\kappa$ B in hepatocarcinogenesis using *in vivo* rodent models of inflammation-induced HCC. These studies have revealed a dual function of NF- $\kappa$ B. First, as an anti-tumorigenic agent in hepatocytes, NF- $\kappa$ B may protect the liver by preventing excessive cell death and thus limiting the compensatory proliferation [49,50]. Second, there may be a pro-tumorigenic function whereby NF- $\kappa$ B activation may support tumor growth by increasing transformed hepatocyte proliferation [51,52]. Hence, the precise function of NF- $\kappa$ B activation during HCC initiation and development remains to be defined.

The JAK/STAT3 pathway may be activated through interaction of the cytokine IL6 with its receptor. IL6 binding activates the phosphorylation of a JAK (mostly JAK2), which in turn will phosphorylate STAT3 on amino acid Y705 [53]. Activated STAT3 forms homodimers and is translocated into the nucleus where it enhances the transcription of several genes belonging mainly to cell survival pathways. Other studies have observed that STAT3 is constitutively activated in a majority of HCC cases (60% or more of the samples analyzed) [54,55]. Additionally, mechanistic studies *in vivo* revealed that STAT3 activation cannot be a consequence of HCC but may actively participate in the progression from hepatitis to an advanced cancer stage. In a mouse model of inflammation induced by specific dietary regimes, HCC incidence is correlated with STAT3 activation [56], and in mice developing a liver inflammatory microenvironment, HCC occurrence is reduced by STAT3 inhibition [49]. These results hint at a pro-tumorigenic role of the JAK/STAT3 pathway in HCC growth, and also in HCC initiation and development.

#### Deregulation of epigenetic mechanisms during hepatocarcinogenesis

In cancer cells, the key cellular processes such as cell survival, cell growth, cell proliferation and apoptosis are deregulated by aberrant gene expression. Alterations in gene expression can be caused by epigenetic deregulation as well as genetic changes (that is, mutations) [7,57].

Epigenetic mechanisms, such as DNA methylation, histone modifications and non-coding RNAs are key regulators of gene activity states; therefore, epigenetic disruptions can affect transcription of the genes that establish and maintain cell identity and proliferation capacity. A number of studies have suggested that epigenetic mechanisms are altered in HCC and may play key roles in hepatocarcinogenesis [11] (Table 2).

#### DNA methylation changes in HCC

DNA methylation is a chemical modification of DNA, involving the addition of a methyl group (-CH<sub>3</sub>) to a nucleotide. In humans, methylation usually occurs at the fifth carbon atom in the nucleotide base cytosine (5-methylcytosine). DNA methylation has been primarily studied in the context of gene transcription and aberrant gene silencing, although it has also been implicated in the silencing of transposable elements [58]. DNA methylation impacts the level of compaction of chromatin, and this affects the interaction between DNA and transcription factors, and consequently influences DNA expression. Global genome hypomethylation is a common phenomenon found in many solid tumors, including HCC [59]. The comparison of global methylation levels in HCC tumors and matched non-tumorigenic liver tissues revealed a significant reduction in total 5-methylcytosine content in tumors [60]. Compared with the surrounding cirrhotic or non-neoplastic tissues, lower levels of methylation at repetitive elements *LINE-1*, *ALU* and *SAT-2* have been observed in liver tumors [61].

Global hypomethylation can contribute to carcinogenesis in two ways. First, in normal liver tissue, as in other healthy tissues, methylation of repetitive elements may contribute to genome integrity by silencing their transcription, thus preventing the activity of potentially harmful mobile genetic elements. DNA hypomethylation could thus explain the chromosome structural alterations and genetic mutations observed in HCC. This hypothesis is corroborated by a study showing that an excess of copies of the heterochromatin sequence 1q12 is correlated with global loss of methylcytosine [62]. Therefore, DNA hypomethylation may alter the interaction between the CpG-rich satellite DNA and chromatin proteins, resulting in heterochromatin decondensation and breakage. Second, global hypomethylation can result in oncogene activation. This notion is supported by oncogene promoter demethylation found during HCC progression [59,63].

In parallel to these genome-wide alterations, regional hypermethylation has been detected in CpG islands of tumor suppressor genes (TSGs) [59,63-65]. These hypermethylated CpG islands result most often in gene silencing. The targeted genes are involved in cell proliferation inhibition (*p16INK4A*, *p21*, *p27*, *RASSF1A*, *SOCS1-3*, *RIZ1*), apoptosis (*CASP8*, *XAF-1*, *ASPP1*, *ASPP2*), cell

adhesion and migration (*E-Cadherin*, *TFPI-2*), and DNA repair (*GSTP1*), and their silencing can promote cell transformation. TSGs and other cancer-associated genes (such as *RASSF1A*, *DOK1* and *CHRNA3*) have been found to be hypermethylated in a high percentage of human samples of HCC [63,66]. Importantly, precancerous lesions in liver, such as fibrosis and cirrhosis, have also been found to exhibit aberrant hypermethylation in TSGs [11]. These observations suggest that TSG hypermethylation may represent a tumor-initiating event in HCC progression. Together, changes in methylation states (both hypermethylation and hypomethylation) appear to play a critical role in liver tumor development, similar to other cancers such as colorectal cancer and Beckwith-Wiedemann syndrome [67].

#### Histone modifications in hepatocellular carcinoma

Chemical modifications on histones (mainly acetylation and methylation on histones H3 or H4) are involved in gene expression through their role in the recruitment of inhibitors or enhancers of transcription. These modifications occur essentially in gene promoters to stimulate or inhibit gene expression.

In a methyl-deficient rodent model of hepatocarcinogenesis, in which hepatocarcinogenesis can be followed from preneoplastic nodules [68], it has been found that levels of H3K9 and H4K20 trimethylation (histone marks associated with repressive and activating transcriptional states, respectively) change during cancer development. In accordance with these changes, upregulation and down-regulation of Suv39h1 and Suv4-20h2, the enzymes responsible for H3K9 and H4K20 methylation, have been observed.

Progressive changes in histone mark patterns (mediated by activation or inactivation of specific histone-modifying complexes) have also been observed in a model of cell reprogramming. Fusion of mouse HCC cells with embryonic stem cells results in the loss of HCC cellular phenotype and reactivation of the tumor suppressor gene *p16INK4A*. Induced differentiation of these reprogrammed cells restores the original HCC phenotype in association with progressive silencing of *p16INK4A* [69]. During differentiation, the *p16INK4A* promoter is promptly 'invaded' by H3K27 trimethylation, accompanied by H3K9 dimethylation at later stages. Finally, histone H3 and H4 deacetylation (commonly associated with inhibition of transcription) are involved in several gene expression alterations in HCC [70,71], and mechanistic studies have revealed that histone deacetylation could also act in association with DNA methylation to induce gene silencing [64,65]. Despite the lack of large-scale studies with human samples, these results suggest that histone modifications may play an important role in hepatocarcinogenesis.



**Table 2. Epigenetic deregulation in hepatocellular carcinoma**

	Upregulation or downregulation	Sample/material studied	Reference(s)
DNA methylation			
Genome-wide	Down	Tumors	[60,62]
Repetitive elements ( <i>LINE-1</i> , <i>SAT-2</i> , <i>ALU</i> )	Down	Tumors	[61]
<i>p16INK4A</i> (TSG)	Up	Cirrhotic tissue, blood, serum, tumors	[59,63,86]
<i>RASSF1A</i> (TSG)	Up	Cell lines, cirrhotic tissue, serum, tumors	[59,63,86]
<i>SOCS1</i> (TSG)	Up	Cell lines, tumors	[55,59,63,102]
<i>E-Cadherin</i> (TSG)	Up	Cirrhotic tissue, tumors	[63,86]
<i>GSTP1</i> (TSG)	Up	Serum, tumors	[59,63,86]
Histone mark			
H3K9 dimethylation	Up	Cell lines	[69]
H3K9 trimethylation	Up	Rodent models	[68]
H3K27 trimethylation	Up	Cell lines	[69]
H4K20 trimethylation	Down	Rodent models	[68]
H3 global acetylation	Down and up	Cell lines	[70]
H4 global acetylation	Down and up	Cell lines	[70]
microRNA			
miR-1	Down	Cell lines, tumors	[73,77]
miR-18	Up	Tumors	[72,75,138]
miR-21	Up	Cell lines, tumors	[72-74,138]
miR-122	Down	Tumors	[73,77,81,138]
miR-199	Down	Tumors	[73-76,138]
miR-221	Up	Cell lines, tumors	[73,74,77,79,81,138]
miR-222	Up	Cell lines, tumors	[77,79]
miR-224	Up	Tumors	[75,78,138]

TSG, tumor suppressor gene.

### MicroRNAs and hepatocellular carcinoma

MicroRNAs are a class of small non-coding RNAs (22 to 25 nucleotides) that repress gene expression by inhibiting the translation of messenger RNAs. MicroRNAs are considered to participate actively in HCC development, and this is supported by several studies in which significant changes in microRNA expression have been observed by comparing HCC tumors with non-cancerous tissues. For example, upregulation of microRNA-18 (miR-18), miR-21, miR-221, miR-222 and miR-224, and downregulation of miR-122, miR-125, miR-130a, miR-150, miR-199 and miR-200 and the let-7 family have been reported in HCC [72-78]. To understand the consequences of microRNA deregulation in hepatocarcinogenesis, several studies aimed to identify and validate target genes of these microRNAs. So far, microRNA alterations identified in liver tumors have mainly been associated with genes involved in cell cycle regulation and cell proliferation. For example, cyclin-dependent inhibitors p27 and p57 and B-cell lymphoma 2 (Bcl-2)-modifying factor are targets of miR-221/222, which are upregulated in HCC [79-81].

miR-1-1 is downregulated in HCC tumors compared with non-cancerous adjacent tissues, and its ectopic expression in HCC cell lines induced cell cycle inhibition and cell death [82]. Finally, miR-122 downregulation in HCC increases production of cyclin G1 [83], and overexpression of miR-22 in HCC cell lines results in increased cell proliferation and higher *de novo* tumor development in immune-compromised athymic or nude mice [84]. Taken together, these results are consistent with the critical role of microRNAs in the regulation of cell proliferation and apoptosis in hepatocytes, and highlight the importance of microRNA alterations in cellular transformation during hepatocarcinogenesis.

### Epigenetic alterations and HCC etiology

Although the evidence for deregulation of epigenetic mechanisms during hepatocarcinogenesis has steadily accumulated over the past decade, their origins remain unclear. As many of these changes have been observed in early stages of carcinogenesis and even in precancerous stages [63,85], it has been proposed that some of these

alterations may be directly induced by exposure to specific risk factors, including HBV and HCV infection, alcohol intake and aflatoxin B1, and their presence may drive the process of hepatocarcinogenesis. For example, methylation in *p16INK4A*, *GSPT1* and *RASSF1A* genes in HCC tumors has been significantly correlated with viral infection [86]. Comparison of methylation profiles between tumor samples associated with HBV infection, HCV infection and alcohol consumption has revealed a specific set of hypermethylated CpG islands for each group [87]. Our recent study also showed a significant association between the methylation pattern in HCC tumors and major risk factors, including HBV infection and alcohol intake [66]. Further studies are needed to test whether major risk factors induce a distinct set of early epigenetic events and whether these changes promote HCC development.

In addition to the well-established risk factors, nutrition deficiency could provide a favorable condition for HCC development. In a rat model of methyl-deficient or lipotrope-deficient diet-induced HCC, global hypomethylation has been found to be associated with tumor development [68]. Similarly, mouse models for alcoholic liver disease and non-alcoholic fat liver disease have provided evidence for microRNA deregulation in the diseased liver [88]. Furthermore, chronic alcohol consumption is also known to cause epigenetic alterations [89]: liver tissue from rats fed with alcohol presented a 40% loss of methylation, and chronic alcohol consumption caused global hypomethylation [90]. Finally, alcohol intake can also influence histone modifications, notably an increase in histone H3 acetylation [91]. In conclusion, viral infection or alcohol consumption seem to induce some specific epigenetic alterations, but the picture is far from being complete. Thus, the exact mechanism by which known risk factors trigger epigenetic changes and the precise gene targets remain to be elucidated.

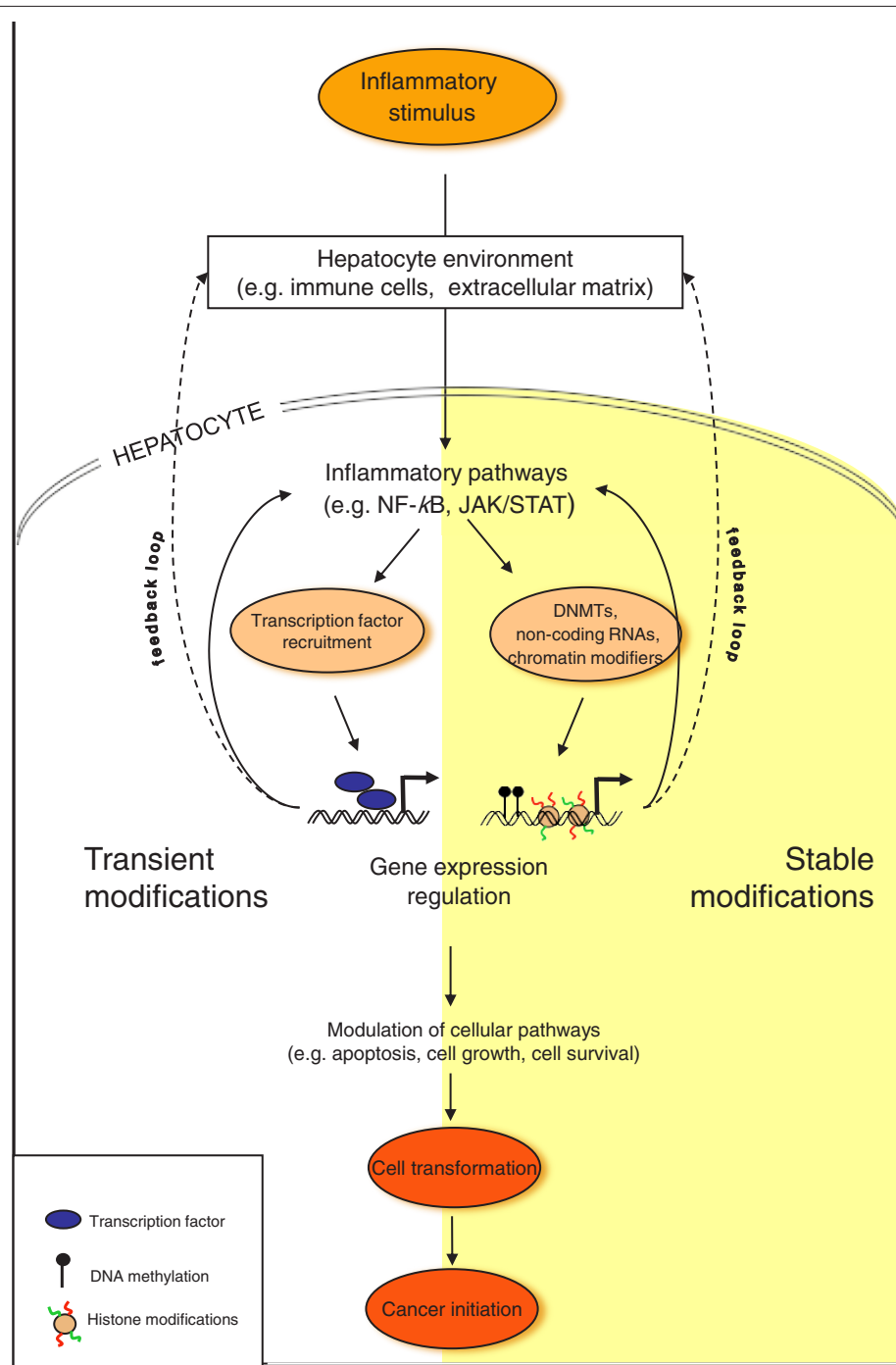
### **Cross-talk between epigenetic mechanisms and inflammatory pathways**

The link between inflammatory pathways and epigenetic mechanisms has been revealed by many recent studies using different model systems. There is growing evidence for a direct mechanistic relationship between the changes induced by inflammation and epigenetic deregulation during tumor development and progression [92,93]. We propose a working hypothesis that epigenome deregulation is an underlying mechanism by which inflammation may promote tumor development. In this scenario, different components of the inflammatory response may induce changes in epigenetic machinery or a so-called epigenetic switch that resets the long-term cellular memory system, a system that normally ensures the stable maintenance of transcriptional patterns and cell phenotype.

