HAL open science

# Deciphering the role of DNA methylation in early cellular transitions

Mathieu Schulz

## To cite this version:

Mathieu Schulz. Deciphering the role of DNA methylation in early cellular transitions. Human genetics. Université Paris sciences et lettres, 2022. English. NNT : 2022UPSLS031 . tel-04016523

**HAL Id: tel-04016523**

**https://theses.hal.science/tel-04016523v1**

Submitted on 6 Mar 2023

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

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

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT

DE L'UNIVERSITÉ PSL

Préparée à l'Institut Curie, Unité de Génétique et Biologie du Développement

# Deciphering the role of DNA methylation in early cellular transitions

## Étude du rôle de la méthylation de l'ADN dans l'émergence des identités cellulaires au cours de l'embryogenèse.


Soutenue par

**Mathieu SCHULZ**

Le 9 Septembre 2022

Dirigée par

**Dr. Déborah BOURC'HIS**


École doctorale n° 515

**Complexité du vivant**

Spécialité

**Épigénétique**

**Biologie du Développement**

Composition du jury :

| | |
|---|---|
| Dr. Céline, VALLOT<br>Institut Curie | *Présidente* |
| Dr. Jamie, HACKETT<br>EMBL | *Rapporteur* |
| Dr. Christa, BUECKER<br>Max Perutz Labs | *Rapportrice* |
| Dr. Edda, SCHULZ<br>Max Planck Institute | *Examinatrice* |
| Pr. Tuncay, BAUBEC<br>Utrecht University | *Examinateur* |
| Pr. Michel, COHEN-TANNOUDJI<br>Institut Pasteur | *Examinateur* |
| Dr. Déborah, BOURC'HIS<br>Institut Curie | *Directrice de thèse* |

institutCurie | PSL

# ACKNOWLEDGMENTS

First of all, I would like to thank all the members of my Jury that have agreed to review this work; Dr. Céline Vallot, (President), Dr. Jamie Hackett, Dr. Christa Buecker (Referees), and Dr. Edda Schulz, Pr. Tuncay Baubec, Pr. Michel Cohen-Tannoudji (Examiners). I have followed your work closely throughout my PhD, and it has been a great source of inspiration that deeply influenced this thesis. Hence, I feel very humbled to have my work evaluated by such important scientists.

*"You'll see, Déborah is the most complete PI I have ever seen"*
—Déborah, according to a former post-doc.

At first, I thought this piece of post-doc wisdom was constrained only on scientific aspects… And of course, there is nothing truer than saying that Deborah is an amazing scientist. At every discussion we had, she would challenge me scientifically, raise unseen questions, providing critical feedback and advices. And it also goes without saying that I am grateful for all the freedom she gave me, allowing me to pursue my own hypotheses and design my own experiments.

But what I think this anonymous post-doc was also trying to tell me, is that Deborah also has all the qualities that makes her a great mentor, and a great person in general. I feel lucky to have been supervised by a PI that let me learn on my own, let me made my own mistakes, yet without letting me go too far in. This way, she taught me how to be to be a self-critical and independent scientist. Also, I feel lucky to have been supervised by a PI that always pushed and helped me to attend to conferences and present my project outside of the institute. I think this is how I became more confident in my work. Then, I also feel lucky to have been in a lab were one of the main focus of the PI is to always make sure the mood is always up and everyone is happy; organizing lab retreat, inviting us to restaurants and bars, supporting us during in difficult times like the pandemic. This way, she taught me to be mindful and empathetic about others. So, thank you for everything Deborah. I'll try for the rest of my career to follow your example and apply all that I have learned with you.

*“Well, why don’t you think about it today and come back ask me tomorrow?”*

—Maxim, best mentor ever.

The question I asked Maxim during my M2 internship that got me that response was something like: “*Max, I am not sure anymore; what are the exact biological questions we are trying to answer?*”.

When I started in Deborah’s Lab, Max was my supervisor during my six months long M2 internship. I learned SO MUCH with Max, he taught me everything from molecular biology to growing embryonic stem cells in tissue culture. He trusted me to set up new tools and start investigating hypotheses that were important for his own future projects, and for that I’ll always be grateful.

But Max also made me two amazing gifts. First, he let me start working on this side project he had with Deborah, which will end up being my PhD research topic. But then, and most importantly, he showed me how to build a scientific story. Max always knew how to explain his research and ideas into a beautiful narrative. And he taught me this in two ways. First by telling me that what makes a good scientist is reading (and Max is a walking Mendeley Reference Management software). Then by teaching me that a good project always starts by asking the good questions. And Max really helped me to try and find my own questions. Max, you are a talented and inspirational scientist, thank you for watching over my first steps as a young researcher.

*“Mathieu, coffee?”*

—Aurélie, without whom nothing would be possible.

Aurélie you are the best bioinformatician in the world. Not only just because you are working on dozens of different types of data (Cut&Run, RNAseq, ATAC, bulk and single cell), for dozens of projects at the same time, while teaching us bioinformatics. But also, because you are the most amazing colleague. You are just the nicest person I know, and I never felt more comfortable with anyone else in the lab talking about life in general, about things that go well and things that go wrong. I am going to miss our morning coffee breaks very much Aurélie, don’t change anything!

*"Euuuh, Mathieu? I think all my figures just disappeared…"*

—Elena, just learning that powerpoint is useless.

Elena started as an amazing master student that I mentored, and she is now on her way to become an amazing PhD student in the Bourc'his lab. I cannot thank you enough Elena, for all that you did for this project. I had the best time working with you, I consider you as friend more than a colleague now, and I know you will accomplish great things in your own PhD project. Elena, you are a true driving force for the lab, we are so lucky to have you. Thanks for everything!

*"C'est qui? C'est pic!"*

—Emeline, the heart of the party.

*"First thing I'll do after this PhD, is applying for Koh Lanta!"*

—Lorraine, the ultimate survivor.

*"We were just driving around the corner to see how fast it could go!"*

—Mathilde, after 'borrowing' a golf cart at the last Labex retreat.

Then, I would like to thank the best team of PhD students you could ever work with: Emeline, Lorraine & Mathilde. Coming to the lab in the morning never felt like a chore, because I knew I'll be hanging around with amazing people. Thank you for all the chats, and all the laughs. You guys were a great source of inspiration for me all these years.

Emeline, I cannot thank you enough for all the help you gave me when you saw that I was struggling, with the tissue culture organization or the liquid nitrogen. It seems there is nothing that you cannot do, and I know that the future of the Bourc'his PhD students' team is assured in your hands.

Lorraine, I always admired your dedication and strength. When you start something, you will always do it 100%, whether it is to achieve a scientific goal, or to organize a good time for the lab. With this kind of commitment, I know that you will become an amazing doctor, and I wish you all the best in your future career.

Mathilde, you are such an inspiration for me, and I always tried to follow your example during my PhD. In addition to be an amazing scientist, you are a mindful person, always taking care of others. Thank you for all the advices you gave me.

Next, I would like to thank all my amazing lab colleagues;

Julian, thank you for all you help in the lab, and with all the *in vivo* part of my PhD project. But also, thank you for being the only one that understand my referenced jokes on Star Wars and Kaamelott. Mélanie, thank you for your help with the Cut&Run experiments. Two full weeks of experiments, singing the dumbest music songs ever, what a ride! Fatima, thank you for all your help with the mice and the surgeries. Without you, we could never have done any chimera experiments. Nico, thank you for being such a funny and nice colleague. Despite your naïve belief in the genetics behind Jurassic Park, I know you'll accomplish great things in the lab. Charlotte & Anissa, thank you for all your encouragement and nice discussions, it was so good to have you in the lab with us, and I'm sure you'll set up the best lab in Rennes! Thank you, Alexis, Germaine and Lolita, for the laughs, the support and interest in my project.

Then, I am also grateful to all the previous member of the Bourc'his Team;

Charlotte D, for being such an inspiring scientist and amazing friend during and after your time with us in the lab. Thank you, Maud, for teaching me how to grow germ cells in a petri dish. Thank you, Margherita, my first master student, for initiating the NPC project, I hope you are thriving in your new PhD student life. I also want to thank Juliane, Joan, Raquel and Tomek, who set up the foundations on which the lab is growing now, thank you for all the advices, help and insightful discussions.

I would have never been able to survive this PhD without the support from my friends and family. I want to thank my parents, and my brother, for their loving support during my entire PhD. I also want to thank all my friends, Johanna, Valentin, Justine, Lancelot, Alexandre, Claire, Bairavie, Hugo, Yoann, and Camille. Thank you for taking care of me, and for getting my mind out of the PhD thesis, you all were like reference points during these years that always allowed me to stay grounded in '*real life*'.

Finally, Clarisse, I feel like the luckiest person in the world for having you at my sides. You supported me every day during all these PhD years, in the good and the bad times. You kept my spirits up, and I would have sink without you. I can't wait for our next adventure together in Canada!

# ABSTRACTS

# ABSTRACT IN ENGLISH

DNA methylation is an epigenetic mark associated with gene repression, playing a central role in development as demonstrated by the early lethality of mouse embryos lacking DNA methylation. During epiblast formation, embryonic stem cells (ESCs) transition from low to high DNA methylation levels, as they exit naïve pluripotency and get primed for somatic lineage commitment. In parallel, emerging primordial germ cells (PGCs) undergo DNA methylation erasure. Whilst early development represents an ideal context to study the interplay between DNA methylation and cell fates, how DNA methylation impacts priming and early lineage induction, notably germline specification, remains unclear.

To address these questions, I took advantage of DNA methylation-free ESCs (*Dnmt*-TKO) submitted to *in vitro* differentiation protocols to probe the role of DNA methylation in cell fate commitment. I found that DNA methylation was dispensable for priming and for neural induction, as TKO cells acquired functional primed features and could be engaged into neurogenesis. Strikingly, DNA methylation appeared unnecessary for adopting the germline fate, and its absence rather extended temporal competency for PGC specification. Based on chromatin profiling results, I propose that DNA methylation allows multi-lineage diversification by decommissioning DNA methylation-sensitive neural and germline enhancers during epiblast formation, therefore tempering these fates as default differentiation routes.

# ABSTRACT EN FRANCAIS

La méthylation de l'ADN est une marque épigénétique associée à la répression des gènes, dont le rôle central dans le développement est illustré par la létalité embryonnaire observée chez les souris sans méthylation de l'ADN. Durant la formation de l'épiblaste, les cellules souches embryonnaire murines (CSE) transitent d'un génome hypométhylé à hyperméthylé, alors qu'elles quittent la pluripotence naïve et s'engagent dans la pluripotence amorcée vers les lignages somatiques. En parallèle, les cellules primordiales germinales (CPG) vont, au contraire, perdre leurs profils de méthylation de l'ADN. Alors que le développement embryonnaire précoce représente une fenêtre idéale pour l'étude des liens entre dynamique de la méthylation de l'ADN et prise d'identité cellulaire, comment la méthylation impacte la transition de pluripotence, et l'adoption des différents lignages embryonnaires, notamment germinal, reste mal compris.

Au cours de ma thèse, j'ai utilisé des CSE dépourvues de méthylation de l'ADN (*Dnmt*-TKO) et des protocoles de différenciations *in vitro*, pour étudier le rôle de la méthylation dans la prise d'identité cellulaire. J'ai pu observer que si la méthylation semble dispensable pour la transition de pluripotence et l'induction du lignage somatique neurale, son absence étend la fenêtre de compétence à adopter un destin germinal. En se basant sur des analyses de cartographies chromatiniennes, je propose que la méthylation contrôle la temporalité des prises d'identité cellulaire neurale et germinale, en désengageant des éléments régulateurs de ces identités, évitant ainsi que ces deux lignées ne deviennent des voies de différentiation par défaut au cours du développement embryonnaire précoce.

# RESUME EN FRANCAIS

La méthylation de l'ADN est une marque épigénétique associée à la répression des gènes, assurant un contrôle des identités cellulaires. Cette marque est critique pour le développement, comme le démontre la létalité embryonnaire des souris mutantes pour les ADN méthyltransférases DNMT1 ou DNMT3A/3B. Durant l'embryogenèse, les profils de méthylation sont profondément reprogrammés. Après la fécondation, les marques de méthylation gamétiques sont effacées alors que le blastocyste se forme, générant des cellules souches pluripotentes naïves au génome hypométhylé, capables de donner naissance à tous les futurs lignages embryonnaires. Alors que l'épiblaste se développe, et que les cellules souches transitent vers un état de pluripotence amorcée, des profils somatiques de méthylation sont établis. Ces profils sont globalement maintenus dans les différents lignages somatiques, mais pas dans les cellules germinales primordiales qui émergent dans l'épiblaste au cours de cette transition, et qui perdent rapidement ces profils de méthylation somatiques nouvellement acquis. Il a été montré que l'absence de méthylation est compatible avec un état de pluripotence naïve, mais affecte les capacités de différenciation somatique des cellules embryonnaires. Cependant, le rôle de la méthylation de l'ADN dans la régulation de la transition de pluripotence naïve-amorcée et l'émergence de la lignée germinale reste peu caractérisé. Élucider les relations entre méthylation de l'ADN et identités cellulaires est central pour comprendre la régulation épigénétique des identités normales et pathologiques.

Durant ma thèse, j'ai interrogé la fonction de la méthylation de l'ADN dans la prise d'identité cellulaire au cours de l'embryogenèse précoce, en générant une lignée de cellules souches embryonnaire murines dépourvues de méthylation par délétion de *Dnmt1*, *Dnmt3a*, *Dnmt3b* (TKO). Mes objectifs étaient d'étudier l'impact de la méthylation de l'ADN **i)** dans l'émergence des lignages somatiques, **ii)** dans la régulation de la transition de pluripotence et **iii)** dans la spécification de la lignée germinale.

**(i)** En réalisant une différenciation somatique non dirigée *in vitro*, j'ai trouvé que les cellules TKO ne survivent pas à cette différentiation, suggérant que la méthylation de l'ADN est centrale dans la coordination de l'émergence des identités somatiques. Au contraire, les cellules TKO en différentiation dirigée vers le lignage neural gardent

leur compétence d'induction somatique vers ce lignage. Cependant, si l'induction du programme neural ne semble pas affectée par l'absence de méthylation de l'ADN, elle apparait centrale pour éviter la réactivation des gènes de pluripotence dans les stades tardifs de la neurogenèse. Ainsi, j'ai montré que si la méthylation de l'ADN peut jouer un rôle crucial dans la coordination des lignages somatiques, elle n'est pas nécessaire pour l'induction du programme neural, mais reste importante pour la consolidation de cette identité.

**(ii)** Pour comprendre comment la méthylation de l'ADN régule les gènes de la pluripotence, j'ai étudié la capacité des cellules TKO à opérer la transition de pluripotence naïve à amorcée au cours d'une différentiation épiblastique *in vitro*. En opposition au dogme établi, il est apparu que la méthylation de l'ADN n'est requise ni pour la sortie de la pluripotence naïve, ni pour l'acquisition d'une identité de pluripotence amorcée. En effet, les cellules TKO ne montrent pas de différence notable dans la régulation des gènes impliqués dans cette transition de pluripotence, et développent des capacités fonctionnelles de cellules amorcées, *in vitro* et *in vivo*. Alors qu'il est connu que la méthylation de l'ADN n'est pas nécessaire pour la mise en place de la pluripotence naïve, je propose que cette notion soit également étendue à la pluripotence amorcée, où il semble que l'absence de méthylation n'impacte pas la transition des états de pluripotence.

**(iii)** Cependant, les cellules sans méthylation de l'ADN sur-expriment les gènes de la lignée germinale au cours de la différentiation épiblastique. En induisant les cellules TKO vers un destin germinal *in vitro*, j'ai trouvé que l'absence de méthylation de l'ADN étend la fenêtre de compétence des cellules souches embryonnaires murines pour la spécification germinale. Pour comprendre comment la méthylation de l'ADN peut contrôler cette fenêtre de compétence germinale, j'ai procédé à des analyses génomiques de profilage de la chromatine. Il est apparu qu'en absence de méthylation de l'ADN, les cellules souches conservent des profils de chromatine permissifs sur des éléments régulateur de l'activité des gènes de la lignées germinale, mais aussi de la lignée neurale. J'ai également trouvé que ces éléments sont la cible de facteurs de transcription sensibles à la méthylation de l'ADN, et qui pourraient être

impliqués dans le maintien de ces profils de chromatine permissifs dans les cellules TKO. Enfin, j'ai pu observer que cet effet est spécifique au destin germinal et neural, puisque les éléments régulateurs d'autres lignages somatiques, comme le mésoderme et l'endoderme, ne montrent pas de rétention marques de chromatine permissive. De plus, en induisant la différentiation *in vitro* des cellules TKO vers les lignées méso/endoderme, j'ai observé que sans méthylation de l'ADN, les cellules souches ne parviennent pas à activer correctement les gènes régulateurs du destin méso/endoderme, mais induisent fortement l'expression de gènes neuraux. De façon similaire, l'adoption d'un destin germinal semble également être accrue dans les cellules TKO, même en absence des signaux inducteur de la lignée germinale.

En résumé, j'ai pu montrer que la méthylation de l'ADN harmonise la coordination entre les différentes identités cellulaires. La méthylation de l'ADN semble critique pour le désengagement de certains éléments régulateurs des destins germinaux et neuraux, possiblement en contrecarrant la liaison de facteurs de transcription sensibles à la méthylation. Ainsi, la méthylation de l'ADN limite l'engagement par défaut des cellules embryonnaires vers un destin neural et germinal, en temporisant le maintien de profils épigénétiques déterministiques.

# TABLE OF CONTENTS

# TABLE OF CONTENTS

**INTRODUCTION** ..... **23**

1. Preface ..... 25

2. DNA methylation: an ancient partner of life ..... 27

**2.1 DNA methylation origin and mechanistic** ..... **27**

2.1.1 DNA methylation across eukaryotic evolution ..... 27

2.1.2 The DNA methylation machinery ..... 29

2.1.3 DNA methylation, a not so static epigenetic mark ..... 34

**2.2 Functions of 5mC as a repressive mark during mammalian development** ..... **36**

2.2.1 DNA methylation in the control of germline genes ..... 37

2.2.2 DNA methylation, a protector of genome against transposable elements ..... 38

2.2.3 DNA methylation in gene imprinting and X chromosome inactivation ..... 41

**2.3 DNA methylation at the cross-road of chromatin pathways** ..... **44**

2.3.1 Nucleosome structure and histone modifications ..... 46

2.3.2 DNA methylation in association with repressive chromatin marks ..... 49

2.3.3 DNA methylation interplays with active chromatin marks ..... 53

3. A tale of early development ..... 58

**3.1 From fertilization to implantation** ..... **58**

3.1.1 Parental gametes and first reprogramming: prepping the embryo ..... 58

3.1.2 Zygotic genome activation: the great reboot ..... 61

3.1.3 To implantation and beyond ..... 65

**3.2 From in vivo to in vitro** ..... **71**

3.2.1 Recapitulating early development in a petri dish ..... 72

3.2.2 Functional assays and tools to test the potentialities of stem cells ..... 78

**3.3 From totipotency to pluripotency** ..... **80**

3.3.1 Totipotency: the ability for all things ..... 80

3.3.2 A switch in pluripotency: the naïve-to primed-pluripotency transition ..... 83

3.3.3 DNA hypomethylation and naïve pluripotency, cause or consequence? ..... 86

3.3.4 DNA methylation in the regulation of pluripotency transitioning ..... 89

4. Lineage specification ..... 93

**4.1 Germline specification: closing the loop** ..... **93**

4.1.1 PGC specification and reprogramming ..... 94

4.1.2 DNA hypomethylation, a hallmark of naivety in pluripotent and germ cells ..... 99
4.1.3 Human and mouse germline epigenetics: so similar, yet so different ..... 101

**4.2 Emergence of somatic lineages—all the rest that matters ..... 104**
4.2.1 Transcriptional and epigenetic control of somatic lineage specification ..... 105
4.2.2 Cracks in the lineage specification program ..... 108

**RESULTS ..... 113**

**Impact of DNA methylation in cell fate decisions ..... 115**
**ABSTRACT ..... 118**
**INTRODUCTION ..... 119**
**RESULTS ..... 121**
**DISCUSSION ..... 141**
**MATERIAL AND METHODS ..... 144**
**REFERENCES ..... 153**
**EXTENDED FIGURES AND TABLES ..... 158**

**DISCUSSION ..... 169**

**1. General Summary ..... 171**

**2. DNA methylation in stem cells: a one-way relationship ..... 173**

**2.1 DNA methylation is dispensable for stem cell identities ..... 173**
2.1.1 Redundancy of repressive epigenetic marks in naivety ..... 173
2.1.2 DNA methylation is not required for priming ..... 174
2.1.3 Different *Dnmt* Knock-outs, different chimerae ..... 174

**2.2 Uncoupling transcription and epigenetic reprogramming ..... 176**
2.2.1 DNA methylation, critical for silencing naivety? ..... 176
2.2.2 DNA methylation at naïve regulatory elements, critical for priming? ..... 177

**3. A head start in cell fate decisions for DNA methylation ..... 180**

**3.1 DNA methylation, an instructor of formative pluripotency ..... 180**
3.1.1 The absence of DNA methylation promotes "*hyper-formativity*"? ..... 180
3.1.2 How DNA methylation and *in vitro* models affect formativity? ..... 181

**3.2 DNA methylation harmonizes differentiation routes ..... 182**
3.2.1 DNA methylation decommissions neural and germline-associated enhancers . 182
3.2.2 DNA methylation counteracts neural and germline fates as default routes ..... 183
3.2.3 Why neural and germline identities constitute default routes in development? . 184

4. Final Perspectives and Conclusions .....186
4.1 Limitations and future perspectives of the study .....186
4.1.1 Be aware of your genetic background and culture conditions ..... 186
4.1.2 Adapting our strategy..... 187
4.1.3 Metabolism, a possible new avenue? ..... 189
4.2 Final perspective and personal appreciation .....190
BIBLIOGRAPHY ..... 191

# INTRODUCTION

# 1. Preface

How organisms are formed is a longstanding question in the history of life sciences. Aristotle in his treaty "On the Generation of Animals" written around 330-322 BC postulated that the development of an organism is a progressive growth process starting from the "germ", and mostly driven by external cues. This is referred to as the "epigenesis model", in opposition to the "preformation model", which assumed that a preformed organism is contained in male or female germ, with its final shape although of smaller size. Centuries later, in 1940, English embryologist Conrad Waddington, in an attempt to conciliate the genotype (with its static nature alike the preformation model), and the phenotype (subject to environmental influences alike the epigenesis model), coined the term "Epigenetics". From this, he drew his idea of an "epigenetic landscape"; a way of illustrating how external signals influences the genotype to produce the phenotype during development *(Allen, 2015)*.

In the mid-XX$^{th}$ century, new mechanisms regulating gene expression are discovered. Among these fundamental studies, the identification of transcription factors by Jacob and Monod *(Jacob and Monod, 1961)*, the Polycomb group of protein by Lewis *(Lewis, 1978)* and DNA methylation by Holliday *(Holliday and Pugh, 1975)* as key actors of cellular memory, shed some light on how genes can relate on external factors to regulate their activity. However, with the advent of molecular biology, and uncertainties on how such processes mechanistically achieve their regulation, came up a period of "genetic reductionism" starting in the early 1980's from which will arise the central molecular dogma of biology enunciated as '**one gene encode for one protein, leading to one phenotype**' *(Pisco et al., 2016)*. However, after the completion of the Human Genome Sequencing Project in the beginning of the XXI$^{st}$ century, came the realization that genes alone do not hold the key to fully understand development. From DNA methylation to the Histone code *(Jenuwein and Allis, 2001)*, and others, these processes became the irrefutable support of what we now call the "epigenetic information". Thus, the word "epigenetics" took on another definition, enunciated notably by Holliday, and later elaborated by Bird and others *(Bird, 2007; Holliday, 1994)*, and is now mostly understood as the "**the processes by which genes can regulate and maintain their activity across cell divisions and**

**development without changes in the DNA sequence**". Thus, epigenetics is now considered as a central way regulating morphogenesis and cell differentiation, as the genetic information remains the same in every cell type **(Figure 1)**.

During my PhD project, focusing on DNA methylation, I aimed at understanding how this epigenetic mark could impact early development using mouse embryonic cells as a model. More specifically, I investigated peri-implantation development, a period when embryos make crucial cell fate decisions while profoundly reshaping their epigenetic landscape. In this Introduction, I will first make on overview of DNA methylation in evolution, with its regulation and known function(s) in development. Then, I will focus on early mouse development, and the emergence of the pluripotent stem cells. During this part, I will emphasize the epigenetic reprogramming that stem cells undergo while they acquire the ability to give rise to the embryonic tissues, alongside the methods and tools available to study this developmental period. Finally, I will describe how the somatic and germline identities are acquired in embryonic cells.

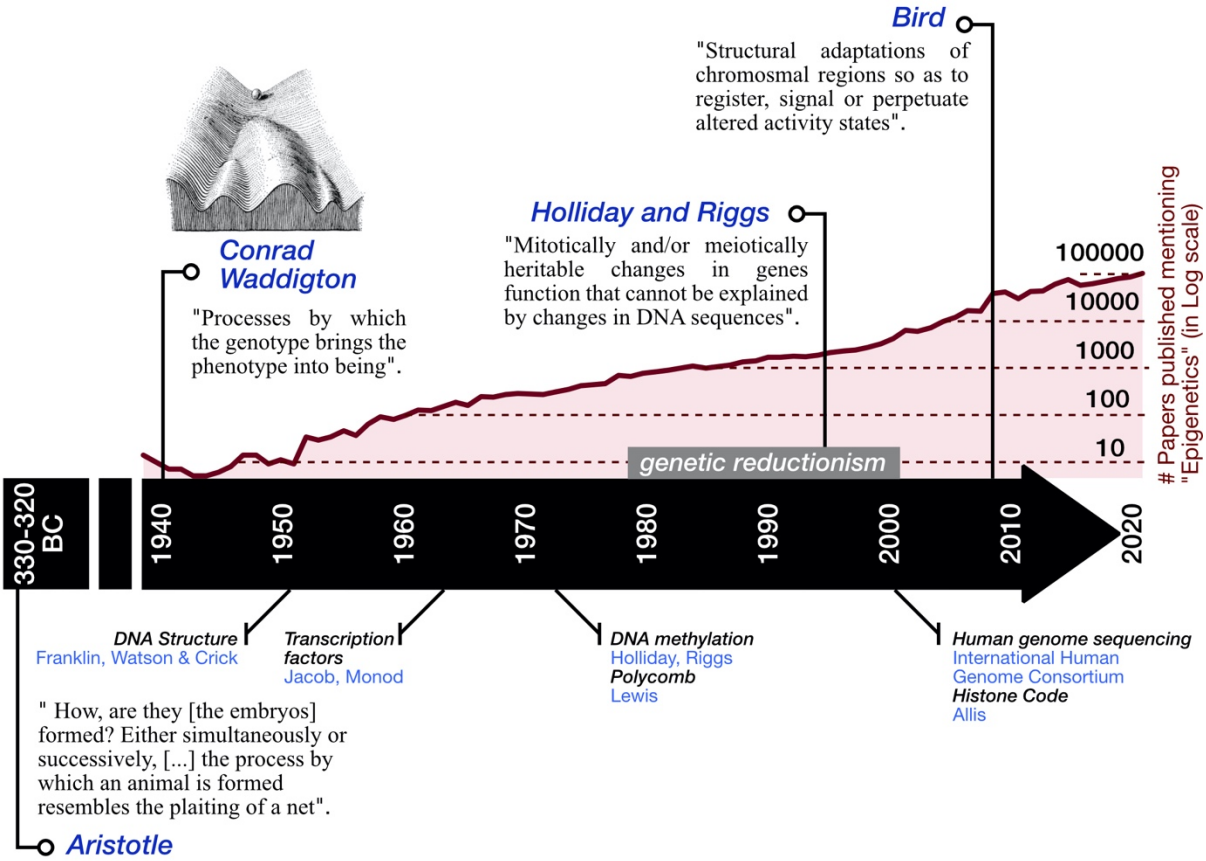



**Fig. 1 I Historical perspective on "Epigenetics"**
Adapted from Pisco, *et al.* 2016 (BioRxv) and Cavalli & Heard, 2019 (Nature). Line curve representing the number of published research articles mentioning "Epigenetics" in Log scale from 1940 to 2020 (Data obtained from https://app.dimensions.ai/discover/publication).

