

Université de Lille

École graduée Biologie Santé de Lille

THÈSE

Pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITÉ DE LILLE

Discipline : Biochimie, Biologie cellulaire et moléculaire, Bio-informatique

Contrôle transcriptionnel de la plasticité hépato-cholangiocytaire

Soutenue publiquement le 12 décembre 2025 par

Ludivine VASSEUR

Composition du jury :

Pr. **Bart STAELS**, INSERM UMR1011, Lille

Président

Dr. **Nathalie THERET**, IRSET/UMR_S1085, Rennes

Rapporteur

Dr. **Renaud DENTIN**, INSERM U1016, Institut Cochin, Paris

Rapporteur

Pr. **Gilles SALBERT**, IGDR, UMR6290 CNRS, Rennes

Examineur

Dr. **Rita MANCO**, UCL Louvain, Belgique

Examineur

Dr. **Jérôme EECKHOUTE**, INSERM UMR1011, Lille

Directeur de thèse

À mon étoile, mon Papy Ly,
J'espère que tu es fier de moi de là-haut.

Remerciement

Au président du jury, Monsieur le Professeur Bart Staels, je vous remercie d'avoir accepté de présider cette thèse. Je souhaite également vous remercier pour votre accueil au sein de l'UMR1011 lors mon stage de Master 1 en 2021 malgré les différentes restrictions sur l'accueil des stagiaires pendant la période COVID. Ce fût un honneur d'apprendre à vos côtés pendant les nombreuses réunions de laboratoires, puis, lors des réunions de suivi en plus petit comité. Merci pour vos partages et retours scientifiques.

À Monsieur le Docteur Dentin Renaud et Madame le Docteur Théret Natalie, je vous remercie d'avoir pris le temps d'évaluer ce travail. Je vous suis reconnaissante d'avoir mis à ma disposition votre temps et votre expertise scientifique pour la lecture de ce manuscrit. Recevez ma profonde gratitude et mon profond respect.

À Monsieur le Professeur Salbert Gilles et Madame le Docteur Manco Rita, merci d'avoir répondu présent pour l'évaluation finale de mon travail de thèse. Je vous remercie d'avance pour la discussion scientifique et le temps que vous prenez pour juger mon travail.

Je remercie Monsieur Salbert pour son accompagnement lors de mes comités de thèse, merci pour vos précieux conseils scientifiques qui ont fait grandir ce projet. Vous m'avez vu évoluer scientifiquement au fil des années lors de mes comités de thèse, et vous avez toujours su avoir les bons mots pour me rassurer et m'accompagner dans mon parcours professionnel. Soyez assuré de ma reconnaissance.

À Monsieur le Docteur Jérôme Eeckhoutte, que dire ... Je vous remercie de m'avoir accueillie pour mon stage de Master 1 en 2021, puis, de m'avoir à nouveau acceptée pour mon stage de Master 2. Merci de m'avoir fait assez confiance pendant mon stage de Master 1 pour déposer une demande de financement de thèse auprès

d'EGID/PreciDIAB afin de me permettre de poursuivre ce projet avec vous. Je ne pourrai jamais suffisamment vous exprimer ma gratitude d'avoir été mon directeur de thèse. Malgré le contexte difficile et les traces du passé, vous avez su garder le cap. Sans vous, cette thèse aurait été différente et je ne sais pas si j'aurais accepté de me lancer. Alors même si le chemin a été long et tumultueux, vous restez un scientifique hors-pair et je n'aurais pas pu demander mieux pour ma direction de thèse. Grâce à vous, j'ai pu grandir scientifiquement, vous m'avez appris à raisonner, à rebondir, à analyser et à re-analyser (parce que finalement la moyenne géométrique est plus fiable que la moyenne arithmétique). J'espère que cette thèse est à la hauteur de vos attentes. Vous m'avez permis de travailler sur un projet novateur et stimulant qui m'a aidé à consolider mes doubles compétences en biologie et en bio-informatique. Votre intelligence, votre bienveillance et votre attention à bien faire, tout en s'adaptant aux différentes personnalités de vos collaborateurs sont vos qualités les plus précieuses. Merci d'avoir été une personne à l'écoute tout au long de ma thèse. Sachez que je vous serai éternellement reconnaissante pour tout ce que vous m'avez apporté.

À Monsieur le Docteur Dubuquoy Laurent, merci d'avoir fait partie de mon comité de suivi de thèse. Merci pour vos précieux conseils scientifiques, merci pour le temps que vous avez pris lors des réunions supplémentaires sur le projet, c'est grâce à vous que nous avons pu avoir accès aux biopsies de patients. Merci également pour votre expertise scientifique et votre œil d'expert sur les différents marquages immunohistologiques réalisés. C'est par votre confiance que ce travail a pris autant de valeur. Vous m'avez également vu évoluer scientifiquement au cours de toutes ces réunions. Merci de m'avoir écoutée, conseillée et orientée pendant mes comités de thèse sur les différentes formations et sur l'après thèse. Soyez sûrs de recevoir ma gratitude.

Je remercie également les organismes financeurs de cette thèse. Ce travail a été financé par l'Institut Européen de Génomique et du Diabète – EGID – (ANR-10-LABX-0046), le Centre National de Médecine de Précision des Diabètes – PreciDIAB – (ANR-18-IBHU-0001) et le Conseil Régional des Hauts-De-France (Convention 22005973). Je remercie également la Fondation pour la Recherche Médicale – FRM – de m'avoir attribué un soutien de fin de thèse (FDT202504020162).

À l'équipe du département de Biochimie de la faculté de pharmacie d'Amiens, merci de m'avoir fait confiance et recrutée en tant qu'Attaché Temporaire de Recherche et d'Enseignement. Merci également à l'équipe MP3CV pour votre accueil chaleureux et mon intégration à l'équipe, notamment à Mesdames Gaëlle Lenglet-Savreux, à Chantal Allou, à Carine Avondo et à Emma Catel-Dobel. Je tiens à remercier particulièrement Monsieur le Professeur Kamel Said, Monsieur le Professeur Romuald Mentaverri et Monsieur le Professeur Youssef Bennis pour leur confiance à ce poste. J'espère pouvoir m'épanouir dans ces fonctions.

Un grand merci, également, à la plateforme de microscopie PLBS, en particulier à Messieurs Bongiovanni Antonino et Jarmouni Yassine. Votre temps et vos conseils précieux, nous ont permis de réussir l'imagerie sur notre modèle de culture tri-dimensionnelle de cellules. À tous ces après-midis passés au microscope ou à scanner les lames dans les salles trop chaudes ou trop froides. Sans vous, nous n'aurions pas réussi à répondre totalement aux commentaires de l'éditeur pour mon papier.

À l'équipe de Rennes, Monsieur le Docteur Loyer Pascal et Madame le Docteur Corlu Anne, vous qui maîtrisez si bien les HepaRGs. Merci d'avoir réalisé les expériences de surexpressions dans cette lignée. Malgré nos courtes réunions, vous avez vite compris notre attente et vous avez su nous envoyer les échantillons rapidement pour que nous puissions soumettre le papier. Avec vos précieux conseils, vous avez (presque) réussi à me réconcilier avec cette lignée qui m'a donné du fil à retordre pendant mon stage de Master 1.

À Madame le Docteur Gnemmi Viviane, votre œil d'experte sur les lames de patients a mis en exergue une dimension supplémentaire à ce projet. Merci de nous avoir donné accès à d'autres échantillons de patients ayant différentes pathologies hépatiques, principalement atteints de cholangiopathies afin de regarder si nous retrouvions notre facteur également dans ces échantillons.

À Monsieur le Docteur Bou Saleh Mohammed, un grand merci pour ton temps et ton expertise. Tu nous as fourni les protocoles détaillés pour les IHC, tu as su nous accompagner à chaque problème et nous donner des protocoles pour le triple marquage lorsque les reviewers nous l'ont demandé. Tu nous as permis d'utiliser certains réactifs et votre cocotte afin d'être sûr que tout se passe bien. Merci également de nous avoir donné quelques conseils sur les HepaRGs qui nous ont permis d'essayer de les appréhender. Ton expertise sur les coupes d'histologies est très impressionnante, merci de nous avoir autant guidés. Je te souhaite le meilleur pour la suite.

À Madame le Docteur Taront Solenne, c'est grâce à toi que le papier contient les organoïdes humains. Merci, malgré les problèmes de matrigel et de milieu, d'avoir lancé cette expérience. Tu as pu finalement réaliser l'infection dans les temps, avant de partir en vacances, pour que je puisse avoir les échantillons à mon retour de vacances.

À Monsieur le Docteur Furlan Alexandro, merci de nous avoir donné des ampoules des BMELs.

À Monsieur le Docteur Lefebvre Philippe, merci pour vos nombreuses questions qui ont fait rebondir ce projet. Merci également d'avoir essayé de me challenger avec vos interrogations et remarques (qui ont toutes fait leur petit bout de chemin). Votre regard extérieur au projet a permis de mettre en avant des aspects un peu discutables au début de ma thèse. Merci également d'avoir amélioré ma condition de travail en me prêtant un de vos écrans « parce que quelqu'un qui fait de la bioinformatique ne peut pas travailler avec un matériel aussi vieux et pixelisé » si je reprends vos propos. Merci aussi pour vos discussions non scientifiques et nos conversations sur les sujets de tous les jours qui m'ont rappelé qu'au-delà d'être une thésarde, je pouvais avoir une vie normale. Vous m'avez considéré comme une personne et non comme une étudiante, alors merci pour ça.

Au groupe des alcooliques Unamines, aux anciens et actuels thésards de JK, Lucie, Ludivine C., Arielle, Arthur, Charlotte B., Claire, Fabienne, Mickael, Sadia, merci

pour nos échanges et discussions. Merci pour nos sorties qui ont un été ma bulle d'air pendant cette thèse. Vous m'avez rapidement intégrée et je vous en serai toujours reconnaissante.

Un merci particulier à Christina et à Chloé, même si vous avez commencé (et fini) votre thèse avec moi, nous avons partagé un bon nombre de formations. Nous avons pu chacune être là dans des moments difficiles de nos parcours. Je vous souhaite plein de bonnes choses pour la suite et de réussir dans les voies que vous avez choisies. Merci pour nos rires en formation, merci pour vos sourires, merci d'avoir partagé un petit bout de vos parcours avec moi.

À Juliette, merci pour ta bonne humeur et ton énergie débordante. Je me souviendrai de toutes nos conversations en salle de culture ou entre 2 couloirs qui remontent le moral. Reste comme tu es, tu es une personne incroyable, je te souhaite plein de bonnes choses pour la suite.

À Kimberly, à ton sourire et ta bonne humeur peu importe la situation. À tes pâtisseries ramenées en salle de pause qui réchauffent nos cœurs, d'ailleurs j'espère qu'on te verra un jour dans l'émission « Le Meilleur Pâtissier » !! Merci pour nos discussions simples sur la vie qui m'ont redonnées le sourire pendant mes moments down. À toutes ces discussions mariage que nous avons eues et qui ont montrées que nous étions sur la même longueur d'onde. Pour reprendre tes propos, je sais que je suis folle de soutenir ma thèse et d'organiser mon mariage en même temps... Restes comme tu es avec ta joie de vivre quotidienne ! Je suis heureuse d'avoir croisé ton chemin et je te souhaite plein de bonnes choses pour la suite.

À Cyril, pour les pauses café et les bières que nous avons pu partager ensemble. Merci d'organiser nos labmeetings. Merci d'avoir su gérer les imprévus sans demander plus d'explications. Je pense que tu es devenu un point de repère pour de nombreuses personnes à JK. Tu aides le service administratif dans l'ombre et tu gères en même temps ta recherche, alors pour ça chapeau. J'espère que nos chemins se recroiseront en dehors du labo, à un prochain afterwork qui sait ?

À tous ceux de l'U1011 et qui ont croisé mon chemin (même si on n'était pas sur le même site), Cécilia, Viktor, Margaux, Bettina, Emily, Doriane, Audrey D., Emily,

Laurent, Marie, Vance, merci pour tous ces échanges que l'on a pu avoir et pour les soirées karaoké annuelles.

À Laura, je pense que tu as été une des premières à louer mes qualités de présentation et la qualité de mes diaporamas. J'espère que j'ai pu au cours de mes présentations te réconcilier avec l'épigénétique. Je te remercie pour les conversations que nous avons pu avoir sur mon après-thèse. Merci pour tes conseils que j'estime précieux. Merci aussi d'avoir eu les mots justes pour que je puisse me poser les bonnes questions et que je fasse quelque chose qui me plaise vraiment. J'espère que tu t'épanouiras dans ce nouveau poste.

À Éloïse, tu es un peu un membre de l'équipe 4 aussi finalement. Je ne sais pas quoi te dire à part merci, merci d'avoir été là. Je pense que sans toi, tout aurait aussi été différent. À nos nombreuses pauses café du vendredi et aux petits déjeuner improvisés pour parler de tout et de rien qui m'ont très souvent remontés le moral. Merci de m'avoir écoutée me plaindre et d'avoir compris ce qui me tracassait.

À l'équipe 4 anciens comme actuels, merci pour nos afterworks, nos restaurants, les pauses café, nos moments gossip. Merci tout simplement d'avoir été là, grâce à vous j'ai pu réaliser ma thèse dans de bonnes conditions, couplé d'entraïdes et de bienveillances. Sachez que sans vous, cette thèse aurait eu un autre goût. J'espère que vous ne m'oubliez pas, sachez que je serai toujours disponible pour venir boire un verre à vos côtés ! Mon départ m'a déchiré le cœur, j'avais besoin d'une nouvelle bulle d'air, mais sans vous c'est quand même différent.

En particulier, je remercie Francesco pour nos discussions à ces afterworks. C'est toujours un plaisir de te croiser et de partager un moment avec toi. Merci pour ton accueil dans l'équipe lors de mon premier jour de Master 1 et merci de m'avoir encadré dans mes premiers pas en salle de culture.

Je tiens également à remercier Marie, même si nos chemins ne se sont pas croisés longtemps, tu as été la première à m'intégrer dans l'équipe. Je suis heureuse de voir que tu t'épanouis chez Genfit.

À Manu, merci de m'avoir fait rire quand je déprimais en salle de pause. Tu as divertie et fait rire l'équipe un bon nombre de fois pendant tout ton contrat ici. Ta bonne humeur qui a nous permis de tenir pendant la période COVID.

À Lora, au final, on aura partagé très peu de temps ensemble dans l'équipe avec ma nouvelle prise de poste. J'espère que tu t'épanouiras dans l'équipe.

À Hélène et surtout à ton courage malgré tout ce que la vie a mis sur ton chemin, tu gardes la tête haute. Tu es impressionnante. Je n'ai pas eu la chance de travailler avec toi, mais humainement tu es une personne incroyable. Fais attention à toi surtout.

À Desiree, now you're the only Jerome's PhD Student !! I hope everything will go well for you and you will find what you want here in the team. I know that your position is complicated due to the consortium and all that happened with its. It was a pleasure to meet you and to share with you the end of my thesis. Don't forget to take some breaks with the team, it will be your fresh air during your PhD contract and be careful with the french administration.

À Loïc, sans toi une partie de ce projet n'existerait pas. Cet atlas de scRNA-Seq c'est le tien. Tu as beaucoup donné pour satisfaire tous les utilisateurs et qu'il soit un maximum « users-friendly » ! Tu t'es vraiment donné à fond pendant 3 ans pour cet atlas, malgré tous les problèmes que tu as pu rencontrer. Depuis que tu es parti, je me sens beaucoup plus seule quand ils parlent des enfants en pause ... mais, je sais que tu t'épanouis avec tes nouvelles missions.

À Clémence, un grand merci pour ton aide et tes conseils. Tu as toujours été une personne souriante, tu m'as aidé à m'intégrer au groupe des alcooliques Unamines. Ta gentillesse et ta volonté de partager tes connaissances ont permis à beaucoup de personnes de s'épanouir dans l'équipe. Tu manques beaucoup à l'équipe, mais je sais que tu t'épanouis dans ton nouveau travail et dans ton rôle de maman.

À Viktorija, we shared the same tastes in music and we have the same passion for cocktails. We always order the same cocktail, just by carelessness, which is always funny. One of my best memories with you will be the Linkin Park concert, for which it was complicated to get tickets. I know that you can count on me to share concerts with you whenever you want!! I hope that you will find a job where you will blossom. Thank you for your kindness and for everything we shared.

À Georgiana, thank you for your support. I'm so proud to help you to start running. It was so emotional to see at your first official run "La course des princesses" in Versailles. You did it so well!! Thank you also for taking care of Cannelle everytime we need to go home.

À Manjula, thank you for being you, thank you for all Indian food you bring to the lab. You were an incredible support since the beginning. You always try to do your best and to help people even if the life did not help you. It was so wonderful to share moment with you. Thank you for helping me and to try to find me a teaching position with your friend in Paris. I was happy to share moment with you. You will always have a particular place in my heart.

À Alexandre, merci pour ton grand cœur et ton côté protecteur. Merci pour l'aide que tu as pu m'apporter grâce à tes connaissances et tous les protocoles que tu as retrouvés dans tes cahiers de labo et qui m'ont bien aidé et débloqué. Je suis heureuse d'avoir partagé pendant de nombreux mois le même bureau que toi, même si j'ai eu froid tout l'hiver avec le vent dans mon cou et chaud l'été parce que le bureau est en plein soleil. Merci de m'avoir laissée accrocher notre petite affiche de hyènes sur la porte. Je te remercie également pour le soutien que tu as été pendant ces années de thèse, et oui, je me souviendrai qu'avant les thésards n'étaient pas forcément payés. Saches que tu es une belle personne et que tu es un puits de savoir scientifique, je suis heureuse d'avoir travaillé avec toi.

À Sandra, ton arrivée à apporter une bulle d'air frais à l'équipe. Tu es arrivée avec ta bonne humeur, ce qui nous a fait un grand bien après les épreuves traversées. Merci pour nos discussions, sauf quand je te demande juste un réactif et que ton côté pipelette prend le dessus. Tu es une belle personne, fait juste attention à ne pas te stresser pour de simples petites choses. Merci également de m'avoir montré que je ne suis pas la seule à avoir autant de paires de chaussures !! Tu mérites le meilleur.

À Ninon, tu es une personne avec un si grand cœur et une très grande scientifique, tu mérites vraiment d'atteindre tes objectifs. Je tiens à te remercier d'avoir tant donné pour ce projet, merci pour l'apport de ton expertise sur la partie EZH2/OGT et merci d'avoir donné autant de ton temps pour mettre en place les protocoles d'IHC. Une grande partie des expériences demandées par les reviewers du papier ont été possibles grâce à toi. Merci pour nos longues conversations pendant les sacrifices pour le projet KAT2B, merci pour nos rires et merci de t'être inquiétée dans les

moments plus difficiles. Je te suis reconnaissante également du temps de relecture de mon introduction pour que je sache si elle était claire pour une personne extérieure au manuscrit.

À Céline, toi qui m'as tout appris depuis mon stage de Master 1. Tu m'as transmis une bonne partie de tes connaissances techniques, ce qui me permet aujourd'hui d'être rigoureuse, pointilleuse et organisée dans mes expériences. Tu m'as également transmis énormément de choses sur la partie *vivo*. Je te serai éternellement reconnaissante pour cela. Tu m'as appris qu'il est possible d'avoir un équilibre entre vie pro/vie privée. Merci pour nos conversations le matin, merci de m'avoir remonté le moral dans les moments les plus difficiles, d'avoir été présente lors du décès de mon papy, parce qu'en plus de m'avoir soutenue, tu m'as aidée pour que je puisse me libérer du temps pour faire les allers-retours. Tu m'as soutenue et encouragée pendant ma prépa semi-marathon au début de l'année. J'espère que l'on aura l'occasion de faire quelques courses ensemble (en plus de celles déjà prévues), je partagerai un dossard avec toi avec plaisir ! Un grand merci pour ton soutien, pour ton écoute dans les bons comme dans les mauvais moments. Tu es une personne formidable au grand cœur. Je voudrais finir par te remercier pour ton soutien pendant ces 3 ans de thèse. Tu as été d'une grande aide sur ce projet, tu as beaucoup donné de ton temps et de ton expertise. Et puis finalement, toi et moi on le sait, on prend sûrement les choses trop à cœur pour certaines personnes, mais nous sommes professionnelles et on se donne à fond. J'espère que l'on restera en contact.

À Julie, merci d'avoir été ma co-encadrante de thèse de l'ombre, même si ton nom n'est pas officiellement sur cette thèse et que tu n'as pas été impliquée dans le projet comme tu l'aurais voulu. Tu m'as montrée que la recherche peut être aussi humaine. C'est à toi que je dois beaucoup dont mon autonomie et ma curiosité pour la bio-informatique en particulier, tu ne m'as pas lâchée, tu m'as permise d'évoluer professionnellement. Je te suis reconnaissante de toutes les heures que tu as passé à me former et m'expliquer comment débloquer les problèmes et comment trouver des solutions. Je te remercie également d'avoir été d'un grand soutien émotionnel pendant ces 3 ans de thèse, sache que sans toi, ce travail n'existerait pas. Merci de m'avoir écoutée dans les bons comme dans les mauvais moments et d'avoir su me donner des conseils, conseils que toi aussi tu devrais appliquer (fait ce que je dis, pas ce que je fais) et de m'avoir fait rire pour me remonter le moral. Je retiendrai toutes nos

discussions et tes conseils pour survivre à la recherche académique, mais aussi les moments que l'on a pu passer ensemble en médiation scientifique lors de la fête de la science. Merci également de t'être battue pour que je puisse partir en congrès pendant ma thèse (mais tout ce qu'on a vécu en Finlande à l'ECCB reste en Finlande). Pour moi, sache-le, tu as été beaucoup plus qu'un support technique pendant ma thèse. Garde ton énergie débordante qui te permet de faire de grandes choses et de te dépasser sur les courses ! J'espère que l'on restera en contact et que vous garderez mon numéro pour les afterworks.

À tous mes amis et à nos soirées qui m'ont permis de près ou de loin à garder le cap pendant ces 3 ans de thèse.

À Adrien, Alice, Myriam, Pierrick, Solène, Pierre, merci pour nos soirées même si on a toujours du mal à trouver une date qui convient à tout le monde. Ces moments ont été mes bouffées d'air frais et mes moments de réconforts quand ça n'allait pas, je sais que vous avez essayé de comprendre et de m'aider alors que je travaille dans un domaine opposé aux vôtres.

À Gaëlle, toi qui es ma plus vieille amie, on se connaît depuis le lycée et on ne s'est jamais lâchée. Tu es une grande force de la nature et la personne la plus empathique que je connaisse. Tu as le cœur sur la main et en te connaissant, on sait pourquoi tu es infirmière. Après ce qu'il t'est arrivé, tu mérites tout le bonheur de monde. Tu as été d'un grand soutien pour moi, pendant toutes mes études. On ne s'est jamais éloignée, même si maintenant on doit s'organiser 3 mois à l'avance pour pouvoir se voir. Mais, ce n'est pas pour autant que notre relation a changé. Tu fais partie intégrante de ma vie et puis c'est grâce à toi que j'ai rencontré Aurélien. Merci d'avoir été là pour nous soutenir avec Thomas sur les courses qu'on a pu faire (malgré le fait que dès que Thomas court avec nous, il pleuve). Merci pour nos week-end à parler de tout et de rien et pour nos journées à Disney qui redonnent le sourire. Je suis heureuse d'avoir une personne comme toi dans ma vie.

Un merci particulier à ma thésarde préférée de Rennes, Claire. À nos appels qui durent des heures où l'on se tenait informées de l'avancée de notre thèse. Ces appels nous ont permis de relativiser et d'être écouté par une personne extérieure qui comprend ce que l'on vit. À nos week-ends aussi où tu m'as notamment fait découvrir

Rennes. Comme je te l'ai dit, on se retrouve pour ta soutenance en mai, tu es dans la dernière ligne droite, courage !

À mes bitches de la danse, comme on se surnomme, parce que vous êtes devenues la famille et que vous avez su m'écouter quand ça n'allait pas. À Angélique, Annabelle, Stéphanie, Flora, Anne, Camille, Sophie, merci pour toutes nos soirées à boire, à manger, à danser et à chanter (même si on le sait, nous sommes meilleures danseuses que chanteuses). Merci pour nos rires et nos journées de répétition, à toutes les fois où l'on change un détail à quelques semaines du spectacle. Parce qu'au final c'est chez Madame qu'on est le mieux, alors, également merci à toi notre Madame d'avoir créé cette famille avec tes grandes. On n'oubliera pas, à ton mariage, tu nous as promis de ne pas t'arrêter.

Un merci particulier à Charlotte, ou grande charlotte, merci d'être toi avec ton franc parlé et tes anecdotes qui n'arrivent qu'à toi, ça fait tellement plaisir d'avoir une personne comme toi dans la vie. Tu m'étonneras toujours. Tu mérites le meilleur, tu es une personne au grand cœur.

À Inès et à Adèle, mes petits bébés de la danse. Vous avez amené une énergie différente à nos cours. Merci pour les moments qu'on a partagés ensemble et ceux à venir. Ines, dépêche-toi d'être diplômée que je devienne ta patiente avec mon genou en vrac. Adèle, je te souhaite plein de bonnes choses à Barcelone, mais, tu nous manques beaucoup ici.

À Charlotte, ma cha, tu fais autant partie de la famille que de mes amies. Tu es pour moi, la sœur que je n'ai jamais eue. Tu étais avec moi quand j'ai eu mon BAC et tu es présente aujourd'hui pour ma soutenance. On a un lien particulier et on est capable de se comprendre en un regard. Merci d'avoir été là dans les moments de down pendant mes études comme dans les moments de joie. Merci de me réconforter quand je ne vais pas bien et de me rappeler que je dois me battre pour moi et pour être heureuse. Tu as été un soutien inconditionnel durant toutes ces études. Merci d'être cette personne qui me motive et m'inspire, mais qui sait me lancer des défis et réveiller mon côté compétiteur. On s'est soutenu dans les pires épreuves et je sais qu'on sera là l'une pour l'autre dans les meilleurs moments de nos vies.

À Christelle et Laurent Eckert, merci d'avoir été là, vous êtes également comme la famille. À toutes ces vacances qu'on a pu passer ensemble et à nos moments de joie, le samedi soir. J'espère fêter cette réussite avec vous également.

À mes beaux-parents, David et Anne-Cécile Renaud, merci de m'avoir accueillie comme si j'étais votre fille. Je sais que vous êtes heureux de me voir réussir aujourd'hui.

À mes parents, je sais que pour vous le monde de la recherche est quelque chose d'abstrait et que vous avez eu du mal à comprendre pourquoi je voulais faire une carrière dans cette filière, mais vous m'avez toujours soutenue. Merci de m'avoir toujours accompagnée et tirée vers le haut, malgré les désaccords que l'on pouvait avoir. Vous avez toujours fait en sorte que toute ma scolarité se passe dans les meilleures conditions possibles. Vous m'avez transmis des valeurs importantes qui sont devenues les piliers de ma personnalité. J'espère que vous êtes fiers de moi, parce que cette réussite aujourd'hui, c'est aussi la vôtre, je vous aime très fort.

À mes grands-parents, je sais que vous ne comprenez pas toujours, ce que je fais, mais, vous faites l'effort d'essayer de comprendre avec des reportages sur la biologie cellulaire. Merci de me soutenir et de m'encourager dans mes projets.

À mon étoile, mon Papy Ly, tu as toujours été un battant, tu t'es battu contre Parkinson pour que cette maladie ne te limite pas dans les choses du quotidien. Tu t'es battu jusqu'au bout pour remarcher et pour montrer à tes 2 petits enfants que tu te battais pour eux. Tu nous as laissés un grand vide ici, mais comme je te l'avais promis, on s'occupe bien de mamie. J'espère que d'en haut, tu es fier de moi, oui papy, je suis arrivée au bout de cette thèse !

À mon frère, Alexandre, tu es le meilleur cadeau que j'ai pu avoir. Tu caches ton côté sensible avec ton caractère, mais je sais que tu as un grand cœur. Même si c'est moi la grande sœur, tu as toujours ce côté protecteur pour moi. Avec le temps, tu apprendras qu'il faut choisir ces combats. Merci d'avoir été là pendant toute cette thèse, même si pour toi c'est un monde abstrait.

À Élise, reste cette personne aussi gentille, souriante et pleine de vie. Tu es discrète, mais, tu sais te faire respecter. Tu es la meilleure personne que j'aurais pu rêver pour mon frère. Merci d'avoir trouvé des solutions de mise en page beaucoup

plus simple que celle que mon frère proposait. Et oui, tu as ta place dans les remerciements parce que tu fais partie de la famille maintenant !

À Aurélien, ma moitié, mon fiancé et dans quelques mois mon mari. Merci d'être là à mes côtés depuis tant d'années. Notre rencontre a changé ma vie, tu es arrivé à une période où je ne croyais plus en moi, après ma première année de classe préparatoire. Sans toi, crois-moi, tout aurait été différent et je n'aurais pas tenu pendant ces années de thèse. Je te remercie d'avoir été un soutien indescriptible pendant ces années et d'avoir tout géré pendant que je rédigeais mon manuscrit de thèse (je te rendrai l'appareil quand tu rédigeras ton mémoire d'expert-comptable). Merci d'avoir compris et d'avoir agi dans l'ombre pour que j'avance dans ma rédaction, même tard le soir. Merci également d'avoir participé à mes répétitions de soutenance, alors que tu ne comprenais rien de ce que je t'expliquais. Tu as été l'épaule sur laquelle je me suis appuyée pendant ces années, tu as su me laisser pleurer pour mieux sécher mes larmes après, et, tu m'as montré qu'au bout du tunnel, il y avait toujours la lumière. Ton soutien indescriptible m'a permis d'être là aujourd'hui. La vie ne nous a pas fait de cadeau parmi lesquels je compte la perte consécutive de nos 2 étoiles (ta mamie Françoise et mon papy Ly) qui nous ont laissées un grand vide, mais, je sais qu'ensemble on est plus fort. Tu m'as aidé à tout surmonter et tu m'as suivi dans mes défis personnels les plus fous dont le semi-marathon de Paris en mars dernier. Tu as également su comprendre que j'avais besoin d'aller à la danse à Amiens, le samedi pour me libérer (même si je sais que tu aimerais qu'on soit plus libre de nos week-end). Tu as été là dans mes moments les plus sombres comme dans les plus joyeux. Je ne vais pas m'étaler plus, le reste tu le sais déjà, je t'aime.

Abstract

The liver's regenerative capacity depends on the plasticity of its epithelial cells. Particularly, the hepatocyte-to-cholangiocytes transdifferentiation process (HCT) has been described as a pro-regenerative mechanism in cholestatic animal models and is a feature of some end-stage human chronic liver diseases. During this process, hepatocytes lose some of their physiological functions, due to a decrease of activity of hepatocyte identity transcription factors and a reduction of chromatin accessibility at their promoters. However, how the cholangiocyte transcriptional program is activated during HCT is poorly defined. Here, we show that HCT does not involve transitioning through a progenitor-like cell state. Our results indicate that HCT involves induction of an archetypal monopolarized epithelial cell transcriptional program, initially repressed in healthy hepatocytes. Indeed, HCT requires the removal of the polycomb-dependant repressive epigenetic mark H3K27me3 to activate transcription factor encoding genes driving the cholangiocyte program. Especially, our study identified the transcription factor GRHL2 as a key player in this process, through direct induction of cholangiocyte/epithelial genes. In line, GRHL2 was detected in human hepatocytes undergoing HCT in human livers from patients with advanced chronic liver disease.

Keywords: Transdifferentiation, Epigenetics, Liver regeneration, Hepatocytes, Polycomb complex, GRHL2

Résumé

Les capacités régénératives du foie sont majoritairement permises par la plasticité de ses cellules épithéliales. En particulier, la transdifférenciation des hépatocytes en cholangiocytes (HCT) est décrite comme un mécanisme de régénération des hépatocytes dans les modèles murins de cholestase et qui semble être également présent dans les pathologies hépatiques de stades avancés. Pendant ce processus, les hépatocytes perdent une partie de leurs caractéristiques physiologiques. Cette perte est causée par une diminution de l'activité des facteurs de transcription hépatocytaires, et une réduction de l'accessibilité de la chromatine au niveau de leurs promoteurs. Cependant, la manière dont le programme transcriptionnel cholangiocyttaire est activé est peu décrit. Cette thèse montre que HCT n'implique pas le passage par un état de cellule progénitrice. Nos résultats indiquent que ce mécanisme passe par l'induction d'un programme transcriptionnel caractéristique d'une cellule épithéliale monopolarisée, réprimé dans un hépatocyte sain. En effet, HCT nécessite la levée de la marque épigénétique de répression H3K27me3, dépendante du complexe Polycomb2, afin d'activer les gènes codants pour les facteurs de transcription du programme cholangiocyttaire. En particulier, notre étude a identifié le facteur de transcription GRHL2 comme un acteur clé de ce processus, induisant de manière directe les gènes cholangiocytaires/épithéliaux. Par ailleurs, ce facteur a été détecté dans les hépatocytes humains en cours d'HCT chez des patients atteints de maladies hépatiques chroniques en phase terminale.

Mots clés : Transdifférenciation, Épigénétique, Régénération hépatique, Hépatocytes, Complexe Polycomb, GRHL2

Sommaire

Remerciement	1
Abstract	15
Résumé	17
Liste des abréviations	25
Liste des figures et des tableaux	31
I. Liste des figures	31
II. Liste des tableaux	33
Avant-propos	35
Introduction	37
I. Organisation structurelle du foie, fonctions et maladies hépatiques	37
1.1. Structure et fonctions physiologiques	37
a. Fonctions physiologiques	37
i. Fonction de détoxification	37
ii. Fonctions métaboliques	38
iii. Autres fonctions physiologiques	39
b. Organisation structurelle du foie	40
i. Lobes et lobules hépatiques	40
ii. Arbres et canaux biliaires	42
c. Caractéristiques des cellules parenchymateuses du foie	43
i. Les cellules épithéliales	43
ii. Autres types cellulaires du foie	45
1.2. Maladies hépatiques chroniques	47
a. Maladies hépatiques d'origine hépatocytaire	48
i. Les MALSDs (Maladies du foie stéatosique associées à un dysfonctionnement métabolique)	48
ii. Les ALDs (Maladies du foie liées à la consommation d'alcool)	54
b. Maladies hépatiques d'origine cholangiocytaire	58
i. Le syndrome d'Alagille (ALGS)	59
ii. La cholangiopathie biliaire primitive (PBC)	60
iii. La cholangiopathie sclérosante primitive (PSC)	62
II. Plasticité des cellules parenchymateuses et régénération hépatique	64
2.1. Origine de la plasticité des cellules parenchymateuses du foie	64
a. Retour aux origines des cellules épithéliales : différenciation à partir de l'hépatoblaste	64
i. Mise en place du bourgeon hépatique lors du développement embryonnaire	64

ii.	Différenciation des hépatoblastes en hépatocytes ou en cholangiocytes.....	66
b.	Plasticité des cellules épithéliales du foie	70
i.	Régénération hépatocytaire.....	70
ii.	Régénération cholangiocytaire	71
2.2.	Régénération cellulaire après une blessure hépatique chronique.....	73
a.	Implication des cellules progénitrices du foie (LPCs)	73
i.	Origine et caractéristiques des LPCs.....	73
ii.	Activation des LPCs en cas de blessure hépatique	75
b.	Plasticité cellulaire et transdifférenciation.....	77
III.	Contrôle transcriptionnel de la plasticité des cellules du parenchyme hépatique.....	80
3.1.	Épigénétique et facteurs de transcription (TFs) : généralités.....	80
a.	Généralités sur les mécanismes épigénétiques.....	80
i.	Structure de la chromatine	80
ii.	Modifications post-traductionnelles (PTMS) des histones	83
iii.	Le cas particulier du complexe Polycomb2 (PrC2)	85
b.	Généralités sur les TFs et régulation transcriptionnelle	87
i.	Fixation à la chromatine et activité des TFs.....	88
ii.	Autres modes de régulation de la transcription : les enhanceurs	90
3.2.	Contrôle transcriptionnel de la plasticité hépatocytaire après une blessure hépatique chronique	92
a.	Modulations épigénétiques et plasticité des cellules épithéliales du foie.....	92
i.	Rôle de l'épigénome dans la régénération hépatique en cas de maladies chroniques.....	92
ii.	Focus sur PrC2 et plasticité cellulaire.....	93
b.	Modulation de l'activité des TFs après la blessure hépatique	94
	Objectif de la thèse	97
	Matériels et méthodes	99
I.	Analyses bio-informatiques	99
1.1.	Ressources	99
a.	Origine et récupération de l'ensemble des données publiques	99
b.	Liste des données provenant de la littérature	100
c.	Outils utilisés lors de cette étude	105
1.2.	Analyse des données de transcriptomiques	108
a.	Analyse de données issues de « Bulk » RNA-Seq	108
b.	Comparaison des transcriptomes en utilisant l'analyse en composante principale (ACP)	109
c.	Analyse d'enrichissement des signatures géniques (GSEA).....	109
d.	Enrichissement et annotation fonctionnelle des gènes	110
i.	Utilisation de l'outil DAVID (Base de données pour l'annotation, la visualisation et la découverte intégrée)	110
ii.	Utilisation de Metascape.....	111
1.3.	Analyse des données de Single-cell RNA-Seq (scRNA-Seq)	111

a.	Analyse des données de scRNA-Seq provenant du « <i>Human Protein Atlas</i> »	111
b.	Analyse des données de single-nuclei RNA-Seq	112
c.	Analyse des données de scRNA-Seq utilisées pour la mise en place de l'atlas murin	112
1.4.	Analyse des données de ChIP-Seq (cistrome et épigénome)	114
a.	Normalisation et génération des signaux issus des données de ChIP-Seq	114
b.	Visualisation des signaux	115
c.	Séparation des gènes selon leurs profils épigénétiques grâce à l'utilisation de SPARK	115
1.5.	Statistiques	115
II.	Procédures expérimentales	116
2.1.	Culture cellulaire	116
a.	Entretien des lignées cellulaires	116
b.	Utilisation de modulateurs pharmacologiques	117
c.	Modulation transitoire de l'expression de <i>GRHL2</i>	117
d.	Culture tri-dimensionnelle (3D) des BMELs	118
e.	Analyse de l'O-GlcNAcylation de EZH2	118
2.2.	Analyses fonctionnelles	119
a.	Culture bactérienne et purification d'ADN plasmidique	119
b.	Extraction d'ARN et RT-qPCR	119
c.	Extraction et révélation des protéines	122
i.	Extraction des protéines totales	122
ii.	Fractionnement de la chromatine	122
iii.	Extrait nucléaire et immunoprécipitation	123
iv.	Simple Western Immunoassays (WES) et Western Blot	124
d.	Immunoprécipitation de la chromatine et séquençage haut débit (ChIP-Seq)	125
e.	Histologie, immuno-histologie (IHC) et immunofluorescence	126
i.	Histologie	126
ii.	Immuno-fluorescence	127
iii.	Marquage des organoïdes issus des BMELs	128
2.3.	Modèles animaux	128
a.	Injection hydrodynamique	128
b.	Surexpression de <i>GRHL2</i> dans les foies de souris par une méthode virale	129
2.4.	Échantillons humains	129
a.	Provenance des échantillons humains	129
b.	Organoïdes humains	130
Résultats	131
I.	L'HCT se produit sans passer par un état transcriptionnel proche d'un hépatoblaste	131
1.1.	Création d'un atlas de données de scRNA-Seq de foies murins pour la caractérisation du processus de l'HCT	131
i.	Mise en place de l'atlas de scRNA-Seq à partir des données de foies murins	131
ii.	Identification du processus d'HCT	132
1.2.	HCT : processus de reprogrammation des cellules matures	133

II. Induction d'une signature de cellule épithéliale mono-polarisée nécessaire à l'HCT, normalement réprimée dans des hépatocytes sains	136
2.1. Activation d'un nouveau programme épithélial lors de l'HCT	136
i. Expression d'une signature épithéliale pendant l'HCT	136
ii. Les hépatocytes, des cellules épithéliales particulières.....	138
2.2. Rôle du complexe PrC2 dans la répression de la signature épithéliale dans un hépatocyte sain	139
a. Caractérisation épigénétique des gènes codants pour les TFs induits lors de l'HCT.....	139
i. Ségrégation des TFs dérégulés pendant l'HCT	139
ii. Différents profils épigénétiques aux promoteurs des TFs induits	141
iii. Répression des promoteurs de C3 indépendante de H3K9me3	142
iv. Transposition des résultats chez l'Homme	144
b. Rôle du complexe PrC2 dans l'HCT	145
III. GRHL2 : un TF épithélial favorisant l'HCT	147
3.1. GRHL2 : régulateur de la différenciation des cellules épithéliales impliqué dans l'HCT	147
a. Identification d'un régulateur transcriptionnel de l'HCT	147
b. GRHL2 : TF actif dans les cholangiocytes et réprimé dans les hépatocytes	149
3.2. L'expression ectopique de GRHL2 entraîne l'activation de la signature épithéliale et le début de l'HCT	151
a. Expression ectopique de <i>GRHL2 in vitro</i> induit l'expression de gènes épithéliaux et la formation de cystes	151
i. Début de formation d'un canal dans les BMELs après l'induction de <i>GRHL2</i>	151
ii. Identification de cibles directes de GRHL2 induites pendant l'HCT.....	153
iii. Induction d'une signature génique cholangiocytaire	156
b. Expression ectopique de <i>GRHL2 in vivo</i> modifie la structure du parenchyme hépatique..	157
i. Modulation de l'expression de <i>GRHL2 in vivo</i> par différentes méthodes	157
ii. Modification du parenchyme hépatique 2 semaines après l'expression ectopique de <i>GRHL2</i> dans les foies de souris	160
3.3. Pertinence physio-pathologique de l'étude de GRHL2 chez des patients montrant des signes de HCT	162
i. Induction de <i>GRHL2</i> dans les foies de patients atteints d'ALD	162
ii. <i>GRHL2</i> dans les MASLDs	164
iii. Activation de <i>GRHL2</i> dans les cholangiopathies	165
Conclusion générale.....	167
Discussion.....	169
I. Impact de l'étude de l'HCT pour les maladies hépatiques	169
1.1. Rôle du micro-environnement dans le mécanisme de l'HCT	169
1.2. Opposition de nos résultats avec le dogme de l'action des LPCs en cas de blessure hépatique.....	170
1.3. Modification du réseau de TFs pendant l'HCT	171
1.4. Mécanismes épigénétiques potentiels sous-jacents à l'HCT.....	171

II. Stabilisation de EZH2 par l'O-GlcNAcylation aux promoteurs des gènes épithéliaux dans les hépatocytes sains	172
III. Mécanisme d'action de GRHL2 lors de l'HCT.....	174
3.1. Action des GRHL en homo-, di- ou en trimère	174
3.2. Partenaires d'interactions de GRHL2	175
3.3. Les différentes isoformes de GRHL2	176
Références.....	177
Annexes	215
I. Communications orales et affichées lors des congrès nationaux et internationaux dans le cadre de mon projet de thèse.....	215
1.1. Communication orale.....	215
1.2. Communication affichée	215
II. Financements obtenus	215
III. Article scientifique	215

Liste des abréviations

α -SMA : Protéine d'actine alpha 2

AAV8 : Virus adénoassocié recombinant de sérotype 8

ABCB11 : Transporteur hépatique d'acide biliaire

Acétyl-CoA : Acétyl co-enzyme A

ACMSD : Amino-carboxy-muconate semi-aldéhyde décarboxylase

ACP : Analyse en composante principale

ADH : Alcool déshydrogénase

ADN : Acide désoxyribonucléique

ADNc : Acide désoxyribonucléique complémentaire

AFP : Alpha feto-protéine

ALAT : Alanine aminotransférase

ALB : Albumine

ALD : Maladie du foie liée à la consommation d'alcool

ALDH : Aldéhyde déshydrogénase

ALGS : Syndrome d'Alagille

APAP : Acétaminophène

APOF : Apolipoprotéine F

ARID1A : Protéine A1 contenant un domaine d'interaction riche en AT

ARN : Acide ribonucléique

ARNm : Acide ribonucléique messenger

ARNt : Acide ribonucléique de transfert

ASAT : Aspartate aminotransférase

ATAC-Seq : Séquençage de la chromatine accessible par la transposase

ATP : Adénosine triphosphate

BdL : Ligation des canaux biliaires

BMEL : Cellule embryonnaire de foie bipotente

Bp : Paires de base d'ADN

BSA : Bovine-sérum albumine

CCl₄ : Tétrachlorure de carbone

CDE : Régime déficient en choline et supplémenté en méthionine

CDS : Régions codantes

CFTR : Régulateur de la conductance transmembranaire de la fibrose

ChIP-Seq : Séquençage de l'ADN immuno-précipité

ChREBP : Protéine de liaison aux hydrates de carbone

CLDN 4 : Claudine 4

CP450 : Cytochrome P450

CpG : 5'-C-phosphate-G-3'

CRM : Module de régulation *cis* de la transcription

CYP2E1 : Isoforme E1 du cytochrome P450

CYP3A4 : Cytochrome P450 3A4

DAVID : Base de données pour l'annotation, la visualisation et la découverte intégrée

DDC : 3,5-Dithoxylcarbonyl-1,4-Dihydrocolline

DE : Expression différentielle

DHS: Sites hypersensibles de la DNase I

DLK1 : Delta Ligand 1 de la voie non canonique de Notch

DMSO : Di-méthylsulfoxyde

DR : Réaction ductulaire

DTT : Dithiothréitol

EED : Facteur de développement de l'ectoderme embryonnaire

EGF : Facteur de croissance épidermal

EGFP : Protéine fluorescente verte

EHF : Facteur de transcription homologue ETS

EPCAM : Molécule d'adhésion des cellules épithéliales

ESRP1 : protéine 1 régulatrice de l'épissage épithélial

EZH1 : Enhancer de zeste 1 du complexe PrC2

EZH2 : Enhancer de zeste 2 du complexe PrC2

F5 : Facteur de coagulation V

FC : Rapport de variation d'un gène d'une condition à une autre

FDA : Agence du département de la santé des services Humains aux États-Unis

FDR : Taux de fausses découvertes

FGF21 : Facteur de croissance fibroblastique 21

FGF7 : Facteur de croissance fibroblastique 7

FMO5 : Flavine contenant la monooxygénase 5

FOXA : Facteur de transcription à boîte en motif à tête de fourche « Forkhead » A

FPKM : Fragments par kilobase de transcrit par million de fragments alignés

FXR : Récepteur farnésoïde X

G6P : Glucose-6-phosphatase

GATA4 : Facteur de transcription à doigt de zinc se fixant sur les motifs « GATA »

GRHL2 : Facteur de transcription de type Grainyhead 2

GSEA : Analyse d'enrichissement de signatures géniques

GSTA3 : Glutathion S-transférase alpha 3

H2A.Z : Variant de l'histone 2A

H3K27ac : Acétylation de la lysine 27 de l'histone 3

H3K27me3 : Tri-méthylation de la lysine 27 de l'histone 3

H3K36me3 : Tri-méthylation de la lysine 36 de l'histone 3

H3K4me1 : Mono-méthylation de la lysine 4 de l'histone

H3K4me3 : Tri-méthylation de la lysine 4 de l'histone 3

H3K9me3 : Tri-méthylation de la lysine 9 de l'histone 3

HBE : Cellule primaire de l'épithélium bronchique

HCC : Hépatocarcinome

HCT : Transdifférenciation des hépatocytes en cholangiocytes

Hep-ID : gènes codants pour les TFs de l'identité hépatocytaire

HepaRG : Cellules progénitrices humaines bipotentes

HepG2 : Cellules immortalisées humaines issues d'un hépatocarcinome

HGF : Facteur de croissance de l'hépatocyte

HGF-MET : Gène HGF-MNNG de transformation HOS

HNF : Facteur nucléaire hépatocytaire

HSC : Cellules stellaires du foie

IGB : Navigateur intégré du génome

IGFII : Facteur de croissance analogue à l'insuline II

IHC : Immuno-histologie

IL1 : Interleukine 1

IL1B : Interleukine 1 bêta

IL6 : Interleukine 6

IL8 : Interleukine 8

IMC : indice de masse corporelle

IP : Immunoprécipitation

iPSC : Cellules souches pluripotentes induites

ISWI : Complexe imitation switch

ITGA3 : Intégrine sous unité alpha 3

ITGB8 : Intégrine sous unité beta 8

JAG1 : Ligand «Jagged » 1 de la voie canonique Notch

JAGGED2 : Ligand «Jagged » 2 de la voie canonique Notch

JMJD3 : Protéine 3 contenant le domaine Jumonji

JUP : Jonction plakoglobine

KCs : Cellules kupffer

KDM2A : Lysine déméthylase 2A

KDM6B : Lysine déméthylase 6B

KO : Invalidation génique

KRT19 : Cytokératine 19

KRT7 : Cytokératine 7

LPC : Cellules progénitrices du foie

LPS : Lipopolysaccharides

LSECs : Cellules endothéliales sinusoïdales du foie

LY6G : Antigène lymphocytaire 6G

MAS : Dysfonction métabolique associée à la stéatose

MASH : Stéatohépatite associée à un dysfonctionnement métabolique

MASLD : Maladie du foie stéatosique associée à un dysfonction métabolique

MBOAT : Facteur de fixation membranaire contenant un domaine O-Acétyltransférase 1

MEC : Matrice extra-cellulaire

MEOS : Système microsomal d'oxydation de l'alcool

MetALD : Maladie métabolique et maladie du foie associé à l'alcool

mirARN : Micros-ARN

MMP : Métalloprotéase

MPH : Hépatocytes primaires murins

MW : Poids moléculaire

NAFLD : Maladie hépatique graisseuse non-alcoolique

NES : Score d'enrichissement normalisé

NOTCH2 : Récepteur 2 de la voie de signalisation Notch

NR1H4 : Récepteur nucléaire de la sous famille 1 groupe H

NSD1 : Enzymes de liaison aux récepteurs nucléaires SET à domaine protéique 1

nTPM : Nombre de transcrits normalisés par million

O-GlcNAcylation : O-linked N-acétylglucosamine

OGT : O-GlcNAc transférase

OMS : Organisation mondiale de la Santé

OSM : Oncostatine M

OVOL1/2 : Répresseur transcriptionnel des domaines Ovo 1 et 2

PBC : Cholangiopathie biliaire primitive

PBS : Phosphate-buffered saline

PCR : Réaction d'amplification par la polymérase

PFA : Paraformaldéhyde

PhX : Hépatectomie partielle

PIC : Cocktail d'inhibiteur de protéase

PNPLA3 : Phospholipase de type patatine contenant le gène 3

Pol II : ARN Polymérase II

PPAR : Récepteur nucléaire activé par les proliférateurs de peroxyosomes

PrC2 : Complexe Polycomb 2

PROX1 : Protéine de la famille Proxpero Homéo-Boxe 1

PSC : Cholangiopathie sclérosante primitive

PTM : Modification post-traductionnelle

PTPRC : Protéine tyrosine phosphatase de type C

qPCR : PCR quantitative

RAB25 : Membre 25 de la famille des oncogène RAS

RbAp46/48 : Protéine de fixation associée aux rétinoblastomes 46 et 48

RE : Réticulum endoplasmique

RIME : Analyses d'immunoprécipitation rapide pour l'étude de la protéine endogène par spectrométrie de masse

Rlog : Normalisation des données sous forme logarithmique régularisée et transformée

RNA-Pol II : ARN Polymérase II

RNA-Seq : Séquençage de l'ARN

ROS : Espèces réactives de l'oxygène

RT : Transcription inverse

RUNX1 : Facteur de transcription à domaine « Runt » 1

scRNA-Seq : Séquençage du transcriptome de cellule unique

SD : Écart-type

SDS : Sodium dodécyl sulfate

SDS-PAGE : Sodium Dodécyl-Sulfate-Polyacrylamide

siARN : petit ARN interférant

SNP : Polymorphisme nucléotidique

snRNA-Seq : méthode de séquençage de l'ARN des noyaux individuels des cellules

SOX9 : Facteur de transcription à boîte SYR 9

SREBP : Protéine de liaison de l'élément régulateur des stérols

srHC-Seq : Séquençage de l'hétérochromatine résistante à la sonication

SUZ12 : Suppresseur de l'homologue Zete 12

SVF : Sérum de veau fœtal

SWI/SNF : Complexe Switch mutant
Sucrose non-fermenté SNF

TAD : Domaine topologiquement associé

TBP : Protéine de fixation au niveau de la boîte TATA

TBS : Solution Salin-Tris

TBS-T : Triton X-100 dilué dans le TBS

TBX3 : Facteur de transcription T-Box 3

TEM : Transition épithélio-mésenchymateuse

TF : Facteur de transcription

TGF β : Facteur de croissance transformant bêta

THRB : Récepteur bêta aux hormones thyroïdiennes

TNF α : Facteur de nécrose tumoral alpha

TSS : Site d'initiation de la transcription

TWEAK : Voie de l'inducteur faible de l'apoptose liée au TNF

UMAP : Carte d'approximation et projection uniforme des variables

UTR : Régions non traduites

UTX : Protéine de répétition tétraricopeptide transcrite de manière ubiquitaire à partir du chromosome X

WES : « Simple Western Immunoassay »

WGBS : Séquençage au bisulfite du génome

WT : Souris non modifiées génétiquement

YAP : Protéine d'association au motif Yes

YFP : Protéine de fluorescence jaune

Liste des figures et des tableaux

I. Liste des figures

Figure 1 : Différentes fonctions physiologiques exercées par le foie.	40
Figure 2 : Organisation structurale du foie humain.....	41
Figure 3 : Schéma montrant l'organisation spatiale du lobule hépatique et la zonation des hépatocytes.	42
Figure 4 : Organisation structurale de l'arbre biliaire.....	43
Figure 5 : Schématisation de l'organisation spatiale des cholangiocytes.	45
Figure 6 : Organisation cellulaire d'un lobule hépatique.....	47
Figure 7 : Différents stades de la pathologie MALSD.	50
Figure 8 : Facteurs de risque pour le développement des MASLDs.....	53
Figure 9 : Facteurs de risque, comorbidités et développement des maladies du foie liées à la consommation d'alcool (ALDs).....	57
Figure 10 : Traits physiques caractéristiques et organes vitaux touchés par l'ALGS.	59
Figure 11 : Développement de la cholangiopathie biliaire primitive (PBC).	61
Figure 12 : Caractéristiques d'une cholangiopathie sclérosante primitive (PSC).	63
Figure 13 : Étape de la formation du bourgeon hépatique.	65
Figure 14 : Différenciation des hépatoblastes en hépatocytes ou en cholangiocytes.....	67
Figure 15 : Étape de formation du réseau biliaire.	68
Figure 16 : Dynamique des modifications épigénétiques lors de la différenciation des hépatocytes.....	69
Figure 17 : Schéma montrant les mécanismes de réaction ductulaire.	72
Figure 18 : Schéma montrant l'origine des cellules progénitrices du foie (LPCs).....	74
Figure 19 : Schéma montrant les modes d'activation et de différenciation des LPCs.	76
Figure 20 : Schéma montrant la transdifférenciation des hépatocytes vers les cholangiocytes.....	79
Figure 21 : Niveaux de condensation de la chromatine.....	81
Figure 22 : Relation fonctionnelle entre régulateurs de la transcription et chromatine.	85
Figure 23 : Composition du complexe PrC2.	86
Figure 24 : Fixation des TFs et initiation de la transcription.....	89
Figure 25 : Contexte épigénétique, recrutement des éléments régulateurs et transcription.....	91
Figure 26 : Réorganisation de la chromatine lors d'une blessure hépatique.	92
Figure 27 : Modification d'activité des TFs après une blessure hépatique.....	95
Figure 28 : Bilan du contexte et principaux axes du projet de thèse.....	98
Figure 29 : Les hépatoblastes, les hépatocytes et les cholangiocytes : trois populations cellulaires bien distinctes.....	132
Figure 30 : Identification de l'HCT dans les cellules du parenchyme hépatique de souris	133

Figure 31 : L'HCT ne passe pas par un état de progéniteur.	134
Figure 32 : Les gènes activés de manière intermédiaire pendant l'HCT ne sont pas caractéristiques des hépatoblastes.	135
Figure 33 : Activation d'une signature épithéliale mono-polarisée lors de l'HCT.	137
Figure 34 : Les hépatocytes, des cellules épithéliales avec un transcriptome particulier.	138
Figure 35 : L'HCT implique l'induction de gènes codants pour des TFs réprimés de manière épigénétique dans des hépatocytes sains.	140
Figure 36 : Les promoteurs du groupe C3 sont retrouvés à la fois dans des domaines d'hétérochromatine et d'euchromatine et leur régulation est indépendante de H3K9me3.	144
Figure 37 : Transposition des analyses épigénétiques sur nos modèles cellulaires et chez l'Homme.	145
Figure 38 : Répression des gènes codants pour les TFs induits pendant l'HCT par le complexe PrC2.	146
Figure 39 : L'HCT implique l'induction du facteur de transcription GRHL2, régulateur de la différenciation des cellules épithéliales.	148
Figure 40 : La surexpression de GRHL2 dans des hépatocytes déclenche l'induction de la signature épithéliale impliquée dans l'HCT.	152
Figure 41 : Identification des gènes directement régulés par GRHL2.	154
Figure 42 : Caractérisation des changements transcriptionnels induits par GRHL2 dans la lignée HepaRG et dans des organoïdes humains.	156
Figure 43 : Stratégie de surexpression de GRHL2 in vivo.	158
Figure 44 : La surexpression de GRHL2 dans les foies de souris modifie la structure du parenchyme hépatique.	161
Figure 45 : Les hépatocytes de patients atteints de maladies hépatiques chroniques en phase terminale de type ALD expriment GRHL2 et les gènes épithéliaux associés à l'HCT.	163
Figure 46 : GRHL2 présent dans des hépatocytes en cours d'HCT chez les patients atteints d'ALD.	164
Figure 47 : Hépatocytes exprimant GRHL2 chez des patients atteints de MASLDs grâce aux données de snRNA-Seq.	165
Figure 48 : Expression de GRHL2 dans des foies de patients atteints de cholangiopathies.	166
Figure 49 : Schéma bilan des événements transcriptionnels ayant lieu pendant l'HCT.	168
Figure 50 : Stabilisation de EZH2 par son l'O-GlcNAcylation.	173
Figure 51 : Expression ectopique de l'isoforme courte de GRHL2.	176

II. Liste des tableaux

Tableau 1 : Focus sur la localisation et l'effet sur la transcription de certaines modifications post-traductionnelles (PTMs) des histones.	84
Tableau 2 : Liste des données publiques utilisées dans cette thèse.....	104
Tableau 3 : Liste des outils bio-informatiques utilisés pour cette thèse.	107
Tableau 4 : Amorces utilisées en RT-qPCR.....	121
Tableau 5 : Liste des anticorps utilisés.	124

Avant-propos

Le foie est un organe majeur pour l'homéostasie énergétique, métabolique ainsi que pour la détoxification de l'organisme. Les maladies du foie, notamment les atteintes chroniques, sont un enjeu majeur de santé publique. En effet, la sédentarité, la suralimentation et la consommation excessive d'alcool sont à l'origine de nombreuses maladies hépatiques. Elles constituent un fardeau lourd pour la santé mondiale (Younossi et al., 2019). Ces maladies ouvrent la voie à des lésions hépatiques graves où la fibrose contribue à précipiter les dysfonctionnements hépatiques (Schwabe et al., 2020).

Le foie se caractérise par sa capacité importante de régénération, possible notamment grâce à la plasticité de ses cellules épithéliales. En effet, les hépatocytes et les cholangiocytes, issus tous les deux du même progéniteur appelé l'hépatoblaste (Tachmatzidi et al., 2021) (W. Li et al., 2020), sont capables de restaurer un pool suffisant de cellules fonctionnelles suite à une agression du foie. Ces cellules épithéliales correspondent à la grande majorité de la masse cellulaire du foie (60 à 80% de la masse du foie sain) (Slim et al., 2014) et sont les principales effectrices des activités de cet organe. La plasticité des cellules épithéliales permet à ces cellules quiescentes de se différencier et de ré-entrer dans le cycle cellulaire en réponse à une blessure (W. Li et al., 2020). Un autre mécanisme de régénération possible, lorsque la prolifération n'est plus suffisante pour compenser la perte de fonction hépatique, est la transdifférenciation. En particulier, en cas de défaillance du système biliaire, les hépatocytes peuvent se transdifférencier en cholangiocytes, et former des canaux biliaires fonctionnels (Schaub et al., 2018).

L'équipe dans laquelle j'ai réalisé mon travail a participé à proposer un concept selon lequel les lésions hépatiques, notamment chroniques, sont liées à une régénération inadaptée du foie (Dubois, Gheeraert, et al., 2020). Effectivement, lors des maladies chroniques, les dommages cellulaires sont plus importants et s'accompagnent d'un environnement inflammatoire persistant qui entraîne une déficience de la régénération des cellules touchées, essentielle à la préservation de la fonction hépatique. Dans ce contexte de souffrance cellulaire, des processus

alternatifs de régénération vont se mettre en place tels que la transdifférenciation entre les cellules épithéliales.

Mes travaux de thèse ont visé à caractériser au niveau moléculaire le processus de transdifférenciation des hépatocytes vers les cholangiocytes. En particulier, ils décrivent les mécanismes de régulation transcriptionnelle impliqués dans ce processus, en mettant notamment en évidence, la manière dont l'épigénome contrôle la plasticité des hépatocytes dans des conditions physiopathologiques. Ces travaux ont permis d'identifier le rôle du complexe répresseur Polycomb et du facteur de transcription « Grainyhead like transcription factor 2 » (*GRHL2*) dans cette reprogrammation hépatocytaire. Ces résultats ont été valorisés sous forme d'un article original, publié en tant que 1^{ère} auteure dans *PloS Biology*. Durant ma thèse, j'ai également présenté mes travaux à travers des communications orales et affichées lors de congrès nationaux et internationaux. Enfin, en tant que porteuse de projet, j'ai rédigé et soumis des demandes de financements ayant conduit à l'attribution d'une bourse de voyage par la société internationale de bio-informatique (« *International Society for Computational Biology* », ISCB) ainsi que d'un soutien de fin de thèse par la Fondation pour la Recherche Médicale (FRM).

Introduction

I. Organisation structurelle du foie, fonctions et maladies hépatiques

1.1. Structure et fonctions physiologiques

a. Fonctions physiologiques

Le foie est un organe impliqué dans de nombreuses fonctions physiologiques comme la dégradation des composés toxiques pour l'organisme, l'homéostasie énergétique, la sécrétion de facteurs de coagulation et la participation à certaines réponses immunitaires (**Figure 1**). Cette partie se concentrera toutefois principalement sur les fonctions de détoxification et de métabolisation des petites molécules (glucose et lipides), dont la dérégulation peut être la cause ^{et/ou} la conséquence de certaines pathologies (Kalra et al., 2025).

i. Fonction de détoxification

Le foie est l'organe qui permet la détoxification des composés issus des voies métaboliques (détoxification dite « endogène ») ainsi que des composés xénobiotiques (détoxification dite « exogène »). Cette fonction est assurée par les hépatocytes (dont la structure est détaillée dans le chapitre **Introduction 1.1.c.i p.43**)

D'un côté, la **détoxification endogène** contribue notamment à la **captation de l'ammonium**, issu du catabolisme des acides aminés de différents tissus, pour la formation de l'urée. Un taux élevé d'ammonium dans le sang peut indiquer une infection hépatique chronique et entraîner des confusions lorsqu'il s'accumule dans le cerveau. La **transformation de l'ammonium en urée** est réalisée via le **cycle de l'ornithine** dans le cytosol et dans la mitochondrie des hépatocytes périportaux (localisation détaillée dans le chapitre **Introduction 1.1.b.i p.40**) (Liska, 1998).

Le foie permet également l'**élimination de la bilirubine**, pigment donnant la couleur jaune à la bile. Cette bilirubine provient du catabolisme final de l'hème des vieux globules rouges et est présente dans l'organisme sous 2 formes : une forme conjuguée avec l'albumine, hydrosoluble et non toxique ; une forme libre, liposoluble

qui peut être neurotoxique si elle s'accumule dans le plasma. La **forme libre** est ainsi captée par les hépatocytes pour être **glycuroconjugée** (grâce à la fixation d'une molécule d'acide glucuronique), **solubilisée et éliminée via les canaux biliaires** (Okólicsányi & Magnenat, 1970).

D'un autre côté, les hépatocytes assurent également la **détoxification des xénobiotiques comme l'alcool ou les médicaments**. Ces molécules vont être, d'abord, oxydées par des enzymes notamment par le cytochrome P450 (CP450) pour devenir hydrophiles. Puis, d'autres enzymes catalysent des réactions de conjugaisons, pour permettre leur élimination par la bile (Apte & Krishnamurthy, 2011).

ii. Fonctions métaboliques

Les hépatocytes ont également un rôle dans le **métabolisme des acides aminés**, issus de la dégradation des protéines. Ces acides aminés vont être pris en charge par les **transaminases hépatiques** comme l'alanine aminotransférase (ALAT) et l'aspartate aminotransférase (ASAT) pour former de nouveaux acides aminés comme l'Alanine ou le L-glutamate. Les acides aminés sont dégradés pour **produire l'ammoniaque**, sont utilisés comme **substrats énergétiques** via les voies de néoglucogenèse et glycogénogenèse, et enfin, peuvent **stimuler la production d'insuline**. Cependant, ils sont **toxiques** pour les **hépatocytes** si leurs taux est trop élevé, ce qui entraîne la mort cellulaire et le relargage de l'ALAT et l'ASAT dans le sang. Ainsi, des taux élevés d'ALAT et d'ASAT dans le sang reflètent la mort hépatocytaire, traduisant une sur-métabolisation des acides aminés dans les voies métaboliques (Paulusma et al., 2022).

De plus, le foie est impliqué dans **l'homéostasie énergétique**. D'une part, il permet le **maintien de la glycémie** : les hépatocytes peuvent **stocker les glucides sous la forme de glycogène et les utiliser en fonction des besoins de l'organisme** (Han et al., 2016). En effet, en période de jeûne, l'augmentation du taux de glucagon et la diminution de celui d'insuline entraînent la dégradation du glycogène hépatique et la libération des stocks de glucose. Au contraire, en période de digestion, le glucose est capté par les hépatocytes et utilisé pour produire de l'énergie sous forme

d'adénosine triphosphate (ATP) via la glycolyse et le cycle de Krebs. Les hépatocytes sont également capables de produire du glucose via la néoglucogenèse à partir de lactate et de glycérol (X. Liu et al., 2017).

D'autres part, les hépatocytes sont l'un des lieux **de stockage des lipides**, produits par différentes voies métaboliques et utilisés comme **réserve énergétique** et substrat pour la **synthèse du cholestérol membranaire**. Ces lipides sont **stockés** dans les hépatocytes sous forme de **gouttelettes lipidiques**, micelles contenant les triglycérides, le cholestérol et des vitamines (X. Liu et al., 2017). En période de jeûne, ces réserves sont mobilisées, libérant des acides gras et le glycérol pour répondre aux besoins de l'organisme.

iii. Autres fonctions physiologiques

Par ailleurs, lors d'une réponse aiguë à une blessure hépatique, une réaction inflammatoire locale a lieu. Les hépatocytes vont **produire et sécréter** différents types de **molécules pro-inflammatoires** dont les cytokines. Ces médiateurs seront **détectés par les cellules immunitaires**, comme les cellules Kupffer (KCs) et les cellules immunitaires circulantes afin de moduler la réponse inflammatoire (J. Gong et al., 2023). Grâce à cette sécrétion, le foie peut être considéré comme ayant un rôle central dans l'homéostasie du système immunitaire : il peut orchestrer les premières phases de la réponse immunitaire en créant un microenvironnement favorable à l'activation des lymphocytes (Lau & Thomson, 2003).

Enfin, le foie a un rôle central dans le **processus de coagulation** qui repose sur un **équilibre de sécrétion de facteurs anti- et pro-coagulants**. Les hépatocytes sécrètent de nombreuses protéines comme le fibrinogène et le facteur V, nécessaires aux différentes phases de l'hémostase (hémostase primaire, coagulation plasmatique et fibrinolyse). Les hépatocytes participent également à **l'élimination des protéines inhibitrices de l'hémostase**. Cependant, cet équilibre est précaire et est facilement perturbé en cas de blessure hépatique, notamment chez les patients atteints de cirrhose (Sinegre & Lebreton, 2014).

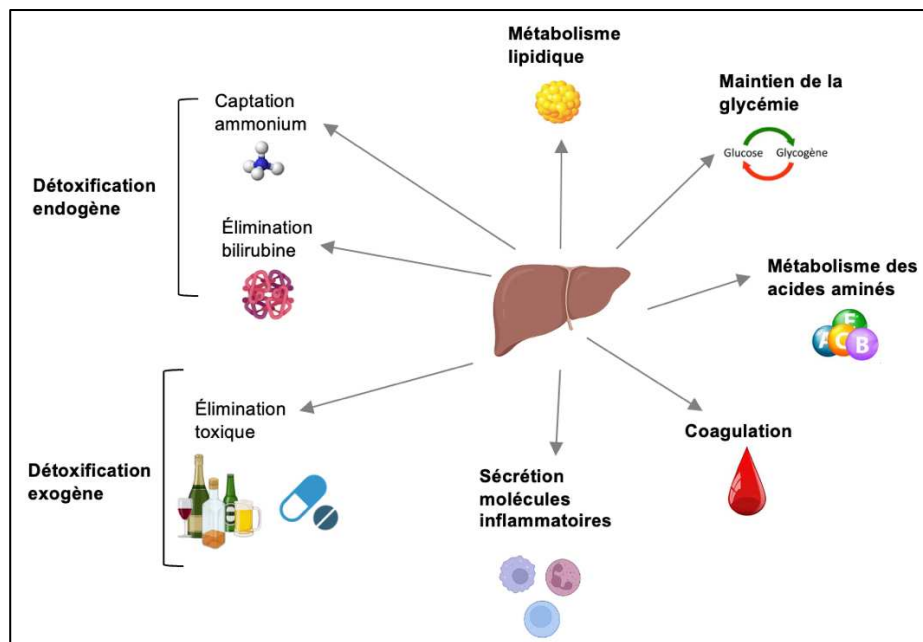


Figure 1 : Différentes fonctions physiologiques exercées par le foie.

b. Organisation structurale du foie

Les différentes fonctions hépatiques sont rendues possibles par une organisation structurale macro- et microscopique des différents types cellulaires du foie, maximisant les différents échanges moléculaires et la communication intra- et inter-organes.

i. Lobes et lobules hépatiques

Chez l'Homme, le foie est composé de 4 lobes : les **lobes droit et gauche**, séparés par le ligament suspenseur, et les **lobes carré et caudal**, séparés par le système vasculaire au niveau du hile du foie (**Figure 2**). Cette zone est à la fois traversée par l'artère hépatique, la veine porte et le canal biliaire relié à la vésicule biliaire.

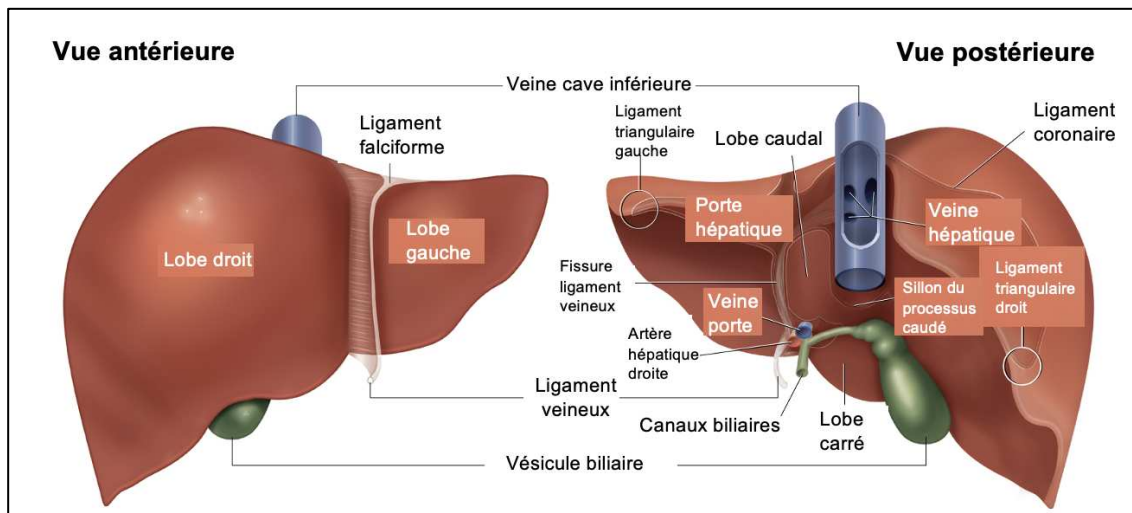


Figure 2 : Organisation structurelle du foie humain.

Détail de l'organisation structurelle du foie montrant les veines, les ligaments et les artères. Adaptée de (Cours US Abdomen - Foie, *n.d.*)

Au niveau microscopique, le foie est composé de **plusieurs unités fonctionnelles** reliées par la veine centrale et composant le parenchyme hépatique. Les **lobules hépatiques** sont organisés en structures fonctionnelles hexagonales, accolées les unes aux autres avec une triade portale à leurs extrémités (Malarkey et al., 2005). Les **triades portales** sont composées d'une sinusoïde hépatique, d'une partie de la veine porte et du canal biliaire (**Figure 3A**). Les sinusoides hépatiques, quant à eux, sont des capillaires bordés des cellules endothéliales sinusoidales du foie (« *Liver sinusoidal endothelial cells* », LSECs) qui vont acheminer le sang vers les hépatocytes pour permettre le passage des grosses molécules (Braet & Wisse, 2002).

Ces sinusoides sont classiquement **divisées en plusieurs zones fonctionnelles** : la position des hépatocytes dans l'axe triade portale par rapport à la veine centrale définit la « **zonation** ». En effet, leur positionnement détermine l'exposition des hépatocytes à l'oxygène et aux nutriments (**Figure 3B**) (Paris & Henderson, 2022). Les hépatocytes proches de la veine porte sont plus exposés aux nutriments et à l'oxygène, ils seront notamment engagés dans la glycogénèse, la néoglucogénèse et l'oxydation des acides gras. Au contraire, les hépatocytes situés autour de la veine centrale, exposés à du sang appauvri en oxygène et en nutriments, seront impliqués dans le métabolisme des xénobiotiques, la cétogénèse et les processus de glycolyse.

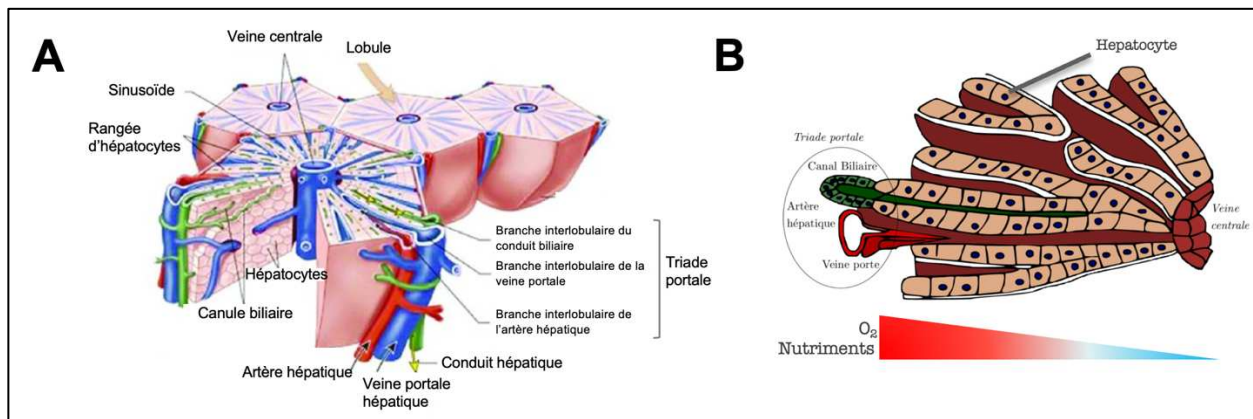


Figure 3 : Schéma montrant l'organisation spatiale du lobule hépatique et la zonation des hépatocytes.

(A) Organisation spatiale des lobules hépatiques, adaptée de (Bricks, 2014). (B) Organisation spatiale des hépatocytes et zonation dans le lobule hépatique, adaptée de (Paris & Henderson, 2022) et (Ashraf et al., 2019)

ii. Arbres et canaux biliaires

Les canaux biliaires correspondent à une série de **différents conduits plus ou moins fins** qui relient le foie, la vésicule biliaire et le duodénum. Ils drainent la bile, produit de sécrétion des hépatocytes, collectée par les canaux biliaires au niveau du canal de Hering (**Figure 5**) (Nagy et al., 2020). L'arbre biliaire est un réseau complexe composé de **canaux interlobulaires et extra-hépatiques** (**Figure 4**). Le réseau biliaire est également très **proche du réseau de capillaires**, ce qui permet une irrigation des cellules épithéliales le composant, ainsi que le **maintien de leurs fonctions et de leurs intégrités** (Strazzabosco & Fabris, 2008). Le canal biliaire a différentes fonctions selon sa position dans l'arbre biliaire, allant du **transport de la bile** à la **modification de sa composition et de sa viscosité**. Par exemple, le canal extra-hépatique est plus large que le canal interlobulaire et permet une **modification de la composition de la bile** plus importante (de Jong et al., 2021).

Du fait de son hétérogénéité et de la complexité de formation du réseau biliaire (détaillé dans le chapitre **Introduction 2.1.a.ii p.66**), de nombreuses malformations peuvent apparaître, causant sur le long terme des problèmes d'évacuation de la bile.

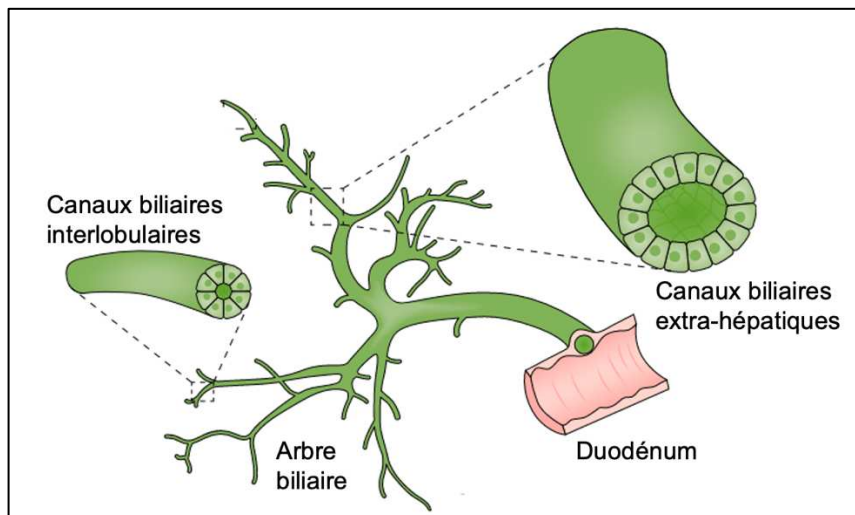


Figure 4 : Organisation structurale de l'arbre biliaire.

Détail de l'organisation macroscopique des canaux biliaires. L'arbre biliaire est un réseau complexe et hétérogène composé de canaux avec différents diamètres. Ces canaux permettant la modification de la composition chimique de la bile. Adaptée de (Banales et al., 2019)

c. Caractéristiques des cellules parenchymateuses du foie

La diversité des fonctions assurées par le foie repose sur une hétérogénéité des types cellulaires (**Figure 6**). Ces cellules peuvent être regroupées en grandes familles que sont : les cellules épithéliales, endothéliales, immunitaires et les fibroblastes. Les analyses récentes d'étude du transcriptome des cellules uniques (« *Single-cell RNA-sequencing* », scRNA-Seq) suggèrent que les échanges intercellulaires jouent un rôle important dans les pathologies hépatiques en influençant de manière locale la plasticité cellulaire (MacParland et al., 2018). En effet, en cas de blessure hépatique, l'action de toutes ces cellules crée un micro-environnement inflammatoire conduisant à des dysfonctionnements hépatiques (Shu et al., 2022).

i. Les cellules épithéliales

Les hépatocytes sont les cellules majoritaires d'un foie mature. Une fraction des hépatocytes est caractérisée par une polyploïdie. Ces cellules possèdent un cytoplasme avec un grand nombre d'organelles, permettant l'activité métabolique de la cellule. Ils contribuent, notamment, au maintien de l'homéostasie énergétique en participant aux fonctions de métabolisation du glucose et des lipides.

Les hépatocytes sont **organisés en plusieurs couches** et possèdent **plusieurs pôles basaux et apicaux afin de maximiser les échanges et l'absorption des nutriments dans le sang** (Slim et al., 2014)(Gissen & Arias, 2015). Ces échanges sont rendus possibles par la proximité des sinusoides hépatiques et de l'espace de Disse, améliorant les échanges des macromolécules (notion de zonation développée dans le chapitre **Introduction 1.1.b.i p.40**).

A contrario, les cholangiocytes ont des caractéristiques de **cellules épithéliales plus classiques**. Ils possèdent un **seul pôle basal/apical** et sont **organisés en monocouches**, reliées entre elles grâce aux jonctions serrées (Buckley & St Johnston, 2022). Les cholangiocytes sont une **population très hétérogène et dynamique**. Ils acquièrent différents niveaux de différenciation selon leurs positions dans l'arbre biliaire (détaillé dans le chapitre **Introduction 1.1.b.ii p.42**). Dans la littérature, il existe différentes nomenclatures pour caractériser les cholangiocytes proches du canal de Hering. D'un côté, certains groupes expliquent que ces cholangiocytes sont qualifiés d'« immatures » car ils ne sont pas totalement différenciés, mais ils présentent un profil génique cholangiocytaire (Banales et al., 2019). D'un autre côté, d'autres groupes expliquent que les cellules du canal de Hering sont des cellules progénitrices du foie (« *Liver progenitor cells* », LPCs), présentes dans un foie adulte et qui peuvent se différencier en hépatocytes ou en cholangiocytes en cas de blessure hépatique (détaillé dans le chapitre **Introduction 2.2.a p.73**) (So et al., 2020).

Ensuite, tout au long de l'arbre biliaire, on retrouvera des cholangiocytes dits « petits » ou « grands » selon le diamètre du canal biliaire et selon leurs caractéristiques (Lou et al., 2025). En effet, les « petits » **cholangiocytes sécrètent des ions pour la solubilisation de la bile** mais ne peuvent pas en modifier sa composition. En revanche, les « grands » cholangiocytes ont une capacité d'absorption et de **modification chimique de la bile (Figure 5)**, et peuvent **réguler le flux biliaire** en contrôlant notamment le volume de la bile. (Hrncir & Gracz, 2023) (Banales et al., 2019). Cette capacité est modulée par divers acteurs dont les acides biliaires (Tabibian et al., 2013). La différence de fonction entre les « petits » et « grands » cholangiocytes s'explique par l'expression progressive de gènes codants

pour des récepteurs de transport des ions, et a pour conséquence l'augmentation du contenu cytoplasmique le long de l'arbre biliaire.

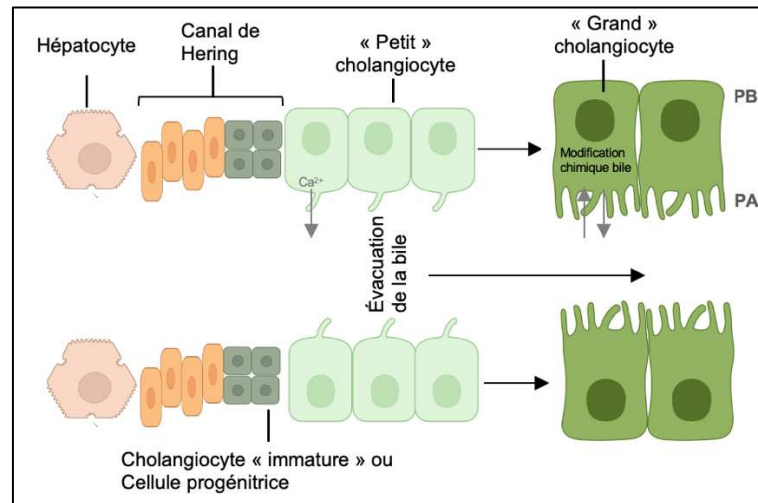


Figure 5 : Schématisation de l'organisation spatiale des cholangiocytes.

Organisation spatiale des cholangiocytes le long de l'arbre biliaire. Les cholangiocytes possèdent différents niveaux de différenciation selon leurs positions dans l'arbre biliaire. D'un côté, les « petits » cholangiocytes sécrètent des ions pour solubiliser la bile. De l'autre, les « grands » cholangiocytes modifient la composition chimique de la bile et peuvent réguler le flux biliaire. Les flèches grises indiquent le sens de l'échange entre la bile et les cholangiocytes. PA indique le pôle apical et PB le pôle basal.

ii. Autres types cellulaires du foie

Du fait de sa vascularisation importante, le foie possède de nombreuses **cellules endothéliales** (LSECs) qui tapissent les sinusoides hépatiques. Ces cellules font le lien entre les traversées d'hépatocytes, l'artère hépatique et la veine porte, permettant d'irriguer l'entière des lobules hépatiques. Les LSECs sont donc à l'interface entre, d'un côté, le sang et les hépatocytes, et de l'autre les cellules stellaires hépatiques (« *Hepatic stellate cells* », HSCs) (Braet & Wisse, 2002). Les LSECs participent à la régénération du foie en **interagissant avec les plaquettes et les cellules inflammatoires** (Poisson et al., 2017), et jouent, également, un rôle dans le maintien de la quiescence des HSCs (DeLeve et al., 2008).

Les **HSCs** sont les **principaux fibroblastes du foie** qui se situent dans l'espace de Disse. Elles sont nommées « cellules étoilées du foie », car leurs extensions cytoplasmiques peuvent couvrir une grande distance dans cet espace

périsinusoïdal. Elles sont **généralement dans un état « quiescent »**, mais, peuvent devenir actives en cas de blessure hépatique. À l'état « quiescent », les HSCs participent au métabolisme de la vitamine A, au maintien de la matrice extracellulaire (MEC) (Puche et al., 2013), et jouent également un rôle important dans l'homéostasie cellulaire du foie (Hansen et al., 2025). En effet, ces cellules sont des **acteurs clés de la niche cellulaire hépatique**, agissant en étroite collaboration avec les hépatocytes. Des études de scRNA-Seq montrent que certaines fonctions des hépatocytes comme la détoxification exogène sont régulées par la présence de ces HSCs « quiescentes ». Ces études soulignent la complexité des échanges multicellulaires nécessaires au maintien de la fonction hépatique (Sugimoto et al., 2025).

Lors de leur **activation**, les HSCs acquièrent de nouvelles propriétés comme la **prolifération et la contraction cellulaire**. L'activation transitoire des HSCs permet de libérer des facteurs de croissances comme le facteur de croissance épidermal (« *Epidermal Growth Factor* », EGF) ou le facteur de croissance de l'hépatocyte (« *Hepatocyte Growth Factor* », HGF) qui participent à la régénération du foie (Friedman, 2008). Les HSCs peuvent s'activer et **adopter plusieurs phénotypes** : un sous-type **myofibroblastique** bien identifié et un sous-type **inflammatoire** encore peu décrit (Geng & Schwabe, 2025). Les HSCs myofibroblastiques produisent et **sécrètent du collagène cicatriciel afin de réorganiser la MEC** en cas de blessure pour activer la réparation tissulaire (R. Schuster et al., 2023). Cependant, une **activation non contrôlée** des HSCs peut conduire à une **fibrose**. Dans les pathologies hépatiques chroniques, le réseau d'échange est perturbé. La suractivation des HSCs conduit, de ce fait, à une désensibilisation des hépatocytes aux signaux de l'homéostasie hépatique, à une augmentation de l'inflammation et *in fine* aggrave l'état hépatique (Bendixen et al., 2024).

Enfin, pour faire face aux agressions, le foie abrite des **cellules immunitaires** dont les KCs, **macrophages résidents du foie**, situés dans les sinusoides hépatiques et ancrés dans l'endothélium sinusoïdal. Ces cellules ont, notamment, un rôle de défense contre les pathogènes qui proviennent du sang et permettent **l'élimination des cellules âgées ou endommagées**. Elles jouent également un rôle dans la dégradation de certaines molécules, potentiellement immunogènes issues de la métabolisation hépatique (Nguyen-Lefebvre & Horuzsko, 2016). Cependant, une

activation chronique des KCs peut être délétère et peut aggraver l'état hépatique dans certaines pathologies (P. Li et al., 2017).

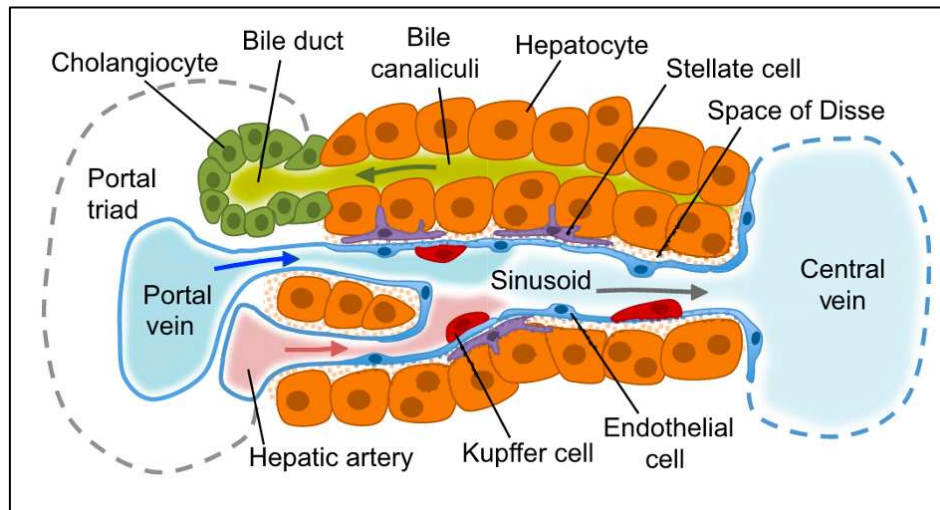


Figure 6 : Organisation cellulaire d'un lobule hépatique.

Les hépatocytes (en orange) sont organisés en sinusoides bordés des cellules endothéliales (en bleu). L'espace de Disse est compris entre les travées d'hépatocytes et les cellules endothéliales. Les cellules stellaires (en violet) se situent dans l'espace de Disse. Les cellules kupffer (en rouge) résident dans les sinusoides d'hépatocytes. Les hépatocytes sécrètent les sels biliaires dont le contenu s'écroulera dans les canaux biliaires, bordés de cholangiocytes (en vert). Le sang s'écoulera suivant les flèches : de la veine porte à la veine centrale. D'après (Gordillo et al., 2015).

1.2. Maladies hépatiques chroniques

Les pathologies hépatiques ont des étiologies multiples et variées, et peuvent toucher différentes régions du foie ainsi que différents types cellulaires. Ces pathologies sont réparties en 2 grandes catégories : les maladies hépatiques aiguës et chroniques. Dans les deux cas, les pathologies causent des dommages cellulaires qui nécessitent une régénération du foie plus ou moins importante afin de préserver les fonctions hépatiques. Cette partie se concentrera sur les principales pathologies hépatiques chroniques d'origine hépatocytaire ou cholangiocytaire.

a. Maladies hépatiques d'origine hépatocytaire

- i. Les MASLDs (Maladies du foie stéatosique associées à un dysfonctionnement métabolique)

Étiologies des MASLDs :

La maladie du foie stéatosique associée à un dysfonctionnement métabolique (« *metabolic dysfunction-associated liver disease* », MASLD) est le terme adopté en 2022 pour remplacer la maladie hépatique grasseuse non-alcoolique (« *non-alcoholic liver disease* », NAFLD) afin de mieux décrire la pathologie (Eslam et al., 2020). C'est une des maladies les plus courantes, avec une prévalence qui augmente au fil des années, en lien notamment avec une **augmentation de l'obésité liée à la sédentarité couplée à la malnutrition** (régime alliant sucre et graisse) (Younossi et al., 2019).

Différents stades de la MASLD :

La MASLD englobe un spectre de gravité allant de la **simple stéatose hépatique** jusqu'au développement de la **stéatohépatite** puis de la **fibrose** et la **cirrhose** qui s'accompagnent d'une inflammation lobulaire (Rotman & Kapuria, 2020) (Figure 7).

Initialement, en réponse à un excès d'apports nutritionnels, le foie présente une **stéatohépatite caractérisée par une accumulation des micelles et de graisse** dans le cytoplasme d'une partie des hépatocytes (Grzych et al., 2023). Cette stéatose résulte d'une **captation des acides gras circulants**, relargués par le tissu adipeux, et d'une **lipogenèse excessive**. Cet excès de lipides et d'acides gras dans le foie alimente une boucle de synthèse *de novo* de ces molécules, régulée par les facteurs de transcription insulino-dépendants, « *Carbohydrate-responsive element-binding protein* » (ChREBP) et « *Sterol regulatory element-binding protein* » (SREBP). Or, les patients atteints de MASLD présentent, généralement, une insulino-résistance, ce qui exacerbe la synthèse lipidique, et donc, induit un stockage plus important des graisses (Kawamura et al., 2022) (Dentin et al., 2006).

Au stade de la stéatohépatite associée à un dysfonctionnement métabolique (« *metabolic dysfunction-associated steatohepatitis* », MASH), l'atteinte hépatique est plus marquée et est liée à une **inflammation** qui se manifeste par une **ballonisation des hépatocytes**. Cette modification morphologique des hépatocytes est le reflet de leurs souffrances, et est associée à une perte de polarité, un volume augmenté, une accumulation de lipides dans un cytoplasme gonflé ainsi qu'une **agrégation des protéines ubiquitinylées** dans des corps de Mallory. Ces derniers sont des **marqueurs précoces de stress cellulaire dans la MASH** et participent à l'activation des HSCs initiant la réparation tissulaire (Kisseleva et al., 2025). De plus, cette inflammation chronique entraîne l'**accumulation d'espèces réactives de l'oxygène** (« *reactive oxygen species* », ROS) dans les hépatocytes et la **présence de cellules inflammatoires** dans les travées d'hépatocytes (Bai et al., 2025).

La stéatohépatite peut ensuite évoluer en **fibrose**, état de **cicatrisation chronique du tissu**. Le foie va alors présenter un **excès de MEC cicatriciel** résultant de la suractivation des HSCs en réponse aux lésions chroniques, et d'une dégradation insuffisante de cette MEC nouvellement produite (Deprince et al., 2020). Afin de compenser la perte de fonction liée à l'évolution de la maladie, le foie met en place des mécanismes de régénération comme la sécrétion de cytokines par les cellules inflammatoires pour permettre la réparation tissulaire (S. Schuster et al., 2018). Cependant, la présence d'une **inflammation chronique** reste délétère pour les cellules et entraîne un **défait de régénération** cellulaire, dit « régénération pathologique », altérant les fonctions physiologiques du foie (mécanisme détaillé dans le chapitre **Introduction 2.2 p.73**).

Enfin, l'inflammation chronique du foie peut aboutir à une **fibrose avancée, la cirrhose**, qui représente un **stade irréversible** de la maladie où la majorité du tissu est décrit comme « non-fonctionnel ». En réponse à la nécrose hépatique, le foie cirrhotique présente des **nodules de régénération**, correspondant à une prolifération d'hépatocytes dysfonctionnels, situés à proximité des bandes de fibrose (Ginès et al., 2021). En lien avec l'augmentation de la nécrose et de la stéatose, le foie est également caractérisé par une **augmentation de la sénescence des hépatocytes**, présentant un arrêt irréversible de la prolifération cellulaire. Ces hépatocytes sénescents entraînent des changements dans le métabolisme du glucose et des

lipides, compromettant davantage la fonction hépatique (Bonnet et al., 2022). Ces altérations **modifient le micro-environnement hépatique** en perturbant les échanges entre les différents types cellulaires hépatiques. Ainsi, les hépatocytes sénescents peuvent sécréter des molécules pro-fibrotiques, activant les HSCs et provoquant un stress des cellules environnantes (Yu et al., 2021).

L'évolution des différents stades peut être lente, ou au contraire très rapide (**Figure 7**). Généralement, la progression de la stéatose hépatique vers la fibrose est lente lorsque le diagnostic a été posé au stade de MASH précoce. À l'inverse, l'évolution est relativement plus rapide vers le stade de cirrhose lorsque la maladie est identifiée à un stade de fibrose avancée et que le patient a de forts risques de développer des maladies cardiovasculaires (Haas et al., 2016).

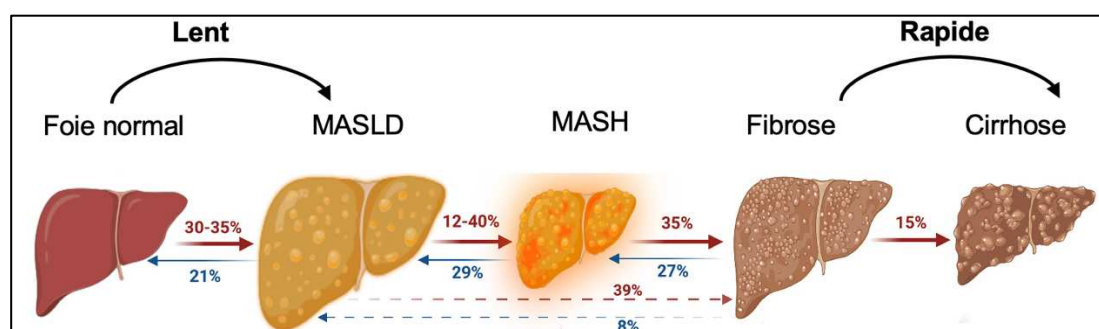


Figure 7 : Différents stades de la pathologie MASLD.

Les flèches indiquent les différentes évolutions possibles. Les chiffres calculés au niveau mondial précisent les prévalences d'évolution. Adaptée de (Lekakis & Papatheodoridis, 2024)

Diagnostic de la MASLD :

L'évolution de la MASLD est **indépendante d'un patient à un autre**, ce qui en fait une maladie complexe à diagnostiquer. Généralement, le diagnostic est **tardif et s'effectue à des stades avancés** de la pathologie puisque les patients ne présentent pas de symptômes aux stades précoces de la maladie. En Europe, le **dépistage est réalisé pour les patients à risque** qui ont un IMC > 30 kg.m². Le diagnostic passe par le dosage des transaminases dans le sang pour évaluer la souffrance hépatique. Puis, une biopsie sera réalisée pour déterminer la présence de certains paramètres de la MASLD comme la fibrose, l'inflammation et la ballonnisation des hépatocytes

grâce à différents marquages (Chalasani et al., 2018). Ces paramètres déterminent un score NAS (« NAFLD activity score ») afin de conclure sur un diagnostic (Grzych et al., 2023).

Facteurs de risques des MASLDs :

Le **vieillessement** constitue un **facteur de risque** pour le développement d'une cirrhose ou d'un hépatocarcinome (HCC), augmentant les risques de fibrose (Ogrodnik, 2021). De nombreux facteurs endogènes liés à l'âge influencent également la progression vers la cirrhose : le ralentissement du métabolisme lipidique et glucidique provoque une accumulation plus importante des lipides dans les hépatocytes et une insulino-résistance accrue (Z. Gong et al., 2017).

La transition de la simple stéatose en MASH est en partie due à une inflammation lobulaire, elle-même causée par des **facteurs intra- et extra-hépatiques**.

D'un côté, l'**accumulation de lipides** favorise le **stress du réticulum endoplasmique** (RE) et l'oxydation accrue des acides gras via le cycle de Krebs, conduisant à une **production excessive de ROS** responsables de stress oxydant, de dommages à l'ADN et de mort cellulaire. Lors de ce stress, les hépatocytes et les KCs vont sécréter des cytokines et des chimiokines tels que l'interleukine-8 (IL8) et l'IL1B, respectivement permettant l'induction des signaux pro-inflammatoires, le recrutement des cellules immunitaires et l'activation de l'inflammasome. Les hépatocytes endommagés vont également sécréter des molécules comme le facteur de croissance fibroblastique 21 (« *Fibroblast growth factor 21* », FGF-21) qui a également un rôle dans l'inflammation hépatocytaire (Krenkel & Tacke, 2017).

Par ailleurs, l'inflammation peut être induite par des facteurs extra-hépatiques, en lien avec les organes périphériques reliés au foie. En particulier, le **tissu adipeux sécrète différentes hormones, cytokines et protéines de la MEC**. Certaines de ces molécules peuvent être un **facteur de sévérité de la MASH** comme un faible taux d'adiponectine et un taux élevé de leptine. De plus, une **altération de la barrière intestinale**, notamment observée en cas d'obésité, provoque le **passage des pathogènes** dans le système circulatoire, ce qui active les macrophages et les cellules

dendritiques. Enfin, les patients atteints de MASH présentent une **modification de la composition des acides biliaires**, altérant la signalisation de récepteurs comme le récepteur farnésioïde X (« *farnesoid X receptor* », FXR) notamment impliqué dans la protection contre la stéatose (**Figure 8**) (Nawrot et al., 2021).

