UNIVERSITY OF STRASBOURG

ED414 – Doctoral School of Life and Health Sciences

THESIS presented by:

Susanna SÄISÄ-BORREILL

INSERM U1113 – IRFAC-SMART – Strasbourg

The Basic Transcriptional Machinery in Intestinal Homeostasis

Defended on: **6 December 2021**

To obtain the degree of: **Doctor of the University of Strasbourg**

Discipline / Specialty: **Molecular and Cellular Aspect of Biology**

THESIS DIRECTORS:

Mr Jean-Noël FREUND	Thesis Director	DR, IRFAC Strasbourg, University of Strasbourg
Ms Isabelle DULUC	Thesis Codirector	CR, IRFAC Strasbourg, University of Strasbourg

JURY MEMBERS:

Ms Audrey FERRAND	External Reporter	CR, IRSD Toulouse, University of Toulouse
Ms Marion LAPIERRE	External Reporter	CR, IRC Montpellier, University of Montpellier
Mr Eric OGIER-DENIS	External Examiner	DR, CRI Paris, Paris Diderot University
Mr Jocelyn CERALINE	Internal Examiner	MCUPH, IGBMC Illkirch, University of Strasbourg

Ce travail a été financé par l'Agence Nationale de la Recherche (ANR)
et par la Fondation pour la Recherche Médicale (FRM)

This work was funded by the French
National Research Agency (ANR) & Foundation for Medical Research (FRM)

S I S U

[see-soo] • noun

FINNISH CULTURAL CONCEPT OF CHARACTER AND PRIDE

Extraordinary and indomitable spirit required to face adversity.

Inexhaustible reserve strength.

Qualities of stoic determination, tenacity of purpose, bravery grit and resilience.

Philosophy that what must be done will be done, no matter what.

REFUSING TO GIVE UP.

ACKNOWLEDGEMENTS

VERSION FRANÇAISE

Finalement après avoir consacré presque 1/6ème de ma vie à cette aventure, ma thèse arrive enfin à sa conclusion. Je suis fière de vous présenter mon projet ! Mais je ne peux pas commencer à-vous conter cette histoire sans remercier toutes les personnes sans qui cela n'aurait pas été possible.

Tout d'abord je voudrais remercier les membres du jury qui ont accepté de juger mon travail de thèse. Sans vous je ne pourrais mettre un point final à cette aventure. J'espère que vous apprécierez mon travail et qu'il vous initiera à mon petit monde de TAF4 et l'intestin. Je suis impatiente de vous rencontrer et vous présenter mon travail lors de ma soutenance ! Je tiens tout particulièrement à remercier Audrey FERRAND pour son aide face à mon mail inattendu. Merci encore d'avoir répondu aussi rapidement, j'espère plus que tout avoir ce post-doc ! Jocelyn, je ne pouvais pas continuer sans te remercier de m'avoir suivi tout au long de ces années. Tu as fait partie de mon comité de suivi de thèse, tu as suivi mon évolution du début jusqu'à la fin et j'espère sincèrement avoir réussi à progresser et à m'être améliorée.

Jean-Noël, mais dans quel traquenard tu m'as fourrée ! Mais quelle idée saugrenue que de me proposer (et me convaincre) de faire une thèse, moi qui étais si bien en tant qu'ingénieur… blague à part merci beaucoup pour toutes ces opportunités. Merci de m'avoir prise en tant qu'ingénieur et de m'avoir proposé de continuer en thèse Tu ne m'as jamais mis la pression et m'as laissée évoluer dans cette autonomie qui m'est si chère. Tu t'es toujours rendu disponible quand j'avais besoin, a toujours répondu à mes questions, soutenu et trouvé des compromis pour que je puisse continuer dans de bonnes conditions. Je sais que je n'ai pas été l'étudiante la plus facile, je partais un peu dans toutes les directions. Je voulais tout faire et même temps, persuadée que je pouvais tout mener de front, sans finalement me rendre compte que je partais trop loin et que je me perdais en route. Mais tu étais là me recadrer à me refocaliser sur l'essentiel et surtout à me laisser mon espace quand je craquais et que j'avais besoin de souffler un coup. Merci pour toutes ces années, j'ai énormément appris à tes côtés et même si je ne suis pas la plus démonstrative, sache que je te suis incroyablement reconnaissante pour tout ce que tu as fait pour moi. Tu es le meilleur directeur de thèse que dont j'aurai pu rêver. Merci du fond du cœur.

Quel serait un bon directeur de thèse dans son fidèle bras droit ! Isabelle que dire… Merci d'être cette incroyable personne que tu es. Je suis une tête de mule butée et bornée mais tu n'as jamais lâché. Tu as fait preuve d'une grande patience avec moi, même lorsque nos points de vue divergeaient sur les détails (presque insignifiants) de mes présentations et que tu retrouvais face à un mur de fierté. La douceur dont tu as fait preuve lors de mon apprentissage « souris » avec toi, me permet aujourd'hui et pour le restant de ma carrière scientifique, de travailler sereinement avec elles. Cependant, il y a une chose que je n'oublierais jamais, c'est le soutien infaillible dont tu as fait preuve tout au long de ma thèse et particulièrement sur la fin. Je t'ai toujours appréciée pour ça, pour cette capacité que tu as à me rassurer, mais sur cette fin de thèse… elle m'a été plus que d'un grand secours. Merci d'avoir été là, merci vraiment, je ne sais pas comment j'aurais fini sans ton aide et ton soutien.

Un laboratoire ne serait rien sans une équipe de choc.

Élisabeth, merci de m'avoir pris sous ton aile à mon arrivée, de m'avoir formée et épaulée tous au long de ces années. J'ai beaucoup appris à tes côtés aussi bien humainement que professionnellement. Tu m'as aidée, conseillée et épaulée tu es les fondations du laboratoire. Andréa, merci d'avoir pris ma relève et d'avoir fini mes expériences quand mon temps était arrivé à son terme. Claire, Isabelle, vous avez été là pour répondre à mes questions et m'aider à trouver des solutions

Émilie, Marine, vous avez été mes ainées de thèse. Vous vous êtes toujours rendues disponibles pour moi et je vous en remercie. Mais vous auriez pu faire un effort supplémentaire et rester plus longtemps, j'ai perdu mes agendas personnels sur pattes avec votre départ ! Boutade mise à part, vous avez été toutes les deux témoins de mes plus grands éclats d'émotions fortes, la colère et le désespoir. A chaque fois vous avez été là pour me calmer et me rassurer pour que finalement je puisse prendre ces situations avec humour. Je vous souhaite le meilleur pour l'avenir !

Véro, ne change jamais. Au-delà du côté professionnel où j'ai adoré échanger avec toi, nos longues discussions philosophiques vont me manquer ! Partager ma vision du monde et en débattre avec toi était à chaque fois une véritable bouffée d'air frais et source d'inspiration. J'espère pouvoir continuer nos discussions à l'avenir, c'est toujours stimulant de partager avec toi.

Ma petite Aurore, tu es la plus grande fierté de ma thèse ! Certes, je suis fière de mon travail de thèse, de ce que j'ai réussir à accomplir, mais pour moi la recherche ce n'est pas une réussite personnelle. C'est avant tout participer à l'augmentation du savoir de l'humanité pour améliorer le monde. La clé pour y arriver c'est le partage. Et j'ai beaucoup partagé avec toi. Tu as été une stagiaire BTS parfaite, curieuse, consciencieuse, volontaire. Tu as assimilé ma méthode de travail sans broncher avec réflexion et facilité. J'ai la sensation de t'avoir transmis cette « flamme de la recherche » qui m'anime. Même si aujourd'hui tu ne travailles pas dans la recherche publique, tu me donnes confiance en l'avenir, qu'il y aura toujours des personnes dans les générations suivantes pour faire de la recherche non pas par intérêt mais par envie d'aider. Merci mon petit rubis.

Aimée, pendant ma rédaction tu m'as dit des mots dans de ce style « tu ne peux pas craquer, je t'admire beaucoup, si tu craques comment ça va se passer pour moi ? ». Ces quelques mots m'ont permis de me reprendre. C'est vrai quoi, j'ai trop la hargne pour abandonner et craquer si proche du but. Merci pour ces mots d'encouragements qui m'ont rappelé que je ne faisais pas ma thèse seulement pour moi mais pour l'humanité et ceux qui ont cru en moi. J'espère avoir été la source d'inspiration que tu espérais, et sache que si j'y suis arrivée, tu y arriveras sans aucun doute !

Amandine, tu es une petite doctorante caractérielle que rien ne peut ébranler lorsqu'elle a une idée en tête. Continue d'avoir cette force de caractère mais n'oublie pas de continuer à t'ouvrir à d'autres possibilités. Tu arriveras au bout de tes objectifs, je n'en doute pas. Merci de m'avoir soutenue pendant ces périodes difficiles et d'avoir trouvé un du temps dans ton planning chargé pour moi.

Aux petits doctorants du bureau des doctorants, j'espère ne pas vous avoir trop effrayé durant ma rédaction de thèse. Prenez en compte que je suis bornée, trop exigeante envers moi-même, capable de dormir en toute circonstance. Merci d'avoir été un support essentiel pendant ma rédaction au quotidien. Merci d'avoir pris sur votre temps pour relire mon introduction et me

proposer vos corrections grammaticales si importantes. Mingyi, Chloé, Sevda je vous souhaite que votre thèse se passe le mieux possible, si jamais vous avez besoin de quoi que ce soit n'oubliez pas, mamie Susanna est là.

Je tiens également à remercier tout le « crew » de l'unité U1113. Léonor notre secrétaire de choc sans qui rien ne tournerai. Chloé pour ces relectures de CV, lettres de motivation, thèse et conseils professionnels. Léo pour tes chocolats, nos discussions et ce souvenir indélébile d'une de mes plus grandes et précieuse amitié. Je remercie aussi tous les autres, tous ce que je ne peux citer plus personnellement car je suis déjà en train d'écrire un roman en trois tomes.

Florence, Floriane, vous ne verrez peut-être jamais ces remerciements, mais si j'en suis arrivée là, c'est grâce à vous et je n'exagère pas. Vous avez été les premières à m'avoir donné confiance en mes capacités et à me pousser au-delà du BTS. Florence, je ne te remercierai jamais assez pour avoir été mon mentor pendant tant d'années, de m'avoir accordé autant de confiance et responsabilité. J'ai pu développer ma passion de la microscopie et mon autonomie grâce à toi. Floriane, la rigueur et le sérieux que j'ai pu observer chez toi sont devenus les fondements de ma méthode de travail. Tu es une personne que j'admire énormément et j'espère que tu arriveras à faire ce que tu souhaites. Merci à toutes les deux, sans votre enseignement je n'aurai jamais eu les compétences et la confiance que j'ai aujourd'hui. Je n'aurai pas trouvé ce poste d'ingénieur à Strasbourg et Jean-Noël ne m'aurait pas proposé de me poursuivre en thèse. Merci du fond du cœur.

Pour partir dans une direction plus personnelle, j'aimerais dire à mes proches que vous avez tous étés là pour moi à votre manière. Mais surtout vous m'avez soutenu et épaulé lorsque j'ai craqué. Je vous dois tellement, je ne sais pas si vous vous en rendez compte, sans vous cette thèse ne serait pas arrivée à son terme ou du moins, certainement pas aussi bien !

Anaïs, Justine, mes comparses du club des maine coons U1113, j'étais obligée de commencer par vous remercier ensemble. Nous avons fait beaucoup ensemble, grâce à vous j'ai eu une salle de bain magnifique, non inondée *tousse*. Nous formons une équipe de choc ! Entre l'Alsacienne maître d'œuvre, la portugaise technique et la finlandaise destructive, nous nous accordons à la perfection ! Il suffit de voir les CV et lettres de motivation qui ressortent lorsque l'on travaille ensemble. Vous êtes plus que des collègues de laboratoire que j'apprécie, vous êtes de véritables amies, parmi mes plus proches amies si j'osais l'avouer ! J'aurais dû plus vous écouter lorsque vous m'aviez affirmé que je devais lever le pied. J'aurai dû aller au-delà de mon illusion d'être invincible. Mais fort heureusement pour moi, vous ne m'avez pas lâchée et grâce à vous j'ai pu éviter le mur, j'ai simplement chuté. Merci de ne pas avoir lâchée, je n'aurais probablement pas rebondi si vous n'aviez pas été d'un soutien infaillible. D'un point de vue plus personnel Anaïs, tu es une personne incroyable ne l'oublie jamais. Tes critiques et conseils m'ont permis d'avancer bien plus que je ne l'aurais cru. Tu m'as recentrée quand je partais dans toutes les directions et n'a jamais hésité à me dire que j'avais tort. Merci pour cette amitié qui a été un des piliers de ma thèse et qui continuera bien au-delà. Justine, ton franc parler va me manquer, et nos sessions de jeux de mots à 1h du matin vont me manquer. Je te remercie d'avoir été là pendant toute ma thèse mais surtout à la fin. Merci de m'avoir tenu compagnie au labo jusqu'à pas d'heure pendant que je finissais mes dernières manips, merci d'avoir corrigé mon manuscrit. J'attends de pouvoir de nouveau travailler avec toi dans le cadre d'un post-doc en Finlande !

Mon frère de cœur, ce que tu m'as manqué lorsque tu es parti ! Te voir partir avant moi a été une véritable déchirure. J'aimerais dire que tu es un traitre et que tu aurais pu rester 1 ans de plus avec moi au labo, mais je sais que tu étais arrivé au bout de ce que pouvait t'apporter la thèse. Gilles, ton énergie et ta joie de vivre ont été un tel phare pour moi pendant ces années de labo. Ce qui est bien, c'est que même maintenant, c'est toujours le cas. La situation a évolué, le phare Gilles et devenu un phare familial. Gilles, Laure, je suis tellement heureuse pour toutes ces années. Je vous suis reconnaissante de la confiance que vous m'avez accordé avec Victoria. Je vous promets de la protéger, de la soutenir et d'être là autant que vous l'avez été pour moi. J'ai beaucoup de mal à exprimer et montrer mes sentiments mais vous êtes ma famille et je serais toujours là pour vous.

Ava, ma bichette. Je n'arrive pas à trouver les mots pour te dire à quel point je te suis reconnaissante. On a tellement partagé et traversé ensemble que tu sais exactement combien je te dois même tu le dénieras encore et toujours, moi je ne l'oublierai jamais. Tu as été mon plus grand soutien sans même le chercher, j'aurai été perdue sans toi. Je suis triste que cette partie de notre histoire s'arrête j'aimais sincèrement travailler avec toi (et vivre dans ton voisinage proche). Mais je reste plus qu'impatiente de découvrir la suite, notre histoire ne fait que commencer. Je te suis autant reconnaissante que suis fière de toi. J'ai eu le privilège d'assister à ton évolution, t'affirmer et te dévoiler. Tu es mon âme-sœur spirituelle, ma famille, continue sur cette lancer ma bichette, trace ta route ! Tu es une personne incroyable, ne sois pas ignorante du monde mais garde ton innocence, je serai là pour te protéger, t'épauler et te rattraper. A tout de suite ma belle.

Antoine, notre relation et nos chemins ont bien évolués tout au long de ma thèse. Tu as été là dès le début, à essayer de comprendre et me soutenir du mieux que tu as pu. Et nous savons tous les deux que ce n'était pas toujours facile avec nos deux caractères opposés, mon sang chaud et mon côté borné. Au-delà d'avoir démystifier le bricolage, j'espère t'avoir apporté autant que tu ne l'as fait pour moi, Tu m'as appris une de mes leçons les plus importante, tu m'as appris que j'en valais la peine. Je suis fière de te compter parmi mes amis proches.

Il était une fois un chauve fan d'animé et couteaux qui devint mon ami. Raphaël, tu es mon ami depuis tant d'année, le lycée cela me parait tellement loin ! Merci de m'avoir aidée sur les détails non scientifiques de mon manuscrit de thèse. Merci d'être resté en ligne avec moi pendant des heures la nuit pour m'aider à rester éveillée et ne pas me sentir seule. Maintenant l'aventure finlandaise nous tends les bras !

Julien, merci d'avoir usé de tes talents pour me faire gagner du temps. Je suis particulièrement fière des illustrations de mon introduction de thèse, mais il est vrai que je n'arrivais plus à trouver le temps de faire les corrections finales. Merci d'avoir répondu présent lors de mon appel à l'aide sorti de nulle part.

Vladimir, je sais que tu es surpris de te voir apparaitre dans mes remerciements, mais comme je te l'ai affirmé, la thèse ce n'est pas un simple chapitre mais un véritable tome, les personnages secondaires ont autant d'importance que les personnages principaux. Merci d'avoir été là pour m'aider avec mes tracas quotidiens de bricolage, merci de m'avoir préparé et livré à manger quand c'était devenu la dernière de mes priorités (même si je n'ai pas eu le choix dans cette histoire !).

Ma famille, je suis tellement fière d'avoir fini. Qui aurait cru que j'irai à ce niveau d'étude, certainement pas moi, je n'aimais pas les études au collège et lycée, c'est arrivé tardivement. Merci d'avoir été là à votre manière. Pierre-Henri ces heures de jeux passées pendant cette

rédaction m'ont permis de me forcer à prendre de petites pauses plus que nécessaires. Jyri, tes corrections et suggestions en anglais ont été plus qu'appréciées, merci d'avoir pris le temps. Christelle, merci de m'avoir recueillie pendant ma fuite stratégique de la canicule qui n'est jamais venue. Changer d'environnement m'a été bénéfique. M'man, merci d'avoir essayé de comprendre et me pousser à continuer. Je sais que tu as fait preuve de patience lorsque je répétais que je voulais arrêter la recherche (discours typique du doctorant épuisé). Tu avais raison, je vais continuer mais cette fois-ci en Finlande. Prépare-moi une petite place provisoire chez toi, j'arrive ! Merci à vous, j'espère vous rendre fière.

Aux Finlandais, Marjatta, Old boy et Peppi, merci de votre soutien direct et indirect. Merci d'avoir pris le temps de lire ma thèse, de m'avoir changé les idées et de m'avoir donné un but quand j'en avais besoin. Vous m'avez aidé à entretenir mon rêve de Finlande qui est devenu mon seul moteur face à l'épuisement.

Pour finir, bien que cela soit une pratique j'imagine peu orthodoxe, je tiens à remercier mes chats. Ils ne comprennent absolument rien à la science et nos discussions ne m'ont beaucoup aidé à élucider TAF4, mais ils ont été ma première ligne de défense. Dans le tumulte émotionnel de la thèse, ils ont accepté toutes mes humeurs et ont été grande source de réconfort au quotidien. Mes boules de poils vous êtes ce que j'ai de plus précieux.

Cette thèse m'a beaucoup appris et poussé plus loin dans mes retranchements. J'en ressors plus grande et j'espère plus sage. Je suis fière de mon travail et heureuse des amitiés inébranlables que j'ai développées. Mais s'il y a une chose que je ne cesserai d'affirmer et reconnaitre c'est que j'ai la chance insolente d'avoir votre soutien. A tous, je vous dis merci, je ne suis pas mécontente d'avoir vécu (et fini !) cette histoire en étant si bien entourée !

ENGLISH VERSION

Finally, after devoting almost 1/6 of my life to this adventure, my thesis is finally coming to a conclusion. I am so proud to present my project to you! But I cannot begin to tell you this tale without thanking all the people without whom it would not have been possible.

First of all, I would like to thank the members of the jury who agreed to judge my thesis work. Without you, I could not put a stop to this adventure. I hope you will enjoy my work and that it will introduce you to my little world of TAF4 and the gut. I look forward to meeting you and presenting my work at my defense! I would especially like to thank Audrey FERRAND for her help with my unexpected email. Thank you again for responding so quickly, I hope more than anything to have this post-doc! Jocelyn, I couldn't go on without thanking you for having followed me throughout these years. You were part of my thesis committee; you followed my evolution from the beginning to the end and I sincerely hope that I managed to progress and improve.

Jean-Noël, what a trap you've put me in! What a crazy idea to propose (and convince me) to do a thesis, me who was so good as an engineer... joking aside, thank you very much for all these opportunities. Thank you for having taken me on as an engineer and for having proposed that I continue with a thesis. You never put pressure on me and let me evolve in this autonomy that I hold so dear. You always made yourself available when I needed, always answered my questions, supported me, and found compromises so that I could continue under good conditions. I know that I was not the easiest student, I was going in all directions. I wanted to do everything at the same time, convinced that I could do everything at once, not realizing in that I was going too far and losing my way. But you were there to help me refocus on the essential and above all to give me my space when I broke down and needed a break. Thank you for all these years, I learned a lot from you and even if I am not the most demonstrative, I am incredibly grateful for everything you did for me. You are the best thesis director I could have ever dreamed of. Thank you with all my heart.

What a good thesis director would be in his faithful right hand! Isabelle what can I say... Thank you for being the incredible person that you are. I am a stubborn and obstinate mule-head, but you never gave up. You were so patient with me, even when we differed on the (almost insignificant) details of my presentations and you had to face the wall of my pride. The gentleness you showed me when I was training with you to work with mice, allows me today and for the rest of my scientific career, to work serenely with them. However, there is one thing I will never forget, and that is the unfailing support you gave me throughout my thesis and especially at the end. I have always appreciated you for that, for your ability to reassure me, but at the end of my thesis... it was more than helpful. Thank you for being there, thank you really, I don't know how I would have finished without your help and support.

A laboratory would be nothing without a great team.

Elisabeth, thank you for taking me under your wing when I arrived, for training me and supporting me throughout these years. I learned a lot at your side, both humanly and professionally. You helped, guided, and supported me; you are the foundation of the laboratory. Andréa, thank you for taking over from me and finishing my experiments when my time was up. Claire, Isabelle, you were there to answer my questions and help me find solutions

Émilie, Marine, you were my thesis elders. You always made yourselves available to me and I thank you for that. But you could have made an extra effort and stayed longer, I lost my personal human diaries with your departure! You have both witnessed my greatest bursts of strong emotions, anger, and despair. Each time you have been there to calm and reassure me so that I can finally take these situations with a sense of humor. I wish you the best for the future!

Véro, never change. Beyond the professional side where I loved to exchange with you, I will miss our long philosophical discussions! Sharing my vision of the world and debating it with you was always a breath of fresh air and a source of inspiration. I hope to continue our discussions in the future, it is always stimulating to share with you.

My little Aurore, you are the greatest pride of my thesis! Of course, I am proud of my thesis work, of what I have achieved, but for me research is not a personal success. It is above all to participate in increasing the knowledge of humanity to improve the world. The key for this is sharing. And I have shared a lot with you. You were a perfect BTS trainee (two-years biotechnology technical degree), curious, conscientious, willing. You assimilated my working method without flinching with thought and ease. I have the feeling that I have passed on to you my "flame of research" that animates me. Even if today, you are not working in public research, you give me confidence in the future, that there will always be people in the following generations to do research not out of interest but out of a desire to help. Thank you, my little ruby.

Aimée, during my essay you said to me words like "you can't crack, I admire you a lot, if you crack how will it happen to me?". These few words helped me to pull myself together. It's true, I have too much spirit to give up and crack so close to my final goal. Thank you for these words of encouragement that reminded me that I was not doing my thesis only for myself but for humanity and those who believed in me. I hope I was the inspiration you were hoping for, and know that if I made it, you will no doubt make it!

Amandine, you are a little characterful doctoral student that nothing can shake when she has an idea in mind. Continue to have this strength of character but don't forget to keep opening yourself to other possibilities. You will achieve your goals; I have no doubt. Thank you for supporting me during these difficult times and for finding time in your busy schedule for me.

To the little doctoral students of the doctoral students' office, I hope I did not scare you too much during my thesis writing. Take into account that I am stubborn, too hard on myself, and able to sleep through anything. Thank you for being an essential support during my daily writing. Thank you for taking time out of your day to proofread my introduction and offer your all-important grammatical corrections. Mingyi, Chloe, Sevda I wish you all the best with your thesis, if you ever need anything don't forget, Granny Susanna is there.

I would also like to thank the crew of the U1113 unit. Léonor, our secretary, for whom nothing would go wrong. Chloé for her proofreading of CVs, cover letters, thesis, and professional advice. Léo for your chocolates, our discussions, and the indelible memory of one of my greatest and most precious friendships. I would also like to thank all the others, all those I cannot mention personally as I am already writing a novel in three volumes.

Florence, Floriane, you may never see these thanks, but if I got here, it is thanks to you, and I am not exaggerating. You were the first to give me confidence in my abilities and to push me

beyond the technical degree. Florence, I can never thank you enough for being my mentor for so many years, for giving me so much confidence and responsibility. I was able to develop my passion for microscopy and my autonomy thanks to you. Floriane, the rigor and seriousness that I observed in you have become the foundation of my working method. You are a person that I admire immensely, and I hope that you will succeed in doing what you wish. Thank you both, without your teaching I would never have had the skills and confidence I have today. I would not have found this engineering job in Strasbourg and Jean-Noël would not have offered me to continue my thesis. Thank you from the bottom of my heart.

To go in a more personal direction, I would like to say to my loved ones that you have all been there for me in your own way. But most of all you have supported me when I broke down. I owe you so much, I don't know if you realize it, without you this thesis would not have been completed or at least, certainly not as well!

Anaïs, Justine, my fellow members of the U113 maine coons club, I had to start by thanking you all. We did a lot together, thanks to you I got a beautiful bathroom, not flooded *cough*. We make a great team! Between the Alsatian master builder, the technical Portuguese, and the destructive Finnish, we fit together perfectly! Just look at the CVs and motivation letters that come out when we work together. You are more than just laboratory colleagues whom I appreciate, you are true friends, among my closest friends if I dare to admit it! I should have listened to you more when you told me that I should take it easy. I should have gone beyond my illusion of being invisible. But fortunately for me, you didn't let go and thanks to you I was able to avoid the wall, I just fell. Thank you for not letting go, I probably wouldn't have bounced back if you hadn't been so supportive. On a more personal note, Anais, you are an amazing person, don't ever forget that. Your criticism and advice have helped me to move forward more than I ever thought possible. You refocused me when I was going in all directions and never hesitated to tell me I was wrong. Thank you for this friendship which has been one of the pillars of my thesis and which will continue well beyond. I will miss you Justine, your outspokenness, and our 1 a.m. wordplay sessions. I thank you for being there throughout my thesis but especially at the end. Thank you for keeping me company in the lab until late while I was finishing my last experiments, thank you for correcting my manuscript. I look forward to working with you again as a post-doc in Finland!

My heart brother, how I missed you when you left! Seeing you go before me was a real heartbreak. I would like to say that you are a traitor and that you could have stayed one more year with me in the lab, but I know that you had reached the end of what the thesis could bring you. Gilles, your energy, and your joy have been such a beacon for me during these years in the lab. The good thing is that even now, this is still the case. The situation has changed, the Gilles beacon has become a family beacon. Gilles, Laure, I am so happy for all these years. I am grateful for the trust you have placed in me with Victoria. I promise to protect her, to support her and to be there as much as you have been for me. I find it very hard to express and show my feelings, but you are my family, and I will always be there for you.

Ava, my sweet doe. I can't find the words to tell you how grateful I am. We have shared and been through so much together that you know exactly how much I owe you, even if you will deny it again and again, I will never forget it. You have been my biggest support without even looking for it, I would have been lost without you. I am sad that this part of our story is over, I truly loved working with you (and living in your close neighbourhood). But I'm still more than eager to find

out what's next, our story is just beginning. I am as grateful as I am proud of you. I have had the privilege of witnessing your evolution, affirming, and revealing yourself. You are my spiritual heart and soul, my family, continue this path my sweetheart, make your way! You are an incredible person, don't be ignorant of the world but keep your innocence, I will be there to protect you, support you and catch you up. See you soon beautiful.

Antoine, our relationship, and our paths have evolved well throughout my thesis. You were there from the beginning, trying to understand and support me as best you could. And we both know that it was not always easy with our two opposite characters, my hot-bloodedness and stubbornness. Beyond demystifying craftsmanship, I hope I have given you as much as you have given me, you have taught me one of my most important lessons, you have taught me that I am worth it. I am proud to count you among my close friends.

Once upon a time, a bald, anime fan became my friend. Raphael, you have been my friend for so many years, high school seems so far away! Thank you for helping me with the non-scientific details of my thesis manuscript. Thank you for staying online with me for hours at night to help me stay awake and not feel lonely. Now the Finnish adventure is upon us little mum!

Julien, thank you for using your talents to save me time. I am particularly proud of the illustrations in my thesis introduction, but it is true that I could not find the time to make the final corrections. Thank you for responding to my call for help out of the blue.

Vladimir, I know you are surprised to see you appear in my thanks, but as I told you, the thesis is not a simple chapter but a real tome, the secondary characters have as much importance as the main characters. Thank you for being there to help me with my daily crafting troubles, thank you for cooking and delivering food when it became the least of my priorities (even though you give me no choice in this matter!).

My family, I am so proud to have finished. Who would have thought I would go to this level of study, certainly not me, I didn't like studying in high school, it came late. Thank you for being there in your own way. Pierre-Henri, those hours of playing during this essay allowed me to force myself to take small breaks that were more than necessary. Jyri, your corrections and suggestions in English were more than appreciated, thanks for taking the time. Christelle, thank you for taking me in during my strategic escape from the heat wave that never dared to came. A change of environment was beneficial. Mom, thank you for trying to understand and push me to keep going. I know you were patient when I kept saying I wanted to stop research (typical exhausted PhD student speech). You were right, I will continue but this time in Finland. Prepare a small temporary place for me at your place, I'm coming! Thanks to all of you, I hope to make you proud.

To the Finns, Marjatta, Old boy and Peppi, thank you for your direct and indirect support. Thank you for taking the time to read my thesis, for changing my mind and for giving me a purpose when I needed it. You have helped me to keep my dream of Finland alive, which has become my only strength to face exhaustion.

Finally, although this is a practice, I imagine is unorthodox, I would like to thank my cats. They understand absolutely nothing about science and our discussions did not help me much in elucidating TAF4, but they were my first line of defense. In the emotional turmoil of the thesis, they accepted all my mood swings and were a great source of comfort on a daily basis. My furballs, you are the most precious thing I have.

This thesis has taught me a lot and pushed me further into my limits. I have come out of it taller and hopefully wiser. I am proud of my work and happy with the steadfast friendships I have developed. But if there is one thing I will never stop affirming and acknowledging, it is that I am insolently lucky to have your support. To all of you, I say thank you, I am happy to have lived (and finished!) this story while being so well surrounded!

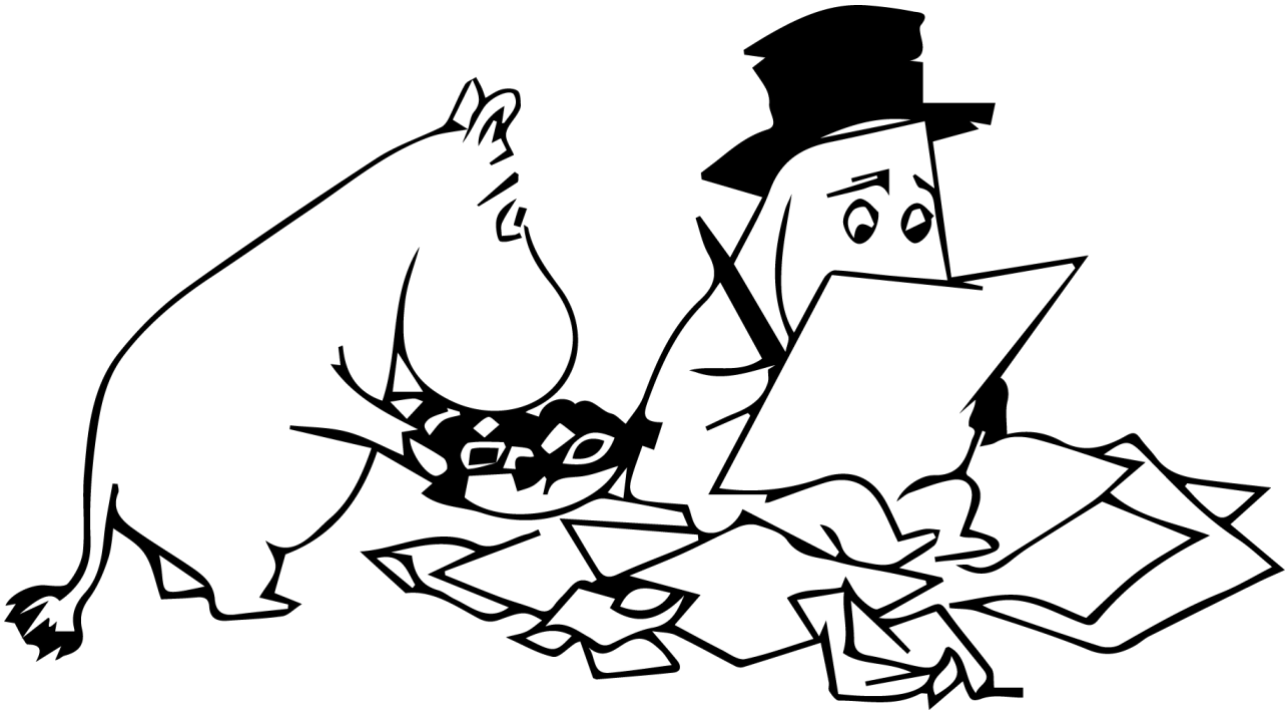


TABLE OF CONTENTS

TABLE OF ILLUSTRATIONS

INTRODUCTION

CHAPITER 1. The Intriguing TAF4 Protein and Gene Transcription Initiation

CHAPITER 2. Tales of the Intestinal Development

CHAPITER 3. The Network of Intestinal Homeostasis

CHAPITER 4. Backstage of the Incredible Intestinal Turn-Over

MY THESIS WORK

MATERIAL AND METHODS

RESULTS

ABBREVIATIONS

A

α-SMA: α-smooth muscle actin

aa: amino acids

AhR: aryl hydrocarbon receptor

ALK-P: alkaline phosphatase, intestinal

ALPI: alkaline phosphatase, intestinal

Aldo: Fructose-Bisphosphate aldolase

AldoA: Fructose-Bisphosphate aldolase A

AldoB: Fructose-Bisphosphate aldolase B

AMP: Adenosine monophosphate

ANPEP: alanyl aminopeptidase, membrane

AP: anterior posterior

APC: adenomatous polyposis coli

ATOH1: Atonal BHLH Transcription Factor 1 also known as Math1

ATP: Adenosine TriPhosphate

AVE: anterior visceral endoderm

B

bHLH: basic helix-loop-helix

Bmi1: also known as the polycomb protein PCGF4

Bmi1: another appellation for RING4

BMP: bone morphogenetic proteins

bp: base pair

BRE: TFIIB recognition element

C

CBC: crypt- base columnar

CBCs: crypt- base columnar cells

CBX: chromobox

CBX/2: chromobox / 2

Cck: cholecystokinin

CD: Crohn's disease

CDHR: cadherin related family

CDHR/2: cadherin related family member / 2

CDX: caudal type homeobox

CDX/2: caudal type homeobox / 2

Chg/A: chromogranin / A

Chg/B: chromogranin / B

cPCR1: canonical PCR1

CpG: cytosine-phosphate-guanine

CRII: carboxy conserved region II

D

DC: **d**endritic **c**ell

Dclk1: **d**ouble**c**ortin-**l**ike **k**inase **1**

DE: **d**efinitive endoderm

Defa: α-defensins

DNA: deoxyribonucleic acid

dpc: days post coitus

DPE: downstream core element

Dpp4: dipeptidyl peptidase 4

Dll -1/-4: Delta-like 1 and 4

DVE: distal visceral endoderm

E

E: embryonic day

ECM: extra cellular matrix

EEC: enteroendocrine cell

EGF: epidermal growth factor

EPI: epiblast

a-EPI: anterior EPI

p-EPI: posterior EPI

ER: endoplasmic **r**eticulum

Erk: extracellular signal-regulated kinases

ESC: embryonic stem cell

EZH2: enhancer of zeste 2 polycomb repressive complex 2 subunit

F

Fabp: fatty acid binding protein

Fabp/1: fatty acid binding protein / 1

FGF/4: fibroblast growth factors / 4

FGFR: fibroblast growth factors receptor

FZD: Frizzled

G

GALT: gut-associated lymphoid tissue

GATA: DNA sequence "GATA"

GATA/6: GATA-binding factor / 6

GIP: gastric inhibitory polypeptide

Gip: Glucose-dependent insulinotropic polypeptide

GLP1: glucagon-like peptide-1

GPR81: G protein-coupled receptors 81

GTF: general transcription factor

GSK-3: glycogen synthase kinase 3

H

H: histone

HAT(s): histone acetyltransferase(s)

HDAC(s): histone deacetylase(s)

HES: hairy and enhancer of split

HES1: hairy and enhancer of split 1

HES-1: hes family bHLH transcription factor 1

HFD: histone fold domain

Hh: hedgehog pathway

HMT: histone methyltransferase

HNF: hepatocyte nuclear factor

HNF4: hepatocyte nuclear factor 4

HNF4A: hepatocyte nuclear factor 4 Alpha

HNF4G: hepatocyte nuclear factor 4 Gamma

Hopx: homeodomain-only protein homeobox

I

IBD: Inflammatory bowel diseases

IECs: essentials **i**ntestinal **e**pithelium **c**ell**s**

IgA: Immunoglobulin A

Ihh: **i**ndian **h**edge**h**og

IL-22: interleukine 22

ILCs: innate lymphoid cell

IMCs: intestinal mesenchyme cells

iNKT: Invariant natural killer T

Inr: initiator element

ISCs: intestinal stem-cells

K

K: lysine (K)

kb: kilobase (1000 base pairs)

KDM2B: lysine (K)-specific demethylase 2B

KLF4: Kruppel-like factor 4

L

LGR5: leucine rich repeat containing G protein-coupled receptor 5

Lyz: lysozymes

LPH: Lactase–phlorizin hydrolase

LPS: Lipopolysaccharides

LRCs: label- retaining cells

Lrig1: Leucine Rich Repeats and Immunoglobulin Like Domains 1

LRP5/6: lipoprotein receptor 5/6

M

MAMP: microbial-associated molecular patterns

MAP-Kinase: mitogen-activated protein kinase

MATH1: another appellation for ATOH1

M cells: Microfold cells

MCs: mesenchyme cells

me/1/2/3: mono-(1) / di-(2) / tri-(3) / methylation

mESC: mouse embryonic stem cell

MGAM: maltase-glucoamylase

Mmp-7: matrix metalloproteinase 7

MUC: mucin

MUC/2: mucin / 2

N

nanog: name derived from Celtic mythology

ncPCR1: non-canonical PCR1

NEUROG3: neurogenin 3

NHR1: Nervy homology region 1

Notch1: notch receptor 1

Nts: neurotensin

O

OCT4: octamer-binding transcription factor 4

Olfm4: Olfactomedin-4

P

ParE: parietal endoderm

PcG: polycomb group

PCGF/5: polycomb group ring finger / 5

PCL: polycomb-like proteins

PD1: death protein 1

PDGF/A: platelet-derived growth factor / A

PDGFRα: PDGF receptor alpha

PEPT1: peptide transporter 1

PHC: polyhomeotic homolog

PHC/1: polyhomeotic homolog / 1

PIC: pre-initiation complex

POU: named after Pit-1, Oct-1/3, and neural Unc-86 transcription factors

Pou2f3: POU class 2 homeobox 3

Ppar1: peroxisome proliferator activated receptor 1

PPs: Peyer's patches

PRC1: polycomb repressive complex 1

PRC2: polycomb repressive complex 2

PREs: PcG response elements

PrE: primitive endoderm

Prox1: Prospero Homeobox 1

PRRs: germline-encoded pattern recognition receptors

PTM: post-translational modification

PXR: pregnane X receptor

PYS: parietal yolk sac

Pyy: peptide tyrosine (Y) tyrosine (Y)

R

RANK: Receptor activator of nuclear factor κ B

RANKL: RANK ligand

RB: retinoblastoma protein

Rbbp4: RB binding protein 4

RegIII: Regenerating islet-derived protein III

RING/1: really interesting new gene / 1

RNA: ribonucleic acid

RNA pol II: RNA polymerase II

RYBP: YY1 binding protein

S

SCFAs: Short-chain fatty acids

Scm: sex comb on midleg homolog

Sct: secretin

SGLTs: facilitative glucose transporter also called GLUTs (sodium independent)

SGLT6: sodium-glucose cotransporter type 6

Shh: sonic hedgehog

SI: Sucrase–isomaltase

Sis: sucrase isomaltase (alpha-glucosidase)

SLC: solute carrier family

SLC/15A/1: sodium-dependent glucose transporters or solute carrier family / 15, member / 1

SMCT2: sodium-coupled monocarboxylate transporters / 2

SOX: SRY-box transcription factor

SOX/9: SRY-box transcription factor / 9

SPDEF: SAM Pointed Domain Containing ETS Transcription Factor

Sst: somatostatin

T

TA-cells: transit- amplifying cells

Tac1: tachykinin 1

TAF: TATA-box binding protein associated factor

TAF/4: TATA-box binding protein associated factor / 4

Taf4a$^{lox/lox}$: VilCre embryos (hereafter Taf4a^{lEndoC})

Taf4a$^{lox/lox}$: VilCre^{ERT2} adult mice (hereafter Taf4a^{lEC})

Taf4a$^{lox/lox}$: Lgr5-GFP-Cre^{ERT2} mice (hereafter Taf4a^{CBC})

TAFH: TAF homology

TAZ: WW domain containing transcription regulator 1

TBP: TATA box binding protein

TE: trophectoderm

TEA: DNA-binding domain named after TEF-1 & abaA

TEAD/4: TEA domain transcription factor / 4

Tert1: telomerase reverse transcriptase

TF: transcription factor

Tff2: trefoild factor 2

TFII: transcription factor polymerase II

TNF: tumor necrosis factor

TREH: trehalase

TRF: TBP-related factor

Tphr1: transferrin receptor

Trpa1: Transient Receptor Potential Cation Channel Subfamily A Member 1

TSS: transcription start site

U

ub1: mono-(1)-ubiquitination

UC: ulcerative colitis

V

VE: visceral endoderm

VFM: ventral folding morphogenesis

VYS: visceral yolk sac

W

WNT: wingless-type MMPTV integration site

WNT/5a: Wnt family member / 5a

Y

YAF2: YY1-associated factor2

YAP: yes-associated protein

YY1: Yin Yang 1 transcription factor

ABSTRACT

SPOILER ALERT !

La machinerie transcriptionnelle de base dans l'homéostasie intestinale

L'épithélium intestinal est un tissu à l'interface entre la lumière intestinale et l'organisme, jouant un rôle majeur de digestion, de barrière et de transport. C'est un tissu en renouvellement cellulaire constant à partir des cellules souches du fond de la crypte intestinale. Celles-ci génèrent des progéniteurs qui à leur tour donnent des cellules filles non proliférantes qui vont se différencier au cours de leur migration le long de la villosité. Le renouvellement cellulaire répond à une organisation topologique et fonctionnelle très précise, passant notamment par les voies Wnt, Notch, YAP et BMP qui contrôlent directement ou indirectement l'expression des gènes dans l'intestin (Medema et al. Nature. 2011 [1]). Ces voies aboutissent à des facteurs de transcription (ex : Tcf4, Hes1, Math1, Smads, Cdx1/2, HNF1/4, GATA5/6, KLF4/5) qui interagissent avec la machinerie transcriptionnelle de base, peu étudiée dans l'homéostasie intestinale.

HNF4a est un récepteur nucléaire impliqué dans le métabolisme et le développement du foie, des cellules β pancréatiques et de l'intestin en contrôlant l'expression de nombreux gènes. L'étude des modèles murins de KO génique conditionnel d'HNF4a dans l'intestin embryonnaire (Garrison et al. World J Gastroenterol. 2006 [2]) ou adulte (Villin-Cre-ER^{T2}, Cattin et al. Mol Cell Biol. 2009 [3]) ont permis d'étudier l'impact d'HFN4α respectivement sur le développement de l'épithélium intestinal et sur l'épithélium adulte. Ces études ont montré l'importance d'HNF4α dans le contrôle de la prolifération, la différenciation cellulaire ainsi que dans le maintien de l'architecture intestinale. L'inactivation d'HNF4α entraine une hyperprolifération due à une augmentation de la voie de signalisation Wnt. Ceci affecte également la formation des cryptes, la maturation des cellules épithéliales et augmente la perméabilité de l'épithélium intestinal. Par ailleurs, HNF4α est également impliqué dans les cancers du côlon, dans certaines maladies inflammatoires et des mutations dans le gène HNF4α sont responsables du diabète de type MODY1 (Maturity Onset Diabetes of the Young type 1). De récents travaux ont permis de mettre en évidence une interaction physique et fonctionnelle entre HNF4α et les sous-unités TAF4-TAF12 de la machinerie transcriptionnelle de base TFIID (Transcription Factor II D) dans le foie (Alpern et al. eLife. 2014 [4]). TFIID est un complexe multiprotéique composé de la TATA-Binding Protein (TBP) et de 13 TBP-Associated Factors (TAFs) dont TAF4 et TAF12 qui forment un hétérodimère.

Mes travaux de thèse visent à étudier le rôle spécifique de la machinerie transcriptionnelle de base dans l'intestin. Pour cela, plusieurs modèles murins d'invalidation génique de TAF4a ont été utilisés pour déduire son rôle (1) dans le développement et la

morphogénèse de l'intestin chez l'embryon ; son implication (2) dans l'homéostasie intestinale chez l'adulte ainsi que son implication dans d'éventuelle (3) pathologies intestinales. Un dernier modèle d'organoïdes a été développé à partir du modèle murin (4) pour étudier plus précisément le rôle de TAF4a dans l'épithélium intestinal sans son microenvironnement.

1) Conséquences de l'inactivation somatique de TAF4a dans l'endoderme embryonnaire

Pour étudier l'implication de TAF4a dans le développement de l'intestin, les souris TAF4a$^{lox/lox}$ dont nous disposons au laboratoire (Mengus et al. EMBO Journal. 2005 [5]) ont été croisées avec le modèle Villine-Cre. Ce modèle permet d'induire la perte d'expression conditionnelle du gène TAF4a à partir de E11-12 dans l'épithélium intestinal. Ce modèle présente une forte mortalité périnatale ; 24h après la naissance aucun souriceau déplétés pour TAF4a n'est viable.

Par immunohistochimie, il a été montré une perte d'expression graduelle de TAF4a au cours du temps, spécifique des cellules de l'épithélium intestinale. Ainsi la perte devient totale vers E17.5 et les fœtus présentent alors un intestin et des villosités plus courts. A E18.5, l'épithélium devient plat avec quelques rare petites villosités boudinées. Cet épithélium plat n'est pas associé à une diminution du nombre de cellules en prolifération, mais présente une augmentation du nombre de cellules en apoptose. L'expression du marqueurs OLFM4 des cellules souches et progéniteurs matures est fortement diminué. L'expression des facteurs de transcriptions impliqués dans l'identité intestinale tels que CDX2 et HNF4α 1-6 ne sont pas altérés tandis que HNF4α 7-9 et HES1 sont faiblement exprimés et de manière irrégulière. Les marquages Mucine 2, Bleu Alcian, Chromogranine A, et Phosphatase Alcaline montrent que les cellules épithéliales sont capables d'arriver à leur stade de différenciation terminale. Toutefois, la différenciation n'est pas parfaite, il y a une forte diminution des cellules matures du lignage sécréteur et absorbant confirmé par comptage des différentes populations cellulaires.

Les analyses immunohistochimiques ont été complétées par des expériences de RNA-Seq pour déterminer les changements de profils d'expression génique au cours du développement. Les résultats de ces expériences montrent qu'il y a plus de gènes down-régulés que de gènes up-régulés (1160 contre 173). L'analyse KEGG des résultats montrent que les voies plus dérégulées sont des voies impliquées dans le métabolisme et l'immunité.

Au cours du développement, TAF4a est donc essentiel à l'homéostasie et la maturation de l'épithélium intestinal.

2) Conséquences de l'inactivation somatique de TAF4a dans l'épithélium de souris adulte

Dû à la mortalité périnatale des souris inactivées pour TAF4a, les souris TAF4alox/lox sont croisées avec le modèle inductible au tamoxifène, Villine-CreERT2 pour permettre l'étude chez l'adulte. La perte d'expression de TAF4a spécifiquement dans l'épithélium intestinal a été validé par immunohistochimie.

Après induction de la perte de TAF4a, les souris adultes présentent une légère diminution de la survie au cours du temps, associé avec une courte période de perte de poids forte. Après cette période, les mâles récupèrent mieux au cours du temps comparés aux femelles qui n'arrivent pas à dépasser leur poids initial. Les souris présentent également des désordres intestinaux, les fèces sont moelles, le caecum est dilaté remplit de gaz, dont la présence est également observée dans l'intestin et le côlon.

L'analyse histologique et immunohistologique de l'iléon, montre que les souris déplétées pour TAF4a n'ont pas d'altération de l'organisation en crypte-villosité de l'épithélium. Elles présentent cependant une hyperprolifération et une augmentation de la vitesse de migration des cellules épithéliales le long des villosités. Comme pour les fœtus, les adultes présentent des défauts de maturation des cellules différenciées et une diminution de l'expression du marqueur OLFM4. Ce résultat s'accompagne avec une altération de l'expression du lysozyme des cellules de paneth de la niche intestinale. De même, l'expression des facteurs de transcription essentiel à l'identité et l'homéostasie intestinale, CDX2, HES1, HNF4α 1-6 et 7-9 ne sont pas altérés.

L'analyse du transcriptome a permis de mettre en évidence le lien entre l'expression des ARNs et les niveaux protéiques observés par immunohistochimie. Comme pour le fœtus, nous retrouvons plus de gènes down-régulés que de gènes up-régulés (769 contre 284). L'analyse KEGG montre que les voies les plus impactées par la perte de TAF4a sont liées à la différentiation ainsi que la réponse immunitaire. Parmi les gènes down-régulés, 296 sont en commun entre les fœtus et l'adulte. Ils sont principalement impliqués dans les voies du métabolisme et l'immunité (analyse KEGG) ; 53 gènes up-régulés sont en commun entre les deux modèles et correspond à des histones.

Pour étudier le lien entre les cellules souches et la différentiation cellulaire, suite à l'invalidation de TAF4a, nous avons utilisé un dernier modèle de souris le modèle LgR5-GFP-CreERT2. Ce modèle permet l'extinction de TAF4a spécifiquement et uniquement dans les cellules souches du fond de la crypte, exprimant LgR5. L'épithélium a été observé 6 à 8 jours après induction au tamoxifène. Comme attendu par ce modèle mosaïque, seules quelques cryptes présentent des cellules souches inactivées pour TAF4a tandis que les cellules de

paneth conservent son expression. Les villosités liées à ces cryptes présentent des cellules inactivées pour TAF4a présentant une activité de la phosphatase alkaline diminuée.

Tous ces résultats mis ensemble montre que l'invalidation de TAF4a perturbe l'équilibre de l'homéostasie intestinale. En passant des cellules de la niche aux des progéniteurs pour aboutir à des défauts de différentiations et maturations des cellules différenciées.

3) Conséquences pathologiques de l'inactivation de TAF4a dans l'épithélium de souris adulte

L'épithélium intestinal permet le lien entre la lumière intestinale et l'organisme. Il a également un rôle de barrière avec la lumière contenant notamment le microbiote. La présence de gaz dans le caecum et de diarrhées suggèrent une altération de ce dernier. L'analyse du RNA 16S des populations bactériennes du caecums de souris TAF4a KO et contrôles montrent une altération de la diversité microbienne. Les souris KO présente une augmentation de population bactérienne pathogènes d'Helicobacter, Clostridium XIVb et XVIII. Ce changement de population bactérienne est associé à une augmentation de la perméabilité paracellulaire non associée à une perméabilité transcellulaire. Le bouleversement du microbiote et de la fonction de barrière nous ont conduit à analyser les données de RNA-seq avec la "méthode de comptage de la population cellulaire du microenvironnement murin". Ce module nous as permis de prédire l'infiltration d'une plus grande population de basophiles et cellules B dérivées dans le mésenchyme sous-jacent. Le marquage CD19 a permis de confirmer une augmentation du nombre de cellules B dérivée dans la *lamina propria*.

L'altération de la fonction de TAF4a induisant une augmentation de prolifération et une altération de la signature des cellules souches, nous avons étudié son l'impact dans la progression tumorale. Pour étudier cela, les souris TAF4alox/lox ::Villine-CreERT2 ont été croisées avec les souris APCΔ14/+ qui développent spontanément des tumeurs dans l'intestin et plus rarement dans le côlon. Les souris mâles TAF4a KO étant en meilleur forme physique sur le long terme que les femelles, ils ont été choisis comme cohorte pour l'étude de la cancérologie. Sur le long terme, les souris déficientes pour TAF4a ne présentent pas d'apparition spontanée de tumeur. En revanche, associé au modèle de prédisposition APCΔ14/+, la perte de fonction totale de TAF4a entraine une très forte augmentation de la tumorigénèse comparé au modèle APCΔ14/+ seul. La survie des souris est très fortement compromise. Les analyses histologiques et immunohistologiques des tumeurs n'indique pas de phénotypes sensiblement différents entre les deux modèles.

4) Le modèle d'organoïde intestinal

Étant donné que la perte de TAF4a dans l'épithélium impact le microenvironnement de ce dernier, nous avons développé deux lignées d'organoïdes d'iléon, l'une provenant de souris contrôle et l'autre de souris inductible TAF4lox/lox ::Villine-CreERT2. Les organoïdes intestinaux sont des structures multicellulaires 3D, cultivées ex vivo reprenant la forme et la composition de l'épithélium intestinal. Les organoïdes sont sphériques avec des « bourgeons », cette partie correspondant aux cryptes intestinales, tandis que la partie sphérique correspond aux villosités. Cette technique est complémentaire à l'étude in vivo et permet d'observer l'impact de TAF4a, uniquement sur l'épithélium isolé, indépendamment du mésenchyme sous-jacent et de la lumière intestinale.

Après induction de la perte de TAF4a, 2h après mise en culture, nous avons observé une forte altération de la structure et de la survie des organoïdes avec moins de 90% de survivants après 8 jours de culture. Les organoïdes perdent leur belle structure bourgeonnante, gardent une petite taille et accumule des débris et de l'ADN dans leur lumière. Par immunofluorescence, les organoïdes présentent une très nette diminution de leur capacité de prolifération (marquage Ki67 et Phospho-Histone H3), ce qui est contraire aux observations faites chez les souris invalidées pour TAF4a. Cette perte ne semble également pas induire d'altérations de la proportion de cellules sécrétrices de mucus (Mucine2) et d'hormones (Chromogranine A). En ce qui concerne la niche intestinale, nous avons mis en évidence une diminution du marquage OLFM4 des cellules souches et progéniteurs tandis que le marquage lysozyme nous permet d'attester de la présence de cellules de paneth. Nous avons noté une très forte augmentation du nombre de cellules en apoptose observées dès 24h après inactivation.

Des analyses du transcriptome (RNA-Seq) ont été effectuées après 3 jours de culture et montrent une diminution de l'expression de 1625 gènes. Sur ces gènes, 182 font partie des 510 gènes de la signature des cellules souches (Muñoz et al., 2012 [6]). L'analyse KEGG a montré que les principales voies impactées font partie de la réplication et de la réparation de l'ADN. Nous avons également pu mettre en évidence la diminution de l'expression d'Ascl2 et TEADII, impliquées dans la régénération de l'intestin par la dédifférenciation des cellules différenciées en cellules souches. Dans les gènes up-régulés, nous retrouvons 1120 gènes, principalement associés les des voies de carcinogénèses chimiques et des voies métaboliques.

L'analyse du profil d'expression génique a été complétée par une analyse ATACseq. Les résultats de cette étude, nous permettent d'observer les régions de la chromatine ouverte et d'en déduire les régions potentiellement activement transcrites. De ce fait, nous avons identifiés dans les contrôles et les mutants environ 118 500 pics associés à des gènes. Sur ces

pics, 93 386 sont communs aux deux conditions donc inchangés, 11 283 sont perdus dans les mutants et 12 831 sont gagnés. Sur les pics modifiés, 1675 correspondent aux régions proximales de la région promotrice de gènes. Nous avons donc concentré nos analyses sur ces derniers. Sur les gènes identifiés, 93 sont perdus aussi bien dans l'analyse de l'ATACseq que celle du RNAseq. Ces gènes font partie de voies de recombinaison homologue et de réparation de l'ADN. Cette expérience nous permet de mettre en évidence une altération de la structure de la chromatine avec une modification des régions ouvertes et fermés, ne traduisant pas forcément de l'expression des gènes. Nous avons également observé que de nombreux pics perdus correspondent à des motifs reconnus par le facteur de transcription HNF4α entre autres. À la suite de ces analyses, nous avons détectés par immunofluorescence une apparition de spots H2AX dans les noyaux dans les KO. Cette protéine est impliquée dans les mécanismes de réparation de l'ADN en participant à un check-point conduisant à l'arrêt du cycle cellulaire et la réparation de l'ADN par cassure double brin.

Pour compléter l'analyse, nous avons induit la perte de TAF4a, cette fois-ci 5 jours après mise en culture des organoïdes afin d'observer l'effet de la perte de TAF4a sur des organoïdes bien établit et différenciés. De même que pour la condition précédente, les organoïdes perdent leur capacité à bourgeonner et les bourgeons déjà présents disparaissent au cours court du temps. Le marquage OLFM4 n'est pas retrouvé et l'apoptose est fortement augmentée. Afin d'étudier l'effet de TAF4a sur les différents types cellulaires, nous avons réalisé une analyse de single-cell sur les organoïdes 3 jours après induction au tamoxifène avant la dégradation des bourgeons. Plusieurs clusters de cellules ont pu être identifiés. Les clusters de cellules souches des contrôles sont fortement réduit chez les KO qui à l'inverse voient le cluster de population non différentiées augmenter. Le cluster des cellules souches restantes, présente des marqueurs apoptotiques. A ce stade de l'analyse, les cellules à mucus et de paneth ne présentent pas d'altération majeures tandis que les cellules absorbantes et entéroendocrines présentent déjà des altérations de différentiation mature.

De nombreuses voies de signalisation sont impliqués dans le maintien de l'homéostasie intestinale. Parmi ces voies, les voies Wnt et Notch possèdent un rôle crucial dans la dynamique de la crypte intestinale. Étant donné que les cellules souches sont fortement altérées par l'absence de TAF4a et que l'expression de plusieurs acteurs de ces voies sont altérés, nous avons essayé de sauver le phénotype des organoïdes en ciblant ces voies. L'utilisation de l'activateur de la voie Wnt, CHIR99021 ; du DAPT atténuant l'action de Notch ou du VPA un activateur des deux voies ; ne permettent pas de sauver le phénotype. Dans la mesure où l'ATACseq et le RNA-seq ont permis de mettre en évidence une altération de la structure de la chromatine et une inhibition de nombreux gènes dont la région est pourtant ouverte, nous nous sommes orientés sur les complexe répressifs des polycombes. Ces derniers sont impliqués dans la régulation et l'inhibition de l'expression des gènes au cours du

développement et des cancers (Piunti and Shilatifard, 2021 [7]). Après traitement avec la drogue inhibitrice EPZ-6438, nous avons pu observer une survie des organoïdes, l'apparition de bourgeon plus tardivement et le sauvetage du phénotype.

Dans son ensemble, ce travail montre que TAF4 est impliqué dans l'homéostasie intestinale en commençant dès son développement. Son action passe par les cellules souches en s'opposant à l'action des polycombes avant terminer par les cellules différenciées. Ce travail a également permis de mettre en lumière le rôle spécifique que peut avoir la machinerie transcriptionnelle de base dans l'intestin.