# 2. DNA methylation: an ancient partner of life

## 2.1 DNA methylation origin and mechanistic

Chemical addition of a methyl group (-CH3) on DNA nucleotides is a highly conserved modification that is found in the three reigns of life. In bacteria, it is found under various forms, like N6-methyl-adenine (m6A), C5-methyl-cytosine (5mC) and N4-methyl-cytosine (4mC) *(Sánchez-Romero and Casadesús, 2020*). Although DNA methylation in the context of bacteria has been initially perceived as a primitive genome defense mechanism against bacteriophages through restriction-modification systems, it has then been broadened to new functions, like transcriptional regulation, virulence, and as a selective force for bacterial evolution *(Cota et al., 2015; Fang et al., 2012; Marinus and Casadesus, 2009)*. In contrast, 5mC and m6A are the only modifications found in eukaryotic genomes. Although it is supposed that the last common eukaryotic ancestor did not bear DNA methylation, these marks seem to have been nevertheless inherited from bacteria during eukaryogenesis *(Talbert et al., 2019)*.

### 2.1.1 DNA methylation across eukaryotic evolution

Genomic analyses have revealed a great variability in the content and profiles of genomic DNA methylation across the eukaryotic kingdom. First of all, 5mC appears to be the predominant form of DNA methylation in eukaryotes. It is mostly found on CG (CpG dinucleotides), but also occurs rarely in a CHG or CHH context (where H= A, T or C) *(Zemach et al., 2010)*. While mCpGs are found in every vertebrate and plant genomes studied so far and play conserved functions like in genomic imprinting or transposon repression (Niederhuth and Schmitz, 2017; Zhang et al., 2018), global levels of 5mC are strikingly variable among eukaryotes' genome. Indeed, many species do not harbor any DNA methylation *(Bewick et al., 2019, 2017)*, such as the model organisms D. *melanogaster*, C. *elegans*, or S. *cerevisiae*, suggesting that this epigenetic modification is not fundamentally required for the survival of some eukaryotic organisms **(Figure 2A)**. Coincidently, DNA methylation-free organisms rely

on alternative mechanisms to regulate gene or transposon expression. Nevertheless, DNA methylation remains critical for vertebrates and flowering plants development, suggesting that in these organisms, DNA methylation might have acquired functions that cannot be fully compensated by other repressive mechanisms *(Schmitz et al., 2019)*.

Paradoxically, species with heavily methylated DNA typically demonstrate CG 'poor' genomes *(Bird, 1980; Schorderet and Gartler, 1992)*. Indeed, the deamination of pyrimidines, including cytosines, is a highly mutagenic process occurring at a rate of approximately 200-300 events per cell per day, converting cytosines into thymines. Moreover, the rate of 5mC deamination is 3-fold higher than the one of unmethylated cytosines. While DNA methylation plays crucial functions, it also comes with evolutionary costs: deamination leads to potential accumulation of miscoding mutations in genes and elevated single nucleotide polymorphisms among individuals, which can be inherited in their offspring if occurring in the germline. Incidentally, it is estimated that a third of mutations found in coding regions of tumoral cells arise from 5mC deamination at CpG sites *(Bellacosa and Drohat, 2015)*. Therefore, in mammalian genomes, CpGs occur at a lesser frequency than expected. Namely, the human genome displays only one fifth of the expected CpG number *(International Human Genome Sequencing Consortium, 2001)*.

In average, vertebrate genomes harbor around 80% of methylated CpGs, with the notable exception of CpG-rich regions, termed CpG islands (CGI) *(de Mendoza et al., 2020)*. These hypomethylated regions are defined as sequences of at least 200bp with a GC content over 50% and a ratio of observed/expected CpGs over 0.6. In mammalian genomes, almost two-third of protein-coding genes bear a CGI in their promoter regions, alike many housekeeping genes *(Gardiner-Garden and Frommer, 1987; Greenberg and Bourc'his, 2019)*. As a consequence, half of CGIs are found around gene promoters in mammalian genomes, while the rest map to gene bodies or intergenic regions where they have been proposed to act as alternative gene promoters and/or, more recently, as tethering elements mediating physical contact between genes and their distal regulatory elements *(Illingworth et al., 2010; Pachano et al., 2021)*.

### 2.1.2 The DNA methylation machinery

The eukaryotic DNA methyltransferases (DNMTs) bear a catalytic domain (MTase) made of 10 motifs that share strong conservation with bacterial methyltransferases. In fact, the predictive value of these motifs is so high that almost all eukaryotic cytosine methyltransferases were identified by the presence of these motifs *(Goll and Bestor, 2005)*. Eukaryotic DNMTs are classified into several lineages **(Figure 2B-C)**:

**The DNMT1 maintenance super family of methyltransferases:**

This first lineage contains the maintenance family of methyltransferases, which are related to DNMT1 and its derivatives, MET1/ CMTs in plants and RID/DIM-2 in fungi (also sometimes known as DNMT4). This family most probably emerged from ancient duplications of an ancestral form of DNMT1, allowing the duplicated copies to acquire new functions in a species-specific manner *(de Mendoza et al., 2020; Huff and Zilberman, 2014)*. The conserved structures in all DNMT1 family members encompass the MTase domain, as well as one or several Bromo-adjacent Homology domains (BAH), whose function are still uncertain but their presence in other chromatin regulators suggests they could act as a platform for protein-protein interactions. The murine DNMT1 was the first DNA methyltransferase to be identified in mammals, in 1988 *(Bestor et al., 1988)*. It contains additional domains including a replication foci-targeted domain (RFD) and a DNA-binding domain (CXXC) **(Figure 3A)** *(Song et al., 2011)*. In order to perpetuate DNA methylation across cell division, DNMT1 is targeted to replication foci by binding the proliferating cell nuclear antigen PCNA via its RFD domain. DNMT1 is then recruited to hemi-methylated DNA by nuclear protein 95 (NP95 or UHRF1), which contains a SET and RING-associated (SRA) domain that preferentially binds to hemi-methylated DNA *(Bostick et al., 2007; Sharif et al., 2007)*. Without this interaction, DNMT1 adopts an auto-inhibitory conformation, preventing any catalytic activity outside these hemi-methylated regions *(Song et al., 2011)*. In addition, more recent data provided novel insights into time-dependent modes of DNMT1 recruitment to the replication machinery, depending on ubiquitylation of the protein PAF15 in early S phase and on Histone H3 in late S phase *(Nishiyama et al., 2020)*. Finally, recent analyses proposed weak *de novo* activity for DNMT1 in mice embryonic stem cells, notably toward some

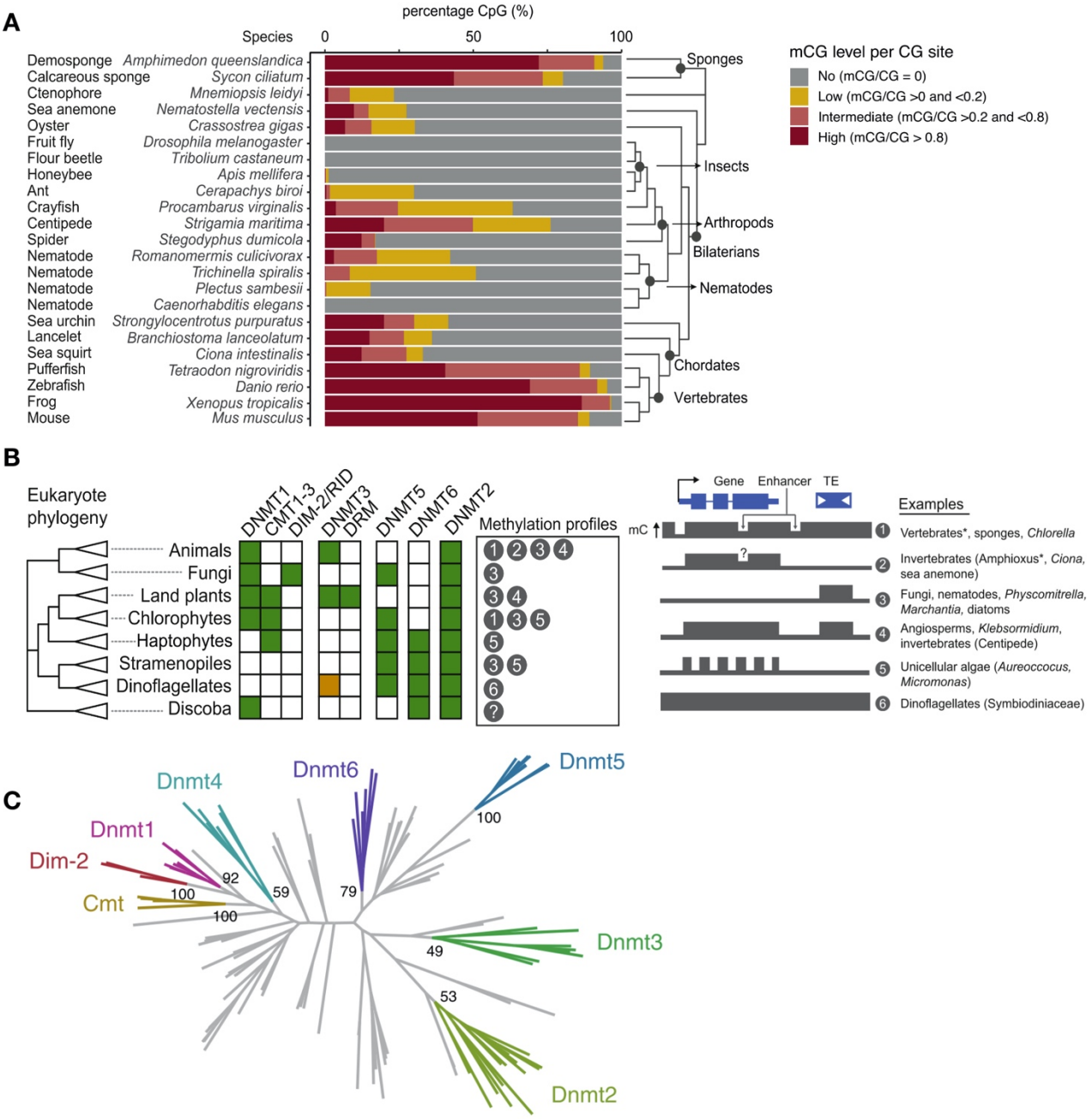



**Fig. 2 I DNA methylation in the eukaryotic kingdom of life**

**A.** Percentage of methylated CpG sites among various and representative eukaryotic species ordered based on phylogenetic relationships (coverage >10X). **B.** DNA methyltransferase (DNMT) distribution among major eukaryotic clades, associated with their DNA methylation profiles, as described on the right. Asterisks indicate species for which the methylation status of enhancers has been characterized (de Mendoza, *et al.* 2020 JMB). **C.** Maximum likelihood family tree for DNMT structures found in eukaryotic (colored branches) and prokaryotic (gray branches) species. Numbers indicate percentage of supported bootstrap for each node in eukaryotes, underlying overall similarity between DNMT families (Huff, 2014 Cell).

classes of transposable elements in mice embryonic stem cells (ESCs) *(Haggerty et al., 2021)*. This is consistent with reminiscent *de novo* activity of purified DNMT1 protein observed on unmethylated DNA substrate in biochemical assays *in vitro* *(Fatemi et al., 2001; Vilkaitis et al., 2005)*. Yet, this conclusion remains controversial: it is not clear how this *de novo* activity of DNMT1 could be directed specifically toward transposable elements. Moreover, remnant activity of DNMT3C—a rodent-specific *de novo* methyltransferase normally expressed in mouse spermatogenesis— at young transposable elements in mouse ESCs, could potentially explain this low *de novo* activity *(Barau et al., 2016; Wang et al., 2020)*. Finally, *in* vivo studies still exclude a role for *de novo* methylation by DNMT1 during early mouse embryogenesis *(Dahlet et al., 2020)*. Interestingly, some species only possess a maintenance DNA methyltransferase **(Figure 2B)**, without known active *de novo* DNA methyltransferase enzymes. This implies that these species maintained their methylation profile across evolutionary times, demonstrating the strong fidelity in the propagation of DNA methylation patterns by the maintenance methyltransferases *(Catania et al., 2020)*. Another sub-family of maintenance DNA methyltransferases, known as DNMT5, is found in many eukaryotic lineages, from fungi to Chlorophytes. Although related to DNMT1 structurally, DNMT5 activity is regulated by a very different mechanism. Briefly, a positive loop between DNMT5 hemi-methylated binding domain stimulates the ATPase activity of its SNF2 domain, catalyzing the methylation of cytosines. This renders DNMT5 activity toward hemi-methylated DNA much more specific than other DNMT1 proteins and with higher fidelity *(Dumesic et al., 2020)*.

**DNMT3 and DNMTX family catalyze *de novo* DNA methylation:**

In contrast to the DNMT1 family, the lineages that contain *de novo* DNA methyltransferases are very diverse and hold very different structure conformations among eukaryotes. The most well-studied *de novo* DNMTs are undoubtedly the DNMT3-related enzymes, present in all eukaryotic lineages. The phylogenetic link between DNMT1 and DNMT3-related proteins remains unclear. However, since their phylogenetic relationship is rooted among prokaryotic MTases, this suggests that DNMT1/DNMT3 families were acquired independently from prokaryotic/viral sequences during evolution *(De Mendoza et al., 2018; Huff and Zilberman, 2014)*.

Some land plants, like angiosperms, have developed homologs of DNMT3s: Domains Rearranged methyltransferases (DRMs). These DRMs, alike DNMT3s, contain a MTase domain, but also display a ubiquitin associated domain that is absent in DNMT3s, and can methylate cytosines in CHH sequence contexts *(Law and Jacobsen, 2010; Yaari et al., 2019)*. Another type of *de novo* DNMTs is the recently characterized DNMTX that has been identified in some fungus species and contains a BAH domain in addition to its catalytic domain. However, DNMTX has been lost in many fungi that now rely only on DNMT5 proteins to maintain their 5mC patterns *(Catania et al., 2020)*. Mammalian genomes also demonstrate strong innovations when looking at *de novo* DNMT3s. After the discovery of two functional *de novo* DNMT3s in the mouse genome, DNMT3A and DNMT3B *(Lei et al., 1996; Okano et al., 1998)*, which are shared among all mammals, a third *Dnmt3*-Like (*Dnmt3L*) gene was identified. It shares strong homology with *Dnmt3A* from which it probably emerged through duplication in placental mammals, despite the absence of catalytic activity *(Aapola et al., 2000; Yokomine et al., 2006)*. Finally, DNMT3C emerged from a duplication of the *Dnmt3B* gene in Muroidea rodents specifically, around 46 to 60 million years ago *(Barau et al., 2016)*.

The typical architecture of mammalian DNMT3 proteins is comprised of a PWWP (Pro-Trp-Trp-Pro) and an ADD (ATRX-DNMT3L-DNMT3A) domains, in addition to the classic MTase domain, as originally reported in DNMT3A and DNMT3B. Yet, DNMT3L lacks the PWWP domain and displays a non-functional MTase domain *(Lyko, 2018)*. Nonetheless, DNMT3L acts as an essential co-factor for *de novo* DNMT3s: biochemical assays have shown interactions through the C-terminal domain of DNMT3L and ADD domain of DNMT3A/B and stimulation of the activity of DNMT3A/B by DNMT3L, which is essential for germline *de novo* DNA methylation specifically *(Bourc'his et al., 2001; Chen et al., 2005; Jia et al., 2007; Suetake et al., 2004)* **(Figure 3A- 3B)**. DNMT3C also lacks a PWWP domain, while showing strong overall homology with DNMT3B, with the notable exception of its N-terminus region *(Molaro et al., 2020)*.

The PWWP domain allows for the recognition of H3K36 di- and tri-methylation (K3K36me2/3), the latter being usually found at gene bodies as it is deposited by SET Domain enzyme 2 (SETD2) alongside the transcriptional machinery *(Baubec et al.,*

*2015; Dhayalan et al., 2010; Krogan et al., 2003; Sun et al., 2005)*. This explains the strong association observed between DNA methylation at gene bodies and high transcriptional activity of these genes. Then, recent work demonstrated differential affinity for the PWWP domains of DNMT3A and 3B, which have a higher affinity for H3K36me2 and H3K36me3, respectively *(Weinberg et al., 2019)*. A hypothetical function for DNA methylation at intragenic regions could be the repression of cryptic promoters and prevention of spurious transcription *(Neri et al., 2017)*. Finally, the PWWP domain is also critical to address the DNMT3s at pericentric heterochromatin that is enriched in H3K36 methylation *(Chen et al., 2004)*.

The ADD domain allows DNMT3s to bind to unmethylated H3K4 residues. If methylated—with H3K4 tri-methylation (H3K4me3) being the most repulsive form—the ADD domain folds into an auto-inhibitory conformation impairing the MTase activity *(X. Guo et al., 2015; Otani et al., 2009)*. As a result of this strong antagonism between H3K4me3 and the *de novo* DNMT3s, regions marked by H3K4me3, like promoters and CpG islands, are usually DNA methylation-poor *(Weber et al., 2007)*. More details about the interplay between DNA methylation, DNMTs and other chromatin marks will be provided later in this Introduction.

Mammalian *de novo* DNMT3s, although similar structurally, play non-redundant functions and are associated to different outcomes when deleted during development. Indeed, while DNMT3A and DNMT3B are critical for proper development *in vivo*, DNMT3B is the predominant catalyzer of DNA methylation deposition during embryogenesis *(Borgel et al., 2010; Yagi et al., 2020)*, while DNMT3A plays a more critical role around developmental genes and in differentiated tissues, including germ cells *(Dura et al., 2021; Shirane et al., 2013)*. Similarly, DNMT3C developed a selective activity toward young transposable elements during spermatogenesis. While the knock-out mice of *Dnmt3L* or *Dnmt3C* are viable but infertile *(Barau et al., 2016; Bourc'his et al., 2001)*, deletion of DNMT3A or DNMT3B demonstrate stronger phenotype. Notably, *Dnmt3A*-KO mice despite being also infertile, show growth retardation, and die four weeks after birth, while *Dnmt3B*-KO mice die at mid-gestation *(Okano et al., 1999; Ueda et al., 2006)*. Similarly, *DNMT3A* or *DNMT3B* mutations in humans are associated with a number of diseases, like

microcephalic dwarfism, acute myeloid leukemia, and immunodeficiency *(Greenberg and Bourc'his, 2019)*.

**DNMT2 and DNMT6 families, the "fake" DNA methyltransferases:**
Although they are associated to the DNA methyltransferase family, no chromatin binding domain can be found in DNMT2/DNMT6 proteins. As a result, these DNMTs do not target DNA sequences but other substrates *(de Mendoza et al., 2020)*. Indeed, DNMT2 is deeply conserved among eukaryotes, being found in all species where DNMT1 and DNMT3 homologs are present, and has been reported to act as a tRNA methyltransferase *(Goll et al., 2006; Goll and Bestor, 2005; Lyko, 2018)*. On that line, a recent study has shown that altering DNMT2 in sperm of mice alters the small non-coding RNA signature affecting paternal epigenetic inheritance *(Yunfang Zhang et al., 2018)*. On the other hand, the DNMT6 proteins are less well characterized, but one study found no change in DNA methylation level when *Dnmt6* is knocked-out or overexpressed in the protozoan *Leishmania donovanil* *(Cuypers et al., 2020)*. Whilst DNMT6 demonstrates no activity at the chromatin level, a role related to RNA methylation, alike DNMT2, cannot be excluded.

### 2.1.3 DNA methylation, a not so static epigenetic mark

DNA methylation is considered as a stable and robust epigenetic mark, notably because of the reliable maintenance activity of DNMT1. However, DNA methylation patterns also present some dynamic properties. Most notably, in mammals, DNA methylation undergoes waves of developmental reprogramming, being almost completely erased in early embryogenesis and gametogenesis.

They are two main processes by which 5mC can be removed. The first is managed via passive dilution of 5mC during cell division and inhibition of the maintenance activity of 5mC by excluding DNMT1 from the replication fork (von Meyenn et al., 2016b). The second mechanism involved in DNA demethylation relates to active enzymatic removal of 5mC **(Figure 3C)**. The Ten-eleven Translocation (TET) proteins—which have been extensively characterized in mammals—are very conserved among eukaryotes, as they originate from lateral gene transfer from

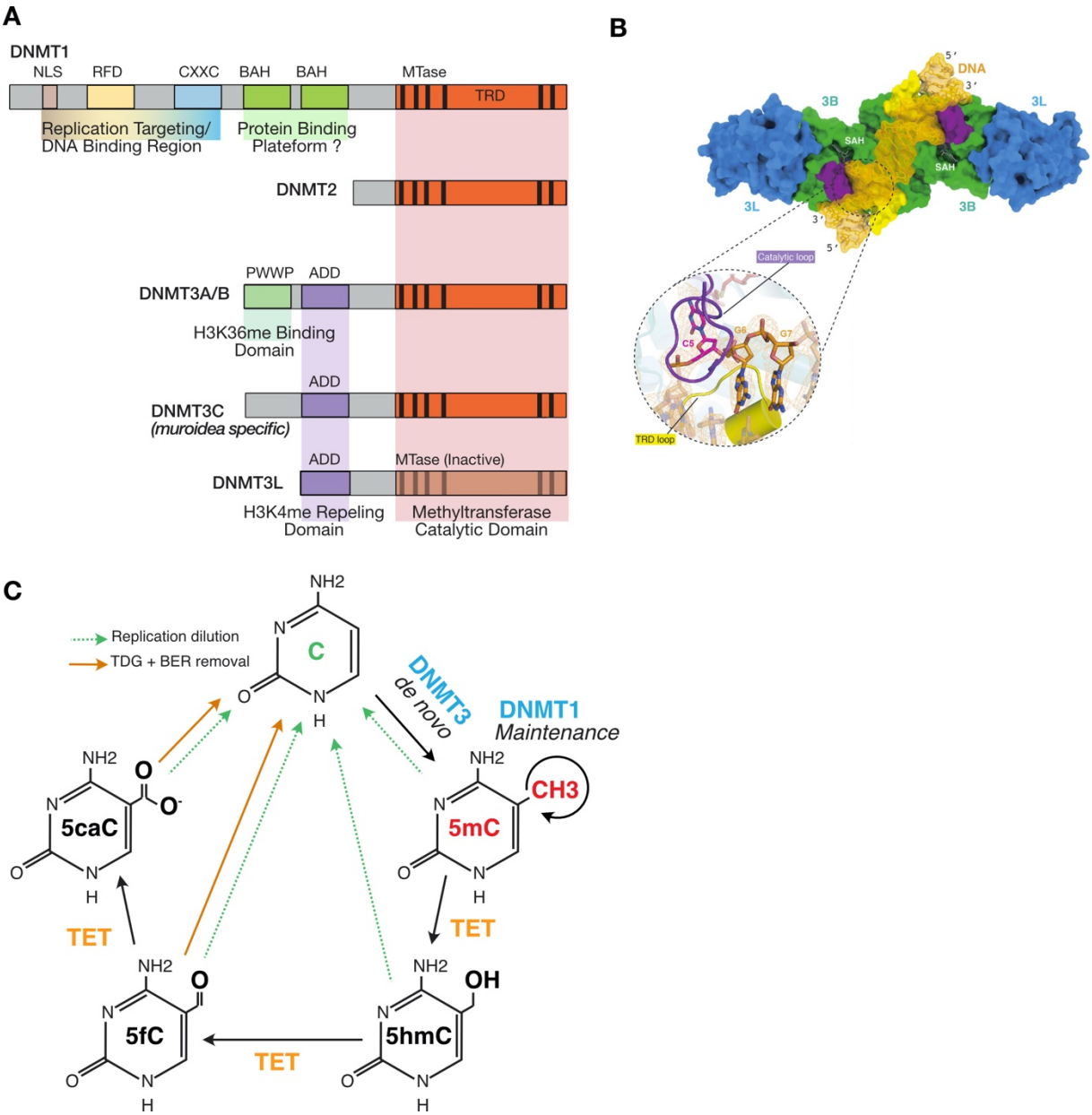



**Fig. 3 I DNMTs and DNA methylation machinery in mammals**

**A.** Architectural structure of mammalian DNA methyltransferases containing; NLS= nuclear localization signal, RFD= Replication foci domain, CxxC= DNA binding domain, BAH= Bromo-adjacent homology domain, MTase= Methyltransferase catalytic domain, PWWP= Pro-Trp-Trp-Pro histone lysine binding domain, ADD= ATRX-DNMT3A-DNMT3L histone lysine binding domain. **B.** 3D Structure of DNMT3B-DNMT3L tetramers binding to the DNA; with a close-up view of the TRD (target recognition domain) and catalytic loops of the MTase on a cytosine residue (C. Lin, *et al.* 2020 Nucleic acids research). **C.** The cycle of DNA methylation; *de novo* established and maintained by DNMT3s and DNMT1, and actively removed through the Ten-eleven translocation (TET) protein or passively by cell division and DNMT1 inactivity (Wu & Zhang 2017 Nat Rev Genetics).

bacteriophages who bear a minimal version of these proteins *(Iyer et al., 2011)*. All mammalian TETs (TET1, 2 and 3) harbor a double stranded ß-helix fold (DSBH) that is responsible for the conversion of 5mC into 5-hydroxymethylcytosine (5hmC), then into 5-formylcytosine (5fC) and finally into 5-carboxylcytosine (5caC) *(Ito et al., 2011)*. These derivative forms of 5mC can then be removed from the DNA by dilution during replication and cell division, as UHRF1 and DNMT1 show lower ability to recognize and maintain these marks on the chromatin *(Hashimoto et al., 2012)*. Conversely, Thymine DNA glycosylase (TDG) pathways can recognize 5fC and 5caC forms and excise these modified bases from the chromatin, followed by the restoration of DNA integrity by the base-excision reparation (BER) pathway *(He et al., 2011; Wu and Zhang, 2017)*. Some links might exist between active and passive DNA demethylation, as a recent study reported that TETs induce *Dppa3/Stella* expression in mouse embryonic stem cells, whose product initiates genome-wide DNA demethylation by excluding UHRF1and DNMT1 from the chromatin in mammals *(Mulholland et al., 2020a)*.

Finally, TET enzymes are also important for mammalian development. They are involved in the regulation of many developmental genes, often in synergy with the DNMTs, thus impacting cell fate decisions in early mammalian embryogenesis *(Charlton et al., 2020; Dai et al., 2016; Dawlaty et al., 2014; Verma et al., 2018)*. Moreover, TETs also play important roles in other contexts like meiosis and the control of transposable elements *(Guallar et al., 2017; Yamaguchi et al., 2012)*. Whilst being central for mammalian development, not all TETs are expressed and likely required simultaneously during development. Notably, while TET1 and TET2 are both present in embryonic stem cells, TET3 is the prominent DNA demethylase enzyme after fertilization *(Mulholland et al., 2020b; Shen et al., 2014)*.

## 2.2 Functions of 5mC as a repressive mark during mammalian development

DNA methylation is usually associated with gene repression when located on regulatory elements. However, the presence of mCpGs does not necessarily induce silencing. I mentioned before the case of CGI promoters that are found on almost 70% of human and mouse genes. These promoter-associated CpG islands remain, in

fact, unmethylated independently of the transcriptional status of the genes *(Weber et al., 2007)*. In contrast, CpG-poor promoters can be sparsely methylated, but this does not impact their transcriptional activity. This suggests that DNA methylation-mediated repression of gene transcription is mostly dependent on CpG density **(Figure 4A)**.

In addition, several transcription factors (TFs) have been found to be sensitive to DNA methylation and can affect transcriptional outcomes. Among them, the Methyl-CpG binding domain (MBD) proteins (MBD1/2/4 and MeCP2) *(Baubec et al., 2013)* can bind to methylated CpGs and induce chromatin compaction through the formation of secondary structures *(Georgel et al., 2003)*, or the recruitment of chromatin modifiers like histone deacetylase, leading to gene silencing *(Ng et al., 1999)*. Reversely, some transcription factors show preferential binding to unmethylated sequences and safeguard transcription by opening the chromatin, like BANP *(Grand et al., 2021)*, or by repelling *de novo* DNA methylation, like NRF1 and the recently discovered QSER1 transcription factor **(Figure 4B- 4C)** *(Dixon et al., 2021; Domcke et al., 2015)*. The exact mechanisms by which DNA methylation mediates its repressive activity are still incomplete. Nonetheless, there are biological processes where DNA methylation-mediated repression remains central: germline-gene repression in somatic lineages, transposable element control, parent-specific imprinting regulation and X chromosome inactivation.