### **Cross-talk between epigenetic mechanisms and inflammatory pathways in HCC**

As discussed earlier, inflammatory processes and epigenetic deregulation are early events in hepatocarcinogenesis. However, it remains unclear whether inflammation and a deregulated epigenome act concomitantly to initiate HCC or if there is a hierarchy and interdependence between them during cancer development and progression. Although cancer is traditionally considered a genetic disease caused by the accumulation of mutations, recent evidence suggests that epigenetic changes may play an important role in cancer development and could also act as precursor events that precede and promote genetic changes [94,95]. Here, we put forward the hypothesis that epigenetic deregulation may be an underlying mechanism by which inflammation promotes HCC development, as shown in Figure 1. In this scenario, different components of the inflammatory response may directly or indirectly induce changes in epigenetic machineries, including those involved in setting and propagating normal patterns of DNA methylation, histone modifications and non-coding RNAs in hepatocytes. Deregulated epigenetic states may contribute to a persistent inflammatory response through altered gene expression states and a positive feedback loop to exacerbate a chronic state of inflammation. In parallel, the deregulated epigenome maintains altered long-term memory systems that promote proliferation and oncogenic transformation (Figure 1). This interdependent and self-reinforcing cross-talk between inflammation and the epigenome maintains and amplifies inflammatory signals resulting in a series of events culminating in the development of liver cancer.

Several recent mechanistic and functional studies have provided support for our model by demonstrating interconnections between inflammatory pathways and epigenetic modifications. For example, chronic inflammation increases the level of ROS in the cytoplasm and high levels of ROS have been reported to induce the expression of *SNAIL* (the gene encoding a master regulator of the process of epithelial-mesenchymal transition), which can in turn recruit DNMTs (DNA-methyltransferases) and HDACs (histone deacetylases) to silence several specific genes [63,96]. *In vivo* alcohol intake or *in vitro* lipopolysaccharide treatment (an inflammatory stimulus) can induce H3K9/S10 phosphorylation at the promoter of cytokine genes [97,98] and these specific histone marks appear to be required for NF- $\kappa$ B recruitment to the gene promoter [99]. Furthermore, it was shown that NF- $\kappa$ B interacts with HDAC-1 [100] and that the capacity of HDAC-1 to inactivate specific genes requires the presence of p50, an NF- $\kappa$ B subunit [101]. In contrast, epigenetic mechanisms can also interfere with inflammation pathways, notably in the activation of the JAK/STAT3



**Figure 1. A hypothetical model depicting cross-talk between activation of inflammatory pathways and epigenome deregulation during liver tumor development.** Different components of the inflammatory response (including transient and stable modifications such as activation of inflammatory pathways nuclear factor (NF)- $\kappa$ B and JAK/STAT) may induce changes in epigenetic machineries (including DNA methylation, histone modifications and non-coding RNAs), resulting in an 'epigenetic switch' that resets the long-term memory system in hepatocytes. The epigenetic switch in turn may contribute to a persistent inflammatory response through altered gene expression states and a positive feedback loop to exacerbate a chronic state of inflammation. In addition, the deregulated epigenome may maintain an altered transcriptional program that promotes proliferation and oncogenic transformation. This interdependent and self-reinforcing cross-talk between inflammation and the epigenome maintains and amplifies inflammatory signals, resulting in a series of events culminating in the development of liver cancer. The epigenetic switch may also be activated in hepatic or liver progenitor cells whose proliferation is stimulated during liver regeneration and repair. Therefore, an inflammatory microenvironment and an epigenetic switch in response to different environmental factors can directly promote activation of liver progenitor cells and their oncogenic transformation. DNMT, DNA methyl transferase.

pathway. Analysis of HCC tumors revealed that aberrant silencing of JAK/STAT inhibitor genes *SOCS-1* and *SOCS-3* by DNA methylation results in constitutive activation of the pathway [55,102]. In addition, the binding of activated STAT3 to its target gene promoter seems to be dependent on histone acetylation status [71]. All these examples support the hypothesis that inflammation and epigenetic mechanisms are not separate but inter-dependent events whose cross-talk may deregulate a wide range of processes resulting in the development of HCC.

#### **Cross-talk between epigenetic mechanisms and inflammatory pathways in liver stem/progenitor cells and liver cancer stem cells**

Research on the process of liver cancer initiation has also focused on cells that might act as the precursors of liver cancer. The observation that HCC cells present specific markers that are common to stem cells and that progression of liver cancer is associated with dedifferentiation (a process by which a specialized, differentiated cell regresses to a simpler, more embryonic, unspecialized form) led to the 'maturation arrest hypothesis', which predicts that liver cancer may arise from stem cells that failed to complete their differentiation [103-107]. Hepatic or liver progenitor cells (LPCs) are adult stem cells that can differentiate into either hepatocytes or cholangiocytes [108,109]. Stem cells are activated when the replication of mature hepatocytes is blocked, in order to take over liver regeneration and repair [110-112]. Several studies have provided evidence to support the hypothesis of an LPC origin for liver cancer [109,113-115]. As exposure to different environmental factors can activate inflammation in liver cells, one current model proposes that the inflammatory microenvironment directly promotes LPC activation and transformation. More specifically, IL6, TNF $\alpha$ , IFN $\gamma$  and TWEAK (TNF-like weak inducer of apoptosis), a member of the TNF family, increased the numbers of rodent LPCs *in vitro* and *in vivo* [116-118]. Moreover, increasing proliferation of LPC by cytokines is not just a side-effect of inflammation-induced cell proliferation, since the proliferative effects of IFN $\gamma$  and TWEAK on LPCs have been shown to be specific to LPCs (when compared with hepatocytes).

Some cytokines and inflammatory pathways have even presented negative effects on LPC proliferation: for example, both IFN $\alpha$  and TGF $\beta$  reduce or block the proliferation of LPCs [119-121]. Deregulation of LPCs during carcinogenesis is likely to be associated with profound and heritable changes in cell fate programming. For these reasons, it has also been proposed that liver cancer may be initiated through the sustained epigenetic reprogramming of LPCs. The stemness/differentiation balance can be regulated by DNA methylation and

bivalent marks (a combination of permissive and repressive histone marks) [95], and proteins responsible for the deposition of these marks are DNMTs and PcG proteins (Polycomb-group proteins), respectively. Therefore, inflammation may contribute to the transformation of LPCs by triggering epigenetic modifications. For example, cytokines such as IL6 or TGF $\beta$  have been shown to influence expression of DNMTs [122-124]. In addition, production of PcG proteins appears to be sensitive to cytokines: in the majority of human HCC samples, activation of JNK1 (a kinase that can be activated by cytokines) correlated with the increase of EZH2 PcG proteins [125]. *In vitro* studies of muscle stem cells confirmed the link between cytokines and PcG production by demonstrating that TNF $\alpha$  promotes the formation of Polycomb repressive complex [126]. In this manner, inflammation can directly affect LPC activation and differentiation by modifying the epigenetic memory system of these cells.

Finally, inflammation and epigenetic interactions can contribute to tumor development and progression by maintaining and expanding 'cancer stem cells' (CSCs). CSCs are defined as a discrete tumor population characterized by two defining properties: self-renewal and the capacity to reconstitute tumor heterogeneity [127]. Recent studies suggest a possible role for inflammation in the activation and maintenance of liver cancer stem cells. In a large series of HCC samples, the hepatic stem-cell-like subtype presented upregulation of the TGF $\beta$  pathway [128]. In healthy liver, LPCs have been found to produce OCT4, NANOG, STAT3 (the well-known core stemness genes) and TBRII (TGF  $\beta$ -receptor type II), although subsequent analysis of cells producing STAT3/OCT4 markers of stemness failed to identify TBRII production [115]. This observation raises the possibility that disruption of the TGF $\beta$  pathway may promote the emergence of CSCs and sustain HCC development and progression [129,130]. One explanation could be that TGF $\beta$  disruption impairs differentiation of LPCs after their activation, thus promoting their transformation into liver cancer stem cells [131]. This hypothesis is further supported by a recent observation that TGF $\beta$  downregulation in liver cancer stem cells enhanced the mitogen-activated protein kinase (MAPK) pathway and thereby conferred resistance to apoptosis [132]. Similarly, other inflammatory pathways, such as the IL6/JAK/STAT pathway, have been demonstrated to be capable of sustaining cancer stem cells. In a mouse model exhibiting disrupted TGF $\beta$  signaling, spontaneous liver cancer development and activation of IL6 signaling were observed [115], whereas in cell lines, the modulation of JAK/STAT was able to promote differentiation and elimination of CSCs [133]. Therefore, CSC activation is flexible and reversible, supporting the possibility that inflammatory



pathways may impact on CSCs through epigenetic mechanisms. A report by You *et al.* [124] supports this hypothesis in the context of liver cancer by showing that TGF $\beta$  influences the expression of *CD133* (a liver cancer stem cell marker) via aberrant DNA methylation at the *CD133* promoter. In addition, it was shown that in breast cancer, microRNAs could directly influence CSC proliferation through modulation of the IL6/JAK/STAT pathway [134]. Taken together, accumulating evidence argues that inflammation may initiate the appearance and/or maintenance of cancer cells with stem cell features in liver and that epigenetic deregulation may be a key underlying mechanism.

Based on the evidence presented, we propose a model (Figure 1) in which various components of the inflammatory response (including activation of inflammatory pathways) induce changes in the epigenome (likely through deregulation of epigenetic machineries such as those mediating DNA methylation, histone modifications and non-coding RNAs). These changes represent an initiating event activating the 'epigenetic switch' that resets the long-term memory system in the target cells (hepatocytes). The epigenetic switch may be defined as stable and mitotically heritable changes in the epigenome that underlie transition in cell phenotype, and are maintained after the initial triggering events have ceased. This epigenetic switch may promote a persistent inflammatory state through both gene expression reprogramming and a feedback loop that amplifies inflammatory signals contributing to chronic inflammation. The deregulated epigenetic states may also maintain an altered transcriptional programme that promotes proliferation and oncogenic transformation. Thus, the self-reinforcing cross-talk between inflammation and the epigenome maintains and amplifies inflammatory signals, resulting in a series of events culminating in the development of HCC.

### Conclusions and further perspectives

Abnormal secretions of cytokines are often observed in the presence of liver diseases and liver cancer. As a consequence, inflammatory pathways governing cell growth, cell cycle and cell survival are deregulated in hepatocytes. These pathways may enhance cytokine expression through positive feedback, thus creating a vicious circle that culminates in liver cancer. Evidence is accumulating to suggest that inflammation contributes to hepatocarcinogenesis through deregulation of the epigenome. Many studies have found a wide range of epigenetic alterations in HCC, consistent with the notion that aberrant DNA methylation, histone acetylation and expression of non-coding RNAs may be responsible for deregulated expression of numerous genes in transformed hepatocytes. Several lines of evidence suggest that inflammation and epigenetic deregulation are not

independent events during hepatocarcinogenesis, but that they may cross-talk and cooperate directly or indirectly to induce cell transformation. In particular, studies on liver stem cells and their potential role in liver cancer strongly suggest that a cross-talk between inflammation and epigenetic mechanisms could be particularly relevant to the initiation of cancer stem cells and early stages of HCC development.

Epigenetic deregulation may provide the missing mechanistic link between inflammation and HCC development. Because the major treatment currently available for HBV and HCV infection is IFN $\alpha$  administration, which has relatively low efficacy [15,135,136], a link between inflammation and epigenetic mechanisms suggests potential new targets for therapeutic intervention. Intrinsic reversibility of epigenetic changes and the recent development of drugs targeting epigenetic deregulation in cancer cells may provide an opportunity for targeting inflammation-epigenome cross-talk in liver cancer. A combination of classical antiviral agents (for example, INF administration) and epigenetic drugs (such as DNMT inhibitors or HDAC inhibitors) may prove particularly efficient for counteracting the synergy between cytokines and epigenome changes.

Despite important progress made in the field, several important questions remain to be addressed before we fully understand the functional impact of the interaction between inflammation and epigenome deregulation in liver tissue and define the precise underlying mechanisms. For example, to what extent is epigenetic deregulation triggered by chronic inflammation? Although activation of inflammatory pathways and disrupted epigenetic states commonly co-exist in liver cancer, hierarchies and the precise order of events that establish and maintain their cross-talk are far from being elucidated. Can inflammatory cytokines trigger epigenetic changes directly or indirectly? Does epigenetic deregulation contribute to chronic inflammation in the infected liver, and if so do positive feedback loops that amplify inflammatory responses exist? Another important question regards the origin of liver cancer stem cells and the role of inflammation and epigenetic mechanisms in their initiation and maintenance. The importance of epigenome reconfiguration in the maintenance of the defining features of stem/progenitor cells provides a mechanistic explanation but also suggests potential targets for intervention. More comprehensive characterization of the epigenome of HCC tumors and liver cancer stem cells may help in answering some of these key questions.

### Abbreviations

CSC, cancer stem cell; DNMT, DNA methyl transferase; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; HDAC, histone deacetylase; IFN, interferon; IL, interleukin; JAK, janus kinase; LPC, liver progenitor cell; NF, nuclear factor; PcG, Polycomb group; ROS, reactive

oxygen species; STAT, signal transducers and activators of transcription; TGF, transforming growth factor; TNF, tumor necrosis factor; TSG, tumor suppressor gene.

#### Competing interests

The authors declare that they have no competing interests.