COMMUNICATIONS SCIENTIFIQUES

- **2019. Poster International.** CRCL Symposium (The Cancer Research Center of Lyon). Lyon. France

  Relevance of the general transcription factor TAF4 in intestinal development, homeostasis, and cancer.

  Susanna Säisä-Borreill, Thomas Kleiber, Elisabeth Martin, Guillaume Davidson, Irwin Davidson, Isabelle Duluc, Jean-Noël Freund

- **2021. Communication.** CECED. Congrès en ligne. France.

  La machinerie transcriptionelle de base dans l'homéostasie intestinale.

  Susanna Säisä-Borreill, Thomas Kleiber, Guillaume Davidson, Stanislas Mondot, Hervé Blottière Elisabeth Martin, Irwin Davidson, , Isabelle Duluc, Jean-Noël Freund

- **2021. Poster.** JFHOD. Congrès en ligne. France.

  Importance du facteur général de transcription TAF4 dans l'homéostasie et la cancérogenèse intestinale.

  Susanna Säisä-Borreill, Thomas Kleiber, Elisabeth Martin, Guillaume Davidson, Irwin Davidson, Isabelle Duluc, Jean-Noël Freund

- **2021. Poster.** Journée des FMTS. Congrès en ligne. France. Prix meilleur poster.

  Importance du facteur général de transcription TAF4 dans l'homéostasie et la cancérogenèse intestinale.

  Susanna Säisä-Borreill, Thomas Kleiber, Elisabeth Martin, Guillaume Davidson, Irwin Davidson, Isabelle Duluc, Jean-Noël Freund

MANUSCRIT EN COURS DE REDACTION

- **The general transcription factor TAF4 opposes epigenetic silencing of intestinal stem cell function by Polycomb**

  Susanna Säisä-Borreill[1], Guillaume Davidson[2], Thomas Kleiber[2], Stanislas Mondot[3], Hervé Blottière[3], Elisabeth Martin[1], Michela Plateroti[1], Gabrielle Mengus[2], Isabelle Duluc[1], Irwin Davidson[2], Jean-Noël Freund[1]

  [1] University of Strasbourg, Inserm, UMR-S1113 / IRFAC, FHU ARRIMAGE, FMTS, 67200 Strasbourg, France

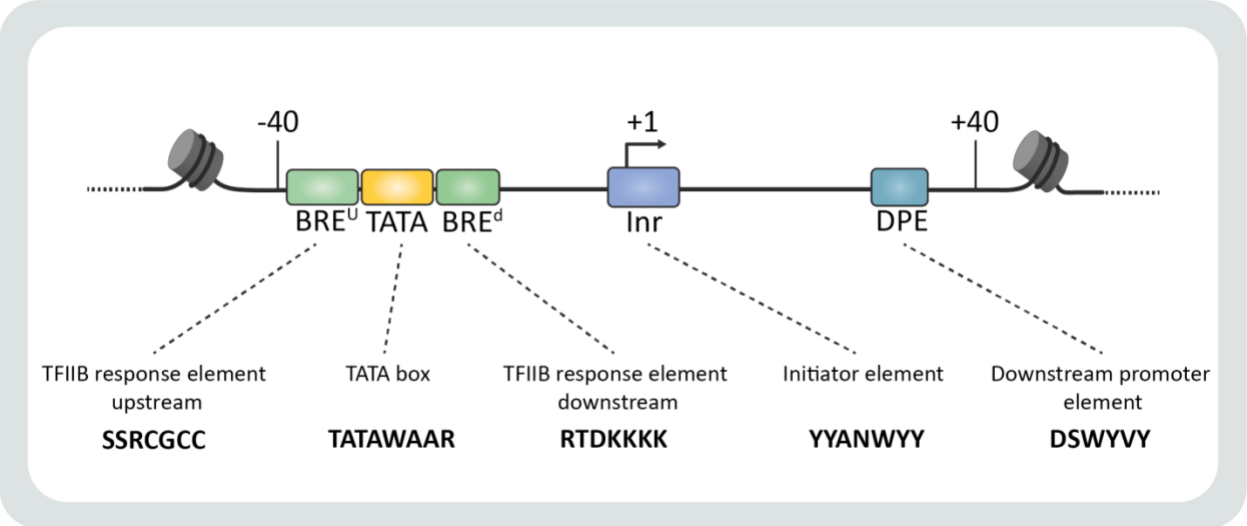
  [2] Department of Functional Genomics and Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/UNISTRA, 1 Rue Laurent Fries, 67404, Illkirch Cédex, France

  [3] University Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350, Jouy-en-Josas, France

INTRODUCTION

CHAPTER 1

The Intriguing TAF4 Protein and Gene Transcription Initiation



Adapted from [8], [9]

Figure 1. Core Promoter of Protein-Coding Genes and Its Most DNA Common Elements

Schematic organization of the core promoter elements in eukaryotic protein-coding genes with their consensus sequences. The +1 gives the position of the transcription start site (TSS).

Key, IUPAC nomenclature. H (A/T/C); K (T/G); M (A/C); N (A/T/C/G), R (A/G); S (C/G); Y (T/C); W (A/T).

The transcription is the process by which DNA is transcribed into RNA under the involvement of numerous complex machineries. In eukaryotes, transcription into messenger RNA (mRNA) of class II genes which encode proteins is a fundamental process finely regulated and highly conserved through evolution. In multi-cellular organisms, cells possess the same DNA, but they have different structures and functions depending on the organ or the stage of development. The transcriptional regulation of class II genes permits to achieve this cell diversity, by defining the biological processes such as proliferation, differentiation and cell identity in time and space. Controlling the gene transcription initiation is the first step in gene expression regulation.

I. Transcription Initiation Involves the Basic Transcriptional Machinery

Class II gene transcription involves more than one hundred actors that assemble on the core promoter of protein-coding genes such as General Transcription Factors (GTFs), RNA polymerase II (RNA pol II), co-activators/repressors and Transcription Factors (TFs).

A. Assembly of the PIC Complex, First Step of Transcription Initiation

Transcription begins with the recruitment of the pre-initiation complex (PIC). The RNA pol II is unable to bind to DNA by itself but needs the assembly of the PIC complex. Besides the RNA pol II, PIC is composed of GTFs and the mediator complex. Transcription initiation starts with the binding of the general transcription factor TFIID at the core promoter that I will elaborate below.

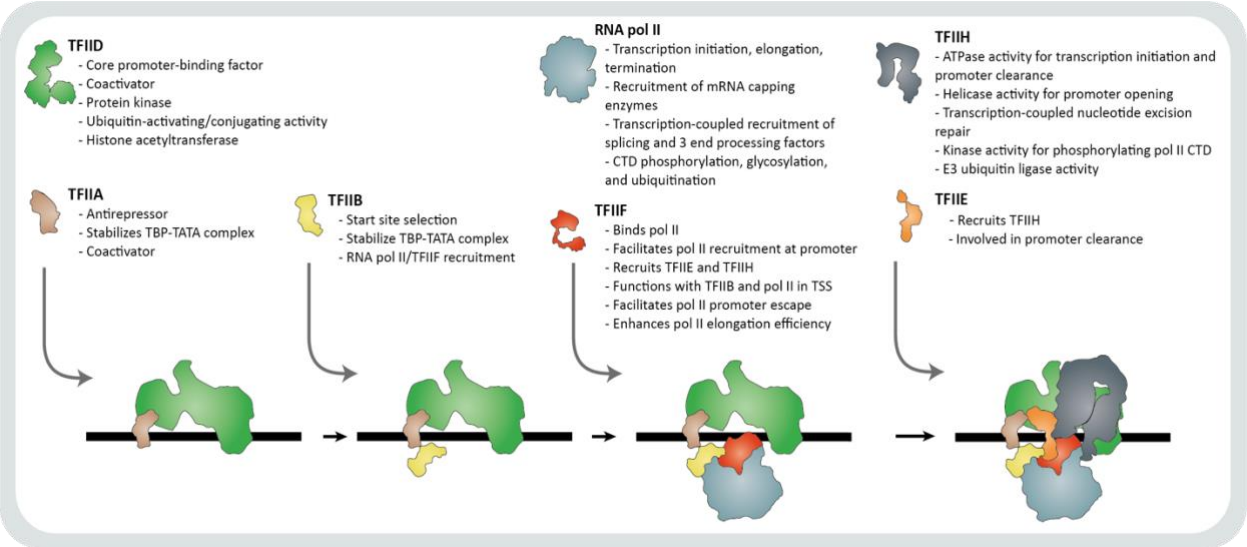
a. The Core promoter, an Assembly Platform

Transcription starts at the site +1, the first transcribed nucleotide; this site is called the Transcription Start Site (TSS). The DNA segment of 50-100 bp around the TSS is named the core promoter. This region of protein-coding genes includes a combination of conserved common DNA elements: TATA box, TFIIB recognition element (BRE), initiator element (Inr) and downstream promoter element (DPE) being the most frequent. The two most common elements present in eukaryotic promoters are the TATA box and Inr (Figure 1) [9]–[11].

The **TATA box** is the first promoter box to have been identified. It can be found in 5-40% of eukaryotic gene promoters. Its localization goes from −33 to −23 bp upstream of the TSS and it is recognized by the TATA box binding protein (TBP) of the TFIID.

The initiator element or **Inr**, encompasses the TSS and can initiate the transcription independently. Like the TATA box, this element is recognized by the TFIID and is detected in 40-70% of eukaryotic promoters, which makes it the prevalent element.

TATA and Inr can be found together in the core promoter or separately in distinct functional gene families. The TATA box is more often found in tissue-specific gene families while Inr is found in ubiquitously expressed genes or "housekeeping" genes.



Inspired from [12], [13]

Figure 2. **Assembly of the Pre-Initiation Complex (PIC) on Core Promoter**

The **BREs** for "TFIIB recognition elements", are bound by TFIIB and detected immediately upstream (BRE^u) or downstream (BRE^d) of the TATA box. Analyses with the Eukaryotic Promoter Database showed that 25% of eukaryotic core promoters potentially contain a BRE^u; and while they work conjointly with the TATA box, the BRE^u is more present in TATA-less core promoters than TATA-containing ones. These elements can have both roles of transcription activator and repressor.

The downstream promoter element (**DPE**) is a binding site for TFIID, found in *Drosophila* and mammalian Inr-containing core promoters. DPE function mostly depends on Inr presence.

b. CpG island

Beside the elements considered as the core, CpG islands are another component of promoters. CpG is the abbreviation for the nucleotide pair liaison, Cytosine-phosphate-Guanine. The CpG site is often surrounded by other CpG, forming a CpG island. It has been reported in human that ~70% of promoters contain a high concentration of CpG [14]. Bioinformatic analysis of human promoters tends to show that most of the TATA+ promoters are CpG-low, like Inr+ and DPE+ promoters. However, BRE+ promoters seem to be more present in CpG-high promoters [15].

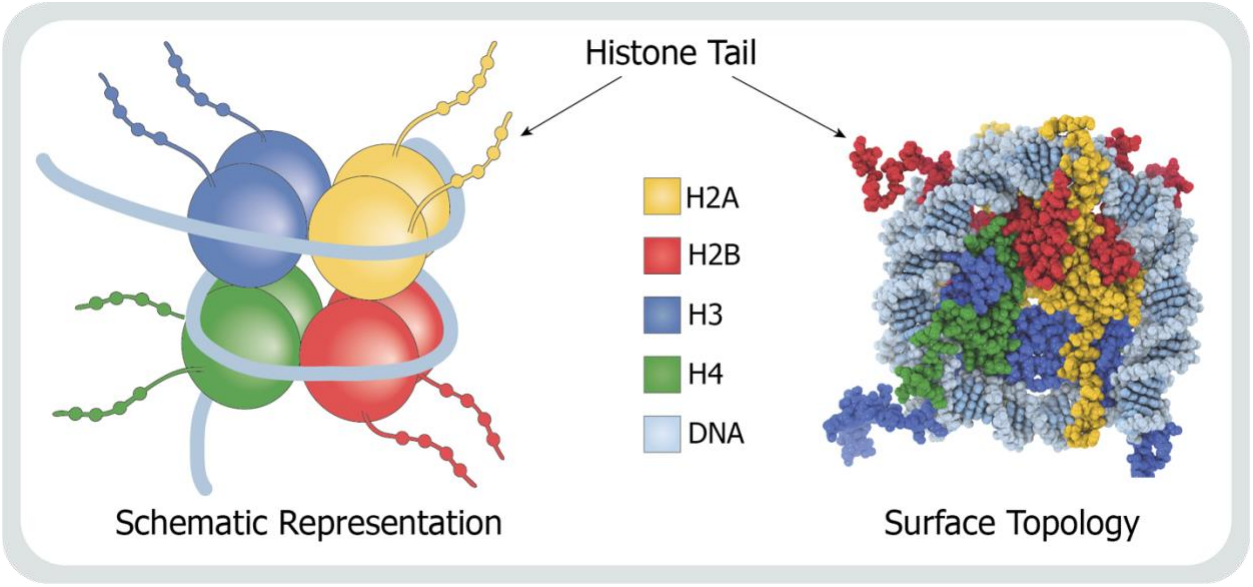
CpG islands are targets of gene expression regulation. In fact, they are subject to epigenetic modifications consisting of cytosine methylation. Methylated CpG are generally associated with transcriptional repression by recruiting gene repressor proteins and reducing interactions between DNA and transcription factors. The epigenetic modulations such as methylation do not involve DNA sequence alterations. It is implicated in gene regulation and can be transmitted to descendants.

c. Assembly of Basal Transcriptional Machinery

Historically, *in vitro* transcription studies helped to identify and isolate the GTFs. They were so called after the chromatographic protein fractions from which they were extracted: TFIIA, TFIIB, TFIID, TFIIE, TFIIF and TFIIH. The RNA pol II corresponds to the TFIIC fraction. Together, TFII fractions formed the PIC. As they were considered as a constant part of the PIC, they were named "general" transcription factors. However, now we know more about the core promoter diversity and the non-ubiquitous expression of some components of TFIID. Due to this variety of PICs, the GTFs play more a "basal" role than a "general" one; it is part of the reason why we talk about a "basal transcriptional machinery" [10].

The **PIC assembly** is strictly necessary to guide the RNA pol II onto the TSS (Figure 2). Each component exerts its own function. First, with the help of TFIIA, TFIID recognizes and binds to the core promoter, then TFIIB is mobilized to stabilize the TFIID-DNA interaction. Successively, TFIIB, which includes the RNA pol II and TFIIF (the RNA Pol II helper) that in turn calls up the TFIIE and TFIIH to finally pave the way for transcription [12], [13].

The **mediator** is a multiprotein complex organized generally with 28 to 30 subunits; most of the mediators are conserved through evolution. Thanks to its numerous subunits, it is involved in collecting, processing, and transferring regulatory information toward the PIC [13].



Inspired from [16], [17]

Figure 3. Nucleosome Structure

On the left schematic representation of a nucleosome with the DNA wrapped around. To the right, nucleosome in top view with bioinformatic space-filling representation [16]. The nucleosome consists of a histone octamer with two pairs of each histone H2A, H2B, H3 and H4. The histone «Tail» represents both C- and N-Terminal tails of histones subjected to post-translational modifications.

The association of the PIC and the mediator is called the **basic transcriptional machinery**. This machinery is necessary for class II gene transcription and can initiate by itself the transcription of constitutive genes essential for basal cell function maintenance, the so-called "housekeeping" genes.

B. Controlling the Transcription Initiation

Transcription by the RNA pol II begins with the assembly of the PIC. Controlling the recruitment of PIC on the DNA means controlling gene transcription. There are plenty of ways to regulate the recruitment of the PIC such as alteration of the gene accessibility by chromatin remodeling, or recruitment/repression of PIC on specific gene promoter by activators, repressors, and transcription factors.

a. Remodeling the Chromatin, Zoom on the Polycomb Group

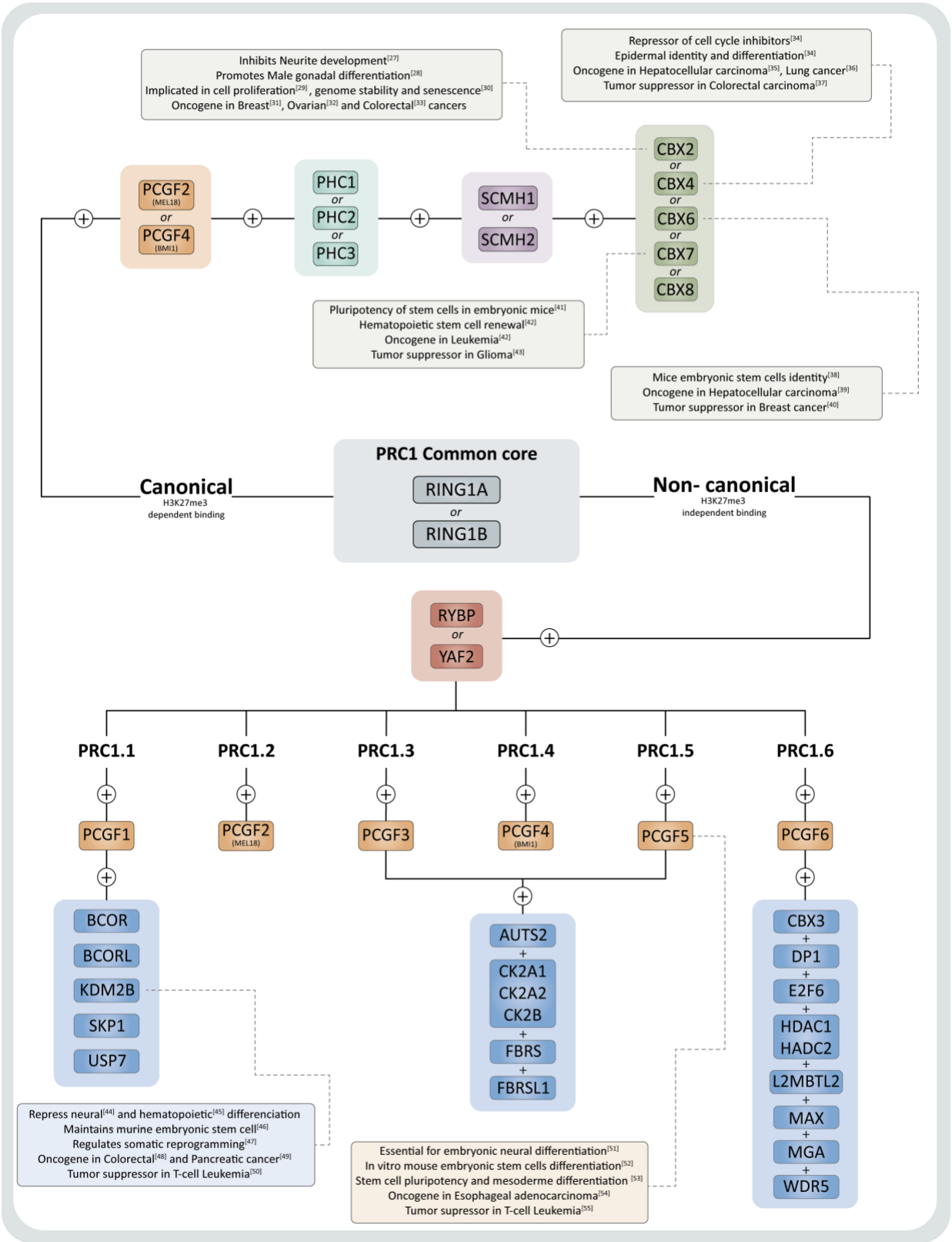
The Chromatin Organization and Remodeling

In eukaryotes, genetic information is stocked into the nucleus of the cells and is organized in a structure called **chromatin**, which permits in human, the compaction of ~2meters of DNA into the cell nucleus less than ~10µm across [20]. Chromatin has different statuses representing the degree of compaction. Firstly, the heterochromatin is strongly compacted and contains non-coding gene regions and unexpressed genes. Secondly, the euchromatin is lightly compacted and contains expressed genes. Chromatin organization is a very dynamic structure that changes according to the cell type or the developmental stage [21].

The Nucleosome is the fundamental unit of chromatin repeating every 160-240 bp along the genome. It is composed of an octamer of histone proteins wrapped in 145-147 bp of DNA. Each octamer comprises heterodimers of H2A-H2B and H3-H4 (Figure 3) [16].

Histone proteins are subject to epigenetic **Post-Translational Modifications** (PTMs), on their C- and N-terminal tails such as methylation, acetylation, and ubiquitination. These modifications can change the electronic charge and/or structure of histone tails, altering the chromatin status and playing an essential role in the chromatin dynamics and gene transcription. PTMs can be temporary or permanent, which allows defining three metaphorical protein categories. The "writer" adds the modifications whose action is opposite to the "eraser" which removes it; and the "reader" recognizes the PTMs and influences gene expression. Through the action of their "writers", "erasers" and "readers", PTMs play different roles in cell fate, cancers, and DNA damage [22].

The histone methylation is one of the major PTMs. The large group of histone methyltransferases (HMTs) is considered as "writer", adding methyl group(s) on lysine or arginine residues. Most of them are added on the lysine (K) of histone H3 and H4, which can be mono-, di-, or trimethylated (me1, me2, me3 respectively). Depending on the methylated lysine residue, they can be considered as gene-activating marks like H3K4, H3K36, or gene-repressing marks such as H3K9, H3K27 [23], [24].



Adapted from [7]

Figure 4. Polycomb Repressive Complex 1 (PRC1) Variants and Key Components

Schematic representation of PRC1 protein composition. The plus sign represents the protein or group of protein necessarily present in complexes. Color boxes represent proteins or protein family complexes. The frames highlight the main functions of some key proteins.

The histone acetylation is the addition of an acetyl group on many lysine histones tails by the "writers" Histone Acetyltransferases (HATs), as is opposed to the "eraser" role of Histone Deacetylases (HDACs). The addition of negatively charged acetyl group changes the histone charge. It decreases the interaction between the normally positively charged histones with the negatively charged DNA, resulting in chromatin opening. This PTM is often viewed as a gene-activating mark [23], [24].

The histone ubiquitination is a very large PTM affecting the histone mass added on lysine residues. Each ubiquitin unit is a polypeptide composed of 76 amino acids, added through the sequential action of "writers" group with different but complementary enzyme activity: E1 activating, E2 conjugating and E3 ligase. This PTM mainly occurs on histone H2A at lysine 118/119 and on histone H2B at lysine 120, into a monoubiquitylated reversible state [22]. H2A ubiquitination is generally considered as a repressive gene expression mark whereas H2B ubiquitination can be either considered as an activate or repressive-gene expression mark.

Through the expression of their "writers", "erasers", and "readers", PTMs regulate cell fate, development, homeostasis, DNA damage responses and pathologies such as cancers [22], [23], [25].

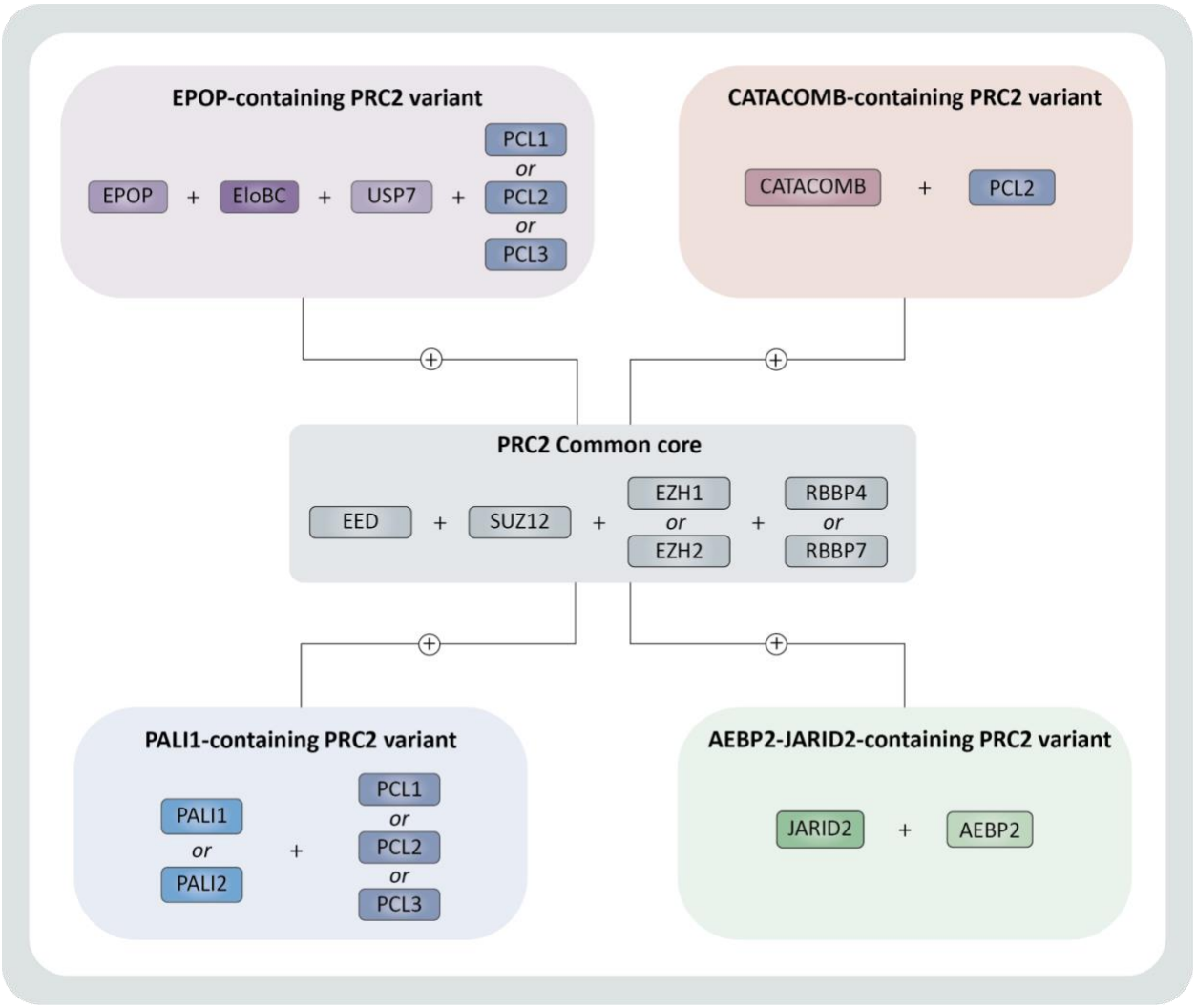
The Polycomb Group (PcG)

Initially identified in *Drosophila*, the **Polycomb Group** (PcG) is a chromatin-associated complex, generally associated with transcription silencing by epigenetic regulation. PcG is clustered in two major families relative to their functional properties: Polycomb Repressive Complex 1 (PRC1) and 2 (PRC2). The composition of these complexes varies according to the cell type [7], [26].

The mammalian **PRC1 family** contains E3 ligase activity, which predominantly mono-ubiquitinates H2A at lysine 119 (H2AK119ub) [7], [26]. PCR1 is divided into two subfamilies, canonical and non-canonical PCR1, both regrouping numerous complexes but sharing two core component proteins: a Really Interesting New Gene (RING1A or RING1B) and a Polycomb Group Ring Finger (PCGF/1-6) protein (Figure 4).

Besides the core component, **Canonical PCR1** (cPRC1) is composed of a chromobox protein CBX (CBX-2/4/6/7/8), a Scm (Sex Comb on Midleg Homolog) Polycomb Group Protein Homolog (SCMH-1/2) and a polyhomeotic homolog (PHC/1-3). CBX proteins are characterized by their "reader" ability to recognized and bind histone PTMs, precisely for H3K27me3 and H3K9me3 PTMs [26]. All these components have different functions. They act as oncogenes or tumor suppressors depending on the tissue. Studies showed that cPRC1 is more essential in somatic cells and in later developmental stages.

The Non-canonical PRC1 (ncPRC1) regroups heterogeneous complexes deviating from the paradigm of the canonical subfamily [7]. On the contrary to cPCR1, ncPCR1 is independent of PCR2 activity on H3K27me3. ncPCR1 possesses six variants named PCR1.1 to PCR1.6 depending on the PCGF proteins they contain. Additionally, each variant includes YY1 binding protein (RYBP) or its homologue YY1-associated factor2 (YAF2).



Adapted from [7].

Figure 5. Polycomb Repressive Complex 2 (PRC2) and Its Variants.

Schematic representation of PRC2 protein composition. Like in for the figure 4, the plus sign represents the protein or group of protein necessarily present in complexes. Colour boxes represent proteins or protein family complexes. The frames highlight the main functions of some key proteins.

PRC1.1 can bind to unmethylated CpG islands through KDM2B and maintains them in unmethylated state, repressing gene expression.

PCR1.2 and PCR1.4 can be found in both canonical and non-canonical PRC1 forms. They seem to be irresponsible of global H2AK119ub levels or transcription regulation in general in mouse embryonic stem-cells.

PRC1.3 and PRC1.5 can compensate each other's functions for H2AK119ub1 level. They can be found in the H3K4me3 region of actively transcribed genes lacking both H3K27me3 and H2AK119ub1, in chromatin forcing binding context. However, in non-forced context they promote H2AK119ub1 and H3K27me3, implying their role in H2AK119ub1 dissemination onto the genome.

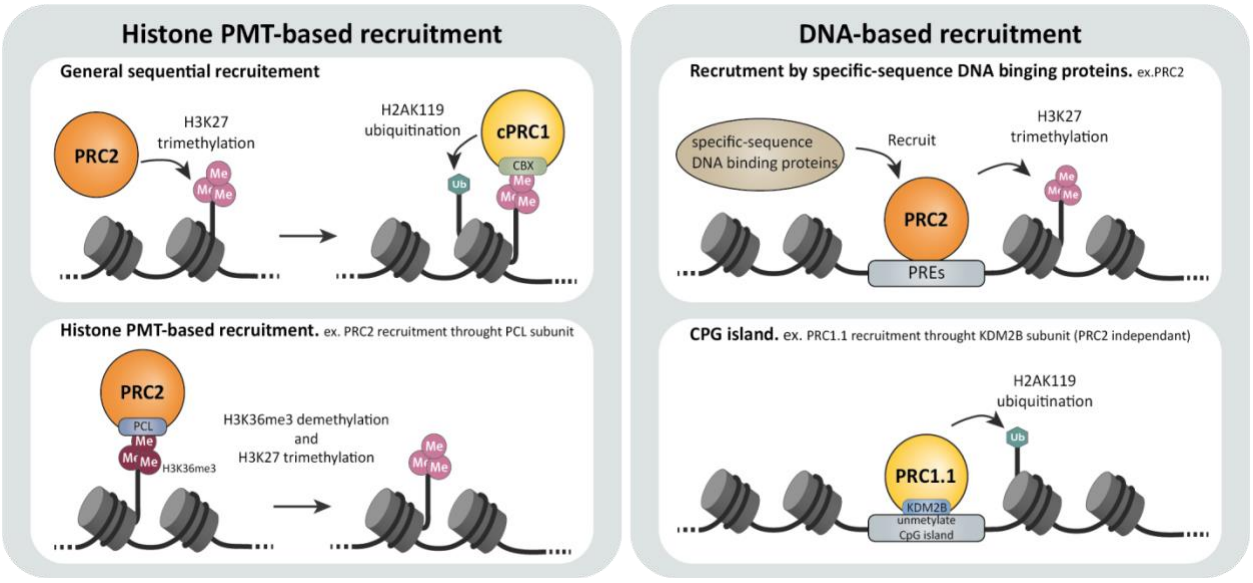
PRC1.6 is the most unconventional one and less described PCR1 variant. It is involved in cell identity presumably by a role in gene silencing.

On the other hand, **PRC2 complex** possesses methyltransferase activity. It catalyzes the mono- to tri- methylation of the histone H3 at lysine 27 (H3K27me1, H3K27me2 and H3K27me3). The core is composed of Suz12, RBBP4/7, Eed (enabling the recognition of trimethylated peptides), and Ezh1/2 that possesses a catalytic SET domain. Ezh1 and Ezh2 have complementary expression pattern. Whereas Ezh1 is largely expressed in adult tissues and non-dividing cells, Ezh2 is expressed in embryonic tissues and proliferative cells [26]. PCR2-Ezh2 is associated with H3K27me3 and gene repression, on the contrary to Ezh1 that has a minor methyltransferase activity and is present on H3K4me3-marked active chromatin. The Core components of PRC2 can partner heterogeneous number of proteins to form four different subunits. These subunits can interact with each other to regulate PRC2 enzymatic activity and recruitment (Figure 5). PCL (polycomb-like) proteins are found in many PRC2 variants. They contain a domain which helps PRC2 to bind to unmethylated CpG island and H3K36me3 *(in vitro)* [7].

The EPOP-containing PRC2 variant is defined by EPOP protein. EPOP seems to have a controversial role depending on the context. *In vitro*, EPOP interacts with H3K27me2 and increases the catalytic activity of PRC2, such as methyltransferases. On the contrary, the depletion of EPOP increases the level of H3K27me3 in mouse embryonic stem-cells [7].

The CATACOMB-containing PRC2 variant contains CATACOMB protein that defines its function. CATACOMB is an endogenous subunit whose expression is induced by DNA demethylation located on CpG island. This subunit interacts with PRC2 to counteract its function, inducing global reduction of H3K27m2 and H3K27me3 levels [7].

The PALI1-containing PRC2 variant contains PALI1 or PALI2. These submits bind to the methyltransferase G9A that methylates H3K9me2. PALI1 and PALI2 play a role in gene repression through PRC2. The depletion of both proteins induces a decrease of H3K27me3 levels in mouse embryonic stem-cells [7].



Created from [7].

Figure 6. Mechanisms of PcG Recruitment to Chromatin.

The figure represents some examples of PcG recruitment to the chromatin, through histone post-translational modifications (PTM) on the left part and through DNA on the right part.

The **top left** represents the proposed **general/canonical sequential recruitment**. PCR2 binds and trimethylated H3K27. H3K27me3 is recognized by the canonical PRC1 (cPRC1) which in turn ubiquitinylated H2AK119. H2A ubiquitination is generally considerate as a repressive-gene mark.

The **bottom left** illustrates the recruitment of PRC2 through its subunit PCL. PCL helps to bind and demethylate H3K36me3. Then, PRC2 trimethylated H3K27 which can be recognized by cPCR1.

On the **top right**, some DNA sequences, called specific polycomb response elements (PREs), are recognized by specific proteins which recruit the PRC2 on the site.

At the **bottom right**, unmethylated CpG island can be directly recognized by PRC1 and PRC2. For example, PCR1.1 recognized them by its KDM2B subunit

The AEBP2-JARID2-containing PRC2 variant contains AEBP2 and JARID2, both catalyzing PRC2 activity. AEBP2 interacts with RBBP4, which inhibits methylation of the activator mark H3K4me3. AEBP2 is found in two isoforms: a short one, associated with PRC2 transcription repression, mostly expressed in embryonic tissues; and a long one, associated with transcription activation, generally found in somatic tissues. JARID2 is important for core component PRC2 recruitment and for specific H3K27me3 deposition [7].

The **PcR proteins are recruited** at specific polycomb response elements (PREs) by specific-sequence DNA binding proteins. The general sequential mode of recruitment is the following: PRC2 is recruited to PREs, which stimulates H3K27 methylation at the PRE site, providing a binding site for PRC1. However, many new alternative mechanisms have been proposed (Figure 6) [7].

As described above, most of the PRC1 and PRC2 variants played a role in transcription repression. On one hand, transcription silencing can be done by blocking access of chromatin, compacting chromatin and/or inhibiting the transcriptional machinery. On the other hand, some variants promote active transcription by blocking PRC1 activity. PcR proteins play an important role in mammalian development and in somatic tissues. The depletion or overexpression of PcG protein induces transcriptional deregulation of the target gene and can lead to diseases such as cancers [7].

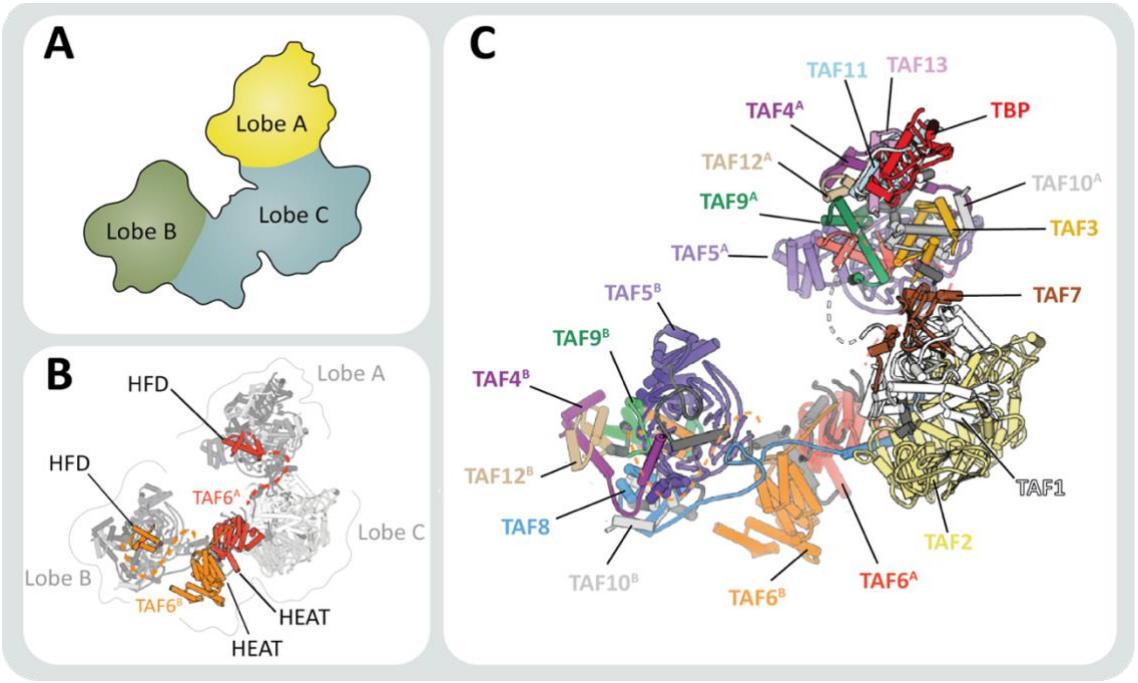
b. Transcription Factors

Transcription factors (TFs) are key proteins to control gene activation and repression. They usually recognize and bind to distinct sequences of 6-12 bp located in enhancers on the DNA. Following this binding, TFs may regulate GTFs recruitment and binding to the DNA, subsequently regulating transcription initiation. In humans, the number of TFs genes is estimated between 1,400 and 1,900, representing 7-9% of all protein-coding genes.

TFs are classified into two major functional groups. The first one is called constitutively active TFs. It contains ubiquitous TF subgroups involved in housekeeping gene transcription, as well as cell type-specific TF subgroups, whose expression is limited in time and space, contributing to determining cell type and differentiation. The second group, called signal-dependent TFs group, contains inactive TFs until the cell receives appropriate intra- or extracellular signals [56].

II. Focus on the General Transcription Factor TFIID

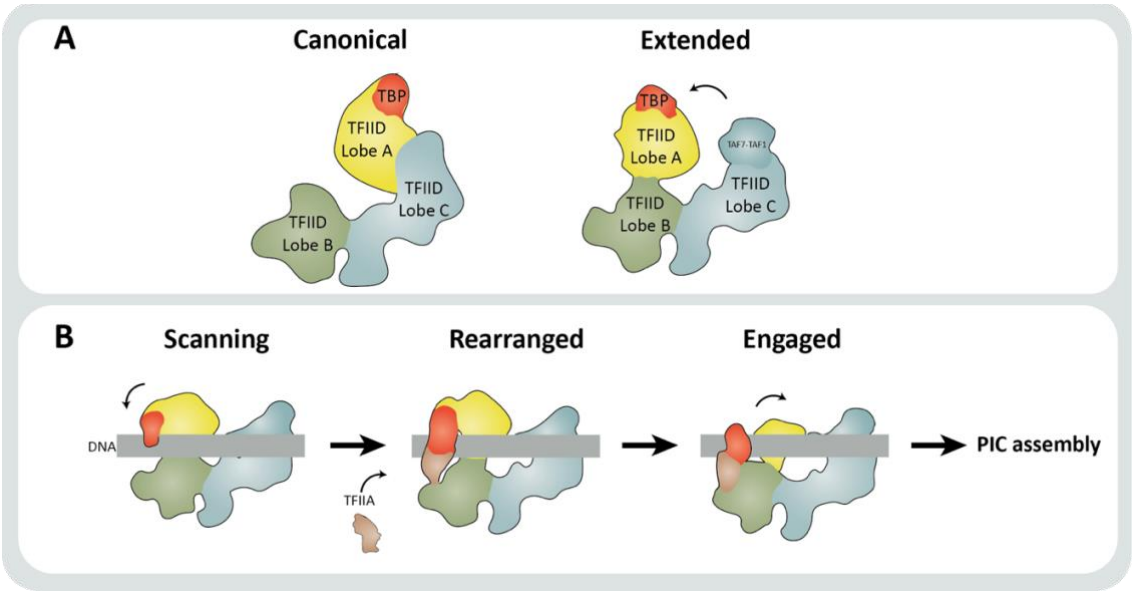
TFIID is a general transcription factor (GTF) composed of a large multiprotein complex (~1.3MDa in total). TFIID recognizes the core promoter sequences, closed chromatin marks and interacts with activators/repressors. Following the recognition and binding to the core promoter, TFIID has the huge task of recruiting and coordinating the pre-initiation complex (PIC). TFIID is essential for transcription initiation.



Inspirate and created with picture from [57]

Figure 7. Structural Organization of Human TFIID

A. Schematic representation of TFIID's three lobes. **B.** Transparent cryo-EM highlighting TAF6 link role between the lobes. The superscript letter represents the belonging lobe of TAFs protein present in two copies in the complex. **C.** Transparent cryo-EM mapping TAF proteins repartition in TFIID canonical form.



Created from [57]

Figure 8. Dynamic States of Human TFIID During Promoter Binding Process

A. In canonical form of TFIID, lobe A along with TBP are near lobe C, whereas is the extended form, they cross space to be near lobe B. The canonical form seems to be an inactive form in contrary to the extended form. **B.** The extended form changes of conformation during promoter binding process. During the scanning state, lobe A rotate toward lobe B, allowing TBP to scan DNA. DNA forms a bridge with TBP. In the rearranged state, TFIIA links TBP to lobe B but keeps a certain flexibility. The addition of TFIIA helps to stabilize TBP-DNA interaction. This stabilization leading to the called engaged state during which Lobe A moves slightly away from TBP following by PIC assembly. Whereas TFIID on the upstream DNA, is flexible and poorly attached in scanning and rearranged state, TFIID on the downstream DNA is strongly attach during all the process. The rigid connection between lobe B and C is important for TBP positioning on DNA.

A. TFIID Architecture

TFIID is a highly conserved complex of 20 polypeptides, composed of TBP and 13 different TBP-associated factors (TAF/1-13). Through their histone-fold domain (HFD), TAFs proteins are organized in heterodimers of TAF6-TAF9, TAF8-TAF10, and TAF4-TAF12. Cryo-transmission Electron Microscopy studies showed that human TFIID is organized into three lobes (A, B and C) (Figure 7.A) [57].

The lobe C has a central role related to its TAF6 dimer. TAF6 forms a homodimer thanks to its HEAT domain. Their N-terminal tails, which contain a HFD domain, get along lobe A and lobe B to form a heterodimer with TAF9, linking both lobe A and B to lobe C (**Figure 7.B**). TAF6 connection is maintained regardless of the conformational states of TFIID. As well, lobe C contains TAF1, -2, -7. It has been proposed that TAF1 associates with TAF2 and TAF7 to form a DNA binding complex, facilitating TFIID assembly [57].

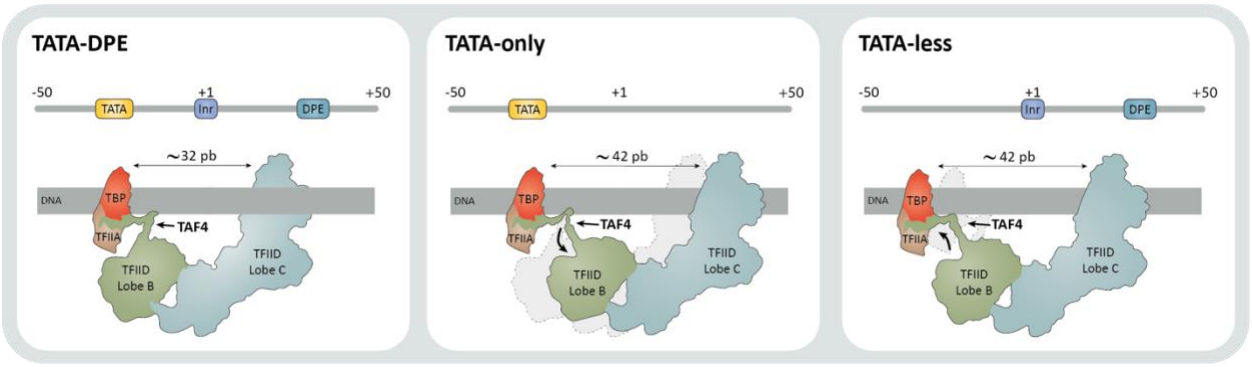
Lobe C left out, **TAF6-TAF9 heterodimer** interacts on one hand with TAF5, and with the heterodimer TAF4-TAF12 on the other hand to form a tetramer. Altogether, these five TAF subunits are linked in some way to each other and define the "**TAF subcomplex**", which is present in both lobe A and lobe B [57].

In lobe B, TAF8-TAF10 heterodimer interacts with the "TAF subcomplex" forming a hexamer-like structure. TAF8 seems to stabilize the connection between lobes B and C [57].

In lobe A, TAF subcomplex interacts with the heterodimer TAF3-TAF10 and TAF11-TAF13 to form an octamer-like structure. While in lobe B TAF10 is associated with TAF8, in lobe A, TAF3 replaces TAF8 in this partnership. The heterodimer TAF11-TAF13 is necessary for the connection between TAF1-TBP and the rest of TFIID (Figure 7.C) [57].

B. Dynamic Change of TFIID through Promoter Binding

TFIID structure plays a crucial role in promoter recognition and binding. Lobes B and C have a relatively rigid core certainly due to the presence of TAF8 in lobe B, on the contrary to the more flexible core of lobe A. TFIID has two distinct states referred to canonical and extended states (Figure 8.A). Only the extended state is found on the DNA [19], [58]. Lobe A is near lobe C in the canonical form whereas it moves and crosses over for lobe B in the extended state. TFIID-state dynamic change is the first step in PIC assembly (Figure 8.B).



Inspirate and created from [19]

Figure 9. Topology Changes of TFIID Rearrangement State According to Promoter Type

To simplify the comprehension, the TFIID Lobe A is not represented. The figure represents different TFIID rearrangements depending on the recognized promoter element. In this figure, the **TATA-DPE promoter type serves as reference**, in order to highlight the difference of topological rearrangments. The **grey shadows** represent the spatial topology of TFIID rearrangement on TATA-DPE promoter.

TFIID and TFIIA work together on promoter binding at multiple sites. Chen et *al., Science 2021* [19] proposed that rearranged TFIID can diverge into different types of tracks in function of the recognized sequence. Adding more flexibility, a TAF4 helix is brought out of lobe B to bind to DNA. This helix helps to coordinate the spatial structure change, connecting TFIIA-TBP to the upstream promoter part and TFIID to the downstream promoter part. Even if TBP plays a principal role in TATA-box binding, it is also implicated in TATA-less promoter binding (Figure 9).

TAFs also recognize core promoter elements to stabilize TFIID to the promoter. TAF1 can bind to Inr, DCE and MTE elements; TAF2 to the Inr and MTE element; and TAF6 and TAF9 to the DPE element. Their interactions are not limited to the core promoter. Some TAFs recognize histone post-translational modifications (PTMs), inducing interactions between TFIID and nucleosomes. TAF3 can bind to the transcriptional active mark H3K4me3 and TAF1 recognizes acetylated histone, often associated with transcriptional active mark. TAF3 and TAF1 through histone PTMs recognition, help to recruit TFIID on gene promoters [59].

C. Diversity and Specificity of TFIID

In addition to its spatial flexibility, many genes require a special structure and composition of TFIID for their transcriptional regulation. Many TFIID subunits possess related factors, paralog and/or variant(s), changing TFIID's composition and structure.

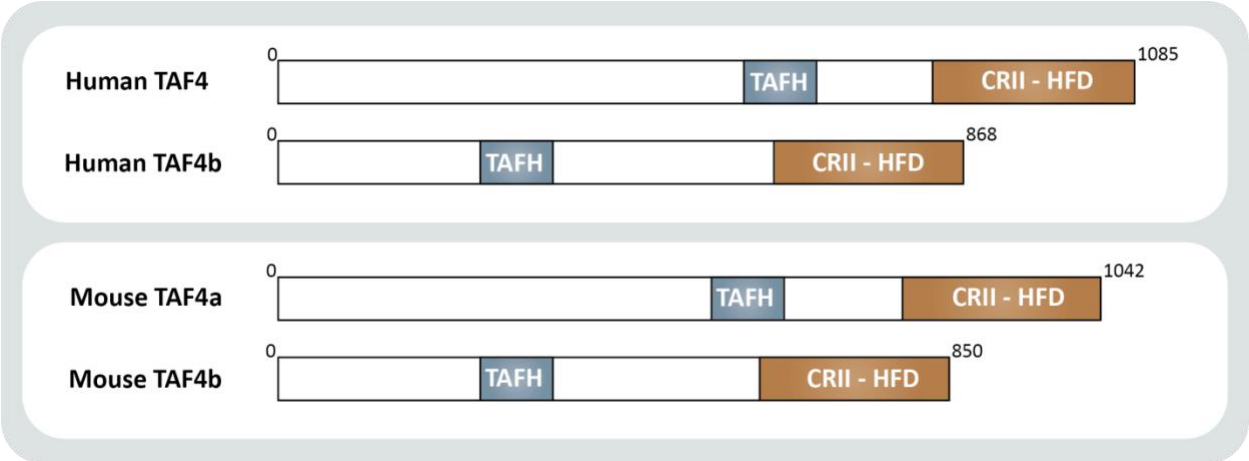
a. TBP and its Related Factors

TBP is not a universal protein, and it possesses many TBP-related factors (TRF). TRF1, TRF2, TRF3 and the recently identified TRF4, replace TBP in an alternative TFIID structure to promote selective transcription [12], [60], [61].

TRF1 is insect specific. It conserves the DNA-binding domain and its ability to interact with TFIIA and TFIIB. Biologically, TFR1 can replace TBP, and it is involved in gene regulation in the central nervous system and in male germ cells [61].

TRF2 is conserved in *Bilateria.* It is the most distant TBP paralog. Interestingly, on the contrary to TBP, TRF1 and TRF3, TRF2 does not have any ability to bind to DNA and it is helpless to recruit canonical TFIID complexes. Still, its interaction with TFIIA and TFIIB remains. Whereas TRF2 is crucial for development and cell differentiation in worms, flies, frogs, and fish, it is not imperative for mouse development but crucial for their fertility [61].

TFR3 is the closest paralog. It is conserved in vertebrates, and largely expressed in mammalian tissues. Unsurprisingly, DNA-binding and TFIIA/TFIIB interaction are conserved. TRF3 can compensate the loss of TBP, and it is implicated in specific gene transcription such as in early embryonic development in zebrafish, *Xenopus*, and in mouse hematopoiesis initiation [61].



Inspired by UniProt database

Figure 10. Protein Domains in TAF4 and TAF4b Amino Acids Sequence in human and mouse

TAFH, or Nervy homology region 1 (NHR1) is a highly conserved domain of 95-100 amino acids. It's an interacting platform with transcription factor regulators. **CRII region** for carboxy conserved region II, is a highly conserved C-terminal region. This region contains a histone fold Domain (HFD) essential for TAF4-TAF12 heterodimerization.

TFR4 is a newly described TBP-related protein. It is only found in *Drosophila,* suggesting a relevant function. It shares a distant TBP domain with TBP and TRF2, and it seems to be implicated in endoplasmic reticulum functions [61].

b. TAFs, Paralogs, Variants and Specific Gene Transcription

Many studies have shown that TAFs have a precise role in diverse specific gene transcription regulation. They can be sensitive to signals, and they interact with specific transcription factors [12]. For example, TAF10 interacts with the transcription factor GATA1 and controls mouse erythropoiesis [62]; TAF1 interacts with the AML1-ETO transcription factor to express specific genes in leukemia [63].

To add more variability and specificity, TAF4 and TAF6 have variants whose functions are still little known, and TAF4, TAF7 and TAF9 have paralogs that are expressed in different cells and tissues. TAF4b, the paralog of TAF4, is highly expressed in mouse gonads [64], [65], and embryonic stem-cells [66]. The paralog of TAF7, TAF7L, is involved in mouse adipogenesis [67] and male germ cell differentiation [68]. TAF9b regulates neuronal gene expression [69], stabilizes p53 and is essential for cell viability [70].

In defiance of its name, the "basal" transcriptional machinery possesses a high variable and specific structure, sensitives to signals. TAFs proteins in TFIID play crucial roles in specific transcription programs and viability.

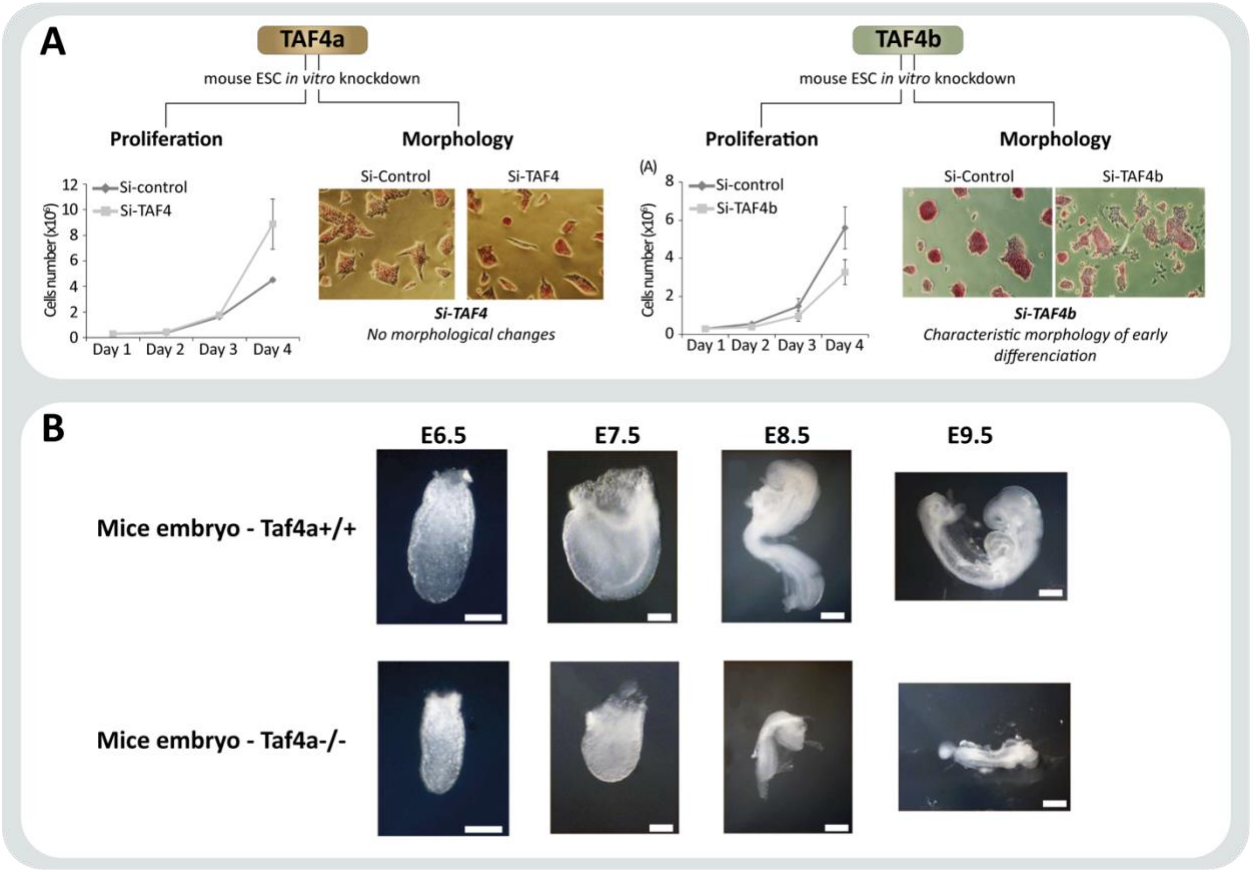
III. Spotlight on the TBP-Associated Factor TAF4

A. Structure of TAF4 Family Protein

The TAF4 family is composed of TAF4 (or TAF4a in mice) and its paralog TAF4b. TAF4 is one of the largest TAF proteins (110 kDa) and it is ubiquitously expressed. Studies in *Drosophila* tissue culture showed that TAF4 is critical for TFIID stability [71]. However, in mouse embryonic fibroblasts, TAF4a is not necessary for TFIID stability and can be replaced by TAF4b [5].

In humans, TAF4 and TAF4b are composed of 1084 and 868 amino acids (aa) respectively. They both possess a TAF homology (TAFH) domain and a CRII region (carboxy conserved region II) (Figure 10).

TAFH, also called Nervy homology region 1 (NHR1) or CRI, is a highly conserved domain of 95-100 amino acids, related to the transcription regulator MTG/EPO family. TAFH is an interacting platform for transcription regulators. TAFH is able to recognize common residue motifs shared by many transcription regulators [72]. Through the TAFH domain, TAF4 could be able to play a role of transcriptional activator or repressor in function of its interactants. Induced expression of truncated forms of TAF4 that contain the TAFH domain but not the CRII region promote cell growth. The TAFH domain seems to be implicated in promoting proliferation pathways [12].



Created from [66], [73]

Figure 11. TAF4 is essential for embryonic development

A. Mouse ESC were transfected *in vitro* with non-targeting Si-control or Si-RNA targeting Taf4a or Taf4b inducing short-term silencing of their proteins. After transfection, mouse ESC with Si-TAF4a have an increase proliferation compared to Si-control, contrary to those with Si-TAF4b. Pictures of transfected cell showed no morphological difference between Si-control and Si-TAF4a mouse ESC. However, those with Si-TAF4b loss their initial pluripotent-cell like morphology to the benefit of differentiated cell morphology at early stage. [58].

B. Pictures of normal or mutant depleted for TAF4a (TAF4a-/-) embryos at different stage of development. Mutants had a smaller overall size at E6.5 and more compacted appearance at E7.5 compared to normal embryos. From E8.5, mutants' overall size no longer seemed to increase anymore and present a shorten body axis. Their head and trunk were underdeveloped. Some developmental process continued to E9.5 like the primitive heart formation. Despite this, mutants were too severely impaired and consequently resorbed. Beyond this stage no mutants were detected indicating that they died quickly after E9.5 [69]. *scale bar: 100µm.*

ESC: embryonic stem cell

E: embryonic stage/days after coitus

Taf4a+/+ : embryo with wild-type Taf4a

Taf4a-/- : mutant embryo with depleted Taf4a

CRII region is the C-terminal part of TAF4 and TAF4b. This region is highly conserved through evolution in the TAF4 family. This region is essential for the interaction with TFIIA [74] and it contains a HFD domain. As described above, HFD is a histone-fold domain necessary for the dimerization between TAF4/TAF4b and TAF12. The expression of a truncated TAF4 with a CRII region and without a TAFH domain slows down cell's proliferation. On the opposite side of TAFH domain, the CRII region seems to be implicated in repressing proliferation pathways [12].