### 2.2.1 DNA methylation in the control of germline genes

Germline genes encompass genes involved in meiosis, general gametogenesis, but also some germline-specific “genome-defense” genes, such as members of the piwi-interacting RNA (piRNA) pathway. The majority of these genes bear a methylated CGI promoter in somatic cells, and are especially sensitive to DNA methylation loss *(Dahlet et al., 2020)*. The temporal repression exerted by DNA methylation on these genes is critical for proper germline development, as many of these DNA methylation-sensitive genes have been found important for the formation of differentiated germ cells, like *Dazl* and *Asz1* *(Haston et al., 2009; Wang et al., 2013)*. Moreover, inability to maintain DNA methylation in the emerging male germline *in vivo* leads to precocious acivation of meiotic genes and decreased germ cell numbers, as exemplified in conditional *Dnmt1*-KO mutant mice *(Hargan-Calvopina et al., 2016)*. But beyond regulating proper

temporal expression during germ cell development, DNA methylation is of crucial importance for silencing germline genes in somatic cells. Indeed, cancer cells frequently adopt an hypomethylated genome and tend to reactivate germline genes, which could promote oncogenic pathways *(Simpson et al., 2005; Van Tongelen et al., 2017)*.

The mechanisms by which DNA methylation is deposited at germline genes are under strong investigation. The most prevalent model postulates that E2F6, a subunit of the non-canonical Polycomb Repressive Complex 1 (PRC1), targets and recruits DNMT3B at meiotic genes, as they have been found to co-interact. This mechanism has been observed first in somatic cells *(Pohlers et al., 2005; Velasco et al., 2010)*. However, *E2f6* knock-out only induces partial reduction of DNA methylation at germline genes in mouse embryos, suggesting that alternative mechanisms address DNMTs at these loci *(Dahlet et al., 2021)*. Potential candidates include SETDB1 and MGA/MAX, which are also involved in the repression of germline genes through the deposition of repressive histone marks H2AK119Ub, H3K27me3 and H3K9me3, in hypomethylated embryonic stem cells *(Endoh et al., 2017; Karimi et al., 2011)*. A recent and thorough reanalysis of published data indeed demonstrated that H2Aub and H3K9me3 (but not H3K27me3) are the primary marks by which germline genes are repressed in embryonic stem cells, while DNA methylation-mediated repression becomes more prevalent as cells differentiate in the epiblast *(Mochizuki et al., 2021)*, where it probably acts as a final repressive “lock” to ensure germline gene silencing in differentiated cells **(Figure 4D)**.

### 2.2.2 DNA methylation, a protector of genome against transposable elements

Transposable elements (TEs) are “mobile” elements that can move around the genome either by a “copy-paste” (retrotransposons) or a “cut-paste” (DNA transposons) mechanism. These elements represent a significant amount of the genome in mammals *(Senft and Macfarlan, 2021)*. In mice, they constitute at least 44% of the genome, with the LINEs, SINEs (Long/Short interspaced nuclear elements) and LTR-elements (Long terminal repeats) representing 19%, 8% and 10% of the genome, respectively *(Mouse genome sequencing consortium, 2002).* Considered for

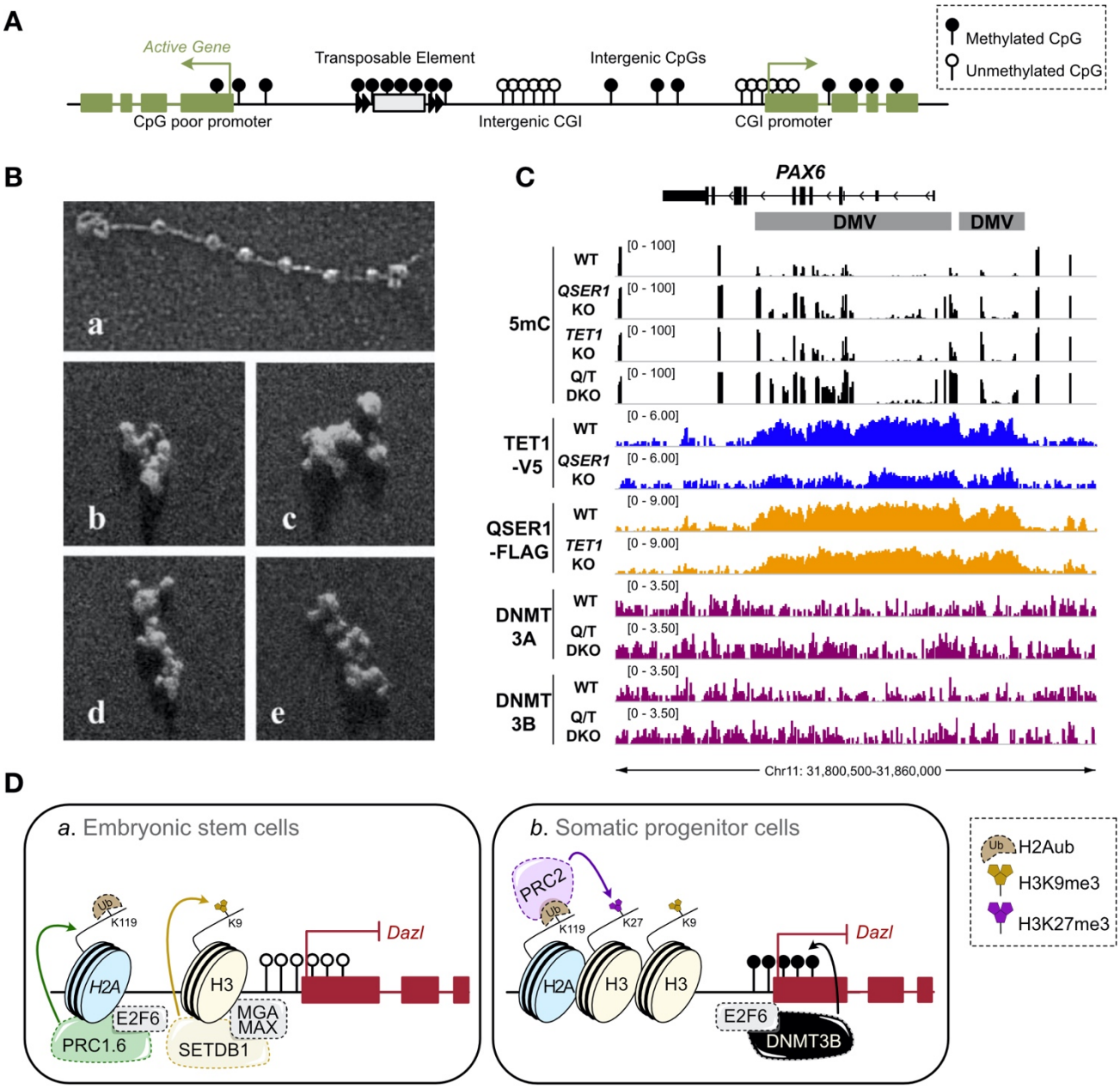



**Fig. 4 I DNA methylation patterning and its control on germline genes**

**A.** DNA methylation landscape in CpG-poor mammalian genomes: most DNA methylation is found on intergenic CpGs, on CpG-poor gene promoters and around transposable elements, while CpG Islands (CGIs) remain mostly unmethylated. **B.** Electron microscopy pictures of 10 nucleosomes associated with DNA. Subpanel (*a*) shows nucleosomal array incubated in native condition. Subpanels (*b-e*) show nucleosomal array forming secondary structures and condensation when exposed to MeCP2 (Georgel, *et al* 2003 Journal of biological chemistry). **C.** Genomic tracks at the *Pax6* locus for DNA methylation (black), TET1 (Blue), QSER1 (Yellow) and DNMTs (Purple), in WT, *Tet1 and Qser1* single or double knock-outs, in differentiated cells. This graph show how binding of QSER1 recruit TET1 and active DNA demethylation, while repulsing DNMTs to promote the formation of hypomethylated valleys (DMVs). Failure to maintain such DMVs comes in harm's way for proper differentiation (Dixon, *et al* 2021 Science). **D.** Multilayered-epigenetic repression of germline genes in differentiating cells toward somatic lineages. (*a*) First, repressive histone marks are deposited around the promoter region of germline genes. (*b*) During differentiation, DNA methylation is then deposited at the promoters as a final "locker" for transcriptional repression.

a long time as “selfish genes” or “junk DNA”, their impact on the genome is considerable, as they can be both detrimental or beneficial for development and evolution **(Figure 5A-B)**. TEs can act as binding platforms for transcription factors, like OCT4 or SOX2 *(Bourque et al., 2008)*. Genome-wide analysis in mouse and human cells has revealed that 20% of TF-binding events are found in TE sequences *(Sundaram et al., 2014)* **(Figure 5C)**. Hence, TEs can be co-opted as cis-regulatory elements for cell-type specific expression of gene networks, like LTR5H elements in human germ cells, and RLTR13D6 elements in mouse embryonic stem cells *(Todd et al., 2019; Xiang et al., 2022)*. Accordingly, expression of specific classes of TEs can be detected at specific stages of development. Examples include elements like MERVL (Mouse endogenous retroviruses type L) in the zygote, Intracisternal A Particles (IAP) family in naïve embryonic stem cells and LINEs in early differentiated pluripotent cells *(Hackett et al., 2017; Kigami et al., 2003; Peaston et al., 2004)*. Other beneficial aspects of TEs lie in the proteins they encode and their co-option for the emergence of new functions for host organisms, such as the RAG1 protein that controls VDJ recombination in vertebrates *(Kapitonov and Jurka, 2005)*, the telomerase activity for chromosome telomere maintenance in *Drosophila* *(Biessmann et al., 1990)*, and the SYNCYTINS involved in placental formation in mammalian species *(Dupressoir et al., 2005)*.

Despite these effects, TEs also represent potential threats for the genome that need to be controlled. Most strikingly, during development, improper repression of TEs is embryonic lethal, as shown in experiments with forced expression of L1 elements *(Jachowicz et al., 2017)* or deletions of genes encoding TE repressors, such as *Setdb1*, *Trim28* (*Kap1*) or *Dnmts* *(En et al., 1993; Matsui et al., 2010)*. Moreover, TE *de novo* insertions can also be linked to cancer formation when located in a tumor-suppressor gene **(Figure 5D)**. As an example, several Alu and LINE insertions were reported in the breast cancer genes *BRCA1/2* as well as in leukemia, skin and colon cancers *(Chénais, 2013; Lee et al., 2012)*.

Coincidently, plants, fungi and vertebrates demonstrate high levels of DNA methylation at transposable elements. But the importance of DNA methylation-mediated repression varies according to classes of TEs and developmental stages. For example, loss of DNA methylation in somatic cells leads to aberrant expression of

IAP elements *(Dunican et al., 2013; Hutnick et al., 2010; Jackson-Grusby et al., 2001)*. On the other hand, in hypomethylated ESCs, whilst DNMTs are addressed to IAPs via a diverse family of specific KRAB (Krüppel-associated box domain) zinc finger proteins interacting with KAP1, compensatory mechanisms like repressive H3K9me3 marks—also recruited to TEs via KAP1– can compensate for DNA methylation loss *(Senft and Macfarlan, 2021)* **(Figure 5E)**. Moreover, Muroidea rodents recently acquired a new *de novo* DNA methyltransferase, DNMT3C that is expressed in male fetal germ cells. DNMT3C selectively represses the promoters of young LINE 1 transposable elements during spermatogenesis *(Barau et al., 2016)*, as *Dnmt3C* mutant mice show sterility phenotype with hypogonadism and azoospermia, and strong reactivation of young L1 transposable elements. Finally, DNA methylation could also temper TE deleterious effects in a more indirect manner than directed repression, with DNA methylation-associated deamination leading to gradual genetic inactivation of TEs on the long run.

### 2.2.3 DNA methylation in gene imprinting and X chromosome inactivation

**X chromosome inactivation (XCI): taming sex chromosomes**

Mammalian females carry two X chromosomes, while males only possess one. To avoid imbalanced dosage of X-linked gene expression, several hypotheses were proposed, like in 1967 when Dr. Susumu Ohno postulated that X-linked genes were to be two-fold up regulated in males *(Ohno, 1967)*. However, in 1961, after observing X-linked mosaicism in female mice that affected coat color, Dr. Mary Lyon proposed a mechanism of random X chromosome inactivation (XCI) for the silencing of one chromosome during early embryogenesis in XX females *(Lyon, 1961)*. Validation of Mary Lyon's theory became evident with the discovery of a long non-coding RNA (ncRNA) called X-inactive specific transcript (*Xist*), in mice and humans, that is expressed from one of the two X chromosomes in females and induces XCI in *cis* by coating the entire chromosome *(Brockdorff et al., 1991; Brown C et al., 1991)*. DNA methylation is then deposited by DNMT3B at later stages of the inactivation process, notably around the CGI promoters of X-linked genes, ensuring a stable lock for proper silencing *(Gendrel et al., 2012)*. Although dispensable at the initial stages of XCI, DNA methylation is critical for its maintenance, as embryos lacking DNMT1 demonstrate

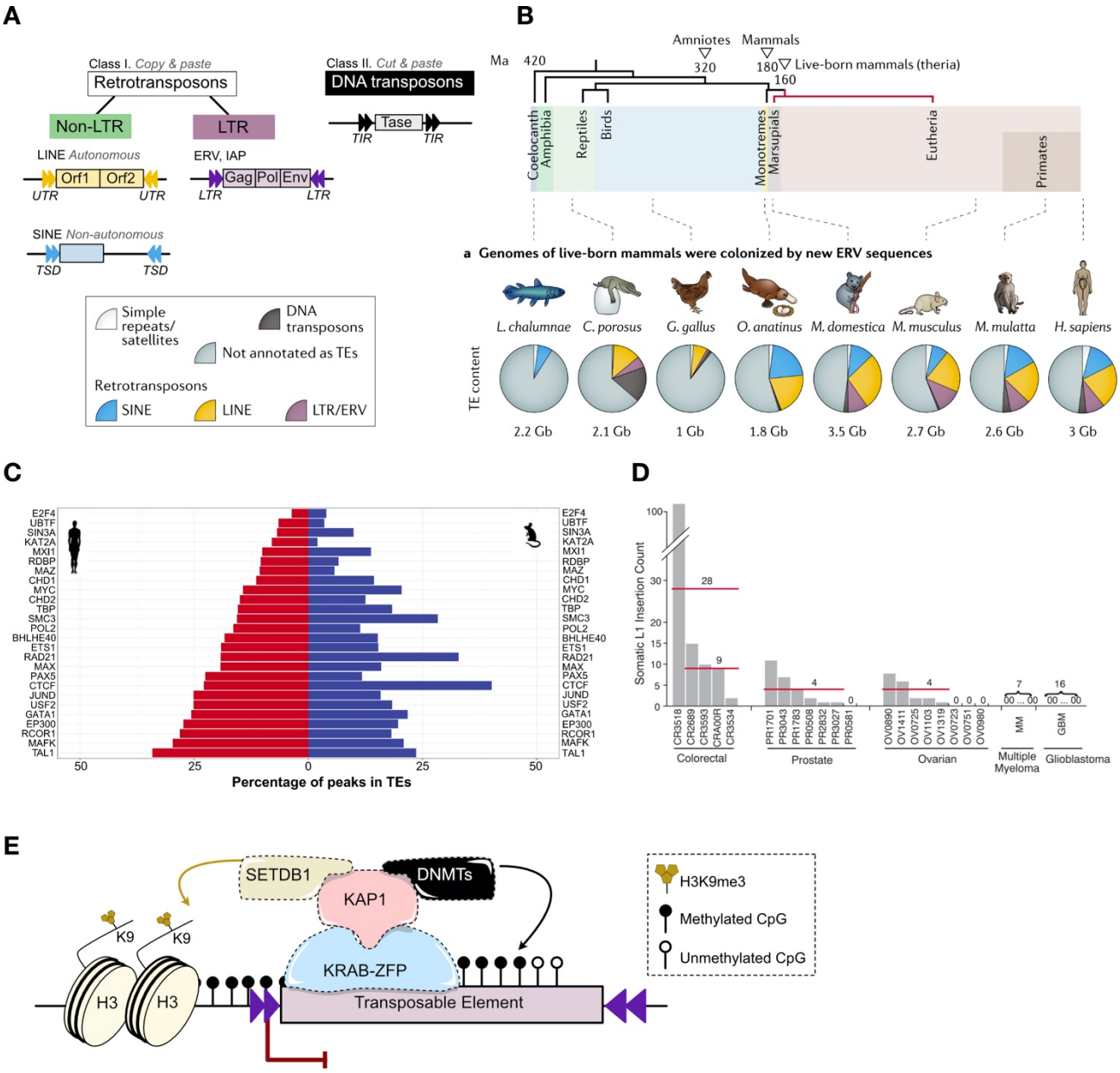



**Fig. 5 I Transposable elements in mammals, their impact and repression**

**A.** Classes of transposable elements (Tase= Transposase, TIR= Tandem inverted repeats, UTR= Untranslated region, TSD= Tandem site duplication, LTR= Long terminal repeat, ERV= Endogenous retroviruses, IAP= Intracisternal A-particule). **B.** Phylogeny of transposable elements in vertebrate genomes, pie charts represent the percentage of the genome (size indicated below in Gb) associated to transposable elements of each main family (Senft and Macfarlan, *et al* 2021 Nat. Rev. Genetics). **C.** Barplots display the percentage of peaks for transcription factors found in a transposable element sequence in humans (red) and mice (blue) (Sundaram, *et al* 2014 Genome research). **D.** Number of somatic L1 insertions found in different Colorectal, Prostatic, Ovarian, Myeloma and Glioblastoma tumors in humans (Lee, *et al* 2012 Science). **E.** Directed repression of transposable elements through KRAB Zinc-fingers proteins, acting as a platform to recruit KAP1 and ensuring multilayered epigenetic repression by DNA methylation and repressive H3K9me3 marks.

reactivation of genes of the inactive X chromosome in embryonic lineages. Accordingly, recent work in an *in vitro* model of mouse neural stem cells has associated 5-aza treatment (an inhibitor of DNMT1) with reactivation of the inactive X chromosome and its associated genes *(Mira-Bontenbal et al., 2022)*. However, the reliance of XCI on DNA methylation is also context-dependent. Indeed, in extra-embryonic tissues, maternally-inherited repressive histones marks deposited by the Polycomb group of proteins (PRC1 and PRC2) are required for paternal X chromosome inactivation **(Figure 6A)** *(Sado et al., 2000; Wang et al., 2001)*.

**Genomic imprinted genes: parentally inherited DNA methylation**

In 1984, groundbreaking work from Surani and Solter demonstrated the non-equivalent contribution of the paternal and maternal genomes for embryonic growth using nuclear transfer in the mouse **(Figure 6B)** *(Barton et al., 1984; Surani et al., 1984)*. Screening the mouse genome for regions with unequal parental effects on phenotypes allowed the first mapping of imprinted chromosomal regions *(Cattanach and Kirk, 1985)*. Since then, over 100 imprinted genes have been identified in mammals *(Bartolomei and Ferguson-Smith, 2011)*, characterized by mono-allelic parent-of-origin expression. This peculiar mode of parental regulation comes from the inheritance of gametic DNA methylation marks established at so-called imprinting control regions (ICRs), which have the extraordinary ability to resist the epigenetic reprogramming that occurs post-fertilization, during early mammalian embryonic development. Most imprinting is driven by maternally inherited ICRs established at CGIs in the oocyte; there are only 3 paternal ICRs in the mouse genome, established during spermatogenesis at intragenic CpG-poor regions *(Schulz et al., 2010)*. The persistence of ICRs during epigenetic reprogramming comes from the presence of motifs for methylation-sensitive KRAB zing finger proteins, like ZFP57 and ZFP445, that allow for the recruitment of the KAP1/DNMT complex and prevent DNA methylation loss **(Figure 6C)** *(Li et al., 2008; Messerschmidt et al., 2012; Quenneville et al., 2011; Takahashi et al., 2019)*.

The regulation of imprints can be lifelong or only transient during development. Screening strategy developed in our lab revealed dozens of potential oocyte-inherited transient imprinted loci in the mouse early embryo, that will lose their imprint in a

tissue-specific manner *(Proudhon et al., 2012)*. One of these transiently imprinted loci is the *Grp1/Zdbf2* locus. Inherited DNA methylation from the oocyte drives paternal allele-specific expression of a long-isoform of *Zdfb2* (*Liz*) in the pre-implantation blastocyst, while the *Zdbf2* promoter is kept repressed by H3K27me3 marks. However, at implantation, *Liz* expression from the paternal allele allows for DNA methylation deposition upstream of *Zdbf2* promoter, alleviating H3K27me3 repression and establishing a lifelong imprint for paternal-specific expression of *Zdbf2*. This transient imprinting is also tissue specific, as in extra-embryonic lineages, *Liz* expression is maintained *(Duffié et al., 2014; Greenberg et al., 2017)*.

Interestingly, the methylation of maternal ICRs in the oocyte is often driven by transposable elements. Indeed, it was found that in mouse oocytes—but also in other species like humans and rats– expression LTR-type transposable elements, termed LIT (LTR-initiated transcriptional units), promote DNA methylation of CGIs located downstream, on their transcriptional path. The expression of such LITs occurs in oocytes specifically, promoting the formation of maternal imprinted loci. For example, *Cdh15* and *AK0080* imprints, that were also identified from the screen conducted in our lab, seem to belong to this class of LIT elements *(Bogutz et al., 2019; Brind'Amour et al., 2018)*.

## 2.3 DNA methylation at the cross-road of chromatin pathways

The folding of DNA in eukaryotic nuclei represents a major challenge for DNA-based processes. As an illustration, human cells need to compact a two-meter-long DNA double-helix into a nucleus of 6µm of diameter. This requires a higher-order organization for the DNA, relying on architectural proteins. The most notable of them are the histones, which associate with the DNA in a nucleoprotein complex, forming the chromatin.

For a long time, there has been debates in the scientific community to know which of the histones or the DNA were the true carriers of the genetic information. Indeed, early observations in the sperm of different fish species revealed that each species bears its own specific type of nuclear proteins, while the DNA composition remains (of course) uniform. Thus, it was argued that "**protein carries the promise**

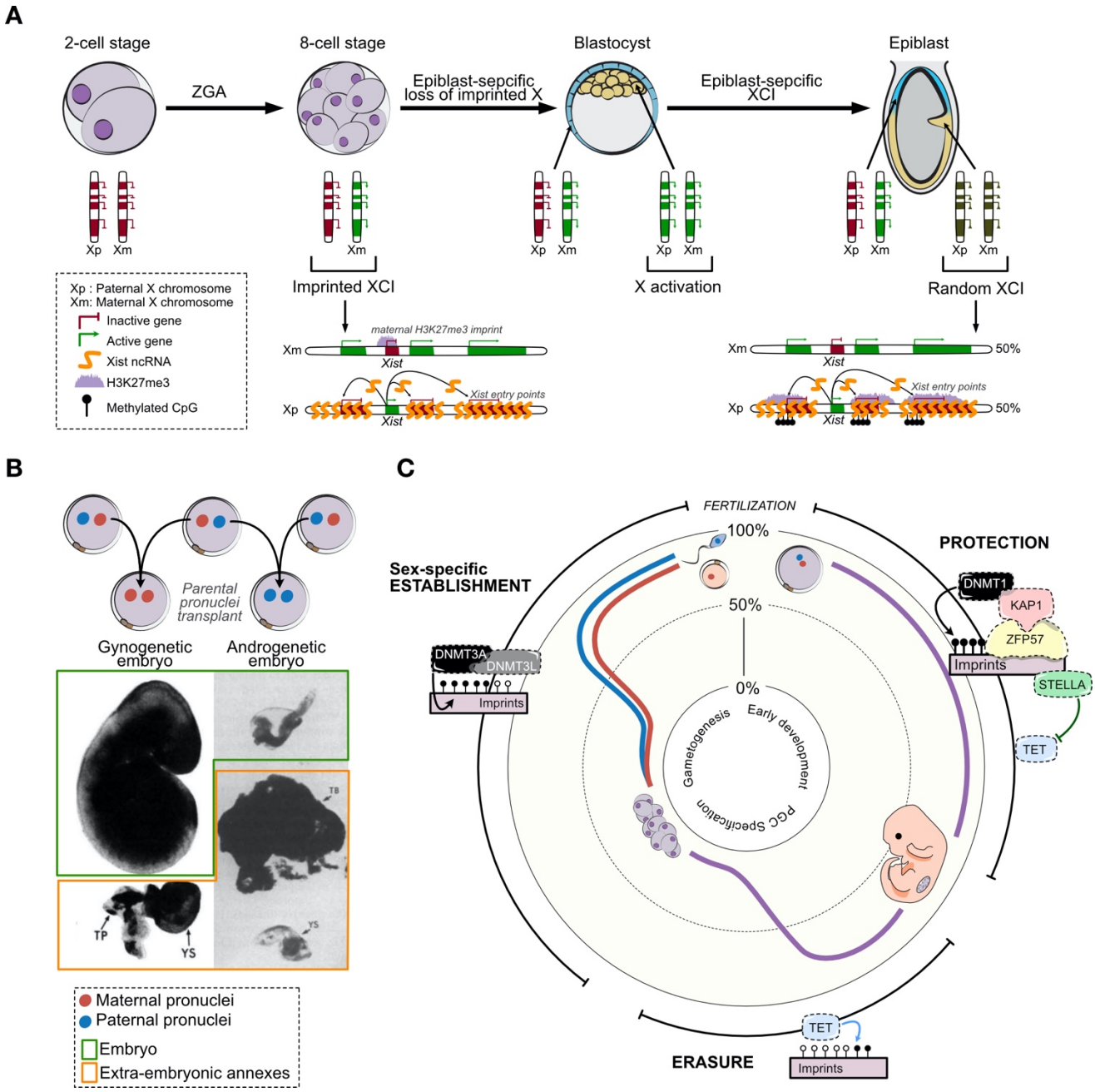



**Fig. 6 I DNA methylation and sexual dimorphism: X chromosome inactivation and imprints**

**A.** X Chromosome inactivation process in early mice development. A maternal-inherited H3K27me3 imprint on *Xist* ensures paternal-only X chromosome inactivation at the 8-cell stage, which will be maintained in extra-embryonic tissues. In the epiblast, the two X become active, and random inactivation will take place in embryonic lineages, with DNA methylation being deposited on the CGI promoters of inactive X-linked genes to ensure a final "lock" of epigenetic repression. **B.** Discovery of the non-equivalent contribution of maternal and paternal genomes by Surani and Barton in 1984. By generating bi-parental embryos, they demonstrated that the paternal genome favors the development of the extra-embryonic annexes (in orange), while the maternal genome promotes the formation of the embryonic tissues (in green), underlying the foundation for the "parental-conflict" theory (Barton *et al,* 1984 Nature and Surani *et al,* 1984 Nature). **C.** The imprinting life-cycle in the germline. Imprints are marked by DNA methylation and established via the DNMT3A-3L complex in a sex-specific manner during gametogenesis. After fertilization, imprints are protected from DNA demethylation by the recruitment of DNMT1 via KAP1-ZFP57 and inhibition of TET activity by STELLA. However, during the arising of germline precursors, imprints are erased via a combination of passive and active TET-mediated DNA demethylation, to allow subsequent sex-specific establishment, thus closing the cycle.

**that there may be the specificities that the geneticist desires within them, [while] the function of the nucleic acids is to "protect" the active side chains of the protein molecule.**" *(Kossel, 1884)*. However, the resolution of the structure of the DNA double-helix by Franklin, Crick and Watson in 1953 provided the molecular framework sustaining the model of a DNA-based carrier of genetic information *(Watson et al., 1953)*. Despite being ruled out as the carriers of genetic information, it has now become evident that post-translational modifications (PTMs) of histones constitute a template for epigenetic information and have an unambiguous impact on the regulation of gene expression. In the next parts, I will describe the structure of the chromatin and introduce histone PTMs, their impact on gene regulation and association with DNA methylation.