#### Acknowledgements

We apologize to authors whose relevant publications were not cited due to space limitations. The work in the Epigenetics Group at the International Agency for Research on Cancer (Lyon, France) is supported by grants from l'Agence Nationale de Recherche Contre le Sida et Hépatites Virales (ANRS, France), l'Association pour la Recherche sur le Cancer (ARC), France, and la Ligue Nationale (Française) Contre le Cancer, France (to ZH). The founders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Published: 31 January 2012

#### References

- Llovet JM, Burroughs A, Bruix J: **Hepatocellular carcinoma.** *Lancet* 2003, **362**:1907-1917.
- Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM: **Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008.** *Int J Cancer* 2010, **127**:2893-2917.
- Lata J: **Chronic liver diseases as liver tumor precursors.** *Dig Dis* 2010, **28**:596-599.
- Brait M, Sidransky D: **Cancer epigenetics: above and beyond.** *Toxicol Mech Methods* 2011, **21**:275-288.
- Jones PA, Baylin SB: **The fundamental role of epigenetic events in cancer.** *Nat Rev Genet* 2002, **3**:415-428.
- Ballestar E, Esteller M: **Epigenetic gene regulation in cancer.** *Adv Genet* 2008, **61**:247-267.
- Rodriguez-Paredes M, Esteller M: **Cancer epigenetics reaches mainstream oncology.** *Nat Med* 2011, **17**:330-339.
- Pogribny IP: **Epigenetic events in tumorigenesis: putting the pieces together.** *Exp Oncol* 2010, **32**:132-136.
- Nephew KP, Huang TH: **Epigenetic gene silencing in cancer initiation and progression.** *Cancer Lett* 2003, **190**:125-133.
- Lima SC, Hernandez-Vargas H, Herceg Z: **Epigenetic signatures in cancer: Implications for the control of cancer in the clinic.** *Curr Opin Mol Ther* 2010, **12**:316-324.
- Herceg Z, Paliwal A: **Epigenetic mechanisms in hepatocellular carcinoma: how environmental factors influence the epigenome.** *Mutat Res* 2011, **727**:55-61.
- Luedde T, Trautwein C: **Intracellular survival pathways in the liver.** *Liver Int* 2006, **26**:1163-1174.
- Tacke F, Luedde T, Trautwein C: **Inflammatory pathways in liver homeostasis and liver injury.** *Clin Rev Allergy Immunol* 2009, **36**:4-12.
- Berasain C, Castillo J, Perugorria MJ, Latasa MU, Prieto J, Avila MA: **Inflammation and liver cancer: new molecular links.** *Ann N Y Acad Sci* 2009, **1155**:206-221.
- Gao B: **Cytokines, STATs and liver disease.** *Cell Mol Immunol* 2005, **2**:92-100.
- Bartsch H, Nair J: **Chronic inflammation and oxidative stress in the genesis and perpetuation of cancer: role of lipid peroxidation, DNA damage, and repair.** *Langenbecks Arch Surg* 2006, **391**:499-510.
- Grivennikov SI, Greten FR, Karin M: **Immunity, inflammation, and cancer.** *Cell* 2010, **140**:883-899.
- Bouchard MJ, Navas-Martin S: **Hepatitis B and C virus hepatocarcinogenesis: lessons learned and future challenges.** *Cancer Lett* 2011, **305**:123-143.
- Gurtsevitch VE: **Human oncogenic viruses: hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis.** *Biochemistry (Mosc)* 2008, **73**:504-513.
- But DY, Lai CL, Yuen MF: **Natural history of hepatitis-related hepatocellular carcinoma.** *World J Gastroenterol* 2008, **14**:1652-1656.
- Bertoletti A, Maini MK, Ferrari C: **The host-pathogen interaction during HBV infection: immunological controversies.** *Antiviral therapy* 2010, **15**(Suppl 3):15-24.
- Brechot C, Kremsdorf D, Soussan P, Pineau P, Dejean A, Paterlini-Brechot P, Tiollais P: **Hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC): molecular mechanisms and novel paradigms.** *Pathologie-biologie* 2010, **58**:278-287.
- Murakami Y, Saigo K, Takashima H, Minami M, Okanoue T, Brechot C, Paterlini-Brechot P: **Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas.** *Gut* 2005, **54**:1162-1168.
- Murakami Y, Minami M, Daimon Y, Okanoue T: **Hepatitis B virus DNA in liver, serum, and peripheral blood mononuclear cells after the clearance of serum hepatitis B virus surface antigen.** *J Med Virol* 2004, **72**:203-214.
- Bouchard MJ, Schneider RJ: **The enigmatic X gene of hepatitis B virus.** *J Virol* 2004, **78**:12725-12734.
- Casemann WH: **Transactivation of cellular gene expression by hepatitis B viral proteins: a possible molecular mechanism of hepatocarcinogenesis.** *J Hepatol* 1995, **22**:34-37.
- Ray RB, Meyer K, Ray R: **Suppression of apoptotic cell death by hepatitis C virus core protein.** *Virology* 1996, **226**:176-182.
- Ray RB, Steele R, Meyer K, Ray R: **Transcriptional repression of p53 promoter by hepatitis C virus core protein.** *J Biol Chem* 1997, **272**:10983-10986.
- Holt AP, Salmon M, Buckley CD, Adams DH: **Immune interactions in hepatic fibrosis.** *Clin Liver Dis* 2008, **12**:861-882, x.
- Viebahn CS, Yeoh GC: **What fires prometheus? The link between inflammation and regeneration following chronic liver injury.** *Int J Biochem Cell Biol* 2008, **40**:855-873.
- Budhu A, Wang XW: **The role of cytokines in hepatocellular carcinoma.** *J Leukoc Biol* 2006, **80**:1197-1213.
- Hsia CY, Huo TI, Chiang SY, Lu MF, Sun CL, Wu JC, Lee PC, Chi CW, Lui WY, Lee SD: **Evaluation of interleukin-6, interleukin-10 and human hepatocyte growth factor as tumor markers for hepatocellular carcinoma.** *Eur J Surg Oncol* 2007, **33**:208-212.
- Nakazaki H: **Preoperative and postoperative cytokines in patients with cancer.** *Cancer* 1992, **70**:709-713.
- Hu RH, Lee PH, Yu SC: **Secretion of acute-phase proteins before and after hepatocellular carcinoma resection.** *J Formos Med Assoc* 1999, **98**:85-91.
- Beckebaum S, Zhang X, Chen X, Yu Z, Frilling A, Dworacki G, Grosse-Wilde H, Broelsch CE, Gerken G, Ciccinnati VR: **Increased levels of interleukin-10 in serum from patients with hepatocellular carcinoma correlate with profound numerical deficiencies and immature phenotype of circulating dendritic cell subsets.** *Clin Cancer Res* 2004, **10**:7260-7269.
- Yuen MF, Norris S, Evans LW, Langley PG, Hughes RD: **Transforming growth factor-beta 1, activin and follistatin in patients with hepatocellular carcinoma and patients with alcoholic cirrhosis.** *Scand J Gastroenterol* 2002, **37**:233-238.
- Kitaoka S, Shiota G, Kawasaki H: **Serum levels of interleukin-10, interleukin-12 and soluble interleukin-2 receptor in chronic liver disease type C.** *Hepatogastroenterology* 2003, **50**:1569-1574.
- Huang YS, Hwang SJ, Chan CY, Wu JC, Chao Y, Chang FY, Lee SD: **Serum levels of cytokines in hepatitis C-related liver disease: a longitudinal study.** *Zhonghua Yi Xue Za Zhi (Taipei)* 1999, **62**:327-333.
- Chia CS, Ban K, Ithnin H, Singh H, Krishnan R, Mokhtar S, Malihan N, Seow HF: **Expression of interleukin-18, interferon-gamma and interleukin-10 in hepatocellular carcinoma.** *Immunol Lett* 2002, **84**:163-172.
- Tangkijvanich P, Thong-Ngam D, Mahachai V, Theamboonlers A, Poovorawan Y: **Role of serum interleukin-18 as a prognostic factor in patients with hepatocellular carcinoma.** *World J Gastroenterol* 2007, **13**:4345-4349.
- Zekri AR, Ashour MS, Hassan A, Alam El-Din HM, El-Shehaby AM, Abu-Shady MA: **Cytokine profile in Egyptian hepatitis C virus genotype-4 in relation to liver disease progression.** *World J Gastroenterol* 2005, **11**:6624-6630.
- Bortolami M, Venturi C, Giacomelli L, Scalera R, Bacchetti S, Marino F, Floreani A, Lise M, Naccarato R, Farinati F: **Cytokine, infiltrating macrophage and T cell-mediated response to development of primary and secondary human liver cancer.** *Dig Liver Dis* 2002, **34**:794-801.
- Song le H, Binh VQ, Duy DN, Kun JF, Bock TC, Kremsner PG, Luty AJ: **Serum cytokine profiles associated with clinical presentation in Vietnamese infected with hepatitis B virus.** *J Clin Virol* 2003, **28**:93-103.
- Karin M: **NF-kappaB as a critical link between inflammation and cancer.** *Cold Spring Harb Perspect Biol* 2009, **1**:a000141.
- He G, Karin M: **NF-kappaB and STAT3 - key players in liver inflammation and cancer.** *Cell Res* 2011, **21**:159-168.
- Assenat E, Gerbal-chaloin S, Maurel P, Vilarem MJ, Pascucci JM: **Is nuclear factor kappa-B the missing link between inflammation, cancer and alteration in hepatic drug metabolism in patients with cancer?** *Eur J Cancer* 2006, **42**:785-792.
- Maeda S, Omata M: **Inflammation and cancer: role of nuclear factor-**

- kappaB activation. *Cancer Sci* 2008, **99**:836-842.
48. Liu P, Kimmoun E, Legrand A, Sauvanet A, Degott C, Lardeux B, Bernuau D: **Activation of NF-kappa B, AP-1 and STAT transcription factors is a frequent and early event in human hepatocellular carcinomas.** *J Hepatol* 2002, **37**:63-71.
49. He G, Yu GY, Temkin V, Ogata H, Kuntzen C, Sakurai T, Sieghart W, Peck-Radosavljevic M, Leffert HL, Karin M: **Hepatocyte IKKbeta/NF-kappaB inhibits tumor promotion and progression by preventing oxidative stress-driven STAT3 activation.** *Cancer Cell* 2011, **17**:286-297.
50. Luedde T, Beraza N, Kotsikoris V, van Loo G, Nenci A, De Vos R, Roskams T, Trautwein C, Pasparakis M: **Deletion of NEMO/IKKgamma in liver parenchymal cells causes steatohepatitis and hepatocellular carcinoma.** *Cancer Cell* 2007, **11**:119-132.
51. Pikarsky E, Porat RM, Stein I, Abramovitch R, Amit S, Kasem S, Gutmovich-Pyest E, Urieli-Shoval S, Galun E, Ben-Neriah Y: **NF-kappaB functions as a tumour promoter in inflammation-associated cancer.** *Nature* 2004, **431**:461-466.
52. Haybaeck J, Zeller N, Wolf MJ, Weber A, Wagner U, Kurrer MO, Bremer J, Iezzi G, Graf R, Clavien PA, Thimme R, Blum H, Nedospasov SA, Zatloukal K, Ramzan M, Ciesek S, Pietschmann T, Marche PN, Karin M, Kopf M, Browning JL, Aguzzi A, Heikenwalder M: **A lymphotoxin-driven pathway to hepatocellular carcinoma.** *Cancer Cell* 2009, **16**:295-308.
53. Heinrich PC, Behrmann I, Haan S, Hermanns HM, Muller-Newen G, Schaper F: **Principles of interleukin (IL)-6-type cytokine signalling and its regulation.** *Biochem J* 2003, **374**:1-20.
54. Zhang CH, Xu GL, Jia WD, Li JS, Ma JL, Ren WH, Ge YS, Yu JH, Liu WB, Wang W: **Activation of STAT3 signal pathway correlates with twist and e-cadherin expression in hepatocellular carcinoma and their clinical significance.** *J Surg Res* 2010 [Epub ahead of print].
55. Calvisi DF, Ladu S, Gorden A, Farina M, Conner EA, Lee JS, Factor VM, Thorgeirsson SS: **Ubiquitous activation of Ras and Jak/Stat pathways in human HCC.** *Gastroenterology* 2006, **130**:1117-1128.
56. Park EJ, Lee JH, Yu GY, He G, Ali SR, Holzer RG, Osterreicher CH, Takahashi H, Karin M: **Dietary and genetic obesity promote liver inflammation and tumorigenesis by enhancing IL-6 and TNF expression.** *Cell* 2010, **140**:197-208.
57. Sinčić N, Herceg Z: **DNA methylation and cancer: ghosts and angels above the genes.** *Curr Opin Oncol* 2011, **23**:69-76.
58. Vaissiere T, Sawan C, Herceg Z: **Epigenetic interplay between histone modifications and DNA methylation in gene silencing.** *Mutat Res* 2008, **659**:40-48.
59. Huang J: **Current progress in epigenetic research for hepatocarcinogenesis.** *Sci China C Life Sci* 2009, **52**:31-42.
60. Lin CH, Hsieh SY, Sheen IS, Lee WC, Chen TC, Shyu WC, Liaw YF: **Genome-wide hypomethylation in hepatocellular carcinogenesis.** *Cancer Res* 2001, **61**:4238-4243.
61. Lee HS, Kim BH, Cho NY, Yoo EJ, Choi M, Shin SH, Jang JJ, Suh KS, Kim YS, Kang GH: **Prognostic implications of and relationship between CpG island hypermethylation and repetitive DNA hypomethylation in hepatocellular carcinoma.** *Clin Cancer Res* 2009, **15**:812-820.
62. Wong N, Lam WC, Lai PB, Pang E, Lau WY, Johnson PJ: **Hypomethylation of chromosome 1 heterochromatin DNA correlates with q-arm copy gain in human hepatocellular carcinoma.** *Am J Pathol* 2001, **159**:465-471.
63. Hamilton JP: **Epigenetic mechanisms involved in the pathogenesis of hepatobiliary malignancies.** *Epigenomics* 2010, **2**:233-243.
64. Zhao J, Wu G, Bu F, Lu B, Liang A, Cao L, Tong X, Lu X, Wu M, Guo Y: **Epigenetic silence of ankyrin-repeat-containing, SH3-domain-containing, and proline-rich-region-containing protein 1 (ASPP1) and ASPP2 genes promotes tumor growth in hepatitis B virus-positive hepatocellular carcinoma.** *Hepatology* 2010, **51**:142-153.
65. Zhang C, Li H, Wang Y, Liu W, Zhang Q, Zhang T, Zhang X, Han B, Zhou G: **Epigenetic inactivation of the tumor suppressor gene RIZ1 in hepatocellular carcinoma involves both DNA methylation and histone modifications.** *J Hepatol* 2010, **53**:889-895.
66. Lambert MP, Paliwal A, Vaissiere T, Chemin I, Zoulami F, Tommasino M, Hainaut P, Sylva B, Scoazec JY, Tost J, Herceg Z: **Aberrant DNA methylation distinguishes hepatocellular carcinoma associated with HBV and HCV infection and alcohol intake.** *J Hepatol* 2011, **54**:705-715.
67. Feinberg AP, Tycko B: **The history of cancer epigenetics.** *Nat Rev Cancer* 2004, **4**:143-153.
68. Pogribny IP, Ross SA, Tryndyak VP, Pogribna M, Poirier LA, Karpinetz TV: **Histone H3 lysine 9 and H4 lysine 20 trimethylation and the expression of Suv4-20h2 and Suv-39h1 histone methyltransferases in hepatocarcinogenesis induced by methyl deficiency in rats.** *Carcinogenesis* 2006, **27**:1180-1186.
69. Yao JY, Zhang L, Zhang X, He ZY, Ma Y, Hui LJ, Wang X, Hu YP: **H3K27 trimethylation is an early epigenetic event of p16INK4a silencing for regaining tumorigenesis in fusion reprogrammed hepatoma cells.** *J Biol Chem* 2010, **285**:18828-18837.
70. Park IY, Sohn BH, Yu E, Suh DJ, Chung YH, Lee JH, Surzycki SJ, Lee YI: **Aberrant epigenetic modifications in hepatocarcinogenesis induced by hepatitis B virus X protein.** *Gastroenterology* 2007, **132**:1476-1494.
71. Won C, Lee CS, Lee JK, Kim TJ, Lee KH, Yang YM, Kim YN, Ye SK, Chung MH: **CADPE suppresses cyclin D1 expression in hepatocellular carcinoma by blocking IL-6-induced STAT3 activation.** *Anticancer Res* 2010, **30**:481-488.
72. Ladeiro Y, Couchy G, Balabaud C, Bioulac-Sage P, Pelletier L, Rebouissou S, Zucman-Rossi J: **MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations.** *Hepatology* 2008, **47**:1955-1963.
73. Jiang J, Gusev Y, Aderca I, Mettler TA, Nagorney DM, Brackett DJ, Roberts LR, Schmittgen TD: **Association of MicroRNA expression in hepatocellular carcinomas with hepatitis infection, cirrhosis, and patient survival.** *Clin Cancer Res* 2008, **14**:419-427.
74. Gramantieri L, Fornari F, Callegari E, Sabbioni S, Lanza G, Croce CM, Bolondi L, Negrini M: **MicroRNA involvement in hepatocellular carcinoma.** *J Cell Mol Med* 2008, **12**:2189-2204.
75. Huang YS, Dai Y, Yu XF, Bao SY, Yin YB, Tang M, Hu CX: **Microarray analysis of microRNA expression in hepatocellular carcinoma and non-tumorous tissues without viral hepatitis.** *J Gastroenterol Hepatol* 2008, **23**:87-94.
76. Ji J, Wang XW: **New kids on the block: diagnostic and prognostic microRNAs in hepatocellular carcinoma.** *Cancer Biol Ther* 2009, **8**:1686-1693.
77. Wong QW, Lung RW, Law PT, Lai PB, Chan KY, To KF, Wong N: **MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1.** *Gastroenterology* 2008, **135**:257-269.
78. Wang Y, Lee AT, Ma JZ, Wang J, Ren J, Yang Y, Tantoso E, Li KB, Ooi LL, Tan P, Lee CG: **Profiling microRNA expression in hepatocellular carcinoma reveals microRNA-224 up-regulation and apoptosis inhibitor-5 as a microRNA-224-specific target.** *J Biol Chem* 2008, **283**:13205-13215.
79. le Sage C, Nagel R, Egan DA, Schrier M, Mesman E, Mangiola A, Anile C, Maira G, Mercatelli N, Ciafrè SA, Farace MG, Agami R: **Regulation of the p27(Kip1) tumor suppressor by miR-221 and miR-222 promotes cancer cell proliferation.** *EMBO J* 2007, **26**:3699-3708.
80. Fornari F, Gramantieri L, Ferracin M, Veronese A, Sabbioni S, Calin GA, Grazi GL, Giovannini C, Croce CM, Bolondi L, Negrini M: **Mir-221 controls CDKN1C/p57 and CDKN1B/p27 expression in human hepatocellular carcinoma.** *Oncogene* 2008, **27**:5651-5661.
81. Gramantieri L, Fornari F, Ferracin M, Veronese A, Sabbioni S, Calin GA, Grazi GL, Croce CM, Bolondi L, Negrini M: **MicroRNA-221 targets Bmf in hepatocellular carcinoma and correlates with tumor multifocality.** *Clin Cancer Res* 2009, **15**:5073-5081.
82. Datta J, Kutay H, Nasser MW, Nuovo GJ, Wang B, Majumder S, Liu CG, Volinia S, Croce CM, Schmittgen TD, Ghoshal K, Jacob ST: **Methylation mediated silencing of MicroRNA-1 gene and its role in hepatocellular carcinogenesis.** *Cancer Res* 2008, **68**:5049-5058.
83. Gramantieri L, Ferracin M, Fornari F, Veronese A, Sabbioni S, Liu CG, Calin GA, Giovannini C, Ferrazzi E, Grazi GL, Croce CM, Bolondi L, Negrini M: **Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma.** *Cancer Res* 2007, **67**:6092-6099.
84. Zhang J, Yang Y, Yang T, Liu Y, Li A, Fu S, Wu M, Pan Z, Zhou W: **microRNA-22, downregulated in hepatocellular carcinoma and correlated with prognosis, suppresses cell proliferation and tumorigenicity.** *Br J Cancer* 2010, **103**:1215-1220.
85. Feo F, Frau M, Tomasi ML, Brozzetti S, Pascale RM: **Genetic and epigenetic control of molecular alterations in hepatocellular carcinoma.** *Exp Biol Med (Maywood)* 2009, **234**:726-736.
86. Katoh H, Shibata T, Kokubu A, Ojima H, Fukayama M, Kanai Y, Hirohashi S: **Epigenetic instability and chromosomal instability in hepatocellular carcinoma.** *Am J Pathol* 2006, **168**:1375-1384.
87. Hernandez-Vargas H, Lambert MP, Le Calvez-Kelm F, Gouysse G, McKay-Chopin S, Tavtigian SV, Scoazec JY, Herceg Z: **Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors.** *PLoS One* 2010, **5**:e9749.
88. Dolganiuc A, Petrascu J, Kodys K, Catalano D, Mandrekas P, Velayudham A,