Par ailleurs, des facteurs environnementaux tels que **l'alimentation et la pollution** contribuent à augmenter les risques de développement de la MASLD, expliquant le nombre croissant d'enfants atteints de la maladie. Selon l'Organisation Mondiale de la Santé (OMS), les personnes ayant un IMC > 30 kg.m² sont exposées au risque de maladies métaboliques, en raison d'un apport calorique trop important et d'un **manque d'activité physique** dans les populations actuelles (Stefan et al., 2025). En effet, un **régime riche en graisse et en sucre**, notamment par la consommation de boissons trop sucrées ou d'aliments d'origine animale riches en acide gras, accentue le risque de développement de stéatose par la stimulation de synthèse lipidique *de novo*. Au même titre, **l'exposition quotidienne à certains polluants** tels que les métaux lourds, les pesticides, les insecticides, ou encore les particules fines de l'air peuvent augmenter les risques de stéatose et d'inflammation (Zheng et al., 2013).

Ces facteurs peuvent être modulés par différentes prédispositions **génétiques et épigénétiques** (**Figure 8**). En outre, certains variants génétiques et mutations peuvent altérer les fonctions de lipolyse et de métabolisme du glucose et en conduisant à un stockage excessif de lipides dans les hépatocytes, un stress cellulaire et une inflammation (Jonas & Schürmann, 2021). De plus, il a également été montré qu'une hyper-méthylation des îlots CpGs de gènes anti-fibrotiques comme ceux codants pour les récepteurs nucléaires activés par les proliférateurs de peroxyosomes (« *Peroxisome proliferator-activated receptors* », PPAR) peut favoriser la fibrose (Zeybel et al., 2015) (Murphy et al., 2013).

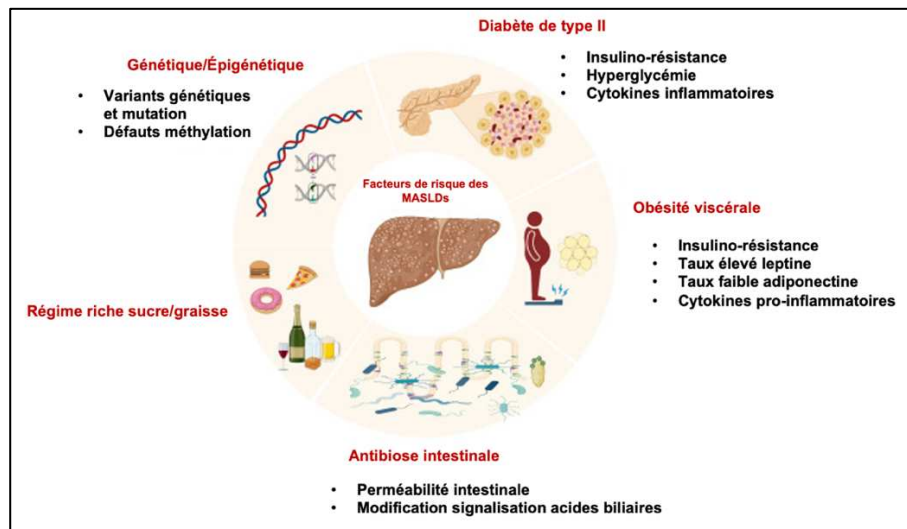


Figure 8 : Facteurs de risque pour le développement des MASLDs.

Adaptée de (Lekakis & Papatheodoridis, 2024)

Traitements des MASLDs :

Les traitements privilégiés pour la MASLD reposent principalement sur des **mesures hygiéno-diététiques**. L'amélioration de l'hygiène de vie reste une stratégie essentielle avant la considération de tout traitement pharmacologique. En effet, l'obésité étant l'un des principaux facteurs de risque de la maladie, un rééquilibrage alimentaire couplé à la pratique régulière d'une activité physique améliore la fibrose hépatique (Canivet & Faure, 2024).

Pour les patients respectant certains critères précis, une **chirurgie bariatrique** peut être réalisée. Des études ont montré que cette intervention permet de réduire la mortalité associée à la MASLD. Néanmoins, cet acte chirurgical n'est pas anodin et peut causer de nombreux effets secondaires (Thereaux et al., 2018).

Enfin, dans les cas extrêmes et notamment en cas de cirrhose ou d'HCC lorsque les stratégies précédentes ne sont plus applicables, le seul traitement possible reste la **transplantation hépatique**. Cet acte chirurgical est soumis à plusieurs critères dont l'âge. Il présente des conséquences lourdes et s'accompagne d'un risque important de **développer une MASLD de manière *de novo*** après la transplantation (Gaggini et al., 2018).

À ce jour, peu de traitements pharmacologiques sont disponibles sur le marché pour contrer efficacement la progression de la MASLD. Le **Resmetirom est la**

première molécule approuvée par l'agence du département de la santé aux États-Unis (« *Food drug administration* », FDA). Ce médicament est un agoniste du récepteur bêta aux hormones thyroïdiennes (« *Thyroid hormone receptor β* », THRB), réduisant le stockage des triglycérides intra-hépatiques (Shakeel et al., 2025). En parallèle, plusieurs molécules ciblant spécifiquement les mécanismes de développement de la pathologie sont actuellement en **phase 3 d'essais cliniques** (Fraile et al., 2021). Parmi elles, on retrouve les **agonistes de PPARA** dont les effets permettent de diminuer l'inflammation lobulaire et la fibrose (Gawrieh et al., 2021).

ii. Les ALDs (Maladies du foie liées à la consommation d'alcool)

Étiologie des ALDs :

Les maladies du foie liées à la consommation d'alcool (« *alcoholic liver disease* », ALD) correspondent à un large spectre de maladies, allant de formes asymptomatiques - parfois pendant de nombreuses années en raison de la grande capacité de compensation fonctionnelle du foie - à des stades avancés comme la cirrhose voire l'HCC, qui peuvent être **mortels à court terme s'il n'y a pas de greffe hépatique (Figure 9)** (B. Gao et al., 2020). Contrairement à la MASLD, l'ADL présente un risque plus important de développer rapidement des formes graves. Selon l'OMS, trois fois plus de patients meurent chaque année des suites d'une cirrhose liée à une consommation excessive d'alcool (supérieure à 5 verres par jour pour les hommes) en comparaison avec le nombre de décès de patients atteints de forme grave de la MASLD.

Mécanismes de métabolisation de l'alcool :

Dans un foie sain, l'oxydation de l'alcool peut être réalisée par 3 voies :

- La voie de l'**Alcool déshydrogénase** (ADH), une enzyme présente dans le cytoplasme des hépatocytes qui va oxyder l'alcool en acétaldéhyde, une molécule toxique provoquant des dommages à l'ADN ;
- Le **système microsomal d'oxydation de l'alcool** (« *microsomal ethanol oxidizing* », MEOS) dont la réaction enzymatique principale a lieu dans le RE

des hépatocytes grâce à l'action de l'Isoforme E1 du cytochrome P450 (CYP2E1). Cette voie va générer des ROS, provoquant un stress oxydatif et une réaction inflammatoire. Bien qu'elle ne métabolise que 10% de l'alcool dans un foie sain, son activité augmente en cas de consommation chronique d'alcool ;

- La voie de la **catalase** où l'oxydation de l'alcool sera permise grâce à l'Aldéhyde déshydrogénase (ALDH), présente dans les mitochondries des hépatocytes. Cette réaction aboutie à la formation de l'acétyl co-enzyme A (acétyl-CoA), qui sera utilisée dans le cycle de Krebs.

La métabolisation de l'alcool provoque une souffrance cellulaire, notamment par la **production de ROS et de dommage à l'ADN**, qui vont activer une réponse inflammatoire (Jiang et al., 2020) (S.-Y. Liu et al., 2021). En parallèle, l'augmentation de la perméabilité intestinale favorise la translocation des lipopolysaccharides (LPSs) vers le foie, où ils se lient à différents récepteurs pour activer les macrophages et les KCs qui vont, eux libérer les interleukines (Mak & Shekhar, 2025). Enfin, la production des ROS par la sur-métabolisation de l'alcool épuise les réserves de glucagon, réduisant l'activité des méthylases de l'ADN et favorisant la prolifération et la tumorigenèse (T. Xu et al., 2019).

À des stades avancés de la maladie, la perte de fonction et la sénescence des hépatocytes amènera à un mécanisme **d'arrêt de prolifération et d'induction d'une régénération dysfonctionnelle de ces cellules** (Sato et al., 2019). Ces mécanismes seront détaillés dans le chapitre **Introduction 2.1 et 2.2** p64 et p73.

Diagnostic des ALDs :

L'ALD est généralement diagnostiquée tardivement : la majorité des patients peuvent présenter une forme asymptomatique pendant plusieurs années. Ce diagnostic repose sur **différents marqueurs cliniques et sanguins** comme la jaunisse et des taux élevés de bilirubine et de transaminase, témoignant de la souffrance hépatique. En effet, la bilirubine est toxique sous sa forme libre et doit être éliminée de l'organisme (Mackowiak et al., 2024). Au niveau histologique, on retrouve,

comme pour la MASLD, une stéatose, une ballonnisation des hépatocytes, une infiltration des cellules immunitaires et des nodules de régénération. Il existe un **score** permettant de connaître la **gravité de l'atteinte hépatique** et permettant de savoir si une greffe est nécessaire : le **score MELD** (« *model for end-stages disease* »). Ce score utilise les valeurs de la bilirubine, de la créatine et d'un indicateur de coagulation (« *international normalized ratio* », INR) afin de définir la gravité de l'atteinte hépatique (Wiesner et al., 2003).

Facteurs de risque des ALDs :

Il existe différents facteurs identifiés de prédisposition à l'échelle du patient et qui influencent le métabolisme de l'alcool et l'étiologie des ALDs (**Figure 9**).

Environ **50% des ALDs présentent des facteurs d'héritabilité**. D'un côté, un **polymorphisme nucléotidique** (« *single nucleotide polymorphism* », SNP) au niveau des **gènes codants pour les enzymes ADH et ALDH** modifie leurs fonctionnalités et la métabolisation de l'alcool, avec notamment des différences marquées entre les populations d'**ethnicités européenne et asiatique**. En effet, chez les populations asiatiques où l'ALDH n'est pas active, l'oxydation de l'alcool se fait par la voie MEOS, ce qui augmente les risques de stress oxydatif, de fibrose et d'inflammation (Charatcharoenwitthaya et al., 2020). D'autre part, plusieurs **mutations** dans des **gènes impliqués dans le métabolisme** des lipides ont été identifiés comme facteurs aggravants du risque de développement de l'ALD (Verhulst et al., 2015). Ainsi, des mutations des gènes de la *phospholipase de type patatine contenant le gène 3* (« *patatin like phospholipase domain containing 3* », PNPLA3) et le *facteur de fixation membranaire contenant un domaine O-Acétyltransférase 1* (« *membrane bound O-Acetyltransferase domain containing 1* », MBOAT) vont provoquer une accumulation de graisse (Grimaudo et al., 2020), une diminution de la synthèse des phospholipides (Stickel & Hampe, 2012) et une augmentation de la concentration des triglycérides dans les hépatocytes (Falleti et al., 2016). Ces facteurs vont ainsi augmenter les risques d'inflammation et de développement d'HCC.

En plus des prédispositions génétiques, les **femmes sont plus susceptibles de développer des ALDs** que les hommes pour une consommation égale d'alcool. En effet, elles développeront ces maladies plus précocement et le risque de mortalité

est plus élevé. Cette différence de prévalence pourrait s'expliquer par le rôle des œstrogènes dans la régulation de la réponse inflammatoire (Wilsnack et al., 2018).

De nombreuses **comorbidités** sont également prises en considération dans l'étiologie de l'ALD. Parmi elles, on peut citer l'**immunodéficience** qui provoque une réponse inflammatoire exacerbée en cas de **cirrhose** ou le **diabète de type II** qui affecte la solubilité de l'alcool. L'association concomitante d'une MASLD et d'une ALD définissent une nouvelle catégorie de pathologies : les **maladies associées au métabolisme et à l'alcool** (« *metabolic and alcoholic associated liver disease* », MetALDs) qui permettent de mieux décrire le spectre de patients atteints d'ALD en soulignant l'**interaction complexe entre prédispositions génétiques et facteurs environnementaux** dans la progression de la maladie (Rinella et al., 2023). Par exemple, la mutation du gène *PNPLA3*, cité précédemment, combinée aux comorbidités propres aux patients MetALD accentuent le risque de développement d'HCC (Galvanin et al., 2025). Ainsi, les patients MetALDs ont plus de risque de développer une fibrose et une forme sévère par rapport aux patients atteints d'une ALD classique (Arab et al., 2025).

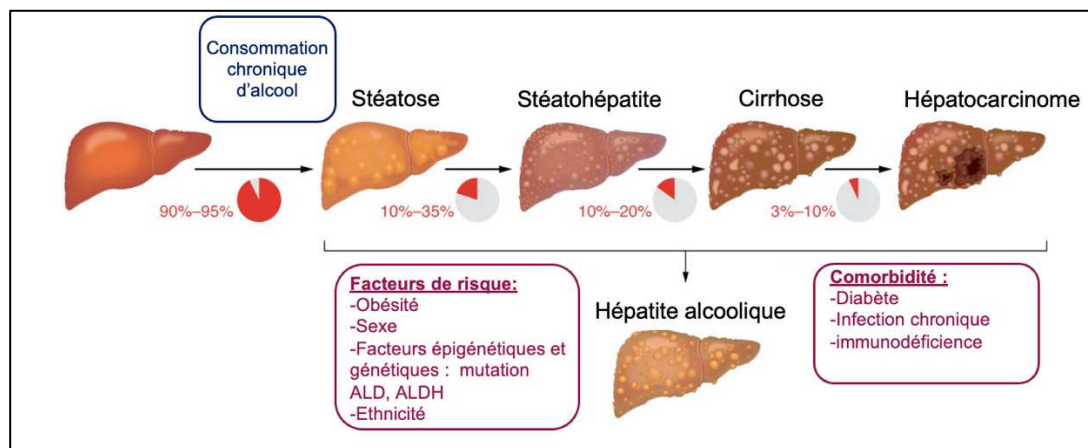


Figure 9 : Facteurs de risque, comorbidités et développement des maladies du foie liées à la consommation d'alcool (ALDs).

Les pourcentages indiquent les risques de progresser d'un stade à un autre si le patient a une consommation chronique d'alcool. Adaptée de (Mackowiak et al., 2024)

Traitements des ALDs :

Les traitements actuels de l'ALD sont uniquement des **traitements de supports** visant à diminuer les causes et les complications de la maladie. Ces traitements reposent, dans un premier temps, sur la diminution de la prise des médicaments hépatotoxiques, et une abstinence d'alcool. Actuellement, le traitement à base de corticoïdes (**corticothérapie**) est considéré comme le traitement de référence dans les pays européens. Il permet d'augmenter l'espérance de vie des patients en limitant l'inflammation et donc la progression de la fibrose (O'Shea et al., 2010). Parallèlement, de **nouvelles thérapies géniques** sont en cours de développement, notamment pour limiter les risques d'évolution de la maladie liés à l'hérédité génique. Par exemple, des petits ARNs interférants (siARNs) ciblant le gène *PNPLA3*, sont en cours de test sur des patients atteints de MetALDs (Galvanin et al., 2025).

Cependant, comme pour la MASLD, la **transplantation hépatique** demeure à ce jour la seule option curative. Elle ne peut toutefois être envisagée que pour une minorité de patients, se trouvant à un stade terminal de la maladie et ayant maintenu une abstinence complète vis-à-vis de l'alcool pendant plusieurs mois.

b. Maladies hépatiques d'origine cholangiocytaire

Les cholangiopathies sont des maladies chroniques du foie touchant les cholangiocytes. Elles sont souvent diagnostiquées à des stades avancés, ainsi, aucun traitement n'est plus envisageable et la transplantation hépatique devient le seul recours. Selon l'OMS, elles sont responsables d'environ 16% des transplantations du foie. En l'absence de diagnostic et de prises de sang précoces, ces maladies peuvent facilement évoluer en cholangiocarcinome. Les cholangiopathies peuvent avoir des étiologies diverses : génétiques, immunitaires, virales ou malignes. Parmi les formes identifiées, les plus connues et étudiées sont la cholangiopathie biliaire primitive (« *primary biliary cholangitis* », PBC), la cholangiopathie sclérosante primitive (« *primary sclerosing cholangitis* », PSC) et le syndrome d'Alagille (« *Alagille syndrom* », ALGS). Dans cette partie, seules ces 3 formes majeures de cholangiopathies seront développées.

i. Le syndrome d'Alagille (ALGS)

Étiologie de l'ALGS :

L'ALGS est une **maladie génétique néonatale à transmission autosomique dominante**, généralement diagnostiquée avant les 3 ans. Elle se caractérise par une atteinte du réseau biliaire, causant une **malformation chronique avec une raréfaction de canaux biliaires intra-hépatiques**. Elle touche également d'autres organes vitaux comme le cœur, le rein, le squelette et les yeux (**Figure 10**) (Saleh et al., 2016). Dans la majorité des cas, la pathologie est liée à des mutations dans les gènes du *ligand «Jagged » 1 de la voie canonique Notch* (« *jagged canonical notch ligand 1* », *JAG1*) ou du *récepteur 2 de la voie de signalisation Notch* («*neurogenic locus notch homolog protein 2*», *NOTCH2*), altérant la voie de signalisation Notch (Turnpenny & Ellard, 2012). Cette voie de signalisation joue un rôle dans la communication intercellulaire, la prolifération et la transdifférenciation cellulaire (Schaub et al., 2018).

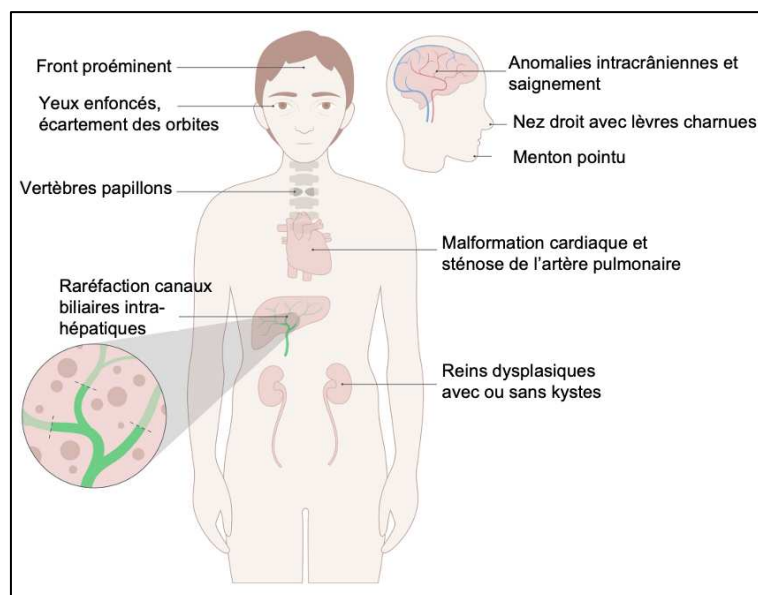


Figure 10 : Traits physiques caractéristiques et organes vitaux touchés par l'ALGS.

L'ALGS est une maladie multifactorielle qui touche plusieurs organes. Elle est caractérisée par des traits physiques particuliers. Au niveau du foie, les patients présentent une raréfaction des canaux biliaires intra-hépatiques. Cette pathologie est causée par une mutation des gènes *JAG1* et *NOTCH2*. Adaptée de (Ayoub et al., 2023)

Diagnostic de l'ALGS :

Le diagnostic de l'ALGS est très complexe et dépend du profil clinique de chaque patient. En effet, cette pathologie **affecte plusieurs organes** et son identification nécessite **l'évaluation de plusieurs facteurs cliniques**. En l'absence d'antécédents génétiques dans la famille, le diagnostic est établi sur la mise en évidence d'atteintes cliniques touchant au moins 3 des 5 organes concernés (**Figure 10**). Au niveau du foie, l'évaluation des marqueurs cholestatiques comme la bilirubine vont permettre d'orienter le diagnostic, mais, le diagnostic définitif repose sur un **test génétique avec les marqueurs JAG1 et NOTCH2** (Ayoub & Kamath, 2020).

Traitement de l'ALGS :

Le traitement de la maladie est, également, très complexe et nécessite plusieurs spécialistes. Concernant le foie, des agents cholérétiques seront utilisés, couplés à des interventions chirurgicales comme des dérivations biliaires et une exclusion iléale. Néanmoins, la greffe hépatique reste la solution la plus durable.

ii. La cholangiopathie biliaire primitive (PBC)

Étiologie de la PBC :

La PBC, connue aussi sous le terme de **cirrhose biliaire**, est une pathologie **multifactorielle** qui se définit par une **inflammation et une destruction des canaux biliaires intra-hépatiques** (Boonstra et al., 2012). Elle résulte de la libération d'antigènes par les cholangiocytes endommagés qui induit l'infiltration de cellules mononucléaires dans les canaux biliaires au niveau du canal de Hering. Ces canaux vont, par la suite, changer de forme et de structure (**Figure 11**) (A. Tanaka, 2020).

La destruction progressive des canaux biliaires interlobulaires est causée par une accumulation d'antigènes anti-mitochondriaux au niveau du pôle apical des cholangiocytes. Cette exposition anormale d'antigènes, combinée à l'induction des gènes codants pour les cytokines et les chimiokines (Imai et al., 1997), va déclencher une réponse auto-immune et entraîner la mort des cholangiocytes (**Figure 11**) (Lleo et al., 2017). Cependant, la raison pour laquelle les antigènes s'accumulent à cet endroit, et spécifiquement dans les cholangiocytes, n'est pas encore bien comprise. Par

ailleurs, la perte de structure des canaux biliaires est liée à une **diminution de l'expression des gènes codants pour les molécules d'adhésion**. Enfin, la présence d'antigènes mitochondriaux dans le système circulatoire va également interférer avec les enzymes impliquées dans le métabolisme des acides aminés (Dubel et al., 1999).

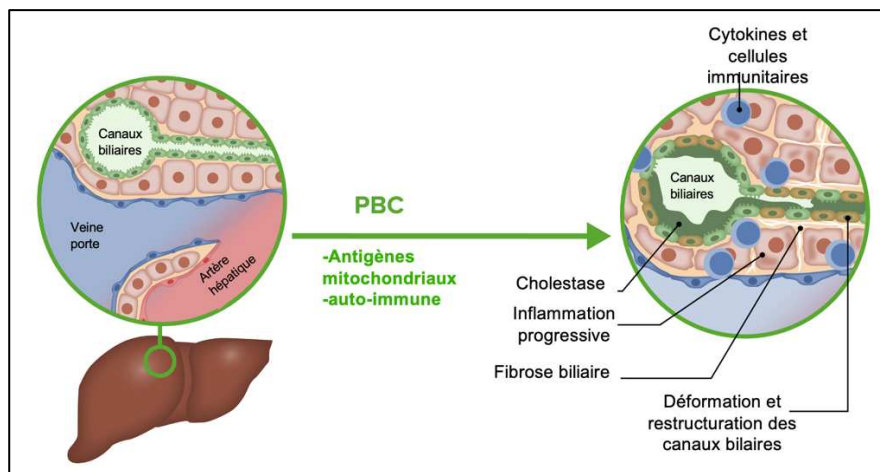


Figure 11 : Développement de la cholangiopathie biliaire primitive (PBC).

La PBC résulte de l'inflammation des canaux biliaires interlobulaires, causée par une perte de tolérance des antigènes mitochondriaux qui s'accumulent au niveau du pôle apical des cholangiocytes. Elle provoque la destruction des canaux biliaires. Adaptée de (Primary Biliary Cholangitis | Concise Medical Knowledge, *n.d.*)

Diagnostic de la PBC :

Les patients atteints de cette maladie ont un risque accru de développer une cholestase et une fibrose hépatique. Contrairement aux autres cholangiopathies, la PBC peut être **diagnostiquée à des stades précoces**. En effet, les patients vont présenter des symptômes cliniques précoces, caractéristiques de la pathologie comme une hyperlipidémie. Pour un diagnostic complet, une analyse sanguine permettra d'évaluer le taux d'antigènes mitochondriaux comme les antigènes de type M2 et certains marqueurs cholestatiques. Si leur taux est élevé, une biopsie sera alors réalisée pour confirmer la présence de marqueurs d'inflammation (Lleo et al., 2017).

Traitements de la PBC :

Les traitements actuels de la PBC sont des **traitements de supports**. On retrouve, principalement, des **molécules régulant la composition de la bile**, ou **ciblant des récepteurs nucléaires de la famille des PPARs**. Ces derniers permettent de diminuer la concentration en cytokines pro-inflammatoires comme IL1B et ralentir la progression de la maladie (*Primary Biliary Cholangitis | Concise Medical Knowledge*, n.d.).

iii. La cholangiopathie sclérosante primitive (PSC)

Étiologie de la PSC :

Une autre forme de cholangiopathie est la PSC, maladie chronique dont **l'évolution est lente et progressive** (Karlsen et al., 2017), et qui **endommage les voies biliaires intra- et extra-hépatiques**. Les voies biliaires sont alors bloquées, provoquant une **accumulation de la bile** dans le foie, altérant également les hépatocytes et provoquant une inflammation. La PSC est caractérisée par une interface pathologique entre inflammation locale, fibrose et cholestase, favorisant l'obstruction biliaire et une atteinte hépatique chronique (**Figure 12**).

L'étiologie de cette pathologie est incertaine, mais une prédisposition génétique est fortement suspectée. Cependant, les gènes impliqués dans la PSC ne sont pas encore bien identifiés (Folseraas et al., 2011). La PSC affecte principalement les hommes et est fréquemment liée à une inflammation de l'intestin de type maladie de Crohn.

Diagnostic de la PSC :

Cette maladie est **multifactorielle**, rendant son diagnostic particulièrement difficile chez les enfants car les symptômes sont confondants avec ceux de la PBC. Sa détection est également compliquée car la maladie est principalement asymptomatique aux premiers stades. Les tests sanguins réalisés sont similaires à ceux utilisés pour la PBC : évaluation des antigènes mitochondriaux, des marqueurs cholestatiques comme la bilirubine, mais aussi, de certains marqueurs de souffrance

hépatique. Le diagnostic repose essentiellement sur **l'exclusion de la PBC après analyse de ces marqueurs**. Par la suite, une imagerie par rayons X, réalisée après une injection d'un produit de contraste marquant les voies biliaires, sera prescrite pour visualiser les voies biliaires et confirmer la présence d'obstructions caractéristiques de la pathologie (Dyson et al., 2018).

Traitements de la PSC :

Les traitements de la PSC sont principalement basés sur l'administration **d'acides biliaires afin de solubiliser la bile et de limiter la cholestase**. Ce traitement peut être couplé ou non à un traitement de fond avec des antibiotiques afin de diminuer la réponse immunitaire et de contrôler l'inflammation.

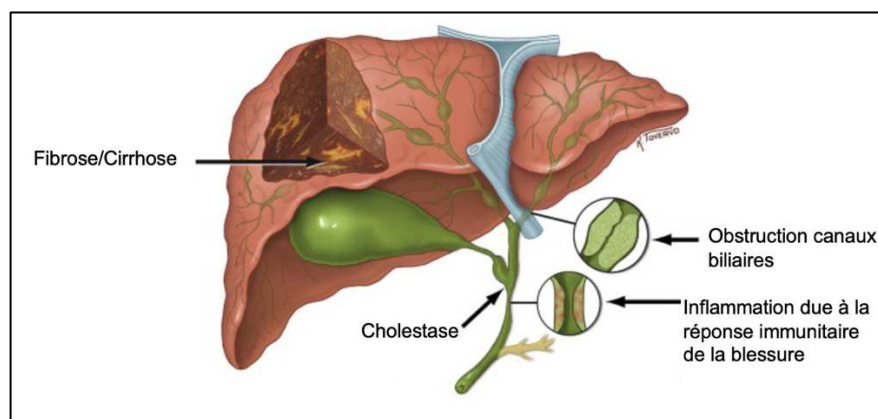


Figure 12 : Caractéristiques d'une cholangiopathie sclérosante primitive (PSC).

La PSC est caractérisée par une obstruction des voies biliaires intra- et extra-hépatiques. Cette pathologie montre une interface entre inflammation locale, fibrose hépatique et cholestase. Adaptée de (Folseraas et al., 2011)

Comme évoqué, la majorité des pathologies hépatiques provoque une souffrance cellulaire souvent liée à une inflammation et à l'incapacité du foie à éliminer les toxines. Le foie a donc développé des processus de régénération pour compenser ces altérations et maintenir les fonctions physiologiques essentielles. Ces processus de régénération reposent principalement sur la plasticité des cellules épithéliales parenchymateuses, capables de répondre aux lésions par des mécanismes d'hyperplasie et d'hypertrophie, assurant ainsi une croissance compensatrice.

Cependant, l'inflammation chronique et la fibrose peuvent déséquilibrer les facteurs impliqués dans la régénération, comme les cytokines. Dans ce cas, les interactions cellulaires entre les différentes cellules du foie sont perturbées, entraînant des mécanismes de régénération incontrôlables et pathologiques.

Pour cette raison, la suite de mes propos va se focaliser sur les mécanismes de régénération à la fois physiologiques et pathologiques des cellules parenchymateuses du foie, dont une connaissance fine est indispensable à une meilleure compréhension des maladies hépatiques.

II. Plasticité des cellules parenchymateuses et régénération hépatique

2.1. Origine de la plasticité des cellules parenchymateuses du foie

a. Retour aux origines des cellules épithéliales : différenciation à partir de l'hépatoblaste

Afin de mieux comprendre les mécanismes de régénération du foie, en particulier en lien avec la plasticité des cellules épithéliales, la partie suivante détaillera les mécanismes de maturation de ces cellules depuis l'embryogenèse.

i. Mise en place du bourgeon hépatique lors du développement embryonnaire

Les étapes initiales du développement embryonnaire du foie sont bien conservées entre les différentes espèces. Le diverticule hépatique **dérive de l'endoderme** provenant de l'axe antérieur primitif de l'embryon. Les cellules de cet endoderme vont migrer vers l'avant de l'embryon, se séparant du mésoderme cardiaque pour former l'endoderme hépatique. Cette spécification hépatique est permise grâce à la **répression de la voie Wnt/ β -caténine** (Gordillo et al., 2015).

Par la suite, l'endoderme hépatique va s'épaissir au niveau ventral, à la jonction entre le mésoderme cardiaque et le septum transversum, formant ainsi le diverticule hépatique. Ce processus s'accompagne de l'**expression des marqueurs** comme le *facteur de transcription à boîte en motif à tête de fourche A1* (« forkhead box A1 », **FOXA1**), le *facteur de transcription à boîte en motif à tête de fourche A2* (« forkhead

box A2 », **FOXA2**) et le facteur de transcription à doigt zinc se fixant sur les motifs « GATA » (« GATA binding protein 4 », **GATA4**). Ces facteurs de transcription (« transcription factors », TFs) peuvent déjà à ce stade **programmer la cellule pour l'expression de gènes hépatocytaires**, par exemple à travers la fixation au niveau de l'enhancer du gène de l'**Albumine** (« Albumin », **ALB**) dans un contexte chromatinien peu favorable (Cascio & Zaret, 1991).

Le diverticule hépatique va former plusieurs couches jusqu'à devenir le **bourgeon hépatique**, composé en grande partie des hépatoblastes (**Figure 13**) (Bort & Zaret, 2009). La formation de ce bourgeon hépatique implique l'expression du facteur nucléaire hépatocytaire beta 1 (« hepatocyte nuclear factor 1 beta », **HNF1B**) et du facteur de transcription de type *Proxpero Homéo-Boxe 1* (« prospero homeobox1 », **PROX1**). Les hépatoblastes sont caractérisés par l'expression des gènes **ALB** et *Alpha-foetoprotéine* (« alpha fetoprotéin », **AFP**).

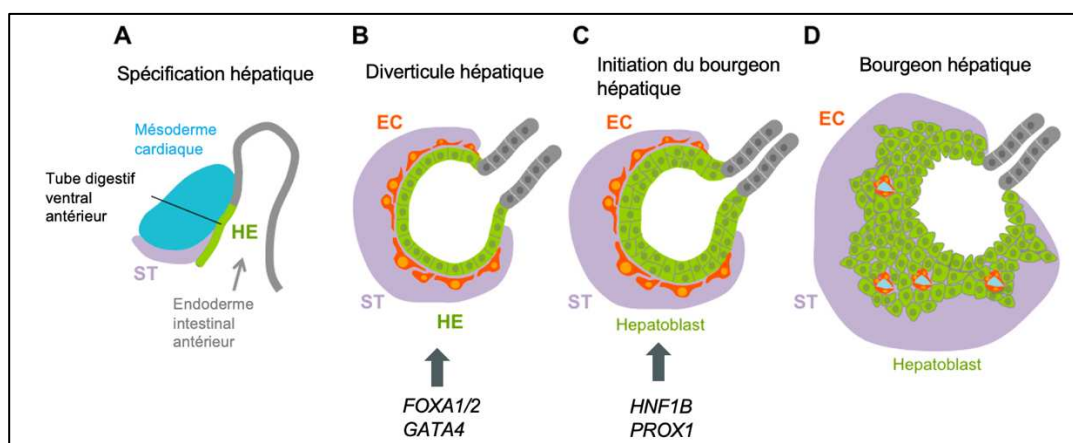


Figure 13 : Étape de la formation du bourgeon hépatique.

(A) Coupe sagittale montrant le lien entre l'endoderme hépatique (HE en vert), le mésoderme cardiaque (en bleu) et le septum transversum (ST, en violet), nécessaires à la spécification hépatique. (B) Mise en place du diverticule hépatique grâce à l'implication de **FOXA1**, **FOXA2** et **GATA4**. Les cellules endothéliales (EC, orange) entourent l'endoderme hépatique épaissi qui entame un processus de bourgeonnement dans le septum transversum. (C) Initiation du bourgeon hépatique permis par l'expression de **HNF1B** et **PROX1**. L'endoderme hépatique forme un épithélium pseudostratifié. (D) Les cellules de l'endoderme (hépatoblastes) prolifèrent et migrent dans le septum transversum pour former le bourgeon hépatique final. Les cellules hématopoïétiques commencent à migrer dans le bourgeon pour permettre la future vascularisation du foie. Adaptée de (Gordillo et al., 2015)

ii. Différenciation des hépatoblastes en hépatocytes ou en cholangiocytes

Les hépatoblastes sont des cellules embryonnaires **non polarisées** capables de se différencier soit en hépatocytes, soit en cholangiocytes selon les signaux qu'elles reçoivent et leur positionnement dans le bourgeon hépatique en développement (Gissen & Arias, 2015).

Différenciation des hépatoblastes en hépatocytes :

Comme décrit précédemment, de **nombreux gènes exprimés lors de la formation des hépatoblastes sont des gènes caractéristiques des hépatocytes**. En effet, la spécification hépatocytaire est le programme de différenciation par défaut des hépatoblastes (Lemaigre, 2009) (L. Yang et al., 2023). Ce processus est marqué par l'expression **progressive de différents marqueurs hépatocytaires** auxquels s'ajoutent l'expression de *récepteur nucléaire de la sous famille 1 groupe H 4* (« *nuclear receptor subfamily 1 group H member 4* », *NR1H4*), codant pour *FXR*, ainsi que l'expression de *PPARA* (L. Yang et al., 2017).

La différenciation des hépatoblastes en hépatocytes est finement contrôlée par de nombreuses voies de signalisation comme la voie de Oncostatin M (OSM) et la voie Wnt (**Figure 14**). Celles-ci interviennent à la fin du processus de différenciation et permettent la maturation cellulaire grâce à l'expression de gènes codants pour des TFs spécifiques aux hépatocytes comme *HNF1A*, *HNF4* et *HNF6* (Si-Tayeb et al., 2010). Ce processus s'accompagne de la polarisation des hépatocytes (Gordillo et al., 2015).

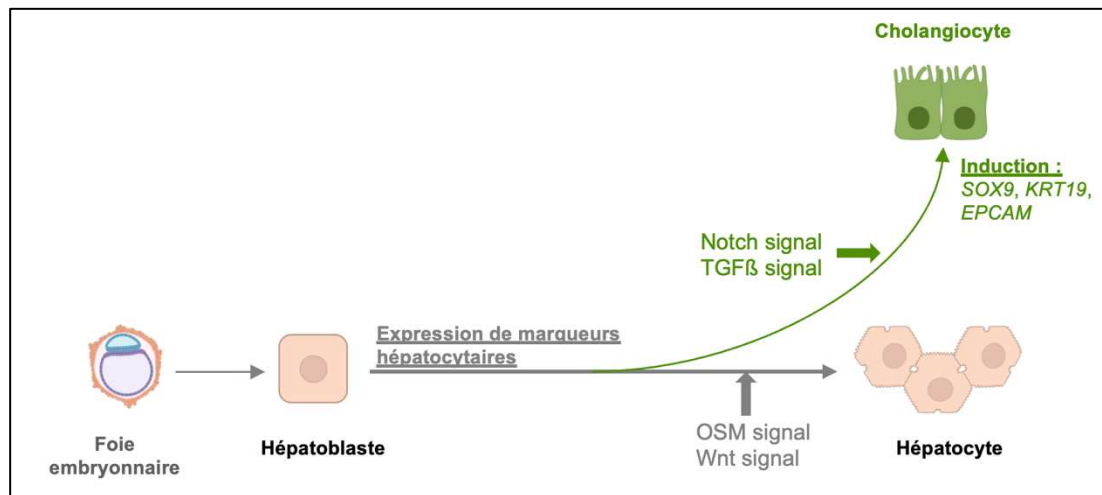


Figure 14 : Différenciation des hépatoblastes en hépatocytes ou en cholangiocytes.

La spécification hépatocytaire est le programme génétique par défaut des hépatoblastes. Cependant, elle est contrôlée par les voies de signalisation OSM et Wnt qui permettent l'expression des gènes codants pour les TFs spécifiques des hépatocytes. D'autre part, l'hépatoblaste se différencie en cholangiocyte lorsqu'il reçoit des signaux des voies Notch et TGFβ, permettant l'expression de marqueurs biliaires.

Différenciation des hépatoblastes en cholangiocytes :

Par ailleurs, lorsque l'hépatoblaste reçoit les signaux des **voies Notch et du facteur de croissance transformant bêta** (« *transforming growth factor β* », TGFβ), **les gènes hépatocytaires sont réprimés** et la cellule s'engage dans la voie de différenciation cholangiocytaire (**Figure 14**) (L. Yang et al., 2023). Ces signaux **induisent l'expression de marqueurs biliaires** tels que la *kératine 19* (« *keratin 19* », **KRT19**) (Tanimizu & Miyajima, 2004), faiblement exprimée dans les hépatoblastes (Lemaigre, 2009) et du gène codant pour *le facteur de transcription à boîte SYR 9* (« *SRY-Box transcription factor 9* », **SOX9**), exprimé uniquement dans les cholangiocytes qui se différencient proches de la veine porte. Il est intéressant de noter que **la maturation et la spécification des cholangiocytes ont lieu avant celles des hépatocytes** (Tanimizu & Mitaka, 2017).

La différenciation des cholangiocytes s'accompagne du **développement progressif du réseau biliaire** qui se déroule en 3 étapes. Ce processus débute par la formation des tubes intra-hépatiques. Tout d'abord, les cholangiocytes immatures, en association avec les hépatoblastes voisins, vont entourer l'espace luminal pour former des canaux biliaires asymétriques. Puis, la différenciation des cholangiocytes permettra d'élaborer des structures symétriques. Enfin, le réseau biliaire deviendra

mature par la prolifération des cholangiocytes et l'augmentation du flux biliaire (**Figure 15**) (Antoniou et al., 2009).

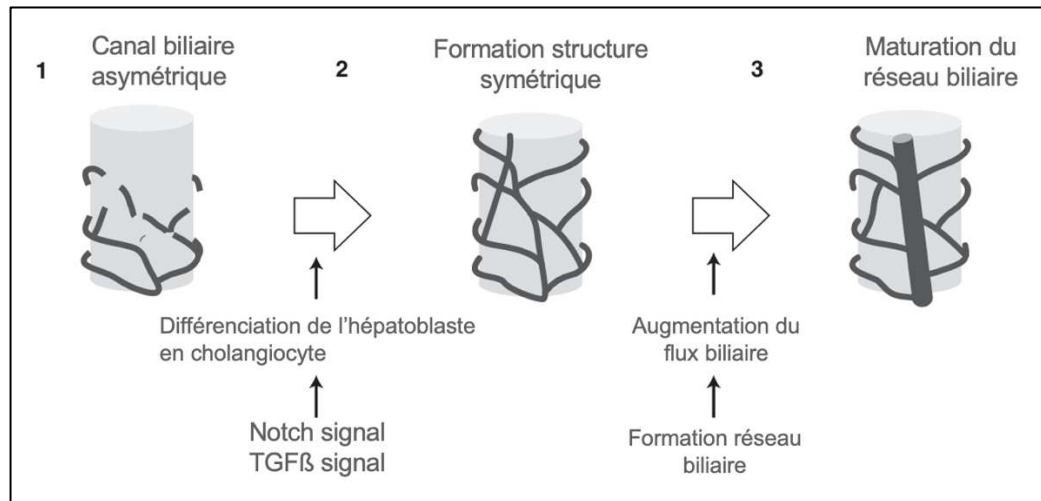


Figure 15 : Étape de formation du réseau biliaire.

Description des 3 étapes de formation du réseau biliaire : (1) différenciation des hépatoblastes en cholangiocytes et formation d'un canal biliaire asymétrique, (2) Formation de structures symétriques (3) maturation du réseau biliaire et prolifération des cholangiocytes. Adaptée de (Tanimizu & Mitaka, 2017)

Modifications épigénétiques contrôlant la différenciation des hépatoblastes :

Cette différenciation des hépatoblastes en hépatocytes ou en cholangiocytes est régulée par différents mécanismes épigénétiques (**Figure 16**). En effet, les hépatoblastes sont caractérisés par des **promoteurs bivalents**, c'est-à-dire possédant des marqueurs épigénétiques associés à des promoteurs à la fois actifs et inactifs comme la triméthylation de la lysine 4 de l'histone 3 (H3K4me3) et la triméthylation de la lysine 27 de l'histone 3 (H3K27me3), respectivement détaillés dans le chapitre **Introduction 3.1.a.ii** p.83. Ces promoteurs bivalents sont notamment situés en **amont des gènes de pluripotences**, c'est-à-dire des gènes exprimés dans les hépatoblastes mais silencieux dans les cellules matures différenciées (H. Kim et al., 2011). En effet, lors de la différenciation cellulaire vers l'hépatocyte ou le cholangiocyte, ces gènes perdent leurs marques épigénétiques activatrices telles que H3K4me3, traduisant leur répression stable dans les cellules différenciées. En parallèle, au cours de leur engagement vers la lignée cholangiocytaire, les promoteurs des gènes spécifiques des hépatocytes acquièrent des marques épigénétiques

répressives comme H3K27me3 et la triméthylation de la lysine 9 de l'histone 3 (H3K9me3) (L. Yang et al., 2023).

Ces changements montrent le rôle crucial de la dynamique des modifications épigénétiques dans le contrôle du destin cellulaire des hépatoblastes. De récentes études montrent qu'une petite partie de gènes présentent des promoteurs bivalents dans les cellules épithéliales hépatiques. Ces gènes seront plus facilement réactivés en cas de blessure hépatiques afin de favoriser la régénération cellulaire (C. Zhang et al., 2021).

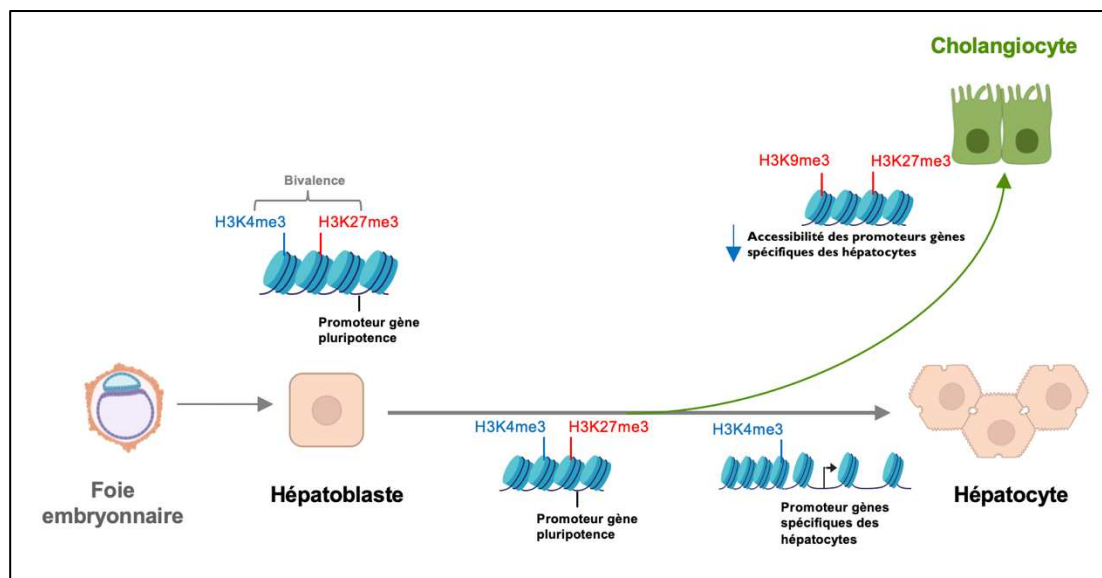


Figure 16 : Dynamique des modifications épigénétiques lors de la différenciation des hépatocytes.

Les hépatoblastes sont caractérisés par la présence de promoteurs bivalents au niveau des gènes de pluripotences. Ces gènes de pluripotences perdent la marque épigénétique d'activation au niveau des promoteurs pendant la différenciation cellulaire. Dans les hépatocytes matures différenciés, on retrouve la présence de marques épigénétiques activatrices au niveau des promoteurs des gènes spécifiques des hépatocytes. Au contraire, ces mêmes gènes sont caractérisés par des marques épigénétiques répressives dans les cholangiocytes matures différenciés.

b. Plasticité des cellules épithéliales du foie

Comme indiqué plus tôt, les hépatocytes et les cholangiocytes conservent un potentiel de plasticité utile à la régénération hépatique.

i. Régénération hépatocytaire

Les hépatocytes ont la **capacité de ré-entrer dans le cycle cellulaire en réponse à une blessure**. Ces mécanismes **reposent sur des changements épigénétiques** qui induisent la perte d'identité cellulaire et l'activation des gènes nécessaires à la prolifération cellulaire (Gadd et al., 2020). Ils ont largement été étudiés dans des modèles d'hépatectomie partielle (« *partial hepatectomy* », PhX) où la régénération des hépatocytes permet de restaurer entièrement la structure et la fonction du foie (Michalopoulos & Bhushan, 2021).

La capacité des hépatocytes à réintégrer le cycle cellulaire lors de la régénération cellulaire, en fonction de leur localisation zonale, fait l'objet de débats dans la littérature. D'un côté, certains auteurs suggèrent que les hépatocytes situés à proximité de la veine porte montrent des capacités de prolifération plus importantes en raison de leurs contacts privilégiés avec l'EGF ou le HGF, signaux nécessaires à la dédifférenciation (Michalopoulos & Bhushan, 2021). Selon cette hypothèse, tous les hépatocytes n'auraient pas la même capacité à réintégrer le cycle cellulaire lors de la régénération hépatique. D'un autre côté, des études suggèrent, au contraire, que tous les hépatocytes, peu importe la zone du foie dans laquelle ils se situent, sont capables de se dédifférencier en cas de blessure hépatique (F. Chen et al., 2020) (Matsumoto et al., 2020).

Dans tous les cas, pendant le processus de régénération, les échanges entre les hépatocytes et les cellules environnantes vont d'abord entraîner la ré-entrée dans le cycle cellulaire, puis, le rétablissement progressif de leur identité et de leurs fonctions. Ce mécanisme est favorisé par la présence d'inflammation, couplée à l'accumulation de LPSs dans le système circulatoire, déclenchant des **mécanismes de dédifférenciation des cellules épithéliales** et notamment des hépatocytes, favorisant ainsi la régénération des cellules endommagées via l'activation de la voie

de signalisation « *Yes-Associated Protein* » (YAP) (Shao et al., 2021). En plus de cette voie de signalisation, des études ont montré que la voie HGF associée au gène de transformation HOS (« *HGF-MNNG HOS transforming gene* », HGF-MET) et celle HGF sont parmi d'autres associées à la mise en place de la régénération hépatique.

ii. Régénération cholangiocytaire

Les cholangiocytes ont, également, une **capacité de prolifération forte** qu'ils soient « petits » ou « grands ». Cette capacité a notamment été étudiée grâce aux modèles de ligation des canaux biliaires (« *bile duct ligation* », BdL) provoquant une accumulation intra-hépatique de la bile (Y. Chen et al., 2022). Ce processus de réaction ductulaire (« *ductular reaction* », DR) permet la prolifération des cholangiocytes et la **néoformation de canaux biliaires, destinée à faciliter l'évacuation de la bile** (Sekiya & Suzuki, 2012). La DR est **associée à une infiltration des cellules immunitaires dans le tissu hépatique environnant**. Il s'agit d'un mécanisme de **protection classique et bénéfique** lorsqu'il est contrôlé (Mavila et al., 2024). La DR est une réaction **hétérogène** qui dépend de la nature de la lésion. Elle peut aller de la simple réorganisation structurelle réversible à une réponse délétère marquée par l'hypoxie et la nécrose des tissus alentours (Desmet, 2011).

La mise en place de la DR dépend des échanges entre les cholangiocytes et les cellules environnantes, qui favorise leur prolifération et différenciation. Ce mécanisme est, notamment, **décrit dans les cholangiopathies où l'accumulation de la quantité de bile non évacuée entraîne une toxicité cellulaire**, favorisant la mise en place de la DR (Pisarello et al., 2015) (Shen et al., 2014). En effet, l'utilisation des modèles murins impliquant l'administration du 3,5-Dithoxylcarbonyl-1,4-Dihydrocolline (DDC) ou BdL a permis de démontrer que la DR implique l'activation de différentes voies de signalisation dont les voies TGF β , inducteur faible de l'apoptose lié au TNF (« *TNF related weak inducer of apoptosis* », TWEAK) et Hippo/YAP (Alvaro et al., 2000) (Bird et al., 2013) (Pepe-Mooney et al., 2019).

Quand la DR devient pathologique :

Lorsque le mécanisme de DR n'est pas finement contrôlé, il peut **entraîner des dysfonctionnements hépatiques** (Mavila et al., 2024). Dans le contexte de maladies hépatiques chroniques, l'accumulation de signaux pro-inflammatoires, couplée à une augmentation du tissu cicatriciel et de la fibrose dans le micro-environnement, favorise l'expansion de la DR (Huang et al., 2023). La DR peut, donc, être **associée à une tentative de régénération inefficace, pouvant devenir pathologique (Figure 17)** et donner lieu à des phénomènes d'anomalies tissulaires.

Ce processus de DR est décrit, dans de nombreuses pathologies hépatiques dont les cholestases, les ALDs et les MALSDs. Il se caractérise par une réorganisation cellulaire accompagnée de la néoformation de canaux biliaires ou de nodules de régénération autour des bandes de fibrose. Certaines cellules de ces zones expriment à la fois des marqueurs biliaires et hépatocytaires, suggérant un mécanisme adaptatif de survie dans le foie extrêmement atteint (Passman et al., 2023). Par ailleurs, la DR est une zone inflammatoire qui contient des facteurs pro-coagulants, pouvant augmenter le risque de maladies cardiovasculaires (Fabris et al., 2024). L'ampleur de la DR est ainsi corrélée avec la sévérité de la pathologie.

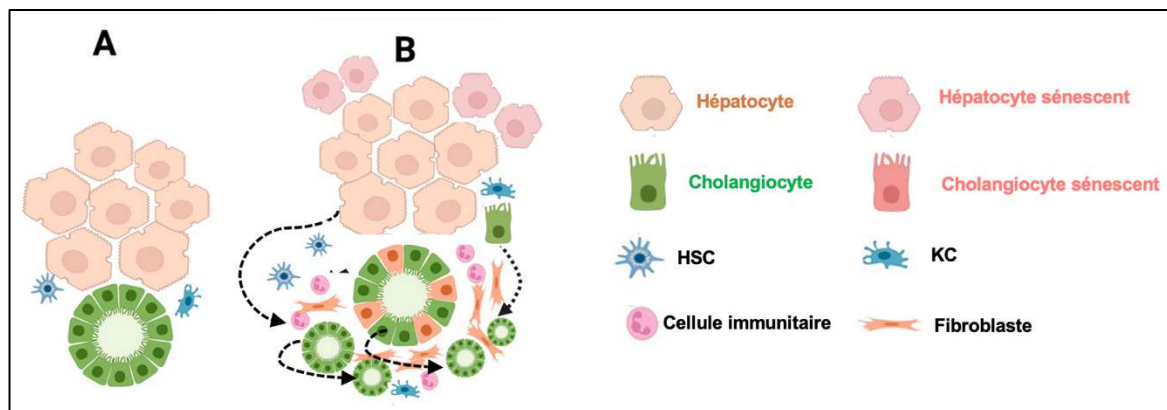


Figure 17 : Schéma montrant les mécanismes de réaction ductulaire.

(A) Schéma de l'organisation cellulaire du foie en condition physiologique **(B)** Dans les conditions de blessure chronique, les cellules endommagées provoquent une infiltration immunitaire et la prolifération des cholangiocytes et des hépatocytes. Adapté de (Mavila et al., 2024)

En cas de maladies chroniques ou en cas d'atteintes aiguës sévères, les hépatocytes et les cholangiocytes peuvent perdre leurs capacités à proliférer et devenir sénescents. Dans ces cas, d'autres mécanismes seront mis en place pour compenser la perte de fonction des cellules. Ces mécanismes impliquent notamment la différenciation de cellules progénitrices du foie ainsi que des phénomènes de transdifférenciation.

2.2. Régénération cellulaire après une blessure hépatique chronique

a. Implication des cellules progénitrices du foie (LPCs)

i. Origine et caractéristiques des LPCs

Un des mécanismes possiblement mis en place, en cas de sénescence des cellules épithéliales, serait **l'activation des LPCs**, également appelées « cellules ovales » du foie. Ces cellules ont été décrites pour la première fois par Faber *et al.* en 1956 dans un modèle d'étude d'HCC (Sancho-Bru *et al.*, 2012). Elles sont **résidentes au niveau du canal de Hering** et font la jonction entre les canaux biliaires et les hépatocytes périportaux (So *et al.*, 2020) pour permettre la régénération cellulaire et tissulaire des épithéliums hépatiques (Michalopoulos, 2007). Le canal de Hering constituerait ainsi **une niche de cellules progénitrices qui émergent lorsque les cellules du foie ne parviennent plus à assurer la régénération** (Miyajima *et al.*, 2014) (X. Li *et al.*, 2021). Ces LPCs prendront le relais en proliférant rapidement, puis, en se différenciant en hépatocytes ou en cholangiocytes (décrit dans le chapitre **Introduction 2.2.a.ii** p.75).

L'origine de ces cellules est discutée dans la littérature, en grande partie à cause du manque de marqueurs spécifiques de ces cellules permettant de les identifier correctement. Bien qu'elles soient localisées principalement dans le canal de Hering, certaines études, utilisant des modèles murins de cholestase bloquant la prolifération hépatocytaire couplés à des techniques de traçage cellulaire, ont montré que les LPCs peuvent provenir des « petits » cholangiocytes (Williams *et al.*, 2014). Par ailleurs, d'autres études mettent en évidence, qu'après la phase de régénération cellulaire, lorsqu'il n'y a plus de lésions hépatiques, les hépatocytes provenant des LPCs

retrouvent leur identité hépatocytaire d'origine. Ces données indiquent que les hépatocytes peuvent également contribuer à la population de LPCs en cas de lésion hépatique chronique (**Figure 18**) (Raven et al., 2017).

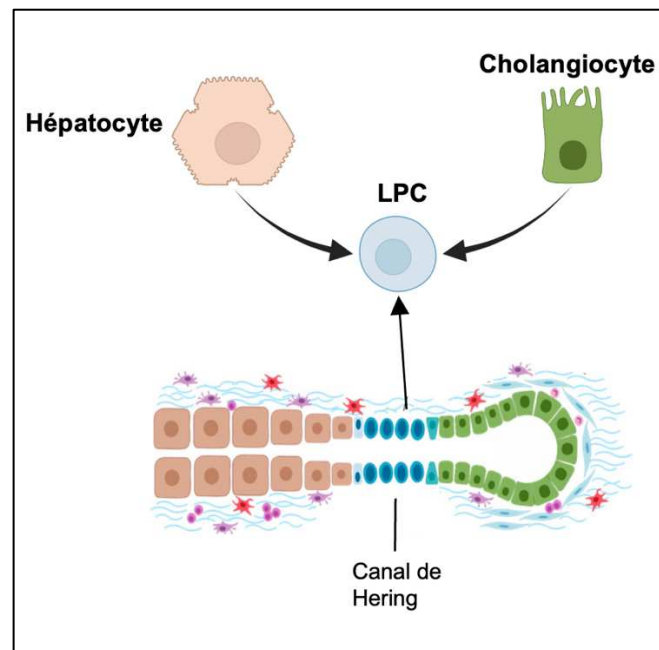


Figure 18 : Schéma montrant l'origine des cellules progénitrices du foie (LPCs).

Les LPCs sont résidentes dans le canal de Hering. Elles prolifèrent rapidement et se différencient en hépatocytes ou en cholangiocytes lorsque ces derniers ne sont plus capables de proliférer. Le manque de marqueurs spécifiques pour ces cellules ne permet pas d'identifier correctement leurs origines. Certaines études montrent qu'elles pourraient provenir des « petits » cholangiocytes ou des hépatocytes.

Dans un foie sain, les LPCs sont présentes mais **restent dans un état quiescent**. Leur signature génique montre un état transitoire entre les hépatocytes et les cholangiocytes puisqu'elles coexpriment des marqueurs hépatocytaires et cholangiocytaires (Pu et al., 2023). Néanmoins, le nombre de LPCs et leurs impacts sur la régénération hépatique est débattue. En effet, certains auteurs mentionnent que cette population cellulaire dormante est minoritaire comparée au parenchyme hépatique, et qu'elles ne s'activent que sous l'effet de différents facteurs dont les signaux pro-inflammatoires (Yan et al., 2023). D'autres chercheurs indiquent qu'elles existent en nombre suffisant pour être rapidement activées et différenciées en cas de blessure chronique (M. Tanaka et al., 2011).

ii. Activation des LPCs en cas de blessure hépatique

Les mécanismes d'activation des LPCs ont été beaucoup étudiés grâce à l'utilisation de divers modèles murins de lésions hépatocytaires et cholangiocytaires (**Figure 19**). Ils montrent que, notamment en cas de blessure chronique, les cellules épithéliales du foie sénescents et d'autres cellules environnantes, telles que les macrophages et les HSCs, vont **libérer différentes molécules permettant l'activation et l'expansion de ces LPCs**. Par exemple, les HSCs peuvent produire et sécréter le FGF7 permettant la prolifération des LPCs (**Figure 19A**). L'ensemble de ces mécanismes est contrôlé par les voies de signalisation Hedgehog, TWEAK et HGF-MET (Jung et al., 2010) (Itoh & Miyajima, 2014). Cependant, une hyperprolifération des LPCs peut, dans certains cas, entraîner la formation d'HCC (X. Li et al., 2021).

Une fois activées, les LPCs se **différencient en hépatocytes ou en cholangiocytes**, selon les **voies de signalisation impliquées dans leur maturation pendant le développement embryonnaire** (détaillé dans le chapitre **Introduction 2.1.a.ii p.66**). Ainsi, l'activation de la voie Wnt/ β -caténine favorise la différenciation des LPCs en hépatocytes, tandis que l'activation de la voie Notch permet la différenciation vers la lignée cholangiocyttaire (**Figure 19B**) (Tirnitz-Parker et al., 2014). Ce mécanisme est également **régulé par une communication avec les cellules environnantes**, notamment les HSCs et les macrophages, qui participent à l'orientation du destin de LPCs (Yimlamai et al., 2014). En effet, en plus des différentes voies de signalisation, les cellules environnantes peuvent libérer des molécules telles que la protéine JAG1 ou la protéine A1 contenant un domaine d'interaction riche en AT (« *AT-rich interaction domain* », ARID1A) qui vont se fixer sur des récepteurs à la surface des LPCs, afin d'orienter ces cellules en hépatocytes ou en cholangiocytes (Tirnitz-Parker et al., 2014).

Il a, par ailleurs, été montré que les LPCs activées jouent un **rôle clé dans la DR**, contribuant à reformer des canaux biliaires lorsque les cholangiocytes entrent en sénescence ou subissent une mort cellulaire (Köhn-Gaone et al., 2016).

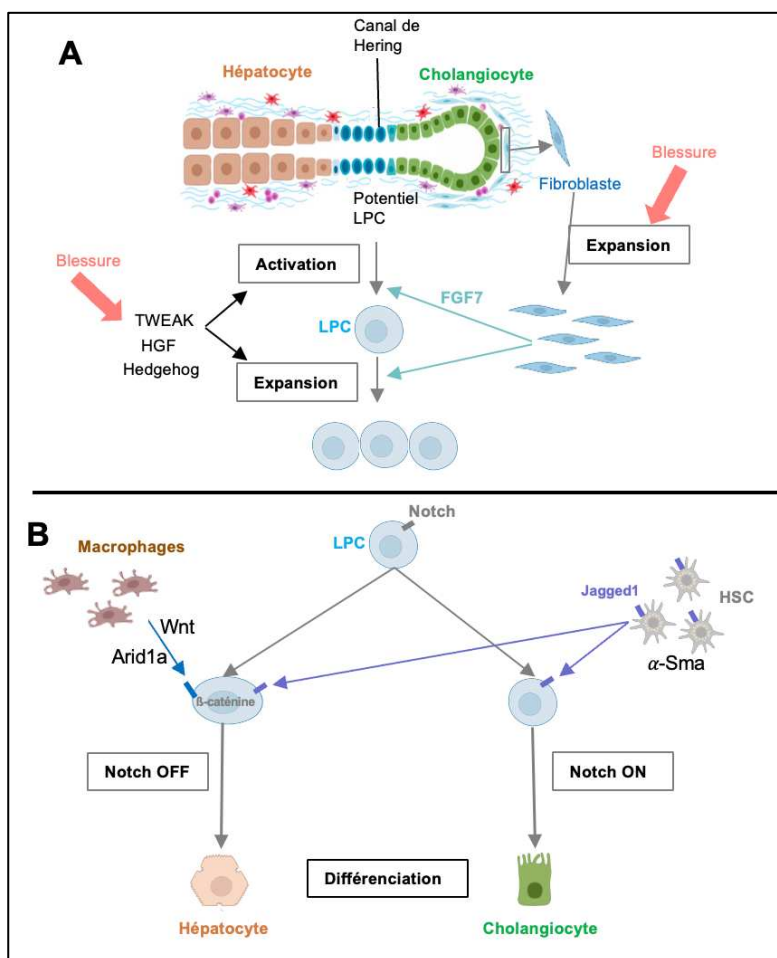


Figure 19 : Schéma montrant les modes d'activation et de différenciation des LPCs.

(A) Activation et expansion des LPCs présents dans le canal Hering lors d'une blessure hépatique. (B) Différenciation des LPCs en hépatocytes ou cholangiocytes. Adaptée de (Tirnitz-Parker et al., 2014)

Durant le processus de différenciation, les LPCs adoptent des **phénotypes intermédiaires**, ayant des caractéristiques à la fois d'hépatocytes et de cholangiocytes, et exprimant **différentes signatures géniques selon leurs états de maturation** (Pu et al., 2023). Différents modèles murins de lésions hépatiques aiguës ou chroniques, comme le régime déficient en choline et supplémenté en éthionine (« *choline-deficient, ethionine-supplemented* », CDE) ou encore l'administration de tétrachlorure de carbone (CCl_4), induisant un stress oxydatif dans les hépatocytes, permettent d'identifier les LPCs grâce à l'expression concomitante de marqueurs biliaires et hépatocytaires tels que *KRT19*, la *molécule d'adhésion des cellules épithéliales* (« *Epithelial cell adhesion molecule* », *EPCAM*), *SOX9* ou encore *HNF4A* (Fujii et al., 2010) (So et al., 2020). Cependant, il peut y avoir un certain **niveau de confusion** puisque certains auteurs considèrent comme progéniteurs des hépatocytes

exprimant le TF SOX9, ce qui complexifie la distinction entre véritables LPCs et hépatocytes réactivant certains programmes développementaux.

Le rôle des LPCs dans la régénération hépatique reste très controversé. En effet, bien que cette population cellulaire soit bien documentée dans la littérature et ancrée comme étant un acteur clé dans les processus de régénération du foie en cas de lésions chroniques, le passage par l'état de progéniteur est discuté. Actuellement, de nombreux auteurs utilisent le terme de « cellules bipotentes », plutôt que celui de « cellules progénitrices » pour désigner des cellules exprimant à la fois des marqueurs hépatocytaires et cholangiocytaires (Pu et al., 2023). L'émergence de nouvelles technologies de marquage et traçage cellulaires, plus spécifiques et sensibles que les techniques de marquages immunologiques classiques, devrait permettre de mieux caractériser les mécanismes de différenciation *in vivo* des LPCs en hépatocytes ou cholangiocytes, ainsi que d'en préciser leur origine cellulaire.

En complément à ce processus de différenciation des LPCs, le mécanisme de transdifférenciation des cellules épithéliales hépatiques matures en réponse à une blessure est également bien documenté.

b. Plasticité cellulaire et transdifférenciation

Un autre mécanisme de régénération possible, lorsqu'il est nécessaire de remplacer un des types cellulaires épithéliaux hépatiques, est la transdifférenciation. En effet, les hépatocytes et les cholangiocytes matures, initialement issus de la même cellule embryonnaire, l'hépatoblaste, ont la capacité de se transdifférencier en l'autre type cellulaire.

Dans cette partie, nous nous focaliserons sur les phénomènes de transdifférenciation des hépatocytes vers les cholangiocytes (« *hepatocyte to cholangiocyte-like transdifferentiation* », HCT), observés en cas de blessure au niveau des canaux biliaires. Ce mécanisme est, notamment, décrit lorsque les mécanismes de renouvellement des cholangiocytes deviennent insuffisants pour maintenir l'intégrité des voies biliaires. Il est à noter que la transdifférenciation des cholangiocytes

en hépatocytes a également été décrite dans d'autres pathologies telles que l'hépatite fulminante, syndrome entraînant une nécrose hépatocytaire (Michalopoulos & Bhushan, 2021).

La transdifférenciation est un mécanisme de **réparation cellulaire reposant sur un changement phénotypique**. En effet, lorsque la DR ne suffit plus à compenser la perte de fonction et que le renouvellement des cholangiocytes n'est plus efficace, les **hépatocytes peuvent former de novo des canaux biliaires** afin d'assurer l'évacuation de la bile. Ce processus a notamment été démontré dans un modèle murin d'ALGS (Schaub et al., 2018). Pendant ce phénomène de transdifférenciation, les hépatocytes **perdent leur programme transcriptionnel d'identité hépatocytaire** (Clerbaux et al., 2017) (Bou Saleh et al., 2021) (Lee et al., 2023). Cependant, les gènes induits au cours de ce processus sont peu décrits dans la littérature. Dans les modèles murins de cholestases, comme par exemple l'administration de DDC, **seule l'induction de gènes comme SOX9 ou EPCAM a été décrite** (Campana et al., 2021). Cette reprogrammation transcriptionnelle s'accompagne d'un **remodelage de la chromatine** (détaillé dans le chapitre **Introduction 3.2.a p.92**) (Merrell et al., 2021) (J. Xu et al., 2022).

Le mécanisme d'HCT est **médié par différentes voies de signalisation**, activées lorsqu'un dysfonctionnement du métabolisme des acides biliaires et un mauvais drainage de la bile sont détectés (Lan et al., 2023). Parmi ces voies figurent la voie Hippo-YAP et TGF β (**Figure 20**) qui sont également impliquées dans la maturation des cholangiocytes. Leur activation témoigne de l'extrême souffrance des cholangiocytes, situation dans laquelle l'HCT devient un mécanisme essentiel de régénération cellulaire (Bou Saleh et al., 2021) (Kamath & Mack, 2019) (Sekiya & Suzuki, 2012).

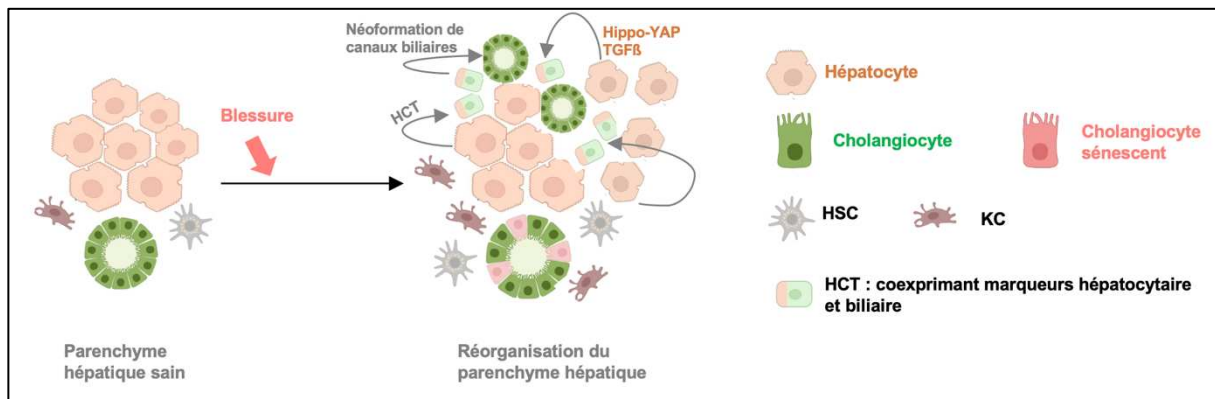


Figure 20 : Schéma montrant la transdifférenciation des hépatocytes vers les cholangiocytes.

Les voies de signalisation menant à ce processus sont indiquées sur les flèches. Les hépatocytes transdifférenciés (HCT) expriment des marqueurs cholangiocytaires et hépatocytaires et peuvent former de nouveaux canaux biliaires.

Les modèles d'études *in vitro* de ces phénomènes n'existent pas, ce qui montre la complexité de modélisation de la transdifférenciation. Ainsi, la compréhension des mécanismes amenant à cette HCT reste limitée. L'utilisation de cellules bipotentes, comme les HepaRG, constitue actuellement l'une des meilleures approches pour identifier les facteurs susceptibles d'orienter la différenciation vers l'une ou l'autre des populations cellulaires (Dianat et al., 2014).

Ce processus est également étudié grâce au scRNA-Seq qui a permis de caractériser les mécanismes potentiellement impliqués dans la différenciation hépatocytaires et de prédire ceux associés à la transdifférenciation (Meng et al., 2025).

Dans certaines pathologies hépatiques, des cellules exprimant à la fois des marqueurs hépatocytaires et cholangiocytaires sont observées sur les biopsies, évoquant le phénomène de l'HCT. En effet, l'identification, au niveau phénotypique et moléculaire, de cellules hybrides a notamment été rapporté chez des patients atteints du syndrome **ALGS ou de PSC** (Andrews et al., 2024). De plus, des analyses transcriptomiques, comme la méthode de séquençage de l'ARN des noyaux individuels des cellules (« *single-nuclei RNA-Seq* », snRNA-Seq) couplé à l'immunomarquage, indiquent également la présence de ces cellules hybrides dans des **stades avancés de pathologies hépatiques telles que les MASLDs ou les ALDs** (Gribben et al., 2024) (Dubuquoy et al., 2015). Ces cellules seraient localisées dans les nodules de régénération (Lin et al., 2010).

La transdifférenciation des cellules épithéliales, et notamment l'HCT, passe par une nécessité de remodelage du transcriptome, impliquant à la fois un remodelage de la chromatine et un changement d'activité des TFs comme décrit ci-après.

III. Contrôle transcriptionnel de la plasticité des cellules du parenchyme hépatique

3.1. Épigénétique et facteurs de transcription (TFs) : généralités

Chez les eucaryotes, le matériel génétique est contenu dans le noyau des cellules sous la forme de chromosomes. La présence de marqueurs épigénétiques permet, selon le contexte, de moduler la transcription de l'ADN afin de maintenir l'identité cellulaire. En effet, cette modulation transcriptionnelle assure la mise en place et le maintien d'un transcriptome spécifique de chaque type cellulaire.

a. Généralités sur les mécanismes épigénétiques

i. Structure de la chromatine

L'unité de base de la chromatine est le nucléosome. Il permet de réguler différents mécanismes notamment la transcription, la réplication et la régulation de la réparation de l'ADN génomique. Cette unité est composée de 146 paires de bases (« *base pairs* », bp) d'ADN enroulées autour d'un octamère de protéines, les histones. Ces histones sont chargées positivement au niveau de leur extrémité N-terminale et sont composées de nombreux résidus de lysine et d'arginine qui permettent leur liaison avec l'ADN. Ainsi, la position de l'ADN sur les octamères d'histones conditionne son accessibilité : **plus la chromatine est compacte, moins elle sera accessible à certains TFs.**

L'octamère d'histone, formant le cœur du nucléosome, est constitué de 2 exemplaires des histones suivantes : H2A, H2B, H3 et H4. Ces octamères d'histones sont ensuite regroupés en chromatosome grâce à un « *linker* », l'histone H1, qui se fixe à l'ADN entre deux nucléosomes et aide à la compaction de l'ADN (Lai & Pugh, 2017). Les chromatosomes se replient en boucles, puis, en fibres afin de s'enrouler de

manière plus étroite et former les chromosomes (**Figure 21**). L'expression génique dépend de plusieurs facteurs incluant :

- La structure locale de la chromatine (état ouvert ou fermé de la chromatine, interactions promoteur/enhancer, présence de domaines topologiquement associés (« *topologically associating domains* », TAD) ;
- Les modifications épigénétiques (méthylation de l'ADN, modification post-traductionnelles des histones) ;
- Les ARNs non-codants, système de régulation complexe et dynamique pouvant déstabiliser les ARNs messagers (ARNm) et limiter la quantité de protéines.

La transcription est, donc, **régulée à différents niveaux grâce à l'action de différents types de modulateurs** qui influencent l'interprétation de l'ADN et assurent le maintien de l'identité cellulaire, c'est-à-dire, le maintien du transcriptome spécifique de la cellule.

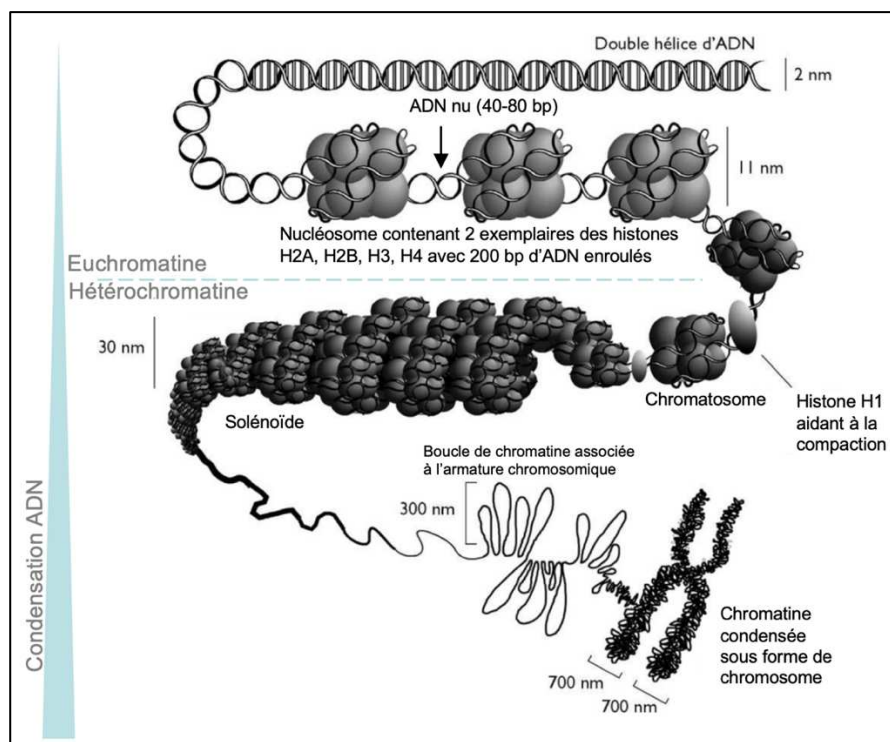


Figure 21 : Niveaux de condensation de la chromatine.

La structure locale de la chromatine permet la régulation de différents mécanismes cellulaires dont la transcription. L'euchromatine est le niveau de condensation le plus faible. On retrouve de l'ADN nu entre 2 nucléosomes. La compaction de l'ADN est possible grâce à l'histone H1, linker. Adaptée de (Legastelois et al., 2017)

Les différents états chromatinien :

La chromatine est une **structure tri-dimensionnelle complexe avec une organisation à différentes échelles**. La position des nucléosomes influence ces différents niveaux d'organisation de manière plus ou moins directe.

Dans les cellules, la chromatine se présente principalement sous **deux niveaux de condensation**, rendant l'ADN plus ou moins **accessible à la machinerie transcriptionnelle**. Ces deux états se distinguent par différents facteurs dont la position et de la densité de nucléosomes, la position dans le noyau et le type de modifications épigénétiques (Z. Chen et al., 2017). Les différents états chromatinien régulent également **l'accessibilité de l'ADN aux protéines de régulation de la transcription**, notamment aux TFs, co-facteurs et enhancers (dont les fonctions sont détaillées dans le chapitre **Introduction 3.1.b p.87**). L'état chromatinien peut également limiter l'ouverture du corps de gène via la fixation de silencer (qui répriment la transcription et favorisent un état chromatinien fermé) ou d'insulateurs (qui bloquent l'interaction promoteur/enhancer) (Chatterjee & Ahituv, 2017).

La modification de la structure et de la position de ces nucléosomes est assurée par **l'activité de deux principaux complexes protéiques** capables de les faire glisser :

- Le complexe Switch mutant (SWI)/ Sucrose non-fermenté (SNF) qui perturbe l'interaction ADN/histone pour remodeler les nucléosomes et changer les séquences transcrites (Bazett-Jones et al., 1999) ;
- Le complexe imitation switch (ISWI) qui régule l'espacement entre les nucléosomes, paramètre essentiel à l'assemblage de la chromatine et à son repliement en structures de plus haut niveau (Vincent et al., 2008).

L'organisation de cette structure est également influencée par la présence des TAD qui peuvent limiter les interactions entre les domaines de *cis*-régulation et les gènes cibles (détaillé dans le chapitre **Introduction 3.1.b.ii p.90**).

Les modifications épigénétiques :

L'épigénome correspond aux **modifications chimiques et structurales au niveau de l'ADN et des histones**. Par exemple, la méthylation des résidus de cytosine réprime fortement la transcription lorsqu'elle se situe sur les séquences répétées du génome contribuant ainsi au maintien de son intégrité (Quadrana et al., 2016). Au contraire, lorsque la méthylation est localisée dans le corps du gène, elle est associée à une expression génique plus importante (Teixeira & Colot, 2009).

Il est actuellement bien admis que les modulations épigénétiques contribuent au développement de nombreuses maladies. Dans la suite de ce manuscrit, seules les modifications post-traductionnelles (« *post-translational modifications* », PTMs) des histones et leur lien avec la régulation de la transcription seront développés.

ii. Modifications post-traductionnelles (PTMS) des histones

Les **histones** peuvent être la cible d'un **grand nombre de PTMs**, notamment au niveau de leur queue N-terminale. Ces modifications peuvent favoriser ou réprimer la transcription selon les résidus qu'elles ciblent.

Ces modifications sont **distribuées sur l'ensemble du génome** afin de réguler l'ouverture de la chromatine et le recrutement des protéines de la machinerie transcriptionnelle. Elles **ciblent différents types de résidus** (lysine, arginine, sérine, ...) (Kouzarides, 2007). Parmi ces PTMs, on retrouve la méthylation, l'acétylation, la phosphorylation, l'ubiquitination, la SUMOylation ou encore la O-linked N-acétylglucosamine (O-GlcNAcylation). Le **degré de modification des résidus peut varier**, par exemple, les lysines peuvent être mono-, di- ou encore tri-méthylées. Ce degré de modification impacte de manière différente l'expression génique. Par exemple, la mono-méthylation de la lysine 4 de l'histone 3 (H3K4me1) est caractéristique des enhancers compétents et corrélée à une expression génique tissu spécifique, tandis que la tri-méthylation de cette lysine (H3K4me3) est retrouvée au niveau des promoteurs et favorise l'initiation du complexe de transcription (Collins et al., 2019, p. 4).

Les PTMs des histones **influencent la fixation, la localisation, l'activité et l'interaction des TFs et des cofacteurs** (Weidemüller et al., 2021). En fonction de l'état chromatinien et des marques PTMs au niveau des promoteurs, la transcription peut être favorisée ou réprimée. Par exemple, la tri-méthylation de H3K27 peut être retrouvée dans l'euchromatine ou l'hétérochromatine et sera associée soit à des promoteurs bivalents pouvant facilement être transcrits, ou à des promoteurs réprimés (**Tableau 1**).

L'ensemble des différentes PTMs des histones associées à un locus de gène constituera **l'état chromatinien de ce locus**. Pour un même locus, cet état peut **varier d'un tissu à un autre** (Bowman & Poirier, 2015). Ainsi, le profil combiné des marques épigénétiques caractéristiques correspondant à un locus de gènes permet **d'identifier des gènes importants pour l'identité de la cellule** (Dubois-Chevalier et al., 2023). De plus, plusieurs PTMs peuvent co-exister simultanément sur une histone, ce qui augmente la complexité à l'échelle du nucléosome. Cependant, même si le rôle individuel des PTMs est bien décrit, l'impact des modifications multiples sur une même histone sur la transcription n'est pas encore bien compris.

Marque d'histone	Région ciblée	Transcription
H3K4me1	Région 3' et 5' – Enhancer compétent	Active
H3K4me3	Promoteur et 5' du gène	Favorise l'initiation du complexe de transcription
H3K9me3	5' et 3' du gène	Réprimée
H3K27ac	Promoteur du gène - Enhancer actif	Active
H3K27me3	Promoteur du gène et 3' du gène	Réprimée
H3K36me3	Corps du gène	Active

Tableau 1 : Focus sur la localisation et l'effet sur la transcription de certaines modifications post-traductionnelles (PTMs) des histones.

Seules les marques d'histones intéressantes pour le projet sont citées dans ce tableau. D'après (Millán-Zambrano et al., 2022)

L'ajout ou l'hydrolyse de ces modifications sont assurés par des protéines dites « *writers* », comme les histones méthyltransférases, ou à l'inverse par des « *erasers* » comme les histones déméthylases. Il existe également des protéines de types « *readers* » ou régulateur de la transcription qui vont reconnaître et se fixer à ces résidus modifiés pour influencer l'état chromatinien (**Figure 22**). **L'action coordonnée**

de ces protéines est impliquée dans la coordination entre épigénome et transcriptome (Tessarz & Kouzarides, 2014) (Zeisel et al., 2021).

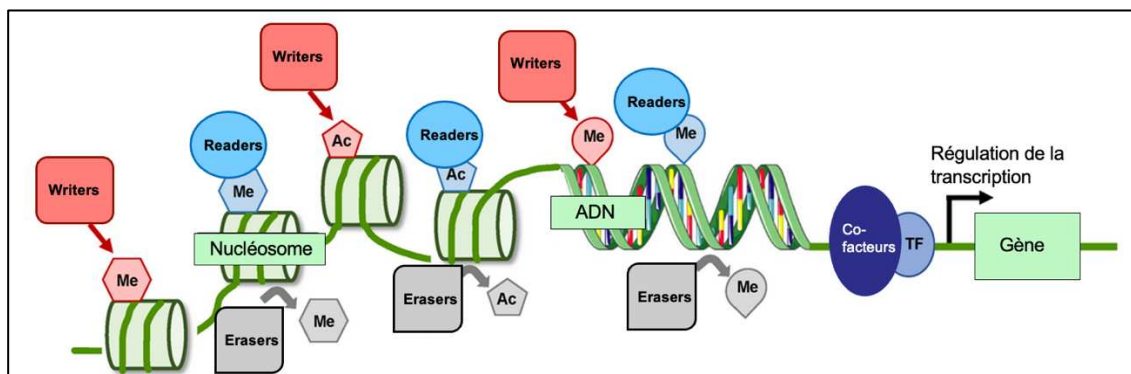


Figure 22 : Relation fonctionnelle entre régulateurs de la transcription et chromatine.

Ac : acétylation ; Me : méthylation ; TF : facteur de transcription ; « writers » : protéine permettant l'ajout des PTMs ; « erasers » : protéines réalisant une hydrolyse des PTMs ; « readers » : régulateurs de la transcription qui reconnaissent les résidus modifiés. Adaptée de (Zeisel et al., 2021)

iii. Le cas particulier du complexe Polycomb2 (PrC2)

Parmi tous ces types de PTMs des histones, la tri-méthylation de la lysine 27 de l'histone 3 (H3K27me3) est une marque épigénétique réversible associée à la répression de la transcription qui a fait l'objet d'une partie de mes travaux de thèse. Elle est apposée par les histones méthyltransférases enhancer de zeste 1 (« *enhancer of zeste 1 polycomb repressive complex 1* », EZH1) et EZH2 appartenant au complexe Polycomb 2 (PrC2). Cependant, il est important de noter que chez les mammifères, il existe deux complexes PrCs : le complexe PrC1 et le complexe PrC2, réprimant tous les deux la transcription avec des mécanismes et des cibles différents.

Le cœur protéique du complexe PrC2 est composé de différentes sous-unités, notamment **une unité catalytique et une unité régulatrice (Figure 23)**. L'unité catalytique contient les protéines EZH1 et EZH2, tandis que l'unité régulatrice est composée des protéines suppresseurs de l'homologue Zeste 12 (« *suppressor of zeste homolog 12* », SUZ12), du facteur de développement de l'ectoderme embryonnaire (« *embryonic ectoderm development* », EED) et des protéines de fixation associées aux rétinoblastomes 46 et 48 (« *retinoblastoma binding protein 46/48* », RbAp46/48).

Cette **unité régulatrice** joue un rôle clé dans le recrutement du complexe **PrC2** sur ses cibles et dans la modulation de son activité enzymatique (Zeng et al., 2022). Cette modulation dépend des partenaires qui s'y associent (Holoch & Margueron, 2017).

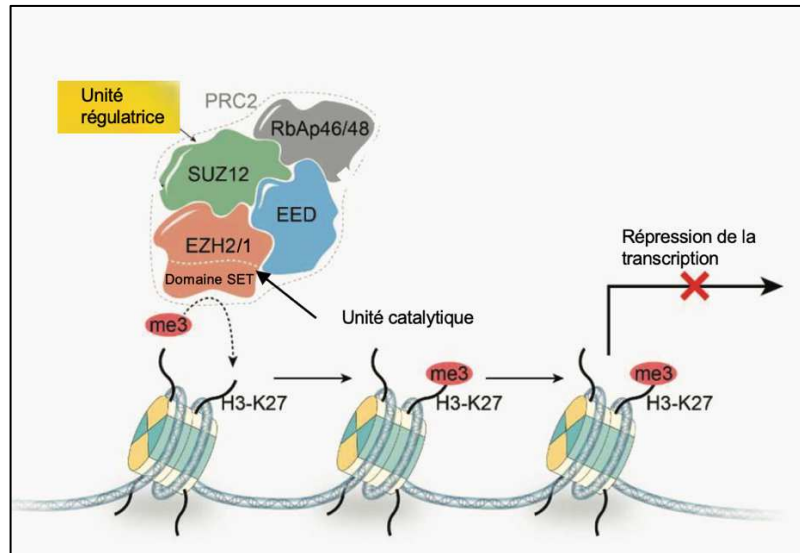


Figure 23 : Composition du complexe PrC2.

Adaptée de (Zeng et al., 2022)

L'**activité enzymatique** de PrC2 repose principalement sur les **enzymes EZH1 et EZH2 qui ont une fonction méthyltransférase redondante** (Chittock et al., 2017). En effet, il a été montré, dans un modèle murin de délétion (« *Knock-out* », KO) des gènes codants pour les enzymes *EZH1* et/ou *EZH2*, que la perte d'activité de l'une des deux enzymes peut être compensée par l'activité de l'autre enzyme (Grindheim et al., 2019). Par ailleurs, l'interaction de ces 2 enzymes avec les autres protéines de l'unité régulatrice - SUZ12, EED ou RbAp46/48 - est indispensable à l'activité de méthyltransférase du complexe (Montgomery et al., 2005) (Pasini et al., 2004). Certaines protéines comme la protéine 3 contenant le domaine Jumonji (« *Jumonji domain-containing protein-3* », JMJD3) et la protéine de répétition tétracopeptide transcrite de manière ubiquitaire à partir du chromosome X (« *Ubiquitously transcribed tetratricopeptide repeat X chromosome* », UTX) ont quant à elles une activité H3K27me3 histone déméthylase (Miller et al., 2010).

Le **recrutement et les mécanismes de régulation de l'activité** du complexe PrC2 ne sont pas encore bien décrits dans la littérature. Certaines études indiquent que le complexe se **fixe préférentiellement au niveau de régions pauvres en îlots de CpGs**, présentant une conformation tri-dimensionnelle particulière de l'ADN, intermédiaire entre euchromatine et hétérochromatine. L'activité du complexe est importante **en l'absence de transcription des gènes** et alors que les régions ne sont **pas occupées par PrC1**. Cette spécificité de ciblage est assurée par l'unité régulatrice de PrC2, qui joue également un rôle crucial dans la fixation à plus long terme du complexe. Ce maintien durable permet de conserver la marque épigénétique de répression sur la chromatine, préservant ainsi l'identité cellulaire au cours des divisions cellulaires (Laugesen et al., 2019) (Kouznetsova et al., 2019).

Enfin, la stabilité et l'activité du complexe PrC2 sont très **dynamiques** et **dépendent de la conformation tri-dimensionnelle de la chromatine** et du **contexte cellulaire** (Guo et al., 2021). Le complexe PrC2 et la marque H3K27me3 sont importants dans les deux états de la chromatine (euchromatine et hétérochromatine). D'un côté, la présence de PrC2 dans l'hétérochromatine permet la répression des gènes selon le statut de la cellule. Ainsi, ce complexe bloque la transcription des gènes qui ne participent pas à l'identité cellulaire. De l'autre, les niveaux de PrC2 peuvent être corrélés avec la présence de H3K4me3, et donc, à la présence de promoteurs bivalents dont la transcription est activée selon l'activité de la cellule. Par exemple, durant le cycle cellulaire, PrC2 joue un rôle important en réprimant différents groupes de gènes cibles selon les phases du cycle (Bemer & Grossniklaus, 2012) (Jadhav et al., 2016).

b. Généralités sur les TFs et régulation transcriptionnelle

La régulation de l'expression génique s'appuie également sur un réseau de TFs qui selon les contextes épigénétiques au niveau des promoteurs de leurs gènes cibles, peuvent favoriser ou inhiber leurs expressions (Pekowska et al., 2011). De plus, cette régulation de la transcription implique aussi différents éléments régulateurs pouvant être proximaux ou distaux du site d'initiation de la transcription (« *transcription start site* », TSS).

i. Fixation à la chromatine et activité des TFs

Initialement, le terme « facteur de transcription » désigne une protéine capable **de se fixer de manière spécifique à une séquence d'ADN pour réguler les niveaux de transcription, en recrutant différentes protéines effectrices secondaires** (Vaquerizas et al., 2009). Les TFs se fixent sur des séquences particulières, les motifs, situés au niveau des régions régulatrices de leurs gènes cibles et interagissent avec le complexe de transcription pour en permettre l'activation ou l'inhibition (**Figure 24**). Cependant, la seule connaissance du motif de liaison du TF n'apporte pas d'information sur son rôle fonctionnel (Zeitlinger, 2020). En effet, déterminer le lien entre les motifs de fixation de ces TFs et ceux présents au niveau des promoteurs des gènes reste un défi majeur essentiel pour comprendre comment les TFs régulent l'expression de leurs gènes cibles (Slattery et al., 2014).

La fixation des TFs est **régulée à différents niveaux** incluant :

- L'accessibilité des motifs de liaison qui dépend de la compaction de la chromatine ;
- La stabilité des TFs influencée par leurs PTMs ;
- Leurs interactions avec des co-facteurs (Lambert et al., 2018).

Pour accéder et se lier à leurs sites de liaison sur l'ADN, certains TFs entrent en compétition, d'autres ont la capacité d'initier le déplacement des nucléosomes (Slattery et al., 2014). Par exemple, certains TFs, comme FOXA1 peuvent initier le déplacement ou modifier la conformation des nucléosomes en recrutant différents remodeleurs de la chromatine ^{et/ou} grâce à la coopération avec d'autres TFs (Iwafuchi-Doi et al., 2016). Une fois liés à l'ADN, les **TFs peuvent agir de manière indirecte sur l'état de condensation de l'ADN** en modifiant certaines marques d'histones, ou de manière directe en recrutant le complexe de transcription (Lambert et al., 2018). Ces actions peuvent avoir des **effets activateurs ou répressifs** sur la transcription des gènes cibles. Par exemple, certains TFs comme la protéine de fixation au niveau de la boîte TATA (« *TATA-box binding protein* », TBP) vont directement activer l'ARN polymérase et le processus de transcription (Reiter et al., 2017).

Un point important est que les éléments de régulations *cis*, qu'ils soient proximaux ou distaux des TSSs, ont des effets différents sur l'assemblage du nucléosome et la fixation des TFs selon les informations reçues par la cellule (processus détaillé dans le chapitre **Introduction 3.1.b.ii** p.90). Selon le contexte cellulaire, les nucléosomes peuvent masquer certains sites de liaison, empêchant ainsi la fixation des TFs sur ces éléments *cis* et modifiant les niveaux d'expression des gènes cibles (Barozzi et al., 2014).

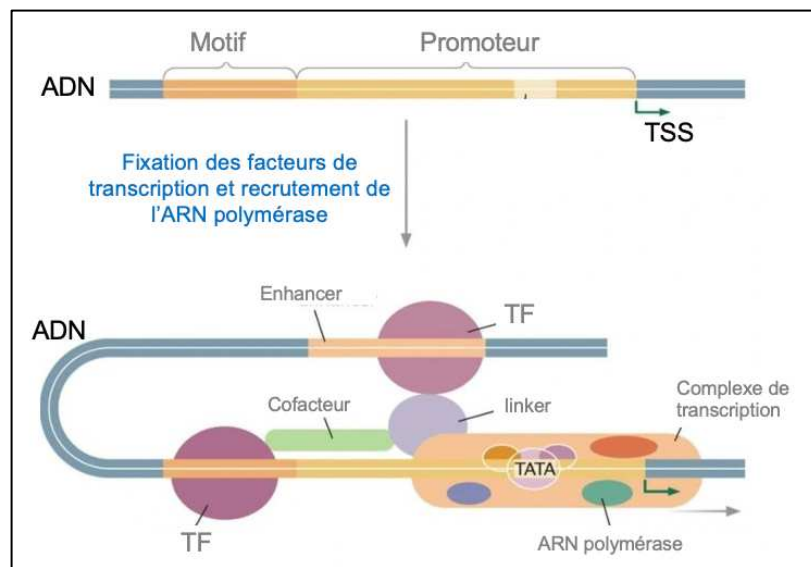


Figure 24 : Fixation des TFs et initiation de la transcription.

TF = facteur de transcription ; TSS = site d'initiation de la transcription. Adaptée de (Reiter et al., 2017)

Les **TFs collaborent**, également, entre eux de différentes manières. Ils peuvent former des complexes di- ou trimériques par interactions protéines-protéines afin de renforcer leur stabilité et leur fixation à l'ADN. Certains TFs modulent indirectement l'état de la chromatine, facilitant ainsi la fixation d'autres TFs sur celle-ci (Mitsis et al., 2020). Cette coopération entre différents TFs est caractéristique d'un état cellulaire.

D'un point de vue conceptuel, la fixation d'un TF permet la modulation de l'expression génique de ses gènes cibles. Or, c'est plutôt **l'activité combinée des TFs** et de leurs gènes cibles – appelée réseau d'activité des TFs – qui **assure le maintien transcriptionnel spécifique** à un type cellulaire donné ainsi que la régulation de la signalisation (Lambert et al., 2018). En effet, certains TFs peuvent être exprimés dans plusieurs types cellulaires tandis que d'autres ont une expression restreinte à un

nombre limité. Ainsi, **chaque type cellulaire** exprime un **réseau de TFs qui lui est spécifique** et d'autres **TFs sont plus ubiquitaires**, présents dans la majorité des tissus (Dubois-Chevalier et al., 2017). La fixation spécifique des TFs est directement impliquée dans la régulation de différentes voies de signalisation. En se liant à des éléments de réponse ou aux promoteurs des gènes cibles, ils assurent la transduction de signaux et modulent l'expression génique dans différents contextes physiopathologiques (Wilkinson et al., 2017). En particulier, dans le foie, le réseau des TFs essentiels à la fonction hépatique présente des interconnexions importantes contrôlant l'identité hépatocytaire et exerçant une boucle de régulation transcriptionnelle. Ces boucles de régulation sont perturbées en cas de blessure hépatique, entraînant une perte de l'identité cellulaire (Dubois-Chevalier et al., 2023).

ii. Autres modes de régulation de la transcription : les enhancers

La régulation transcriptionnelle passe aussi par l'intermédiaire de différents éléments régulateurs pouvant être proximaux ou distaux du TSS. Ces deux types d'éléments se distinguent par des PTMs d'histones différentes (Medina-Rivera et al., 2018). En effet, **les éléments régulateurs distaux, les enhancers actifs**, sont caractérisés par des marques de type H3K4me1 et H3K27ac, alors que les **éléments proximaux**, c'est-à-dire les promoteurs, sont principalement associés aux marques H3K4me3 et H3K27ac. Cependant, certaines études ont montré que les niveaux de H3K4me3 peuvent également être liés à une activité enhancer (Pekowska et al., 2011).

Le regroupement des éléments régulateurs est permis par le **repliement de la chromatine sous la forme de « hub »**, permettant le rapprochement des enhancers, les TFs et les co-facteurs dans un contexte épigénétique favorable à la transcription (**Figure 25**) (Shukla et al., 2022). En effet, l'ouverture de la chromatine et le rapprochement des enhancers et des promoteurs permet une transcription du gène cible (Field & Adelman, 2020). Ces éléments de régulation agissent ensemble : les enhancers permettent d'activer les promoteurs, notamment grâce à la fixation des TFs, qui eux favorisent la transcription du gène cible lorsque le contexte épigénétique est favorable (Andersson, 2015). Lorsque l'architecture de la chromatine est modifiée,

particulièrement au cours de la différenciation cellulaire, l'activité des enhancers et des promoteurs est donc également modifiée. De plus, l'activité des modules de régulation *cis* de la transcription (« *cis-regulatory modules* », CRMs) peut être modifiée en cas de mutations dans les régions d'ADN fixant les enhancers ou les promoteurs, ce qui peut mettre en péril la régulation transcriptionnelle (Chatterjee & Ahituv, 2017).

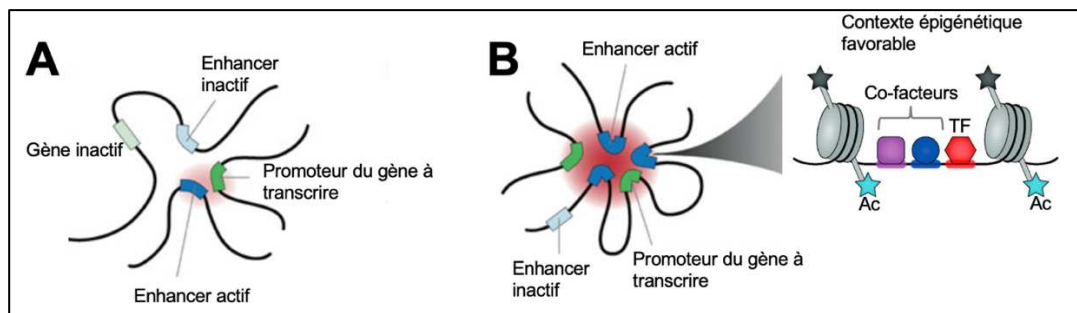


Figure 25 : Contexte épigénétique, recrutement des éléments régulateurs et transcription.

(A) Schématisation d'un enhancer en interaction avec un promoteur grâce à la structure chromatinienne favorable
(B) Schématisation d'une architecture regroupant différents acteurs co-recrutés pour mettre la transcription du gène cible. Ce regroupement s'accompagne d'un contexte épigénétique favorable à la transcription, notamment avec la présence de marques épigénétiques d'activation de la transcription. Ac : Acétylation. Adaptée de (Andersson, 2015)

En plus des enhancers, certaines régions de fixation des éléments de la transcription telles que les sites « *CCCTC-binding factor* » (CTCF) ou les insulateurs, jouent un rôle important dans l'expression spécifique de certains gènes. En effet, selon le contexte épigénétique, l'activité des enhancers permet la fixation des TFs et cofacteurs aux insulateurs qui sont spécifiques à un type cellulaire et ont une influence sur la transcription (Heintzman et al., 2009).