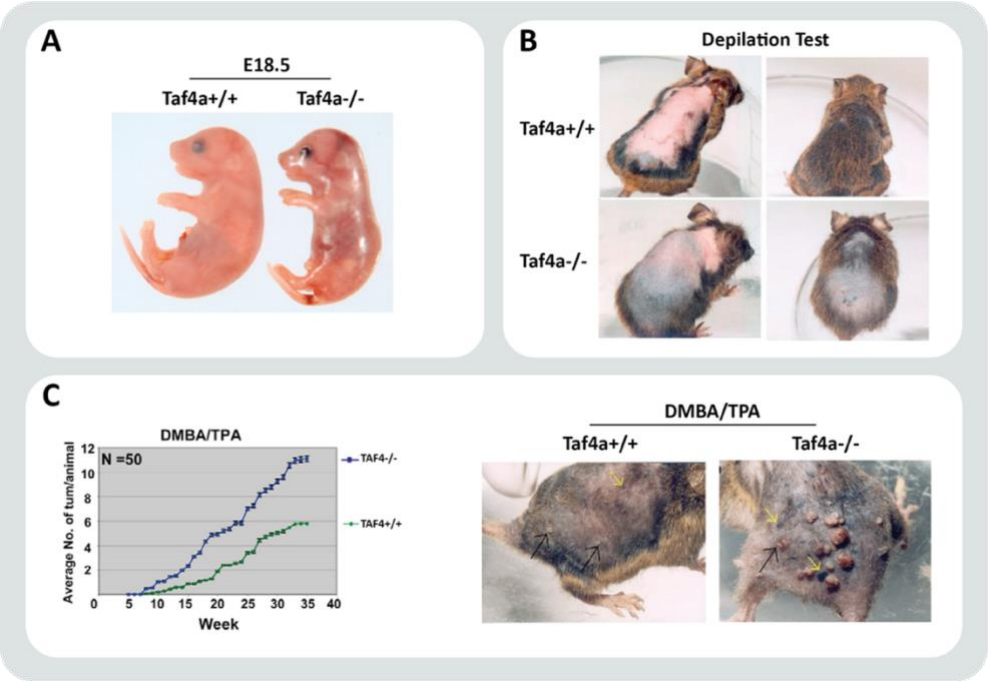
B. TAF4 is Essential for Embryonic Development

Using an *in vitro* model of **mouse embryonic stem-cell (ESC) and RNA interference (SiRNA)** method, *Bahat et al. 2013* [66] reported that TAF4a and TAF4b play a different transcriptional regulation role. TAF4b cooperates with the DNA-binding transcription factor Oct4 implicated in embryonic stem-cell pluripotency network. This interaction permits to regulate genes, promoting the pluripotency, proliferation, and cell progression of ESC, whereas TAF4a inhibits them and does not interact with Oct2. In this model, the results showed that TAF4 and TAF4b exert opposite effects. TAF4a maintains the differential potential of mouse ESC, instead of TAF4b, which maintains their pluripotency (Figure 11.A).

On the other hand, using an *in vitro* model of **mouse ESC** as well, **but a short hairpin RNA (shRNA)** method, *Pijnappel et al. 2013* [75], reported that both TAF4a and TAF4b are essential for maintaining pluripotency of ESC. TAF4a is also part of the pluripotency network as the binding site for Oct4 is found ~10-Kb upstream of the mouse *Taf4a* gene and the knockdown of Oct4 reduced *Taf4a* expression.

These two works display a contradictory role of TAF4 and TAF4b in the pluripotency maintenance of ESC. This difference may be due to the difference in the method used and the efficiency for TAF4/TAF4b knockdown.

Studies done by *Langer et al., 2016* [73], highlighted the role of **Taf4a** during ***in vivo* embryogenesis,** using a mouse model. While the embryos lacking TBP [76], TAF7 [77] or TAF10 [78] showed an early death at 4.5-6.5 days post coitus (E4.5-E6.5), embryos lacking *Taf4a* (*Taf4a-/-*) showed a lethality later at E9.5. As Taf4a and Taf4b expression overlaps in the early stage, the longer survival may be explained by the redundancy of Taf4b, which can maintain TFIID integrity and early-stage functions. However, *Taf4a-/-* embryos showed many defects and atrophies (Figure 11.B). *Taf4b*-lacking mice (*Taf4b-/-*) are viable and showed defects only in adults. Grouped together, these results indicated that while Taf4a can fully replace Taf4b in embryogenesis, Taf4b cannot replace Taf4a in later embryogenesis stages. *Taf4a* and *Taf4b* are also dispensable during the early stages of embryogenesis. ***Taf4a-/-* embryos** showed an increase of cell death apoptosis, no foregut invagination and gut tube formation, misplacement of heart and head. Rather than being implicated in ESC pluripotency, TAF4 seems to act later in the general course of cell differentiation and specialization.



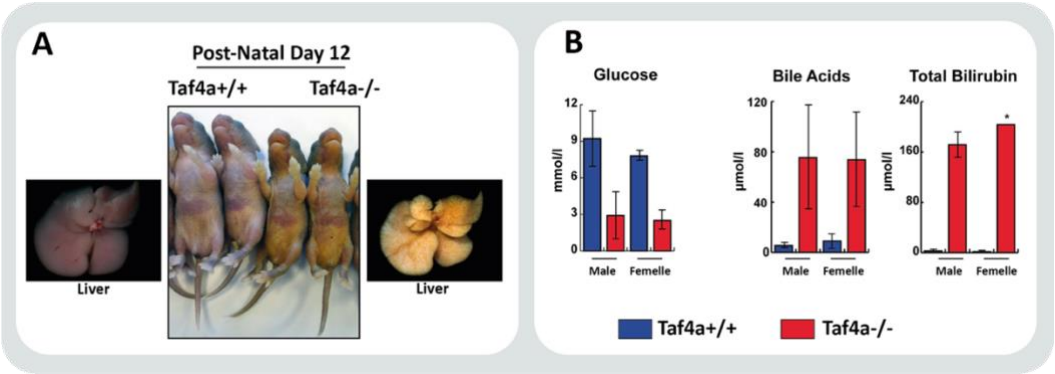
Created from [79]

Figure 12. TAF4 in Skin Development, Adult Hair cycle and skin tumorigenesis

Taf4a was knock-out specifically in basal keratinocytes using cre-lox system under K14 promoter.

A. Gross morphology of E18.5 fetus in the constitutive model. The skin of mutant is translucid. Their skin barrier is defective causing early post-natal death. **B.** 6 weeks after inducing Taf4a depletion in the conditional model, mice were depilated to test hair cycle ability. Pictures were taken on day 0 (on the left) and day 20 (on the right) post-depilation. Mutant furs do not return to their original state. They have an impair hair cycle. **C.** 3 weeks after Taf4a depletion, mice were treated with combination of mutagens DMTA and TPA to induce carcinogenesis. The left picture represents the evolution of tumor number over time. On the left, pictures of mice were taken 20 weeks after treatment. Black arrows, epidermal tumors; yellow arrows, melanocytic growth. Mutants presents more and bigger tumors.

Taf4a+/+ : fetus/mice with wild-type Taf4a **Taf4a-/- :** mutant fetus/mice with depleted Taf4a



` Created from [4]

Figure 13. TAF4 is Essential During Post-Natal Hepatocyte Differentiation

Taf4a was conditionally knock-out specifically in post-natal hepatocytes using the cre-lox system under albumin promoter.

A. Gross pictures of mice and their liver at post-natal day 12. Compared to normal mice (on the left), mutant mice (on the right) show a retarded growth, reduced body weight, have severe jaundice, and their liver were yellow with bile deposit. Mutants had a severely altered liver function. **B.** Blood plasmas were analyzed at day 10. Glucose levels revealed a severe hypoglycemia in mutant, probably a cause of the growth impair and early death. The increase of plasma bile acids and bilirubin confirmed altered liver function and corelated with the jaundice.

Taf4a+/+ : mice with wild-type Taf4a **Taf4a-/- :** mutant mice with depleted Taf4a

C. TAF4 Plays a Role in Tissue Specialization and Homeostasis

a. TAF4 in Skin Development, Hair Cycle and Tumorigenesis

The specific conditional inactivation of *Taf4a* in basal keratinocytes during **mouse development** causes early post-natal death [79]. Fetus analyzed at E18.5 just before the birth presents a shiny transparent skin, thinner altered epidermis, and defective skin barrier function (Figure 12.A). The defective barrier is represented by down expression of essential junction proteins inducing water loss and skin permeability.

On the contrary, the controlled induced loss of *Taf4a* in basal keratinocytes in **adult mouse** models does not impair their survival but leads to fur loss. Mice are unable to regrow back the fur efficiently; the anagen phase of hair cycle is severely impaired (Figure 12.B). Histological analysis of *Taf4a*-deleted mouse epidermis revealed hyperplasia, meaning that *Taf4a* is implicated in control of keratinocyte proliferation, stimulating the MAP kinase (Mitogen-Activated Protein Kinase) pathway, and activating the EGF (Epidermal Growth Factor) signaling. Adult mice present more than 500 down-regulated genes and ~350 up-regulated genes. Some members of the mitogen EGF family promoting keratinocyte proliferation are up-regulated, as well as members of the MAP-kinase pathway. In the down-regulated side, several MAP-kinase inhibitors are found.

Additionally, the loss of *Taf4a* induces expression of markers associated with oncogenic transformation, yet mice do not develop spontaneous **skin tumors**. However, in absence of *Taf4a*, mice present less resistance to tumorigenesis. When they are exposed to the chemicals inducing carcinogenesis, mice present a higher number and a larger size of melanocyte tumors (Figure 12.C).

In skin context, TAF4a is a regulator of keratinocyte proliferation, which is essential for fetus skin development, hair cycle and it possesses a tumor suppressor activity.

b. TAF4 Regulates Gene Hepatocyte Expression in Association with HFN4A

Early post-natal *Taf4a* inactivation specifically induced in hepatocytes leads to death of the mice by day 15 [4]. Their livers showed defective organization and function, defective bile duct formation, disorganized hepatocyte epithelium, cell junction loss, and consequently metabolic abnormalities (Figure 13). All these observations indicate a delay in hepatocyte maturation. The inactivation reveals that the liver presents a deficit of post-natal gene activation. Among them, 296 genes are regulated by the nuclear transcription factor Hepatocyte Nuclear Factor 4 Alpha **(HNF4A),** especially affecting genes implicated in metabolic functions. HNF4A is a transcription factor critical for hepatocyte gene expression. In the case of some genes activated by HNF4A before the birth and inactivation, the loss of *Taf4a* brings a diminution of the PIC formation, without necessarily impacting the occupancy of HNF4A in the regulatory elements of lost genes.

In hepatocytes, TAF4a is necessary to stabilize HNF4A on regulatory elements and the following PIC formation, suggesting that TFIID and HNF4A link through the proven interaction of HNF4A and TAF4a, to activate genes implicated in hepatocyte function.

c. TAF4 is Essential for Pancreatic Beta-Cells Identity and Insulin Pathways

Specific inactivation of *Taf4a* in adult murine beta-cells helps to define its specific role in the insulin pathway [80]. Expression of beta-cell genes' functions is impacted, leading to an increase of glycemia, lower plasma insulin levels and defective insulin maturation and secretion in response to glucose. After a few weeks, the phenotype stabilizes, thus permitting mouse survival. However, after 10 weeks, a few new beta-cells expressing TAF4a are detected in the pancreatic Langerhans islets. The origin of these new TAF4a-positive beta-cells may be the ones escaping from knockout but mostly they can come from trans-differentiated cells. Beta-cells have an altered chromatin accessibility and a deregulation of several transcription factors binding to beta-cell genes.

Analyzes of pancreatic Langerhans islet with single-cell RNA-seq method allow to determine that *Taf4a* inactivation leads to a stronger endoplasmic reticulum stress in beta-cells, alpha-cells, and delta-cells, suggesting a defective communication between beta-cells and their neighbours. Also, some beta-cells seem to quickly lose the expression of their specific identity genes and start to express both alpha and/or beta-cell identity genes. This result suggests that the loss of *Taf4a* in beta-cells induces cell reprogramming and trans-differentiation into alpha-cells and delta-cells.

In beta-cells, TAF4a is essential for chromatin accessibility in terms of cell identity and insulin pathway.

D. TAF4 and its Variants, a Balance Story

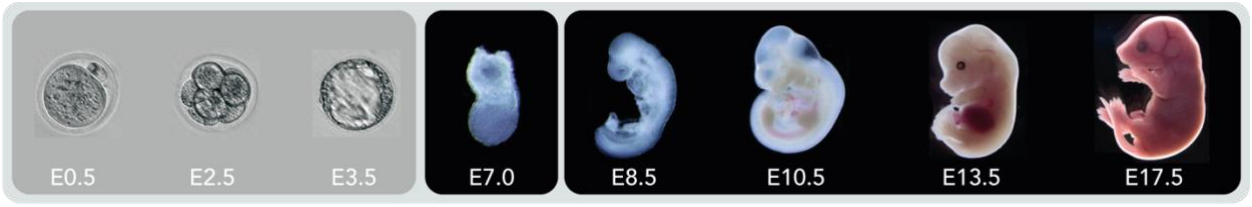
Alternative splicing plays a key role in organism complexity by generating multiple variants from one single gene. Currently, it has been estimated that there are between 80,000 and 400,000 proteins for 20,000 protein-coding genes in humans. In mice, five variants of Taf4 have been identified, comprising a full-length version and TAFH domain-truncated variants [12].

Even if Taf4 variants are not well known yet, some observations have been made. The expression pattern of the full length Taf4 and its variants can differ. While the full length Taf4 and some of its variants are invariably expressed in tissues, other variants seem to be expressed in a cell-specific manner. The full length Taf4 is mostly observed in stem-cells of tissues unlike its variants that seem to be more observed in differentiated cells. As TAF4 protein isoforms generated from alternatives variants have alternative protein structures, they can have a huge impact in TFIID integrity and functionality. In humans, the expression of the full length Taf4 is essential for neural progenitor pluripotency but dispensable for differentiated cells where alternative variants with impaired TAFH domain stimulate neuronal differentiation [81]. As well, in *in vitro* human mesenchymal stem-cells, the expression of TAFH domain-truncated variants represses proliferation and promotes differentiation into chondrocytes [82].

Taken all together, TAF4 and its variants seem to play an essential role in the balance between pluripotency and differentiation state of the cells. They add more complexity and adaptability to TFIID complexes in function of tissue and cell contexts.

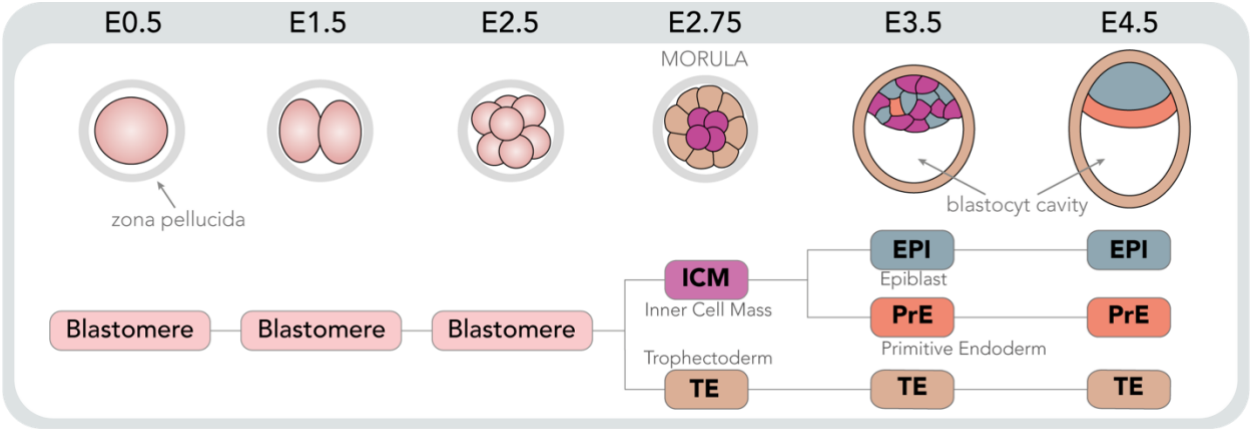
CHAPTER 2

Tales of the Intestinal Development



Pictures picked from [embryology.med.unsw.edu.au & hhmi.org]

Figure 14. **Gross Morphology of Mouse Embryo Through Development**



Inspirated and adapted from [83]–[85]

Figure 15. **Early Development of Pre-Implantation Mouse Embryo**

Schematic drawing of morphological and lineage fate of pre-implantation mouse embryo. Blastomeres are divided by cleavage doubling cell number while embryo keeps its size. The Embryo is encapsuled by zona pellucida, a glycoprotein barrier limiting embryo expansion without disturbing its cells movements. At E2.5, blastomere cells polarized and specialized into ICM or TE cells in a stage called Morula. Over time, the embryo takes a new form called blastocyst and loses its zona pellucida. ICM cells specialize in EPI or PrE while a blastocyst cavity appears and grows. Embryo implantation occurs around E4.5.

The development of an organism from a one-cell embryo engages a myriad of complex timed stages, involving cell proliferation, maturation, migration, tissue morphogenesis and growth. Beyond the pleasure of elucidating life complexity, understanding and decoding embryogenic events is of great help for other fields. For example, studying transcription factor functions and consequences of their perturbation contributes to cancer and regenerative medicine. Mouse embryonic development is a canonical model for studying mammalian embryonic cell fates and properties. Mice share with human similar genome organization, genetic regulation and 99% of murine genes have human homologs. As they often harbor the same functions and mutation-induced phenotypes, mouse models are largely used for human comprehension. In this chapter, we will begin the saga of the intestine through the history of murine embryonic development.

I. From one cell to a complete mouse, an adventure full of stages

The incredible generation of a full mouse organism is governed by many actors **(Figure 14).** It is a complex dynamic story of proliferation, cell fate decision, signaling pathway crosstalk and transcription factor expression, suppression, down regulation, or re-expression.

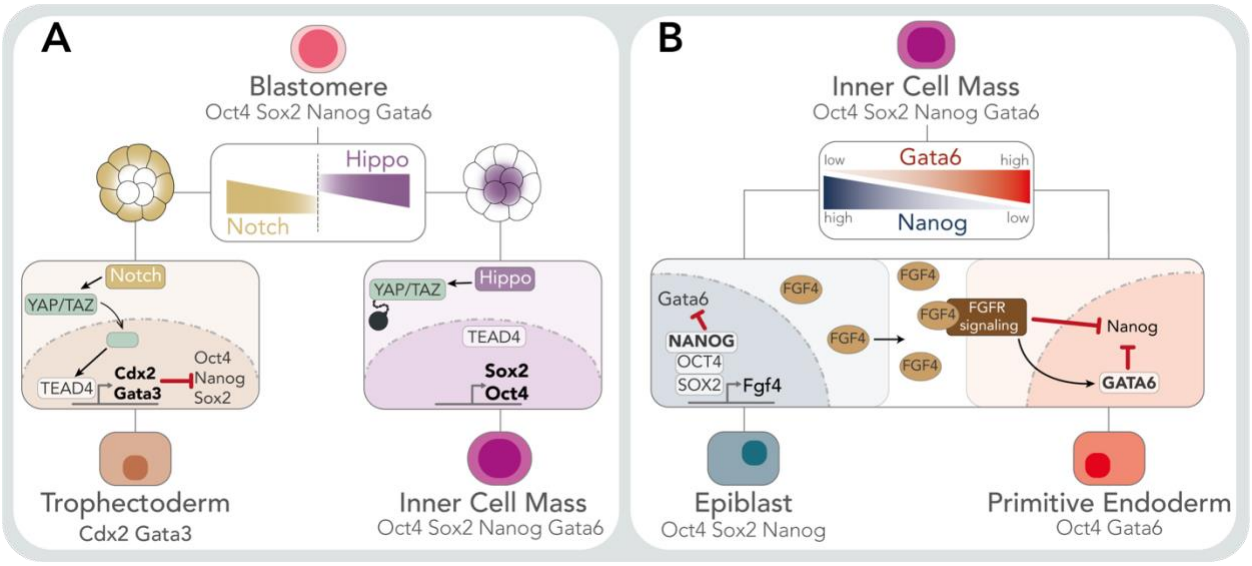
A. Forming a Full-Grown Organism, a Quest with Many Steps

In mice, the stage of embryonic development is timed from the mating day. As the exact moment of fertilization cannot be determined and the post-coitum plug is only observed half-day later, the timing starts at the stage of 0.5 days. Embryonic stages are described, characterized, and referenced in respect of the number of days post-coitum (dpc) also called embryonic day (E) [86].

a. Fertilization and Implantation, the Beginning of the Story

Morphological Changes and Cellular Specialization

The story begins with the **one-cell mouse embryo** or zygote, after the fertilization of the oocyte. Following fertilization, embryo cells are subjected to a succession of cleavage divisions that each time double the number of cells (Figure 15). From the first cleavage, the zygote generates blastomere cells. In the early stages, each cleavage of blastomere cells increases contacts and compaction between them. At E2.5, the compaction induced by the cleavage of the 8-cell blastomere into the 16-cell one signs the beginning of the **morula stage**. At this stage, blastomere cells acquire specialization to generate two different cell types, the inner cell mass (ICM), composed of undifferentiated cells in the embryo center, and the outside cells or trophectoderm (TE), which are essential for the implantation of the embryo into the uterine mucosa [87].



Inspirated and adapted from [83], [88], [89]

Figure 16. The Game of Pathways in the Preimplantation Embryo

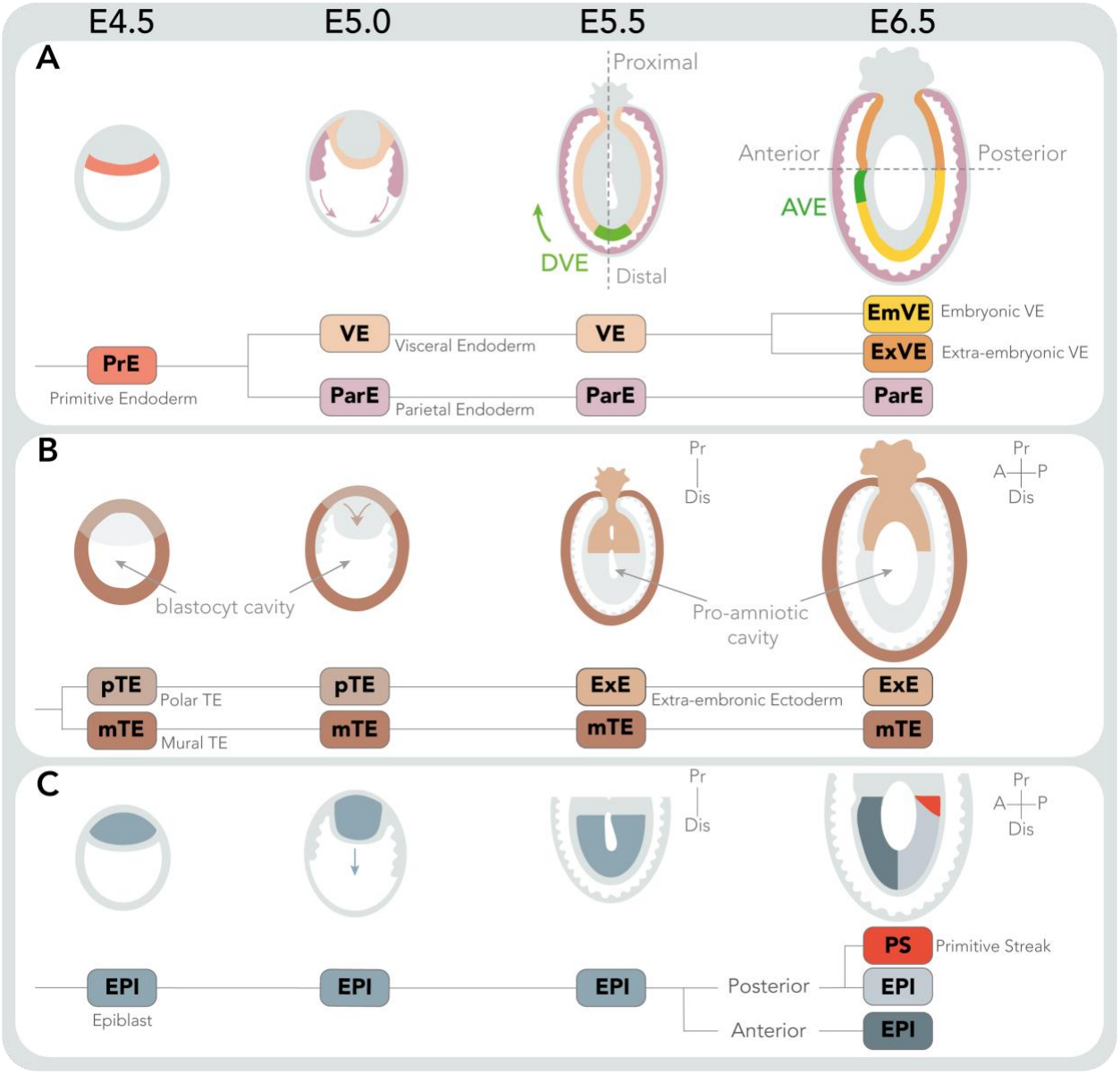
A. Outside blastomere, Hippo pathway is suppressed, and Notch is activated enhancing YAP/TAZ activity. YAP/TAZ effectors translocate into the nucleus, interact with TEAD4, inducing expression of CDX2 and GATA3 that inhibit Oct4, Nanog and Sox2 expression, promoting trophectoderm differentiation. Meanwhile, inside the blastocyst, Hippo is activated, sequestering YAP/TAP in cytoplasm, breaking TEAD4-expressed genes. Consequently, transcription repression of SOX2 and OCT4 is pulling off. **B.** Later, IMC cells chose to express higher quantity of NANOG, which represses Gata6 expression, inducing Fgf4 expression and specializes cells into Epiblast. FGF4 protein secreted by epiblast is captured by receptor of FGF signalling pathways by neighbouring cells. Activation of FGF pathways promotes GATA6 expression. Both FGF pathways and GAT6 repress Nanog expression, promoting Primitive Endoderm specialization.

Around E3.0, the embryo enters the blastocyst **stage** and settles into the uterus. The TE layer interacts with the uterine epithelium, mediating the implantation that occurs around E4.5. Then, the blastocyst cavity appears and expands. During this stage, ICM begins a new specialization into epiblast (EPI) which will form the whole mouse body and the primitive endoderm (PrE), an extra-embryonic
lineage. Within blastocyst progression, EPI cells are spatially separated from extra-embryonic lineages as PrE migrates toward the blastocyst cavity. By the end of the stage at E4.5, blastocyst is composed of the trophectoderm (TE) and the primitive endoderm (PrE) that envelop epiblast (EPI), the only embryonic tissue [87]. The implantation defines the beginning of morphogenetic and cell identity changes.

Mechanism and signaling pathways involved in implantation

Behind these morphological changes and cell fates, **many actors** are involved, implicating a close collaboration between transcription factors and cell communication thought signaling pathways (Figure 16). The first decision of cell lineage fate is related to differential activation and crosstalk of Hippo, YAP/TAZ and Notch signaling pathways. Lastly, Cdx2 and Gata3 expression is restricted in outside cells, while Sox2 and Oct4 are restricted to inside cells, giving the rise of TE and ICM lineage respectively [88], [90], [91]. During the second lineage specialization, ICM cell fate is mainly driven by the earliest downregulation of two transcription factors, Nanog and Gata6, in combination with the Fgf4, ligand of the fibroblast growth factor signaling. ICM cells choosing to express *Nanog* down-regulate *Gata6*, leading to the EPI cell fate. In contrast, ICM cells choosing to express *Gata6* down-regulate *Nanog* and are specialized into PrE cells. The duality of Nanog and Gata6 expression gives a "salt and pepper" pattern [88], [90], [91].

To simplify, signaling pathways play a key role in tissue/organ development and homeostasis. They are implicated in communication between cells and their environment, through the binding of signaling ligands to their specific receptors. In response, they induce specific cascade reactions to modulate gene transcription and activity.



Inspirated and adapted from [84], [85], [91]

Figure 17. Mouse Embryo Transformation into Egg Cylinder Form

Schematic drawing of morphological and lineage changes in post-implantation mouse embryo. Over time, the embryo grows and trades its blastocyst form for in profile of the "egg cylinder" hollow cup-shaped form.

A. Fate of Primitive endoderm (PrE).

B. trophectoderm (TE) fate is subdivided into polar trophectoderm and mural trophectoderm.

C. Epiblast fate.

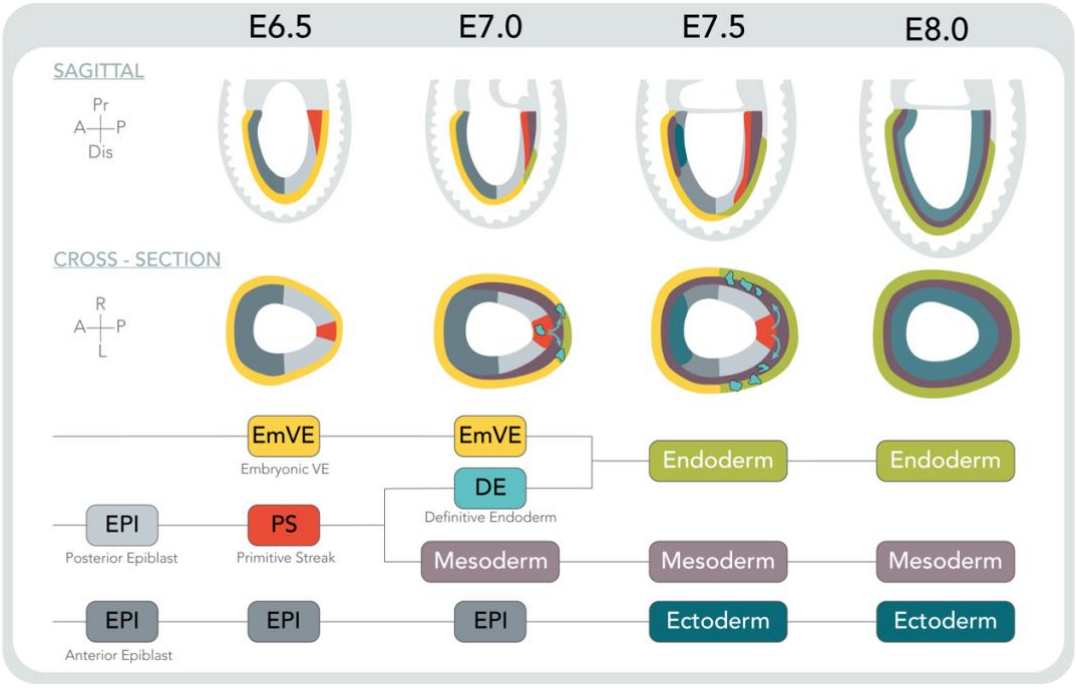
DVE, distal visceral endoderm**. AVE,** anterior visceral endoderm.

A, anterior; **P**, posterior; **Dis**, distal; **Pr**, proximal.

b. The Gastrulation or the Difficult Task to Define the Body Axis while Preparing for Organogenesis

Morphological Changes and Cellular Specializations

Following implantation, the embryo quickly grows up in 24 hours, drastically changes its shape to adopt a cylindrical form referred to as the "egg cylinder" and initiates **antero-posterior (AP) patterning** (Figure 17). To reach these changes by E5.5, embryo cells will be subject to many stages of maturation and migration. Part of PrE cells migrates along the mural TE layer and processes differentiation into parietal endoderm (ParE) cells. Meanwhile, PrE cells between EPI cells and the blastocyst cavity mature into visceral endoderm (VE). Polar TE cells that enclose the blastocyst cavity mature to develop the extra-embryonic ectoderm (ExE). To go further into the changes, EPI expands towards the blastocyst cavity and takes the form of cylindrical epithelium. ExE cells follow EPI cells, entering the cavity, whereas the VE layer is pushed down and expands to enclose EPI and ExE cells. By E5.5, some distal VE cells specify into distal visceral endoderm (DVE) that migrates and matures towards the EPI-ExE junction, forming the anterior visceral endoderm (AVE). At E6.5, the migration is completed, and the position of AVE cells determines the embryonic anterior pole. At this time, the new differentiated VE enveloping the ExE cells mature again to form the extra-embryonic VE (exVE) and the VE enveloping the EPI matures into the embryonic VE (emVE). Subsequently, at the opposite of the anterior site, posterior EPI (p-EPI) cells start to form the **primitive streak** (PS), defining the embryonic posterior pole. The PS is a transient structure arising from EPI cells that have undergone a series of events and epithelial-to-mesenchymal transition (EMT). It is a dynamic structure that proliferates and grows through gastrulation until it is resolved [89].

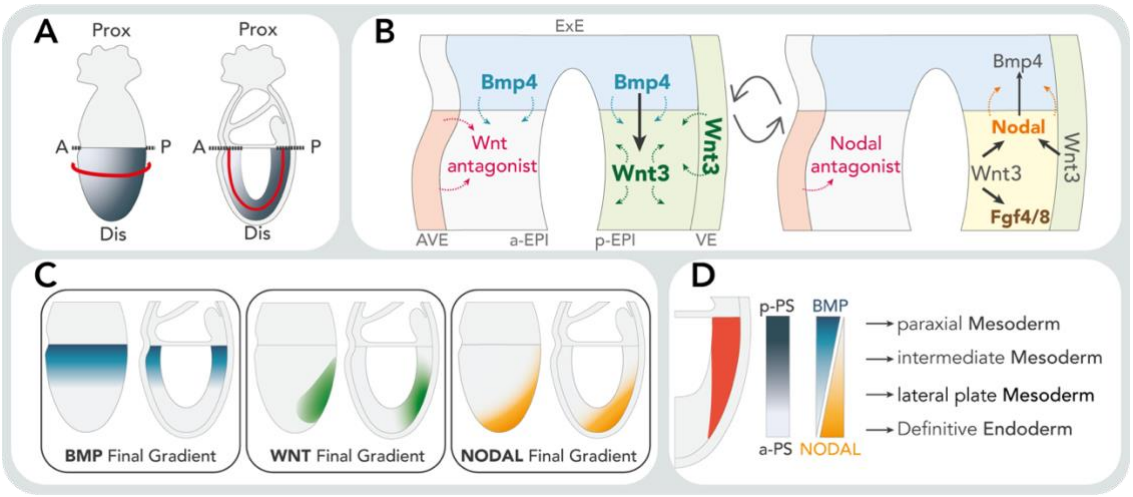


Inspirated and adapted from [84], [85], [92]

Figure 18. Gastrulation and Germ-Layers in Mouse Development

Schematic drawing of morphological changes and lineages in mouse gastrulation. Sagittal and cross-section illustrate germ-layers general progression through mouse gastrulation. The mesoderm appears from the primitive streak (PS) on posterior side, spreads between epiblast (EPI) and embryonic visceral (EmVE) layer, progressing to anterior side. Later, definitive endoderm (DE) cells derived from PS, crosses mesoderm germ-layer and inserts between EmVE cells forming the endoderm. In parallel, anterior epiblast cells specialize and form the ectoderm. By the end of gastrulation, the embryo has the form of a 3-layers cup.

Axis labels: **A**, anterior; **P**, posterior; **Dis**, distal; **Pr**, proximal; **L**, left; **R**, right.



Inspirated from [84], [93], [94]

Figure 19. Signaling Pathway Landscape, Gastrulation and Primitive Streak Fate

A. Simplified representation of anterior and posterior embryo gradient orientation in epiblast mass, through external view (left) and sagittal views (right). **B.** Two-part simplified representation of signaling crosstalk during gastrulation. Dotted arrows represent signaling proteins production, secretion, and diffusion. Solid arrows represent the stimulation of signaling proteins targets. **C.** Schematic representation of signaling pathways landscape in EPI mass, by the end of gastrulation. Signaling landscape is defined through the conjoint action of source cells, antagonist secretions and EPI cells speciation. **D.** Schematic representation of BMP and NODAL gradient concentrations in primitive streak, and their impact on primitive streak fate.

Dis, Distal; **Prox**, Proximal; **A**, Anterior; **P**, Posterior; **VE**, Visceral Endoderm; **ExE**, Extraembryonic Ectoderm; **AVE**, Anterior Visceral Endoderm; **a-EPI**, anterior Epiblast; **p-EPI**, posterior Epiblast; **a-PS**, anterior Primitive Streak; **p-PS**, posterior Primitive Streak

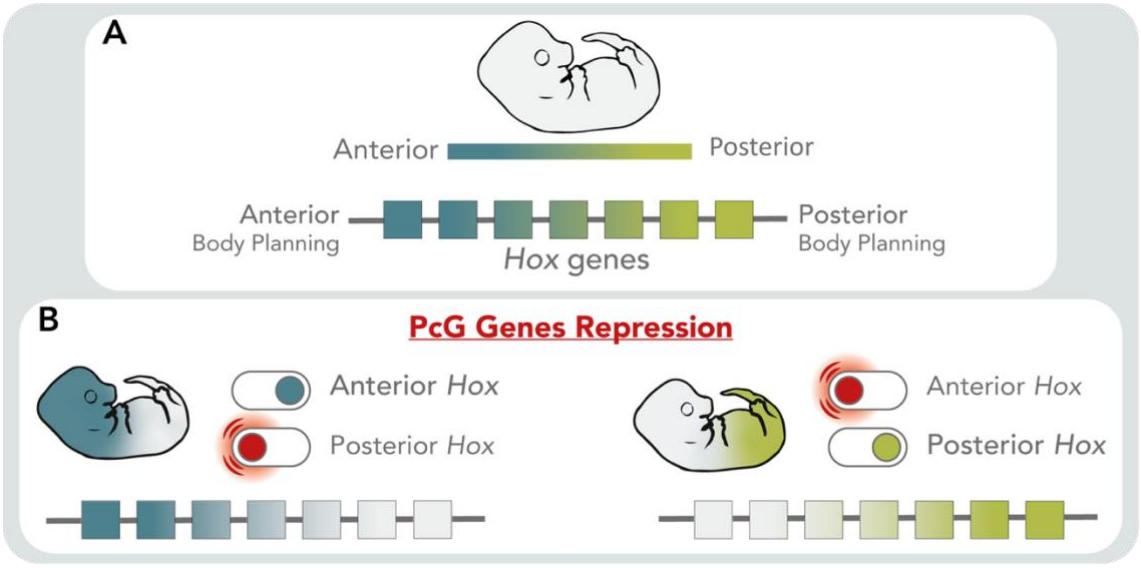
At E.6.5, **PS** is detected, testifying embryo entry into the gastrulation stage (Figure 18). During gastrulation, pluripotent EPI cells transform and generate the three germ-layers: the mesoderm, the endoderm, and the ectoderm. PS is derived from p-EPI cells and is the origin of the mesoderm and endoderm germ-layers. Firstly, on the PS posterior part, cells are mobilized, specialized, and invade the space between EPI mass and emVE, forming the mesoderm germ-layer. Then, in the PS anterior part, cells specialize into definitive endoderm (DE), migrate through the mesoderm layer, and are finally interspersed between emVE cells. The combination of emVE and DE cells forms the endoderm germ-layer. The mesoderm and endoderm germ-layers continue to grow and extend towards the anterior side. During this stage, the remaining EPI cells, the anterior ones (a-EPI), maturate into the ectoderm, the last germ-layer [93].

Mechanisms and Signaling Pathways Involved

Interestingly, with embryo AP definition, a huge, organized **landscape of signaling pathways** is setting up [84], [93], [94]. Cell fate is driven by the differential signaling pathways exposure in terms of concentration and duration to these. Bmp, Wnt, Nodal and Fgf signaling pathways start at gastrulation on the posterior side at the junction of EPI cells with ExE cells, impacting PS and germ-layer fate (Figure 19). Bmp signals are expressed and secreted by ExE cells, captured by neighbour EPI. Then, in response, p-EPI cells express the Wnt signal. Produced by p-EPI as well as VE cells, the WNT signal launches *Nodal* and *Fgf* signal expression by p-EPI. Finally, secreted NODAL promotes *Bmp* expression by ExE cells, feeding the signaling loop. Meanwhile, at the opposite site, AVE cells express antagonists of Nodal and Wnt signaling, limiting their activity to the posterior side. This limitation forces PS apparition to take place on the proximal posterior side and exposes it to signaling gradients. These gradients are at the origin of mesoderm and definitive endoderm cells. The endoderm germ-layer develops around stronger Nodal sources than the mesoderm germ-layer.

Organogenesis Preparation

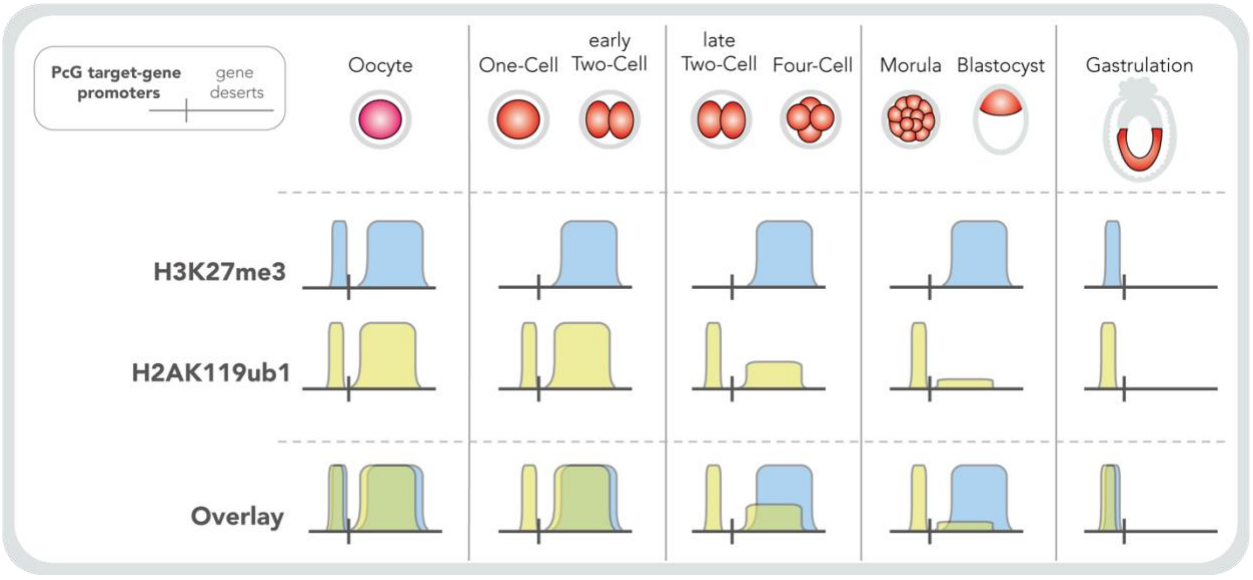
At the end of gastrulation, the embryo spares no effort to **prepare for organogenesis**. Cells in the newly formed germ-layers start right away to differentiate into progenitor cells of various organs and tissues. Organogenesis is the final stage of the embryonic development that begins at E8.5 and occurs until the end of gestation. It represents the development of all organs from the germ-layers and associated transformations until embryo maturation [93]. The ectoderm gives rise to epidermis and organs/tissues linked to neural parts like the brain, the spinal cord, and the peripheral nervous system. Meanwhile, postural structures generated from the paraxial mesoderm give bones, skeletal muscles, and ligaments. The reproductive system, kidneys, and non-epithelial parts of the urinary tract come from the intermediate mesoderm while the lateral plate mesoderm generates adipose tissue, blood, lymph, heart, spleen, and smooth muscles. Finally, the endoderm generates the epithelia of the gut



Adapted from [7]

Figure 20. Polycomb Group Proteins and Their Role in Antero-Posterior Patterning Through Hox Genes Repression

A. Schematic representation of antero-posterior gradient in mouse embryo and differential influence of *Hox* genes. *Hox* genes are represented by a square box. Blue colour represents *Hox* essential for anterior patterning, while green colour represents *Hox* essential for posterior patterning. **B**. Schematic representation of *Hox* regulation by polycomb group (PcG). On the anterior side, PcG represses expression of *Hox* implicated in posterior patterning. Anterior *Hox* are the only ones expressed. Inversely, on posterior side, PcG is implicated in anterior *Hox* repression. PcG is an essential mechanism to define anterior and posterior identity.



Adapted from [95]

Figure 21. Early Mouse Embryo Histone Modifications and Polycomb

Schematic representation of histone modifications H3K27me3 and H2AK119ub1 mapping, through early mouse embryo stages. After fertilization, H3K27me3 and H2AK119ub1 oocyte symmetries are depleted and restored after blastocyst implantation. H3K27me3 remains stable in protein-coding gene deserts along early mouse development until implantation, while H2AK119ub1 decreases over the time. Histone modifications on PcG target-gene promoters show that H3K27me3 disappears from fertilization to implantation, contrary to H2AK119ub1 that remains stable. As PCR2 is implicated in H3K27me3 of PcG target-gene promoters and PRC1 in their H2AK119, these observations highlight the dynamic and importance of PcG in mouse development. The absence of symmetry illustrates the independent PRC2-action that PRC1 may have.

H3K27me3, Histone-3 (H3) lysine-27 (K27) tri-methylation (me3); **H2AK119ub1**, Histone-2A (H2A) lysine-119 (K199) mono-ubiquitylation (ub1)**; PcG,** polycomb group**.**

tube and associated organs such as digestive tract, pancreas, and liver, that will be described in more details later.

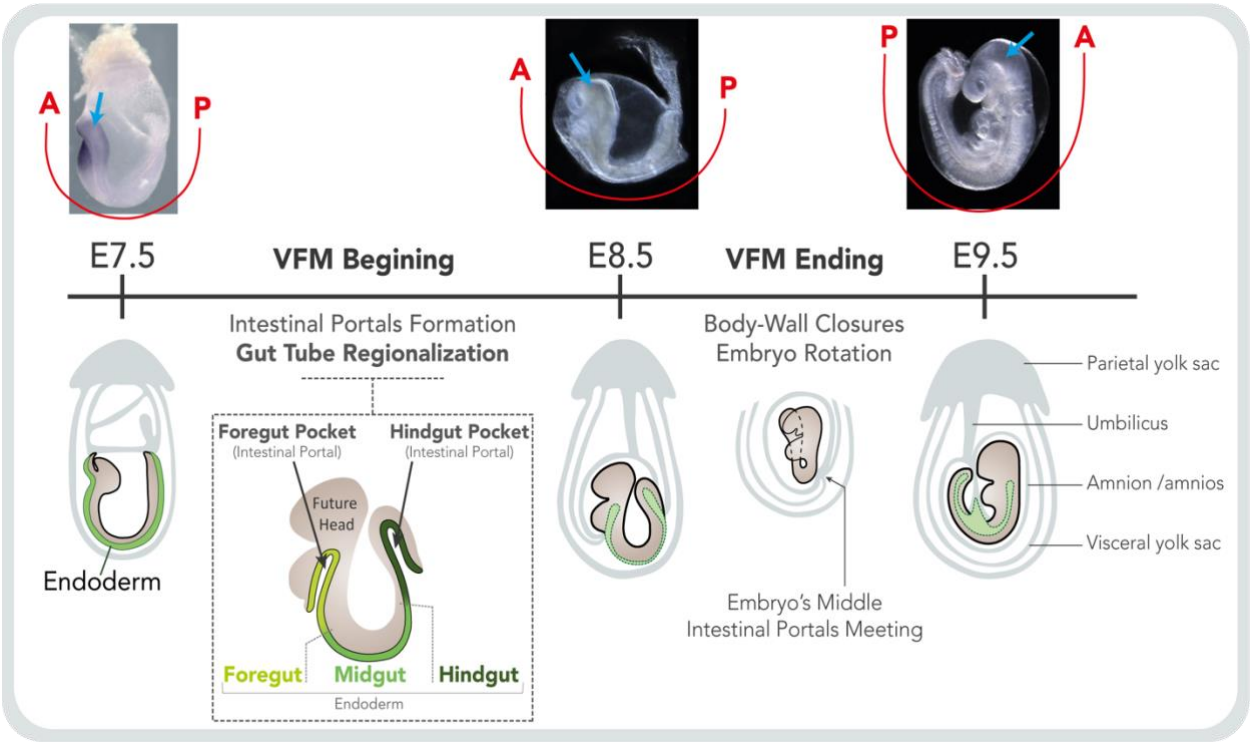
c. The Polycomb Group, Molecular Actors of the Development.

Mammalian development is highly regulated and involves many mechanisms for gene regulation apart from signaling pathways and transcription factors. It could be expected that histone modifications are essential for the development and the **polycomb group** (PcG) is involved in this process. PcGs are implicated in *Nanog*, *Oct4* and *Sox2* transcription regulation, playing a role in embryonic stem cell pluripotency and differentiation [96], [97]. It also has been reported that PcGs may repress expression of *Hox* genes, a group of genes coding for homeobox transcription factors, responsible for the body plan (Figure 20). Through its repression activity on *Hox* genes, PcGs are involved in antero-posterior patterning [7].

As a reminder, polycomb repressive complex 2 (PRC2) tri-methylates (me3) histone (H) 3, at its lysine (K) 23. H3K23me3 permits the recruitment of canonical PRC1 (cPRC1), that subsequently adds one ubiquitin group (ub1) to H2A on its K119. On the contrary, non-canonical PRC1 (ncPRC1) variants work independently of PRC2's methyl-depositing activity. The **biological dynamics** of H3K27me3 and H2AK119ub1 by PRC1 and PRC2 is essential for mouse development. It has been correlated by mapping studies from gamete stage before fertilization until gastrulation. Both histone modifications overlap in different time with PRC-targeted genes but also in the so-called gene deserts (Figure 21). The gene deserts define genome regions that do not contain any protein-coding gene. They represent around 20% and 25% of the genome in mice and humans respectively. Even if they do not encode for proteins, they have essential regulatory functions like distal enhancers involved in transcription initiation regulation [95].

Even if all subunits are not required, both polycomb complexes **PRC1 and PRC2** are essential for transcriptional repression in mouse development. Mouse mutants for complete PRC2 core depletion die generally during gastrulation progression. On the other hand, mouse mutants lacking PRC1 core are even challenged with more drastic lethality. PRC1 inevitably involves the core of RING1a or RING1b proteins which play non-redundant functions. Separately, *Ring1a*-null mice reach birth despite high anterior penetrance, while null-mice *Ring1b*-null mice arrest during gastrulation around E6.5. However, mouse mutants for both *Ring1a* and *Ring1b* (lacking PRC1 activity) arrest quickly at the two-cell stage [95].

CBX proteins are only found in canonical PRC1 variants. As some studies have shown, CBX proteins seem to compensate each other. None of them are individually necessary to carry out early mouse development. However, they have distinct roles and functions in late development. *Cbx2*-null mice present a reduced weight, gonad alteration specific to their gender and most of them die 4 weeks after birth. *Cbx4*-null mice have growth retardation associated with lethality before weaning. For Cbx6 and



Adapted from [98], Pictures taken from [99]

Figure 22. Following the Endoderm During Ventral Folding Morphogenesis

Ventral Folding Morphogenesis (VFM) begins on the anterior and posterior part of germ-layers. Endoderm forms pocket on its anterior and posterior sides and starts to regionalize into foregut, midgut, and hindgut. Endoderm Pockets progress toward each other to finally meet at the embryo middle. Then, embryo rotates, closing its body walls and enclosing it into extra-embryonic membranes. At the end, endoderm forms the gut tube inside the embryo.

Cbx7, their implications are trickier. *Cbx6*-null mice are born with several organ alterations, yet, depending on the mouse embryonic stem cells (mESC). These alterations are more or less sharp, as some mice can indeed reach adulthood. Lastly, *Cbx7*-null mice reach adulthood and may or may not develop liver/lung tumors or small defects [7].

Kdm2b is a specific subunit of ncPRC1.1. Through its reader ability for CpG islands, KDM2b is responsible for the PRC2-independent recruitment of ncPRC1.1 onto DNA. *Kdm2b*-null mice present lethality at mid-gestation. To go further, mice bearing partial *Kdm2b* depletion may reach birth, however they carry many defects in the posterior axial skeleton and represent a minor part of the litter [7].

These examples highlight the role of PcG in development and antero-posterior patterning. PcG variants seem to have more distinct roles in tissue/organ specialization.

B. The Endoderm Germ-Layer, the Final Side-Quest of Intestinal Organogenesis

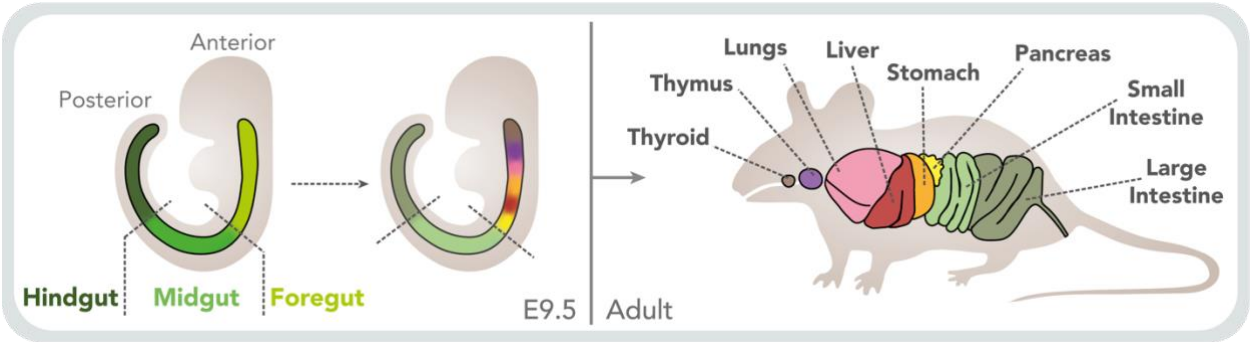
a. Novel of Ventral Folding Morphogenesis and Gut Tube Formation

By the end of gastrulation, the embryo has a cup-like shape composed by the three germ-layers. The endoderm forms the outside layer of the embryo, followed by the mesoderm and the ectoderm. Next, to transform a three-layer cup into a mouse fetus, the embryo processes ventral folding morphogenesis (VFM) [94].

VFM closes the body wall and leads to the recognizable fetal position. VFM begins with germ-layer inversion, resulting in endoderm internalization. It starts on the proximal part, in a way which looks like a very large sock that is turned over. On the most anterior side and posterior side, endoderm folding forms structures called intestinal portals, or also referred to as foregut pocket (anterior) and hindgut pocket (posterior). As VFM progresses, intestinal portals extend to each other, closing the embryo body wall like two zippers of a pencil case that finally meet each other in the middle of the embryo. At this point, the endoderm forms a primitive gut tube that remains open at midgut, where it is connected to the yolk sac. Finally, the embryo rotates and turns, leading to the fetal position and finishing body wall closure while locking up the embryo into extra-embryonic membranes (Figure 22) [94].

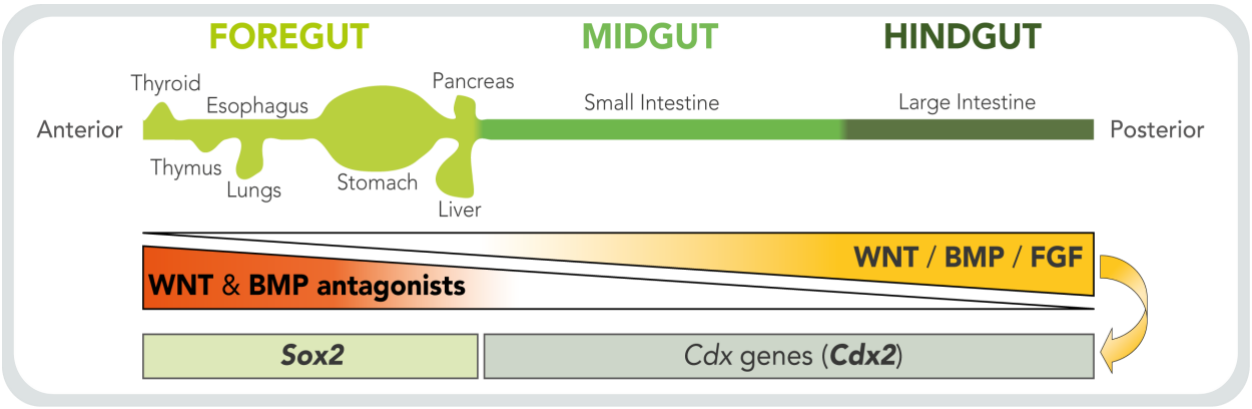
b. Gut tube patterning, Key of Endoderm's Organogenesis

Germ-layer movements during gastrulation and VFM put the endoderm in proximity to the differential mesoderm that surrounds it. Crosstalk between the endoderm and the mesoderm is one of the cornerstones of endoderm patterning just as the expression of regional transcription factors [100], [101]. By the end of gastrulation, the endoderm is composed of naïve cells that are still plastic and not yet engaged in specific organ fate. However, through gut tube formation, the endoderm is patterned by serial signal factors from the differential mesoderm along the antero-posterior axis. Gut tube patterning begins around E8.0 and is complete by E9.5, when both intestinal portals are joined.



Adapted from [94]

Figure 23. **Gut Tube Patterning and Organ Territories**



Adapted from [102]

Figure 24. **Gut Patterning and Signaling Pathways**

Mesenchyme secretes factors setting up a signaling landscape along the gut tube. WNT, BMP and FGF are higher in hindgut and midgut promoting *Cdx* genes expression, essential for intestinal fate. Meanwhile, the expression of WNT and BMP antagonists in the anterior part permits Sox2 gene expression, promoting anterior fate and patterning.

By this time, the gut tube is patterned into foregut, midgut, and hindgut regions that are distinguished from each other according to the expression of regional transcription factors. Foregut epithelium is characterized by the expression of *Sox2*. It will give rise to the esophagus and stomach along with trachea, lungs, liver, thyroid, and pancreas. On the contrary, midgut and hindgut will give rise to the small and large intestines, and express *Cdx* genes (caudal type homeobox), homeotic genes of ParaHox group implicated in midgut/hindgut identity (Figure 23).

The Mesoderm secretes several signaling factors, setting up a signaling landscape along the gut [100], [101]. Yet it was clearly established that Wnt, Fgf and Bmp signaling pathways are implicated in posterior regionalization, promoting midgut/hindgut identity and repressing foregut fate (Figure 24). The early temporal and spatial gradient of the Wnt pathway through its effector WNT3 is one of the signals helping to set up the different intestinal regions. The midgut endoderm, which is exposed to a lower level for a shorter time, sets the anterior domain of the small intestine. Meanwhile, the hindgut that is exposed to Wnt at a higher level for a longer time sets the posterior domain of small intestine, cecal and colonic gene expression. Wnt plays a critical role in intestinal endoderm specification by inducing *Cdx2* gene expression. *Cdx2* is the "mastermind" behind the intestinal identity, whose deletion results in a complete loss of intestinal gene expression and structure. FGF4, a signaling factor of the Fgf pathway, promotes *Cdx* gene expression, inducing intestine identity. In addition to their role in promoting intestinal specialization, both WNT3 and FGF4 repress foregut fate. In parallel, the Bmp pathway appears to be involved as well. BMP4 effectors promote midgut/hindgut *Hox* genes expression and the inhibition of BMP2 and BMP4 effectors in the anterior part promotes foregut fate [102].