- Szabo G: **MicroRNA expression profile in Lieber-DeCarli diet-induced alcoholic and methionine choline deficient diet-induced nonalcoholic steatohepatitis models in mice.** *Alcohol Clin Exp Res* 2009, **33**:1704-1710.
89. Mandrekar P: **Epigenetic regulation in alcoholic liver disease.** *World J Gastroenterol* 2011, **17**:2456-2464.
90. Lu SC, Martinez-Chantar ML, Mato JM: **Methionine adenosyltransferase and S-adenosylmethionine in alcoholic liver disease.** *J Gastroenterol Hepatol* 2006, **21**(Suppl 3):S61-64.
91. Park PH, Lim RW, Shukla SD: **Involvement of histone acetyltransferase (HAT) in ethanol-induced acetylation of histone H3 in hepatocytes: potential mechanism for gene expression.** *Am J Physiol Gastrointest Liver Physiol* 2005, **289**:G1124-1136.
92. Niwa T, Ushijima T: **Induction of epigenetic alterations by chronic inflammation and its significance on carcinogenesis.** *Adv Genet* 2010, **71**:41-56.
93. Niwa T, Tsukamoto T, Toyoda T, Mori A, Tanaka H, Maekita T, Ichinose M, Tatematsu M, Ushijima T: **Inflammatory processes triggered by *Helicobacter pylori* infection cause aberrant DNA methylation in gastric epithelial cells.** *Cancer Res* 2010, **70**:1430-1440.
94. Feinberg AP, Ohlsson R, Henikoff S: **The epigenetic progenitor origin of human cancer.** *Nat Rev Genet* 2006, **7**:21-33.
95. Hernandez-Vargas H, Sincic N, Ouzounova M, Herceg Z: **Epigenetic signatures in stem cells and cancer stem cells.** *Epigenomics* 2009, **1**:20.
96. Lim SO, Gu JM, Kim MS, Kim HS, Park YN, Park CK, Cho JW, Park YM, Jung G: **Epigenetic changes induced by reactive oxygen species in hepatocellular carcinoma: methylation of the E-cadherin promoter.** *Gastroenterology* 2008, **135**:2128-2140, 2140 e2121-2128.
97. Yamamoto Y, Verma UN, Prajapati S, Kwak YT, Gaynor RB: **Histone H3 phosphorylation by IKK-alpha is critical for cytokine-induced gene expression.** *Nature* 2003, **423**:655-659.
98. Saccani S, Pantano S, Natoli G: **p38-Dependent marking of inflammatory genes for increased NF-kappa B recruitment.** *Nat Immunol* 2002, **3**:69-75.
99. Anest V, Hanson JL, Cogswell PC, Steinbrecher KA, Strahl BD, Baldwin AS: **A nucleosomal function for IkkappaB kinase-alpha in NF-kappaB-dependent gene expression.** *Nature* 2003, **423**:659-663.
100. Zhong H, May MJ, Jimi E, Ghosh S: **The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1.** *Mol Cell* 2002, **9**:625-636.
101. Elsharkawy AM, Oakley F, Lin F, Packham G, Mann DA, Mann J: **The NF-kappaB p50:p50:HDAC-1 repressor complex orchestrates transcriptional inhibition of multiple pro-inflammatory genes.** *J Hepatol* 2010, **53**:519-527.
102. Niwa Y, Kanda H, Shikauchi Y, Saiura A, Matsubara K, Kitagawa T, Yamamoto J, Kubo T, Yoshikawa H: **Methylation silencing of SOCS-3 promotes cell growth and migration by enhancing JAK/STAT and FAK signalings in human hepatocellular carcinoma.** *Oncogene* 2005, **24**:6406-6417.
103. Alison MR, Islam S, Lim S: **Stem cells in liver regeneration, fibrosis and cancer: the good, the bad and the ugly.** *J Pathol* 2009, **217**:282-298.
104. Wu PC, Fang JW, Lau VK, Lai CL, Lo CK, Lau JY: **Classification of hepatocellular carcinoma according to hepatocellular and biliary differentiation markers. Clinical and biological implications.** *Am J Pathol* 1996, **149**:1167-1175.
105. Yoon DS, Jeong J, Park YN, Kim KS, Kwon SW, Chi HS, Park C, Kim BR: **Expression of biliary antigen and its clinical significance in hepatocellular carcinoma.** *Yonsei Med J* 1999, **40**:472-477.
106. Uenishi T, Kubo S, Yamamoto T, Shuto T, Ogawa M, Tanaka H, Tanaka S, Kaneda K, Hirohashi K: **Cytokeratin 19 expression in hepatocellular carcinoma predicts early postoperative recurrence.** *Cancer Sci* 2003, **94**:851-857.
107. Yamashita T, Forgues M, Wang W, Kim JW, Ye Q, Jia H, Budhu A, Zanetti KA, Chen Y, Qin LX, Tang ZY, Wang XW: **EpCAM and alpha-fetoprotein expression defines novel prognostic subtypes of hepatocellular carcinoma.** *Cancer Res* 2008, **68**:1451-1461.
108. Alison MR, Lovell MJ: **Liver cancer: the role of stem cells.** *Cell Prolif* 2005, **38**:407-421.
109. Libbrecht L: **Hepatic progenitor cells in human liver tumor development.** *World J Gastroenterol* 2006, **12**:6261-6265.
110. Roskams T: **Progenitor cell involvement in cirrhotic human liver diseases: from controversy to consensus.** *J Hepatol* 2003, **39**:431-434.
111. Yang S, Koteish A, Lin H, Huang J, Roskams T, Dawson V, Diehl AM: **Oval cells compensate for damage and replicative senescence of mature hepatocytes in mice with fatty liver disease.** *Hepatology* 2004, **39**:403-411.
112. Roskams TA, Libbrecht L, Desmet VJ: **Progenitor cells in diseased human liver.** *Semin Liver Dis* 2003, **23**:385-396.
113. Knight B, Tirnitz-Parker JE, Olynyk JK: **C-kit inhibition by imatinib mesylate attenuates progenitor cell expansion and inhibits liver tumor formation in mice.** *Gastroenterology* 2008, **135**:969-979, 979 e961.
114. Zulehner G, Mikula M, Schneller D, van Zijl F, Huber H, Sieghart W, Grasl-Kraupp B, Waldhor T, Peck-Radosavljevic M, Beug H, Mikulits W: **Nuclear beta-catenin induces an early liver progenitor phenotype in hepatocellular carcinoma and promotes tumor recurrence.** *Am J Pathol* 2010, **176**:472-481.
115. Tang Y, Kitisin K, Jogunoori W, Li C, Deng CX, Mueller SC, Ransom HW, Rashid A, He AR, Mendelson JS, Jessup JM, Shetty K, Zasloff M, Mishra B, Reddy EP, Johnson L, Mishra L: **Progenitor/stem cells give rise to liver cancer due to aberrant TGF-beta and IL-6 signaling.** *Proc Natl Acad Sci U S A* 2008, **105**:2445-2450.
116. Knight B, Yeoh GC, Husk KL, Ly T, Abraham LJ, Yu C, Rhim JA, Fausto N: **Impaired preneoplastic changes and liver tumor formation in tumor necrosis factor receptor type 1 knockout mice.** *J Exp Med* 2000, **192**:1809-1818.
117. Brooling JT, Campbell JS, Mitchell C, Yeoh GC, Fausto N: **Differential regulation of rodent hepatocyte and oval cell proliferation by interferon gamma.** *Hepatology* 2005, **41**:906-915.
118. Yeoh GC, Ernst M, Rose-John S, Akhurst B, Payne C, Long S, Alexander W, Croker B, Grail D, Matthews VB: **Opposing roles of gp130-mediated STAT-3 and ERK-1/2 signaling in liver progenitor cell migration and proliferation.** *Hepatology* 2007, **45**:486-494.
119. Nguyen LN, Furuya MH, Wolfrum LA, Nguyen AP, Holdren MS, Campbell JS, Knight B, Yeoh GC, Fausto N, Parks WT: **Transforming growth factor-beta differentially regulates oval cell and hepatocyte proliferation.** *Hepatology* 2007, **45**:31-41.
120. Preisegger KH, Factor VM, Fuchs-bichler A, Stumptner C, Denk H, Thorgeirsson SS: **Atypical ductular proliferation and its inhibition by transforming growth factor beta1 in the 3,5-diethoxycarbonyl-1,4-dihydrocollidine mouse model for chronic alcoholic liver disease.** *Lab Invest* 1999, **79**:103-109.
121. Lim R, Knight B, Patel K, McHutchison JG, Yeoh GC, Olynyk JK: **Antiproliferative effects of interferon alpha on hepatic progenitor cells in vitro and in vivo.** *Hepatology* 2006, **43**:1074-1083.
122. Braconi C, Huang N, Patel T: **MicroRNA-dependent regulation of DNA methyltransferase-1 and tumor suppressor gene expression by interleukin-6 in human malignant cholangiocytes.** *Hepatology* 2010, **51**:881-890.
123. Meng F, Wehbe-Janek H, Henson R, Smith H, Patel T: **Epigenetic regulation of microRNA-370 by interleukin-6 in malignant human cholangiocytes.** *Oncogene* 2008, **27**:378-386.
124. You H, Ding W, Rountree CB: **Epigenetic regulation of cancer stem cell marker CD133 by transforming growth factor-beta.** *Hepatology* 2010, **51**:1635-1644.
125. Chang Q, Zhang Y, Beezhold KJ, Bhatia D, Zhao H, Chen J, Castranova V, Shi X, Chen F: **Sustained JNK1 activation is associated with altered histone H3 methylations in human liver cancer.** *J Hepatol* 2009, **50**:323-333.
126. Palacios D, Mozzetta C, Consalvi S, Caretti G, Saccone V, Proserpio V, Marquez VE, Valente S, Mai A, Forcales SV, Sartorelli V, Puri PL: **TNF/p38alpha/polycomb signaling to Pax7 locus in satellite cells links inflammation to the epigenetic control of muscle regeneration.** *Cell Stem Cell* 2010, **7**:455-469.
127. Visvader JE, Lindeman GJ: **Cancer stem cells in solid tumours: accumulating evidence and unresolved questions.** *Nat Rev Cancer* 2008, **8**:755-768.
128. Yamashita T, Ji J, Budhu A, Forgues M, Yang W, Wang HY, Jia H, Ye Q, Qin LX, Wauthier E, Reid LM, Minato H, Honda M, Kaneko S, Tang ZY, Wang XW: **EpCAM-positive hepatocellular carcinoma cells are tumor-initiating cells with stem/progenitor cell features.** *Gastroenterology* 2009, **136**:1012-1024.
129. Amin R, Mishra L: **Liver stem cells and tgf-Beta in hepatic carcinogenesis.** *Gastrointest Cancer Res* 2008, **2**:S27-30.
130. Kitisin K, Pishvaian MJ, Johnson LB, Mishra L: **Liver stem cells and molecular signaling pathways in hepatocellular carcinoma.** *Gastrointest Cancer Res* 2007, **1**:S13-21.
131. Mishra L, Banker T, Murray J, Byers S, Thenappan A, He AR, Shetty K, Johnson L, Reddy EP: **Liver stem cells and hepatocellular carcinoma.** *Hepatology* 2009, **49**:318-329.
132. Ding W, Mouzaki M, You H, Laird JC, Mato J, Lu SC, Rountree CB: **CD133+ liver cancer stem cells from methionine adenosyl transferase 1A-deficient mice demonstrate resistance to transforming growth factor (TGF)-beta-induced**

- apoptosis. *Hepatology* 2009, **49**:1277-1286.
133. Yamashita T, Honda M, Nio K, Nakamoto Y, Takamura H, Tani T, Zen Y, Kaneko S: **Oncostatin m renders epithelial cell adhesion molecule-positive liver cancer stem cells sensitive to 5-Fluorouracil by inducing hepatocytic differentiation.** *Cancer Res* 2010, **70**:4687-4697.
  134. Iliopoulos D, Hirsch HA, Struhl K: **An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation.** *Cell* 2009, **139**:693-706.
  135. Deny P, Zoulim F: **Hepatitis B virus: from diagnosis to treatment.** *Pathol Biol (Paris)* 2010, **58**:245-253.
  136. Muller C: **Chronic hepatitis B and C - current treatment and future therapeutic prospects.** *Wien Med Wochenschr* 2006, **156**:391-396.
  137. Huang WX, Huang P, Link H, Hillert J: **Cytokine analysis in multiple sclerosis by competitive RT - PCR: A decreased expression of IL-10 and an increased expression of TNF-alpha in chronic progression.** *Mult Scler* 1999, **5**:342-348.
  138. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, Shimotohno K: **Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues.** *Oncogene* 2006, **25**:2537-2545.

doi:10.1186/gm307

**Cite this article as:** Martin M, Herceg Z: From hepatitis to hepatocellular carcinoma: a proposed model for cross-talk between inflammation and epigenetic mechanisms. *Genome Medicine* 2012, **4**:8.