3.2. Contrôle transcriptionnel de la plasticité hépatocytaire après une blessure hépatique chronique

a. Modulations épigénétiques et plasticité des cellules épithéliales du foie

i. Rôle de l'épigénome dans la régénération hépatique en cas de maladies chroniques

La reprogrammation génique s'accompagne d'une **modification de la compaction de la chromatine**, afin de faciliter l'accès du complexe de transcription, comme démontré par des analyses de séquençage de la chromatine accessible à la transposable (« *assay for transposase-accessible chromatin using sequencing* », ATAC-Seq) sur un modèle murin de CCL₄. En particulier, les régions chromatiniennes impliquées dans la **prolifération cellulaire seront plus ouvertes** que celles liées aux fonctions métaboliques (Z. Li & Sun, 2024) (Arechederra et al., 2020).

Ces changements sont permis par une modification de l'activité du complexe SWI/SNF. En particulier, la protéine A1 contenant un domaine d'interaction riche en AT (« *AT rich interactive domain 1A* », ARID1A) maintient les promoteurs dans un état activable, permettant le recrutement de YAP et leur activation en cas de blessure hépatique (**Figure 26**) (W. Li et al., 2019).

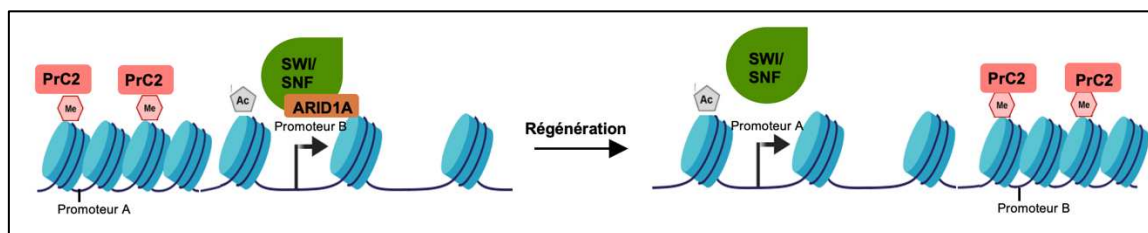


Figure 26 : Réorganisation de la chromatine lors d'une blessure hépatique.

Ac : Acétylation ; Me : méthylation ; PrC2 : Complexe Polycomb 2

Ainsi, l'induction des gènes, nécessaires à la régénération hépatique, est facilitée en cas de blessure hépatique (Bou Saleh et al., 2021) (Verma et al., 2021). Cependant, couplés aux signaux d'inflammation chroniques, il semblerait que ces changements transcriptionnels bloquent la transcription de gènes nécessaires aux fonctions

hépatiques. Cette redirection du complexe de transcription peut donc devenir délétère pour le foie (Caldez et al., 2018).

Au cours de la réparation tissulaire, la redistribution des marques épigénétiques entraîne l'installation de marques d'activation sur les gènes cholangiocytaires, illustrant la plasticité des cellules épithéliales et leur capacité de réponse en cas de blessure (Huyghe et al., 2024).

En plus, ce remodelage épigénétique entraîne une modification de l'activité des enzymes « writers » et « erasers » et la distribution de certaines marques épigénétiques, notamment H3K27ac et H3K27me3, tout au long du génome (Verma et al., 2021). En particulier, la variation de la quantité de donneurs de groupements méthyles affecte l'activité des histone méthyltransférases (Reid et al., 2017). Il existe, donc, un lien étroit entre modifications métaboliques et reprogrammation épigénétique dans les hépatocytes au cours de la régénération du foie. Couplée à cette modification, il a été montré que certains patients atteints d'ALD ont une déficience de certaines enzymes nécessaires à la méthylation des histones. Ainsi, au-delà du remodelage classique lié à la blessure hépatique, l'activité anormale de ces enzymes entraîne soit une suractivation, soit une altération de l'expression de gènes spécifiques, et *in fine* une activation excessive de certaines voies de signalisation, impliquées dans la régénération cellulaire (Mandrekar, 2011). En particulier, dans les cas de DR, cette déficience favorise l'activation de la voie Hippo/YAP impliquée dans l'HCT (Aloia et al., 2019).

ii. Focus sur PrC2 et plasticité cellulaire

Comme expliqué précédemment, les cibles réprimées par le complexe PrC2 varient au cours du cycle cellulaire, montrant sa **dynamique et sa capacité d'adaptation aux besoins changeants de la cellule** (Oliviero et al., 2016). Ainsi, son intégrité assure le maintien des programmes transcriptionnels spécifiques à chaque type cellulaire (German & Ellis, 2022).

Dans un hépatocyte sain, le complexe PrC2 régule l'expression des gènes associés à la maturation des hépatocytes et prévient la fibrogenèse en agissant au niveau des promoteurs de ces gènes situés dans l'euchromatine (Grindheim et al., 2019). Une altération de son activité entraîne une **perte de fonction hépatique, une augmentation de la fibrose et l'apparition de nodules de régénération** (Grindheim et al., 2019). Il a été montré que les souris constitutivement délétées des gènes *EZH1* et *EZH2* dans le foie perdent leur capacité régénérative comparées aux souris contrôles (N. Yang et al., 2023). Après un traitement chronique de CCl₄, les foies des souris dépourvues de l'unité catalytique de PrC2 présentent des lésions hépatiques plus importantes et une régénération compromise (Bae et al., 2015).

Cependant, le rôle de PrC2 lors des changements transcriptionnels au cours de l'HCT n'est pas bien documenté.

L'épigénétique, dans son ensemble, joue un rôle important dans la plasticité cellulaire et les mécanismes de régénération. Ce processus implique non seulement les modifications de l'ADN et des histones, mais aussi l'action concertée de TFs capables de remodeler les programmes géniques.

b. Modulation de l'activité des TFs après la blessure hépatique

En réponse aux signaux inflammatoires, présents en cas de blessure hépatique, l'activité de certains TFs est modifiée (S. Hu et al., 2025). En effet, l'inflammation chronique et la présence de toxine **perturbent la fixation et la localisation des TFs hépatocytaires**. Par conséquent, le réseau spécifique de TFs hépatiques est également modifié, entraînant la perte du phénotype hépatocytaire avec la **diminution de l'expression des TFs nécessaires au maintien de l'identité hépatocytaire** (**Figure 27A**) (Feng et al., 2023) (Sun et al., 2016) (Dubois, Gheeraert, et al., 2020).

Ce changement d'activité des TFs est en lien avec la diminution de **l'accessibilité de la chromatine au niveau des CMRs hépatocytaires** (Snykers et al., 2009). Plus spécifiquement, la perte d'activité des TFs hépatocytaires comme HNF4A limite l'ouverture de la chromatine et baisse les niveaux de H3K27ac aux

promoteurs des gènes spécifiques des hépatocytes, responsables de la différenciation des cellules (Dubois, Staels, et al., 2020).

De plus, comme les TFs ont une action hiérarchique et interactive, principalement supportée par les TFs « essentiels » comme FOXA2, la modification d'activité de ces protéines entraîne un **changement d'activité des boucles de rétroactions pour permettre la prolifération cellulaire** (Dubois-Chevalier et al., 2018) (Iwafuchi-Doi et al., 2016) (Kymizi et al., 2006).

Parallèlement, le **stress cellulaire** et le **remodelage des nucléosomes** permettent la mise en place d'un phénotype cellulaire mixte, caractérisé par **l'expression simultanée de marqueurs cellulaires** qui ne sont normalement pas exprimés en même temps dans une cellule saine (**Figure 27B**) (Huyghe et al., 2024). Ainsi, dans les stades avancés de certaines pathologies hépatiques comme les MASLDs ou les ALDs, certains hépatocytes expriment des TFs comme SOX4 ou KRT19, absents à l'état basal dans un foie sain (Gribben et al., 2024) (Dubuquoy et al., 2015).

Ces changements activent des voies de signalisation nécessaires à la régénération cellulaire (Davies et al., 2023). Par exemple, le gène codant pour le TF à domaine « Runt » 1 (« *RUNX family transcription factor 1* », *RUNX1*) est induit, initiant la réparation cellulaire (Huyghe et al., 2024).

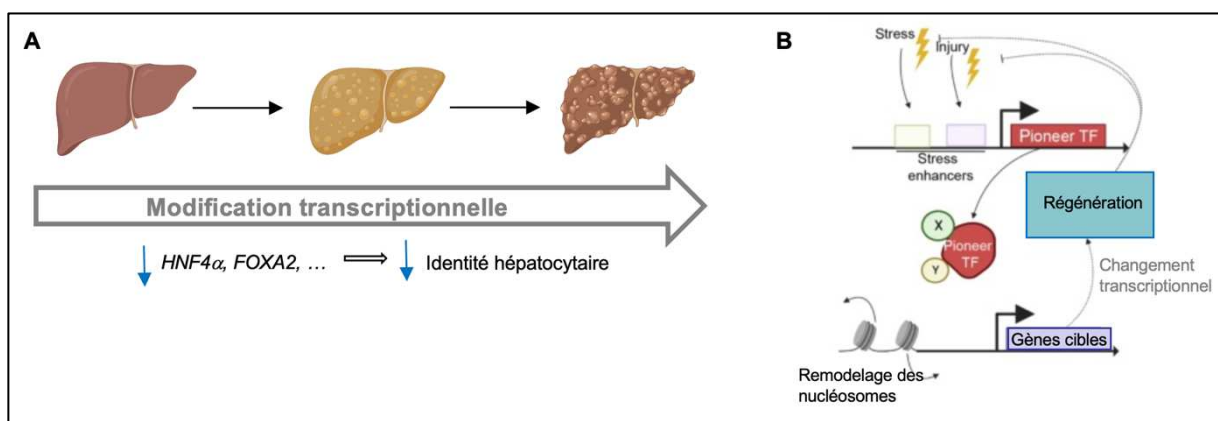


Figure 27 : Modification d'activité des TFs après une blessure hépatique.

(A) Boucles de rétroactions modifiant l'activité des TFs en cas de blessure hépatique. (B) Dérégulation transitoire des TFs essentiels à cause de blessure ou d'un stress cellulaire entraînant une identité cellulaire mixte (noté X et Y).

Y). Ces événements contribuent à un remodelage des nucléosomes permettant l'activation de gènes impliqués dans la régénération. D'après (Huyghe et al., 2024) et (Sun et al., 2016)

À l'inverse, les mécanismes transcriptionnels impliqués dans l'activation génique au cours de l'HCT demeurent peu documentés. Seuls quelques marqueurs ont été identifiés dans la littérature comme représentatifs de ce processus.

Objectif de la thèse

En résumé, le foie est un organe central de l'homéostasie énergétique, impliqué dans de nombreuses fonctions. Ces fonctions sont acquises pendant l'organogenèse, lorsque les hépatoblastes se différencient en hépatocytes sous l'influence de TFs spécifiques à chaque stade. Un autre type cellulaire dérivé des hépatoblastes est le cholangiocyte. Ensemble, les hépatocytes et les cholangiocytes constituent les cellules épithéliales du foie, avec des caractéristiques fonctionnelles différentes. Cependant, du fait de leurs fonctions, ces cellules épithéliales sont exposées à un grand nombre d'agressions, aiguës ou chroniques, pouvant entraîner une défaillance de leurs fonctions. Pour éviter la défaillance totale de l'organe, le foie possède des capacités de régénération importantes, essentiellement permises par la plasticité de ses cellules épithéliales.

Les processus de régénération cellulaire sont sous le contrôle d'un réseau complexe de TFs qui agissent par l'intermédiaire de CRMs hépatocytaires, caractérisés par une ouverture de la chromatine et la présence de marques d'acétylation des histones comme H3K27ac. Or, la blessure hépatique entraîne une diminution de l'activité de ces TFs, une réduction de l'accessibilité des CRMs hépatocytaires et *in fine* une perte d'identité hépatocytaire. À l'inverse, les TFs induits pendant le mécanisme de l'HCT sont peu documentés. Seule l'augmentation des niveaux de certains TFs comme SOX9 et HNF1B a été rapportée dans la littérature, les constituants comme marqueurs de l'HCT.

Compte-tenu de l'importance du phénomène de l'HCT en particulier dans les maladies hépatiques chroniques et de la faible compréhension des mécanismes transcriptionnels sous-jacents, l'objectif de mon projet de thèse est de caractériser ces événements transcriptionnels impliqués dans cette reprogrammation hépato-cholangiocytaire en :

- Identifiant les TFs impliqués dans l'HCT et en définissant leurs importances fonctionnelles ;
- Caractérisant les mécanismes épigénétiques impliqués dans l'activation de ces TFs ;

- Déterminant la pertinence des TFs et des mécanismes identifiés vis-à-vis de la dérivation hépato-cholangiocytaire dans certaines maladies hépatiques chroniques humaines (**Figure 28**).

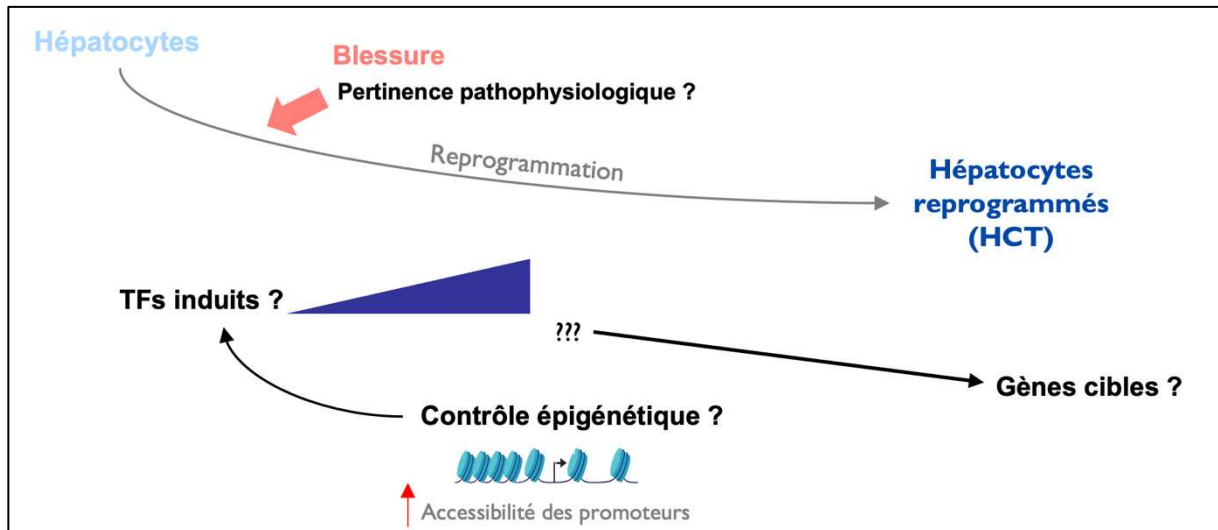


Figure 28 : Bilan du contexte et principaux axes du projet de thèse.

La littérature décrit que le processus d'HCT s'accompagne de la diminution de l'expression du programme spécifique hépatocytaire. L'objectif de mon projet de thèse est de caractériser les événements transcriptionnels induits lors de l'HCT ainsi que les mécanismes épigénétiques sous-jacents cette activation. De plus, mon travail a permis de mettre en évidence le rôle clé d'un facteur de transcription capable d'activer un programme épithélial de cellules monopolarisées, réprimé dans un hépatocyte sain.

Ainsi, dans un premier temps, l'équipe a élaboré un atlas regroupant plusieurs données de scRNA-Seq, qui a permis de montrer que le phénomène de l'HCT se produit sans passer par un état cellulaire proche d'une cellule embryonnaire (hépatoblaste). Les liens entre l'HCT et l'induction d'une signature épithéliale monopolarisée ont été évalués grâce à des analyses bio-informatiques. Par ailleurs, la réanalyse de données de RNA-Seq, couplée à la mise en place d'un modèle *in vitro* d'inhibiteurs pharmacologiques d'EZH1/EZH2, indiquent le rôle clé de PrC2 dans l'inhibition de l'expression des gènes du programme épithélial dans l'hépatocyte sain. Enfin, l'impact de l'induction d'un gène codant pour le TF *GRHL2* ainsi que l'activation d'un programme épithélial mono-polarisé ont été évalués avant d'étudier la pertinence de l'étude dans certaines pathophysiologies comme l'ADL et certaines cholangiopathies.

Matériels et méthodes

I. Analyses bio-informatiques

1.1. Ressources

a. Origine et récupération de l'ensemble des données publiques

Les données publiques utilisées pour mes travaux de thèse ont été téléchargées à partir des sites « *Gene expression Omnibus* » (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) (Edgar et al., 2002), ENCODE (<https://www.encodeproject.org>) (Davis et al., 2018), Roadmap Épigénome (<http://www.roadmapepigenomics.org>) (Bernstein et al., 2010), CistromDB (<http://cistrome.org/db/#/>) (Taing et al., 2024), ou encore à partir du « *Human Protein Atlas* » (<https://www.proteinatlas.org>) (Karlsson et al., 2021). Les données sont récapitulées dans le **Tableau 2**.

La liste des gènes codants pour les TFs murins a été obtenue sur la base de données AnimalTFDB 3.0 (<https://guolab.wchscu.cn/AnimalTFDB/#/>) (H. Hu et al., 2019). Les cofacteurs ont été retirés de cette liste.

Les coordonnées des sites d'initiation de la transcription (TSS) proviennent de l'étude de (J. Wang et al., 2018). Pour chaque gène, le TSS le plus actif dans les foies de souris a été considéré. Les coordonnées de TSS murins (mm10, coordonnées définies avec une fenêtre de +/- 500 bp autour du TSS) ont été converties grâce à l'outil « *LiftOver* » de UCSC (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>) (Perez et al., 2025) afin d'obtenir les coordonnées humaines correspondantes (hg38).

La liste de gènes dénotée Epith-Signature provient de (Breschi et al., 2020) en sélectionnant le type cellulaire « Epithelial » dans la table supplémentaire S7. Les « EnsemblID » ont été convertis en « Official Gene Symbol » grâce à l'outil « *Biotool* » (https://www.biotools.fr/human/ensembl_symbol_converter).

Les listes de gènes exprimés uniquement dans les hépatocytes ou les cholangiocytes, dénotés Chol-Enriched ou Hep-enriched, ont été téléchargées sur le site du « *Human Protein Atlas* » (<https://www.proteinatlas.org>) (Dusart et al., 2023).

b. Liste des données provenant de la littérature

Le tableau ci-dessous récapitule l'ensemble des données publiées qui ont été utilisées tout au long de ce projet de thèse. Il rassemble le type d'analyse, une description de l'étude, les tissus ou types cellulaires utilisés, les numéros d'accèsions et les références bibliographiques associés à chaque type de donnée utilisée.

Numéro d'accèsion	Description	Types cellulaires / tissus	Référence
<u>Données de scRNA-Seq utilisées pour l'atlas</u>			
CRA002445	Foie embryonnaire (E11, E11.5, E13, E14.5, E16, E17.5)	Hépatoblastes	Wang et al. Comparative of cell lineages differentiation during hepatogenesis in humans and mice at the single-cell transcriptome level. Cell Res 2020
GSE171993	Foie juste après naissance (D1, D3, D7, D21, D56)	Postnatal Hépatocytes et Cholangiocytes	Liang et al. Temporal analysis of postnatal liver development and maturation by single-cell transcriptomic. Dev. Cell 2022
GSE151309	Postnatal (D14); PHx (24, 48 and 96 h) et Foie adulte contrôle	(Postnatal) Hépatocytes	Chembazhi et al. Cellular plasticity balances the metabolic and proliferation dynamics of a regenerative liver. Genome Res 2021
https://doi.org/10.5281/zenodo.6035873	APAP en une seule injection (0, 24, 48, 72, 96, 168 hrs)	Hépatocytes ; Cholangiocytes	Ben-Moshe et al. The Spatiotemporal program of zonal regeneration following acute injury. Cell Stem Cell 2022
GSE136679	Contrôle et APAP e une seule injection (6, 24, 48, 96 hrs) and PHx (3, 48, 120 hrs)	Hépatocytes	Walesky et al. Functional compensation precedes recovery of tissue mass following liver injury. Nat Communication 2020
GSE186554	Contrôle + Injection poly(I:C) 7h + LPS 3 or 12h	Hépatocytes	Sun et al. Transcriptional switch of hepatocytes initiates macrophage recruitment and T-cell suppression in endotoxemia. J Hepatol 2022
GSE192742	Contrôle et régime occidental (24 or 36 semaines)	Hépatocytes ; Cholangiocytes	Guilliams et al. Spatial proteogenomics reveals distinct and evolutionarily conserved macrophage niches. Cell 2022
GSE158866	Contrôle et PhX 48h	Hépatocytes	Chen et al. Single-cell omics alysis reveals functional diversification of hepatocytes during liver regeneration. JCI Insight 2020
GSE125688	Contrôle et 1 semaine de DDC (0.1%)	Hépatocytes ; Cholangiocytes	Pepe-Mooney et al. ngle-Cell Analysis of the Liver Epithelium Reveals Dynamic Heterogeneity and an Essential Role for YAP in Homeostasis and Regeneration. Cell Stem Cell 2019
GSE157698	Contôle et 2-6 semaines de DDC (0.1%) en incluant les	Hépatocytes	Merrell et al. Dynamic transcriptional and epigenetic changes drive cellular plasticity in the liver. Hepatology 2021

	hépatocytes marqués YFP		
<u>Données de RNA-Seq provenant de foie de souris</u>			
GSE111447	Données de PRO-Seq normalisées par la taille du gène	Foie	Wang et al. Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. BMC Genomics 2018
GSE2222597	Cellules isolées	Hépatocytes, cholangiocytes	Zummo et al. A time- and space-resolved nuclear receptor atlas in mouse liver. Journal of Molecular Endocrinology 2023
GSE156894	Cellules isolées de souris sous régime 0,1% DDC pendant 4 semaines	Hépatocytes, cholangiocytes, HCT	Merrell & Peng et al. Dynamic transcriptional and epigenetic changes drive cellular plasticity in liver. Hepatology 2021
GSE108315	Cellules isolées provenant de souris délétées de certains effecteurs de la voie NOTCH	Hépatocytes	Schaub et al. De Novo formation of the biliary system by TGFβ-mediated hepatocyte transdifferentiation. Nature 2018
GSE118757	Cellules isolées provenant de souris délétées dans le foie de manière spécifiques de EZH1 et EZH2	Hépatocytes	Grindheim et al. Polycomb repressive complex 2 proteins Ezh1 and Ezh2 regulate timing of postnatal maturation and fibrosis by repressing genes with euchromatic promoters. Gastroenterology 2021
GSE281717	Surexpression de GRHL2 de manière ectopique dans les BMELs	BMELs	Cette étude
<u>Données de RNA-Seq provenant du « Human Protein Atlas »</u>			
GSE155960	/	Tissu adipeux	Hildreth et al. Single cell sequencing of human white adipose tissue identifies new cell states in health and obesity. Nature 2021
GSE159929 (GSM4850584 - GSM4850583 - GSM4850589 - GSM4850580 - GSM4850590)	/	Moelle épinière – Œsophage – Ganglion lymphatique – Estomac - Testicule	He et al. Single-cell transcriptome profiling an adult human atlas of 15 major organ. Genome Biology 2020
Allen brain		Cerveau	Allen brain map
GSE164898	Uniquement les données provenant des cellules glandulaires et myoépithéliales du sein	Sein	Bhat-Nakshatri et al. A single cell atlas of healthy tissues reveals clinically relevant clusters of breast epithelial cells. Cell Reports 2021
Fig11 981034	Uniquement les données des cellules respiratoires basales, cellules club et ionocytes	Bronches	Lukassen et al. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. The EMBO journal 2020
GSE116222	Uniquement les cellules du goblet intestinal, cellules de paneth et entérocytes distaux	Colon	Parikh et al. Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature 2019
GSE111976	Uniquement les données des cellules ciliées, glandulaires et lumbales	Endomètre	Wang et al. Single-cell transcriptomic atlas of the human endometrium during the menstrual cycle. Nature Medicine 2020
GSE137537	/	Œil	Menon et al. Single-cell transcriptomic atlas of the human retina identifies cell

			types associated with age-related macular degeneration. Nature Communication 2019
GSE178101	Uniquement les cellules sécrétoires	Trompe de Fallope	Ulrich et al. Cellular heterogeneity of human fallopian tubes in normal and hydrosalpinx disease states identified using scRNA-seq. Development Cell 2022
GSE109816	/	Coeur	Wang et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. Nature cell biology 2020
GSE131685	Uniquement les cellules du canal collecteur, cellules tubulaires distales et cellules tubulaires proximales	Rein	Liao et al. Single-cell RNA sequencing of human kidney. Scientific data 2020
GSE115469	Uniquement hépatocytes et cholangiocytes	Foie	MacParland et al. Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. Nature Communication 2018
Tabula sapiens	Pour les poumons : les cellules alvéolaires de type 1 et 2	Poumons – Prostate – Rectum – Muscle squelettique – Langue – Système vasculaire	THE TABULA SAPIENS CONSORTIUM. The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. 2022
E-MTAB-8381	Uniquement les données des cellules granuleuses	Ovaire	Wagner et al. Single-cell analysis of human ovarian cortex identifies distinct cell populations but no oogonial stem cells. Nature Communications 2020
GSE131886	Uniquement les cellules canalaire et glandulaires exocrines	Pancréas	Quadrir et al. Single-cell resolution analysis of the human pancreatic ductal progenitor cell niche. PNAS 2020
GSE112845	/	Cellules mononuclées du sang (PBMC)	Chen et al. PBMC fixation and processing for Chromium single-cell RNA sequencing. Journal of Translational Medicine 2018
E-MTAB-6701	/	Placenta	Vento-Tormo et al. Single-cell reconstruction of the early maternal–fetal interface in humans. Nature 2018
GSE125970	Uniquement les données des canaux et glandes salivaires	Glandes salivaires	Wang et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. The Journal of Experimental Medicine 2020
GSE143704	Uniquement les cellules kératinocytes suprabasales et basales	Peau	De Michelis et al. A reference single-cell transcriptomic atlas of human skeletal muscle tissue reveals bifurcated muscle stem cell populations. Skeletal Muscle 2020
GSE130973	Uniquement les entérocytes proximaux	Intestin	Solé-Boldo et al. Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. Communications Biology 2020
GSE125970	/	Rate	Wang et al. Single-cell transcriptome analysis reveals differential nutrient absorption functions in human intestine. The Journal of Experimental Medicine 2020

GSE120508	Uniquement les cellules mésothéliales	Thymus	Guo et al. The adult human testis transcriptional cell atlas. Cell Research 2018
<u>Données de RNA-Seq de foie humains</u>			
GSE167308	Données de groupes de patients des groupes contrôles et AH ont été utilisés	Hépatocytes microdisséqués	Bou Saleh et al. Loss of hepatocyte identity following aberrant YAP activation: A key mechanism in alcoholic hepatitis. Journal of Hepatology 2021
GSE281717	Inhibition de EZH1 et EZH2 par des composés pharmacologiques	HepG2	Cette étude
	Surexpression de GRHL2 de manière ectopique dans les BMELs	HepG2	Cette étude
GSE119168	Utilisation des hépatocytes contrôles et des hépatocytes Krt7+	Hépatocytes microdisséqués de patients AH	Aguilar-Bravo et al. Hepatocyte dedifferentiation profiling in alcohol-related liver disease identifies CXCR4 as a driver of cell reprogramming. Journal of Hepatology 2023
<u>Données de ChIP-Seq murin</u>			
ENCODE (GSE37074 - GSE31039 - GSE29184 - GSE31039)	DHS H3K4me3 H3K27ac H3K36me3	Foie de souris (DHS et H3K36me3) Hépatocytes isolés	ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012
CistromDB GSM3358216	H3K27me3	Hépatocytes isolés	Grindheim et al. Polycomb repressive complex 2 proteins Ezh1 and Ezh2 regulate timing of postnatal maturation and fibrosis by repressing genes with euchromatic promoters. Gastroenterology 2021
GSE48375	RNA-Polymerase II	Foie de souris	Gilardi et al. Genome-Wide Analysis of SREBP1 Activity around the Clock Reveals Its Combined Dependency on Nutrient and Circadian Signals. PLOS Genetics 2014
GSE185705	EZH2	Foie de souris	Yang et al. A hyper-quiescent chromatin state formed during aging is reversed by regeneration. Molecular Cell 2023
GSM751035	H3K9me3	Foie de souris	Tennant et al. Identification and analysis of murine pancreatic islet enhancers. Diabetologia 2013
GSE188027	Méthylation de l'ADN	Foie de souris	ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature 2012
PMID34226551	Domaines chromatinien définis en utilisant la combinaison de présence ou l'absence des signaux d'ATAC-Seq H2AZ et de ChIP-Seq H3K4me3, H3K27me3 et H3K9me3	Domaines regroupant les structures chromatiennes de foie de souris	Zhang et al. Chromatin states shaped by an epigenetic code confer regenerative potential to the mouse liver. Nat Commun 2021
GSE119218	Hétérochromatine, Euchromatine et état intermédiaire	Domaines définis par sonification-	Grindheim et al. Polycomb Repressive Complex 2 Proteins EZH1 and EZH2 Regulate Timing of Postnatal

		résistante chromatine (srHC-Seq)	Hepatocyte Maturation and Fibrosis by Repressing Genes With Euchromatic Promoters in Mice. Gastroenterology 2019
GSE53627	H3K27me3	ChIP-Seq H3K27me3 dans les foies de souris délétés constitutivement de EZH1 et EZH2	Bae et al. The methyltransferases enhancer of zeste homolog (EZH) 1 and EZH2 control hepatocyte homeostasis and regeneration. FASEB j. 2015
Données de ChIP-Seq humain			
ENCODE (ENCFF211ANT - ENCFF538BJP- GSM733737)	H3K27me3 H3K27ac H3K4me3	Foie	BI Human Reference Epigenome Mapping Project
Roadmap consortium (GSM621675- GSM669931)	H3K4me3 H3K26me3	Foie	Bernstein et al. The NIH Roadmap Epigenomics Mapping Consortium aims to produce a public resource of epigenomic maps for stem cells and primary ex vivo tissues selected to represent the normal counterparts of tissues and organ systems frequently involved in human disease. Nature Biotechnology 2010
GSE46194	GRHL2 ChIP-Seq réalisé sur les cellules HBE issues de 3 donneurs	Cellules épithéliales bronchiales humaines	Gao et al. Evidence for multiple roles for grainyhead-like 2 in the establishment and maintenance of human mucociliary airway epithelium. Proceedings of the National Academy of Sciences 2013
GSE71019	GRHL2	OVCA429	Chung et al. GRHL2-miR-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification. Scientific Reports 2016
GSE281717	H3K27ac	HepG2	Cette étude
	H3K27me3	HepG2	Cette étude
	Surexpression de GRHL2 dans	HepG2	Cette étude
Listes de gènes utilisés dans cette étude			
Table S7 - PMID32759341	Epith-Signature (Signature épithéliale défini par des données transcriptomiques dans 105 types de cellules primaires)	Humain	Breshi et al. A limited set of transcriptional programs define major cell types. Genome Research 2020
Human Protein Atlas	Chol-enriched genes (Définis parmi tous les types cellulaires de l'atlas)	Humain	Dusart et al. A tissue centric atlas of cell type transcriptome enrichment signatures. bioRxiv. 2023
GO:00613339	Établissement ou maintien de la monopolarité		GoTerm
GO:0090162	Établissement ou maintien du phénotype épithélial		GoTerm

Tableau 2 : Liste des données publiques utilisées dans cette thèse.

c. Outils utilisés lors de cette étude

Le tableau ci-dessous récapitule l'ensemble des outils bio-informatiques utilisés pour l'analyse des différents types de données réanalysées lors de ma thèse. Il indique les versions et les références de chaque outil utilisé. Les fonctions et les paramètres seront détaillés par la suite.

Nom de l'outil	Version	Référence
BedTools	2.30.0	Quinlan et al. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010
BiocManager	1.30.23	Morgan M, Ramos M (2025). BiocManager: Access the Bioconductor Project Package Repository.
Boot	1.3-30	Angelo Canty, B. D. Ripley (2024). boot: Bootstrap R (S-Plus) Functions
Bowtie2	2.5.3	Langmead et al. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012
Car	3.1-2	Fox J, Weisberg S (2019). An R Companion to Applied Regression, Third edition. Sage, Thousand Oaks CA
CellRank	2.0.0	Weiler et al. CellRank 2: unified fate mapping in multiview single-cell data. Nat Methods 2024
Chromap	0.1.3.R-256	Zhang et al. Fast alignment and preprocessing of chromatin profiles with Chromap. Nature Communications 2021
Circlize	0.4.16	Gu et al. "circlize implements and enhances circular visualization in R." Bioinformatics. 2014
ComplexHeatmap	2.15.4	Gu et al. "Complex heatmaps reveal patterns and correlations in multidimensional genomic data." Bioinformatic 2016
DAVID	Dec.2021 v20223q4	Sherman et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Research 2022
DeepTools	3.5.1	Ramirez et al. Deeptool2 : A next generation web server for Deep-Sequencing data analysis. Nucleic Acids Research 2016
DescTools	0.99.54	Signorell et al. DescTools: Tools for Descriptive Statistics (2025)
DeSeq2	1.42.1	Love et al. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 2014
DevTools	2.4.5	Wickham H, Hester J, Chang W, Bryan J (2022). devtools: Tools to Make Developing R Packages Easier. https://devtools.r-lib.org/ , https://github.com/r-lib/devtools .
Dplyr	1.1.4	Wickham et al. "The Split-Apply-Combine Strategy for Data Analysis." Journal of Statistical Software 2011
Epitools	0.5-10.1	Brown LD et al. Estimation d'intervalle pour une proportion. Science statistique 2001
Factoextra	1.0.7	Kassambra et al. Factoextra : Extract and vizualized the results of multivariate data

		analysis. https://github.com/kassambara/factoextra
FactoMineR	2.11	Lê et al. FactoMineR : An <i>R</i> -Package for Multivariate Analysis. J. Stat. Soft 2008
Galaxy		Afgan et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Research 2018
Ggplot2	3.5.1	Wickham et al. ggplot2: Elegant Graphics for Data Analysis. Springer 2009
Ggrepel	0.9.5	Slowikowski K (2024). ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. https://ggrepel.slowkow.com/ , https://github.com/slowkow/ggrepel .
Ggubr	0.6.0	Kassambara A (2023). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.6.0, https://rpkgs.datanovia.com/ggpubr/
Gplots	3.1.3.1	Warnes et al. gplots : arious R reprogramming tools for plotting data. https://cran.r-project.org/web/packages/gplots/index.html
GSEA	4.3.2	Subramanian et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. PNAS 2005
Hisat2	2.1.0	Kim et al. HISAT : a fast spliced aligner with low memory requirements. Nature Methods 2015
HtSeq count	0.9.1	Anders et al. HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 2014
Integrate Genome Browser	10.0.1	Freese et al. Integrated genome browser: visual analytics platform for genomics. Bioinformatics 2016
Integrative Genome Viewer	2.17.4	Thorvaldsdóttir <i>et al.</i> Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform, 2013
Macs2	2.2.9.1	Zhang et al. Model-based Analysis of ChIP-Seq (MACS). Genome Biology 2008
Metascape	3.5.20240901	Zhou et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications 2019
Mulcomp	1.4-25	Hothorn et al. (2008). "Simultaneous Inference in General Parametric Models." Biometrical Journal
Plotly	4.10.4	Sievert C (2020). Interactive Web-Based Data Visualization with R, plotly, and shiny.
Plotprofile	3.5.1.0.0	Ramirez et al. Deeptool2 : A next generation web server for Deep-Sequencing data analysis. Nucleic Acids Research 2016
Prism	10.1.0	"What you can conclude when two error bars overlap (or don't)?" GraphPad Software, Inc 2008
R	3.4.1	R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org
Rccp	1.0.12	Eddelbuettel et al. (2025). Rcpp: Seamless R and C++ Integration

Reshape2	1.4.4	Wickham et al. Reshaping Data with the reshape Package. Journal of Statistical Software 2007
Rmisc	1.5.1	Ryan M. Hope (2012). Rmisc: Ryan Miscellaneous. R package version 1.5.1, https://cran.r-project.org/web/packages/Rmisc . Accessed 13 Jun. 2025.
Rstatix	0.7.2	Kassambara A (2023). rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2, https://rpkgs.datanovia.com/rstatix/ .
RStudio	2022.12.0+355	Posit team (2025). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/ .
Salmon	1.10.1	Srivastava et al. Alevin Efficiently estimates accurate gene abundances from dscRNA-Seq data. Genome Biology 2019
ScVelo	0.2.5	Bastidas-Ponce et al. Comprehensive single cell mRNA profiling reveals a detailed roadmap for pancreatic endocrinogenesis. Development, June 2019. URL: http://dx.doi.org/10.1242/dev.173849 , doi:10.1242/dev.173849.
Sets	1.0-25	Meyer D, Hornik K (2023). sets: Sets, Generalized Sets, Customizable Sets and Intervals. R package version 1.0-25, https://CRAN.R-project.org/package=sets
Seurat	4.0.3	Hao et al (2023). "Dictionary learning for integrative, multimodal and scalable single-cell analysis." Nature Biotechnology. doi:10.1038/s41587-023-01767-y, https://doi.org/10.1038/s41587-023-01767-y
Slingshot	2.10.0	Street et al (2018). "Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics." BMC Genomics, 477. https://doi.org/10.1186/s12864-018-4772-0 .
SPARK	1.3.0	Nielsen et al. Spark: A navigational paradigm for genomic data exploration. Genome Research 2012
Sra Tools	2.11.0	Leinonen et al. The Sequence read archive. Nucleic acids research 2010
STACAS	2.0	Andreatta et al. "Semi-supervised integration of single-cell transcriptomics data", Nature Communications (2024) - https://www.nature.com/articles/s41467-024-45240-z
Stats	4.3.1	R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/
Sva	3.50.0	Leek et al (2025). sva: Surrogate Variable Analysis. doi:10.18129/B9.bioc.sva, R package version 3.56.0, https://bioconductor.org/packages/sva .
Tidyverse	2.0.0	Wickham et al. Welcome to the Tidyverse. Journal of Open Source Software 2019
Usethis	2.2.3	Wickham H et al (2025). <i>usethis: Automate Package and Project Setup</i> . R package version 3.1.0, https://usethis.r-lib.org .

Tableau 3 : Liste des outils bio-informatiques utilisés pour cette thèse.

1.2. Analyse des données de transcriptomiques

a. Analyse de données issues de « *Bulk* » RNA-Seq

Toutes les données de RNA-Seq ont été (ré)analysées de la même manière. Les fichiers SRA des données issues de la littérature ont été téléchargés en utilisant l'outil « SRA Toolkit » (Leinonen et al., 2011). Les fichiers FastQ ont été analysés avec une instance locale de Galaxy (Afgan et al., 2018). Les séquences, ou « *reads* », ont été alignées sur les génomes de référence humain (Hg38) ou murin (mm10) avec l'outil « Hisat2 » (D. Kim et al., 2015). Puis, le comptage de ces « *reads* » a été réalisé en utilisant « HtSeq » avec les paramètres suivants : « *mode* » = *UNION*, « *stranded* » = *No*, « *minimum alignment quality* » = *10*, « *features type* » = *Exon*, « *attributes* » = *GeneID* (Anders et al., 2015). La normalisation et l'expression différentielle ont été faites avec l'outil DeSeq2 (Love et al., 2014) dans l'interface R (R Core Team, 2023). Lors de cette étape, un pré-filtrage des données a été réalisé pour enlever les gènes contenant moins de 10 « *reads* » toutes conditions confondues (en utilisant le paramètre : « *rowSums(counts(dds))* » ≥ 10). Les changements d'expression statistiquement significatifs ont été définis en tenant compte des p-values corrigées avec les seuils indiqués sous les figures. Les valeurs de « *fold change* » utilisées dans nos analyses proviennent de DeSeq2.

Comme le fichier de sortie de DeSeq2 reporte les données normalisées sous la forme de données dites logarithmiques régularisées transformées (Rlog), ne permettant pas la comparaison entre différents gènes une normalisation supplémentaire a été appliquée. Pour cela, la taille des transcrits associés à chaque gène a été calculée en moyennant la taille de tous les transcrits associés à un gène par l'outil Ensembl (GRCm39 ou GRCh.38.p13) (Cunningham et al., 2022). La normalisation a été réalisée en utilisant la formule suivante :

Expression normalisée = Rlog – log (Moyenne des tailles de transcrits en kb)

Les résultats des analyses sur les données de transcriptomiques de RNA-Seq ont été visualisés en utilisant des graphiques à points, des diagrammes circulaires ou des boîtes à moustaches grâce à l'outil ggplot2 (Wickham, 2009). Des cartes de chaleurs ont également été générées avec ComplexHeatmap (Gu et al., 2016).

b. Comparaison des transcriptomes en utilisant l'analyse en composante principale (ACP)

Le transcriptome des hépatocytes HCT [décrit comme « DDC-rep » dans l'étude d'origine (Merrell et al., 2021)] des foies des souris sous régime DDC a été comparé à celui d'hépatocytes et de cholangiocytes issus de souris contrôles (Zummo et al., 2023). Tous les gènes détectés dans les analyses ont été considérés. Afin d'enlever les artéfacts pouvant avoir un impact sur l'analyse des données, une correction de biais (« *batch correction* ») a été réalisée en utilisant la fonction Combat de l'outil sva sur R (Leek et al., 2012). Pour cela, le transcriptome des hépatocytes de l'étude de (Zummo et al., 2023) a été utilisé comme référence. L'analyse a été réalisée avec les paramètres suivants : « *mean.only = T* », « *par.prior = F* », « *prior.plot = F* ».

Après cette correction, une analyse en composante principale (« *principal component analysis* », ACP) a été réalisée en utilisant la fonction PCA de l'outil « FactomineR » (Lê et al., 2008) dans l'interface R avec le transcriptome des hépatocytes et des cholangiocytes de l'étude de (Zummo et al., 2023). Les autres données ont été considérées comme des individus supplémentaires. Seules les deux premières composantes ont été utilisées pour projeter les données et montrent que les hépatocytes des 2 études se superposent.

Le même processus a été utilisé pour réaliser l'ACP sur les gènes codants pour les TFs.

c. Analyse d'enrichissement des signatures géniques (GSEA)

Les analyses d'enrichissement des signatures géniques ont été réalisées avec le logiciel GSEA (Subramanian et al., 2005), développé par le Broad Institut. Le fichier, contenant les niveaux d'expression génique (.gct), a été généré en utilisant les données d'expressions sous la forme de Rlog, issus de la normalisation DeSeq2, et permet le classement des gènes par le logiciel GSEA. Les différentes listes de gènes utilisées sont indiquées au-dessus de chaque analyse. Pour ces analyses, 1 000 permutations de gènes ont été réalisées et les paramètres suivants ont été utilisés :

pondération du score statistique (« *weighted* »), classement des gènes par différence de classe, pas transformation dans l'annotation des gènes utilisés (« *No collapse* »).

En plus des graphiques indiquant les enrichissements dans l'une ou l'autre des conditions, le logiciel fournit les scores d'enrichissement normalisés (NES) et du taux de fausses découvertes (FDR) qui ont été reportés dans les figures.

d. Enrichissement et annotation fonctionnelle des gènes

i. Utilisation de l'outil DAVID (Base de données pour l'annotation, la visualisation et la découverte intégrée)

Le logiciel DAVID (base de données pour l'annotation, la visualisation et la découverte intégrée) a été utilisé (Sherman et al., 2022). Cet outil permet de regrouper les annotations partageant un nombre important de gènes. Il donne un score d'enrichissement des gènes selon les grandes fonctions retrouvées dans les bases de données. Pour ces analyses, les 3 000 gènes dont l'expression varie le plus entre la condition contrôle et la condition HCT ($\log_2FC < 0$ ou $\log_2FC > 0$ & $q\text{-value} < 0,05$) ont été considérés. Les analyses ont été réalisées en utilisant les paramètres suivants de l'outil « Functional Annotation Clustering » : « *Similar Term Overlap = 5* », « *Similar Threshold = 0,5* » et « *EASE threshold = 0,05* »

Seuls les 5 clusters avec les plus hauts taux d'enrichissements ont été considérés. Les résultats ont été représentés de deux manières différentes : des diagrammes en barres et des nuages de mots. Ces graphiques ont été réalisés respectivement grâce au logiciel GraphPad PRISM, et via le site internet WordArt (<https://wordart.com>) avec les paramètres : sans répétition de mots, taille des mots définie selon le nombre de répétitions, forme générale circulaire, disposition horizontale).

ii. Utilisation de Metascape

Lorsque le nombre de gènes à analyser était moins important, le site Metascape a été utilisé pour les analyses fonctionnelles (Zhou et al., 2019). Une fois l'analyse réalisée, les tableaux ont été enregistrés et exportés. Pour chaque analyse, seuls les 5 premiers groupes avec les annotations les plus statistiquement significatives (\log_{10} p-value) ont été utilisés. Les diagrammes en barre ont, ensuite, été réalisés en utilisant l'outil ggplot2 (Wickham, 2009) de l'interface R (R Core Team, 2023).

1.3. Analyse des données de Single-cell RNA-Seq (scRNA-Seq)

a. Analyse des données de scRNA-Seq provenant du « *Human Protein Atlas* »

Les données de scRNA-Seq provenant du « *Human Protein Atlas* » ont été téléchargées sur le site (<https://www.proteinatlas.org/about/download>). Les données d'expression, issues d'une analyse « *pseudo-bulk* », ont été obtenues sous la forme de nTPM (Nombre de transcrits normalisés par million) (Karlsson et al., 2021). L'expression génique dans les différents groupes de cellules (épithélial, mésenchymateux, endothélial, ...) a été défini en calculant la moyenne des nTPMs de chaque cellule (hépatocytes, cholangiocytes, fibroblastes, ...) appartenant à ce groupe, défini sur le site du « *Human Protein Atlas* ».

Le « *tau index* », permettant d'évaluer la spécificité de l'expression d'un gène, a été calculé en utilisant l'outil « Tspex » (<https://tspec.lge.ibi.unicamp.br>) sur des données ayant subies une transformation logarithmique (\log_2) (Camargo et al., 2024). Les données d'expression d'un gène pour chaque groupe cellulaire ont été définies par la médiane, calculée à partir des différentes cellules de chaque groupe.

Enfin, le clustering hiérarchique, basé sur l'expression de Epith-Signature dans chaque type cellulaire du groupe de cellules épithéliales, a été réalisé avec les fonctions « hclust » et « dist » de l'outil « stat » présents par défaut dans R (R Core Team, 2023) avec la méthode Ward.D2. Les distances entre chaque type cellulaire ont été calculées avec une distance euclidienne.

b. Analyse des données de single-nuclei RNA-Seq

Les données de (Gribben et al., 2024) ont été exploitées à l'aide de l'application R Shiny fournie par les auteurs (https://www.mohorianulab.org/shiny/vallier/LiverPlasticity_GribbenGalanakis2024/), mais aussi, en téléchargeant les données à partir de GEO et en les réanalysant à l'aide de l'outil Seurat. Dans les 2 cas, seules les cellules annotées en tant que hépatocytes ont été utilisées. Les « *odd ratios* » et les expressions géniques ont été calculés à partir des hépatocytes provenant du lobe droit de chaque patient. Ainsi chaque patient n'a contribué qu'à un seul échantillon dans l'ensemble des données. Les rapports de chance, appelés « *odds ratio* », ont été calculés avec la fonction « *odds ratio* » de l'outil « *epitools* » (Brown et al., 2001).

c. Analyse des données de scRNA-Seq utilisées pour la mise en place de l'atlas murin

Ces analyses, ainsi que la mise en place de l'atlas, ont été réalisées par 2 ingénieurs de l'équipe (Dr. Julie CHEVALIER et Loïc GUILLE). Les données utilisées sont récapitulées dans le **Tableau 2**. Le détail des paramètres et les codes utilisés sont disponibles sur <https://zenodo.org/records/13897656>

Brièvement, chaque jeu de données a individuellement été filtré en utilisant l'outil Seurat (Hao et al., 2024) afin de garder dans chacun les cellules exprimant entre 1 000 et 5 000 gènes et moins de 20% de gènes mitochondriaux. L'expression génique dans les cellules a été normalisée avec l'outil *sctransform*, permettant de réaliser une ACP qui a été utilisée pour les UMAPs et pour le clustering de Louvain (Hafemeister & Satija, 2019). Les clusters ont été étiquetés sur la base de gènes marqueurs s'exprimant dans les cellules, ce qui a permis de supprimer les clusters de cellules n'appartenant pas à la lignée hépato-cholangiocytaire. Les différents groupes ont été définis en utilisant les expressions géniques suivantes :

-*Alb* et *Apoc3* pour les hépatocytes ;

-*Sox9* et *Epcam* pour les cholangiocytes ;

- *Hnf4a*, *Afp*, *Dlk1*, *Cd13*, *Fgb*, *Fxyd1* and *Gjb1* pour les hépatoblastes ;

- *Gypa, Vim, Flt1, Pecam1, Itgam, Clec4f, Eln, Cd248, Upk1b, Dcn, Lrat, Des, Pdgfrb, Lox, Acta2, Col1a* pour les cellules non-parenchymateuses ;

Une correction des biais a été réalisée en utilisant l'outil STACAS (Andreatta et al., 2024) pour créer un atlas. La UMAP a été générée sur ces données intégrées provenant de STACAS et visualisées grâce au logiciel ISCEBERG (« *Integrative Single Cell Expression Browser for Exploration of RNASeq data using Graphics* ») (Guille et al., 2022).

Afin d'identifier les trajectoires des différentes cellules, des analyses de vélocité ont été réalisées avec l'outil ScVelo (Bergen et al., 2020). Pour cela, les hépatocytes contrôles de l'étude de (Pepe-Mooney et al., 2019) ont été ajoutés aux cellules sélectionnées (cellules provenant des clusters 2, 8 et 15 de la résolution 0,8 de l'atlas). Puis, les « *reads* » des cellules ont été réalignés sur le génome GRCm39 v.108 en utilisant l'outil Salmon (Srivastava et al., 2019) afin de générer les matrices de comptage des gènes épissés ou non. Ces matrices ont subi une correction des biais pour supprimer l'effet « batch » (précédemment défini par STACAS) et ont été utilisées pour le calcul de vélocité avec l'outil ScVelo. Les résultats ont, ensuite, été analysés avec l'outil CellRank2 (Weiler et al., 2024) afin de calculer les probabilités de transitions d'un état cellulaire à un autre.

Enfin, les analyses « *pseudo-time* » ont été réalisées avec l'outil « *slingshot* » (Street et al., 2018), puis, la trajectoire HCT a été sélectionnée pour être combinée aux résultats de CellRank2.

1.4. Analyse des données de ChIP-Seq (cistrome et épigénome)

a. Normalisation et génération des signaux issus des données de ChIP-Seq

Après vérification de la qualité des fichiers FastQ, les séquences ont été alignées sur le génome de référence Hg38 en utilisant l'outil Chromap (H. Zhang et al., 2021) avec les paramètres prédéfinis pour le ChIP, incluant : suppression des doublons de PCR, « *error threshold* » = 8, « *min-num-seeds* » = 2, « *max-seed-frequency* » = 500 – 1000, « *max-num-best-mapping* » = 1, « *max-insert-size* » = 1 000, « *MAPQ-threshold* » = 10, « *min-read-length* » = 30, « *bc-error-threshold* » = 1, « *bc-probability-threshold* » = 0,90.

Les signaux, issus du ChIP-Seq des marques d'histones H3K27me3 et H3K27ac, ont été obtenus en utilisant la fonction BamCoverage de l'outil DeepTools (Ramírez et al., 2016) avec les paramètres suivants : « *bin-size* » = 25, « *reads extension* » = 200 et la méthode de normalisation CPM.

Pour le ChIP-Seq identifiant les cibles de GRHL2, les fichiers Bam provenant des 4 réplicats biologiques ont été analysés avec l'outil « *Irreproducible Discovery Rate* » (IDR). Cet outil permet d'identifier les régions chromatinienne liées aux TFs de manière reproductibles et concordantes entre les réplicats biologiques (<https://github.com/nboley/idr>). Le score IDR a été réalisé en suivant le pipeline provenant d'ENCODE installé depuis <https://github.com/ENCODE-DCC/chip-seq-pipeline2/tree/master> (version 2.2.1) (Hitz et al., 2023). Les 10 000 meilleurs pics, identifiés grâce au score de l'IDR obtenu en utilisant les données provenant des cellules transfectées avec un plasmide vide comme contrôle, ont été utilisés dans nos analyses. Les gènes cibles de GRHL2 et leurs annotations fonctionnelles ont été réalisés grâce à l'outil Cistrome-GO (S. Li et al., 2019). Les modes « solo » ou « ensembl » (prédiction combinée des gènes cibles grâce aux sites de fixation de GRHL2 et les gènes modulés par GRHL2 définis grâce au RNA-Seq) ont été utilisés avec une distance de 1 kb du pic.

b. Visualisation des signaux

Une fois les coordonnées des TSSs obtenues, elles ont été étendues dans une fenêtre de +/- 500 bp pour réaliser les analyses. Tous les signaux ont été visualisés grâce à l'outil DeepTool (Ramírez et al., 2016). Dans un premier temps, la fonction Compute Matrix a été utilisée avec comme point de référence le TSS du gène étendu sur une fenêtre d'analyse de +/- 2500 bp et une taille de sections (« *bin* ») de 500 bp. Une fois la matrice générée, les signaux moyens au niveau des promoteurs des gènes d'intérêt ont été produits en utilisant l'outil PlotProfile.

Les signaux de ChIP-Seq au niveau de gènes individuels en utilisant « *Integrated Genome Browser* » (IGB) (Freese et al., 2016) ou grâce à l'outil « *Intergative Genome Viewer* » (IGV) (Thorvaldsdóttir et al., 2013).

c. Séparation des gènes selon leurs profils épigénétiques grâce à l'utilisation de SPARK

Afin d'identifier les groupes de gènes ayant des profils épigénétiques différents, l'outil SPARK a été utilisé (Nielsen et al., 2012). Le regroupement de gènes a été réalisé en utilisant une fenêtre de 5 kb centrée sur le TSS et séparée en 10 sections. Le mode de normalisation global a été utilisé, c'est-à-dire que le signal tout le long du génome a été considéré pour la normalisation.

1.5. Statistiques

Les analyses statistiques ont été réalisées avec le logiciel R (R Core Team, 2023) ou avec le logiciel GraphPad Prism. Les paramètres utilisés (type de test statistique, correction, nombre d'échantillons) pour chaque analyse sont indiqués dans les légendes des figures. Une correction pour tests multiples a été appliquée en calculant la FDR, présentée sous forme de *q*-values. Sauf indication contraire, les différences statistiques ont été considérées significatives lorsque la *q*-value était inférieure à 0,05, et sont indiqués par *. Bien que les analyses statistiques aient été réalisées sur l'ensemble des données, dans certains cas, seuls les résultats ayant un intérêt pour cette étude sont présentés.

Tous les diagrammes en barres présentent la moyenne \pm SD (écart-type). Les diagrammes en barres sont composés d'une boîte montrant le 1^{er} et le 3^{ème} quartile avec la médiane comme ligne et le minimum et maximum comme extrémités.

II. Procédures expérimentales

2.1. Culture cellulaire

a. Entretien des lignées cellulaires

Les HepG2, lignée cellulaire issue d'un hépatoblastome, ont été cultivées dans du milieu Willam's E (Gibco, 22551022) supplémenté par 10% de sérum de veau fœtal (SVF) (Life, SVF30106.03), 20 mU/mL d'insuline bovine (Sigma, I5550), 50 mM de dexaméthasone (Sigma, D1756), 100 U/mL de pénicilline/streptomycine (Gibco, 15140). Les plaques ont été recouvertes de gélatine 1g/L (Sigma, G1890).

Les cellules embryonnaires hépatiques bipotentes issues de foie de souris (BMELs), décrites dans (Strick-Marchand & Weiss, 2002), ont été cultivées dans du milieu RPMI 1640 Glutamax (Gibco, 61870-044) supplémenté par 10% de SVF (EuroBio, CVSVF00-001), 30 mg/mL du facteur de croissance analogue à l'insuline II (IGFII) (PreproTech, 100-12-100µg), 50 ng/mL EGF murin (PreproTech, 315-09-1mg), 10 mU/mL d'insuline bovine et 100 U/mL de pénicilline/streptomycine. Les BMELs poussent sur des plaques recouvertes de Collagène de type I.

Les HepaRG, cellules progénitrices bipotentes humaines isolées à partir d'un HCC et décrites par (Gripon et al., 2002), ont étéensemencées à une densité de $2,6 \times 10^4$ cellules/cm² et sont maintenues dans le milieu Willam's E supplémenté par 10% SVF, 100 U/mL de pénicilline/streptomycine, 5 µg/mL d'insuline et 5×10^{-5} d'hydrocortisone hemisuccinate (Sigma, H2270). Après 15 jours, 2% de DMSO (Diméthylsulfoxyde) a été ajouté dans le milieu pendant 2 semaines pour obtenir une différenciation en hépatocytes. Les cellules ont ensuite été détachées grâce à la trypsine et re-suspendue dans le milieu de culture sans DMSO avant l'électroporation.

Toutes les cellules ont été cultivées à 37°C dans une atmosphère humide avec 5% de CO₂.

b. Utilisation de modulateurs pharmacologiques

L'inhibition pharmacologique de EZH1 et EZH2 a été réalisée en traitant les cellules pendant 4 jours avec une combinaison de 1 μ M d'UNC1999 (Target Mol, T3057) (Kusakabe et al., 2021) et 1 μ M de Valemetostat Tosylate (MedChem Express, 1809336-93-3) (Duan et al., 2020). La condition contrôle correspond à des cellules traitées avec 0.2% de DMSO.

Pour vérifier l'effet de la O-GINAcylation sur la stabilité de EZH2, les HepG2 ont été traitées pendant 24h avec 5 μ M l'OSMI-1 (Sigma-Aldrich, 1681056610), puis, les cellules ont été incubées pendant 0, 2, 4, 5 ou 8h avec un mélange d'OSMI-1 et de 50 μ g/mL de cycloheximide (Sigma-Aldrich, C7698). Le DMSO 0.1% a été utilisé comme contrôle.

L'inhibition de l'activité de l'O-GlcNAc transférase (OGT) a aussi été effectuée en utilisant un siARN ciblant l'OGT. Les cellules ont été traitées avec 20nM de siARN pendant 48h en utilisant le JetPrime (PolyPlus transfection, 10100046) et en respectant les instructions du fabricant. Un siARN sans cible a été utilisé comme contrôle.

c. Modulation transitoire de l'expression de *GRHL2*

Les cellules (HepG2 ou BMELs) ont été transfectées en utilisant le JetPrime (PolyPlus transfection, 10100046) en respectant les instructions du fabricant. Le plasmide *GRHL2* a été acheté auprès de VectorBuilder. Il permet l'expression du cDNA de *GRHL2* (NM_024915.4) fusionné à un FLAG, le tout sous le contrôle du promoteur *EEF1A1* (*EF1A*). Cette construction permet en parallèle l'expression de la protéine fluorescente (EGFP) qui a été utilisée pour contrôler l'efficacité de la transfection. Le vecteur pcDNA3 (Invitrogen™) a été utilisé comme vecteur contrôle. Après 24h de transfection, le milieu de culture a été rafraîchi. Les analyses ont été réalisées 48h après la transfection.

Dans les HepaRG différenciés, la transfection a été réalisée par électroporation grâce au système d'électroporation Néon (Thermo Fisher Scientific) comme décrit précédemment dans (Laurent et al., 2013). En résumé, 10⁶ HepaRG ont été électroporées avec 2 μ g de plasmide contrôle ou *GRHL2* par un courant de 1500 volt pendant 20 millisecondes. Ces cellules ont, ensuite, été réensemencées à faible

densité ($2,6 \times 10^4$ cellules/cm²) dans le milieu de culture Willam's E sans DMSO. Le milieu a été renouvelé au bout de 24h. Enfin, 48h après l'électroporation, les cellules ont été grattées avec du Phosphate-buffered saline (PBS) et culottées par centrifugation pendant 5 min à 400g.

d. Culture tri-dimensionnelle (3D) des BMELs

Après 48h de transfection, les BMELs ont été comptées pour être mises en culture dans un système en trois dimensions. Chaque dôme cellulaire a été formé en mélangeant 5 µL de suspension cellulaires (20 000 cellules pour les 5 µL) avec 30 µL de matrigel (Corning® Matrix for Organoid Culture Phenol Red-free LDVEC-Free, 356255) dans un tube pré-refroidi à 4°C. Après homogénéisation de la solution en utilisant des cônes pré-refroidis également, 30 µL de la solution ont été ajoutés au centre de chaque puits de la plaque Ibidi (Ibidi, 80806), placée à 37°C depuis la veille. Après 15 à 30 min de solidification à 37°C du matrigel, 300 µL de milieu de culture des BMELs a été ajouté dans chaque puit. Le milieu a été renouvelé tous les 2 jours.

e. Analyse de l'O-GlcNAcylation de EZH2

Afin de définir si EZH2 est bien O-GlcNAcylé, les HepG2 ont été traitées à confluences avec 2µM de Thiamet-G pendant 24h. Une fois le traitement réalisé, les plaques ont été rincées deux fois avec du PBS froid et les cellules ont été récupérées et centrifugées 5min à 400g à 4°C. Le surnageant a été jeté et le culot re-suspendu dans un tampon hypotonique (50 mM NaF, 1mM sodium orthovanadate, PIC et 2µM de Thiamet-G). Après 5 minutes d'incubation sur glace, les tubes ont été centrifugés 5 min à 600 g à 4°C et le surnageant a été jeté. La fraction nucléaire a été re-suspendue et incubée 30 min à 4°C avec le tampon de lyse (50 nM NaF, 1mM sodium orthovanadate, PIC et 2 µM Thiamet-G). Les échantillons ont ensuite subi une sonication pendant 10 min (Bioruptor ; 30s ON/ 30s OFF) et ont été centrifugés à vitesse maximale pendant 5 min à 4°C. La fraction soluble a été conservée. Le marquage des protéines O-GlcNAcyliées par la GalNaz a été réalisé avec 200 µg de protéines nucléaires enrichis en utilisant le système de marquage enzymatique O-GlcNAc Click-iT™ et le kit de détection des protéines biotinylées Click-iT™ en respectant les recommandations du fabricant. Comme contrôle négatif, les protéines

ont été soumises à la réaction de marquage enzymatique Click-iT™ mais sans le donneur de substrat UDP-GalNaz.

Les protéines marquées ont, ensuite, été incubées avec 60 µL de billes magnétiques Dynabeads™ M-280 Streptavidin pendant 1h30 à 4°C dans 1% Triton X-100 dilué dans le PBS sous agitation. Les billes ont été lavées 3 fois avec un tampon PBS avec 1% Triton X-100. Les protéines retenues sur les billes ont finalement été éluées dans 40µL de Laemmli 6X pour analyse par Western-Blot.

2.2. Analyses fonctionnelles

a. Culture bactérienne et purification d'ADN plasmidique

Les transformations bactériennes ont été réalisées en utilisant des bactéries DH5α (Invitrogen™, 11563117) ou des bactéries Top10 (Invitrogen™, C404003). Une fois les bactéries décongelées, 2µL de plasmide ont été ajoutés dans la suspension. Les tubes ont été placés 20 min sur glace, puis, 1 min à 42°C, puis 2 min sur glace. Enfin, 100 µL de milieu S.O.C. (Invitrogen™, 15544034) a été ajouté à la suspension avant incubation 1h à 37°C sous agitation. Les bactéries ont, ensuite, été étalées sur boîte de pétri pour les laisser pousser sur la nuit. Une fois la transformation réalisée, les bactéries ont été pré-cultivées dans 4mL de milieu enrichi Terrific Broth (Invitrogen™, 22711-022) pendant 6h à 37°C. Puis, la pré-culture a été transférée dans 400mL de milieu à 37°C toute la nuit. La purification de l'ADN plasmidique a été réalisée avec le kit NucleoBond® Xtra Maxi Plus (Machery-Nagel, 740416.50) en prenant le protocole « *low copies* » dans les instructions du fabricant.

b. Extraction d'ARN et RT-qPCR

Les tissus ont été homogénéisés en utilisant un Minilys et des billes de céramique de 1,4 mm (Bertin Technology) (3 x 15s, 5 000 rpm). L'ARN total des cellules et des tissus a été extrait en utilisant le kit Nucleospin® RNA (Machery-Nagel, 740955) selon les instructions du fabricant.

Les ARNs ont été dosés à l'aide du NanoDrop (Thermo Fisher Scientific), puis, rétro-transcrits en ADN complémentaire (ADNc) avec le kit « *High Capacity Reverse*

transcription » (Applied Biosystem, 4368814). Les analyses par PCR quantitative (qPCR) ont été réalisées grâce à la technologie Takyon™ Fast Advanced Master Mix (Applied Biosystem, A4612) avec le mode « *Fast* » sur le QuantStudio3™ (Applied Biosystem). La spécificité de l'amplification a été vérifiée en enregistrant les courbes de dissociation, et, l'efficacité de tous les couples d'amorces a été approuvée lorsque celle-ci était supérieure à 95%. Les niveaux d'ARNm ont été normalisés grâce à la moyenne d'expression de plusieurs gènes de ménage (sauf pour les organoïdes humain et les échantillons de patients où seul *TBP* a été utilisé). Les « *Fold Change* » ont été calculés en utilisant la méthode des seuils ($\Delta\Delta Ct$). Les séquences des amorces utilisées sont énumérées dans le **Tableau 4**. L'expression de *GRHL2* et de *EHF* a été mesurée à l'aide de sondes TaqMan™ Fast Advanced Master Mix (Applied Biosystem, 4444557) associées à un fluorophore FAM (6-carboxyfluorescéine). L'expression obtenue a été normalisée en comparant les résultats à l'expression de l'ARN ribosomal 18S N5 (« 18S ribosomal N5 », 18S), gène de référence, dont la sonde est couplée au fluorophore VIC (2'-chloro-7'-phenyl-1,4-dichloro-6-carboxy-fluorescein). Ces qPCR ont été réalisées avec le kit selon les instructions du fabricant.

Nom du gène	Séquence Sens	Séquence Antisens
Humain		
ALB (Albumin)	CCAAGAAAGTACCCCAAGTGTC	GGTTCAGGACCACGGATAGAT A
CLDN4 (Claudine 4)	ACTGCCTGGAGGATGAAGA	GACACCGGCACTATCACCAT
CYP3A4 (Cytochrome P450 Family 3 Subfamily A Member 4)	CATTCTCATCCCAATTCTTGAG GT	CCACTCGGTGCTTTTGTGTAT CT
EHF (ETS Homologous factor)	Sonde Taqman Hs000171917_m1 (ThermoFisher Scientific)	
ESRP1 (Epithelial Splicing regulatory protein 1)	GCTCTGTCAACTCCGTCTCA	GCCCTCCGACAGTTTAACCA
GRHL2 (Grainhead like transcription factor 2)	Sonde Taqman Hs00227745_m1 (ThermoFisher Scientific)	
ITGA3 (Integrin subunit alpha 3)	ATATGGGCGCGGTGTTCTTG	GTCCTGCCACCCATCATTGT
ITGB8 (Integrin subunit beta 8)	GGCATAGTGTTGCCCAATGA	GTGAGGGGTGTTCCATGGTT
JUP (Junction plakoglobin)	CGAGCTCAGTTCGCTGTCC	CGAGCTCAGTTCGCTGTCC
KRT7 (Keratin 7)	TGCCTACATGAGCAAGGTGG	GCTCTGTCAACTCCGTCTCA
RAB25 (Member of Ras Oncogene Family)	TCGTCATGCTCGTGGGTAAC	GCAGTCCATTGTTTCAGCGA
RPLP0 (Rplp0 ribosomal protein lateral stalk subunit P0)	CATGCTCAACATCTCCCCCTTC TCC	GGGGCTCAACATCTCCCCCT TCTCC
TBP (TAT-Box Binding Protein)	TTCGGAGAGTTCTGGGATTG	GGATTATATTCGCGTTTCG
Souris		
Abcdbl11 (ATP Binding Cassette Subfamily B Member 11)	GTCTGACTCAGTGATTCTTCGC	GAGCAATGCGCACACACTTC
Acmsd (aminocarboxymuconate semialdehyde decarboxylase)	CCAAGCTCTTCCACAGTTCC	CCTTCGAGGGTATCTGGCAA
Cldn4 (Claudine 4)	CAGTGCAAGATGTACGACTC	CCTACCACTGAGAGAAGCAT

Ehf (ETS Homologous factor)	CATCTCAAGTGGGAACGGCCA	GTCATGGAGATCTGAGCCCG
Esrp1 (Epithelial Splicing regulatory protein 1)	TCAGTGTTTCAGCTGAGGAGA	AGGGAGGAGACAGGCATGG
Fmo5 (flavin containing monooxygenase 5)	CAATGCTGGGTGAGCGGTGA	GCCTCCGCCCTTCATCTGT
Grhl2 (Grainhead like transcription factor 2)	Sonde Taqman Mm00518236 (ThermoFisher Scientific)	
Gsta3 (glutathione S-transferase, alpha 3)	CCCTTAACTGGGAAGGGCGT	ACCATATGCAAGAGTTTCTGC TGT
Itga3 (Integrin subunit alpha 3)	CGGAACATCACCTGCTGGCTT A	TGGCAATGTGTCTCTGGCAT
Itgb8 (Integrin subunit beta 8)	AGGTCCAGCCTTGGTTCTTG	AGCCACATCTATTGTGTTTAC T
Jup (Junction plakoglobin)	CGAGCTCAGTTCGCTGTCC	CGAGCTCAGTTCGCTGTCC
Krt7 (Keratin 7)	GCGGAGATGAACCGCTCTAT	TCTAAGTTGGCAGCTGGTT
Ly6g (lymphocyte antigen 6 complex, locus G)	TGATGGATTTTGCCTTGCTCTG	GTATTGTCCAGAGTAGTGGGG C
Ptpcr (Protein Tyrosine Phosphatase Receptor Type C)	AGGCGACAGAGGTGTCTGAT	TTTCTCCAACTGATTGTATTC CA
Rab25 (Member of Ras Oncogene Family)	AAAAGCTGAGAGTTGAGGCCA	CCGATCAGCACCTTAAA
RNA18S5 (18S Ribosomal N5)	CATTCTTGGCAAATGCTTTCG	CGCCGCTAGAGGTGAAAATTC T
Rplp0 (Rplp0 ribosomal protein lateral stalk subunit P0)	CATGCTCAACATCTCCCCCTTC TCC	GGGGCTCAACATCTCCCCCT TCTCC
Alb (Albumin)	GCTACGGCACAGTGCTTGCT	TCCACGAGAGTTGGGGTTGA
Cyp3a11 (Cytochrome P450 Family 11 Subfamily A Member 1)	TGCTGACAAACAAGCAGGGAT GGACCTGG	GTCCCATATCGGTAGAGGAGC ACCAAGCTG

Tableau 4 : Amorces utilisées en RT-qPCR.

Les analyses de RT-qPCR indiquent que l'expression de *GRHL2* est faible ou non détectée à l'état basal dans les modèles utilisés au cours de cette thèse, sauf pour les organoïdes humains.

Concernant les analyses de RNA-Seq, l'intégrité et la quantité des ARNs ont été évalués à l'aide du Bioanalyseur Agilent 2100 (Agilent Technology) avant que les échantillons soient utilisés pour la préparation de la librairie et le séquençage.

c. Extraction et révélation des protéines

i. Extraction des protéines totales

Pour l'extraction des protéines totales, les cellules ont été lavées 2 fois dans du PBS froid avant d'être grattées et culottées par centrifugation à 400g pendant 5 min. Les culots cellulaires ont été re-suspendus dans un tampon salin [25 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.5% NP40, 0.9 mM Sodium Butyrate (Sigma, 303410-5g), 0.1 mM Sodium Orthovanadate (Sigma, 450243-10g), un cocktail d'inhibition de dé-acétylation (MCE, HY-K0030), 83 U/μL Benzonase (Sigma, E1014-25KU) et un cocktail d'inhibiteur de protéase (PIC)]. Après 10 minutes d'incubation sur glace, les échantillons ont été fragmentés par sonication 2 fois pendant 10 minutes (10 cycles de 30s ON/30s OFF) avec le Biorupteur Pico (Diagenode). Les tubes ont, ensuite, été centrifugés 5 min à vitesse maximum et la partie soluble a été récupérée. Un aliquot de surnageant a été pris pour déterminer la concentration protéique à l'aide du kit de dosage de protéines Pierce™ BCA (Bicinchoninic Acid Assay). Du Laemmli 6X [175mM Tris-HCL pH 6.8, 15% glycérol, 5% Sodium docetyl sulfate (SDS), 300nM Dithiothreitol (DTT) et 0.01% de bleu de bromophénol] a été ajouté dans tous les surnageants.

ii. Fractionnement de la chromatine

Les tissus provenant des foies humains, issus de biopsies, ont été homogénéisés avec le Polytron (Fisher Scientific, IKA® T10 basic ULTRA-TURRAC®) dans le tampon A [10 mM HEPES pH 7.5, 10 mM KCl, 1.5mM MgCl₂, 340 mM Sucrose, 10% glycérol, 1mM DTT, PIC] et incubés 10 min à 4°C. Les échantillons ont, ensuite, été centrifugés pendant 5 min à 1 300g à 4°C. Les surnageants ont été transférés dans de nouveaux tubes avant d'être à nouveau centrifugés pendant 5 min à 1 300g à 4°C afin d'enlever tout le matériel insoluble. Les culots cellulaires ont, ensuite, été lavés avec le tampon A et incubés par la suite avec la solution B [3 mM EDTA, 0.2mM EGTA, 1 mM DTT et PIC] pendant 30 min à 4°C. Les échantillons ont été centrifugés à 1 700g pendant 5 min à 4°C et le culot contenant la chromatine a ensuite été lavé une première fois avec la solution B, puis avec du PBS froid, avant d'être re-suspendu dans le tampon C [50 mM Tris-HCl pH 8.0, 1mM MgCl₂ et 83 U/μL Benzonase]. Les

échantillons ont été incubés pendant 20 min à 4°C avant le prélèvement d'un aliquot pour la quantification de concentration en protéines et l'ajout de Laemmli 6x.

iii. Extrait nucléaire et immunoprécipitation

Pour les foies murins, le lobe postérieur droite a systématiquement été utilisé pour l'extraction protéique. Les échantillons de foie murin ont été filtrés avec des filtres de 0,7µm dans 1mL de PBS froid et transférés dans des tubes Eppendorf de 1.5 mL. Les échantillons ont, ensuite, été centrifugés à 400g pendant 5min à 4°C. Le culot cellulaire a été re-suspendu dans 500 µL de solution hypotonique [20 nM Tris-HCl pH 8.0, 10 nM NaCl, 3mM MgCl₂, 0.4% NP40, PIC, 5mM Sodium Butyrate, cocktail d'inhibiteur de dé-acétylase, PIC]. Après 20 minutes d'incubation sur glace, les échantillons ont été fragmentés par sonication pendant 10 min (10 cycles de 30s ON/30s OFF) avec le Biorupteur Pico et centrifugés à vitesse maximum pendant 5 minutes à 4°C. La fraction soluble (extrait nucléaire) a été collectée et transférée dans un nouveau tube. 500 µg de l'extrait nucléaire ont été utilisés pour réaliser l'immunoprécipitation (IP) en ajoutant 2 volumes de tampon de dilution [25nM Tris-HCl pH8.0, 500nM NaCl, 1mM EDTA, 0.5% NP40, PIC, 5mM sodium butyrate, PIC, cocktail d'inhibiteur de dé-acétylase, PIC]. Les échantillons ont été incubés avec 2µg d'anticorps [anti-flag et anti-IgG listés dans le **Tableau 5**] toute la nuit sous rotation à 4°C. Les billes magnétiques Dynal (Sigma, A9418) (5µg de protéine A et 5µg de protéine G par IP) ont été lavées 3 fois avec de la BSA-PBS (Bovine-sérum albumine). Puis, les billes ont été bloquées dans 500 µL de PBS-BSA (5mg/mL) toute la nuit sous rotation à 4°C avant d'être ajoutées aux échantillons. Après 4h d'incubation à 4°C, les échantillons ont été lavés 4 fois pendant 5min sous rotation avec un tampon de lavage froid [25nM Tris-HCl pH8.0, 150nM NaCl, 1mM EDTA, 0.2% NP40, PIC, 5mM sodium butyrate, PIC, cocktail d'inhibiteur de dé-acétylase, PIC]. Enfin, les échantillons ont été élués dans 40 µL de Laemmli.

iv. Simple Western Immunoassays (WES) et Western Blot

Lorsque les protéines ont été utilisées pour un Western Blot, 20 µg de protéines préalablement dénaturées ont été chargées sur un gel d'acrylamide (15% SDS-PAGE), puis, un transfert sur membrane de nitrocellulose (Thermo Fisher Scientific, 88018) a été réalisé. Après 1h de saturation des membranes dans 5% de BSA ou de lait en poudre dilué dans le Tris-Buffer Salin (TBS) 1x (Euromedex, ET220) et Tween 0.01% (Biotech, CAS#9005-84-5), les membranes ont été incubées avec les anticorps primaires dilués dans la même solution à 4°C toute la nuit, sous agitation (listés dans le **Tableau 5**). Le lendemain, après plusieurs lavages, les membranes ont été incubées avec les anticorps secondaires correspondants pendant 1h à température ambiante sous agitation. Après lavages, les membranes ont été révélées à l'aide des kits de détection SuperSignal™ West Dura Extended Duration Substrate (Thermo Fisher Scientific, 34075) pour détecter la chimioluminescence grâce à l'iBright™ CL1500 Imaging System (Thermo Fisher Scientific).

Nom de la cible	Fournisseurs, référence (hôte)	Utilisation	Dilution (saturation)
<u>Anticorps primaires</u>			
Immunoglobine G (IgG)	Santa Cruz sc2025 (Mouse)	IP	2µg
Antiflag	Sigma F3165 (Mouse)	WES ChIP-Seq	1/50 1.5 µg
H3	Cell Signaling CST36385 (Mouse)	WB WES	1/1000 1/50
H3K27me3	Cell Signaling CST9733 (Rabbit)	WB WES ChIP-Seq	1/1000 1/400 2µg
HSP90	Biolegend CAT#662802 (Mouse)	WB WES	1/1000 1/400
GRHL2	Sigma HPA004820 (Rabbit)	WES IHC ChIP-Seq	1/50 1/500 1.5µg
H3K27ac	Active Motif 39685 (Mouse)	ChIP-Seq	2µg
Albumine	Bethyk Laboratory A80-229A (Chèvre)	IHC	1/50
Kératine 7 (Krt7)	Daco Clone OVTL12/30 (Mouse)	IHC	1/50
<u>Anticorps secondaires</u>			
HRP Conjugate anti-rabbit	Sigma CAT#A4416	WB	1/5000
HRP Conjugate anti-mouse	Sigma CAT#A05545	WB	1/5000
Donkey anti-Goat IgG Cross-Adsorbed, Alexa Fluor 647	Invitrogen A-21447	IHC	1/250

Tableau 5 : Liste des anticorps utilisés.

WB = Western Blot ; IHC = Immuno-histologie ; WES = Simple Western Immuno Assay ; IP = immunoprécipitation

Alternativement, les protéines ont été analysées en utilisant le système WES (« *Simple Western Immunoassay* ») (Protein Simple) selon le protocole du fabricant. Dans ce cas, les échantillons ont été dilués avec du tampon 0.1X (Protein Simple) afin d'obtenir des concentrations finales de protéines entre 0,4 et 0,5 mg/mL selon le type de protéine cible. La séparation a été effectuée en utilisant les capillaires 12-230 kDa. Les anticorps primaires utilisés sont référencés dans le **Tableau 5**. Les anticorps secondaires utilisés sont propres au système WES et ont été achetés auprès du fabricant (Protein Simple, PS-MK14 et PS-MK-15). Les données ont été analysées à l'aide du logiciel Compass (Protein Simple). Les quantifications ont été obtenues en utilisant l'aire sous le pic de la protéine d'intérêt et en la comparant à l'aire sous le pic de la protéine de ménage ou au signal obtenu avec le kit « *Total protein assay* » suivant les instructions du fabricant (Protein Simple, DM-TP01).

d. Immunoprécipitation de la chromatine et séquençage haut débit (ChIP-Seq)

Les HepG2 (5×10^6) ont été transfectées avec le plasmide codant pour *GRHL2* ou le plasmide contrôle comme décrit précédemment. Après 48h de transfection, les cellules ont été lavées avec le PBS puis fixées par une incubation de 10 min à température ambiante avec 1% de formaldéhyde (Thermo Fisher Scientific, 28908), suivie d'une incubation de 5 min avec 1.25 mM de glycine. Après 2 lavages au PBS froid, les cellules ont été culotées par 5 min de centrifugation à 400g. Les culots cellulaires ont été incubés pendant 10 min dans [0,25% Triton X-100, 10 mM EDTA, 10 mM HEPES, 0.5 mM EGTA, PIC], suivi d'une centrifugation de 5 minutes à 600g à 4°C. Ensuite, les culots ont été incubés pendant 10 min dans [0.2 M NaCl, 1 mM EDTA, 10 mM HEPES, 0.5 mM EGTA, PIC], suivi de 5 minutes de centrifugation à 600g à 4°C. Les noyaux cellulaires ont été re-suspendus dans 600 µL de tampon de lyse [50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS, PIC] et fragmentés par sonication (45 cycles de 30s ON/30s OFF). La chromatine (30 µg pour les inputs, H3K27ac et H3K27me3 et 200 µg pour GRHL2) a été diluée 10 fois dans le tampon de dilution [20 mM TrisHCl pH 8.0, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl], et, incubée à 4°C toute la nuit sous agitation avec les anticorps (comme décrit dans le **Tableau 5** avec un mélange de quantité égale d'anti-Flag et d'anti-GRHL2). Le lendemain, des billes

de sépharose de protéines A et G (GE Healthcare) ont été ajoutées et les échantillons ont été incubés 4h à 4°C, sous agitation, en présence de 70 µg/mL d'ARNt de levure (Sigma, R5636). Les échantillons ont, ensuite, été lavés 3 fois avec du tampon RIPA [50 mM HEPES, 1 mM EDTA, 0.7% Na Deoxycholate, 1% NP40, 500 mM LiCl] contenant 10 µg/mL d'ARNt de levure, puis, une fois dans un tampon TE [10 mM Tris-HCl pH 8, 0.1 mM EDTA]. L'ADN a, ensuite, été élué dans [100 mM NaHCO₃, 1% SDS] et incubé sur la nuit à 65°C. La purification d'ADN a, enfin, été réalisée à l'aide du kit NucleoSpin® Gel and PCR Clean-UP (Macherey-Nagel, 740609). L'ADN a été soumis à un séquençage haut-débit et analysé comme décrit dans le chapitre **Matériel et Méthode 1.4.a p.114**.

e. Histologie, immuno-histologie (IHC) et immunofluorescence

i. Histologie

Les échantillons de foie ont été fixés avec 4% de paraformaldéhyde (PFA ; échantillons murins) ou avec un tampon de formaline (échantillons humains) et inclus en paraffine. Des coupes de tissus de 4µm d'épaisseur ont été marquées à l'hématoxyline/ éosine ou avec un anticorps spécifique, en utilisant les étapes décrites ci-après. Pour cela, les coupes ont été soumises à une étape de démasquage dans un solution de citrate pH 6.0 pendant 21 min à 100°C à l'aide du TimeRetriever Pressure Cooker (Bio SB) afin de permettre l'immunomarquage. Ensuite, les sections ont été incubées de manière séquentielle dans une solution de 0.1% de Triton X-100 avec de TBS (TBS-T) pendant 10 min à 4°C ; dans une solution de blocage [Bloxall® Endogenous Blocking Solution (Vector Laboratories)] pendant 10 min à température ambiante ; dans 7% de sérum de chèvre dilué dans TBS-T pendant 15 min à température ambiante ; 15 min à température ambiante dans un mélange de 5% BSA et 5% de lait dilué dans le TBS-T ; toute la nuit à 4°C dans la solution contenant l'anticorps primaire GRHL2 dilué au 1:300 (**Tableau 5**) dans 5% BSA et 5% de lait dans du TBS-T. Entre chaque étape, les coupes ont été lavées avec du TBS-T. Le lendemain après lavage, les coupes ont été incubées 30 min à température ambiante avec une solution HRP Impress® anti-souris IgG (Vector Laboratories) suivi d'un lavage avec du TBS-T. Le signal a été détecté avec du diaminobenzine [DAB (Cell Signaling)]. Enfin, les sections ont été contrastées avec de l'hématoxyline (solution

Grill n°1, Sigma-Aldrich) et montées dans un milieu contenant de l'aquatex (Sigma-Aldrich). Les coupes fixées ont été imagées avec l'Axioscan Z1 slide scanner (Zeiss). Les images ont été analysées avec le logiciel ZEN (Zeiss).

ii. Immuno-fluorescence

Pour le co-marquage GRHL2 et ALB, la procédure précédente a été adaptée. Pour le marquage ALB, au lieu d'utiliser 7% de sérum de chèvre dans le TBS-T et le DAB, les coupes de tissus ont été incubées pendant 15 min à température ambiante dans une solution de blocage [« *Animal-Free Blocking Solution* » (Cell Signaling) 1x], puis, dans l'anticorps primaire ALB dilué au 1:50 (**Tableau 5**). Ensuite, l'anticorps secondaire ImmPRESS® HRP anti-chèvre IgG (Vector Laboratories) a été incubé pendant 10 min à température ambiante avec une solution Alexa Fluor™ 594 Tyramide SuperBoost™ (Invitrogen™) selon les instructions du fabricant. À la suite de ce marquage, les étapes similaires ont été répétées pour GRHL2 en utilisant un anticorps secondaire ImmPRESS® HRP anti-souris IgG (Vector Laboratories) et une solution Alexa Fluor™ 488 Tyramide SuperBoost™ (Invitrogen™). Les coupes fixées ont été imagées avec l'Axioscan Z1 slide scanner après un contraste des noyaux avec le DAPI (62248, Thermo Scientific, 1:10000).

De même pour le triple marquage ALB, KRT7 et GRHL2, le protocole précédent a également été adapté. Après une incubation dans une solution de blocage [« *Animal-Free blocking solution* » 1X diluée dans TBS-T], les coupes ont été incubées avec respectivement : l'anticorps anti-KRT7 dilué au 1:50 (**Tableau 5**), l'anticorps secondaire IgG Impress® HRP Polymer cheval anti-souris et l'anticorps fluorescent Alexa Fluor™ 488 Tyramide SuperBoost™. La détection de GRHL2 a été réalisée après une étape de démasquage en utilisant le même système de détection que pour le double marquage (anticorps anti-GRHL2 dilué au 1:300 avec une détection Alexa Fluor™ 594 Tyramide SuperBoost™). Enfin, le marquage de ALB a été réalisé en incubant les coupes avec l'anticorps anti-ALB (1:50) suivi de l'anticorps secondaire Alexa Fluor™ 647 Tyramide SuperBoost™ (A21447, Invitrogen, 1:250). Les noyaux ont été marqués avec du DAPI et les sections ont été scannées à l'Axioscan.

iii. Marquage des organoïdes issus des BMELs

Après 4 jours de culture dans le matrigel, les organoïdes ont été fixés en utilisant 4% de PFA pendant 20 min à température ambiante, lavés avec du PBS et quenchés avec une solution de NH_4Cl 50 mM pendant 15 min à température ambiante. Ensuite, les cellules ont été perméabilisées à l'aide d'une solution 0.1% Triton X-100 pendant 5 min, et incubées pendant 1h dans la solution de blocage [PBS + 1% BSA], le tout à température ambiante. Cette incubation a été suivie du marquage de l'actine avec la phalloïdine (Invitrogen, A21428, 2633537) diluée au 1:1000 pendant 1h, puis celui des noyaux avec le Hoechst dilué au 1:5000 dans la solution PBS + 1% BSA. Après chaque étape, les échantillons ont été lavés trois fois avec le PBS. Les agrégats cellulaires ont été révélés soit par imagerie confocale (Zeiss, LSM700) ou en utilisant le microscope CellDiscoverer 7 (Zeiss). Les images ont été analysées à l'aide du logiciel ZEN.

2.3. Modèles animaux

Toutes les souris mâles (C57BL/6J) ont été achetées chez Charles River à l'âge de 11 semaines. Elles ont été logées dans des cages standards dans une pièce à température contrôlée (22-24°C) avec des cycles de lumière/obscurité de 12h. Les souris ont reçu de l'eau et un régime alimentaire normal *ab libitum*. Les souris ont pu s'acclimater sur le site pendant au moins 1 semaine avant toute expérience.

a. Injection hydrodynamique

Les injections hydrodynamiques ont été réalisées par voies intraveineuses dans la queue des souris (injections de 5s) en utilisant 30 μg de plasmide (codant pour *GRHL2* ou de plasmide contrôle) dilués dans une solution saline [Chlorure de sodium à 0.9% p/v ; le volume final a été ajusté au poids de chaque souris (0,1 mL/g a été utilisé)]. Après l'injection hydrodynamique, les souris ont été surveillées de près pendant 2 heures. Les foies ont été prélevés 48 h après les injections.

b. Surexpression de *GRHL2* dans les foies de souris par une méthode virale

Les virus (AAV8 ; 4.10^{11} GC/mL) ont été achetés auprès de VectorBuilder et expriment le même cDNA que les plasmides utilisés dans les expériences *in vitro* avec cellules en culture (NM_024915.4).

Après anesthésie, les souris ont été soumises à une injection rétro-orbitale de 100µL de solution virale stérile diluée dans 0.9% de NaCl (n=4 par groupe). Les souris ont été surveillées pendant 1h à la suite de l'injection. Les foies ont été prélevés 1 ou 2 semaines après l'injection.

2.4. Échantillons humains

a. Provenance des échantillons humains

Les foies de patients avec cirrhose en lien avec une consommation d'alcool proviennent de la cohorte TargetOH (« Comparison of Inflammatory Profiles and Regenerative Potential in Alcoholic Liver Disease ») (Louvet et al., 2015) (Bou Saleh et al., 2021; Louvet et al., 2017). Les échantillons dits contrôles sont issus d'une partie non cancéreuse de résections hépatiques réalisés chez des patients atteints de métastases au foie, mais dont la tumeur primaire n'est pas hépatique.

Les échantillons humains de foie ALD proviennent de patients ayant subi une greffe de foie dans le service de gastrologie à l'hôpital Huriez de Lille. Les échantillons ont directement été fixés pour les analyses histologiques ou congelés directement dans l'azote liquide pour les analyses géniques et protéiques. L'analyse a été autorisée par le comité éthique de l'hôpital de Lille. Tous les patients ont été informés de l'étude et ont donné leur consentement.

b. Organoïdes humains

Les organoïdes humains ont été réalisés en partenariat avec le laboratoire INFINITE, en particulier par les Dr. Solenne Dezitter et Dr. Mohammed Bou Saleh.

Brièvement, les tissus humains cryoconservés ou frais ont été découpés en petits morceaux et re-suspendus dans un milieu de lavage (1% FBS, 15mM HEPES dans du DMEM/F-12) et incubés 1 à 2 minutes sur glace. Après un second lavage, les morceaux de tissus ont été incubés 30 minutes à 37° dans un milieu DMEM/F-12 contenant 2.5mg/ml Collagenase Type IV, 10µg/ml DNase I et 15mM HEPES. Une fois la digestion réalisée, les cellules ont été centrifugées à 300g pendant 5 minutes et re-suspendu dans du Matrigel (Corning) et ensemencées au centre de chaque puits d'une plaque de culture 24 puits. Après une incubation de 20 minutes à 37°C, un milieu d'expansion a été ajouté [DMEM/F-12 supplémenté avec 1% P/S, 1% GlutaMAX, HEPES 10 mM, 1:50 supplémenté avec B27 sans vitamine A, 1:100 supplément N2, 1 mM N-acetylcysteine, 10% milieu Rspo1-conditioned (volume pour volume), 10 mM nicotinamide, 10 nM [Leu15]-gastrine recombinante humaine I, 50 ng/ml EGF humain, 100 ng/ml FGF10 humain, 25 ng/ml HGF humain, 10 µM Forskoline, 5 µM A83-01, 25 ng/ml Noggin humain, 30% de milieu Wnt3a-conditioned (volume pour volume) and 10 µM Rho kinase (ROCK) inhibiteur (Y-27632)]. Le milieu a été changé tous les 3 ou 4 jours jusqu'à la formation des organoïdes (7 à 10 jours).

Une fois les organoïdes formés, le milieu d'expansion a été remplacé par le milieu de différenciation [milieu d'initiation sans Noggin, Wnt3a and Rho kinase (ROCK) inhibiteur (Y-27632)]. À ce moment-là, les virus [AAV3-GRHL2 (VectorBuilder, NM_024915.4) ou virus contrôle] ont été ajoutées dans le milieu de différenciation à une concentration de 1.10^{10} par puit. Le milieu a été changé après 5 jours, puis, tous les 3 jours pendant 15 jours.

Résultats

I. L'HCT se produit sans passer par un état transcriptionnel proche d'un hépatoblaste

1.1. Création d'un atlas de données de scRNA-Seq de foies murins pour la caractérisation du processus de l'HCT

i. Mise en place de l'atlas de scRNA-Seq à partir des données de foies murins

Afin de comprendre les mécanismes et les changements d'expression géniques qui ont lieu pendant le phénomène de l'HCT, nous avons intégré différentes données de scRNA-Seq obtenues à partir de foies postnatals et adultes murins sains et de différents modèles de blessures hépatiques (aiguës ou chroniques). Ces données ont été comparées à celles de la différenciation à partir de l'hépatoblaste, via l'utilisation de données issues de foies embryonnaires. Ensemble, tous ces éléments permettent de capturer l'hétérogénéité hépato-cholangiocytaire (détail des données utilisées dans le **Tableau 2**).

La visualisation avec l'aide de carte de projection uniforme des variables (« *Uniform manifold approximation and projection* », UMAP) et l'étiquetage des cellules indique que les hépatoblastes, les hépatocytes et les cholangiocytes apparaissent comme des populations distinctes (**Figure 29A-B**). En effet, une fois l'intégration réalisée, ces 3 populations cellulaires ont été identifiées en regardant l'expression de différents marqueurs :

- *L'Afp* et le gène *Delta* à encoche non-canonique du ligand 1 de Notch (« *Delta like non-canonical notch Ligand 1* », *Dlk1*) pour les hépatoblastes ;
- Le *facteur de coagulation V* (« *Coagulation factor 5* », *F5*) et le gène codant pour l'apolipoprotéine *F* (« *apolipoprotein F* », *ApoF*) pour les hépatocytes ;
- *Krt19* et *Epcam* pour les cholangiocytes.

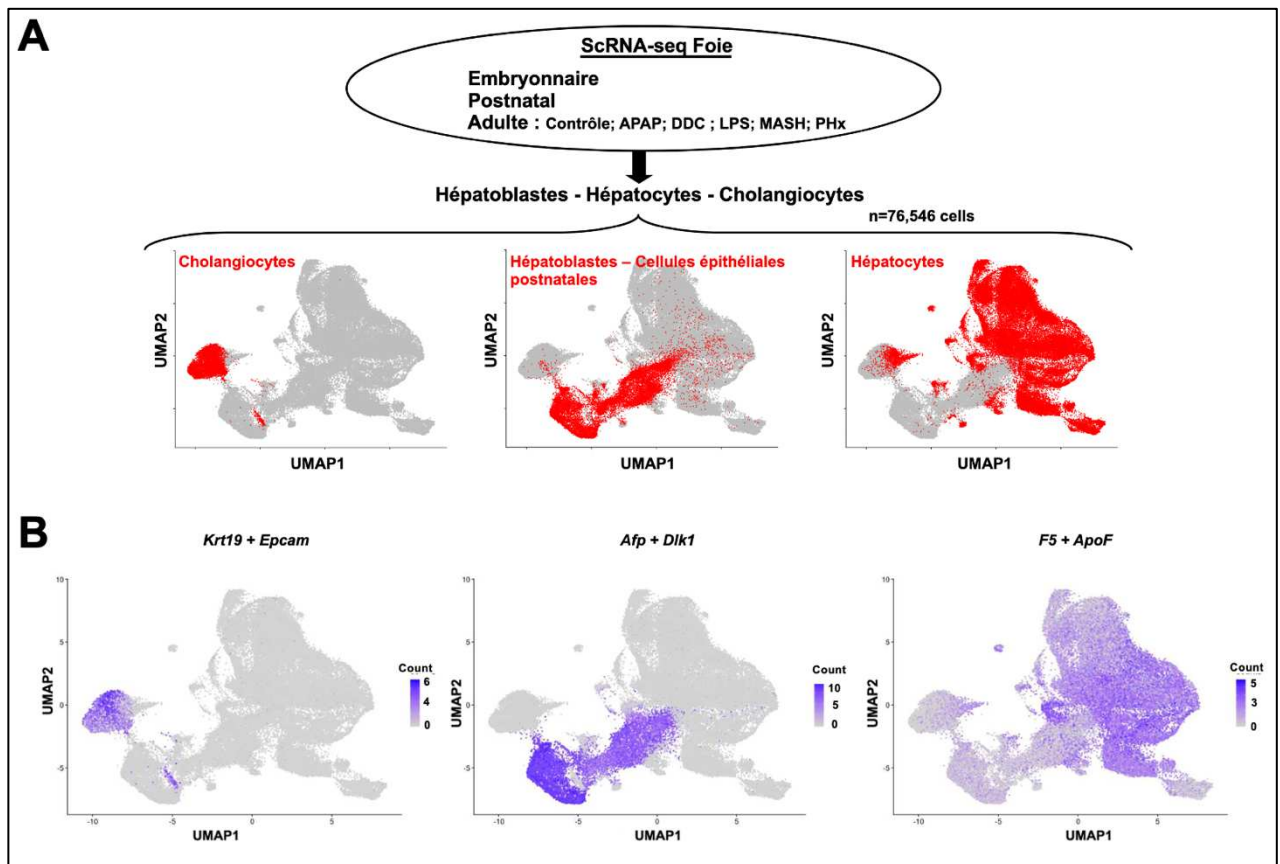


Figure 29 : Les hépatoblastes, les hépatocytes et les cholangiocytes : trois populations cellulaires bien distinctes.

(A) Aperçu de l'ensemble des données utilisées pour générer l'atlas de scRNA-Seq par l'utilisation des jeux de données issues de foies murins, ainsi que les projections UMAP des 76 546 cellules individuelles obtenues et le positionnement des hépatoblastes, hépatocytes et cholangiocytes. (B) UMAP de (A) utilisée pour projeter l'expression des différents marqueurs : *Afp* et *Dlk1* pour les hépatoblastes, *F5* et *ApoF* pour les hépatocytes et *Krt19* et *Epcam* pour les cholangiocytes.

ii. Identification du processus d'HCT

La méthode de séparation utilisée a permis d'assimiler un groupe de cellules correspondant aux cholangiocytes (cluster 1). Ce dernier est connecté à deux autres groupes de cellules : un contenant les hépatoblastes et des cellules épithéliales postnatales (cluster 2), et un autre groupe avec majoritairement des hépatocytes (cluster 3) (**Figure 30A-B**).

De plus, en regardant l'origine des hépatocytes du clusters 3, on voit que la grande majorité (95%) proviennent de souris sous régime DDC (**Figure 30C**) (Merrell et al., 2021). Ces derniers incluent les hépatocytes exprimant la protéine de fluorescence jaune (« *Yellow fluorescent protein* », YFP), utilisée pour le traçage cellulaire de l'étude (**Figure 30D**). Ces résultats suggèrent que le cluster 3 capture

l'HCT, connu pour être induit par le modèle murin DDC. Ces données montrent, également, que l'HCT n'est pas lié à la reprogrammation des hépatocytes arrivant dans d'autres modèles murins d'hépatotoxicité comme dans le modèle de PhX ou celui d'acétaminophène (« *acetaminophen* », APAP) (**Figure 30C-D**).