### 2.3.1 Nucleosome structure and histone modifications

In eukaryotes, DNA is wrapped around an octamer composed of two copies of each histone variant (H2A, H2B, H3 and H4), and this association is further stabilized by the linker histone H1. This nucleoprotein complex (DNA-histone octamer-H1) is called the "nucleosome" and is repeated every 200±40 base pairs (bp). It forms a "*bead-on-a-string*" structure, of around 10 nm in diameter, that can be furthermore compacted into higher order structures, like the 30 nm chromatin fiber, or the chromosomes, resulting in an overall compaction of the DNA by a factor of 30-40 **(Figure 7A-7B)** *(Luger et al., 2012, 1997)*. However, this 30 nm chromatin fiber structure remains difficult to observe *in vivo*, and its formation appears to be dependent on many different parameters like salt concentration, length of DNA and histone PTMs, for example *(Hansen, 2002; Tremethick, 2007)*. As a consequence, some studies suggest that the 10 nm chromatin fiber represents the "natural" state of the chromatin and that higher degrees of compaction are achieved by the compaction of smaller dynamic units of 2-10 nucleosomes *(Krietenstein and Rando, 2020; Maeshima et al., 2014)*.

The compaction state of the chromatin and its impact on the activity of the genome has been studied for a long time. In 1928, after observing that some chromosomal domains remained in a compacted state after mitosis, German geneticist Emil Heitz referred to these regions as "*heterochromatin*" and suggested that they were in an inactive state compared to the rest of the genome. This is in

contrast with "*euchromatin*" regions that appear more loose in the interphasic cell nucleus **(Figure 7C)** *(Brown, 1966)*. Indeed, it is now well-known that heterochromatic regions are transcriptionally silent and gene poor, and are further divided into constitutive heterochromatin—around centromeres and telomeres—and facultative heterochromatin—regions that can be in different states of condensation depending on the cell type. On the other hand, euchromatin contains actively transcribed gene-rich domains *(Allis and Jenuwein, 2016)*.

Interestingly, whilst the vast majority of proteins regulating the eukaryotic chromatin landscape are also found in prokaryotes, octameric nucleosomes are found exclusively in archaea and eukaryotes. The origin of this discrepancy is still strongly debated, some arguing that eukaryotic nucleosomes could have been inherited from lateral gene transfer from archaea or giant viruses in the last eukaryotic common ancestor (LECA) *(Erives, 2017; Talbert et al., 2019)*. Nevertheless, the integration of nucleosomes into the eukaryotic genome shifted the fundamental logic of transcription regulation compared to prokaryotes. Indeed, as the latter are depleted of chromatin, there is no evident structural restriction for the RNA polymerase to bind promoters, whereas nucleosomes can act as a barrier for transcription factors or even polymerases in eukaryotes *(Struhl, 1999)*. Moreover, post-translational modifications (PTMs) of the histone tails can also directly affect the transcriptional machinery, allowing the promoters to exist in intermediate states, beyond the "*black and white*" silenced and active states found in prokaryotes.

The most well-known and studied PTMs are found on the N-terminal tails of histones H3 and H4. This might be due to the fact that in non-dividing cells, the turn-over of H3 and H4 is lower than the one observed for H2A and H2AB, which display less PTMs *(Jamai et al., 2007)*. This is consistent with the lesser evolutionary conservation found for H2B proteins in eukaryotes *(Malik and Henikoff, 2003)*. Despite this, histone tails have critical functions for organisms, as their removal in yeast leads to lethality *(Ling et al., 1996)*, as they represent around 28% of the histone DNA sequence and as they act as a platform for protein interactions. They are also the substrate of different PTMs that are dynamically regulated by chromatin modifiers and protein domains among which we can find "writers" that deposit/remove PTMs, and

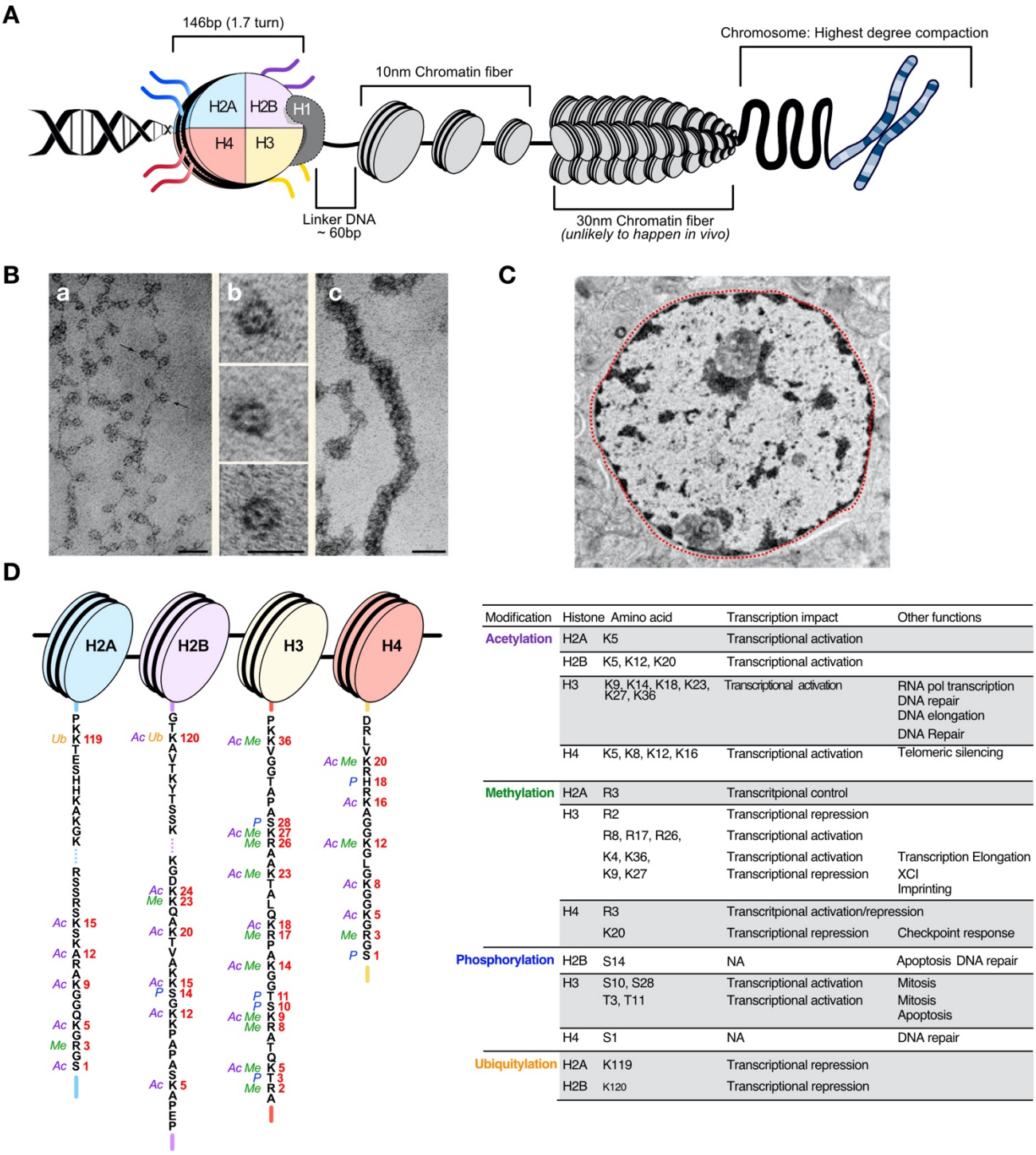



**Fig. 7 I Chromatin structure and histone PTMs**

**A.** Schematic representation of the increasingly higher order of compaction of the DNA. First wrapped around nucleosomes every 200bp in average, the DNA is further condensed into chromatin fibers of 10nm and 30nm of diameter. Chromosomes formed during mitosis or meiosis represent the highest compaction form of the DNA. **B.** (*a*) Ionic spread of nucleosomes "bead on a string", size marker 30nm, (*b*) isolated mononucleosomes, size marker 10nm, (*c*) Chromatin fiber, size marker 50nm (Olins, *et al* 2003 Nat Rev Mol Cell biology). **C.** Electron microscopy pictures of domains of heterochromatin (dark regions) and euchromatin (lighter regions) inside the nucleus of the cell (circled in red). From "medcell.med.yale.edu". **D.** Non-exhaustive list of known histone modifications with their associated function on transcription and biological processes. Adapted from Abcam, Cell signaling and Servier Medical Art.

"readers" that can recognize a specific PTM and can regulate gene expression **(Figure 7D)** *(Luger and Richmond, 1998)*.

Early work by Vincent Allfrey led in 1964 proposed that acetylation of lysine residues of histone tails is associated with active transcription *(Allfrey et al., 1964)*. Mechanistically, histone acetylation is a very plastic mark and is controlled by many histone acetyltransferases (HATs) and deacetylases (HDACs) *(Fioravanti et al., 2020)*. The deposition of acetylation promotes nucleosome unfolding and chromatin de-condensation by antagonizing the positively charged lysine, which decreases the interaction between histone tails and the negatively charged DNA *(Hong et al., 1993; Norton et al., 1989)*. Such chromatin "*relaxation*" allows the recruitment of the transcription machinery and thus, histone acetylation strongly correlates with a "relaxed" chromatin state and gene activity *(Lee et al., 1993)*.

Lysine methylation is another common PTM found on histone tails. In contrast to histone acetylation, methylation of distinct lysine residues can be associated with either gene activity or gene repression *(Janssen and Lorincz, 2021)*. No impact on the histone tail charge is known for lysine methylation and thus, it is thought that these modifications do not have a direct effect on the chromatin compaction. Nonetheless, lysine methylation can be recognized by chromatin readers and other proteins bearing specific domains. Among them, we briefly mentioned the PWWP domain of DNMT3B that is attracted to H3K36me3 *(Baubec et al., 2015)*.

### 2.3.2 DNA methylation in association with repressive chromatin marks

**H3K9 methylation:**

The methylation of the ninth lysine residue of H3 can occur at mono-, di- or tri-methylated state. Among the Lysine methyltransferases (KMTs) responsible for the deposition of H3K9me, we can find G9A and GLP that catalyze H3K9me1/me2, SUV39H1/H2 that tri-methylates residues marked by H3K9me1, and ESET/SETDB1 that deposits de novo H3K9me3. Overall, H3K9 methylation is usually associated with the formation of constitutive heterochromatin at gene-poor regions, most notably at pericentromeric regions and on telomeres *(Nicetto and Zaret, 2019)*. Compaction of chromatin through H3K9me is mediated by the recruitment of the architectural protein HP1 that recognizes methylated H3K9 through its chromodomain *(Nielsen et al.,*

*2002)*. The exact mechanism by which HP1 and H3K9 methylation promote condensation remains unresolved, but it is known that HP1 can act as a platform to recruit architectural chromatin proteins like BRG1 that modulates heterochromatin structures *(Nielsen, 2002)*. Recent studies predicted that HP1 could generate phase-separated droplets, hence promoting a model of heterochromatin formation through liquid-phase separation *(Larson et al., 2017; Strom et al., 2017)***.** In addition, H3K9 methylation is also recruited around transposable elements and imprinted genes, thanks to a variety of specific KRAB-zinc finger proteins that can recruit KAP1, as mentioned previously *(Almeida et al., 2022; Matsui et al., 2010)*. All in all, the general view is that the methylation of H3K9 is strongly correlated with repression and silencing.

DNA and H3K9 methylation have complex relationships. Indeed, H3K9me3/2 show strong positive correlation with DNA methylation in the context of early embryonic development. Methylation of H3K9me2/3 can promote DNA methylation maintenance **(Figure 8A)**, through UHRF1 binding and subsequent enhancement of DNMT1 activity *(Liu et al., 2013; Rothbart et al., 2012)*. Accordingly, loss of H3K9me2 in ESCs is accompanied by DNA methylation dissolution due to impaired targeting of UHRF1 at the replication fork (von Meyenn et al., 2016b). In contrast, loss of DNMTs does not appear to perturb the deposition of H3K9me3 at pericentromeric regions *(Tsumura et al., 2006)*. Finally, H3K9me can also act as a protector against DNA demethylation. In early embryos, the presence of DPPA3 (or STELLA/PGC7) on H3K9me2-marked regions protects the maternal genome from TET3-induced removal of DNA methylation *(Nakamura et al., 2012)*.

In addition to enhancing DNA methylation maintenance, H3K9me can also interact with the *de novo* DNA methylation machinery. For example, mouse ESCs with *G9a*-KO or treated with a G9A/GLP inhibitor demonstrate hypomethylation at gene promoters and imprinted genes, as G9A/GLP have been found to recruit DNMT3s through their ankyrin domains *(Epsztejn-Litman et al., 2008; Zhang et al., 2016)*. Loss of DNA methylation in single *G9a* mutants is most evident around germline genes and TEs *(Dong et al., 2008)*. Then, upon deletion of KAP1, and alike to what has been observed with G9A/GLP KO, TEs and imprinted loci are reactivated and lose both DNA and H3K9 methylation in mouse ESCs *(Rowe et al., 2010)*.

Interestingly, at periods of development associated with a DNA hypomethylated genome, like in ESCs, H3K9me3 can partially compensate for DNA methylation loss, and participate in maintaining epigenetic repression of TEs *(Walter et al., 2016)*. Moreover, it was proposed that H3K9me3 accumulated at the promoters of germline genes to induce their repression in the early DNA hypomethylated embryo, prior to DNA methylation establishment during epiblast differentiation *(Mochizuki et al., 2021)*. In contrast, during the methylation phase of the male germline from embryonic days E13.5 to E16.5, H3K9 and DNA methylation seem to display a more antagonistic relationship: domains marked by H3K9me3 tend to acquire DNA methylation with a delayed kinetics *(Shirane et al., 2020)*.

Taken together, it appears that H3K9 methylation is especially critical for the maintenance of DNA methylation, at imprinted, germline genes and on TEs **(Figure 4D -5E)** and to act in synergy with DNA methylation to ensure their stable repression. But these central effects seem to be restricted to developmental time windows whereby DNA methylation is especially dynamic or even absent. In contrast, loss of H3K9 methyltransferases in differentiated cells, where DNA methylation patterns are established and stable, does not result in DNA hypomethylation or re-activation of TEs.

**H3K27 methylation:**

The methylation of H3K27 is mediated by the Polycomb group of proteins (PcG), a complex of proteins that was first identified in a *Drosophila* mutant for body segmentation *(Lewis, 1978)*. The high conservation of the PcG complex, from unicellular eukaryotes to plants and animals, suggests an ancient origin that might have appeared in LECA *(Talbert et al., 2019)*. In mammals, the Polycomb repressive complex 2 (PRC2) is responsible for methylating H3K27. The PRC2 methyltransferase ability is found in the SET domain of EZH2 and EZH1, the latter having a lower catalytic activity than the former *(Margueron et al., 2008)*. On the side note, the PRC1 complex catalyzes mono-ubiquitination of H2AK119, and can recruit PRC2 and promote H3K27me3 deposition, as demonstrated in mouse ESCs *(Blackledge et al., 2014)*. Similarly to H3K9me3, H3K27me3 is associated with gene repression and (facultative) heterochromatin *(Trojer and Reinberg, 2007)*. In particular, H3K27me3 is involved in a

form of non-canonical genomic imprinting, as it can be inherited by the maternal to the embryonic genome. A classic example of such H3K27me3-mediated imprints is the early imprinted X chromosome in mice, which we previously described **(Figure 6A)**. However, the conservation of such non-canonical imprints remains debated and warrants further research. For example, after fertilization, maternal patterns of H3K27me3 are erased in humans and there is to date no evidence for early imprinted XCI, in contrast to mice. Yet, other studies have observed a re-establishment of H3K27me3 maternal domains at the morula stage in human embryos, with asymmetries in the two parental alleles *(Lu et al., 2021; Xia et al., 2019; Zhang et al., 2019)*.

During early embryonic development, and in cultured ESCs in particular, domains of H3K27me3 can be found in bivalent patterns with H3K4me3 around developmental genes *(Bernstein et al., 2006)*. Interestingly, H3K27me3 demonstrates strong antagonistic relationship with DNA methylation **(Figure 8B-8E)**, notably at unmethylated CpG islands *(Ku et al., 2008; Lynch et al., 2012; Tanay et al., 2007)*. Such antagonism at Polycomb domains, including CGIs, is most visible in cancer and in differentiated cells where epigenetic switches can occur between DNA and H3K27 methylation at CGIs, antagonizing one another *(Mohn et al., 2008; Zhang et al., 2008)*. In mouse ESCs as well, it was observed that loss of DNA methylation at highly methylated CpGs promotes gain of H3K27me3, while a concomitant loss of H3K27me3 can be observed at CpG sites with usually lower DNA methylation levels *(Brinkman et al., 2012; Hagarman et al., 2013; Reddington et al., 2013)*. Yet, in other developmental contexts, like in prenatal prospermatogonia, DNA methylation loss, through DNMT3A or DNMT3L-KO, does not affect H3K27me3 levels *(Dura et al., 2022; Shirane et al., 2020)*. Finally, it appears that in some contexts, H3K27me3 and DNA methylation can co-exist. Indeed, outside from CGIs and gene promoters in differentiated cells, large H3K27me3 domains can be found in DNA methylated regions *(Hawkins et al., 2010; Pauler et al., 2009)*. Another example of overlap between PRC2/H3K27me3 and DNA methylation worth mentioning can be found at CGIs in extra-embryonic tissues specifically *(Smith et al., 2017)*.

Functionally, *in vitro* molecular assays monitoring for the activity of PRC2 when incubated with hypomethylated or methylated DNA template, revealed that DNA

methylation presence did not prevent the deposition of H3K27me3 by PRC2 *(Cooper et al., 2014)*. This suggests that the methylation of cytosines is not an inhibitor of PRC2 activity. Conversely, mutations of the PWWP domain of DNMT3A allow for deposition of DNA methylation at CGIs marked with H3K27me3. It has been suggested that this mutant could allow such recruitment to H3K27me3-marked domains by recognizing H2AK119Ub deposited by PRC1 *(Sendžikaitė et al., 2019; Weinberg et al., 2021)*. Recently, is was shown that the N-terminus domain of the long isoform of DNMT3A (DNTM3A1) is also involved in the recognition of H2AUb and recruitment of DNMT3A1 at H3K27me3 sites *(Gu et al., 2022)*. Moreover, it has been shown that the PRC2 complex can directly recruit regulators of DNA methylation, notably the TET enzymes, and sequester DNMT3L, which out-competes the *de novo* activity of DNMT3A and DNMT3B at Polycomb domains *(Neri et al., 2013b, 2013a)*. Finally, protectors of PRC2/H3K27me3-marked regions also exist to prevent the deposition of DNA methylation. Among them, KDM2B (also known as FBXL10) can recognize unmethylated CpGs through its CxxC domain and prevents DNA methylation, as its loss results in hypermethylation at two thirds of PRC2-targeted CGIs in mouse ESCs *(Blackledge et al., 2014; Boulard et al., 2015)*. In addition, the presence of H3K4me3 at H3K4me3-H3K27me3 bivalent genes repels the *de novo* DNMT3s, as this mark is sensed by their ADD domains. The strong antagonism between H3K4me3 and DNA methylation will be developed in more details later on, but provides another extra layer of antagonism between H3K27me3 and DNA methylation at developmental genes. In the end, it appears that the antagonism between H3K27me3 and DNA methylation is only moderately dependent on the methylation state of the cytosine itself, but rather relies more on a balance between the recruitment of PRC2 and the repulsion of *de novo* DNA methylation machinery at CpG-rich, H3K4me3-marked regions.

### 2.3.3 DNA methylation interplays with active chromatin marks

**H3K4 methylation:**

The deposition of H3K4 methylation in mammals is mediated by a large family of KMTs forming three major complexes that vary by the composition of their subunits: SETD1A/B, MLL1/2 and MLL3/4 *(Hughes et al., 2020)*. The latter is responsible for H3Kme1 deposition, while the formers catalyze H3K4me2/3, notably around CGI

promoters. Indeed, SETD1A is directly addressed to CGIs and the chromatin by a protein with a Zinc-finger CXXC domain—a CpG methylation sensitive domain—known as CPF1. On the other hand, MLL1 and 2 proteins bear a ZF-CxxC domain on their own *(Brown et al., 2017; Hu et al., 2017; Milne et al., 2010)*.

The accumulation H3K4 methylation at the transcription start site (TSS) of genes has been observed in mice, human and zebrafish, and associates most of the time with active genes, or at least, genes poised for further activation. In human ESCs, nearly 80% of H3K4me3-marked genes are expressed *(Zhao et al., 2007)*. With the notable exception of broad repressive domains of non-canonical H3K4me3 found in the oocyte, this mark is usually considered as an active epigenetic mark *(Liu et al., 2016)*. In embryonic stem cells, H3K4me3 is found to colocalize with H3K27me3 at the CpG dense promoters of bivalent developmental genes **(Figure 8E)**. The prevalent model stipulates that this dual marking poises genes for further activation later on, during differentiation. However, recent work suggests that the purpose of H3K4me3 mostly relies on avoiding permanent silencing of developmental genes, notably via DNA methylation, since bivalent genes are not more expressed or expressed faster than genes marked only with H3K4me3 or H3K27me3 during development *(Bernstein et al., 2006; Kumar et al., 2021)*.

As described above, H3K4me3 and DNA methylation share a strong anti-correlation, to the point where the status of H3K4 methylation is a better predictor of DNA methylation than the CpG density of underlying DNA sequences *(Meissner et al., 2008)*. As described briefly above, the ADD domain of the DNMT3s selectively interacts with unmethylated forms of H3K4, alleviating its auto-inhibitory conformation, and allowing catalytic activity of the *de novo* DNA methyltransferases to be released **(Figure 8C)** *(Ooi et al., 2007; Zhang et al., 2010)*. The repulsive effect of H3K4 methylation on DNMT3 activity is proportional to the number of methyl groups present on K4, with H3K4me3 showing a 43-fold decreased affinity for DNMT3 ADD domain compared to unmethylated H3K4. The importance of the ADD domain for the regulation of DNMT3 targeting at H3K4me regions has been furthermore highlighted in a *Dnmt*-KO mutant cell line, with ectopic expression of DNMT3A with a modified ADD domain insensitive to H3K4 methylation status *(Noh et al., 2015)*. In such context, DNA methylation is found at CGI regions marked by H3K4 methylation.

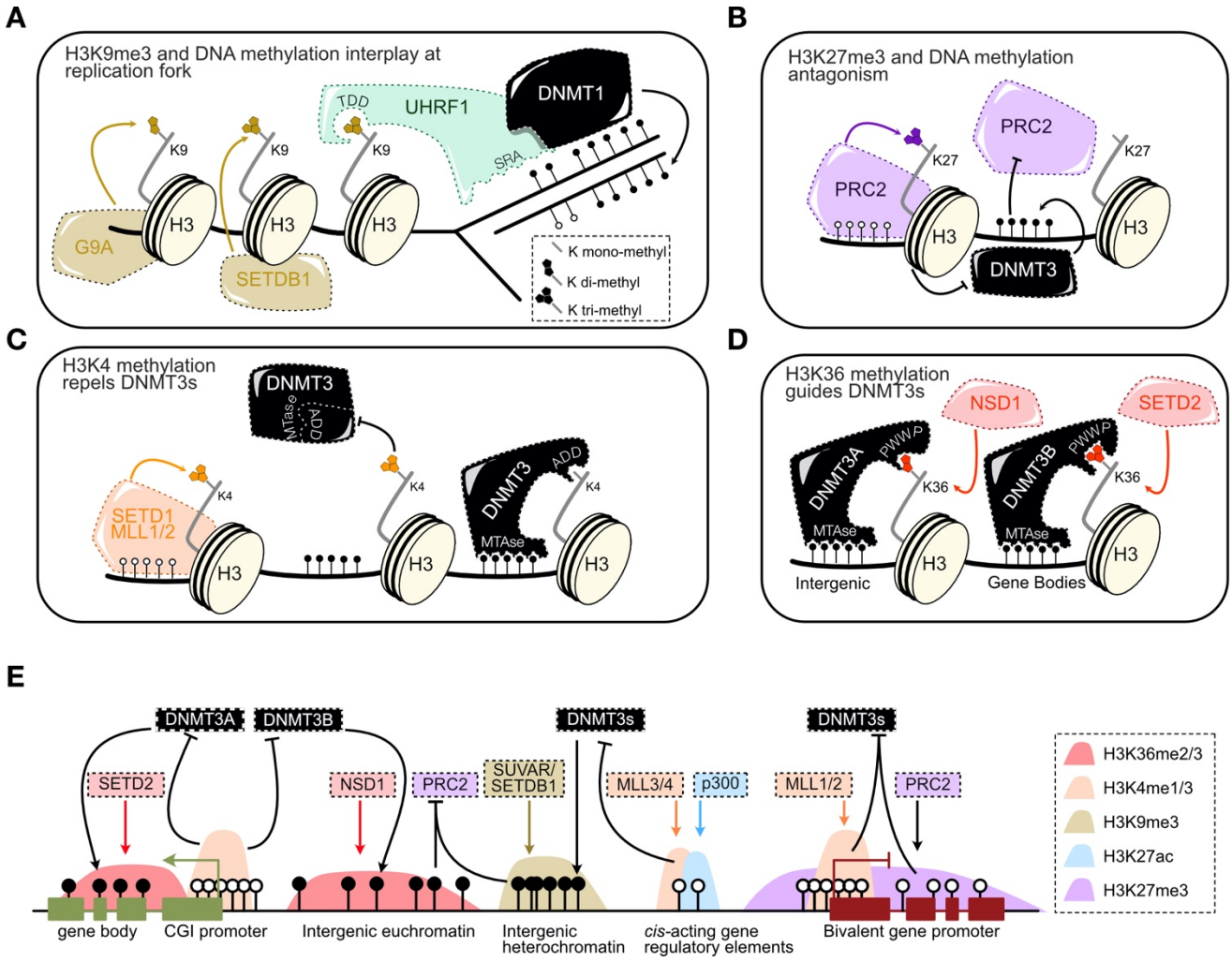



**Fig. 8 I Interplays between DNA methylation and histone marks**

**A.** Schematics of the mode of recruitment of DNMT1 by H3K9me3 at replication forks through UHRF1. **B.** Mutual antagonism between PRC2-mediated H3K27me3 and DNA methylation. **C.** Inhibition of DNMT3 activity by unmethylated H3K4 by allowing a self-inhibitory conformation of DNMT3s between the ADD and MTase domains. **D.** H3K36 methylation guidance of DNMT3A and DNMT3B at intergenic regions and gene bodies. **E.** Cross-talk between histone marks and their writers with DNA methylation and its machinery Adapted from Janssen, *et al* 2021(Nat Rev Genetics).

**H3K36 methylation:**

It is inexact to consider H3K36 methylation as a literal *'active'* mark despite its deposition made by the SETD2 at actively transcribed genes, in synergy with RNA polymerase II, from yeasts to humans *(Krogan et al., 2003; Sun et al., 2005)*. Initially, it was shown that enrichment of H3K36me3 at gene bodies is recognized by the PWWP domain of DNMT3A, which is recruited there and deposits DNA methylation *(Dhayalan et al., 2010)*. Similarly, mutation of the PWWP domain of DNMT3B induces a loss of DNA methylation at gene bodies marked with H3K36me3, similarly to a loss of SETDB2 *(Baubec et al., 2015)*. However, it appears that DNMT3A might have a preferential affinity for H3K36me2, which is deposited by NSD1 at intergenic regions *(Weinberg et al., 2019; Xu et al., 2020)*. Thus, DNA and H3K36 methylation generally demonstrate a positive correlation, with H3K36 methylation playing an important role in guiding *de novo* DNA methylation **(Figure 8D)** *(Shirane et al., 2020)*. However, the function of H3K36 methylation-based guiding of DNA methylation at intergenic regions and gene bodies remains elusive. Two major hypotheses exist. First, DNA methylation at gene bodies could facilitate transcriptional elongation and splicing events, either by preventing CTCF binding—that slows down polymerase elongation—or by promoting the binding of MeCP2 and HP1 *(Maunakea et al., 2013; Shukla et al., 2011; Yearim et al., 2015)*. Otherwise, it has been proposed that DNA methylation found at inter- and intragenic regions could repress cryptic promoters. This function has been reported to occur in a limited extent in mouse ESCs lacking DNMT3B *(Neri et al., 2017)*. However, this effect was not observed in *Dnmt*-triple KO ESCs *(Karimi et al., 2011)*. Intriguingly, a recent study uncovered a new possible non-canonical function of H3K36me3 *(Barral et al., 2022)*. In mouse ESCs, over 3000 heterochromatic regions show dual marking by H3K9me3 and H3K36me3. In these regions, the deposition of H3K36me3 seems to be mediated by NSD1, in synergy with SETDB1. Indeed, the loss of SETDB1 induces both H3K9me3 and K36me3 loss. With the loss of SETDB1, these dual H3K9me3-H3K36me3 heterochromatin regions acquire features of enhancers, suggesting that H3K36 methylation could act as a bookmark for poised enhancers, which are kept silenced by H3K9me3.