### **ANNEXE III: Research Paper Manuscript**

“Transdifferentiation of cancer cells to cancer stem-like cells by Transforming Growth Factor Beta (TGF- $\beta$ ) is associated with DNA methylome reconfiguration”

Current status : under revision in CANCER RESEARCH



# **Transdifferentiation of cancer cells to cancer stem-like cells by Transforming Growth Factor Beta (TGF- $\beta$ ) is associated with DNA methylome reconfiguration**

## **Authors**

Marion Martin<sup>1</sup>, Marie-Pierre Cros<sup>1</sup>, Geoffroy Durand<sup>2</sup>, Florence Le Calvez-Kelm<sup>2</sup>, Hector Hernandez-Vargas<sup>1\*</sup>, Zdenko Herceg<sup>1\*</sup>

\*equal senior author contribution

<sup>1</sup>. *Epigenetics Group. International Agency for Research on Cancer (IARC).*

<sup>2</sup>. *Genetic Cancer Susceptibility Group. International Agency for Research on Cancer (IARC).*

*150 rue Albert-Thomas, 69008 Lyon, France.*

## **Correspondence to:**

- Zdenko Herceg, Tel : +33 4 72 73 83 98. Fax: +33 4 72 73 83 29. Email: herceg@iarc.fr,  
and
- Hector Hernandez-Vargas, Tel: +33 4 72 73 83 48. Fax: +33 4 72 73 83 29. Email:  
vargash@iarc.fr

**Running title:** TGF- $\beta$  induces a cancer stem-like methylome

**Keywords:** HCC, cancer stem cells, CD133, DNA methylation, TGF- $\beta$  pathway

**Word count.** Abstract: 191; Text: 5867.

**Number of Figures:** 7.



## **Conflict of Interest Disclosure Statement**

The authors declare no competing financial interests.

## **Financial support**

This work was supported by la Ligue National (Française) Contre le Cancer, l'Association pour le Recherche Contre le Cancer (l'ARC), and the Agence Nationale de Recherches sur le SIDA et les Hépatites Virales (ANRS).

## Abstract

Distinct subpopulations of neoplastic cells within tumors, including hepatocellular carcinoma (HCC), display pronounced ability to initiate new tumors and induce metastasis. Recent evidence suggests that signals from transforming growth factor beta (TGF- $\beta$ ) may increase the survival of these so called cancer stem cells (CSC) leading to poor HCC prognosis. However, how TGF- $\beta$  establishes and modifies the key features of these cell subpopulations is not fully understood. In the present report we describe the unique DNA methylome of CD133-expressing putative liver CSCs. Next, we show that TGF- $\beta$  is able to induce CSCs in liver cancer cell lines in a way that is stable and persistent across cell division. This epigenetic process is associated with genome-wide changes in DNA methylation, affecting the DNA methylation machinery itself. The nature of these changes is non-random and is partially reflected at the transcriptional level. Our study reveals a self-perpetuating crosstalk between TGF- $\beta$  signaling and the DNA methylation machinery, which can be relevant in the establishment of cellular phenotypes. This is the first indication of the ability of TGF- $\beta$  to induce genome-wide changes in DNA methylation, resulting in a stable switch to a liver CSC epigenetic program.

## Introduction

Hepatocellular carcinoma (HCC) is the major form of primary liver cancer (1), and typically originates in a background of chronic inflammation caused by various factors, such as alcohol consumption, or viral infection (hepatitis B and hepatitis C) (2). Inflammation is an essential part of the wound-healing response to those risk factors. However, chronic inflammation favors the accumulation of mutations and epigenetic aberrations in hepatocytes, thereby promoting malignant transformation (3,4). This process is mediated by chemokines, cytokines, and growth factors secreted by the stromal components of the liver microenvironment (4). Among those secreted factors, the transforming growth factor beta (TGF- $\beta$ ) has been shown to have a key role that is cell-type dependent and variable during the hepatocarcinogenesis process (5). In established HCC, TGF- $\beta$  overexpression is associated with poor prognosis (6–8). However, characterization of the tumor cells targeted by TGF- $\beta$  in HCC is still lacking.

As has been shown for other human malignancies, a subpopulation of cancer cells in HCC is known to display a higher tumorigenic potential (9–11). These so called cancer stem cells (CSCs), are defined by their self-renewal and differentiation capacity, and have been isolated based on their expression of several cell markers (EpCAM, CD133, CD90, CD44, CD24, CD13, and OV6) (9). Of these, the surface marker CD133/Prominin1[PROM1] has been the most consistently reported. CD133 is a transmembrane protein whose function is only partially known (12)(13). Regardless of its function, CD133 may represent a marker of a distinct cell subpopulation with defined characteristics. The functional characterization of these cells will increase our understanding of the mechanisms involved in promoting and sustaining liver cancer progression.

Several recent reports suggest a link between TGF- $\beta$  signaling and liver CSCs. Firstly, signaling pathways identified in liver cancer, including TGF- $\beta$ , are active in isolated liver CSCs (14). Secondly, TGF- $\beta$ -induced EMT generates self-renewing stem cells, a process also implicated in a higher risk of

tumor metastasis, as invasiveness and self-renewal are shared features of stem cells, CSCs and metastatic cells (15,16). Finally, a recent study showed that TGF- $\beta$  is able to induce the expression of CD133 in liver cancer cell lines together with an increased tumor initiating ability in mice (17). Together, these studies point towards a specific role for TGF- $\beta$  in inducing a CSC program in HCC. DNA methylation, together with other epigenetic mechanisms, is able to stably modify the cell phenotype through cellular division (18). Because of the relative stability of DNA methylation marks, DNA methylation is a strong candidate mechanism to translate the presence of TGF- $\beta$  in the cellular microenvironment into persistent changes in phenotype. However, there is still limited evidence of a link between exposure to components of the tumor microenvironment and the induction of stable changes in DNA methylation in target cells.

In this study, we first defined the DNA methylome profile of CD133+ putative liver cancer stem cells. We then tested the hypothesis that DNA methylation is involved in the induction of liver CSCs by TGF- $\beta$ . In testing this hypothesis, we showed that TGF- $\beta$  function is intimately linked to the DNA methylation machinery in this context, and that this may represent a key process in the establishment of chronic exposure imprints in liver cancer cells.

## Results

### ***CD133- and CD133+ liver cancer cells differentially express DNA methylation genes***

CD133 is an established marker of CSCs in different types of human malignancies, including HCC (12). To test the notion that this marker distinguishes a cell subpopulation with a distinct DNA methylation program, we characterized two non-related liver cancer cell lines. In a first step, we estimated the frequency of CD133 expressing cells in Huh7 and HepG2 liver cancer cells using fluorescence-activated cell sorting (FACS) against all common CD133 isoforms (12). The expression of CD133 was

evident in both cell lines, with a mean of 5% (SD=2%) in HepG2, and 25% (SD=13%) in Huh7 (Figure 1a). Expression of the surface protein positively correlated with CD133 expression at the mRNA level (data not shown). This low to moderate percentage of cells expressing CD133 is consistent with a CSC marker, and contrasts with the extreme values of expression that we observed for other molecules such as CD90, CD44 or EpCAM (Figure S1).

Because the stemness transcription factors were shown to be differentially expressed in subpopulations of liver cancer stem cells (19), we next studied the mRNA expression of well-defined stemness transcription factors in liver cancer cell populations enriched for CD133+ cells. To this end, we enriched Huh7 and HepG2 cells for CD133+ cells using magnetic cell sorting (MACS). Efficiency of CD133 enrichment by MACS was approximately 3- to 4-fold for both cell lines (70-80% of CD133+ cells in Huh7, and 20-25% in HepG2, after two MACS columns). In spite of this variability between cell lines, we observed a consistent and significant overexpression of the three stemness genes *NANOG*, *POU5F1* (Oct4), and *SOX2* (Figure 1b). In addition, mRNA obtained from populations enriched in CD133+ cells at intermediate levels (average of 53% of CD133+ cells in Huh7, and 8% in HepG2, after one single MACS column) displayed the expected intermediate levels of mRNA expression in both cell lines, suggesting that the overexpression of stemness genes was specifically dependent on the CD133+ fraction of cells (Figure 1b).

To serve as a basis for exploring a potentially different methylation program in CD133+ liver cancer cells, we studied the expression of genes coding for relevant players of the DNA methylation machinery. This included genes involved in maintenance of DNA methylation (*DNMT1*), de novo DNA methylation (*DNMT3A* and *DNMT3B*) and DNA demethylation (*TET1* and *TET2*). Notably, *DNMT3A* was consistently and significantly overexpressed in both Huh7 and HepG2 cells progressively enriched for CD133 (Figure 1c). In addition, *DNMT3B* was overexpressed in HepG2 CD133-enriched cells, while *TET2* displayed opposite differential expression in CD133-enriched Huh7 and HepG2 cells (Figure 1c).

Together, these data suggests that in at least two independent liver cancer cell lines, CD133 marks a specific subpopulation characterized by a marked expression of genes related to stemness properties. Functionally, expression of this marker has been associated with an increased tumor-initiating ability and ability to grow in non-attachment conditions, a well known surrogate measure of CSC-like activity. We found that MACS-sorted CD133+ Huh7 cells were able to form spheres under non-attachment conditions, in contrast to their CD133- counterpart (Figure S1). This was not the case with HepG2 cells, where no sphere formation was observed, possibly due to the lower enrichment of CD133+ cells that was attained using MACS. In addition, the consistent overexpression of de novo DNA methylation genes (*DNMT3A* in both cell lines, and *DNMT3B* in HepG2) favors the idea of a unique DNA methylation program.

#### ***A differential DNA methylome defines CD133- and CD133+ liver cancer cells***

The above results support the hypothesis of a phenotypic and functional uniqueness of CD133+ cells. These cells also display a higher expression of de novo DNMTs, and this may be reflected in a differential configuration of their DNA methylome. To study this possibility, we performed a genome-wide DNA methylome analysis in FACS-sorted CD133- and CD133+ fractions from Huh7 and HepG2 cells (Figure 2a). DNA isolated from these fractions was interrogated with the Illumina Infinium 450k bead array, which covers different genomic features of interest in addition to all human bona fide CpG islands (20). We first performed unsupervised analyses and found that parental cell line was the main factor defining DNA methylation variation (Figure S2). Therefore, our main analysis compared CD133- to CD133+ fractions accounting for cell of origin (see Materials and Methods). The class comparison analysis resulted in 823 differentially methylated probes [corresponding to 472 annotated genes] at significant p value ( $p < 0.001$ ), although relatively high FDRs ( $FDR = 0.58$ ), probably due to sample variability and cell line differences. Therefore, for downstream data mining, we increased the stringency of the analyses by further filtering the significant list to keep only those CpG

sites where the average differential methylation was at least 5% between the two groups in both cell lines. The resulting 608 differentially methylated probes correspond to 394 RefSeq genes, and represent those CpG sites significantly hypo or hypermethylated in CD133+ cells in both cell lines, relative to their negative counterpart (Table S1). Most of these probes (n=510, 84%) were hypomethylated in CD133+ cells, while 98 (16%) were hypermethylated (Figure 2b). An important proportion of differentially methylated loci (45%) were not related to CpG islands (CGI) or their neighboring shelves and shores (“open sea” probes in Figure 2c). For those probes matching annotated genes, we found a significant overrepresentation of differentially methylated loci in the body of the genes (45%). This distribution relative to gene position and CpG island status was similar for hypomethylated sites, while hypermethylated sites were even more enriched in both, open sea (64%) and gene body (57%) probes (data not shown). Supporting the quality of the dataset was the finding of one CpG site within the CD133 (Prominin1 [*PROM1*]) locus among this list of differentially methylated sites. This CpG site was hypomethylated in CD133+ subpopulations from both cell lines, by 4.4% and 8% in Huh7 and HepG2 cells, respectively (Figure 2d).

After having identified differentially methylated CpGs and the genes associated with these sites, we next aimed to identify the pathways that are specifically altered in CD133+ cells. To this end, we performed pathway analysis considering methylome profiles of both cell lines together or independently. Notably, in both cases there was enrichment in pathways previously associated with cancer stem cell activity, such as Jak-STAT, Notch, Wnt and Akt (Table S2). Other pathways included actin cytoskeleton, focal adhesion, and cell adhesion. In addition, there was a significant overrepresentation of inflammatory pathways, such as NFkB, p38, TNF, and TGF- $\beta$  signaling pathways.

In summary, our data show that CD133+ liver cancer cells display a unique DNA methylome. In spite of the cell line specific profiles, we were able to produce a common CD133+ methylome signature, which includes the *PROM1* gene itself. In addition, the methylome of CD133+ cells is characterized by

a global reduction in DNA methylation, with an overrepresentation of non-CpG sites. For those differentially methylated sites related to annotated genes (and mainly found in the gene bodies), there was an association with CSC- and inflammation-related pathways. These findings suggest that DNA methylation makes an important contribution to defining the phenotype and functionality of this cell subpopulation.