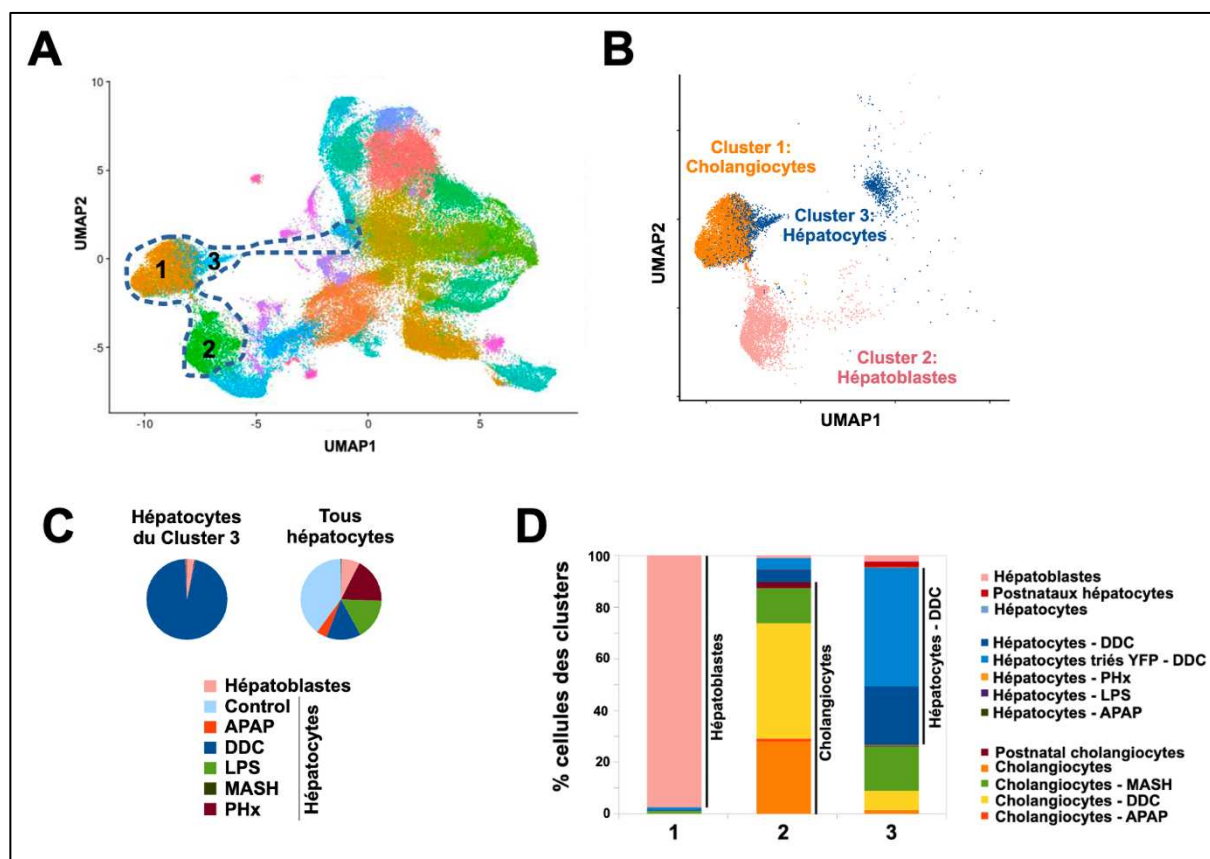


Figure 30 : Identification de l'HCT dans les cellules du parenchyme hépatique de souris

(A) UMAP de la **Figure 29A** montrant le résultat du clustering et indiquant les 3 clusters sélectionnés pour l'étude. (B) Zoom de la projection de UMAP de la **Figure 29A** mettant en évidence les 3 clusters sélectionnés pour l'analyse HCT et ceux pour la différenciation cholangiocytaire. Les cellules retrouvées dans ces 3 clusters sont mises en évidence dans le panel (D). (C) Diagramme circulaire montrant l'origine des hépatocytes retrouvés dans les 3 clusters en comparaison à tous les hépatocytes analysés dans la **Figure 29A** (A). (D) Données des origines des hépatocytes des clusters 1-3. Les bars horizontales ont été ajoutées pour indiquer les différents clusters.

1.2. HCT : processus de reprogrammation des cellules matures

En utilisant les cellules des groupes 1, 2 et 3, définis précédemment, nous avons effectué des analyses de vélocité de l'ARN, calculées sur la base du ratio des ARNs épissés et non-épissés dans les cellules. Les analyses montrent que deux mécanismes peuvent mener à la formation de cholangiocytes : le processus de différenciation cellulaire des hépatoblastes vers les cholangiocytes, et la

reprogrammation des hépatocytes, le processus de HCT. Ces deux phénomènes restent indépendants, montrant que lors de l'HCT les cellules ne semblent pas retrouver un transcriptome proche de l'hépatoblaste (**Figure 31A**). De la même manière, les analyses de probabilité de transition, réalisées à l'aide de CellRank2 (Weiler et al., 2024), montrent que le sous-groupe d'hépatocytes HCT, issu des souris sous régime DDC, a plus de chance de transiter vers un cholangiocyte, et qu'aucun hépatocyte du groupe HCT ne montre une probabilité de transition vers l'hépatoblaste (**Figure 31B**).

En outre, les analyses de « *pseudo-time* » réalisées ont permis de capturer les gènes dont l'expression varie au cours du processus HCT. Les résultats de cette analyse montrent que l'expression du marqueur biliaire *Epcam* augmente, tandis que celle du marqueur hépatocytaire *Apoc3* diminue. De manière cohérente avec les analyses précédentes, le marqueur hépatoblastique *Afp* n'est pas détecté, donc, pas induit de manière transitoire dans ce phénomène (**Figure 31C**).

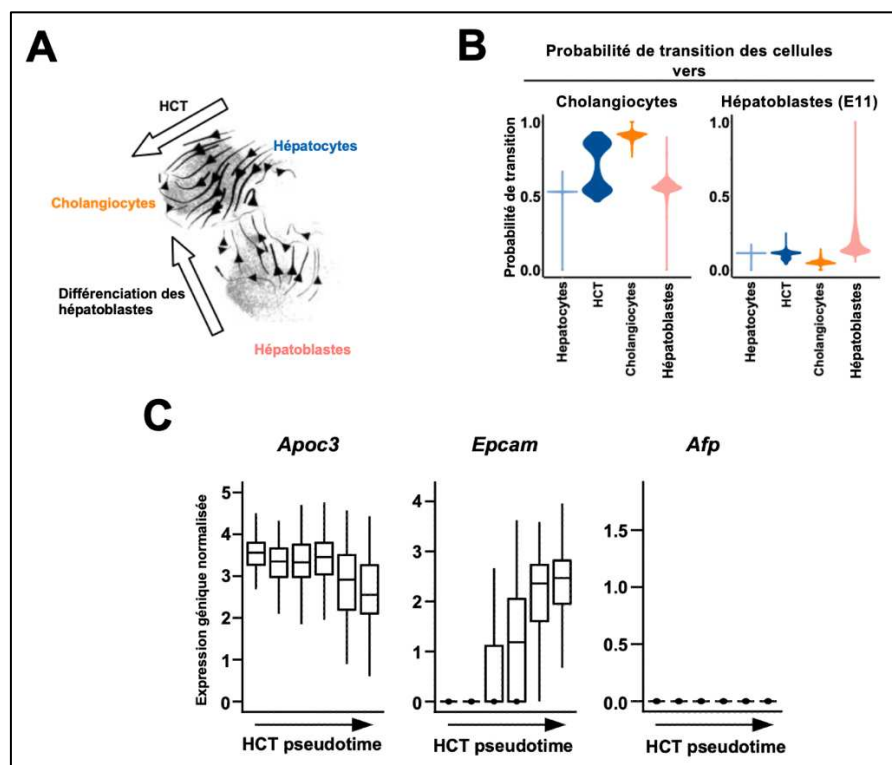


Figure 31 : L'HCT ne passe pas par un état de progéniteur.

(A) Analyse de vélocité réalisée sur les cellules de la **Figure 30B**. (B) Diagramme en violon montrant les probabilités de transition des cellules individuelles à partir des groupes indiqués vers les cholangiocytes adultes (à gauche) ou

vers les hépatoblastes (E11, à droite). Ces données ont été obtenues à partir de CellRank2. L'HCT correspond aux hépatocytes des souris sous régime DDC. **(C)** Diagramme en boîte montrant l'expression normalisée des différents marqueurs dans les sous-ensembles des cellules hépatocytaires pendant le processus HCT, définis grâce à l'analyse « *pseudo-time* ».

De la même manière, cette analyse « *pseudo-time* » montre que peu de gènes ont un pic d'expression dans un état intermédiaire de l'HCT (environ 10%) (**Figure 32A**). Ces gènes activés de manière intermédiaires dans le processus de l'HCT ne sont pas caractérisés par une expression spécifique dans les hépatoblastes, confirmant que l'HCT ne passe pas par l'expression d'une signature génique ressemblant à des cellules embryonnaires (**Figure 32B**).

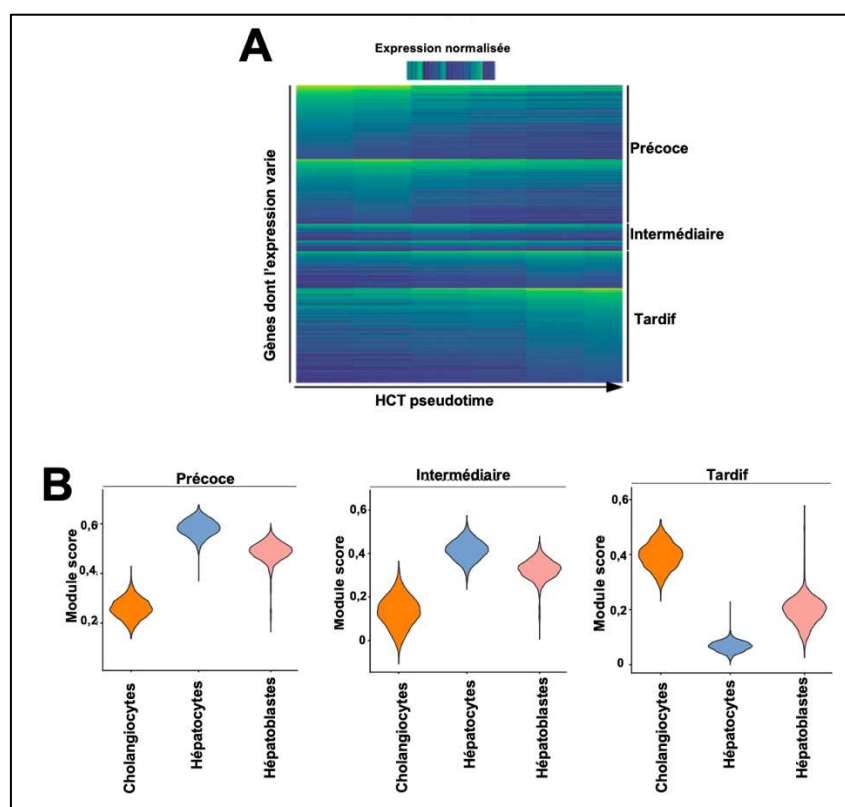


Figure 32 : Les gènes activés de manière intermédiaire pendant l'HCT ne sont pas caractéristiques des hépatoblastes.

(A) Carte de chaleur montrant l'expression normalisée des gènes, subissant des changements géniques pendant l'HCT. Les gènes sont classés en fonction de leur pic d'expression génique, qui peut être précoce, intermédiaire ou tardive. **(B)** Les cellules utilisées pour l'analyse de vélocité ont ensuite été réutilisées pour l'analyse des gènes précoces, intermédiaires ou tardifs du panneau (A) dans les hépatoblastes, hépatocytes ou cholangiocytes.

Toutes ces données indiquent que le mécanisme de l'HCT n'entraîne pas l'acquisition de manière transitoire pour les hépatocytes d'un état ressemblant à celui de cellules progénitrices.

II. Induction d'une signature de cellule épithéliale mono-polarisée nécessaire à l'HCT, normalement réprimée dans des hépatocytes sains

2.1. Activation d'un nouveau programme épithélial lors de l'HCT

Afin de caractériser les changements transcriptionnels ayant lieu lors de l'HCT, une comparaison du transcriptome d'hépatocytes sains à celui d'hépatocytes ayant transités vers un cholangiocyte a été réalisée. Pour cela, une ACP a été réalisée en utilisant les gènes communs aux études de (Merrell et al., 2021) et de (Zummo et al., 2023) (méthode détaillée dans le chapitre **Matériels et méthodes 1.2.b p.109**). Comme attendu, les résultats vérifient que les hépatocytes en cours d'HCT, issus des souris DDC, ont un transcriptome global proche de celui des cholangiocytes (**Figure 33A**).

i. Expression d'une signature épithéliale pendant l'HCT

Par la suite, les gènes ont été classés selon leurs taux de variation d'expression (« *Fold Change* », FC) entre les hépatocytes et les hépatocytes HCT (hépatocytes en cours de transdifférenciation vers un cholangiocyte). Les 3 000 gènes ayant une variation d'expression significative ($q\text{-value} < 0,05$), ont été utilisés pour réaliser une analyse fonctionnelle (analyse d'enrichissement des voies biologiques). Pour cela, l'outil DAVID, permettant de regrouper les gènes selon les termes d'enrichissements similaires et de réaliser une classification de ces gènes, a été utilisé. De manière surprenante, les gènes induits pendant l'HCT montrent des termes d'enrichissement caractéristiques de l'archétype des cellules épithéliales tels que les jonctions serrées et la présence de cils. (**Figure 33B**). Au contraire pour les gènes réprimés pendant ce processus, les annotations retrouvées sont en lien avec les fonctions hépatocytaires (**Figure 33C**).

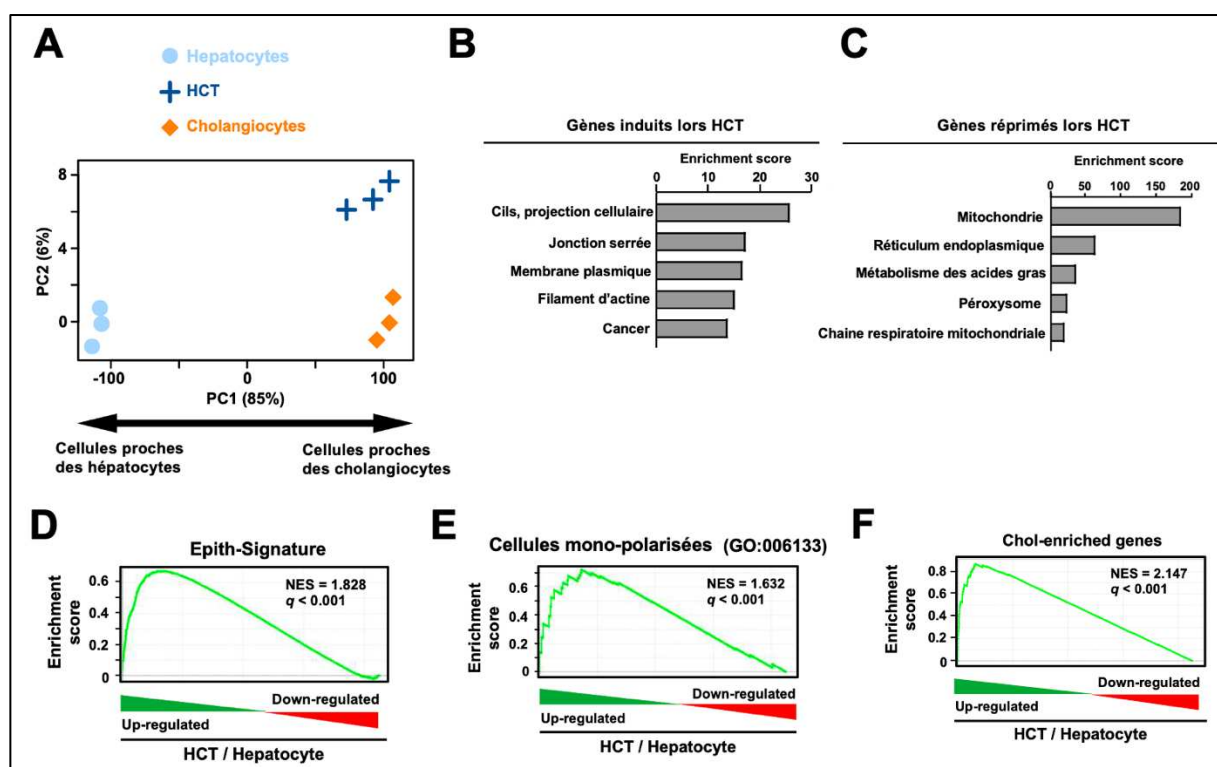


Figure 33 : Activation d'une signature épithéliale mono-polarisée lors de l'HCT.

(A) ACP des changements transcriptionnels lors de HCT. Les données issues de (Merrell *et al.*, 2021), obtenues par traçage cellulaire, issues des souris traitées avec du DDC, ont été projetées dans les dimensions d'une ACP réalisée à partir d'hépatocytes et de cholangiocytes des souris contrôles de (Zummo *et al.*, 2023). (B) (C) Annotation fonctionnelle des gènes induits (B) ou réprimés (C) lors de l'HCT réalisée grâce à l'outil DAVID (q-value < 0.05, top 3 000 gènes). Les gènes différentiellement exprimés ont été obtenus en comparant les données issues des hépatocytes HCT de souris traitées avec DDC aux hépatocytes sains dans les données de transcriptomiques de (Merrell *et al.*, 2021). Les termes enrichis ont été regroupés au sein de différents groupes. Ici, seuls les 5 premiers groupes sont affichés et associés avec leur score d'enrichissement. (D) (E) (F) Graphiques d'enrichissement issues de l'analyse GSEA réalisée en utilisant comme liste de gènes soit Epith-Signature (Breschi *et al.*, 2020) (D), une liste de termes géniques associés à « l'établissement ou le maintien d'une cellule mono-polarisée » (GO:0061339) (E) ou une liste de gènes de enrichis dans les cholangiocytes (F) et comme données géniques les changements transcriptionnels présentés en (B). Le NES définit le score d'enrichissement.

Pour aller plus loin, les analyses d'enrichissement de signature génique (« *gene set enrichment analysis* », GSEA) ont été réalisées en utilisant une signature de gène caractérisant les cellules épithéliales (dénommée Epith-Signature par la suite) (Breschi *et al.*, 2020), une liste de gènes caractéristiques des cellules épithéliales mono-polarisées (GO : 006133) et une liste de gènes enrichis dans les hépatocytes, issus du « *Human Protein Atlas* » (Dusart *et al.*, 2023). Les résultats ont indiqué que ces trois listes sont préférentiellement exprimées dans les hépatocytes en cours de l'HCT en comparaison aux hépatocytes sains (**Figure 33D-F**).

ii. Les hépatocytes, des cellules épithéliales particulières

En comparant l'expression des gènes contenus dans la Epith-Signature parmi tous les types cellulaires identifiés comme épithéliaux dans le « Human Protein Atlas », les hépatocytes sains sont caractérisés par une expression faible de ces gènes contrairement aux cholangiocytes, qui eux expriment de manière forte ces gènes comme les autres cellules épithéliales (**Figure 34A**).

Pour aller plus loin, nous avons voulu voir si cette différence était liée au fait que les cellules peuvent présenter des transcriptomes différents. Ainsi, seuls les gènes ayant un indice de spécificité (« *tau index* ») élevé (supérieur 0,9) dans chaque cellule épithéliale ont été considérés. Dans ce cas, les hépatocytes sont les cellules ayant le nombre de gènes spécifiques le plus important comparé aux autres cellules épithéliales, incluant les cholangiocytes (**Figure 34B**). Ces résultats montrent bien que les hépatocytes ont un transcriptome différent d'une cellule épithéliale classique.

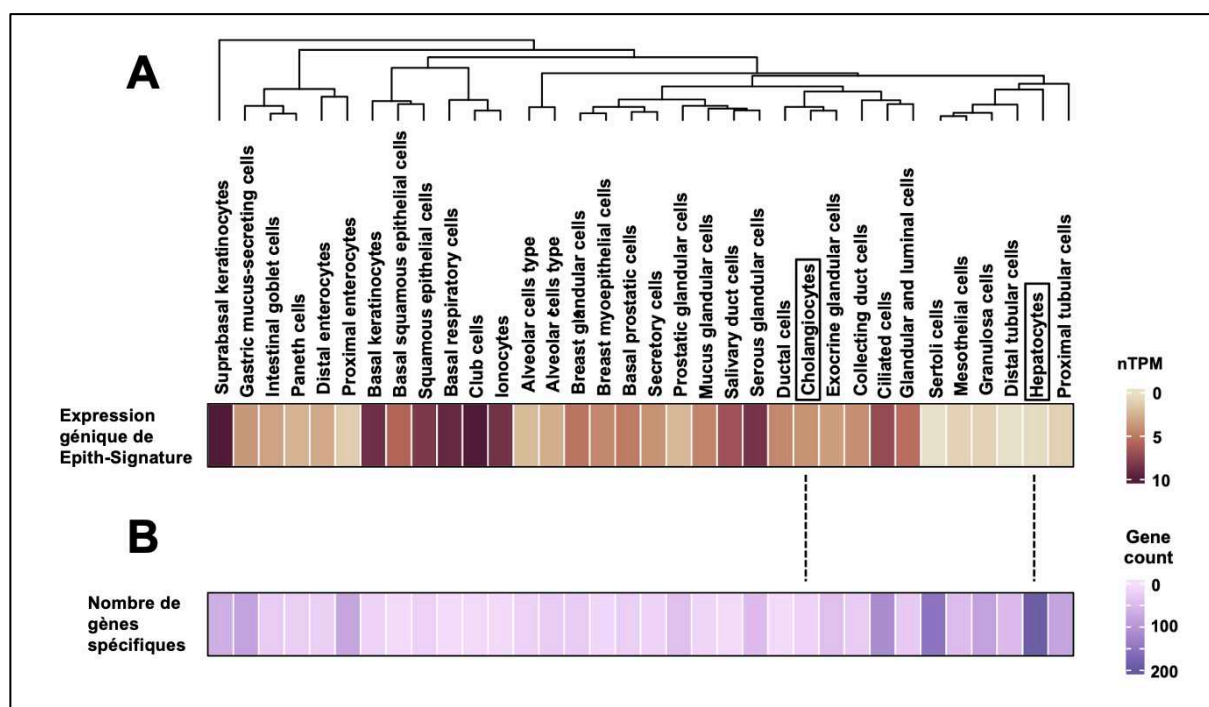


Figure 34 : Les hépatocytes, des cellules épithéliales avec un transcriptome particulier.

(A) Regroupement hiérarchique de différents types de cellules épithéliales humaines. La carte de chaleur montre la médiane de l'expression génique de Epith-Signature dans les différentes cellules épithéliales, utilisée pour la classification hiérarchique. (B) Nombre de gènes spécifiquement exprimés dans les cellules épithéliales en utilisant un index de spécificité (τ index > 0.9). Les types cellulaires sont rangés selon le dendrogramme représenté en (G).

Ensemble, ces données montrent que l'événement principal de l'HCT est l'induction d'une signature caractéristique des cellules épithéliales mono-polarisées qui n'est pas exprimée dans un hépatocyte sain.

2.2. Rôle du complexe PrC2 dans la répression de la signature épithéliale dans un hépatocyte sain

a. Caractérisation épigénétique des gènes codants pour les TFs induits lors de l'HCT

Pour aller plus loin dans l'analyse des changements transcriptionnels intervenants pendant l'HCT, la même ACP qu'en **Figure 33A** a été réalisée mais en sélectionnant uniquement les gènes codants pour les TFs (**Figure 35A**). Les mêmes résultats sont retrouvés :

- L'axe 1 sépare les hépatocytes des cholangiocytes ;
- Les hépatocytes HCT ont un transcriptome proche de celui des cholangiocytes.

Ainsi, l'étude des changements d'expression des gènes codants pour les TFs est suffisante pour caractériser les mécanismes ayant lieu pendant l'HCT. De cette façon, parmi les gènes exprimés de manière différentielle ($q\text{-value} < 0,05$) dans les hépatocytes ayant subi une HCT, les gènes codants pour les TFs ont été sélectionnés, permettant de mettre en évidence, respectivement, 115 gènes réprimés et 423 gènes induits.

i. Ségrégation des TFs dérégulés pendant l'HCT

Afin de confirmer la dérégulation des TFs induits dans l'HCT, l'expression de ces derniers a été évaluée dans un autre modèle murin, un modèle murin du syndrome ALGS où le processus a également été identifié (Schaub et al., 2018). Concordant avec notre précédente identification, 82% des gènes codants pour les TFs, identifiés comme induits pendant le processus de l'HCT, sont également induits dans les hépatocytes transdifférenciés issus de ce modèle (**Figure 35B**). Ces données valident le rôle de ces TFs pendant l'HCT. De plus, une partie des gènes codants pour les TFs

réprimés pendant l'HCT correspond aux Hep-ID TFs, TFs hépatocytaires spécifiques identifiés dans (Dubois-Chevalier et al., 2023), montrant la perte d'identité hépatocytaire au cours du phénomène (**Figure 35C**).

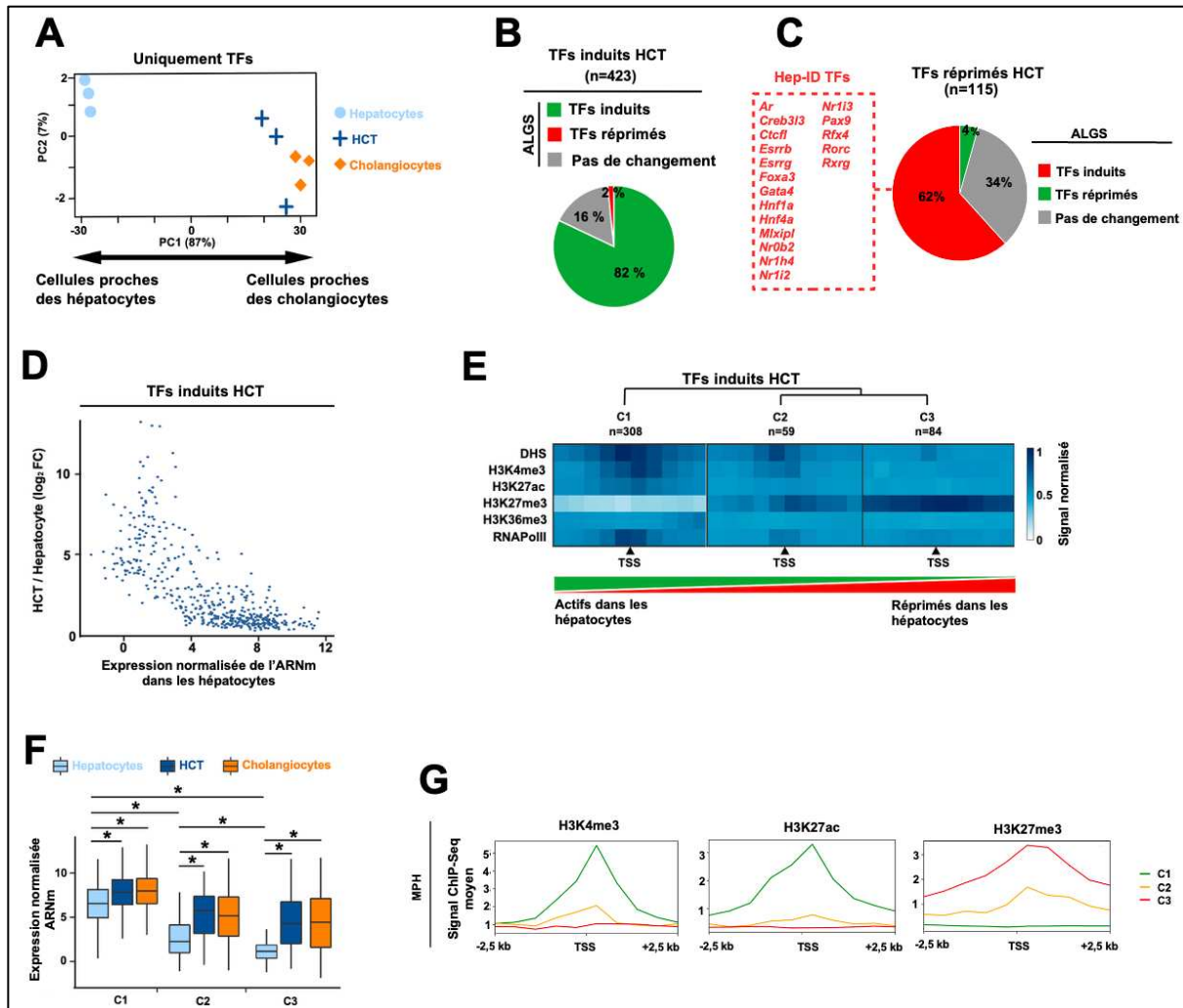


Figure 35 : L'HCT implique l'induction de gènes codants pour des TFs réprimés de manière épigénétique dans des hépatocytes sains.

(A) ACP réalisée comme dans **Figure 33A** uniquement sur les gènes codants pour des TFs. (B) (C) Gènes codants pour des TFs induits (B) ou réprimés (C) lors de l'HCT obtenus en comparant les hépatocytes tracés en cours d'HCT dans les souris sous régime DDC avec les hépatocytes triés sains dans les données de RNA-Seq de (Merrell et al., 2021) [q -value $< 0,05$ et $\log_2FC > 0$ ou < 0]. Cette liste de gènes a été mise en parallèle d'un autre modèle murin d'hépatocytes en cours d'HCT, le modèle ALGS de (Schaub et al., 2018) [q -value $< 0,150$ et $\log_2FC > 0$ ou < 0]. Les gènes avec une q -value $> 0,150$ n'ont pas de différence d'expression génique (« Pas de changement »). (D) Comparaison de l'expression basale des gènes codants pour les TFs induits dans des hépatocytes sains avec leurs niveaux d'induction dans les hépatocytes HCT. (E) Résultats de l'analyse SPARK au niveau des promoteurs des gènes codants pour les TFs induits lors du processus d'HCT permettant de comparer les signaux moyens d'accessibilité de la chromatine (DHS), de l'ARN polymérase II (RNAPol II) ou des PTMs des histones indiquées (H3K4me3, H3K27ac, H3K27me3 et H3K36me3) et de définir 3 clusters différents. Le nombre de gènes codants pour des TFs dans chaque groupe est indiqué. Le dendrogramme montre la méthode de regroupement hiérarchique utilisée pour définir les groupes. Les flèches indiquent le TSS. (F) Boîtes à moustaches de l'expression des gènes codants pour les TFs des clusters C1-3 dans des hépatocytes ou cholangiocytes sains ou dans des hépatocytes en cours d'HCT. Les différences statistiquement significatives ont été évaluées par un test ANOVA two-way avec une comparaison multiple de Dunett utilisée en test post-hoc. * $q < 0,05$ (G) Signal moyen de ChIP-Seq pour les

marques d'histones indiquées aux promoteurs des gènes codants pour les TFs des groupes C1-3 dans les hépatocytes primaires. Le signal est montré sur une fenêtre de +/- 2,5 kb autour du TSS.

Pour caractériser les gènes codants pour les TFs induits pendant l'HCT, une corrélation entre leurs expressions basales dans des hépatocytes sains et leurs inductions dans les hépatocytes HCT a été réalisée. Ces TFs montrent une disparité d'expression dans un hépatocyte sain. Cependant, un sous-ensemble de TFs, ayant une expression basale très faible dans un hépatocyte sain, contient les TFs les plus induits pendant l'HCT (**Figure 35D**). Ainsi, pendant le phénomène de reprogrammation hépato-cholangiocytaire, la levée des marqueurs de la répression au niveau de la chromatine semble être nécessaire pour permettre leurs expressions.

ii. Différents profils épigénétiques aux promoteurs des TFs induits

La caractérisation des promoteurs des gènes codants pour les TFs induits pendant l'HCT a été réalisée à l'aide de données cistromiques et épigénomiques dans des hépatocytes sains. Pour cela, nous avons utilisé le logiciel SPARK, qui sépare les gènes selon les profils épigénétiques au niveau de leurs TSSs (Nielsen et al., 2012). Différents types de modulateurs épigénétiques, provenant de foies de souris saines ou d'hépatocytes isolés de foies sains, ont été incrémentés dans le logiciel :

- L'accessibilité de la chromatine [sites hypersensibles de la DNase I (« *DNase I hypersensitivity* », DHS)] ;
- Des marques PTMs d'histones en lien avec l'activation génique [H3K4me3 et H3K27ac] ;
- Des PTMs d'histones caractéristiques de l'élongation de la transcription [la triméthylation de la lysine 36 de l'histone 3 (H3K26me3) et l'ARN polymérase II (« *RNA-polymerase II* », RNA-Pol II)] ;
- Une marque de répression de la transcription [H3K27me3].

Trois groupes de promoteurs possédants des profils épigénétiques divergents ont ainsi pu être identifiés (**Figure 35E**). Le groupe contenant le plus de gènes (C1) comporte des promoteurs actifs dans un foie sain avec des signaux d'accessibilité de la chromatine et des niveaux forts de H3K4me3 et H3K27ac, tandis que les groupes C2

et C3 présentent des promoteurs, respectivement, peu actifs mais compétents sur le plan transcriptionnel et réprimés. En effet, les promoteurs situés dans C2 montrent des niveaux modérés pour les marques d'activation et de répression de la chromatine (H3K4me3 et H3K27me3), faisant penser à des promoteurs bivalents. Les promoteurs contenus dans C3 sont caractérisés par des signaux forts de H3K27me3 et n'ont aucun signal d'activation de la transcription (**Figure 35E-G**). En lien avec les statuts épigénétiques retrouvés au niveau des promoteurs, les TFs des groupes C2 et C3 ont des expressions basales faibles dans les hépatocytes, mais sont plus fortement induits pendant l'HCT en comparaison du groupe C1 (**Figure 35F**).

iii. Répression des promoteurs de C3 indépendante de H3K9me3

Parmi les différentes marques PTMs des histones, la tri-méthylation de la lysine 9 de l'histone 3 (H3K9me3) est importante dans l'extinction des gènes d'un programme transcriptionnel alternatif. Cette marque peut donc influencer l'identité cellulaire (Nicetto & Zaret, 2019). Afin d'évaluer son rôle dans notre étude, particulièrement au niveau des promoteurs des gènes induits lors de l'HCT, la même analyse des profils épigénomiques des promoteurs avec le logiciel SPARK a été réalisée en ajoutant cette marque. Malgré cette incrémentation, les résultats restent inchangés : la séparation des promoteurs des gènes codants pour les TFs induits pendant l'HCT ne semble pas dépendante de H3K9me3 (**Figure 36A**).

Or, cette marque de répression de la transcription est majoritairement présente dans un état de condensation fermé de la chromatine et permet de rendre l'ADN inaccessible à la transcription. Nous avons, donc, voulu identifier les différents statuts chromatinien de nos promoteurs, nos groupes C1-3. Pour cela, les groupes ont été croisés avec une étude biophysique d'identification des régions de la chromatine dans des hépatocytes murins. Dans cette étude, les auteurs utilisent une méthode de séquençage de l'hétérochromatine résistante à la sonication à la suite d'un gradient de sucrose (« *sequencing of sonication resistant heterochromatin* », srHC-Seq). Cette méthode permet de segmenter la chromatine en trois parties :

- L'hétérochromatine, régions non fragmentées résistantes à la sonication qui apparaissent dans le bas du gradient ;

- L'euchromatine, régions fragmentées retrouvées dans le haut du gradient ;
- Un domaine intermédiaire ne correspondant ni à de l'hétérochromatine, ni à de l'euchromatine.

Les résultats indiquent que, contrairement aux groupes C1 et C2, les promoteurs du groupe C3 ne sont pas distingués par leurs états de condensation de la chromatine : ils sont retrouvés à la fois dans l'hétérochromatine et l'euchromatine (**Figure 36B**). Afin de confirmer ces résultats, une autre méthode de segmentation de la chromatine a été utilisée, provenant de (C. Zhang et al., 2021). Cette dernière catégorise les promoteurs selon la présence ou non de différentes PTMs au niveau des histones, comme H3K27me3, H3K9me3, H3K4me3 ou encore le variant de l'histone 2A (H2A.Z), et combine les résultats avec des signaux provenant de séquençage de la chromatine accessible par la transposase (« *assay for transposase-accessible chromatin* », ATAC-Seq). De la même manière qu'avec la méthode biophysique, les promoteurs de C3 sont retrouvés à la fois dans des domaines d'hétérochromatine et d'euchromatine (**Figure 36C**). Ces résultats pourraient expliquer que la marque de répression H3K9me3 n'est pas impliquée dans la répression des gènes du groupe C3 dans un hépatocyte sain.

De plus, l'analyse de la méthylation de l'ADN, réalisée grâce à la méthode de séquençage au bisulfite du génome (« *genome bisulfite sequencing* », WGBS) montre que tous nos groupes présentent une hypo-méthylation de l'ADN autour de leurs TSSs (**Figure 36D**). Or, la présence de H3K9me3 est normalement corrélée avec la méthylation des promoteurs (C. Zhang et al., 2021).

Toutes ces données soulignent le rôle de H3K27me3 dans la répression des gènes du destin cellulaire alternatif dans les deux compartiments chromatiniens de manière indépendante de H3K9me3.

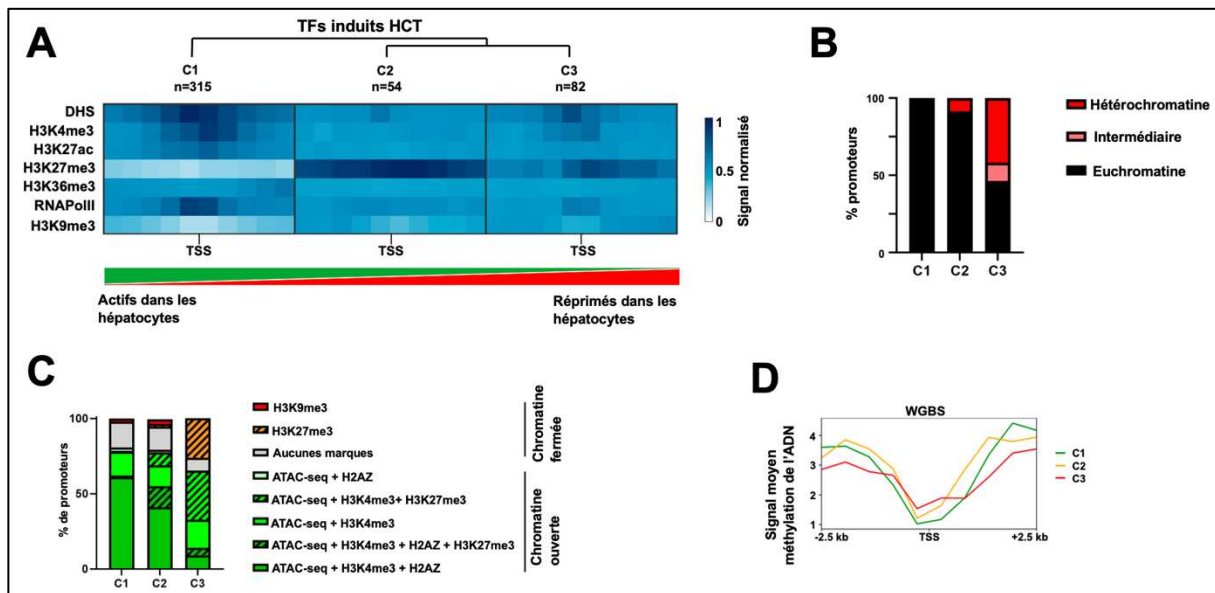


Figure 36 : Les promoteurs du groupe C3 sont retrouvés à la fois dans des domaines d'hétérochromatine et d'euchromatine et leur régulation est indépendante de H3K9me3.

(A) Résultats de l'analyse SPARK comme dans **Figure 35D** en ajoutant les signaux de H3K9me3 pour les sous-groupes identifiés. Le dendrogramme montre la méthode de regroupement hiérarchique des clusters. Les flèches indiquent le TSS. (B) Les domaines des promoteurs des groupes C1-3 ont été caractérisés grâce à l'étude de (Grindheim et al., 2019) par la méthode de srHC-Seq. Les domaines intermédiaires correspondent aux régions situées ni dans l'hétérochromatine ni dans l'euchromatine. (C) Distribution des promoteurs des groupes C1-3 au sein des différents domaines de chromatine en utilisant l'absence ou la présence d'ATAC-Seq et des marques de ChIP-Seq H2AZ, H3K9me3, H3K4me3, H3K27me3 par le biais de ChromHMM (C. Zhang et al., 2021). (D) Signal moyen de méthylation de l'ADN au niveau des promoteurs C1-3 dans des foies murins. Les signaux ont été obtenus à partir de la méthode WGBS. Le signal est montré sur une fenêtre de +/- 2,5 kb autour du TSS.

iv. Transposition des résultats chez l'Homme

Enfin, nous avons vérifié que les marques épigénétiques retrouvées dans les foies de souris sont concordantes avec celles retrouvées dans une lignée cellulaire d'hépatocarcinome humain (« *hepatocellular carcinoma* », HepG2) et chez l'Homme. L'analyse des données de ChIP-Seq de H3K27ac, H3K4me3 et H3K27me3 dans ces cellules et les foies humains sains valident la similarité des profils épigénétiques avec nos résultats précédents (**Figure 37A-B**).

En lien avec les différents profils épigénétiques, l'analyse des données d'expression géniques dans les HepG2 et chez l'Homme vérifie que les gènes du groupe C1 sont plus exprimés dans les hépatocytes que ceux du groupe C3 (**Figure 37C-D**). Ces résultats valident, ainsi, la robustesse de notre classification de gènes codants pour des TFs et la pertinence de leur étude dans nos modèles cellulaires et chez l'Homme.

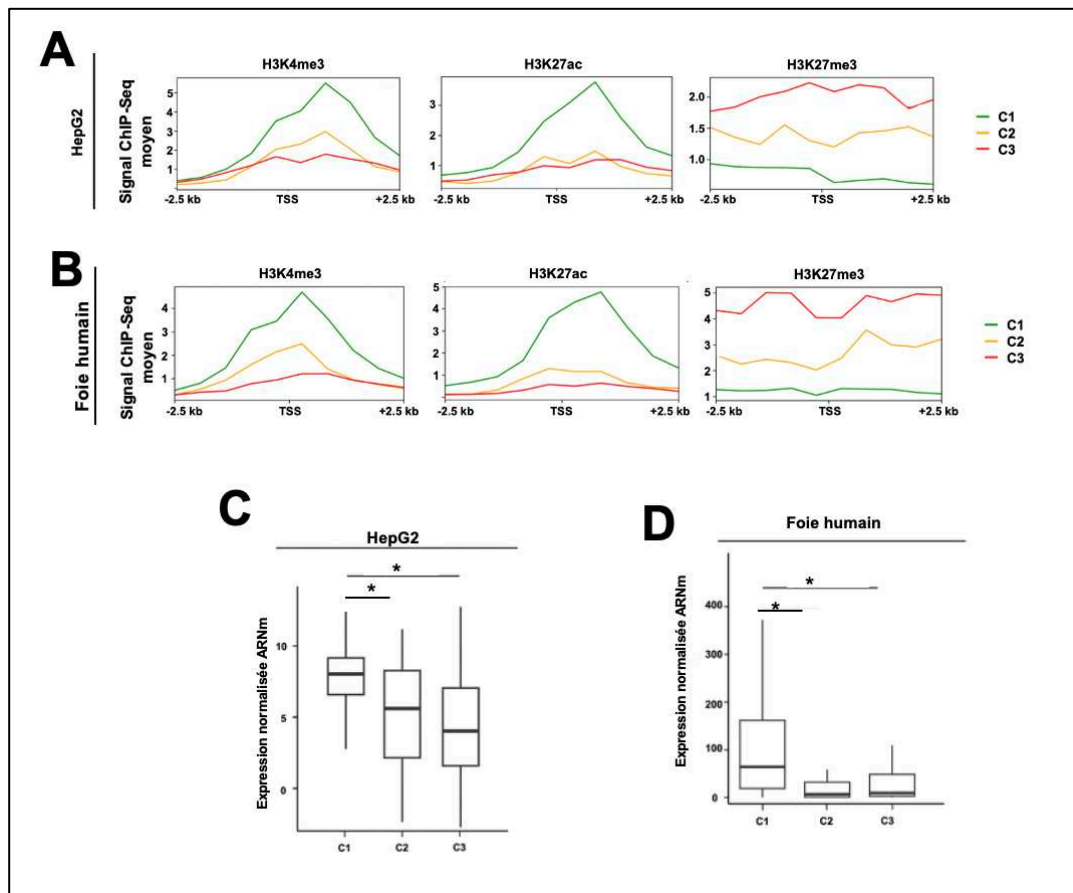


Figure 37 : Transposition des analyses épigénétiques sur nos modèles cellulaires et chez l'Homme.

(A)(B) Signal moyen de ChIP-Seq pour les marques d'histones indiquées aux promoteurs des gènes codants pour les TFs des groupes C1-3 dans les HepG2 (A) ou dans des foies de patients sains (B). Le signal est montré sur une fenêtre de +/- 2,5 kb autour du TSS. (C)(D) Boîte à moustaches de l'expression des gènes codants pour les TFs des clusters C1-3 dans les HepG2 (C) ou dans des foies sains de patients (D). Les différences statistiques ont été évaluées à l'aide d'un test de Kruskal-Wallis suivi d'un test de correction multiple de Dunn. * $q < 0,05$.

b. Rôle du complexe PrC2 dans l'HCT

Comme expliqué, l'activité enzymatique du complexe PrC2 repose, principalement, sur l'activité des enzymes EZH1 et EZH2. Afin de vérifier la présence du complexe PrC2 au niveau des promoteurs des groupes C1-3, les données de ChIP-Seq de EZH2, provenant de l'étude de (N. Yang et al., 2023), ont été utilisées. Les résultats vérifient que la présence de H3K27me3 sur les promoteurs des gènes codants pour les TFs du groupe C3 est en lien avec le recrutement de EZH2 (**Figure 38A**).

De plus, une délétion des enzymes EZH1 et EZH2 (EZH1/2 KO) dans des foies de souris valide que la diminution du signal de H3K27me3 au promoteur des gènes des groupes C2 et C3 est corrélée avec l'induction transcriptionnelle de ces gènes,

selon les données de transcriptomiques issues de (Grindheim et al., 2019) (**Figure 38B-C**). Cette perte de la marque de répression permet, également, l'induction des signatures géniques Epith-Signature et Chol-enriched (**Figure 38D**).

Afin de valider ces résultats, un modèle *in vitro* d'inhibition de l'activité de EZH1 et EZH2 a été mis en place. Les HepG2 ont été traitées avec des inhibiteurs de ces 2 enzymes : le Valemetostat Tosylate (Duan et al., 2020) et l'UNC1999 (Kusakabe et al., 2021), dénotés EZHi. Après 4 jours de traitement avec les molécules utilisées à 1 μ M chacune, une diminution de la marque H3K27me3 est retrouvée dans les cellules (**Figure 38E**). L'analyse des données de RNA-Seq, issues de ce traitement, montre que la perte de H3K27me3 est liée à l'induction des gènes codants pour les TFs des groupes C2-3.

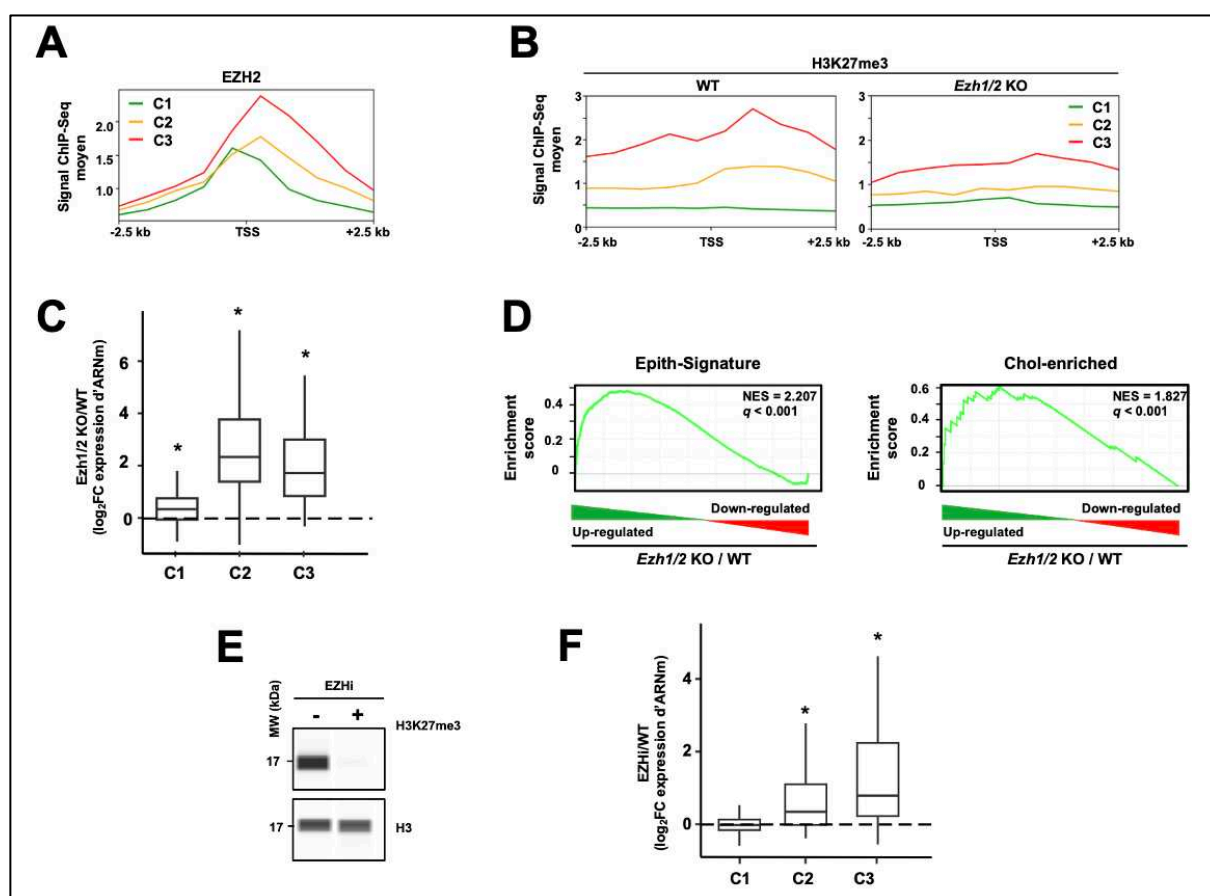


Figure 38 : Répression des gènes codants pour les TFs induits pendant l'HCT par le complexe PrC2.

(A) Signal moyen de ChIP-Seq contre EZH2 aux promoteurs des gènes codants pour les TFs des groupes C1-3 dans les foies de souris saines. Le signal est montré sur une fenêtre de +/- 2,5 kb autour du TSS. (B) Signal moyen de ChIP-Seq de la marque H3K27me3 aux promoteurs des gènes codants pour les TFs des groupes C1-3 dans le foie de souris contrôle (WT) ou constitutivement délétées de *Ezh1* et *Ezh2* (*Ezh1/2* KO). Le signal est montré sur une fenêtre de +/- 2,5 kb autour du TSS. (C) Boîte à moustache de la modulation d'expression des gènes des groupes C1-3 obtenues par analyse RNA-Seq des foies de souris constitutivement délétées de *Ezh1* et *Ezh2*

(*Ezh1/2* KO, n=3) comparées aux foies de souris contrôles (WT, n=2). Les différences statistiquement significatives, afin de déterminer si la médiane des log2FC est différente de 0, ont été évaluées par un t-test one-sided one-sample Wilcoxon suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$. **(D)** Graphiques d'enrichissement issus de l'analyse GSEA réalisée en utilisant comme liste de gènes soit Epith-Signature (Breshi *et al*, 2020) ou une liste de gènes de enrichis dans les cholangiocytes et les données de changements transcriptionnels présentés en (C). Le NES définit le score d'enrichissement. **(E)** Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de H3K27me3 et H3 dans des HepG2 traités pendant 4 jours avec ou sans inhibiteurs de EZH1/2 (EZH1). Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire **(F)** Boîte à moustaches de la modulation d'expression des gènes des groupes C1-3 obtenus par analyse RNA-Seq à partir des cellules traitées ou non avec des inhibiteurs de EZH1/2 ou 0,2% de DMSO comme contrôle (n=4, réplicats biologiques indépendants). Les différences statistiquement significatives, afin de déterminer si la médiane des log2FC est différente de 0, ont été évaluées par un t-test one-sided one-sample Wilcoxon suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$.

Toutes ces données indiquent que le phénomène de l'HCT implique l'induction de gènes codants pour des TFs, normalement réprimés par la marque H3K27me3, dépendant du complexe PrC2, dans un hépatocyte sain.

III. GRHL2 : un TF épithélial favorisant l'HCT

3.1. GRHL2 : régulateur de la différenciation des cellules épithéliales impliqué dans l'HCT

a. Identification d'un régulateur transcriptionnel de l'HCT

Nous avons, ensuite, voulu démontrer le rôle fonctionnel des TFs induits pendant l'HCT. Nos analyses se focaliseront sur les gènes codants pour les TFs du groupe C3 puisqu'en plus d'être réprimés dans les hépatocytes sains et induits pendant l'HCT, le groupe C3 présente un indice de spécificité plus élevé que les autres groupes de gènes (**Figure 39A**). Pour aller plus loin, nous avons supposé que ce groupe C3 comportait des TFs clés nécessaires à la mise en place d'un programme transcriptionnel épithélial alternatif pendant l'HCT. Ainsi, chaque TF du groupe C3 a été analysé de manière individuelle en regardant :

- Une classification décroissante correspondante au nombre de cellules épithéliales exprimant chaque TF (allant des facteurs exprimés dans toutes les cellules épithéliales, soit ubiquitaires, à ceux exprimés dans peu de cellules épithéliales, soit plus spécifiques) ;
- Son induction pendant l'HCT ;
- Sa présence dans les listes de gènes Epith-Signature et Chol-enriched.

De cette manière, différents TFs émergent de l'analyse comme étant exprimés dans peu de cellules épithéliales et fortement induits pendant l'HCT. On retrouve, notamment, le *facteur de transcription de type Grainyhead 2* (« *Grainyhead like transcription factor 2* », *GRHL2*) ou encore le *facteur de transcription homologue ETS* (« *ETS homologous factor* », *EHF*) (**Figure 39B**). Or, *GRHL2* est décrit dans la littérature comme ayant un rôle essentiel dans l'acquisition par une cellule du phénotype épithélial (Jacobs et al., 2018) (A. F. Chen et al., 2018). De plus, *GRHL2* semble réguler l'expression des gènes codants pour *Ehf* et pour des *répresseurs transcriptionnels des domaines Ovo 1 et 2* (« *ovo like transcriptionnals repressors 1/2* », *Ovo1/2*). Ensemble, ces TFs sont suspectés d'inhiber la TEM pour maintenir un phénotype épithélial dans les cellules cancéreuses (Watanabe et al., 2019). *GRHL2* ressort, donc, comme un régulateur principal de la différenciation des cellules épithéliales.

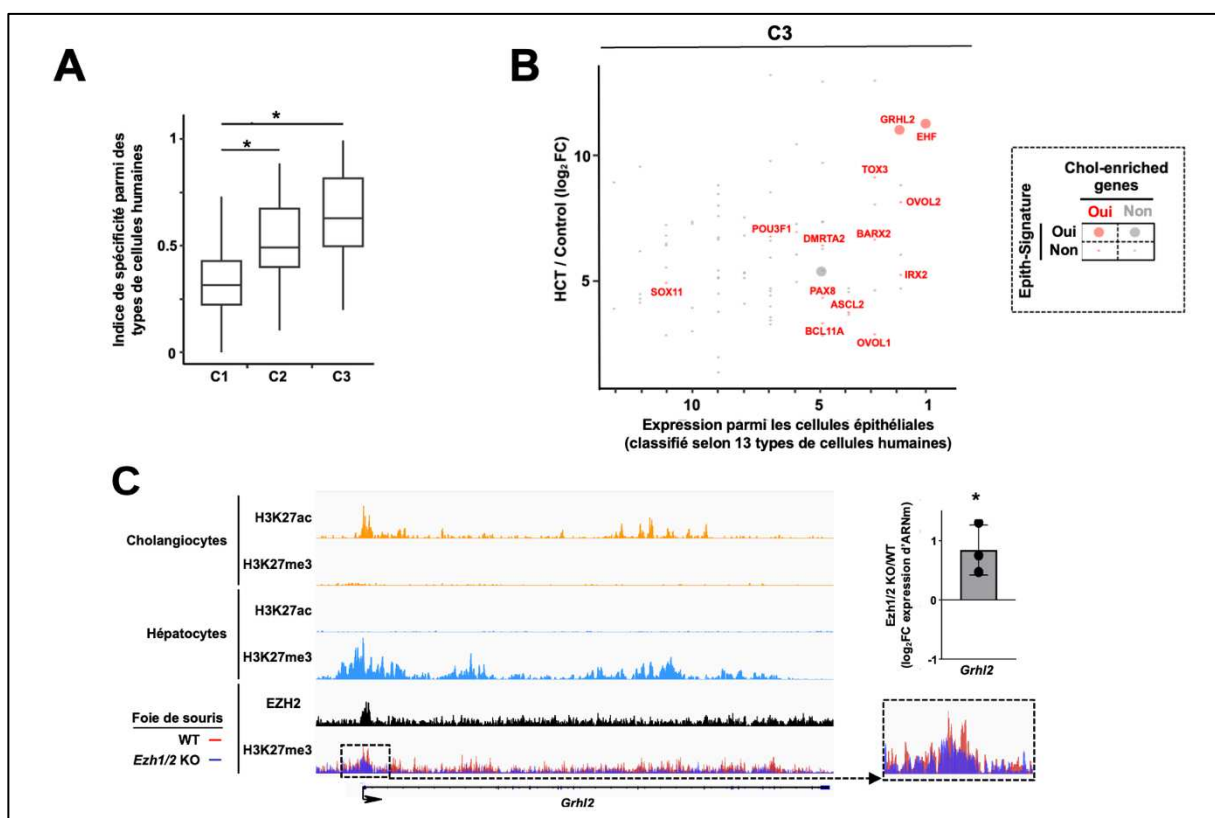


Figure 39 : L'HCT implique l'induction du facteur de transcription GRHL2, régulateur de la différenciation des cellules épithéliales.

(A) Boîte à moustache montrant l'indice de spécificité des gènes codants pour les TFs des groupes C1-3. Treize différents types cellulaires ont été utilisés pour le calcul (Cellules épithéliales et adipocytes, sang et cellules immunitaires, cellules endocrines, cellules endothéliales, cellules germinales, cellules gliales, cellules mésenchymateuses, cellules neuronales, cellules de la pigmentation, cellules trophoblastiques, cellules

indifférenciées). Ce calcul a été réalisé grâce à l'outil tspx. L'expression génique dans une lignée cellulaire donnée a été définie comme l'expression médiane des cellules individuelles associées à la lignée cellulaire. Les différences statistiques ont été évaluées à l'aide d'un test de Kruskal-Wallis suivi d'un test de comparaison multiple de Dunn comme test post-hoc. * q-value < 0,05. **(B)** Diagramme de dispersion des gènes codants pour les TFs du groupe C3 montrant leurs expressions dans les cellules épithéliales sur la base des données du panneau A (les gènes ont été classés par ordre décroissant d'expression) et leurs niveaux d'induction dans l'HCT. De plus, la présence ou non de ces gènes codants pour les TFs dans Epith-Signature a été mise en évidence avec de points plus gros, tandis que leurs présences dans la liste de gènes Chol-enriched a été mis en évidence à l'aide de points rouges. **(C)** L'outil « Integrate genomic viewer » (Thorvaldsdóttir et al., 2013) a été utilisé pour visualiser les marques de ChIP-Seq indiquées (H3K27me3 et H3K27ac) dans les hépatocytes ou les cholangiocytes au niveau du promoteur de *GRHL2*. Ces marques ont été traitées de manière similaires et visualisées avec la même échelle dans les 2 types de cellules. Les signaux de ChIP-Seq de H3K27me3 dans les foies de souris contrôles (WT) ou EZH1/2 KO ont été superposés. Un agrandissement est présenté à droite, accompagné d'une boîte à moustache montrant les changements géniques de *GRHL2* dans les foies de ces animaux calculés à partir des données de RNA-Seq de (Grindheim et al., 2019).

b. GRHL2 : TF actif dans les cholangiocytes et réprimé dans les hépatocytes

De plus, en lien avec nos résultats précédents, *GRHL2* est réprimé par la marque H3K27me3 et le complexe PrC2 dans des hépatocytes sains, tandis qu'il possède une marque d'activation de la transcription, H3K27ac, au niveau de son promoteur dans les cholangiocytes. Par ailleurs, dans les foies de souris EZH1/2 KO, la marque de répression H3K27me3 est moins présente que dans les foies de souris contrôles (« *wild-type* », WT). Cette perte de marque de répression au niveau de son promoteur est corrélée avec l'induction de l'expression de *Grhl2* (**Figure 39C**). Ainsi, ce TF clé de l'HCT possède toutes les caractéristiques d'un gène activé pendant ce processus.

Pour ces raisons, nous avons décidé de focaliser la suite de nos analyses sur l'étude du rôle du TF GRHL2 dans l'HCT. Cependant, avant de continuer la description des résultats, une description des différents rôles connus de ce facteur est nécessaire.

GRHL2 dans la littérature scientifique :

- Appartient à famille des facteurs de transcription de type Grainyhead, ayant un rôle dans la **morphogenèse et la différenciation cellulaire** (Kars & Muyan, 2018).
- Facteur clé du développement embryonnaire, en particulier, impliqué dans le **maintien et l'intégrité de l'organisation des cellules épithéliales** dont l'établissement des jonctions serrées (Aue et al., 2015) : sa suppression génique dans des embryons conduit à une désorganisation cellulaire et une perte de l'intégrité des cellules (Forrester-Gauntlett et al., 2023).
- Régule la **plasticité des cellules épithéliales** : des niveaux d'expressions important de *Grhl2*, dans des modèles murins de cancer résistants à la chimiothérapie, sont en lien avec une **TEM plus importantes** (W. Chen et al., 2016). Ce TF est associé à la régulation de la TEM, induisant l'expansion des cellules épithéliales (Nishino et al., 2017)
- Dualité d'action entre GRHL2 et le mirARN 122 observée en cas de fibrose hépatique où l'inhibition du mirARN et la fixation de GRHL2 exacerbent la fibrose et augmente la TEM (Y. Wang et al., 2020) (Chung et al., 2016)
- **Orchestre la différenciation cellulaire** : active la voie de signalisation TGFB et la phosphorylation de certains récepteurs de la voie MPAK/ERK, impliqués dans la transduction des signaux de détermination cellulaire (A. F. Chen et al., 2018) (Y. Wang et al., 2020).
- **Répression dans la mise en place des programmes cellulaires alternatifs**, notamment retrouvés dans les cellules cancéreuses (Jacobs et al., 2018) (A. F. Chen et al., 2018).
- Gènes régulés par GRHL2 : *Ovol1/2* (Watanabe et al., 2019) et *Ehf* (Sakamoto et al., 2021). Sa fixation induit l'expression de gènes comme la *claudine 4* (« *claudin 4* », *Cldn4*) ou le gène codant pour le *membre 25 de la famille des oncogène RAS* (« *member RAS oncogene family 25* », *Rab25*) (Senga et al., 2012).

3.2. L'expression ectopique de GRHL2 entraîne l'activation de la signature épithéliale et le début de l'HCT

a. Expression ectopique de *GRHL2* *in vitro* induit l'expression de gènes épithéliaux et la formation de cystes

i. Début de formation d'un canal dans les BMELs après l'induction de *GRHL2*

Une stratégie expérimentale de surexpression de *GRHL2* dans des cellules embryonnaires bipotentes de foie de souris (BMELs), ayant la capacité de se différencier en hépatocytes ou en cholangiocytes (Strick-Marchand & Weiss, 2002), a été mise en place. Le but de cette stratégie étant de vérifier que ce TF permet l'induction d'une signature génique rappelant celle activée lors de l'HCT. Une méthode de surexpression transitoire a été utilisée en mettant au contact les cellules avec un plasmide permettant l'expression de *GRHL2* ou un plasmide vide pendant 48h (**Figure 40A**). L'analyse d'enrichissement des voies biologiques, issue des données de RNA-Seq de cette expérience, montre que la surexpression de *GRHL2* dans les BMELs induit préférentiellement l'expression de gènes liés au destin cellulaire dont la différenciation des cellules épithéliales (**Figure 40B**). De même, les analyses GSEA réalisées prouvent que GRHL2 stimule les signatures géniques Epith-Signature et Chol-enriched (**Figure 40C**).

Pour aller plus loin, après transfection transitoire, les BMELs ont été cultivées dans le matrigel en trois dimensions. L'organisation cellulaire retrouvée pointe vers un début de la mise en place d'une organisation cellulaire autour d'un cyste dans les cellules transfectées avec le plasmide GRHL2 (**Figure 40D-E**). Ce type de structure correspond à un anneau cellulaire organisé autour d'un lumen. Ces structures sont, notamment, retrouvées dans les cholangiocytes.

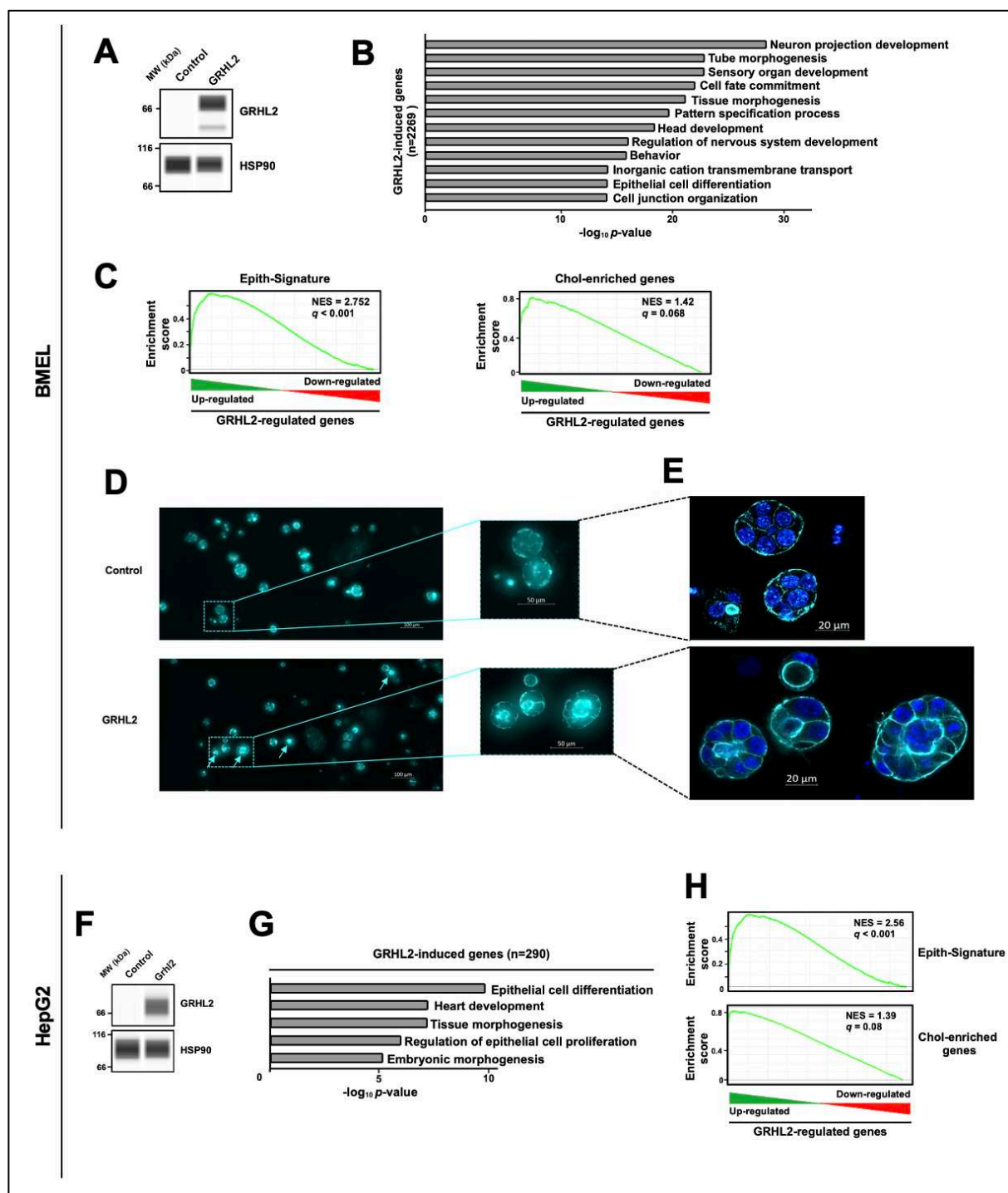


Figure 40 : La surexpression de *GRHL2* dans des hépatocytes déclenche l'induction de la signature épithéliale impliquée dans l'HCT.

(A) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de HSP90 ou *GRHL2* dans les BMELs transfectées avec un plasmide d'expression de *GRHL2* ou avec un plasmide vide pendant 48h. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. (B) Diagramme en barre montrant les différents termes d'enrichissement obtenus après analyse Metascape (Top12) des gènes induits après surexpression de *GRHL2* dans les BMELs ($\log_2 FC > 0$ et q -value $< 0,05$) définis après l'analyse RNA-Seq réalisée sur 4 réplicats biologiques indépendants. (C) Graphique d'enrichissement issu de l'analyse GSEA réalisée en utilisant comme liste de gènes soit Epith-Signature (Breshi *et al*, 2020) ou une liste de gènes de enrichis dans les cholangiocytes et les données les changements transcriptionnels présentés en (B). Le NES défini le score d'enrichissement. (D) BMELs transfectées avec un plasmide d'expression de *GRHL2* ou un plasmide vide comme contrôle durant 48h, puis, cultivées dans du matrigel. Les images sont représentatives des agrégats obtenus. Le marquage de la Phalloïdine est montré. Les flèches indiquent les cellules avec un cystes. (E)

Exemple de cellules avec le début de formation d'un cyste obtenu par imagerie confocale. Un traitement d'image LSM a été utilisé pour lisser les signaux. Pour l'image contrôle, un filtre médian a également été appliqué. Les marquages de la phalloïdine et du DAPI sont montrés. **(F)** Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de HSP90 ou GRHL2 dans les HepG2 transfectés avec un plasmide d'expression de *GRHL2* ou avec un plasmide vide pendant 48h. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. **(G)** Diagramme en barre montrant les différents termes d'enrichissement obtenus après analyse Metascape (Top 5) des gènes induits après surexpression de *GRHL2* dans les HepG2 ($\log_2FC > 0$ et $q\text{-value} < 0,05$) définis après l'analyse RNA-Seq réalisée sur 4 réplicats biologiques indépendants **(H)** Graphique d'enrichissement issu de l'analyse GSEA réalisée en utilisant comme liste de gènes soit Epith-Signature (Breshi *et al*, 2020) ou une liste de gènes enrichis dans les cholangiocytes et les données de changements transcriptionnels présentés en (G). Le NES définit le score d'enrichissement.

ii. Identification de cibles directes de GRHL2 induites pendant l'HCT

Pour définir les gènes directement régulés par GRHL2 en lien avec l'HCT, la transfection transitoire avec le plasmide contenant ce TF a été réalisée dans des cellules hépatocytaires, les HepG2 (**Figure 40F**). Même si la reprogrammation du transcriptome cellulaire est moins importante que dans les BMELs, les analyses d'enrichissement des voies biologiques, réalisées à partir des données de RNA-Seq, montrent que GRHL2 induit l'expression de gènes caractéristiques de l'archétype d'une cellule épithéliale (**Figure 40G**). Ainsi, les analyses GSEA pointent également vers l'activation des signatures géniques Epith-Signature et Chol-enriched (**Figure 40H**). Cependant, les résultats ne montrent pas l'inactivation des gènes codants pour les Hep-ID TFs.

Pour mieux caractériser la manière dont GRHL2 induit la reprogrammation transcriptionnelle des cellules, le cistrome de ce TF a été défini en utilisant des données de ChIP-Seq obtenues à partir de cellules HepG2 transfectées. Les 10 000 principaux sites de fixation de GRHL2 dans les HepG2 ont été comparés à ceux retrouvés dans des cellules exprimant de manière endogène ce TF, les cellules primaires de l'épithélium bronchiale (« *primary human bronchial epithelial* », HBEs) (X. Gao *et al.*, 2013) et les cellules issues d'une lignée cellulaire immortalisée de cancer ovarien de type A (« *ovarian epithelial cancer A cell-line* », OVCA429s) (Holding *et al.*, 2019). Même si l'expression forcée de *GRHL2* pourrait entraîner une liaison à la chromatine et une expression non spécifique, la comparaison de fixation avec ces autres types cellulaires indique que : l'expression transitoire de *GRHL2* dans les HepG2 conduit à la fixation de sa protéine sur des régions chromatiniennes similaires à la fixation de sa protéine endogène dans les autres modèles cellulaires (**Figure 41A**).

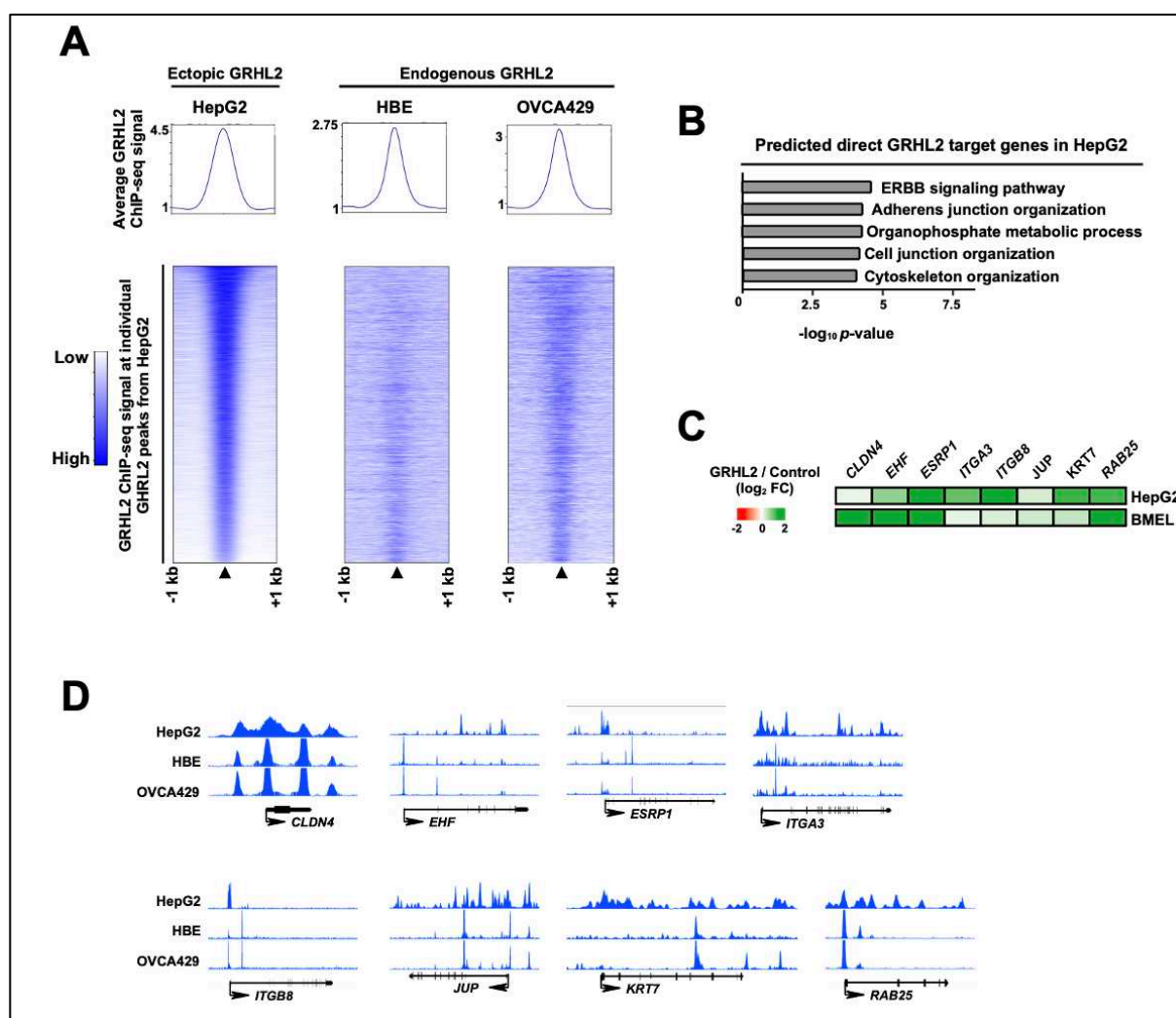


Figure 41 : Identification des gènes directement régulés par GRHL2.

(A) 10 000 sites principaux de fixation de GRHL2 identifiés dans les HepG2 ont été utilisés pour comparer le signal avec les lignées cellulaires indiquées, exprimant de manière endogène *GRHL2*. Les HBEs sont des cellules primaires de l'épithélium bronchial et les OVCA429 sont une lignée cellulaire de cancer ovarien de type A. Le signal individuel moyen (au-dessus, fixé arbitrairement à 1) et le signal au niveau des sites de liaison individuels (en bas, carte de chaleur) sont représentés dans une fenêtre de 2 kb centrée sur les pics que GRHL2. Les sites de liaisons individuels ont été classés en fonction du signal de ChIP-Seq de GRHL2. Ce dernier a été défini sur la base de 4 réplicats biologiques indépendants. (B) Diagramme en barre montrant l'analyse d'enrichissement des voies biologiques (Top 5) pour les gènes prédits comme étant directement régulés par GRHL2 dans les cellules HepG2 grâce à l'outil CistromGO (S. Li et al., 2019) sur la base des sites de liaison de GRHL2 identifiés. (C) Carte de chaleur montrant les changements d'expression des différents gènes indiqués à partir des données de RNA-Seq dans les BMELs ou les HepG2 après surexpression de notre facteur (mêmes données retrouvées dans [Figure 40](#)) (D) Visualisation grâce au navigateur intégré du génome, IGV, afin de visualiser les profils de ChIP-Seq des HepG2, HBE et OVCA429 au niveau des promoteurs des gènes cibles de GRHL2 (*CLDN4*, *EHF*, *ESRP1*, *ITGA3*, *ITGB8*, *JUP*, *KRT7* et *RAB25*).

De plus, l'utilisation de l'outil CistromGO (S. Li et al., 2019), associant les pics du signal GRHL2 avec les potentiels gènes cibles, confirme que les signaux de GRHL2 sont retrouvés au niveau de gènes dont les termes d'enrichissement des voies

biologiques sont compatibles avec une reprogrammation transcriptionnelle vers un programme de cellule épithéliale tel que l'adhésion cellulaire et la mise en place de jonctions (**Figure 41B**).

Pour identifier les cibles directes de GRHL2 pertinentes vis-à-vis du processus de l'HCT, les pics de signaux, issus du ChIP-Seq de GRHL2 au niveau des promoteurs des différents gènes, ont été sélectionnés sur la base de :

- L'activation transcriptionnelle médiée par ce TFs grâce à la présence d'un site de fixation ;
- Leurs inductions pendant l'HCT ;
- Leurs présences dans les listes de gènes Epith-Signature et Chol-enriched.

Ces analyses ont permis de pointer différents gènes régulés de manière direct par GRHL2, en plus des cibles déjà décrites comme *EHF*, *CLDN4* et *RAB25*, tels que les gènes codants pour la *protéine 1 régulatrice de l'épissage épithélial* (« *Epithelial splicing regulatory protein 1* », *ESRP1*), l'*intégrine sous unité alpha 3* (« *Integrin subunit alpha3* », *ITGA3*), l'*intégrine sous unité beta 8* (« *Integrin subunit beta 8* », *ITGB8*), la *jonction plakoglobine* (« *junction plakoglobin* », *JUP*) et la *cytokératine 7* (« *keratin 7* », *KRT7*). La littérature indique que tous ces gènes ont un rôle dans l'établissement et le maintien de l'identité des cellules épithéliales et la mise en place de l'épithélium (Senga et al., 2012) (Romagnoli et al., 2020) (Xia et al., 2023) (Buckley & St Johnston, 2022) (Karantza, 2011) (Warzecha et al., 2010).

L'action de ces gènes par GRHL2 est confirmée par leurs inductions dans les modèles *in vitro* de transfection transitoire avec le plasmide codant pour ce TFs (**Figure 41C-D**).

iii. Induction d'une signature génique cholangiocytaire

Des cellules progénitrices humaines (HepaRG) ont été cultivées, puis, enrichies en hépatocytes grâce à l'ajout de 0.2% de DMSO dans le milieu de culture, 15 jours après confluence des cellules. Après orientation des cellules, la transfection a été réalisée par la méthode d'électroporation (**Figure 42A**). Cette transfection transitoire valide l'induction des gènes cibles de GRHL2 (*CLDN4*, *ITGA3*, *JUP*, *RAB25*) dans ces cellules (**Figure 42B**).

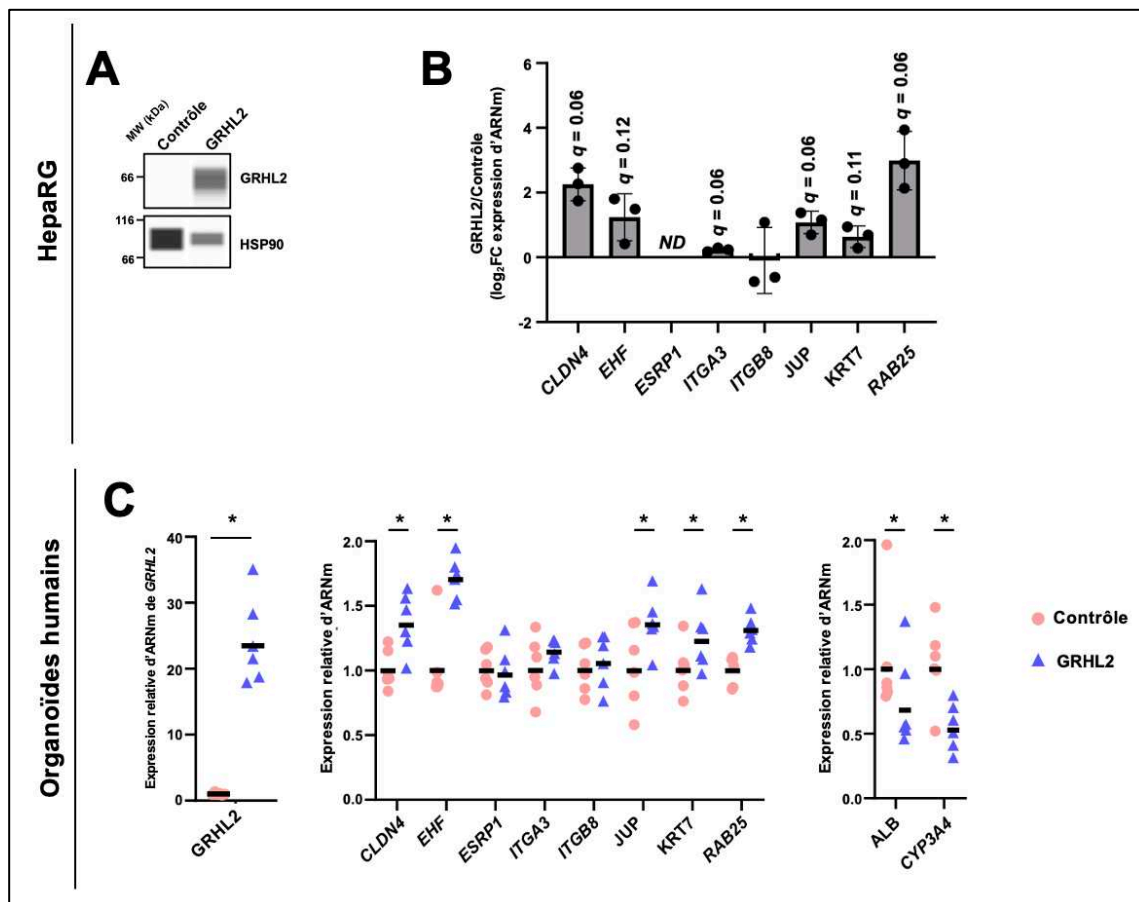


Figure 42 : Caractérisation des changements transcriptionnels induits par GRHL2 dans la lignée HepaRG et dans des organoïdes humains.

(A) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de HSP90 ou GRHL2 dans les HepaRG différenciées en hépatocytes, puis, électroporées pour réaliser la transfection avec un plasmide d'expression de *GRHL2* ou avec un plasmide vide pendant 48h. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. (B) Résultats des analyses de RT-qPCR montrant l'expression des gènes cibles de GRHL2 dans les HepaRG différenciées en hépatocytes, puis, électroporées pour réaliser la transfection avec un plasmide d'expression de *GRHL2* ou avec un plasmide vide pendant 48h (n=3). Les log₂FC d'expression génique dans les cellules exprimant *GRHL2* par rapport aux contrôles (transfectées avec un plasmide vide) sont montrés à l'aide d'un diagramme en barres qui affiche la moyenne \pm écart-type (SD) ainsi que les réplicats biologiques. Les différences statistiquement significatives, afin de déterminer si la médiane des log₂FC sont différentes de 0, ont été évaluées par un t-test bilatéral suivi d'une correction pour test multiple de Benjamini-Hochberg. * q < 0 ; 05. (C) Résultats des analyses de RT-qPCR montrant l'expression relative de *GRHL2* (droite),

de ses gènes cibles (milieu) et de gènes hépatocytes (droite) dans les organoïdes humains transduits avec *GRHL2* ou un AAV3 codant pour la GFP (n=6). Les données individuelles ainsi que la moyenne géométrique sont montrées. Les différences statistiques ont été évaluées à l'aide d'un test one-sided Mann-Witney U test suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$.

Enfin, l'expression de *GRHL2* a, également, été induite dans des organoïdes humains. Ces derniers ont été obtenus en cultivant des cellules primaires de foie humain, provenant d'une zone non tumorale après résection de métastase, triées EPCAM positives. Ces cellules ont été cultivées dans du Matrigel, avant l'induction de leur différenciation comme dans (Broutier et al., 2016). Le gène codant pour le TF *GRHL2* a été exprimé dans ces organoïdes à l'aide de virus de type AAV3. Les résultats valident l'induction des cibles directes de *GRHL2*, auxquels s'ajoute une diminution de certains marqueurs hépatocytaires comme *ALB* ou le gène codant pour le cytochrome *P450 3A4* (« Cytochrome *P450 Family 3 Subfamily A Member 4* », *CYP3A4*) (**Figure 42C**).

Ainsi, l'utilisation de différents modèles cellulaires a permis de montrer que *GRHL2* active l'expression de gènes épithéliaux. Ce dernier induit aussi une organisation cellulaire en cystes, retrouvés dans les cholangiocytes.

b. Expression ectopique de *GRHL2 in vivo* modifie la structure du parenchyme hépatique

Ensuite, nous avons voulu analyser l'impact de l'expression de *Grhl2 in vivo* ainsi que l'effet de son activation dans le foie.

i. Modulation de l'expression de *GRHL2 in vivo* par différentes méthodes

Dans un premier temps, l'expression de *Grhl2* a été induite dans les foies de souris en utilisant la méthode d'injection hydrodynamique. Pour cela, un plasmide codant pour *GRHL2* a été injecté avec une solution saline dans les queues de souris afin de permettre son absorption par le foie (**Figure 43A**). Même si l'expression ectopique de *GRHL2* n'est pas optimale (15 à 20% d'hépatocytes maximum captant le plasmide), les gènes cibles du TFs ont été activés (**Figure 43B**).

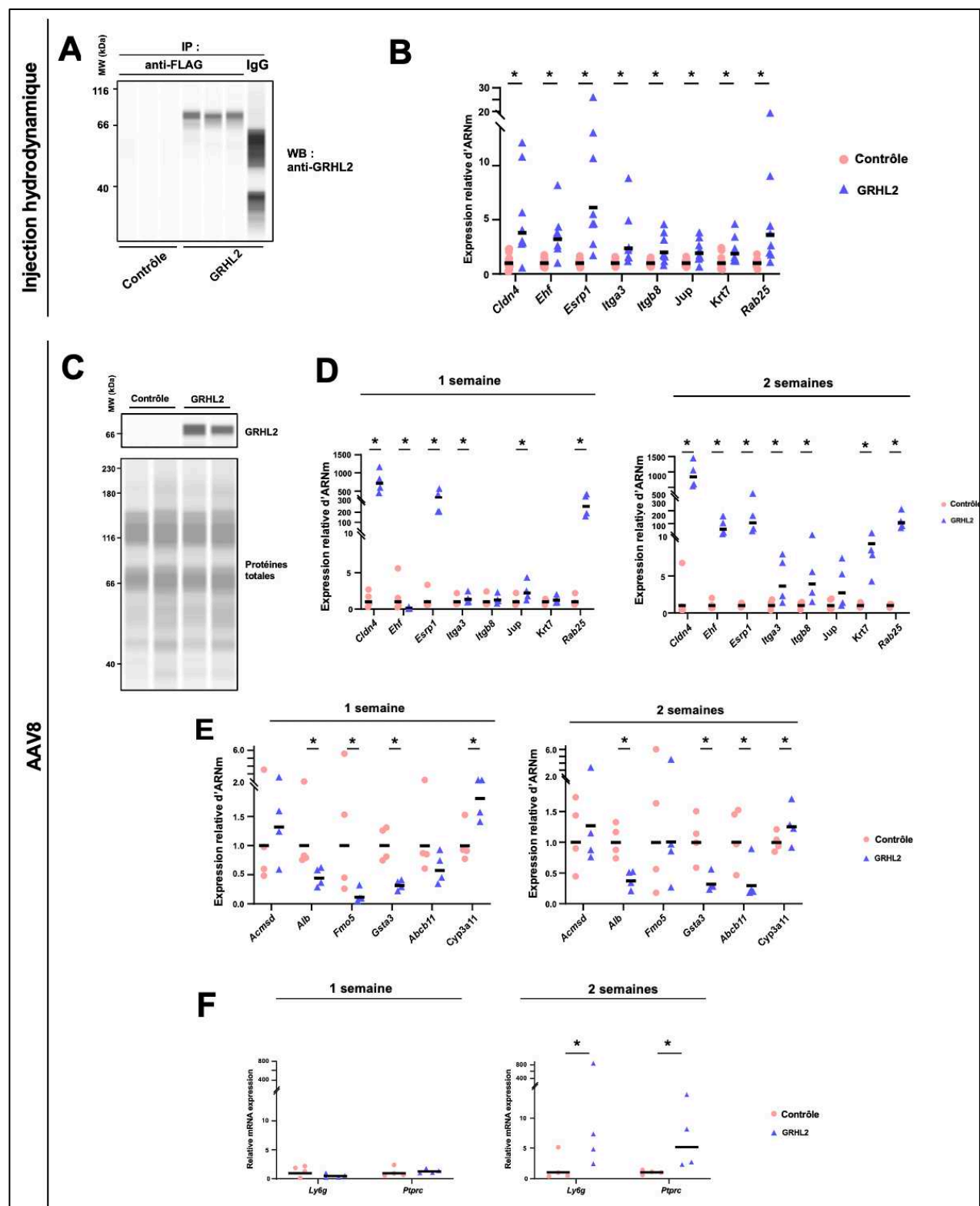


Figure 43 : Stratégie de surexpression de *GRHL2* *in vivo*.

(A) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de GRHL2 dans les foies de souris ayant été sujettes à une injection hydrodynamique avec le plasmide codant pour *Flag-GRHL2* ou un plasmide contrôle. Les extraits nucléaires de chaque foie de souris (3 par groupes) ont été immunoprécipités avec un anticorps anti-flag, puis révélés à l'aide d'un anticorps anti-GRHL2. MW : poids moléculaire. (B) Résultats des

analyses de RT-qPCR montrant l'expression relative des gènes indiqués dans les foies de souris sujettes à une injection hydrodynamique avec le plasmide codant pour *Flag-GRHL2* (n=8) ou un plasmide contrôle (n=9). Les données individuelles ainsi que la moyenne géométrique sont montrées. Les différences statistiques ont été évaluées à l'aide d'un test one-sided Mann-Witney U test suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$. **(C)** Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de GRHL2 dans les foies de souris injectées avec *GRHL2* ou *GFP* (contrôle) à l'aide d'AAV8. Les extraits nucléaires ont été préparés après une semaine d'injection et révélés à l'aide d'un anticorps anti-GRHL2. Un extrait total des protéines est utilisé comme contrôle de charge. **(D)** Résultats des analyses de RT-qPCR montrant l'expression relative des gènes indiqués dans les foies de souris 1 semaine ou 2 semaines après l'injection des virus AAV8 *GRHL2* (n=4) ou *GFP* (n=4). Les données individuelles ainsi que la moyenne géométrique sont montrées. Les différences statistiques ont été évaluées à l'aide d'un test one-sided Mann-Witney U test suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$. **(E)** Résultats des analyses de RT-qPCR montrant l'expression relative des gènes hépatocytaires dans les foies de souris 1 semaine ou 2 semaines après l'injection du virus AAV8 *GRHL2* (n=4) ou *GFP* (n=4). Les données ont subi la même analyse que dans le panneau (D). **(F)** Résultats des analyses de RT-qPCR montrant l'expression relative de marqueurs inflammatoires dans les foies de souris 1 semaine ou 2 semaines après du virus AAV8 *GRHL2* (n=4) ou *GFP* (n=4). Les données ont suivi la même analyse que dans le panneau (D).

Afin d'atteindre une portion plus importante d'hépatocytes, une autre méthode de surexpression de *GRHL2 in vivo* a été utilisée. Les AAV8 codants pour *GRHL2* ou la *GFP* ont été injectés en rétro-orbitale. Une semaine après l'injection, GRHL2 est bien présent dans les foies de souris et un début d'induction de ses cibles est retrouvé (**Figure 43C**). Cependant, deux semaines après l'induction, l'expression des gènes épithéliaux est plus forte, notamment pour *Cldn4*, *Ehf*, *Esrp1*, *Itga3*, *Itgb8*, *Jup*, *Krt7* et *Rab25* (**Figure 43D**). Au contraire, l'expression des marqueurs hépatocytaires est diminuée deux semaines après l'injection. On retrouve, par exemple, la diminution de l'expression du gène codant pour l'*amino-carboxy-muconate semi-aldéhyde décarboxylase* (« aminocarboxymuconate semialdehyde decarboxylase », *Acmsd*), l'*Alb*, la *flavine contenant la monooxygénase 5* (« flavin containing monooxygenase 5 », *Fmo5*), ou encore le *glutathion S-transférase alpha 3* (« glutathione S-transferase alpha 3 », *Gsta3*) (**Figure 43E**). De la même manière, on retrouve une diminution du gène codant pour le *transporteur hépatique d'acide biliaire* (« bile salt export pump », *Bsep* ou *Abcb11*). Au contraire, l'expression de *Cyp3a11* (variant murin de *Cyp3a4*) est augmentée, caractérisant une altération du métabolisme hépatique des acides biliaires (**Figure 43E**).

ii. Modification du parenchyme hépatique 2 semaines après l'expression ectopique de *GRHL2* dans les foies de souris

En lien avec les changements transcriptionnels dans les souris ayant reçu une injection virale de *GRHL2*, nous avons regardé les structures du parenchyme hépatique. Après une semaine, le signal *GRHL2* est homogène dans les hépatocytes, ces derniers montrent un phénotype classique avec une organisation cellulaire en travée hépatocyttaire (**Figure 44A-B**). Au contraire, deux semaines après l'injection, les hépatocytes sont hétérogènes avec une structure et une taille différente. On retrouve une altération du parenchyme hépatique (**Figure 44B-C**). Cette caractéristique structurelle, couplée à la présence de cellules infiltrantes, fait penser à des événements de DR qui sont le reflet de la souffrance cellulaire (**Figure 44C** et **Figure 43F**). La présence de cette infiltration est également retrouvée au niveau génique avec une augmentation des gènes codants pour l'*antigène lymphocytaire 6G* (« *Lymphocyte antigen 6 member G* », *Ly6g*) et le *récepteur de la protéine tyrosine phosphatase de type C* (« *Protein tyrosine phosphate receptor type C* », *Ptpcr*) (**Figure 43F**). De plus, dans certaines zones, les cellules présentent de nombreux événements de mitose, indiquant la prolifération hépatique en réponse aux dommages cellulaires (**Figure 44C**).

Ainsi, les résultats indiquent que *GRHL2* favorise l'HCT par l'activation des gènes caractéristiques d'un programme épithélial.

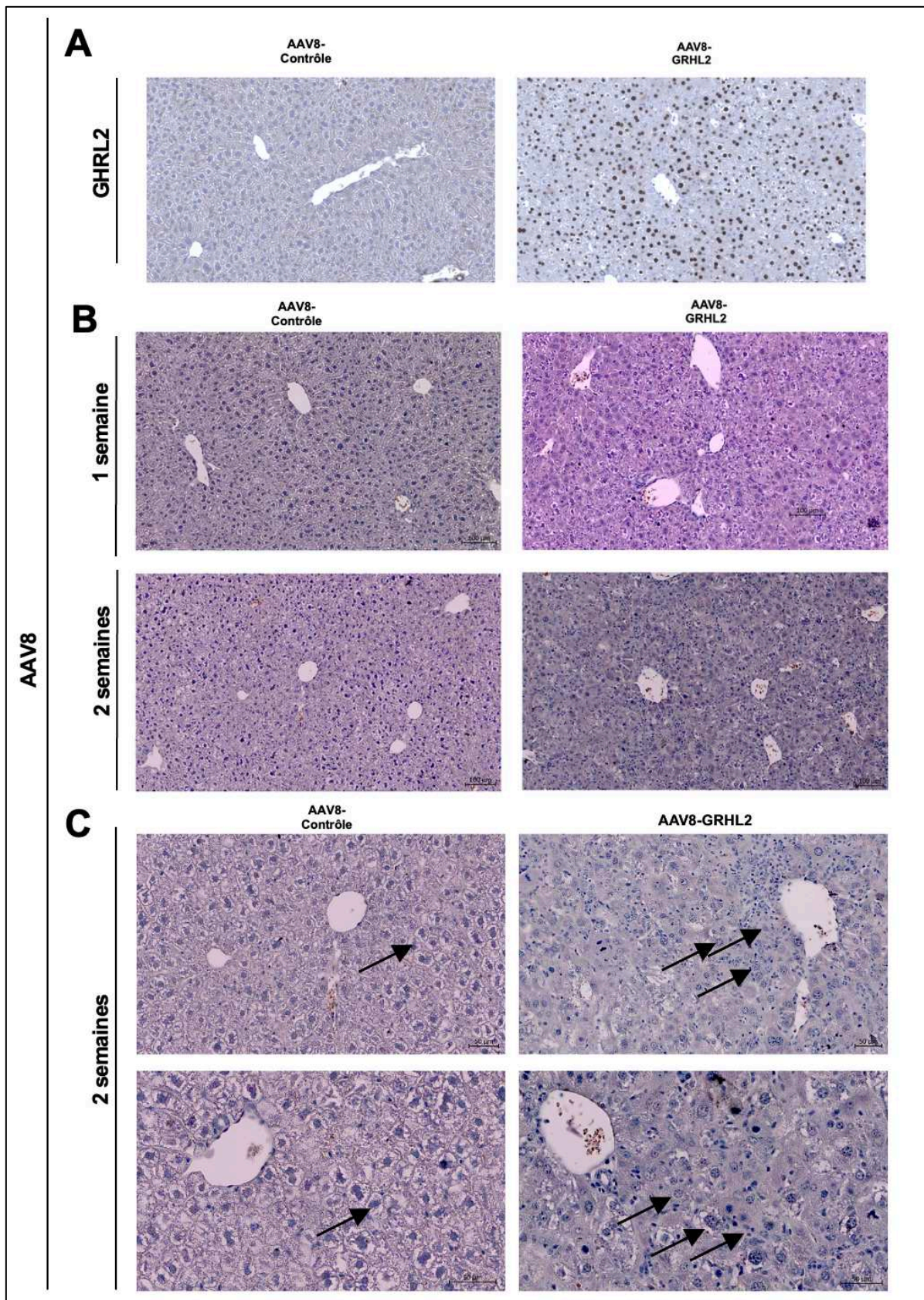


Figure 44 : La surexpression de *GRHL2* dans les foies de souris modifie la structure du parenchyme hépatique.

(A) Immunomarquage de GRHL2 dans les foies de souris une semaine après une injection de *GRHL2* ou de *GFP* (contrôle) de virus AAV8. (B) Marquage des foies de souris 1 ou 2 semaines après injection des virus AAV8-*GFP*

(contrôle) ou AVV8-*GRHL2* à l'hématoxyline et éosine. **(C)** Différentes zones issues du marquage hématoxyline et éosine sur les foies de souris 2 semaines après injection virale. Les flèches pointent vers une organisation caractéristique des hépatocytes dans les foies de souris contrôles en opposition aux formes hétérogènes observées dans les foies de souris surexprimant *GRHL2*.

3.3. Pertinence physio-pathologique de l'étude de *GRHL2* chez des patients montrant des signes de HCT

De récentes études montrent que certains foies de patients atteints de maladies hépatiques chroniques en stade terminale sont caractérisés par la présence d'hépatocytes avec des signes d'HCT, c'est-à-dire, des hépatocytes perdant l'expression des marqueurs de l'identité hépatocytaire et exprimant des marqueurs cholangiocytaires (Gribben et al., 2024) (Andrews et al., 2024).

i. Induction de *GRHL2* dans les foies de patients atteints d'ALD

Dans un premier temps, des échantillons issus de patients atteints d'une cirrhose décompensée liée à l'ALD ont été utilisés afin d'évaluer l'expression du gène codant pour *GRHL2* et son lien avec l'HCT. Dans cette partie, des fragments d'échantillons de foies dit « contrôles » proviennent de patients ayant subi une résection hépatique en raison de tumeurs métastatiques hépatiques. Les analyses sur ces échantillons montrent que *GRHL2* est induit dans les échantillons de patients atteints d'ALD au niveau de l'ARNm et de la protéine (**Figure 45A-B**). Cette activation est en lien avec une diminution des niveaux d'H3K27me3 et de EZH2 dans ces foies (**Figure 45B**).