**H3K27 acetylation:**

Acetylation of histone residues has long been associated with gene activation. Indeed, hypoacetylation is observed on the inactive X chromosome of mammalian females or silent mating types in yeasts, while lysine hyperacetylation is observed in the hyperactive male X chromosome in *Drosophila* *(Bone et al., 1994; Braunstein et al., 1996; Jeppesen and Turner, 1993)*. We already described in a previous section the mechanism behind gene activity and histone acetylation. Concerning the relationship between DNA methylation and H3K27ac, it was mostly characterized at enhancers. Namely, in mouse ESCs, enhancers with higher level of DNA methylation will gain H3K27ac signal upon loss of DNMTs *(King et al., 2016)*. Similarly, in *Dnmt3A*-KO male germ cells in the mouse, ectopic H3K27ac-enriched regions are gained at regions that lost DNA methylation, coinciding with motif sites for methyl-sensitive transcription factors *(Dura et al., 2022*).Yet, the co-occurrence of both DNA methylation and H3K27ac can happen at enhancers in some cancer cells. In the HCT116 cell line, over 38% of enhancers are heavily DNA methylated, and in this model, loss of DNA methylation concomitantly reduces H3K27ac levels, suggesting that there is no inherent antagonism preventing DNA methylation and H3K27ac to co-localize *(Charlet et al., 2016)*. However, this is in sharp contrast to mouse ESCs, as the presence of H3K27ac is anti-correlated with DNA methylation variations at gene regulatory elements **(Figure 8E)**, with enhancers enriched for H3K27ac being more hypomethylated than enhancers enriched for H3K4me1 only *(Alajem et al., 2021)*. As H3K27ac deposition is mediated by p300 and the transcription machinery, the reduction of H3K27ac when DNA methylation is present has been associated with eviction of the Mediator complex and limited promoter-enhancers contacts *(Song et al., 2019)*.

# 3. A tale of early development

## 3.1 From fertilization to implantation

Fertilization refers to the fusion of the haploid maternal and paternal gametes, forming the one celled diploid zygote that marks the beginning of a new life. We generally picture development as a "*down-hill*" process, as illustrated with the Waddington landscape, in which cells progressively lose their potency (i.e., from pluripotency, to multipotency, to progenitor cells), and become more specialized. In this regard, fertilization stands out as an "*up-hill*" trajectory in the development of sexually reproducing species, as the encounter of highly specialized parental gametes will give rise to the totipotent zygote. Indeed, in mammals, primordial germ cells do not arise just after the formation of the zygote, but rather emerge from the stem cells of the epiblast that have already undergone lineage restriction *(Chen et al., 2019; Hancock et al., 2021; Mochizuki et al., 2018)*. This resetting of the development line happens in a matter of hours, and comes at the cost of significant changes at every level of the cellular organization, including morphological, metabolic, transcriptomic and epigenetic. Each of these processes will have great impact not only for the proper formation of the zygote, but also in the following first cell fate decisions.

### 3.1.1 Parental gametes and first reprogramming: prepping the embryo

In 1962, Sir J.B. Gurdon demonstrated the importance of the content of the egg in developmental reprogramming when he transplanted the nucleus of a differentiated epithelial cell of a tetrapod into an enucleated egg, which resumed normal development *(John B. Gurdon, 1962)*. Similar somatic nuclear transfer experiments were conducted in mammals, and while poorly efficient (only 16% of recovered transplanted embryos in average) *(Campbell et al., 1996)*, these groundbreaking assays demonstrated for the first time that reprogramming relied not solely on the genetic information but also on the contents found in the egg. Although both parental gametes bring extra layers of information to the future embryo beyond their halved genomes, it appears that the maternal gamete, the oocyte, is the main carrier of factors that regulate early zygotic development.

Indeed, as the sperm genome is mostly organized around protamines, histones represent only a fraction of the paternally inherited epigenome. Although exact numbers are still debated, the estimation lays around 5% of retained histones in mouse spermatozoa, and 10% in humans *(Brykczynska et al., 2010; Hammoud et al., 2009)*. However, it has been noticed that many of the conserved histones are found at the promoter of developmental genes and could have a potential impact on the viability of the offspring. Most notably, in men sperm, H3K4me3 is found in domains colocalizing with CGIs around genes associated with early embryogenesis. Hence, transient overexpression during spermatogenesis of the H3K4 demethylase KDM1A induces severe morphological defects, most notably craniofacial abnormalities, in the offspring of exposed males for three generations, providing a possible mechanism of paternal transgenerational epigenetic inheritance *(Lambrot et al., 2021; Siklenka et al., 2015)*. In addition, DNA methylation patterns and non-coding RNAs (ncRNAs) can also be transmitted from the paternal gamete to the zygote. For example, impaired DNA methylation profiles in the sperm—induced via metabolite/diet (like folate deficiency or low protein diet), ageing or stress-response—can be transmitted in the embryo and are associated with impaired regulation of embryonic development *(Lambrot et al., 2013)*, but also impaired offspring metabolism *(Carone et al., 2010)*, increased incidence of neurologic disorders *(Jenkins et al., 2014)* and fear response *(Dias and Ressler, 2014)*. Similarly, ncRNAs isolated from the sperm of males exposed to traumatic stress or dietary stress (either high-fat or protein restriction), and injected into unexposed oocytes or zygotes, induce metabolic alterations in mouse offspring *(Chen et al., 2016; Gapp et al., 2014; Sharma et al., 2016)*. Whilst these mechanisms could act as potential vector for paternal epigenetic inheritance, evidences are still sparse. Moreover, the epigenetic reprogramming that occurs after fertilization is mostly driven and controlled by maternal factors.

Then, the content of the oocyte plays a key role not only for proper early zygotic development, but also for the paternal genome reorganization since maternal factors regulate the organization of the paternal chromatin after fertilization. Notably, maternal carry-over of histones 3.3 (H3.3) will replace paternal protamines and participate in the activation of the paternal genome later in development *(Kong et al., 2018; Van Der Heijden et al., 2005)*. Another example is the maternal protein nucleoplasmin 2

(NPM2), whose deletion leads to enucleolated embryos and improper development due to impaired centromere clustering and disorganization of centromeric/pericentromeric heterochromatin *(Ogushi et al., 2017)*.

DNA methylation is extensively remodeled during the first period of embryonic development **(Figure 9A)**. Notably, active DNA demethylation via maternal TET3 targets the hypermethylated paternal genome *(Gu et al., 2011; Peat et al., 2014)*, while the oocyte methylome is protected from TET3 action via DPPA3 binding at maternal domains marked by H3K9 methylation *(Nakamura et al., 2012, 2007)*. Accordingly, reduction of DPPA3 in oocytes of obese mice is associated with precocious loss of DNA methylation in the maternal pronucleus *(Han et al., 2018)*. DNA demethylation of the maternal genome is thought to primarily occur through passive dilution, with the eviction of DNMT1 and UHRF1 from the maternal genome until the eight-cell stage *(Howell et al., 2001; Maenohara et al., 2017)*. While TET1 and TET2 play minimal role for early development *(Dawlaty et al., 2011; Li et al., 2011)*, TET3-deficient embryos demonstrate partial loss of DNA methylation on the maternal genome, suggesting that TET3 only partially mediate actine DNA demethylation in the parental genomes *(Guo et al., 2014; Shen et al., 2014)*.

In addition to DNA methylation, other chromatin marks are also reprogrammed in the early zygote **(Figure 9A)**. This reprogramming occurs mostly during the two-cell stage and is driven by maternal-inherited histone post-translational modifying enzymes *(Wu and Dean, 2020)*. Briefly, the broad repressive H3K4me3 non-canonical domains inherited from the oocyte disappear at this stage. During the two cell stage, *cis*-regulatory elements important for zygotic genome activation (ZGA) gain H3K27ac and 80% of ZGA genes conserve at their promoters canonical H3K4me3 peaks inherited from the oocyte and sperm genomes *(Dahl et al., 2016)*. Then, H3K27me3-marked gene promoters in the oocyte are also lost before ZGA *(Zheng et al., 2016)*, while maternal-inherited large intergenic domains of H3K27me3 are still present in the two-cell stage embryo and regulate *Xist* and placental development in the mouse *(Inoue et al., 2017)*. Finally, while levels of H3K9me3 are globally lost immediately after fertilization, some regions reacquire H3K9me3 at the 2-cell stage, notably maternal and paternal ICRs, some TEs previously marked by DNA methylation, and oocyte-inherited broad domains of H3K27me3 *(Shi et al., 2019; C. Wang et al., 2018).* The

purpose of such erasure of repressive epigenetic marks in the following hours after fertilization is not yet fully understood. Maybe this epigenetic reprogramming is necessary to reset the contrasting maternal and paternal epigenomes—and therefore a source of genomic imprinting and parental imbalance in gene expression—as species that do not possess ICRs, such as the zebrafish, also do not undergo global, but rather locus-specific DNA demethylation. Another possibility is that this epigenetic "relaxation" supports the transition from a silenced to a strongly active zygotic genome, rebooting the zygote with the ability to give rise to both embryonic and extra-embryonic lineages *(Potok et al., 2013; Schroeder et al., 2015; Yu Zhang et al., 2018)*.

### 3.1.2 Zygotic genome activation: the great reboot

The zygotic, or embryonic, genome activation (ZGA) still represents a black box in mammalian development, notably because of the limited amount of material available at that stage. While in zebrafish, many of the pioneer transcription factors initiating ZGA largely overlap with the pluripotency gene network, like *Nanog*, *Oct4* and *Sox2* homolog genes *(Lee et al., 2013)*, knowledge of the regulators of ZGA in mammals remains scare. So far, knock-out of maternal factors in the oocyte or the one-cell stage embryo provided the strongest evidence for identifying central regulators of mouse ZGA in the course of development. Depletion of maternal *Yap*, *Cdx2* or *Hsf1* are examples of factors whose loss impact ZGA and further induces developmental delays or arrest before implantation *(Bierkamp et al., 2010; Jedrusik et al., 2015; Yu et al., 2016)*. Alternatively, screening strategies have been deployed in cultured mouse ESCs to look for promoters of ZGA-associated genes—as a small fraction of the ESC population displays 2-cell like cell (2CLC) features—in what could be qualified as a "retro-engineering" approach. Such experiments unveiled new potential regulators for mouse ZGA, such as *Dux*, *Patz1*, *Smacar5*, *Pou2f2*, *Myc*, *Dnmt1* or *Dppa2* and *Dppa4* *(Alda-Catalinas et al., 2020; Eckersley-Maslin et al., 2019; Fu et al., 2019)*. Although these *in vitro* strategies alleviate the material limitation to search for ZGA regulators, they still need to be validated in the context of development, as the relevance of

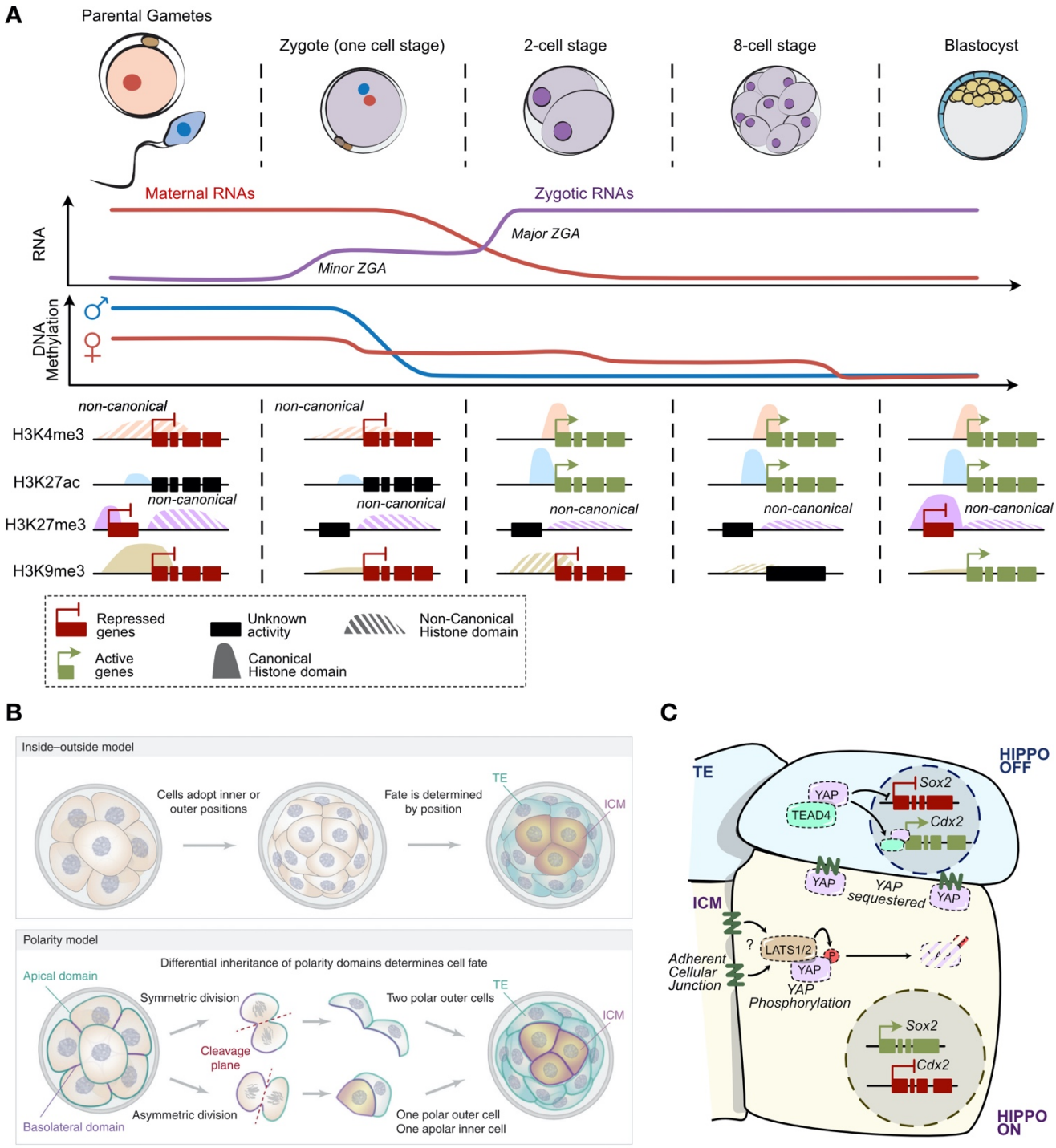



**Fig. 9 I Zygotic genome activation, embryonic epigenetic reprogramming and cell fate choice**

**A.** Graphical model of zygotic genome activation and epigenetic reprogramming after fertilization. Briefly, as maternal RNAs are degraded, zygotic transcripts are produced in waves, concomitantly with the erasure of parentally inherited DNA methylation profiles. Repressive histone marks are transiently erased before the 2-cell stage and non-canonical domains, mostly maternally inherited, are maintained until the blastocyst stage. **B.** Inside-outside model: position of the blastomere within the zygote dictates lineage allocation / Polarity model: Inheritance of apical domain govern trophectoderm fate (White *et al*, 2020 CSH Perspectives in Biology). **C.** Representation of the role of the HIPPO pathway in ICM/TE fate decisions. Blastomeres in the inner cell mass, via cell-cell junctions, activate the HIPPO pathway and prevent YAP from activating CDX2 and from inhibiting SOX2 expression, allowing for ICM specification.

*Dppa2/4* for the acquisition of 2CLC features in mouse ESCs remains debated *in vivo* *(Kubinyecz et al., 2021)*.

It is known that ZGA occurs in waves of transcription. First hints of transcription arise mostly from the paternal pronucleus during its DNA demethylation phase, defining a minor wave of ZGA during the first cell cycle *(Aoki et al., 1997)*. This is followed by the major wave of ZGA that is initiated after the first cell division, in the two-cell stage zygote *(Fu et al., 2020; Hamatani et al., 2004)*. Many models are proposed to explain the timing of ZGA in early embryos.

One of the particularities of vertebrate species is the size of their oocyte, which can often be thousands of times bigger than somatic cells. In early development, the embryo engages in cellular divisions without changing its global volume, which results in an increase of the nuclear to cytoplasmic ratio, the "N/C ratio". As such, one model for the regulation of ZGA proposed that increase in N/C ratio can titrate away repressors and repressive histone marks from the DNA and induce ZGA. Indeed, in fast developing vertebrates like newt, frogs or fishes, reducing the N/C ratio, through artificial increase of the cytoplasm or by generating haploid embryos, delays ZGA, while tetraploid embryos conversely undergo ZGA one cell cycle earlier *(Kane and Kimmel, 1993; Kobayakawa and Kubota, 1981; Masui and Wang, 1998)*. In contrast, in mice, inhibiting cell division but not DNA replication, which results in N/C ratio increase, does not affect the compaction, transcription or alkaline phosphatase activity in the early embryo *(Pratt et al., 1981)*. Therefore, while it appears that the N/C ratio is not the main regulator of ZGA in mammals, chromatin changes and transcriptional activators might play a more central role. Although investigations at the level of single gene locus are limited, they remain critical to understand how chromatin changes and transcriptional activators can promote and regulate ZGA and preset the first cell fate decisions.

Indeed, following ZGA, comes the first cell fate choice in mammalian development: the specification of the embryonic Inner Cell Mass (ICM) and the extra-embryonic trophectoderm (TE). This fate choice constitutes the first symmetry-breaking event in the early embryo. Initial views figured that the specification of these two lineages emerge randomly among blastomeres, as they appear mostly homogenous at this stage *(Motosugi et al., 2005)*. In opposition, the idea of

heterogeneity across blastomeres that could pre-determine subsequent cell fate decisions has also been enunciated long ago. For example, it is known that upon separating the first two blastomeres at the 2-cell stage, the two generated hemi-embryos contain different ratios of ICM/TE cells *(Katayama et al., 2010)*. Moreover, only 30% of sister-blastomeres display equivalent potential and were able to form a functional epiblast *(Casser et al., 2017)*. In addition, single cell transcriptomic analysis revealed that *Carm1* can be differentially expressed between sister-blastomeres, and promotes a self-reinforcing positive loop that biases one blastomere toward the adoption of a specific fate *(Shi et al., 2015).* Another factor, the ncRNA *LincGET* also demonstrates this bistable pattern, and its overexpression promotes the adoption of the ICM fate by binding to the *Carm1* locus, and by enhancing the deposition of H3R26me2 at the *Sox2* gene, a critical ICM fate regulator *(J. Wang et al., 2018)*. However, how this inter-blastomere heterogeneity can modulates developmental potentialities and the distribution of specifying factors remains to be explained. Although molecular heterogeneity could arise at the microscale within organelle compartments as early as in the oocyte, it could also involve stochastic events and morphogenetic forces generating this heterogeneity *de novo (Chen et al., 2018)*.

In fact, at the morula stage (8-cell stage), compaction of the cells bring new morphological constraints to the embryo. In 1967 was proposed the “outside-inside” or “positional” model *(Tarkowski and Wróblewska, 1967)*, which postulated that outer cells were biased to give rise to the TE, while inner cells would be specified into the ICM **(Figure 9B)**. This model was supported by experiments in which transplanting blastomeres in the most external side of the embryo would predominantly give rise to the TE *(Hillman et al., 1972)*. In contrast to this, the “polarization” or “polarity” model enunciated by Johnson in 1981 *(Johnson and Ziomek, 1981)* gave a more prominent role to the acquisition of an apical domain for future lineage allocation **(Figure 9B)**. The inheritance of the apical membrane during a symmetrical division by the two daughter cells promotes their allocation to the TE, while asymmetric divisions give rise to one apical outsider cell and one internalized cell that preferentially contribute to the ICM *(Alarcon, 2010; Plusa et al., 2005)*. These morphological cues can be translated into signaling cues that control the genetic network behind ICM and TE fates. Most notably, the HIPPO pathway and its effectors YAP and TEAD4 are sensitive to cellular

junctions **(Figure 9C)**. In the case of internalized cells, as cell-to-cell junctions are established, the HIPPO pathway is activated, and the YAP protein is phosphorylated and excluded from the nucleus, allowing for the transcription of *Sox2*, the earliest specifier of the ICM fate. In contrast, in the apical outsider cells, inactivation of the HIPPO pathway allows for YAP relocation to the nucleus where it promotes the expression of *Cdx2,* the first specifier of the TE. This likely requires the synergic action of TEAD4, as *Tead4* -/- embryos fail to express *Cdx2* *(Nishioka et al., 2008; Wicklow et al., 2014)*.

However, it has been observed that in some cases, asymmetric division from a polar blastomere can give rise to an apolar blastomere that, if not relocalized in the inside of the embryo, will remain in the outer layer with an active HIPPO pathway *(Anani et al., 2014; Korotkevich et al., 2017)*. After activation, such cells will be internalized, which suggests that there is a switch in the hierarchy between the polarity and position information in the morula. First, the polarity information is inherited from the parent blastomere depending on the symmetry of the cell division. In this case, the polarity information acts upstream of the HIPPO pathway, to direct the fate of the blastomere. Then, for outside apolar blastomeres, the positional information becomes prominent through the cellular junctions to relocalize the blastomere or induce the formation of a polar membrane *(Toyooka, 2020)*.

### 3.1.3 To implantation and beyond

The formation of the blastocyst, three days after fertilization in mice, marks the segregation between the external layer of TE cells and the ICM cells that remain internal but are now spatially constricted by the expansion of a fluidic cavity named the blastocoel.

Just prior to implantation of the blastocyst into the uterine wall, the ICM cells undergo their second fate decision that will separate the embryonic epiblast (Epi) cells from the extra-embryonic primitive endoderm (PrE) cells. The origin of the primitive endoderm and epiblast differentiation has been extensively studied. Induction of the primitive endoderm is mainly driven via FGF signaling pathways *(Chazaud et al., 2006; Kang et al., 2013)*. Indeed, the knock-out of *Fgfr1/2* leads to a fully epiblastic ICM *(Kang et al., 2017)*, while exposing the embryos to recombinant FGF4 protein leads to

a complete PrE conversion of the ICM *(Yamanaka et al., 2010)*. These pathways are in direct control of NANOG and GATA6, two critical and antagonizing specifiers of the Epi and PrE, respectively. Initially co-expressed in ICM cells, differential exposure to FGF signal initiates the reinforcement of the expression of *Nanog* or *Gata6* and further commitment towards Epi or PrE fate and the segregation of the two lineages within the ICM *(Guo et al., 2010; Plusa et al., 2008)*. More specifically, imbalanced expression between *Nanog*/*Gata6* promotes the expression of ETV and SPRY factors that inhibit the production of FGF receptors, reinforce NANOG expression, and stimulate the production of FGF4. The excretion of FGF4 exposes neighboring cells to FGF signaling, which promotes PrE fate via *Gata6* expression reinforcement *(Kang et al., 2017)*. However, the nature of the initial trigger that imbalances *Nanog* and *Gata6* in the ICM remains poorly understood **(Figure 10A)**. Some work proposed that internalized cells at the morula stage would be primed towards a PrE fate, as they already inactivated the HIPPO pathway and the SOX2 pluripotency factor. Moreover, it has also been observed that early-internalized cells express higher FGF4 levels, while late-internalized cells express higher levels of the FGF receptor *(Krupa et al., 2014; Mihajlovic et al., 2015; Morris et al., 2013)*. This is in support of a unifying model in which the HIPPO pathway, as well as spatial localization of the blastomeres in the morula, not only prefigure the specification of the ICM and TE, but also prime the segregation of the ICM into the Epiblast and Primitive endoderm in the blastocyst.

After implantation, the blastocyst develops into the “egg cylinder” from 4 days to 7 days post-fertilization, with the epiblast cells giving rise to embryonic lineages at its center. During this development, extensive cell death is observed in the epiblast. Early views proposed that the observed cell death in the egg cylinder is associated with the formation of the epiblast lumen, assumption that was supported from observed cavitation in 3D-grown embryoid bodies *in vitro* *(Coucouvanis and Martin, 1995)*. Yet, recent studies suggest that this apoptotic wave is dispensable for epiblastic morphogenic events *(Bedzhov and Zernicka-Goetz, 2014)*. In contrast some proposed that this increase in apoptotic cells during epiblast differentiation is rather associated with a negative selection that eliminates cells unfit for further lineage specification *(Sancho et al., 2013)*. The formation of the lumen and the “egg cylinder” primarily relies on the formation of a rosette-like structure from actomyosin-

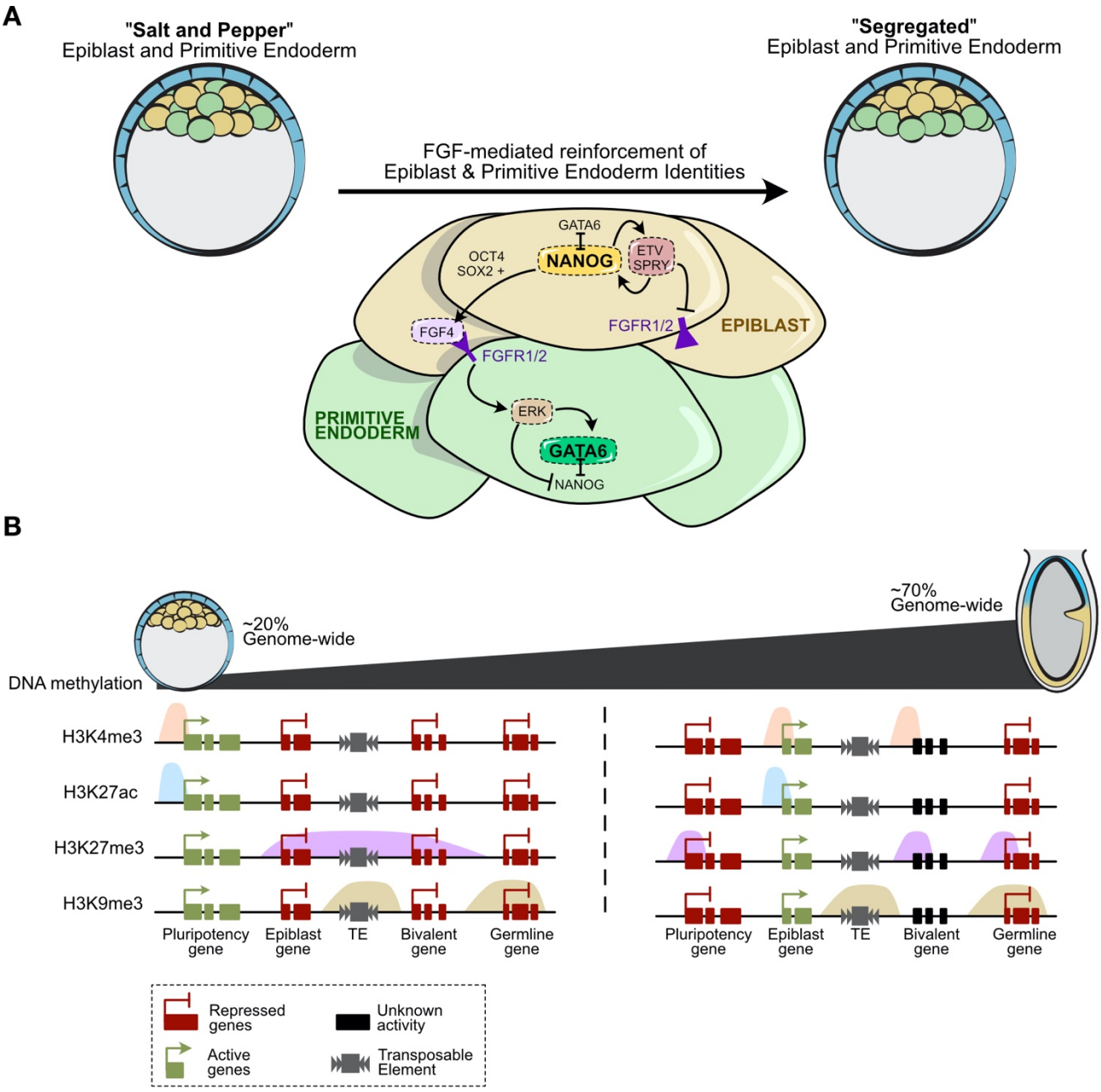



**Fig. 10 I Epiblast induction and Epigenetic reprogramming during implantation**

**A.** Graphical model of induction of the Epiblast (Epi) and the Primitive endoderm (PrE). Briefly, balance of antagonizing activity of NANOG (Epiblast specifier) and GATA6 (PrE-Specifier) is tipped by the FGF pathway. As NANOG expression is reinforced, receptors of FGF are lost, and while exposure to FGF4 emanating from Epiblast cells induces GATA6 expression in PrE cells. **B.** Epigenetic reprogramming undergone by naïve epiblast cells in the ICM until the formation of the egg-cylinder primed epiblast cells. Active histone marks, like H3K27ac and H3K4me3, are lost at pluripotency genes and are redistributed on early post-implantation genes. Moreover, H3K27me3 broad domains found in the naïve epiblast of the blastocyst are reshaped, and peaks of H3K27me3 are deposited at bivalent regions and lineage specific loci, like germline genes. H3K9me3 continues to ensure repression of transposable elements notably, which is strengthen by *de novo* DNA methylation.

dependent polarization of epiblast cells. Then the apical domains of opposed epiblast cells separate via repulsive anti-adhesive molecules, like podocalyxin *(Neagu et al., 2020)*. Interestingly, such wave of cellular death and rosette formation can also be observed with *in vitro* models of epiblast differentiation *(Turco et al., 2012)*. Indeed, the derivation of cultured ESCs from the blastocyst and of differentiated epiblast-like stem cells (EpiLC/EpiSCs) allows to recapitulate peri-implantation development in culture. Whilst we will describe later in this Introduction these *in vitro* methods, the modeling of epiblast differentiation has led to better comprehension of peri-implantation development. This has been especially true for uncovering regulators of stem cell differentiation, like the role of autocrine secretion of FGF4 and induction of the FGF/ERK pathway *(Kunath et al., 2007; Stavridis et al., 2007)*, but also for studying the regulation of pluripotency, or the epigenetic reprogramming that takes place during early development, prior to lineage specification.