II. Intestine Organogenesis, the Final Quest in Three Acts

A. Act One: Intestine Elongation and Regionalization

One of the most important facts about the intestine is its incredible length. Even if intestinal elongation will continue as long as the mouse increases in size, even after birth, events occurring during development are critical to establish a proper length and loop structures. At E9.5, intestinal gut expresses Cdx2 and is organized in a pseudostratified epithelium that proliferates rapidly and uniformly prior to E14.5, increasing intestinal length, circumference, and luminal area. Cell proliferation is an important factor in gut elongation. The **crosstalk** between the mesenchyme and the epithelium is again essential for intestinal lengthening [101].

Before villus formation, the **Wnt signaling pathway** controls the pseudostratified intestine proliferation through its member Wnt5a expressed by mesenchymal cells. The Wnt pathway controls proliferation and the epithelium integration of post-mitotic cells following a precise gradient. Mis-regulation of Wnt5a leads to a shorter length, perturbed midgut, truncated caecum, and integration failure of post-mitotic cells into the epithelium, resulting in a dense epithelium of 3 to 4 cell layers

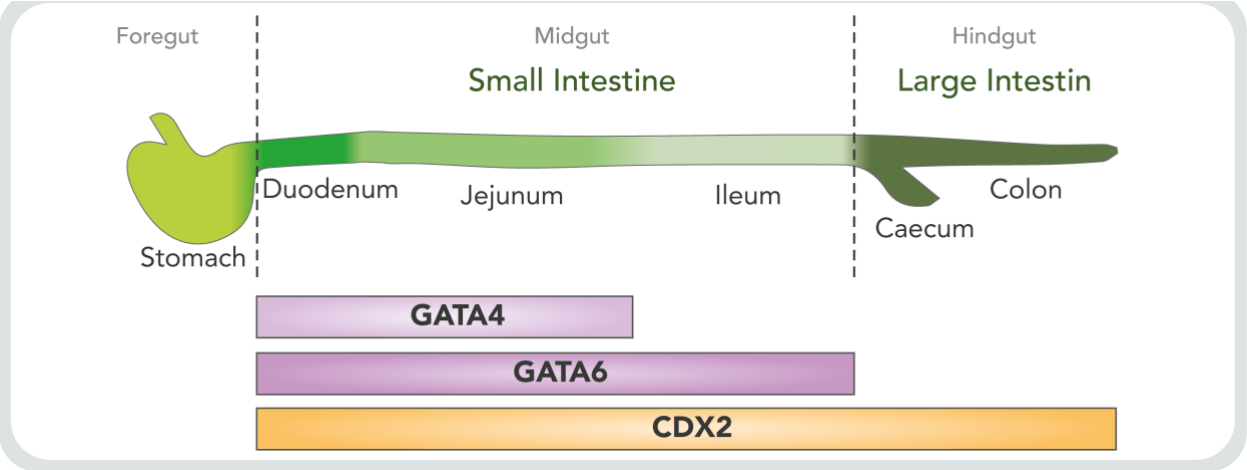
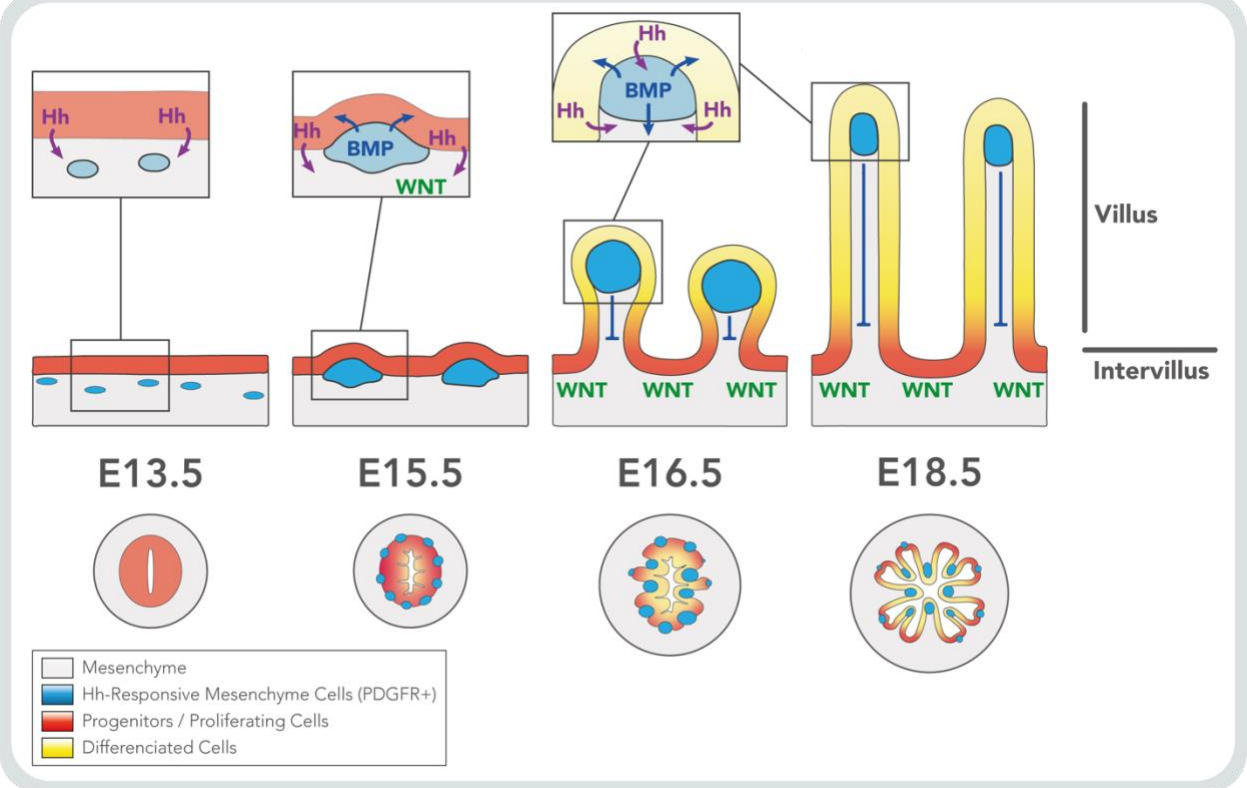


Figure 25. Embryonic Intestinal Regionalization
Schematic representation of intestinal regionalization and some transcription factors implicated.



Adapted from [101]

Figure 26. Villus Morphogenesis of Mouse Small Intestine
Schematic representation of developing intestinal epithelium through sagittal view or cross section view. Between E9.5 and E13.5, the early developmental intestine is flat and composed by uniform progenitors. AtE13.5, epithelial cells secrete Hedgehog signals (Hh). Hh-responsive mesenchyme cells aggregate into clusters under the intestinal epithelium. At E15.5, clusters constrict physically the epithelium between them, consequently forming villus-bumps that will growth until they form the complete mature villi at E18.5. In response to Hh-signals, clustered cells secrete BMP signals that restrict WNT signals, compartmenting progenitors, and proliferative cells in intervillus domains.

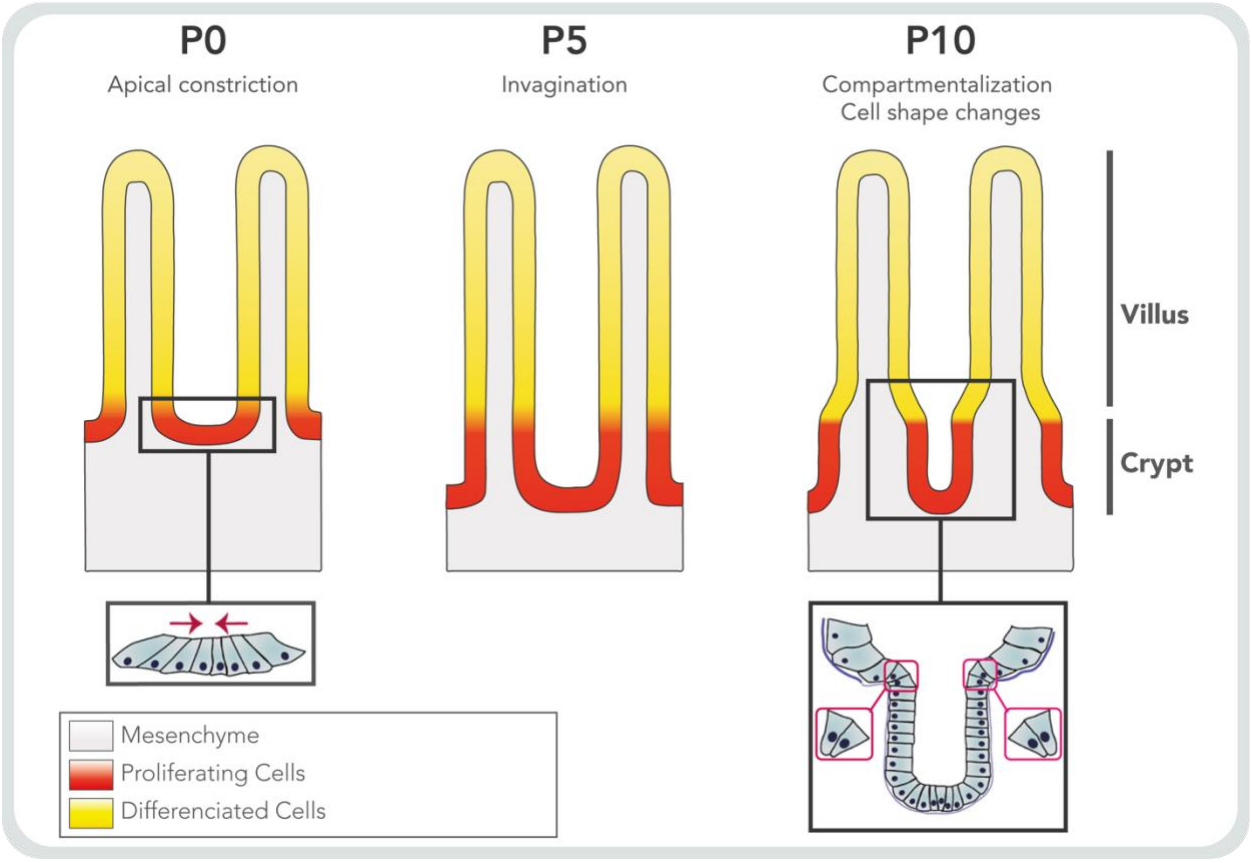
instead of a perfect single cell layer pseudostratified epithelium [103]. Besides Wnt, the Hedgehog (Hh) and Fgf pathways, through Fgf9 activity, are implicated in intestinal elongation. Alteration of these pathways impairs intestinal length. However, precise mechanisms involved are still unknown. Along with the signaling pathways, transcription factors like Gata4 are also implicated by regulation of cell cycle mediators.

During elongation, the intestine starts to regionalize, in order to define specialized parts with distinct purposes (Figure 25). Right after the stomach, duodenum secretes a variety of enzymes for food breaking down, while the following jejunum and ileum are specialized into specific molecule absorption. The Gata transcription factor is a transcription factor family known to be involved in small intestine regionalization. Gata6 is expressed along the small intestine. Furthermore, it is important for ileum specialization by repressing the colonic fate. Gata4 is expressed in both duodenum and jejunum, and enough by itself to determine the jejunum fate [101].

B. Act two: Villi Raising Signals and Epithelium Maturation

Villi are finger-like structures that start to rise at E13.5-14.5, from the flat pseudostratified epithelium into the lumen. They increase the absorbing surface and at the end of intestinal maturation, they will contain all the differentiated cells essential for absorption and metabolism. To form villi and specify into the intestine epithelium, the small intestine endoderm needs to associate with its very old partner in crime, the mesenchyme. Crosstalk between them is critical for villus formation and intestine epithelium specialization, implicating several signaling pathways like Bmp, Hedgehog (Hh) and Wnt pathways (Figure 26).

The epithelium participates in the crosstalk by secreting Hh signals, regulating mesenchyme maturation and promoting aggregation of sub-epithelial mesenchymal clusters. In return, Hh-responsive mesenchymal cells secrete several factors like BMP2/4 and WNT5A. Cluster cells underlying future villi are characterized by the expression of PDGFRα (Platelet-Derived Growth Factor Alpha), a receptor of PDGFRA signaling effector secreted by epithelial cells. BMP proteins are expressed by the cluster cells. BMP2 and BMP4 are critical for controlling cluster distribution and size [104]. Even if the exact mechanisms involved in this process are still unknown, signals secreted by cluster cells induce overlying epithelial cell changes. They lose their pseudostratified organization for a polarized columnar organization and seem to stop proliferating through Bmp pathway influence. Later, when villi start to form, Bmp secretion restricts the proliferating signal Wnt5a in intervillous domains, the area between the villi. Besides their signaling function, the mesenchyme clusters physically constrict the epithelium which consequently forms villus bumps that will grow until they form full mature finger-like villus structures that become differentiated domains. By E18.5, all mature differentiated cell populations are represented in villi while intervillous domains contain progenitors.



Adapted from [105]

Figure 27. Crypt Morphogenesis of Mouse Small Intestine

From the first post-natal day (P) of mice, the flat intervillus epithelium layer starts to constrict, subsequently, the forming crypt invaginates into the mesenchyme. Finally, the differentiated cells located at the crypt-villus junction change their shape for a "wedge-like" one. At this point, the crypt-villus axis has adopted its final form.

C. Act Three: Crypt Morphogenesis, the Final Act of Intestine Maturation

While the villi are the domain of differentiated cells and stand in the lumen, crypts are the domain of stem-cell niches, taking the form of an organized cavity embedded in the underlying mesenchyme. In mice, crypts are formed after birth, during the first week of life, from intervillous cells.

Until recently, the crypt **morphogenesis hypothesis** was that mesenchymal cells initiated upward movements from crypt-villus junctions, developing the crypt architecture. This model excluded downward movements of intervillous cells. However, Sumigray *et al.* **updated this model** with a new model of crypt morphogenesis [105]. They followed crypt morphogenesis from post-natal day 1 (P1) to P10 by using a combination of 3D microscopy imaging and transcriptomic analysis. They showed that the crypts are formed from intervillous invagination into the mesenchyme. This invagination is led by the constriction of cells by myosin-II apical protein which forces cell direction towards the mesenchyme. This process is complete at P10 (Figure 27).

In parallel with the morphogenesis model, the BMP pathway has been revealed to be essential for crypt morphogenesis. Inhibition of this pathway induces abnormal formation of crypts perpendicular to the villus axis in a high number [106].

Finally, the end of crypt morphogenesis defines the **mature state** of intestinal development. By that time, the intestine takes its final shape, expressing all intestinal cell types that are compartmentalized with the signaling pathways along the crypt-villus axis. First chapters of the intestine life story end here with its maturation. As the page is turned, the new chapters of homeostasis and adulthood begin.

III. Transcription Factors, "Round Table Knights" of the Intestine

Through all the chapters of its story, the intestine works closely with its associates. In addition to its old partner in crime the mesenchyme, the intestinal epithelium expresses many different transcription factors. These transcription factors are like the intestine's "round table"; they may have different degrees of importance and roles, but all together, they are essential to the quests of intestinal development, maturation, homeostasis, and protection. Some of these key transcription factors are presented below.

A. *Cdx* genes, Intestine's "Knight Masters"

As described earlier, ***Cdx* genes** (for caudal homeobox genes) are a group of transcription factors expressed in the nucleus and containing homeobox domain. This family controls the spatial development of embryo posterior by regulation of *Hox* genes among others.

Cdx2 is known to be the great "mastermind" behind intestinal function. As discussed previously, its expression begins at E8.5 in the endoderm and it is critical to define both

midgut and foregut fate as well as intestinal identity. CDX2 is expressed uniformly along the crypt-villus axis along the intestinal tract and implicated in intestinal identity and survival. Although, its critical roles vary in function of the developmental stage. At E8.5-E9.5, *Cdx2* is critical for intestinal identity, as its depletion leads to an atrophic intestine with anterior conversion and colon deletion [107]. Yet, intestines depleted for *Cdx2* at E13.5 are not fully converted or are amputated. Moreover, they present a reduced villus length and partial stratification into gastric/esophageal units [108]. In the adult stage, *Cdx2* is essential for cell differentiation and maintains stem cell/progenitor populations and proliferation [109], [110]. As well, it is essential to maintain intestine and villus length [110]. Chromatin accessibility profiles differ in embryonic and adult intestine. Part of *Cdx2* master mind activity depends on the temporal accessibility of its target genes [111]. Additionally, *Cdx2* has also been reported to prevent H3K27me3 in tissue-specific zones. This preventive activity seems to be turned against the polycomb complex PRC2, as half of the genes repressed in *Cdx2*-null intestines and non-targets of Cdx2 are restored [112]. One final important point is that Cdx2 has an intestinal protective role of tumor suppressor [113].

Cdx1 expression begins at E7.5 and continues through adulthood along the intestinal tract, overlapping Cdx2 expression. Even so, Cdx1 is gradually expressed along the crypt-villus axis, with a stronger expression in the crypts. Cdx1 is unnecessary for mouse survival. As Cdx1 activity is redundant to Cdx2, intestine development/homeostasis of fully invalidated mice for *Cdx1* are unaffected [114]. However, *Cdx1* is still implicated in intestinal identity as its forced expression in gastric cells induces expression of stem-cell marks and differentiated intestinal markers [115].

B. Transcription Factor: "Squires" of the Intestine

a. HNF4 proteins

HNF4 proteins (Hepatocyte nuclear factor 4) are a group of transcription factors belonging to the nuclear receptor superfamily. This family includes **HNF4α** that is uniformly expressed in the stomach, small intestine and colon. **HNF4γ** is only expressed in the developing and mature colon. Both *Hnf4* factors are dispensable for early intestine specialization but are essential for the small intestine maturation. *Hnf4γ* compensates *Hnf4α* loss but deletion of both of them impairs intestinal maturation and post-natal survival [116]. *Hnf4α* is also critical for colon development [2]. In adulthood, *Hnf4α* is implicated in proliferation regulation, cell differentiation and intestinal barrier function. Its inactivation induces minor alterations, nevertheless there is an increase of proliferation, an impaired balance between differentiated cells and defects of cell-cell junctions leading to barrier permeability [3]. Interestingly, *Hnf4* factors control chromatin accessibility in maturating fetal intestine (>E16.5) [116].

b. SOX9

Sox9 belongs to the SOX gene family (SRY-related HMG-box genes), with an expression restricted to intervillous domains during developmental stage and in crypts in the mature intestine. In the mature intestine, *Sox9* is not critical for mouse survival, unessential for small intestine architecture but essential for colon architecture. *Sox9* is expressed in stem cells, progenitors, and some few differentiated cells [117]. It is implicated in cell proliferation control, cell differentiation [118], and reserved stem cell maintenance [119].

c. GATA, a family business

Gata family. Among the Gata members, Gata4, Gata5 and Gata6 are expressed in the developing endoderm. However, while Gata4 and Gata6 are essential for gastro-intestinal development, Gata5 is unnecessary. Gata4 is expressed in the proximal part of the small intestine and Gata6 in the small and large intestine.

Gata4 is expressed in the upper part of the crypts and along the villi. It is known to be critical for controlling jejunum fate, villus length, and enterocyte maturation in both developing [120], [121] and mature [122] proximal intestines. Finally, unlike the mature intestine, *Gata4* is required for epithelial cell proliferation and small intestinal growth [121].

On the other hand, **Gata6** is expressed along the crypt-villus axis and after the beginning of the act 2 (villi formation), it is no more essential for jejunum development, during which it has a minor implication in cell differentiation. Still, *Gata6* is more implicated in ileum development in terms of villus length, proliferation, and differentiated cell balance/identity [120]. In the mature intestine, *Gata6* has the same implication as in the developing intestine [123]. Additionally, *Gata6* is also involved in intestinal barrier [124] and necessary for colon proliferation, cell differentiation and maturation [125].

CHAPTER 3

The Network of Intestinal Homeostasis

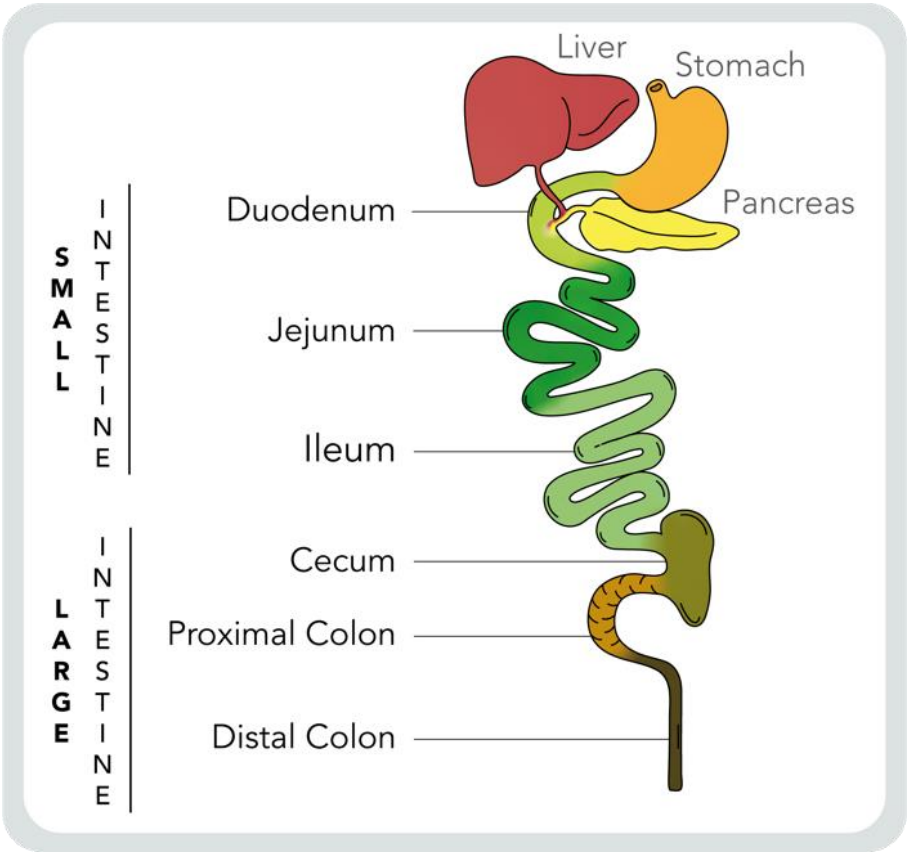


Figure 28. Mature Mouse Intestine Segments and Direct associated Organ Partner

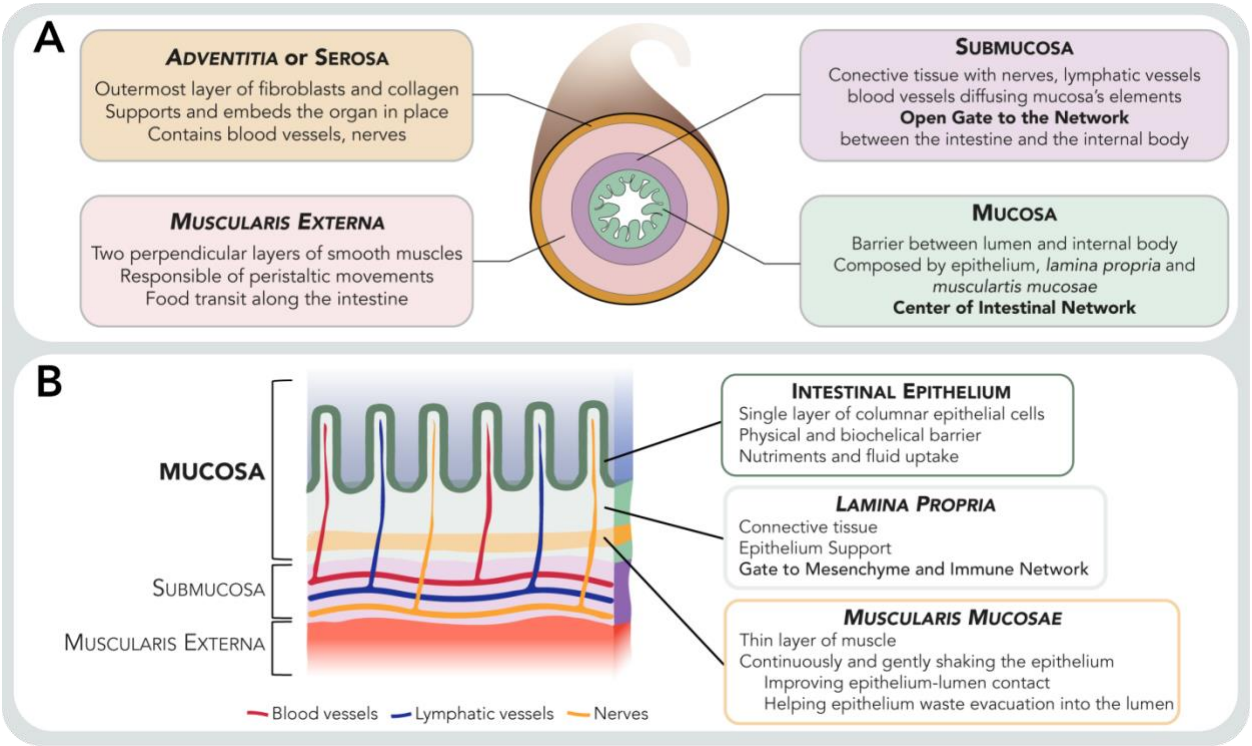


Figure 29. Structure of the Intestine and Intestinal Epithelium along the gut

The intestine, or the gut, is the part of the gastrointestinal tract that extends from the duodenum, under the pyloric junction with the stomach, to the rectum. It is composed of the small intestine and the large intestine or colon. It represents a dynamic platform at the interface between the organism and the intestinal cavity, called the lumen. It is essential for digestion, nutrient, and water absorption, in addition to its role of protective barrier against the microbiota and harmful or undesirable compounds present in the lumen. When the intestine reaches its maturity, all the features and complex networks essential for its homeostasis have been put in place. Here, through a window of the intestinal epithelium microenvironment, we will take the time to explore the beauty of this highly engineered and balanced network behind the intestine physiology.

I. Intestinal Epithelium Physiology and its Closest Network Partners

A. General Intestine Organization and Functions

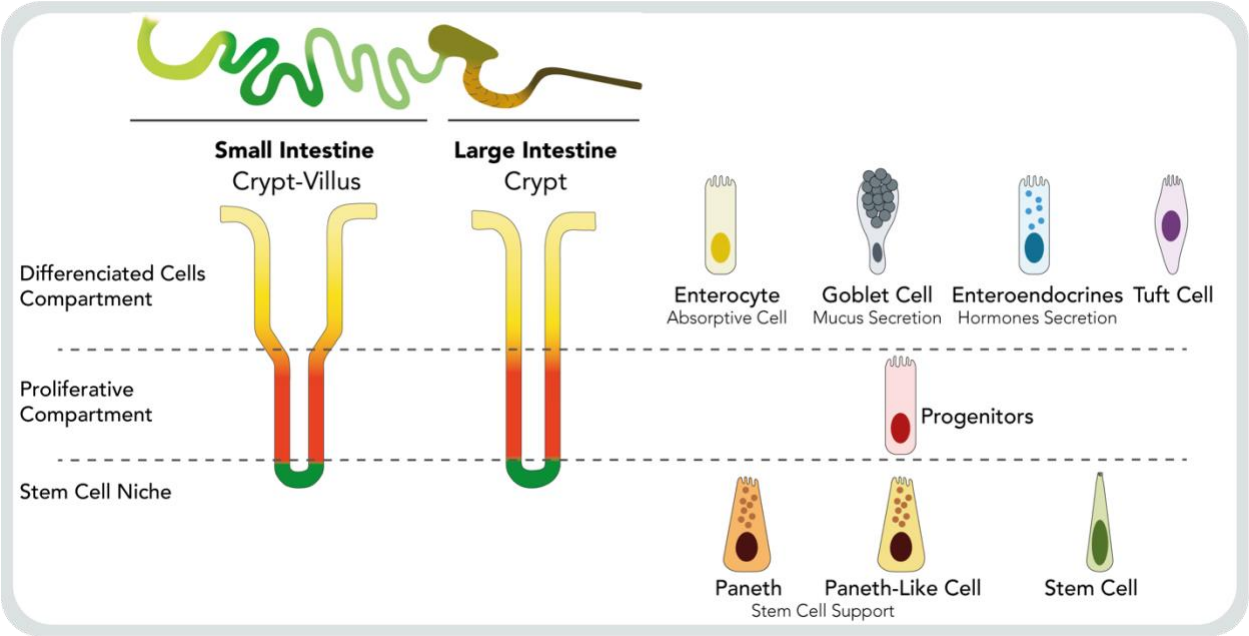
a. Regionalization and Functions of Intestine Segments

The small and large intestines have their own functions in food processing. The small intestine is mostly specialized in absorption of lumen components, while the large intestine is responsible for waste processing and water absorption. To work efficiently, they are organized into specialized segments with distinct physiological functions (Figure 28) [126].

After food breaking down in the stomach, smaller food particles enter the duodenum, the first and shortest segment of the small intestine. The duodenum contributes to transforming food into nutrients thanks to the liver and the pancreas, which release bile and enzyme cocktails respectively. The jejunum then comes into action by absorbing water and important nutrients like amino acids, carbohydrates, and fatty acids. Finally, the most distal segment of the small intestine absorbs the remaining nutrients and previously unabsorbed digested products like vitamin B12 and bile acids [127]. Also, as it has a high abundance of organized immune structures, the ileum is an important actor of the intestinal immune system. In the end of the digestive system, the colon is composed by the caecum, the proximal colon, and the distal one. It is essential for absorption of remaining water and feces processing.

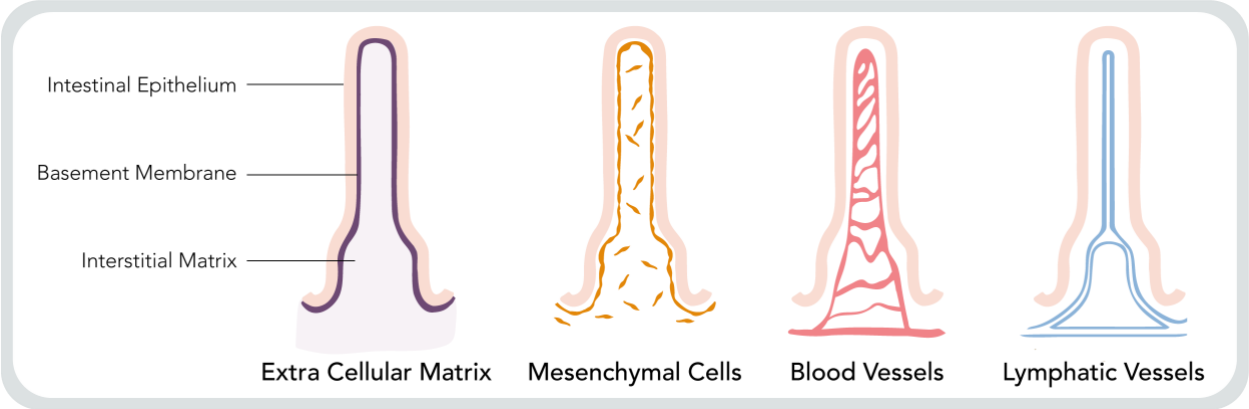
b. General Intestine Structure

Along the intestinal tract, segments take the shape of a tube composed of a succession of layers, *adventitia*, *muscularis externa*, submucosa, and mucosa (Figure 29.A). The *Adventitia* is the outermost layer of relaxed fibroblasts and collagen surrounding an organ. This layer connects the organ with the rest of the organism with vessels and nerve plexuses passing through it. The *Muscularis externa* is composed of two layers of smooth muscles, responsible for independent peristaltic movements. The submucosa is a conjunctive tissue containing blood vessels, lymphatic vessels, and nerves. The mucosal barrier is the most internal layer physically separating the lumen and the organism. It comprises the lumen, mucus layer and mucosa. The mucosa is itself composed of the *muscularis mucosae, lamina propria*, and is innermost within the intestinal epithelium, the nerve center of intestinal functions and network. Each component of the mucosa plays its own physiological functions (Figure 29.B).



Inspirate from [128]

Figure 30. Intestinal Epithelium Structure



Inspirate from [129]

Figure 31. Composition of the *Lamina Propria*, An Open Gate of the Intestinal Networks

c. Architecture of the Intestinal Epithelium in function of the Intestinal Segment

The intestinal epithelium is a high turnover tissue in constant renewal, at the interface between the intestinal lumen and the body, playing a major role in digestion, barrier, and transport. It is composed of many different cell types grouped in three distinct conceptual domains: the domains of differentiated cells, of proliferative cells and the stem cell niche [128].

Along the intestine, the epithelium adapts its architecture, and the characteristics of its differentiated cells depend on the intestinal segments and their underlying physiological functions. As the small intestine is specialized in nutrients absorption, it adopts the structure of crypt-villus axis to increase the absorptive surface and contact with the lumen. On the other hand, as the major function of the colon is rather waste processing, it takes the structure of long crypts without villus, which facilitates dealing with sharp dry contents (Figure 30) [128].

Cellular populations of the small intestine epithelium are specialized to deal with lumen components, whereas the absorptive cells, and the enteroendocrine cells for hormone secretion in response to food. In comparison, the colon has a higher proportion of goblet cells that secrete a protective mucus to preserve the epithelium against luminal components. The intestinal epithelium architecture, physiology, cellular renewal, and differentiation respond to a very precise topological and functional organization involving a precise balanced network.

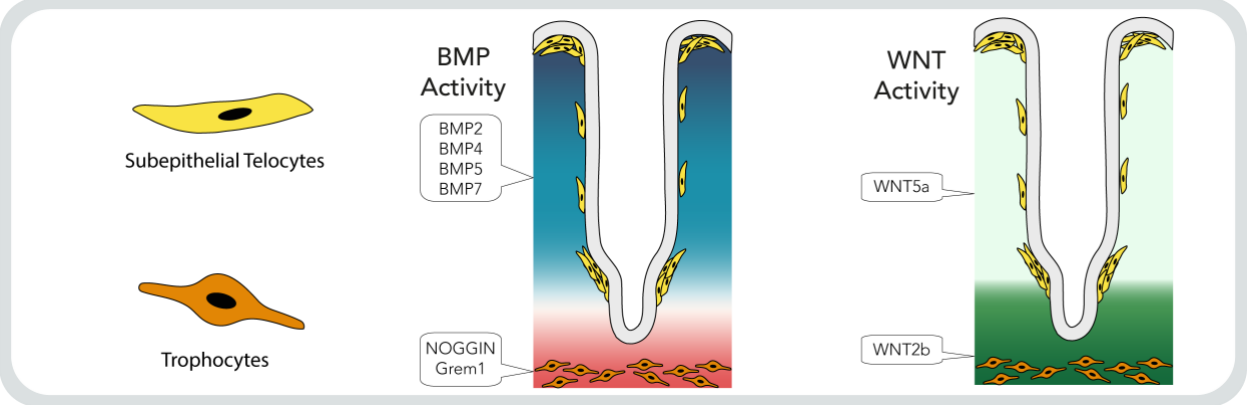
B. The *Lamina Propria*, the Closest Network Partner for Intestinal Epithelium Physiology

a. The *Lamina Propria*, a Dynamic Platform of Network Actors

In the mature intestine, the epithelium lies on the *lamina propria*. Due to their proximity, the *lamina propria* is the first and essential physiological associate of the intestinal network.

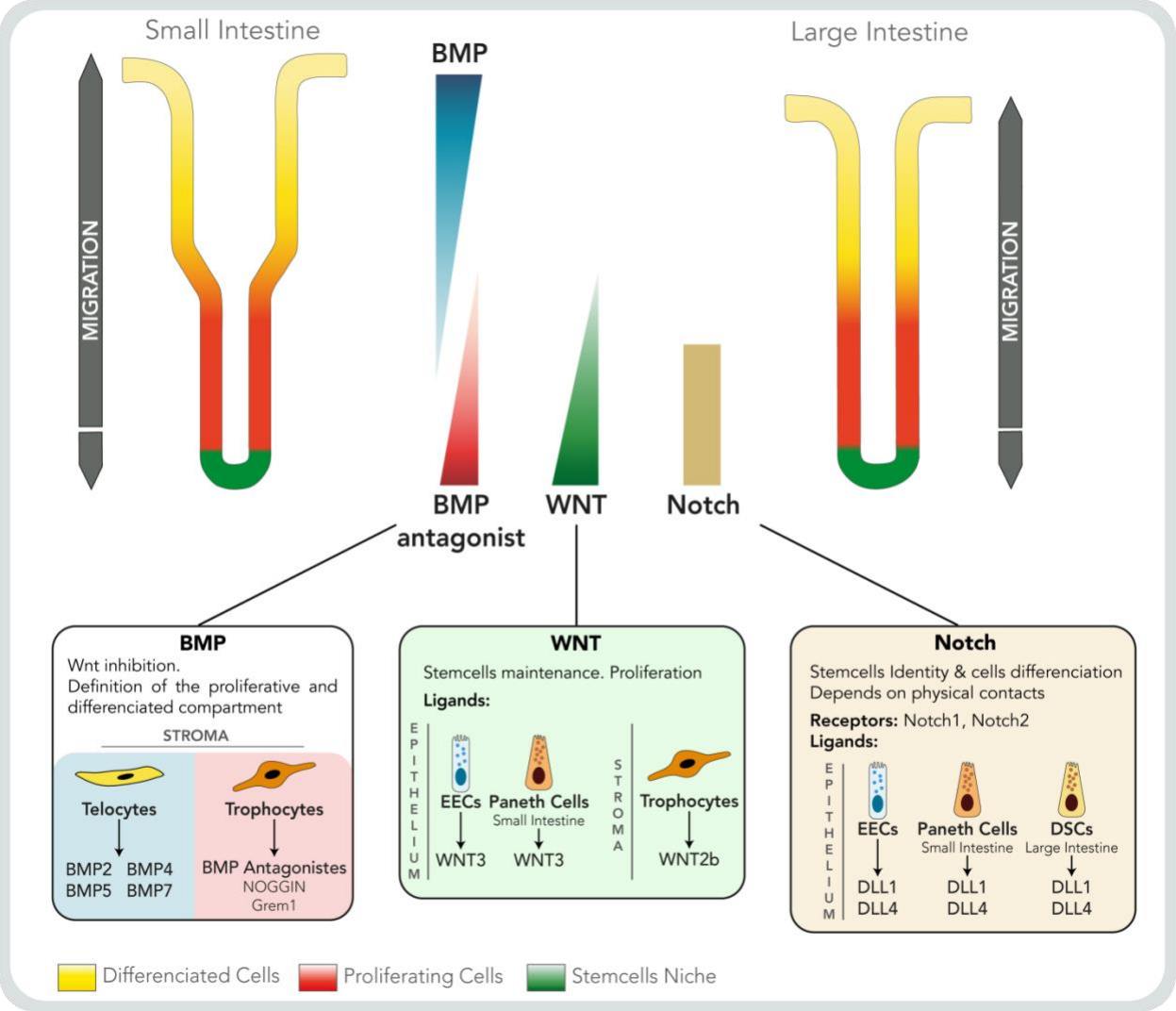
The architecture of the *lamina propria* is consolidated by an extracellular matrix (ECM), forming a support structure for resident cells [130]. The *lamina propria* ECM forms two distinct structures, at basement membrane and an interstitial matrix. The basement membrane is close to epithelial cells, providing an adhesion support and acting in organization and differentiation. Interestingly, its main components, the type IV collagen and laminins, are all secreted by the epithelial and mesenchymal cells [131].

Lastly, the *lamina propria* may be considered as a dynamic platform of the intestinal network. Indeed, it contains blood and lymphatic vessels that connect and expose the intestinal network to the internal body [130]. As well, the *lamina propria* contains mesenchymal cells and active moving immune cells (Figure 31). We will see later that these cells are critical for intestinal epithelium physiology by setting-up a signaling landscape.



Inspirate from [132]

Figure 32. Presentation of the Telocytes and Trophocytes Mesenchymal Cells.



Inspirate from [128]

Figure 33. Signaling Landscape in Intestinal Epithelium

a. Mesenchymal Cells of the *Lamina Propria*, Old Partners

Since the development, the architecture, regionalization, and functions of the intestinal epithelium are shaped by its interactions with mesenchyme cells (MCs). These interactions are essential for the
compartmentalization of the intestinal epithelium cells (IECs) and the balance between proliferation and differentiation [130].

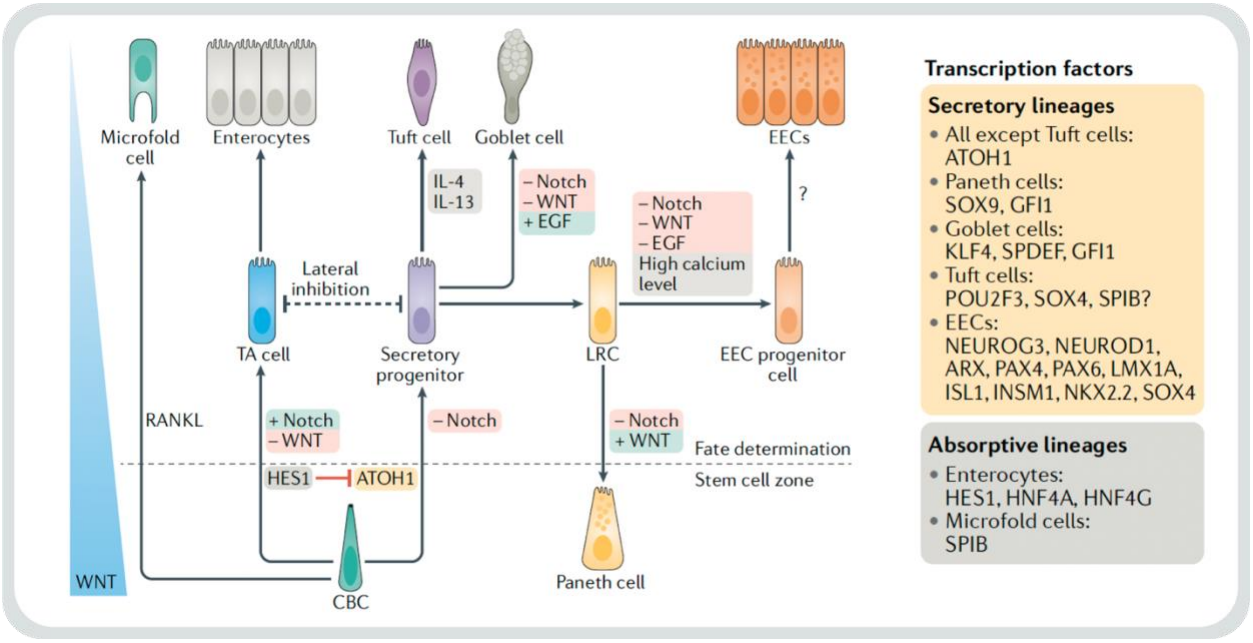
MCs are derived from the mesoderm; they are non-endothelial and non-hematopoietic cells. In the *lamina propria*, MCs have different features and functions. They are separated in distinct groups like the subepithelial telocytes and the trophocytes (Figure 32). Identified in 2020, these MCs play a crucial role in the *lamina propria* signaling landscape [133], [134]. The subepithelial telocytes accumulate at the crypt-villus junction and on the top of the villi. They secrete BMP ligands, Wnt5a ligands and, IECs pro-differentiation factors. On the other hand, the trophocytes, located at the *lamina propria* and just below the crypts, secrete BMP antagonists and Wnt2b protein, promoting IECs proliferation. The combination of these secretions confine BMP activity at the villi and sets an important Wnt activity in the crypts. These observations prove that MCs are crucial for the IECs microenvironment and the intestinal epithelium balance. Even if the *lamina propria* signaling gradient is known, the specific identity and contribution of mesenchyme cells are not fully elucidated yet.

Recent studies highlight a new implication of MCs in cell migration. As explained before, intestinal epithelium is in constant renewal. IECs migrated from the bottom of the crypts to the top of the villi. Cell migration is a critical part of intestinal epithelium physiology. This cell migration seems to involve not only the passive force exerted by mitotic cells but also active ones [135]. These forces may be driven by MCs [135], more precisely a subepithelial group of MCs expressing the muscle protein α-SMA (α-smooth muscle actin): the myofibroblasts [136].

C. Intestinal Epithelium Cell Fates Commitment and its Partner Network

a. Signaling Landscape, Transcription Factors and Cell Fates

As introduced before, the intestinal epithelium is organized into different cell compartments defined by a precise signaling landscape along the crypt-villus axis. This landscape is sets up by the secretion of signaling pathway ligands by mesenchymal cells and IECs themselves. The main signaling pathways involved in cell fate, proliferation or inhibition of proliferation are presented in Figure 33 [137].



Picture from [128]

Figure 34. **Intestinal Cells Fate, Signaling Pathways and Transcription Factors**

The intestinal epithelium is characterized by a high turnover rate and diversity of differentiated cells. This turnover is enabled by a rapid proliferation of stem cells at the base of the crypts, which generates transit-amplifying cells (TA cells) or secretory progenitors, that will mature into the differentiated lineage. Through their maturation, differentiated cell migrated from the crypt base to the tip top of villus, where they are eliminated into the lumen by a process called anoïkis. Mature cells only live a few days, migrate, and are dismissed at the tip top of the villus, except the case of Paneth which may live for 1-2 months, migrate to the base of the crypts, and are eliminated by infiltrated macrophages at the end of their long lifetime.

The small intestine comprises six differentiated cell lineages divided in two categories, the absorptive and the secretory lineages. The generation of these cell lineages starts when the stem-cells
or crypt-based columnar cells (CBCs) are pushed out of the stem-cell niche at the bottom of the crypts. The first decision regarding cell lineage is imposed by Notch signaling (Figure 34).

CBCs that keep the Notch signals maintain the expression of Hes1 (Hairy and enhancer of split 1), a transcription factor that represses pro-secretory transcription factors [137]. Loss of the Wnt signaling is an essential part of enterocyte differentiation as the activation of the Wnt pathway stops enterocyte differentiation [138]. TA-cells further mature into **enterocytes** as they migrate along the villi. The transcription factors Hnf4α and Hnf4γ in addition to their implication in intestinal development (cf previous chapter), are also essential for enterocyte differentiation [3], [139].

M cells are another type of absorptive cell. They are specific to the Peyer's patches, a secondary immune structure within the intestinal mucosa. All epithelial cells express the receptor RANK (receptor activator of NF-κB) for RANKL (RANK ligand), a ligand of the TNF (tumor necrosis factor) signaling pathway secreted by stroma cells of the Peyer's patches [140]. RANKL is sufficient to induce M cell differentiation, as treatment of IECs with RANKL induces M cells differentiation outside the Peyer's patches [140].

However, in CBCs that lose Notch signaling, Hes1 expression is no longer maintained to repress prosecretory factors. Consequently, Math1 (also known as Atoh1) transcription factor is expressed and CBCs mature into **secretory progenitors**. At this point, secretory progenitors will give rise to four different lineages in function of the signals. Wnt pathway is the second

pathway involved in cell fate. Secretory progenitors that do not receive sufficient WNT signals will mature into either goblet cells or enteroendocrine cells (EECs).

The differentiation into **goblet cells** depends on the EGF signaling pathway and the transcription factors Klf4 [141] and Spdef [142]. Inhibition of the EGF pathway *in vitro* induces a strong shift of the goblet cell lineage toward the EECs lineage [143]. *Klf4* is negatively regulated by Notch signaling through HES1, while *Spdef* is under the transcriptional control of MATH1. Both of these transcription factors are essential for promoting cell fate commitment and maturation to the goblet cell lineage.

EECs and Paneth cells are derived from the same intermediate mature secretory progenitors, the slowly dividing label-retaining cells (LRCs). In a reduced EGF and Wnt signaling context, LRCs mature into **EEC progenitors** that express Neurog3 (Neurogenin 3). Neurog3 is downregulated by Notch targeted HES1 in absorptive lineages. This transcription factor is crucial for EEC lineage, for depletion of *Neurog3* induces loss of all different EECs mature types [144]. Precise upstream signals implicated in enteroendocrine subtypes specification are still not well defined. On the other hand, LRCs that retain enough Wnt activity, mature into **Paneth cells** after they quit the stem cell niche. Wnt targets two important genes for Paneth cells, *Sox9* and *EphB3*. Sox9 is a transcription factor directly targeted by the Wnt pathway and critical for Paneth cell differentiation. Impaired *Sox9* expression leads to loss of Paneth cells [118], [145]. EphB3 for ephrin type-B receptor 3, is a receptor for the ligand Ephrin-B, highly present in the stem-cell niche. Unlike the migration pattern of other cells, these receptors permit the migration of Paneth cells from the bottom of the crypts to the stem-cell niche.

Finally, the last secretory cells deriving from the secretory progenitors are the tuft cells. **Tuft cell maturation** is independent of Notch, Math1, Neurog3, and Sox9, while it is involved in the immune network. Innate lymphoid immune cells in the stroma secret the cytokines IL-4 and IL-13 essential for tuft cell differentiation. Pou2f3 is also a critical transcription factor for tuft cell differentiation [146] and its expression is stimulated by IL-13 *in vitro* [147].

b. Chromatin Accessibility, Polycomb Group and Cell Fates

Along the intestine life, the transcriptional program induced by specific transcription factors is the central element for cell identity, differentiation, and maturation. Indeed, around 4000 to

6000 genes are differentially expressed in all IEC types [148]. As the regulation genes prone to transcription is the essential part of the transcriptional program, this regulation involves chromatin dynamics and accessibility, which depend on the spatial time and lineage specificity. The dynamics of chromatin accessibility and cell fates has recently been investigated. Even if this investigation is in its early stage and needs to go further, some features of IECs chromatin accessibility have been highlighted.

In the intestine, Cdx2 and Hnf4 are important transcription factors for intestinal identity, development, and homeostasis. Cdx2-targeted genes expression is dependent on chromatin accessibility. For example, in the intestinal stem cells (ISCs) and differentiated cells, CDX2 interacts in total with 10,733 binding sites, but only 1372 sites are common between them [148]. Yet, CDX2 is unable to induce chromatin changes to access to its target genes [111]. However, it has been reported that it can prevent the addition of the transcriptional repressive mark H3K27me3. This preventing function is general to the genome as the covered chromatin areas do not necessarily contain its target genes [112]. Half of these repressive marks are catalyzed by the polycomb complex PRC2 [112]. As well, HNF4α and HNF4γ are involved in the control of chromatin accessibility as their loss induces chromatin accessibility shifts [116].

The polycomb group is an active actor of chromatin accessibility, mostly leading to gene repression through the deposition of H3K27me3 and H2AK119ub1. The polycomb group is implicated in early embryonic development and is composed of the groups PRC1 and PRC2. PRC1 is implicated in H2AK119ub1 histone modification and PRC2 in H3K27me3 modifications.

PRC1 repressive activity is implicated in ISC function and identity, notably by maintaining Wnt activity through the repression of Wnt-inhibiting genes and non ISC-specific-genes [149]. Even if H2AK119ub1 is lost when PRC1 is impaired, H3K27me3 remains unaffected, indicating an independent function of both marks and their associated polycomb complex.

In contrast, PRC2 seems not to be implicated in ISC renewal but rather in progenitor proliferation and differentiated lineage balance. In physiological context, PRC2 is dispensable for the stem-cell niche but essential for TA-cell proliferation and inhibition of the secretory lineage by the repression of Math1 and Gfi1 [150]. PRC2 H3K27me3 activity permit to determine enterocyte fate [151] and M cell fate [152]. M cell fate is determined through the final positive regulation by PRC2 of *Atoh8*, a newly identified actor for M cells differentiation [152], [153].

D. The Differentiated Cells, Critical Part of Intestine and Organism Physiology

Differentiated IECs have differential transcriptomic and proteomic programs correlated to their plethora of physiological functions. Explaining all their incredible functions and programs could cover too many tomes of literature. Instead, in this part, I will highlight some of their essential functions.

a. The Goblet cells, Protective Shield Against Lumen Components

Among the secretory cells, the goblet cells secrete mucus for coating, lubricating, and contributing to intestinal homeostasis. They protect the epithelium surface against mechanical, chemical, and biological attacks [154].

The Mucus is organized in a one- or two-layer systems with the inner layer in contact with ECs and the outer layer more diluted. They have different properties between small and large intestines. In the small intestine, the one-layer mucus has large pores to allow the penetration of large particles of nutrients or bacteria. Nevertheless, even if the mucus covers the crypts and villi, antibacterial signals secreted by Paneth cells and enterocytes keep bacteria in the mucus on the top of the villi. On the contrary, colonic mucus is organized in two layers. The inner layer is organized to block the access of bacteria, while pore size of the outer layer increases to allow the penetration of bacteria.

Goblet cells secrete two major protein families implicated in mucus structure: Mucins (MUC) and Trefoil factors (TFF). Among the **mucins**, the secreted mucin MUC2 is the main component of intestinal mucus and forms the mucus skeleton. Some mucins as MUC4 and MUC20 are membrane tethered. MUC4 has been reported to be involved in anti-adhesive protection mechanisms and physiological features like cell apoptosis modulation or cell differentiation and proliferation depending on the cell context [155].

The **TFF family** is important for gut mucosal protection, homeostasis, and "restitution", a process of rapid repair of epithelia. In the inflammation context, TFFs enters a "loop" with immune cells. Immune cells secrete cytokine signals, increasing TFFs expression, which in turn express pro-inflammatory signals to influence immune cells [156]. TFFs expressed in the intestine are critical for epithelial regeneration after extensive inflammation [157]. Little is known about the actual physiological function of MUC20, however, it may be implicated in signaling pathways like HGF/MET [158]. Its overexpression is associated with a poor outcome in colorectal cancer [159].

b. The Paneth Cells, Atypical Differentiated Cells for Support and Defense

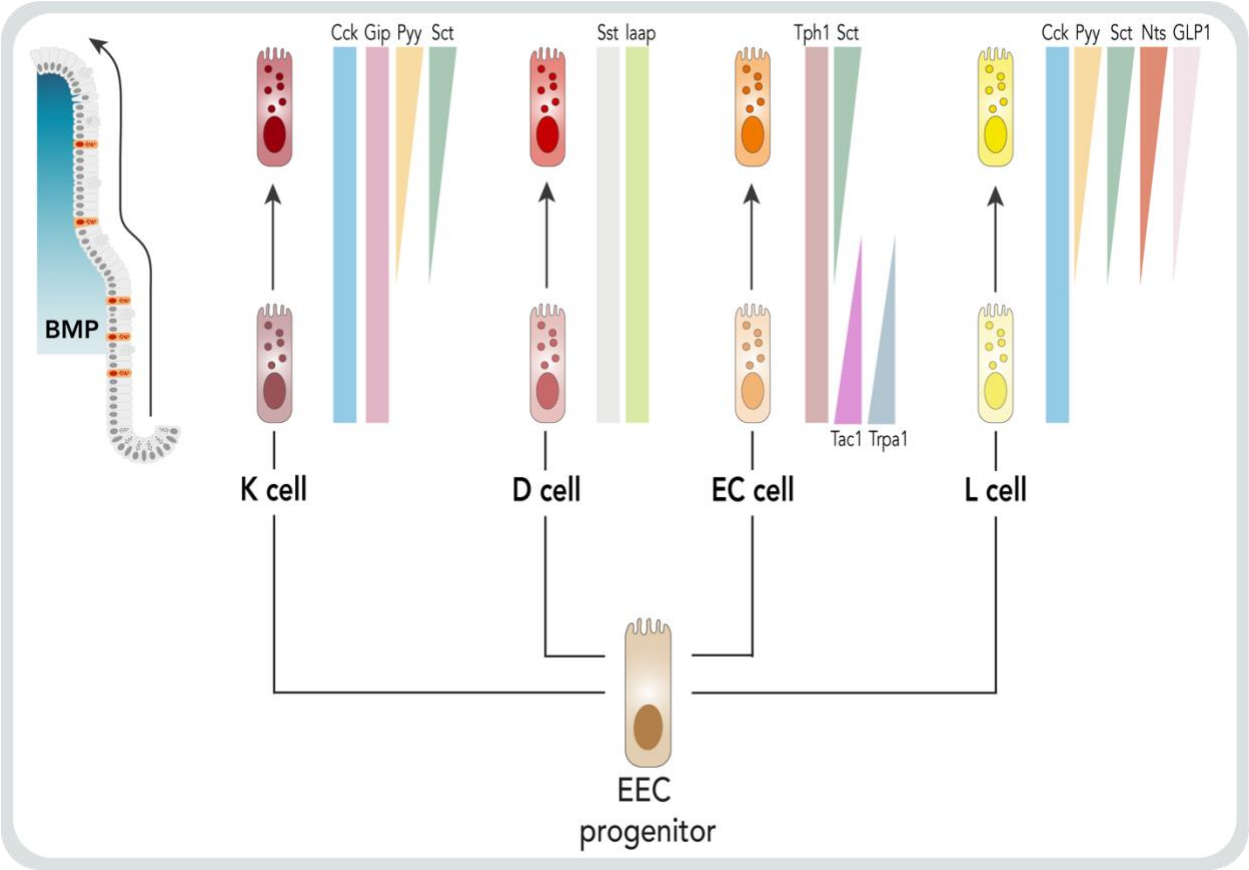
In contrast to other differentiated cells, Paneth cells move downwards as they mature in the stem-cell niche and live for 1-2 months. Limited to 5 to 15 cells per crypt and interspersed between ICSs, they are only present in the small intestine. They are in higher number in the ileal segment. They have both functions of IECs support and antibacterial secretion [161].

Paneth cells secrete trophic factors in order to support IECs and the neighbouring crypt base cells. These factors belong EGF, Notch and Wnt signaling pathways, promoting stem-cell maintenance and proliferation [161].

The small intestine mucosa is constantly in contact with bacteria and their products. To protect the epithelium, Paneth cells continuously secrete antimicrobial peptides to regulate the intestinal microbiota. These peptides, such as lysozyme, α-defensins (Defa), RegIII, matrix metalloproteinase 7 (Mmp-7), are contained in prominent granules which are secreted in the lumen [161]. Antimicrobial production is sensitive to stimuli. In case of pathogenic aggression, their production increases to prevent bacterial invasion into the crypts. Most of the antimicrobial peptides act by impairing bacterial membrane or cell wall integrity/synthesis [161]. Defensins are the major antimicrobial family only secreted by Paneth cells in mice. They are activated by Mmp-7 cleavage. In the context of α-defensin overexpression (Defa5) or deficiency (Mmp-7 depletion), mice present a great alteration in commensal microbiota balance of the small intestine [162]. Thus, Paneth cells are essential to the control of commensal microbiota composition in the small intestine.

c. The Enteroendocrine Cells, Intestinal Messengers

Enteroendocrine cells (EECs) represent a minority of cell populations (<1%). It is a heterogeneous lineage of hormone-producing lineage in response to ingested nutrients. The intestine is the largest hormone-producing organ, via the EECs that produce more than 30 different hormonal peptides. Their ability to secrete hormones opens the intestinal network to the internal body to regulate physiological processes related to metabolism, like insulin release and appetite.



Inspirate from [160]

Figure 35. Model of Enteroendocrine Cells differentiation and their Hormones Repertoires along the BMP signaling

Cck. Cholecystokinin**; Gip,** Glucose-dependent insulinotropic polypeptide; **Pyy,** peptide tyrosine tyrosine; **Sct,** secretin, **Sst,** somatostatin; **Tphr1,** transferrin receptor; **Tac1**, Tachykinin-1; **Trpa1,** Transient Receptor Potential Cation Channel Subfamily A Member 1, **Nts,** neurotensin.