Pour vérifier que *GRHL2* est induit de manière spécifique dans les hépatocytes, nous avons ensuite analysé des données de RNA-Seq obtenues à partir de zones microdisséquées enrichies en hépatocytes (Bou Saleh et al., 2021). Ces résultats confirment que l'augmentation de l'expression de *GRHL2* est corrélée à l'induction de l'Epith-Signature et des différents groupes définis précédemment comme induits lors de l'HCT (**Figure 45C**).

De plus, l'immunomarquage des coupes de patients atteints d'ALD montre que, contrairement aux échantillons contrôles où *GRHL2* est retrouvé uniquement dans les cholangiocytes, le marquage se situe dans les néoductules et dans les hépatocytes

proches des nodules de régénération au niveau des zones fibrotiques (**Figure 45D**). Ces résultats sont validés par le co-marquages avec ALB (**Figure 46A**).

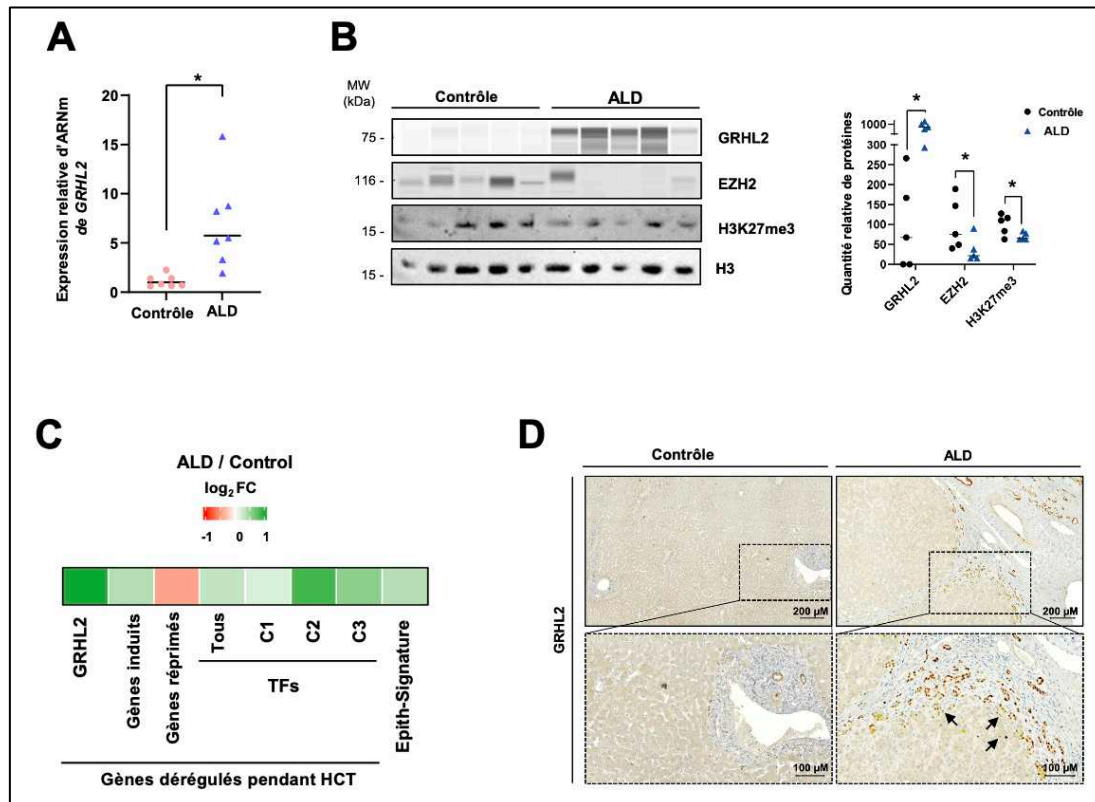


Figure 45 : Les hépatocytes de patients atteints de maladies hépatiques chroniques en phase terminale de type ALD expriment *GRHL2* et les gènes épithéliaux associés à l'HCT.

(A) Résultats de l'analyse RT-qPCR montrant l'expression relative d'ARNm de *GRHL2* dans les échantillons de patients contrôles ou atteints d'ALD (n=7). Les données individuelles ainsi que la moyenne géométrique sont montrées. Les différences statistiques ont été évaluées à l'aide d'un test one-sided Mann-Witney U test suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$. (B) Analyse protéomique réalisée en utilisant le système de WES ou une méthode de Western-Blot pour regarder les niveaux de *GRHL2*, *EZH2*, *H3K27me3* et *H3* dans les échantillons de patients contrôles ou atteints d'ALD, provenant d'un fractionnement de la chromatine. Les différences statistiques ont été évaluées à l'aide d'un test one-sided Mann-Witney U test suivi d'une correction pour test multiple de Benjamini-Hochberg. * $q < 0,05$. MW : poids moléculaire. (C) Carte de chaleur issue des données de RNA-Seq utilisant des zones enrichies en hépatocytes microdisséquées provenant de zones contrôles et de patients atteints ALD (Bou Saleh et al., 2021) (n=5 et n=7 pour respectivement les échantillons contrôles et ALD). Les données montrent les différences d'expression des différentes signatures géniques et de *GRHL2* entre ces 2 conditions (la médiane du \log_2FC est indiquée). (D) Immunomarquage de *GRHL2* dans les échantillons contrôles et les échantillons de foie de patients atteints d'ALD. Les résultats sont représentatifs des marquages réalisés sur 5 patients différents. Les images agrandies de la zone délimitée sont montrées en bas. Les flèches pointent les hépatocytes *GRHL2*-positifs.

De plus, le triple marquage de ces échantillons avec ALB, KRT7 et *GRHL2* indique la présence de cellules triples-positives dans les zones de cicatrisation de la fibrose, correspondant à la mise en place de l'HCT dans ce micro-environnement spécifique des foies cirrhotiques (**Figure 46B**) (Hammond et al., 2025).

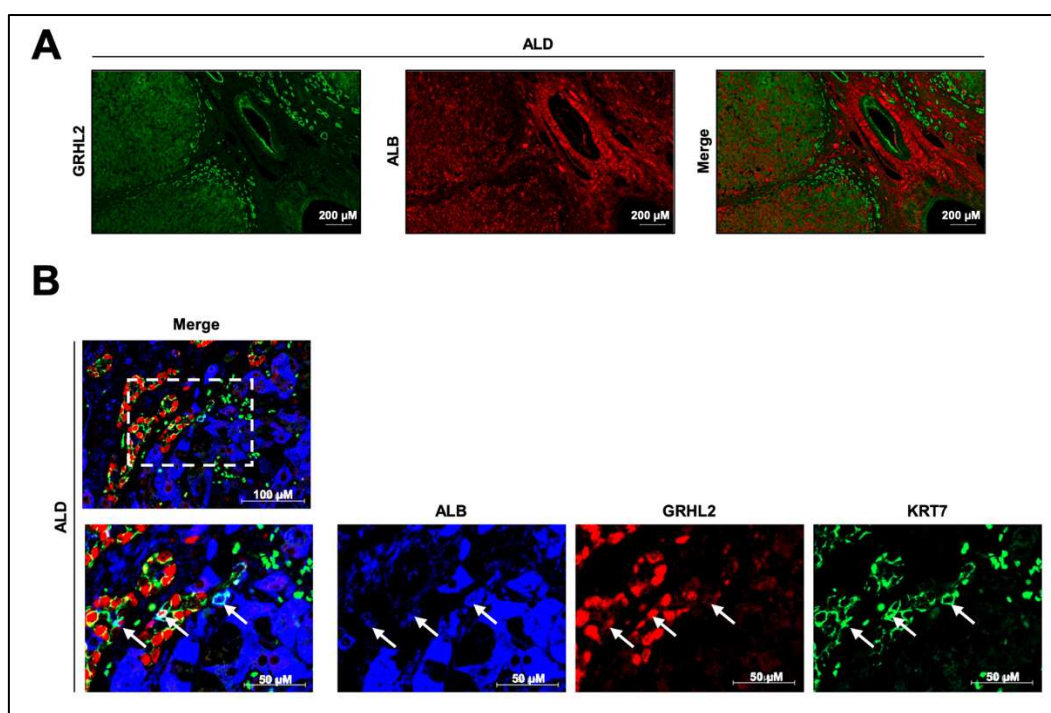


Figure 46 : GRHL2 présent dans des hépatocytes en cours d'HCT chez les patients atteints d'ALD.

(A) Co-immunomarquage de GRHL2 (en vert) et d'un marqueur hépatocytaire (ALB, en rouge) dans les foies de patients atteints d'insuffisance hépatique liée à l'ALD, réalisé sur la série de coupe présentée en **Figure 45D**. (B) Triple immunomarquage de GRHL2 (en rouge), le marqueur hépatocytaire ALB (en bleu) et d'un marqueur cholangiocytaire (en vert) réalisé sur la série de coupe présentée en **Figure 45D**.

ii. GRHL2 dans les MASLDs

Par ailleurs, l'utilisation de données de séquençage des noyaux individuels des cellules (« *single-nuclei RNA-Seq* », snRNA-Seq), issus de (Gribben et al., 2024) indique que *GRHL2* est exprimé dans les hépatocytes provenant de patients de stades avancés (**Figure 47A-B**). En effet, le calcul des rapports appelés « *Odds ratio* » entre les hépatocytes contrôles et les hépatocytes issus de différents stades de la maladie montre une plus forte prévalence de cellules GRHL2-positives dans les stades avancés de la pathologie (**Figure 47C**). Ces cellules positives expriment, également, les gènes cibles de GRHL2 validant la mise en place de l'HCT et de la signature épithéliale dans les hépatocytes des stades avancés de MASLDs (**Figure 47D**).

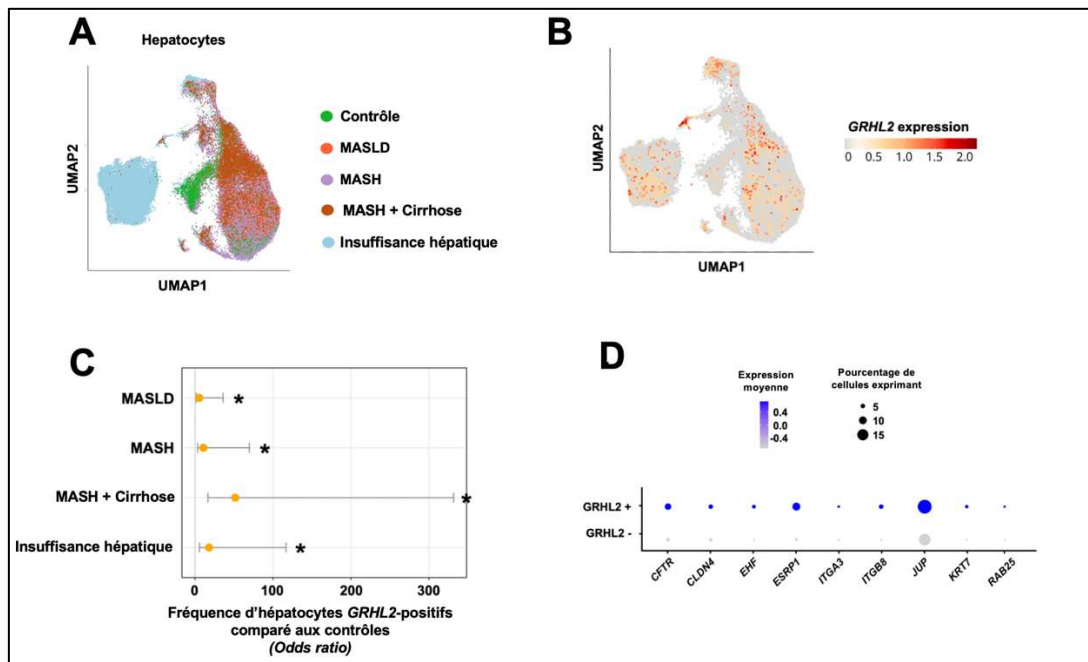


Figure 47 : Hépatocytes exprimant *GRHL2* chez des patients atteints de MASLDs grâce aux données de snRNA-Seq.

(A-B) Données de (Gribben et al., 2024) utilisées pour regarder l'expression de *GRHL2* dans des hépatocytes issus de foies de patients atteints de MASLDs. (A) indique les différents stades de la maladie dont proviennent les hépatocytes tandis que (B) montre l'expression de *GRHL2* dans les noyaux des cellules. (C) La fraction d'hépatocytes exprimant *GRHL2* à chaque stade de la MASLD a été comparé aux hépatocytes sains par le calcul d'un rapport de côtes. Le diagramme en barre montre les valeurs par rapport aux côtes (point orange) ainsi que les intervalles de confiance à 95%. Les différences statistiques ont été évaluées à l'aide de tests bilatéraux de Fisher, puis, une correction de tests multiples de Benjamini-Hochberg. (D) Graphique en bulle indiquant l'expression des gènes indiqués dans les hépatocytes *GRHL2*-positifs ou *GRHL2*-négatifs (c'est-à-dire tous les hépatocytes du panneau A).

iii. Activation de *GRHL2* dans les cholangiopathies

Enfin, la présence de *GRHL2* dans des pathologies du foie émanantes de cholangiocytes (PBC et PSC) a été vérifiée à l'aide d'immunomarquage. Les résultats montrent que, comme pour les pathologies hépatiques émergentes des hépatocytes, *GRHL2* est retrouvé dans les zones de fibrose, autour des nodules de régénération (Figure 48).

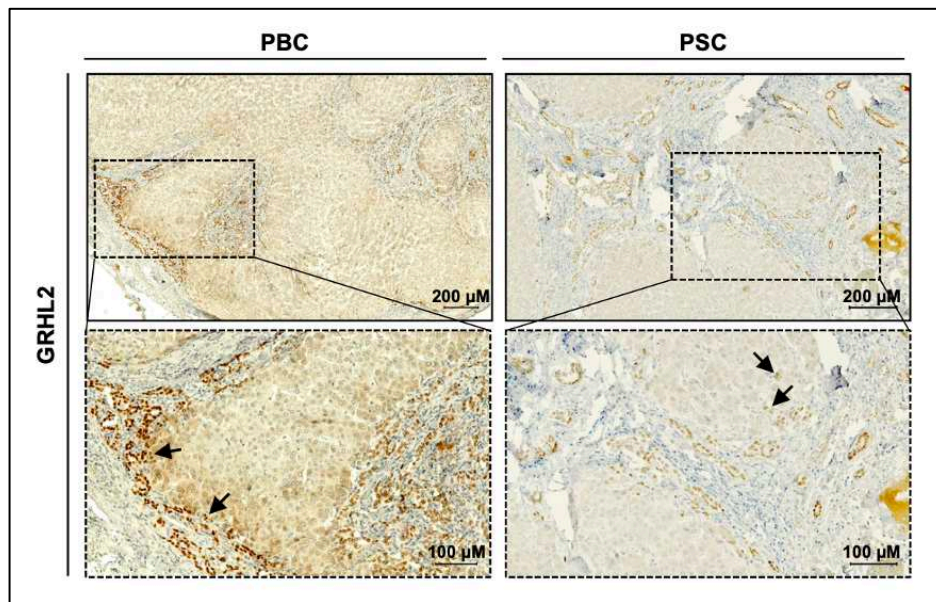


Figure 48 : Expression de *GRHL2* dans des foies de patients atteints de cholangiopathies.

Immunomarquage de *GRHL2* dans les foies de patients atteints de PSC et de PBC. Les résultats sont représentatifs des marquages réalisés sur 3 séries de patients. Les images agrandies de la zone délimitée sont montrées en bas. Les flèches pointent les hépatocytes *GRHL2*-positifs.

Ces résultats pointent vers une induction de *GRHL2* dans les pathologies hépatiques à des stades avancés montrant des signes d'HCT, peu importe l'origine de la pathologie (hépatocytaire ou cholangiocytaire).

Conclusion générale

Cette étude s'est intéressée à la caractérisation des événements transcriptionnels ayant lieu pendant la reprogrammation hépatocytaire. En effet, la compréhension de la plasticité des hépatocytes est importante pour intégrer les mécanismes de régénération hépatique, notamment dans les maladies touchant les voies biliaires (PSC, PBC, ALGS). Mais, ce processus est également décrit comme une des caractéristiques des hépatocytes dans des pathologies de type ALD ou MASLD à des stades avancés. Ce travail éclaire sur le phénomène de transdifférenciation pertinent dans les contextes physiopathologiques et jusqu'ici mal caractérisés.

Nos données apportent une valence moléculaire à la différence phénotypique fondamentale distinguant les hépatocytes et les cholangiocytes, seuls ces derniers étant caractérisés par un phénotype archétypal de cellules épithéliales. Ainsi, la reprogrammation hépatocytaire s'accompagne de l'activation d'un programme transcriptionnel épithélial, normalement réprimé dans les hépatocytes sains. La répression de ce programme dans un hépatocyte à l'état basal est maintenue par la présence de la marque épigénétique H3K27me3 grâce à l'activité des enzymes EZH1 et EZH2.

En particulier, ce projet a mis en évidence le rôle clé du TF GRHL2 dans les mécanismes de l'HCT, induisant de manière directe l'expression de la signature épithéliale dans les hépatocytes en cours de reprogrammation. En effet, ce TF a été retrouvé dans des échantillons de patients atteints de diverses pathologies hépatiques (**Figure 49**).

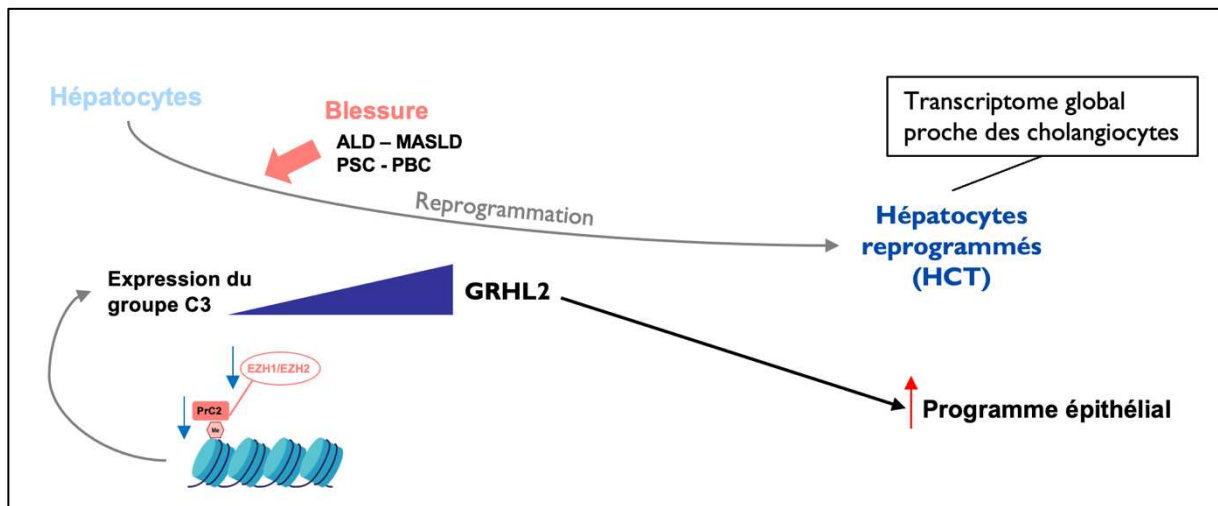


Figure 49 : Schéma bilan des événements transcriptionnels ayant lieu pendant l'HCT.

Pendant le processus d'HCT, la levée de la répression de H3K27me3 au niveau du promoteur de certains gènes permet leur activation. En particulier, nous avons identifié l'induction du gène codant pour le facteur de transcription *GRHL2*. Ce TF permet l'activation d'un programme épithélial dans les hépatocytes, absent dans un hépatocyte sain mais qui est présent dans les cholangiocytes.

Discussion

I. Impact de l'étude de l'HCT pour les maladies hépatiques

L'HCT permet au foie de renouveler les cholangiocytes lorsque les mécanismes de régénération sont compromis. Ainsi, la compréhension des mécanismes moléculaires ayant lieu lors de l'HCT est important pour la médecine régénérative, notamment les pathologies de type ALGS, PBC et PSC. Cependant, en plus des cholangiopathies, l'HCT est décrit dans une partie des hépatocytes issus de stades avancés des maladies de types ALD et MASLD. Malheureusement, dans ce contexte, on ne sait pas encore si le processus de l'HCT est un signe de régénération ou un signe d'accélération de l'insuffisance hépatique.

1.1. Rôle du micro-environnement dans le mécanisme de l'HCT

Un des points importants qui n'a pas été évoqué dans ce manuscrit est le rôle du micro-environnement dans le développement de l'HCT. Durant ce processus, les hépatocytes perdent certaines de leurs caractéristiques telles que la perte des jonctions intercellulaires et leurs interactions avec la MEC. Ces modifications activent la voie de signalisation Hippo/YAP, déjà décrite comme une des voies permettant la mise en place du programme cholangiocytaire (Ahmad et al., 2022) (Bou Saleh et al., 2021). En parallèle, l'altération de la polarisation des hépatocytes, au cours de la dédifférenciation, modifie les signaux de destin cellulaire. De plus, les cytokines telles que le TGFB jouent un rôle dans l'induction du processus, notamment dans les maladies hépatiques liées à la consommation d'alcool et pourraient également contribuer à l'HCT (Hammond et al., 2025) (Lan et al., 2023). Enfin, lors du processus de l'HCT, en lien avec la souffrance cellulaire, on peut supposer que les HSCs vont s'activer et libérer de l'acide stérolique afin de former les gouttelettes lipidiques, ce qui pourrait également favoriser le processus de l'HCT (Mak et al., 2023).

Ces données montrent que ce processus s'accompagne d'un réseau complexe de différents signaux. Ainsi, définir comment ces signaux interviennent dans la mise en place du mécanisme de l'HCT et comment ils sont liés à l'activation des différents

TFs dont GRHL2 permettrait d'avoir une compréhension plus fine de la reprogrammation.

1.2. Opposition de nos résultats avec le dogme de l'action des LPCs en cas de blessure hépatique

De nombreuses études antérieures posent l'hypothèse que les hépatocytes acquièrent un transcriptome de cellule embryonnaire en cas de blessure hépatique (He et al., 2022). Cette hypothèse repose sur la coexpression de marqueurs spécifiques hépatocytaires et cholangiocytaires dans certains hépatocytes, exprimant notamment le TF SOX9 (W. Li et al., 2019) (Tarlow et al., 2014). Cependant, dans les données de (L. Li et al., 2023), seuls un sous-ensemble de gènes induits dans les cellules dites « progénitrices » montrent des niveaux d'expression plus importants dans des hépatoblastes comparés aux hépatocytes. De plus, l'étude de (L. Li et al., 2023) montre que l'activation des LPCs passe par l'activation d'enhancers différents que ceux activés pendant la différenciation des hépatoblastes vers les cholangiocytes. Ainsi, les mécanismes de régulation transcriptionnelle impliqués dans l'HCT semblent être différents de ceux transitant par l'activation des LPCs.

Enfin, bien que le traçage de la lignée hépatocyttaire chez l'Homme ne soit pas possible, les données de scRNA-Seq de foies de patients atteints de MASLD, publiées dans (Gribben et al., 2024), mettent en évidence l'absence de cellules présentant un transcriptome proche d'une cellule embryonnaire. Ces résultats sont donc concordants avec notre étude montrant que l'expression de *GRHL2* est suffisante pour induire les gènes impliqués dans l'HCT dans les hépatocytes matures. Ensemble, nos données et celles de (Gribben et al., 2024) établissent que l'HCT est un mécanisme direct de reprogrammation des cellules matures qui ne passe pas par un retour à un état cellulaire embryonnaire. Ces résultats sont en accord avec les mécanismes de transdifférenciation décrits dans d'autres types cellulaires comme ceux mis en évidence dans des études sur la reprogrammation des fibroblastes cardiaques (Ieda et al., 2010)

1.3. Modification du réseau de TFs pendant l'HCT

Comme expliqué, l'action combinée des TFs permet de maintenir un état transcriptionnel propre à chaque cellule. Lors de l'HCT, on retrouve une perte du réseau des HepID, entraînant la perte de l'identité hépatocytaire (Dubois-Chevalier et al., 2023). Notre étude a montré que GRHL2 joue un rôle clé pendant l'HCT. Cependant, le rôle de ce TF est probablement englobé dans le changement du réseau de TFs pendant ce mécanisme, permettant *in fine* l'activation de l'identité cholangiocytaire. En effet, notre étude confirme l'induction du TF SOX4, en plus de celle de GRHL2, pendant l'HCT, validant une modification dans le réseau de TF (Katsuda et al., 2024). L'expression de ces deux TFs contribue aux changements transcriptionnels ayant lieu dans les hépatocytes pendant l'HCT. Malheureusement, la surexpression combinée de SOX4 et de GRHL2 dans nos modèles cellulaires ne montre pas de modifications dans l'induction des gènes épithéliaux.

En plus des changements transcriptionnels, l'HCT peut s'accompagner, sur le long terme, de la formation *De Novo* de canaux biliaires. Or, la mise en place de ces canaux est en lien avec l'évacuation et la modification de la bile, impliquant l'expression de gènes notamment du *régulateur de la conductance transmembranaire de la formation de la fibrose cystique* (« *cystic fibrosis transmembrane conductance regulator* », CFTR). Cependant, dans nos modèles, CFTR n'est pas induit par GRHL2, montrant que les voies de signalisation impliquées dans la fonctionnalité des cholangiocytes ne sont pas induites dans nos modèles.

1.4. Mécanismes épigénétiques potentiels sous-jacents à l'HCT

Notre étude montre que l'induction de la signature épithéliale s'accompagne d'une modification de l'épigénome dans les hépatocytes, notamment une levée de la marque de répression H3K27me3. Cependant, il semblerait que H3K27me3 ne soit pas la seule marque épigénétique impliquée dans ce processus. En effet, des études récentes ont pointé le rôle des enzymes de liaison aux récepteurs nucléaires SET à domaine protéique 1 (« *Nuclear receptor binding SET Domain Protein 1* », NSD1) et de la lysine déméthylase 2A (« *Lysine demethylase 2A* », KDM2A), catalysant l'apposition de la marque H3K36me3 dans l'HCT (Sussman et al., 2025). Cette étude

n'indique pas le rôle des enzymes EZH1 et EZH2 et ne montre aucun lien entre la levée de H3K27me3 et l'HCT. La différence avec nos résultats peut être expliquée par le fait que les gènes codants pour *EZH1* et *EZH2* ont été éteints séparément. Ces deux enzymes étant redondantes dans les hépatocytes, la suppression combinée de ces dernières est nécessaire pour déclencher la levée de la répression de H3K27me3 (Bae et al., 2015).

En outre, notre étude montre que la diminution des niveaux de H3K27me3 au niveau des promoteurs des gènes comme *GRHL2* permet l'activation d'un réseau de gènes épithéliaux. Néanmoins, les mécanismes menant à cette levée de la marque de répression ne sont pas établis. Plusieurs hypothèses pourraient être envisagées comme une dégradation des protéines du complexe PrC2, ou alors, une activation des déméthylases, causée soit par le contexte inflammatoire, soit par un remodelage de la chromatine. Dans les deux cas, une inhibition transitoire de l'activité du complexe PrC2 active l'expression des gènes codants pour des TFs, délétères pour l'activité hépatique (Holoch et al., 2021). Il est, donc, nécessaire de mieux comprendre les événements impliquants le remodelage épigénétique et leurs lien fonctionnel avec la modification d'activité des TFs pendant le mécanisme de l'HCT.

II. Stabilisation de EZH2 par l'O-GlcNAcylation aux promoteurs des gènes épithéliaux dans les hépatocytes sains

Une autre hypothèse en lien avec la perte de H3K27me3 dans les hépatocytes pourrait être une perte de la stabilité des enzymes EZH1 ^{et/ou} EZH2. Bien que peu documenté, le recrutement et l'activité enzymatique du PrC2 peuvent être **modulés par différentes modifications post-traductionnelles, notamment au niveau de EZH2**. En effet, EZH2 est la cible de différentes modifications comme la O-GlcNAcylation (Chu et al., 2014) qui permet de guider cette protéine vers les promoteurs cibles et d'augmenter sa stabilité. De plus, des études sur la différenciation des hépatoblastes montrent que la perte de stabilité de EZH2 favorise la différenciation des cholangiocytes (L. Yang et al., 2023).

Afin de vérifier si cette PTM est présente dans notre modèle cellulaire, des expériences de « Click-Chemistry » ont été réalisées. Les résultats montrent que EZH2 est bien O-GlcNAcylé dans les HepG2 (**Figure 50A**). Or, l'apposition de cette PTM est régulée par la O-GlcNAc transférase (OGT). Ainsi, pour établir le lien entre EZH2 et la O-GlcNAcylation, les cellules ont été traitées avec un petit ARN interférant (siARN) ciblant l'OGT. Cette expérience a montré que la diminution de l'activité de l'OGT, diminue les niveaux d'O-GlcNAc et de EZH2 (**Figure 50B**). La stabilité d'EZH2 a été testée en utilisant la cyclohexamine, afin de bloquer la synthèse protéique. De cette manière, le temps de demi-vie de la protéine EZH2 a été évalué en présence ou non d'un inhibiteur de l'OGT. Les résultats ont montré que le temps de demi-vie de la protéine est plus faible sans cette PTM. Cette dernière est donc stabilisée par la O-GlcNAcylation (**Figure 50C**).

Ensemble, ces résultats montrent que dans notre modèle cellulaire et dans les hépatocytes, EZH2 est stabilisé par la présence de l'O-GlcNAc qui permettrait le maintien de la marque de répression H3K27me3 aux promoteurs des gènes épithéliaux.

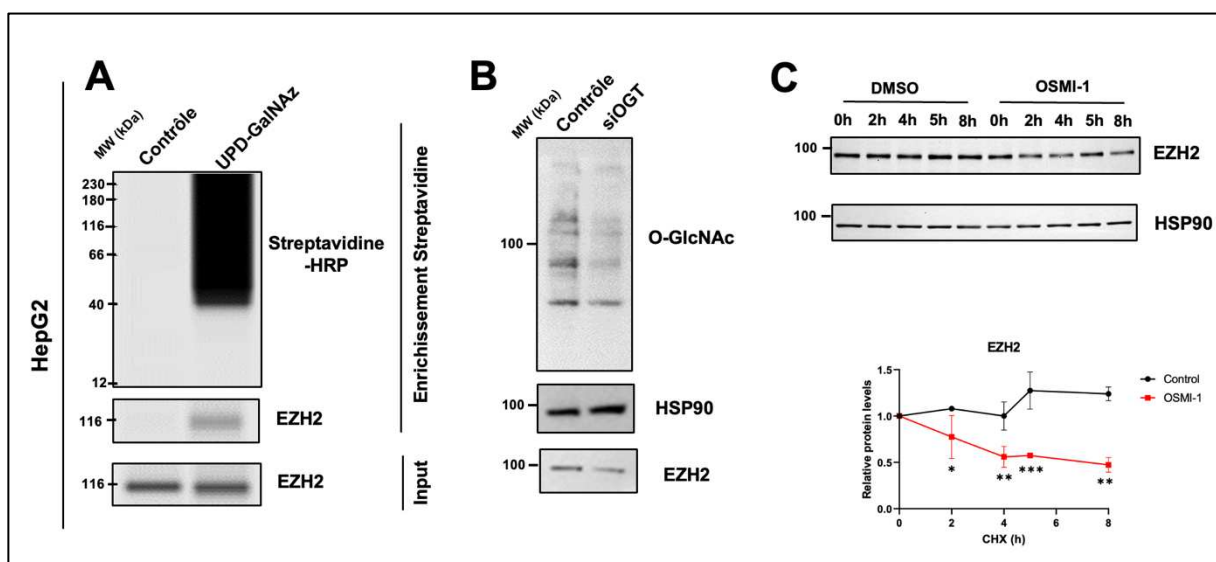


Figure 50 : Stabilisation de EZH2 par son l'O-GlcNAcylation.

(A) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de EZH2 et de Streptavidine dans les HepG2 traitées avec du Thiamet-G pendant 24h puis ayant été marquées par la GlcNaz afin d'évaluer les protéines O-GlcNAcylées. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. (B) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de EZH2, HSP90 et du O-GlcNAc dans les HepG2 infectées avec un siOGT ou un siARN vide. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. (C) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de EZH2, et HSP90 dans les HepG2 traitées ou non avec de l'OSMI-1 pendant 24h puis avec la cyclohexamine pendant les temps indiqués pour

bloquer la synthèse protéique. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. Les différences statistiques ont été évaluées à l'aide d'une ANOVA bidirectionnelle, suivi d'un test post-hoc de Sidak.

De plus, en croisant les gènes codants pour les TFs induits dans l'HCT et ceux montrant une perte de H3K27me3 dans les foies de souris OGT KO, issues de (Ortega-Prieto et al., 2023), on retrouve *Grhl2*. On peut, donc, supposer que l'activation de l'expression de ce gène est permise par une perte de la stabilité de EZH2, via une diminution de l'activité de l'OGT.

Au final, il est possible que pendant le processus de l'HCT, la perte de la O-GlcNAcylation diminue la stabilité de EZH2 dans les hépatocytes, permettant la diminution des niveaux de H3K27me3 et l'activation des gènes épithéliaux, mais sa relevance physio-pathologique n'est pas encore bien déterminée. Actuellement, l'équipe cherche à définir si ce changement d'O-GlcNAcylation accompagne l'HCT dans des contextes physiopathologiques.

III. Mécanisme d'action de GRHL2 lors de l'HCT

3.1. Action des GRHL en homo-, di- ou en trimère

Même si ma thèse s'est focalisée sur le rôle de GRHL2, il est important de noter que ce TF fait partie de la famille des GRHL, un groupe de TFs impliqués dans la formation de structure épithéliale. Or, l'activité de ces protéines est possible sous la forme d'homo-, di- ou trimères (Miles et al., 2017).

Dans le modèle murin DDC, l'expression de *Grhl1* et *Grhl3* est légèrement plus importante dans les hépatocytes reprogrammés (respectivement $\log_2FC = 1,85$ et $q\text{-value} < 0,001$ & $\log_2FC = 0,09$ et $q\text{-value} = 0,78$). De plus, dans notre méthode de regroupement basé sur les signaux épigénétiques, ces deux TFs sont retrouvés respectivement dans le cluster C2 et C1. Cette classification se traduit une expression des gènes codants pour ces TFs dans les hépatocytes sains ($\log_2FPKM(Grhl1) = 4,35$; $\log_2FPKM(Grhl3) = 1,34$). Ainsi, à l'opposé de GRHL2, les gènes codants pour les TFs de GRHL1 et GRHL3 sont actifs ou du moins compétents dans les hépatocytes. Ainsi, on peut supposer que lors de l'HCT, GRHL1 et GRHL3 pourraient permettre de

modifier le statut épigénétique au niveau du promoteur de *GRHL2* pour permettre sa transcription, comme démontré dans (Masalha et al., 2021). Puis, dans un second temps, l'action combinée de ces TFs pourrait permettre de stabiliser la fixation de GRHL2 au niveau des promoteurs de ses gènes cibles.

Pour vérifier cette hypothèse, il serait, donc, intéressant de regarder dans nos modèles d'études, les partenaires d'interactions de GRHL2 par des analyses de co-immunoprécipitation. Cette méthode pourrait permettre de comprendre si l'activité de GRHL2 est permise par une action en homo-, di-ou en trimère avec les autres protéines de la famille.

3.2. Partenaires d'interactions de GRHL2

Enfin, un des points à élucider serait de comprendre comment les gènes cibles de GRHL2 sont activés. Est-ce que la fixation de GRHL2 permet de modifier le contexte épigénétique au niveau de leurs promoteurs et de diminuer la présence de H3K27me3, ou à l'inverse, est-ce lors de l'HCT la modification de l'activité de EZH2 diminue les niveaux de H3K27me3 pour faciliter la fixation de GRHL2 ?

Dans ce sens, regarder le co-recrutement des histones déméthylases et de GRHL2 dans un modèle murin de surexpression de *GRHL2* pourrait être intéressant. En effet, quelques études rapportent que la fixation de TFs de la famille des GRHL, en particulier GRHL2, permet le recrutement de certaines déméthylases de H3K27me3 comme la lysine déméthylase 6B (« *lysine demethylase 6B* », KDM6B) qui peut diminuer l'activité du complexe PrC2 et donc diminuer les niveaux de H3K27me3 (W. Chen et al., 2012). Cependant, dans les modèles *in vivo* d'étude, KDM6B ne semble pas être fortement induit : les FCs sont faibles et non significatifs dans les modèles d'expression ectopique de *GRHL2* dans les HepG2 et les BMELs. Restant des modèles *in vitro*, tous les signaux ne sont pas présents, ce qui pourrait expliquer ce résultat.

3.3. Les différentes isoformes de GRHL2

De plus, chez l'Homme il existe deux isoformes de *GRHL2* issues d'un épissage alternatif : un isoforme courte tronquée de 16 acides aminés et l'isoforme longue que nous avons utilisée dans ce projet de thèse. Or, il est décrit que ces deux isoformes pourraient avoir des rôles différents, notamment en cas de maladie hépatique en lien avec l'alcool où l'isoforme courte de GRHL2 serait de manière plus importante au niveau protéique, mais aussi génique (Satishchandran et al., 2018, p. 122). Or, dans nos échantillons de patients atteints d'ALD, il semblerait que GRHL2 ne soit détecté qu'à une seule taille, celle correspondante à l'isoforme longue (**Figure 45B**).

L'isoforme courte de *GRHL2* a été surexprimée dans des HepG2 afin de regarder si elle peut permettre l'activation des gènes épithéliaux. Les résultats indiquent que les deux isoformes de GRHL2 peuvent activer les gènes épithéliaux comme *RAB25*, *CLDN4* et *OVOL2* (**Figure 51**). Toutefois, cette expérience ne permet pas de conclure sur l'efficacité d'activation des gènes épithéliaux par les différentes isoformes de GRHL2. Des analyses supplémentaires seraient nécessaires afin de préciser le rôle de ces isoformes dans le contexte de l'HCT.

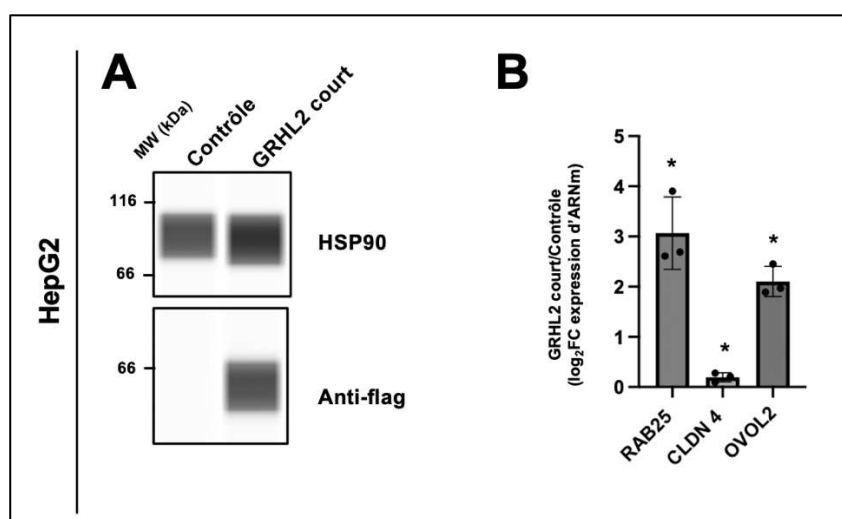


Figure 51 : Expression ectopique de l'isoforme courte de GRHL2.

(A) Analyse protéique réalisée en utilisant le système de WES pour regarder les niveaux de HSP90 ou d'anti-flag dans les HepG2 transfectées avec un plasmide contrôle (vide) ou un plasmide exprimant l'isoforme courte de *GRHL2* pendant 48h. Ces données sont représentatives des résultats de 3 réplicats biologiques indépendants. MW : poids moléculaire. **(B)** Résultats des analyses de RT-qPCR montrant l'expression des gènes cibles de *GRHL2* dans les HepG2 transfectés avec un plasmide contrôle (vide) ou un plasmide exprimant l'isoforme court de *GRHL2* pendant 48h (n=3). Les log₂FC d'expression génique dans les cellules exprimant *GRHL2* par rapport aux contrôles (transfectées avec un plasmide vide) sont montrés à l'aide d'un diagramme en barre qui affiche la moyenne ± écart-type (SD) ainsi que les réplicats biologiques. Les différences statistiquement significatives, afin de déterminer si la médiane des log₂FC est différente de 0, ont été évaluées par un t-test bilatéral suivi d'une correction pour test multiple de Benjamini-Hochberg. * $p < 0.05$.

Références

- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., & Blankenberg, D. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, 46(W1), W537–W544. <https://doi.org/10.1093/nar/gky379>
- Ahmad, U. S., Uttagomol, J., & Wan, H. (2022). The Regulation of the Hippo Pathway by Intercellular Junction Proteins. *Life*, 12(11), 1792. <https://doi.org/10.3390/life12111792>
- Aloia, L., McKie, M. A., Vernaz, G., Cordero-Espinoza, L., Aleksieva, N., van den Ameele, J., Antonica, F., Font-Cunill, B., Raven, A., Aiese Cigliano, R., Belenguer, G., Mort, R. L., Brand, A. H., Zernicka-Goetz, M., Forbes, S. J., Miska, E. A., & Huch, M. (2019). Epigenetic remodelling licences adult cholangiocytes for organoid formation and liver regeneration. *Nature Cell Biology*, 21(11), Article 11. <https://doi.org/10.1038/s41556-019-0402-6>
- Alvaro, D., Gigliozi, A., & Attili, A. F. (2000). Regulation and deregulation of cholangiocyte proliferation. *Journal of Hepatology*, 33(2), 333–340. [https://doi.org/10.1016/S0168-8278\(00\)80377-3](https://doi.org/10.1016/S0168-8278(00)80377-3)
- Anders, S., Pyl, P. T., & Huber, W. (2015). HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. <https://doi.org/10.1093/bioinformatics/btu638>
- Andersson, R. (2015). Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model. *BioEssays*, 37(3), 314–323. <https://doi.org/10.1002/bies.201400162>
- Andreatta, M., Hérault, L., Gueguen, P., Gfeller, D., Berenstein, A. J., & Carmona, S. J. (2024). Semi-supervised integration of single-cell transcriptomics data. *Nature Communications*, 15(1), 872. <https://doi.org/10.1038/s41467-024-45240-z>
- Andrews, T. S., Nakib, D., Perciani, C. T., Ma, X. Z., Liu, L., Winter, E., Camat, D., Chung, S. W., Lumanto, P., Manuel, J., Mangroo, S., Hansen, B., Arpinder, B., Thoeni,

- C., Sayed, B., Feld, J., Gehring, A., Gulamhusein, A., Hirschfield, G. M., ... MacParland, S. (2024). Single-cell, single-nucleus, and spatial transcriptomics characterization of the immunological landscape in the healthy and PSC human liver. *Journal of Hepatology*, 80(5), 730–743. <https://doi.org/10.1016/j.jhep.2023.12.023>
- Antoniou, A., Raynaud, P., Cordi, S., Zong, Y., Tronche, F., Stanger, B. Z., Jacquemin, P., Pierreux, C. E., Clotman, F., & Lemaigre, F. P. (2009). Intrahepatic Bile Ducts Develop According to a New Mode of Tubulogenesis Regulated by the Transcription Factor SOX9. *Gastroenterology*, 136(7), 2325–2333. <https://doi.org/10.1053/j.gastro.2009.02.051>
- Apte, U., & Krishnamurthy, P. (2011). Detoxification Functions of the Liver. In S. P. S. Monga (Ed.), *Molecular Pathology of Liver Diseases* (Vol. 5, pp. 147–163). Springer US. https://doi.org/10.1007/978-1-4419-7107-4_11
- Arab, J. P., Díaz, L. A., Rehm, J., Im, G., Arrese, M., Kamath, P. S., Lucey, M. R., Mellinger, J., Thiele, M., Thursz, M., Bataller, R., Burton, R., Chokshi, S., Francque, S. M., Krag, A., Lackner, C., Lee, B. P., Liangpunsakul, S., MacClain, C., ... Mathurin, P. (2025). Metabolic dysfunction and alcohol-related liver disease (MetALD): Position statement by an expert panel on alcohol-related liver disease. *Journal of Hepatology*, 82(4), 744–756. <https://doi.org/10.1016/j.jhep.2024.11.028>
- Arechederra, M., Berasain, C., Avila, M. A., & Fernández-Barrena, M. G. (2020). Chromatin dynamics during liver regeneration. *Seminars in Cell & Developmental Biology*, 97, 38–46. <https://doi.org/10.1016/j.semcdb.2019.03.004>
- Ashraf, U. M., Sanchez, E. R., & Kumarasamy, S. (2019). COUP-TFII Revisited: Its Role in Metabolic Gene Regulation. *Steroids*, 141, 63–69. <https://doi.org/10.1016/j.steroids.2018.11.013>
- Ayoub, M. D., Bakhsh, A. A., Vandriel, S. M., Keitel, V., & Kamath, B. M. (2023). Management of adults with Alagille syndrome. *Hepatology International*, 17(5), 1098–1112. <https://doi.org/10.1007/s12072-023-10578-x>
- Ayoub, M. D., & Kamath, B. M. (2020). Alagille Syndrome: Diagnostic Challenges and Advances in Management. *Diagnostics*, 10(11), Article 11. <https://doi.org/10.3390/diagnostics10110907>

- Bae, W. K., Kang, K., Yu, J. H., Yoo, K. H., Factor, V. M., Kaji, K., Matter, M., Thorgeirsson, S., & Helmighausen, L. (2015). The methyltransferases enhancer of zeste homolog (EZH) 1 and EZH2 control hepatocyte homeostasis and regeneration. *The FASEB Journal*, 29(5), 1653–1662. <https://doi.org/10.1096/fj.14-261537>
- Bai, S., Zou, Y., Zhang, W., Yu, J., Qie, Z., Liu, Z., Yu, P., Xiang, C., & Teng, Y. (2025). Novel Aurone Derivative Ameliorates MASH Lipid Metabolism via the AMPK-ACC-PPAR α Axis. *International Journal of Molecular Sciences*, 26(22), 11099. <https://doi.org/10.3390/ijms262211099>
- Banales, J. M., Huebert, R. C., Karlsen, T., Strazzabosco, M., LaRusso, N. F., & Gores, G. J. (2019). Cholangiocyte pathobiology. *Nature Reviews Gastroenterology & Hepatology*, 16(5), 269–281. <https://doi.org/10.1038/s41575-019-0125-y>
- Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., & Natoli, G. (2014). Coregulation of Transcription Factor Binding and Nucleosome Occupancy through DNA Features of Mammalian Enhancers. *Molecular Cell*, 54(5), 844–857. <https://doi.org/10.1016/j.molcel.2014.04.006>
- Bazett-Jones, D. P., Côté, J., Landel, C. C., Peterson, C. L., & Workman, J. L. (1999). The SWI/SNF Complex Creates Loop Domains in DNA and Polynucleosome Arrays and Can Disrupt DNA-Histone Contacts within These Domains. *Molecular and Cellular Biology*, 19(2), 1470–1478. <https://doi.org/10.1128/MCB.19.2.1470>
- Bemer, M., & Grossniklaus, U. (2012). Dynamic regulation of *Polycomb* group activity during plant development. *Current Opinion in Plant Biology*, 15(5), 523–529. <https://doi.org/10.1016/j.pbi.2012.09.006>
- Bendixen, S. M., Jakobsgaard, P. R., Hansen, D., Hejn, K. H., Terkelsen, M. K., Bjerre, F. A., Thulesen, A. P., Eriksen, N. G., Hallenborg, P., Geng, Y., Dam, T. V., Larsen, F. T., Wernberg, C. W., Vijayathurai, J., Scott, E. A. H., Marcher, A.-B., Detlefsen, S., Grøntved, L., Dimke, H., ... Ravnskjaer, K. (2024). Single cell-resolved study of advanced murine MASH reveals a homeostatic pericyte signaling module. *Journal of Hepatology*, 80(3), 467–481. <https://doi.org/10.1016/j.jhep.2023.11.001>
- Bergen, V., Lange, M., Peidli, S., Wolf, F. A., & Theis, F. J. (2020). Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*, 38(12), 1408–1414. <https://doi.org/10.1038/s41587-020-0591-3>

- Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M. A., Beaudet, A. L., Ecker, J. R., Farnham, P. J., Hirst, M., Lander, E. S., Mikkelsen, T. S., & Thomson, J. A. (2010). The NIH Roadmap Epigenomics Mapping Consortium. *Nature Biotechnology*, 28(10), 1045–1048. <https://doi.org/10.1038/nbt1010-1045>
- Bird, T. G., Lu, W.-Y., Boulter, L., Gordon-Keylock, S., Ridgway, R. A., Williams, M. J., Taube, J., Thomas, J. A., Wojtacha, D., Gambardella, A., Sansom, O. J., Iredale, J. P., & Forbes, S. J. (2013). Bone marrow injection stimulates hepatic ductular reactions in the absence of injury via macrophage-mediated TWEAK signaling. *Proceedings of the National Academy of Sciences*, 110(16), 6542–6547. <https://doi.org/10.1073/pnas.1302168110>
- Bonnet, L., Alexandersson, I., Baboota, R. K., Kroon, T., Oscarsson, J., Smith, U., & Boucher, J. (2022). Cellular senescence in hepatocytes contributes to metabolic disturbances in NASH. *Frontiers in Endocrinology*, 13. <https://doi.org/10.3389/fendo.2022.957616>
- Boonstra, K., Beuers, U., & Ponsioen, C. Y. (2012). Epidemiology of primary sclerosing cholangitis and primary biliary cirrhosis: A systematic review. *Journal of Hepatology*, 56(5), 1181–1188. <https://doi.org/10.1016/j.jhep.2011.10.025>
- Bort, R., & Zaret, K. S. (2009). Embryonic Development of the Liver. In I. M. Arias (Ed.), *The Liver* (1st ed., pp. 17–25). Wiley. <https://doi.org/10.1002/9780470747919.ch2>
- Bou Saleh, M., Louvet, A., Ntandja-Wandji, L. C., Boleslawski, E., Gnemmi, V., Lassailly, G., Truant, S., Maggiorotto, F., Ningarhari, M., Artru, F., Anglo, E., Sancho-Bru, P., Corlu, A., Argemi, J., Dubois-Chevalier, J., Dharancy, S., Eeckhoutte, J., Bataller, R., Mathurin, P., & Dubuquoy, L. (2021). Loss of hepatocyte identity following aberrant YAP activation: A key mechanism in alcoholic hepatitis. *Journal of Hepatology*, 75(4), 912–923. <https://doi.org/10.1016/j.jhep.2021.05.041>
- Bowman, G. D., & Poirier, M. G. (2015). Post-Translational Modifications of Histones That Influence Nucleosome Dynamics. *Chemical Reviews*, 115(6), 2274–2295. <https://doi.org/10.1021/cr500350x>
- Braet, F., & Wisse, E. (2002). Structural and functional aspects of liver sinusoidal endothelial cell fenestrae: A review. *Comparative Hepatology*, 1(1), 1. <https://doi.org/10.1186/1476-5926-1-1>

- Breschi, A., Muñoz-Aguirre, M., Wucher, V., Davis, C. A., Garrido-Martín, D., Djebali, S., Gillis, J., Pervouchine, D. D., Vlasova, A., Dobin, A., Zaleski, C., Drenkow, J., Danyko, C., Scavelli, A., Reverter, F., Snyder, M. P., Gingeras, T. R., & Guigó, R. (2020). A limited set of transcriptional programs define major cell types. *Genome Research*, 30(7), 1047–1059. <https://doi.org/10.1101/gr.263186.120>
- Bricks, T. (2014). *Développement d'un dispositif microfluidique ayant pour objectif l'étude des effets de premiers passages intestinaux et hépatiques* [Phdthesis, Université de Technologie de Compiègne]. <https://theses.hal.science/tel-01235621>
- Broutier, L., Andersson-Rolf, A., Hindley, C. J., Boj, S. F., Clevers, H., Koo, B.-K., & Huch, M. (2016). Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nature Protocols*, 11(9), 1724–1743. <https://doi.org/10.1038/nprot.2016.097>
- Brown, L. D., Cai, T. T., & DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science*, 16(2), 101–133. <https://doi.org/10.1214/ss/1009213286>
- Buckley, C. E., & St Johnston, D. (2022). Apical–basal polarity and the control of epithelial form and function. *Nature Reviews Molecular Cell Biology*, 23(8), 559–577. <https://doi.org/10.1038/s41580-022-00465-y>
- Caldez, M. J., Van Hul, N., Koh, H. W. L., Teo, X. Q., Fan, J. J., Tan, P. Y., Dewhurst, M. R., Too, P. G., Talib, S. Z. A., Chiang, B. E., Stünkel, W., Yu, H., Lee, P., Fuhrer, T., Choi, H., Björklund, M., & Kaldis, P. (2018). Metabolic Remodeling during Liver Regeneration. *Developmental Cell*, 47(4), 425–438.e5. <https://doi.org/10.1016/j.devcel.2018.09.020>
- Camargo, A. P., Vasconcelos, A. A., Fiamenghi, M. B., Pereira, G. A. G., & Carazzolle, M. F. (2024). *tspex: A tissue-specificity calculator for gene expression data*. <https://doi.org/10.21203/rs.3.rs-51998/v1>
- Campana, L., Esser, H., Huch, M., & Forbes, S. (2021). Liver regeneration and inflammation: From fundamental science to clinical applications. *Nature Reviews Molecular Cell Biology*, 22(9), 608–624. <https://doi.org/10.1038/s41580-021-00373-7>

- Canivet, C. M., & Faure, S. (2024). Diagnostic et évaluation de l'hépatopathie stéatosique métabolique. *La Revue de Médecine Interne*, 45(1), 41–47. <https://doi.org/10.1016/j.revmed.2023.10.438>
- Cascio, S., & Zaret, K. S. (1991). Hepatocyte differentiation initiates during endodermal-mesenchymal interactions prior to liver formation. *Development*, 113(1), 217–225. <https://doi.org/10.1242/dev.113.1.217>
- Chalasani, N., Younossi, Z., Lavine, J. E., Charlton, M., Cusi, K., Rinella, M., Harrison, S. A., Brunt, E. M., & Sanyal, A. J. (2018). The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology*, 67(1), 328. <https://doi.org/10.1002/hep.29367>
- Charatcharoenwitthaya, P., Liangpunsakul, S., & Piratvisuth, T. (2020). Alcohol-Associated Liver Disease: East Versus West. *Clinical Liver Disease*, 16(6), 231–235. <https://doi.org/10.1002/cld.920>
- Chatterjee, S., & Ahituv, N. (2017). Gene Regulatory Elements, Major Drivers of Human Disease. *Annual Review of Genomics and Human Genetics*, 18(1), 45–63. <https://doi.org/10.1146/annurev-genom-091416-035537>
- Chen, A. F., Liu, A. J., Krishnakumar, R., Freimer, J. W., DeVeale, B., & Blelloch, R. (2018). GRHL2-Dependent Enhancer Switching Maintains a Pluripotent Stem Cell Transcriptional Subnetwork after Exit from Naive Pluripotency. *Cell Stem Cell*, 23(2), 226–238.e4. <https://doi.org/10.1016/j.stem.2018.06.005>
- Chen, F., Jimenez, R. J., Sharma, K., Luu, H. Y., Hsu, B. Y., Ravindranathan, A., Stohr, B. A., & Willenbring, H. (2020). Broad Distribution of Hepatocyte Proliferation in Liver Homeostasis and Regeneration. *Cell Stem Cell*, 26(1), 27–33.e4. <https://doi.org/10.1016/j.stem.2019.11.001>
- Chen, W., Xiao Liu, Z., Oh, J.-E., Shin, K.-H., Kim, R. H., Jiang, M., Park, N.-H., & Kang, M. K. (2012). Grainyhead-like 2 (GRHL2) inhibits keratinocyte differentiation through epigenetic mechanism. *Cell Death & Disease*, 3(12), e450–e450. <https://doi.org/10.1038/cddis.2012.190>
- Chen, W., Yi, J. K., Shimane, T., Mehrzarin, S., Lin, Y.-L., Shin, K.-H., Kim, R. H., Park, N.-H., & Kang, M. K. (2016). Grainyhead-like 2 regulates epithelial plasticity and

stemness in oral cancer cells. *Carcinogenesis*, 37(5), 500–510. <https://doi.org/10.1093/carcin/bgw027>

Chen, Y., Gao, W.-K., Shu, Y.-Y., & Ye, J. (2022). Mechanisms of ductular reaction in non-alcoholic steatohepatitis. *World Journal of Gastroenterology*, 28(19), 2088–2099. <https://doi.org/10.3748/wjg.v28.i19.2088>

Chen, Z., Li, S., Subramaniam, S., Shyy, J. Y.-J., & Chien, S. (2017). Epigenetic Regulation: A New Frontier for Biomedical Engineers. *Annual Review of Biomedical Engineering*, 19(Volume 19, 2017), 195–219. <https://doi.org/10.1146/annurev-bioeng-071516-044720>

Chittock, E. C., Latwiel, S., Miller, T. C. R., & Müller, C. W. (2017). Molecular architecture of polycomb repressive complexes. *Biochemical Society Transactions*, 45(1), 193–205. <https://doi.org/10.1042/BST20160173>

Chu, C.-S., Lo, P.-W., Yeh, Y.-H., Hsu, P.-H., Peng, S.-H., Teng, Y.-C., Kang, M.-L., Wong, C.-H., & Juan, L.-J. (2014). O-GlcNAcylation regulates EZH2 protein stability and function. *Proceedings of the National Academy of Sciences*, 111(4), 1355–1360. <https://doi.org/10.1073/pnas.1323226111>

Chung, V. Y., Tan, T. Z., Tan, M., Wong, M. K., Kuay, K. T., Yang, Z., Ye, J., Muller, J., Koh, C. M., Guccione, E., Thiery, J. P., & Huang, R. Y.-J. (2016). GRHL2-miR-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification. *Scientific Reports*, 6(1), 19943. <https://doi.org/10.1038/srep19943>

Clerbaux, L.-A., Hul, N. V., Gouw, A. S. H., Manco, R., Español-Suñer, R., & Leclercq, I. A. (2017). Relevance of the CDE and DDC Mouse Models to Study Ductular Reaction in Chronic Human Liver Diseases. In *Experimental Animal Models of Human Diseases—An Effective Therapeutic Strategy*. IntechOpen. <https://doi.org/10.5772/intechopen.69533>

Collins, B. E., Greer, C. B., Coleman, B. C., & Sweatt, J. D. (2019). Histone H3 lysine K4 methylation and its role in learning and memory. *Epigenetics & Chromatin*, 12(1), 7. <https://doi.org/10.1186/s13072-018-0251-8>

Cours US Abdomen—Foie. (n.d.). Retrieved 26 January 2025, from <https://echo.interniste.com/echographie-foie-us-fr-c2-s0>

- Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., ... Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, 50(D1), D988–D995. <https://doi.org/10.1093/nar/gkab1049>
- Davies, A., Zoubeidi, A., Beltran, H., & Selth, L. A. (2023). The Transcriptional and Epigenetic Landscape of Cancer Cell Lineage Plasticity. *Cancer Discovery*, 13(8), 1771–1788. <https://doi.org/10.1158/2159-8290.CD-23-0225>
- Davis, C. A., Hitz, B. C., Sloan, C. A., Chan, E. T., Davidson, J. M., Gabdank, I., Hilton, J. A., Jain, K., Baymuradov, U. K., Narayanan, A. K., Onate, K. C., Graham, K., Miyasato, S. R., Dreszer, T. R., Strattan, J. S., Jolanki, O., Tanaka, F. Y., & Cherry, J. M. (2018). The Encyclopedia of DNA elements (ENCODE): Data portal update. *Nucleic Acids Research*, 46(D1), D794–D801. <https://doi.org/10.1093/nar/gkx1081>
- de Jong, I. E. M., van den Heuvel, M. C., Wells, R. G., & Porte, R. J. (2021). The heterogeneity of the biliary tree. *Journal of Hepatology*, 75(5), 1236–1238. <https://doi.org/10.1016/j.jhep.2021.04.016>
- DeLeve, L. D., Wang, X., & Guo, Y. (2008). Sinusoidal endothelial cells prevent rat stellate cell activation and promote reversion to quiescence. *Hepatology*, 48(3), 920–930. <https://doi.org/10.1002/hep.22351>
- Dentin, R., Benhamed, F., Hainault, I., Fauveau, V., Foufelle, F., Dyck, J. R. B., Girard, J., & Postic, C. (2006). Liver-Specific Inhibition of ChREBP Improves Hepatic Steatosis and Insulin Resistance in ob/ob Mice. *Diabetes*, 55(8), 2159–2170. <https://doi.org/10.2337/db06-0200>
- Deprince, A., Haas, J. T., & Staels, B. (2020). Dysregulated lipid metabolism links NAFLD to cardiovascular disease. *Molecular Metabolism*, 42, 101092. <https://doi.org/10.1016/j.molmet.2020.101092>
- Desmet, V. J. (2011). Ductal plates in hepatic ductular reactions. Hypothesis and implications. I. Types of ductular reaction reconsidered. *Virchows Archiv*, 458(3), 251–259. <https://doi.org/10.1007/s00428-011-1048-3>
- Dianat, N., Dubois-Pot-Schneider, H., Steichen, C., Desterke, C., Leclerc, P., Raveux, A., Combettes, L., Weber, A., Corlu, A., & Dubart-Kupperschmitt, A. (2014). Generation

of functional cholangiocyte-like cells from human pluripotent stem cells and HepaRG cells. *Hepatology*, 60(2), 700–714. <https://doi.org/10.1002/hep.27165>

Duan, R., Du, W., & Guo, W. (2020). EZH2: A novel target for cancer treatment. *Journal of Hematology & Oncology*, 13(1), 104. <https://doi.org/10.1186/s13045-020-00937-8>

Dubel, L., Tanaka, A., Leung, P. S., Van De Water, J., Coppel, R., Roche, T., Johanet, C., Motokawa, Y., Ansari, A., & Gershwin, E. M. (1999). Autoepitope Mapping and Reactivity of Autoantibodies to the Dihydrolipoamide Dehydrogenase–Binding Protein (E3bp) and the Glycine Cleavage Proteins in Primary Biliary Cirrhosis. *Hepatology*, 29(4), 1013–1018. <https://doi.org/10.1002/hep.510290403>

Dubois, V., Gheeraert, C., Vankrunkelsven, W., Dubois-Chevalier, J., Dehondt, H., Bobowski-Gerard, M., Vinod, M., Zummo, F. P., Güiza, F., Ploton, M., Dorchies, E., Pineau, L., Boulinguez, A., Vallez, E., Woittrain, E., Baugé, E., Lalloyer, F., Duhem, C., Rabhi, N., ... Eeckhoutte, J. (2020). Endoplasmic reticulum stress actively suppresses hepatic molecular identity in damaged liver. *Molecular Systems Biology*, 16(5). <https://doi.org/10.15252/msb.20199156>

Dubois, V., Staels, B., Lefebvre, P., Verzi, M., & Eeckhoutte, J. (2020). Control of Cell Identity by the Nuclear Receptor HNF4 in Organ Pathophysiology. *Cells*, 9(10), 2185. <https://doi.org/10.3390/cells9102185>

Dubois-Chevalier, J., Dubois, V., Dehondt, H., Mazrooei, P., Mazuy, C., Sérandour, A. A., Gheeraert, C., Guillaume, P., Baugé, E., Derudas, B., Hennuyer, N., Paumelle, R., Marot, G., Carroll, J. S., Lupien, M., Staels, B., Lefebvre, P., & Eeckhoutte, J. (2017). The logic of transcriptional regulator recruitment architecture at *cis*-regulatory modules controlling liver functions. *Genome Research*, 27(6), 985–996. <https://doi.org/10.1101/gr.217075.116>

Dubois-Chevalier, J., Gheeraert, C., Berthier, A., Boulet, C., Dubois, V., Guille, L., Fourcot, M., Marot, G., Gauthier, K., Dubuquoy, L., Staels, B., Lefebvre, P., & Eeckhoutte, J. (2023). An extended transcription factor regulatory network controls hepatocyte identity. *EMBO Reports*, n/a(n/a), e57020. <https://doi.org/10.15252/embr.202357020>

Dubois-Chevalier, J., Mazrooei, Parisa, Lupien, Mathieu, Staels, Bart, Lefebvre, Philippe, & Eeckhoutte, J. (2018). Organizing combinatorial transcription factor

recruitment at cis-regulatory modules. *Transcription*, 9(4), 233–239. <https://doi.org/10.1080/21541264.2017.1394424>

Dubuquoy, L., Louvet, A., Lassailly, G., Truant, S., Boleslawski, E., Artru, F., Maggiotto, F., Gantier, E., Buob, D., Leteurtre, E., Cannesson, A., Dharancy, S., Moreno, C., Pruvot, F.-R., Bataller, R., & Mathurin, P. (2015). Progenitor cell expansion and impaired hepatocyte regeneration in explanted livers from alcoholic hepatitis. *Gut*, 64(12), 1949–1960. <https://doi.org/10.1136/gutjnl-2014-308410>

Dusart, P., Öling, S., Struck, E., Norreen-Thorsen, M., Zwahlen, M., Von Feilitzen, K., Oksvold, P., Botic, M., Iglesias, M., Renne, T., Odeberg, J., Pontén, F., Lindskog, C., Uhlén, M., & Butler, L. (2023). *A tissue centric atlas of cell type transcriptome enrichment signatures*. <https://doi.org/10.1101/2023.01.10.520698>

Dyson, J. K., Beuers, U., Jones, D. E. J., Lohse, A. W., & Hudson, M. (2018). Primary sclerosing cholangitis. *The Lancet*, 391(10139), 2547–2559. [https://doi.org/10.1016/S0140-6736\(18\)30300-3](https://doi.org/10.1016/S0140-6736(18)30300-3)

Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207–210. <https://doi.org/10.1093/nar/30.1.207>

Eslam, M., Newsome, P. N., Sarin, S. K., Anstee, Q. M., Targher, G., Romero-Gomez, M., Zelber-Sagi, S., Wai-Sun Wong, V., Dufour, J.-F., Schattenberg, J. M., Kawaguchi, T., Arrese, M., Valenti, L., Shiha, G., Tiribelli, C., Yki-Järvinen, H., Fan, J.-G., Grønbaek, H., Yilmaz, Y., ... George, J. (2020). A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *Journal of Hepatology*, 73(1), 202–209. <https://doi.org/10.1016/j.jhep.2020.03.039>

Fabris, L., Campello, E., Cadamuro, M., & Simioni, P. (2024). The evil relationship between liver fibrosis and cardiovascular disease in metabolic dysfunction-associated fatty liver disease (MAFLD): Looking for the culprit. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1870(3), 166763. <https://doi.org/10.1016/j.bbadis.2023.166763>

Falletti, E., Cussigh, A., Cmet, S., Fabris, C., & Toniutto, P. (2016). *PNPLA3* rs738409 and *TM6SF2* rs58542926 variants increase the risk of hepatocellular carcinoma in alcoholic cirrhosis. *Digestive and Liver Disease*, 48(1), 69–75. <https://doi.org/10.1016/j.dld.2015.09.009>

- Feng, R., Liebe, R., & Weng, H.-L. (2023). Transcription networks in liver development and acute liver failure. *Liver Research*, 7(1), 47–55. <https://doi.org/10.1016/j.livres.2022.11.010>
- Field, A., & Adelman, K. (2020). Evaluating Enhancer Function and Transcription. *Annual Review of Biochemistry*, 89(Volume 89, 2020), 213–234. <https://doi.org/10.1146/annurev-biochem-011420-095916>
- Folseraas, T., Melum, E., Franke, A., & Karlsen, T. H. (2011). Genetics in primary sclerosing cholangitis. *Best Practice & Research Clinical Gastroenterology*, 25(6), 713–726. <https://doi.org/10.1016/j.bpg.2011.09.010>
- Forrester-Gauntlett, B., Peters, L., & Oback, B. (2023). Grainyhead-like 2 is required for morphological integrity of mouse embryonic stem cells and orderly formation of inner ear-like organoids. *Frontiers in Cell and Developmental Biology*, 11. <https://doi.org/10.3389/fcell.2023.1112069>
- Fraile, J. M., Palliyil, S., Barelle, C., Porter, A. J., & Kovaleva, M. (2021). Non-Alcoholic Steatohepatitis (NASH) & nash; A Review of a Crowded Clinical Landscape, Driven by a Complex Disease. *Drug Design, Development and Therapy*, 15, 3997–4009. <https://doi.org/10.2147/DDDT.S315724>
- Freese, N. H., Norris, D. C., & Loraine, A. E. (2016). Integrated genome browser: Visual analytics platform for genomics. *Bioinformatics*, 32(14), 2089–2095. <https://doi.org/10.1093/bioinformatics/btw069>
- Friedman, S. L. (2008). Hepatic Stellate Cells: Protean, Multifunctional, and Enigmatic Cells of the Liver. *Physiological Reviews*, 88(1), 125–172. <https://doi.org/10.1152/physrev.00013.2007>
- Fujii, T., Fuchs, B. C., Yamada, S., Lauwers, G. Y., Kulu, Y., Goodwin, J. M., Lanuti, M., & Tanabe, K. K. (2010). Mouse model of carbon tetrachloride induced liver fibrosis: Histopathological changes and expression of CD133 and epidermal growth factor. *BMC Gastroenterology*, 10, 79. <https://doi.org/10.1186/1471-230X-10-79>
- Gadd, V. L., Aleksieva, N., & Forbes, S. J. (2020). Epithelial Plasticity during Liver Injury and Regeneration. *Cell Stem Cell*, 27(4), 557–573. <https://doi.org/10.1016/j.stem.2020.08.016>

- Gaggini, M., Carli, F., Rosso, C., Buzzigoli, E., Marietti, M., Della Latta, V., Ciociaro, D., Abate, M. L., Gambino, R., Cassader, M., Bugianesi, E., & Gastaldelli, A. (2018). Altered amino acid concentrations in NAFLD: Impact of obesity and insulin resistance. *Hepatology*, 67(1), 145. <https://doi.org/10.1002/hep.29465>
- Galvanin, C., Buch, S., Nahon, P., & Trépo, E. (2025). PNPLA3 in Alcohol-Related Liver Disease. *Liver International*, 45(1), e16211. <https://doi.org/10.1111/liv.16211>
- Gao, B., Xiang, X., Leggio, L., & Koob, G. F. (2020). Alcoholic Liver Disease. In I. M. Arias, H. J. Alter, J. L. Boyer, D. E. Cohen, D. A. Shafritz, S. S. Thorgeirsson, & A. W. Wolkoff (Eds.), *The Liver* (1st ed., pp. 682–700). Wiley. <https://doi.org/10.1002/9781119436812.ch53>
- Gao, X., Vockley, C. M., Pauli, F., Newberry, K. M., Xue, Y., Randell, S. H., Reddy, T. E., & Hogan, B. L. M. (2013). Evidence for multiple roles for grainyhead-like 2 in the establishment and maintenance of human mucociliary airway epithelium. *Proceedings of the National Academy of Sciences*, 110(23), 9356–9361. <https://doi.org/10.1073/pnas.1307589110>
- Gawrieh, S., Nouredin, M., Loo, N., Mohseni, R., Awasty, V., Cusi, K., Kowdley, K. V., Lai, M., Schiff, E., Parmar, D., Patel, P., & Chalasani, N. (2021). Saroglitazar, a PPAR- α/γ Agonist, for Treatment of NAFLD: A Randomized Controlled Double-Blind Phase 2 Trial. *Hepatology*, 74(4), 1809. <https://doi.org/10.1002/hep.31843>
- Geng, Y., & Schwabe, R. F. (2025). Hepatic stellate cell heterogeneity: Functional aspects and therapeutic implications. *Hepatology*, 10.1097/HEP.0000000000001386. <https://doi.org/10.1097/HEP.0000000000001386>
- German, B., & Ellis, L. (2022). Polycomb Directed Cell Fate Decisions in Development and Cancer. *Epigenomes*, 6(3), Article 3. <https://doi.org/10.3390/epigenomes6030028>
- Ginès, P., Krag, A., Abraldes, J. G., Solà, E., Fabrellas, N., & Kamath, P. S. (2021). Liver cirrhosis. *The Lancet*, 398(10308), 1359–1376. [https://doi.org/10.1016/S0140-6736\(21\)01374-X](https://doi.org/10.1016/S0140-6736(21)01374-X)
- Gissen, P., & Arias, I. M. (2015). Structural and functional hepatocyte polarity and liver disease. *Journal of Hepatology*, 63(4), 1023–1037. <https://doi.org/10.1016/j.jhep.2015.06.015>

- Gong, J., Tu, W., Liu, J., & Tian, D. (2023). Hepatocytes: A key role in liver inflammation. *Frontiers in Immunology*, 13. <https://doi.org/10.3389/fimmu.2022.1083780>
- Gong, Z., Tas, E., Yakar, S., & Muzumdar, R. (2017). Hepatic lipid metabolism and non-alcoholic fatty liver disease in aging. *Molecular and Cellular Endocrinology*, 455, 115–130. <https://doi.org/10.1016/j.mce.2016.12.022>
- Gordillo, M., Evans, T., & Gouon-Evans, V. (2015). Orchestrating liver development. *Development*, 142(12), 2094–2108. <https://doi.org/10.1242/dev.114215>
- Gribben, C., Galanakis, V., Calderwood, A., Williams, E. C., Chazarra-Gil, R., Larraz, M., Frau, C., Puengel, T., Guillot, A., Rouhani, F. J., Mahbubani, K., Godfrey, E., Davies, S. E., Athanasiadis, E., Saeb-Parsy, K., Tacke, F., Allison, M., Mohorianu, I., & Vallier, L. (2024). Acquisition of epithelial plasticity in human chronic liver disease. *Nature*. <https://doi.org/10.1038/s41586-024-07465-2>
- Grimaudo, S., Pipitone, R. M., Pennisi, G., Celsa, C., Cammà, C., Di Marco, V., Barcellona, M. R., Boemi, R., Enea, M., Giannetti, A., Spatola, F., Marchesini, G., Craxì, A., & Petta, S. (2020). Association Between *PNPLA3* rs738409 C>G Variant and Liver-Related Outcomes in Patients With Nonalcoholic Fatty Liver Disease. *Clinical Gastroenterology and Hepatology*, 18(4), 935-944.e3. <https://doi.org/10.1016/j.cgh.2019.08.011>
- Grindheim, J. M., Nicetto, D., Donahue, G., & Zaret, K. S. (2019). Polycomb Repressive Complex 2 Proteins EZH1 and EZH2 Regulate Timing of Postnatal Hepatocyte Maturation and Fibrosis by Repressing Genes With Euchromatic Promoters in Mice. *Gastroenterology*, 156(6), 1834–1848. <https://doi.org/10.1053/j.gastro.2019.01.041>
- Gripon, P., Rumin, S., Urban, S., Le Seyec, J., Glaise, D., Cannie, I., Guyomard, C., Lucas, J., Trepo, C., & Guguen-Guillouzo, C. (2002). Infection of a human hepatoma cell line by hepatitis B virus. *Proceedings of the National Academy of Sciences*, 99(24), 15655–15660. <https://doi.org/10.1073/pnas.232137699>
- Grzych, G., Bernard, L., Lestrelin, R., Tailleux, A., & Staels, B. (2023). État des lieux sur la physiopathologie, le diagnostic et les traitements de la stéato-hépatite non alcoolique (NASH). *Annales Pharmaceutiques Françaises*, 81(2), 183–201. <https://doi.org/10.1016/j.pharma.2022.09.005>

- Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, 32(18), 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313>
- Guille, L., Johanns, M., Zummo, F., Staels, B., Lefebvre, P., Eeckhoutte, J., & Dubois-Chevalier, J. (2022). *ISCEBERG: Interactive Single Cell Expression Browser for Exploration of RNAseq data using Graphics* (Version v1.0.1) [Computer software]. Zenodo. <https://doi.org/10.5281/zenodo.6563734>
- Guo, Y., Zhao, S., & Wang, G. G. (2021). Polycomb Gene Silencing Mechanisms: PRC2 Chromatin Targeting, H3K27me3 ‘Readout’, and Phase Separation-Based Compaction. *Trends in Genetics*, 37(6), 547–565. <https://doi.org/10.1016/j.tig.2020.12.006>
- Haas, J. T., Francque, S., & Staels, B. (2016). Pathophysiology and Mechanisms of Nonalcoholic Fatty Liver Disease. *Annual Review of Physiology*, 78(1), 181–205. <https://doi.org/10.1146/annurev-physiol-021115-105331>
- Hafemeister, C., & Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biology*, 20(1), Article 1. <https://doi.org/10.1186/s13059-019-1874-1>
- Hammond, N. L., Murtuza Baker, S., Georgaka, S., Al-Anbaki, A., Jokl, E., Simpson, K., Sanchez-Alvarez, R., Athwal, V. S., Purssell, H., Siriwardena, A. K., Spiers, H. V. M., Dixon, M. J., Bere, L. D., Jones, A. P., Haley, M. J., Couper, K. N., Bobola, N., Sharrocks, A. D., Hanley, N. A., ... Piper Hanley, K. (2025). Spatial gene regulatory networks driving cell state transitions during human liver disease. *EMBO Molecular Medicine*, 17(6), 1452–1474. <https://doi.org/10.1038/s44321-025-00230-6>
- Han, H.-S., Kang, G., Kim, J. S., Choi, B. H., & Koo, S.-H. (2016). Regulation of glucose metabolism from a liver-centric perspective. *Experimental & Molecular Medicine*, 48(3), e218–e218. <https://doi.org/10.1038/emm.2015.122>
- Hansen, D., Jensen, J. E. R., Andersen, C. A. T., Jakobsgaard, P. R., Havelund, J., Lauritsen, L., Mandacaru, S., Siersbaek, M., Shackleton, O. L., Inoue, H., Brewer, J. R., Schwabe, R. F., Blagoev, B., Færgeman, N. J., Salmi, M., & Ravnskjaer, K. (2025). Hepatic stellate cells regulate liver fatty acid utilization via plasmalemma vesicle-associated protein. *Cell Metabolism*, 37(4), 971-986.e8. <https://doi.org/10.1016/j.cmet.2025.01.022>

Hao, Y., Stuart, T., Kowalski, M. H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., & Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nature Biotechnology*, 42(2), 293–304. <https://doi.org/10.1038/s41587-023-01767-y>

He, Q., Cui, L., Yuan, X., Wang, M., & Hui, L. (2022). Cell identity conversion in liver regeneration after injury. *Current Opinion in Genetics & Development*, 75, 101921. <https://doi.org/10.1016/j.gde.2022.101921>

Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., Ye, Z., Lee, L. K., Stuart, R. K., Ching, C. W., Ching, K. A., Antosiewicz-Bourget, J. E., Liu, H., Zhang, X., Green, R. D., Lobanenkov, V. V., Stewart, R., Thomson, J. A., Crawford, G. E., ... Ren, B. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459(7243), 108–112. <https://doi.org/10.1038/nature07829>

Hitz, B. C., Jin-Wook, L., Jolanki, O., Kagda, M. S., Graham, K., Sud, P., Gabdank, I., Strattan, J. S., Sloan, C. A., Dreszer, T., Rowe, L. D., Podduturi, N. R., Malladi, V. S., Chan, E. T., Davidson, J. M., Ho, M., Miyasato, S., Simison, M., Tanaka, F., ... Cherry, J. M. (2023). The ENCODE Uniform Analysis Pipelines. *bioRxiv: The Preprint Server for Biology*, 2023.04.04.535623. <https://doi.org/10.1101/2023.04.04.535623>

Holding, A. N., Giorgi, F. M., Donnelly, A., Cullen, A. E., Nagarajan, S., Selth, L. A., & Markowitz, F. (2019). VULCAN integrates ChIP-seq with patient-derived co-expression networks to identify GRHL2 as a key co-regulator of ERα at enhancers in breast cancer. *Genome Biology*, 20(1), 91. <https://doi.org/10.1186/s13059-019-1698-z>

Holoch, D., & Margueron, R. (2017). Mechanisms Regulating PRC2 Recruitment and Enzymatic Activity. *Trends in Biochemical Sciences*, 42(7), 531–542. <https://doi.org/10.1016/j.tibs.2017.04.003>

Holoch, D., Wassef, M., Lökvist, C., Zielinski, D., Aflaki, S., Lombard, B., Héry, T., Loew, D., Howard, M., & Margueron, R. (2021). A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nature Genetics*, 53(12), Article 12. <https://doi.org/10.1038/s41588-021-00964-2>

Hrncir, H. R., & Gracz, A. D. (2023). Cellular and Transcriptional Heterogeneity in the Intrahepatic Biliary Epithelium. *Gastro Hep Advances*, 2(1), 108–120. <https://doi.org/10.1016/j.gastha.2022.07.015>

Hu, H., Miao, Y.-R., Jia, L.-H., Yu, Q.-Y., Zhang, Q., & Guo, A.-Y. (2019). AnimalTFDB 3.0: A comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Research*, 47(D1), D33–D38. <https://doi.org/10.1093/nar/gky822>

Hu, S., Ai, Y., Hu, C., Cassim Bawa, F. N., & Xu, Y. (2025). Transcription factors, metabolic dysfunction-associated fatty liver disease, and therapeutic implications. *Genes & Diseases*, 12(3), 101372. <https://doi.org/10.1016/j.gendis.2024.101372>

Huang, W.-J., Qiu, B.-J., Qi, X.-S., Chen, C.-Y., Liu, W.-M., Zhou, S., Ding, M., Lu, F.-F., Zhao, J., Tang, D., Zhou, X., Fu, G.-B., Wang, Z.-Y., Ma, H.-Q., Wu, Y.-L., Wu, H., Chen, X.-S., Yu, W.-F., & Yan, H.-X. (2023). CD24+LCN2+ liver progenitor cells in ductular reaction contributed to macrophage inflammatory responses in chronic liver injury. *Cell & Bioscience*, 13(1), Article 1. <https://doi.org/10.1186/s13578-023-01123-2>

Huyghe, A., Trajkova, A., & Lavial, F. (2024). Cellular plasticity in reprogramming, rejuvenation and tumorigenesis: A pioneer TF perspective. *Trends in Cell Biology*, 34(3), 255–267. <https://doi.org/10.1016/j.tcb.2023.07.013>

Ieda, M., Fu, J.-D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B. G., & Srivastava, D. (2010). Direct Reprogramming of Fibroblasts into Functional Cardiomyocytes by Defined Factors. *Cell*, 142(3), 375–386. <https://doi.org/10.1016/j.cell.2010.07.002>

Imai, T., Hieshima, K., Haskell, C., Baba, M., Nagira, M., Nishimura, M., Kakizaki, M., Takagi, S., Nomiyama, H., Schall, T. J., & Yoshie, O. (1997). Identification and Molecular Characterization of Fractalkine Receptor CX3CR1, which Mediates Both Leukocyte Migration and Adhesion. *Cell*, 91(4), 521–530. [https://doi.org/10.1016/S0092-8674\(00\)80438-9](https://doi.org/10.1016/S0092-8674(00)80438-9)

Itoh, T., & Miyajima, A. (2014). Liver regeneration by stem/progenitor cells: ITOH AND MIYAJIMA. *Hepatology*, 59(4), 1617–1626. <https://doi.org/10.1002/hep.26753>

Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J. A., Mahony, S., Pugh, B. F., Lee, D., Kaestner, K. H., & Zaret, K. S. (2016). The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Molecular Cell*, 62(1), 79–91. <https://doi.org/10.1016/j.molcel.2016.03.001>

- Jacobs, J., Atkins, M., Davie, K., Imrichova, H., Romanelli, L., Christiaens, V., Hulselmans, G., Potier, D., Wouters, J., Taskiran, I. I., Paciello, G., González-Blas, C. B., Koldere, D., Aibar, S., Halder, G., & Aerts, S. (2018). The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nature Genetics*, 50(7), 1011–1020. <https://doi.org/10.1038/s41588-018-0140-x>
- Jadhav, U., Nalapareddy, K., Saxena, M., O'Neill, N. K., Pinello, L., Yuan, G.-C., Orkin, S. H., & Shivdasani, R. A. (2016). Acquired tissue-specific promoter bivalency is a basis for PRC2 necessity in adult cells. *Cell*, 165(6), 1389–1400. <https://doi.org/10.1016/j.cell.2016.04.031>
- Jiang, Y., Zhang, T., Kusumanchi, P., Han, S., Yang, Z., & Liangpunsakul, S. (2020). Alcohol Metabolizing Enzymes, Microsomal Ethanol Oxidizing System, Cytochrome P450 2E1, Catalase, and Aldehyde Dehydrogenase in Alcohol-Associated Liver Disease. *Biomedicines*, 8(3), Article 3. <https://doi.org/10.3390/biomedicines8030050>
- Jonas, W., & Schürmann, A. (2021). Genetic and epigenetic factors determining NAFLD risk. *Molecular Metabolism*, 50, 101111. <https://doi.org/10.1016/j.molmet.2020.101111>
- Jung, Y., Witek, R. P., Syn, W.-K., Choi, S. S., Omenetti, A., Premont, R., Guy, C. D., & Diehl, A. M. (2010). Signals from dying hepatocytes trigger growth of liver progenitors. *Gut*, 59(5), 655–665. <https://doi.org/10.1136/gut.2009.204354>
- Kalra, A., Yetiskul, E., Wehrle, C. J., & Tuma, F. (2025). Physiology, Liver. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK535438/>
- Kamath, B., & Mack, C. (2019). From Hepatocyte to Cholangiocyte: The Remarkable Potential of Transdifferentiation to Treat Cholestatic Diseases. *Hepatology*, 69(4), 1828. <https://doi.org/10.1002/hep.30250>
- Karantza, V. (2011). Keratins in health and cancer: More than mere epithelial cell markers. *Oncogene*, 30(2), 127–138. <https://doi.org/10.1038/onc.2010.456>
- Karlsen, T. H., Folseraas, T., Thorburn, D., & Vesterhus, M. (2017). Primary sclerosing cholangitis – a comprehensive review. *Journal of Hepatology*, 67(6), 1298–1323. <https://doi.org/10.1016/j.jhep.2017.07.022>

- Karlsson, M., Zhang, C., Méar, L., Zhong, W., Digre, A., Katona, B., Sjöstedt, E., Butler, L., Odeberg, J., Dusart, P., Edfors, F., Oksvold, P., von Feilitzen, K., Zwahlen, M., Arif, M., Altay, O., Li, X., Ozcan, M., Mardinoglu, A., ... Lindskog, C. (2021). A single-cell type transcriptomics map of human tissues. *Science Advances*, 7(31), eabh2169. <https://doi.org/10.1126/sciadv.abh2169>
- Kars, G., & Muyan, M. (2018). GRHL2 (grainyhead like transcription factor 2). *Atlas of Genetics and Cytogenetics in Oncology and Haematology*, 5. <https://doi.org/10.4267/2042/68899>
- Katsuda, T., Sussman, J. H., Ito, K., Katznelson, A., Yuan, S., Takenaka, N., Li, J., Merrell, A. J., Cure, H., Li, Q., Rasool, R. U., Asangani, I. A., Zaret, K. S., & Stanger, B. Z. (2024). Cellular reprogramming in vivo initiated by SOX4 pioneer factor activity. *Nature Communications*, 15(1), 1761. <https://doi.org/10.1038/s41467-024-45939-z>
- Kawamura, S., Matsushita, Y., Kurosaki, S., Tange, M., Fujiwara, N., Hayata, Y., Hayakawa, Y., Suzuki, N., Hata, M., Tsuboi, M., Kishikawa, T., Kinoshita, H., Nakatsuka, T., Sato, M., Kudo, Y., Hoshida, Y., Umemura, A., Eguchi, A., Ikenoue, T., ... Nakagawa, H. (2022). Inhibiting SCAP/SREBP exacerbates liver injury and carcinogenesis in murine nonalcoholic steatohepatitis. *The Journal of Clinical Investigation*, 132(11). <https://doi.org/10.1172/JCI151895>
- Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, 12(4), 357–360. <https://doi.org/10.1038/nmeth.3317>
- Kim, H., Jang, M.-J., Kang, M.-J., & Han, Y.-M. (2011). Epigenetic signatures and temporal expression of lineage-specific genes in hESCs during differentiation to hepatocytes in vitro. *Human Molecular Genetics*, 20(3), 401–412. <https://doi.org/10.1093/hmg/ddq476>
- Kisseleva, T., Ganguly, S., Murad, R., Wang, A., & Brenner, D. A. (2025). Regulation of Hepatic Stellate Cell Phenotypes in Metabolic Dysfunction–Associated Steatohepatitis. *Gastroenterology*, 169(5), 797–812. <https://doi.org/10.1053/j.gastro.2025.03.010>
- Köhn-Gaone, J., Gogoi-Tiwari, J., Ramm, G. A., Olynyk, J. K., & Tirnitz-Parker, J. E. E. (2016). The role of liver progenitor cells during liver regeneration, fibrogenesis, and

carcinogenesis. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 310(3), G143–G154. <https://doi.org/10.1152/ajpgi.00215.2015>

Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, 128(4), 693–705. <https://doi.org/10.1016/j.cell.2007.02.005>

Kouznetsova, V. L., Tchekanov, A., Li, X., Yan, X., & Tsigelny, I. F. (2019). Polycomb repressive 2 complex—Molecular mechanisms of function. *Protein Science*, 28(8), 1387–1399. <https://doi.org/10.1002/pro.3647>

Krenkel, O., & Tacke, F. (2017). Liver macrophages in tissue homeostasis and disease. *Nature Reviews Immunology*, 17(5), 306–321. <https://doi.org/10.1038/nri.2017.11>

Kusakabe, Y., Chiba, T., Oshima, M., Koide, S., Rizq, O., Aoyama, K., Ao, J., Kaneko, T., Kanzaki, H., Kanayama, K., Maeda, T., Saito, T., Nakagawa, R., Kobayashi, K., Kiyono, S., Nakamura, M., Ogasawara, S., Suzuki, E., Nakamoto, S., ... Kato, N. (2021). EZH1/2 inhibition augments the anti-tumor effects of sorafenib in hepatocellular carcinoma. *Scientific Reports*, 11(1), 21396. <https://doi.org/10.1038/s41598-021-00889-0>

Kyrmizi, I., Hatzis, P., Katrakili, N., Tronche, F., Gonzalez, F. J., & Talianidis, I. (2006). Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes & Development*, 20(16), 2293. <https://doi.org/10.1101/gad.390906>

Lai, W. K. M., & Pugh, B. F. (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews Molecular Cell Biology*, 18(9), 548–562. <https://doi.org/10.1038/nrm.2017.47>

Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, 175(2), 598–599. <https://doi.org/10.1016/j.cell.2018.09.045>

Lan, T., Tai, Y., Zhao, C., Xiao, Y., Yang, Z., Zhang, L., Gan, C., Dai, W., Tong, H., Tang, C., Huang, Z., & Gao, J. (2023). Atypical cholangiocytes derived from hepatocyte-cholangiocyte transdifferentiation mediated by COX-2: A kind of misguided liver regeneration. *Inflammation and Regeneration*, 43(1), Article 1. <https://doi.org/10.1186/s41232-023-00284-4>

- Lau, A. H., & Thomson, A. W. (2003). Dendritic cells and immune regulation in the liver. *Gut*, 52(2), 307–314. <https://doi.org/10.1136/gut.52.2.307>
- Laugesen, A., Højfeldt, J. W., & Helin, K. (2019). Molecular mechanisms directing PRC2 recruitment and H3K27 methylation. *Molecular Cell*, 74(1), 8–18. <https://doi.org/10.1016/j.molcel.2019.03.011>
- Laurent, V., Glaise, D., Nübel, T., Gilot, D., Corlu, A., & Loyer, P. (2013). Highly Efficient SiRNA and Gene Transfer into Hepatocyte-Like HepaRG Cells and Primary Human Hepatocytes: New Means for Drug Metabolism and Toxicity Studies. In I. R. Phillips, E. A. Shephard, & P. R. Ortiz de Montellano (Eds.), *Cytochrome P450 Protocols* (pp. 295–314). Humana Press. https://doi.org/10.1007/978-1-62703-321-3_25
- Lê, S., Josse, J., & Husson, F. (2008). **FactoMineR**: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1). <https://doi.org/10.18637/jss.v025.i01>
- Lee, S.-H., So, J., & Shin, D. (2023). Hepatocyte-to-cholangiocyte conversion occurs through transdifferentiation independently of proliferation in zebrafish. *Hepatology*, 77(4), 1198. <https://doi.org/10.1097/HEP.0000000000000016>
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, 28(6), 882–883. <https://doi.org/10.1093/bioinformatics/bts034>
- Legastelois, R., Jeanblanc, J., Vilpoux, C., Bourguet, E., & Naassila, M. (2017). Mécanismes épigénétiques et troubles de l'usage d'alcool: Une cible thérapeutique intéressante? *Biologie Aujourd'hui*, 211(1), 83–91. <https://doi.org/10.1051/jbio/2017014>
- Leinonen, R., Sugawara, H., Shumway, M., & on behalf of the International Nucleotide Sequence Database Collaboration. (2011). The Sequence Read Archive. *Nucleic Acids Research*, 39(suppl_1), D19–D21. <https://doi.org/10.1093/nar/gkq1019>
- Lekakis, V., & Papatheodoridis, G. V. (2024). Natural history of metabolic dysfunction-associated steatotic liver disease. *European Journal of Internal Medicine*, 122, 3–10. <https://doi.org/10.1016/j.ejim.2023.11.005>

- Lemaigre, F. P. (2009). Mechanisms of Liver Development: Concepts for Understanding Liver Disorders and Design of Novel Therapies. *Gastroenterology*, 137(1), 62–79. <https://doi.org/10.1053/j.gastro.2009.03.035>
- Li, L., Cui, L., Lin, P., Liu, Z., Bao, S., Ma, X., Nan, H., Zhu, W., Cen, J., Mao, Y., Ma, X., Jiang, L., Nie, Y., Ginhoux, F., Li, Y., Li, H., & Hui, L. (2023). Kupffer-cell-derived IL-6 is repurposed for hepatocyte dedifferentiation via activating progenitor genes from injury-specific enhancers. *Cell Stem Cell*, 30(3), 283-299.e9. <https://doi.org/10.1016/j.stem.2023.01.009>
- Li, P., He, K., Li, J., Liu, Z., & Gong, J. (2017). The role of Kupffer cells in hepatic diseases. *Molecular Immunology*, 85, 222–229. <https://doi.org/10.1016/j.molimm.2017.02.018>
- Li, S., Wan, C., Zheng, R., Fan, J., Dong, X., Meyer, C. A., & Liu, X. S. (2019). Cistrome-GO: A web server for functional enrichment analysis of transcription factor ChIP-seq peaks. *Nucleic Acids Research*, 47(W1), W206–W211. <https://doi.org/10.1093/nar/gkz332>
- Li, W., Li, L., & Hui, L. (2020). Cell Plasticity in Liver Regeneration. *Trends in Cell Biology*, 30(4), 329–338. <https://doi.org/10.1016/j.tcb.2020.01.007>
- Li, W., Yang, L., He, Q., Hu, C., Zhu, L., Ma, X., Ma, X., Bao, S., Li, L., Chen, Y., Deng, X., Zhang, X., Cen, J., Zhang, L., Wang, Z., Xie, W.-F., Li, H., Li, Y., & Hui, L. (2019a). A Homeostatic Arid1a-Dependent Permissive Chromatin State Licenses Hepatocyte Responsiveness to Liver-Injury-Associated YAP Signaling. *Cell Stem Cell*, 25(1), 54-68.e5. <https://doi.org/10.1016/j.stem.2019.06.008>
- Li, W., Yang, L., He, Q., Hu, C., Zhu, L., Ma, X., Ma, X., Bao, S., Li, L., Chen, Y., Deng, X., Zhang, X., Cen, J., Zhang, L., Wang, Z., Xie, W.-F., Li, H., Li, Y., & Hui, L. (2019b). A Homeostatic Arid1a-Dependent Permissive Chromatin State Licenses Hepatocyte Responsiveness to Liver-Injury-Associated YAP Signaling. *Cell Stem Cell*, 25(1), 54-68.e5. <https://doi.org/10.1016/j.stem.2019.06.008>
- Li, X., Lin, J., pan, Y., Cui, P., & Xia, J. (2021). Identification of a Liver Progenitor Cell-Related Genes Signature Predicting Overall Survival for Hepatocellular Carcinoma. *Technology in Cancer Research & Treatment*, 20, 15330338211041425. <https://doi.org/10.1177/15330338211041425>

- Li, Z., & Sun, X. (2024). Epigenetic regulation in liver regeneration. *Life Sciences*, 353, 122924. <https://doi.org/10.1016/j.lfs.2024.122924>
- 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., & Alison, M. R. (2010). The Histogenesis of Regenerative Nodules in Human Liver Cirrhosis. *Hepatology*, 51(3), 1017. <https://doi.org/10.1002/hep.23483>
- Liska, D. J. (1998). The detoxification enzyme systems. *Alternative Medicine Review: A Journal of Clinical Therapeutic*, 3(3), 187–198.
- Liu, S.-Y., Tsai, I.-T., & Hsu, Y.-C. (2021). Alcohol-Related Liver Disease: Basic Mechanisms and Clinical Perspectives. *International Journal of Molecular Sciences*, 22(10), Article 10. <https://doi.org/10.3390/ijms22105170>
- Liu, X., Wang, H., Liang, X., & Roberts, M. S. (2017). Hepatic Metabolism in Liver Health and Disease. In *Liver Pathophysiology* (pp. 391–400). Elsevier. <https://doi.org/10.1016/B978-0-12-804274-8.00030-8>
- Lleo, A., Marzorati, S., Anaya, J.-M., & Gershwin, M. E. (2017). Primary biliary cholangitis: A comprehensive overview. *Hepatology International*, 11(6), 485–499. <https://doi.org/10.1007/s12072-017-9830-1>
- Lou, C., Lan, T., Xu, S., Hu, X., Li, J., Xiang, Z., Lin, S., Fan, X., Chen, J., & Xu, X. (2025). Heterogeneity and plasticity of cholangiocytes in liver injury: A journey from pathophysiology to therapeutic utility. *Gut*. <https://doi.org/10.1136/gutjnl-2025-334763>
- Louvet, A., Labreuche, J., Artru, F., Boursier, J., Kim, D. J., O'Grady, J., Trépo, E., Nahon, P., Ganne-Carrié, N., Naveau, S., Diaz, E., Gustot, T., Lassailly, G., Cannesson-Leroy, A., Canva-Delcambre, V., Dharancy, S., Park, S. H., Moreno, C., Morgan, T. R., ... Mathurin, P. (2015). Combining Data From Liver Disease Scoring Systems Better Predicts Outcomes of Patients With Alcoholic Hepatitis. *Gastroenterology*, 149(2), 398-406.e8. <https://doi.org/10.1053/j.gastro.2015.04.044>
- Louvet, A., Labreuche, J., Artru, F., Bouthors, A., Rolland, B., Saffers, P., Lollivier, J., Lemaître, E., Dharancy, S., Lassailly, G., Canva-Delcambre, V., Duhamel, A., & Mathurin, P. (2017). Main drivers of outcome differ between short term and long term in severe alcoholic hepatitis: A prospective study. *Hepatology*, 66(5), 1464. <https://doi.org/10.1002/hep.29240>

- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Mackowiak, B., Fu, Y., Maccioni, L., & Gao, B. (2024). Alcohol-associated liver disease. *The Journal of Clinical Investigation*, 134(3). <https://doi.org/10.1172/JCI176345>
- MacParland, S. A., Liu, J. C., Ma, X.-Z., Innes, B. T., Bartczak, A. M., Gage, B. K., Manuel, J., Khuu, N., Echeverri, J., Linares, I., Gupta, R., Cheng, M. L., Liu, L. Y., Camat, D., Chung, S. W., Seliga, R. K., Shao, Z., Lee, E., Ogawa, S., ... McGilvray, I. D. (2018). Single cell RNA sequencing of human liver reveals distinct intrahepatic macrophage populations. *Nature Communications*, 9(1), 4383. <https://doi.org/10.1038/s41467-018-06318-7>
- Mak, K. M., & Shekhar, A. C. (2025). Lipopolysaccharide, arbiter of the gut–liver axis, modulates hepatic cell pathophysiology in alcoholism. *The Anatomical Record*, 308(3), 975–1004. <https://doi.org/10.1002/ar.25562>
- Mak, K. M., Wu, C., & Cheng, C. P. (2023). Lipid droplets, the Holy Grail of hepatic stellate cells: In health and hepatic fibrosis. *The Anatomical Record*, 306(5), 983–1010. <https://doi.org/10.1002/ar.25138>
- Malarkey, D. E., Johnson, K., Ryan, L., Boorman, G., & Maronpot, R. R. (2005). New Insights into Functional Aspects of Liver Morphology. *Toxicologic Pathology*, 33(1), 27–34. <https://doi.org/10.1080/01926230590881826>
- Mandrekar, P. (2011). Epigenetic regulation in alcoholic liver disease. *World Journal of Gastroenterology: WJG*, 17(20), 2456–2464. <https://doi.org/10.3748/wjg.v17.i20.2456>
- Masalha, M., Ben-Dov, I. Z., Ram, O., Meningher, T., Jacob-Hirsch, J., Kassem, R., Sidi, Y., & Avni, D. (2021). H3K27Ac modification and gene expression in psoriasis. *Journal of Dermatological Science*, 103(2), 93–100. <https://doi.org/10.1016/j.jdermsci.2021.07.003>
- Matsumoto, T., Wakefield, L., Tarlow, B. D., & Grompe, M. (2020). In Vivo Lineage Tracing of Polyploid Hepatocytes Reveals Extensive Proliferation during Liver Regeneration. *Cell Stem Cell*, 26(1), 34–47.e3. <https://doi.org/10.1016/j.stem.2019.11.014>

- Mavila, N., Siraganahalli Eshwaraiah, M., & Kennedy, J. (2024). Ductular Reactions in Liver Injury, Regeneration, and Disease Progression—An Overview. *Cells*, 13(7), Article 7. <https://doi.org/10.3390/cells13070579>
- Medina-Rivera, A., Santiago-Algarra, D., Puthier, D., & Spicuglia, S. (2018). Widespread Enhancer Activity from Core Promoters. *Trends in Biochemical Sciences*, 43(6), 452–468. <https://doi.org/10.1016/j.tibs.2018.03.004>
- Meng, L., Du, M., Li, H., Kong, F., Yang, J., Dong, R., Zheng, S., Chen, G., Shen, Z., & Wang, J. (2025). Single-cell transcription reveals hepatocyte-to-cholangiocyte reprogramming and biliary gene profile in biliary atresia. *Hepatology Communications*, 9(6), e0710. <https://doi.org/10.1097/HC9.0000000000000710>
- Merrell, A. J., Peng, T., Li, J., Sun, K., Li, B., Katsuda, T., Grompe, M., Tan, K., & Stanger, B. Z. (2021). Dynamic transcriptional and epigenetic changes drive cellular plasticity in the liver. *Hepatology*, hep.31704. <https://doi.org/10.1002/hep.31704>
- Michalopoulos, G. K. (2007). Liver regeneration. *Journal of Cellular Physiology*, 213(2), 286–300. <https://doi.org/10.1002/jcp.21172>
- Michalopoulos, G. K., & Bhushan, B. (2021). Liver regeneration: Biological and pathological mechanisms and implications. *Nature Reviews Gastroenterology & Hepatology*, 18(1), 40–55. <https://doi.org/10.1038/s41575-020-0342-4>
- Miles, L. B., Dworkin, S., & Darido, C. (2017). Alternative splicing and start sites: Lessons from the Grainyhead-like family. *Developmental Biology*, 429(1), 12–19. <https://doi.org/10.1016/j.ydbio.2017.06.018>
- Millán-Zambrano, G., Burton, A., Bannister, A. J., & Schneider, R. (2022). Histone post-translational modifications—Cause and consequence of genome function. *Nature Reviews Genetics*, 23(9), 563–580. <https://doi.org/10.1038/s41576-022-00468-7>
- Miller, S. A., Mohn, S. E., & Weinmann, A. S. (2010). Jmjd3 and UTX Play a Demethylase-Independent Role in Chromatin Remodeling to Regulate T-Box Family Member-Dependent Gene Expression. *Molecular Cell*, 40(4), 594–605. <https://doi.org/10.1016/j.molcel.2010.10.028>
- Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G. P., & Eliopoulos, E. (2020). Transcription factors and evolution: An integral part of gene

expression (Review). *World Academy of Sciences Journal*, 2(1), 3–8.
<https://doi.org/10.3892/wasj.2020.32>

Miyajima, A., Tanaka, M., & Itoh, T. (2014). Stem/Progenitor Cells in Liver Development, Homeostasis, Regeneration, and Reprogramming. *Cell Stem Cell*, 14(5), 561–574. <https://doi.org/10.1016/j.stem.2014.04.010>

Montgomery, N. D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S. J., Otte, A. P., & Magnuson, T. (2005). The Murine Polycomb Group Protein Eed Is Required for Global Histone H3 Lysine-27 Methylation. *Current Biology*, 15(10), 942–947. <https://doi.org/10.1016/j.cub.2005.04.051>

Murphy, S. K., Yang, H., Moylan, C. A., Pang, H., Dellinger, A., Abdelmalek, M. F., Garrett, M. E., Ashley-Koch, A., Suzuki, A., Tillmann, H. L., Hauser, M. A., & Diehl, A. M. (2013). Relationship Between Methylome and Transcriptome in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology*, 145(5), 1076–1087. <https://doi.org/10.1053/j.gastro.2013.07.047>

Nagy, P., Thorgeirsson, S. S., & Grisham, J. W. (2020). Organizational Principles of the Liver. In I. M. Arias, H. J. Alter, J. L. Boyer, D. E. Cohen, D. A. Shafritz, S. S. Thorgeirsson, & A. W. Wolkoff (Eds.), *The Liver* (1st ed., pp. 1–13). Wiley. <https://doi.org/10.1002/9781119436812.ch1>

Nawrot, M., Peschard, S., Lestavel, S., & Staels, B. (2021). Intestine-liver crosstalk in Type 2 Diabetes and non-alcoholic fatty liver disease. *Metabolism*, 123, 154844. <https://doi.org/10.1016/j.metabol.2021.154844>

Nguyen-Lefebvre, A. T., & Horuzsko, A. (2016). *Kupffer Cell Metabolism and Function*.