The chromatin landscape found in embryonic stem cells of the ICM illustrates the occurrence of a *tabular rasa,* epigenetically speaking. The genome is mostly hypomethylated, retaining only around 20% of the parentally inherited DNA methylation patterns, notably because of the absence of DNMT3s and the sequestration of DNMT1 outside of the nucleus. This methylation is mostly present on ICRs and on young and/or active TES, notably IAPs in the mouse *(Messerschmidt et al., 2012; Wang et al., 2014)*. In addition to genomic hypomethylation, ESCs also have higher chromatin accessibility *(Argelaguet et al., 2019a)*, and demonstrate long range interactions, and more diffuse heterochromatic domains than differentiated cells *(Bonev et al., 2017; Meshorer et al., 2006).* Similarly, ESCs demonstrate high levels of active H3K27 acetylation and H3K4 methylation, especially around regulatory elements of pluripotency-associated genes *(Atlasi et al., 2019; Kurimoto et al., 2015)*. These observations led the to “open chromatin” model, which hypothesized that such openness allows for embryonic stem cells to be more plastic in their response to various differentiation cues, while supporting their pluripotency. Upon epiblast differentiation, embryonic cells progressively restrict this “open” chromatin landscape, as they initiate lineage specification. This transition from the undifferentiated pluripotent embryonic stem cells in the blastocyst to differentiated epiblast stem cells in the egg cylinder is also accompanied by a transition in their

pluripotency state. This transition starts from a ground-state, naïve pluripotency, with its own specific transcriptional program, to a primed pluripotency state where cells are biased to give rise to somatic versus germline lineages. We will cover the regulation of this priming later in this Introduction, but first, we will see how this switch in pluripotency states is concomitant with drastic chromatin changes.

First, DNA methylation is reacquired genome-wide in primed post-implantation epiblast stem cells **(Figure 10B)**. The bulk of DNA methylation establishment is primarily deposited by DNMT3B, which catalyzes DNA methylation at gene bodies, germline genes, genes of the inactive X chromosome in females, pericentromeric regions and others. DNMT3A and DNMT3L are also expressed during epiblast priming, but DNMT3A seems less active than DNMT3B and mostly methylates Polycomb targets later in the differentiation, while the lack of DNMT3L does not perturb *de novo* DNA methylation during this period *(Yagi et al., 2020)*. Also, DNA methylation is strongly acquired around regulatory elements of genes expressed in undifferentiated pluripotent embryonic stem cells, which suggests that DNA methylation might participate in their repression and coincidently, facilitates epiblast priming *(Bell et al., 2020; Kalkan et al., 2017)*. By E6.5, primed epiblast cells demonstrate around 70% of methylated CpGs in average, while extra-embryonic tissues, like the visceral endoderm (VE), only display intermediate levels of methylation (~50% of mCpGs). This is accompanied by reciprocal differences in promoter methylation, where epiblast markers and developmental genes (*Oct4, Nanog, HoxB2, Hoxb3*) are hypermethylated in the VE, while markers of VE (*Apoa1, Apoa4, Apob, Apoc2*) are hypermethylated in the epiblast. This suggests different mechanisms regulating DNA methylation and a potential role for DNA methylation in reinforcing extra-embryonic versus embryonic lineage identities. In addition, in extra-embryonic tissues, DNA methylation levels can be correlated with the establishment of chromatin compartments A and B. In the embryonic epiblast, however, it does not seem to be the case, as chromatin compartments appear to be established at E3.5 when the genome is hypomethylated and are conserved in the epiblast at E7.5, after the methylation of the genome, suggesting that DNA methylation deposition poorly contributes to chromatin compartmentalization during epiblast formation *(Yu Zhang et al., 2018)*.

Histone marks are also deeply reshaped during naïve-to-primed pluripotency transition, mostly around regulatory elements like enhancers and promoters. Concomitantly to the gain of DNA methylation observed around genes regulating embryonic stem cells, active histone marks are rearranged. Although levels of histone acetylation globally decrease in association with increased expression of histone deacetylases, locus-specific examination revealed that H3K27ac is lost around pluripotency genes and gained at post-implantation epiblast-specific genes *(Buecker et al., 2014; Gökbuget and Blelloch, 2019; Shirane et al., 2016)*. This reorganization is mostly driven by transcription factors like FOXD3, OTX2 and OCT4, which play a central role in the decommission of enhancers during differentiation *(Buecker et al., 2014; Krishnakumar et al., 2016; Respuela et al., 2016; Yang et al., 2014)*. In contrast, H3K9 methylation levels increase during epiblast differentiation. Domains of H3K9me3 are especially retained at the LTR regions of TEs, and some genes gain repressive H3K9me3 domains, most notably germline genes *(C. Wang et al., 2018)*. This is probably associated with progressive repression of other developmental genes. Indeed, as the embryonic cells engage into lineage specification, the highest number of H3K9me3-marked genes is found in progenitor cells of the endoderm and mesoderm germ layers *(Nicetto et al., 2019).* Similarly, the deposition of H3K9me2, mediated by G9A, also increases during epiblast differentiation *(Lienert et al., 2011; Wen et al., 2009)*. While DNA methylation and H3K9 methylation are known to work synergistically, the hierarchy between the two repressive marks is still debated in the context of repressing developmental genes in somatic germ layers. Finally, while H3K27me3 accumulates around developmental genes in the epiblast *(Wang et al., 2018)*, the number of bivalent promoters, marked by both H3K27me3 and H3K4me3, significantly increases in differentiated epiblast cells compared to blastocyst embryonic cells *(Liu et al., 2016)* **(Figure 10B)**. These bivalent regions appear anti-correlated with DNA methylation in the epiblast, since they are generally hypomethylated. This appears to be regulated by active DNA demethylases, as in the E6.5 epiblast of TET1/2 double-KO mice, gain of DNA methylation is associated with decreased H3K4me3 in bivalent regions specifically, whereas H3K4me3 levels are unaffected around housekeeping genes. Beyond the E6.5 epiblast, in somatic

progenitor cells, this bivalency is resolved with genes becoming either activated and retaining H3K4me3, or inactivated with H3K27me3 *(Xiang et al., 2019)*.

## 3.2 From *in vivo* to *in vitro*

The first iteration of *in vitro* cells or tissue culture-based methods using nutrient-rich liquid media was enunciated by Gottlieb Haberlandt, an Austrian botanist, in 1902 *(Haberlandt, 1902)*. He developed the idea that cells not only represent a structural unit of living organisms, but also a functional unit at the physiological level. He further argued that such culture methods would allow to investigate the potentialities of the cells, and their relationships to one another. Notably, he drew the basis of the modern definition of totipotency by postulating that plant cells retain the intrinsic ability to generate any tissue and that "**in this way, artificial embryos from vegetative cells could be successfully cultivated**" *(Jourdin, 2008)*. These predictive thoughts were followed by the first animal cell cultures established by A. Carrel in 1911 *(Carrel and Burrows, 1911)*, and then by the first plant cell cultures, established by P.A. White in 1939 *(White, 1939)*.

Nowadays, cultured models allow scientists to investigate and challenge the relationships between cells, elaborating notably on their differentiation abilities, and expanding our fundamental knowledge on how cells can grow an organism. Moreover, with the rise of 3D culture systems and the development of organoids, we are getting closer and closer to Haberlandt's original vision of an *in vitro* grown artificial embryo. This has been well illustrated by recent landmark studies in 2021, which managed to artificially support the growth of a mouse embryo to the equivalent of 11 days of gestation *(Aguilera-Castrejon et al., 2021)*, and allowed for the development of artificial human blastocysts *(Yu et al., 2021)*. In this part, we will cover the *in vitro* methods and tools developed to study peri-implantation development, and how they allow to alleviate the limitations posed by *in vivo* approaches, alongside with new culture models and the new ethical questions they raise. Indeed, the ability of embryonic stem cells to self-renew and to be propagated in a dish under appropriate culture conditions, after derivation from the embryo at a certain time in development, results in an *in vitro* culture that can be considered as a "capture" of this particular developmental state.

### 3.2.1 Recapitulating early development in a petri dish

**The culture of embryonic stem cells resembling ICM Epiblast:**

Historically, ESCs were first cultured in serum-based medium complemented with Leukemia Inhibitory Factor (LIF). They were considered to be similar to *in vivo* E4.0 ICM pluripotent stem cells, despite demonstrating heterogenous expression of pluripotency genes like *Nanog*, and notably, higher DNA methylation levels, which are closer to somatic differentiated cells *(Singer et al., 2014; Williams et al., 1988)*. On one hand, LIF supports self-renewal of ESCs through the activation of the JAK/STAT3 signaling pathway, which induces expression of the naïve pluripotency gene network. It could also promote DNA hypomethylation by simultaneously enhancing TET activity and DNMT degradation *(Wulansari et al., 2021; Ye et al., 2013)*. On the other hand, serum contains differentiation-promoting and differentiation-inhibiting cytokines like FGF and BMP, respectively, thus inducing a "metastable state" for the stem cells *(Hackett and Surani, 2014)*.

Later on, this conflicted identity was resolved by using serum-free conditions, with the addition of GSK3 and MEK/ERK inhibitors (2i), supplemented with LIF, that was found to better nurture an equivalent to the naïve state of the pluripotent cells of the ICM *(Ying et al., 2008)*. Indeed, 2i ESCs harbor a hypomethylated genome similar to what is observed *in vivo* prior to implantation. While 70 to 80% of CpG dinucleotides are methylated in serum/LIF ESCs in average, 2i cells demonstrate 20% of methylation genome-wide *(Habibi et al., 2013; Stadler et al., 2011)*. The two inhibitors act on two fronts **(Figure 11)**. First, PD blocks FGF signaling by inhibiting the MEK/ERK signaling cascade, which prevents differentiation *(Deathridge et al., 2019; Hamilton et al., 2019)*. Then, CHIR activates WNT by inhibiting GSK3, which indirectly stimulates transcription of pluripotency genes *(Wray et al., 2011; Zhang et al., 2020)*. How these two inhibitors favor pluripotency is currently under strong scrutiny. Nevertheless, key findings have demonstrated that they promote DNA hypomethylation by jointly repressing DNMTs and inducing TET DNA demethylases *(Sim et al., 2017)*. In addition, 2i-cultured ESCs share many expressed genes with their *in vivo* equivalent from the ICM, like core pluripotency triad (CPT) genes (*Oct4*, *Nanog* and *Sox2*) as well as naïve-specific markers like *Klf4, Rex1 (also known as Zfp42), Esrrb*, *Dppa3* and others *(De Los Angeles et al., 2015; Dunn et al., 2014; Tesar*

*et al., 2007)*. Under these conditions, *in vitro* naïve pluripotent ESCs also recapitulate the "permissive" chromatin landscape found *in vivo*. Most notably, levels of active histone marks H3K4me3 and H3K27ac are globally high. Additional features of naïvety encompass two active X chromosomes in female ESCs, and the activation of the distal enhancer of *Oct4* to drive its expression *(Aoto et al., 2006; Azuara et al., 2006; Choi et al., 2016; Schulz et al., 2014)*. However, these characteristics are lost upon epiblast differentiation.

**The capture of differentiated epiblast stem cells:**

Cultures of primed epiblast stem cells can be derived from E6.0-7.5 post-implantation embryos (EpiSCs), or alternatively obtained from ESCs cultured in presence of FGF2 and Activin A (*in vitro* EpiSCs or Epiblast like-cells, EpiLCs) **(Figure 11)**. Induction of the FGF pathway has been shown to mediate the extinction of naïve pluripotency-promoting factors, including the export of the KLF4 protein from the nucleus *(Dhaliwal et al., 2018; Hamilton et al., 2019; Kunath et al., 2007)*. In opposition to ESCs, EpiSCs display a methylated genome and silenced naïve genes, while the CPT network remains expressed. In addition, EpiSCs—whether derived *in vivo* or *in vitro*—express markers of the early post-implanted epiblast, like *Dnmt3B*, *Otx2*, *Fgf5*, *Oct6,* and early somatic lineage markers such as *Sox17* *(Acampora et al., 2013; Morgani et al., 2017)*. The delineation between ESCs and EpiSCs/EpiLCs is furthermore supported by the fact that EpiSCs poorly reprogram and revert back into naïve ESCs when re-introduced in a blastocyst, suggesting a true change in cell identity and differentiation competence between the two pluripotent states *(Brons et al., 2007; Ohtsuka et al., 2012)*.

Recent efforts have been made to capture alternative states of pluripotent stem cells *in vitro,* that could represent new "transient state of pluripotency" found in the E5.0-5.5 epiblast, a concept that we will develop later. It has been demonstrated that culturing ESCs in ABC/L media (ActivinA, BMP4, CHIR and LIF) induces the formation of "advanced pluripotent stem cells" (ASCs). Identified as an intermediate state between 2i ESCs and EpiSCs, ASCs demonstrate a higher contribution to chimerae and germline *in vivo* than EpiSCs *(Bao et al., 2018)*. These ASCs cells display high DNA methylation, but keep higher expression of naïve pluripotency genes than

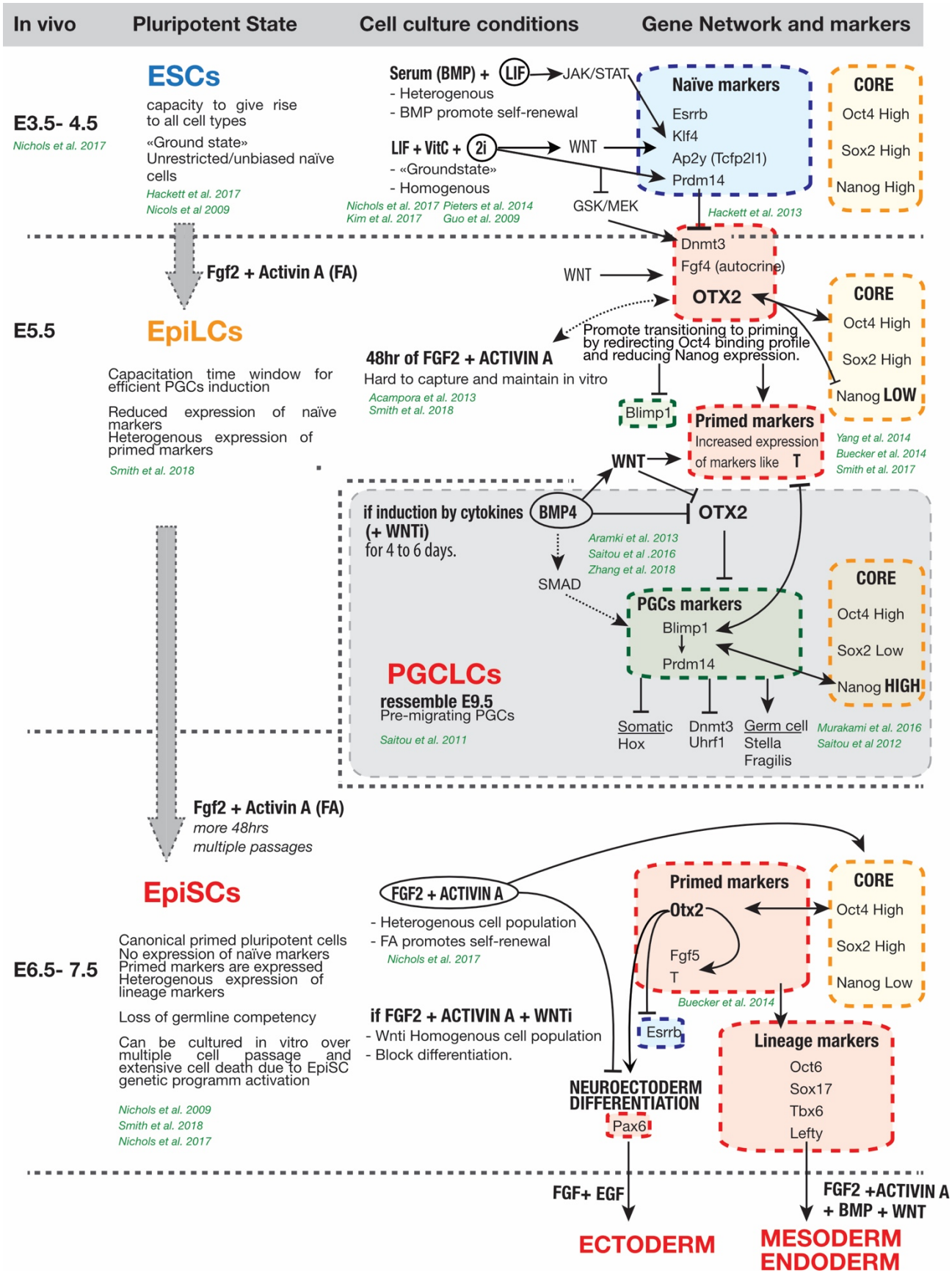



**Fig. 11 I A guideline through *in vitro* early development**

Graphical model of *in vitro* methods to recapitulate early development, highlighting relationships between signaling pathways and major inducers of each fate, with subsequent impact on transcriptional programs.

serum/LIF ESCs, since they demonstrate hypomethylated regulatory elements for the naïve markers *Rex1*, *Essrb*, *Nr0b1*, *Tlc1*, like 2i ESCs. Counterintuitively, they show higher expression of mesoderm and primitive streak markers. Overall, the methylation level of ASCs remains lower than in EpiSCs—72-74% and 80%, respectively—and is especially lower at intergenic regions. Moreover, they transcriptionally resemble the E5.5 epiblast *(Bao et al., 2018; B. Wu et al., 2020)*. Further investigation on the unconventional ASC methylome is needed to understand how DNA methylation could play a role in regulating their particular potential. Finally, the novel characterization of rosette stem cells (RSCs) *in vivo,* present in the E5.1 epiblast, and their derivation *in vitro* constitute another breakthrough in the identification of new intermediate pluripotent states. These cells rely *in vitro* on LIF and the inhibition of both WNT and MEK pathways, thus resulting in a population of cells that co-express naïve markers and the post-implantation marker *Otx2*. RSCs harbor intermediate levels of DNA methylation (2% of 5mC genome-wide) and are prone to rapid differentiation compared to 2i ESCs, while having higher chimera contribution than EpiSCs *(Neagu et al., 2020)*. Examples of intermediate pluripotent cell cultures are listed in **Table 1**.

**Recapitulating germline specification *in vitro*:**

Recent advances have allowed the faithful recapitulation of PGC-like cell (PGCLC) specification *in vitro,* by shortly differentiating ESCs into Epiblast-like cells (EpiLCs) for two days using FGF2 and ActivinA, and then exposing EpiLCs to a cytokine cocktail (notably comprised of bone morphogenetic proteins, BMPs) for four days **(Figure 11)**. Strikingly, PGCLC differentiation can be used to generate fertile offspring, highlighting their proximity to *in vivo* PGCs *(Hayashi et al., 2011)*. Moreover, PGCLCs transcriptionally re-activate some naïve markers like *Dppa3* and *Prdm14*, and repress *Dnmt* genes, *Otx2* and other epiblast markers, while maintaining core pluripotency markers expressed in a comparable fashion to *in vivo* PGCs *(Bleckwehl and Rada-Iglesias, 2019)*.

*In vitro* PGCLCs also partially reshape their epigenetic landscape. DNA methylation for example, after reaching a level of 70-80% of CpG methylation in EpiLCs **(Table 2)**, decreases to 50% and 40% in PGCLCs after four and six days of culture, respectively, while DNA methylation levels in terminal *in vivo* PGCs reaches

around 7% at E13.5 in male *(Seisenberger et al., 2012; Shirane et al., 2016)*. DNA methylation in PGCLCs is mostly lost in intergenic regions, gene bodies and to some extent, at some repeats, but remains high at the promoters of germline genes *(Hargan-Calvopina et al., 2016; Shirane et al., 2016)*. However, this *in vitro* epigenetic reprogramming does not appear to be functionally sufficient, since the generation of germline stem cells-like cells (GSCLCs) from PGCLCs, show impaired ability to enter spermatogenesis, likely due to the incomplete erasure of DNA methylation at the promoters of pluripotency and meiotic genes during PGCLC specification *(Ishikura et al., 2016)*. Despite this, GSCLCs share similar morphology with their *in vivo* counterparts (Spermatogonia Stem cells), and constitute a great *in vitro* model to study germline development.

**What about a totipotent-like cell culture model?**

Seminal work from T. McFarlan in 2012 allowed for the identification of a subpopulation of 2C-like cells (2CLCs) in ESC cultures. Their strategy for their identification relied on the activation of a MERVL reporter *(Macfarlan et al., 2012)*, since these retrotransposons are transiently up-regulated in the totipotent 2-cell stage embryo, and drive the transcription of 3% of mRNAs during ZGA by acting as alternative promoters *(Kigami et al., 2003; Peaston et al., 2004)*. Using this strategy, they observed that up to 3% of ESCs in culture spontaneously acquire features of 2C-like cells: an active chromatin signature, expression of the totipotency-associated *Dux* transcription factor, and low expression of pluripotency factors like *Oct4*, *Sox2* and *Nanog*. Following this discovery, many studies have attempted to increase the fraction of 2CLCs within the ESC population by modulating the transitioning rate. Chemical treatments that favor chromatin opening or spliceosome repression have been shown to increase the acquisition of totipotency-like features *(Shen et al., 2021; Yang et al., 2022)*.

Recently, human eight-cell stage like cells (8CLCs) were generated using media containing chemicals promoting histone acetylation and inhibiting histone methylation. These 8CLCs express totipotent genes like *Zscan4* and *Dux*, and downregulate pluripotent genes like *Oct4* and *Sox2*. Moreover, they also demonstrate an open chromatin state *(Mazid et al., 2022)*. Interestingly, *Dppa3* was found to be

critical in humans for the formation of 8CLCs, in contrast to mice where this gene is dispensable in establishing 2CLCs. This suggests that mechanisms promoting totipotency between primates and rodents might have diverged in evolution. Furthermore, the authors used a newly developed human *in vitro* reconstructed artificial blastoids, to demonstrate the ability of 8CLCs to give rise to both embryonic and extra-embryonic tissues in humans. To conclude, all these *in vitro* methods aim at recapitulating early development and allow monitoring these transitions, notably by studying the transcriptional program or the reprogramming the chromatin landscape to infer cell identity. However, functional assays are also critical for characterizing the different embryonic cell types.

**Table 1. Example of Alternative/intermediate mouse pluripotent stem cell from *in vitro* cultures**

| | ***Origin*** | ***Active Signaling pathways*** | ***Pluripotency genes*** | ***Lineage markers*** | ***Potentialities*** | ***Epigenome*** | ***Refs*** |
|---|---|---|---|---|---|---|---|
| ***AFSCs*** | EpiSCs with ActA, Bmp4, CH and LIF **(ABC/L media)** | NODAL<br>ERK/MERK<br>BMP<br>JAK/STAT | *Oct4, Nanog, Sox2*<br><br>Naïve markers are repressed | *Blimp1*<br>*Sox17*<br>*Eomes*<br>*Lefty*<br>*Brachury (T)* | Teratoma formation<br><br>Integration into ICM of blastocysts | X chromosome inactivated<br><br>70% mCpGs | S.Bao 2018 Cell Research<br><br>B.Wu 2020 Stem cell Reports |
| ***ASCs*** | ESCs cultured with ActA, Bmp4, CH and LIF **(ABC/L media)** | BMP<br>WNT<br>JAK/STAT<br>NODAL | *Oct4, Nanog, Sox2*<br><br>Naïve markers alike ESCs | Mesoderm markers Up (*Eomes, Hoxa* …)<br><br>Meiotic genes up | Chimerae<br><br>Contribution to the Germline<br><br>Contribution to extra-embryonic | 70% mCpGs<br><br>2 active X chromosomes | S.Bao 2018 Cell Research<br><br>B.Wu 2020 Stem cell Reports |
| ***RSCs*** | ICM cells cultured with LIF, IPWI2, MEKi **(LIM media)** | JAK/STAT | *Oct4, Nanog, Sox2*<br><br>Naïve markers *Klf4, Esrrb* | *Otx2* induced<br><br>*Oct6* repressed | Rapid priming *in vitro* | 3% of mC genome-wide<br><br>Increase in bivalent domains (H3K27me3/H3K4 me3)<br><br>Accumulation of pericentric H3K27me3<br><br>2 active X chromosomes | Neagu 2020 Nat. Cell Biology |
| ***FTW-ESCs***<br><br>***(INTPSCs)*** | ICM cells or ESCs cultured with FGF2, ActivinA and CHIR in N2B27 **(FTW media)** | MEK/ERK<br>TGF-b/SMAD<br>WNT | *Oct4, Sox2*<br><br>Low *Nanog*, *Rex1* and *Klf4* | High *Dnmt3A, Dnmt3L, Fgf5, Lef1* | Germline competence *in vivo* and *in vitro*<br><br>Chimera contribution (< ESCs) | 2 active X chromosomes<br><br>Broad bivalent domains | L.Yu 2020 Cell Stem Cells<br><br>(*Tsukiyama and Ohinata, 2014)* |
| ***FSCs*** | E5.5 Epiblast or from ESCs cultures with ActivinA, WNTi XAV939, RARi BMS493 **($A_{IO}$XR media)** | MEK/ERK, NODAL, | Low *Nanog*<br>*Oct4, Sox2*<br><br>Low *Klf4* | High *Otx2*<br><br>Minimal *Brachury* (*T*) | Germline competence *in vitro* and *in vivo*.<br><br>Rapid somatic differentiation *in vitro*<br><br>All lineages in chimerae | Increased bivalent domains<br><br>Open chromatin alike EpiLCs | Kinoshita 2020 Cell Stem Cell |

### 3.2.2 Functional assays and tools to test the potentialities of stem cells

Although transcriptome and chromatin profiling can inform cell identity at a given time, it says little about the potentialities of the cell. Evaluation of cellular potential is critical when studying the ability of ESCs for differentiation. To test this, differentiation and contribution experiments remain the go-to approaches, although with different levels of confidence for interpretation *(De Los Angeles et al., 2015)*.