Originally classed in function of their hormone-products, Beumer et al., proposed in 2018 a new simplified taxonomy of the EECs, based on single-cell analyses of EECs along the crypt-villus axis [163]. EECs change their hormonal profile as they migrate along the villus. These changes occur under the influence of the BMP signaling gradient, particularly that of BMP4, which is initiated by mesenchymal cells (Figure 35).

d. The Tuft Cells, a Poorly Understood Subtle Sensory Cell Population

Tuft cells constitute a small proportion of IECs, around 0,5% of the IECs. These cells are more prevalent in the distal part of both small and large intestines. Like enterocytes, tuft cells have microvilli on their apical part and are considered part of the brush border. They are morphologically characterized by a highly organized brush border forming a distinct "tufted" mass compared to other cells. Like enteroendocrine cells, they are considered as a sensory cell type that plays a key role in type 2 immunity to helminth infection. Tuft cells are important for the communication with immune-resident cells of the *lamina propria* by the expression of cytokines in inflammatory contexts [164].

Even if tuft cells arise from secretory progenitors, the classification of tuft cells as a secretory cell lineage is still under discussion. They are the only "secretory" cells that do not depend on *Math1* expression, nor the other major secretory transcription factors like Neurog3, Sox9 or Spdef. On the contrary, they depend on their own transcription factor Pou2f3. Recently, two-distinct populations of mature tuft cells have been described, both expressing the Doublecortin Like Kinase 1 (DCLK1), a protein of the tuft signature. However, the Tuft-1 cluster includes more neuronal development related genes while the Tuft-2 cluster is more enriched in immune-related genes [165].

Despite recent discoveries, tuft cell function is still poorly elucidated. Current studies reveal their potential implication in gastrointestinal pathogen clearing, infection resistance strengthening, mucosal healing after damages and immune response in allergy contexts [164].

e. The Enterocytes, the Intestinal Army for Nutrient Absorption and Metabolism

Enterocytes are the most abundant differentiated cells in the intestine. They represent 80–90% of all differentiated cells and are efficient in the absorption of ingested nutrients and water from the lumen. As seen before, most of the nutrient absorption occurs in the small intestine, while the colon is specialized in the absorption of reimagining water and electrolytes. Enterocytes are polarized. Their nucleus is kept at their basal pole and, their cell pole in contact with the *lamina propria*. At their apical pole, they present the brush border, a structure of

membrane folds forming the microvilli, that greatly enlarges the absorptive surface in contact with the lumen.

Enterocytes have a key role in **absorption and metabolism**. They are responsible for the absorption of lipids, carbohydrates, amino acids, and other nutrients like vitamins, electrolytes, and iron. Nutrient absorption is most often covered by transporters characterized and classified according to their targets and features. Many of these transporters are members of the huge solute carrier (SLC) group of 430 members organized into 66 subgroups [166].

Enzymatic Activities

The enterocytes have metabolic functions through the expression of digestive and metabolic enzymes. Enterocytes synthetize digestive enzymes in their microvilli of the brush border (BB), in contact with the luminal compartment They are responsible of the last luminal digestion prior absorption like the oligosaccharidases, or a protective activity like the case of alkaline phosphatase.

Secretion of **oligosaccharidases** allows the terminal hydrolysis of saccharide units from the digested food. There are three α-glucosidases and one β-glycosidase. **Maltase-glucoamylase** (MGAM) hydrolyzes the terminal units of α-glucose in two distinct sites, the maltase and glucoamylase sites. While the **sucrase–isomaltase** (SI), through its two subunits, catalyzes sucrose and isomaltase into fructose/glucose and maltase respectively. The **trehalase** (TREH) converts the disaccharide trehalose into glucose. Lastly, **Lactase–Phlorizin Hydrolase** (LPH) is the only β-glycosidase. It hydrolyzes lactose and cellobiose/cellotriose [167].

Alkaline phosphatase (ALK-P) is also known as alkaline phosphate intestinal in the intestinal context. ALK-P is a "swiss knife" enzyme essential for gut homeostasis. It is a secreted enzyme classically used to observe the enterocyte maturity. It has a protective role, as it regulates gut microbiota, promotes gut barrier integrity, and detoxifies bacterial-derived endotoxins which otherwise will induce gut inflammation. It is also considered as a regulator of lipid absorption. Indeed, it has been shown to enhance CD36 activity by dephosphorylating it [168]. Interestingly, ALK-P is implicated in some intestine internal body networks as for the gut-pancreas axis to mention just one. In the case of gut-pancreas axis, disruption of its physiological functions has been correlated to injury of the pancreatic β-cells, hyperinsulinemia, and glucose intolerance. In this axis, it has a protective role against harmful metabolic function like in the case of diabetes [168].

The Aminopeptidase N (ANPEP or CD13) and the dipeptidyl peptidase-4 (Dpp4) are both expressed at the membrane of enterocytes microvilli of the brush border. They are responsible

of the last hydrolysis of ingested proteins [169] and of the control of lipid metabolism [170] respectively. Additionally, DDP4 may be find in a secreted form and contrary to its general function in the organism, Dpp4 in enterocytes has no impact on the glucose metabolism [171].

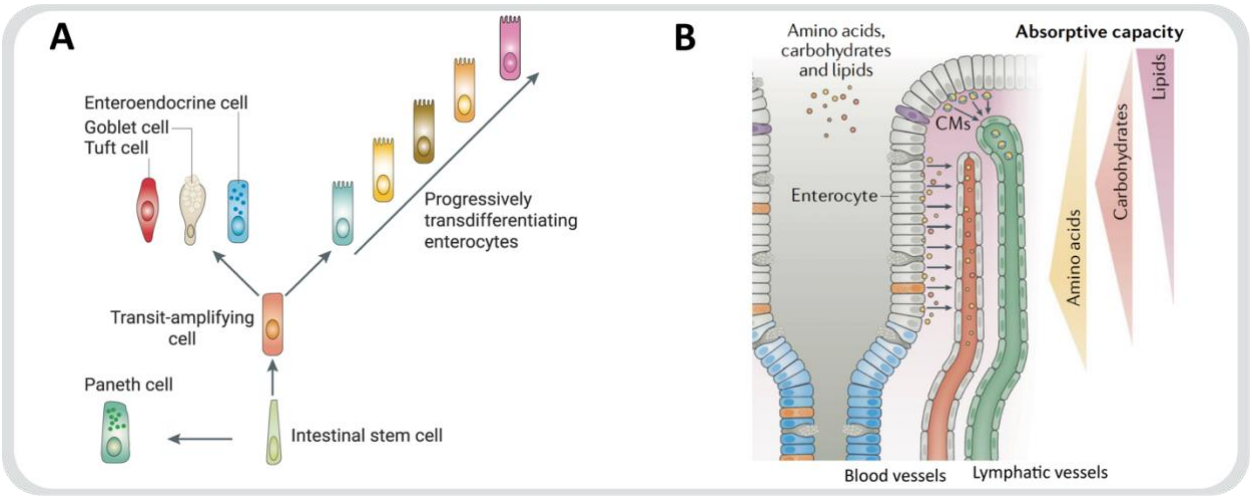
Besides the secreted enzymes in the BB, enterocytes have a plethora of non-secreted metabolic enzymes. Fructose-bisphosphate **aldolase** (Aldo) is one of the most characteristic enzymes families of enterocytes. This family, through its member AldoB (Aldolase B) [172], is preferentially localized in the nucleus and AldoA, found in both nucleus and cytoplasm [173], catalyzes cleavage of single fructose molecule.

Nutrient Absorption and Transport

Carbohydrate transporters are separated in two categories: the sodium-dependent glucose cotransporters (SGLTs), and the facilitative glucose transporters (GLUTs). The **SGLT family** uptakes and transports glucose from the apical cell membrane. Its members joined their action with the sodium. They actively transport glucose against the concentration gradient. This family implicates members of the SLC5a subgroup. On the other hand, **GLUT family** members use their carbohydrate-target gradient for their transport activity. In the small intestine, these members are known for fructose (GLUT2, GLUT5) and probably glucose transport (GLUT7, GLUT8, GLUT9, GLUT12) [174].

Amino acid transporters (AAT) mediate the transfer and transport of amino acids in or out the intracellular organelles. Generally, in the small intestine, AATs are sodium-dependent and members of the SLC1a, SLC6a, SLC7a, SLC16a, and SLC38a subgroups [174].

Hydrolyzed lipids are taken in by enterocytes, which packaged them into a structure called chylomicron [175]. These structures are re-secreted by the enterocytes and go through lymphatic vessels that transport them to the rest of the intern body. Lipids are taken by the enterocytes through two distinct ways. When the concentration of lipids is higher in the lumen, they diffuse passively into enterocytes. On the contrary, when the intracellular concentration of lipids is higher than in the lumen, intake of lipids requires the activity of ATP-depending transporters. Many proteins are implicated in transporting lipids such as the fatty acid-binding protein family (FABP) that seems to be implicated in the intracellular trafficking of lipids. FAPB2 directs the lipids to the pathways of triacylglycerol resynthesis [175].



Picture from [128], [176]

Figure 36. Model of Enterocyte Progression and Transdifferentiation and their Absorptive Capacity

A. Schematic Representation of the sequential state change of the enterocytes as they migrate along the villus **B.** Representation of enterocyte transdifferentiation consequences on their functional absorption specialization along the villus. Subsequentially, they specialized in amino acids, carbohydrates, and lipids absorption. Chylomicrons (CMs) are transported by the lymph node while other nutrients as sugars and amino acids are transported by blood vessels.

CMs, chylomicrons.

Enterocytes are also responsible for **vitamin absorption.** B9 Vitamin or folates cannot be synthesized *de novo*. It is the source of methyl group essential for methionine synthesis. Folates are essential for tissue physiology. Its deficit induces many symptoms as anemia, mental illness, and cancers. Proton-coupled folate transporter (PCFT) (also known as SLC46A1), is a folate-uptaking transporter at the brush border. This transporter is expressed along the gut, with its higher level in the duodenum and jejunum and its lowest in the colon [132], [177].

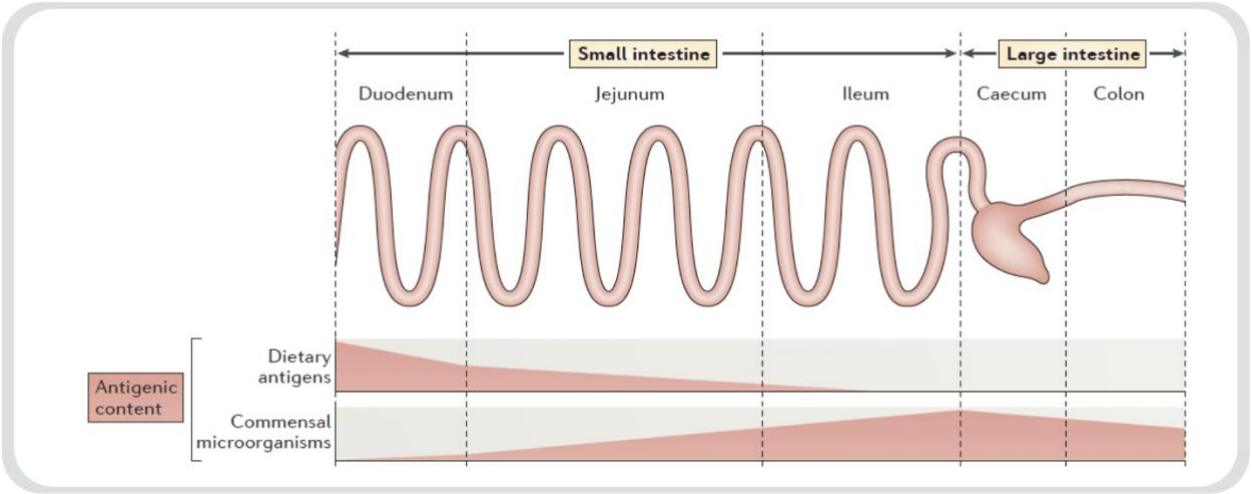
Other transporters are involved in specific features like PEPT1 (SLC15a1) or SLC46A3. PEPT1 is expressed in the duodenum, jejunum, and ileum. It has a flexible substrate pattern, as it can uptake more than 8000 different di- and tripeptides. PEPT1 is a great transporter for drug uptake [174], [178]. SLC46a3 is a transporter found in lysosomal membranes and is essential for transporting and delivering small molecules from the lysosomes to the cytoplasm. This transporter has been found to be essential for drug delivery into the cytoplasm and may be used as a predictive response marker for cancers and other diseases [179].

Enterocytes and "Zonated" Maturation

In 2018, Moor et al. [128] uncovered a surprising progressive trans-differentiation of enterocytes. Using a combination of microdissection and single-cell RNA-seq analyze, these authors found out that transcriptomic program of enterocytes is spatially heterogeneous and "zonated". Meaning that enterocytes do not mature into one step as terminally differentiated cells type but continuously differentiate as they migrate along the villus axis [128].

The absorptive function of a specific nutrients is compartmentalized distinctly in function of villus zones, potentially for a more efficient nutrient uptake. Enterocytes present subtle variations, witnessing their nutrient program. The engagement level of amino acids transporters and carbohydrates transporters decreases as the enterocytes migrate to the top of the villi, while the level of apolipoproteins increases. These observations indicate that bottom- and mid-villus enterocytes are specialized in absorption of amino acids and carbohydrates, whereas villi-top cells are engaged in lipid secretion, like chylomicrons (Figure 36) [128].

In addition to their nutrient absorption function, enterocytes provide a protective role. Next to the protective secreted enzymes like the ALK-P, enterocytes at the bottom of the villi help to protect crypt-resident cells. In addition to the guardian protective function of Paneth cells, they also express inflammasome components and secrete proteins belonging to the antibacterial Reg family, as Reg3a, Reg3b and Reg3d, considered to be enterocyte-specific in the homeostasis context [128].



Picture from [180]

Figure 37. **Repartition of the Global Microbiota Population along the Intestine**

f. The M Cells, Observation Tower of the Intestine

Microfold cells (M cells) are a unique subset of IECs absorptive lineage overhanging the Peyer's patches (PPs). PPs are lymphoid follicle structures found in the intestine, with higher numbers in the ileum. PPs are composed of the M cells that cover an aggregate form of the gut-associated lymphoid tissue (GALT). GALT is the largest peripherical immune tissue of the body. They keep an eye on the microbiota and the apparition of pathogenic microorganisms [181].

M cells are specialized in sampling luminal contents by phagocytosis. Then, they transport antigens to the underlying dendritic cells, macrophages, and lymphocytes to control the immune system [181].

II. Microbiota and Immune System, Players of the Intestinal Network

A. Dear Microbiota, Intestine's Lifetime Friend

The intestinal microbiome is a dynamic diversified community of bacteria, fungi, viruses, and yeasts. Bacteria compose a major part of the microbiome, regroup under the term microbiota. They have a symbiotic relationship with their host. The host-organism provides a habitat rich in nutrients and the intestinal microbiota supports the host through various physiological functions like metabolism functions and protection by preventing the implantation of pathogenic bacteria.

Acquired at birth, microbiota is an integral part of the organism that develops along the host life until its death. Commensal bacteria are a dynamic symbiotic community distributed along the intestine in function of segment features. Factors such as lifestyle and nutrition change influence the microbiota composition, resulting in disturbance of microbial metabolites and products.

a. Microbiota Landscape Along the Gut

Microbiota repartition along the intestine is heterogeneous (Figure 37). Localization, bacteria family, and density are hypothesized to vary due to factors like chemical gradient (pH), oxygen level, nutrient availability, and immunity effectors.

In the small intestine, the flow transit is faster with metabolism like simple sugars and amino acids. The duodenum microenvironment has the fastest flow with bile acids, pancreatic secretions, and high oxygen concentration, limiting bacterial density and diversity. As the flow progresses, oxygen quantity decreases. The jejunum supports colonization of Gram-positive

aerobic and facultative anaerobic bacteria, while the Ileum supports a predominance of aerobic species with a population like what can be found in the proximal large intestine in its most close part to the caecum. The large intestine has a very slow flow, anaerobic environment, and fermentation metabolism of complex polysaccharides from undigested plant fibers from ingested-plant or host mucus structure. This environment promotes anaerobic bacteria [182].

b. Metabolic Functions of the Microbiota

Microbiota is essential to metabolize elements undigested by the host. Many bacteria families are specialized in element metabolization. Meaning that bacteria species are distributed in function of their metabolism specificity and diet influence the composition of microbiota, correlated in some case to increase risk of diseases [182].

Undigested dietary fibers are fermented by microbiota, resulting in gases (methane, hydrogen, carbon dioxide), short-chain fatty acids (SCFAs), alcohols (methanol and ethanol) and small organic acids like lactate and succinate. SCFAs are more prevalent in the caecum, which is the major center of fiber metabolism, then 90-95% of them are absorbed in the colon. SCFAs serve as an energy source and have other features. For example, acetates passes into the blood to serve as an energy source for peripheral tissues, lipogenesis, and cholesterol synthesis in the liver; butyrates that serves as an energy source for colonic epithelial cells [182].

Microbiota break down undigested proteins into peptides and amino acids but also other metabolites by secreting proteases and peptidases. Among the metabolites, we can find some neuroactive compounds and other metabolites that may increase risk of pathology like cancers and chronic inflammation [182].

c. Microbiota and the Intestinal Epithelium Cells

As introduced before, IECs and microbiota communicate and influence each other. This part will illustrate some of the impact of communication between IECs and microbiota.

Following ALK-P and AMPs production by enterocytes, the Paneth cells produce AMPs to prevent bacterial invasion and contact with IECs. Microbiota and the epithelium are also involved in the general intestinal homeostasis. The interaction between bacteria and IECs has been proved to be a key regulator of intestinal permeability. Bacteria regulate the tight junction expression, assembly and trans-epithelial permeability [183].

Microbiota may directly influence epithelial genes transcription. HNF4α activity and transcription network are suppressed by microbiota through histone modifications which

impact chromatin accessibility [184]. Microbiota is also able to affect binding sites of transcription factors like suppression of enhancer sites for GATA4 [184]. EECs secrete peptides to communicate with different systems of the inner body like nervous system, pancreas, liver, and adipose tissue. They are crucial to regulate the energy homeostasis and hormone secretion in function of the diet and its downstream metabolic processes. As the microbiota vary in function of the diet, it is not surprising to find that microbiota influences EECs. Production of SCFAs by microbiota from undigested food influence EECs hormones production associated with satiety and lipid metabolism. Propionate promotes the secretion of satiety hormones PYY and GLP-1 while acetate stimulates reduces appetite and sensitive insulin hormone [185].

d. Dysbiosis, Consequence of an Unbalanced Microbiota

The Intestine microbiota is implicated in many physiological processes and host health. To mention a few examples, intestinal microbiota composition is associated with host's mental health (anxiety, depression) [186], suspected in glucose metabolism of diabetes type-II [187] and diabetes type-I [188].

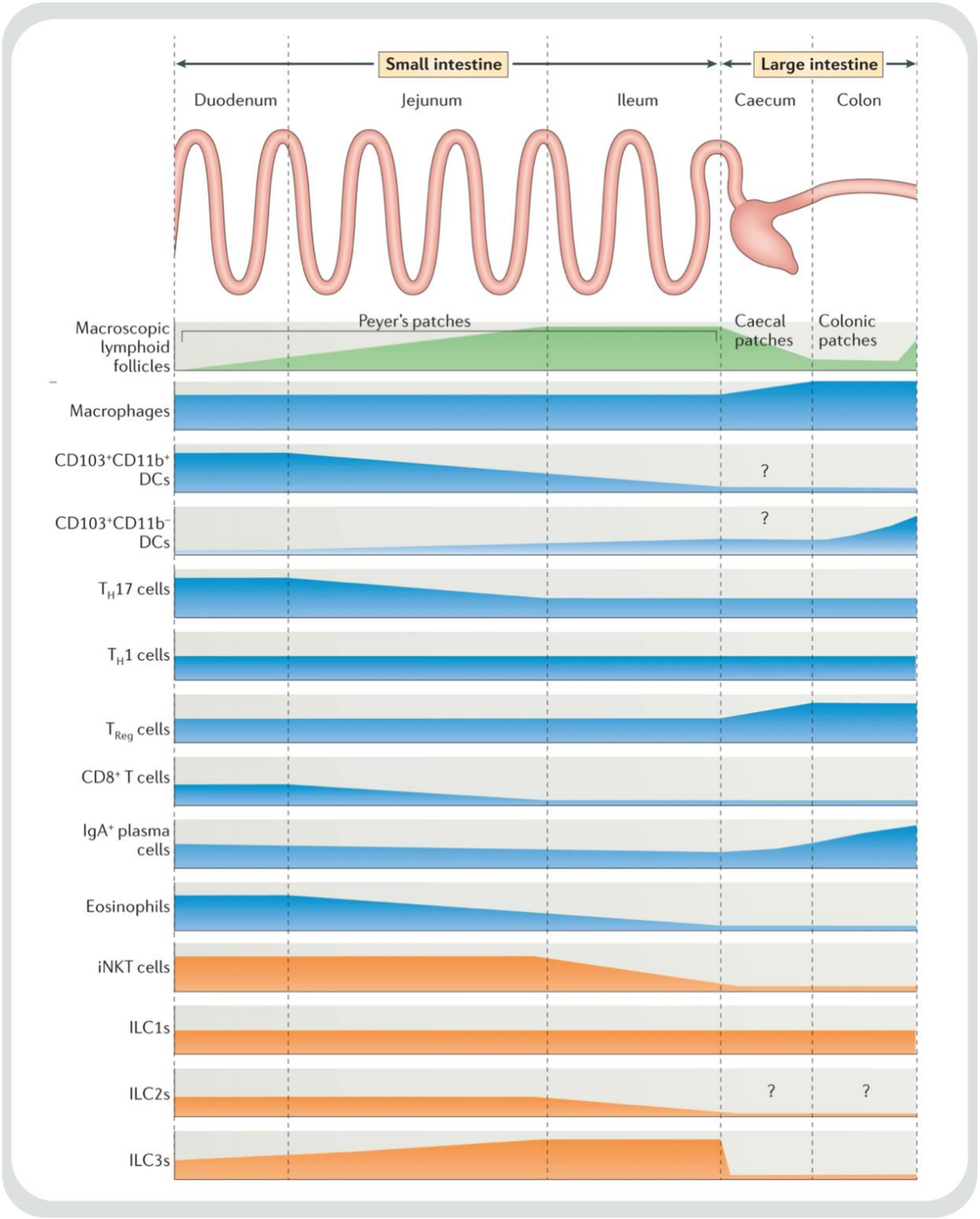
Unbalance of the microbiota is called dysbiosis which is characterized by an alteration of microbiota and homeostasis. Dysbiosis disorganizes the tight junctions and hence increases the gut permeability, and leading to inflammation, insulin resistance or also fatty liver [185].

Many factors like high-fat diet may induce a dysbiosis by favoring some bacteria families, impacting host's health through metabolite production. For example, choline is a metabolite predominantly obtained by bacteria metabolization of diet rich in red meat and eggs. High concentration of choline is associated with risk of cardiovascular diseases [189].

Alterations of IECs like Paneth cells are also another factors of dysbiosis. As Paneth cells secrete AMPs controlling microbiota homeostasis, dysfunction of these cells unsurprisingly leads to dysbiosis [190].

B. Immune System, Guardian of the Intestine

The intestine is permanently exposed to a variety of antigens and stimuli coming from food, microbiota, and pathogens. As a result, the intestine has a unique and fine-tuned immune system to unsure monitoring and appropriate responses to stimuli. The challenge is to maintain a tolerance to harmless antigen and dealing the pathogens with the innate and adaptative immune system.



Picture from [180]

Figure 38. Distribution of the Immune Stromal Cells along the Intestine

Representation of the regionalization of immune cells along the gut. Question marks represent the regions which are not yet characterized.

DC, dendritic cell; **ILC**, innate lymphoid cell; **iNKT**, invariant natural killer T; **T_H,** T helper.

a. Immune organization in the Intestine

The lamina propria, epithelium and gut-associated lymphoid tissue (GALT) like Peyer's patches (PPs), are great sites for the intestinal immune system and distributed differently in function of segments (Figure 38).

The *lamina propria* contains B cells, T cells, and numerous innate immune cells like macrophages and dendritic cells (DCs). On the contrary, the epithelium mostly contains T cells. Taken together, the epithelium and *lamina propria* of the intestine are the largest reserve of T cells and macrophages of the organism [180].

Intestinal macrophages are the most abundant leukocytes in the intestinal *lamina propria*. Macrophages are essential for phagocytosis and degradation of microorganisms and dead cells. They also sample the lumen content, promote epithelial barrier integrity and produce a large amount of IL-10 to prevent inflammation [191].

Intestinal DCs are antigen presenting cells that constantly sample the intestinal contents and are associated with M cells in PPs. DCs mediate the tolerance to food antigens and gut microbiota. DCs population is heterogeneous and regulates T cell response [192].

In the intestine, both **CD4 and CD8 T cells** are found in the *lamina propria*. CD4 cells differentiate into different subsets of T helper and T regulator cells in function of stimuli in the intestinal segments. T cells mediate tolerance for food antigen and commensal microbiota, protect the organism by clearing infected cells and modulate immune responses to infection [193].

Intestinal B cells are present in GALT and *lamina propria*. GALT contains a various subsets of B cells that are surrounded by other cells like DCs, T follicular helper cells and macrophages. The *lamina propria* contains the highest density of plasma B cells of the organism. IgA expressing plasma cells are the predominant subset of plasma cells. The proportion of IgA plasma cells increases progressively along the intestine length. B cells are essential for immune response and control of microbiota [194].

Lastly, **innate lymphoid cells** (ILCs) are a novel described resident immune cell population. The family seems to act as an intestinal immunity coordinator, maintains intestinal homeostasis, and barrier integrity [180].

b. Immune Cells and Intestinal Epithelium cells

Some IECs can transport microbial antigens and metabolites to the immune cells to modulate the immune response. M cells that cover Peyer's patches are specialized in antigen internalization and presentation to resident immune cells. Goblet cells present antigen to the dendritic cells of the *lamina propria* by formation of goblet cell-associated antigens passages carrying antigens samples. Finally, enterocytes also participate in antigen presentation by presenting lipid antigens to the natural killer T cells.

In turn, immune cells influence IECs. In response to microbial metabolites, ILCs secrete cytokines like the cytokine interleukin 22 (IL-22) to regulate IECs. IL-22 promotes IECs, homeostasis, repair, and production of AMPs to control growth of commensal bacteria. Many other cytokines and immune cells have essential function in IECs homeostasis, like promoting differentiation into tuft and goblet cells, and reinforcing tight junctions [195].

c. Regulation of the Microbiota by Immune Cells

Intestinal immune cells control directly or indirectly the microbiota. Innate immune cells are part of microbiota regulation. These cells keep an eye on the microbiota through their germline-encoded pattern recognition receptors (PRRs). These receptors recognize microbial-associated molecular patterns (MAMP), molecules typical of a microbial signature like lipoproteins, carbohydrates, and nucleic acids. These PRRs are present on some IECs and immune cells like DCs and macrophages of the *lamina propria*. This sensing surveillance is essential to control immunity and, recognition of pathogen MAMP, leading to activation of cellular defense mechanisms like production of AMP, secretion of pro-inflammatory cytokines and recruitment of immunes cells [196]. Several innate immune cells are able to produce IL-22, promoting the expression and secretion of AMPs by IECs. Secretion of IL-22 is essential to keep microbiota homeostasis. Absence of IL-22 leads to alteration of microbiota composition, which increases susceptibility to colitis inflammation. In the homeostasis context, IL-22 is produced by DCs and ILCs [180].

Adaptive immune system is not left behind. One of the crucial ways of microbiota homeostasis regulation is coating of the mucus by immunoglobulin A. B cells of the intestinal immune system secrete IgA into the mucus. This antibody joins its force with AMPs in the mucus to keep the bacteria in touch. It is also able to target specific bacteria and bacteria function like flagella production to counteract pathogenic situation. A subset of T Cells, the T follicular helper cells, promotes IgA selection and is associated with microbiota control through the inhibitory receptor programmed cell death protein 1 (PD1). Deficiency of PD1 reduces bacteria coated with IgA, resulting in a decrease of the commensal *Bifidobacterium* species and an increase of pathogenic *Enterobacteriaceae*. Invariant natural killer T (iNKT) cells are also known to be implicated in microbiota composition in homeostasis. Furthermore, after injury, a subset of intraepithelial lymphocytes T cells secret a plethora of cytokines and AMPs to prevent dissemination and invasion of bacteria in the mucosa [180].

III. Consequences of an Unbalanced Intestinal Network

Unbalance of the intestinal network leads to increased risks of pathology and diseases like inflammatory bowel diseases and colorectal cancers.

A. Unbalance of Microbiota and Immune System, a Story of Inflammatory Bowel Diseases

a. IBD Disorder and Epidemiology

Inflammatory bowel diseases (IBD) are chronic inflammation relapsing-remitting disorder of the intestinal tract. IBDs include ulcerative colitis (UC) and Crohn's disease (CD). UC involves inflammation and ulcers in the colon and rectum part, whereas CD is distinguished by inflammation of different areas of the intestine, in random and discontinuous spots. They are characterized by intestinal lesions and barrier disruptions, causing abdominal pain, severe diarrheal, and tiredness. Nowadays, IBD increases the risk of colorectal cancer and there are no curative treatments for IBD but treatments of the symptoms which may lead to long-term remission.

Epidemiological analyses show that the overall incidence and prevalence of IBD slightly continue to increase over the time and vary between regions. In 2017 the global level of IBD prevalence was 84.3 per 100,000 persons, while it was 136.6 per 100,000 in Western Europe with the second-highest prevalence in the United Kingdom (449.6 per 100,000) [197]. Today, it is estimated that 0.3% of the European population are diagnosed with IBD, corresponding to 2.5-3 million patients [198]. Causes of IBD are still not well understood. However, the current hypothesis supports that it is a combination of genetic predispositions, host immune system and environmental factors.

b. Dysbiosis, Immune Cells and Genetic Factors, Actors of IBD pathogenicity

Other studies recognized that IBDs are associated with a combination of diet, genetic predisposition, dysbiosis and immune dysregulations. IBD are characterized by a dysfunctional intestinal barrier, which increases the intestine permeability, and exposes lumen contents to immune cells of the *lamina propria*. Furthermore, the clear unbalance of the intestinal microbiota-immune network is part of IBD pathogenicity with an aberrant mucosal immune response towards the native microbiota [199].

IBDs present a clear dysbiosis with a decrease of microbiota diversity and an increase of *Enterobacteriaceae*. Various studies suggested that microbiota dysbiosis is a key factor of IBD pathology However, whether intestinal dysbiosis is a cause or consequences of IBDs is unclear and still under debate [199]. Still, as intestinal dysbiosis involvement has been highlighted in IBDs pathogenesis it can play an obvious role in host health, host mental health, and metabolic disorders. At least, it opens the way to future treatments. Change of microbiota composition ultimately leads to change of microbiota-associated functions like production of metabolites issue from fermentation and imbalance of immune cells, which are suspected to play a role in the IBD pathogenesis. Notably, there is a decrease in species producing butyrate, a SCFA associated with intestinal homeostasis regulation and inflammation reduction.

Therefore, IBD have a high number of genetic factors improving the risk by inefficient microbiota handling through defects of recognition, killing function mucosa barrier integrity and immune regulation. NOD2 is one of the genetic factors identified in IBD. NOD2 is an intracellular PPR essential to recognize the peptidoglycan as a component of bacterial cell wall, implicated in bacteria handling and regulation [199].

B. Colorectal Cancer, a Pathogenic Intestinal Network Loop

a. Colorectal Cancer, Multifactorial Heath Burden

Colorectal cancer (CRC) is currently one of the most common cancers in the world. In 2018, CRC was ranked as third place in terms of incidence and second in terms of mortality. CRC concerns 1.8 million new cases and 880,000 deaths per year [200]. CRC is a major health burden that is unlikely to decrease, as projections to 2030 indicate an increase of incidence to 2.2 million of new cases and 1.1 million deaths [201]. CRC is a disease of complex origin, involving genetic and environmental factors. Despite the strong influence of hereditary components, most of the CRCs are sporadic and not inherited. Factors such as diet, smoking, obesity, diabetes, alcohol abuse and IBD, heavily increase the risk of CRC.

CRC develops slowly and symptoms appear after it has reached a considerable size which can block the feces, causing pain, cramping, and bleeding. CRC then leads to loss of appetite and iron deficiency anemia. About 40% of the CRCs take place in the proximal colon, 22% in the distal part and 28% in the rectum. Early stages of CRCs are often diagnosed during routine colonoscopies and about 20 % of them are already associated with metastasis.

b. Gene-Expression Disorders, Key of CRC pathogenicity

CRCs implicate alterations of many pathways of the intestinal network. Currently, it is assumed that most of the CRCs take their origin in stem-cells or stem-cell like in the colonic crypts. Most of these cells become cancerogenic through the cascade and accumulation of genetic and epigenetic alterations, activating oncogenes and inactivating tumor-suppressor genes.

The most frequently altered gene in CRC is the Adenomatous Polyposis Coli (*APC*) gene, with 85% of sporadic CRCs showing an APC inactivation/decrease. Loss of APC takes off the inhibition of the Wnt/β-catenin pathway and affects chromosome segregation during cell division. Consequently, this induces a genes activation/inactivation cascade. Many other genes are altered in CRCs, as well as the critical transcription factors Cdx2 and HNF4α. In CRCs, about 20-30% of cases show downregulation of the intestinal mastermind Cdx2, which is associated to a very poor outcome [202]. HNF4α, another transcription factor deeply implicated in intestinal homeostasis has a controversial role in tumorigenesis. To fully grasp its controversial role, we need to examine its role in more detail than before. *HNF4α* is transcribed from two differential promoters named P1 and P2 that will give rise to at least nine variants

(HNF4α -1 to -9). These two promoters are differentially activated depending on the state of the cell. P1 is active in differentiation functions of IECs, while P2 has proliferation function. The controversial role of HNF4α originates from these promoters. P1-HNF4α appears to have a protective role against colitis and acts as tumor suppressor in CRCs, in contrast to P2-HNF4α which does not have the same protective power [203].

Epigenetic modifications have been recognized as fundamental mechanisms of CRCs initiation and progression. Alteration of the histone modification pattern, such as histone methylation, is a pivotal feature of chromosome accessibility and genes regulation. Upregulation of H3K27me3 is linked to a better survival, on the contrary to the upregulation of H3K37me2 which is associated to a poor survival. On the other hand, decrease of both H3K9me3 and H4K20me3 is associated to a shorter survival and high tumor recurrence [160].

Polycomb group, as described above, is a regulator of histone modifications that plays an important role in CRCs. Through its CBX components, variants of canonical PRC1 are involved in CRCs. CBX2 is overexpressed in CRCs, in contrast to CBX7 which is downregulated, but both are associated with poor prognosis [33], [204], [205]. Furthermore, dysregulation of CBX members in CRCs is correlated to differential infiltration of immune cell in the tumor [206]. PRC2 is involved in CRCs as well. Its core member EZH2 has been reported to play a role in IBD and in the onset of the subsequent CRCs [207]. Therefore, EZH2 is overexpressed in CRCs and may contribute to the development of CRC by promoting cell proliferation through epigenetic alterations, which leads to the downregulation of p21, a cell cycle inhibitor [208].

c. The Infinity Loop of Colorectal Cancer Pathogenic Intestinal Network

Like most solid tumors, CRCs are complex tumors included in a complex microenvironment. CRCs microenvironment involves all the *lamina propria* elements, the immune cells and the microbiota of the intestinal network, that work together in an activation and stimulation infinity loop.

Intestinal mesenchyme cells (IMCs) in the *lamina propria* interact with the IECs and cancer cells, contributing to the tumorigenesis. In homeostasis context, fibroblasts are nearly quiescent mesenchyme cells [209]. However, they may be activated by many factors like inflammation or mechanical stress turning them into myofibroblast. Activation of the myofibroblasts changes their morphology and metabolism according to their new functions of wound healing and tissue remodeling. Then myofibroblasts decrease by apoptosis or return to their quiescent state. Failure of these diminution leads to pathology such as chronic inflammation, a risk factor of cancer. Also, they represent a major part of pathologic activated fibroblasts identified as cancer-associated fibroblast (CAFs). CAFs can be detected early during tumor development and are associated to poor prognosis. They contribute to tumor initiation, progression, metastasis, and inflammation. Additionally, activated IMCs are activated by inflammatory signals from tumor cells and tumoral immune cells. IMCs regulate proliferation,

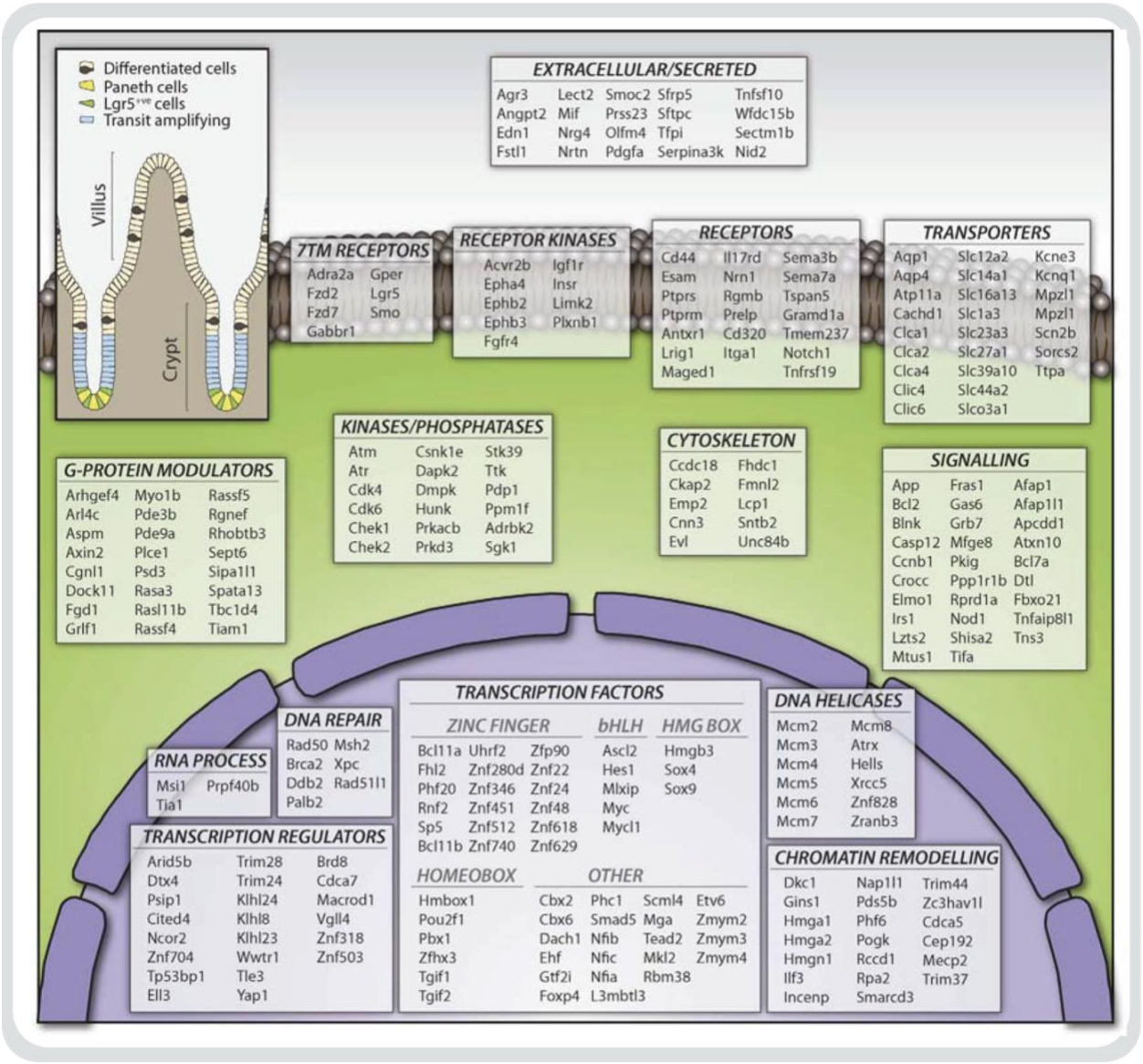
survival, and differentiation of cancer cells and their neighbouring IECs, promoting angiogenesis and tumor vascularization. Both CAFs and IMCs modulate immune response and inflammation [209].

Immune networks may be both of good and bad prognosis, depending on the immune cell types infiltrating the tumor. T cells have an important anti-tumor function, so infiltration by CD8+ and CD4+ cells are associated to a good prognosis. They inhibit metastasis and are inversely correlated to tumor recurrence. T cells are activated by DCs, which have a better prognosis when they have infiltrated the tumors. However, they have been also associated to a bad prognosis. In the tumor microenvironment that is enriched in immunosuppressive cytokines (secreted by IMCs, CAFs, Immune cells, IECs), they can mature into bad immune actors. Then, they secrete immunosuppressive cytokines, which inactivate the infiltrated NK (natural killer). Many other immune cells have been linked to a poor prognosis, like tumor-associated macrophages and the multiple origin myeloid-derived suppressor cells. Generally, these friendly tumor immune cells, secrete immunosuppressive and inflammatory cytokines affecting mesenchymal cells and microbiota [210].

Numerous studies showed that the **intestinal microbiota** is associated to CRC pathogenicity. Firstly, many CRCs arise from IBD, where the dysbiosis plays a major role. Following the evolution of IBD to CRC, it is unsurprising that the dysbiosis continues to be involved in tumor progression. Indeed, in all-cause CRC, patients present a dysbiosis with a reduced diversity and richness of commensal bacteria. Today, many bacterial species have been identified as risk factors for CRCs and contributors of the tumor progression, for example *Fusobacterium nucleatum*, pro-inflammatory bacteria associated with a worsen outcome. *Enterococcus faecalis* that induces DNA damage in IECs, is enriched in CRC patients. Although, dysbiosis is associated to an unbalanced production of beneficial metabolites along the generation of pro-carcinogenesis microbial products [211]. Last important fact is that inflammation cannot develop CRCs without germ-bacteria [212]. Generally, microbiota produces genotoxins causing DNA damages and inflammatory signals affecting immune cells and IMCs, unfortunately alimenting the loop of intestinal pathogenic network.

CHAPTER 4

Backstage of the Incredible Intestinal Turn-Over



Picture taken from [6]

Figure 39. Genes of the Intestinal Stem-Cell Signature and Classification by their Molecular Function

Intestinal homeostasis is dependent on the continuous supply of the intestinal epithelium cell types. The intestinal epithelium is characterized by a rapid cell turnover resulting from a continuous supply from the dividing stem-cells. Inside the crypt, the dividing stem-cells give rise to progenitors that are pushed outside the crypt and mature into differentiated cell types. The stem cells reside at the bottom of the crypt in the intestinal stem-cell niche, a fine monitored and protected environment.

I. Intestinal Stem-Cell Identity and Life

A. The Intestinal Stem Cells, who are they?

The intestinal stem-cells (ISCs) are located in the bottom of the crypt. The currently defined stemness cells of the intestinal epithelium include the crypt columnar cells and the cells at the +4 position.

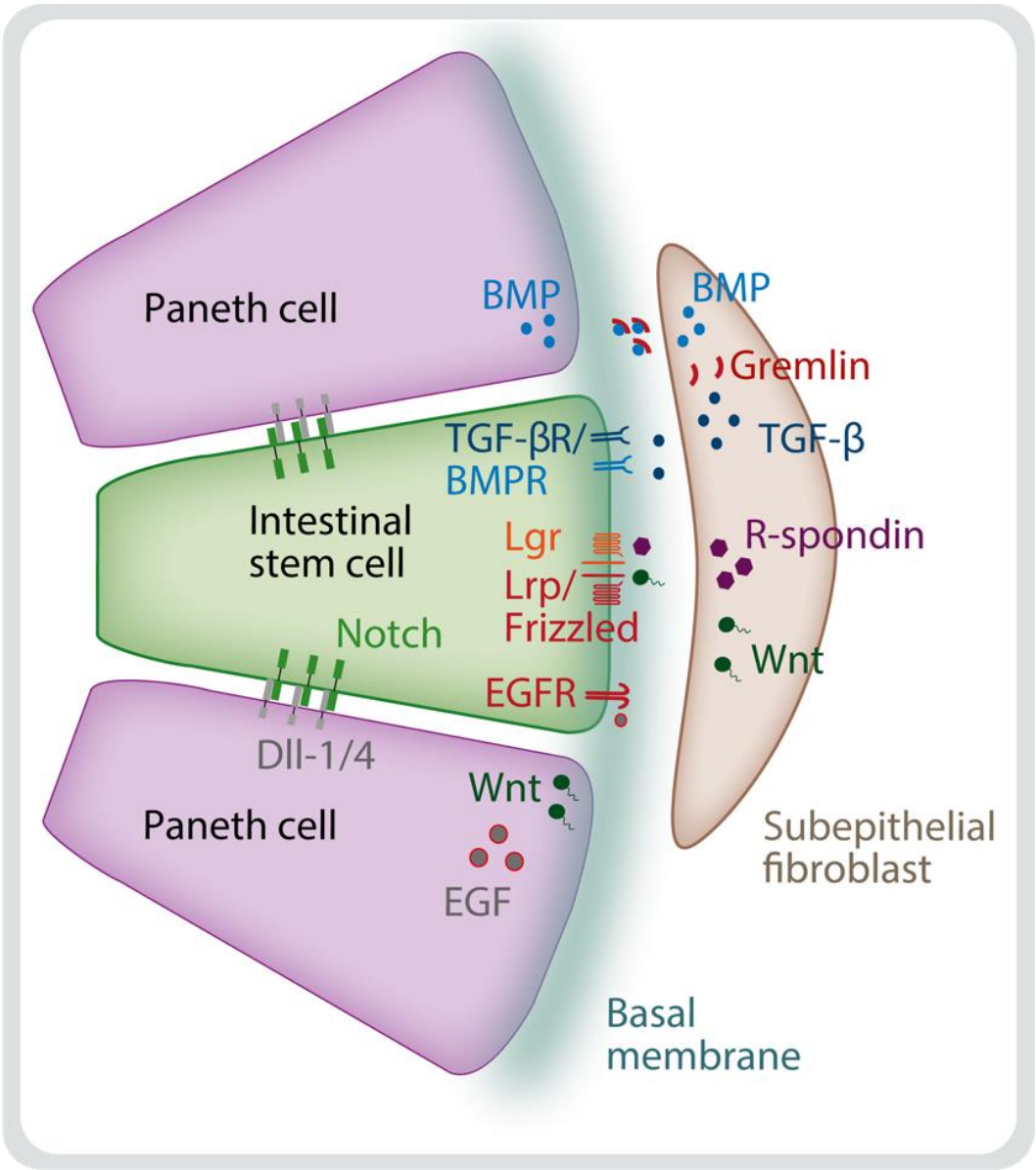
a. The Crypt Columnar Cells, the Continuous Proliferating Stemness Cells

The general concept of stem cells defines them as cells with self-renewal, multipotencies and differentiation capacity. The continuously dividing cells have been observed and described at the bottom of the crypt at the position +1 and +3, sandwiched between Paneth cells. These cells have been identified as ISCs and named crypt base columnar (CBC) cells [6]. They are characterized by a high specific expression of the protein "Leucine Rich Repeat Containing G Protein-Coupled Receptor 5" (LGR5). *Lgr5* is a target of Wnt pathways and encodes a R-spondin receptor. LGR5+ CBC cells are stem cells with a long-term self-renewal capacity and ability to differentiate into multiple intestinal cell lineages. When CBCs are passively pushed outside the stem-cell niche into the progenitor area, they lose their stemness and start to differentiate [213].

In 2012, Muñoz et al.,[6] used a combination of FACS-sorted cells, transcriptomic and proteomic profiling to analyse LgR5+ ISCs within their direct daughters. Those profiling reveal 384 genes of unique stem cell signature and 278 proteins enriched in ISCs. Correlation between transcriptomic and proteomic data permits defining 510 genes of LGR5+ CBCs cells. Genes of CBC signature are named and functionally classified in the Figure 39 [6].

b. The +4 Cells, Stem-Cells or not Stem-Cells, that is the Question

The name "+4 cells", refers to the cell position +4 from the bottom of the crypt. This position is localized just above the most upstream Paneth cell, at the limit between stem-cell and progenitor area. These +4 cells in the crypt seem to be another stem-cell type in a quiescent state. Their contribution in the intestinal homeostasis is still debated.



Picture taken from [214]

Figure 40. **Overview of the signaling Pathways Involved in Intestinal Stem-Cell Life**

Indeed, at the time of their discovery, they matched the stem-cell characteristic dogma at that time. Studies have defined the genes *Bmi1* (also known as *PCGF4*), *Tert1*, *Hopx* and *Lrig1* as specific markers of +4 cells. Cells expressing these markers are able to reconstitute the intestinal epithelium from the CBC cells to the differentiated cells. However, evolution of the technologies allows to show

that the defined +4 cells specific markers are not only expressed in +4 cells but also in LgR5+ CBC cells. Taken together, these studies and discoveries fueled the debate of the true nature and homeostasis contribution of these cells. Therefore, +4 cells are now enough documented to be considered as reserve stem-cells which replenished the CBCs cells when necessary. They have a high resistance to DNA damage, refractory to Wnt stimulations [213], [215], [216].

Compared to the CBC cells, +4 cells do not express target genes of canonical Wnt pathway like *LgR5*, *Ascl2* or *Ccnd1*, while they express higher expression of genes associated to cell cycle inhibitors, as *Cdkn1a*, is observed [213], [215], [216].

B. Supports of the Intestinal Stem Cells

The microenvironment of the stem-cell niche is essential for stem-cell maintenance, proliferation, and protection. As introduced in the previous chapter, the stem-cell niche microenvironment involves Paneth cells, intestinal mesenchyme cells (IMCs), immune cells, and microbiota.

a. The Crypt Signaling Pathways and Intestinal Stem Cell Life

Signaling landscape is set up by the Wnt, Notch, BMP, and Hippo paracrine signals providing by mesenchyme cells and Paneth cells. These pathways mediate ISC maintenance, proliferation, and differentiation (Figure 40) [213].

Wnt ligands activate the Wnt canonical pathways that are critical for ISCs maintenance, and more generally, in the crypt homeostasis. Binding of Wnt ligand on the ISCs co-receptor complex composed by Frizzled (FZD) and Low-Density Lipoprotein Receptor 5/6 (LRP5/6), induces nucleus translocation of **β-catenin**. The transcriptional coactivator β-catenin has many target-genes essential for ISC maintenance such as *LGR5, Ascl2* or *Ccnd1*. In the case of Wnt ligands absence, β-catenin accumulates into the cytoplasm and are addressed to degradation by adenomatous polyposis coli protein (APC). Wnt ligands are expressed by Paneth cells and IMCs surrounding the crypt. Due to the redundancy of Wnt ligand production, Paneth-cells

can be depleted *in vivo* without any phenotype. Wnt pathway is restricted to the lower crypt cells and forms a signaling gradient along the crypt [213].

Notch signaling is crucial for Stem-cell maintenance and proliferation. Notch signaling is dependent on cell-cell contact, indeed cells producing Notch-ligand present it to adjacent responsive cells expressing a Notch receptor. Paneth-cells express and present the Notch ligands Delta-like 1 and 4 (Dll -1/-4) to the adjacent CBCs, which express Notch1 receptor. Activation of Notch signaling on receptive cells induces the expression of Hes1 and Olfactomedin-4 (Olfm4) [213]. Olfm4 is a negative regulator of Wnt/ β-catenin signaling in ISCs and progenitor cells which may have anti-apoptotic activity and a protective role in IBD [217].

EGF ligands are produced by Paneth cells and CBCs highly express the EGF receptor ERBB1 along with their negative regulator LRIG1. LRIG1 controls the EGF signaling as its depletion induces uncontrolled EGF signaling, increasing proliferation and drastic crypt expression. Even if the link between ERBB1 and LRIG1 have been clearly established, mechanisms underlying this duality are not yet entirely understood. EGF signaling controls ISCs and is non-essential for their identity [213].

b. Complex Interaction between Immune Cells and Homeostasis Intestinal Stem Cells

As introduced in the last chapter, immune cells of the *lamina propria* are part of the intestinal network in homeostasis and in the context of injuries. Moreover, they are more deeply implicated in intestinal homeostasis as presented, recent studies highlighted their critical implication in ISCs balanced homeostasis. They may act in ISCs homeostasis directly or indirectly by acting on other cell types essential to ISCs.

Thus, T cells contribute to ISCs homeostasis through cytokines secretion. Treg promote ISCs renewal by secreting IL-10, in contrast to Th17 that secrete IL-17 and promotes ISCs differentiation and reduces ISCs renewal. Both Th1 and Th2 have a similar effect as they both secrete cytokines suppressing ISCs renewal and directing specifically ISCs differentiation into respectively Paneth cells and Tuft cells [218].

The newly identified resident innate lymphoid cells (ILCs) are also implicated in stem cell homeostasis. ILC subgroup secreting IL-22 reduces Wnt and Notch signaling, promotes Paneth-cells and progenitors proliferation, unlike to their impact on LgR5+ CBC population that decreases by apoptosis [219].

The macrophage subgroup expressing Colony-Stimulating Factor 1 Receptor (CSF1R) is restricted to the crypt and closely associated with ISCs. Recently, this group has been shown to be an important actor of ISCs homeostasis and Paneth-cell differentiation. Their ablation led to a decrease of LGR5+ CBC cells and affect Paneth-cell differentiation [220].

Even if some studies highlighted the implication of resident immune cells of the *lamina propria* in IEC homeostasis, Intestine-Immune Network is only scratched on the surface and seems far more complex than estimated.

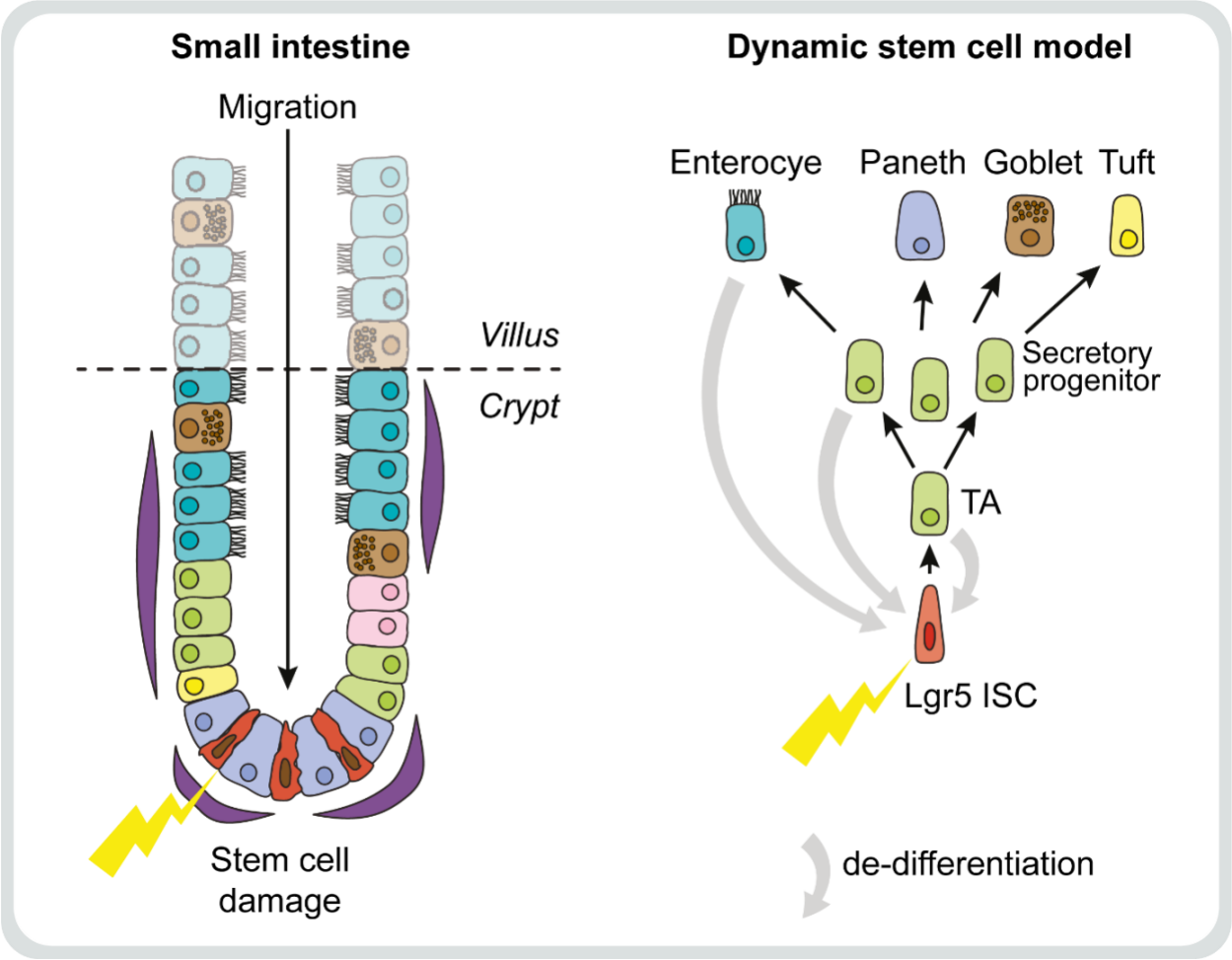
c. Impact of Microbiota Metabolites on Intestinal Stem-Cell Homeostasis

Present from birth to death, the microbiota is a part of organism life, particularly for intestinal homeostasis. Bacteria secrete metabolites that play an important role in intestinal networks. Microbiota metabolites affect also ISCs homeostasis, as presented some few examples are detailed below.

Several studies demonstrated that Short-Chain Fatty Acid (SCFAs) metabolites exert an effect on stem cells and seem to promote their proliferation. Therefore, SCFAs studies presented conflictual results, certainly due to the differential effect of each SCFA and variability of diet that strongly affect SCFA production. Lactate metabolite tends to have the same proliferating effect by regulating Wnt3 on Paneth-cells and IMCs that express the specific lactate receptor GPR81 [176].

Metabolites containing an active chemical indole or indole group-like are metabolized into indole derivative and play various essential biological functions acting in ISCs homeostasis. Microbiota indole- derivative metabolites are present in a high quantity compared to other metabolites, they seem to have an important function in intestinal homeostasis [176].

Tryptophan is an indole amino acid metabolized by the commensal microbiota to indole-derived metabolites known to be ligands of the xenobiotic sensor receptor aryl hydrocarbon receptor (AhR) and Pregnane X Receptor (PXR). The AhR signaling pathway is known to crosstalk with the Wnt/β-catenin signaling pathway essential for ISCs renewal. Depletion of AhR in the mouse present colonic hyperplasia and higher number of tumors. Further, microbiota indole-derivative metabolite indole-3-carbinol (I3C) binds AhR, reducing ISCs proliferation by promoting Wnt receptor degradation [176].



Picture taken from [221]

Figure 41. Current Model of Intestinal Stem-Cell Regeneration

ISC, intestinal stem cell; **TA**, transit amplifying cell

II. Intestinal Superpowers of Plasticity and Organoid Generation

A. Intestinal Plasticity, or The Superpower of Regeneration

a. Plastic cells, what that?