### ***TGF- $\beta$ , but not IL-6, induces CD133 expression in a stable fashion***

It has been reported that TGF- $\beta$  exposure increases the percentage of CD133+ cells in the Huh7 cell line (17), although the underlying mechanism remains largely unknown. We thus aimed to investigate whether TGF- $\beta$  may induce CD133+ cells in liver cancer cell lines through changes in DNA methylation. Importantly, both Huh7 and HepG2 cells, express the receptor for TGF- $\beta$  (TGFBR1) at similar levels, and respond to TGF- $\beta$  by phosphorylating the receptor-dependent SMAD3 (Figure S3a and S3b). In addition to TGF- $\beta$ , we performed a set of parallel experiments with the pro-inflammatory cytokine interleukin 6 (IL-6), which has also been associated with HCC risk (21). To this end, we selected commonly used cytokine concentrations that induced morphological changes after 4 days of treatment in both cell lines (in the case of TGF- $\beta$ ), but did not have any effect on cell viability (Figure S3c and S3d). As expected, TGF- $\beta$  exposure during 4 days induced an almost three-fold and two-fold increase in the percentage of CD133+ cells in Huh7 and HepG2 cells, respectively (Figure 3a). Interestingly, IL-6 treatment also induced a significant increase in CD133 positivity in both cell lines, although the increase was comparatively mild (approximately 50% increase) (Figure 3a). Next, we analyzed the persistence of the effect in CD133 expression induced by both cytokines. To this end, we treated both cell lines as in the previous experiment. After 4 days, cell culture medium was replaced by standard medium, and cells were left in culture for additional 4 days. Cells were collected and screened for CD133 expression using FACS. Notably, only cells treated with TGF- $\beta$  showed a persistent increase in the percentage of CD133+ cells, of similar magnitude to the increase



observed at day 4 (Figure 3a). Importantly, only TGF- $\beta$  exposure was able to induce a significant increase in the expression of CD133 at the transcriptional level in both cell lines (8 and 6 fold increase for Huh7 and HepG2, respectively) (Figure S3).

TGF- $\beta$  is a member of a large family of pleiotropic cytokines that signal through a receptor complex comprising a diversity of type I and a type II serine/threonine kinases. The recombinant TGF- $\beta$ 1 used in our assays is expected to bind the activin receptor-like kinase (ALK)5 (the TGF-beta type I receptor) (22). To rule out unspecific effects of this treatment, we used the small molecule inhibitor SB-431542, which targets ALK5 and ALK5-related type I receptors, with no effect on other family members that, for example, recognize bone morphogenetic proteins (BMPs) (23). By using this specific inhibitor of TGF- $\beta$  pathway, we were able to abrogate the effect of TGF- $\beta$  in inducing CD133 expression (as well as the morphological changes) in both cell lines (Figure 3b and S3e). Therefore, the ability to induce CD133<sup>+</sup> cells is specific and fully dependent on TGF- $\beta$  type I receptor signaling in both, Huh7 and HepG2 cells (Figure 3b).

Together, these findings suggest that TGF- $\beta$  is able to specifically and stably induce CD133 expression (in contrast to the milder and transient effect of IL-6), an observation consistent with epigenetically-induced phenotype persistence.

#### ***De novo induction of CD133 by TGF- $\beta$ is associated to an increased expression of DNMT3 genes***

The increase in CD133 positivity induced by TGF- $\beta$  can be due to a switch in the expression of CD133, or an increased rate of growth specifically in the smaller CD133<sup>+</sup> fraction of cells. To distinguish between these two possibilities, we repeated the previous experiment in cells negative for CD133 expression, selected by depletion of CD133<sup>+</sup> cells using MACS (see Materials and Methods). In both cell lines, TGF- $\beta$  was able to significantly induce a population of CD133<sup>+</sup> cells, evident after 4 days of

treatment (Figure 4a). Also in this case, we replaced the medium after 4 days, and let the cells grow in the absence of cytokines for additional 4 days. After these additional 4 days, the increase in CD133 positive fraction was even higher, relative to the one observed at day 4, for both cell lines (Figure 4a). Importantly, although there was a spontaneous induction of a CD133+ fraction in Huh7 cells (from 0 to 20% after 4 days), this percentage did not significantly change at day 8, and is similar to what is found in untreated Huh7 cells in basal conditions. This indicates that there is a balance between the CD133 negative and positive fractions in this cell line. In contrast, the surface expression of CD133 remained close to zero in HepG2 control cells. This finding indicates that TGF- $\beta$  is able to induce the expression of CD133 surface protein, and not an increased proliferation of CD133+ cells. This is also supported by the expected lower rate of proliferation of cells treated with TGF- $\beta$  (Figure S4). Similar to our previous experiment, under these conditions IL6 only showed a transient effect. Interestingly, this effect of IL6 was only seen in Huh7 cells after 4 days, potentially linked to the spontaneous induction of CD133+ cells in this cell line.

After having shown that TGF- $\beta$  is able to induce a de novo fraction of CD133+ cells, we asked whether this effect correlated with a differential expression of DNA methylation players, as we have shown that CD133+ cells overexpress *DNMT3* genes in basal culture conditions (Figure 1c). All DNMTs and *TET2* displayed a significant increase in mRNA expression in at least one of the two cell lines, while *TET1* was underexpressed after 4 days of release from TGF- $\beta$  exposure (Figure 4b). As shown for the basal CD133-expressing cells (i.e. those isolated from untreated HCC cell lines), the most consistent finding was the statistically significant overexpression of *DNMT3A* in both cell lines after TGF- $\beta$  treatment. Of note, in none of the conditions of study IL-6 exposure was able to induce statistically significant changes at the mRNA expression level of genes related to DNA methylation/demethylation (Figure 4b).

Combined, these data shows the ability of TGF- $\beta$  (in contrast to IL-6) to induce a stable de novo fraction of CD133-expressing cells in two independent liver cancer cell lines. This induction correlates

with a functional characteristic of basal CD133+ cells, which is the increased ability to grow under non-attachment cell culture conditions (Figure 4c). Moreover, the differential expression of de novo DNMTs induced by TGF- $\beta$  indicates that the expression of CD133 may be a marker of a more general expression program that defines this “transdifferentiated” cell subpopulation.

### ***Transdifferentiation to CD133+ cells correlates with a methylome reconfiguration***

Having shown that CD133+ cells display a unique DNA methylome, and that TGF- $\beta$  is able to induce a de novo CD133+ fraction of cells, we decided to study the DNA methylome induced by TGF- $\beta$  exposure. To this end, we used the same Infinium 450k platform to interrogate DNA methylation changes induced by 4 days of TGF- $\beta$  exposure in both, Huh7 and HepG2 cells (Figure 5a). In addition, to define the epigenetic persistence of TGF- $\beta$  effects, we included the DNA from cells released 4 days into normal cell culture medium after the TGF- $\beta$  treatment. As described above for the DNA methylation profile of CD133-expressing cells, the methylome of Huh7 and HepG2 cells are clearly distinguishable, independently of the experimental conditions (Figure S5). However, in addition to cell type-specific changes we were able to observe genome-wide changes induced by TGF- $\beta$  in a cell type-independent fashion. To define a TGF- $\beta$ -induced DNA methylation signature, we focused on those loci that were significantly hypo or hypermethylated in both cell lines. In addition, we were interested in those changes that were persistent through cell division and stable in the absence of TGF- $\beta$ . Therefore, we selected significant loci (FDR<0.05) that were differentially methylated at both, 4 days of treatment and 4 additional days after release. Finally, we selected those CpG sites that reached an average difference of at least 5% at day 8 (4 days post-release) (Figure 5b). 568 probes fulfill all criteria, with 100 hypomethylated after TGF- $\beta$  exposure (18%) and a great majority hypermethylated (n=468, 82%) (Table S3 and S4). In addition, differentially methylated sites were classified into different clusters according to their pattern of expression (Figure 5b). Four out of five clusters represented probes consistently hypermethylated after TGF- $\beta$  treatment in both cell lines,

and corresponding to most of the differentially methylated probes (Figure 5c). These loci included both de novo DNMTs, *DNMT3A* (one CpG site) and *DNMT3B* (two CpG sites). Differentially methylated sites also included TGF- $\beta$ -related and chromatin-related genes, such as *CDKN1B*, *COL1A1*, *TRRAP*, *FGFR3*, *HDAC7*, *ARID3A*, and *KDM6B*. One cluster corresponded to probes significantly hypomethylated after TGF- $\beta$  exposure, including loci such as *CALD1*, *BMP1*, *IL18*, and *IRAK2*. In a similar way to the CD133 methylome, we found enrichment of differentially methylated probes in open sea and gene body regions (60% and 43%, respectively) (Figure 5d). Furthermore, a significant proportion of differentially methylated probes were mapped to known enhancer sites (40%). This enrichment in enhancer sites reached 73% of significant probes when considering only those probes hypomethylated after TGF- $\beta$  treatment. A selection of differentially methylated loci, including DNMTs, was validated using an independent quantitative method, bisulfite DNA pyrosequencing (Figure 5e).

Our data shows that the effect of TGF- $\beta$  in liver cancer cell lines comes along with a remarkable reconfiguration of the DNA methylome at multiple loci. This reconfiguration is stable and common to two independent cell lines, and affects a significant proportion of enhancer regions, potentially linked to gene expression changes. The TGF- $\beta$  methyl-sensitive signature described here includes DNA methylation players themselves and a significant enrichment of TGF- $\beta$  pathway loci, indicating a potential role for DNA methylation in establishing a TGF- $\beta$ -induced phenotype switch in these cells.

***TGF- $\beta$ -induced methylome matches the basal CD133+ methylome and is reflected on mRNA expression***

To gain a better insight on the consequences of TGF- $\beta$ -induced methylome reconfiguration on the phenotype, we performed a whole genome expression analysis in both, Huh7 and HepG2 cells. We chose the 8-days time point (4 days of TGF- $\beta$  treatment + 4 days post-release), considered in our

model as the one defining long-term, stable changes induced by this cytokine (Figure 6a). Bead array expression analysis showed an expected profile of gene expression in both cell lines, including known TGF- $\beta$  targets (Table S5). In addition, “type I transforming growth factor beta receptor binding” was the first gene ontology category at the molecular function level (Table S5). However, when intersecting the expression (n=1032) and methylation (n=242) significant gene lists, there was no significant overlap (26 common genes) (Figure 6b and 6c). Interestingly, a majority of overlapping genes (17 out of 26) were positively correlated between mRNA expression and DNA methylation (Figure 6c and Table S6). This was the case for key TGF- $\beta$  pathway targets such as *BMP1*, and de novo DNA methylation factor *DNMT3B*. Most of these correlations were validated by qRT-PCR analyses (Figure S6).

Although overlapping differentially expressed and differentially methylated genes did not show a marked correlation, the effect of a specific methylation change on gene transcription is known to depend on the genomic location (24). Therefore, we plotted all expression and methylation data, and analyzed separately CpG island and non-CpG island sites. As expected, no obvious correlation can be seen when plotting simultaneously all genes, independently of genomic location (Figure 6d). However, hypermethylated sites within CpG islands were positively correlated with gene expression. Genes displaying this behavior included *SMARCD3*, *COL12A1*, *CDKN1A*, and *TIMP2* (Figure 6d). Notably, most island CpG sites positively correlating with gene expression were located in the body of the genes. In contrast, CpG sites in non-CGI locations were either negatively (e.g. *PROM1*, *KRT86*, *IGFBP3*, *DLK1*) or positively (e.g. *VIM*, *HSPB8*, *CAV1*, *SERPIN2*) correlated with gene expression (Figure 6d). This was independent of the location of the CpG site within the gene.

Our two independent genome-wide experiments have shown that basal CD133+ cells from two liver cancer cell lines display a common methylome signature, and that TGF- $\beta$  is able in turn to induce a common reconfiguration of the methylome. As TGF- $\beta$  stably induces a de novo fraction of CD133+ cells, we asked whether the DNA methylation changes induced by TGF- $\beta$  were similar to the basal

CD133+ methylation profile, as obtained by FACS sorting from non-treated cell cultures. To answer this question, we studied the overlap between the two signatures (i.e. CD133+ and TGF- $\beta$ ) defined above, common to Huh7 and HepG2 cells. At p values <0.001, the CD133+ signature corresponds to 472 annotated genes, while the TGF- $\beta$  signature represents 1774 genes. We observed a significant overlap of 117 genes when intersecting both signatures (Figure 6d and Table S2). This overlap is highly significant ( $p < 1.5 \times 10^{-29}$ ) and represents 3 times more common sites than expected by chance. Pathways enriched in these 117 overlapping genes included mTOR signaling, Notch signaling, Wnt signaling and Focal Adhesion using two independent algorithms (Table S7). This result suggests that basal CD133+ cells and TGF- $\beta$ -induced CD133+ cells share a common methylome, potentially involved in sustaining their functional characteristics. It also shows that a subset of TGF- $\beta$  methyl-sensitive CpG sites is not co-methylated in CD133+ cells.

Based on the assumption that CD133+ cells induced by TGF- $\beta$  are epigenetically similar to CD133+ cells in basal conditions, we next asked for the location and magnitude of the TGF- $\beta$  effect on the methylome that is independent on those changes that define the CD133+ subpopulation. To answer this question we used all our bead array data, and modeled the main components of methylome variation in a linear regression (as described in Materials and Methods). Assuming that two known factors are able to modify the methylome based on our own results (the cell line of origin and the CD133-status), we included these two variables in the model. In addition, we included the potential effect of TGF- $\beta$ , independent of the other two factors. A panel of differentially methylated sites resulted from this analysis, and they represent CD133-independent changes induced by TGF- $\beta$  (Table S5). Therefore, this analysis suggests that TGF- $\beta$  has an additional effect on the methylome, independent on the induction of CD133+ cells.

In summary, TGF- $\beta$  -induced methylome resembles the basal CD133+ methylome and is partially reflected at the transcriptional level. A subset of TGF- $\beta$  methyl-sensitive loci positively correlates with gene expression.

## Discussion

In the present report, we comprehensively describe the DNA methylome of putative liver cancer stem cells. We used two non-related HCC cell lines to isolate pure populations of CD133- and CD133+ cells for DNA methylome assays. As has been previously reported, CD133+ cells isolated from liver cancer cell lines (including those used in the present study), are functionally distinct cells with increased ability to induce tumors in animal models (14). These findings are in line with clinical studies reporting poor prognosis for those HCC cases displaying higher proportions of CD133-expressing cells. Therefore, the signature identified here, may represent an established cellular program that defines the main characteristics of these cells.

Recently, the prognostic implications of TGF- $\beta$  pathway activation in HCC have been linked to the ability of this signaling pathway to induce metastatic behavior in a fraction of HCC cells (25). An additional link between CSCs and TGF- $\beta$  in HCC has been the recent demonstration that TGF- $\beta$  is able to increase the proportion of CD133+ cells in vitro (17). Here, we were able to reproduce and extend those previous observations. We showed that TGF- $\beta$  is able to increase CD133 expression at the protein and mRNA level in two non-related HCC cell lines. The effect induced by TGF- $\beta$  is stable, as opposed to the transient effect of the proinflammatory cytokine IL6. We show that this effect depends on specific signaling through TGF- $\beta$  type I receptor and is independent of increased cell proliferation of CD133+ cells. By using CD133-depleted cellular fractions, we show that TGF- $\beta$  is able to induce de novo expression of CD133. Furthermore, this induction of CD133 cells correlates with an increased ability to grow in non-attachment conditions, a surrogate functional assay for stem/CSC properties.