A classical method to test the level of potency of embryonic stem cells is to apply differentiation protocols *in vitro* into different lineages. This approach presents a number of advantages. First, it is very modular, as a large panel of differentiation protocols now exist to induce specific embryonic or extra-embryonic lineages, going from early progenitors to terminally differentiated cells. Moreover, these protocols can be directed or not, in monolayers or in three dimensions, and can give rise to one or several lineages at the same time, allowing scientist to test many different conditions. Second, these methods are usually fast to perform and high-throughput, allowing numerous downstream applications.

Contribution experiments usually give the highest level of confidence on the potency of cells. Historically, the formation of teratocarcinoma was an early method to test the potency of stem cells *in vivo*. It involves injecting embryonic stem cells in immune-deficient mice, and monitoring the development of a tumoral mass containing different tissues deriving from the three-germ layers *(Stevens, 1958)*. However, the histological analysis of teratomas is poorly quantitative and prone to misinterpretation. In comparison, the generation of chimeric mice, either by aggregation of stem cells with host morulae, or by direct injection into the ICM of host blastocysts, not only allows to assess the contribution to all three germ layers, but also their ability for germline transmission *(Artus and Hadjantonakis, 2011; Brook and Gardner, 1997)*. Chimeric mice and teratoma experiments were pivotal in the discernment of naïve and primed pluripotent stem cells, as the latter are poorly competent for germline transmission *(Brons et al., 2007)*. However, these methods can be difficult to put in place and do not generally display a very high yield. In this regard, methods aiming at developing artificial embryos or organs—also called embryoids and organoids—could reconcile *in vitro* and *in vivo* approaches, opening new avenues for research.

Seminal work by Evans and Martin in 1975 established the first model of artificial embryonic development by aggregating ESCs into 3D embryoid bodies that can induce the three embryonic germ layers and recapitulate basic antero-posterior axis *(Martin and Evans, 1975; ten Berge et al., 2008)*. Work conducted in Martinez Arias Lab established the first protocol of "gastruloids". These are also formed by aggregation of ESCs, and can grow alongside the three axes of development: Antero-posterior, Left-right, and Dorso-ventral. Thus, gastruloids can be defined as "**Three dimensional aggregates of pluripotent stem cells that, under appropriate culture conditions, develop an embryo-like organization with three orthogonal axes and a precise distribution of the multiple derivatives of the three germ layers**" *(Van Den Brink et al., 2014; van den Brink and van Oudenaarden, 2021)*. Further improvement of this protocol allowed gastruloids to recapitulate—with striking similarity compared to *in vivo* embryos— *i)* the segmentation of the posterior axis, with colinear expression of *Hox* genes and somitogenesis *(Beccari et al., 2018; Veenvliet et al., 2020)*, and *ii)* anterior development with the formation of neural tissues *(Duval et al., 2019; Girgin et al., 2021)*. Thus, gastruloid formation can decipher the specific properties of the stem cells. Indeed, it was found that EpiSCs are less competent to aggregate and develop into gastruloids, which is reminiscent of their poor ability to generate chimerae *(Cermola et al., 2021)*.

Recent application of these protocols to human embryonic—or induced pluripotent—stem cells has allowed the development of human gastruloids recapitulating mesodermal and neural-like tissues *(Libby et al., 2021; Moris et al., 2020; Olmsted and Paluh, 2021)*. The reconstruction of human gastruloids opens many opportunities in the medical field for studying genetic diseases, creating new organs, improving *in vitro* fertilization and others. However, the moral and ethical status of human embryoids, especially now that neural tissues can be formed, must be addressed. Using induced pluripotent stem cells (iPSCs) and limiting gastruloid formation to a limited number of specific tissues allow circumventing most ethical concerns and are probably sufficient for modeling and studying most diseases. However, such use of gastruloids requires scientists to unify their terminology to clarify the nature of the reconstructed tissues with their developmental time equivalent *(Matthews et al., 2021; Nicolas Rivron et al., 2018)*. In many countries, the growth of

human embryos *in vitro* is limited to the equivalent of 14 days of gestation, and a debate on whether this limit should be extended exists among researchers *(Appleby and Bredenoord, 2018; Blackshaw and Rodger, 2021; Hyun et al., 2016)*. The question whether reconstructed human embryos should also fall into this limit obviously needs to be addressed in concertation with ethicists and policymakers in a public manner, as informed consent of patients will be necessary for the conduction of reproductive studies on reconstructed human germline *(van den Brink and van Oudenaarden, 2021)*.

## 3.3 From totipotency to pluripotency

The purpose of peri-implantation development, from fertilization to the formation of the differentiating epiblast, is to initiate the formation of a new organism. But all the different stages in this period of development do not share the same potential for generating this new organism. In reality, cells of the embryo become more and more restricted in their differentiation abilities during development. The zygote and the blastomeres of the 2-cell stage embryo are able to give rise to both embryonic and extra-embryonic lineages, being then referred as "totipotent" cells, whereas cells that contribute only to the three embryonic germ layers (endoderm, mesoderm and ectoderm) are considered as "pluripotent", like naïve embryonic stem cells in the blastocyst and primed epiblast stem cells *(Riveiro and Brickman, 2020)*. Concomitantly to this restriction in differentiation potential, is the chromatin remodeling that takes place in early development, notably the reprogramming of DNA methylation profiles. Thus, the influences of epigenetic changes on developmental abilities are under strong scrutiny, and evidence mostly comes from *in vitro* models of early development. In the following part, I will cover the regulation of totipotency and pluripotency states, and how chromatin changes, especially DNA methylation, affect cellular potentialities.

### 3.3.1 Totipotency: the ability for all things

As described previously, investigating the main regulators of ZGA or totipotency is challenging due to the limited amount of available cells. As a consequence, most of

the work attempting to identify the transcriptomic signature of totipotency relies on 2C-like cell cultures.

Despite representing cells with the highest potential, mammalian totipotent cells share few common factors with pluripotent cells **(Figure 12A)**. For example, despite being a core factor in the control of pluripotency, OCT4 plays no apparent role in totipotency. This is exemplified by the maternal deletion of *Oct4*, which is compatible with normal development after fertilization. Moreover, levels of *Oct4* and other pluripotency factors like *Sox2*, *Prdm14*, *Rex1* are relatively low in 2CLCs and their loss does not affect their emergence in ESC cultures *(Rodriguez-Terrones et al., 2018; Wu et al., 2013)*. One important factor for totipotent cells is ZSCAN4 **(Figure 12B)**. It is a regulator of telomere elongation, and has been shown to regulate development timing and totipotency establishment, as its knock-down delays 2-to-4-cell stage progression, and it is also required later on for blastocyst formation. In addition, *Zscan4* knock-down in ESCs induces chromosomal abnormalities and reduces growth *(Falco et al., 2007; Zalzman et al., 2010)*. Furthermore, the expression of MERVL retrotransposons, specifically at the 2-cell stage *in vivo* and in cultured 2CLCs, serve as alternative promoters of totipotency specific transcripts *(Macfarlan et al., 2012; Peaston et al., 2004).* Then, the formation of the "open chromatin" landscape could also serve as a marker for totipotent stem cells. Accordingly, the knock-out of the H3K9me3 methyltransferase *Setdb1* or the H3K9me3 reader *Kap1* both promote ESC to 2CLC transition *in vitro* *(K. Wu et al., 2020)*. In regard to DNA methylation, as it is erased after fertilization and that DNMT1 has been found to prevent 2CLC transition *in vitro* *(Fu et al., 2019)*, it is probable that DNA methylation plays no role in totipotency. Indeed, the triple knock-out of the DNA methyltransferases *Dnmt1*, *Dnmt3A* and *Dnmt3B* does not affect the number 2CLCs found in ESC populations *in vitro* *(Eckersley-Maslin et al., 2016).* In contrast, inducing pluripotent stem cells (iPSCs)—which requires an equally dramatic reprogramming with the induction of pluripotency from a fully differentiated somatic cell—requires DNA methylation removal, mostly via TET2 that has been shown to be central for efficient reprogramming *(Caldwell et al., 2021; Pei et al., 2014)*. This rather suggests a restricting role for DNA methylation in the acquisition of totipotency and pluripotency.

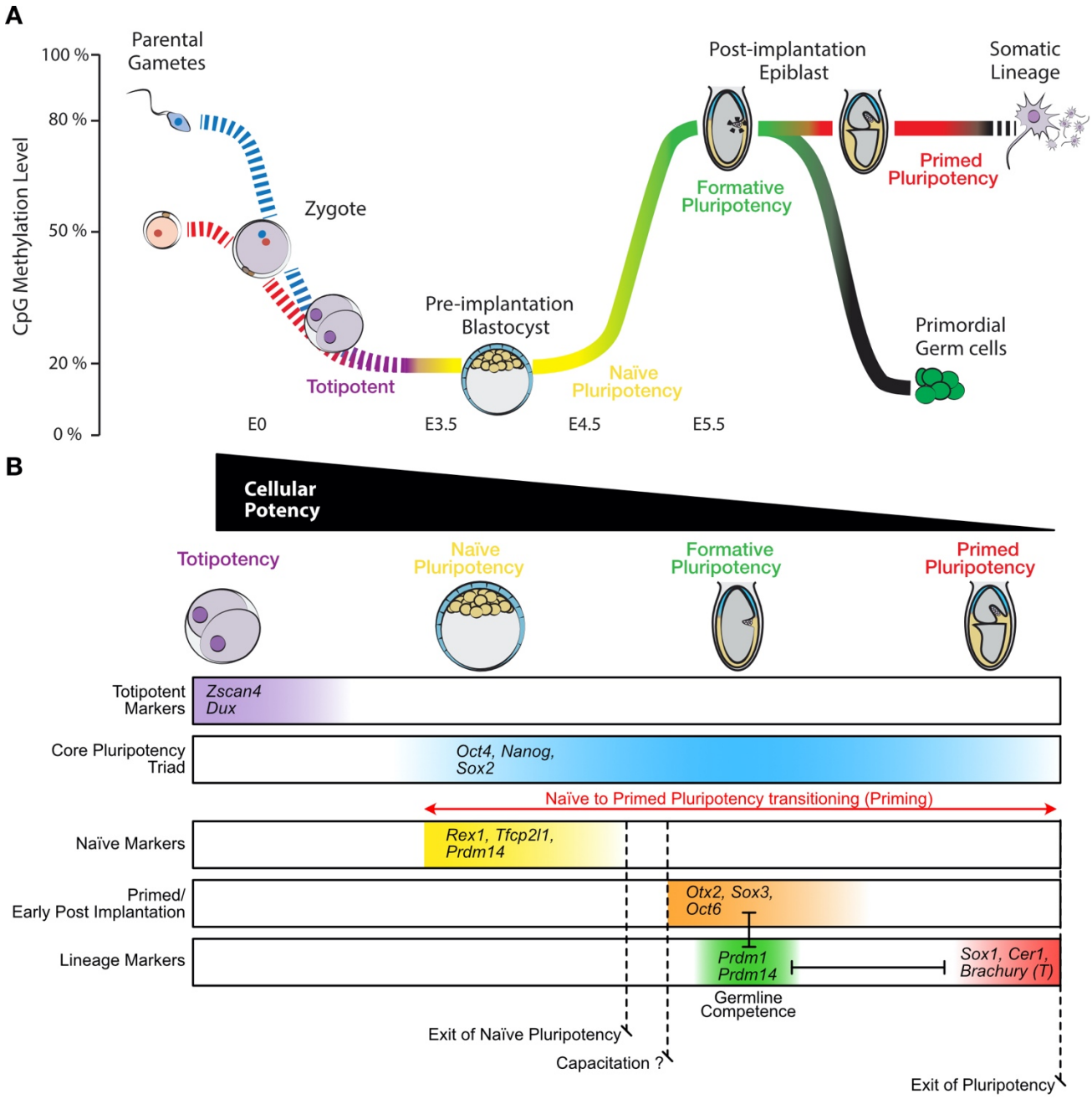



**Fig. 12 I DNA methylation dynamics and Naïve to Primed pluripotency transitioning**

**A.** Schematic of the interplay between changes in DNA methylation and in cellular potency during early mouse embryogenesis, prior to lineage specification. **B.** Progression of potency, for the 2-cell totipotent embryo to the primed epiblast. *In vitro* studies suggest that progression through the different phases of pluripotency is sequential. Yet, it is possible that during *in vivo* development, the different stages of pluripotency overlap with each other and could coexist in the same embryo at the same time of development. In particular, it remains to be demonstrated that complete dissolution of naïve pluripotency factors occurs *in vivo* prior to capacitation of stem cells for priming.

Finally, little information exists on the transition from totipotency to pluripotency. When this transition occurs is also unclear *in vivo*. While a single 2-cell stage blastomere is sufficient for the formation of an entire organism, 4-cell and 8-cell stage blastomeres display variable efficiency in their ability to contribute to extra-embryonic tissues and require the support of carrier host-blastocysts *(Piotrowska-Nitsche et al., 2005)*. This 'incomplete totipotency' has been associated with different expression levels of *Oct4*, hence suggesting a progressive transition period where totipotency fades away, and pluripotency is initiated.

### 3.3.2 A switch in pluripotency: the naïve-to primed-pluripotency transition

The ground state of pluripotency is defined around the identity of embryonic stem cells found in the preimplantation blastocyst. These cells are in a naïve pluripotent state—which refers to their unrestricted ability to give rise to all embryonic lineages—but also present a specific transcriptomic and chromatin signature. Thus, there is a slight semantic difference between "ground state" and "naïve" pluripotent cells. Indeed, while ground state cells are always in a naïve pluripotency state, not all naïve stem cells are ground state. This is best illustrated by ESCs cultured in serum/LIF which are "naïve", as they can contribute to all embryonic lineages in chimerae, but are not "ground state" as they have an hypermethylated genome and co-express naïve and primed markers *(Hackett and Surani, 2014)*.

As we previously mentioned, naïve pluripotent stem cells express high levels of general core pluripotency triad (CPT) *Oct4-Nanog-Sox2*. These genes are central to establish and regulate pluripotency. Loss of OCT4 or SOX2 activity is associated with impaired differentiation toward the mesendoderm or neural lineages, respectively, while NANOG loss in ICM cells cannot sustain a viable epiblast *in vivo*, and promotes precocious differentiation *in vitro* *(Silva et al., 2009; Thomson et al., 2011)*. The reinforcement of a naïve 'ground state' identity is achieved by induction of specific naïve markers, like ESRRB, REX1, and others **(Figure 12B)**. Such induction is mediated by activation of specific *cis*-regulatory elements via the CPT. Notably, OCT4 binds, often in cooperation with NANOG and SOX2, to naïve genes *(Buecker et al., 2014)* and also promotes their own expression *(Chew et al., 2005; Rodda et al., 2005)*. Then, some naïve markers, like ESRRB, initiates a self-reinforcing positive loop

for the maintenance of naïve pluripotency *(Festuccia et al., 2021; Sevilla et al., 2021)*. While sitting at the top of the pluripotency regulatory network, the cooperativity and many shared targets between the CPT and naïve genes were initially thought to convey redundancy and plasticity to maintain pluripotency. However, it is noteworthy that some pluripotency factors can also have specific properties. For example, while loss of NANOG is dispensable for maintenance of ESCs, its expression allows the maintenance of pluripotency in a LIF-independent manner *(Chambers et al., 2003; Silva et al., 2009; Q.-L. Ying et al., 2003)*. Indeed, signaling pathways play a critical role in establishing and maintaining pluripotency, most notably, the JAK-STAT, WNT and BMP pathways that have been reviewed elsewhere **(Figure 15A)** *(Pieters and van Roy, 2014; Weinberger et al., 2016)*.

Similarly, the exit of naïve pluripotency and induction of primed pluripotency—also called naïve-to-primed transitioning, or priming—seems to be initiated by external signals driving differentiation, such as FGF and NODAL **(Figure 11-12B)**. Such signaling pathways are initially repressed in naïve pluripotent cells *(Grabole et al., 2013; Kunath et al., 2007; Mulas et al., 2017)*. In addition, a differential regulation from the CPT can also participate in the initiation of priming. Early work comparing gene regulatory elements (GREs) between ESCs and EpiLCs *in vitro*, have identified a unique switch of enhancers during this transition. This was marked by the redistribution of the core pluripotency OCT4 factor at developmental genes by the early post-implantation marker OTX2 *(Yang et al., 2014)*. Interestingly, the regulation of *Oct4* itself seems to be DNA methylation-dependent during this transition: its expression is driven by a distal enhancer (DE) in ESCs and a proximal enhancer (PE) during epiblast differentiation. While being both DNA methylation-free in ESCs, the *Oct4* DE becomes methylated as cells progress toward the epiblast, although it may be a consequence rather than a cause of the switch. In addition, in the future germline, the expression of *Oct4* is driven initially by the DE, then by both demethylated DE and PE elements *(Choi et al., 2016)*. Relying on the example set by *Oct4*, the question whether the DNA methylation status of GREs plays a role in developmental progression can be raised. Several studies have indeed demonstrated that DNA methylation is acquired at naïve gene enhancers, concomitantly with their repression upon priming *(Kalkan et al., 2017; Petell et al., 2016)*.

*Nanog* is transiently down-regulated during the formation of epiblast cells, and this short down-regulation seems critical for proper lineage specification *(Chambers et al., 2007; Hayashi et al., 2011)*. Investigating the gene regulatory network behind priming has helped uncovering the role of many factors involved in the regulation of this transition: some act as sensors of differentiation signals like ID1 *(Malaguti et al., 2019)*, others by mediating the rewiring of CPT during epiblast differentiation like ETV5, RBPJ, and TCF3 *(Kalkan et al., 2019)*, or by activating specific primed genes like ZIC3 or FOXD3 *(Krishnakumar et al., 2016; Respuela et al., 2016; Yang et al., 2018)*. Many of these priming regulators, like ETV and RBPJ, seem to be conserved and shared from mice to primates *(Lackner et al., 2021)*. Importantly, OTX2 has been found to play an increasing role in the regulation of priming, and can be considered as a master-regulator of this transition **(Figure 12B)**. As previously described, OTX2 redirects OCT4 binding to epiblast-specific elements, participating in the decommission of naïve enhancers during differentiation. But OTX2 is also sensitive to differentiation signals, like BMP and WNT, which regulate its repressive role on germline specification, and allows for somatic fate adoption by downregulating *Nanog* at the exit of pluripotency, hence supporting the somatic program (Acampora et al., 2013; Di Giovannantonio et al., 2021; Krishnakumar et al., 2016; Zhang et al., 2018).

Naïve and primed states represent the two extremities of the pluripotency window, following totipotency and preceding lineage specification **(Figure 12A)**. Yet, it has become more and more evident that pluripotency progresses in a step-wise manner toward priming, rather than existing in a oscillating metastable state *(Nichols and Smith, 2009)*. In a paper published in 2017, Austin Smith proposed that this intermediate transient step-wise progression corresponds to the 'executive' phase that allows stem cells to gain their competence for differentiation. He termed this phase "formative" pluripotency, which is distinct from naïve and primed pluripotency, and is supported by a specific transcriptomic signature (with dissolution of naïve genes, and still low lineage marker expression), partial X chromosome inactivation and intermediate DNA methylation levels *(Smith, 2017)*. However, such a formative state might exist only subjectively, as transitioning across pluripotency states appears like a continuum. Thus, at the exit of naïve pluripotency, stem cells might transition through successive discrete and transient states before reaching primed pluripotency,

each of these intermediate states could therefore be identified as an eligible formative pluripotent state. Yet, under this description, ASCs and RSCs that we previously described, do not completely fulfill the definition of formative pluripotency as they co-express both naïve network and early-primed genes. While *in vitro* it is clear that the dissolution of naïve genes is a requirement for proper differentiation *(Mulas et al., 2017)*, this distinction is not yet evident *in vivo*. It is possible that *in vivo*, formative pluripotency might also apply to stem cells undergoing simultaneous exit of naïve pluripotency and induction of priming. Overall, the development of intermediate pluripotent stem cells is an important step stone to further elucidate the impact of cell signaling and chromatin regulation on differentiation and stem cell renewal.

### 3.3.3 DNA hypomethylation and naïve pluripotency, cause or consequence?

*In vivo* studies have demonstrated that repressive chromatin marks are critical for post-implantation development and differentiation: the knock-out of DNMTs and histone modifying proteins often results in embryonic lethality at the onset of gastrulation *(Lei et al., 1996; O'Carroll et al., 2001)*. In the mouse epiblast, double-KO of *Dnmt3A* and *Dnmt3B* leads to precocious expression of germline genes and some developmental genes, thus highlighting a potential role for DNA methylation in maintaining the integrity of transcription *in vivo* *(Dahlet et al., 2020)*. However, the lethality and growth retardation in these embryos do not become evident until E8.5, a moment in mouse development when embryonic stem cells have already lost their pluripotency and have started acquiring a somatic identity with high DNA methylation levels. Thus, the question of how critical is DNA methylation for establishing pluripotency and regulating differentiation is central.

The use of *in vitro* models can help answering this question. First, most differences between hypomethylated 2i and hypermethylated serum/LIF ESC genomes coincide with gene bodies, non-CpG island promoters and some TEs (notably SINEs and LINEs), whereas imprinting control regions (ICRs) and IAP elements remain resistant to DNA demethylation in 2i ESCs. Despite these differences, embryonic stem cells cultured in serum/LIF or 2i media are both considered as naïve pluripotent as they transcriptionally resemble ICM cells and can

contribute to a similar extent to chimeric embryos *(Ficz et al., 2013; Jamie A. Hackett et al., 2013; Hackett and Surani, 2014)*. The latter suggests that the genomic methylation levels of mouse embryonic stem cells do not majorly impact their differentiation potential. In addition, the generation of *Dnmt* triple-KO (TKO) ESCs demonstrated that pluripotent cells can be maintained and propagated in naïve conditions in complete absence of DNA methylation **(Figure 13A-13B-13C)**, as they express naïve markers similarly to their WT counterparts *(Tsumura et al., 2006)*. Similarly, penta-KO ESCs (*Tet1*, *Tet2*, *Tet3*, *Dnmt3A*, *Dnmt3B*) demonstrate no major impact on the transcriptome, including pluripotency gene expression *(Wang et al., 2020)*. This demonstrates that DNA methylation probably does not play a role in the direct regulation of naïve gene networks in ESCs.

Therefore, the possibility of other repressive chromatin marks acting as compensatory mechanism to the lack of DNA methylation was quickly raised after the generation of *Dnmt*-TKO ESCs. Yet, TKO ESCs with KO of *Eed* and *Suv39h*, responsible for the deposition of repressive histone modifications H3K27me3 and H3K9me3, respectively, can be cultured in 2i medium for over a week without demonstrating tendency for precocious differentiation *(Walter et al., 2016)*. These *in vitro* results give a hint that DNA methylation and other repressive marks are dispensable for the stable propagation of ESCs in naïve culture conditions, thus prompting the idea of genomic hypomethylation as being a consequence rather than a cause of naïve pluripotency **(Figure 13D)**. Finally, while both *Dnmt*-TKO and *Dnmt3A/3B*-DKO ESCs do not contribute to post-implantation embryonic lineages, they can colonize the ICM of the blastocyst, once more demonstrating that DNA methylation is not required for pre-implantation development and naïve pluripotency *in vivo* *(Kinoshita et al., 2021; Sakaue et al., 2010)*. Such high tolerance of embryonic stem cells to the removal of epigenetic regulators could be explained in part by their dispensability in the maintenance of naïve stem cell state, but also by the strong selective pressure at this stage for the stem cells. As the blastocyst stage really represents a bottleneck in development, it is possible that embryonic stem cells have developed a strong tolerance for epigenetic insults, and put in place many redundant mechanisms allowing for the continuation of embryogenesis.

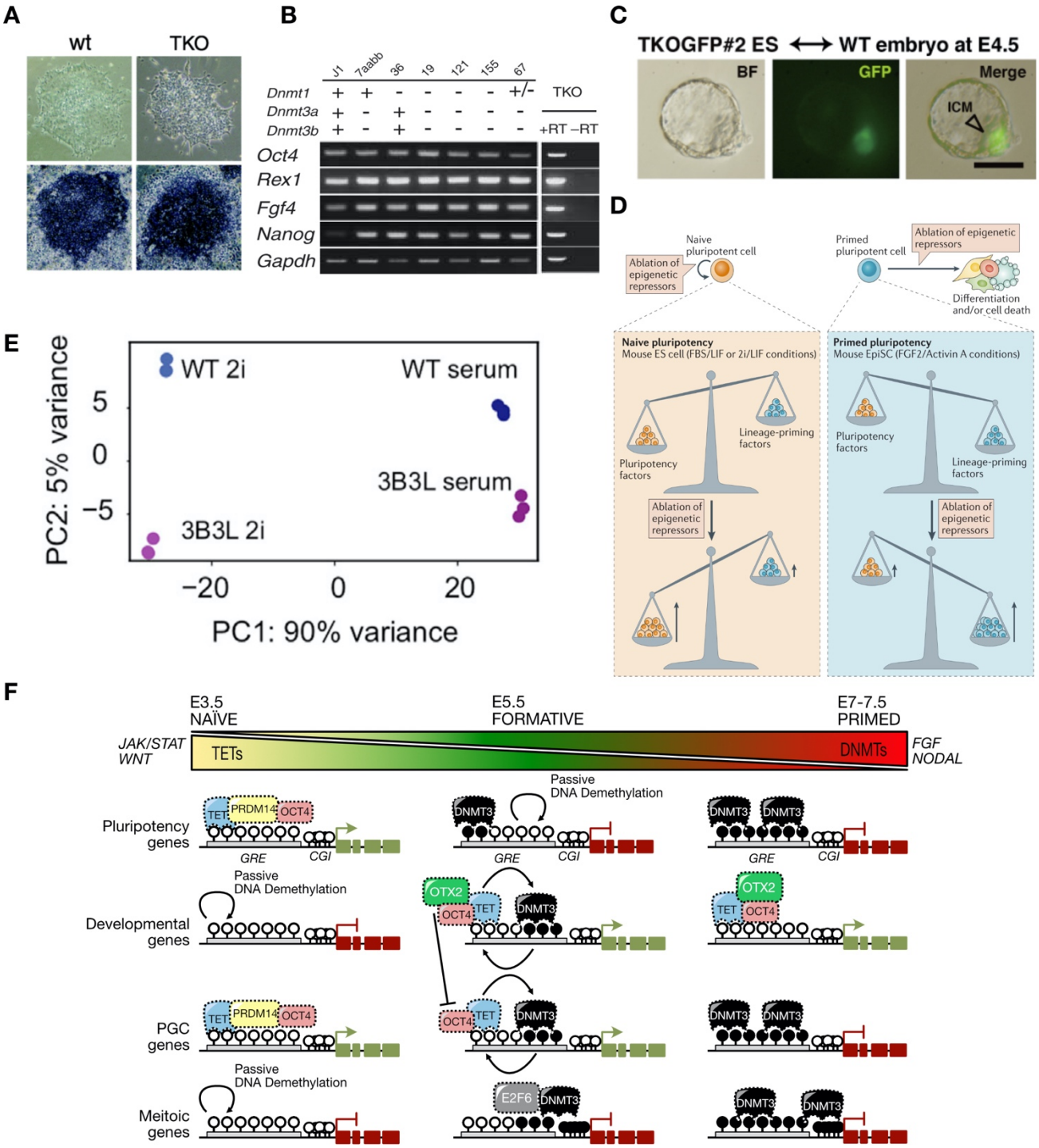



**Fig. 13 I Naïve-to-Primed pluripotency transitioning**

**A.** Brightfield images of WT and *Dnmt*-Triple KO ESCs (Top) and stained for Alkaline Phosphatase activity (Bottom) (Tsumura *et al,* 2006 Genes to Cells). **B.** RT-PCR of pluripotency markers in WT and TKO ESCs relative to the housekeeping gene *Gapdh*. **C.** Contribution of TKO cells bearing a GFP transgene 24h after injection in a E3.5 blastocyst ICM. TKO ESCs were found in the epiblast of the inner cells, highlighting that DNA methylation is dispensable for epiblast ICM formation *in vivo* (Sakaue *et al,* 2010 Current Biology). **D.** Graphical representation of the differential impact of repressive chromatin marks in naïve and primed pluripotent stem cells. Whilst the removal of epigenetic repressors promotes naïvety, in primed cells it induces precocious differentiation (Weinberger *et al,* 2016 Nat. Rev. Molecular Cell Biology). **E.** PCA analysis of the transcriptome of WT cells cultures in 2i or Serum/LIF media, alongside TKO cells expressing ectopic DNMT3B/3L, inducing high methylation. This graph highlights that the prime source of variability arises from the culture condition rather than the global level of methylation (McLaughlin *et al,* 2019 Cell Reports). **F.** Graphical representation of the variation of DNA methylation levels at different gene regulatory elements during naïve-to-primed pluripotency transition. The co-expression of TETs and DNMTs at the formative phase favors oscillation in DNA methylation levels around some gene loci.