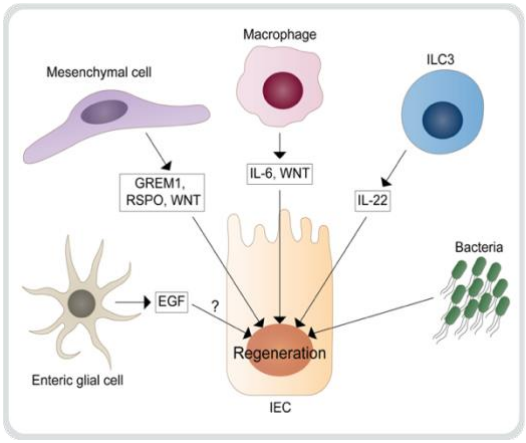
LgR5+ CBCs are essential to maintains intestinal epithelium (IE) balance and turn over. However, these incredible stem cells are sensible to epithelium injuries, induces by irradiations, chemicals products or pathogens aggression. After injuries, the intestinal epithelium undergoes the process of regeneration. Specific induced loss of CBCs in mice model do not impact IE homeostasis as a new pool of LgR5+ cells burst rapidly. Yet, at this point, renewal of the LgR5+ CBCs pool are critical for the epithelium regeneration. Our current knowledges on this matter define that +4 reserve stem cells, short-lived progenitors and even some subset of differentiated cells, can dedifferentiate act as a new source of CBCs. This ability of dedifferentiation is due to a high plasticity of intestinal cells [221].

At the beginning, it was hypothesized that quiescent +4 stem cells activated and convert into LgR5+ ISCs to replaced them. However, new studies controverse the idea of quiescent +4 stem cells. Indeed, this population seems to be more a subset of post-mitotic cells fated to a secretory cell linage. Several groups describe at least two +4 cells subgroup expressing related gene from enteroendocrine or tuft lineage. In addition, some secretory, goblet cell and enterocyte lineage precursors have also the capacity to replace ISCs loss Figure 41 [221].

b. Mechanisms Driving their Plasticity

Even if the mechanism of cell plasticity in regeneration context are still not fully apprehended, some mechanism and signals in the stem cell niche.

Epigenetic and chromatin accessibility is one of the mechanisms implicated. Through their differentiation, cells undergo through chromatin accessibility remodeling. In context of regeneration, the dedifferentiated cells need to reacquire expression ISC genes they lost. Chromatin accessibility of ISCs, absorptive and secretory progenitors have been analyzed and compare. ISCs and absorptive progenitors have a high accessibility of their chromatin compared to the secretory progenitors that have less chromatin accessibly. However, dedifferentiated secretory progenitor loss their difference in chromatin accessibility compared to ISCs. Absorptive progenitors seem to need less chromatin reorganization to dedifferentiate into ISCs. Meanwhile one study demonstrates that the polycomb complex PRC2 is essential for stemness and progenitor identity, and for the progenitor dedifferentiation ability after LgR5+ CBCs loss. Chromatin dynamic and epigenetic modification seems to be involved in the dedifferentiation ability for at least progenitors cells [148].



Adapted from [222]

Figure 42. Actors of the Intestinal Network Promote Intestinal Stem-Cell Regeneration

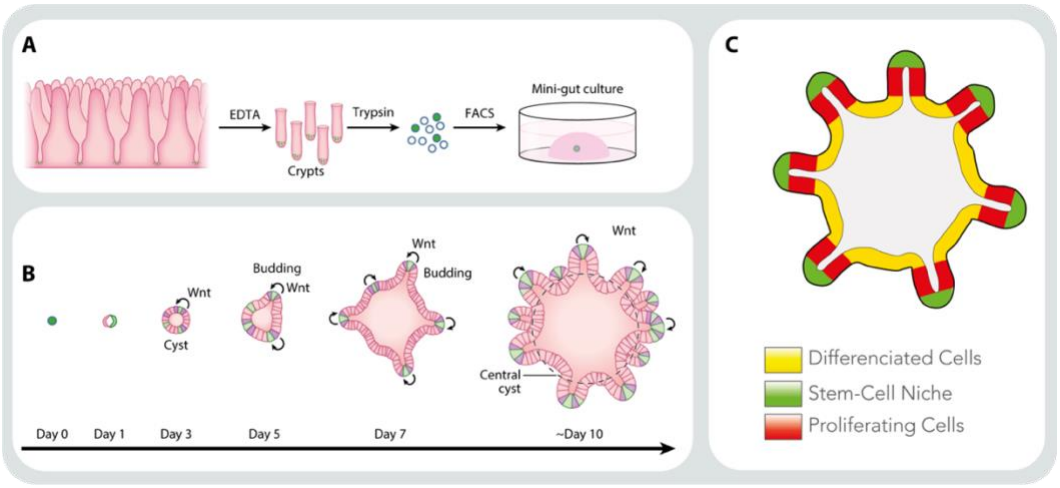
Overview of the cellular actors in the intestinal network and the components they secreted to promote intestinal stem-cell regeneration.

IEC, intestinal epithelium cell

RSPO, R-Spondin

GREM1, Gremlin 1

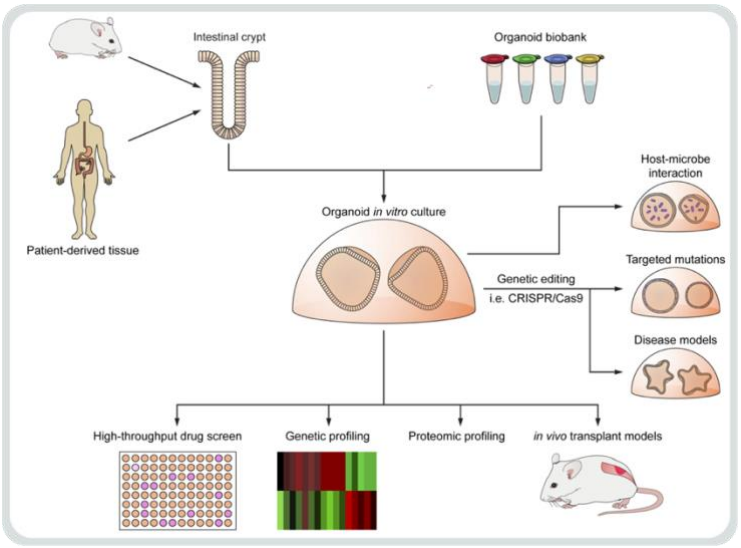
IL, Interleukine



Adapted from [214]

Figure 43. Mini Gut Culture or How LGR5+ Stem Cells Give Rise to Intestinal Organoids

A. Culture of intestinal crypt or isolate LGR5+ cells permit to generate organoids in *in vitro* culture when embedded in Matrigel and put in medium containing EGF, Noggin and R-spondin. **B.** Single Intestinal Cells form a "cyst" structure that will growth to form a complete organoid. By day 5, organoids start to form budding.



Adapted from [223]

Figure 44. Small tour of the Possible Applications of Intestinal Organoid Technology

Intestinal organoids may be established from mice or human patient-derived tissue and used to establish organoid bio bank. At that point, they may be used for functional studies in the mice or even drug screening as example.

As expected, the supportive network of ISCs in implicated in their plasticity. Blockage of the Wnt block damage repair and regeneration. Furthermore, Ascl2 is a target gene of the Wnt pathways and essential for the regenerative response. Notch is a second pathways playing a pivotal role in regeneration. After damage Hes1 a gene-target of Notch is increase. Blocking Notch pathways compromise the regeneration. Lastly, even if the Yap/Taz signaling (target of Hippo pathways), is dispensable for intestinal homeostasis, it is critical for the regeneration [222].

Not left behind, the actors of intestinal networks are actively implicated in regeneration and plasticity. Through the secretion of signals Immune cells, mesenchymal cells and bacteria promote the regeneration (

Figure 42) [222].

B. Organoids Challenging the LG5+ Intestinal Stem Cell

ISCs self-renewal capacity have been challenged. In 2009, Sato et al,. establish a mini-gut culture system with mouse ISCs. One single isolated LgR5+ ISCs are able to form *in vitro* a 3D structure reproducing the intestinal epithelium features and keep them in culture for more than a year. The 3D mini-gut organoids form an intestinal villus structure-like and contains all the differentiated cells whose localization reflects their *in vivo* counterparts. ISCs and Paneth cells are localized in the "budding" organoid's part, reflecting the *in vivo* crypt, while mature enterocytes only express *in vivo* in the villus, are found on spherical central part Figure 43 [214].

The 3D mini-gut organoids culture is possible by reproducing intestinal factors composition. Several critical factors expressed by stroma cells and critical intestinal homeostasis are added to the medium. On one hand, ISC/progenitor maintenance and proliferation are promoting by R-spondin, EGF factors. On the other, the pro-differentiator factor Noggin is added (a bone morphogenetic protein inhibitor). *In vivo*, the basement membrane is part of intestinal homeostasis. It is critical for organoids culture and provide by the Matrigel rich on lamina protein [214].

The development of organoids has provided a new useful method to studies the intestinal epithelium. As for example, they are used to elucidate mechanisms behind stem-cell, regeneration, crypt-villus morphogenesis, or intestinal network (immune cell, microbiota, *lamina-propria*) (Figure 44) [223].

OBJECTIVES

The intestine is a complex tissue at the interface between the intestinal lumen and the body, playing a major role in digestion, nutrient transport. It exerts a protective barrier between the host and the external environment. The physiology of the intestine involves a robust and consistent balance. Disruptions of its equilibrium leads to a cascade of alterations, causing pathologies such as the chronic inflammation of the intestine (IBD: Crohn disease and recto-colitis) and the colorectal cancer, two major public health problems in the world. Understanding the functioning of the intestine is essential for deciphering the subtleties of its networks and for getting the tools to fight the intestinal pathologies.

The intestinal epithelium is the heart of the intestinal physiology, function, and network. It is a tissue with a precise topological and functional organization. It is characterized by constant cell renewal operated by the stem cells located at the bottom of the intestinal crypt. The intestinal stem cells are multipotent cells capable of self-renewal and generating all the differentiated cell types of the intestinal epithelium. Each cell type (in terms of differentiation, maturation, and physiological functions of differentiated cells) is characterized by a specific cell-type gene expression profile. These different profiles are mainly controlled by transcription factors, proteins that govern the target-gene expression.

Transcription factors are intensively studied in the context of homeostasis and pathology. However, we tend to forget how they work: linking between diverse signals and the basic transcriptional machinery to induce gene transcription. The basic transcriptional machinery is a highly conserved key element in tissues physiology. This machinery is responsible for the initiation of the transcription, specifically via its multi-protein complex: Transcription Factor II D (TFIID) component. TFIID is a multi-protein complex consisting of the TATA-Binding Protein (TBP) and a variable combination of 13 TBP-Associated Factors (TAFs).

Contrary to the received wisdom, this "basic" machinery has specific functions in tissues through its TAFs components. Some TAF proteins has been demonstrated to be able to respond to signals and to have tissue-specific functions. In this way, TAF4 is one of the hugest TAF proteins that have specific function in skin, liver, pancreas, and embryo. For example, recent work has demonstrated a physical and functional interaction between Taf4 and Hnf4α (hepatocyte nuclear factor 4 alpha) in the liver.

Taken together, all these data highlight the potential of Taf4-specificity in the intestine. The goal of my thesis was to elucidate the transcriptional role of Taf4 in intestinal functions. I have addressed this issue by a loss-of-function approach, studying Taf4 gene invalidation specifically in the mouse intestine with a conditional knock-out mouse model.

My thesis work is composed of 3 axes:

1- role of Taf4 in the **intestinal development** and morphogenesis

2- role of Taf4 in the **intestinal homeostasis** in adult mice and in organoid model

3- impact of Taf4 on **pathologies** such as cancer

MY THESIS WORK

MATERIAL AND METHODS

1. Mice and Treatments

a. Mice Models

Mice experiments were performed in the certified animal facility (#G/H-67-482-21) according to the protocol approved by the French Ministry of Agriculture under the permit APAFIS #14197. *Taf4a$^{lox/lox}$* (Mengus et al., 2005) [5], *VilCre* and *VilCre^{ERT2}* (El Marjou et al., 2004) [224], *Lgr5-GFP-Cre^{ERT2}* (Barker et al., 2007) [225], and *Apc$^{Δ14/+}$* (Colnot et al., 2004) [226] mice have been described. Animals were genotyped by PCR on tail DNA with the following primers:
***Taf4a$^{lox/lox}$* allele:** CTAGTTACTGCTCTGCACAAT / GTGCTCCATGACTCTGGCAAG / CAGCCAAAGCTACATAATAAGT
***VilCre* and *VilCre^{ERT2}* alleles:** CAAGCCTGGCTCGACGGCC / CGCGAACATCTTCAGGTTCT
***Lgr5-GFP-Cre^{ERT2}* allele:** CTGCTCTCTGCTCCCAGTCT / ATACCCCATCCCTTTTGAGC / GAACTTCAGGGTCAGCTTGC
***Apc^{wt}* allele:** CTGTTCTGCAGTATGTTATCA / CTATGAGTCAACACAGGATTA
***Apc^{Δ14}* allele:** CTGTTCTGCAGTATGTTATCA / TATAAGGGCTAACAGTCAATA

b. Inactivation of *Taf4a* gene

For conditional inactivation of the *Taf4a* gene, *Taf4a$^{lox/lox}$::VilCre^{ERT2}* mice, *Taf4a$^{lox/lox}$::Lgr5-GFP-Cre^{ERT2}* mice or *Taf4a$^{lox/lox}$::VilCre^{ERT2}::Apc$^{Δ14/+}$* mice aged 2-3 months received intraperitoneal injections of 1.6 mg Tamoxifen (Tam, Sigma-Aldrich) in corn oil, once daily for 4 days. Controls, either wild types or *Taf4a$^{lox/lox}$* or *VilCre^{ERT2}* or *Lgr5-GFP-Cre^{ERT2}* or *Apc$^{Δ14/+}$*, also received Tam.

c. BrdU Pulse-Chase

BrdU pulse-chase labelling experiments were performed on *Taf4a$^{lox/lox}$::VilCre^{ERT2}* and control mice treated 10 days earlier with Tam and injected (day 0) with a single dose of 1 mg BrdU (Sigma Aldrich). Animals were euthanized at day 1, 2 or 3 after BrdU administration.

d. Paracellular and Transcellular Permeability

Paracellular and transcellular permeability was determined by measuring after gavage the serum levels of fluorescein isothiocyanate conjugated dextran (FITC-Dextan 4.4 kDa, Sigma-Aldrich) or D-Xylose (Sigma-Aldrich), respectively. Briefly, 10 days after Tam treatment, *Taf4a$^{lox/lox}$::VilCre^{ERT2}* and control *Taf4a$^{lox/lox}$* mice were starved during 4 hours and then force-fed with 500 mg/kg FITC-Dextran or 2 g/kg D-Xylose in phosphatase-buffered saline (pH 7.4). Blood samples were taken 3 hours after gavage by submandibular collection.

The serum levels of FITC-Dextran and D-Xylose were respectively measured by direct spectrophotofluometry and with the D-Xylose Kit (Chrondrex, 6601) coupled with spectrophotometry, using the spectrophometer TriStar^2 Multimode reader LB 942 (BERTHOLD Technologies).

2. Organoid Cultures and Treatments

Organoid cultures were established from ileal crypts of 4-month-old *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} and control *Taf4a*$^{lox/lox}$ mice, not treated with Tam. Ileal fragments were incubated in Gentle Cell Dissociation Reagent (STEMCELL™, 07174) for 15 min, then ~50 crypts were embedded in 20 µl of Matrigel® (Corning®, #356231) in 48-wells plates (Greiner Bio-one, 677180) and grown in 250 µl Mouse IntestiCult™ Organoid Growth Medium (STEMCELL™, 06005). IntestiCult™ medium was changed every 2 days. Organoids were passed every week after mechanical breakage with a 200 µl pipette and dilution at 1:4 for maintenance and RNA extraction or at 1:8 for immunolabelling and kinetics studies. Experiments were performed on organoids established for at least 5 passages.

For early ***Taf4a*** **gene invalidation**, organoids were plated for 2 hours and then treated for 72 hours with 1 µM (Z)-4-hydroxytamoxifen (H-Tam, Sigma-Aldrich, H7904) in Ethanol or with Ethanol alone for control experiments. For late *Taf4a* gene invalidation, organoids were grown for 5 days and then treated for 72 hours with H-Tam in EtOH or EtOH. During the 3 days of treatment, IntestiCult™ medium with fresh H-Tam in EtOH or EtOH alone was changed every day.

When indicated, **EPZ6438** 1 µM (MedChemExpress, HY-13803) in DMSO was added to the culture medium together with H-Tam, then changed with fresh medium every day during the 3 days of H-Tam treatment and then every 2 days up to day 15 of culture.

3. Immunostaining of Tissue Samples and Organoids

a. Antibodies

Antibodies used for this study are listed in the Table 1.

b. Organoids Immunofluorescence

Organoids grown in IntestiCult™ medium and Matrigel in 8-wells Lab-Tek® Chamber Slide™ (Dominique Dutscher) were directly fixed for 30 min with 4% paraformaldehyde, 60 mM PIPES, 25 mM HEPES, 10 mM EGTA, 2 mM magnesium acetate and permeabilized for 30 min in 0.5% Triton X100 (Euromedex). After blockade for 2 hours at 37°C in 5% BSA (Euromedex), primary antibodies were added and incubated 2 hours at room temperature followed by overnight at 4°C. Secondary antibodies were incubated for 2 hours at 37°C. Nuclei were stained with 40,6-diamidino-2-phenylindole dihydrochloride (DAPI) and actin was revealed by Phalloïdin-TRITC (Sigma-Aldrich, P1951). Samples were mounted in ProLong™ Gold Antifade Mountant (Life Technologie™, P36930) and analyzed with an Axio Zoom.V16 microscope (Zeiss) for stereomicroscopy or with an Axio Imager M2 microscope coupled to a Hamamatsu's camera Orca Flash 4v3 using the ApoTome.2 function (Zeiss) for optical sectioning.

	Antobody			Dilution	
	Product	Compagny	Species	Tissue	Organoids
Aldolase B+C	ab75751	Abcam	Rabbit	1:500	
b-catenin	610154	BD Transduction	Mouse	1:500	
Biotin Anti-BrdU	2284	Abcam	Sheep	1:500	
Caspase3 Clived	D175	Cell Signaling	Rabbit	1:250	1:250
Cdx2	ab76541	Abcam	Rabbit	1:10000	
ChromograninA	20086	Imminostar	Rabbit	1:1000	
Dclk1	ab37994	Abcam	Rabbit	1:250	
EphB2	AF-467	R&D system	Goat	1:500	1:250
H2AX	JBW301	Millipore	Mouse		1:500
Hes1	119885	Cell Signaling		1:5000	
Hnf4α 1-6	PP-K9218-00	R&D system	Mouse	1:2000	
Hnf4α 7-9	PP-H6939-00	R&D system	Mouse	1:500	
Ki67	RM-9106-S	Thermofisher	Rabbit	1:500	1:500
Lyzozyme	A099	DAKO	Rabbit	1:500	1:500
Muc2	sc-15334	Santa Cruz	Rabbit	1:1000	1:1000
Olfm4	39141S	Cell Signaling	Rabbit	1:1000	1:500
phospho-Erk1/2	4370	Cell Signaling	Rabbit	1:500	
Sox9	AB5535	Millipore	Rabbit	1:1000	
Taf4a	sc-136092	Santa Cruz	Mouse	1:500	

Table 1. Antibodies

4. Bacterial 16S RNA Analysis

The luminal content was collected in the cecum of 8 5-month-old *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} and 8 control *Taf4a*$^{lox/lox}$ mice treated with Tam at the age of 3 months. DNA extraction was performed using NucleoSpin® DNA Stool kit (Macherey-Nagel – 740472.5) adapted with a mechanical lysis step (Fastprep – 6.5m.sec-1 for 2 min). The V3-V4 region of the 16S rRNA gene was amplified with the primers TACGGRAGGCAGCAG / ATCTTACCAGGGTATCTAATCCT according to GeT-PlaGe platform protocol (INRA). Sequencing was performed on Illumina MiSeq system using 2*300 bp paired-end mode. For sequence data analysis, remaining adapter/primer sequences were trimmed, and reads were checked for quality (≥20) and length (≥200 bp) using Cutadapt (Martin, 2011) [227]. Reads were further corrected for known sequencing errors using SPAdes (Bankevich et al., 2012) [228] and then merged using PEAR (Zhang et al., 2014) [229]. Operational Taxonomic Unit (OUT) were identified using a Vsearch pipeline (Rognes et al., 2016) [230] set up to dereplicate, cluster, chimera check the merged

reads. OTU taxonomical classification was performed using classifier from the RDPTools suit (Cole et al., 2009) [231]. Statistical tests were performed using Wilcoxon Rank Sum Test for group comparison. Multiple tests were corrected using the False Discovery Rate method (*q*-value) as required.

5. Bulk RNA Preparation and RNAseq

RNA was extracted on the one hand from the small intestine of 3 E17.5 *Taf4a*$^{lox/lox}$*::VilCre* and 3 controls *Taf4a*$^{lox/lox}$ mouse embryonic littermates, and on the other hand from 3 adults *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} and 3 controls *Taf4a*$^{lox/lox}$ mice 10 days after Tam administration. RNA was also extracted from 3 wells of H-Tam-treated *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} and control *Taf4a*$^{lox/lox}$ organoids at day 3 of culture, and from 3 wells of *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} organoids at day 15 of culture, namely organoids being treated with H-Tam from day 1 to 3 and with EPZ6438 from day 1 to 15.

RNA preparation used Tri Reagent (Euromedex) and the quality was analyzed using nanoRNA chips on a Bioanalyser 2100 (Agilent Technologies). Complementary DNA was generated and linearly amplified from 3 ng total RNA using the Ovation RNA-seq V2 system (NuGEN technologies Inc., Leek, The Netherlands), according to the manufacturer's instructions. The amplified cDNA was then purified using Agencourt AMPure XP beads (Beckman-Coulter, Villepinte, France) in a 1.8:1 bead to sample ratio and fragmented by sonication using a Covaris E220 instrument (with duty cycle: 10%, maximum incident power: 175 watts and cycles/burst: 200 for 120 s). The RNA-seq libraries were generated from 100 ng fragmented cDNA using the Ovation Ultralow v2 library system (NuGEN technologies Inc., Leek, The Netherlands) according to the manufacturer's instructions, with 6 PCR cycles for library amplification. The final libraries were verified for quality and quantified using capillary electrophoresis before sequencing on an Illumina Hi-Seq4000.

Reads were preprocessed to remove adapter and low-quality sequences (Phred quality score below 20). After this preprocessing, reads shorter than 40 bases were discarded for further analysis. These preprocessing steps were performed using Cutadapt version 1.10 (Martin, 2011) [227]. Reads were mapped to rRNA sequences using Bowtie version 2.2.8 (Langmead et al., 2009) [232], and reads mapping to rRNA sequences were removed for further analysis. Reads were mapped onto the mm9 assembly of Mus musculus genome using STAR version 2.5.3a (Dobin et al., 2013) [233]. Gene expression quantification was performed from uniquely aligned reads using Htseq-count version 0.6.1p1 (Anders et al., 2015) [234], with annotations from Ensembl version 67 and "union" mode. Only non-ambiguously assigned reads have been retained for further analyses. Read counts have been normalized across samples with the median-of-ratios method (Anders and Huber, 2010) [235], to make these counts comparable between samples. Comparisons of interest were performed using the Wald

test for differential expression and implemented in the Bioconductor package DESeq2 version 1.16.1 (Love et al., 2014) [236]. Genes with high Cook's distance were filtered out and independent filtering based on the mean of normalized counts was performed. *p*-values were adjusted for multiple testing using the Benjamini and Hochberg method (Benjamini and Hochberg, 1994) [237]. Heatmaps were generated with R-package pheatmap v1.0.12. Deregulated genes were defined as genes with Log2(FoldChange) >1 or <-1 and adjusted *p*-value <0.05.

6. ATACseq

ATAC-seq was performed from 20,000 cells of H-Tam-treated *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} and control *Taf4a*$^{lox/lox}$ organoids at day 3 of culture. Sequenced reads were mapped to the mouse genome assembly mm9 using Bowtie (Langmead et al., 2009) [232] with the following arguments: "-m 1 --strata --best -y -S -l 40 -p 2".

After sequencing, peak detection was performed using the MACS software (Zhang et al., 2008) [238] v2.1.1.20160309 with arguments "--nomodel --shift -100 --extsize 200 --broad". Peaks were annotated with Homer (Heinz et al., 2010) [239] using the GTF from ENSEMBL v67. Peak intersections were computed using Bedtools (Quinlan and Hall, 2010) [240]. Global Clustering was done using seqMINER (Ye et al., 2011) [241]. De novo motif discovery was performed using the MEME suite (Bailey et al., 2015) [242]. Footprinting signatures were calculated using TOBIAS v0.5.1 (Bentsen et al., 2020) [243], differential footprinting scores were plotted with R-package ggplot2 (Wickham, 2016) [244].

7. Single-Cell RNAseq (sc-RNAseq)

Taf4a$^{lox/lox}$*::VilCre*^{ERT2} and control *Taf4a*$^{lox/lox}$ organoids treated with H-Tam from day 5 to 8 were dissociated at the end of the H-Tam treatment with Accutase (A6964, Sigma) at 27 °C for 5 min and the cells were suspended in culture medium. Cells were then sorted by flow cytometry to select live cells. Four hundred cells were captured using 10X Genomics Chromium Analyzer. After sequencing, raw reads were processed using CellRanger (v 3.1) to align on the mm10 mouse genome, remove unexpressed genes and quantify barcodes and UMIs. Data were then analyzed in R (v3.6.3) using Seurat v3.1.4 (Stuart et al., 2019) [245]. First cells were filtered, only cells with feature count ranging from 200 to 6000 and with percentage of mitochondrial reads <15% were kept for the analysis. Then counts were normalized with the "LogNormalize" method and scaled to remove unwanted sources of variation. Clustering was performed on variable features using the 20 most significant principal components and a resolution of 0.9. Analysis of regulome was performed using SCENIC v1.1.2.2 (Aibar et al., 2017) [246].

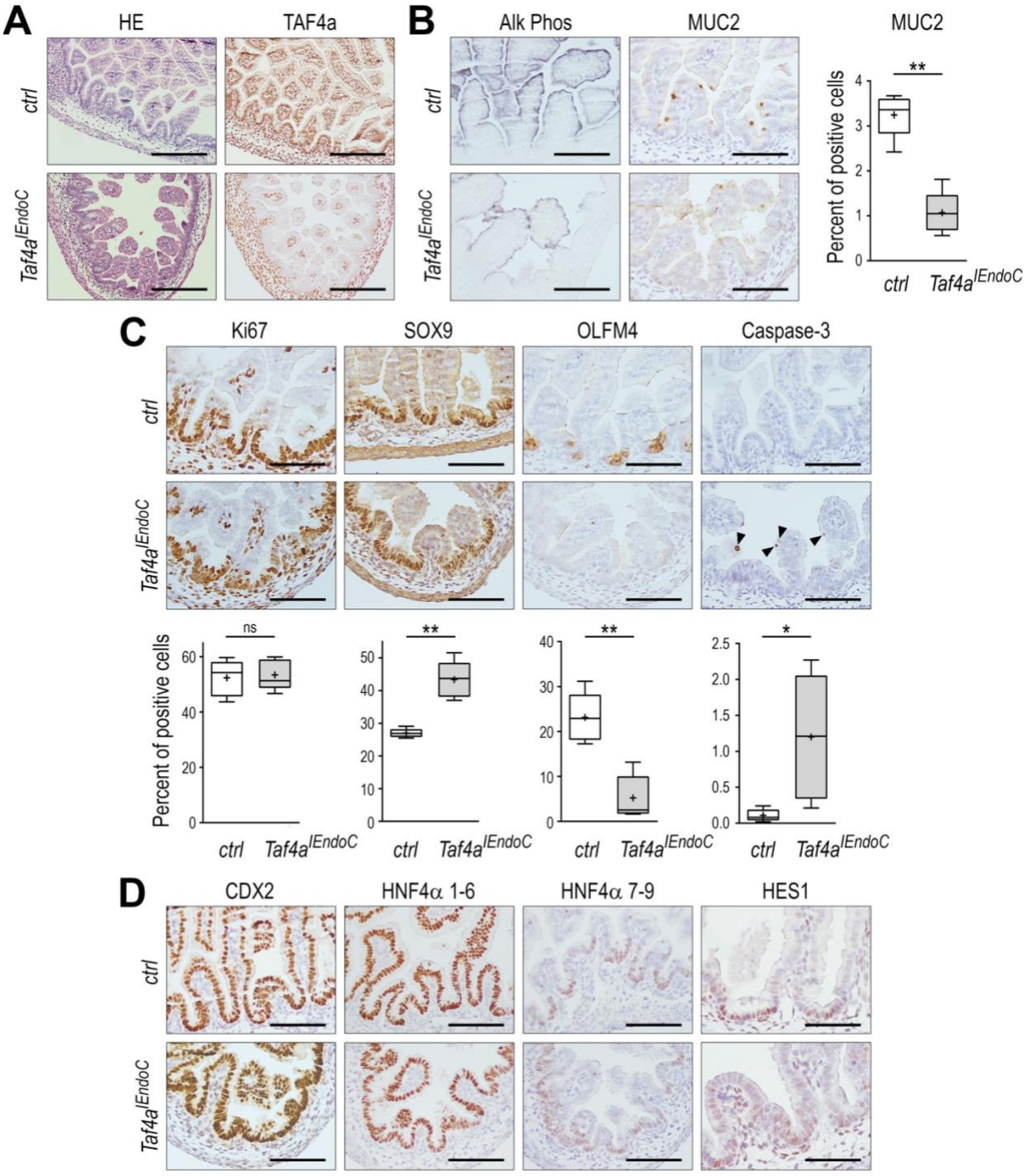


Figure 45. Effect of *Taf4a* inactivation in the gut endoderm of E17.5 fetuses

A. Morphology (HE) and immunohistochemical detection of the TAF4a protein in E17.5 control (ctrl) and *Taf4a*^{IEndoC} littermates. **B**. Same as above for Alkaline phosphatase activity and MUC2 immunostaining. **C**. Same as above for Ki67, SOX9, OLFM4 and activated Caspase-3. **D**. Same as above for CDX2, the isoforms 1-6 and 7-9 of HNF4a, and HES1. Bars are 200 µm in A and 100 µm in B, C and D. ns, no significant; * $p<0.05$; ** $p<0.01$.

RESULTS

1. Morphogenetic Consequences of *Taf4a* Inactivation in the Embryonic Intestinal Endoderm

Taf4a gene inactivation in the presumptive intestinal endoderm around 11-12 days post coitum (dpc) in *Taf4a*$^{lox/lox}$*::VilCre* embryos (hereafter *Taf4a*^{lEndoC}) led to perinatal lethality. The depletion in TAF4a protein was progressive and almost complete at 17.5 dpc, resulting at this stage in impaired intestinal morphogenesis with short and less dense villi (Figure 45.A). Altered alkaline phosphatase activity and MUC2 expression suggested perturbed cell differentiation, whereas cell proliferation visualized by Ki67 labeling was unchanged (Figure 45.B).

In E18.5 *Taf4a*^{lEndoC} fetuses, the dynamic morphogenetic process rapidly worsened as the mucosa became flat with only few bulged villi (Figure 46). Alkaline phosphatase activity and MUC2 expression were strongly reduced. Nevertheless, despite villi involution, presumptive crypt regions were preserved as evidenced by the labelling of SOX9 and Ki67 which further demonstrated the proliferative capacity retained by these cells. However, OLFM4 marking the emergence of adult-type stem cells in the intervilli regions was barely detectable. By contrast, important transcription factors for intestinal identity and/or differentiation, i.e. CDX2 and HNF4α 1-6, were unaltered, whereas HES1 and HNF4α 7-9 became very low and irregular in the regions presenting the most altered phenotype. In addition, apoptotic cells were displayed by activated Caspase-3 staining, mainly at the level of remnant villi. The apoptotic process was already engaged at the level of villi at E17.5 together with the failure to efficiently turn on OLFM4 in the intervilli regions, whereas the patterns of HES1 and HNF4α 7-9 were not yet altered at this stage (Figure 45.C,D).

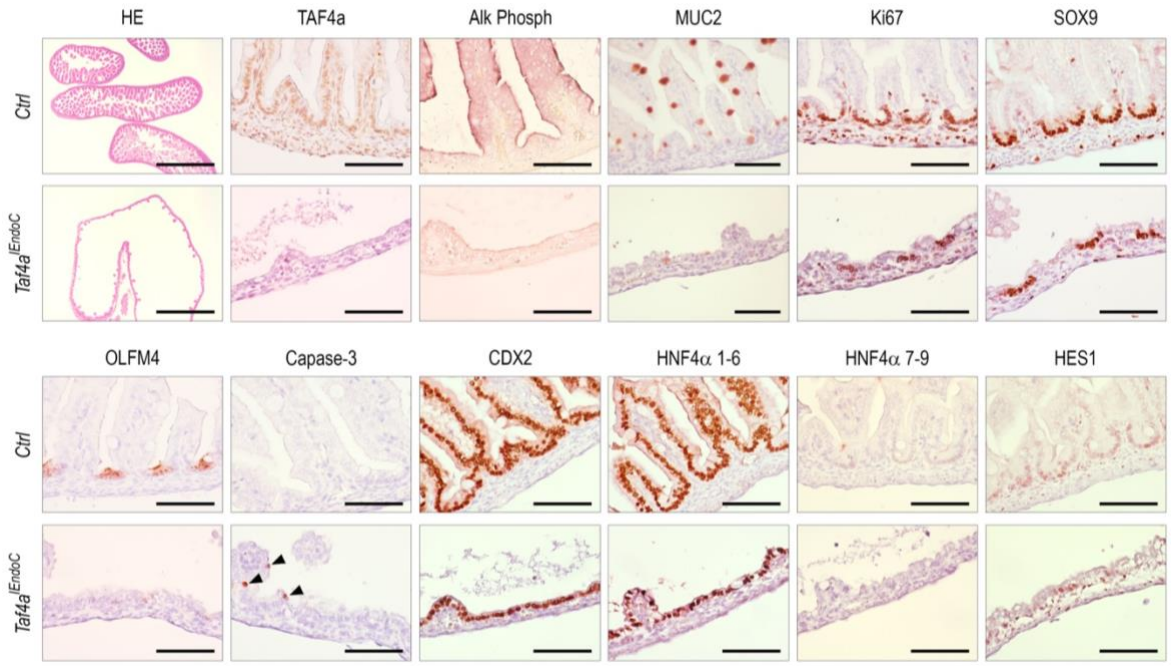


Figure 46. Morphological alteration resulting from *Taf4a* inactivation in the gut endoderm of E18.5 fetuses

Morphology (HE) and immunohistological detection of the indicated proteins in E18.5 control (ctrl) and *Taf4a*^{lEndoC} littermates. Bars are 50 µm except for HE and MUC2 where they represent 500 µm and 100 µm, respectively.

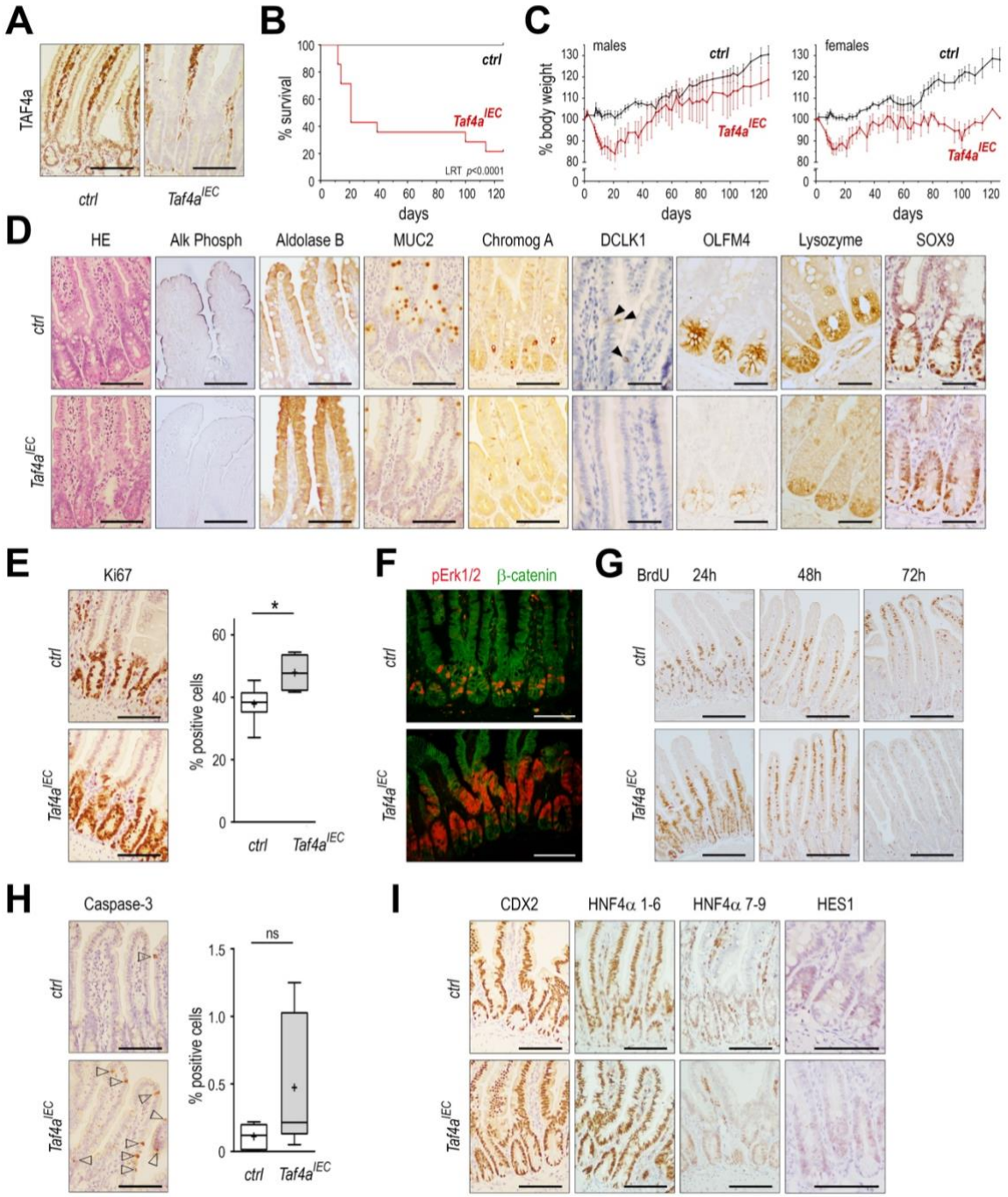


Figure 47. Homeostasis defects induced by *Taf4a* inactivation in the adult gut epithelium

A. Immunodetection of the TAF4a protein in the intestine of adult *Taf4a*^{IEC} and control *Taf4a*$^{lox/lox}$ (ctrl) mice 10 days after Tam injection. Bars are 100 µm. **B**. Overall survival of adult *Taf4a*^{IEC} (n=28) *vs* control *Taf4a*$^{lox/lox}$ mice (n=22) after Tam injection. **C**. Evolution of the body weight in males (n=7) and females (n=7) *Taf4a*^{IEC} mice after Tam injection, compared to control sex-matched animals. LRT: Logrank test. **D**. Histology (HE) and immunodetection of the indicated proteins in the ileum of *Taf4a*^{IEC} and control mice 10 days after Tam injection. Bars are 100 µm for HE, Alk Phos, Aldolase B, MUC2 and Chromog A, and 50 µm for DCKL1, OLFM4, Lysozyme and SOX9. **E**. Ki67 immunostaining and cell counts in the ileum of Tam-injected *Taf4a*^{IEC} and control mice. Bars are 100 µm. Boxes extend from the 25th-75th percentile and whiskers represent mean to max. * p<0.05. **F**. Immunofluorescent staining of phosph-Erk1/2 (red) and b-catenin (green) in Tam-injected *Taf4a*^{IEC} and control mice. Bars are 100 µm. **G**. BrdU detection in the jejunal mucosa of Tam-treated *Taf4a*^{IEC} and control mice at the indicated days after the single injection of BrdU. Bars are 200 µm. **H**. Activated-Capsase3 immunodetection and cell counts in *Taf4a*^{IEC} and control mice 10 days after Tam injection. Boxes extend from the 25th-75th percentile and whiskers represent mean to max. Bars are 100 µm. ns, not significant. **I**. Immunodetection of CDX2, of the isoforms 1-6 and 7-9 of HNF4a, and of HES1 in the intestine of Tam-treated *Taf4a*^{IEC} and control mice 10 days after Tam administration. Bars: 100 µm.

2. Consequences of *Taf4a* Inactivation on the Dynamic Homeostasis of the Adult Intestinal Epithelium

Taf4a gene inactivation in the adult intestinal epithelium by Tam injection to 2-3 month-old *Taf4a*$^{lox/lox}$*::VilCre*^{ERT2} mice (hereafter *Taf4a*^{IEC}) (Figure 47.A) compromised overall survival (Figure 47.B). During the few days after gene invalidation, mice transiently lost weight and then recovered, better males than females, but remained below the controls (Figure 47.C). *Taf4a*-invalidated mice suffered from diarrhea and exhibited a swollen cecum distended by gas.

Although the general morphology of the crypt-villus unit was not affected, alterations occurred at the cellular level in both proliferation and differentiation compartments (Figure 47.D). In the villi, enterocytes expressed lower Alkaline Phosphatase activity but increased Aldolase levels. The number of MUC2-positive goblet cells was reduced and enteroendocrine cells expressed less Chromogranin A. DCLK1 labeling of tuff cells was also strongly reduced. In the crypts, OLFM4 expressed by Crypt Base Columnar (CBCs) stem cells was perturbed along with Lysozyme in the adjacent Paneth cells, whereas SOX9 was unaltered. Functionally, these changes associated with increased cell proliferation visualized with Ki67, higher phospho-Erk1/2 staining in the proliferative compartment, accelerated cell turnover assessed by BrdU pulse-chase labeling, and increased apoptosis displayed by activated Caspase-3 (Figure 47.E-H). However, the patterns of the transcription factors CDX2, HNF4α 1-6 and 7-9, and HES1 were unaltered (Figure 47.I).

To strengthen the relationship between *Taf4a* invalidation in stem cells and the defective differentiation of mature villi cells, *Taf4a*$^{lox/lox}$*::Lgr5-GFP-Cre*^{ERT2} mice (hereafter *Taf4a*^{CBC}) were generated to selectively invalidate *Taf4a* in the CBCs. Adult mice were treated with Tam and analyzed 6 days later. As expected from the mosaic expression of the *Lgr5-GFP-Cre*^{ERT2} allele in CBCs, only few crypts exhibited TAF4a protein depletion in stem cells whereas adjacent Paneth cells still expressed the protein because of their slow turnover (Figure 48.A). Accordingly, *Taf4a*^{CBC} mice showed villi lined by TAF4a-positive cells with few ribbons of TAF4a-negative cells originating from the TAF4a-depleted crypts. Noteworthy, these TAF4a-negative cells in the villi exhibited less apical alkaline phosphatase activity then adjacent TAF4a-expressing cells, indicating that *Taf4a* invalidation in the stem cells perturbed the terminal differentiation of their progeny migrating upwards in the villi (Figure 48.B).

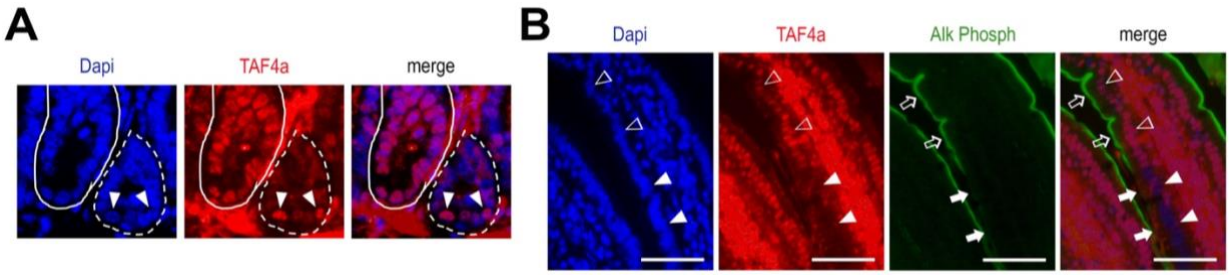


Figure 48. Effect of *Taf4a* Inactivation in Adult CBC Cells

A. TAF4a protein immunodetection in *Taf4a*^{CBC} mice 6 days after Tam treatment. The regular line encircles a crypt in which the *Taf4a* gene has not been invalidated. The adjacent crypt encircled with the dotted line shows a *Taf4a*-invalidated crypt where Paneth cells with a half-life of several dozens of days are the only cells that still express the TAF4 protein (white arrowheads). **B.** Nuclei (blue), TAF4a (red) and Alkaline Phosphatase (green) staining of villi of *Taf4a*^{CBC} mice 6 days after Tam injection. White arrowheads and open arrowheads respectively show TAF4a-depleted and TAF4a-expressing nuclei, and white arrows and open arrows respectively point to alkaline phosphatase absent and present at the apical pole of enterocytes. Bars are 50 μm.

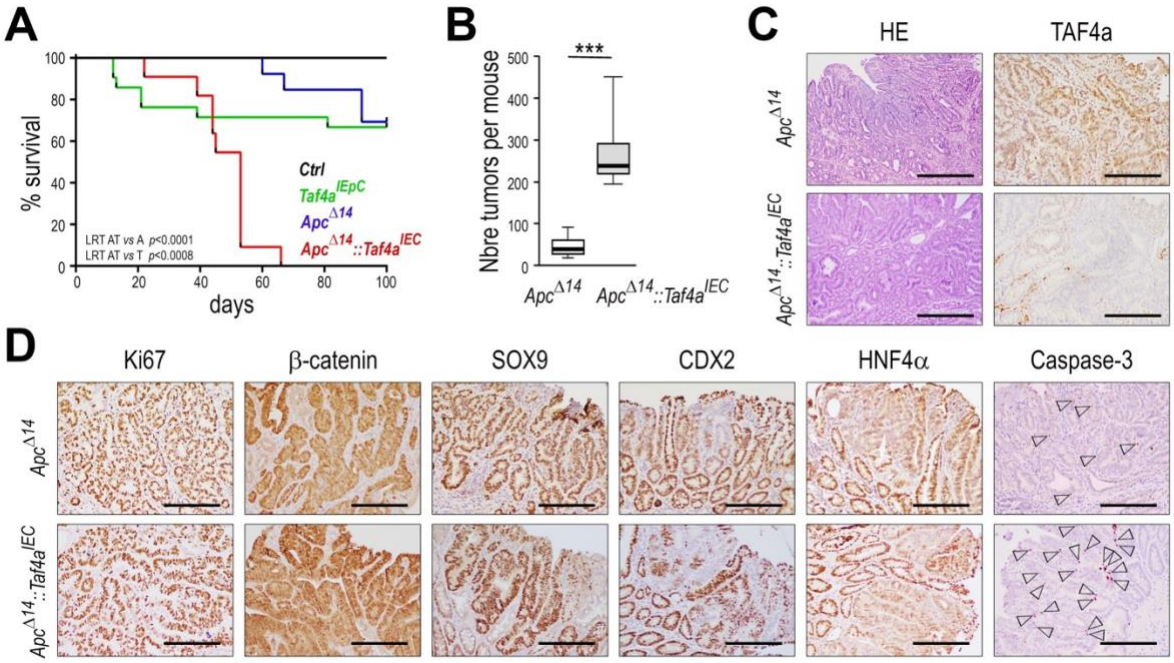


Figure 49. Impact of *Taf4a* Inactivation on Intestinal Tumor Development

A. Overall survival of adult $Apc^{+/D14}$::$Taf4a^{IEC}$, $Apc^{+/D14}$, $Taf4a^{IEC}$ and $Taf4a^{lox/lox}$ (ctrl) males after Tam injection; n>10 for each genotype. LRT AT *vs* A: Logrank test between $Apc^{+/D14}$::$Taf4a^{IEC}$ and $Apc^{+/D14}$ mice; LRT AT *vs* T: Logrank test between $Apc^{+/D14}$::$Taf4a^{IEC}$ and $Taf4a^{IEC}$ mice. **B**. Tumor number in the small intestine of Tam-treated $Apc^{+/D14}$::$Taf4a^{IEC}$ and $Apc^{+/D14}$ mice. *** $p<0.0001$. **C**. Histology (HE) and immunodetection of TAF4a in ileal tumors of $Apc^{+/D14}$::$Taf4a^{IEC}$ and $Apc^{+/D14}$ mice. Bars are 400 µm for HE and 200 µm for TAF4a. **D**. Same as C for the indicated proteins, showing similar immunohistological patterns except for increased apoptosis labelled by activated Caspase3 (arrowheads) in $Apc^{+/D14}$::$Taf4a^{IEC}$ tumors. Bars are 200 µm.

In line with the increased cell proliferation, *Taf4a* invalidation exacerbated tumor burden in the tumor prone model of $Apc^{+/\Delta 14}$ mice, while $Taf4a^{IEC}::Apc^{+/\Delta 14}$ tumors were histologically similar to those of $Apc^{+/\Delta 14}$ mice, except for higher rate of apoptosis (Figure 49).

Altogether, these data highlight that *Taf4a* invalidation perturbs the dynamic homeostasis of the stem cell niche and transit amplifying cell compartment ultimately leading to impaired differentiation of both absorptive and secretory cell lineages.

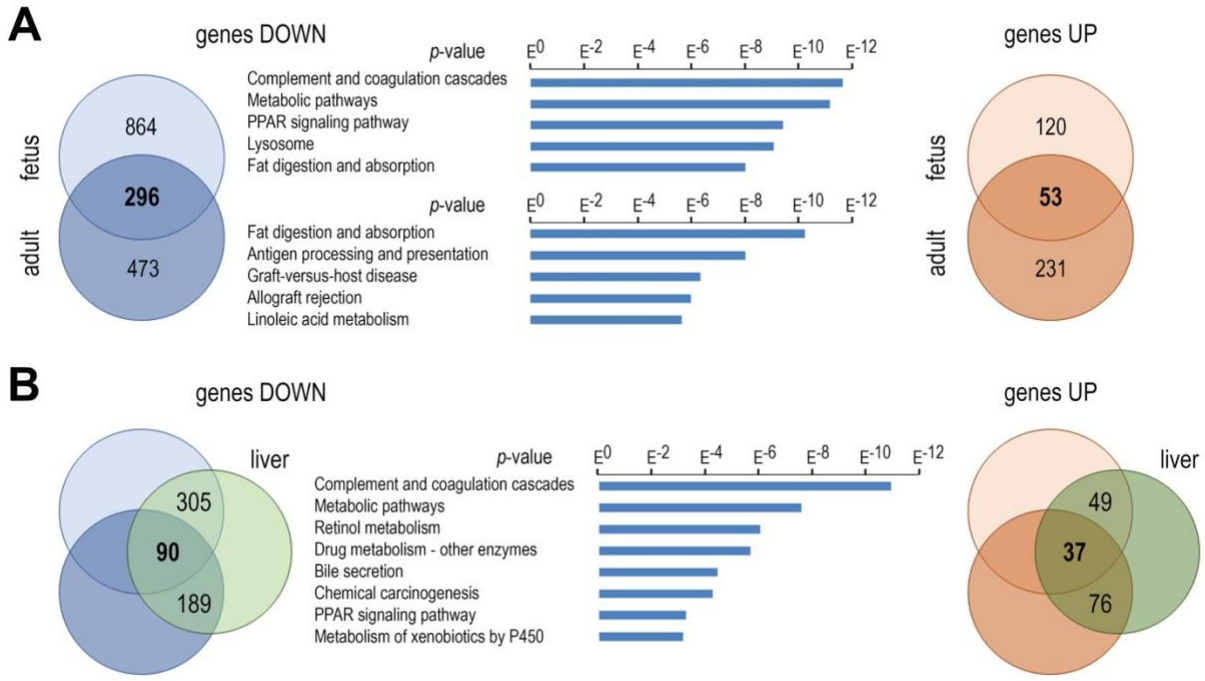


Figure 50. Gene Expression Changes After *Taf4a* Inactivation

A. Venn diagrams representing the down- and up-regulated genes in E17.5 *Taf4a*^{lEndoC} fetuses and adult *Taf4a*^{lEC} mice compared to their respective controls. KEGG ontology enrichment is shown for the downregulated genes and ordered according to the p-value. **B.** Same as in A for the down- and up-regulated genes enriched in common in the intestine of E17.5 T*af4a*^{lendoC} fetuses and adult *Taf4a*^{lEC} mice, and in the *Taf4a*-null liver of 12 days suckling mice.

3. Gene Expression Changes Resulting from *Taf4a* Inactivation in the Fetal and Adult Intestine

The molecular outcome of the loss of function of TAF4a was investigated by RNAseq in E17.5 fetuses and in adult mice. RNAseq performed on the ileum of adult *Taf4a*^{IEC} and control littermates 10 days after Tam administration revealed the deregulation of a limited number of genes (1057) ($|\log_2(FC)|>1$, p-value<0.05; **Table 2**). Many genes involved in enterocyte functions were downregulated, including those coding for digestive enzymes (ALPI, ANPEP, TREH), fatty acids binding proteins (FABP2, -6, -7), and more than 40 soluble carrier family members among which the transporters/cotransporters of Na/Glucose (SLC5a11/SGLT6, SLC5a12/SMCT2), oligopeptides (SLC15a1/PEPT1), short-chain fatty acids (SLC16a3), amino acids (SLC7a8, SLC7a9, SLC7a15, SLC38a3) and folate (SLC46a1, SLC46a3). However, the α-glucosidase genes *Sis* and *Mgam* and the aldolase gene *Aldob* were upregulated. The expression of goblet cells mucin genes *Muc4* and *Muc20* was decreased, like enteroendocrine genes encoding precursors and regulatory peptides (*Chga*, *Chgb*, *Cck*, *Gip*, *Pyy*, *Sct* and *Sst*). The same was observed with *Dclk1* and *Pou2f3* for tuff cells. Thus, *Taf4a* inactivation perturbed the terminal differentiation of all mature epithelial cell types migrating upwards the villi. In addition, it also altered gene expression in the stem cell niche, as exemplified by the reduced expression of genes of the CBC signature (*Agr3*, *Ciita*, *Esrrg*, *Fras1*, *H2-Eb1*, *Hk2*, *Lect2*, *Olfm4*, *Rdh16*, *Sdsl*, *Sectm1b*, *Tifa*, *Tnfsf10*, *Tns4*, *Vnn1*) (Muñoz et al., 2012)[6] as well as Paneth cell genes (*Defa26*, *Dafa-rs1*, *Defb1*, *Defb37*, *Lyz1*, *Mmp7*).

Comparing the transcriptional changes resulting from *Taf4a* invalidation in E17.5 *Taf4a*^{lendoC} fetuses and adult *Taf4a*^{IEC} mice revealed a higher number of downregulated than upregulated genes at both stages (respectively 1160 *vs* 173 in fetuses and 769 *vs* 284 in adults) (Figure 50.A, Tables 2-3). The top 5 pathways identified by KEGG for downregulated genes in fetuses and adults were related to metabolism, fat digestion and complement function / immune system. Previously, we have reported the effect of *Taf4a* invalidation in another endoderm-derived organ, the liver of mouse neonates, which also perturbed metabolic pathways (Alpern et al., 2014)[4]. Among the downregulated genes in the liver, 395 were shared with the fetal gut and 279 with the adult gut, 90 of them being common in the 3 situations (Figure 50.B). KEGG pathway analysis linked these 90 genes to various aspects of metabolism, to the function of the complement and to chemical carcinogenesis (Figure 50.B). Inversely, no specific KEGG pathway was associated to the upregulated genes in common in the intestine of *Taf4a*-null fetuses and adult mice, if one excludes from the list the high number of histone genes (respectively 21 in fetuses and 28 in adults), a feature already reported in the liver.

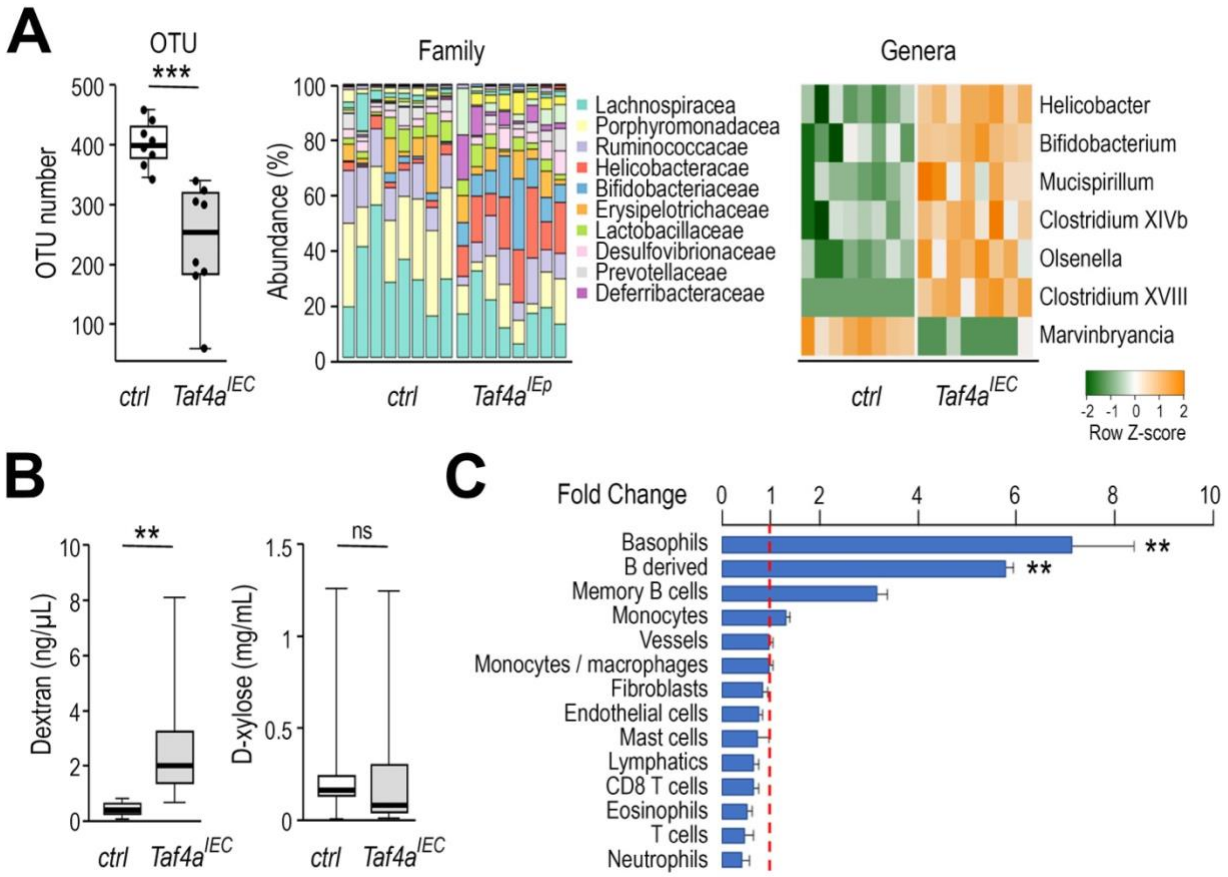


Figure 51. Microenvironment Changes Resulting from *Taf4a* Inactivation

A. Microbiota. Number of clusters of similar sequence variants of the 16S rDNA marker gene sequence (I: Operational taxonomic units), relative abundance of microbiota families, and genera abundance in the cecum of Tam-treated *Taf4a*^{IEC} (n=8) and *Taf4a*$^{lox/lox}$ control (n=8) mice; *** q<0.001. **B**. Epithelial barrier activity. Blood concentration of FITC-Dextran (left) and D-Xylose (right) in Tam-treated *Taf4a*^{IEC} (n=14) and *Taf4a*$^{lox/lox}$ control (Ctrl) (n=13) mice 3 hours after gavage; ** p<0.01. **C**. Microenvironment. Stromal cell population evaluation using the mMCP method, expressed as the proportion of each cell type in *Taf4a*^{IEC} mice compared to controls. ** p<0.01.