Nicetto, D., & Zaret, K. S. (2019). *Role of H3K9me3 heterochromatin in cell identity establishment and maintenance*. 10.

Nielsen, C. B., Younesy, H., O'Geen, H., Xu, X., Jackson, A. R., Milosavljevic, A., Wang, T., Costello, J. F., Hirst, M., Farnham, P. J., & Jones, S. J. M. (2012). Spark: A navigational paradigm for genomic data exploration. *Genome Research*, 22(11), 2262–2269. <https://doi.org/10.1101/gr.140665.112>

Nishino, H., Takano, S., Yoshitomi, H., Suzuki, K., Kagawa, S., Shimazaki, R., Shimizu, H., Furukawa, K., Miyazaki, M., & Ohtsuka, M. (2017). Grainyhead-like 2 (GRHL2)

regulates epithelial plasticity in pancreatic cancer progression. *Cancer Medicine*, 6(11), 2686–2696. <https://doi.org/10.1002/cam4.1212>

Ogrodnik, M. (2021). Cellular aging beyond cellular senescence: Markers of senescence prior to cell cycle arrest in vitro and in vivo. *Aging Cell*, 20(4), e13338. <https://doi.org/10.1111/acer.13338>

Okólicsányi, L., & Magnenat, P. (1970). Transfer of bilirubin through the liver of Gunn rats. *Experientia*, 26(7), 733–734. <https://doi.org/10.1007/BF02232513>

Oliviero, G., Brien, G. L., Waston, A., Streubel, G., Jerman, E., Andrews, D., Doyle, B., Munawar, N., Wynne, K., Crean, J., Bracken, A. P., & Cagney, G. (2016). Dynamic Protein Interactions of the Polycomb Repressive Complex 2 during Differentiation of Pluripotent Cells*. *Molecular & Cellular Proteomics*, 15(11), 3450–3460. <https://doi.org/10.1074/mcp.M116.062240>

Ortega-Prieto, P., Parlati, L., Benhamed, F., Regnier, M., Cavalcante, I., Montabond, M., Onifarasoaniaina, R., Favier, M., Pavlovic, N., Magusto, J., Cauzac, M., Pagesy, P., Gautheron, J., Desdouets, C., Guilmeau, S., Issad, T., & Postic, C. (2023). O-GlcNAc transferase acts as a critical nutritional node for the control of liver homeostasis. *JHEP Reports*, 100878. <https://doi.org/10.1016/j.jhepr.2023.100878>

O'Shea, R. S., Dasarathy, S., McCullough, A. J., & Gastroenterology, P. G. C. of the A. A. for the S. of L. D. and the P. P. C. of the A. C. of. (2010). Alcoholic Liver Disease. *Hepatology*, 51(1), 307. <https://doi.org/10.1002/hep.23258>

Paris, J., & Henderson, N. C. (2022). Liver zonation, revisited. *Hepatology*, 76(4), 1219–1230. <https://doi.org/10.1002/hep.32408>

Pasini, D., Bracken, A. P., Jensen, M. R., Denchi, E. L., & Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *The EMBO Journal*, 23(20), 4061–4071. <https://doi.org/10.1038/sj.emboj.7600402>

Passman, A. M., Haughey, M. J., Carlotti, E., Williams, M. J., Cereser, B., Lin, M.-L., Devkumar, S., Gabriel, J. P., Gringeri, E., Cillo, U., Russo, F. P., Hoare, M., ChinAleong, J., Jansen, M., Wright, N. A., Kocher, H. M., Huang, W., Alison, M. R., & McDonald, S. A. C. (2023). Hepatocytes undergo punctuated expansion dynamics from a periportal stem cell niche in normal human liver. *Journal of Hepatology*, 79(2), 417–432. <https://doi.org/10.1016/j.jhepr.2023.03.044>

Paulusma, C. C., Lamers, W. H., Broer, S., & van de Graaf, S. F. J. (2022). Amino acid metabolism, transport and signalling in the liver revisited. *Biochemical Pharmacology*, 201, 115074. <https://doi.org/10.1016/j.bcp.2022.115074>

Pekowska, A., Benoukraf, T., Zacarias-Cabeza, J., Belhocine, M., Koch, F., Holota, H., Imbert, J., Andrau, J., Ferrier, P., & Spicuglia, S. (2011). H3K4 tri-methylation provides an epigenetic signature of active enhancers. *The EMBO Journal*, 30(20), 4198–4210. <https://doi.org/10.1038/emboj.2011.295>

Pepe-Mooney, B. J., Dill, M. T., Alemany, A., Ordovas-Montanes, J., Matsushita, Y., Rao, A., Sen, A., Miyazaki, M., Anakk, S., Dawson, P. A., Ono, N., Shalek, A. K., van Oudenaarden, A., & Camargo, F. D. (2019). Single-Cell Analysis of the Liver Epithelium Reveals Dynamic Heterogeneity and an Essential Role for YAP in Homeostasis and Regeneration. *Cell Stem Cell*, 25(1), 23-38.e8. <https://doi.org/10.1016/j.stem.2019.04.004>

Perez, G., Barber, G. P., Benet-Pages, A., Casper, J., Clawson, H., Diekhans, M., Fischer, C., Gonzalez, J. N., Hinrichs, A. S., Lee, C. M., Nassar, L. R., Raney, B. J., Speir, M. L., van Baren, M. J., Vaske, C. J., Haussler, D., Kent, W. J., & Haeussler, M. (2025). The UCSC Genome Browser database: 2025 update. *Nucleic Acids Research*, 53(D1), D1243–D1249. <https://doi.org/10.1093/nar/gkae974>

Pisarello, M. J. L., Loarca, L., Ivanics, T., Morton, L., & LaRusso, N. (2015). MicroRNAs in the Cholangiopathies: Pathogenesis, Diagnosis, and Treatment. *Journal of Clinical Medicine*, 4(9), Article 9. <https://doi.org/10.3390/jcm4091688>

Poisson, J., Lemoine, S., Boulanger, C., Durand, F., Moreau, R., Valla, D., & Rautou, P.-E. (2017). Liver sinusoidal endothelial cells: Physiology and role in liver diseases. *Journal of Hepatology*, 66(1), 212–227. <https://doi.org/10.1016/j.jhep.2016.07.009>

Primary Biliary Cholangitis | Concise Medical Knowledge. (n.d.). Retrieved 20 February 2025, from <https://www.lecturio.com/concepts/primary-biliary-cholangitis/>

Pu, W., Zhu, H., Zhang, M., Pikiolk, M., Ercan, C., Li, J., Huang, X., Han, X., Zhang, Z., Lv, Z., Li, Y., Liu, K., He, L., Liu, X., Heim, M. H., Terracciano, L. M., Tchorz, J. S., & Zhou, B. (2023). Bipotent transitional liver progenitor cells contribute to liver regeneration. *Nature Genetics*, 55(4), 651–664. <https://doi.org/10.1038/s41588-023-01335-9>

Puche, J. E., Saiman, Y., & Friedman, S. L. (2013). Hepatic Stellate Cells and Liver Fibrosis. *Comprehensive Physiology*, 3.

Quadrana, L., Bortolini Silveira, A., Mayhew, G. F., LeBlanc, C., Martienssen, R. A., Jeddelloh, J. A., & Colot, V. (2016). The *Arabidopsis thaliana* mobilome and its impact at the species level. *eLife*, 5, e15716. <https://doi.org/10.7554/eLife.15716>

R Core Team. (2023). *R: A Language and Environment for Statistical Computing* (Version 4.3.1) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, 44(W1), W160–W165. <https://doi.org/10.1093/nar/gkw257>

Raven, A., Lu, W.-Y., Man, T. Y., Ferreira-Gonzalez, S., O'Duibhir, E., Dwyer, B. J., Thomson, J. P., Meehan, R. R., Bogorad, R., Koteliansky, V., Kotelevtsev, Y., French-Constant, C., Boulter, L., & Forbes, S. J. (2017). Cholangiocytes act as facultative liver stem cells during impaired hepatocyte regeneration. *Nature*, 547(7663), 350–354. <https://doi.org/10.1038/nature23015>

Reid, M. A., Dai, Z., & Locasale, J. W. (2017). The impact of cellular metabolism on chromatin dynamics and epigenetics. *Nature Cell Biology*, 19(11), 1298–1306. <https://doi.org/10.1038/ncb3629>

Reiter, F., Wienerroither, S., & Stark, A. (2017). Combinatorial function of transcription factors and cofactors. *Current Opinion in Genetics & Development*, 43, 73–81. <https://doi.org/10.1016/j.gde.2016.12.007>

Rinella, M. E., Lazarus, J. V., Ratzliff, V., Francque, S. M., Sanyal, A. J., Kanwal, F., Romero, D., Abdelmalek, M. F., Anstee, Q. M., Arab, J. P., Arrese, M., Bataller, R., Beuers, U., Boursier, J., Bugianesi, E., Byrne, C. D., Narro, G. E. C., Chowdhury, A., Cortez-Pinto, H., ... Newsome, P. N. (2023). A multisociety Delphi consensus statement on new fatty liver disease nomenclature. *Journal of Hepatology*, 79(6), 1542–1556. <https://doi.org/10.1016/j.jhep.2023.06.003>

Romagnoli, M., Bresson, L., Di-Cicco, A., Pérez-Lanzón, M., Legoix, P., Baulande, S., de la Grange, P., De Arcangelis, A., Georges-Labouesse, E., Sonnenberg, A.,

- Deugnier, M.-A., Glukhova, M. A., & Faraldo, M. M. (2020). Laminin-binding integrins are essential for the maintenance of functional mammary secretory epithelium in lactation. *Development*, 147(4), dev181552. <https://doi.org/10.1242/dev.181552>
- Rotman, Y., & Kapuria, D. (2020). Non-alcoholic Fatty Liver Disease: Mechanisms and Treatment. In I. M. Arias, H. J. Alter, J. L. Boyer, D. E. Cohen, D. A. Shafritz, S. S. Thorgeirsson, & A. W. Wolkoff (Eds.), *The Liver* (1st ed., pp. 670–681). Wiley. <https://doi.org/10.1002/9781119436812.ch52>
- Sakamoto, K., Endo, K., Sakamoto, K., Kayamori, K., Ehata, S., Ichikawa, J., Ando, T., Nakamura, R., Kimura, Y., Yoshizawa, K., Masuyama, K., Kawataki, T., Miyake, K., Ishii, H., Kawasaki, T., Miyazawa, K., & Saitoh, M. (2021). EHF suppresses cancer progression by inhibiting ETS1-mediated ZEB expression. *Oncogenesis*, 10(3), Article 3. <https://doi.org/10.1038/s41389-021-00313-2>
- Saleh, M., Kamath, B. M., & Chitayat, D. (2016). Alagille syndrome: Clinical perspectives. *The Application of Clinical Genetics*, 9, 75–82. <https://doi.org/10.2147/TACG.S86420>
- Sancho-Bru, P., Altamirano, J., Rodrigo-Torres, D., Coll, M., Millán, C., José Lozano, J., Miquel, R., Arroyo, V., Caballería, J., Ginès, P., & Bataller, R. (2012). Liver progenitor cell markers correlate with liver damage and predict short-term mortality in patients with alcoholic hepatitis. *Hepatology*, 55(6), 1931–1941. <https://doi.org/10.1002/hep.25614>
- Satishchandran, A., Ambade, A., Rao, S., Hsueh, Y.-C., Iracheta-Vellve, A., Tornai, D., Lowe, P., Gyongyosi, B., Li, J., Catalano, D., Zhong, L., Kodys, K., Xie, J., Bala, S., Gao, G., & Szabo, G. (2018). MicroRNA 122, Regulated by GRLH2, Protects Livers of Mice and Patients From Ethanol-Induced Liver Disease. *Gastroenterology*, 154(1), 238–252.e7. <https://doi.org/10.1053/j.gastro.2017.09.022>
- Sato, K., Marzioni, M., Meng, F., Francis, H., Glaser, S., & Alpini, G. (2019). Ductular Reaction in Liver Diseases: Pathological Mechanisms and Translational Significances. *Hepatology*, 69(1), 420–430. <https://doi.org/10.1002/hep.30150>
- Schaub, J. R., Huppert, K. A., Kurial, S. N. T., Hsu, B. Y., Cast, A. E., Donnelly, B., Karns, R. A., Chen, F., Rezvani, M., Luu, H. Y., Mattis, A. N., Rougemont, A.-L., Rosenthal, P., Huppert, S. S., & Willenbring, H. (2018). De novo formation of the biliary

system by TGF β -mediated hepatocyte transdifferentiation. *Nature*, 557(7704), 247–251. <https://doi.org/10.1038/s41586-018-0075-5>

Schuster, R., Younesi, F., Ezzo, M., & Hinz, B. (2023). The Role of Myofibroblasts in Physiological and Pathological Tissue Repair. *Cold Spring Harbor Perspectives in Biology*, 15(1), a041231. <https://doi.org/10.1101/cshperspect.a041231>

Schuster, S., Cabrera, D., Arrese, M., & Feldstein, A. E. (2018). Triggering and resolution of inflammation in NASH. *Nature Reviews Gastroenterology & Hepatology*, 15(6), 349–364. <https://doi.org/10.1038/s41575-018-0009-6>

Schwabe, R. F., Tabas, I., & Pajvani, U. B. (2020). Mechanisms of Fibrosis Development in Nonalcoholic Steatohepatitis. *Gastroenterology*, 158(7), 1913–1928. <https://doi.org/10.1053/j.gastro.2019.11.311>

Sekiya, S., & Suzuki, A. (2012). Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. *The Journal of Clinical Investigation*, 122(11), 3914–3918. <https://doi.org/10.1172/JCI63065>

Senga, K., Mostov, K. E., Mitaka, T., Miyajima, A., & Tanimizu, N. (2012). Grainyhead-like 2 regulates epithelial morphogenesis by establishing functional tight junctions through the organization of a molecular network among claudin3, claudin4, and Rab25. *Molecular Biology of the Cell*, 23(15), 2845–2855. <https://doi.org/10.1091/mbc.E12-02-0097>

Shakeel, L., Shaukat, A., & Akilimali, A. (2025). Resmetirom: A Breakthrough in the Treatment of Metabolic Dysfunction–Associated Steatotic Liver Disease (MASLD). *Health Science Reports*, 8(6), e70920. <https://doi.org/10.1002/hsr2.70920>

Shao, C., Yang, X., Jing, Y., Hou, X., Huang, Y., Zong, C., Gao, L., Liu, W., Jiang, J., Ye, F., Shi, J., Zhao, Q., Li, R., Zhang, X., & Wei, L. (2021). The stemness of hepatocytes is maintained by high levels of lipopolysaccharide via YAP1 activation. *Stem Cell Research & Therapy*, 12(1), Article 1. <https://doi.org/10.1186/s13287-021-02421-7>

Shen, W., Chen, G., Dong, R., Zhao, R., & Zheng, S. (2014). MicroRNA-21/PTEN/Akt axis in the fibrogenesis of biliary atresia. *Journal of Pediatric Surgery*, 49(12), 1738–1741. <https://doi.org/10.1016/j.jpedsurg.2014.09.009>

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: A web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, 50(W1), W216–W221. <https://doi.org/10.1093/nar/gkac194>

Shu, W., Yang, M., Yang, J., Lin, S., Wei, X., & Xu, X. (2022). Cellular crosstalk during liver regeneration: Unity in diversity. *Cell Communication and Signaling : CCS*, 20, 117. <https://doi.org/10.1186/s12964-022-00918-z>

Shukla, V., Cetnarowska, A., Hyldahl, M., & Mandrup, S. (2022). Interplay between regulatory elements and chromatin topology in cellular lineage determination. *Trends in Genetics*, 38(10), 1048–1061. <https://doi.org/10.1016/j.tig.2022.05.011>

Sinegre, T., & Lebreton, A. (2014). Hémostase et cirrhose. *Revue Francophone Des Laboratoires*, 2014(463), 65–72. [https://doi.org/10.1016/S1773-035X\(14\)72525-0](https://doi.org/10.1016/S1773-035X(14)72525-0)

Si-Tayeb, K., Lemaigre, F. P., & Duncan, S. A. (2010). Organogenesis and Development of the Liver. *Developmental Cell*, 18(2), 175–189. <https://doi.org/10.1016/j.devcel.2010.01.011>

Slattery, M., Zhou, T., Yang, L., Dantas Machado, A. C., Gordân, R., & Rohs, R. (2014). Absence of a simple code: How transcription factors read the genome. *Trends in Biochemical Sciences*, 39(9), 381–399. <https://doi.org/10.1016/j.tibs.2014.07.002>

Slim, C. L., van IJzendoorn, S. C., Lázaro-Diéguéz, F., & Müsch, A. (2014). The special case of hepatocytes: Unique tissue architecture calls for a distinct mode of cell division. *BioArchitecture*, 4(2), 47–52. <https://doi.org/10.4161/bioa.29012>

Snykers, S., Henkens, T., De Rop, E., Vinken, M., Fraczek, J., De Kock, J., De Prins, E., Geerts, A., Rogiers, V., & Vanhaecke, T. (2009). Role of epigenetics in liver-specific gene transcription, hepatocyte differentiation and stem cell reprogramming. *Journal of Hepatology*, 51(1), 187–211. <https://doi.org/10.1016/j.jhep.2009.03.009>

So, J., Kim, A., Lee, S.-H., & Shin, D. (2020). Liver progenitor cell-driven liver regeneration. *Experimental & Molecular Medicine*, 52(8), 1230–1238. <https://doi.org/10.1038/s12276-020-0483-0>

Srivastava, A., Malik, L., Smith, T., Sudbery, I., & Patro, R. (2019). Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biology*, 20(1), Article 1. <https://doi.org/10.1186/s13059-019-1670-y>

- Stefan, N., Yki-Järvinen, H., & Neuschwander-Tetri, B. A. (2025). Metabolic dysfunction-associated steatotic liver disease: Heterogeneous pathomechanisms and effectiveness of metabolism-based treatment. *The Lancet Diabetes & Endocrinology*, 13(2), 134–148. [https://doi.org/10.1016/S2213-8587\(24\)00318-8](https://doi.org/10.1016/S2213-8587(24)00318-8)
- Stickel, F., & Hampe, J. (2012). Genetic determinants of alcoholic liver disease. *Gut*, 61(1), 150–159. <https://doi.org/10.1136/gutjnl-2011-301239>
- Strazzabosco, M., & Fabris, L. (2008). Functional Anatomy of Normal Bile Ducts. *The Anatomical Record*, 291(6), 653–660. <https://doi.org/10.1002/ar.20664>
- Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E., & Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*, 19(1), 477. <https://doi.org/10.1186/s12864-018-4772-0>
- Strick-Marchand, H., & Weiss, M. C. (2002). Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos: Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos. *Hepatology*, 36(4), 794–804. <https://doi.org/10.1053/jhep.2002.36123>
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43), 15545–15550. <https://doi.org/10.1073/pnas.0506580102>
- Sugimoto, A., Saito, Y., Wang, G., Sun, Q., Yin, C., Lee, K. H., Geng, Y., Rajbhandari, P., Hernandez, C., Steffani, M., Qie, J., Savage, T., Goyal, D. M., Ray, K. C., Neelakantan, T. V., Yin, D., Melms, J., Lehrich, B. M., Yasaka, T. M., ... Schwabe, R. F. (2025). Hepatic stellate cells control liver zonation, size and functions via R-spondin 3. *Nature*, 640(8059), 752–761. <https://doi.org/10.1038/s41586-025-08677-w>
- Sun, X., Chuang, J.-C., Kanchwala, M., Wu, L., Celen, C., Li, L., Liang, H., Zhang, S., Maples, T., Nguyen, L. H., Wang, S. C., Signer, R. A. J., Sorouri, M., Nassour, I., Liu, X., Xu, J., Wu, M., Zhao, Y., Kuo, Y.-C., ... Zhu, H. (2016). Suppression of the SWI/SNF Component *Arid1a* Promotes Mammalian Regeneration. *Cell Stem Cell*, 18(4), 456–466. <https://doi.org/10.1016/j.stem.2016.03.001>

- Sussman, J. H., Cure, H. W., Yuan, S., Ito, K., Asangani, I. A., Garcia, B. A., Stanger, B. Z., & Katsuda, T. (2025). In vivo CRISPR screening reveals epigenetic regulators of hepatobiliary plasticity. *Genes & Development*, genesdev;gad.352420.124v1. <https://doi.org/10.1101/gad.352420.124>
- Tabibian, J. H., Masyuk, A. I., Masyuk, T. V., O'Hara, S. P., & LaRusso, N. F. (2013). Physiology of Cholangiocytes. In Y. S. Prakash (Ed.), *Comprehensive Physiology* (1st ed., pp. 541–565). Wiley. <https://doi.org/10.1002/cphy.c120019>
- Tachmatzidi, E. C., Galanopoulou, O., & Talianidis, I. (2021). Transcription Control of Liver Development. *Cells*, 10(8), Article 8. <https://doi.org/10.3390/cells10082026>
- Taing, L., Dandawate, A., L'Yi, S., Gehlenborg, N., Brown, M., & Meyer, C. A. (2024). Cistrome Data Browser: Integrated search, analysis and visualization of chromatin data. *Nucleic Acids Research*, 52(D1), D61–D66. <https://doi.org/10.1093/nar/gkad1069>
- Tanaka, A. (2020). Current understanding of primary biliary cholangitis. *Clinical and Molecular Hepatology*, 27(1), 1. <https://doi.org/10.3350/cmh.2020.0028>
- Tanaka, M., Itoh, T., Tanimizu, N., & Miyajima, A. (2011). Liver stem/progenitor cells: Their characteristics and regulatory mechanisms. *Journal of Biochemistry*, 149(3), 231–239. <https://doi.org/10.1093/jb/mvr001>
- Tanimizu, N., & Mitaka, T. (2017). Epithelial Morphogenesis during Liver Development. *Cold Spring Harbor Perspectives in Biology*, 9(8), a027862. <https://doi.org/10.1101/cshperspect.a027862>
- Tanimizu, N., & Miyajima, A. (2004). Notch signaling controls hepatoblast differentiation by altering the expression of liver-enriched transcription factors. *Journal of Cell Science*, 117(15), 3165–3174. <https://doi.org/10.1242/jcs.01169>
- Tarlow, B. D., Pelz, C., Naugler, W. E., Wakefield, L., Wilson, E. M., Finegold, M. J., & Grompe, M. (2014). Bipotential Adult Liver Progenitors Are Derived from Chronically Injured Mature Hepatocytes. *Cell Stem Cell*, 15(5), 605–618. <https://doi.org/10.1016/j.stem.2014.09.008>
- Teixeira, F. K., & Colot, V. (2009). Gene body DNA methylation in plants: A means to an end or an end to a means? *The EMBO Journal*, 28(8), 997–998. <https://doi.org/10.1038/emboj.2009.87>

- Tessarz, P., & Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. *Nature Reviews Molecular Cell Biology*, 15(11), 703–708. <https://doi.org/10.1038/nrm3890>
- Thereaux, J., Lesuffleur, T., Czernichow, S., Basdevant, A., Msika, S., Nocca, D., Millat, B., & Fagot-Campagna, A. (2018). Association Between Bariatric Surgery and Rates of Continuation, Discontinuation, or Initiation of Antidiabetes Treatment 6 Years Later. *JAMA Surgery*, 153(6), 526–533. <https://doi.org/10.1001/jamasurg.2017.6163>
- Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178–192. <https://doi.org/10.1093/bib/bbs017>
- Tirnitz-Parker, J. E. E., Olynyk, J. K., & Ramm, G. A. (2014). Role of TWEAK in coregulating liver progenitor cell and fibrogenic responses. *Hepatology*, 59(3), 1198–1201. <https://doi.org/10.1002/hep.26701>
- Turnpenny, P. D., & Ellard, S. (2012). Alagille syndrome: Pathogenesis, diagnosis and management. *European Journal of Human Genetics*, 20(3), 251–257. <https://doi.org/10.1038/ejhg.2011.181>
- Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: Function, expression and evolution. *Nature Reviews Genetics*, 10(4), 252–263. <https://doi.org/10.1038/nrg2538>
- Verhulst, B., Neale, M. C., & Kendler, K. S. (2015). The heritability of alcohol use disorders: A meta-analysis of twin and adoption studies. *Psychological Medicine*, 45(5), 1061–1072. <https://doi.org/10.1017/S0033291714002165>
- Verma, S., Purohit, J. S., Arora, A., Sinha, S., & Chaturvedi, M. M. (2021). Liver regeneration: Metabolic and epigenetic regulation. *Hepatoma Research*, 7(0), N/A–N/A. <https://doi.org/10.20517/2394-5079.2020.122>
- Vincent, J. A., Kwong, T. J., & Tsukiyama, T. (2008). ATP-dependent chromatin remodeling shapes the DNA replication landscape. *Nature Structural & Molecular Biology*, 15(5), 477–484. <https://doi.org/10.1038/nsmb.1419>
- Wang, J., Zhao, Y., Zhou, X., Hiebert, S. W., Liu, Q., & Shyr, Y. (2018). Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. *BMC Genomics*, 19(1), 633. <https://doi.org/10.1186/s12864-018-5016-z>

- Wang, Y., Zeng, Z., Guan, L., & Ao, R. (2020). GRHL2 induces liver fibrosis and intestinal mucosal barrier dysfunction in non-alcoholic fatty liver disease via microRNA-200 and the MAPK pathway. *Journal of Cellular and Molecular Medicine*, 24(11), 6107–6119. <https://doi.org/10.1111/jcmm.15212>
- Warzecha, C. C., Jiang, P., Amirikian, K., Dittmar, K. A., Lu, H., Shen, S., Guo, W., Xing, Y., & Carstens, R. P. (2010). An ESRP-regulated splicing programme is abrogated during the epithelial–mesenchymal transition. *The EMBO Journal*, 29(19), 3286–3300. <https://doi.org/10.1038/emboj.2010.195>
- Watanabe, K., Liu, Y., Noguchi, S., Murray, M., Chang, J.-C., Kishima, M., Nishimura, H., Hashimoto, K., Minoda, A., & Suzuki, H. (2019). OVOL2 induces mesenchymal-to-epithelial transition in fibroblasts and enhances cell-state reprogramming towards epithelial lineages. *Scientific Reports*, 9(1), 6490. <https://doi.org/10.1038/s41598-019-43021-z>
- Weidemüller, P., Kholmatov, M., Petsalaki, E., & Zaugg, J. B. (2021). Transcription factors: Bridge between cell signaling and gene regulation. *PROTEOMICS*, 21(23–24), 2000034. <https://doi.org/10.1002/pmic.202000034>
- Weiler, P., Lange, M., Klein, M., Pe’er, D., & Theis, F. (2024). CellRank 2: Unified fate mapping in multiview single-cell data. *Nature Methods*, 21(7), 1196–1205. <https://doi.org/10.1038/s41592-024-02303-9>
- Wickham, H. (2009). Polishing your plots for publication. In H. Wickham (Ed.), *Ggplot2: Elegant Graphics for Data Analysis* (pp. 139–155). Springer. https://doi.org/10.1007/978-0-387-98141-3_8
- Wiesner, R., Edwards, E., Freeman, R., Harper, A., Kim, R., Kamath, P., Kremers, W., Lake, J., Howard, T., Merion, R. M., Wolfe, R. A., & Krom, R. (2003). Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology*, 124(1), 91–96. <https://doi.org/10.1053/gast.2003.50016>
- Wilkinson, A. C., Nakauchi, H., & Göttgens, B. (2017). Mammalian Transcription Factor Networks: Recent Advances in Interrogating Biological Complexity. *Cell Systems*, 5(4), 319–331. <https://doi.org/10.1016/j.cels.2017.07.004>

- Williams, M. J., Clouston, A. D., & Forbes, S. J. (2014). Links Between Hepatic Fibrosis, Ductular Reaction, and Progenitor Cell Expansion. *Gastroenterology*, 146(2), 349–356. <https://doi.org/10.1053/j.gastro.2013.11.034>
- Wilsnack, R. W., Wilsnack, S. C., Gmel, G., & Kantor, L. W. (2018). Gender Differences in Binge Drinking. *Alcohol Research : Current Reviews*, 39(1), 57–76.
- Xia, M., Wu, M., Li, Y., Liu, Y., Jia, G., Lou, Y., Ma, J., Gao, Q., Xie, M., Chen, Y., He, Y., Li, H., & Li, W. (2023). Varying mechanical forces drive sensory epithelium formation. *Science Advances*, 9(44), eadf2664. <https://doi.org/10.1126/sciadv.adf2664>
- Xu, J., Kausalya, P. J., Ong, A. G. M., Goh, C. M. F., Mohamed Ali, S., & Hunziker, W. (2022). ZO-2/Tjp2 suppresses Yap and Wwtr1/Taz-mediated hepatocyte to cholangiocyte transdifferentiation in the mouse liver. *Npj Regenerative Medicine*, 7(1), 1–13. <https://doi.org/10.1038/s41536-022-00251-6>
- Xu, T., Li, L., Hu, H.-Q., Meng, X.-M., Huang, C., Zhang, L., Qin, J., & Li, J. (2019). MicroRNAs in alcoholic liver disease: Recent advances and future applications. *Journal of Cellular Physiology*, 234(1), 382–394. <https://doi.org/10.1002/jcp.26938>
- Yan, Z.-J., Chen, L., & Wang, H.-Y. (2023). To be or not to be: The double-edged sword roles of liver progenitor cells. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1878(3), 188870. <https://doi.org/10.1016/j.bbcan.2023.188870>
- Yang, L., Wang, W., Qiu, W., Guo, Z., Bi, E., & Xu, C. (2017). A single-cell transcriptomic analysis reveals precise pathways and regulatory mechanisms underlying hepatoblast differentiation. *Hepatology*, 66(5), 1387–1401. <https://doi.org/10.1002/hep.29353>
- Yang, L., Wang, X., Yu, X.-X., Yang, L., Zhou, B.-C., Yang, J., & Xu, C.-R. (2023). The default and directed pathways of hepatoblast differentiation involve distinct epigenomic mechanisms. *Developmental Cell*, S1534580723003325. <https://doi.org/10.1016/j.devcel.2023.07.002>
- Yang, N., Occean, J. R., Melters, D. P., Shi, C., Wang, L., Stransky, S., Doyle, M. E., Cui, C.-Y., Delannoy, M., Fan, J., Slama, E., Egan, J. M., De, S., Cunningham, S. C., de Cabo, R., Sidoli, S., Dalal, Y., & Sen, P. (2023). A hyper-quiescent chromatin state formed during aging is reversed by regeneration. *Molecular Cell*. <https://doi.org/10.1016/j.molcel.2023.04.005>

Yimlamai, D., Christodoulou, C., Galli, G. G., Yanger, K., Pepe-Mooney, B., Gurung, B., Shrestha, K., Cahan, P., Stanger, B. Z., & Camargo, F. D. (2014). Hippo Pathway Activity Influences Liver Cell Fate. *Cell*, 157(6), 1324–1338. <https://doi.org/10.1016/j.cell.2014.03.060>

Younossi, Z. M., Golabi, P., de Avila, L., Paik, J. M., Srishord, M., Fukui, N., Qiu, Y., Burns, L., Afendy, A., & Nader, F. (2019). The global epidemiology of NAFLD and NASH in patients with type 2 diabetes: A systematic review and meta-analysis. *Journal of Hepatology*, 71(4), 793–801. <https://doi.org/10.1016/j.jhep.2019.06.021>

Yu, H., Jiang, X., Dong, F., Zhang, F., Ji, X., Xue, M., Yang, F., Chen, J., Hu, X., & Bao, Z. (2021). Lipid accumulation-induced hepatocyte senescence regulates the activation of hepatic stellate cells through the Nrf2-antioxidant response element pathway. *Experimental Cell Research*, 405(2), 112689. <https://doi.org/10.1016/j.yexcr.2021.112689>

Zeisel, M. B., Guerrieri, F., & Levrero, M. (2021). Host Epigenetic Alterations and Hepatitis B Virus-Associated Hepatocellular Carcinoma. *Journal of Clinical Medicine*, 10(8), Article 8. <https://doi.org/10.3390/jcm10081715>

Zeitlinger, J. (2020). Seven myths of how transcription factors read the cis-regulatory code. *Current Opinion in Systems Biology*, 23, 22–31. <https://doi.org/10.1016/j.coisb.2020.08.002>

Zeng, J., Zhang, J., Sun, Y., Wang, J., Ren, C., Banerjee, S., Ouyang, L., & Wang, Y. (2022). Targeting EZH2 for cancer therapy: From current progress to novel strategies. *European Journal of Medicinal Chemistry*, 238, 114419. <https://doi.org/10.1016/j.ejmech.2022.114419>

Zeybel, M., Hardy, T., Robinson, S. M., Fox, C., Anstee, Q. M., Ness, T., Masson, S., Mathers, J. C., French, J., White, S., & Mann, J. (2015). Differential DNA methylation of genes involved in fibrosis progression in non-alcoholic fatty liver disease and alcoholic liver disease. *Clinical Epigenetics*, 7(1), Article 1. <https://doi.org/10.1186/s13148-015-0056-6>

Zhang, C., Macchi, F., Magnani, E., & Sadler, K. C. (2021). Chromatin states shaped by an epigenetic code confer regenerative potential to the mouse liver. *Nature Communications*, 12(1), 4110. <https://doi.org/10.1038/s41467-021-24466-1>

- Zhang, H., Song, L., Wang, X., Cheng, H., Wang, C., Meyer, C. A., Liu, T., Tang, M., Aluru, S., Yue, F., Liu, X. S., & Li, H. (2021). Fast alignment and preprocessing of chromatin profiles with Chromap. *Nature Communications*, 12(1), 6566. <https://doi.org/10.1038/s41467-021-26865-w>
- Zheng, Z., Xu, X., Zhang, X., Wang, A., Zhang, C., Hüttemann, M., Grossman, L. I., Chen, L. C., Rajagopalan, S., Sun, Q., & Zhang, K. (2013). Exposure to ambient particulate matter induces a NASH-like phenotype and impairs hepatic glucose metabolism in an animal model. *Journal of Hepatology*, 58(1), 148–154. <https://doi.org/10.1016/j.jhep.2012.08.009>
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C., & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, 10(1), 1523. <https://doi.org/10.1038/s41467-019-09234-6>
- Zummo, F. P., Berthier, A., Gheeraert, C., Vinod, M., Bobowski-Gérard, M., Molendi-Coste, O., Pineau, L., Jung, M., Guille, L., Dubois-Chevalier, J., Dombrowicz, D., Staels, B., Eeckhoute, J., & Lefebvre, P. (2023). A time- and space-resolved nuclear receptor atlas in mouse liver. *Journal of Molecular Endocrinology*, 71(1). <https://doi.org/10.1530/JME-23-0017>

Annexes

I. Communications orales et affichées lors des congrès nationaux et internationaux dans le cadre de mon projet de thèse

1.1. Communication orale

- 24^{ème} Journée Andrée Verbert (JAV) 2024 de l'École doctorale de Lille – Lille, France – 25/11/2024 : « **Transcriptional control of hepato-cholangiocyte plasticity** »
- EGIR-MASLD 2024 (« *European Group for the study of Inuslino-resistance* ») – Lille, France – du 04/12/2024 au 06/12/2024 : « **Characterization of transcriptional factors induced upon hepatocyte reprogramming** »
- Café Scientifique du laboratoire INFINITE – Lille, France – 15/05/2025 : « **Transcriptional control of hepato-cholangiocyte plasticity** »
- EGID Annual Think-Tank – Lille, France – 30/06/2025 : « **Transcriptional control of hepato-cholangiocyte plasticity** »

1.2. Communication affichée

- ECCB 2024 ("European conference on computational biology") – Turku, Finlande – du 17/09/2024 au 20/09/2024 : « **Characterization of transcriptional factors induced upon hepatocyte reprogramming** »

II. Financements obtenus

- Lauréate d'une bourse de voyage pour assister au congrès ECCB 2024 (500€)
- Financement de fin de thèse auprès de la Fondation pour la Recherche Médicale (FRM) (23k) (FDT202504020162)

III. Article scientifique

Vasseur L, Gheeraert C*, Dubois-Chevalier J*, Very N, Guille L, Bou Saleh M, Boulet C, Sobolewski C, Loyer P, Berthier A, Legrand N, Corlu A, Gnemmi V, Lasailly G,

Leteurtre E, Galinously D, Bongiovanni A, Taront S, I Toft N, Grøntved L, Tulasne D, Furlan A, Ntandja-Wandji L, Staels B, Lefebvre P, Eeckhoutte J : « **Derepression of the epithelial transcription factor GRHL2 promotes direct hepatocyte-to-cholangiocyte transdifferentiation** », PLoS Biology, Dec 2025 (DOI: 10.1371/journal.pbio.3003547)

RESEARCH ARTICLE

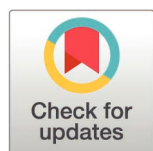
Derepression of the epithelial transcription factor GRHL2 promotes direct hepatocyte-to-cholangiocyte transdifferentiation

Ludvine Vasseur¹, Céline Gheeraert¹*, Julie Dubois-Chevalier¹*, Ninon Very¹, Loïc Guille¹, Mohamed Bou Saleh², Clémence Boulet¹, Cyril Sobolewski², Pascal Loyer³, Alexandre Berthier¹, Noémie Legrand², Anne Corlu³, Viviane Gnemmi⁴, Guillaume Lasailly², Emmanuelle Leteurtre⁴, Dmitry Galinowsky¹, Antonino Bongiovanni⁵, Solenne Taront², Nicolaj I. Toff⁶, Lars Grøntved⁶, David Tulasne⁷, Alessandro Furlan^{7,8}, Line Carolle Ntandja-Wandji², Bart Staels¹, Philippe Lefebvre¹, Laurent Dubuquoy², Jérôme Eeckhoutte^{1*}

1 Univ. Lille, Inserm, CHU Lille, Institut Pasteur de Lille, U1011-EGID, F-59000, Lille, France, **2** Univ. Lille, Inserm, CHU Lille, U1286 – INFINITE – Institute for Translational Research in Inflammation, F-59000, Lille, France, **3** Inserm, INRAE, Univ Rennes, Institut NUMECAN (Nutrition Métabolismes et Cancer) UMR-A 1341, UMR-S 1317, F-35000, Rennes, France, **4** Univ. Lille, CNRS, Inserm, CHU Lille, UMR9020-U1277 – CANTHER – Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000, Lille, France, **5** Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, US 41 – UAR 2014 – PLBS, Lille, France, **6** Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense M, Denmark, **7** Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, UMR9020 – U1277 – Canther – Cancer Heterogeneity, Plasticity and Resistance to Therapies, Lille, France, **8** Tumorigenesis and Resistance to Treatment Unit, Centre Oscar Lambret, F-59000, Lille, France

* These authors contributed equally to this work.

* jerome.eeckhoutte@inserm.fr



OPEN ACCESS

Citation: Vasseur L, Gheeraert C, Dubois-Chevalier J, Very N, Guille L, Bou Saleh M, et al. (2025) Derepression of the epithelial transcription factor GRHL2 promotes direct hepatocyte-to-cholangiocyte transdifferentiation. PLoS Biol 23(12): e3003547. <https://doi.org/10.1371/journal.pbio.3003547>

Academic Editor: Joo-Hyeon Lee, Memorial Sloan Kettering Cancer Center, UNITED STATES OF AMERICA

Received: January 24, 2025

Accepted: November 21, 2025

Published: December 12, 2025

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pbio.3003547>

Copyright: © 2025 Vasseur et al. This is an open access article distributed under the terms

Abstract

The liver's regenerative capacity is underscored by the plasticity potential of adult hepatocytes. In this context, hepatocyte-to-cholangiocyte transdifferentiation (HCT) has been ascribed with pro-regenerative functions in animal models and is a feature of end-stage human chronic liver diseases. While dampened activities of hepatocyte identity transcription factors (TFs) underlay HCT, how the cholangiocyte transcriptional program is implemented is poorly defined. Here, we identify that HCT does not involve transitioning through a hepatoblast-like transcriptional program. Furthermore, we show that HCT primarily involves induction of the archetypal transcriptional program of monopolarized epithelial cells initially repressed in hepatocytes. Indeed, HCT requires relieving H3K27me3-mediated and polycomb-dependent epigenetic silencing of epithelial TF encoding genes including *Grainyhead Like Transcription Factor 2* (GRHL2). Ectopic expression of GRHL2 in hepatocytes, including in vivo in the adult mouse liver, induces epithelial genes reminiscent of those activated during HCT. Finally, GRHL2 is detected in human hepatocytes undergoing HCT as evidenced using samples from end-stage chronic liver diseases. Hence, HCT is a process chiefly characterized by induction of a conventional epithelial transcriptional program

of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All individual numerical values that underlie originally generated data in this study data are available in the [S1 Data](#) file. All ChIP-seq and RNA-seq data generated in this study have been deposited into the Gene Expression Omnibus SuperSeries under accession number GSE281717. Codes and data related to integrated mouse liver single-cell RNA-seq data have been deposited into zenodo (<https://zenodo.org/records/13897656> and <https://zenodo.org/records/13897729>, respectively).

Funding: This work was supported by the Association Française pour l'Etude du Foie (AFEF to NV), Fondation de l'Université de Lille (to MBS), Fondation pour la Recherche Médicale (FRM grants EQU202203014645 to PL, EQU202003010299 to BS and EQU20200310299 to LD) and the Agence Nationale de la Recherche (ANR) grants "HSCreg" (ANR 21-CE14-0032-01 to JE) and "MEDICAL" (ANR 21 CE17-0016-02 to LD) and "European Genomic Institute for Diabetes" E.G.I.D. (ANR 10 LABX-0046), a French State fund managed by ANR under the frame program Investissements d'Avenir I-SITE ULNE/ANR-16-IDEX-0004 ULNE. This work also received support from Métropole Européenne de Lille (MEL) and Université de Lille (Investissements d'Avenir I-SITE ULNE/ANR-16-IDEX-0004 ULNE) (V-Chaire Industrielle-23-003 to JE). LV received support from EGID (ANR 10 LABX-0046), PreciDIAB (ANR-18-IBHU-0001) and Conseil Régional des Hauts-de-France (convention 22005973) as well as from Fondation Recherche Médicale (FDT202504020162). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: AAV3, adeno-associated viruses serotype 3; AAV8, adeno-associated viruses serotype 8; ALB, albumin; ALD, alcohol-related liver diseases; ALGS, Alagille syndrome; APAP, acetaminophen; BMEL, bipotential mouse embryonic liver; ChIP, chromatin immunoprecipitation; DAVID, database for annotation, visualization, and integrate

originally lacking in hepatocytes promoted by derepression of the master epithelial TF GRHL2.

Introduction

Hepatocytes, the predominant liver cell type, play a crucial role in the organ's homeostatic and secretory functions. Indeed, hepatocytes are key regulators of energy and iron homeostasis, and bear critical functions related to detoxification, coagulation, and the acute phase response. These unique hepatocyte functionalities are acquired during cellular differentiation, which is itself driven by establishment of a hepatocyte-specific transcriptional program. This phenomenon is orchestrated by transcriptional regulators, primarily hepatocyte identity transcription factors (Hep-ID TFs), which are hepatocyte-specific/enriched TFs [1–3].

Hepatocytes originate from hepatoblasts, which can alternatively differentiate into cholangiocytes. Cholangiocytes line the lumen of bile ducts and control bile volume and composition [4]. Together, hepatocytes and cholangiocytes constitute the two types of epithelial cells found in the liver [5]. However, unlike cholangiocytes, which form a conventional monolayer of cuboidal or columnar epithelial cells with a single apico-basal axis (hereafter defined as monopolarized epithelial cells), hepatocytes are characterized by multiple apical (and basal) surfaces, a unique feature among epithelia [6–8]. The liver possesses unique regenerative capacities where epithelial cell plasticity is instrumental while the existence of adult liver stem cells is highly disputed [9–11]. Indeed, the plasticity of adult liver epithelial cells allows them to reciprocally transdifferentiate [9,10]. In the mouse, this has been identified as a pro-regenerative mechanism operating when repair through proliferation of the injured cell type is not sufficient and/or impaired [9].

For example, in a mouse model mimicking the phenotype of Alagille syndrome (ALGS) with a deficient intrahepatic biliary system at birth, hepatocytes transdifferentiate to fully mature cholangiocytes to form functional bile ducts [12]. Lineage tracing has provided compelling evidence for transdifferentiation of liver epithelial cells in such conditions [12]. In humans, end stages of chronic liver diseases, such as alcohol-related liver diseases (ALD) [13] and metabolic dysfunction-associated steatotic liver disease (MASLD) [11], are characterized by the presence of cells co-expressing hepatocyte and cholangiocyte marker genes. In line, transcriptomic analyses of hepatocytes from humans with end-stage ALD and MASLD revealed instances of hepatocyte-to-cholangiocyte-like cell transdifferentiation [11,14]. HCT is also observed in livers from patients with primary sclerosing cholangitis (PSC) [15]. In this context, hepatocyte-derived cells contribute to the ductular reaction, a histologically defined phenomenon of hyperplasia of bile duct-like structures [16].

Despite its pathophysiological importance, HCT remains a poorly understood phenomenon. In particular, how HCT is orchestrated at the molecular level to induce a cholangiocyte-like transcriptional program remains ill-defined.

discovery; DDC, 3,5-diethoxycarbonyl-1,4-dihydrocollidine; DHS, DNase I Hypersensitivity; DMSO, dimethyl sulfoxide; DTT, dithiothreitol; GRHL2, Grainyhead Like Transcription Factor 2; GSEA, gene set enrichment analysis; HBE, human bronchial epithelial; HCT, hepatocyte-to-cholangiocyte transdifferentiation; Hep-ID TFs, hepatocyte identity transcription factors; MASLD, metabolic dysfunction-associated steatotic liver disease; MPH, mouse primary hepatocytes; NPC, non-parenchymal cells; nTPM, normalized gene expression in transcripts per million; PBC, primary biliary cholangitis; PBS, phosphate-buffered saline; PCA, principal component analysis; PFA, paraformaldehyde; PHx, partial hepatectomy; PRC2, polycomb repressive complex 2; PSC, primary sclerosing cholangitis; qPCR, Quantitative PCR; RT, room temperature; scRNA-seq, single-cell RNA-seq; SDS, sodium docetyl sulfate; TBS, Tris-Buffered Saline; TFs, transcription factors; TSS, transcriptional start sites; UMAP, uniform manifold approximation and projection.

Results

HCT occurs directly without transitioning through a hepatoblast-like transcriptional state

In several previous studies, HCT was proposed to involve dedifferentiation of hepatocytes into so-called bipotent liver progenitor-like cells [17–19]. To further compare gene expression changes involved in HCT with those defining developmental cholangiocyte differentiation from progenitors, we integrated single-cell RNA-seq (scRNA-seq) data obtained from embryonic, postnatal, and adult mouse livers [20,21]. In order to broadly capture the hepatocyte heterogeneity landscape, data from adult mouse livers comprised a wide range of injury models: acute exposure to acetaminophen (APAP) [22,23] or lipopolysaccharide [24], chronic exposure to 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) inducing cholestasis [25,26], western diet-induced metabolic dysfunction-associated steatohepatitis [27] and partial hepatectomy (PHx) [23,28,29] (S1 Table). Cells from the liver epithelial lineage were recovered from each dataset and batch correction was performed before visualization using uniform manifold approximation and projection (UMAP) (see [Materials and methods](#) for details). This approach resulted in the segregation of hepatoblasts from mature cholangiocytes or hepatocytes (Figs 1A and S1A). For instance, clustering analysis identified a cluster of cholangiocytes (denoted cluster 1; S1B Fig). Interestingly, this cluster was, however, connected to two other clusters comprising, on the one hand, hepatoblasts/postnatal epithelial cells (cluster 2) and, on the other hand, mostly adult hepatocytes (cluster 3) (Figs 1B and S1B). Interrogating the origin of hepatocytes from cluster 3 revealed that they almost exclusively stemmed from DDC-treated mice (95%; Fig 1C) including yellow fluorescent protein-labeled hepatocytes used for lineage tracing in this study [25] (S1C Fig). This is consistent with cluster 3 capturing HCT, which is known to be prominently induced in the chronic DDC-induced cholestasis model, most probably due to exhaustion of the cholangiocyte regenerative potential [9,30]. These data also indicated that HCT was unrelated to hepatocyte reprogramming occurring in other conditions such as APAP intoxication or PHx. This implies that induction of SOX9 in a subset of hepatocytes, commonly observed across these injury models (S1D Fig) and interpreted in the literature as hepatocytes acquiring biliary features [31], is in fact associated with different context-specific hepatocyte reprogramming events not necessarily involving transition towards a cholangiocyte-like cellular state. Using cells from clusters 1, 2, and 3, we performed RNA velocity analyses, which indicated that the developmental differentiation of cholangiocytes either from hepatoblasts or through HCT are independent (Figs 1D and S1E). In line, refined analysis of cellular trajectory and fate transition probability using CellRank2 [32] indicated that a subset of hepatocytes from DDC-treated mice (called HCT) was characterized by an extremely high probability to become cholangiocytes (Fig 1E and 1F). Further arguing against HCT involving hepatocytes to transiently adopt a progenitor-like transcriptional program, no hepatocytes from DDC-treated mice displayed significant fate probability towards hepatoblasts (Figs 1F and S1E). Finally, the fraction of variable genes showing peak expression at intermediate stages during HCT was limited (9%; S1F Fig). Moreover, these genes

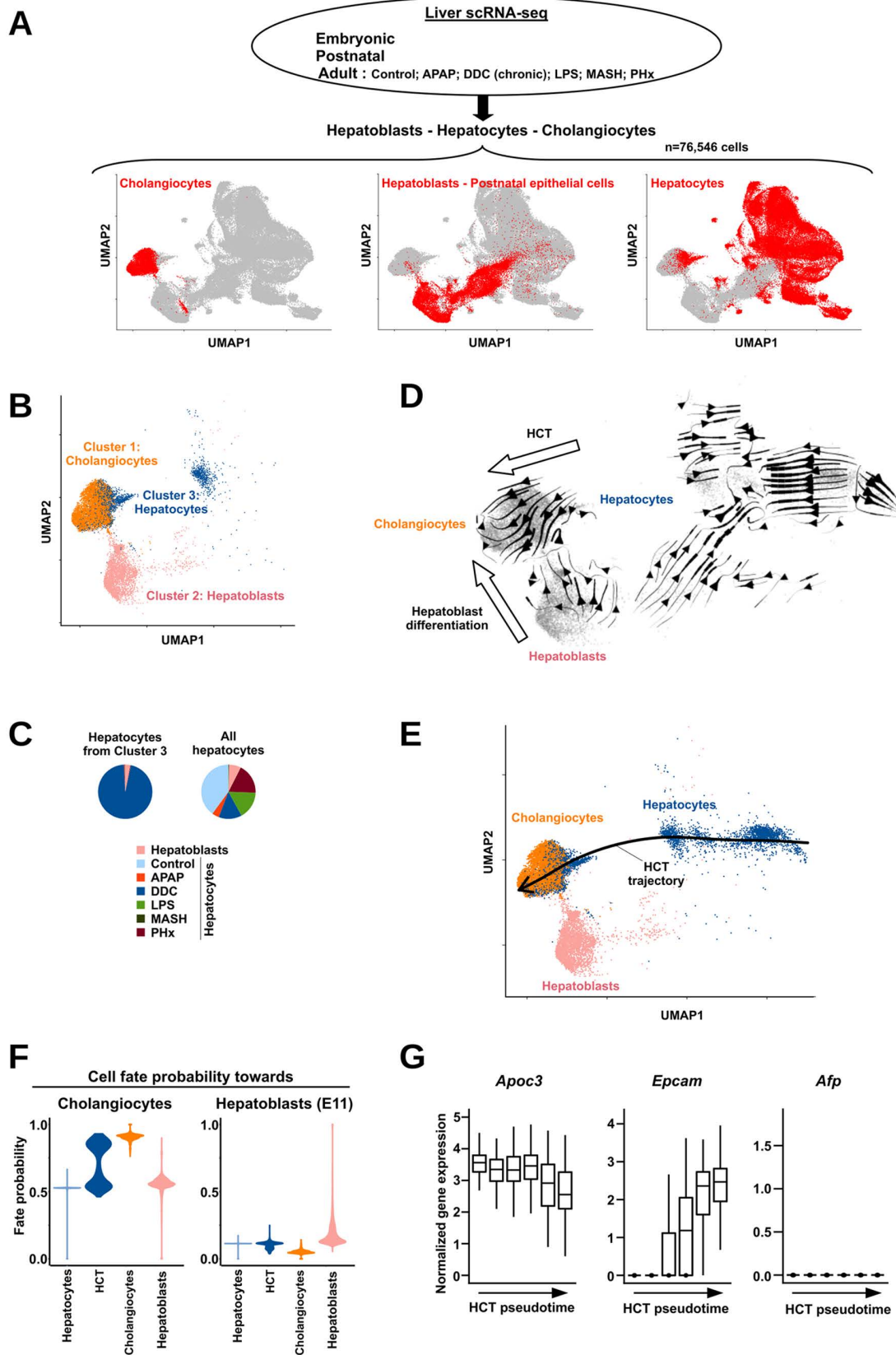


Fig 1. Characterization of HCT using a scRNA-seq atlas of mouse parenchymal liver cells. (A) Overview of datasets used to generate our mouse liver scRNA-seq atlas together with the UMAP projection of the obtained 76,546 individual cells and positioning of hepatoblasts (embryonic and postnatal cells), hepatocytes and cholangiocytes. Congruent expression of marker genes is shown in [S1A Fig](#). (B) Data from A were used for cell clustering ([S1B Fig](#)). Shown here is a zoom of the UMAP projection highlighting the 3 clusters selected to further study HCT vs. developmental cholangiocyte differentiation. The indicated main cell type in clusters 1–3 stems from analyses shown in [S1C Fig](#). (C) Pie charts showing the sample of origin of hepatocytes found in cluster 3 compared to all hepatocytes originally analyzed (panel A). (D) Results of RNA velocity analyses showing the cell transitions associated respectively with developmental differentiation of cholangiocytes from hepatoblasts and HCT. A magnification of the area containing cholangiocytes is shown in [S1E Fig](#). (E) HCT trajectory inferred from pseudotime analyses is shown on the UMAP from panel A. The indicated directionality is in line with the RNA velocity data shown in panel D. (F) Violin plots showing the cell fate transition probabilities of individual cells from the indicated groups towards adult cholangiocytes (left) or towards hepatoblasts (embryonic day 11; right) obtained using CellRank [142]. HCT refers to hepatocytes from mice chronically exposed to DDC. (G) Boxplots showing the normalized expression of the indicated marker genes in discrete hepatocyte cell subsets along the course of HCT as defined using pseudotime analysis (see [Materials and methods](#) for details).

<https://doi.org/10.1371/journal.pbio.3003547.g001>

were enriched for biological pathways shared with both hepatocytes and cholangiocytes ([S1G Fig](#)) and were not characterized by heightened expression in hepatoblasts ([S1H Fig](#)). Indeed, hepatocytes in intermediate stages in the HCT process still express hepatocyte markers [e.g., *Apoc3* (Apolipoprotein C-III) or *Alb* (Albumin)] and start expressing cholangiocyte markers [e.g., *Epcam* (*Epithelial cell adhesion molecule*)] while lacking expression of typical progenitor genes [e.g., *Afp* (*Alpha fetoprotein*)] ([Figs 1G and S1I](#)).

Altogether, these data indicate that HCT occurs through a process, which does not involve transient acquisition of a progenitor-like cell state.

HCT is mainly characterized by induction of the transcriptional signature of monopolarized epithelial cells in hepatocytes

In order to characterize the transcriptional reprogramming underlying HCT, we compared the transcriptome of lineage-traced hepatocytes from DDC-treated mice with that from healthy hepatocytes and cholangiocytes obtained using bulk RNA-seq of sorted cells (see [Materials and methods](#)) [25,33]. When projected onto a space defined by the first two principal components from a principal component analysis (PCA) conducted on the transcriptomes of hepatocytes and cholangiocytes, hepatocytes undergoing HCT localized near cholangiocytes ([Figs 2A and S2A](#)). As expected, these data therefore verified that hepatocytes undergoing HCT in the liver of DDC-treated mice have a transcriptome similar to that of cholangiocytes. In line, gene set enrichment analysis (GSEA) indicated that genes characterized by enriched expression in cholangiocytes among the different liver cell-types ($n=57$ obtained from [34] and denoted as Chol-enriched genes; [S2 Table](#)) were significantly biased towards up-regulated genes during HCT ([Fig 2B](#)). To unbiasedly assess the nature of the transcriptional reprogramming, enrichment analyses for biological or molecular pathways and their clustering by similarity was next performed using the database for annotation, visualization, and integrate discovery (DAVID) [35]. In line with a transdifferentiation process, genes down-regulated in hepatocytes undergoing HCT, when compared to healthy hepatocytes, were enriched for terms related to normal hepatocyte functions ([S2B Fig](#)). When up-regulated genes were considered, enriched terms pointed to archetypal epithelial cell features such as cilium and tight junctions ([Figs 2C and S2C](#)). GSEA further indicated that genes characterizing the human epithelial cell lineage identified in [36] ($n=837$; hereafter denoted as Epith-Signature. See [S2 Table](#)) displayed significantly greater expression in hepatocytes undergoing HCT ([Fig 2D](#)). Similar results were obtained when using genes characteristic of the monopolar cell polarity ([S2D Fig](#)). Moreover, genes from the Epith-Signature were enriched in the subset of HCT up-regulated genes characterized by low basal expression in healthy hepatocytes and strong induction in HCT (p -value < 0.0001 , chi-squared test; [Fig 2E](#)). Among those genes were commonly used cholangiocyte markers such as *Krt7* (*Keratin 7*), *Krt19* (*Keratin 19*), and *Epcam* ([Fig 2E](#); over 60% of the Chol-enriched genes are comprised within the Epith-Signature). Consistent with these findings, comparing expression of genes from the Epith-Signature in different human cells from the epithelial lineage obtained from the Human Protein Atlas [37] indicated that, unlike cholangiocytes, hepatocytes were characterized by low expression levels

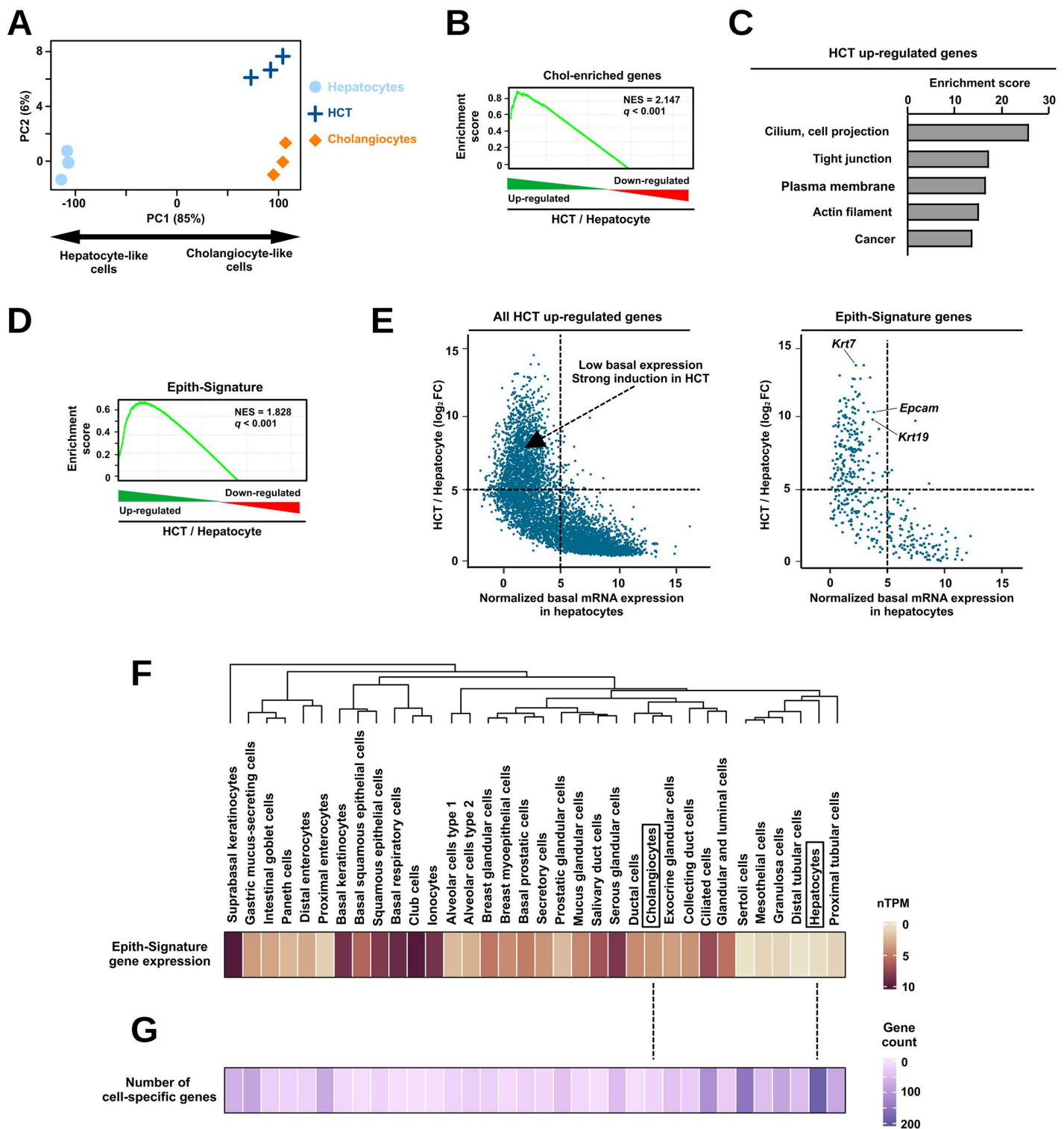


Fig 2. Activation of a *bona fide* epithelial gene transcriptional program characterizes HCT. (A) Transcriptomic changes induced in HCT were assessed by projecting datasets obtained from lineage-traced hepatocytes undergoing HCT from DDC-treated mice (denoted HCT) within a 2-dimensional space generated through a PCA performed using the transcriptome of healthy mouse cholangiocytes and hepatocytes (see the [Materials and methods](#) section for details). As indicated, the first principal component (PC1) is sufficient to separate hepatocyte-like from cholangiocyte-like cells as indicated. (B)

Enrichment plot obtained using GSEA performed with the Chol-enriched genes [34] as the gene set and transcriptomic changes induced by HCT obtained by comparing bulk RNA-seq data from lineage-traced hepatocytes undergoing HCT from DDC-treated mice compared to that of control hepatocytes from healthy livers from [25]. NES stands for normalized enrichment score. (C) Functional annotation of genes up-regulated during HCT (q -value < 0.05; top 3,000 genes were further selected) performed using the DAVID tool [35]. Deregulated genes were obtained by mining data from [25] as described for panel B. Enriched terms were clustered by the DAVID tool and were named according to words repetitively used in the name of individual enriched terms within each cluster (S2C Fig). Top 5 clusters are displayed together with their enrichment scores (i.e., best enrichment score among individual terms within each cluster). (D) Enrichment plot obtained using GSEA performed with the Epith-Signature [36] as the gene set and transcriptomic changes induced by HCT defined as in panel B from [25]. NES stands for normalized enrichment score. (E) HCT up-regulated genes (left) or genes from the Epith-Signatures (right) were plotted according to their basal expression levels in healthy hepatocytes and their induction upon HCT. Dotted lines were used to highlight a subset of genes displaying low basal expression and strong induction during HCT. Commonly used cholangiocyte markers (i.e., *Krt7*, *Krt19*, *Epcam*) are individually highlighted. (F) Expression of genes from the Epith-Signature [36] was used to perform hierarchical clustering of the indicated individual human epithelial cell types (S2E Fig). The dendrogram issued from these analyses is displayed on the left of the heatmap, which shows the median expression of genes from the Epith-Signature in each individual epithelial cell types. Cholangiocytes and hepatocytes are highlighted using frames. nTPM stands for normalized transcripts per million. (G) Heatmap showing the number of genes specifically expressed in a given cell type among all epithelial cell types as defined using a stringent tau specificity index cut-off, i.e., $\tau > 0.9$. Ranking of epithelial cell types is similar to that of panel F.

<https://doi.org/10.1371/journal.pbio.3003547.g002>

(Figs 2F and S2E). Further indication that the transcriptome of hepatocytes differs from that of other epithelial cells, including cholangiocytes, was provided when mining the number of genes with high expression selectivity among individual epithelial cells [i.e., genes with a tau specificity index > 0.9 when all individual epithelial cell types are considered [38]. Indeed, hepatocytes were characterized by the greatest number of specifically expressed genes, i.e., with low expression in the other individual epithelial cells, while cholangiocytes had very few (Fig 2G).

Altogether, these data indicate that a main event underlying HCT is induction of a conventional monopolarized epithelial cell gene signature, which is typically absent in healthy hepatocytes.

Epithelial lineage TFs are repressed by the polycomb repressive complex 2 in hepatocytes

Mining transcriptional changes underlying HCT using a procedure similar to Fig 2A but applied only to TF-encoding genes generated similar results (Figs 3A and S3A). This revealed that changes in TF-encoding gene expression were fully informative with regards to capturing HCT, which prompted us to further characterize deregulated TFs as a mean to define drivers of HCT. Calling differentially expressed genes (q -value < 0.05) in hepatocytes that underwent HCT compared to healthy ones identified 115 down- and 423 up-regulated TF-encoding genes, respectively (S3 Table). As expected, down-regulated TFs comprised many Hep-ID TFs (S3B Fig) consistent with loss of hepatocyte identity upon HCT. With regards to up-regulated TFs, 82% of them were also significantly induced in cholangiocyte-like cells derived from hepatocytes in the mouse model of ALGS lacking an intrahepatic biliary system [12] (Fig 3B), consistent with a role for these identified up-regulated TFs in HCT.

The up-regulated TFs displayed low basal expression levels in healthy hepatocytes, and were also the most strongly induced upon HCT (Fig 3C). These data suggest that a subset of up-regulated TFs might be originally repressed in hepatocytes. We therefore next characterized the promoters of up-regulated TFs in healthy hepatocytes by monitoring levels of chromatin accessibility [DNase I Hypersensitivity (DHS)] as well as levels of modified histones prototypically linked to gene activation [histone H3 lysine 4 trimethylation (H3K4me3) and lysine 27 acetylation (H3K27ac)], transcription (H3K36me3 and RNA Polymerase II) or repression (H3K27me3). We used DHS-seq and ChIP-seq data obtained from healthy mouse livers [hepatocytes overwhelmingly contribute to chromatin-based signals obtained using whole liver [39–41] or isolated mouse primary hepatocytes (MPH) (S4 Table). Using the Spark software [42], we identified 3 main clusters of promoters with divergent epigenetic patterns (Fig 3D). The largest cluster (C1) comprised promoters already active in normal liver with accessible chromatin and strong H3K4me3 and H3K27ac signals, while C2 and C3 gradually displayed features of inactive but transcriptionally competent or repressed promoters, respectively. In particular, promoters within C2 displayed moderate levels of both H3K4me3 and H3K27me3, i.e., potential bivalent promoters [43], while promoters in C3 were characterized by lack of any feature of gene activation concomitant with strong repressive H3K27me3 signals (Figs 3D and S4A). Mining single-nucleus ATAC sequencing data [44] confirmed the differences in

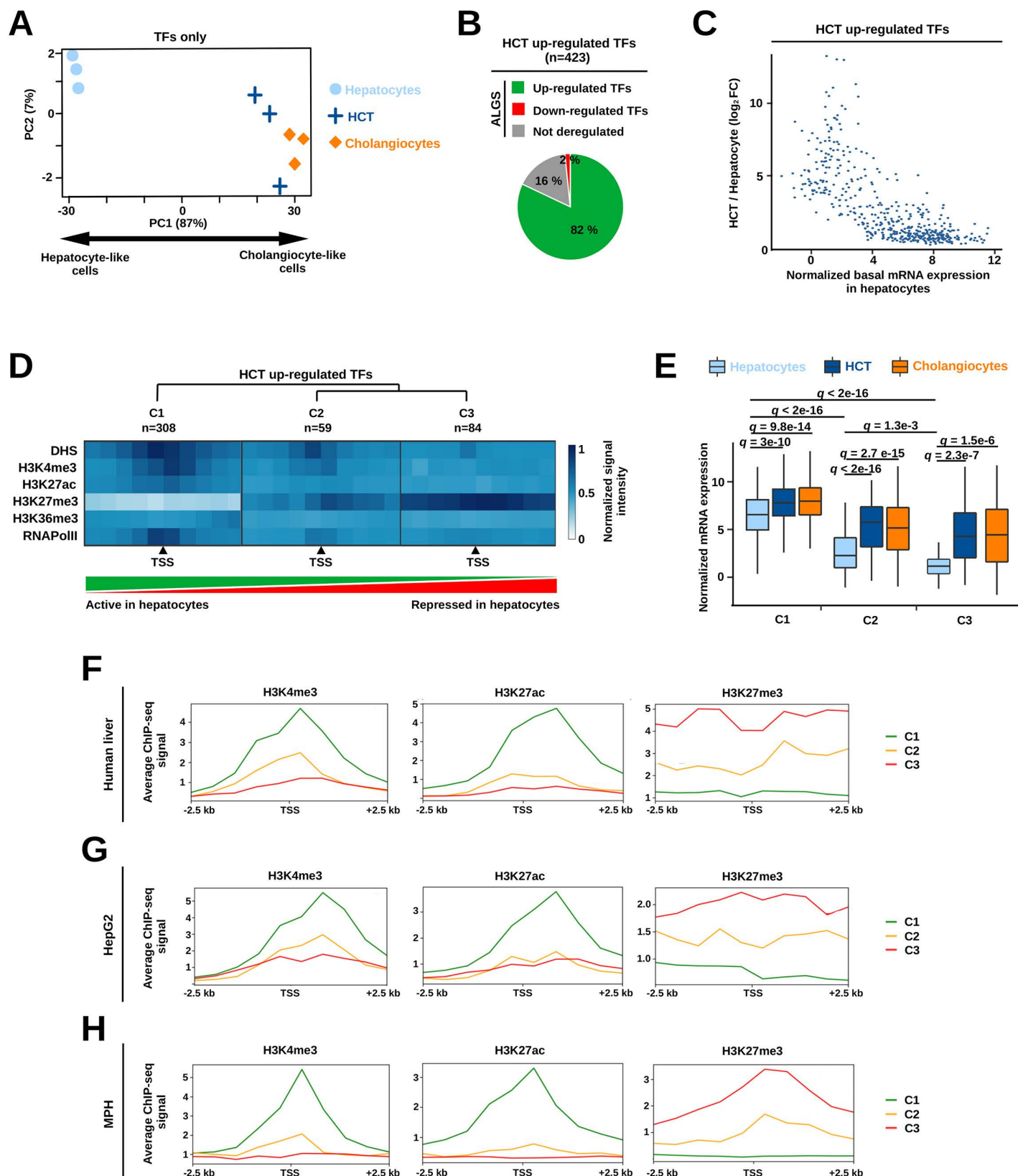


Fig 3. HCT involves induction of TF encoding genes epigenetically repressed in normal hepatocytes. (A) Analyses performed as in Fig 2A except only TF-encoding genes were considered. (B) TF-encoding genes up-regulated during HCT were obtained by comparing bulk RNA-seq data from lineage-traced hepatocytes undergoing HCT from DDC-treated mice compared to that of control hepatocytes from healthy livers from [25] (q -value < 0.05 and $\log_2 FC > 0$). The obtained gene list was monitored for changes in hepatocytes undergoing HCT in another mouse model, i.e., model of ALGS lacking an intrahepatic biliary system at birth [12]. In this second dataset, genes with q -value < 0.15 and $\log_2 FC < 0$ or > 0 were defined as down- or up-regulated, respectively. Genes with q -value > 0.15 were considered as “Not deregulated”. (C) Induction of TF-encoding genes upon HCT was compared to their basal expression levels in healthy mouse hepatocytes. A scatter plot showing individual TF-encoding genes is shown. (D) Average signal for chromatin accessibility (DHS) and indicated modified histones or RNA polymerase II (RNAPolII) at TF-encoding gene promoters for the main 3 clusters identified by Spark [42]. The number of TF-encoding genes comprised in each cluster (C1 to C3) is indicated. Dendrogram of hierarchical clustering is also shown. The arrowheads indicate the position of the transcriptional start site (TSS) at the center of the 5kb windows used for these analyses. (E) Box plot showing expression of TF-encoding genes from clusters C1–3 in mouse healthy hepatocytes or cholangiocytes as well as hepatocytes undergoing HCT from [25]. Statistical significance was assessed using two-way ANOVA with Dunnett’s multiple comparison post hoc test. (F–H) Average ChIP-Seq signals for indicated modified histones at promoters from clusters C1–3 (defined in panel D) in MPH (E), human liver (F), or HepG2 cells (G). Low signal was arbitrarily set to 1. Signals are shown in windows of ± 2.5 kb around the genes TSS.

<https://doi.org/10.1371/journal.pbio.3003547.g003>

promoter accessibility between the 3 clusters and did not reveal any difference between periportal and pericentral hepatocytes (S4B Fig). In line with the chromatin status of their promoters, TF-encoding genes from C2, and even more so from C3, displayed low basal expression in mouse hepatocytes but showed greater induction during HCT when compared to those from C1 (Fig 3E). Adding ChIP-seq data for the heterochromatin mark H3K9me3 to Spark did not significantly modify the clustering of up-regulated TF promoters (S5A Fig). In line, genes from C3 were not found within mouse liver chromatin domains primarily characterized by the presence of H3K9me3 [45] (S5B Fig). Consistently, promoters from C3 did not display DNA methylation, which is often found at H3K9me3 repressed genes [46] (S5C Fig). Biophysically identified heterochromatin regions in mouse hepatocytes, defined through sequencing of sonication-resistant heterochromatin [47], indicated that C3 promoters were found both in euchromatin and heterochromatin (S5D Fig). These findings further emphasize the role of H3K27me3 in repressing genes of alternative cell fates in these two chromatin compartments, independently of H3K9me3 [47,48].

We next aimed to define whether observations in mice are consistent with promoter epigenetic patterns found in humans using ChIP-seq data for H3K4me3, H3K27ac and H3K27me3 from human healthy livers [49] and human HepG2 cells [this study and [50]; S4 Table]. We found epigenetic patterns in human livers (Fig 3F) and HepG2 cells (Fig 3G) reminiscent of those observed in mouse hepatocytes (Fig 3D and 3H), which was in line with greater mRNA expression of C1 genes (S6A and S6B Fig). These data both verified the robustness of our clusters of TF-encoding gene promoters and indicated its human relevance.

Deposition of H3K27me3 is under the control of the polycomb repressive complex 2 (PRC2) and more particularly of the enzymes EZH1 and 2 (Enhancer Of Zeste 1 and 2 Polycomb Repressive Complex 2 Subunits) [51,52]. In line, mining ChIP-seq data from [53] (S4 Table) indicated that presence of H3K27me3 at TF-encoding gene promoters in mouse liver is linked to stronger recruitment of EZH2, especially for promoters from C3 (Fig 4A). Moreover, *Ezh1/2* deletion in mouse livers [47] (S4 Table) triggered both a decrease in H3K27me3 (Fig 4B) at the C2 and C3 genes and their transcriptional induction (Fig 4C). This translated into induction of the Epith-Signature and Chol-enriched gene sets (Fig 4D). Finally, treatment of HepG2 cells with pharmacological inhibitors of EZH1/2 (UNC1999 and valemestostat tosylate together referred to as EZHi), which decreased H3K27me3 levels (Fig 4E), also triggered concomitant derepression of TF-encoding genes from C2 and C3 as measured by RNA-seq analysis (Fig 4F).

Altogether, these data indicate that HCT involves the induction of TF-encoding genes normally repressed through polycomb-dependent H3K27me3 in hepatocytes.

Forced expression of the epithelial lineage TF GRHL2 promotes HCT

We next aimed to demonstrate a functional role for up-regulated TFs in HCT. In addition to being repressed in normal hepatocytes and highly induced upon HCT, TF-encoding genes from C3 displayed the highest tau lineage-specificity

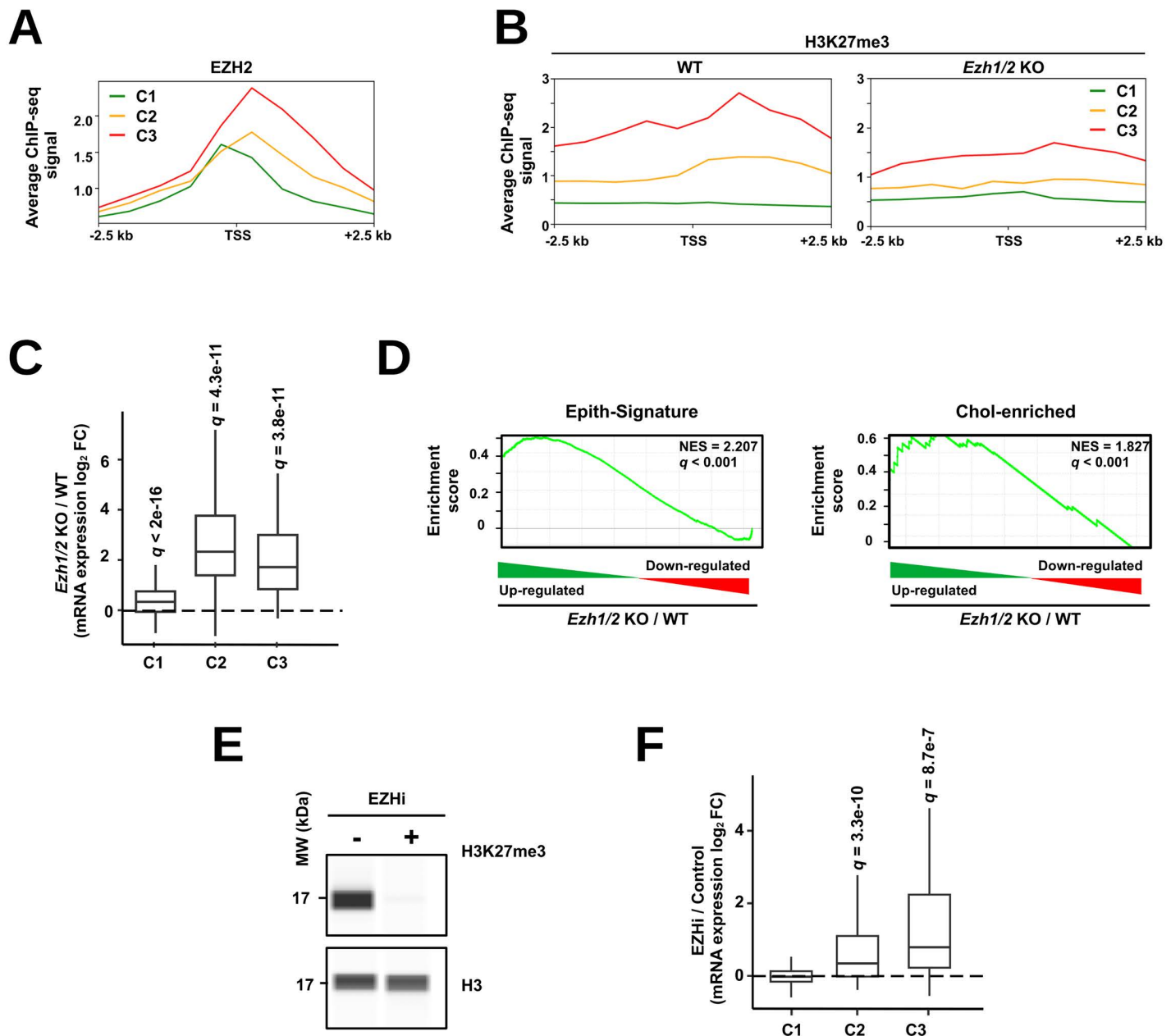


Fig 4. Polycomb-mediated repression of TF encoding genes up-regulated during HCT. (A) Average EZH2 ChIP-Seq signal at promoters from clusters C1–3 (defined in Fig 3) in the mouse liver. Low signal was arbitrarily set to 1. Signal is shown in a window of ± 2.5 kb around the genes TSS. (B) Average H3K27me3 ChIP-Seq signal at promoters from clusters C1–3 (defined in Fig 3) in the liver of wild-type (WT) mice or mice deleted for the *Ezh1* and *Ezh2* genes (*Ezh1/2* KO). Signal is shown in a window of ± 2.5 kb around the genes TSS. (C) Box plot showing transcriptional modulation of TF-encoding genes expression for clusters C1–3 (defined in Fig 3) obtained from RNA-seq performed on the livers of mice with deletion of both the *Ezh1* and *Ezh2* genes (*Ezh1/2* KO; $n=3$) and wild-type mice (WT; $n=2$) as the control [47]. One-sided one-sample Wilcoxon t test with Benjamini–Hochberg correction was used to determine if the median log₂ FC was significantly above 0, i.e., any statistically significant increase in gene expression between livers of *Ezh1/2* ko and WT mice. (D) Enrichment plots obtained using GSEA performed with the Epith-Signature [36] or Chol-enriched genes [34] as the gene sets and transcriptomic changes induced by *Ezh1/2* deletion in the mouse liver (similar to panel C). NES stands for normalized enrichment score. (E) Immunoblotting performed using the Wes system to monitor levels of H3K27me3 and H3 in HepG2 cells treated or not with EZH1/2 inhibitors (EZHi). Data are representative of those obtained in 3 independent biological replicates. MW, molecular weight. (F) Box plot showing transcriptional modulation of TF-encoding genes expression for clusters C1–3 (defined in Fig 3) obtained from RNA-seq performed on HepG2 cells treated with EZH1/2 inhibitors or 0.2% DMSO as the control ($n=4$ biological independent replicates). One-sided one sample Wilcoxon t test with Benjamini–Hochberg correction was used to determine if the median log₂ FC was significantly above 0, i.e., any statistically significant increase in gene expression between EZHi-treated and control cells. The original data underlying this figure can be found at the Gene Expression Omnibus (GSE281717).

<https://doi.org/10.1371/journal.pbio.3003547.g004>

indexes calculated from gene expression in 13 human cell lineages (including the epithelial lineage) obtained from [37] (Fig 5A). We therefore reasoned that C3 comprises key TFs involved in the establishment of the alternative *bona fide* epithelial transcriptional program characterizing HCT. We therefore monitored individual TFs from C3 by simultaneously comparing: *i*) their expression enrichment in epithelial cells (ranking based on highest to lowest expression levels among aforementioned human cell lineages); *ii*) their level of induction upon HCT and *iii*) their presence in the Epith-Signature and Chol-enriched gene set. The TFs Grainyhead Like Transcription Factor 2 (GRHL2) and EHF (ETS Homologous Factor) emerged from these analyses as they were enriched in epithelial cells and cholangiocytes and strongly induced in HCT (Fig 5B). Interestingly, these TFs, together with OVOL1/2 (Ovo Like Transcriptional Repressors 1 and 2), had previously been suspected to inhibit epithelial–mesenchymal transition in cancer cells [54,55] with a hierarchically dominant position for GRHL2, which directly induces expression of EHF and OVOL2 [56,57] (Fig 5B). Moreover, GRHL2 is considered a master regulator of epithelial cell differentiation [58,59]. In line with our previous characterization of genes from C3, *Grhl2* displays H3K27me3-mediated repression in hepatocytes (Fig 5C). We next set out to experimentally define if GRHL2 could induce expression of genes reminiscent of those activated upon HCT in bipotential mouse embryonic liver (BMEL) cells, which have the ability to differentiate both in hepatocytes and cholangiocytes [60,61]. Ectopic expression of GRHL2 (Figs 5D and S7A Fig) triggered induction of a cholangiocyte-like transcriptional program. Indeed, RNA-seq data mining using biological pathway enrichment analyses indicated that expressing GRHL2 in BMEL cells increased the expression of genes related to cell fate commitment including epithelial cell differentiation (Figs 5E and S7B and S5–S6 Tables). In line, GSEA pointed to stimulation of expression of the Epith-Signature and Chol-enriched genes (Fig 5F). Accordingly, when BMEL were grown in three-dimensional Matrigel, GRHL2 promoted acquisition of cell aggregates with a cystic organization (Figs 5G, 5H, and S8) typically adopted by cholangiocytes [62,63].

To more directly assess the ability of GRHL2 to promote HCT, we performed similar experiments in the HepG2 hepatic cell line (Figs 6A and S7C). Although the global reprogramming of the transcriptome was less pronounced than in BMEL cells (S7B and S7D Fig and S7 and S8 Tables), GRHL2 again promoted expression of an archetypal epithelial transcriptional program as judged both through biological pathway enrichment analyses (Fig 6B) and GSEA using the Epith-Signature and Chol-enriched gene sets (Figs 6C and S7E). This was not associated with a decrease in hepatocyte-enriched gene expression (S7F Fig and S2 Table). To further characterize how GRHL2 induces transcriptional reprogramming, we defined its cistrome using chromatin immunoprecipitation (ChIP)-seq in HepG2 cells. Importantly, the GRHL2 binding sites we identified in HepG2 cells (top 10,000 peaks) also displayed recruitment of endogenous GRHL2 previously defined in primary human bronchial epithelial (HBE) cells [64] or an ovarian epithelial cancer cell line [65–67] (S4 Table), albeit with lower enrichment (Fig 6D). This indicated that, while forced TF expression can lead to spurious chromatin binding and gene regulation, chromatin binding of ectopically expressed GRHL2 in hepatocytes recapitulates that of the endogenously expressed protein [64,68]. Moreover, the CistromeGO tool [69] indicated that GRHL2 peaks from HepG2 cells were enriched near genes related to terms consistent with transcriptional remodeling towards an archetypal epithelial cell program such as cell adhesion and junction (Fig 6E). We next inspected GRHL2 ChIP-seq data at genes selected on the basis of their induction upon HCT and GRHL2 ectopic expression as well as their presence in the Epith-Signature and/or cholangiocyte-enriched gene list. In addition to the previously identified GRHL2 target genes *EHF*, *CLDN4* (*Claudin 4*) and *RAB25* (*RAB25, Member RAS Oncogene Family*) [56,62], these analyses also pointed to direct regulation of genes such as *ESRP1* (*Epithelial splicing regulatory protein 1*), *ITGA3* (*Integrin Subunit Alpha 3*), *ITGB8* (*Integrin Subunit Beta 8*), *JUP* (*Junction Plakoglobin*), and *KRT7* (Fig 6F and 6G and S9 Table). These genes have a proven role in establishment and maintenance of epithelial cell identity and epithelium integrity [62,70–74]. The sensitivity of these genes to GRHL2-mediated up-regulation was further supported by their induction upon moderate ectopic expression of GRHL2 (S9 Fig). As an additional model of human hepatocytes, we used differentiated HepaRG cells [75,76]. GRHL2 overexpression in these cells (Fig 6H) also triggered induction of *CLDN4*, *EHF*, *JUP*, *KRT7*, and *RAB25* (Fig 6I). We also induced GRHL2 expression in human organoids issued from EPCAM-positive cells expanded

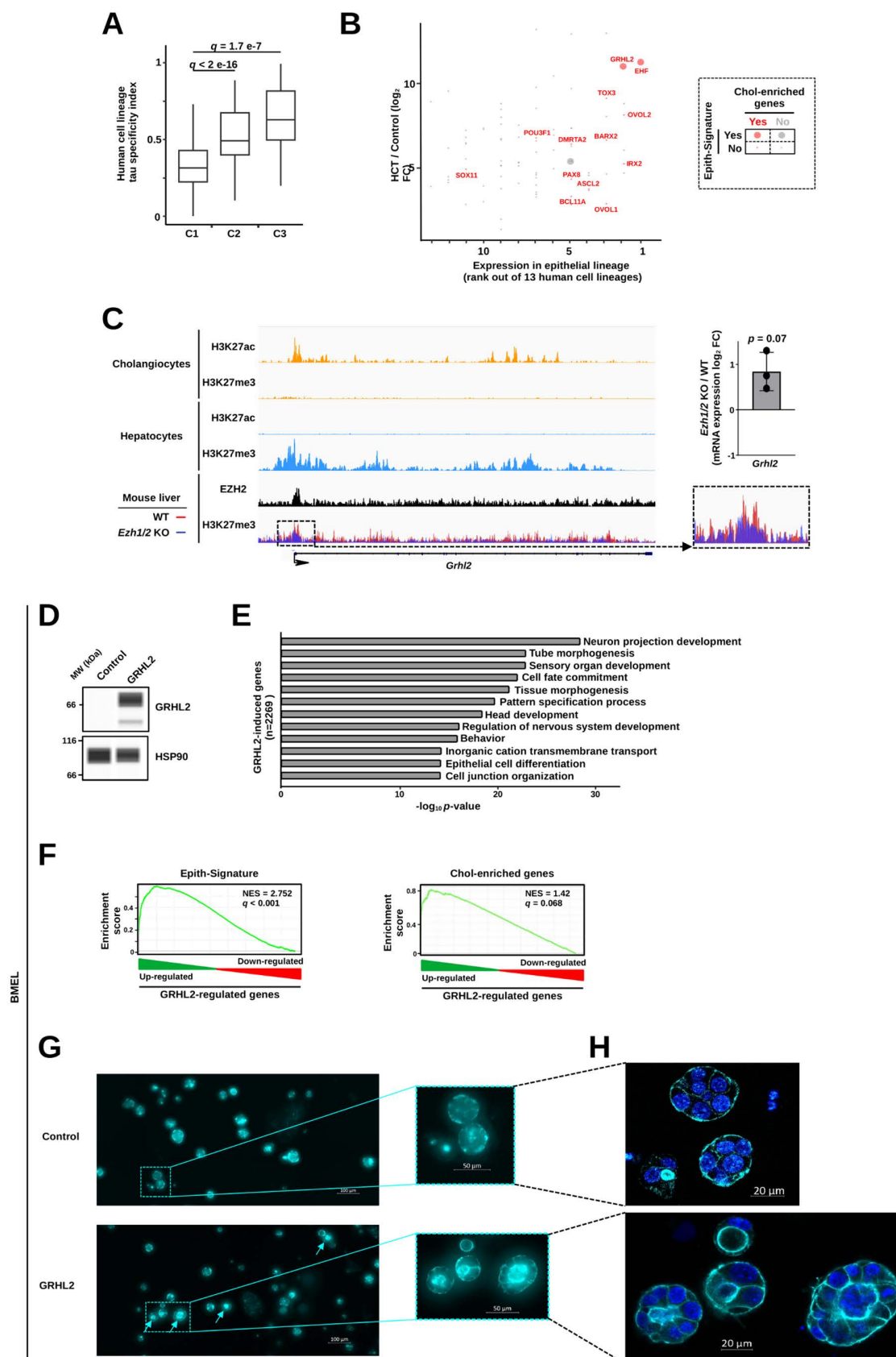


Fig 5. HCT involves induction of the master epithelial lineage TF GRHL2. (A) Box plot showing the tau specificity indexes of TF-encoding genes from clusters C1–3 defined in Fig 3. Thirteen human cell lineages (epithelial cells and adipocytes, blood and immune cells, endocrine cells, endothelial cells, germinal cells, glial cells, mesenchymal cells, muscle cells, neuronal cells, pigment cells, trophoblast cells and undifferentiated cells) were used to compute the tau specificity index using the *tspec* tool [136]. Gene expression in a given cell lineage was defined as the median expression of individual cell types comprised in that lineage. Statistical significance was assessed using a Kruskal–Wallis test with Dunn’s multiple comparison post hoc test. (B) Scatter plot of individual TF-encoding genes from cluster C3 (as defined in Fig 3) showing their expression in the epithelial lineage (based on data from panel A; lineages were ranked from highest to lowest expression levels for each gene) and their induction levels upon HCT (similar to those used in Fig 3). Additionally, presence of individual TF-encoding genes within the Epith-Signature was highlighted using large dots while identification as a cholangiocyte-enriched gene among liver cell types was highlighted using red dots. (C) Integrated Genomics Viewer (IGV) [141] was used to visualize the indicated ChIP-seq profiles from mouse primary cholangiocytes, hepatocytes or liver at the *Grhl2* gene. H3K27ac and H3K27me3 ChIP-seq data from cholangiocytes and hepatocytes were processed similarly and visualized on the same scale in the two cell types. H3K27me3 ChIP-seq signals in the mouse liver of WT and *Ezh1/2* KO mice were overlaid and a magnification of the *Grhl2* promoter area is shown on the right together with a bar graph showing changes in *Grhl2* expression in the liver these animals computed from RNA-seq data [47]. Mean \pm SD together with individual biological replicates are shown. Two-sided one-sample *t* test was used to determine if the mean of \log_2 FC was statistically different from 0. (D) Immunoblotting performed using the Wes system to monitor GRHL2 and HSP90 levels in BMEL cells transfected with a GRHL2 expression plasmid or an empty control construct. Data are representative of those obtained in 3 independent biological replicates. MW, molecular weight. (E) Bar plot showing biological term enrichments (top 12) within genes upregulated upon GRHL2 ectopic expression in BMEL cells, i.e., genes with \log_2 FC > 0 and *q*-value < 0.05 as defined through RNA-seq analyses on 4 independent biological replicates. Data were obtained using Metascape [126]. (F) Enrichment plots obtained using GSEA performed with the Epith-Signature [36] or Chol-enriched genes [34] as the gene sets and transcriptomic changes induced by GRHL2 ectopic expression in BMEL cells (RNA-seq data from 4 independent biological replicates used in panel E). NES stands for normalized enrichment score. (G) BMEL cells were transfected with a GRHL2 expression plasmid or an empty control construct and grown in Matrigel. Representative images of obtained aggregates stained with phalloidin are shown. The arrows point to cystic cell aggregates. (H) Cell aggregates highlighted in the zoomed areas from panel G were further analyzed using confocal microscopy in order to better display the differences in their organization highlighting the cyst structures adopted by BMEL cells overexpressing GRHL2. Images were acquired using LSM confocal microscopy. To smooth the images, LSM image processing was applied, and due to the low signal in control images, a median filter was also used to improve image quality. The original data underlying this figure can be found at the Gene Expression Omnibus (GSE281717).

<https://doi.org/10.1371/journal.pbio.3003547.g005>

in three-dimensional Matrigel and induced to differentiate according to [77]. We found that this procedure gives rise to a population of cells expressing both hepatocyte and cholangiocyte markers [78]. GRHL2 triggered moderate induction of several of its target genes concomitant with reduced expression of the *ALB* and *CYP3A4* (*Cytochrome P450 family 3 subfamily A member 4*) hepatocytes markers (Fig 6J).

Finally, the impact of forced GRHL2 expression in mouse hepatocytes in vivo was assessed following hydrodynamic tail vein injection of the GRHL2-encoding plasmid (Figs 7A and S10A) or injection of recombinant adeno-associated viruses (AAV8) (Figs 7B and S10B). Reminiscent of our in vitro findings, we observed that GRHL2 induces expression of the cholangiocyte-enriched epithelial genes *Cldn4*, *Ehf*, *Esrp1*, *Itga3*, *Itgb8*, *Jup*, *Krt7*, and *Rab25* when expressed in hepatocytes in vivo (Fig 7C and 7D). Concomitantly, expression of hepatocyte marker genes such as *Alb*, *Gsta3* (*Glutathione S-transferase, alpha 3*) or *Fmo5* (*Flavin containing monooxygenase 5*) was reduced (Fig 7E). Expression of the hepatic bile acid exporter *Abcb11* (*ATP-binding cassette, sub-family B member 11*; *Bsep*) was also diminished while that of *Cyp3a11* was increased, changes characterizing mice with altered hepatic bile acid metabolism [79,80]. While the liver architecture was not affected after 1 week (S10C Fig), dramatic histological alterations to the hepatocyte parenchyma, characterized by gained heterogeneity in cell shape and size (Figs 7F and S10C), and to immune cell infiltration (S10D Fig) was observed after 2 weeks.

Taken together, these results indicate that GRHL2 promotes direct HCT through activation of genes characteristic of an archetypal epithelial cell program.