Both, basal CD133-expressing cells and TGF- $\beta$ -induced CD133+ cells, expressed increased levels of the de novo DNA methylation transcripts, *DNMT3A* and *DNMT3B*. This led us to further explore the

ability of TGF- $\beta$  to induce DNA methylation changes at the genome-wide level. We were able to show cell line-independent changes in DNA methylation induced by TGF- $\beta$  in a stable fashion suggesting an epigenetic mechanism involved in the establishment of a stemness cellular program. The methylome of TGF- $\beta$ -treated cells significantly overlapped with the methylome of CD133+ cells in basal conditions. This overlap indicates that TGF- $\beta$  not only induces a CSC marker, but also a defined DNA methylation profile. Furthermore, the finding that TGF- $\beta$  has additional effects on DNA methylation, cell line- and CD133-independent, was somewhat unexpected. This differential methylation may indicate an ongoing activation of the DNA methylation machinery leading to CD133 cell transdifferentiation. Further studies at longer time points and analyses of isolated CD133- cells may shed light on the ability of TGF- $\beta$  to imprint a DNA methylation signature without inducing CD133 expression.

Notably, although TGF- $\beta$  is known to induce DNA methylation changes at discrete loci (26–28), little evidence existed to date of a genome-wide level of TGF- $\beta$ -methylsensitive activity. Specifically, several previous reports were focused on chromatin changes associated with epithelial-mesenchymal transition (EMT). EMT is a developmental process that involves actin cytoskeleton reorganization and loss of apical–basal polarity and cell-to-cell contact, and like other developmental processes it involves epigenetic reprogramming (29). Pathological EMT processes, such as tumor metastasis, maintenance of cancer stemness, and fibrosis are also subject to epigenetic reprogramming. However, both at physiological and pathological levels, EMT-related epigenetic reprogramming has been mainly linked to widespread changes of histone marks or histone modifiers (30–32). Interestingly, gene-specific changes in DNA methylation have been correlated with the ability to maintain epigenetic silencing of critical EMT genes (27). In other words, DNA methylation seems to be involved in the process of fixing the switch between epithelial and mesenchymal phenotypes. This is in line with a model of persistent changes in DNA methylation induced by TGF- $\beta$  (Figure 7). Indeed, our experimental design was intended to reproduce an epigenetic process, by selecting only those changes in DNA methylation that survived cell division. Whether this effect of



TGF- $\beta$  is specific of transformed cells will require further studies. Nevertheless, the uncovering of the mechanisms involved in TGF- $\beta$  induction of cells with cancer stem cell behavior may have important consequences.

In summary, our data support and reinforce several previous studies that have pointed to an association between CSCs, and TGF- $\beta$ . In addition, we provide a mechanistic insight into the process that may lead to the stable induction of CSCs. Our study demonstrates that a key cytokine involved in HCC progression, TGF- $\beta$ , is able to transdifferentiate liver non-stem cancer cells into cancer stem cells. The results reported here are in agreement with a model in which DNA methylation plays a pivotal role in establishing the cellular program of liver cancer stem cells (Figure 7). The dynamics of a related process has recently been shown for CD44+ breast cancer stem cells (33). However, in our model, the effect of TGF- $\beta$  is persistent (as compared to the effect of another cytokine, IL6) and therefore epigenetically acquired. The evidence of an active interplay between TGF- $\beta$  and the DNA methylation machinery provided by our study supports this notion.

## **Materials and Methods**

### ***Cell culture and treatments***

Huh7 and HepG2 cells (American Type Culture Conditions) were cultured in DMEM medium (Gibco) at 37°C and 5% CO<sub>2</sub>, and were regularly tested for mycoplasma contamination (MycoAlert detection kit, Lonza).

For cytokine treatments, cells were plated and allowed to adhere before adding medium containing 10ng/ml final of IL-6 or TGF- $\beta$ 1 (recombinant human, Peprotech). For inhibition experiments, cells were treated with 2  $\mu$ M SB-431542 (Sigma-Aldrich) alone or in combination with TGF- $\beta$ 1.

For spheres formation assay, hepatosphere medium was prepared as previously reported (34).

Spheres were counted after 5 or 6 days.

### ***Fluorescence Activated Cell Sorting (FACS)***

Cells were labeled with antibodies against CD44, CD133 (AC133), EpCAM, CD90 or TGFBR II (Table S8).

Secondary antibodies were conjugated alternatively with FITC, Cy3 or Alexa750.

To study cell cycle progression, bromodeoxyridine (BrdU) (Sigma) incorporation and DNA content were simultaneously assessed, as previously described (35). Fluorescent events were captured using FACS instrument (FACSCanto II, BD Biosciences), and analyzed using BD FACSdiva 6.0 (BD Biosciences software) and WinMDI software (version 2.9).

### ***Magnetic Activated cell sorting (MACS)***

Huh7 and HepG2 cells were depleted or enriched for CD133<sup>+</sup> cells using magnetic-activated cell sorting (MACS, Miltenyi Biotec), with some adaptations to the manufacturer's instructions. Cells incubated 30 min at 4°C with FcR blocking reagent, followed by 15 min incubation with MicroBeads conjugated to monoclonal anti human CD133 antibodies. After washing, cell suspension was applied onto a pre-rinsed LS column placed in the magnetic field of a MACS separator. For CD133<sup>+</sup> depletion, flow-through the LS column containing unlabelled cells was collected. For CD133<sup>+</sup> enrichment, the

column was removed from the separator and placed on a 15 ml collection tube. Labeled cells were collected by firmly pushing the plunger in the column. To increase purity of CD133+ cells, the eluted fraction was enriched a second time over a new LS column. For each experiment, aliquots were kept to test by FACS the efficiency of the enrichment.

### ***Bisulfite modification and pyrosequencing***

To quantify the percentage of methylated cytosine in individual CpG sites, we performed bisulfite pyrosequencing, as previously described (36). For samples processed for Infinium bead arrays, the conversion was performed on 600 ng of DNA using the EZ DNA methylation Kit (Zymo Research) and modified DNA was eluted in 16 ul of water. Quality of modification was checked by PCR using modified and unmodified primers for *GAPDH* gene. Pyrosequencing assays (primers for PCR, sequencing primers and regions) are described in Table S9.

### ***Bead array methylation assays***

Methylation profiles of the different samples were analyzed using the 450K Infinium methylation bead arrays (Illumina, San Diego, USA). Briefly the Infinium Humanmethylation450 beadchip interrogates more than 480,000 methylation sites (20). The analysis on the bead array was conducted following the recommended protocols for amplification, labeling, hybridization and scanning. Each methylation analysis was performed in duplicate (for CD133+ versus CD133- samples) or in triplicate (for all other methylome analyses).

### ***Whole genome expression array***

Total RNA was isolated using the TRIzol Reagent (Invitrogen) according to the manufacturer's instructions. RNA quantity and quality were assessed with a ND-8000 spectrophotometer and bioanalyzer. 500 ng of total RNA was used for each Human HT-12 Expression BeadChips (Illumina), as previously described (37). 10 candidate genes were selected for validation using quantitative RT-PCR.

Four different housekeeping genes (*HPRT1*, *GAPDH*, *SFRS4* and *TBP1*) were alternatively used for internal control. The different primers used are listed in Table S10.

### ***Immunoblotting***

Protein extraction and immunoblotting was performed as previously described (37). Immunostaining was performed with anti-SMAD3, anti-phosphorylated SMAD3 and anti-tubulin/actin for loading control.

### ***Bioinformatics Analysis***

Raw methylation and expression bead array data was exported from Genome Studio (version 2010.3, Illumina) into BRB-ArrayTools software (version 4.3.1, developed by Dr. Richard Simon and the BRB-ArrayTools Development Team). Data was normalized and annotated using the R/Bioconductor package “lumi” (38). Class comparison between groups of bead arrays was done computing a t-test separately for each gene using the normalized log-transformed beta values. Only those probes with p values <0.001 and FDR<0.05 were considered significant for most analyses, except CD133- vs.

CD133+ comparison, where only the p value threshold was used. To define a “stable” methylome signature induced by TGF- $\beta$ , we performed a control vs. TGF- $\beta$  class comparison blocking by cell line status (Huh7 or HepG2), and including day 4 and day 8 of exposure (day 8 corresponding to 4 days of exposure to TGF- $\beta$  + 4 additional days with control medium). The analysis performed is an analysis of variance for a randomized block design. Two linear models are fit to the methylation data for each gene. The full model includes class variable and the block variable, and the reduced model includes only the block variable. Likelihood ratio test statistics are used to investigate the significance of the difference between the classes.

Using Infinium annotation data, Infinium sites (cytosines) were classified according to their relation to CpG islands and to the closest annotated gene. Sites unrelated to any annotated gene were classified as intergenic. Sites not related to CpG islands (CGI), CGI shores or CGI shelves, were

classified as Open sea sites. WebGestalt (WEB-based GENE SeT Analysis Toolkit) web application was used for gene set enrichment analyses, including Gene Ontology, and pathways (39).

Additional R/Bioconductor packages and R/Bioconductor scripts were used for modeling the effect of TGF- $\beta$  and CD133 expression in a linear regression. Data loading and preprocessing was performed with the “watermelon” package, removing low quality probes based on detection P value (40). This was followed by batch correction using the ComBat function of the “sva” package (41) and linear modeling using “limma” (42).

### ***Statistical Analysis***

BRBArrayTools and R/Bioconductor packages were used for bead array analyses, as described above. For other comparisons, means and differences of the means with 95% confidence intervals were obtained using GraphPad Prism (GraphPad Software Inc.). Two-tailed student t test was used for unpaired analysis comparing average expression between classes. P values < 0.05 were considered statistically significant.

## **List of abbreviations**

BMP: bone morphogenetic protein; BrdU: Bromodeoxyuridine; CGI: CpG island; CSC: cancer stem cell; EMT: epithelial-mesenchymal transition; FACS: fluorescence-activated cell sorting; HCC: hepatocellular carcinoma; IL6: Interleukin 6; TGF- $\beta$ : Transforming growth factor beta.

## **Competing interests' statement**

The authors declare no competing financial interests.

## **Authors' contributions**

M.M. and H.H. designed all experiments. M.M. and M.P.C. performed the experiments. F.L. supervised and coordinated the Illumina methylation and expression bead arrays. G.D. supervised and coordinated the Illumina methylation and expression bead arrays. H.H. performed the bioinformatics analyses and wrote the manuscript. Z.H. and H.H. coordinated the project. All authors discussed the results and manuscript text.

## **Acknowledgments**

We thank Fabienne Barbet from the ProfileXpert platform for the scan of the Illumina 450K arrays (service de Génomique & Microgénomique, Université Lyon 1, SFR santé LYON-EST, UCBL-INSERM US 7-CNRS UMS 3453, 8 av. Rockefeller 69373 Lyon Cedex 08)). This work was supported by la Ligue

National (Française) Contre le Cancer, l'Association pour le Recherche Contre le Cancer (l'ARC), and the Agence Nationale de Recherches sur le SIDA et les Hépatites Virales (ANRS).

## References

1. Ferlay J, Shin H-R, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. cancer* 2010;127:2893–917.
2. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet* 2012;379:1245–55.
3. Martin M, Herceg Z. From hepatitis to hepatocellular carcinoma: a proposed model for cross-talk between inflammation and epigenetic mechanisms. *Genome Med* 2012;4:1–13.
4. Hernandez-Gea V, Toffanin S, Friedman SL, Llovet JM. Role of the microenvironment in the pathogenesis and treatment of hepatocellular carcinoma. *Gastroenterology* 2013;144:512–27.
5. Massagué J. TGFbeta in Cancer. *Cell* 2008;134:215–30.
6. Bedossa P, Peltier E, Terris B, Franco D, Poynard T. Transforming growth factor-beta 1 (TGF-beta 1) and TGF-beta 1 receptors in normal, cirrhotic, and neoplastic human livers. *Hepatology* 1995;21:760–6.
7. Abou-Shady M, Baer HU, Friess H, Berberat P, Zimmermann A, Graber H, et al. Transforming growth factor betas and their signaling receptors in human hepatocellular carcinoma. *Am. J. Surg* 1999;177:209–15.
8. Mima K, Okabe H, Ishimoto T, Hayashi H, Nakagawa S, Kuroki H, et al. CD44s regulates the TGF- $\beta$ -mediated mesenchymal phenotype and is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Res* 2012;72:3414–23.
9. Ji J, Wang XW. Clinical implications of cancer stem cell biology in hepatocellular carcinoma. *Semin. Oncol* 2012;39:461–72.
10. Ma S. Biology and clinical implications of CD133(+) liver cancer stem cells. *Exp. cell Res* 2013;319:126–32.
11. Visvader JE, Lindeman GJ. Cancer stem cells: current status and evolving complexities. *Cell stem cell* 2012;10:717–28.
12. Grosse-Gehling P, Fargeas CA, Dittfeld C, Garbe Y, Alison MR, Corbeil D, et al. CD133 as a biomarker for putative cancer stem cells in solid tumours: limitations, problems and challenges. *J.*



Pathol 2013;229:355–78.

13. Mak AB, Nixon AML, Kittanakom S, Stewart JM, Chen GI, Curak J, et al. Regulation of CD133 by HDAC6 promotes  $\beta$ -catenin signaling to suppress cancer cell differentiation. *Cell reports* 2012;2:951–63.
14. Marquardt JU, Thorgeirsson SS. Stem cells in hepatocarcinogenesis: evidence from genomic data. *Semin. liver Dis* 2010;30:26–34.
15. Arzumanyan A, Reis HMGPV, Feitelson MA. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nat. Rev. Cancer* 2013;13:123–35.
16. Mani SA, Guo W, Liao M-J, Eaton EN, Ayyanan A, Zhou AY, et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008;133:704–15.
17. You H, Ding W, Rountree C. Epigenetic Regulation of Cancer Stem Cell Marker CD133 by Transforming Growth Factor-beta. *HEPATOLOGY* 2010;51:1635–44.
18. Bergman Y, Cedar H. DNA methylation dynamics in health and disease. *Mol. Biol* 2013;20:274–81.
19. Wang XQ, Ng RK, Ming X, Zhang W, Chen L, Chu ACY, et al. Epigenetic regulation of pluripotent genes mediates stem cell features in human hepatocellular carcinoma and cancer cell lines. *PloS one* 2013;8.
20. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics* 2011;98:288–95.
21. Ohishi W, Cologne JB, Fujiwara S, Suzuki G, Hayashi T, Niwa Y, et al. Serum interleukin-6 associated with hepatocellular carcinoma risk: A nested case-control study. *Int. J. cancer. J* 2014;134:154–63.
22. Callahan JF, Burgess JL, Fornwald JA, Gaster LM, Harling JD, Harrington FP, et al. Identification of novel inhibitors of the transforming growth factor beta1 (TGF-beta1) type 1 receptor (ALK5). *J. Med. Chem* 2002;45:999–1001.
23. Inman GJ, Nicolás FJ, Callahan JF, Harling JD, Gaster LM, Reith AD, et al. SB-431542 is a potent

and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol* 2002;62:65–74.

24. Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, et al. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* 2013;23:555–67.

25. Sasaki A, Kamiyama T, Yokoo H, Nakanishi K, Kubota K, Haga H, et al. Cytoplasmic expression of CD133 is an important risk factor for overall survival in hepatocellular carcinoma. *Oncol. reports* 2010;24:537–46.

26. Dumont N, Wilson MB, Crawford YG, Reynolds PA, Sigaroudinia M, Tlsty TD. Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. *Proc. Natl. Acad. Sci* 2008;105:14867–72.

27. Papageorgis P, Lambert AW, Ozturk S, Gao F, Pan H, Manne U, et al. Smad signaling is required to maintain epigenetic silencing during breast cancer progression. *Cancer Res* 2010;70:968–78.