4. Pathophysiological consequences of *Taf4a* inactivation in the adult gut

Along with digestive and metabolic functions, the gut epithelium plays an important role of barrier at the interface between the luminal microbiota and the mucosal immune system. Diarrhea and the swollen cecum after *Taf4a* gene invalidation led to hypothesize changes in the microbiota, which was demonstrated by comparing the bacterial composition of the cecal content of Tam-treated *Taf4a*$^{\Delta IEC}$ and control *Taf4a*$^{lox/lox}$ mice (Figure 51.A). Indeed, the α-diversity was reduced with fewer Operational Taxonomic Units (OTUs) in mutants compared to controls. The microbiota of *Taf4a*^{IEC} mice showed higher abundance of *Helicobacteriaceae*, *Bifidobacteriaceae* and *Deferribacteraceae* families but decay of *Porpyromonadaceae*. At genus level, seven genera displayed differential levels: *Bifidobacterium*, *Clostridium XIVb* and *XVIII*, *Helicobacter*, *Mucispirillum* and *Olsenella* were significantly more abundant in mutants, whereas controls harbored more *Marvinbryantia*. OTUs included in *Helicobacter* genera were close to *H. typhlonius* (100%), *H. ganmani* (98%) and an uncultured species (JRPC – 100%), and those included into *Bifidobacterium* genera close to *B. pseudolongum* (98%) and *B. animalis* (100%); *M schaedleri* (100%) was the main strain among *Mucispirillum* genera (Figure 51.A).

Interestingly, the change in microbiota composition upon *Taf4a* invalidation was accompanied by increase paracellular permeability of the epithelium measured *in vivo* by FITC-Dextran luminal-to-blood transfer, whereas transcellular permeability measured with D-Xylose was unchanged (Figure 51.B).

These observations prompted us to investigate the cellular composition of the mucosa. For this purpose, RNAseq data were analyzed with the murine Microenvironment Cell Population counter method (Petitprez et al., 2020)[247] to estimate the abundance of tissue-infiltrating immune and stromal cell populations. As shown in Figure 51.C, the proportion of basophils and B-derived cells was predicted to increase in *Taf4a*^{IEC} mice in contrast to T cells and neutrophils.

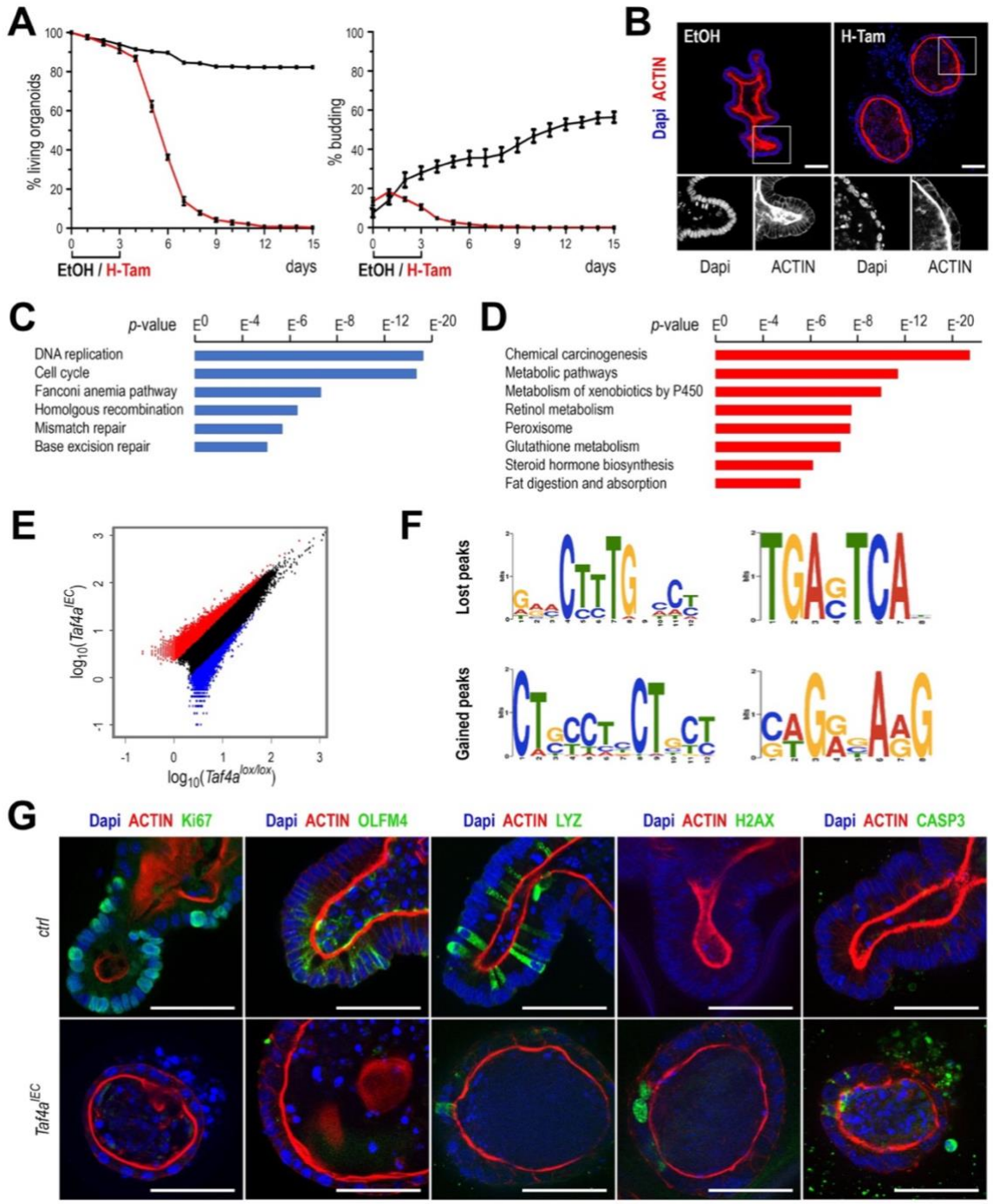


Figure 52. Effect of *Taf4a* Inactivation on Organoid Morphogenesis

A. Survival and budding activity after early *Taf4a* gene inactivation. *Taf4a*^{IEC} organoids were plated and treated 2 hours later with H-Tam in EtOH (red line) or EtOH alone (black line) for 3 days. 50 to 100 3D structures were counted at each time point. **B**. Morphology of the *Taf4a*^{IEC} organoids at day 5 of culture, i.e. 2 days after the treatment with EtOH (left) or H-Tam (right). Boxed regions are enlarged below. Nuclei are stained with Dapi (blue) and the actin network with Phalloidin (red). Bars are 50 µm. **C**. Top 6 enriched KEGG pathways in downregulated genes in H-Tam treated *Taf4a*^{IEC} *vs* *Taf4a*$^{lox/lox}$ organoids at day 3 of culture. **D**. Top 8 enriched KEGG pathways in upregulated genes in H-Tam treated *Taf4a*^{IEC} *vs* *Taf4a*$^{lox/lox}$ organoids. **E**. Scattered plots of the ATACseq peaks in H-Tam treated *Taf4a*^{IEC} and *Taf4a*$^{lox/lox}$ organoids. Blue and red dots respectively represent lower (lost) and higher (gained) peaks in *Taf4a*^{IEC} compared to *Taf4a*$^{lox/lox}$ organoids. **F**. Most frequent DNA binding sites found in the lost and gained peaks, respectively. **G**. Immunofluorescence detection of the indicated proteins in *Taf4a*$^{lox/lox}$ (ctrl) and *Taf4a*^{IEC} organoids treated with H-Tam at day 5 of culture, i.e. 2 days after treatment with H-Tam). Bars are 50 µm.

5. Impact of *Taf4a* Inactivation on Organoid Morphogenesis

Since the loss of function of TAF4a in the intestinal epithelium *in vivo* impacts not only epithelial homeostasis but also the composition of the microbiota and the repertoire of stromal immune cells, crypt-derived 3D organoids were developed to investigate the impact of *Taf4a* invalidation in the gut epithelium *ex vivo* in a constant acellular microenvironment. For this purpose, organoids were established from the ileum of Tam-free $Taf4a^{IEC}$ and control $Taf4a^{lox/lox}$ mice. As expected, organoids of both genotypes generated typical 3D structures with crypt-like buddings lined by a single polarized epithelium when cultured in standard culture medium. The same applied for $Taf4a^{lox/lox}$ organoids treated with Hydroxytamoxifen (H-Tam) during the first 3 days of culture. On the contrary, treating $Taf4a^{IEC}$ organoids with H-Tam resulted in altered 3D structures characterized by a flat cuboidal epithelium, the reduction of buddings outgrowth, and the accumulation of cell debris and DNA within the lumen and even outside the spheroids, as illustrated at day 5 of culture (Figure 52.A-B). Ultimately, these disorganized spheroids degenerated, indicating compromised morphogenesis and survival upon loss of *Taf4a*.

The molecular basis of the defects resulting from *Taf4a* inactivation was investigated by RNAseq performed at day 3 of culture (at the end of the H-Tam treatment), a stage at which the cellular phenotype is still mild but molecular alterations expected to be already ongoing. Compared with control H-Tam-treated $Taf4a^{lox/lox}$ organoids, treated $Taf4a^{IEC}$ organoids showed 3271 deregulated genes (**Table 4**). KEGG annotation of the 1625 downregulated genes identified DNA replication / cell cycle and DNA repair as the major pathways affected by the loss of TAF4a (Figure 52.C). Noteworthy, 182 out of the 510 genes of the intestinal stem cell signature belong to the list of downregulated genes, including receptor genes involved in the control of the dynamic function of CBC: *Fzd7*, *Fzd2*, *Lgr5*, *Notch1* and *Tnfrsf19*. The corresponding ligand genes expressed by Paneth cells of the stem cell niche were also downregulated: *Wnt3*, *Wnt5a*, *Wnt9b*, *Dll3*, *Jag2*, *Egfl8*. Additional genes encoding differentiation markers of mature intestinal cell types were perturbed, among which *Defa17/20/21/24/26* and *Mmp7* for Paneth cells, *Muc4/20* and *Tff2* for goblet cells, *Chga/b* and *Cck* for enteroendocrine cells, and *Aldoa* and *Fabp5* for enterocytes. However, other marker genes were upregulated such as *Cdhr2/5*, *Dpp4*, *Fabp1/2/6*, *Mgam*, *Sis*, *Vill* and *Lyz1*. KEGG analysis of the upregulated genes identified pathways related to chemical carcinogenesis and metabolic processes (Figure 52.D).

ATACseq performed in parallel with RNAseq provided a broad picture of chromatin remodeling linked to the loss of TAF4a (Figure 52.E). Cumulating whole genome peaks from H-Tam treated control $Taf4a^{lox/lox}$ and mutant $Taf4a^{IEC}$ organoids gave a total of 118500 peaks, 94386 of them being unchanged between both conditions whereas 11283 were lost and 12831 gained upon *Taf4a* invalidation (**Table 5**). Lost peaks were enriched in DNA-binding sequences for nuclear receptors (HNF4a/g, PPARa/g, RXRa/g, NR2C2, NR2F6) and for JUN- or FOS-containing dimers, whereas gained peaks were enriched in DNA-binding sequences for Zinc finger transcription factors (GATAs, SP2/4, CTCF) (Figure 52.F). 1675 of the differential peaks mapped into proximal promoter gene regions (-1000 to +100 bp from transcription start sites), 93 of them being associated with downregulated genes identified by RNAseq in $Taf4a^{IEC}$ organoids. KEGG pathway analysis associated these 93 genes to homologous recombination (p-value=0.0056) and DNA replication (p-value=0.0086).

Immunofluorescence staining at day 5 of culture (2 days after the end of H-Tam treatment) corroborated the molecular results (Figure 52.G). Indeed, compared with control H-Tam-treated *Taf4a*$^{lox/lox}$ organoids, *Taf4a* invalidation in *Taf4a*^{IEC} organoids led to a strong reduction of cell proliferation visualized by Ki67 staining. The decline in cell proliferation was associated with the loss of the stem cell marker OLFM4 and the decay of Lysozyme for Paneth cells of the stem cell niche. In addition, considering the number of genes involved in DNA replication and repair that exhibited downregulation and promoter conformation changes, chromatin integrity was analyzed by gH2AX immunostaining. Noteworthy, a strong punctate staining was observed selectively in *Taf4a*^{IEC} organoids indicating DNA damage. In line with this, activated Caspase-3 labeling apoptotic bodies was detected in *Taf4a*^{IEC} but not control *Taf4a*$^{lox/lox}$ organoids. Taken together, these data indicate that the loss of TAF4a triggers defects in DNA integrity and replication leading to cell apoptosis, which perturbs the dynamic function of the stem/progenitor compartment and elicits organoid degeneration and death.

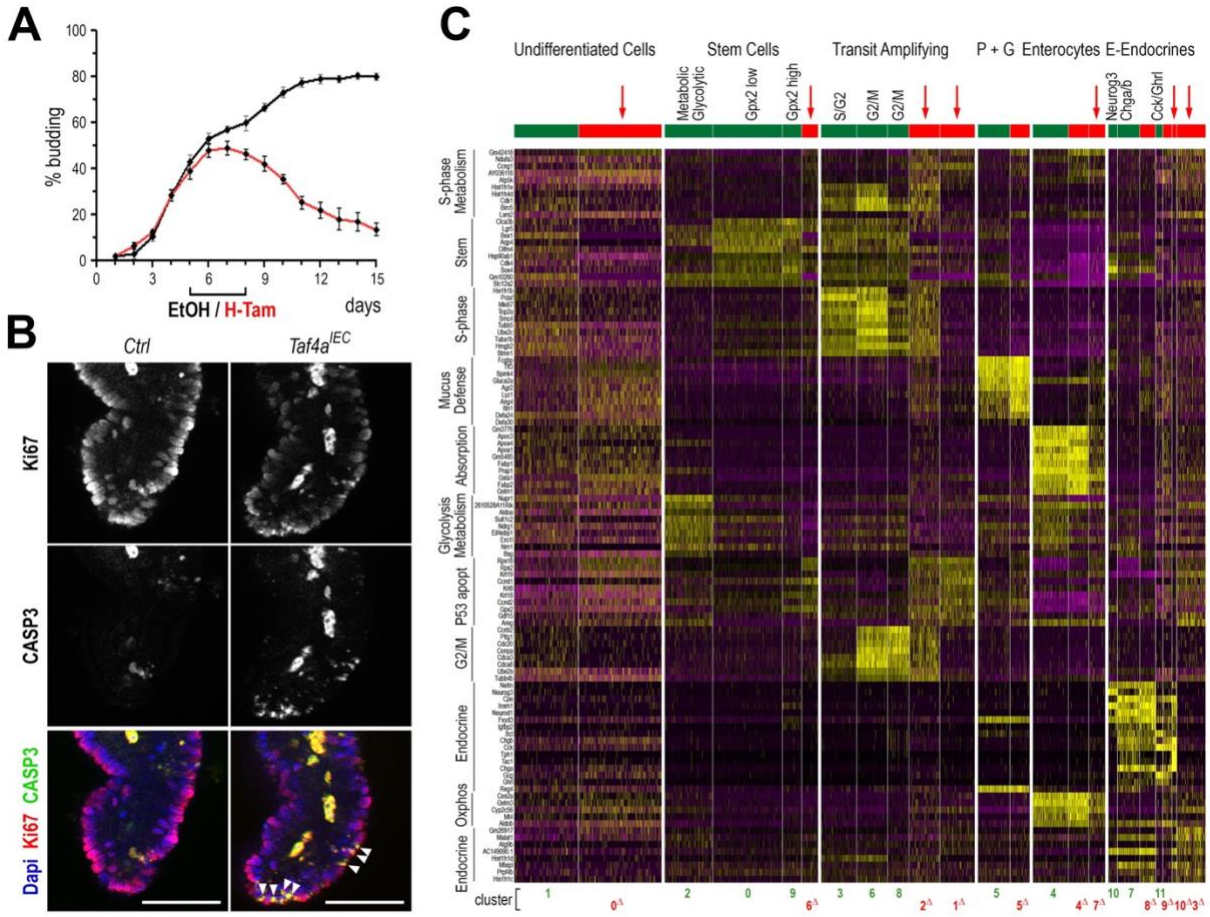


Figure 53. Effect of *Taf4a* Inactivation on Organoid Homeostasis

A. Budding activity after late *Taf4a* gene inactivation. *Taf4a*^{IEC} organoids were plated and treated with H-Tam (red line) or EtOH (black line) at days 5 to 7 of culture. 50 to 100 3D structures were counted at each time point. **B.** Immunodetection of proliferating cells (Ki67) and apoptotic bodies (activated Caspase-3) at day 8 of culture in H-Tam treated *Taf4a*^{IEC} *vs* *Taf4a*$^{lox/lox}$ (ctrl) organoids. White arrowheads show co-staining. Bars are 50 µm. **C.** Heatmap of the top markers identified by sc-RNAseq at day 7 of culture in *Taf4a*^{IEC} and *Taf4a*$^{lox/lox}$ organoids treated with H-Tam between days 5 to 7. Cell clusters numbered 0 to 11 were from control *Taf4a*$^{lox/lox}$ organoids (in green), and those numbered 0^{D} to 10^{D} from *Taf4a*^{IEC} organoids (in red). P + G are Paneth and goblet cells; E-Endocrine are enteroendocrine cells.

6. Impact of *Taf4a* Inactivation on Organoid Maintenance

Having demonstrated the negative effect of the loss of TAF4a on organoid morphogenesis, we addressed its consequence on organoid maintenance, when buddings are already formed comprising stem cells and their niche. For this purpose, *Taf4a*^{IEC} and control *Taf4a*$^{lox/lox}$ organoids were grown for 5 days in the absence of H-Tam, and then H-Tam was added to the culture medium for 3 days. Under this condition, unlike *Taf4a*$^{lox/lox}$ organoids *Taf4a*^{IEC} organoids progressively loose buddings generated during the first 5 days of culture, stopped growing but remained generally alive up to day 15 in culture (Figure 53.A). However, they could no longer be passaged. As illustrated in Figure 53.B, immunofluorescence staining of Ki67 demonstrated reduced cell proliferation in *Taf4a*^{IEC} compared with *Taf4a*$^{lox/lox}$ organoids. Moreover, activated Caspase3 immunodetection revealed the presence of apoptotic bodies at the level of Ki67-positive cells, suggesting that proliferative cells enter apoptosis in the absence of TAF4a. Single-cell RNAseq was performed at day 8 of culture (at the end of the H-Tam treatment) to decipher the cellular impact of *Taf4a* invalidation (Figure 53.C). Cell clustering comparison between H-Tam treated *Taf4a*^{IEC} and *Taf4a*$^{lox/lox}$ organoids revealed that the population of active stem cells was strongly reduced after TAF4a depletion (cluster 6^{D} *vs* clusters 0, 2 and 9) and that the remaining stem cells acquired an apoptotic signature. Similarly, the population of transit amplifying cells was also reduced while the remaining proliferative cells also underwent apoptosis (clusters 1^{D}, 2^{D} *vs* 3, 6, 8). At this stage of culture, there was no major effect on Paneth and goblet cells (cluster 5^{D} *vs* 5), whereas the population of enterocytes showed the emergence of poorly differentiated cells (clusters 4^{D}, 7^{D} *vs* 4). The pattern of enteroendocrine cells was also perturbed, characterized by a reduction of precursor cells and the presence of poorly differentiated cells (clusters 3^{D}, 8^{D}, 9^{D}, 10^{D} *vs* 7, 10, 11). Finally, in line with the decline of stem and transit amplifying cell populations and with the appearance of poorly differentiation cells, the proportion of non-proliferative undifferentiated cells increased in *Taf4a*^{IEC} organoids (clusters 0^{D} *vs* 1).

Therefore, the absence of TAF4a not only affects organoid morphogenesis and buddings formation, but it also compromises the maintenance of buddings by perturbing stem cells and their niche.

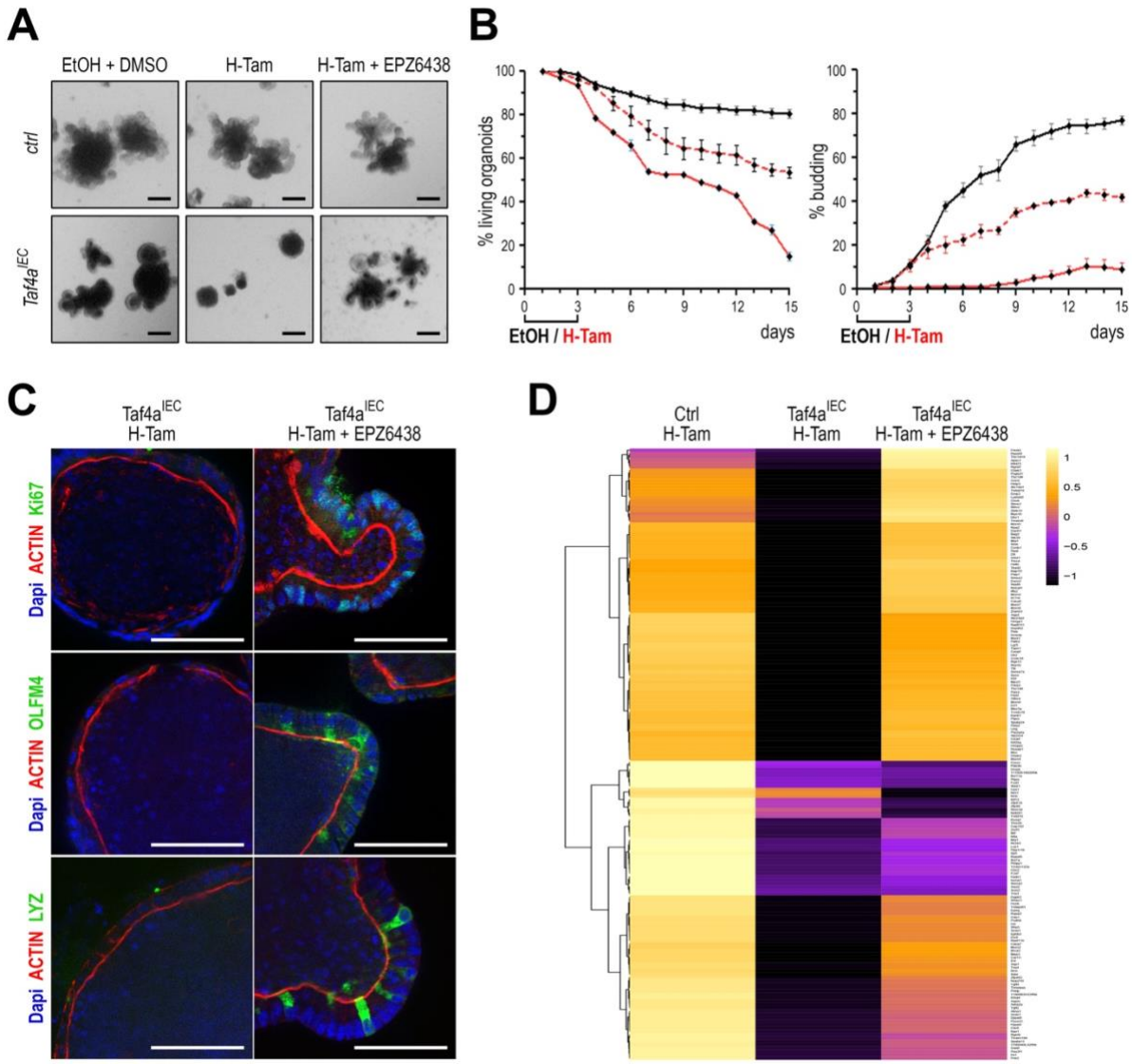


Figure 54. Rescue of *Taf4a* Inactivated Organoids by Polycomb Complex Inhibition

A. Morphology at day 15 of culture of *Taf4a*$^{lox/lox}$ (ctrl) and *Taf4a* organoids treated with EtOH + DMSO, or with H-Tam + DMSO (2h to 72h after plating), or with H-Tam (2h to 72h after plating) + EPZ6438 (2h to 15 days after plating). Bars are 200 µm. **B**. Survival and budding activity in *Taf4a*^{IEC} organoids treated with EtOH + DMSO (black line), with H-Tam + DMSO (red line), or with H-Tam + EPZ6438 (red dotted line). 50 to 100 3D structures were counted at each time point. **C**. Immunostaining of the indicated proteins in *Taf4a*^{IEC} organoids treated H-Tam + DMSO or with H-Tam + EPZ6438 at day 15 of culture. Bars are 50 µm. **D**. Heatmap of genes of the stem cell signature in H-Tam-treated *Taf4a*$^{lox/lox}$ (ctrl) and *Taf4a*^{IEC} organoids at day 3 of culture and of H-Tam + EPZ6438 treated *Taf4a*^{IEC} organoids at day 15 of culture.

7. Rescue of *Taf4a*-Depleted Organoids

The fact that *Taf4a* invalidation caused chromatin remodeling and transcription repression of a number of genes of the cell proliferation and stem cells signatures prompted us to look at Polycomb-repressive complexes (PRC) which are important regulators of epigenetic silencing in developmental processes and cancer (Piunti and Shilatifard, 2021)[7]. Interestingly, examining the RNAseq data of H-Tam-treated *Taf4a*^{IEC} *vs* control *Taf4a*$^{lox/lox}$ organoids revealed that several genes encoding core or associated members of the PRC1 complexes were downregulated (*Pcgf5*, *Cbx2, -5, -6, -7*, *Kdm2d*, *Pcgf5*, *Timeless*) whereas genes of the PRC2 complex remained essentially unaltered except *Rbbp7* and *Phf1* which were respectively down- and upregulated. Based on this observation, we investigated whether PRC inhibition could rescue *Taf4a*-depleted organoids. For this purpose, *Taf4a*^{IEC} and control *Taf4a*$^{lox/lox}$ organoids were treated as above with H-Tam and EPZ6438, a selective inhibitor of the histone methyltransferase activity of the EZH2 component of PRC2, was added at the same time and maintained throughout the culture up to day 15. As shown in Figure 54.A-B and in Figure 55.A, when *Taf4a* invalidation was induced immediately (2 hours) after plating, EPZ6438 allowed survival and budding formation in *Taf4a*^{DIEC} organoids. Similarly, EPZ6438 treatment also counteracted buddings degeneration induced by *Taf4a* invalidation at day 5 of culture (Figure 55.B). Immunofluorescence staining at day 15 of culture in *Taf4a*^{IEC} organoids treated with H-Tam + EPZ6438 after plating showed that the restored buddings were populated with proliferating cells (Ki67), stem cells (OLFM4) and Paneth cells (Lyz) (Figure 54.C). In parallel, we established at day 15 of culture the transcriptome of *Taf4a*^{IEC} organoids treated with H-Tam + EPZ6438 and compared it with that of *Taf4a*-depleted *Taf4a*^{DIEC} organoids after the 3-days treatment with H-Tam (Table 6). Inhibition of PRC2 by EPZ6438 restored the expression of 58.4% (948/1623) of the genes downregulated by *Taf4a* inactivation and conversely it reduced the expression of 57.5% (947/1646) of the genes upregulated upon loss of *Taf4a*. In particular, EPZ6438 reestablished the expression of 115 over the 182 genes of the intestinal stem cell signature that were downregulated in the absence of TAF4a (Figure 54.D). Altogether, these results unveil a major role played by PRC in the global alteration of gene expression resulting from *Taf4a* invalidation, which can be override by inhibiting PRC2 activity.

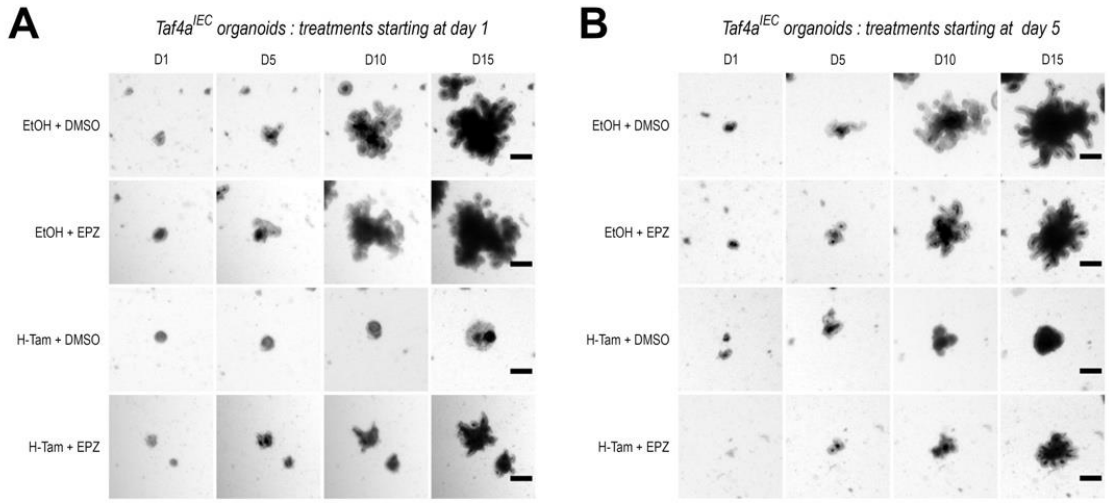


Figure 55. Kinetics of Morphological Defects in *Taf4a*^{IEC} Organoids Following *Taf4a* Inactivation and Rescue by EPZ6438.

Pictures show the evolution of the same organoid in each condition from days 1 to 15 of culture. Treatments started at day 1 (**A**) or at day 5 (**B**) of culture. Bars are 200 µm.

DISCUSSION

The goal of this work was to investigate the role of the TAF4a component of the general transcriptional machinery in the gut. For this purpose, we used gene invalidation to deplete TAF4a from the TFIID complex and addressed its consequences in the intestinal epithelium in vivo during gut morphogenesis, in adult gut homeostasis and ex vivo in organoid models. The results obtained in animals show (i) that *Taf4a* knockout compromises the formation of the crypt-villus axis in embryos, (ii) that it perturbs the dynamic homeostasis of the adult gut epithelium by altering the stem cell activity and the balance between cell proliferation and differentiation, and (iii) that it modifies the cell migration along the crypt villus. These changes impact on the composition of the microbiota and, through increased paracellular permeability, the cellular composition of the stroma. Consequently, it has deleterious consequences in an intestinal tumor-prone context. The results obtained in organoids show (i) that TAF4a is required for organoid morphogenesis and maintenance, (ii) for stem cells emergence and (iii) to ensure the transcription of genes of the intestinal stem cell signature. Moreover, (i) in the absence of TAF4a, gene silencing is mediated by Polycomb complexes and (ii) inhibition of the trimethylation activity of the EZH2 component of Polycomb Repressive Complex-2 compensates for TAF4a deficiency as regards stem cell emergence and transcription of genes of the stem cell signature.

I. Involvement of TAF4a in the Murine Intestinal Development and Homeostasis

A. Taf4 and the General Organization of the Intestinal Epithelium

a. Taf4 and the Intestinal Endoderm Development

To explore Taf4a function in intestinal development, we used the constitutive model *Taf4a*$^{lox/lox}$*::VilCre* (*Taf4a*^{IEndoC}) The inactivation to Taf4a presumably starts around 11-12 days post-coïtum under the control of the promotor of the *Villin* gene. We observed that in this mouse model, TAF4a invalidation was progressive during the intestinal development certainly due to a long half-life of the protein Taf4. The endoderm showed a partial TAF4a expression at E14.4, which progressively became totally absent at E18.5, before birth. It is noteworthy that this progressive process of Taf4 loss was concomitant to the stage of villi formation during intestinal development. Indeed, under physiological conditions in control mice, villi morphogenesis starts around E14.5 through endodermal mesenchymal cell interactions, resulting in villi containing mature differentiated cells at E18.5. Proliferating cells are confined in the intervillus area, between the villi.

Taf4a depletion in the intestinal endoderm leads to post-natal death of the neonates, associated to an altered epithelium with shorter and less dense villi. Moreover, it seems that the defect in villi morphogenesis parallels the decrease of TAF4a levels. This dysgenesis of villi is also associated with a decrease of the markers of differentiation of the different cell lineages and an increase in cells undergoing apoptosis.

Despite the loss of the villous structure, intervillus areas with proliferating cells are still detected, expressing the proliferation markers SOX9 and KI67. However, these cells present an alteration of the level and pattern of the HNF4 7-9 isoforms presumably associated with proliferation [203]. They also lost Olfm4 expression, a regulator of Wnt/β-catenin signaling in ISCs and progenitor cells which may have anti-apoptotic activity [217].

Taken together, these results show that Taf4a plays a critical role in intestinal development and maturation and then for the survival of mice at birth. This is linked to the drastic change in the pattern of gene expression affecting general features of the digestive cells and their ability to communicate with their microenvironment.

b. Taf4a General functions in Adult Intestinal Epithelium Cell Homeostasis

In order to study the functions of *Taf4a* in adult intestinal homeostasis, we developed the conditional model $Taf4a^{lox/lox}::VilCre^{ERT2}$ ($Taf4a^{IEC}$), to by-pass post-natal lethality of *Taf4a* invalidation in the endoderm. The intestine is a long-segmented tract with differential function and organization. Analyses of TAF4 functions in intestinal homeostasis has been presented here on the ileum but we also performed the same analyses in the duodenum and jejunum with similar results. The ileum segment presents all the physiological function of the intestine, a crypt-villus organization and all the differentiated cell, with the highest proportion of enteroendocrine cells and goblet cells along the small intestine. Furthermore, it is also the region of the small intestine with the highest content in microbiota and in cells of the immune repertoire.

Like in embryos, the controlled induced loss of *Taf4a* in adults also greatly impaired intestinal physiology, yet without major morphological effect but leading to a reduced survival of the animals, predominantly females. Indeed, only about 20% of mice did survive by day 120 after the induction of *Taf4a* depletion. The phenotype was associated with body weight-loss.

The results we have obtained show that the loss of function of TAF4a alters both proliferating and differentiated area of the gut epithelium. Strikingly, the reduced activity of

stem cells suggested by the reduced immunolabeling of OLFM4 was compensated by increased cell proliferation and followed by a higher turnover. However, the presence of apoptotic cells along the villi indicated abnormal cell differentiation. This is in line with the reduced expression of markers of differentiated cells including enterocytes, goblet cells, enteroendocrine and tuft cells.

c. Refinement of Taf4a Implication by Transcriptomic Analysis

The RNAseq analyzes performed in embryos and adult animals allowed analyzing the general transcriptomic changes resulting from the loss of TAF4a. We observed that at both developmental stages, more genes are down-regulated by *Taf4a* invalidation (1160 for $Taf4a^{IendoC}$ and 769 for $Taf4a^{IEC}$) than up-regulated (173 and 284). About ¼th of the deregulated genes were shared at both developmental stages indicating a minority of commune functions regulated by TAF4a during embryogenesis and homeostasis compared to diverse functions. However, an intriguing result is that only a limited number of genes exhibited an altered expression among the whole number of expressed genes. The same was also observed by loss of *Taf4a* in the liver and in the pancreas, suggesting that the TFIID complex can function even in the absence of TAF4a for the transcription of many genes. Another gene, *Taf4b*, can compensate the absence of *Taf4a*, but it does not seem to happen in the gut since we did not observe any obvious expression of this factor in the *Taf4a*-deficient intestinal epithelium of embryos and adult mice.

Our results showed that many genes implicated in enterocyte metabolic and absorption functions are drastically decreased. Among them are glucose, oligopeptides, lipid, amino acid transporters and more interestingly folate transporter SLC46a1 (also known as PCFT). Folate or B9 vitamin cannot be synthetized *de novo* and known to be essential for tissue physiology, mental health. Mice conditionally depleted for SLC46a1 present a continuous folate-deficit causing growth retardation, anemia, ineffective hematopoiesis with around 90% of lethality, 10-12 weeks after their birth [248].

Among the down-related gene in enteroendocrine cells, many are hormone related like *Gip.* Gip (glucose-dependent insulinotropic peptide) is a postprandial hormone known to stimulate pancreatic β cells mass and their production of insulin in response to glucose. Importantly, this hormone is also inducing lipid storage and fat mass production by upregulating the lipogenesis [249]. This later point may be correlated to the strongly reduced fat mass observed in *Taf4a* depleted adult mice. `

Altogether, these data suggest a global alteration of intestinal genetic program leading to defects of nutrient digestion and absorption, explaining the phenotype observed on mouse health and survival rate of Taf4^{IEC} mice in the long term.

B. Taf4a and the Crypt Base Columnar Stem Cells

Taf4a^{IendoC} and *Taf4a*^{IEC} mice present a reduced OLM4 expression in both RNA and protein levels. This protein normally expressed by embryonic stem cells and adult Crypt Base Columnar (CBCs) stem cells is likely to control cell proliferation and survival as it has proliferation inhibitor and anti-apoptotic functions [217], which makes sense with our observations. Paneth cells are the only differentiated cells of the stem cell niche with a stem cell support function, but they are also the differentiated cells with the higher lifespan in the gut epithelium. RNAseq data did not identify any significant deregulation of Paneth cell genes involved in the maintenance of stem cells like Wnt's or the Notch ligands. It remains therefore to elucidate if defects in Paneth cell functions may contribute to the reduced number of stem cells observed after *Taf4a* invalidation. Conversely, several Paneth cell genes involved in antimicrobial activities were downregulated, which could be related to the modification observed at the level of the microbiota.

To strengthen our analyze of the adult CBCs, we generated the *Taf4a*$^{lox/lox}$*::Lgr5-GFP-Cre*^{ERT2} (*Taf4a*^{CBC}) mouse model where *Taf4a* invalidation is selectively targeted in the adult Lgr5+ CBCs stem cells. As expected, the inactivation of *Taf4a* was mosaic and also penetrant since *Taf4a*-inactivated crypts of *Taf4*^{CBC} mice gave rise to progenitor and differentiated cells that populate both crypts and villi. In addition, like in *Taf4a*^{IEC} mice, we did not observe any apoptotic figure at the level of the *Taf4a*-inactivated crypts of *Taf4*^{CBC}-mice (Caspase-3 staining), suggesting that the reduced expression of OLFM4 after *Taf4a* invalidation rather corresponds to a reduction of stem cell activity than of stem cell number *in vivo* in mice.

II. Pathological Consequences of *Taf4a* Inactivation in Adult Intestinal Epithelium

A. Permeability and the cellular microenvironment: Microbiota, and Immune Cells

In *Taf4a*^{IEC} mice, enterocytes expressed lower level of Alkaline Phosphatase, Paneth cells had impaired expression of lysozyme and α-defensins, and the expression of Mucin 4 and 20 was altered in goblet cells. All of these proteins are implicated in the control of the microbiota and in the protection of the epithelium. Alteration in their expression is often associated with

an unbalance of the intestinal cellular micro-environment involving notably the microbiota and immune cells as observed in *Taf4a*^{IEC} mice. *Taf4a*^{IEC} mice also presented diarrhea with a swollen caecum distended by gaze, all signs of a dysbiosis. The dysbiosis was confirmed by the analysis of the bacterial 16S RNA which unveiled a decrease of microbiota diversity associated with an increase of pathogenic bacteria such as *Helicobacteriaceae*. This dysbiosis is accompanied by a predicted change of intestinal immune cell population showing higher infiltration by basophils and B-derived cells, and by an increased permeability of the epithelial barrier. Interestingly, only the paracellular permeability, but not the transcellular permeability of the epithelium, was altered upon loss of function of TAF4a. The reason for the increased paracellular permeability remains however elusive. Indeed, none of the genes encoding junctional complexes were downregulated whether it be genes for adherens, gap or tight junctions or desmosomes. Further analysis is required by microscopy to address if TAF4a inactivation could alter the incorporation of some protein(s) in the junctional complexes. Nevertheless, it is interesting to realize that the inactivation of only one component of the general transcriptional machinery was able to affect not only the epithelial cells themselves but also their non-epithelial cellular environment. Change in the microbiota population may likely be due to impair interactions between epithelial cells and bacteria. Another possibility is that the altered digestive and absorptive functions of the gut epithelium following loss of TAF4a activity modify the molecular composition of the lumen, thereby fostering the growth of specific subclasses of bacteria leading to dysbiosis.

Based on all these observations, experiments are ongoing to investigate if the loss of function of TAF4a has any impact on the sensitivity to an inflammatory setting using the model of intestinal inflammation induced by DSS in the drinking water.

B. Taf4 and Colorectal Cancer

To explore the role of *Taf4a* in gut tumorigenesis, we crossed *Taf4a*^{IEC} mice with *Apc*$^{+/D14}$ mice, the latter spontaneously developing tumors in the small intestine and more rarely in the colon. For this part of the project, we choose to perform the experiment on a cohort of male mice which exhibited a longer lifespan than females after *Taf4a* invalidation. We observed a strong increase in tumor burden although adenocarcinoma exhibited a similar histology in *Taf4a*^{IEC}*::Apc*$^{+/D14}$ mice as in *Apc*$^{+/D14}$ mice. At this level of investigation, we suggest that the increased number of tumors reflects the increased proliferation in the intestinal epithelium following loss of function of TAF4a. However, we cannot rule out any effect of the change of

the microbiota and/or of the repertoire of immune cells that could promote tumor initiation and growth.

Although, the results in mice show a clear impact of the loss of TAF4a on intestinal tumorigenesis, data in human from the "The Human Protein Atlas" (https://www.proteinatlas.org/ENSG00000130699-TAF4/pathology/colorectal+cancer/COAD) are mitigated. Indeed, even if patients with a "high" level of TAF4a expression in the tumor exhibit an about 10% better survival rate at 5 years than patients with a "low" level of TAF4a, this difference is not statistically significant (p=0.11).

III. Implication of Taf4a in Intestinal Epithelial Cell in a cell-free Microenvironment

At the interface of the external medium (the lumen) and the inner of the body, the intestinal epithelium is continuously targeted by and reciprocally controls signals coming from the lamina propria, lumen, microbiota, immune cells, and others. The phenotype of *Taf4a* knockout animals provides one more example of this complex crosstalk. Therefore, to study the impact of the loss of *Taf4a* in the intestinal epithelium in a stable and acellular microenvironment, we generated ileum organoids from *Taf4*^{IEC} and control mice.

The model of organoids provided very important results. First, the most significant result obtained from *Taf4* invalidation *ex vivo* revealed that the phenotype observed in organoids showed similarities but also differences with the phenotype in mice. The main difference is that *Taf4a*-deficient crypt-villi remain viable in mice whereas *Taf4a*-knockout organoids collapse. In line with this, cell proliferation is maintained in mice whereas proliferative cells become apoptotic in organoids. Although we don't have any clear explanation for this, it is likely that signals coming from the stroma, and missing in the organoid condition, may help epithelial survival in mice. This putative action of the stroma could act on the whole epithelium or specifically on the stem cell compartment by recruiting accessory/reserve stem cells. Another difference between mice and organoids observed by RNAseq is that the expression of cell differentiation markers is strongly compromised in mice whereas it is preserved in organoids. However, this raw result is most likely a consequence of the protocol we used. Indeed, in mice RNAseq were performed 10 days after the induction of *Taf4a* invalidation, thus after 2~cycles of cell turnover. On the contrary, it was done at day 3 of culture in organoids, meaning that the formation of organoids is just ongoing and differentiated cells from the previous culture of organoids are still in place.

In contrast with differences, the most important similarity between mice and organoids is the reduction/loss of active stem cells after loss of function of *Taf4a*. We conclude from this major result that TAF4a is required for the biology of active CBC/Lgr5+ stem cells in the gut. This is visualized both at the protein level by immunofluorescence detection of OLFM4 and at the level of the RNA signature of Lgr5+ cells. Moreover, organoids demonstrate that the complex crosstalk between stem cells and their niche is perturbed since the receptors of morphogens/growth factors expressed by stem cells are downregulated, together with that of their ligands produced by the neighboring Paneth cells. In line with this, an important set of downregulated genes encodes factors associated with DNA replication, cell cycle and DNA repair. Finally, the last argument supporting the notion that TAF4a is required for intestinal stem cells and their proliferative progeny comes from the single-cell RNAseq analysis showing a strong reduction in the number of stem and proliferative cells, and the emergence of apoptotic and non-proliferative undifferentiated cells. Together, these results demonstrate that TAF4a is crucial for the functional organization of the stem cell compartment with the niche, which results in the absence of budding outgrowth in *Taf4a*-deficient organoids. Finally, in the absence of TAF4a accessory/reserve cells putatively present in organoids cannot overcome the loss of CBC deficiency.

Directly or indirectly, TAF4a has a strong impact on the functional organization of the chromatin. On the one hand, ATACseq data illustrate that this factor controls the dynamics and accessibility of several thousands of chromatin loci in close vicinity or more distant to protein coding genes. Although the presence of TAF4a-dependent ATACseq site in the proximity of protein coding genes is expected according to the role of TFIID in the general transcription machinery, distal sites can be explained by the fact that other regions of the genome, like for instance enhancer regulatory sequences, can also be transcribed [250].

Beside the modification of chromatin accessibility visualized by ATAQseq, the most prominent result of this work is that the inhibition of the Polycomb Repressive Complex 2 allows to reverse the deleterious phenotype of *Taf4a* invalidation in organoids. This concerns organoid survival, crypt-like budding formation, emergence of CBCs and Paneth cells in the buddings, and re-expression of genes of the stem cell signature. Polycomb group complexes are multiprotein complexes acting as histone modifiers whose complementary action decreases chromatin accessibility and leads to gene silencing. PRC1 is implicated in early embryonic development, intestinal stem-cell functions and identity [149], while PRC2 is an actor of progenitor identity and differentiated lineage balance [150][151][152]. Our results show that

several components of the PRC1 complex, involved in maintaining the repressed stage of chromatin by Histone 2A monoubiquitination, are downregulated by *Taf4a* invalidation, unlike the components of the PRC2 complex endowed with trimethylation activity of Lysine 27 of Histone 3 that represses transcription. Previous studies have shown that PRC1 can dissociate the RNA polymerase II preinitiation complex [251], indicating a functional link between Polycomb complexes and the general transcription machinery. Our results identified a new functional link in that inhibiting the trimetylation activity of the EZH2 component of PRC2 widely restores the pattern of gene expression altered by TAF4a deficiency. Interestingly, this pattern of gene expression is restored even if TAF4a is still absent from the TFIID complex. It suggests that TAF4a is not directly required to ensure gene transcription by RNA polymerase II but rather to keep chromatin open and accessible for the transcription machinery.

In summary, this work shed new light on the genetic regulation of the development, homeostasis, and pathology of the intestinal epithelium. We identified the TAF4a component of TFIID as a major actor of these processes through the regulation of a limited panel of genes involved in stem cell biology, the stem cell niche, and the proliferation and differentiation compartments. At the molecular level, we identified an antagonistic effect between TAF4a acting as a factor that promotes open chromatin accessible for the general transcription machinery, and PRC2 that defines inhibitory histone marks to silence gene transcription by PRC1. Finally, we also report the impact of TAF4a in pathological settings, especially in cancer. It opens new avenues to investigate new developments of PRC2 inhibitors as therapeutic agents against cancer [252].

REFERENCES

[1] J. P. Medema et L. Vermeulen, « Microenvironmental regulation of stem cells in intestinal homeostasis and cancer », Nature, vol. 474, n° 7351, p. 318-326, juin 2011, **doi: 10.1038/nature10212.**

[2] W. D. Garrison, M. A. Battle, C. Yang, K. H. Kaestner, F. M. Sladek, et S. A. Duncan, « Hepatocyte nuclear factor 4alpha is essential for embryonic development of the mouse colon », Gastroenterology, vol. 130, n° 4, p. 1207-1220, avr. 2006, **doi: 10.1053/j.gastro.2006.01.003.**

[3] A.-L. Cattin et al., « Hepatocyte Nuclear Factor 4α, a Key Factor for Homeostasis, Cell Architecture, and Barrier Function of the Adult Intestinal Epithelium », Mol. Cell. Biol., vol. 29, n° 23, p. 6294-6308, déc. 2009, **doi: 10.1128/MCB.00939-09.**

[4] D. Alpern et al., « TAF4, a subunit of transcription factor II D, directs promoter occupancy of nuclear receptor HNF4A during post-natal hepatocyte differentiation », eLife, vol. 3, p. e03613, sept. 2014, **doi: 10.7554/eLife.03613.**

[5] G. Mengus et al., « TAF4 inactivation in embryonic fibroblasts activates TGF beta signalling and autocrine growth », EMBO J., vol. 24, n° 15, p. 2753-2767, août 2005, **doi: 10.1038/sj.emboj.7600748.**

[6] J. Muñoz et al., « The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent "+4" cell markers », EMBO J., vol. 31, n° 14, p. 3079-3091, juin 2012, **doi: 10.1038/emboj.2012.166.**

[7] A. Piunti et A. Shilatifard, « The roles of Polycomb repressive complexes in mammalian development and cancer », Nat. Rev. Mol. Cell Biol., vol. 22, n° 5, p. 326-345, mai 2021, **doi: 10.1038/s41580-021-00341-1.**

[8] C. Yang, E. Bolotin, T. Jiang, F. M. Sladek, et E. Martinez, « Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters », Gene, vol. 389, n° 1, p. 52-65, mars 2007, **doi: 10.1016/j.gene.2006.09.029.**

[9] A. L. Roy et D. S. Singer, « Core Promoters in Transcription: Old Problem, New Insights », Trends Biochem. Sci., vol. 40, n° 3, p. 165-171, mars 2015, **doi: 10.1016/j.tibs.2015.01.007.**

[10] Y. M. Danino, D. Even, D. Ideses, et T. Juven-Gershon, « The core promoter: At the heart of gene expression », Biochim. Biophys. Acta BBA - Gene Regul. Mech., vol. 1849, n° 8, p. 1116-1131, août 2015, **doi: 10.1016/j.bbagrm.2015.04.003.**

[11] L. Vo Ngoc, Y.-L. Wang, G. A. Kassavetis, et J. T. Kadonaga, « The punctilious RNA polymerase II core promoter », Genes Dev., vol. 31, n° 13, p. 1289-1301, 01 2017, **doi: 10.1101/gad.303149.117.**

[12] J. Kazantseva et K. Palm, « Diversity in TAF proteomics: consequences for cellular differentiation and migration », Int. J. Mol. Sci., vol. 15, n° 9, p. 16680-16697, sept. 2014, **doi: 10.3390/ijms150916680.**

[13] C. Carlberg et F. Molnár, « The Basal Transcriptional Machinery », in Mechanisms of Gene Regulation, C. Carlberg et F. Molnár, Éd. Dordrecht: Springer Netherlands, 2014, p. 37-54. **doi: 10.1007/978-94-007-7905-1_3.**

[14] S. Saxonov, P. Berg, et D. L. Brutlag, « A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters », Proc. Natl. Acad. Sci. U. S. A., vol. 103, n° 5, p. 1412-1417, janv. 2006, **doi: 10.1073/pnas.0510310103.**

[15] N. I. Gershenzon et I. P. Ioshikhes, « Synergy of human Pol II core promoter elements revealed by statistical sequence analysis », Bioinformatics, vol. 21, n° 8, p. 1295-1300, avr. 2005, **doi:**

10.1093/bioinformatics/bti172.

[16] R. K. McGinty et S. Tan, « Nucleosome structure and function », Chem. Rev., vol. 115, n° 6, p. 2255-2273, mars 2015, **doi: 10.1021/cr500373h.**

[17] B. Portha, A. Fournier, M. D. Kioon, V. Mezger, et J. Movassat, « Environnement périnatal : marques épigénétiques et déterminisme des maladies métaboliques de l'adulte. Preuves de concept », Cah. Nutr. Diététique, vol. 48, n° 6, p. 282-297, déc. 2013, **doi: 10.1016/j.cnd.2013.07.002.**

[18] M. C. Thomas et C.-M. Chiang, « The General Transcription Machinery and General Cofactors », Crit. Rev. Biochem. Mol. Biol., vol. 41, n° 3, p. 105-178, janv. 2006, **doi: 10.1080/10409230600648736.**

[19] X. Chen et al., « Structural insights into preinitiation complex assembly on core promoters », Science, vol. 372, n° 6541, p. eaba8490, avr. 2021, **doi: 10.1126/science.aba8490.**

[20] H. B. Sun, J. Shen, et H. Yokota, « Size-Dependent Positioning of Human Chromosomes in Interphase Nuclei », Biophys. J., vol. 79, n° 1, p. 184-190, juill. 2000, **doi: 10.1016/S0006-3495(00)76282-5.**

[21] D. Li, X. Shu, P. Zhu, et D. Pei, « Chromatin accessibility dynamics during cell fate reprogramming », EMBO Rep., vol. 22, n° 2, p. e51644, févr. 2021, **doi: 10.15252/embr.202051644.**

[22] A. Hosseini et S. Minucci, « Chapter 6 - Alterations of Histone Modifications in Cancer », in Epigenetics in Human Disease (Second Edition), vol. 6, T. O. Tollefsbol, Éd. Academic Press, 2018, p. 141-217. **doi: 10.1016/B978-0-12-812215-0.00006-6.**

[23] J. E. Audia et R. M. Campbell, « Histone Modifications and Cancer », Cold Spring Harb. Perspect. Biol., vol. 8, n° 4, p. a019521, avr. 2016, **doi: 10.1101/cshperspect.a019521.**

[24] Y. Zhang et al., « Overview of Histone Modification », Adv. Exp. Med. Biol., vol. 1283, p. 1-16, 2021, **doi: 10.1007/978-981-15-8104-5_1.**

[25] A. Jambhekar, A. Dhall, et Y. Shi, « Roles and regulation of histone methylation in animal development », Nat. Rev. Mol. Cell Biol., vol. 20, n° 10, p. 625-641, oct. 2019, **doi: 10.1038/s41580-019-0151-1.**

[26] S. Aranda, G. Mas, et L. D. Croce, « Regulation of gene transcription by Polycomb proteins », Sci. Adv., vol. 1, n° 11, p. e1500737, déc. 2015, **doi: 10.1126/sciadv.1500737.**

[27] X. Gu et al., « CBX2 Inhibits Neurite Development by Regulating Neuron-Specific Genes Expression », Front. Mol. Neurosci., vol. 11, p. 46, 2018, **doi: 10.3389/fnmol.2018.00046.**

[28] S. A. Garcia-Moreno, Y.-T. Lin, C. R. Futtner, I. M. Salamone, B. Capel, et D. M. Maatouk, « CBX2 is required to stabilize the testis pathway by repressing Wnt signaling », PLOS Genet., vol. 15, n° 5, p. e1007895, mai 2019, **doi: 10.1371/journal.pgen.1007895.**

[29] N. Coré, F. Joly, A. Boned, et M. Djabali, « Disruption of E2F signaling suppresses the INK4a-induced proliferative defect in M33-deficient mice », Oncogene, vol. 23, n° 46, p. 7660-7668, oct. 2004, **doi: 10.1038/sj.onc.1207998.**

[30] C. Baumann, X. Zhang, et R. De La Fuente, « Loss of CBX2 induces genome instability and senescence-associated chromosomal rearrangements », J. Cell Biol., vol. 219, n° 11, p. e201910149, nov. 2020, **doi: 10.1083/jcb.201910149.**

[31] S. Zheng, P. Lv, J. Su, K. Miao, H. Xu, et M. Li, « Overexpression of CBX2 in breast cancer promotes tumor progression through the PI3K/AKT signaling pathway », Am. J. Transl. Res., vol. 11, n° 3, p. 1668-1682, 2019. **PMID: 30972192**

[32] L. J. Wheeler et al., « CBX2 identified as driver of anoikis escape and dissemination in high grade serous ovarian cancer », Oncogenesis, vol. 7, n° 11, p. 92, nov. 2018, **doi: 10.1038/s41389-018-0103-1.**

[33] H. Zhou, Y. Xiong, Z. Liu, S. Hou, T. Zhou, et Y. Xiong, « Expression and prognostic significance of CBX2 in colorectal cancer: database mining for CBX family members in malignancies and vitro analyses », Cancer Cell Int., vol. 21, n° 1, p. 402, juill. 2021, **doi: 10.1186/s12935-021-02106-4.**

[34] A. N. Mardaryev et al., « Cbx4 maintains the epithelial lineage identity and cell proliferation in the developing stratified epithelium », J. Cell Biol., vol. 212, n° 1, p. 77-89, janv. 2016, **doi: 10.1083/jcb.201506065.**

[35] B. Wang et al., « Chromobox Homolog 4 Is Correlated with Prognosis and Tumor Cell Growth in Hepatocellular Carcinoma », Ann. Surg. Oncol., vol. 20, n° 3, p. 684-692, déc. 2013, **doi: 10.1245/s10434-013-3171-7.**

[36] C. Hu et al., « CBX4 promotes the proliferation and metastasis via regulating BMI-1 in lung cancer », J. Cell. Mol. Med., vol. 24, n° 1, p. 618-631, janv. 2020, **doi: 10.1111/jcmm.14771.**

[37] X. Wang et al., « CBX4 Suppresses Metastasis via Recruitment of HDAC3 to the Runx2 Promoter in Colorectal Carcinoma », Cancer Res., vol. 76, n° 24, p. 7277-7289, déc. 2016, **doi: 10.1158/0008-5472.CAN-16-2100.**

[38] A. Santanach et al., « The Polycomb group protein CBX6 is an essential regulator of embryonic stem cell identity », Nat. Commun., vol. 8, n° 1, p. 1235, nov. 2017, **doi: 10.1038/s41467-017-01464-w.**

[39] J. Wang, H. He, Q. Jiang, Y. Wang, et S. Jia, « CBX6 Promotes HCC Metastasis Via Transcription Factors Snail/Zeb1-Mediated EMT Mechanism », OncoTargets Ther., vol. 13, p. 12489-12500, 2020, **doi: 10.2147/OTT.S257363.**

[40] H. Deng et al., « CBX6 is negatively regulated by EZH2 and plays a potential tumor suppressor role in breast cancer », Sci. Rep., vol. 9, n° 1, p. 197, janv. 2019, **doi: 10.1038/s41598-018-36560-4.**

[41] L. Morey et al., « Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells », Cell Stem Cell, vol. 10, n° 1, p. 47-62, janv. 2012, **doi: 10.1016/j.stem.2011.12.006.**

[42] J. Jung et al., « CBX7 Induces Self-Renewal of Human Normal and Malignant Hematopoietic Stem and Progenitor Cells by Canonical and Non-canonical Interactions », Cell Rep., vol. 26, n° 7, p. 1906-1918.e8, févr. 2019, **doi: 10.1016/j.celrep.2019.01.050.**

[43] Z. Bao et al., « CBX7 negatively regulates migration and invasion in glioma via Wnt/β-catenin pathway inactivation », Oncotarget, vol. 8, n° 24, p. 39048-39063, juin 2017, **doi: 10.18632/oncotarget.16587.**

[44] P. Domizi, F. Malizia, L. Chazarreta-Cifre, L. Diacovich, et C. Banchio, « KDM2B regulates choline kinase expression and neuronal differentiation of neuroblastoma cells », PloS One, vol. 14, n° 1, p. e0210207, 2019, **doi: 10.1371/journal.pone.0210207.**

[45] J. Andricovich, Y. Kai, W. Peng, A. Foudi, et A. Tzatsos, « Histone demethylase KDM2B regulates lineage commitment in normal and malignant hematopoiesis », J. Clin. Invest., vol. 126, n° 3, p. 905-920, mars 2016, **doi: 10.1172/JCI84014.**

[46] J. He, L. Shen, M. Wan, O. Taranova, H. Wu, et Y. Zhang, « Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes », Nat. Cell Biol., vol. 15, n° 4, p. 373-384, avr. 2013, **doi: 10.1038/ncb2702.**

[47] Z. Zhou et al., « Kdm2b Regulates Somatic Reprogramming through Variant PRC1 Complex-Dependent Function », Cell Rep., vol. 21, n° 8, p. 2160-2170, nov. 2017, **doi: 10.1016/j.celrep.2017.10.091.**

[48] J. G. P. Sanches et al., « The Role of KDM2B and EZH2 in Regulating the Stemness in Colorectal Cancer Through the PI3K/AKT Pathway », Front. Oncol., vol. 11, p. 637298, 2021, **doi:**

10.3389/fonc.2021.637298.