Induction of GRHL2 in hepatocytes undergoing HCT in human livers

We and others have reported that livers of patients with end-stage chronic liver diseases are characterized by populations of hepatocytes displaying signs of HCT, i.e., concomitant loss of hepatocyte identity gene expression and induction of cholangiocyte markers [11,13–15,81,82]. To assess GRHL2 expression and its link to HCT in human livers, we first used samples from patients with ALD-related decompensated cirrhosis. Fragments of healthy liver

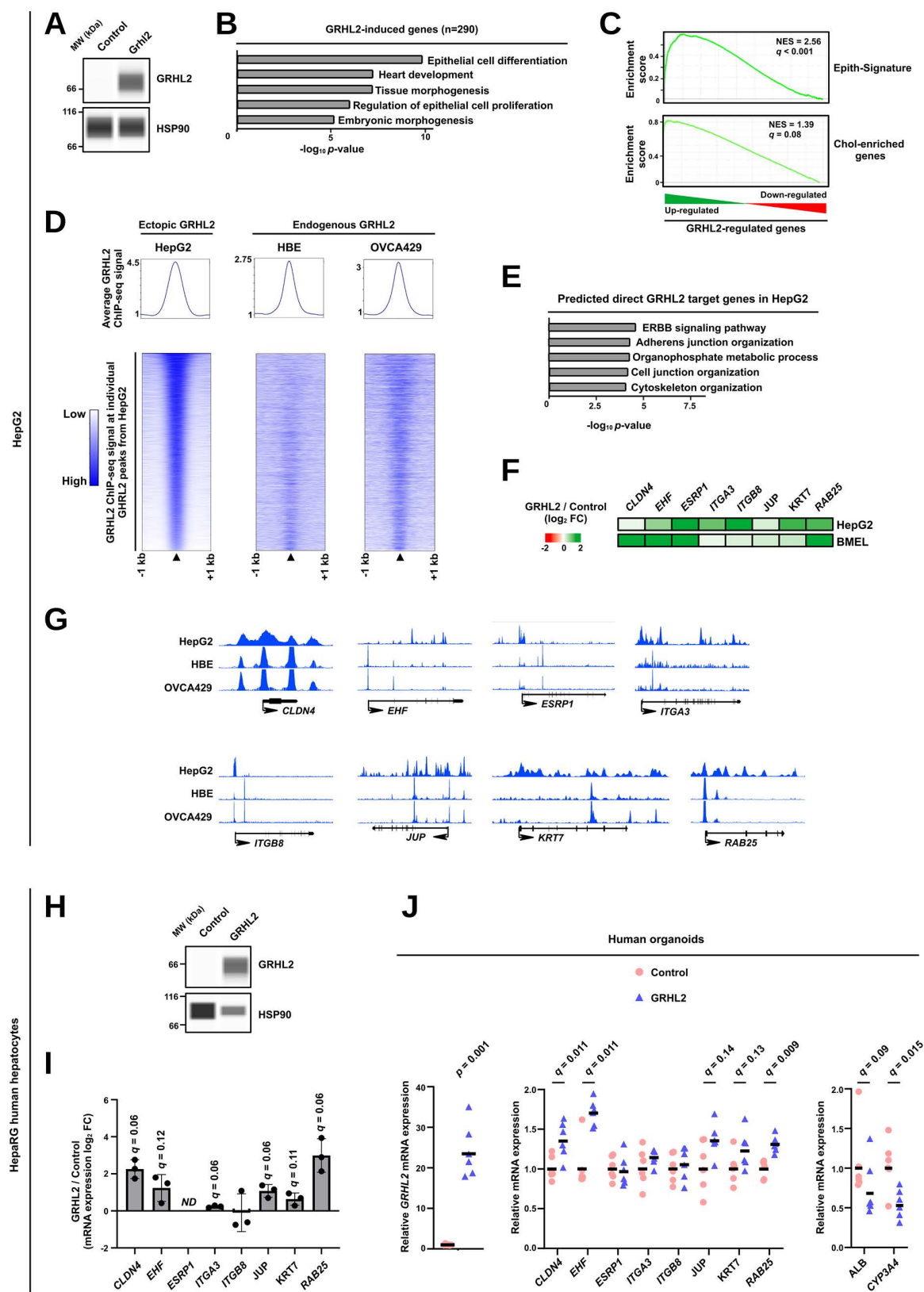


Fig 6. GRHL2 ectopic expression in hepatocytes triggers induction of epithelial genes involved in HCT. (A) Immunoblotting performed using the Wes system to monitor GRHL2 and HSP90 levels in HepG2 cells transfected with a GRHL2 expression plasmid or an empty control construct. Data are representative of those obtained in 3 independent biological replicates. MW, molecular weight. (B) Bar plot showing biological term enrichments (top 5) within genes upregulated upon GRHL2 ectopic expression in HepG2 cells, i.e., genes with \log_2 FC > 0 and q -value < 0.05 as defined through RNA-seq analyses on 4 independent biological replicates. Data were obtained using Metascape [126]. (C) Enrichment plot obtained using GSEA performed with the Epith-Signature [36] or Chol-enriched genes [34] as the gene sets and transcriptomic changes induced by GRHL2 ectopic expression in HepG2 cells (RNA-seq data from 4 independent biological replicates used in panel B). NES stands for normalized enrichment score. (D) The top 10,000 GRHL2 binding sites identified in HepG2 cells were used to monitor GRHL2 ChIP-seq signal from the indicated cells. Human bronchial epithelial (HBE) are primary human bronchial epithelial cells while OVCA429 is an ovarian epithelial cancer cell line. Both average ChIP-seq signal (top; background signal was arbitrarily set to 1) and signal at individual binding sites (heatmap at the bottom) are shown in 2 kb regions centered on GRHL2 peaks. Individual binding sites were ranked according to the GRHL2 ChIP-seq signal from HepG2 cells. GRHL2 binding sites and ChIP-seq signal in HepG2 cells were defined from 4 independent biological replicates (see Materials and methods). (E) Bar plot showing biological term enrichments (top 5) within genes predicted to be directly targeted by GRHL2 in HepG2 cells by the CistromeGO tool [69] on the basis of identified GRHL2 binding sites. (F) Heatmap showing changes in expression of the indicated genes in the RNA-seq data obtained from HepG2 or BMEL (Fig 5) cells ectopically expressing GRHL2. (G) Integrated Genome Browser (IGB) [140] was used to visualize GRHL2 ChIP-seq profiles from HepG2, HBE and OVCA429 (same datasets as those used in panel D) at the indicated target genes (*CLDN4*, *EHF*, *ESRP1*, *ITGA3*, *ITGB8*, *JUP*, *KRT7*, and *RAB25*). (H) Immunoblotting performed using the Wes system to monitor GRHL2 and HSP90 levels in HepaRG cells differentiated into hepatocytes and electroporated with a GRHL2 expression plasmid or an empty control construct. Data are representative of those obtained in 4 independent biological replicates. MW, molecular weight. (I) RT-qPCR data showing the expression of the indicated GRHL2 target genes in HepaRG cells differentiated into hepatocytes and electroporated with a GRHL2 expression plasmid or an empty control construct ($n=3$). \log_2 FC in gene expression in cells expressing GRHL2 compared to control cells (transfected with an empty plasmid) are shown in the bar graph, which displays means \pm SD together with individual biological replicates. Two-sided one-sample t test with Benjamini–Hochberg correction was used to determine if the mean of \log_2 FC was statistically different from 0. ND stands for not detected. The numerical values of all biological replicates can be found in the S1 Data file. (J) RT-qPCR data showing the relative mRNA expression of *GRHL2* (left), its target genes (middle) or hepatocyte markers (right) in human organoids transduced with GRHL2 or GFP encoding AAV3 ($n=6$). Individual data together with the geometric means (horizontal bars) are shown. Fold changes were obtained using the average of the control group, arbitrarily set to 1 for each analyzed gene, as the reference. One-sided Mann–Whitney U test with Benjamini–Hochberg correction for multiple testing was performed to determine whether gene expression was significantly induced upon GRHL2 ectopic expression. The numerical values of all biological replicates can be found in the S1 Data file. The original data underlying this figure can be found at the Gene Expression Omnibus (GSE281717).

<https://doi.org/10.1371/journal.pbio.3003547.g006>

samples from patients subjected to liver resection due to hepatic tumors were used as control in these experiments. First, we found that GRHL2 expression was induced in patients with liver failure both at the mRNA and protein levels (Fig 8A and 8B). This was associated with decreased EZH2 and H3K27me3 levels in those livers (Fig 8B). To define that GRHL2 was specifically induced in hepatocytes, we monitored its mRNA levels in RNA-seq-based transcriptomic data previously obtained using microdissection of hepatocyte-enriched areas from those livers [14]. Compared to control samples, GRHL2 expression was induced in ALD-related liver failure (\log_2 FC = 1.38, q -value = 0.11), which correlated with the induction of genes defined as up-regulated during HCT as well as genes from the Epith-Signature (Fig 8C). Moreover, unlike control tissues which stained positive for GRHL2 in cholangiocytes only, GRHL2 immunostaining in livers from ALD patients was positive not only in neoductules but also in hepatocytes found at the border of the regenerative nodules next to the fibrotic bands (Figs 8D, S11A, and S11B). Immunofluorescence co-staining of GRHL2 and albumin (ALB) showed a gradual increase in GRHL2 together with a gradual loss of ALB staining in hepatocytes bordering the fibrotic scar (S11C Fig). Additional co-staining for KRT7 further revealed triple-positive cells at the scar interface (Figs 8E, S11D, and S11E), in line with HCT occurring in this specific micro-environment in human cirrhotic livers [83]. Single-nuclei RNA-seq data analysis of human livers at different stages of MASLD [11] showed increased GRHL2-expressing hepatocytes in end-stages of the disease (i.e., cirrhosis and liver failure) when compared to healthy control livers (S12A–S12D Fig). Moreover, GRHL2 expression was found in hepatocytes acquiring mixed expression of hepatocyte and cholangiocyte identity genes consistent with its HCT-promoting activities (S12E Fig). Finally, GRHL2 immunostaining in samples from patients with liver failure due to primary biliary cholangitis (PBC) or PSC also revealed GRHL2-positive hepatocytes (Figs 8F and S13).

Overall, these data indicate that induction of GRHL2 characterizes HCT in human livers.

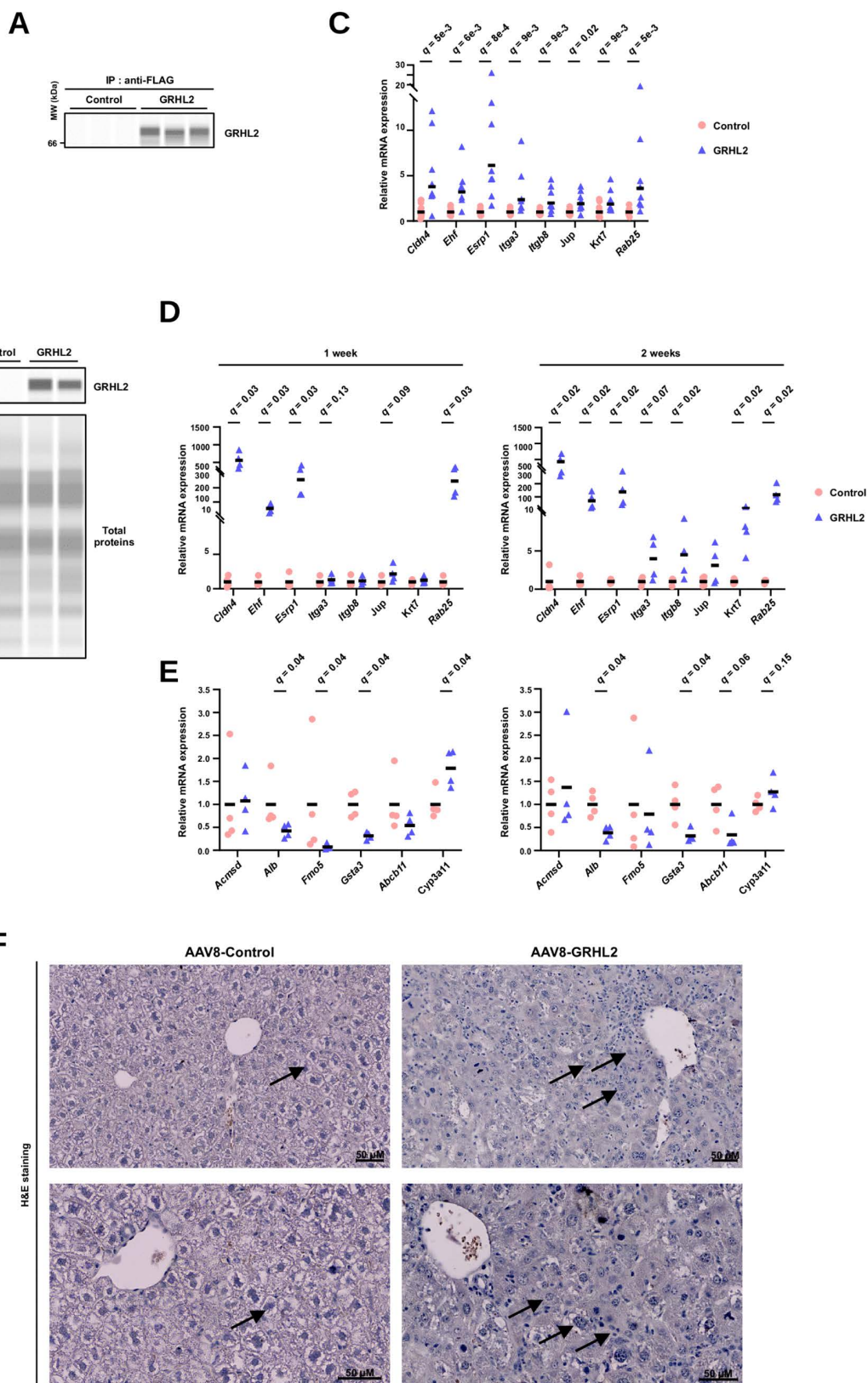


Fig 7. GRHL2 ectopic expression triggers HCT in vivo in the mouse liver. (A) Immunoblotting performed using the Wes system to monitor GRHL2 levels in the liver of mice subjected to hydrodynamic injection with the flag-tagged GRHL2-encoding plasmid or an empty construct (control). Nuclear extracts from individual mouse livers (3 mice per experimental group) were subjected to immunoprecipitation with an antibody against flag and immunoblotted using an anti-GRHL2 antibody. S10 Fig shows the immunoprecipitation specificity control. MW, molecular weight. (B) Immunoblotting performed using the Wes system to monitor GRHL2 levels in the liver of mice injected with GRHL2 or GFP (control) encoding AAV8. Nuclear extracts from individual mouse livers (2 mice per experimental group) were prepared after 1 week and subjected to immunoblotting using an anti-GRHL2 antibody. A total protein assay was used for loading control. MW, molecular weight. (C) RT-qPCR data showing the relative mRNA expression of the indicated GRHL2 target genes in the liver of mice subjected to hydrodynamic injection with the flag-tagged GRHL2-encoding plasmid ($n=8$ mice) or an empty construct (control; $n=9$ mice). For each gene, data for individual mice are shown together with the geometric mean (horizontal bar). Fold changes were obtained using the average of the control group, arbitrarily set to 1 for each analyzed gene, as the reference. One-sided Mann–Whitney U test with Benjamini–Hochberg correction for multiple testing was performed to determine whether gene expression was significantly induced upon GRHL2 ectopic expression. The numerical values of all individual mice can be found in the S1 Data file. (D) RT-qPCR data showing the relative mRNA expression of the indicated GRHL2 target genes in the liver of mice 1 or 2 weeks after being injected with GRHL2 ($n=4$ mice) or GFP (control; $n=4$ mice) encoding AAV8. For each gene, data for individual mice are shown together with the geometric mean (horizontal bar). Fold changes were obtained using the average of the control group, arbitrarily set to 1 for each analyzed gene, as the reference. One-sided Mann–Whitney U test with Benjamini–Hochberg correction for multiple testing was performed to determine whether gene expression was significantly induced upon GRHL2 ectopic expression. The numerical values of all individual mice can be found in the S1 Data file. (E) RT-qPCR data showing the relative mRNA expression of the indicated hepatocyte marker genes in the liver of mice 1 or 2 weeks after being injected with GRHL2 ($n=4$ mice) or GFP (control; $n=4$ mice) encoding AAV8. Data were plotted and analyzed as in panel D. The numerical values of all individual mice can be found in the S1 Data file. (F) Representative images obtained using liver slices from mice injected with GRHL2 or GFP (control) encoding AAV8 stained with hematoxylin and eosin (H&E) after 2 weeks. The arrows show the archetypal hepatocyte structure in the control livers which contrasts with the heterogenous cell shapes observed in GRHL2 overexpressing livers. Additional images are shown in S10C Fig.

<https://doi.org/10.1371/journal.pbio.3003547.g007>

Discussion

Hepatocyte plasticity is of primary importance for the regenerative potential of the liver. As such HCT assists the liver when coping with compromised or exhausted renewal potential of cholangiocytes [12,84]. Thus, understanding the molecular basis of HCT is of primary interest for regenerative medicine in the field of bile duct diseases including ALGS and intrahepatic biliary injuries such as PBC and PSC [85]. Beyond such conditions, HCT has also been described as a feature of end-stage disease-associated hepatocytes in ALD and MASLD [11,14,15]. In this specific context, whether HCT is indicative of ongoing regeneration in cirrhotic livers or is a phenomenon contributing to precipitate liver failure remains to be established.

Changes to the hepatocyte microenvironment are main drivers of HCT. For instance, cytokines such as transforming growth factor beta are instrumental in inducing this process [12,86,87], which occurs in human ALD-related cirrhosis at the scar interface (this study and [83]). Here, we show that HCT is linked to expression of GRHL2 in hepatocytes lining the fibrotic bands in livers of patients with end-stage ALD, PBC, and PSC. This implies that loss of hepatocyte intercellular junctions and interaction with stiff extracellular matrices likely also contribute to HCT. Indeed, this can lead to activation of the Hippo/YAP signaling pathway, a well-established driver of cholangiocyte differentiation and bile duct development [14,88,89]. Moreover, alterations in the unique multi-polarization of hepatocytes could also contribute to reprogram their transcriptomes. Indeed, modulation of the cellular polarity during differentiation processes is not merely a consequence, but rather actively contributes to cell fate decisions [90,91]. Hence, HCT is most probably triggered by a complex interplay of signals. Defining how these different mechanisms operate in concert to drive HCT and, in particular, how they are functionally connected with the activities of TFs such as GRHL2 will further improve our understanding of this process.

Previous reports have posited that hepatocytes acquire a progenitor-like hepatobiliary identity in injured livers [84]. However, this assumption has usually been based on the co-expression of a limited number of hepatocyte- and cholangiocyte-specific markers including notably SOX9 [18,19]. However, careful examination of data reported in [17] indicates that only a limited subset of the genes induced in so-called “progenitor-like cells” have stronger expression levels in hepatoblasts than in hepatocytes. Moreover, this study also pointed to gene induction in these cells involving transcriptional enhancers different from the ones active in hepatoblasts. Although relying on data issued from hepatocyte lineage

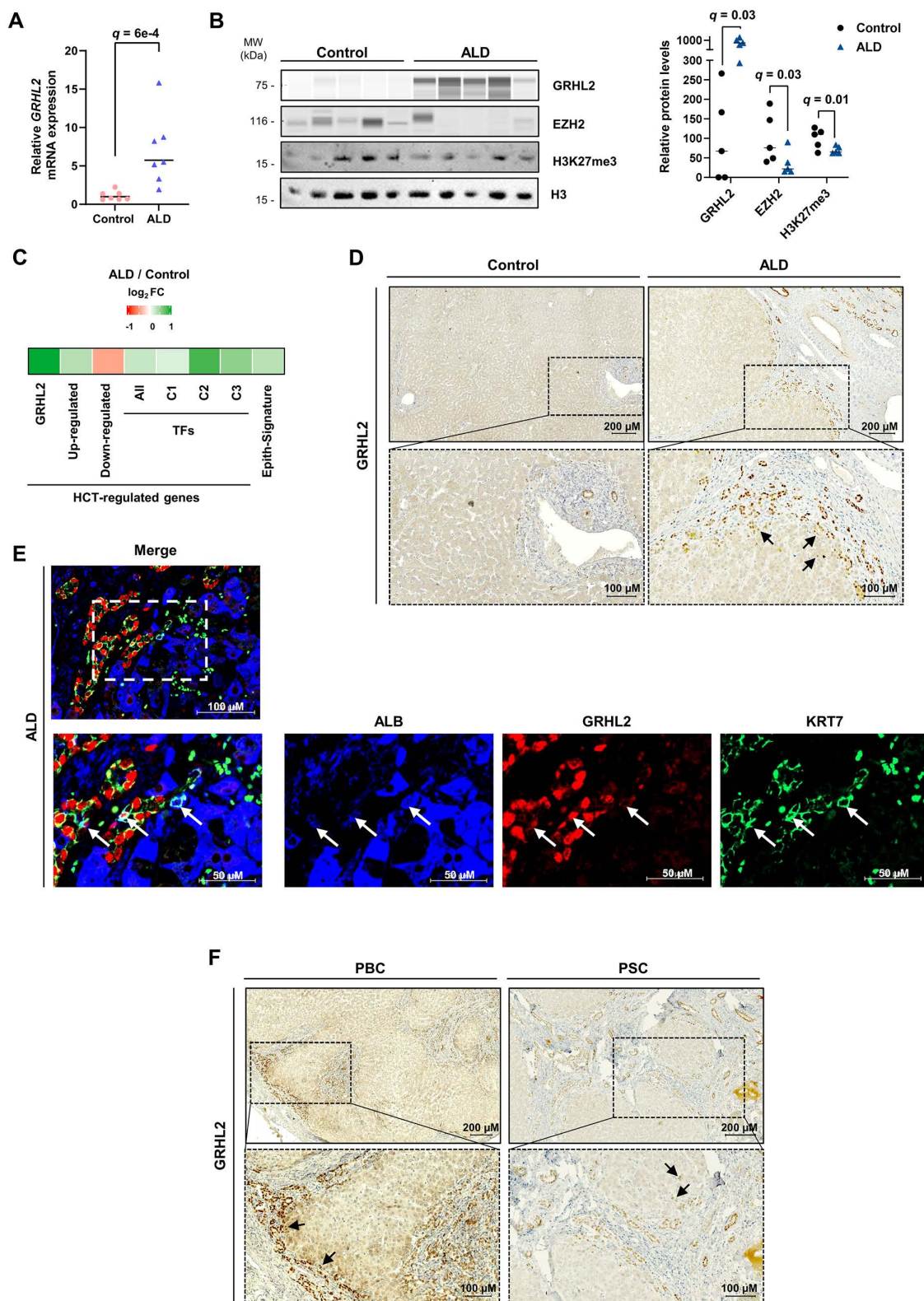


Fig 8. Hepatocytes in end-stage human chronic liver diseases express GRHL2 and epithelial genes associated with HCT. (A) RT-qPCR data showing relative *GRHL2* mRNA expression in human samples from control and ALD-related liver failure ($n=7$). Data from individual samples are shown

together with the geometric mean (horizontal bar). Fold changes were obtained using the average of the control group, arbitrarily set to 1. One-sided Mann–Whitney *U* test was performed to determine whether GRHL2 expression is significantly higher in samples from the liver failure group. The numerical values of all individual samples can be found in the [S1 Data](#) file. **(B)** Immunoblotting performed using the Wes system to monitor GRHL2, EZH2, H3K27me3, and H3 levels in human samples from control and ALD-related liver failure (*n*=5). The bar graph on the right shows results of signal quantifications normalized to H3. One-tailed Mann–Whitney *U* test with Benjamini–Hochberg correction for multiple testing was used to define whether protein levels were significantly higher in samples from the liver failure group. MW, molecular weight. The numerical values of all individual samples can be found in the [S1 Data](#) file. **(C)** RNA-seq analyses using microdissected hepatocyte-enriched areas from control and ALD-related human liver failure from our previous study [14] (*n*=5 and 7 for the control and liver failure groups, respectively) were used to monitor differential expression of *GRHL2* or the indicated previously defined groups of genes (median log2 FC is reported). **(D)** Representative immunostaining of GRHL2 in human livers from the control and ALD-related liver failure groups. Results obtained with liver samples from 5 patients are provided in [S11A](#) and [S11B Fig](#). Zoomed images of the area delimited by the dotted rectangles are shown at the bottom. Arrows point to examples of GRHL2-positive hepatocytes. **(E)** Co-immunostaining of the hepatocyte marker Albumin (ALB; in blue), GRHL2 (in red), and the cholangiocyte marker KRT7 (in green) in ALD-related human liver failure (Patient ALD#1 also used in panel D and in [S11C–S11E Fig](#)). Zoomed images of the area delimited by the dotted rectangle in the top image are shown at the bottom. Arrows indicate triple positive cells for ALB, GRHL2, and KRT7. Additional fields are shown in [S11C–S11E Fig](#). **(F)** Representative immunostaining of GRHL2 in human livers from patients with PBC- or PSC-related liver failure. Results obtained with liver samples from 3 patients with PBC or PSC are provided in [S13 Fig](#), respectively. Zoomed images of the area delimited by the dotted rectangles are shown at the bottom. Arrows point to examples of GRHL2-positive hepatocytes.

<https://doi.org/10.1371/journal.pbio.3003547.g008>

tracing is not possible in humans, scRNA-seq analysis of livers from patients with MASLD-associated liver failure pointed to HCT occurring in the absence of cells displaying a progenitor-like state [11]. These findings are consistent with our study, which also indicates that ectopic GRHL2 expression is sufficient to directly induce HCT genes in mature hepatocytes. Altogether, these data establish that HCT is a direct conversion process not requiring reversion to a progenitor-like cellular state.

Our study points to induction of a transcriptional program characterizing archetypal epithelial cells with a single apico-basal axis as the primary gene expression reprogramming event underlying HCT. In this context, GRHL2 induces epithelial genes such as genes involved in tight junctions and promotes acquisition of cyst-like structures. GRHL2 induced expression of some classical cholangiocyte markers (e.g., *KRT7*), but not all commonly used ones, indicating GRHL2 is most likely not sufficient to drive full HCT. GRHL2's role in promoting HCT is most probably incorporated into a network of collaborating TFs. For instance, GRHL2 promoted expression of the epithelial transcriptional regulators *EHF* and *ESRP1* [92,93]. Accordingly, spatial transcriptomics in human cirrhotic livers indicated that areas where HCT occurs are enriched for expression of GRHL2 and several of its target genes we identified in our study including *ESRP1* [83]. Besides forming the bile ducts, a subset of cholangiocytes is also equipped with the ability to modify the hepatocyte-secreted bile and control its flow [94]. This involves genes such as *Cystic Fibrosis Transmembrane Conductance Regulator* (*CFTR*), which were not significantly induced by GRHL2 in our experimental models. Additional regulatory pathways such as those induced by retinoic acid may be required to further modulate cholangiocyte functionalities [95]. Besides *GRHL2*, we also confirmed the previously described induction of *SOX4* (*SRY-Box Transcription Factor 4*) ([S3 Table](#)). In addition to promoting HCT by operating as a pioneer factor, *SOX4* also directly represses the hepatocyte transcriptional program [96]. We also observed reduced expression of hepatocyte markers upon GRHL2 ectopic expression in vivo in mouse hepatocytes together with profound alteration of the hepatic parenchyma structure. These data suggest that GRHL2 contributes to the repression of hepatocyte identity although this warrants further investigation. Finally, GRHL2-induced changes in gene expression in HCT may also result from the regulation of non-coding RNAs such as *MIR122* [97,98].

Induction of the cholangiocyte transcriptional program involves epigenetic remodeling events including removal of H3K27me3. A recent assessment of the epigenetic control of HCT identified a role for enzymes controlling H3K36me3 (NSD1 and KDM2A) but not for EZH and H3K27me3 [99]. However, *Ezh1* and *Ezh2* were silenced individually, which most probably explains these results since these two enzymes are redundant in hepatocytes and combined deletion of the two genes is required to trigger loss of H3K27me3 [100]. In ALD-related liver failure samples, we found decreased EZH2 and H3K27me3 levels, whose functional consequences is supported by a recent study showing that even a transient PRC2 inhibition is sufficient to trigger irreversible induction of TF encoding genes [101]. A role for H3K27me3 in controlling

hepatoblast-to-cholangiocyte differentiation has been reported in a study showing that EZH2 inhibition promotes cholangiocyte differentiation [102]. This is in line with our findings that loss of EZH2-mediated repression is involved in HCT even if HCT does not per se phenocopy the developmental process. Hence, a better understanding of the precise kinetics of epigenetic remodeling events and their functional connection to TF activities along the course of HCT is warranted.

Materials and methods

Ethics statement

Work performed in this study using human samples is authorized by the Lille ethical committee (Lille University Hospital), and informed consent was obtained from all subjects. The research was carried out according to the World Medical Association Declaration of Helsinki. All animal studies were performed in compliance with EU specifications regarding the use of laboratory animals and have been approved by the Nord-Pas de Calais Ethical Committee (CEEA-075; approval numbers 44535-2023030722295795 v6 and 30418-20210824091242 v8).

Cell culture, transfection, and treatments

HepG2 cells were cultured in William's E medium (Gibco, 22551022) supplemented with 10% FBS (Life, SVF30160.03), 20 mU/mL bovine insulin (Sigma, I5550), 50 mM dexamethasone (Sigma, D1756), and 100 U/mL penicillin-streptomycin (Gibco, 15140). Plates were coated with gelatin 1 g/L (Sigma, G1890). BMEL cells, described in [60], were cultured in RPMI 1640 Glutamax (Gibco, 61870-044), supplemented with 10% FBS (EuroBio, CVFSVF00-01), 30 mg/mL IGFII (PeproTech, 100-12-100 µg), 50 ng/mL mouse EGF (PeproTech, 315-09-1 mg), 10 µg/mL bovine insulin, and 100 U/mL penicillin-streptomycin. BMEL cells were grown on type I collagen-coated plates. All cells were grown at 37°C in a humidified atmosphere with 5% CO₂.

HepaRG cells previously described in [75] were seeded at the density of 2.6×10^4 cells/cm² and maintained in a Williams E medium supplemented with 10% FCS, 100 units/mL penicillin, 100 µg/ml streptomycin, 5 µg/mL insulin, and 5×10^{-5} M hydrocortisone hemisuccinate. After 15 days, cells were maintained in the same medium supplemented with 2% dimethyl sulfoxide (DMSO) for 2 more weeks to obtain complete hepatocyte differentiation. Differentiated cells were then detached with trypsin and resuspended in Williams E medium without DMSO prior to electroporation.

RT-qPCR analyses indicated that *GRHL2* expression was low or not detectable in all these cell lines when grown in basal conditions.

Pharmacological inhibition of EZH1 and EZH2 was obtained by treating cells with a combination of 1 µM of UNC1999 (Target Mol, T3057) [103] and 1 µM of Valemetostat Tosylate (MedChem Express, 1809336-93-3) [104] for 4 days. DMSO 0.02% was used as control.

Plasmids were transfected either in HepG2 cells or BMEL cells using JetPrime (Polyplus transfection) according to the manufacturer's instruction. The GRHL2 encoding construct was purchased from VectorBuilder. It allowed for expression of the GRHL2 cDNA (NM_024915.4) fused to a flag tag under the control of an *EEF1A1* (*EF1A*) promoter. This bicistronic construct also allowed for expression of enhanced green fluorescent protein, which was leveraged to monitor transfection efficiency using green fluorescence microscopy. The pcDNA3 vector (Invitrogen) was used as control. Cell culture medium was refreshed after 24 h and cells were harvested 48 h after transfection.

Plasmids were electroporated in differentiated HepaRG cells using the Neon electroporation system (Thermo Fisher Scientific) as previously described [105]. Briefly, 10^6 HepaRG cells were electroporated with 2 µg of control or GRHL2 encoding plasmids with a single pulse of 1,500 volts for 20 ms. The cells were then plated at low density (2.6×10^4 cells/cm²) in Williams E medium without DMSO. After 24 hours, the medium was renewed. Forty-eight hours after electroporation, the cells were harvested in ice-cold phosphate-buffered saline (PBS), scraped and pelleted by centrifugation at 400g during 5 min.

Three-dimensional cell culture

BMEL cells transfected using JetPrime as detailed hereabove were 48 h later used for three-dimensional culture in Matrigel (Corning, 356255). 20,000 cells were resuspended in 5 μ L of culture medium and gently mixed with 30 μ L of Matrigel using a pre-chilled pipette tip, taking care to avoid introducing air bubbles. Matrigel-embedded cells were deposited at the center of pre-warmed Ibidi plates (IBIDI, 80806). Plates were incubated at 37°C for 20 min to allow the Matrigel to solidify before adding complete medium to each well.

RNA extraction and gene expression analyses using real-time PCR

Tissues were homogenized using Minilys and 1.4 mm Ceramic beads (Bertin Technology) (3 \times 15 s, 5,000 rpm). Total RNA was extracted from cell lines and tissues using the Nucleospin RNA kit (Macherey-Nagel, 740955) following the manufacturer's protocol. RNA was reverse-transcribed using high-capacity cDNA Reverse Transcription Kit (Applied Biosystem). Quantitative PCR (qPCR) was performed on a QuantStudio 3 (Applied Biosystem) with the Fast program (Applied Biosystem) using the PowerTrack SYBR Green Master Mix (Applied Biosystem). The specificity of amplification was checked by recording the dissociation curves, and the PCR efficiency was verified to be above 95% for each primer pair. mRNA levels were normalized to the average expression of housekeeping genes (*TBP* only for human livers and organoids) and the fold induction was calculated using the $\Delta\Delta$ CT method. The sequences of primers used are listed in [S10 Table](#). To measure *GRHL2* and *EHF* mRNA expression levels, duplex real-time PCR using the TaqMan technology (Applied Biosystem) was performed. FAM (6-carboxyfluorescein)-labeled probes against *GRHL2* (ThermoFisher, Hs00227745) and VIC (2'-chloro-7'-phenyl-1,4-dichloro-6-carboxy-fluorescein)-labeled probes against ribosome RNA 18S (housekeeping control) were used together with the TaqMan Fast Advanced Master Mix (Applied Biosystem) according to the manufacturer's instructions.

Alternatively, RNA integrity and quantity were evaluated using the Agilent 2100 Bioanalyser (Agilent Technologies) before samples were used for library preparation and high-throughput sequencing.

Protein extraction

Total cellular extracts. For total protein extraction, cells were washed twice with ice-cold PBS, scraped and pelleted by centrifugation at 400g during 5 min. Cell pellets were resuspended in high salt lysis buffer [25 mM Tris-HCl pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.5% NP40, 0.9 mM Na Butyrate, 0.1 mM Na^+ Orthovanadate (Sigma, 450243), deacetylase inhibitor cocktail (MCE; HY-K0030), 83 U/ μ L Benzonase (Sigma, E1014-25KU) and protease inhibitor cocktail (PIC) (Roche; 63592800)]. After 10 min of incubation at 4°C, samples were sonicated twice for 10 min (10 cycles of 30 s ON/30 s OFF) with Bioruptor Pico (Diagenode). Samples were next centrifuged at 16,000g for 5 min at 4°C and supernatants were transferred to new tubes. An aliquot was used to determine protein concentration using the Pierce BCA protein assay kit (Thermo scientific) and Laemmli 6 \times (175 mM Tris-HCl pH 6.8, 15% glycerol, 5% Sodium docetyl sulfate (SDS), 300 mM Dithiothreitol (DTT), and 0.01% Bromophenol blue) was added.

Chromatin fractionation. Human liver tissues were homogenized in Buffer A (10 mM HEPES pH 7.5, 10 mM KCl, 1.5 mM MgCl_2 , 340 mM sucrose, 10% glycerol, 1 mM DTT, PIC) and Triton X-100, 0.1% using Polytron (Fisher Scientific, IKA T10 basic ULTRA-TURRAX), followed by 10 min incubation at 4°C. Then, samples were centrifuged at 1,300g for 5 min at 4°C and supernatants were transferred to new tubes subjected to centrifugation at 1,300g for 5 min at 4°C in order to remove any insoluble material. Cell pellets were washed with buffer A once and subsequently incubated in solution B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, and PIC) for 30 min at 4°C. Samples were centrifuged at 1,700g for 5 min at 4°C and supernatants were discarded. Chromatin pellets were washed a first time with solution B and a second time with PBS, and then resuspended in buffer C (50 mM Tris-HCl pH 8.0, 1 mM MgCl_2 , and 83 U/ μ L benzonase). Samples were incubated for 20 min at 4°C before protein quantification and addition of Laemmli buffer 6 \times as indicated hereabove.

Nuclear extract and Immunoprecipitation. For mouse experiment, the posterior right lobe was systematically used for protein extraction. Liver samples were filtered through 0.7 μ m filters in 1mL of cold PBS and transferred onto 1.5mL Eppendorf. Samples were centrifuged at 400g at 4°C during 5min and supernatants were removed. Pellets were suspended into 500 μ L of hypotonic buffer (20mM of Tris-HCl pH 8.0, 10mM of NaCl, 3mM of $MgCl_2$, 0.4% NP40, 5mM of sodium butyrate (Sigma, 303410), PIC, phosphatase, and deacetylase inhibitor cocktails and incubated 5min at 4°C. Samples were spin at 600g at 4°C during 5min and supernatants were transfer into new tubes (cytoplasmic fraction). Pellets were suspended in 200 μ L of nucleus lysis buffer (25mM of Tris pH 8.0, 500mM of NaCl, 1mM of EDTA, 0.5% of NP40, PIC, 5mM of sodium butyrate, PIC, phosphatase, and deacetylase inhibitor cocktails. After 20min of incubation at 4°C, samples were sonicated for 10min (10 cycles of 30s ON/30s OFF) with Bioruptor Pico and then centrifugate at max speed during 5min at 4°C. Soluble fractions (nuclear extract) were collected and transfer into new tubes. 500 μ g of nuclear extract was used per immunoprecipitation and 2 volume of dilution buffer (25mM of Tris pH 8.0, 500mM of NaCl, 1mM of EDTA, 0.5% of NP40, PIC, 5mM of sodium butyrate, PIC, phosphatase, and deacetylase inhibitor cocktails) was added. Samples were incubated with 2 μ g of antibody [anti-IgG and anti-flag (listed in [S11 Table](#))] overnight at 4°C under rotation. Dynal magnetic beads (Sigma, A9418) (5 μ L of protein A and 5 μ L of protein G per IP) were washed 3 times in PBS-BSA. Beads were block in 500 μ L of PBS-BSA (5mg/mL) overnight at 4°C under rotation before been added into samples. After 4h of incubation at 4°C, samples were washed 4 times with cold washing buffer (25mM Tris-HCl pH 8.0, 150mM NaCl, 1mM EDTA, 0.2% NP40, PIC, 5mM of sodium butyrate, PIC, phosphatase, and deacetylase inhibitor cocktails), each time 5min under rotation at room temperature. Samples were eluted in 40 μ L of Laemmli.

Western blotting

Twenty μ g of proteins were separated by 15% SDS-PAGE and immunodetected using primary antibodies (listed in [S11 Table](#)). Detection was achieved using HRP-conjugated secondary antibodies (Sigma-Aldrich). Images were acquired with an iBrigh CL1500 Imaging System (Thermo Fisher Scientific).

Simple western immunoassays (Wes)

Protein extracts were analyzed using a Wes system (Protein Simple) according to the manufacturer's instructions. Samples were diluted in 0.1 \times sample buffer 2 (Protein Simple). Final protein concentrations used in our assays ranged from 0.4 to 0.5mg/mL. Separation was performed using the 12–230kDa capillary cartridges. Primary antibodies used are listed in [S11 Table](#). Secondary antibodies were provided by the manufacturer (PS-MK14 and PS-MK15, Protein Simple). Data were analyzed using the Compass software (Protein Simple). Quantifications were obtained using the area under the peak of the protein of interest. Total protein loading control was performed using the chemiluminescent total protein assay (Protein Simple, DM-TP01).

Chromatin immunoprecipitation (ChIP)

HepG2 (5×10^6 cells) were transfected with GRHL2 plasmid using JetPrime (Polyplus transfection) according to the manufacturer's instruction. Cells were collected 48h after transfection, washed with PBS, and fixed with 1% formaldehyde (Thermo Scientific; 28908) for 10min at room temperature (RT) before quenching by incubation with 125mM of glycine for 5min. After 2 washes with ice-cold PBS, cells were scraped and pelleted by centrifugation at 600 g for 5min at 4°C. Cell pellets were incubated for 10min in 0.25% Triton X-100, 10mM EDTA, 10mM HEPES, 0.5mM EGTA, and span 5min at 600g at 4°C. Supernatants were removed and cell pellets were resuspended and incubated 10min in 0.2 M NaCl, 1mM EDTA, 10mM HEPES, 0.5mM EGTA, followed by 5min of centrifugation at 4°C at 600g. Nuclei were then resuspended in 600 μ L of lysis buffer (50mM Tris-HCl pH 8.0, 10mM EDTA, 1% SDS) and sonicated (45 cycles 30s ON/30s OFF). All buffers were supplemented with PIC (Roche). Chromatin (30 μ g for input, H3K27ac and H3K27me3 ChIP and 200 μ g for

GRHL2 ChIP) was diluted 10-fold in dilution buffer (20mM Tris HCl pH 8.0, 1% Triton X-100, 2mM EDTA, 150mM NaCl) and incubated overnight at 4°C with 2 µg of anti-H3K27me3 antibody, 2 µg of anti-H3K27ac antibody or 1.5 µg of anti-flag antibody (Sigma, F1804) combined with 1.5 µg of anti-GRHL2 antibody (antibodies are listed in [S11 Table](#)). Protein A/G sepharose beads (GE Healthcare) were added and samples were incubated for 4 h at 4°C under agitation in the presence of 10 µg/mL yeast tRNA (Sigma-Aldrich, R5636). Beads were washed three times with RIPA buffer (50mM HEPES, 1mM EDTA, 0.7% Na Deoxycholate, 1% NP40 and 500mM LiCl) containing 10 µg/mL yeast tRNA and once with TE buffer (10mM Tris-HCl pH 8.0, 1mM EDTA). DNA was then eluted in 100mM NaHCO₃ containing 1% SDS and incubated overnight at 65°C for reverse cross-linking. DNA purification was performed using Nucleospin PCR and Gel purification kit (Macherey Nagel, 740609). ChIPed DNA and input samples were subjected to high-throughput sequencing and were further analyzed as described hereafter.

Histology, immunohistochemistry and immunofluorescent assays

Liver samples were fixed with 4% paraformaldehyde (PFA; mouse samples) or buffered formalin (human samples) and embedded in paraffin. 4 µm-thick tissue sections were stained with hematoxylin and eosin or underwent an antigen retrieval step with sodium citrate solution (pH 6.0) during 21 min at 110°C into TintoRetriever Pressure Cooker (Bio SB) for immunostaining. Sections were then sequentially incubated with 0.1% Triton X-100 in Tris-Buffered Saline (TBS) for 10min at 4°C, Bloxall Endogenous Blocking solution (Vector Laboratories) for 10 min at RT, 7% goat serum in 0.05% Tween-20 in TBS (TBS-T) for 15min at RT, 15min at RT with 5% milk and 3% BSA in TBS-T and primary GRHL2 antibody ([S11 Table](#)) diluted 1:300 in 5% milk and 3% BSA in TBS-T overnight at 4°C. After each step, sections were washed in TBS-T. Subsequently, tissue sections were incubated 30 min at RT with ImmPRESS HRP Polymer anti-mouse IgG Reagent (Vector Laboratories), followed by a TBS wash. Signal was revealed with diaminobenzidine [DAB (Cell Signaling)], and sections were counterstained with hematoxylin (solution Gill N°1, Sigma-Aldrich) and mounted in Aquatex (Sigma-Aldrich). Stained slices were imaged with an Axioscan Z1 slide scanner (Zeiss). Images processing was performed using ZEN software (Zeiss).

For GRHL2 and ALB co-staining, the procedure was adapted as follows. For ALB labeling, instead of using 7% goat serum in TBS-T and DAB, sections were incubated for 15 min at RT with Animal-Free Blocking solution 1× (Cell Signaling) in TBS-T, primary anti-ALB antibody diluted 1:50 ([S11 Table](#)), ImmPRESS HRP Polymer horse anti-goat IgG reagent (Vector Laboratories) and 10 min at RT with Alexa Fluor 594 Tyramide SuperBoost reagent (Invitrogen) according to the manufacturer's instructions. Following ALB labeling, similar steps were repeated for GRHL2 staining after an antigen retrieval step using primary anti-GRHL2 antibody (1:300), ImmPRESS HRP Polymer horse anti-mouse IgG reagent (Vector Laboratories), and Alexa Fluor 488 Tyramide SuperBoost reagent (Invitrogen). Stained slices were imaged using Axioscan Z1 slide scanner after nuclei counterstaining with DAPI (62248, Thermo Scientific, 1:10,000).

For triple immunofluorescent staining of GRHL2, ALB, and KRT7, the procedure was further adapted. After blocking with Animal-Free Blocking solution 1× in TBS-T, sections were incubated with primary anti-KRT7 antibody diluted 1:50 ([S7 Table](#)), ImmPRESS HRP Polymer horse anti-mouse IgG reagent and Alexa Fluor 488 Tyramide SuperBoost reagent. GRHL2 detection was then performed following an antigen retrieval step, using the same HRP-base amplification system (primary anti-GRHL2 antibody 1:300 and Alexa Fluor 594 Tyramide SuperBoost dye). Finally, ALB staining was carried out using the primary anti-ALB antibody (1:50), followed by Alexa Fluor 647-conjugated donkey anti-goat IgG (A21447, Invitrogen, 1:250). Nuclei were counterstained with DAPI and sections were scanned using an Axioscan Z1 slide scanner.

BMEL cells grown in Matrigel for 4 days were fixed using 4% PFA during 20 min at RT, washed with PBS and quenched with NH₄Cl 50mM 15 min at RT. Cells were then permeabilized with 0.1% Triton X-100 for 5 min and samples were incubated for 1 h at RT with a blocking solution [PBS + BSA 1%] followed by 1 h incubation with phalloidin (Invitrogen AlexaFluor Plus A30107, 2379366) diluted 1:1000 and 5 min with Hoechst diluted 1:5,000 in PBS + BSA 1%. Samples

were washed with PBS between each step. Cell aggregates were imaged using Celldiscoverer 7 (Zeiss) or confocal microscopy (Zeiss, LSM700) and analyzed using the ZEN software.

Animal studies

Male mice (C57BL/6J) were purchased from Charles River at 6–11 weeks of age and housed in standard cages in a temperature-controlled room (22–24°C) with 12 hours light-dark cycles and provided with water and standard diet ad libitum. Mice were allowed to acclimate for at least 1 week prior to any experiment.

Hydrodynamic plasmid transfection was performed using intravenous tail vein injections (5-s injections) of 30 µg of plasmid diluted in saline solution (0.9% w/v sodium chloride; final volume was adjusted to mouse body weight, i.e., 0.1 mL/g was used). After hydrodynamic injection, mice were closely monitored during 2 hours. Livers were collected 48 hours later.

Intravenous viral injections using a 26G syringe were performed via the retro-orbital sinus with 4×10^{10} viral genomes of AAV8 suspended in 100 µL of sterile 0.9% NaCl per animal. Recombinant AAV8 vectors enabling the expression of GRHL2 (NM_024915.4) under the *EF1A* promoter, or GFP-expressing control viruses, were used (VectorBuilder). Tissue samples were collected 1 or 2 weeks after injection for subsequent analyses.

Human samples

Liver samples with alcohol-related decompensated cirrhosis were obtained from the TargetOH cohort (“Comparison of Inflammatory Profiles and Regenerative Potential in Alcoholic Liver Disease”; ClinGov NCT03773887; and DC-2008-642) [14,106,107]. Controls consisted in liver samples obtained from the non-cancerous part of liver resections performed in patients with hepatic tumors. All human liver samples were obtained from patients who underwent liver transplantation at Huriez Hospital’s Liver Unit (Lille, France). The liver samples were immediately fixed for histology or frozen for RNA and protein extraction.

Human liver organoid culture and AAV transduction

Human liver organoids were derived from histologically normal liver tissue obtained from patients undergoing PHx for metastasis excision. Tissue samples were collected from macroscopically healthy regions distant from metastatic lesions. Organoid isolation, expansion, and differentiation were performed as described in [77,78]. Briefly, liver cells were dissociated through enzymatic digestion and EPCAM-positive cells sorted out using magnetic-activated cell sorting using anti-EpCAM microbeads (Miltenyi Biotec, 130-061-101) according to the manufacturer’s instructions. Cells were embedded in 20 µL domes of growth factor–reduced Matrigel (Corning, 356231) and cultured for two weeks in HepatiCult Organoid Growth Medium (STEMCELL Technologies, 100-0385) to enable organoid formation and expansion. Expanded organoids were dissociated using gentle trypsinization and re-seeded at a density of 1,000 cells per Matrigel dome. After reformation, organoids were transduced with recombinant AAV serotype 3 (AAV3) vectors encoding either GRHL2 (NM_024915.4) or GFP under the *EF1A* promoter (VectorBuilder). Transduction was performed in differentiation medium at a dose of 1×10^5 viral genomes (vg) per organoid-forming cell. Medium was changed on the third day post-infection. Organoids were maintained in differentiation medium for an additional 12 days before being harvested and processed for total RNA extraction.

Origin and recovery of public datasets

Public transcriptomic and functional genomics data used in this study were downloaded from Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) [108], ENCODE [109], Roadmap epigenome (<http://www.roadmapepigenomics.org>) [110], CistromeDB (<http://cistrome.org/db/#/>) [111], or the Human Protein atlas (<https://www.proteinatlas.org>) [37]. Details are provided in S1 and S4 Tables.

The list of mouse TF encoding genes was obtained from the AnimalTFDB 3.0 database [112].

Gene transcriptional start sites (TSS) used in this study were obtained from [113]. For each individual gene, the most active TSS in the mouse liver was considered. Mouse TSS coordinates (mm10; coordinates corresponding to windows of ± 500 bp around the TSS) were used in the LiftOver tool of the UCSC genome browser (<https://genome.ucsc.edu/cgi-bin/hgLiftOver>) [114] to get the corresponding human coordinates (hg38).

The Epith-Signature gene list was obtained from [36] by selecting “Epithelial” as the cell type in their [S7 Table](https://genome.cshlp.org/content/30/7/1047/suppl/DC1) (<https://genome.cshlp.org/content/30/7/1047/suppl/DC1>). Ensembl ID were converted to official gene symbols using the ENSEMBL Gene ID Converted from biotools.fr (https://www.biotools.fr/human/ensembl_symbol_converter).

Hepatocyte- and cholangiocyte-enriched genes were retrieved from the Human Protein Atlas website (<https://www.proteinatlas.org>) [34].

RNA-seq data processing and transcriptomic analyses

Bulk RNA-Seq data analyses. For public transcriptomic data, SRA files were first downloaded and converted into FastQ files using the SRA toolkit [115]. FastQ files were uploaded and analyzed on a local instance of Galaxy [116]. Reads were mapped either to the human genome (hg38) or to the mouse genome (mm10) using Hisat2 (version 2.0.1) [117]. Read counting was performed using the count function of the HTSeq tool (version 0.9.1) using the following parameters: mode=UNION, Stranded=No, Minimum alignment quality=10, Feature types=Exon, ID attributes=Gene ID [118]. Normalization of gene expression data and differential expression analyses were performed in R [119] using the DESeq2 (version 1.42.1) package [120]. Pre-filtering was applied to remove genes with low counts, i.e., genes with cumulated reads in all samples below 10. Significant changes in gene expression were defined considering corrected *p*-values using cut-offs indicated in figure legends. Fold Changes used in our analyses were those provided by DESeq2. Normalized gene expression data were obtained by dividing Rlog normalized gene expression data obtained from DESeq2 by transcript lengths. Transcript lengths had previously been calculated by averaging the lengths of all transcripts associated to a given gene symbol in Ensembl (GRCm39 or GRCh.38.p13) [121].

Results of bulk RNA-seq transcriptomic data analyses were visualized using dot plots, box plots or pie charts generated using the ggplot2 R package (version 3.5.1) [122] or using heatmaps generated using the ComplexHeatmap R package (version 2.15.4) [123].

[S12 Table](#) provides the list and version of all bioinformatical tools used in this study.

Comparison of transcriptomes using principal component analysis (PCA). The transcriptome of hepatocytes undergoing HCT (denoted as “reprogrammed hepatocytes” in the initial study) in livers of mice treated with DDC was compared to that of hepatocytes and cholangiocytes from control mice [25]. All genes, pending their expression was detected in all datasets, were considered in these analyses. In order to strengthen the robustness of our analyses, transcriptomes of hepatocytes and cholangiocytes sorted from healthy mouse livers obtained in our laboratory were also considered [33]. Batch correction was performed using the Combat function of the sva R package (version 3.50.0) [124] using the transcriptome of hepatocytes from [33] as the reference dataset and the following parameters: mean.only=T, par.prior=T, prior.plots=F. Next, a PCA was computed using the PCA function of the FactomineR package [125] on the transcriptome of hepatocytes and cholangiocytes from [33], the other datasets being considered as supplemental individuals. The 2 first principal components were used to project individual datasets showing that healthy hepatocytes from the two initial different studies were overlapping. This entire process was similarly performed using only TF encoding genes.

Gene ontology and pathway enrichment analyses. Enrichment analysis for biological or molecular pathways were performed using Metascape (<https://metascape.org>) [126] or Database for Annotation, Visualization, and Integrate Discovery (DAVID 2021, DAVID Knowledgebase v2023q4) [35]. For analyses using DAVID, the top 3,000 deregulated genes were selected based on corrected *p*-values when required. Functional annotation clustering was performed using

default parameters except for similarity term overlap and similarity threshold which were set to 5 and 0.5, respectively. In order to name the clusters, enriched GO terms and pathways from each cluster were used for word counting using WordArt (<https://wordart.com>).

Gene set enrichment analyses. Gene set enrichment analyses (GSEA) were performed using the GSEA software (version 4.3.2) from the Broad Institute [127] and gene sets from S2 Table. We used 1,000 gene-set permutations, no collapse, and the following additional parameters: enrichment statistic=weighted, gene ranking metric= difference of classes. Ranking was performed by the GSEA software using Rlog normalized gene expression data obtained from DeSeq2. In addition to enrichment plots, figures also provide NES and *q*-value, which are the normalized enrichment score and the false discovery rate provided by the GSEA software, respectively.

Mouse liver single-cell RNA-seq analyses

Pre-processing of individual datasets. Data were analyzed with R (version 4.3.1) using raw count matrices obtained as.rds files. Details about all datasets re-analyzed are provided in S1 Table. First, each dataset was filtered using Seurat (version 4.0.3) [128] to keep only cells with more than 1,000 and less than 5,000 expressed genes (excepted for GSE186554 for which the authors had already selected cells comprising 1,500–4,500 expressed genes) and with less than 20% of mitochondrial counts. Details of the Seurat functions used together with all parameters can be retrieved in the supplemental Rscript “1_PreprocessingExample.R” (available at <https://zenodo.org/records/13897656>), which provides the code used for one of the re-analyzed study. Then data were normalized with *sctransform* [129] followed by a PCA used to compute UMAP. When required, Louvain clustering was performed in order to remove any remaining cell cluster not within the hepatocyte/cholangiocyte lineages. This involved monitoring expression of the following marker genes: *Alb* and *Apoc3* for hepatocytes, *Sox9* and *Epcam*, for cholangiocytes, *Hnf4a*, *Afp*, *Dlk1*, *Cd13*, *Fgb*, *Fxyd1* and *Gjb1*, *Id3* for hepatoblast, *Gypa*, *Vim*, *Flt1*, *Pecam1*, *Itgam*, *Clec4f*, *Eln*, *Cd248*, *Upk1b*, *Dcn*, *Lrat*, *Des*, *Pdgfrb*, *Lox*, *Acta2*, *Col1a1* for non-parenchymal cells (NPC). Clusters characterized by expression of NPC marker genes were discarded (see S1 Table for details regarding clustering resolution and numbers of selected cells for each dataset). Finally, to prevent potential bias arising from an unbalanced cell number across datasets and experimental conditions, two datasets were subsampled so that cells from injured livers did not exceed the maximal number of healthy hepatocytes provided by a single study (*n*=8,964) (S1 Table).

Integration of the different datasets. Data were log-normalized and the 2000 most variable genes in each dataset were defined with the *vst* method using Seurat. Next, STACAS (version 2.0) [130] was used on the 500 common most variable genes selected by *SelectIntegrationFeatures* function with *nfeatures*=500. The *FindAnchors* and *IntegrateData* functions were used with the following parameters: *dims*=1:30, *alpha*=0.8, *coverage*=0.5, *semisupervised*=false. Efficient batch correction was validated on basis of the LISI and ASW scores following guidelines from STACAS. To evaluate integration with the ASW score, a unifying nomenclature where cells were labeled using cell type identity and experimental condition (e.g., healthy control, DDC etc) was used. All codes and parameters are provided in “Rscript 2_Integration_and_Focus.R” available at <https://zenodo.org/records/13897656>.

Selection of cells to specifically investigate HCT and trajectory analyses. Finally, PCA, UMAP, and Louvain clustering were computed on the integrated data provided by STACAS. Data were visualized using ISCEBERG [Interactive Single Cell Expression Browser for Exploration of RNAseq data using Graphics; [131]]. Cells from clusters 2, 8, and 15 (resolution=0.8; S1B Fig) were further considered after removal of studies contributing less than 5% of all cells from these 3 clusters. In order to compute RNA velocity, healthy hepatocytes from [26], condition: « control »] were added to the selected cells. Reads from all selected cells were mapped to Ensembl (GRCm39 v108) using *salmon-alevin* (*salmon* version 1.10.1) [132] to generate spliced and unspliced transcript count matrices. These matrices were batch corrected to match the aforementioned STACAS computed integration. RNA velocity was next computed using *ScVelo* (v0.2.5 under Python 3.9.5) [133] and data uploaded into *CellRank2* (version 2.0.0) [32] to compute cell fate probabilities.

Pseudotime analyses were performed using slingshot (version 2.10.0) [134] and next also uploaded to CellRank2 in order to specifically select cells defining a trajectory of HCT. The batch corrected gene expression data from cells belonging to this HCT trajectory from DDC-treated mice [25] were used to plot gene expression along pseudotime. Normalization was performed with the “LogNormalize” function of Seurat (i.e., for each cell, gene counts were divided by the total number of counts, multiplied by a scale factor set to 10000 and then log-transformed). All codes and parameters are provided in “Rscript 3_Velocity_CellRank.R” available at <https://zenodo.org/records/13897656>. Gene set expression analyses were performed using the AddModuleScore function of Seurat in ISCEBERG.

Human single-nuclei RNA-seq data analyses

Data from [11] were mined using the R Shiny application provided by the authors at https://www.mohorianulab.org/shiny/vallier/LiverPlasticity_GribbenGalanakis2024/ or using Seurat and data downloaded from GEO (GSE202379). In all instances, only cells labeled as hepatocytes were considered. Odds ratios and gene expression were computed using hepatocytes originating only from the right lobe so that each patient contributed only one sample to the dataset. Odds ratio were computed using median-unbiased estimation thanks to the oddsratio function of the epitools (version 0.5-10.1) [135].

Mining transcriptomic data of human primary cells from the Human Protein Atlas

Data from scRNA-seq were downloaded from the Human Protein Atlas website (<https://www.proteinatlas.org/about/download>) and corresponded to normalized gene expression in transcripts per million (nTPM) issued from pseudo-bulk analyses [37]. Gene expression in different cell lineages (epithelial, mesenchymal, endothelial, etc.) was computed as the mean of nTPM from the individual cell types associated to the different cell lineages according to the nomenclature provided by the Human Protein Atlas.

The Tau index was computed to assess gene expression lineage-specificity using tspex (<https://tspex.lge.ibi.unicamp.br>) [136] after expression data had been log₂-transformed. Expression data for cell lineages were here defined as the median expression from all individual cell types linked to a given lineage.

Hierarchical clustering based on expression of the Epith-Signature was performed using the hclust (version) and dist functions of the R package stats version 4.3.1 [119] with the following method parameters: Ward.D2 and Euclidean.

ChIP-Seq data analysis

After initial quality controls of FastQ files, reads were mapped to hg38 using Chromap version 0.1.3-R256 [137] with preset parameters for “chip” including: remove PCR duplicates; error threshold = 8; min-num-seeds = 2; max-seed-frequency = 500–1,000; max-num-best-mappings = 1; max-insert-size = 1,000; MAPQ-threshold = 10; min-read-length = 30; bc-error-threshold = 1; bc-probability-threshold = 0.90.

For H3K27ac and H3K27me3 ChIP-seq, signal tracks were obtained using the bamCoverage function of the DeepTools version 3.5.1 [138] with the following parameters: bin size = 25; normalize = CPM; read extension = 200.

For GRHL2 ChIP-seq, Bam files ($n = 4$ biological replicates) were analyzed using the Irreproducible Discovery Rate (IDR), which allows to identify TF binding peaks that are reproducible and rank-concordant across replicates (<https://github.com/nboley/idr>). This was performed using the IDR pipeline from ENCODE (version 2.2.1) installed from <https://github.com/ENCODE-DCC/chip-seq-pipeline2/tree/master> [139]. ChIP-seq data from mocked transfected cells were used as controls. Since all replicates passed the reproducibility tests of IDR, the top 10,000 peaks based on the IDR score and the average signal from the 4 replicates were used in subsequent analyses. GRHL2 target genes and their enrichment for specific biological terms or pathways were predicted using the CistromeGO tool [69]. The solo mode or the ensemble mode, i.e., combined prediction of target genes using both the GRHL2 binding peaks and genes modulated by GRHL2 as defined using RNA-seq were used with a half-decay distance of 1 kb automatically defined by the CistromeGO algorithm.

All ChIP-seq and chromatin accessibility signals analyses were made looking at gene TSS and monitored using the DeepTools version 3.5.1 [138]. First, the compute matrix tool was used with the following parameters: TSS set as the reference point, binSize of 500 bp, $\pm 2,500$ bp window around the TSS. Then, issued matrices were used to generate average signal plots using the PlotProfile tool. ChIP-seq signals at individual genes were visualized using the Integrated Genome Browser (IGB) (version 10.0.1) [140] or the Integrative Genomics Viewer (IGV) (version 2.17.4) [141].

In order to identify groups of genes with distinct epigenetic patterns around their TSS, the SPARK software (version 1.3.0) was used [42]. SPARK clustering was performed using a window of 2.5 kb centered around gene TSS split into 10 bins and signal normalization set as global, i.e., signal across the whole genome is considered for normalization.

Statistical analyses

Statistical analyses were performed using R 4.3.1 [119] or GraphPad Prism. The number of samples and statistical tests used are indicated in the figure legends. Corrections for multiple testing was performed by computing false discovery rates reported as q -values. Exact p - or q -values are indicated unless they exceed 0.15. All bar graphs show means \pm SD (standard deviations). Box plots are composed of a box from the 25th to the 75th percentile with the median as a line and min to max as whiskers.

Supporting information

S1 Fig. Additional characterization of the mouse parenchymal liver cells scRNA-seq atlas and of the cellular trajectory involved in HCT.

(PDF)

S2 Fig. Additional characterization of the gene transcriptional changes occurring in HCT.

(PDF)

S3 Fig. Additional characterization of deregulated TF-encoding genes during HCT.

(PDF)

S4 Fig. Average DHS-seq, H3K36me3 and RNAPolIII ChIP-seq signals at the promoters of genes from clusters C1–3 in the mouse liver.

(PDF)

S5 Fig. Role of H3K9me3 in the repression of HCT genes in normal hepatocytes.

(PDF)

S6 Fig. Basal expression of TF-encoding genes from the C1–3 clusters.

(PDF)

S7 Fig. Additional validation of GRHL2 ectopic expression and characterization of induced transcriptional changes in BMEL and HepG2 cells.

(PDF)

S8 Fig. Additional representative images of BMEL cell aggregates.

(PDF)

S9 Fig. Dose-dependent induction of GRHL2 target genes in HepG2 cells.

(PDF)

S10 Fig. Additional validation of GRHL2 ectopic expression in the mouse liver of mice subjected to hydrodynamic tail vein injection.

(PDF)

S11 Fig. Additional immunostainings of GRHL2 in human end-stage chronic liver diseases.

(PDF)

S12 Fig. Characterization of GRHL2-expressing hepatocytes in human MASLD using snRNA-seq data.

(PDF)

S13 Fig. Additional immunostainings of GRHL2 in human PBC and PSC samples.

(PDF)

S1 Table. Details regarding single-cell RNA-seq datasets used in Fig 1.

(XLSX)

S2 Table. Gene sets used in this study.

(XLS)

S3 Table. List of genes and TF-encoding genes differentially expressed between HCT and Hepatocytes.

(XLSX)

S4 Table. Transcriptomic and functional genomic data used in this study.

(XLS)

S5 Table. List of genes down-regulated upon ectopic expression of GRHL2 in BMEL cells ($q \leq 0.05$).

(TXT)

S6 Table. List of genes up-regulated upon ectopic expression of GRHL2 in BMEL cells ($q \leq 0.05$).

(TXT)

S7 Table. List of genes down-regulated upon ectopic expression of GRHL2 in HepG2 cells ($q \leq 0.05$).

(TXT)

S8 Table. List of genes up-regulated upon ectopic expression of GRHL2 in HepG2 cells ($q \leq 0.05$).

(TXT)

S9 Table. Features of the selected GRHL2 target genes.

(XLSX)

S10 Table. Primer sequences used in this study.

(XLS)

S11 Table. Antibodies used in this study.

(XLSX)

S12 Table. Bioinformatical tools used in this study.

(XLS)

S1 Data. All individual numerical values that underlie the experimental data generated specifically for this study.

(XLSX)

S1 Raw Images. Uncropped images issued from western blotting or simple western immunoassays.

(PDF)

Acknowledgments

The authors thank Gilles Salbert (Univ Rennes) and Réjane Paumelle (Univ Lille) for discussions, Jean-Sébastien Annicotte and Joel Haas for help with animal experimentation and Solène Audry for technical assistance. The authors are indebted to Yassine Jarmouni and the BioImaging Center Lille (BICeL; Plateformes Lilloises en Biologie et Santé (PLBS) – UMS 2014 – US 41, Univ. Lille). We acknowledge Bilille platform – US 41 – UAR 2014 – PLBS and the i-tensive computing center of the University of Lille for the computing resources (Cloud Bilille).

Author contributions

Conceptualization: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Laurent Dubuquoy, Jérôme Eeckhoutte.

Formal analysis: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Mohamed Bou Saleh, Jérôme Eeckhoutte.

Funding acquisition: Bart Staels, Philippe Lefebvre, Laurent Dubuquoy, Jérôme Eeckhoutte.

Investigation: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Loïc Guille, Mohamed Bou Saleh, Clémence Boulet, Cyril Sobolewski, Pascal Loyer, Alexandre Berthier, Noémie Legrand, Solenne Taront, Jérôme Eeckhoutte.

Methodology: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Loïc Guille, Mohamed Bou Saleh, Clémence Boulet, Cyril Sobolewski, Pascal Loyer, Alexandre Berthier, Noémie Legrand, Anne Corlu, Viviane Gnemmi, Antonino Bongiovanni, Nicolaj I Toft, Lars Grøntved.

Project administration: Bart Staels, Philippe Lefebvre, Laurent Dubuquoy, Jérôme Eeckhoutte.

Resources: Viviane Gnemmi, Guillaume Lasailly, Emmanuelle Leteurtre, David Tulasne, Alessandro Furlan, Line Carolle Ntandja-Wandji, Laurent Dubuquoy.

Software: Ludivine Vasseur, Julie Dubois-Chevalier, Dmitry Galinowsky, Antonino Bongiovanni.

Supervision: Bart Staels, Laurent Dubuquoy, Jérôme Eeckhoutte.

Validation: Ludivine Vasseur, Céline Gheeraert.

Visualization: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Loïc Guille, Mohamed Bou Saleh, Viviane Gnemmi, Dmitry Galinowsky, Nicolaj I Toft, Lars Grøntved, Jérôme Eeckhoutte.

Writing – original draft: Ludivine Vasseur, Jérôme Eeckhoutte.

Writing – review & editing: Ludivine Vasseur, Céline Gheeraert, Julie Dubois-Chevalier, Ninon Very, Mohamed Bou Saleh, Pascal Loyer, Alexandre Berthier, Anne Corlu, Viviane Gnemmi, Emmanuelle Leteurtre, Bart Staels, Philippe Lefebvre, Laurent Dubuquoy, Jérôme Eeckhoutte.

References

1. Tachmatzidi EC, Galanopoulou O, Talianidis I. Transcription control of liver development. *Cells*. 2021;10(8):2026. <https://doi.org/10.3390/cells10082026> PMID: 34440795
2. Gérard C, Tys J, Lemaigre FP. Gene regulatory networks in differentiation and direct reprogramming of hepatic cells. *Semin Cell Dev Biol*. 2017;66:43–50. <https://doi.org/10.1016/j.semcdb.2016.12.003> PMID: 27979774
3. Dubois V, Staels B, Lefebvre P, Verzi MP, Eeckhoutte J. Control of cell identity by the nuclear receptor HNF4 in organ pathophysiology. *Cells*. 2020;9(10):2185. <https://doi.org/10.3390/cells9102185> PMID: 32998360
4. Banales JM, Huebert RC, Karlsen T, Strazzabosco M, LaRusso NF, Gores GJ. Cholangiocyte pathobiology. *Nat Rev Gastroenterol Hepatol*. 2019;16(5):269–81. <https://doi.org/10.1038/s41575-019-0125-y> PMID: 30850822

5. Tanimizu N, Mitaka T. Epithelial morphogenesis during liver development. *Cold Spring Harb Perspect Biol.* 2017;9(8):a027862. <https://doi.org/10.1101/cshperspect.a027862> PMID: 28213465
6. Morales-Navarrete H, Nonaka H, Scholich A, Segovia-Miranda F, de Back W, Meyer K, et al. Liquid-crystal organization of liver tissue. *Elife.* 2019;8:e44860. <https://doi.org/10.7554/eLife.44860> PMID: 31204997
7. Gissen P, Arias IM. Structural and functional hepatocyte polarity and liver disease. *J Hepatol.* 2015;63(4):1023–37. <https://doi.org/10.1016/j.jhep.2015.06.015> PMID: 26116792
8. Slim CL, van IJzendoorn SCD, Lázaro-Díéguez F, Müsch A. The special case of hepatocytes: unique tissue architecture calls for a distinct mode of cell division. *Bioarchitecture.* 2014;4(2):47–52. <https://doi.org/10.4161/bioa.29012> PMID: 24769852
9. Gadd VL, Aleksieva N, Forbes SJ. Epithelial plasticity during liver injury and regeneration. *Cell Stem Cell.* 2020;27(4):557–73. <https://doi.org/10.1016/j.stem.2020.08.016> PMID: 32971004
10. Chen F, Schönberger K, Tchorz JS. Distinct hepatocyte identities in liver homeostasis and regeneration. *JHEP Rep.* 2023;5(8):100779. <https://doi.org/10.1016/j.jhepr.2023.100779> PMID: 37456678
11. Gribben C, Galanakis V, Calderwood A, Williams EC, Chazarra-Gil R, Larraz M, et al. Acquisition of epithelial plasticity in human chronic liver disease. *Nature.* 2024;630(8015):166–73. <https://doi.org/10.1038/s41586-024-07465-2> PMID: 38778114
12. Schaub JR, Huppert KA, Kurial SNT, Hsu BY, Cast AE, Donnelly B, et al. De novo formation of the biliary system by TGFβ-mediated hepatocyte transdifferentiation. *Nature.* 2018;557(7704):247–51. <https://doi.org/10.1038/s41586-018-0075-5> PMID: 29720662
13. Dubuquoy L, Louvet A, Lassailly G, Truant S, Boleslawski E, Artru F, et al. Progenitor cell expansion and impaired hepatocyte regeneration in explanted livers from alcoholic hepatitis. *Gut.* 2015;64(12):1949–60. <https://doi.org/10.1136/gutjnl-2014-308410> PMID: 25731872
14. Bou Saleh M, Louvet A, Ntandja-Wandji LC, Boleslawski E, Gnemmi V, Lassailly G, et al. Loss of hepatocyte identity following aberrant YAP activation: a key mechanism in alcoholic hepatitis. *J Hepatol.* 2021;75(4):912–23. <https://doi.org/10.1016/j.jhep.2021.05.041> PMID: 34129887
15. Andrews TS, Nakib D, Perciani CT, Ma XZ, Liu L, Winter E, et al. Single-cell, single-nucleus, and spatial transcriptomics characterization of the immunological landscape in the healthy and PSC human liver. *J Hepatol.* 2024;80(5):730–43. <https://doi.org/10.1016/j.jhep.2023.12.023> PMID: 38199298
16. Sato K, Marzioni M, Meng F, Francis H, Glaser S, Alpini G. Ductular reaction in liver diseases: pathological mechanisms and translational significances. *Hepatology.* 2019;69(1):420–30. <https://doi.org/10.1002/hep.30150> PMID: 30070383
17. Li L, Cui L, Lin P, Liu Z, Bao S, Ma X, et al. Kupffer-cell-derived IL-6 is repurposed for hepatocyte dedifferentiation via activating progenitor genes from injury-specific enhancers. *Cell Stem Cell.* 2023;30(3):283–299.e9. <https://doi.org/10.1016/j.stem.2023.01.009> PMID: 36787740
18. Li W, Yang L, He Q, Hu C, Zhu L, Ma X, et al. A Homeostatic Arid1a-dependent permissive chromatin state licenses hepatocyte responsiveness to liver-injury-associated YAP signaling. *Cell Stem Cell.* 2019;25(1):54–68.e5. <https://doi.org/10.1016/j.stem.2019.06.008> PMID: 31271748
19. Tarlow BD, Pelz C, Naugler WE, Wakefield L, Wilson EM, Finegold MJ, et al. Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell Stem Cell.* 2014;15(5):605–18. <https://doi.org/10.1016/j.stem.2014.09.008> PMID: 25312494
20. Wang X, Yang L, Wang Y-C, Xu Z-R, Feng Y, Zhang J, et al. Comparative analysis of cell lineage differentiation during hepatogenesis in humans and mice at the single-cell transcriptome level. *Cell Res.* 2020;30(12):1109–26. <https://doi.org/10.1038/s41422-020-0378-6> PMID: 32690901
21. Liang Y, Kaneko K, Xin B, Lee J, Sun X, Zhang K, et al. Temporal analyses of postnatal liver development and maturation by single-cell transcriptomics. *Dev Cell.* 2022;57(3):398–414.e5. <https://doi.org/10.1016/j.devcel.2022.01.004> PMID: 35134346
22. Ben-Moshe S, Veg T, Manco R, Dan S, Papinutti D, Lifshitz A, et al. The spatiotemporal program of zonal liver regeneration following acute injury. *Cell Stem Cell.* 2022;29(6):973–989.e10. <https://doi.org/10.1016/j.stem.2022.04.008> PMID: 35659879
23. Walesky CM, Kolb KE, Winston CL, Henderson J, Kruff B, Fleming I, et al. Functional compensation precedes recovery of tissue mass following acute liver injury. *Nat Commun.* 2020;11(1):5785. <https://doi.org/10.1038/s41467-020-19558-3> PMID: 33214549
24. Sun X, Wu J, Liu L, Chen Y, Tang Y, Liu S, et al. Transcriptional switch of hepatocytes initiates macrophage recruitment and T-cell suppression in endotoxemia. *J Hepatol.* 2022;77(2):436–52. <https://doi.org/10.1016/j.jhep.2022.02.028> PMID: 35276271
25. Merrell AJ, Peng T, Li J, Sun K, Li B, Katsuda T, et al. Dynamic transcriptional and epigenetic changes drive cellular plasticity in the liver. *Hepatology.* 2021;74(1):444–57. <https://doi.org/10.1002/hep.31704> PMID: 33423324
26. Pepe-Mooney BJ, Dill MT, Alemany A, Ordovas-Montanes J, Matsushita Y, Rao A, et al. Single-cell analysis of the liver epithelium reveals dynamic heterogeneity and an essential role for YAP in homeostasis and regeneration. *Cell Stem Cell.* 2019;25(1):23–38.e8. <https://doi.org/10.1016/j.stem.2019.04.004> PMID: 31080134
27. Guillems M, Bonnardel J, Haest B, Vanderborcht B, Wagner C, Remmerie A, et al. Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell.* 2022;185(2):379–396.e38. <https://doi.org/10.1016/j.cell.2021.12.018> PMID: 35021063
28. Chembazhi UV, Bangru S, Hernaez M, Kalsotra A. Cellular plasticity balances the metabolic and proliferation dynamics of a regenerating liver. *Genome Res.* 2021;31(4):576–91. <https://doi.org/10.1101/gr.267013.120> PMID: 33649154
29. Chen T, Oh S, Gregory S, Shen X, Diehl AM. Single-cell omics analysis reveals functional diversification of hepatocytes during liver regeneration. *JCI Insight.* 2020;5(22):e141024. <https://doi.org/10.1172/jci.insight.141024> PMID: 33208554
30. Pose E, Sancho-Bru P, Coll M. 3,5-Diethoxycarbonyl-1,4-Dihydrocollidine diet: a rodent model in cholestasis research. *Methods Mol Biol.* 2019;1981:249–57. https://doi.org/10.1007/978-1-4939-9420-5_16 PMID: 31016659

31. Shang T, Jiang T, Cui X, Pan Y, Feng X, Dong L, et al. Diverse functions of SOX9 in liver development and homeostasis and hepatobiliary diseases. *Genes Dis.* 2023;11(4):100996. <https://doi.org/10.1016/j.gendis.2023.03.035> PMID: 38523677
32. Weiler P, Lange M, Klein M, Pe'er D, Theis F. CellRank 2: unified fate mapping in multiview single-cell data. *Nat Methods.* 2024;21(7):1196–205. <https://doi.org/10.1038/s41592-024-02303-9> PMID: 38871986
33. Zummo FP, Berthier A, Gheeraert C, Vinod M, Bobowski-Gérard M, Molendi-Coste O, et al. A time- and space-resolved nuclear receptor atlas in mouse liver. *J Mol Endocrinol.* 2023;71(1):e230017. <https://doi.org/10.1530/JME-23-0017> PMID: 36988391
34. Dusart P, Öling S, Struck E, Norreen-Thorsen M, Zwahlen M, von Feilitzen K, et al. A tissue centric atlas of cell type transcriptome enrichment signatures. *bioRxiv.* 2023.
35. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* 2022;50(W1):W216–W21. <https://doi.org/10.1093/nar/gkac194>
36. Breschi A, Muñoz-Aguirre M, Wucher V, Davis CA, Garrido-Martín D, Djebali S, et al. A limited set of transcriptional programs define major cell types. *Genome Res.* 2020;30(7):1047–59. <https://doi.org/10.1101/gr.263186.120> PMID: 32759341
37. Karlsson M, Zhang C, Méar L, Zhong W, Digre A, Katona B, et al. A single-cell type transcriptomics map of human tissues. *Sci Adv.* 2021;7(31):eab2169. <https://doi.org/10.1126/sciadv.abh2169> PMID: 34321199
38. Yanai I, Benjamin H, Shmoish M, Chalifa-Caspi V, Shklar M, Ophir R, et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics.* 2005;21(5):650–9. <https://doi.org/10.1093/bioinformatics/bti042> PMID: 15388519
39. Dubois-Chevalier J, Mazrooei P, Lupien M, Staels B, Lefebvre P, Eeckhoutte J. Organizing combinatorial transcription factor recruitment at cis-regulatory modules. *Transcription.* 2018;9(4):233–9. <https://doi.org/10.1080/21541264.2017.1394424> PMID: 29105538
40. Duncan AW, Taylor MH, Hickey RD, Hanlon Newell AE, Lenzi ML, Olson SB, et al. The ploidy conveyor of mature hepatocytes as a source of genetic variation. *Nature.* 2010;467(7316):707–10. <https://doi.org/10.1038/nature09414> PMID: 20861837
41. Roh HC, Tsai LT, Lyubetskaya A, Tenen D, Kumari M, Rosen ED. Simultaneous transcriptional and epigenomic profiling from specific cell types within heterogeneous tissues in vivo. *Cell Rep.* 2017;18(4):1048–61. Epub 2017/01/26. <https://doi.org/10.1016/j.celrep.2016.12.087> PMID: 28122230; PMCID: PMC5291126.
42. Nielsen CB, Younesy H, O'Geen H, Xu X, Jackson AR, Milosavljevic A, et al. Spark: a navigational paradigm for genomic data exploration. *Genome Res.* 2012;22(11):2262–9. <https://doi.org/10.1101/gr.140665.112> PMID: 22960372
43. Kumar D, Cinghu S, Oldfield AJ, Yang P, Jothi R. Decoding the function of bivalent chromatin in development and cancer. *Genome Res.* 2021;31(12):2170–84. <https://doi.org/10.1101/gr.275736.121> PMID: 34667120
44. Bravo González-Blas C, Matetovici I, Hillen H, Taskiran II, Vandepoel R, Christiaens V, et al. Single-cell spatial multi-omics and deep learning dissect enhancer-driven gene regulatory networks in liver zonation. *Nat Cell Biol.* 2024;26(1):153–67. <https://doi.org/10.1038/s41556-023-01316-4> PMID: 38182825
45. Zhang C, Macchi F, Magnani E, Sadler KC. Chromatin states shaped by an epigenetic code confer regenerative potential to the mouse liver. *Nat Commun.* 2021;12(1):4110. <https://doi.org/10.1038/s41467-021-24466-1> PMID: 34226551
46. Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. *Nat Rev Mol Cell Biol.* 2015;16(9):519–32. <https://doi.org/10.1038/nrm4043> PMID: 26296162
47. Grindheim JM, Nicetto D, Donahue G, Zaret KS. Polycomb repressive complex 2 proteins EZH1 and EZH2 regulate timing of postnatal hepatocyte maturation and fibrosis by repressing gene expression at promoter regions in euchromatin in mice. *Gastroenterology.* 2019. Epub 2019/01/29. <https://doi.org/10.1053/j.gastro.2019.01.041> PMID: 30689973
48. Becker JS, McCarthy RL, Sidoli S, Donahue G, Kaeding KE, He Z, et al. Genomic and proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Mol Cell.* 2017;68(6):1023–1037.e15. <https://doi.org/10.1016/j.molcel.2017.11.030> PMID: 29272703
49. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH roadmap epigenomics mapping consortium. *Nat Biotechnol.* 2010;28(10):1045–8. <https://doi.org/10.1038/nbt1010-1045> PMID: 20944595
50. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57–74. <https://doi.org/10.1038/nature11247> PMID: 22955616
51. Chittock EC, Latwiel S, Miller TCR, Müller CW. Molecular architecture of polycomb repressive complexes. *Biochem Soc Trans.* 2017;45(1):193–205. <https://doi.org/10.1042/BST20160173> PMID: 28202673
52. Holoch D, Margueron R. Mechanisms regulating PRC2 recruitment and enzymatic activity. *Trends Biochem Sci.* 2017;42(7):531–42. <https://doi.org/10.1016/j.tibs.2017.04.003> PMID: 28483375
53. Yang N, Ocean JR, Melters DP, Shi C, Wang L, Stransky S, et al. A hyper-quiescent chromatin state formed during aging is reversed by regeneration. *Mol Cell.* 2023;83(10):1659–1676.e11. <https://doi.org/10.1016/j.molcel.2023.04.005> PMID: 37116496
54. Watanabe K, Liu Y, Noguchi S, Murray M, Chang J-C, Kishima M, et al. OVOL2 induces mesenchymal-to-epithelial transition in fibroblasts and enhances cell-state reprogramming towards epithelial lineages. *Sci Rep.* 2019;9(1):6490. <https://doi.org/10.1038/s41598-019-43021-z> PMID: 31019211
55. Werner S, Frey S, Riethdorf S, Schulze C, Alawi M, Kling L, et al. Dual roles of the transcription factor grainyhead-like 2 (GRHL2) in breast cancer. *J Biol Chem.* 2013;288(32):22993–3008. <https://doi.org/10.1074/jbc.M113.456293> PMID: 23814079

56. Aue A, Hinze C, Walentin K, Ruffert J, Yurtdas Y, Werth M, et al. A grainyhead-like 2/Ovo-Like 2 pathway regulates renal epithelial barrier function and lumen expansion. *J Am Soc Nephrol*. 2015;26(11):2704–15. <https://doi.org/10.1681/ASN.2014080759> PMID: 25788534
57. Wang Z, Coban B, Wu H, Chouaref J, Daxinger L, Paulsen MT, et al. GRHL2-controlled gene expression networks in luminal breast cancer. *Cell Commun Signal*. 2023;21(1):15. <https://doi.org/10.1186/s12964-022-01029-5> PMID: 36691073
58. Chen AF, Liu AJ, Krishnakumar R, Freimer JW, DeVeale B, Blelloch R. GRHL2-dependent enhancer switching maintains a pluripotent stem cell transcriptional subnetwork after exit from naive pluripotency. *Cell Stem Cell*. 2018;23(2):226–238.e4. <https://doi.org/10.1016/j.stem.2018.06.005> PMID: 30017589
59. Jacobs J, Atkins M, Davie K, Imrichova H, Romanelli L, Christiaens V, et al. The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat Genet*. 2018;50(7):1011–20. <https://doi.org/10.1038/s41588-018-0140-x> PMID: 29867222
60. Strick-Marchand H, Weiss MC. Inducible differentiation and morphogenesis of bipotential liver cell lines from wild-type mouse embryos. *Hepatology*. 2002;36(4 Pt 1):794–804. <https://doi.org/10.1053/jhep.2002.36123> PMID: 12297826
61. Duplaquet L, Leroy C, Vincent A, Paget S, Lefebvre J, Vanden Abeele F, et al. Control of cell death/survival balance by the MET dependence receptor. *Elife*. 2020;9:e50041. <https://doi.org/10.7554/eLife.50041> PMID: 32091387
62. Senga K, Mostov KE, Mitaka T, Miyajima A, Tanimizu N. Grainyhead-like 2 regulates epithelial morphogenesis by establishing functional tight junctions through the organization of a molecular network among claudin3, claudin4, and Rab25. *Mol Biol Cell*. 2012;23(15):2845–55. <https://doi.org/10.1091/mbc.E12-02-0097> PMID: 22696678
63. Marsee A, Roos FJM, Verstegen MMA, HPB Organoid Consortium, Gehart H, de Koning E, et al. Building consensus on definition and nomenclature of hepatic, pancreatic, and biliary organoids. *Cell Stem Cell*. 2021;28(5):816–32. <https://doi.org/10.1016/j.stem.2021.04.005> PMID: 33961769
64. Gao X, Vockley CM, Pauli F, Newberry KM, Xue Y, Randell SH, et al. Evidence for multiple roles for grainyhead-like 2 in the establishment and maintenance of human mucociliary airway epithelium.[corrected]. *Proc Natl Acad Sci U S A*. 2013;110(23):9356–61. <https://doi.org/10.1073/pnas.1307589110> PMID: 23690579
65. Paltoglou S, Das R, Townley SL, Hickey TE, Tarulli GA, Coutinho I, et al. Novel androgen receptor coregulator GRHL2 exerts both oncogenic and antimetastatic functions in prostate cancer. *Cancer Res*. 2017;77(13):3417–30. <https://doi.org/10.1158/0008-5472.CAN-16-1616> PMID: 28473532
66. Holding AN, Giorgi FM, Donnelly A, Cullen AE, Nagarajan S, Selth LA, et al. VULCAN integrates ChIP-seq with patient-derived co-expression networks to identify GRHL2 as a key co-regulator of ERα at enhancers in breast cancer. *Genome Biol*. 2019;20(1):91. <https://doi.org/10.1186/s13059-019-1698-z> PMID: 31084623
67. Chung VY, Tan TZ, Tan M, Wong MK, Kuay KT, Yang Z, et al. GRHL2-miR-200-ZEB1 maintains the epithelial status of ovarian cancer through transcriptional regulation and histone modification. *Sci Rep*. 2016;6:19943. <https://doi.org/10.1038/srep19943> PMID: 26887977
68. Jacobs J, Atkins M, Davie K, Imrichova H, Romanelli L, Christiaens V, et al. The transcription factor Grainy head primes epithelial enhancers for spatiotemporal activation by displacing nucleosomes. *Nat Genet*. 2018;50(7):1011–20. <https://doi.org/10.1038/s41588-018-0140-x> PMID: 29867222
69. Li S, Wan C, Zheng R, Fan J, Dong X, Meyer CA, et al. Cistrome-GO: a web server for functional enrichment analysis of transcription factor ChIP-seq peaks. *Nucleic Acids Res*. 2019;47(W1):W206–11. <https://doi.org/10.1093/nar/gkz332> PMID: 31053864
70. Xia M, Wu M, Li Y, Liu Y, Jia G, Lou Y, et al. Varying mechanical forces drive sensory epithelium formation. *Sci Adv*. 2023;9(44):eadf2664. <https://doi.org/10.1126/sciadv.adf2664> PMID: 37922362
71. Romagnoli M, Bresson L, Di-Cicco A, Pérez-Lanzón M, Legoix P, Baulande S, et al. Laminin-binding integrins are essential for the maintenance of functional mammary secretory epithelium in lactation. *Development*. 2020;147(4):dev181552. <https://doi.org/10.1242/dev.181552> PMID: 31988184
72. Buckley CE, St Johnston D. Apical-basal polarity and the control of epithelial form and function. *Nat Rev Mol Cell Biol*. 2022;23(8):559–77. <https://doi.org/10.1038/s41580-022-00465-y> PMID: 35440694
73. Karantz V. Keratins in health and cancer: more than mere epithelial cell markers. *Oncogene*. 2011;30(2):127–38. <https://doi.org/10.1038/onc.2010.456> PMID: 20890307
74. Warzecha CC, Jiang P, Amirikian K, Dittmar KA, Lu H, Shen S, et al. An ESRP-regulated splicing programme is abrogated during the epithelial-mesenchymal transition. *EMBO J*. 2010;29(19):3286–300. <https://doi.org/10.1038/emboj.2010.195> PMID: 20711167
75. Gripon P, Rumin S, Urban S, Le Seyec J, Glaise D, Canine I, et al. Infection of a human hepatoma cell line by hepatitis B virus. *Proc Natl Acad Sci U S A*. 2002;99(24):15655–60. <https://doi.org/10.1073/pnas.232137699> PMID: 12432097
76. Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, et al. Transdifferentiation of hepatocyte-like cells from the human hepatoma HepaRG cell line through bipotent progenitor. *Hepatology*. 2007;45(4):957–67. <https://doi.org/10.1002/hep.21536> PMID: 17393521
77. Broutier L, Andersson-Rolf A, Hindley CJ, Boj SF, Clevers H, Koo B-K, et al. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nat Protoc*. 2016;11(9):1724–43. <https://doi.org/10.1038/nprot.2016.097> PMID: 27560176
78. Ntandja Wandji L, Bousaleh M, Sobolewski C, El Amrani M, Bray F, Boleslawski E, et al. Aberrant YAP activation in hepatocytes influences myofibroblast activation to induce a specific fibrosis profile in alcohol-related hepatitis. *JHEP Rep*. Forthcoming 2025.
79. Zhang Y, Li F, Patterson AD, Wang Y, Krausz KW, Neale G, et al. Abcb11 deficiency induces cholestasis coupled to impaired β-fatty acid oxidation in mice. *J Biol Chem*. 2012;287(29):24784–94. <https://doi.org/10.1074/jbc.M111.329318> PMID: 22619174

80. Cho J-Y, Matsubara T, Kang DW, Ahn S-H, Krausz KW, Idle JR, et al. Urinary metabolomics in Fxr-null mice reveals activated adaptive metabolic pathways upon bile acid challenge. *J Lipid Res.* 2010;51(5):1063–74. <https://doi.org/10.1194/jlr.M002923> PMID: [19965603](#)
81. Hyun J, Sun Z, Ahmadi AR, Bangru S, Chembazhi UV, Du K, et al. Epithelial splicing regulatory protein 2-mediated alternative splicing reprograms hepatocytes in severe alcoholic hepatitis. *J Clin Invest.* 2020;130(4):2129–45. <https://doi.org/10.1172/JCI132691> PMID: [31945016](#)
82. Aguilar-Bravo B, Ariño S, Blaya D, Pose E, Martínez García de la Torre RA, Latasa MU, et al. Hepatocyte dedifferentiation profiling in alcohol-related liver disease identifies CXCR4 as a driver of cell reprogramming. *J Hepatol.* 2023;79(3):728–40. <https://doi.org/10.1016/j.jhep.2023.04.013> PMID: [37088308](#)
83. Hammond NL, Murtuza Baker S, Georgaka S, Al-Anbaki A, Jokl E, Simpson K, et al. Spatial gene regulatory networks driving cell state transitions during human liver disease. *EMBO Mol Med.* 2025;17(6):1452–74. <https://doi.org/10.1038/s44321-025-00230-6> PMID: [40281306](#)
84. He Q, Cui L, Yuan X, Wang M, Hui L. Cell identity conversion in liver regeneration after injury. *Curr Opin Genet Dev.* 2022;75:101921. Epub 2022/06/02. <https://doi.org/10.1016/j.gde.2022.101921> PMID: [35644120](#)
85. Kamath B, Mack C. From hepatocyte to cholangiocyte: the remarkable potential of transdifferentiation to treat cholestatic diseases. *Hepatology.* 2019;69(4):1828–30. <https://doi.org/10.1002/hep.30250> PMID: [30179266](#)
86. Argemi J, Latasa MU, Atkinson SR, Blokhin IO, Massey V, Gue JP, et al. Defective HNF4alpha-dependent gene expression as a driver of hepatocellular failure in alcoholic hepatitis. *Nat Commun.* 2019;10(1):3126. <https://doi.org/10.1038/s41467-019-11004-3> PMID: [31311938](#)
87. Lan T, Tai Y, Zhao C, Xiao Y, Yang Z, Zhang L, et al. Atypical cholangiocytes derived from hepatocyte-cholangiocyte transdifferentiation mediated by COX-2: a kind of misguided liver regeneration. *Inflamm Regen.* 2023;43(1):37. <https://doi.org/10.1186/s41232-023-00284-4> PMID: [37452426](#)
88. Ahmad US, Uttagomol J, Wan H. The regulation of the hippo pathway by intercellular junction proteins. *Life (Basel).* 2022;12(11):1792. <https://doi.org/10.3390/life12111792> PMID: [36362947](#)
89. Molina L, Nejak-Bowen K, Monga SP. Role of YAP1 signaling in biliary development, repair, and disease. *Semin Liver Dis.* 2022;42(1):17–33. <https://doi.org/10.1055/s-0041-1742277> PMID: [35073587](#)
90. Lamba A, Zernicka-Goetz M. The role of polarization and early heterogeneities in the mammalian first cell fate decision. *Curr Top Dev Biol.* 2023;154:169–96. <https://doi.org/10.1016/bs.ctdb.2023.02.006> PMID: [37100517](#)
91. Bebelman MP, Belicova L, Gralinska E, Jumel T, Lahree A, Sommer S, et al. Hepatocyte differentiation requires anisotropic expansion of bile canaliculi. *Development.* 2024;151(22):dev202777. <https://doi.org/10.1242/dev.202777> PMID: [39373104](#)
92. Reehorst CM, Nightingale R, Luk IY, Jenkins L, Koentgen F, Williams DS, et al. EHF is essential for epidermal and colonic epithelial homeostasis, and suppresses Apc-initiated colonic tumorigenesis. *Development.* 2021;148(12):dev199542. <https://doi.org/10.1242/dev.199542> PMID: [34180969](#)
93. Bebee TW, Park JW, Sheridan KI, Warzecha CC, Cieply BW, Rohacek AM, et al. The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development. *Elife.* 2015;4:e08954. <https://doi.org/10.7554/eLife.08954> PMID: [26371508](#)
94. Hrnčir HR, Gracz AD. Cellular and transcriptional heterogeneity in the intrahepatic biliary epithelium. *Gastro Hep Adv.* 2023;2(1):108–20. <https://doi.org/10.1016/j.gastha.2022.07.015> PMID: [36593993](#)
95. Ogawa M, Jiang J-X, Xia S, Yang D, Ding A, Laselva O, et al. Generation of functional ciliated cholangiocytes from human pluripotent stem cells. *Nat Commun.* 2021;12(1):6504. <https://doi.org/10.1038/s41467-021-26764-0> PMID: [34764255](#)
96. Katsuda T, Sussman JH, Ito K, Katznelson A, Yuan S, Takenaka N, et al. Cellular reprogramming in vivo initiated by SOX4 pioneer factor activity. *Nat Commun.* 2024;15(1):1761. <https://doi.org/10.1038/s41467-024-45939-z> PMID: [38409161](#)
97. Satishchandran A, Ambade A, Rao S, Hsueh YC, Iracheta-Vellve A, Tornai D, et al. MicroRNA 122, regulated by GRLH2, protects livers of mice and patients from ethanol-induced liver disease. *Gastroenterology.* 2018;154(1):238–52 e7. Epub 2017/10/11. <https://doi.org/10.1053/j.gastro.2017.09.022> PMID: [28987423](#); PMCID: [PMC5742049](#).
98. Tanimizu N, Kobayashi S, Ichinohe N, Mitaka T. Downregulation of miR122 by grainyhead-like 2 restricts the hepatocytic differentiation potential of adult liver progenitor cells. *Development.* 2014;141(23):4448–56. <https://doi.org/10.1242/dev.113654> PMID: [25406394](#)
99. Sussman JH, Cure HW, Yuan S, Ito K, Asangani IA, Garcia BA, et al. In vivo CRISPR screening reveals epigenetic regulators of hepatobiliary plasticity. *Genes Dev.* 2025;39(9–10):603–16. <https://doi.org/10.1101/gad.352420.124> PMID: [40169232](#)
100. Bae WK, Kang K, Yu JH, Yoo KH, Factor VM, Kaji K, et al. The methyltransferases enhancer of zeste homolog (EZH) 1 and EZH2 control hepatocyte homeostasis and regeneration. *FASEB J.* 2015;29(5):1653–62. <https://doi.org/10.1096/fj.14-261537> PMID: [25477280](#)
101. Holoch D, Wassef M, Lövkvist C, Zielinski D, Aflaki S, Lombard B, et al. A cis-acting mechanism mediates transcriptional memory at Polycomb target genes in mammals. *Nat Genet.* 2021;53(12):1686–97. <https://doi.org/10.1038/s41588-021-00964-2> PMID: [34782763](#)
102. Yang L, Wang X, Yu X-X, Yang L, Zhou B-C, Yang J, et al. The default and directed pathways of hepatoblast differentiation involve distinct epigenomic mechanisms. *Dev Cell.* 2023;58(18):1688–1700.e6. <https://doi.org/10.1016/j.devcel.2023.07.002> PMID: [37490911](#)
103. Kusakabe Y, Chiba T, Oshima M, Koide S, Rizq O, Aoyama K, et al. EZH1/2 inhibition augments the anti-tumor effects of sorafenib in hepatocellular carcinoma. *Sci Rep.* 2021;11(1):21396. <https://doi.org/10.1038/s41598-021-00889-0> PMID: [34725436](#)
104. Duan R, Du W, Guo W. EZH2: a novel target for cancer treatment. *J Hematol Oncol.* 2020;13(1):104. <https://doi.org/10.1186/s13045-020-00937-8> PMID: [32723346](#)