28. Ding W, Dang H, You H, Steinway S, Takahashi Y, Wang HG, et al. miR-200b restoration and DNA methyltransferase inhibitor block lung metastasis of mesenchymal-phenotype hepatocellular carcinoma. *Oncogenesis* 2012;1.

29. Wu C-Y, Tsai Y-P, Wu M-Z, Teng S-C, Wu K-J. Epigenetic reprogramming and post-transcriptional regulation during the epithelial-mesenchymal transition. *Trends Genet* 2012;28:454–63.

30. Peinado H, Cano A, Esteller M, Ballestar E. Snail Mediates E-Cadherin Repression by the Recruitment of the Sin3A/Histone Deacetylase 1 (HDAC1)/HDAC2 Complex. *Mol. Cell. Biol.* 1AD;24:306–19.

31. Hou Z, Rauscher FJ, Longmore GD, Langer EM, Yan K-P, Ayyanathan K, et al. The LIM Protein AJUBA Recruits Protein Arginine Methyltransferase 5 To Mediate SNAIL-Dependent Transcriptional Repression. *Mol. Cell. Biol.* 5AD;28:3198–207.

32. Abell AN, Jordan NV, Huang W, Prat A, Midland AA, Johnson NL, et al. MAP3K4/CBP-

regulated H2B acetylation controls epithelial-mesenchymal transition in trophoblast stem cells. *Cell stem cell* 2011;8:525–37.

33. Chaffer CL, Marjanovic ND, Lee T, Bell G, Kleer CG, Reinhardt F, et al. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* 2013;154:61–74.

34. Haraguchi N, Ishii H, Mimori K, Tanaka F, Ohkuma M, Kim HM, et al. CD13 is a therapeutic target in human liver cancer stem cells. *J. Clin. Investig* 2010;120:3326–39.

35. Hernández-Vargas H, Palacios J, Moreno-Bueno G. Molecular profiling of docetaxel cytotoxicity in breast cancer cells: uncoupling of aberrant mitosis and apoptosis. *Oncogene* 2007;26:2902–13.

36. Hernandez-Vargas H, Lambert M-P, Le Calvez-Kelm F, Gouysse G, McKay-Chopin S, Tavtigian SV, et al. Hepatocellular carcinoma displays distinct DNA methylation signatures with potential as clinical predictors. *PLoS ONE* 2010;5.

37. Ouzounova M, Vuong T, Ancey P-B, Ferrand M, Durand G, Le-Calvez Kelm F, et al. MicroRNA miR-30 family regulates non-attachment growth of breast cancer cells. *BMC Genomic* 2013;14.

38. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008;24:1547–8.

39. Wang J, Duncan D, Shi Z, Zhang B. WEB-based GENE SeT AnaLysis Toolkit (WebGestalt): update 2013. *Nucleic acids Res* 2013;41:W77–W83.

40. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomic* 2013;14.

41. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012;28:882–3.

42. Wettenhall JM, Smyth GK. limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics* 2004;20:3705–6.

## Figure Legends

### **Figure 1. CD133- and CD133+ liver cancer cells differentially express genes involved in DNA methylation establishment and maintenance.**

**A.** Liver cancer cell lines (Huh7, and HepG2) were assessed for surface expression of CD133 by flow cytometry. The left panel shows a representative histogram for each of the cell lines (black histogram), with background (secondary antibody) represented by the empty histogram. The average expression  $\pm$  SD, from at least 3 independent assays, is shown in the right panel. **B.** The same cell lines were sorted using MACS (as described in Materials and Methods) and RNA was extracted to study the expression of stemness genes (*NANOG*, *POU5F1/Oct4*, and *SOX2*) by qRT-PCR. Intermediate levels of CD133 enrichment for Huh7 and HepG2 cells, and increase from left to right within each panel. A representative experiment of at least three independent MACS assays per cell line is shown. **C.** RNA samples isolated from the MACS-sorted cell populations (as in B) were subjected to qRT-PCR to measure the expression levels of the genes involved in DNA methylation or demethylation. Expression was normalized to the housekeeping gene *GAPDH*. (\*) indicates P value < 0.05.

### **Figure 2. A differential DNA methylome defines CD133- and CD133+ liver cancer cells.**

**A.** Huh7 and HepG2 cells were FACS sorted using CD133 antibody. Gates used to select negative and positive fractions are depicted in the upper panels. Duplicates of each fraction were used for Infinium 450k bead array DNA methylation analyses. **B.** AVG\_Beta values obtained from the bead array assay were plotted for on significant CpG site within the *CD133(PROM1)* promoter. The difference in methylation between CD133- and CD133+ cells (delta\_Beta) is indicated for each cell line. **C.** Median methylation (and distribution) for all differentially methylated loci ( $P < 0.001$ ) distinguishing CD133- versus CD133+ in both cell lines. **D.** Significant loci were distributed according to CpG island relationship as Island, north shore, south shore, north shelf, south shelf, and “Open sea”, and are represented in the upper pie chart. The lower pie chart represents the distribution of

significant loci in relation to annotated genes (within 200 or 1500 bp from the TSS, first exon, 3' or 5' UTRs, and gene body).

**Figure 3. TGF- $\beta$ , but not IL-6, induces CD133 expression in a stable fashion.**

**A.** Experimental design is indicated in the upper panel. Huh7 and HepG2 cells were grown in control culture conditions (depicted in gray text and lines), or exposed to 10 ng/ml IL-6 (red) or 10 ng/ml TGF- $\beta$  (blue) for 4 days. Cells plated in parallel, had their medium replaced by control culture medium and were left in culture for an additional 4 days. FACS expression of surface CD133 protein is shown for day 0, day 4, and day 8 (4 days treatment + 4 days post-release) for all conditions. Histograms are shown for one representative replicate in the middle panel. Mean  $\pm$  SD is shown for three biological replicates in the lower panel barplots. **B.** TGF- $\beta$  type I receptor antagonist SB 431542 was used at 2 $\mu$ M, alone or in combination with 10 ng/ml of TGF- $\beta$ , and DMSO used as control. CD133 expression was studied by FACS after 4 consecutive days of exposure to each experimental condition. (\*) indicates P value < 0.05 relative to all other conditions, for both cell lines. Representative phase contrast images are shown in the lower panels.

**Figure 4. de novo CD133 induction by TGF- $\beta$  correlates with overexpression of DNMT3 genes.**

**A.** the experiment in Figure 3A was repeated after MACS-sorting to enrich in CD133<sup>-</sup> cells, as depicted in the upper panel. Levels of CD133 expression were close to 0%, as shown in the upper histograms for both, Huh7 and HepG2 cells. Mean from three replicates is shown in the lower panels. (\*) indicates P value < 0.05 relative to control conditions. **B.** sphere formation assays were performed in non-attachment plates, after exposure to TGF- $\beta$  during 4 days. Spheres were counted after 6 days of growth in hepatosphere medium w/o TGF- $\beta$ . **C.** Huh7 (left panels) and HepG2 (right panel) cells were treated as in (A), RNA was extracted and qRT-PCR was performed for genes involved in DNA methylation or demethylation. Expression was normalized to housekeeping gene *GAPDH*. (\*) indicates P value < 0.05 relative to non-treated cells at the corresponding time point.

**Figure 5. Transdifferentiation to CD133+ cells correlates with a methylome reconfiguration.**

**A.** Huh7 and HepG2 cells were treated with TGF- $\beta$  for 4 days, or 4+4 post-release days, as described above. Biological triplicates were used to assess DNA methylation changes with Infinium 450k bead arrays. **B.** heatmap represents all probes differentially methylated (FDR<0.05) between control and TGF- $\beta$  treated cells, in both cell lines, and both time points. Blue indicates lower methylation, and red indicates higher methylation. The numbers on the right point to 5 different probe clusters selected according to their behavior across all samples. A fraction of each cluster is depicted in more details in **(C)**, to illustrate some of the significant genes within each category. CpG sites corresponding to *DNMT3* loci are indicated with a red asterisk. **D.** Significant loci were distributed according to CpG island relationship as Island, north shore, south shore, north shelf, south shelf, and Open sea, and are represented in the upper pie chart. Middle pie chart represents the distribution of significant loci in relation to annotated genes (within 200 or 1500 bp from the TSS, first exon, 3' or 5' UTRs, and gene body). Lower pie chart represents the fraction of differentially methylated probes annotated to a known UCSC enhancer. **E.** A selection of significant loci were validated by pyrosequencing (as described in Materials and Methods), in both cell lines. (\*) indicates P value < 0.05 relative to non-treated cells at the corresponding time point.

**Figure 6. The TGF- $\beta$  -induced methylome matches the basal CD133+ methylome and is reflected in mRNA expression.**

**A.** Whole genome expression analysis was performed after 4 days of TGF- $\beta$  exposure (+4 days post-release) in both cell lines, as described in Methods. RNA from control and treated conditions was interrogated with Illumina expression bead arrays. **B.** Venn diagrams illustrate the overlap between expression and methylation after TGF- $\beta$  exposure. 26 overlapping genes are listed in the right panel **(C)**, with red indicating increased expression/methylation, and green indicating reduced expression/methylation after TGF- $\beta$ . **D.** Correlation between methylation and expression at the

genomic regional level in Huh7 cells. Panels show the correlation of delta\_beta (methylation) in the x axis and fold-change (expression) in the y axis. Upper panels correspond to all RefSeq genes without any filter, or separately for CpG-island (CGI) or non-CGI related sites. Lower panels show the same analysis after filtering for differentially methylated and differentially expressed genes. Examples of specific genomic regions (i.e. TSS200, TSS1500, or Gene Body) are listed below the lower panels. The same analysis in HepG2 cells is shown in Figure S6. **E.** A similar correlation by genomic features was done between the two methylomes, CD133+ and TGF- $\beta$  (described in figures 2 and 5, respectively). Correlations and  $r^2$  values are shown within the panels. **F.** Overlap between the two methylation signatures, CD133+, and TGF- $\beta$  exposure, was done using only differentially methylated genes. The relative enrichment (representation factor) and significance are also shown.

**Figure 7. Model of epigenetic switch from non-stem cancer cell to CSC induced by TGF- $\beta$  in liver cancer cells.**

Within liver cancer cell lines a minority of cells express the surface marker CD133. These so called cancer stem cells (CSC) are depicted in red. Non-stem cancer cells are shown in blue. In the cell culture, CSCs can be induced spontaneously (discontinuous arrow) or under the effect of cytokines such as interleukin-6 (IL6) and transforming growth factor beta (TGF- $\beta$ ). The effect of TGF- $\beta$  is higher in magnitude and persistent through cell division (thicker and one-directional arrow). In our model, this epigenetic effect of TGF- $\beta$  is mediated by changes in DNA methylation mediated by de novo DNMTs, DNMT3A and DNMT3B.

## Supplementary Figures

### Supplementary Figure 1.

**A.** Percentage of positive cells for candidate liver cancer stem cell markers, in two unrelated liver cancer cell lines, Huh7 and HepG2. **B.** sphere formation assay comparing CD133- and CD133+ cells in Huh7 cells. After MACS purification, cells were plated in non-attachment plates, and their growth as spheres was quantified after 6 days. Only structures grown in suspension, with refractory well-defined limits, were counted as spheres. Mean and SD from 3 technical replicates is shown on the left panel. One representative image of each condition is shown on the right panel.

### Supplementary Figure 2.

Unsupervised clustering and heatmaps of differentially methylated loci (CD133- vs CD133+) after Infinium 450k bead array analysis. Common analyses for two cell lines (Huh7 and HepG2) is shown in **(A)**. Single cell analyses are shown for Huh7 **(B)** and HepG2 **(C)**. The color scale represents less methylated loci in blue, and highly methylated loci in red.

### Supplementary Figure 3.

**A.** FACS analysis of TGFBR11 expression in Huh7 and HepG2 in basal conditions. Percentage of positive cells relative to background secondary antibody is shown in each chart. **B.** western blot for SMAD proteins was performed for the two cell lines, in control conditions, or after stimulation with TGF- $\beta$  during 4 days. **C.** representative phase contrast images of Huh7 and HepG2 cells left untreated or exposed to IL-6 or TGF- $\beta$  during 4 days. **D.** viability was assessed by trypan blue exclusion in cells treated or not with IL6 or TGF- $\beta$  during the indicated time points. Percentage of trypan positive cells is represented on the bar plots. **E.** Representative phase contrast images of Huh7 and HepG2 cells treated from 1-3 days with the indicated conditions: mock, DMSO, TGF $\beta$  receptor I inhibitor (SB-431542), TGF- $\beta$  alone or in combination with SB-431542 inhibitor. All conditions were performed in triplicate culture wells.



#### **Supplementary Figure 4.**

**A.** BrdU uptake was used to estimate the proliferation index of both cell lines in different culture conditions, and after two time points. FACS analysis was performed in combination with propidium iodide staining to separate the cells by cell cycle stage. **B.** mRNA expression of CD133 in the same conditions described for Figure 4a. **C.** Non-attachment growth assay was performed after 4 days post-release from a 4 day treatment with TGF- $\beta$ . Sphere formation was assessed 6 days after culture with hepatosphere medium. (\*) indicates P value < 0.05 relative to non-treated.

#### **Supplementary Figure 5.**

**A.** experiment design to study differential methylation in response to TGF- $\beta$ . Unsupervised clustering (**B**) and heatmaps (**C**) of differentially methylated loci after treatment with TGF- $\beta$  in two HCC cell lines, Huh7 and HepG2. Analysis was performed as described in Materials and Methods.

#### **Supplementary Figure 6.**

**A.** Correlation between methylation and expression at the genomic regional level in Huh7 cells. Panels show the correlation of delta\_beta (methylation) in the x axis and fold-change (expression) in the y axis. Upper panels correspond to all RefSeq genes without any filter, or separately for CpG-island (CGI) or non-CGI related sites. Lower panels show the same analysis after filtering for differentially methylated and differentially expressed genes. Examples of specific genomic regions (i.e. TSS200, TSS1500, or Gene Body) are listed below the lower panels. The same analysis in HepG2 cells is shown in (**B**). **C.** A selection of significant genes was validated by qRT-PCR in both cell lines. (\*) indicates P value < 0.05 relative to non-treated.

## **Supplementary Tables**

### **Supplementary Table 1.**

Differentially methylated sites in CD133- vs. CD133+ cells, based on Infinium 450k data.

### **Supplementary Table 2.**

Gene set enrichment analyses using BRBArray Tools, and comparing the methylomes of CD133- and CD133+ cells in two cell lines, Huh7 and HepG2.

### **Supplementary Table 3.**

List of differentially methylated sites in response to TGF- $\beta$  and in two cell lines, Huh7 and HepG2 (TGF- $\beta$  signature).

### **Supplementary Table 4.**

Pathway analysis of TGF- $\beta$  methyl-sensitive sites in two cell lines, Huh7 and HepG2.

### **Supplementary Table 5.**

Genes differentially expressed (including gene ontology and pathway analysis) in response to TGF- $\beta$ .

### **Supplementary Table 6.**

List of genes inversely correlated at expression and methylation levels in response to TGF- $\beta$ .

### **Supplementary Table 7.**

List of genes overlapping two methylome signatures, CD133 and TGF- $\beta$ . Significant genes from both differential methylomes and common to Huh7 and HepG2 were used for pathway analysis.

### **Supplementary Table 8.**

List of antibodies used for characterization of liver cancer stem cells.

**Supplementary Table 9.**

List of pyrosequencing assays.

**Supplementary Table 10.**

List of primers used for qRT-PCR.