[49] M. Quan et al., « Lysine demethylase 2 (KDM2B) regulates hippo pathway via MOB1 to promote pancreatic ductal adenocarcinoma (PDAC) progression », J. Exp. Clin. Cancer Res. CR, vol. 39, n° 1, p. 13, janv. 2020, **doi: 10.1186/s13046-019-1489-0.**

[50] Y. Isshiki et al., « KDM2B in polycomb repressive complex 1.1 functions as a tumor suppressor in the initiation of T-cell leukemogenesis », Blood Adv., vol. 3, n° 17, p. 2537-2549, sept. 2019, **doi: 10.1182/bloodadvances.2018028522.**

[51] M. Yao et al., « PCGF5 is required for neural differentiation of embryonic stem cells », Nat. Commun., vol. 9, n° 1, p. 1463, mai 2018, **doi: 10.1038/s41467-018-03781-0.**

[52] Y. Meng, Y. Liu, E. Dakou, G. J. Gutierrez, et L. Leyns, « Polycomb group RING finger protein 5 influences several developmental signaling pathways during the in vitro differentiation of mouse embryonic stem cells », Dev. Growth Differ., vol. 62, n° 4, p. 232-242, mai 2020, **doi: 10.1111/dgd.12659.**

[53] W. Zhao et al., « Polycomb group RING finger proteins 3/5 activate transcription via an interaction with the pluripotency factor Tex10 in embryonic stem cells », J. Biol. Chem., vol. 292, n° 52, p. 21527-21537, déc. 2017, **doi: 10.1074/jbc.M117.804054.**

[54] X. W. Wang, W. Wei, W. Q. Wang, X. Y. Zhao, H. Guo, et D. C. Fang, « RING finger proteins are involved in the progression of barrett esophagus to esophageal adenocarcinoma: a preliminary study », Gut Liver, vol. 8, n° 5, p. 487-494, sept. 2014, **doi: 10.5009/gnl13133.**

[55] S. Nagel, C. Pommerenke, C. Meyer, M. Kaufmann, H. G. Drexler, et R. A. F. MacLeod, « Deregulation of polycomb repressor complex 1 modifier AUTS2 in T-cell leukemia », Oncotarget, vol. 7, n° 29, p. 45398-45413, juill. 2016, **doi: 10.18632/oncotarget.9982.**

[56] C. Carlberg et F. Molnár, « Transcription Factors », in Mechanisms of Gene Regulation, Dordrecht: Springer Netherlands, 2014, p. 55-70. **doi: 10.1007/978-94-007-7905-1_4.**

[57] A. B. Patel et al., « Structure of human TFIID and mechanism of TBP loading onto promoter DNA », Science, vol. 362, n° 6421, 21 2018, **doi: 10.1126/science.aau8872.**

[58] A. B. Patel, B. J. Greber, et E. Nogales, « Recent insights into the structure of TFIID, its assembly, and its binding to core promoter », Curr. Opin. Struct. Biol., vol. 61, p. 17-24, avr. 2020, **doi: 10.1016/j.sbi.2019.10.001.**

[59] A. C. Schier et D. J. Taatjes, « Structure and mechanism of the RNA polymerase II transcription machinery », Genes Dev., vol. 34, n° 7-8, p. 465-488, avr. 2020, **doi: 10.1101/gad.335679.119.**

[60] S. H. C. Duttke, « Evolution and diversification of the basal transcription machinery », Trends Biochem. Sci., vol. 40, n° 3, p. 127-129, mars 2015, **doi: 10.1016/j.tibs.2015.01.005.**

[61] O. Parra-Marín, K. López-Pacheco, R. Hernández, et I. López-Villaseñor, « The highly diverse TATA box-binding proteins among protists: A review », Mol. Biochem. Parasitol., vol. 239, p. 111312, sept. 2020, **doi: 10.1016/j.molbiopara.2020.111312.**

[62] P. Papadopoulos et al., « TAF10 Interacts with the GATA1 Transcription Factor and Controls Mouse Erythropoiesis », Mol. Cell. Biol., vol. 35, n° 12, p. 2103-2118, juin 2015, **doi: 10.1128/MCB.01370-14.**

[63] Y. Xu et al., « TAF1 plays a critical role in AML1-ETO driven leukemogenesis », Nat. Commun., vol. 10, n° 1, p. 4925, oct. 2019, **doi: 10.1038/s41467-019-12735-z.**

[64] K. J. Grive et al., « TAF4b Regulates Oocyte-Specific Genes Essential for Meiosis », PLoS Genet., vol.

12, n° 6, p. e1006128, juin 2016, **doi: 10.1371/journal.pgen.1006128.**

[65] L. A. Lovasco, E. A. Gustafson, K. A. Seymour, D. G. de Rooij, et R. N. Freiman, « TAF4b is required for mouse spermatogonial stem cell development », Stem Cells Dayt. Ohio, vol. 33, n° 4, p. 1267-1276, avr. 2015, **doi: 10.1002/stem.1914.**

[66] A. Bahat, R. Kedmi, K. Gazit, I. Richardo-Lax, E. Ainbinder, et R. Dikstein, « TAF4b and TAF4 differentially regulate mouse embryonic stem cells maintenance and proliferation », Genes Cells Devoted Mol. Cell. Mech., vol. 18, n° 3, p. 225-237, mars 2013, **doi: 10.1111/gtc.12030.**

[67] H. Zhou et al., « Dual functions of TAF7L in adipocyte differentiation », eLife, vol. 2, p. e00170, janv. 2013, **doi: 10.7554/eLife.00170.**

[68] H. Zhou et al., « Taf7l cooperates with Trf2 to regulate spermiogenesis », Proc. Natl. Acad. Sci. U. S. A., vol. 110, n° 42, p. 16886-16891, oct. 2013, **doi: 10.1073/pnas.1317034110.**

[69] F. J. Herrera, T. Yamaguchi, H. Roelink, et R. Tjian, « Core promoter factor TAF9B regulates neuronal gene expression », eLife, vol. 3, p. e02559, juill. 2014, **doi: 10.7554/eLife.02559.**

[70] M. Frontini et al., « TAF9b (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9 », Mol. Cell. Biol., vol. 25, n° 11, p. 4638-4649, juin 2005, **doi: 10.1128/MCB.25.11.4638-4649.2005.**

[71] K. J. Wright, M. T. Marr, et R. Tjian, « TAF4 nucleates a core subcomplex of TFIID and mediates activated transcription from a TATA-less promoter », Proc. Natl. Acad. Sci. U. S. A., vol. 103, n° 33, p. 12347-12352, août 2006, **doi: 10.1073/pnas.0605499103.**

[72] J. N. Davis, L. McGhee, et S. Meyers, « The ETO (MTG8) gene family », Gene, vol. 303, p. 1-10, janv. 2003, **doi: 10.1016/s0378-1119(02)01172-1.**

[73] D. Langer et al., « Essential role of the TFIID subunit TAF4 in murine embryogenesis and embryonic stem cell differentiation », Nat. Commun., vol. 7, p. 11063, mars 2016, **doi: 10.1038/ncomms11063.**

[74] M. Guermah, Y. Tao, et R. G. Roeder, « Positive and negative TAF(II) functions that suggest a dynamic TFIID structure and elicit synergy with traps in activator-induced transcription », Mol. Cell. Biol., vol. 21, n° 20, p. 6882-6894, oct. 2001, **doi: 10.1128/MCB.21.20.6882-6894.2001.**

[75] W. W. M. P. Pijnappel et al., « A central role for TFIID in the pluripotent transcription circuitry », Nature, vol. 495, n° 7442, p. 516-519, mars 2013, **doi: 10.1038/nature11970.**

[76] I. Martianov, S. Viville, et I. Davidson, « RNA polymerase II transcription in murine cells lacking the TATA binding protein », Science, vol. 298, n° 5595, p. 1036-1039, nov. 2002, **doi: 10.1126/science.1076327.**

[77] A. Gegonne et al., « The general transcription factor TAF7 is essential for embryonic development but not essential for the survival or differentiation of mature T cells », Mol. Cell. Biol., vol. 32, n° 10, p. 1984-1997, mai 2012, **doi: 10.1128/MCB.06305-11.**

[78] W. S. Mohan, E. Scheer, O. Wendling, D. Metzger, et L. Tora, « TAF10 (TAF(II)30) is necessary for TFIID stability and early embryogenesis in mice », Mol. Cell. Biol., vol. 23, n° 12, p. 4307-4318, juin 2003, **doi: 10.1128/MCB.23.12.4307-4318.2003.**

[79] A. Fadloun et al., « The TFIID subunit TAF4 regulates keratinocyte proliferation and has cell-autonomous and non-cell-autonomous tumour suppressor activity in mouse epidermis », Dev. Camb. Engl., vol. 134, n° 16, p. 2947-2958, août 2007, **doi: 10.1242/dev.005041.**

[80] T. Kleiber, G. Davidson, G. Mengus, I. Martianov, et I. Davidson, « Single cell transcriptomics reveal trans-differentiation of pancreatic beta cells following inactivation of the TFIID subunit Taf4 », Cell Death Dis., vol. 12, n° 8, p. 790, août 2021, **doi: 10.1038/s41419-021-04067-y.**

[81] J. Kazantseva, K. Tints, T. Neuman, et K. Palm, « TAF4 controls differentiation of human neural progenitor cells through hTAF4-TAFH activity », J. Mol. Neurosci. MN, vol. 55, n° 1, p. 160-166, janv. 2015, **doi: 10.1007/s12031-014-0295-6.**

[82] J. Kazantseva, A. Kivil, K. Tints, A. Kazantseva, T. Neuman, et K. Palm, « Alternative splicing targeting the hTAF4-TAFH domain of TAF4 represses proliferation and accelerates chondrogenic differentiation of human mesenchymal stem cells », PloS One, vol. 8, n° 10, p. e74799, 2013, **doi: 10.1371/journal.pone.0074799.**

[83] C. Chazaud et Y. Yamanaka, « Lineage specification in the mouse preimplantation embryo », Dev. Camb. Engl., vol. 143, n° 7, p. 1063-1074, avr. 2016, **doi: 10.1242/dev.128314.**

[84] S. M. Morgani, J. J. Metzger, J. Nichols, E. D. Siggia, et A.-K. Hadjantonakis, « Micropattern differentiation of mouse pluripotent stem cells recapitulates embryo regionalized cell fate patterning », eLife, vol. 7, p. e32839, févr. 2018, **doi: 10.7554/eLife.32839.**

[85] S. Nowotschin et al., « The emergent landscape of the mouse gut endoderm at single-cell resolution », Nature, vol. 569, n° 7756, p. 361-367, mai 2019, **doi: 10.1038/s41586-019-1127-1.**

[86] M. Mittnenzweig et al., « A single-embryo, single-cell time-resolved model for mouse gastrulation », Cell, vol. 184, n° 11, p. 2825-2842.e22, mai 2021, **doi: 10.1016/j.cell.2021.04.004.**

[87] Y. Kojima, O. H. Tam, et P. P. L. Tam, « Timing of developmental events in the early mouse embryo », Semin. Cell Dev. Biol., vol. 34, p. 65-75, oct. 2014, **doi: 10.1016/j.semcdb.2014.06.010.**

[88] Z. Wu et K.-L. Guan, « Hippo Signaling in Embryogenesis and Development », Trends Biochem. Sci., vol. 46, n° 1, p. 51-63, janv. 2021, **doi: 10.1016/j.tibs.2020.08.008.**

[89] M. N. Shahbazi et M. Zernicka-Goetz, « Deconstructing and reconstructing the mouse and human early embryo », Nat. Cell Biol., vol. 20, n° 8, p. 878-887, août 2018, **doi: 10.1038/s41556-018-0144-x.**

[90] T. Rayon et al., « Notch and Hippo Converge on Cdx2 to Specify the Trophectoderm Lineage in the Mouse Blastocyst », Dev. Cell, vol. 30, n° 4, p. 410-422, août 2014, **doi: 10.1016/j.devcel.2014.06.019.**

[91] C. Bassalert, L. Valverde-Estrella, et C. Chazaud, « Primitive Endoderm Differentiation: From Specification to Epithelialization », Curr. Top. Dev. Biol., vol. 128, p. 81-104, 2018, **doi: 10.1016/bs.ctdb.2017.12.001.**

[92] A. C. H. McDonald et J. Rossant, « Gut endoderm takes flight from the wings of mesoderm », Nat. Cell Biol., vol. 16, n° 12, p. 1128-1129, déc. 2014, **doi: 10.1038/ncb3077.**

[93] E. S. Bardot et A.-K. Hadjantonakis, « Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate », Mech. Dev., vol. 163, p. 103617, sept. 2020, **doi: 10.1016/j.mod.2020.103617.**

[94] S. Nowotschin et A.-K. Hadjantonakis, « Chapter Fifteen - Guts and gastrulation: Emergence and convergence of endoderm in the mouse embryo », in Current Topics in Developmental Biology, vol. 136, L. Solnica-Krezel, Éd. Academic Press, 2020, p. 429-454. **doi: 10.1016/bs.ctdb.2019.11.012.**

[95] J. Richard Albert et M. V. C. Greenberg, « The Polycomb landscape in mouse development », Nat. Genet., vol. 53, n° 4, p. 427-429, avr. 2021, **doi: 10.1038/s41588-021-00833-y.**

[96] S. Kaufhold, H. Garbán, et B. Bonavida, « Yin Yang 1 is associated with cancer stem cell transcription factors (SOX2, OCT4, BMI1) and clinical implication », J. Exp. Clin. Cancer Res. CR, vol. 35, p. 84, mai 2016, **doi: 10.1186/s13046-016-0359-2.**

[97] V. Kashyap et al., « Regulation of stem cell pluripotency and differentiation involves a mutual

regulatory circuit of the NANOG, OCT4, and SOX2 pluripotency transcription factors with polycomb repressive complexes and stem cell microRNAs », Stem Cells Dev., vol. 18, n° 7, p. 1093-1108, sept. 2009, **doi: 10.1089/scd.2009.0113.**

[98] M. P. Dobreva et al., « Periostin as a biomarker of the amniotic membrane », Stem Cells Int., vol. 2012, p. 987185, 2012, **doi: 10.1155/2012/987185.**

[99] M. Madabhushi et E. Lacy, « Anterior Visceral Endoderm Directs Ventral Morphogenesis and Placement of Head and Heart via BMP2 Expression », Dev. Cell, vol. 21, n° 5, p. 907-919, nov. 2011, **doi: 10.1016/j.devcel.2011.08.027.**

[100] A. M. Zorn et J. M. Wells, « Vertebrate endoderm development and organ formation », Annu. Rev. Cell Dev. Biol., vol. 25, p. 221-251, 2009, **doi: 10.1146/annurev.cellbio.042308.113344.**

[101] S. Wang, K. D. Walton, et D. L. Gumucio, « Chapter Two - Signals and forces shaping organogenesis of the small intestine », in Current Topics in Developmental Biology, vol. 132, D. M. Wellik, Éd. Academic Press, 2019, p. 31-65. **doi: 10.1016/bs.ctdb.2018.12.001.**

[102] C. Davenport, U. Diekmann, I. Budde, N. Detering, et O. Naujok, « Anterior–Posterior Patterning of Definitive Endoderm Generated from Human Embryonic Stem Cells Depends on the Differential Signaling of Retinoic Acid, Wnt-, and BMP-Signaling », STEM CELLS, vol. 34, n° 11, p. 2635-2647, 2016, **doi: 10.1002/stem.2428.**

[103] K. D. Walton, A. M. Freddo, S. Wang, et D. L. Gumucio, « Generation of intestinal surface: an absorbing tale », Dev. Camb. Engl., vol. 143, n° 13, p. 2261-2272, juill. 2016, **doi: 10.1242/dev.135400.**

[104] K. D. Walton et al., « Villification in the mouse: Bmp signals control intestinal villus patterning », Dev. Camb. Engl., vol. 143, n° 3, p. 427-436, févr. 2016, **doi: 10.1242/dev.130112.**

[105] K. D. Sumigray, M. Terwilliger, et T. Lechler, « Morphogenesis and Compartmentalization of the Intestinal Crypt », Dev. Cell, vol. 45, n° 2, p. 183-197.e5, avr. 2018, **doi: 10.1016/j.devcel.2018.03.024.**

[106] A.-P. G. Haramis et al., « De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine », Science, vol. 303, n° 5664, p. 1684-1686, mars 2004, **doi: 10.1126/science.1093587.**

[107] N. Gao, P. White, et K. H. Kaestner, « Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2 », Dev. Cell, vol. 16, n° 4, p. 588-599, avr. 2009, **doi: 10.1016/j.devcel.2009.02.010.**

[108] S. Grainger, J. G. A. Savory, et D. Lohnes, « Cdx2 regulates patterning of the intestinal epithelium », Dev. Biol., vol. 339, n° 1, p. 155-165, mars 2010, **doi: 10.1016/j.ydbio.2009.12.025.**

[109] E. J. Stringer et al., « Cdx2 determines the fate of postnatal intestinal endoderm », Development, vol. 139, n° 3, p. 465-474, févr. 2012, **doi: 10.1242/dev.070722.**

[110] A. Hryniuk, S. Grainger, J. G. A. Savory, et D. Lohnes, « Cdx function is required for maintenance of intestinal identity in the adult », Dev. Biol., vol. 363, n° 2, p. 426-437, mars 2012, **doi: 10.1016/j.ydbio.2012.01.010.**

[111] N. Kumar et al., « The lineage-specific transcription factor CDX2 navigates dynamic chromatin to control distinct stages of intestine development », Dev. Camb. Engl., vol. 146, n° 5, p. dev172189, mars 2019, **doi: 10.1242/dev.172189.**

[112] M. Saxena, A. K. S. Roman, N. K. O'Neill, R. Sulahian, U. Jadhav, et R. A. Shivdasani, « Transcription factor-dependent "anti-repressive" mammalian enhancers exclude H3K27me3 from extended genomic domains », Genes Dev., vol. 31, n° 23-24, p. 2391-2404, déc. 2017, **doi:**

10.1101/gad.308536.117.

[113] C. Balbinot et al., « The Cdx2 homeobox gene suppresses intestinal tumorigenesis through non-cell-autonomous mechanisms », J. Exp. Med., vol. 215, n° 3, p. 911-926, mars 2018, **doi: 10.1084/jem.20170934.**

[114] C. Bonhomme et al., « Cdx1, a dispensable homeobox gene for gut development with limited effect in intestinal cancer », Oncogene, vol. 27, n° 32, p. 4497-4502, juill. 2008, **doi: 10.1038/onc.2008.78.**

[115] Y. Fujii et al., « CDX1 confers intestinal phenotype on gastric epithelial cells via induction of stemness-associated reprogramming factors SALL4 and KLF5 », Proc. Natl. Acad. Sci. U. S. A., vol. 109, n° 50, p. 20584-20589, déc. 2012, **doi: 10.1073/pnas.1208651109.**

[116] L. Chen et al., « HNF4 factors control chromatin accessibility and are redundantly required for maturation of the fetal intestine », Dev. Camb. Engl., vol. 146, n° 19, p. dev179432, oct. 2019, **doi: 10.1242/dev.179432.**

[117] E. J. Formeister, A. L. Sionas, D. K. Lorance, C. L. Barkley, G. H. Lee, et S. T. Magness, « Distinct SOX9 levels differentially mark stem/progenitor populations and enteroendocrine cells of the small intestine epithelium », Am. J. Physiol.-Gastrointest. Liver Physiol., vol. 296, n° 5, p. G1108-G1118, mai 2009, **doi: 10.1152/ajpgi.00004.2009.**

[118] P. Bastide et al., « Sox9 regulates cell proliferation and is required for Paneth cell differentiation in the intestinal epithelium », J. Cell Biol., vol. 178, n° 4, p. 635-648, août 2007, **doi: 10.1083/jcb.200704152.**

[119] K. C. Roche, A. D. Gracz, X. F. Liu, V. Newton, H. Akiyama, et S. T. Magness, « SOX9 Maintains Reserve Stem Cells and Preserves Radioresistance in Mouse Small Intestine », Gastroenterology, vol. 149, n° 6, p. 1553-1563.e10, nov. 2015, **doi: 10.1053/j.gastro.2015.07.004.**

[120] E. M. Walker, C. A. Thompson, B. M. Kohlnhofer, M. L. Faber, et M. A. Battle, « Characterization of the developing small intestine in the absence of either GATA4 or GATA6 », BMC Res. Notes, vol. 7, p. 902, déc. 2014, **doi: 10.1186/1756-0500-7-902.**

[121] B. M. Kohlnhofer, C. A. Thompson, E. M. Walker, et M. A. Battle, « GATA4 regulates epithelial cell proliferation to control intestinal growth and development in mice », Cell. Mol. Gastroenterol. Hepatol., vol. 2, n° 2, p. 189-209, mars 2016, **doi: 10.1016/j.jcmgh.2015.11.010.**

[122] M. A. Battle et al., « GATA4 is essential for jejunal function in mice », Gastroenterology, vol. 135, n° 5, p. 1676-1686.e1, nov. 2008, **doi: 10.1053/j.gastro.2008.07.074.**

[123] E. Beuling et al., « GATA Factors Regulate Proliferation, Differentiation, and Gene Expression in Small Intestine of Mature Mice », Gastroenterology, vol. 140, n° 4, p. 1219-1229.e2, avr. 2011, **doi: 10.1053/j.gastro.2011.01.033.**

[124] F. Laudisi et al., « GATA6 deficiency leads to epithelial barrier dysfunction and enhances susceptibility to gut inflammation », J. Crohns Colitis, p. jjab145, août 2021, **doi: 10.1093/ecco-jcc/jjab145.**

[125] E. Beuling et al., « GATA6 is required for proliferation, migration, secretory cell maturation, and gene expression in the mature mouse colon », Mol. Cell. Biol., vol. 32, n° 17, p. 3392-3402, sept. 2012, **doi: 10.1128/MCB.00070-12.**

[126] R. Farré, M. Fiorani, S. Abdu Rahiman, et G. Matteoli, « Intestinal Permeability, Inflammation and the Role of Nutrients », Nutrients, vol. 12, n° 4, p. E1185, avr. 2020, **doi: 10.3390/nu12041185.**

[127] Y. Matsumoto et al., « Distinct intestinal adaptation for vitamin B12 and bile acid absorption revealed in a new mouse model of massive ileocecal resection », Biol. Open, vol. 6, n° 9, p. 1364-1374, sept. 2017, **doi: 10.1242/bio.024927.**

[128] A. E. Moor et al., « Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the

Intestinal Villus Axis », Cell, vol. 175, n° 4, p. 1156-1167.e15, nov. 2018, **doi: 10.1016/j.cell.2018.08.063.**

[129] D. W. Powell, I. V. Pinchuk, J. I. Saada, X. Chen, et R. C. Mifflin, « Mesenchymal cells of the intestinal lamina propria », Annu. Rev. Physiol., vol. 73, p. 213-237, 2011, **doi: 10.1146/annurev.physiol.70.113006.100646.**

[130] N. McCarthy, J. Kraiczy, et R. A. Shivdasani, « Cellular and molecular architecture of the intestinal stem cell niche », Nat. Cell Biol., vol. 22, n° 9, p. 1033-1041, sept. 2020, **doi: 10.1038/s41556-020-0567-z.**

[131] S. Pompili, G. Latella, E. Gaudio, R. Sferra, et A. Vetuschi, « The Charming World of the Extracellular Matrix: A Dynamic and Protective Network of the Intestinal Wall », Front. Med., vol. 8, p. 610189, avr. 2021, **doi: 10.3389/fmed.2021.610189.**

[132] M. Visentin, N. Diop-Bove, R. Zhao, et I. D. Goldman, « The intestinal absorption of folates », Annu. Rev. Physiol., vol. 76, p. 251-274, 2014, **doi: 10.1146/annurev-physiol-020911-153251.**

[133] N. McCarthy et al., « Distinct Mesenchymal Cell Populations Generate the Essential Intestinal BMP Signaling Gradient », Cell Stem Cell, vol. 26, n° 3, p. 391-402.e5, mars 2020, **doi: 10.1016/j.stem.2020.01.008.**

[134] M. Shoshkes-Carmel et al., « Subepithelial telocytes are an important source of Wnts that supports intestinal crypts », Nature, vol. 557, n° 7704, p. 242-246, mai 2018, **doi: 10.1038/s41586-018-0084-4.**

[135] D. Krndija et al., « Active cell migration is critical for steady-state epithelial turnover in the gut », Science, vol. 365, n° 6454, p. 705-710, août 2019, **doi: 10.1126/science.aau3429.**

[136] V. Fernández-Majada et al., « Subepithelial Myofibroblasts Are Critical Regulators of Intestinal Epithelial Restoration », mai 2021. **doi: 10.1101/2021.05.28.446131.**

[137] J. Beumer et H. Clevers, « Cell fate specification and differentiation in the adult mammalian intestine », Nat. Rev. Mol. Cell Biol., sept. 2020, **doi: 10.1038/s41580-020-0278-0.**

[138] X. Yin, H. F. Farin, J. H. van Es, H. Clevers, R. Langer, et J. M. Karp, « Niche-independent high-purity cultures of Lgr5+ intestinal stem cells and their progeny », Nat. Methods, vol. 11, n° 1, p. 106-112, janv. 2014, **doi: 10.1038/nmeth.2737.**

[139] L. Chen et al., « A reinforcing HNF4-SMAD4 feed-forward module stabilizes enterocyte identity », Nat. Genet., vol. 51, n° 5, p. 777-785, mai 2019, **doi: 10.1038/s41588-019-0384-0.**

[140] K. A. Knoop et al., « RANKL Is Necessary and Sufficient to Initiate Development of Antigen-Sampling M Cells in the Intestinal Epithelium », J. Immunol., vol. 183, n° 9, p. 5738-5747, nov. 2009, **doi: 10.4049/jimmunol.0901563.**

[141] J. P. Katz et al., « The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon », Dev. Camb. Engl., vol. 129, n° 11, p. 2619-2628, juin 2002. **doi: 10.1242/dev.129.11.2619**

[142] A. Gregorieff et al., « The ets-domain transcription factor Spdef promotes maturation of goblet and paneth cells in the intestinal epithelium », Gastroenterology, vol. 137, n° 4, p. 1333-1345.e1-3, oct. 2009, **doi: 10.1053/j.gastro.2009.06.044.**

[143] O. Basak, J. Beumer, K. Wiebrands, H. Seno, A. van Oudenaarden, et H. Clevers, « Induced Quiescence of Lgr5+ Stem Cells in Intestinal Organoids Enables Differentiation of Hormone-Producing Enteroendocrine Cells », Cell Stem Cell, vol. 20, n° 2, p. 177-190.e4, févr. 2017, **doi: 10.1016/j.stem.2016.11.001.**

[144] M. Jenny et al., « Neurogenin3 is differentially required for endocrine cell fate specification in the

intestinal and gastric epithelium », EMBO J., vol. 21, n° 23, p. 6338-6347, déc. 2002, **doi: 10.1093/emboj/cdf649.**

[145] Y. Mori-Akiyama et al., « SOX9 is required for the differentiation of paneth cells in the intestinal epithelium », Gastroenterology, vol. 133, n° 2, p. 539-546, août 2007, **doi: 10.1053/j.gastro.2007.05.020.**

[146] F. Gerbe et al., « Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth parasites », Nature, vol. 529, n° 7585, p. 226-230, janv. 2016, **doi: 10.1038/nature16527.**

[147] N. Goto et al., « Lineage tracing and targeting of IL17RB+ tuft cell-like human colorectal cancer stem cells », Proc. Natl. Acad. Sci. U. S. A., vol. 116, n° 26, p. 12996-13005, juin 2019, **doi: 10.1073/pnas.1900251116.**

[148] J. Rispal, F. Escaffit, et D. Trouche, « Chromatin Dynamics in Intestinal Epithelial Homeostasis: A Paradigm of Cell Fate Determination versus Cell Plasticity », Stem Cell Rev. Rep., vol. 16, n° 6, p. 1062-1080, déc. 2020, **doi: 10.1007/s12015-020-10055-0.**

[149] F. Chiacchiera et al., « Polycomb Complex PRC1 Preserves Intestinal Stem Cell Identity by Sustaining Wnt/β-Catenin Transcriptional Activity », Cell Stem Cell, vol. 18, n° 1, p. 91-103, janv. 2016, **doi: 10.1016/j.stem.2015.09.019.**

[150] F. Chiacchiera, A. Rossi, S. Jammula, M. Zanotti, et D. Pasini, « PRC2 preserves intestinal progenitors and restricts secretory lineage commitment », EMBO J., vol. 35, n° 21, p. 2301-2314, nov. 2016, **doi: 10.15252/embj.201694550.**

[151] M. Oittinen et al., « Polycomb Repressive Complex 2 Enacts Wnt Signaling in Intestinal Homeostasis and Contributes to the Instigation of Stemness in Diseases Entailing Epithelial Hyperplasia or Neoplasia », Stem Cells Dayt. Ohio, vol. 35, n° 2, p. 445-457, févr. 2017, **doi: 10.1002/stem.2479.**

[152] J. J. George et al., « Polycomb Repressive Complex 2 Regulates Genes Necessary for Intestinal Microfold Cell (M Cell) Development », Cell. Mol. Gastroenterol. Hepatol., vol. 12, n° 3, p. 873-889, 2021, **doi: 10.1016/j.jcmgh.2021.05.014.**

[153] J. J. George, L. Martin-Diaz, M. J. T. Ojanen, R. Gasa, M. Pesu, et K. Viiri, « PRC2 Regulated Atoh8 Is a Regulator of Intestinal Microfold Cell (M Cell) Differentiation », Int. J. Mol. Sci., vol. 22, n° 17, p. 9355, août 2021, **doi: 10.3390/ijms22179355.**

[154] P. Paone et P. D. Cani, « Mucus barrier, mucins and gut microbiota: the expected slimy partners? », Gut, vol. 69, n° 12, p. 2232-2243, déc. 2020, **doi: 10.1136/gutjnl-2020-322260.**

[155] K. L. Carraway, G. Theodoropoulos, G. A. Kozloski, et C. A. Carothers Carraway, « Muc4/MUC4 functions and regulation in cancer », Future Oncol. Lond. Engl., vol. 5, n° 10, p. 1631-1640, déc. 2009, **doi: 10.2217/fon.09.125.**

[156] W. Hoffmann, « Trefoil Factor Family (TFF) Peptides and Their Links to Inflammation: A Re-evaluation and New Medical Perspectives », Int. J. Mol. Sci., vol. 22, n° 9, p. 4909, mai 2021, **doi: 10.3390/ijms22094909.**

[157] H. Mashimo, D. C. Wu, D. K. Podolsky, et M. C. Fishman, « Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor », Science, vol. 274, n° 5285, p. 262-265, oct. 1996, **doi: 10.1126/science.274.5285.262.**

[158] S.-T. Chen, T.-C. Kuo, Y.-Y. Liao, M.-C. Lin, Y.-W. Tien, et M.-C. Huang, « Silencing of MUC20 suppresses the malignant character of pancreatic ductal adenocarcinoma cells through inhibition of the HGF/MET pathway », Oncogene, vol. 37, n° 46, p. 6041-6053, nov. 2018, **doi: 10.1038/s41388-018-0403-0.**

[159] X. Xiao et al., « Role of MUC20 overexpression as a predictor of recurrence and poor outcome in colorectal cancer », J. Transl. Med., vol. 11, p. 151, juin 2013, **doi: 10.1186/1479-5876-11-151.**

[160] J. Qin, B. Wen, Y. Liang, W. Yu, et H. Li, « Histone Modifications and their Role in Colorectal Cancer (Review) », Pathol. Oncol. Res., vol. 26, n° 4, p. 2023-2033, 2020, **doi: 10.1007/s12253-019-00663-8.**

[161] H. K. Chung, L. Xiao, K. C. Jaladanki, et J.-Y. Wang, « Regulation of Paneth Cell Function by RNA-Binding Proteins and Noncoding RNAs », Cells, vol. 10, n° 8, p. 2107, août 2021, **doi: 10.3390/cells10082107.**

[162] N. H. Salzman et al., « Enteric defensins are essential regulators of intestinal microbial ecology », Nat. Immunol., vol. 11, n° 1, p. 76-83, janv. 2010, **doi: 10.1038/ni.1825.**

[163] J. Beumer et al., « Enteroendocrine cells switch hormone expression along the crypt-to-villus BMP signalling gradient », Nat. Cell Biol., vol. 20, n° 8, p. 909-916, août 2018, **doi: 10.1038/s41556-018-0143-y.**

[164] A. Banerjee, E. T. McKinley, J. von Moltke, R. J. Coffey, et K. S. Lau, « Interpreting heterogeneity in intestinal tuft cell structure and function », J. Clin. Invest., vol. 128, n° 5, p. 1711-1719, mai 2018, **doi: 10.1172/JCI120330.**

[165] A. L. Haber et al., « A single-cell survey of the small intestinal epithelium », Nature, vol. 551, n° 7680, p. 333-339, nov. 2017, **doi: 10.1038/nature24489.**

[166] E. Perland et R. Fredriksson, « Classification Systems of Secondary Active Transporters », Trends Pharmacol. Sci., vol. 38, n° 3, p. 305-315, mars 2017, **doi: 10.1016/j.tips.2016.11.008.**

[167] D. Hooton, R. Lentle, J. Monro, M. Wickham, et R. Simpson, « The Secretion and Action of Brush Border Enzymes in the Mammalian Small Intestine », Rev. Physiol. Biochem. Pharmacol., vol. 168, p. 59-118, 2015, **doi: 10.1007/112_2015_24.**

[168] F. Kühn et al., « Targeting the Intestinal Barrier to Prevent Gut-Derived Inflammation and Disease: A Role for Intestinal Alkaline Phosphatase », Visc. Med., vol. 37, n° 5, p. 383-393, oct. 2021, **doi: 10.1159/000515910.**

[169] W. Kramer et al., « Aminopeptidase N (CD13) is a molecular target of the cholesterol absorption inhibitor ezetimibe in the enterocyte brush border membrane », J. Biol. Chem., vol. 280, n° 2, p. 1306-1320, janv. 2005, **doi: 10.1074/jbc.M406309200.**

[170] E. M. Varin et al., « Hematopoietic cell– versus enterocyte-derived dipeptidyl peptidase-4 differentially regulates triglyceride excursion in mice », JCI Insight, vol. 5, n° 16, p. e140418, **doi: 10.1172/jci.insight.140418.**

[171] E. E. Mulvihill et al., « Cellular Sites and Mechanisms Linking Reduction of Dipeptidyl Peptidase-4 Activity to Control of Incretin Hormone Action and Glucose Homeostasis », Cell Metab., vol. 25, n° 1, p. 152-165, janv. 2017, **doi: 10.1016/j.cmet.2016.10.007.**

[172] C. Gregori, A. Porteu, S. Lopez, A. Kahn, et A. L. Pichard, « Characterization of the aldolase B intronic enhancer », J. Biol. Chem., vol. 273, n° 39, p. 25237-25243, sept. 1998, **doi: 10.1074/jbc.273.39.25237.**

[173] F. Ye, Y. Chen, L. Xia, J. Lian, et S. Yang, « Aldolase A overexpression is associated with poor prognosis and promotes tumor progression by the epithelial-mesenchymal transition in colon cancer », Biochem. Biophys. Res. Commun., vol. 497, n° 2, p. 639-645, mars 2018, **doi: 10.1016/j.bbrc.2018.02.123.**

[174] C. Chen, Y. Yin, Q. Tu, et H. Yang, « Glucose and amino acid in enterocyte: absorption, metabolism and maturation », Front. Biosci. Landmark Ed., vol. 23, p. 1721-1739, mars 2018, **doi: 10.2741/4669.**

[175] C.-W. Ko, J. Qu, D. D. Black, et P. Tso, « Regulation of intestinal lipid metabolism: current concepts and relevance to disease », Nat. Rev. Gastroenterol. Hepatol., vol. 17, n° 3, p. 169-183, mars 2020, **doi: 10.1038/s41575-019-0250-7.**

[176] P. Y. Xing, S. Pettersson, et P. Kundu, « Microbial Metabolites and Intestinal Stem Cells Tune Intestinal Homeostasis », Proteomics, vol. 20, n° 5-6, p. e1800419, mars 2020, **doi: 10.1002/pmic.201800419.**

[177] J. L. Parker et al., « Structural basis of antifolate recognition and transport by PCFT », Nature, vol. 595, n° 7865, p. 130-134, juill. 2021, **doi: 10.1038/s41586-021-03579-z.**

[178] B. Spanier, « Transcriptional and functional regulation of the intestinal peptide transporter PEPT1 », J. Physiol., vol. 592, n° 5, p. 871-879, 2014, **doi: 10.1113/jphysiol.2013.258889.**

[179] K. J. Hamblett et al., « SLC46A3 Is Required to Transport Catabolites of Noncleavable Antibody Maytansine Conjugates from the Lysosome to the Cytoplasm », Cancer Res., vol. 75, n° 24, p. 5329-5340, déc. 2015, **doi: 10.1158/0008-5472.CAN-15-1610.**

[180] A. M. Mowat et W. W. Agace, « Regional specialization within the intestinal immune system », Nat. Rev. Immunol., vol. 14, n° 10, p. 667-685, oct. 2014, **doi: 10.1038/nri3738.**

[181] H. Ohno, « Intestinal M cells », J. Biochem. (Tokyo), vol. 159, n° 2, p. 151-160, févr. 2016, **doi: 10.1093/jb/mvv121.**

[182] A. Adak et M. R. Khan, « An insight into gut microbiota and its functionalities », Cell. Mol. Life Sci., vol. 76, n° 3, p. 473-493, févr. 2019, **doi: 10.1007/s00018-018-2943-4.**

[183] B. Allam-Ndoul, S. Castonguay-Paradis, et A. Veilleux, « Gut Microbiota and Intestinal Trans-Epithelial Permeability », Int. J. Mol. Sci., vol. 21, n° 17, p. 6402, sept. 2020, **doi: 10.3390/ijms21176402.**

[184] J. M. Davison, C. R. Lickwar, L. Song, G. Breton, G. E. Crawford, et J. F. Rawls, « Microbiota regulate intestinal epithelial gene expression by suppressing the transcription factor Hepatocyte nuclear factor 4 alpha », Genome Res., vol. 27, n° 7, p. 1195-1206, janv. 2017, **doi: 10.1101/gr.220111.116.**

[185] D. Woźniak, W. Cichy, J. Przysławski, et S. Drzymała-Czyż, « The role of microbiota and enteroendocrine cells in maintaining homeostasis in the human digestive tract », Adv. Med. Sci., vol. 66, n° 2, p. 284-292, sept. 2021, **doi: 10.1016/j.advms.2021.05.003.**

[186] C. A. Simpson, C. Diaz-Arteche, D. Eliby, O. S. Schwartz, J. G. Simmons, et C. S. M. Cowan, « The gut microbiota in anxiety and depression - A systematic review », Clin. Psychol. Rev., vol. 83, p. 101943, févr. 2021, **doi: 10.1016/j.cpr.2020.101943.**

[187] M. Gurung et al., « Role of gut microbiota in type 2 diabetes pathophysiology », EBioMedicine, vol. 51, p. 102590, janv. 2020, **doi: 10.1016/j.ebiom.2019.11.051.**

[188] H. Han et al., « Gut Microbiota and Type 1 Diabetes », Int. J. Mol. Sci., vol. 19, n° 4, p. 995, mars 2018, **doi: 10.3390/ijms19040995.**

[189] T.-C. D. Shen, « Diet and Gut Microbiota in Health and Disease », Nestle Nutr. Inst. Workshop Ser., vol. 88, p. 117-126, 2017, **doi: 10.1159/000455220.**

[190] N. H. Salzman et C. L. Bevins, « Dysbiosis--a consequence of Paneth cell dysfunction », Semin. Immunol., vol. 25, n° 5, p. 334-341, nov. 2013, **doi: 10.1016/j.smim.2013.09.006.**

[191] C. C. Bain et A. Schridde, « Origin, Differentiation, and Function of Intestinal Macrophages », Front. Immunol., vol. 9, p. 2733, 2018, **doi: 10.3389/fimmu.2018.02733.**

[192] A. J. Stagg, « Intestinal Dendritic Cells in Health and Gut Inflammation », Front. Immunol., vol. 9, p.

2883, 2018, **doi: 10.3389/fimmu.2018.02883.**

[193] H. Ma, W. Tao, et S. Zhu, « T lymphocytes in the intestinal mucosa: defense and tolerance », Cell. Mol. Immunol., vol. 16, n° 3, p. 216-224, mars 2019, **doi: 10.1038/s41423-019-0208-2.**

[194] J. Spencer et L. M. Sollid, « The human intestinal B-cell response », Mucosal Immunol., vol. 9, n° 5, p. 1113-1124, sept. 2016, **doi: 10.1038/mi.2016.59.**

[195] A. T. Soderholm et V. A. Pedicord, « Intestinal epithelial cells: at the interface of the microbiota and mucosal immunity », Immunology, vol. 158, n° 4, p. 267-280, déc. 2019, **doi: 10.1111/imm.13117.**

[196] M. H. Kogut, A. Lee, et E. Santin, « Microbiome and pathogen interaction with the immune system », Poult. Sci., vol. 99, n° 4, p. 1906-1913, avr. 2020, **doi: 10.1016/j.psj.2019.12.011.**

[197] M. Zhao, L. Gönczi, P. L. Lakatos, et J. Burisch, « The Burden of Inflammatory Bowel Disease in Europe in 2020 », J. Crohns Colitis, n° jjab029, févr. 2021, **doi: 10.1093/ecco-jcc/jjab029.**

[198] W. Y. Mak, M. Zhao, S. C. Ng, et J. Burisch, « The epidemiology of inflammatory bowel disease: East meets west », J. Gastroenterol. Hepatol., vol. 35, n° 3, p. 380-389, 2020, **doi: 10.1111/jgh.14872.**

[199] I. Khan et al., « Alteration of Gut Microbiota in Inflammatory Bowel Disease (IBD): Cause or Consequence? IBD Treatment Targeting the Gut Microbiome », Pathogens, vol. 8, n° 3, p. 126, août 2019, **doi: 10.3390/pathogens8030126.**

[200] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, et A. Jemal, « Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries », CA. Cancer J. Clin., vol. 68, n° 6, p. 394-424, 2018, **doi: 10.3322/caac.21492.**

[201] M. Arnold, M. S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, et F. Bray, « Global patterns and trends in colorectal cancer incidence and mortality », Gut, vol. 66, n° 4, p. 683-691, avr. 2017, **doi: 10.1136/gutjnl-2015-310912.**

[202] J. M. Bae, T. H. Lee, N.-Y. Cho, T.-Y. Kim, et G. H. Kang, « Loss of CDX2 expression is associated with poor prognosis in colorectal cancer patients », World J. Gastroenterol. WJG, vol. 21, n° 5, p. 1457-1467, févr. 2015, **doi: 10.3748/wjg.v21.i5.1457.**

[203] K. Chellappa et al., « Opposing roles of nuclear receptor HNF4α isoforms in colitis and colitis-associated colon cancer », eLife, vol. 5, p. e10903, mai 2016, **doi: 10.7554/eLife.10903.**

[204] P. Pallante et al., « The loss of the CBX7 gene expression represents an adverse prognostic marker for survival of colon carcinoma patients », Eur. J. Cancer Oxf. Engl. 1990, vol. 46, n° 12, p. 2304-2313, août 2010, **doi: 10.1016/j.ejca.2010.05.011.**

[205] X. Zheng et al., « Critical evaluation of Cbx7 downregulation in primary colon carcinomas and its clinical significance in Chinese patients », BMC Cancer, vol. 15, n° 1, p. 145, mars 2015, **doi: 10.1186/s12885-015-1172-6.**

[206] Q. Li, Y. Pan, Z. Cao, et S. Zhao, « Comprehensive Analysis of Prognostic Value and Immune Infiltration of Chromobox Family Members in Colorectal Cancer », Front. Oncol., vol. 10, p. 582667, sept. 2020, **doi: 10.3389/fonc.2020.582667.**

[207] J. Zhou et al., « Targeting EZH2 histone methyltransferase activity alleviates experimental intestinal inflammation », Nat. Commun., vol. 10, n° 1, p. 2427, juin 2019, **doi: 10.1038/s41467-019-10176-2.**

[208] M. Ohuchi et al., « Increased EZH2 expression during the adenoma-carcinoma sequence in colorectal cancer », Oncol. Lett., vol. 16, n° 4, p. 5275-5281, oct. 2018, **doi: 10.3892/ol.2018.9240.**

[209] V. Koliaraki, C. K. Pallangyo, F. R. Greten, et G. Kollias, « Mesenchymal Cells in Colon Cancer », Gastroenterology, vol. 152, n° 5, p. 964-979, avr. 2017, **doi: 10.1053/j.gastro.2016.11.049.**

[210] E. Picard, C. P. Verschoor, G. W. Ma, et G. Pawelec, « Relationships Between Immune Landscapes, Genetic Subtypes and Responses to Immunotherapy in Colorectal Cancer », Front. Immunol., vol.

11, p. 369, mars 2020, **doi: 10.3389/fimmu.2020.00369.**

[211] Y. Cheng, Z. Ling, et L. Li, « The Intestinal Microbiota and Colorectal Cancer », Front. Immunol., vol. 11, p. 3100, 2020, **doi: 10.3389/fimmu.2020.615056.**

[212] J. C. Arthur et al., « Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota », Science, vol. 338, n° 6103, p. 120-123, oct. 2012, **doi: 10.1126/science.1224820.**

[213] H. Gehart et H. Clevers, « Tales from the crypt: new insights into intestinal stem cells », Nat. Rev. Gastroenterol. Hepatol., vol. 16, n° 1, p. 19-34, janv. 2019, **doi: 10.1038/s41575-018-0081-y.**

[214] S. Date et T. Sato, « Mini-gut organoids: reconstitution of the stem cell niche », Annu. Rev. Cell Dev. Biol., vol. 31, p. 269-289, 2015, **doi: 10.1146/annurev-cellbio-100814-125218.**

[215] T. Takahashi et A. Shiraishi, « Stem Cell Signaling Pathways in the Small Intestine », Int. J. Mol. Sci., vol. 21, n° 6, p. 2032, janv. 2020, **doi: 10.3390/ijms21062032.**

[216] M. Yousefi, L. Li, et C. J. Lengner, « Hierarchy and Plasticity in the Intestinal Stem Cell Compartment », Trends Cell Biol., vol. 27, n° 10, p. 753-764, oct. 2017, **doi: 10.1016/j.tcb.2017.06.006.**

[217] R. Kuno et al., « Notch and TNF-α signaling promote cytoplasmic accumulation of OLFM4 in intestinal epithelium cells and exhibit a cell protective role in the inflamed mucosa of IBD patients », Biochem. Biophys. Rep., vol. 25, p. 100906, mars 2021, **doi: 10.1016/j.bbrep.2020.100906.**

[218] M. Biton et al., « T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation », Cell, vol. 175, n° 5, p. 1307-1320.e22, nov. 2018, **doi: 10.1016/j.cell.2018.10.008.**

[219] J.-M. Zha et al., « Interleukin 22 Expands Transit-Amplifying Cells While Depleting Lgr5+ Stem Cells via Inhibition of Wnt and Notch Signaling », Cell. Mol. Gastroenterol. Hepatol., vol. 7, n° 2, p. 255-274, 2019, **doi: 10.1016/j.jcmgh.2018.09.006.**

[220] A. Sehgal, D. S. Donaldson, C. Pridans, K. A. Sauter, D. A. Hume, et N. A. Mabbott, « The role of CSF1R-dependent macrophages in control of the intestinal stem-cell niche », Nat. Commun., vol. 9, n° 1, p. 1272, mars 2018, **doi: 10.1038/s41467-018-03638-6.**

[221] F. de Sousa e Melo et F. J. de Sauvage, « Cellular Plasticity in Intestinal Homeostasis and Disease », Cell Stem Cell, vol. 24, n° 1, p. 54-64, janv. 2019, **doi: 10.1016/j.stem.2018.11.019.**

[222] J. H. Hageman, M. C. Heinz, K. Kretzschmar, J. van der Vaart, H. Clevers, et H. J. G. Snippert, « Intestinal Regeneration: Regulation by the Microenvironment », Dev. Cell, vol. 54, n° 4, p. 435-446, août 2020, **doi: 10.1016/j.devcel.2020.07.009.**

[223] M. Almeqdadi, M. D. Mana, J. Roper, et Ö. H. Yilmaz, « Gut organoids: mini-tissues in culture to study intestinal physiology and disease », Am. J. Physiol. Cell Physiol., vol. 317, n° 3, p. C405-C419, sept. 2019, **doi: 10.1152/ajpcell.00300.2017.**

[224] F. el Marjou et al., « Tissue-specific and inducible Cre-mediated recombination in the gut epithelium », Genes. N. Y. N 2000, vol. 39, n° 3, p. 186-193, juill. 2004, **doi: 10.1002/gene.20042.**

[225] N. Barker et al., « Identification of stem cells in small intestine and colon by marker gene Lgr5 », Nature, vol. 449, n° 7165, p. 1003-1007, oct. 2007, **doi: 10.1038/nature06196.**

[226] S. Colnot et al., « Colorectal cancers in a new mouse model of familial adenomatous polyposis: influence of genetic and environmental modifiers », Lab. Investig. J. Tech. Methods Pathol., vol. 84, n° 12, p. 1619-1630, déc. 2004, **doi: 10.1038/labinvest.3700180.**

[227] M. Martin, « Cutadapt removes adapter sequences from high-throughput sequencing reads », EMBnet.journal, vol. 17, n° 1, Art. n° 1, mai 2011, **doi: 10.14806/ej.17.1.200.**

[228] A. Bankevich et al., « SPAdes: a new genome assembly algorithm and its applications to single-cell

sequencing », J. Comput. Biol. J. Comput. Mol. Cell Biol., vol. 19, n° 5, p. 455-477, mai 2012, **doi: 10.1089/cmb.2012.0021.**

[229] J. Zhang, K. Kobert, T. Flouri, et A. Stamatakis, « PEAR: a fast and accurate Illumina Paired-End reAd mergeR », Bioinforma. Oxf. Engl., vol. 30, n° 5, p. 614-620, mars 2014, **doi: 10.1093/bioinformatics/btt593.**

[230] T. Rognes, T. Flouri, B. Nichols, C. Quince, et F. Mahé, « VSEARCH: a versatile open source tool for metagenomics », PeerJ, vol. 4, p. e2584, 2016, **doi: 10.7717/peerj.2584.**

[231] J. R. Cole et al., « The Ribosomal Database Project: improved alignments and new tools for rRNA analysis », Nucleic Acids Res., vol. 37, n° Database issue, p. D141-145, janv. 2009, **doi: 10.1093/nar/gkn879.**

[232] B. Langmead, C. Trapnell, M. Pop, et S. L. Salzberg, « Ultrafast and memory-efficient alignment of short DNA sequences to the human genome », Genome Biol., vol. 10, n° 3, p. R25, 2009, **doi: 10.1186/gb-2009-10-3-r25.**

[233] A. Dobin et al., « STAR: ultrafast universal RNA-seq aligner », Bioinforma. Oxf. Engl., vol. 29, n° 1, p. 15-21, janv. 2013, **doi: 10.1093/bioinformatics/bts635.**

[234] S. Anders, P. T. Pyl, et W. Huber, « HTSeq--a Python framework to work with high-throughput sequencing data », Bioinforma. Oxf. Engl., vol. 31, n° 2, p. 166-169, janv. 2015, **doi: 10.1093/bioinformatics/btu638.**

[235] S. Anders et W. Huber, « Differential expression analysis for sequence count data », Genome Biol., vol. 11, n° 10, p. R106, 2010, **doi: 10.1186/gb-2010-11-10-r106.**

[236] M. I. Love, W. Huber, et S. Anders, « Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 », Genome Biol., vol. 15, n° 12, p. 550, 2014, **doi: 10.1186/s13059-014-0550-8.**

[237] Y. Benjamini et Y. Hochberg, « Controlling the false discovery rate: a pratical and powerful approach to multiple testing », J. R. Stat. Soc., vol. 57, p. 289-300, 1994, **doi: 10.1111/j.2517-6161.1995.tb02031.x**

[238] Y. Zhang et al., « Model-based analysis of ChIP-Seq (MACS) », Genome Biol., vol. 9, n° 9, p. R137, 2008, **doi: 10.1186/gb-2008-9-9-r137.**

[239] S. Heinz et al., « Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities », Mol. Cell, vol. 38, n° 4, p. 576-589, mai 2010, **doi: 10.1016/j.molcel.2010.05.004.**

[240] A. R. Quinlan et I. M. Hall, « BEDTools: a flexible suite of utilities for comparing genomic features », Bioinforma. Oxf. Engl., vol. 26, n° 6, p. 841-842, mars 2010, **doi: 10.1093/bioinformatics/btq033.**

[241] T. Ye et al., « seqMINER: an integrated ChIP-seq data interpretation platform », Nucleic Acids Res., vol. 39, n° 6, p. e35, mars 2011, **doi: 10.1093/nar/gkq1287.**

[242] T. L. Bailey, J. Johnson, C. E. Grant, et W. S. Noble, « The MEME Suite », Nucleic Acids Res., vol. 43, n° W1, p. W39-49, juill. 2015, **doi: 10.1093/nar/gkv416.**

[243] M. Bentsen et al., « ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation », Nat. Commun., vol. 11, n° 1, p. 4267, août 2020, **doi: 10.1038/s41467-020-18035-1.**

[244] H. Wickham, ggplot2: Elegant Graphics for Data Analysis, 2nd ed. 2016. Cham: Springer International Publishing : Imprint: Springer, 2016. **doi: 10.1007/978-3-319-24277-4.**

[245] T. Stuart et al., « Comprehensive Integration of Single-Cell Data », Cell, vol. 177, n° 7, p. 1888-

1902.e21, juin 2019, **doi: 10.1016/j.cell.2019.05.031.**

[246] S. Aibar et al., « SCENIC: single-cell regulatory network inference and clustering », Nat. Methods, vol. 14, n° 11, p. 1083-1086, nov. 2017, **doi: 10.1038/nmeth.4463.**

[247] F. Petitprez et al., « The murine Microenvironment Cell Population counter method to estimate abundance of tissue-infiltrating immune and stromal cell populations in murine samples using gene expression », Genome Med., vol. 12, n° 1, p. 86, oct. 2020, **doi: 10.1186/s13073-020-00783-w.**

[248] K. V. Salojin et al., « A mouse model of hereditary folate malabsorption: deletion of the PCFT gene leads to systemic folate deficiency », Blood, vol. 117, n° 18, p. 4895-4904, mai 2011, **doi: 10.1182/blood-2010-04-279653.**

[249] A. M. Martin, E. W. Sun, et D. J. Keating, « Mechanisms controlling hormone secretion in human gut and its relevance to metabolism », J. Endocrinol., vol. 244, n° 1, p. R1-R15, janv. 2020, **doi: 10.1530/JOE-19-0399.**

[250] R. Ye, C. Cao, et Y. Xue, « Enhancer RNA: biogenesis, function, and regulation », Essays Biochem., vol. 64, n° 6, p. 883-894, déc. 2020, **doi: 10.1042/EBC20200014.**

[251] L. Lehmann, R. Ferrari, A. A. Vashisht, J. A. Wohlschlegel, S. K. Kurdistani, et M. Carey, « Polycomb repressive complex 1 (PRC1) disassembles RNA polymerase II preinitiation complexes », J. Biol. Chem., vol. 287, n° 43, p. 35784-35794, oct. 2012, **doi: 10.1074/jbc.M112.397430.**

[252] R. Duan, W. Du, et W. Guo, « EZH2: a novel target for cancer treatment », J. Hematol. Oncol.J Hematol Oncol, vol. 13, n° 1, p. 104, juill. 2020, **doi: 10.1186/s13045-020-00937-8.**

Susanna SÄISÄ-BORREILL

The Basic Transcriptional Machinery in Intestinal Homeostasis

Résumé

Les facteurs de transcription régulateurs interagissent avec la machinerie transcriptionnelle de base pour initier la transcription des gènes. Contrairement aux facteurs régulateurs, la machinerie transcriptionnelle de base a été peu étudiée dans des contextes tissulaires spécifiques, dont l'intestin. L'épithélium intestinal est un tissu complexe, à l'interface entre la lumière intestinale et l'organisme, jouant un rôle majeur de digestion, barrière et transport des nutriments.

Durant ma thèse, j'ai étudié le rôle de la machinerie transcriptionnelle de base dans l'homéostasie intestinale, en me focalisant sur l'un de ses composants majeurs, la protéine TAF4 « TATA-box binding protein Associated Factor 4 ». A l'aide de modèles murins et d'organoïdes intestinaux d'invalidation de Taf4 j'ai pu mettre en lumière son rôle dans l'homéostasie de l'intestin l aux stades embryonnaire et adulte. Son action s'exerce sur les cellules souches et les cellules différenciées, en s'opposant notamment aux complexes polycomb, impliqués dans la régulation négative des gènes en modifiant la structure chromatinienne.

Mots-clefs : TAF4, Polycomb, Intestin, Cellules Souches, Développement, Homéostasie, Cancers.

Abstract

Regulatory transcription factors interact with the basic transcriptional machinery to initiate gene transcription. In contrast to regulatory factors, the basic transcriptional machinery has been little studied in specific tissue contexts, including the intestine. The intestinal epithelium is a complex tissue at the interface between the intestinal lumen and the body, playing a major role in digestion, barrier, and nutrient transport.

During my thesis, I studied the role of the basic transcriptional machinery in intestinal homeostasis, focusing on one of its major components, the TAF4 protein (TATA-box binding protein Associated Factor 4). Using mouse models and intestinal organoids of *Taf4* invalidation, I was able to highlight its role in the homeostasis of the intestine at the embryonic and adult stages. Its action is exerted on stem cells and differentiated cells, in particular by opposing polycomb complexes, involved in the negative regulation of genes by modifying the chromatin structure.

Keywords: TAF4, Polycomb, Intestine, Stem-Cells, Development, Homeostasis, Cancers.