105. Laurent V, Fraix A, Montier T, Cammas-Marion S, Ribault C, Benvegna T, et al. Highly efficient gene transfer into hepatocyte-like HepaRG cells: new means for drug metabolism and toxicity studies. *Biotechnol J*. 2010;5(3):314–20. <https://doi.org/10.1002/biot.200900255> PMID: [20213646](#)
106. Louvet A, Labreuche J, Artru F, Bouthors A, Rolland B, Saffers P, et al. Main drivers of outcome differ between short term and long term in severe alcoholic hepatitis: a prospective study. *Hepatology*. 2017;66(5):1464–73. <https://doi.org/10.1002/hep.29240> PMID: [28459138](#)
107. Louvet A, Labreuche J, Artru F, Boursier J, Kim DJ, O'Grady J, et al. Combining data from liver disease scoring systems better predicts outcomes of patients with alcoholic hepatitis. *Gastroenterology*. 2015;149(2):398–406.e8; quiz e16–7. <https://doi.org/10.1053/j.gastro.2015.04.044> PMID: [25935634](#)
108. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002;30(1):207–10. <https://doi.org/10.1093/nar/30.1.207> PMID: [11752295](#)
109. Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res*. 2018;46(D1):D794–801. <https://doi.org/10.1093/nar/gkx1081> PMID: [29126249](#)
110. Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518(7539):317–30. <https://doi.org/10.1038/nature14248> PMID: [25693563](#)
111. Taing L, Dandawate A, L'Yi S, Gehlenborg N, Brown M, Meyer CA. Cistrome Data Browser: integrated search, analysis and visualization of chromatin data. *Nucleic Acids Res*. 2024;52(D1):D61–D6. Epub 2023/11/17. <https://doi.org/10.1093/nar/gkad1069> PMID: [37971305](#); PMCID: [PMC10767960](#).
112. Hu H, Miao Y-R, Jia L-H, Yu Q-Y, Zhang Q, Guo A-Y. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res*. 2019;47(D1):D33–8. <https://doi.org/10.1093/nar/gky822> PMID: [30204897](#)
113. Wang J, Zhao Y, Zhou X, Hiebert SW, Liu Q, Shyr Y. Nascent RNA sequencing analysis provides insights into enhancer-mediated gene regulation. *BMC Genomics*. 2018;19(1):633. <https://doi.org/10.1186/s12864-018-5016-z> PMID: [30139328](#)
114. Haeussler M, Zweig AS, Tyner C, Speir ML, Rosenbloom KR, Raney BJ, et al. The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res*. 2019;47(D1):D853–8. <https://doi.org/10.1093/nar/gky1095> PMID: [30407534](#)
115. Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res*. 2011;39(Database issue):D19–21. <https://doi.org/10.1093/nar/gkq1019> PMID: [21062823](#)
116. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative bio-medical analyses: 2018 update. *Nucleic Acids Res*. 2018;46(W1):W537–44. <https://doi.org/10.1093/nar/gky379> PMID: [29790989](#)
117. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60. <https://doi.org/10.1038/nmeth.3317> PMID: [25751142](#)
118. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9. <https://doi.org/10.1093/bioinformatics/btu638> PMID: [25260700](#)
119. R_Core_Team. R: a language and environment for statistical computing. 4.3.1 ed. Vienna, Austria: R Foundation for Statistical Computing; 2023.
120. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: [25516281](#)
121. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amodè MR, Armean IM, et al. Ensembl 2022. *Nucleic Acids Res*. 2022;50(D1):D988–95. <https://doi.org/10.1093/nar/gkab1049> PMID: [34791404](#)
122. Wickham H. ggplot2: elegant graphics for data analysis. Springer-Verlag New York; 2016.
123. Gu Z. Complex heatmap visualization. *Imeta*. 2022;1(3):e43. <https://doi.org/10.1002/imt2.43> PMID: [38868715](#)
124. 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(6):882–3. <https://doi.org/10.1093/bioinformatics/bts034> PMID: [22257669](#)
125. Lê S, Josse J, Huisson F. FactoMineR: an R package for multivariate analysis. *J Stat Software*. 2008;25(1):1–18.
126. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523. <https://doi.org/10.1038/s41467-019-09234-6> PMID: [30944313](#)
127. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50. <https://doi.org/10.1073/pnas.0506580102> PMID: [16199517](#)
128. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048> PMID: [34062119](#)
129. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol*. 2019;20(1):296. <https://doi.org/10.1186/s13059-019-1874-1> PMID: [31870423](#)
130. Andreatta M, Hérault L, Gueguen P, Gfeller D, Berenstein AJ, Carmona SJ. Semi-supervised integration of single-cell transcriptomics data. *Nat Commun*. 2024;15(1):872. <https://doi.org/10.1038/s41467-024-45240-z> PMID: [38287014](#)
131. Guille L, Johanns M, Zummo FP, Staels B, Lefebvre P, Eeckhoutte J, et al. ISCEBERG: interactive single cell expression browser for exploration of RNAseq data using graphics (v1.0.1). Zenodo. 2022. <https://doi.org/10.5281/zenodo.6563734>

132. Srivastava A, Malik L, Smith T, Sudbery I, Patro R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biol.* 2019;20(1):65. <https://doi.org/10.1186/s13059-019-1670-y> PMID: [30917859](https://pubmed.ncbi.nlm.nih.gov/30917859/)
133. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol.* 2020;38(12):1408–14. <https://doi.org/10.1038/s41587-020-0591-3> PMID: [32747759](https://pubmed.ncbi.nlm.nih.gov/32747759/)
134. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N, et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics.* 2018;19(1):477. <https://doi.org/10.1186/s12864-018-4772-0> PMID: [29914354](https://pubmed.ncbi.nlm.nih.gov/29914354/)
135. Aragon T. epitools: Epidemiology Tools. R package version 05–101. 2020.
136. Camargo AP, Vasconcelos AA, Fiamenghi MB, Pereira GAG, Carazzolle MF. tspex: a tissue-specificity calculator for gene expression data. Springer Science and Business Media LLC; 2020. <https://doi.org/10.21203/rs.3.rs-51998/v1>
137. Zhang H, Song L, Wang X, Cheng H, Wang C, Meyer CA, et al. Fast alignment and preprocessing of chromatin profiles with Chromap. *Nat Commun.* 2021;12(1):6566. <https://doi.org/10.1038/s41467-021-26865-w> PMID: [34772935](https://pubmed.ncbi.nlm.nih.gov/34772935/)
138. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* 2016;44(W1):W160–5. <https://doi.org/10.1093/nar/gkw257> PMID: [27079975](https://pubmed.ncbi.nlm.nih.gov/27079975/)
139. Hitz BC, Jin-Wook L, Jolanki O, Kagda MS, Graham K, Sud P, et al. The ENCODE uniform analysis pipelines. *bioRxiv.* 2023;:2023.04.04.535623. <https://doi.org/10.1101/2023.04.04.535623> PMID: [37066421](https://pubmed.ncbi.nlm.nih.gov/37066421/)
140. Freese NH, Norris DC, Loraine AE. Integrated genome browser: visual analytics platform for genomics. *Bioinformatics.* 2016;32(14):2089–95. <https://doi.org/10.1093/bioinformatics/btw069> PMID: [27153568](https://pubmed.ncbi.nlm.nih.gov/27153568/)
141. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 2013;14(2):178–92. <https://doi.org/10.1093/bib/bbs017> PMID: [22517427](https://pubmed.ncbi.nlm.nih.gov/22517427/)
142. Lange M, Bergen V, Klein M, Setty M, Reuter B, Bakhti M, et al. CellRank for directed single-cell fate mapping. *Nat Methods.* 2022;19(2):159–70. <https://doi.org/10.1038/s41592-021-01346-6> PMID: [35027767](https://pubmed.ncbi.nlm.nih.gov/35027767/)

Fig.S1

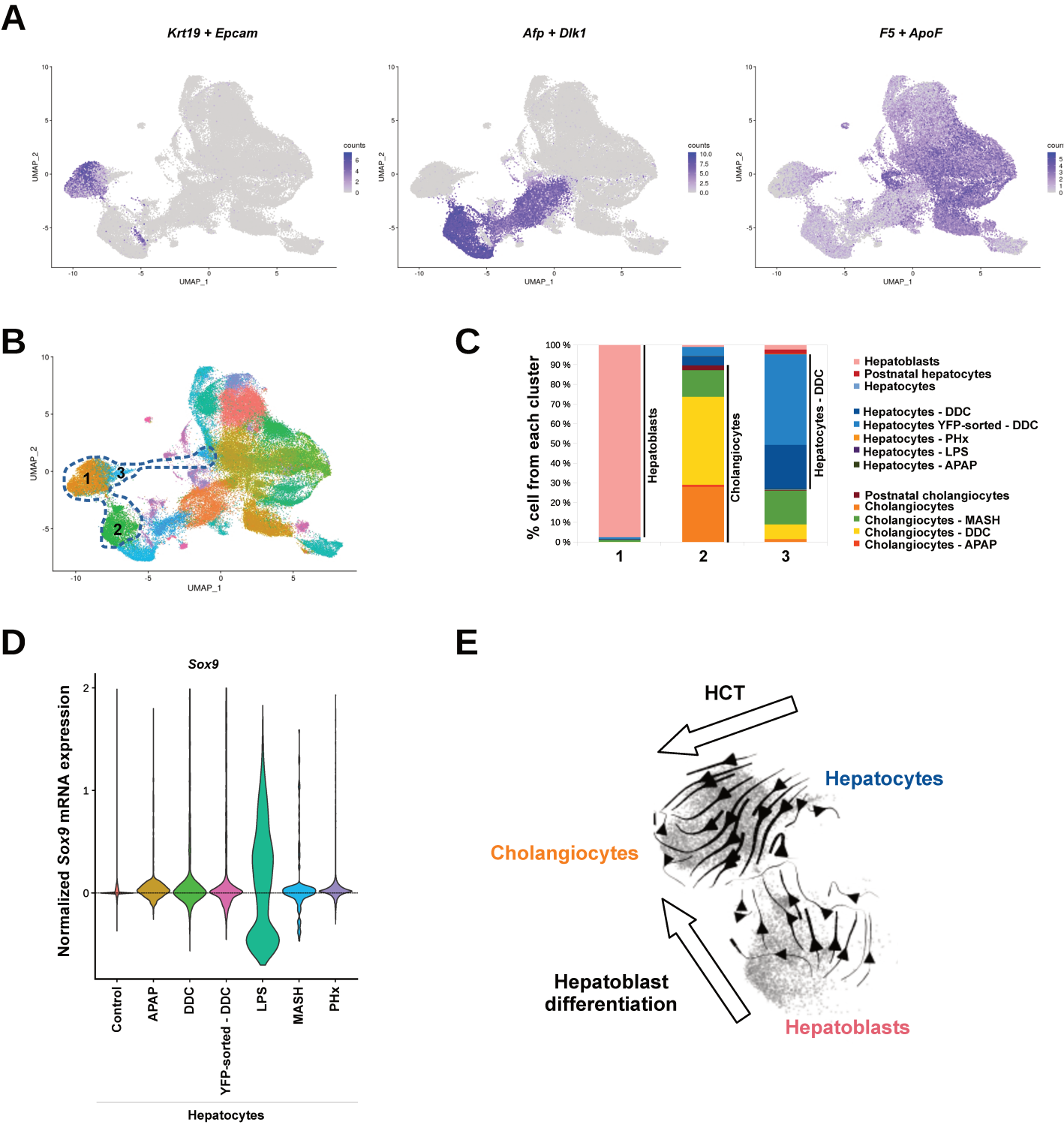
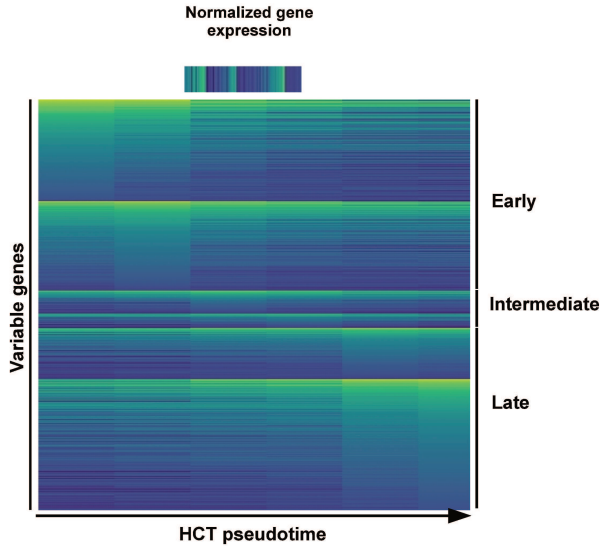
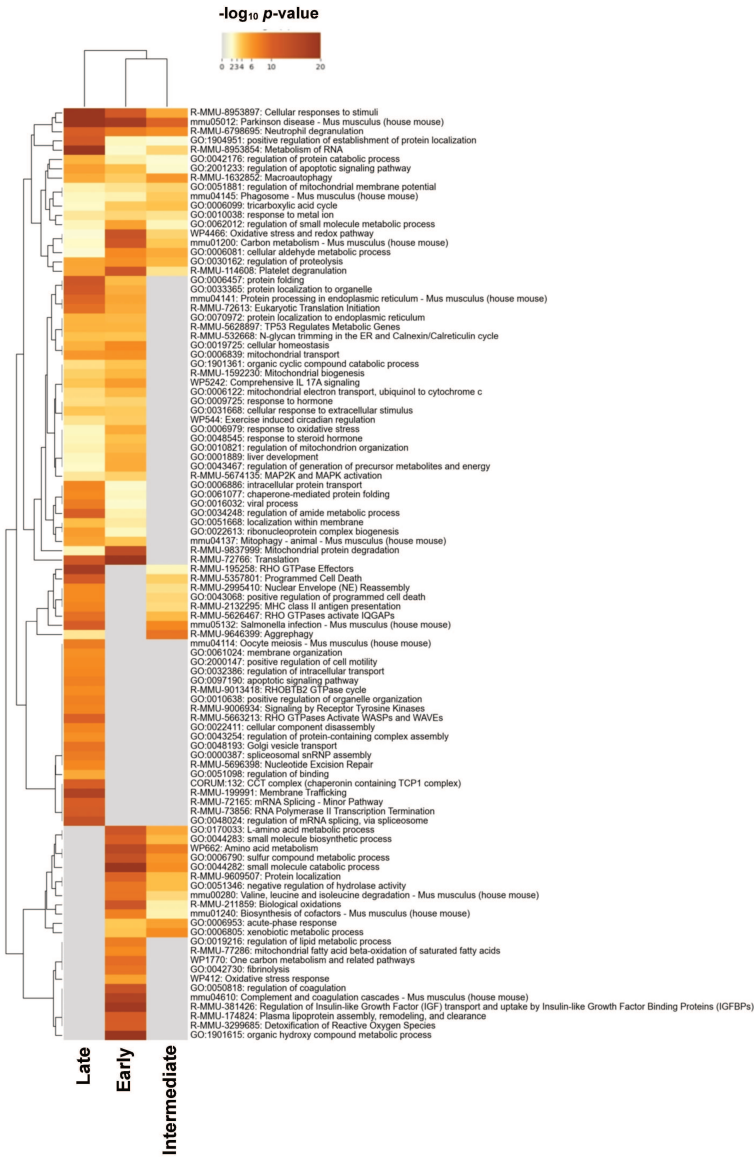


Fig.S1

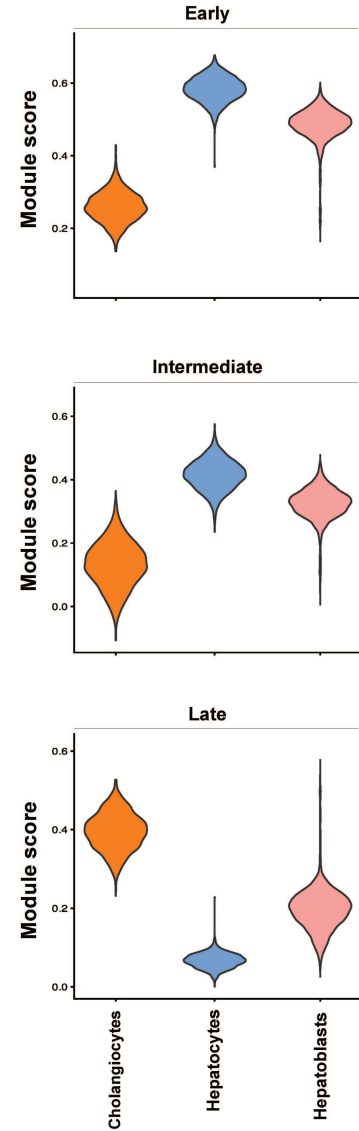
F



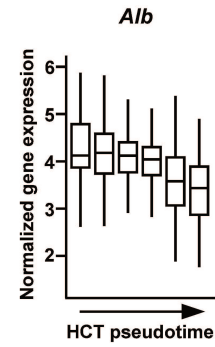
G



H



I



Supplementary Fig.1: Additional characterization of the mouse parenchymal liver cells scRNA-seq atlas and of the cellular trajectory involved in HCT

(A) UMAP from Fig.1A was used to project expression of the indicated markers genes: *Afp* and *Dlk1* for hepatoblasts, *F5* and *ApoF* for hepatocytes and *Krt19* and *Epcam* for cholangiocytes. Cumulative expression of the 2 marker genes was used.

(B) UMAP from Fig.1A was used to display the result of cell clustering and the 3 selected clusters from Fig.1B were further highlighted.

(C) Dataset of origin of hepatocytes from clusters 1-3. Horizontal bars were added to highlight the main cell type in each one of the 3 clusters.

(D) Violin plots showing the normalized mRNA expression levels of *Sox9* in the indicated subsets of mature hepatocytes.

(E) An area from Fig.1D centered on cholangiocytes is shown to highlight cell transitions predicted by RNA velocity analyses and associated with hepatoblast differentiation or HCT.

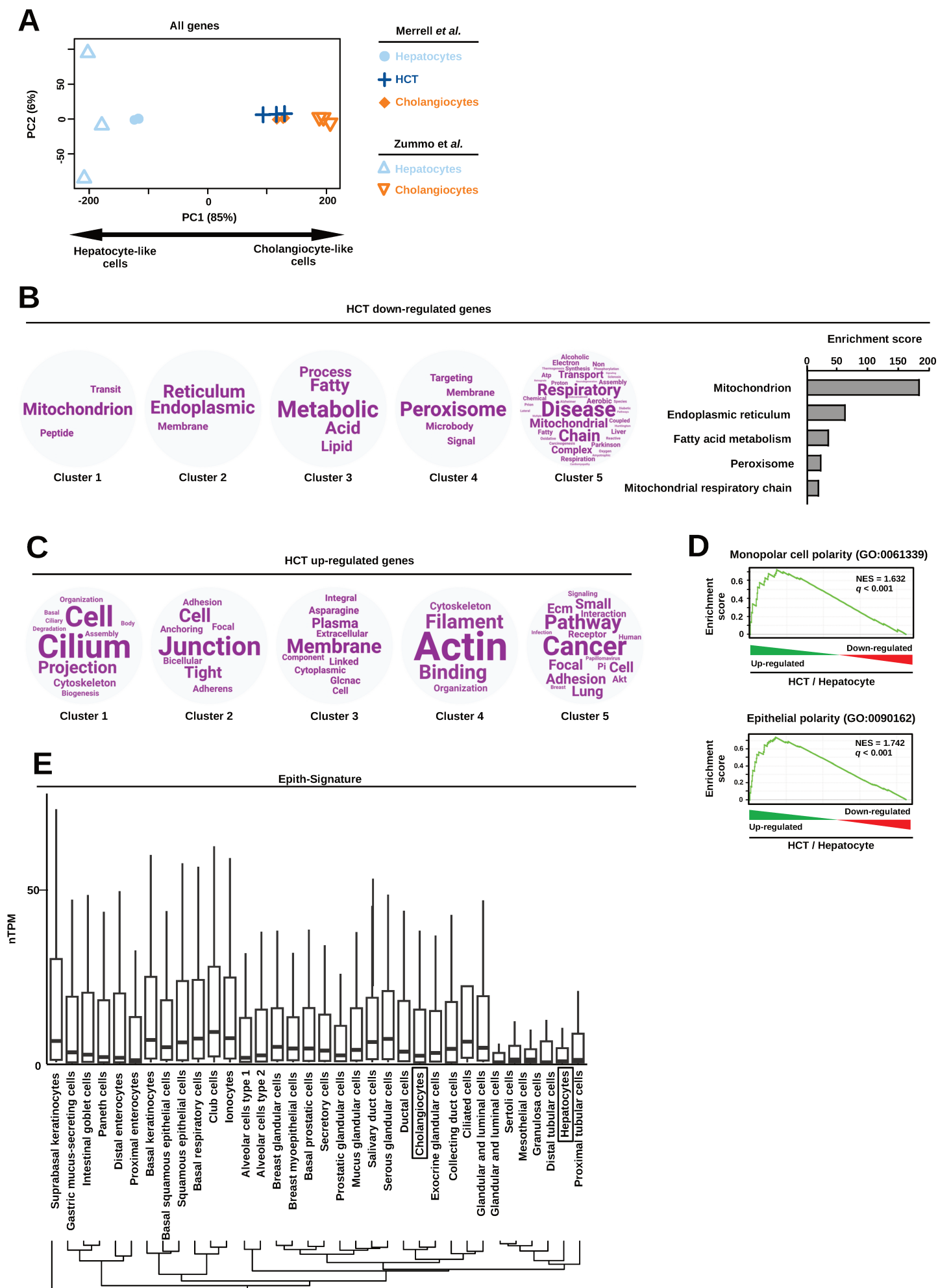
(F) Heatmap showing the normalized genes expression of variable genes along the course of HCT pseudotime. Genes were ranked according to the time at which they were the most highly expressed ranging from early to late. This allowed to discriminate 3 groups of genes indicated as early, intermediate and late.

(G) Heatmap showing the results of biological term enrichments analyses performed using Metascape for early, intermediate and late genes identified in panel F. The dendrogram on the left shows the hierarchical clustering based on $-\log_{10} p$ -values as provided by Metascape.

(H) Cells used for RNA velocity analyses were subsequently utilized to mine expression of early, intermediate and late genes identified in panel F in hepatoblasts, healthy hepatocytes and cholangiocytes.

(I) Boxplot showing the normalized expression of *Alb* in discrete hepatocyte cell subsets along the course of HCT as defined using pseudotime analysis similarly to Fig.1G.

Fig. S2



Supplementary Fig.2: Additional characterization of the gene transcriptional changes occurring in HCT

(A) Visualization of all RNA-seq datasets used in the PCA displayed in [Fig.2A](#). In addition to data from (Merrell et al. 2021) shown in [Fig.2A](#), the healthy hepatocytes and cholangiocytes data from (Zummo et al. 2023) used to initially perform the PCA (see Materials and Methods) are also shown here.

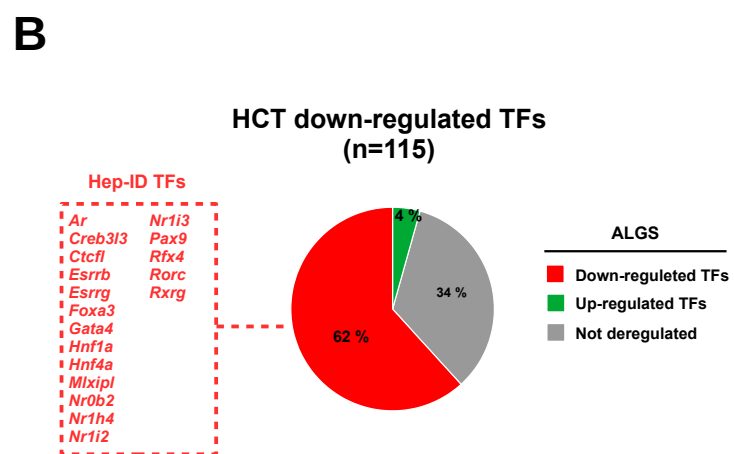
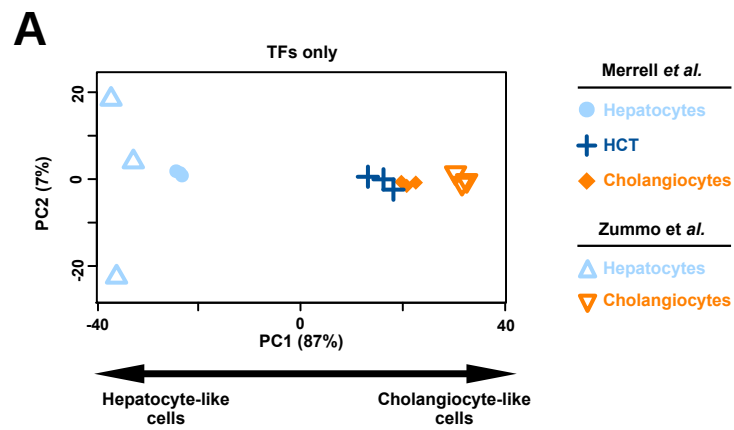
(B) Functional annotation of genes down-regulated during HCT performed similarly to [Fig.2B](#). Word clouds show the results of word counting performed using the name of individual enriched terms in each of the 5 top clusters identified by the DAVID tool. The size of each word is proportional to its number of occurrences. These analyses were used to name these 5 clusters, whose enrichment scores (defined as the best enrichment score among individual terms within each cluster) are shown using a bar graph (right).

(C) Word clouds showing the results of word counting performed using the name of individual enriched terms in each of the 5 top clusters identified by the DAVID tool when mining HCT-up-regulated genes. The size of each word is proportional to its number of occurrences. These analyses were used to name the 5 clusters in [Fig.2B](#).

(D) Enrichment plots obtained using GSEA performed with the gene ontology terms “Establishment or maintenance of monopolar cell polarity” (GO:0061339) or “Establishment of epithelial cell polarity” (GO:0090162) as the gene sets and transcriptomic changes induced by HCT defined as in [Fig.2B](#) from (Merrell et al. 2021). NES stands for normalized enrichment score.

(E) Box plot showing the expression of genes from the Epith-Signature (Breschi et al. 2020) in each indicated individual epithelial cell type. The dendrogram at the bottom shows the results of hierarchical clustering analysis performed on these data. Cholangiocytes and hepatocytes are highlighted using frames. nTPM stands for normalized transcripts per million.

Fig.S3

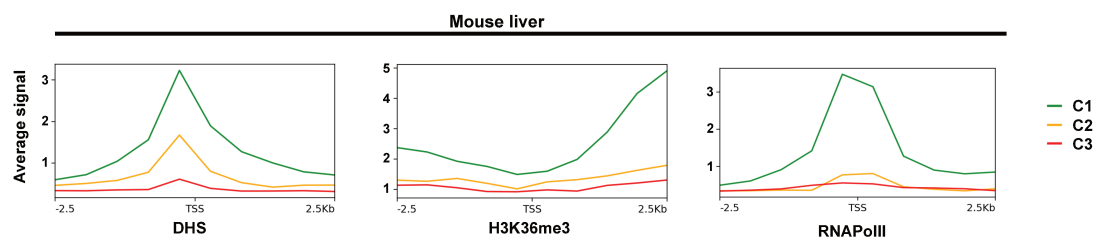


Supplementary Fig.3: Additional characterization of deregulated TF-encoding genes during HCT

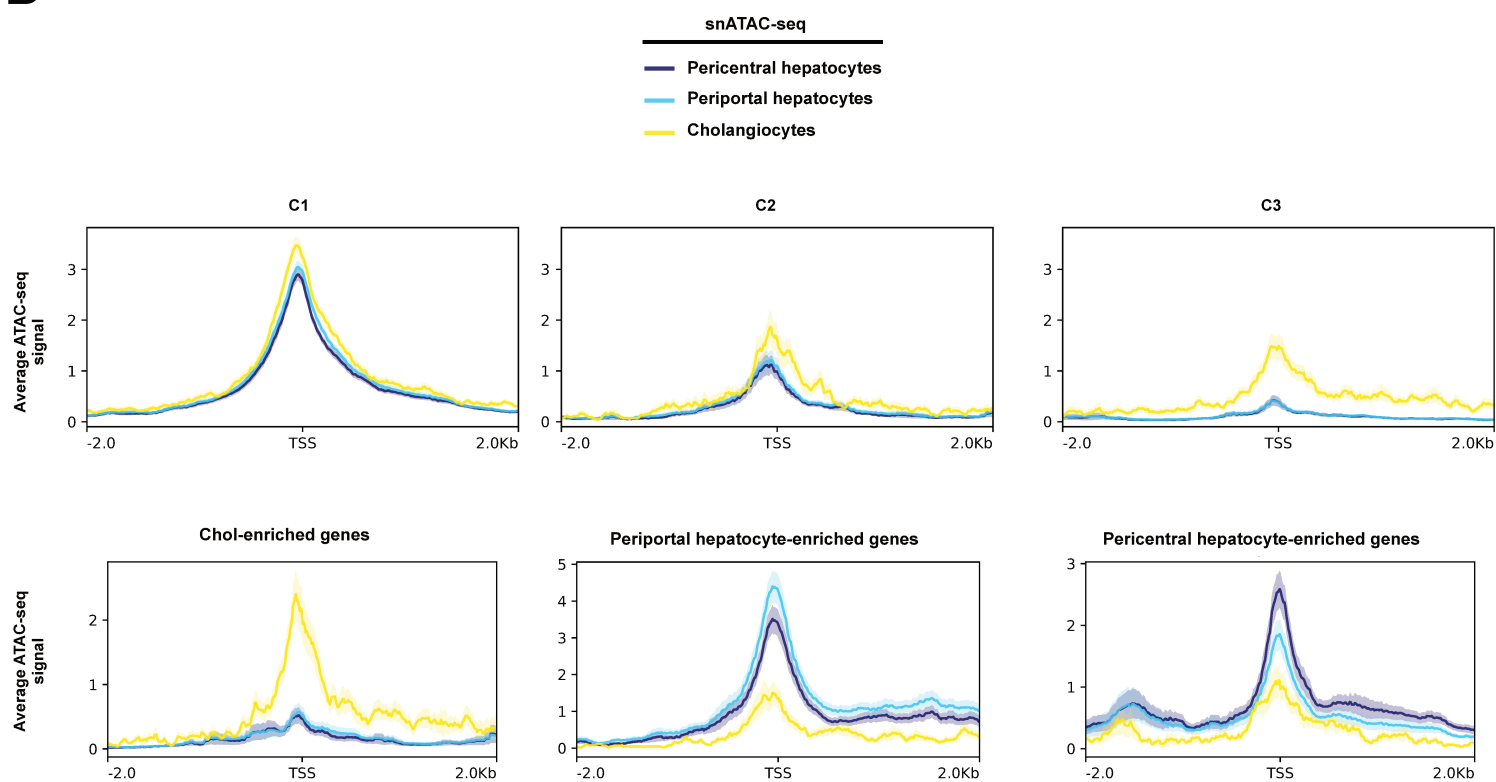
(A) Visualization of all RNA-seq datasets used in the PCA displayed in [Fig.3A](#). In addition to data from (Merrell et al. 2021) shown in [Fig.3A](#), the healthy hepatocytes and cholangiocytes data from (Zummo et al. 2023) used to initially perform the PCA (see Materials and Methods) are also shown here.

(B) TF-encoding genes down-regulated during HCT were obtained by comparing bulk RNA-seq data from lineage-traced hepatocytes undergoing HCT from DDC-treated mice compared to that of control hepatocytes from healthy livers from (Merrell et al. 2021) (q -value < 0.05 and $\log_2 \text{FC} < 0$). The obtained gene list was monitored for changes in hepatocytes undergoing HCT in another mouse model, i.e. model of ALGS lacking an intrahepatic biliary system at birth (Schaub et al. 2018). In this second dataset, genes with q -value < 0.15 and $\log_2 \text{FC} < 0$ or > 0 were defined as down- or up-regulated, respectively. Genes with q -value > 0.15 were considered as “Not deregulated”. Hepatocyte identity (Hep-ID) TFs, as defined in our previous study (Dubois-Chevalier et al. 2023), which were down-regulated both in DDC- and ALGS-related HCT are indicated in red.

A



B

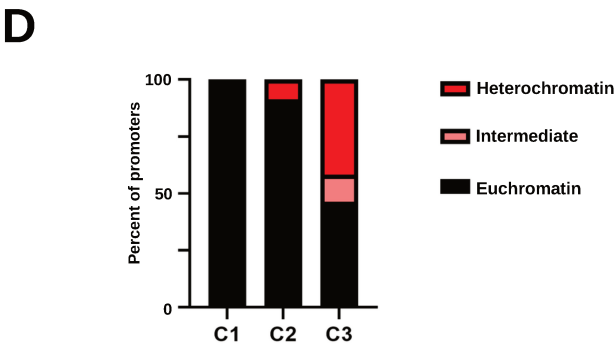
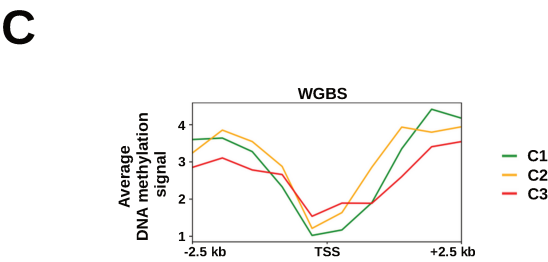
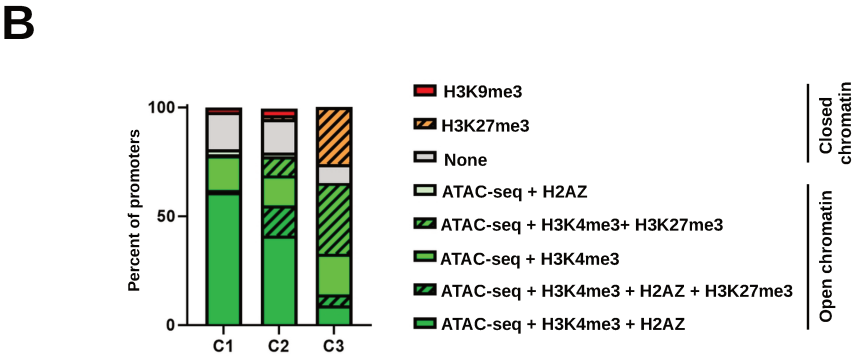
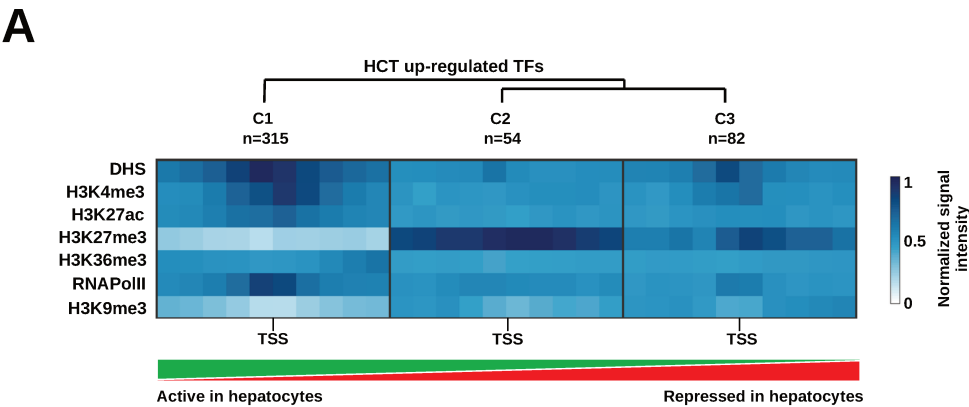


Supplementary Fig.4: Average DHS-seq, H3K36me3 and RNAPolIII ChIP-seq signals at the promoters of genes from clusters C1-3 in the mouse liver

(A) Average DHS-seq or ChIP-Seq signals for H3K36me3 or RNA polymerase II (RNAPolIII) at promoters from clusters C1-3 (defined in [Fig.3D](#)) in the mouse liver. Low signal was arbitrarily set to 1. Signals are shown in windows of +/- 2.5 kb around the genes TSS.

(B) Average ATAC-seq signals at promoters from clusters C1-3 (defined in [Fig.3D](#); top) or at promoters of genes whose expression is enriched in cholangiocytes, periportal or pericentral hepatocytes (bottom) ([Supplementary Tables 3 and 4](#)). ATAC-seq signal at individual genes was obtained by aggregating single-nucleus ATAC-seq data from the mouse liver for periportal, pericentral and cholangiocytes ([Supplementary Table 4](#)) (Bravo Gonzalez-Blas et al. 2024). Average signals +/-SE found at the indicated gene promoters are shown in windows of +/- 2 kb around the genes TSS.

Fig. S5



Supplementary Fig.5: Promoters from C3 are found in both euchromatin and heterochromatin domains mostly independently of labelling with H3K9me3

(A) Spark-based clustering of promoters from C3 performed as in [Fig.3D](#) except H3K9me3 ChIP-seq data from the mouse liver were also used. Average signal for chromatin accessibility (DHS) and indicated modified histones or RNA polymerase II (RNAPolII) at TF-encoding gene promoters for the main 3 clusters identified by Spark is shown. The arrows indicate the position of the transcriptional start site (TSS) at the center of the 5 kb windows used for these analyses. Overall, the 3 main clusters obtained are consistent with those reported in [Fig.3D](#) indicating H3K9me3 ChIP-seq signal is not a main driver of promoter heterogeneity.

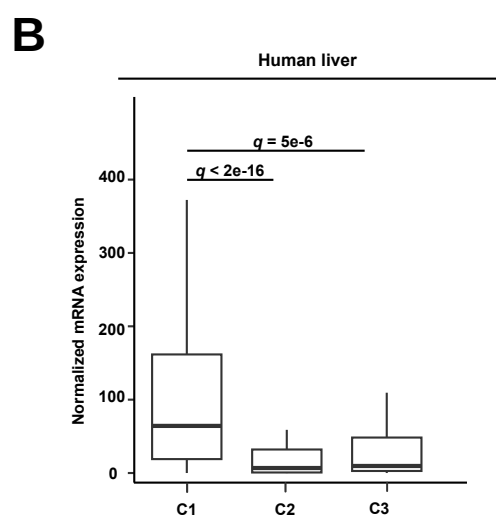
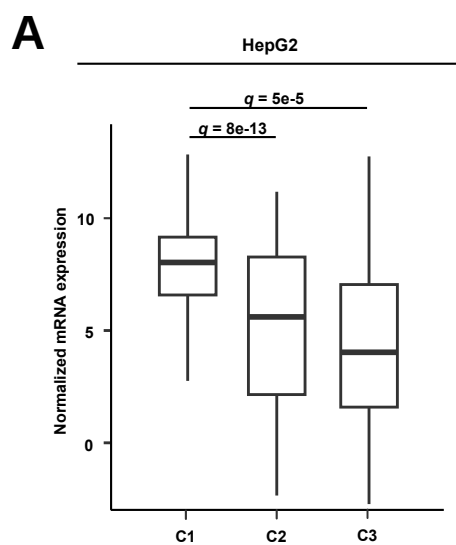
(B) Distribution of promoters from C1-3 within mouse liver chromatin domains defined using modeling the combinatorial presence or absence of signals for ATAC-seq and H2AZ, H3K4me3, H3K27me3 and H3K9me3 ChIP-seq using ChromHMM (Zhang et al. 2021). Enriched signals in the different chromatin domains are shown in the legend and correspondence with the chromatin state numbering used in the original manuscript is as follows: ATAC-seq + H3K4me3 + H2AZ (S2); ATAC-seq + H3K4me3 + H2AZ + H3K27me3 (S2 with H3K27me3); ATAC-seq + H3K4me3 (S1); ATAC-seq + H3K4me3 + H3K27me3 (S1 with H3K27me3); ATAC-seq + H2AZ (S3); H3K27me3 (S6); H3K9me3 (S5).

(C) Average DNA methylation signal at promoters from clusters C1-3 in the mouse liver. Signal is derived from whole-genome shotgun bisulfite sequencing (WGBS; [Supplementary Table 4](#)) and is shown in a window of +/- 2.5 kb around the genes TSS. Overall genes from all clusters display DNA hypomethylation around their TSS.

(D) Individual promoters from C1-3 were monitored for their presence in euchromatin or heterochromatin in the mouse liver as defined in (Grindheim et al. 2019) using sonication-resistant heterochromatin sequencing (srHC-seq). Intermediate domains include regions not called as hetero/euchromatic and hetero/euchromatin double-positive regions. The percentage

of promoters in euchromatin, intermediate domains and heterochromatin is shown for the 3 clusters.

Fig. S6

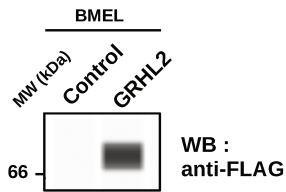


Supplementary Fig.6: Basal expression of TF-encoding genes from the C1-3 clusters

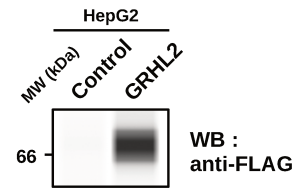
(A) Box plot showing the basal expression of TF-encoding genes from clusters C1-3 in HepG2 cells. Statistical significance was assessed using a Kruskal-Wallis test with Dunn's multiple comparison post hoc test.

(B) Box plot showing the basal expression of TF-encoding genes from clusters C1-3 in the human liver [control samples from our previous RNA-seq analyses (Bou Saleh et al. 2021) were used here]. Statistical significance was assessed using a Kruskal-Wallis test with Dunn's multiple comparison post hoc test.

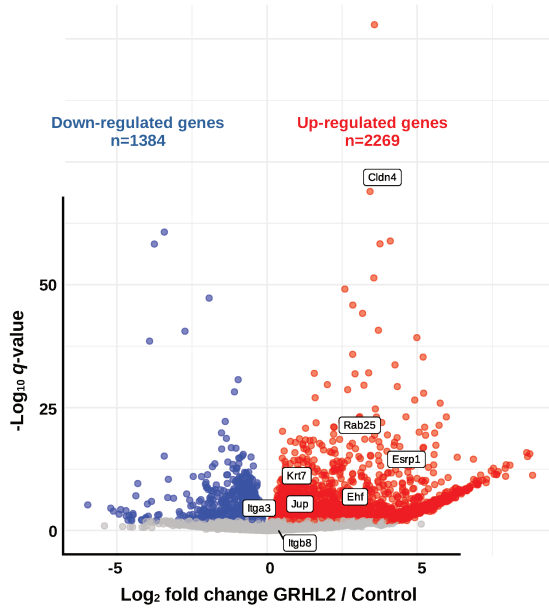
A



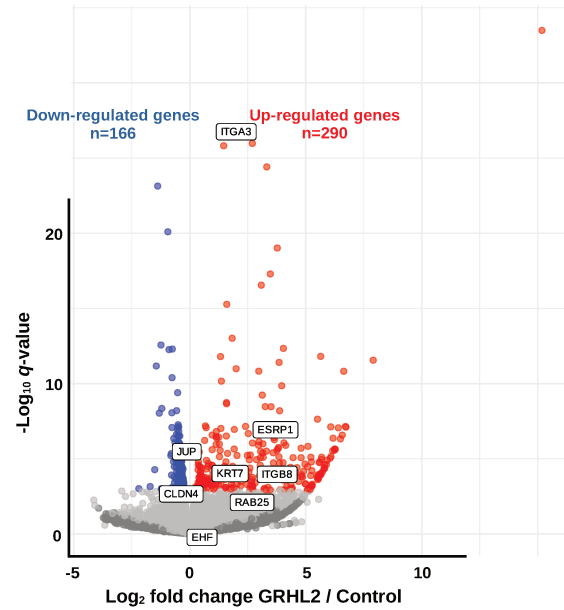
C



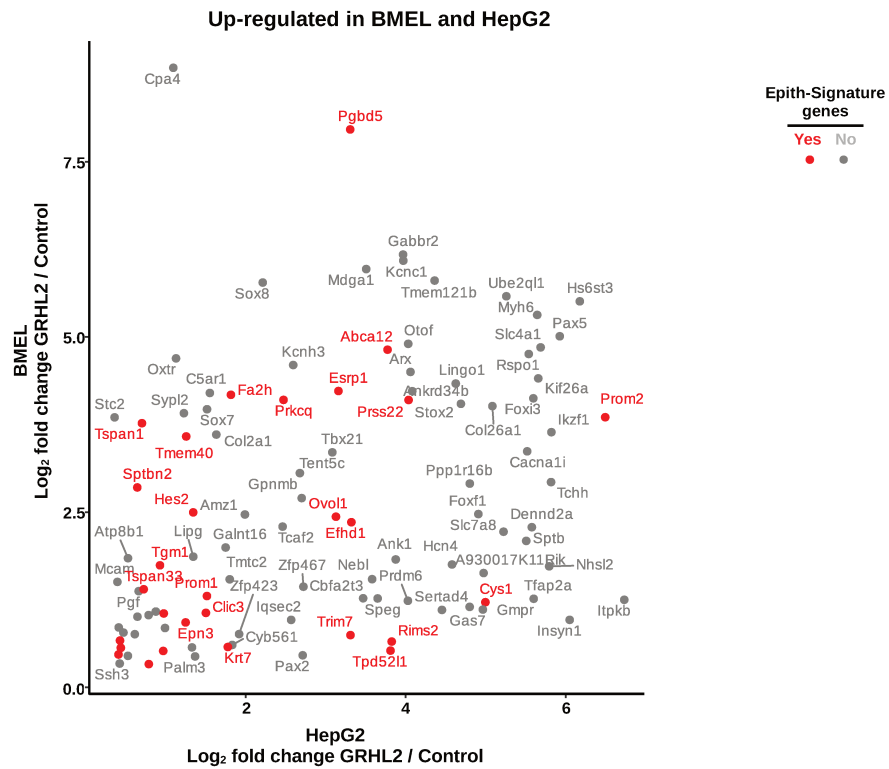
B



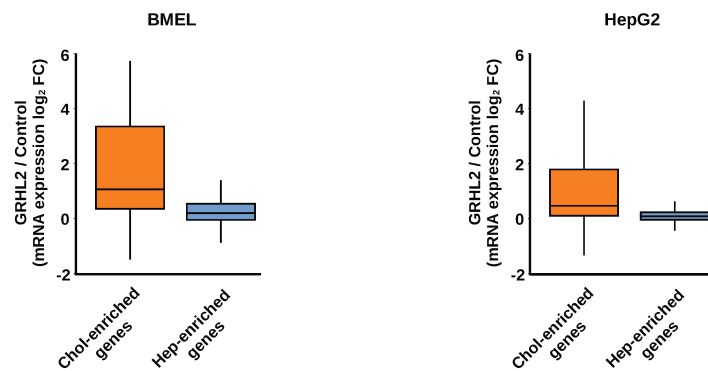
D



E



F



Supplementary Fig.7: Additional validation of GRHL2 ectopic expression and characterization of induced transcriptional changes in BMEL and HepG2 cells

(A) Since the construct allowing for GRHL2 ectopic expression included a Flag tag, we performed additional Wes immunoblotting assays on the samples used in [Fig.5C](#) with an anti-Flag antibody to monitor GRHL2 levels in transfected BMEL cells. MW, molecular weight. WB, Wes immunoblotting.

(B) Volcano plots showing transcriptomic changes induced by GRHL2 ectopic expression in BMEL cells (RNA-seq data from 4 independent biological replicates used in [Fig.5](#)). Significantly regulated genes ($q < 0.05$) are displayed in blue and red for down- and up-regulated genes, respectively.

(C) Since the construct allowing for GRHL2 ectopic expression included a Flag tag, we performed additional Wes immunoblotting assays on the samples used in [Fig.6A](#) with an anti-Flag antibody to monitor GRHL2 levels in transfected HepG2 cells. MW, molecular weight. WB, Wes immunoblotting.

(D) Volcano plots showing transcriptomic changes induced by GRHL2 ectopic expression in HepG2 cells (RNA-seq data from 4 independent biological replicates used in [Fig.6](#)). Significantly regulated genes ($q < 0.05$) are displayed in blue and red for down- and up-regulated genes, respectively.

(E) Scatter plot of individual genes significantly up-regulated upon GRHL2 ectopic expression in both BMEL and HepG2 cells (defined hereabove) showing their log₂ fold changes in these two models. Additionally, presence of individual genes within the Epith-Signature was highlighted using red dots.

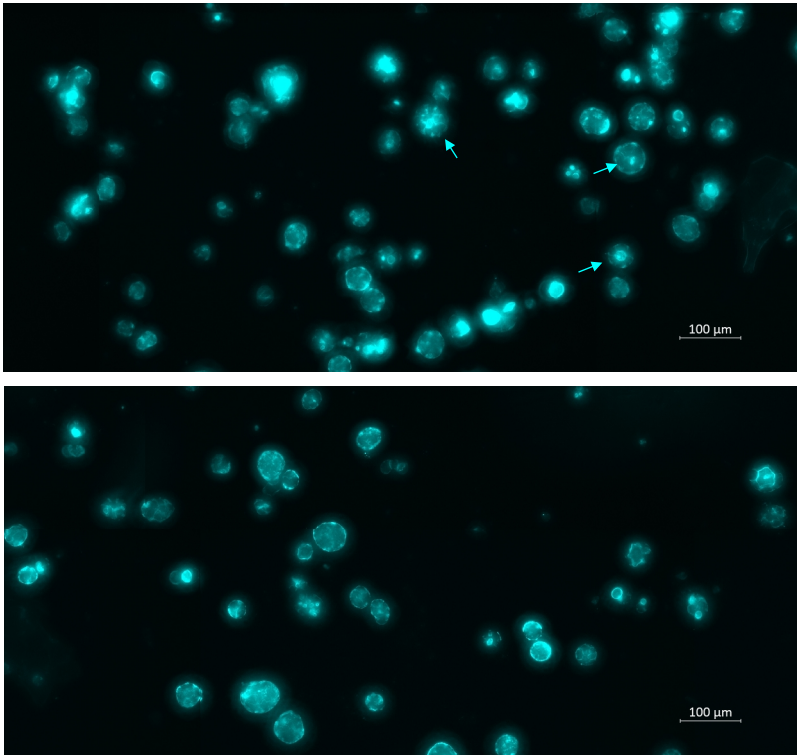
(F) Distribution of the log₂ fold changes for genes of the cholangiocyte-enriched (Chol-enriched) or hepatocyte-enriched (Hep-enriched) lists in BMEL or HepG2 cells ectopically expressing GRHL2.

The original data underlying this figure can be found at the Gene Expression Omnibus (GSE281717).

Fig.S8

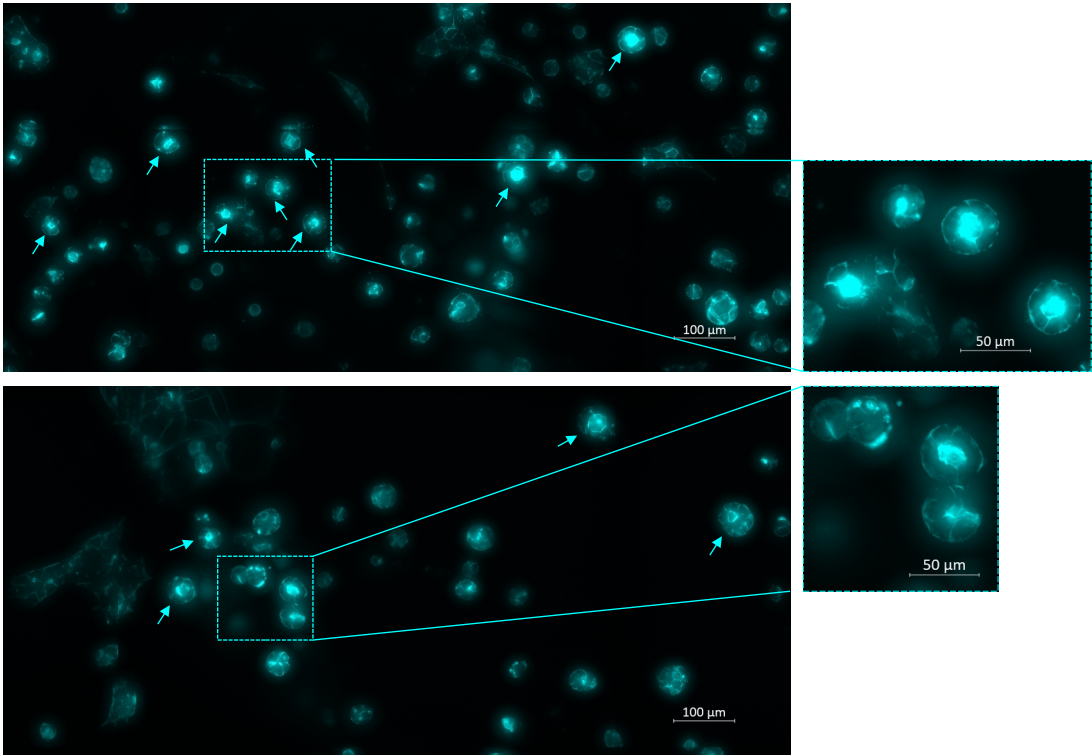
A

Control



B

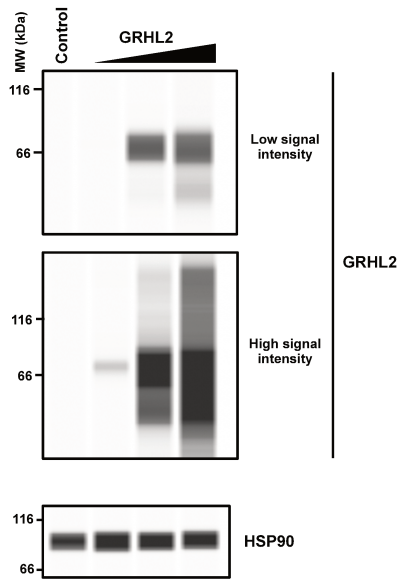
GRHL2



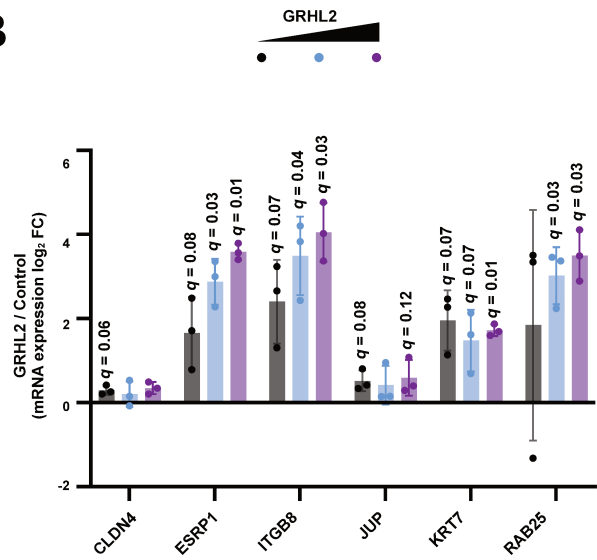
Supplementary Fig.8: Additional representative images of BMEL cell aggregates

BMEL cells were transfected with a GRHL2 expression plasmid or an empty control construct and grown in Matrigel. Additional representative images of obtained aggregates stained with phalloidin are shown. The arrows point to cystic cell aggregates.

A



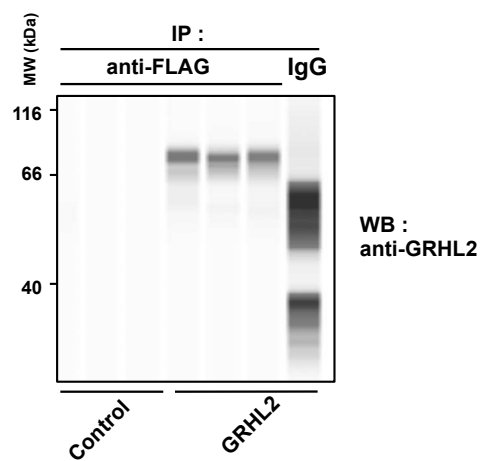
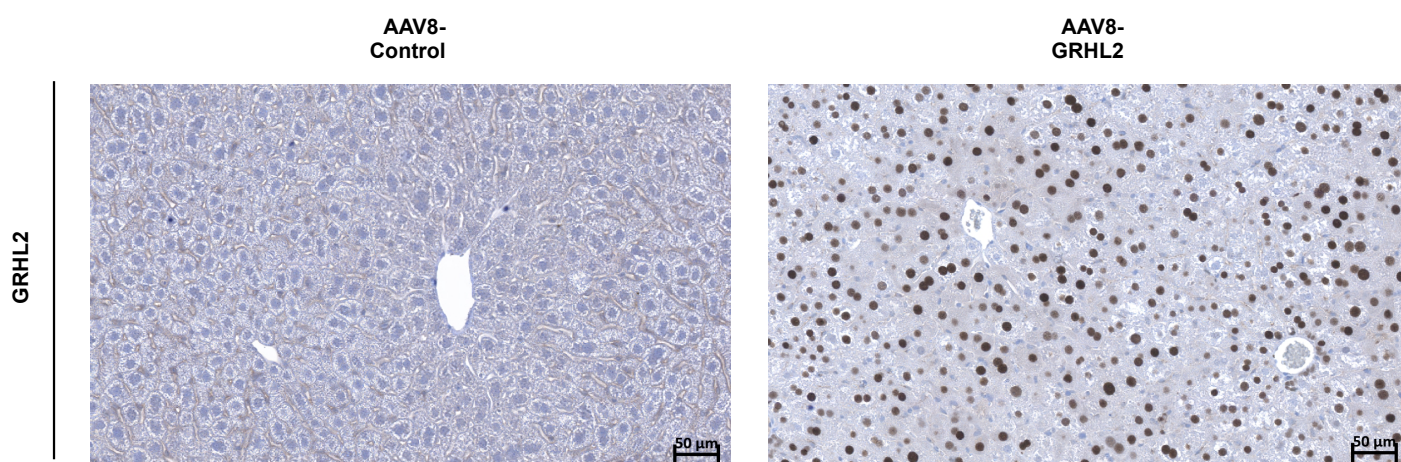
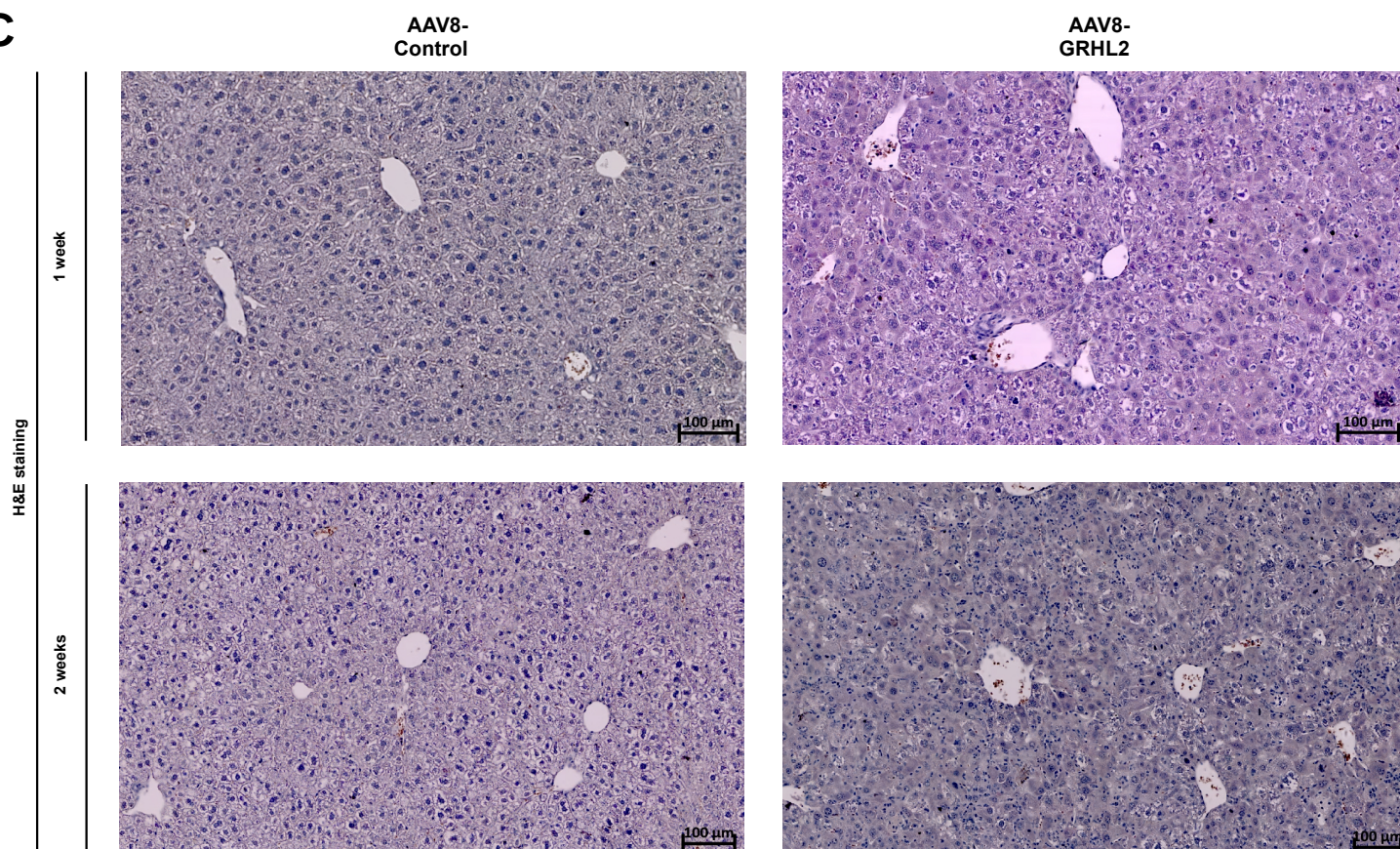
B



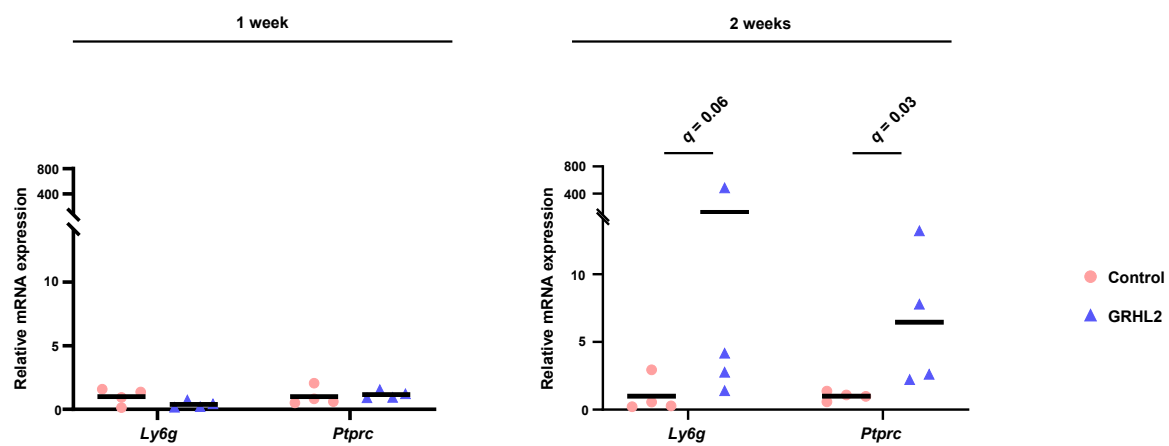
Supplementary Fig.9: Dose-dependent induction of GRHL2 target genes in HepG2 cells

(A) Immunoblotting performed using the Wes system to monitor GRHL2 and HSP90 levels in HepG2 cells transfected with increasing amounts of a GRHL2 expression plasmid or an empty control construct. For GRHL2, reconstructed lane views from the same Wes run using low and high sensitivity settings is shown. Data are representative of those obtained in 2 independent biological replicates. MW, molecular weight.

(B) RT-qPCR data showing the expression of the indicated GRHL2 target genes in HepG2 cells transfected as in panel G (n=3). Log₂ FC in gene expression in cells expressing GRHL2 compared to control cells (transfected with an empty plasmid) are shown in the bar graph, which displays means \pm SD together with individual biological replicates. Two-sided one-sample *t*-test with Benjamini-Hochberg correction was used to determine if the mean of log₂ FC was statistically different from 0. The numerical values of all biological replicates can be found in the S1 data file.

A**B****C**

D



Supplementary Fig.10: Additional validation of GRHL2 ectopic expression in the mouse liver

(A) Immunoblotting performed using the Wes system to monitor GRHL2 levels in the liver of mice subjected to hydrodynamic injection with the flag-tagged GRHL2-encoding plasmid or an empty construct (control). Nuclear extracts from individual mouse livers (3 mice per experimental group) were subjected to immunoprecipitation with an antibody against flag and immunoblotted using an anti-GRHL2 antibody. As a specificity control, immunoprecipitation with non-immune IgG was also conducted. MW, molecular weight. WB, Wes immunoblotting.

(B) Representative immunostaining of GRHL2 in livers from mice injected with GRHL2 or GFP (control) encoding AAV8 after 1 week.

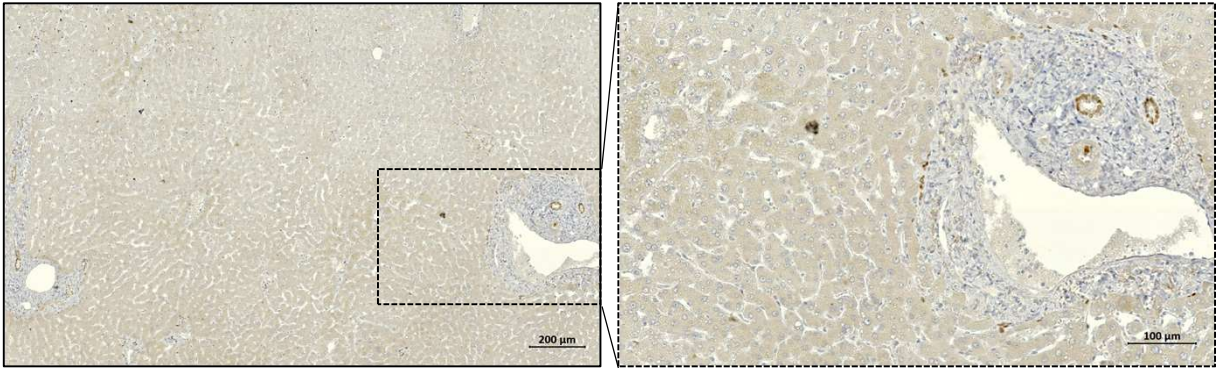
(C) Representative images obtained using liver slices from mice injected with GRHL2 or GFP (control) encoding AAV8 stained with hematoxylin and eosin after 1 or 2 weeks.

(D) RT-qPCR data showing the relative mRNA expression of *Ly6g* and *Ptprc* (*Cd45*) used as indicators of immune cell infiltration in the liver of mice 1 or 2 weeks after being injected with GRHL2 (n=4 mice) or GFP (control; n=4 mice) encoding AAV8. Data were plotted and analyzed as in [Fig.7D](#). The numerical values of all individual mice can be found in the S1 data file.

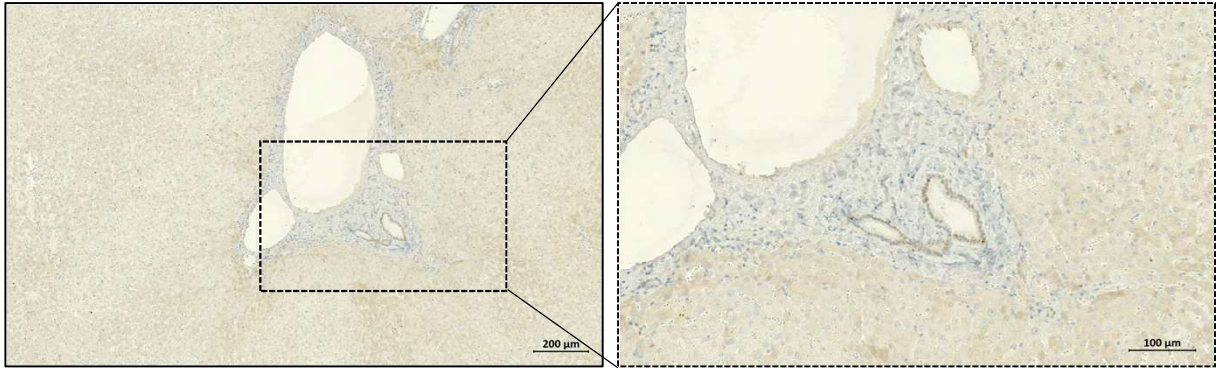
A

GRHL2

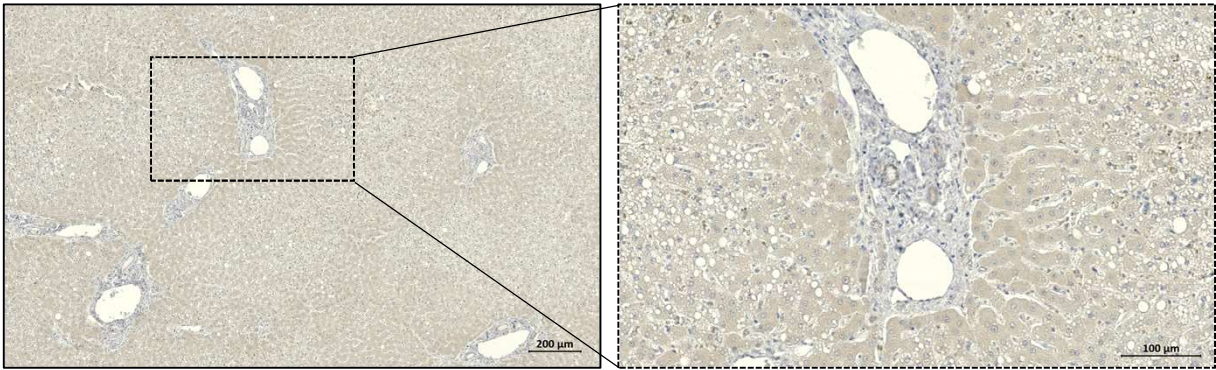
Control #1



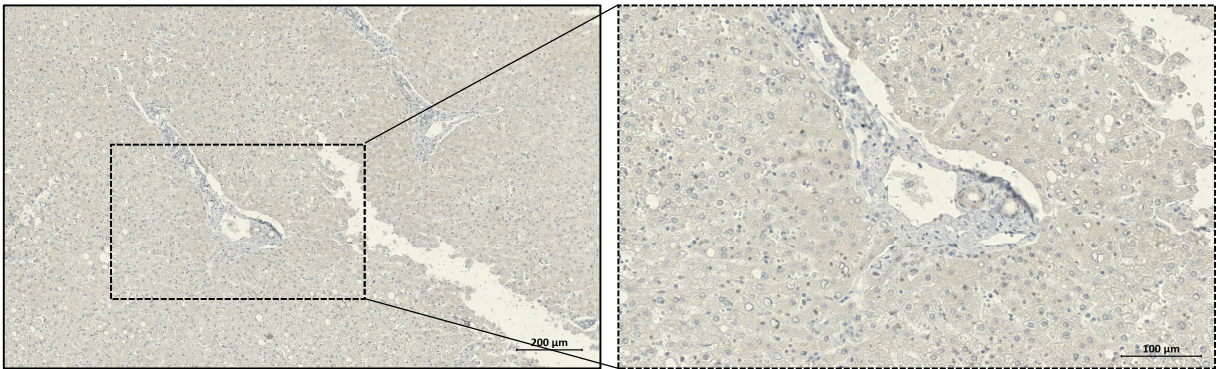
Control #2



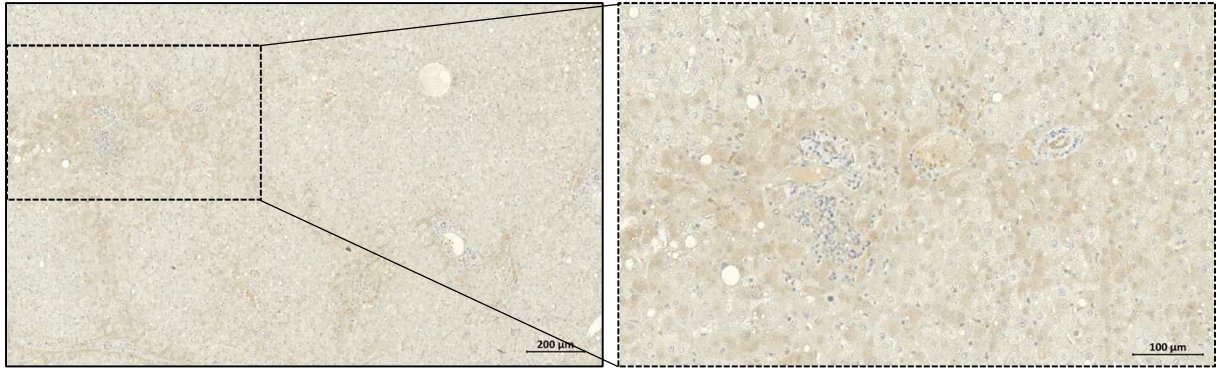
Control #3



Control #4



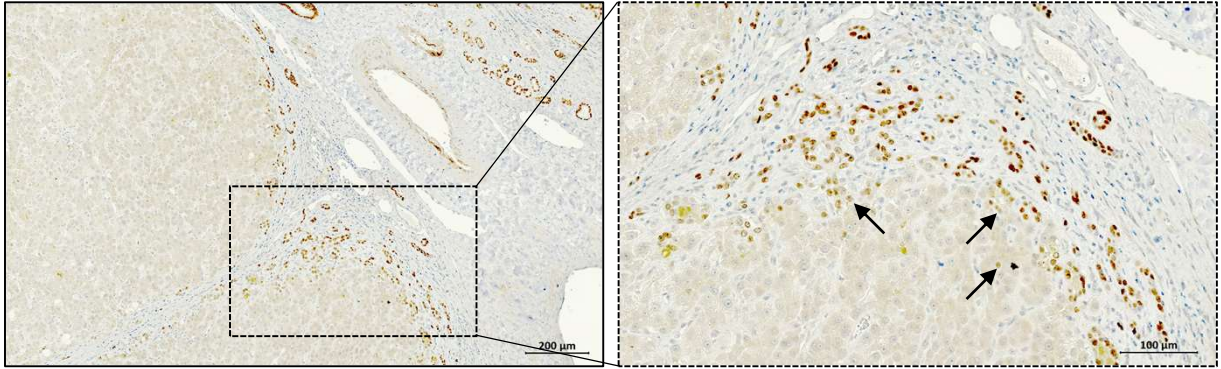
Control #5



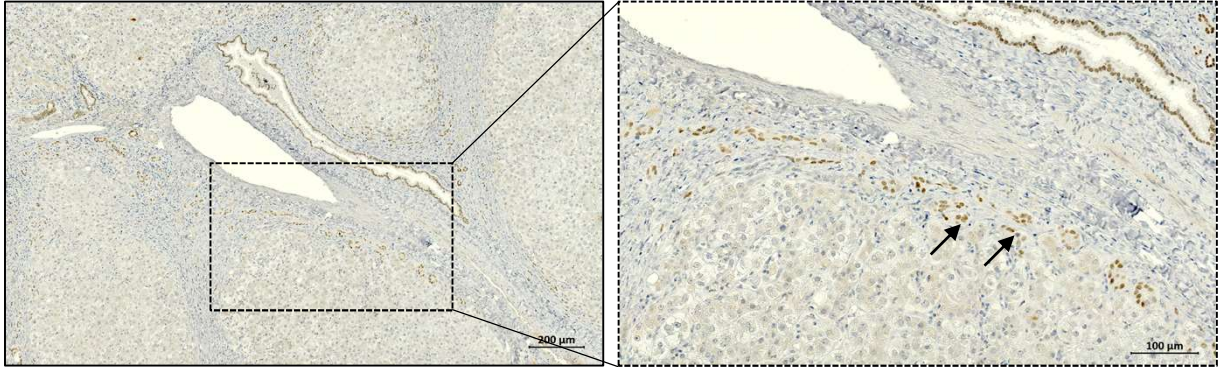
B

GRHL2

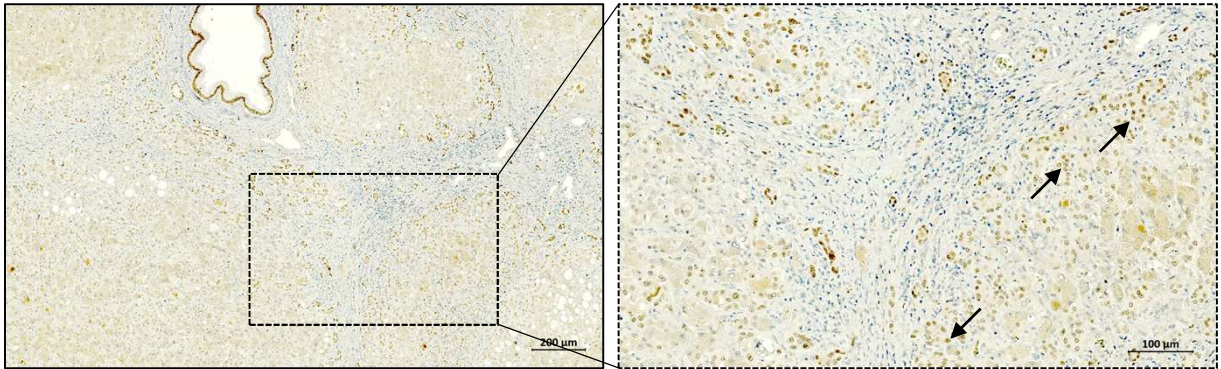
ALD #1



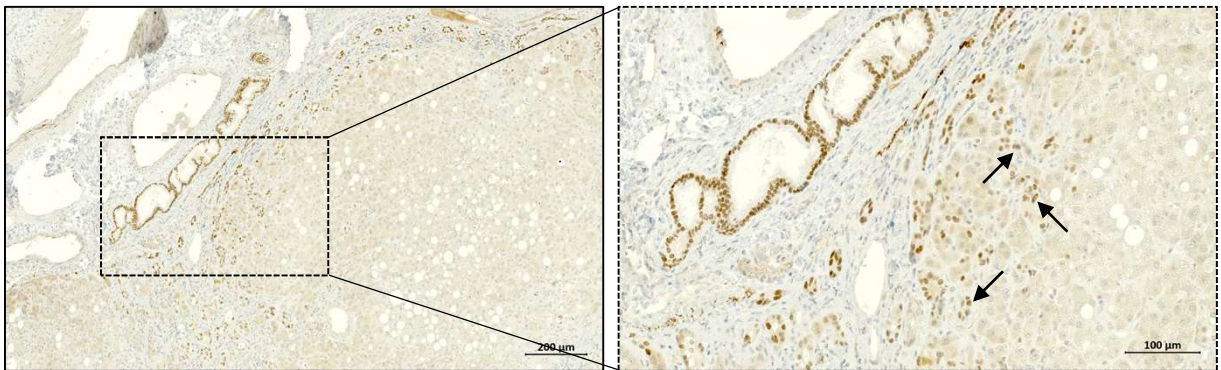
ALD #2



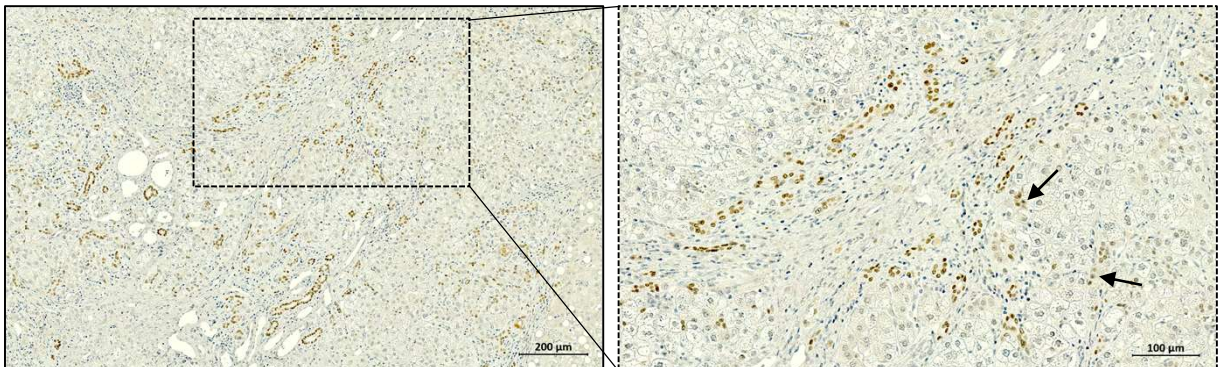
ALD #3



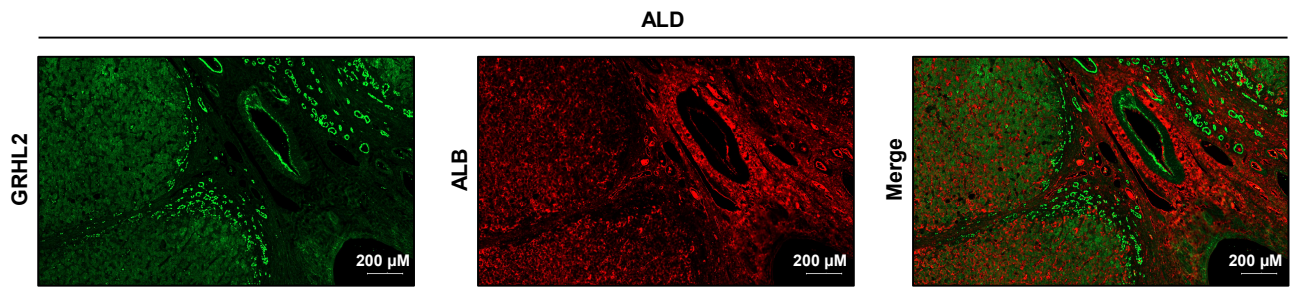
ALD #4



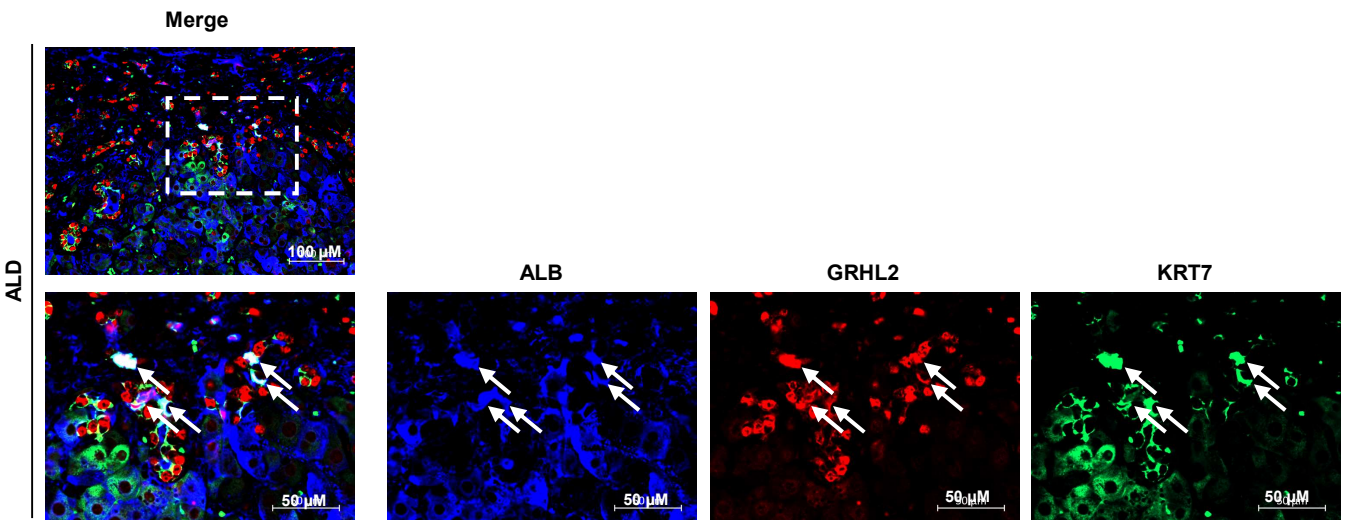
ALD #5



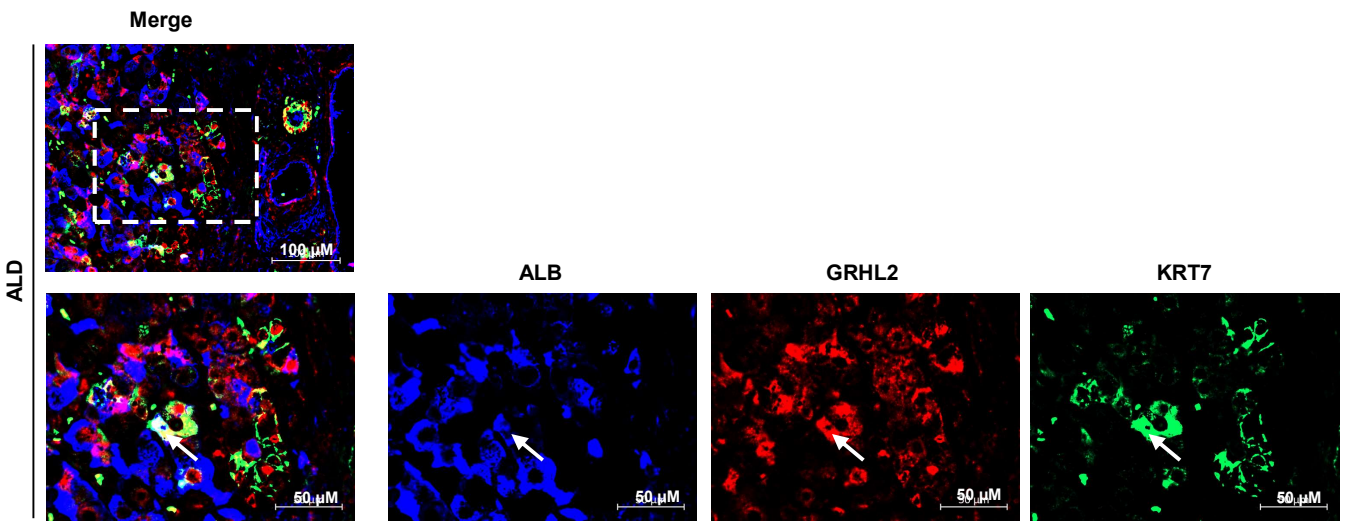
C



D



E



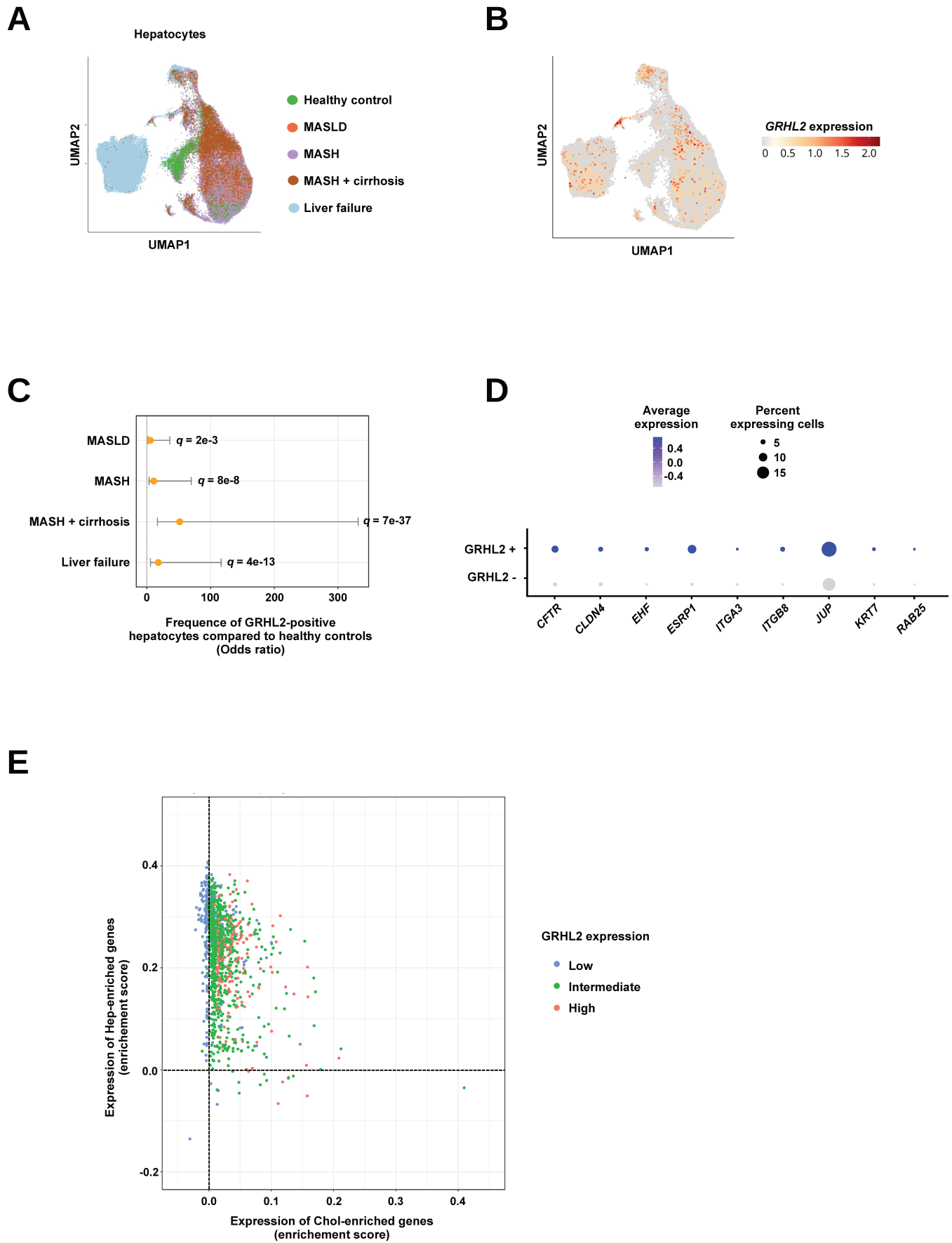
Supplementary Fig.11: Additional immunostainings of GRHL2 in human liver failure

(A-B) Immunostaining of GRHL2 in human livers from the control and ALD-related liver failure groups (n=5). Control #1 and ALD #1 are those shown in [Fig.8D](#). Zoomed images of the area delimited by the dotted rectangles is shown on the right. Arrows point to examples of GRHL2-positive hepatocytes.

(C) Co-immunostaining of GRHL2 (in green) and the hepatocyte marker Albumin (ALB; in red) in ALD-related human liver failure performed using a serial section of that shown in [Fig.8D](#).

(D-E) Co-immunostaining of the hepatocyte marker ALB (in blue), GRHL2 (in red) and the cholangiocyte marker KRT7 (in green) in ALD-related human liver failure. Panels D and E show different fields where triple positive cells were observed (indicated by arrows). Zoomed images of the area delimited by the dotted rectangle in the top image are shown at the bottom.

Fig.S12



Supplementary Fig.12: Characterization of GRHL2-expressing hepatocytes in human MASLD using snRNA-seq data

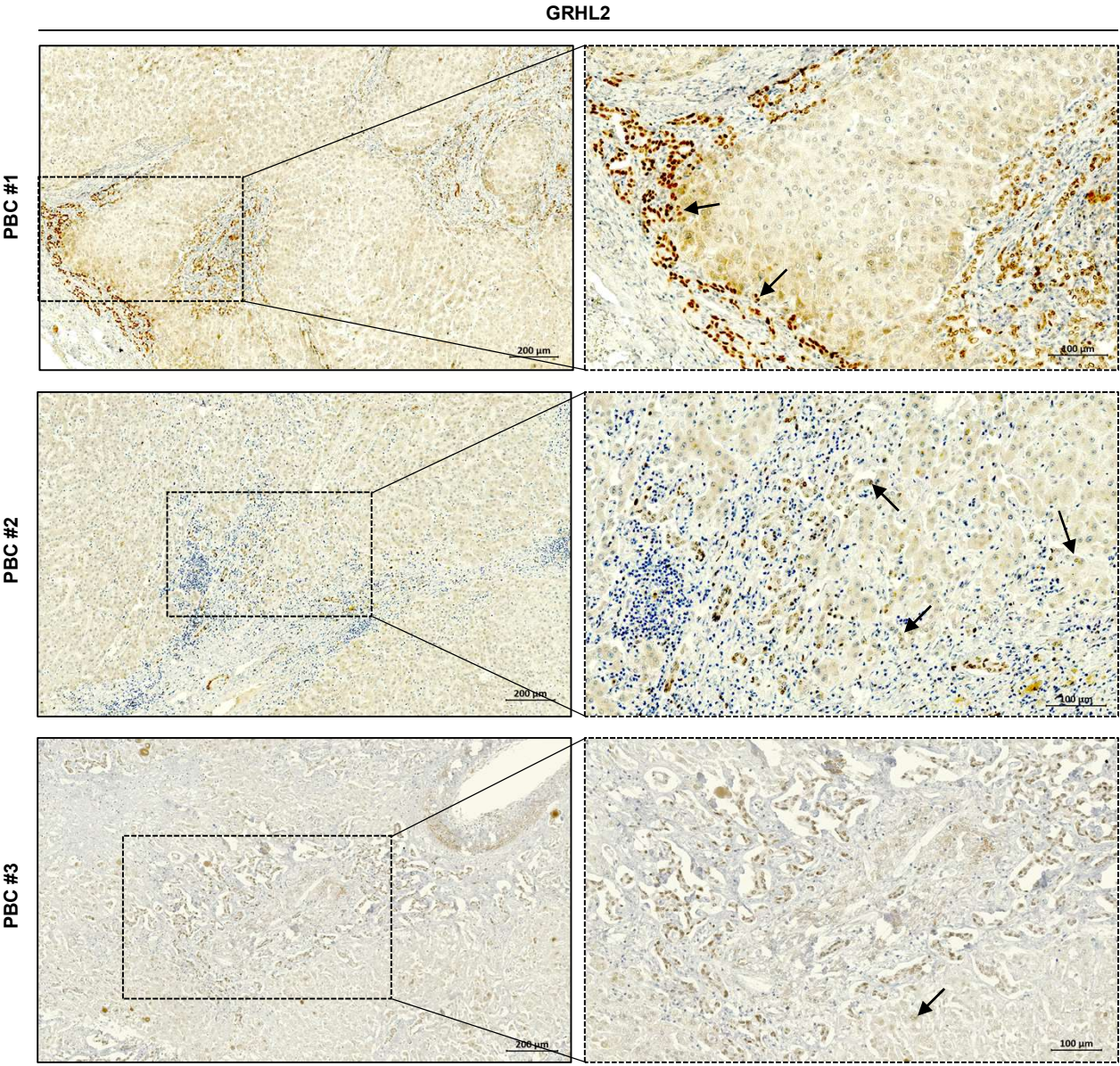
(A-B) Data from (Gribben et al. 2024) were used to monitor *GRHL2* expression in hepatocytes obtained from human livers of MASLD patients. Panel A shows the different stages of the disease hepatocyte originate from while panel B shows *GRHL2* expression in single nuclei.

(C) The fraction of *GRHL2*-expressing hepatocytes at each disease stage was retrieved and odds ratios were computed using hepatocytes from healthy controls as the reference. The forest plot shows odds ratio values (orange dots) together with 95 % confidence intervals. Statistical significance was assessed using two-sided Fisher exact tests with with Benjamini-Hochberg correction.

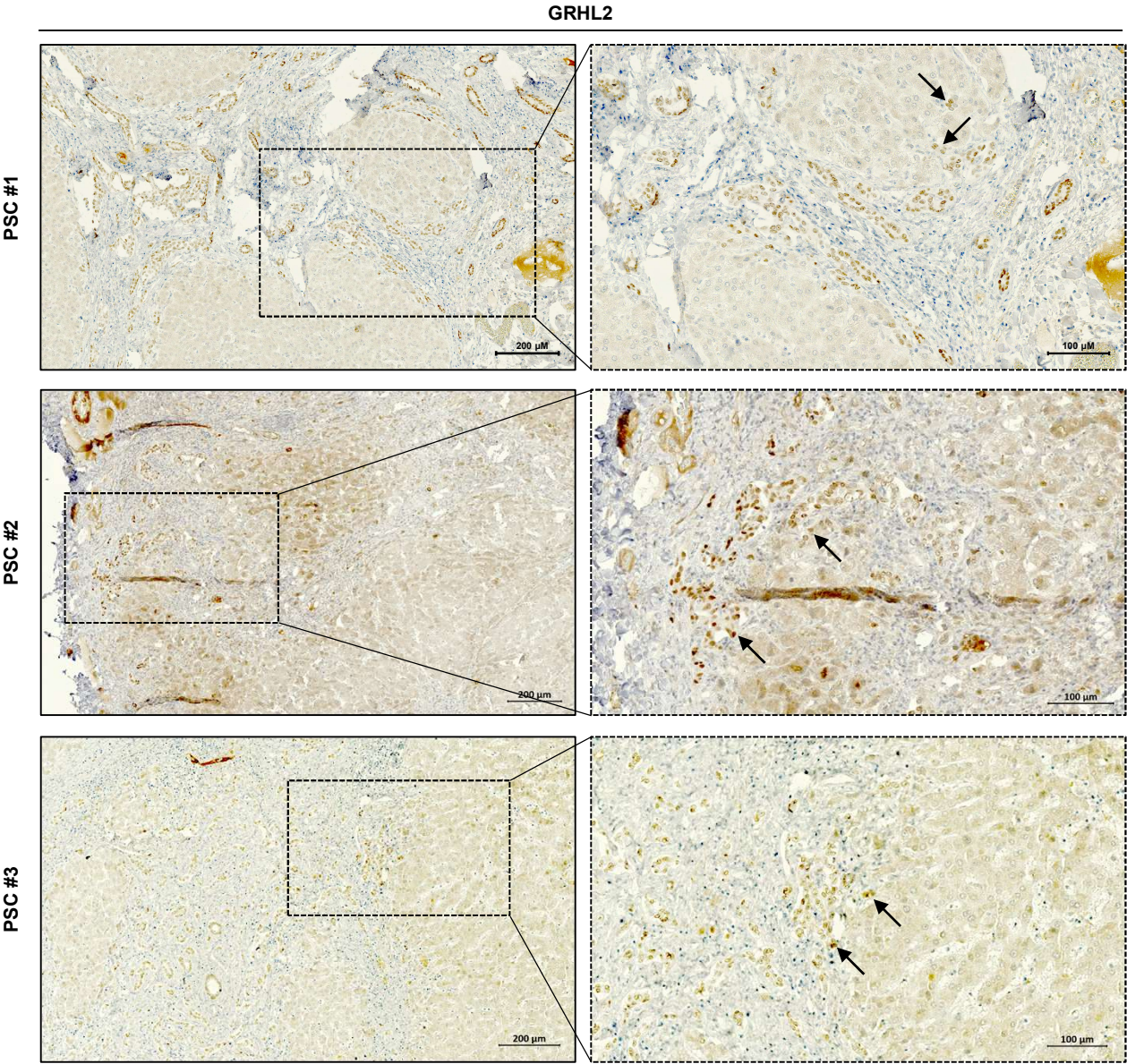
(D) Bubble plots showing expression of the indicated genes in hepatocytes negative (*GRHL2* -) or positive (*GRHL2* +) for *GRHL2* expression (i.e. any hepatocytes from panel A where *GRHL2* expression was detected).

(E) *GRHL2*-positive hepatocytes were assessed for expression of the Chol-enriched or Hep-enriched gene sets using the AddModuleScore function of Seurat. Individual cells are colored according to *GRHL2* expression levels (i.e. tertiles defined as low, intermediate and high expression).

A



B



Supplementary Fig.13: Additional immunostainings of GRHL2 in human PBC and PSC samples

(A-B) Immunostaining of GRHL2 in human livers from donors with PBC- or PSC-related liver failure (n=3 each). PBC #1 and PSC #1 are those shown in [Fig.8F](#). Zoomed images of the area delimited by the dotted rectangles is shown on the right. Arrows point to examples of GRHL2-positive hepatocytes.

References

- Bou Saleh M, Louvet A, Ntandja-Wandji LC, Boleslawski E, Gnemmi V, Lassailly G, Truant S, Maggiotto F, Ningarhari M, Artru F et al. 2021. Loss of hepatocyte identity following aberrant YAP activation: a key mechanism in alcoholic hepatitis. *J Hepatol* **75**: 912-923.
- Bravo Gonzalez-Blas C, Matetovici I, Hillen H, Taskiran, II, Vandepoel R, Christiaens V, Sansores-Garcia L, Verboven E, Hulselmans G, Poovathingal S et al. 2024. Single-cell spatial multi-omics and deep learning dissect enhancer-driven gene regulatory networks in liver zonation. *Nat Cell Biol* **26**: 153-167.
- Breschi A, Muñoz-Aguirre M, Wucher V, Davis CA, Garrido-Martín D, Djebali S, Gillis J, Pervouchine DD, Vlasova A, Dobin A et al. 2020. A limited set of transcriptional programs define major cell types. *Genome Research* **30**: 1047-1059.
- Dubois-Chevalier J, Gheeraert C, Berthier A, Boulet C, Dubois V, Guille L, Fourcot M, Marot G, Gauthier K, Dubuquoy L et al. 2023. An extended transcription factor regulatory network controls hepatocyte identity. *EMBO Rep* doi:10.15252/embr.202357020: e57020.
- Gribben C, Galanakis V, Calderwood A, Williams EC, Chazarra-Gil R, Larraz M, Frau C, Puengel T, Guillot A, Rouhani FJ et al. 2024. Acquisition of epithelial plasticity in human chronic liver disease. *Nature* doi:10.1038/s41586-024-07465-2.
- Grindheim JM, Nicetto D, Donahue G, Zaret KS. 2019. Polycomb Repressive Complex 2 Proteins EZH1 and EZH2 Regulate Timing of Postnatal Hepatocyte Maturation and Fibrosis by Repressing Gene Expression at Promoter Regions in Euchromatin in Mice. *Gastroenterology* doi:10.1053/j.gastro.2019.01.041.
- Merrell AJ, Peng T, Li J, Sun K, Li B, Katsuda T, Grompe M, Tan K, Stanger BZ. 2021. Dynamic transcriptional and epigenetic changes drive cellular plasticity in the liver. *Hepatology* doi:10.1002/hep.31704.
- Schaub JR, Huppert KA, Kurial SNT, Hsu BY, Cast AE, Donnelly B, Karns RA, Chen F, Rezvani M, Luu HY et al. 2018. De novo formation of the biliary system by TGFβ-mediated hepatocyte transdifferentiation. *Nature* doi:10.1038/s41586-018-0075-5.
- Zhang C, Macchi F, Magnani E, Sadler KC. 2021. Chromatin states shaped by an epigenetic code confer regenerative potential to the mouse liver. *Nat Commun* **12**: 4110.
- Zummo FP, Berthier A, Gheeraert C, Vinod M, Bobowski-Gerard M, Molendi-Coste O, Pineau L, Jung M, Guille L, Dubois-Chevalier J et al. 2023. A time- and space-resolved nuclear receptor atlas in mouse liver. *J Mol Endocrinol* **71**.