Ultimately, in a more recent study, ectopic induction of DNMT3B and its co-factor DNMT3L in 2i ESCs (3B3L ES cells) resulted in a global gain of DNA methylation (to a similar level as observed in serum/LIF ESCs). In spite of these changes in DNA methylation—and in chromatin 3D organization at Polycomb loci, like at the *HoxD* locus—little effect was detected in the transcriptome of these 3B3L ESCs when compared to WT ESCs **(Figure 13E)** *(McLaughlin et al., 2019)*. More interestingly, the 3B3L ESC transcriptome mirrors the one of WT ESCs. This observation has been supported by another study, in which the use of an alternative titrated 2i medium (t2i) did not impact the transcriptional profile of ESCs that remained comparable to traditional 2i ESCs, despite inducing 60% of mCpG genome-wide *(Gretarsson and Hackett, 2020)*. Overall, this suggests that naïve ESCs do not rely on DNA methylation to maintain naïve pluripotency identity, and reversely, DNA methylation does not affect pluripotency. It seems that repressive marks, in general, are overridden by a network of transcription factors and signaling pathways that regulate both DNA methylation levels and pluripotent stem cell identity. Indeed, while it has become evident that repressive DNA methylation participates in the stable propagation of differentiated cell identities across cell divisions, its necessity is less obvious in the context of pluripotent stem cells, which rapidly change their identity and need to remain plastic enough to respond to various signaling cues *(Festuccia et al., 2017)*. However, while dispensable for the naïve state, it remains possible that DNA methylation might participate in priming the epigenetic landscape during epiblast differentiation.

### 3.3.4 DNA methylation in the regulation of pluripotency transitioning

**DNA methylation and exit of naïve pluripotency: take it or leave it**

According to the formative pluripotency hypothesis raised by Austin Smith *(Smith, 2017),* the capacitation of embryonic stem cells for differentiation must be preceded by exiting naïve pluripotency. The transition from naïve to primed pluripotency is characterized, notably, by the establishment of DNA methylation on the hypomethylated ESC genome. Whether this gain in 5mC plays a role in exiting naïve pluripotency and/or in regulating formative pluripotency is an important question.

*In vivo*, the naïve pluripotency gene network starts being down-regulated at E4.5 and is completely extinguished at E5.5, while the genome rises up to 60% of CpG methylation at E5.5 and 80% at E6.5 *(Auclair et al., 2014; Mohammed et al., 2017)*. *In vitro,* the exit of naïve pluripotency occurs in an asynchronous manner. Using a *Rex1*-GFP reporter ESC cell line, it was found that repression of the naïve program shows high inter-cellular heterogeneity until 16h after releasing ESCs from the 2i medium towards EpiLC differentiation. Complete repression of the naïve gene network is only observed after 36 to 48h. Similarly, DNA methylation is progressively acquired as ESCs exit naïve pluripotency *in vitro (Kalkan et al., 2017).*

A more recent study identified that super-enhancers (SEs) in ESCs display low to intermediate DNA methylation levels. The discovery of permanently unmethylated (PU) and differentially methylated (DM) subunits in these super-enhancers demonstrated that the DM regions gained DNA methylation in EpiLCs and EpiSCs, and interacted with the promoters of repressed naïve genes, including *Rex1*, *Essrb* and *Prdm14,* during EpiLC differentiation. In contrast, PU subunits in SEs more frequently associated with early post-epiblast markers, like *Otx2* and *Lefty1*, but also with the core pluripotency genes, *Oct4*, *Sox2* and *Nanog (Bell et al., 2020)*. These changes in DNA methylation levels within PU and DM subunits were also observed *in vivo.* In addition, *Dnmt3A*; *Dnmt3B*-DKO ESCs demonstrate a delayed exit from naïve pluripotency associated with lack of DNA methylation at the *Nanog* promoter *(Li et al., 2017)*. Overall, these studies suggest that DNA methylation might participate in the regulation of naïve pluripotency exit by repressing naïve genes through their regulatory elements **(Figure 13F)**. However, these results must be put into perspective: some studies found that *Dnmt*-TKO ESCs demonstrate partial repression of pluripotency genes and induce somatic genes to a certain extent during somatic differentiation *(Schmidt et al., 2012)*, while others on the same *Dnmt*-TKO line found normal pluripotency gene repression but inability for somatic differentiation *(Sakaue et al., 2010)*. Thus, how the lack of DNA methylation effectively affects the exit from naïve pluripotency, and especially, the acquisition of primed features, remains unclear.

**DNA methylation heterogeneity in formative cells: to be or not to be (methylated)?**

Due to its transient nature *in vivo*, the development of long-term *in vitro* cultures of formative cells or "intermediate pluripotent stem cells" is of particular interest in stem cell biology and consequently, represents a rapidly growing field of research. This would provide tremendous help in the analysis of the role played by DNA methylation in the control of early differentiation and further cell fate choices **(see Table1)**.

One way of studying the impact of DNA methylation on priming is by looking at ESCs in serum/LIF. Indeed, while being metastable, ESCs maintained in these conditions demonstrate strong heterogeneity in transcription that has been linked to variable DNA methylation levels, especially at enhancers of key pluripotency genes, like *Esrrb* and *Rex1* *(Angermueller et al., 2016; Clark et al., 2018)*. Interestingly, by using *Rex1* as a reporter for measuring 'naïvety', two populations of cells were identified in serum/LIF cultures: *Rex1*-low cells with higher methylation levels at a subset of naïve promoters and higher *Dnmt3B* expression, and *Rex1*-high cells demonstrating less methylation at promoters and associated with strong *Tet1* and *Prdm14* expression *(Singer et al., 2014)*. These two states in serum/LIF are associated with strong bursts of transcription, contributing to cell-to-cell transcriptional heterogeneity, whereas 2i medium seems to promote only *Rex1*-high populations and reduces gene expression variations. Moreover, the absence of DNA methylation in *Dnmt*-TKO cells promotes higher expression of *Nanog* and *Essrb*, compared to WT. Overall, these studies suggest that various DNA methylation levels, notably at GRE of the naïve gene network, participate in the emergence of naïve and primed sub-populations of ESCs in serum/LIF and in the transitioning between the two states. Indeed, a strong increase in DNA methylation at enhancers during priming and epiblast formation precedes differentiation in the three germ layers *(Argelaguet et al., 2019b)*. These observations were supported by *in vivo* analyses, also revealing similar DNA methylation heterogeneity at regulatory elements of pluripotency genes in the embryo *(Bell et al., 2020; Song et al., 2019)*.

Recent theoretical modelling predicted that DNA methylation oscillations during exit of naïve pluripotency could be CpG density-dependent and mainly due to a turner-over between DNMTs that start to be expressed upon exit of naïve pluripotency and TET enzyme activity *(Ginno et al., 2020; Rulands et al., 2018)*.

Notably, CpG-poor regions, like enhancers, demonstrate high DNA methylation oscillations at the formative stage in the E5.5 epiblast *in vivo* and upon priming *in vitro*. This phenomenon has been reviewed recently *(Parry et al., 2020)*. However, the exact role played by TETs in DNA methylation variations remains discussed, some arguing that TETs may play a role in a subset of targets but are not the main driver of DNA methylation heterogeneity at most regulatory elements, genome-wide **(Figure 13F)** *(Bell et al., 2020)*. Indeed, the TET enzymes seem to play an important role in driving DNA methylation heterogeneity at somatic enhancers *(Charlton et al., 2020)*. However, DNA methylation heterogeneity at the super-enhancers of pluripotency factors is more likely linked to a balance between *de novo* methylation and passive DNA demethylation *(Song et al., 2019)*. Whether TET enzymes are central players in DNA methylation variation is yet unresolved, but their role remains critical for proper development *(Dawlaty et al., 2014; Luo et al., 2020)*.

Linking this epigenetic heterogeneity with transcriptional variations could be of interest since transcriptional noise has been thought to help the cells breaking symmetry during differentiation, as observed upstream of differentiation in the hematopoietic lineage, as well as in the uncommitted epiblast *in vivo* *(Chang et al., 2008; Eldar and Elowitz, 2010; Mohammed et al., 2017)*. However, the E5.5 epiblast *in vivo* and EpiLCs *in vitro* seem more transcriptionally homogenous than ESCs or EpiSCs, which is in sharp contrast with the epigenetic heterogeneity observed at these intermediate stages *(Bleckwehl et al., 2021; Mohammed et al., 2017)*.

Overall, the progressive transcriptional repression of naïve pluripotency markers during the transition from naïve pluripotency to priming could be linked to variations in DNA methylation acquisition at distal regulatory elements, thus allowing a plastic response of ESCs to external signaling cues. However, the mechanisms by which such oscillations in DNA methylation impact transcription in general, and how they impact differentiation markers in primed EpiSCs are not entirely resolved yet.

# 4. Lineage specification

## 4.1 Germline specification: closing the loop

In a sense, any cells in which a genetic mutation will be passed on to the next generation, is a cell belonging to the germline. In this regard, from the zygote to the gametes, the germline is associated with "**cells that have the potential to give rise to cells of the next generation**" *(Nicholls and Page, 2021)*. In animals with sexual reproduction strategies, there is two main routes for the emergence of the germline and the future gametes. In many organisms, ranging from flies, to fishes and birds, the germline is determined and set aside very early in development, as it is 'preformed' by maternally inherited determinants. In contrast, mammals derive their germline from the soma later in development, following '*induction'* from the neighboring embryonic and extra-embryonic cells. From an evolutionary stand point, the induction model of germline specification is considered as the ancestral trait, while preformation has been acquired independently in different species *(Ortega-Recalde and Hore, 2019)*. Interestingly, many species in which the germline is preformed, like the zebrafish, do not demonstrate global epigenetic reprogramming in their life-cycle.

The mammalian primordial germ cells (PGCs) that arise from the soma reactivate pluripotency genes and erase DNA methylation *(Leitch and Smith, 2013)*. This strong epigenetic reprogramming in the germline, could be one component of what is known as the "Weismann Barrier". In 1892, German embryologist August Weismann developed in his book "*Germ-Plasm, a theory of heredity*", the idea of "**insulated units of heredity**", where the germline is insulated from somatic and environmental influences. In his opinion, random mutations in the germline are the only source of variation for natural selection to apply on *(Weismann, 1893)*. Although, later in his life, Weismann expressed more nuanced opinions on this point, the concept of the barrier is currently being revisited, notably in light of epigenetic regulation *(Winther, 2001)*. In this part, I will cover the emergence of the germline and the soma from pluripotent stem cells. More precisely, I will question the role of DNA methylation in controlling germline emergence, and in insulating somatic and germline fates from each other.

### 4.1.1 PGC specification and reprogramming

**Primordial germ cell induction *in vivo***

The irrefutable demonstration that mouse PGCs arise via induction was made in 1996, when Zhou and colleagues demonstrated that cells from the distal or posterior axis of the E6.5 epiblast could both be induced into PGCs, when transplanted next to the extra-embryonic ectoderm (ExE) *(Tam and Zhou, 1996)*. This seminal work proved that short range signals from extra-embryonic tissues are determinant in the specification of the germline *in vivo*, but also that all epiblast cells at E6.5 share the same potential to adopt the germline fate. It was later found that BMP4, secreted by the ExE, was an essential driver of PGC specification and mesoderm development in the posterior end of the primitive streak in the epiblast: *Bmp4* null embryos are completely devoid of PGCs, lack other mesodermal tissues and display significant developmental abnormalities *(Lawson et al., 1999)* **(Figure 14A)**. Similarly, BMP8B, secreted also by the ExE, and BMP2 signals, emitted from the adjacent extra-embryonic endoderm, also cooperate with BMP4 to induce germline induction, since both BMP8B-BMP2 null mutant mice have impaired PGC number *(Ying et al., 2000; Ying and Zhao, 2001)*. The dissection of *in vivo* signaling cascades were crucial to define conditions for *in vitro* PGCLC cultures, as many of the molecules used for *in vitro* germline derivation are in fact BMPs and cytokines supporting pluripotent stem cell cultures *(Hayashi et al., 2011)*.

Binding of BMP4 to its receptors on epiblast cells mobilizes SMAD proteins that will then initiate the transcriptional program of the germline *(Aubin et al., 2004; Saitou and Yamaji, 2010; Senft et al., 2019)*. The expression of *Fragilis* (*Ifitm3*) and *Stella* (*Dppa3*) where initially thought to first mark the nascent germline, as they are found expressed around E6.5-7.0 in response to BMP signals. Moreover, it was observed that *Fragilis*/*Dppa3*-positive cells were also repressing *Hox* genes, furthermore supporting the importance of these two genes in the separation of the germline form the somatic program *(Saitou et al., 2002; Tanaka et al., 2004)*. However, later studies discovered the essential roles played by *Blimp1*/*Prdm1* and *Prdm14* in the specification of the germline *(Ohinata et al., 2005; Yamaji et al., 2008)*. While these two genes start being expressed after exposure to BMPs at E6.25-6.5, repression of *Otx2* by BMPs is equally critical, as *Oxt2* repress *Blimp1* and *Prdm14*, hence

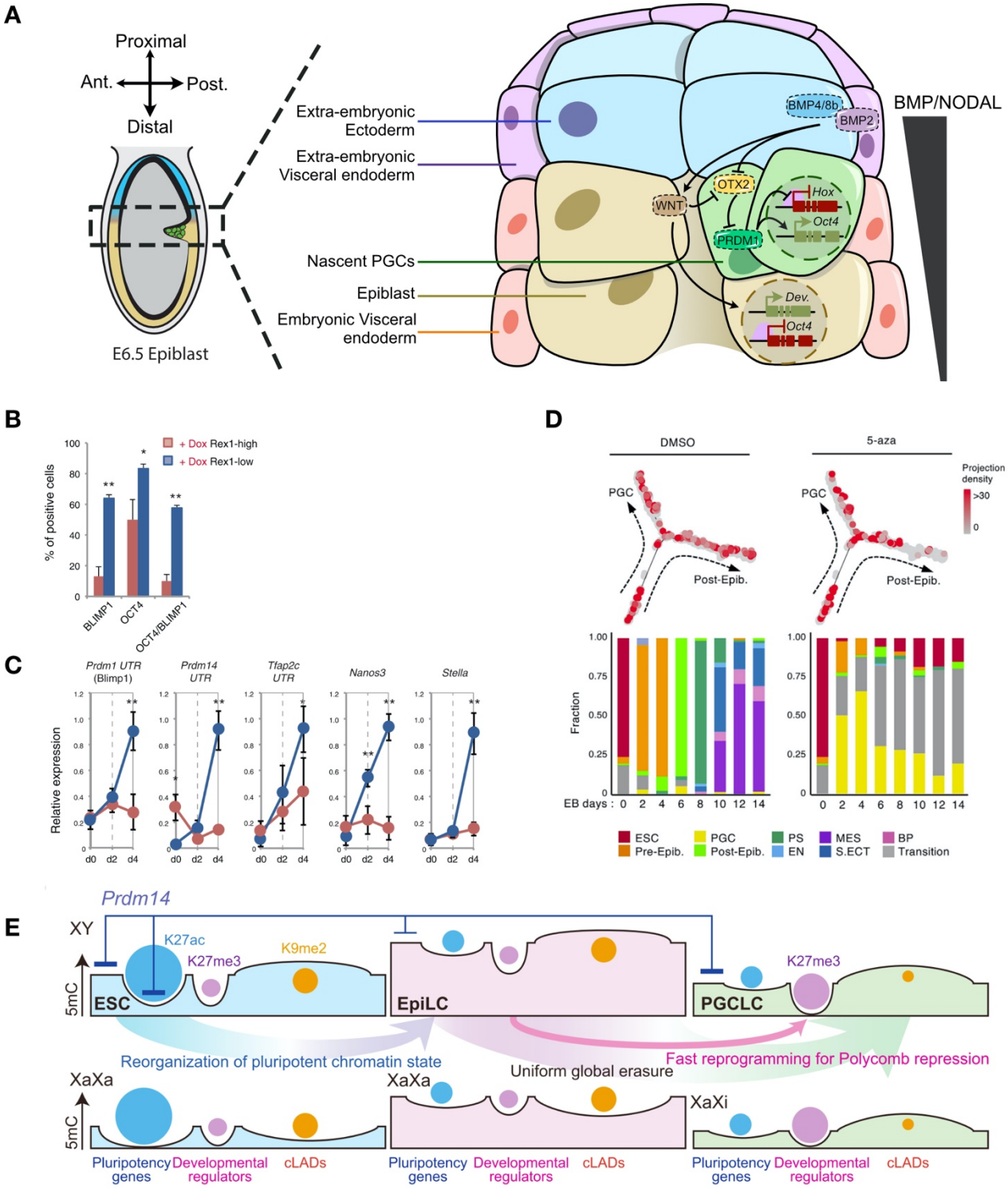



**Fig. 14 I Regulation of germline specification**

**A.** Graphical representation of the inductive signal for germline specification in the E6.5 Epiblast. BMP signals coming from the adjacent extra-embryonic tissues, in synergy with WNT, induce germline development by repressing regulators of somatic lineages like *Otx2*, *Hox* genes, and promoting epigenetic reprogramming, like acquisition of H3K27me3 on somatic genes. In other epiblast cells, gradual dilution of BMP signal turn the fates around, now allowing the emergence of somatic fate. **B, C.** Using a *Rex1* reporter to monitor for exit of naïve pluripotency, Mulas and colleagues found that only cells expressing low level of this naïve pluripotent markers, effectively undergo PGC specification (Mulas *et al*, 2017 Stem Cell Report). **D.** In contrast, Kim and colleagues found that embryoid bodies treat with DNA methylation inhibitor (5-aza) display more PGCs that arise from stem cells that have not completed their exit from naïve pluripotency (Kim *et al*, 2020 Cell Reports). **E.** Graphical representation of the epigenetic reprogramming form ESC to PGCs. Briefly, *Prdm14* drive DNA hypomethylation, while pluripotency genes lose H3K27ac and H3K27me3 is reshuffled around developmental regulators. Female ESCs more or less follow the same dynamics, but initially display one inactive X chromosome during PGC specification (Shirane *et al,* 2016 Developmental Cell).

consolidating the somatic transcriptional program (Acampora et al., 2013; Yang et al., 2014; Q.-L. Ying et al., 2003; Zhang et al., 2018). Moreover, *Blimp1*-positive cells demonstrate repression of *Hox* genes, and reactivation of pluripotency core triad like PGCs (*Oct4*, *Nanog* and *Sox2*) **(Figure 14A)**, and also repress *de novo* DNA methyltransferases *(Kurimoto et al., 2008)*.

**DNA methylation, a regulator of the germline fate?**

As mentioned above, we can note that PGCs and naïve pluripotent ESCs share many similarities. As they arise, the PGCs display DNA methylation levels similar to the epiblast stem cells from which they arise. However, they quickly start losing DNA methylation, combining passive and active DNA demethylation. Repression of DNMTs and expression of TET enzymes, lead the PGCs to reach levels of 5mC below those observed in ICM cells *(Seisenberger et al., 2012)*. Thus, specification of PGCs is of particular interest to study the functional role of DNA methylation in development as the epigenetic reprogramming they undergo is thought to be central to establish naivety and generate functional gametes for the next generation *(Bao et al., 2009; Messerschmidt et al., 2014)*. Interestingly, DNA methylation loss is also observed *in vitro* during PGCLC differentiation **(Table 2)**, although to a lesser extent than the reprogramming happening in PGCs *in vivo* *(Shirane et al., 2016)*.

Yet, it was observed early on that ESCs poorly contributed to the germline *in vitro*, even in the presence of positive inductive signals or forced expression of key PGC markers, suggesting that low genomic methylation alone is not sufficient for PGC specification *(Hayashi et al., 2011; Nakaki et al., 2013)*. Indeed, cells with the lowest expression of the naïve marker *Rex1* demonstrated higher potential for differentiation **(Figure 14B – 14C)**, notably toward the germline *(Mulas et al., 2017)*. This was furthermore supported by single cell analysis *in vitro*, supporting a model in which PGCLCs do not arise from a subpopulation of EpiLCs that keep low expression of naïve genes, despite the strong similarities between PGCLC and ESC transcriptomes *(Bleckwehl et al., 2021)*.

**Table 2. DNA methylation levels of pluripotent and germ cells *in vivo* and their cultured relatives**

| **Stage** | **Equivalent stage** | **CpG methylation (%)** | | **Ref.** |
|---|---|---|---|---|
| ***In vivo*** | ***In vitro*** | ***In vivo*** | ***In vitro*** | |
| ICM | 2i-grown ESC | 20% | 25% | Wang *et al*. 2014<br>Von Mayenn *et al.* 2016 |
| E6.5 epiblast and nascent PGC | D2 EpiLC | 71% | 75% | Shirane *et al*. 2016<br>Seisenberg *et al.* 2012 |
| / | D4 PGCLC | / | 50% | |
| E9.5 PGC | D6 PGCLC | 30% | 37% | |
| E11.5 PGC | / | 15% | / | |
| E13.5 PGC | / | 10% | / | |
| SSC/GSC | GSCLC | 81% | 82% | Ishikura *et al*. 2016 |

However, whether the dissolution of naïve pluripotency is necessary, or not, for germline induction is not entirely resolved, as two recent studies pointed a possible origin of PGCLCs from residual naïve gene network activity in some stem cells during *in vitro* EpiLC and Embryoid body differentiation *(Aramaki et al., 2021; Kim et al., 2020)*. Most notably, it was shown that treating Embryoid bodies with 5-azacytidine (an inhibitor of DNMTs) promoted germline differentiation, and DNA methylation loss was notably observed around naïve pluripotency genes **(Figure 14D)**. Therefore, the function of DNA methylation in EpiLCs could be to promote PGC differentiation by coordinating the exit from naïve pluripotency, maybe via its repressive activity at pluripotency genes.

Indeed, as described above, at the time of PGC specification, DNA methylation is concomitantly established in the epiblast, notably around key regulators of naïve pluripotency, in parallel with a loss of active histone marks. However, and in opposition with primed epiblast cells that keep high DNA methylation levels after their specification at E6.5, PGCs start erasing again DNA methylation in a two steps mechanism. First, PGCs mostly dilute epiblast-inherited DNA methylation by passive demethylation, resulting in a genome-wide loss of DNA methylation until E9.5. Then, in a second step, DNA methylation is lost at specific targets like X chromosome CGIs,

and some late demethylated genes, notably meiotic genes *(Seisenberger et al., 2012)*. Naïve pluripotency genes, like *Nanog*, *Essrb* and *Tcl1*, are upregulated starting at E11.5, as their GREs demethylation is orchestrated by PRDM14 and TET1 **(Figure 14E)**. Finally, naïve markers decline in expression in PGCs until E16.5, in opposition to late germline genes (meiotic genes) that start to be expressed at E13.5 (like *Mael*, *Stra8*), by active TET-driven DNA demethylation *(Seisenberger et al., 2012; Yamaguchi et al., 2012)*.

However, despite a drastic decline in DNA methylation levels from the emergence of PGCs until E13.5 when they reach 5% of CpG methylation, there are few transcriptomic differences across this period, once again demonstrating the uncoupling of DNA methylation status and transcription at these early stages. This has been supported recently in a study where the conditional KO of *Dnmt1* in PGCs did not affect their specification in the early epiblast, but rather induced PGC loss at later stages. This was accompanied by spurious expression of meiotic genes, clearly supporting a critical role for DNA methylation in migratory PGCs, notably in the regulation of meiotic genes *(Hargan-Calvopina et al., 2016)*. Indeed, DNA methylation is gained on the CGI promoters of meiotic genes during epiblast differentiation, and these remain resistant to DNA demethylation until the second erasure wave in PGCs which is completed at E13.5. In contrast, these late germline gene maintain high methylation at their CGIs in somatic genes *(Maatouk et al., 2006; Shirane et al., 2016)*.

The mechanisms by which DNA methylation is deposited at these late germline genes to prevent their expression during epiblast differentiation have already been covered in this Introduction and rely on the recruitment of DNMT3B by the PRC1 sub-unit containing E2F6. To what extent DNA methylation plays a critical role in the exit of naïve pluripotency and the specification of the germline at E6.5 remains an open question. However, more and more evidences highlight a function for DNA methylation in the control of meiotic germline genes in later stage of PGCs, which allows proper germline specialization after initial PGC specification.

### 4.1.2 DNA hypomethylation, a hallmark of naivety in pluripotent and germ cells

Reprogramming of the PGC epigenome results in the restoration of developmental potency of the germline, which is best illustrated by the re-activation of pluripotency markers like *Nanog* and *Stella*, and repression of the *Dnmt* genes *(Bleckwehl and Rada-Iglesias, 2019; Seisenberger et al., 2013a; Yabuta et al., 2006)*. Interestingly, the genomic DNA methylation profiles in PGCs are similar to those observed in ESCs, with the notable exception of imprinted control regions (ICRs) that lose 5mC in the germline *(Shirane et al., 2016)*. Additionally, ESCs and PGCs also share a similar metabolic state. Indeed, during naïve-to-primed transition, ESCs gradually switch from an oxidative to a glycolytic metabolism. However, as PGCs arise, the future germ cells demonstrate enhanced production of the metabolite a-ketoglutarate and return to a oxidative metabolic state, alike naïve ESCs *(Tischler et al., 2018)*. It has been recently shown that oxidative metabolism via mitochondrial STAT3 and $\alpha$-KG promotes DNA hypomethylation in ESCs **(Figure 15A)** *(Betto et al., 2021)*. Whether this applies to PGC hypomethylation is unknown.

Many studies have focused on explaining how DNA hypomethylation is mechanistically achieved in PGCs and naïve ESCs, and most of them have highlighted the central role played by PRDM14. PRDM14 is a shared zinc-finger transcription factor between ESCs and PGCs, and is critical to regulate DNA demethylation, naïve pluripotency and germline specification. *In vitro*, ectopic expression of PRDM14 in primed EpiSCs and EpiLCs promotes the expression of naïve pluripotency markers through the recruitment of the core pluripotency factor OCT4 *(Gillich et al., 2012; Okashita et al., 2016)*, whereas loss of PRDM14 in serum/LIF ESCs induces differentiation *(Grabole et al., 2013)*. Mechanistically, PRDM14 has been shown to promote DNA hypomethylation in *in vitro* ESCs by coordinating its action with the TET enzymes *(Nakaki and Saitou, 2014)*. Indeed, it has been found that PRDM14 can be co-immunoprecipitated with TET1 and TET2 and stimulates their binding at some key target sites, notably *Dppa3.* Recently, DPPA3/STELLA has been proposed as a central actor of genome-wide passive DNA demethylation by excluding UHRF1 from chromatin **(Figure 15A)**, therefore reducing DNMT1-dependent DNA methylation maintenance activity in ESCs *in vitro* (Mulholland et al., 2020a; Okashita et al., 2014;

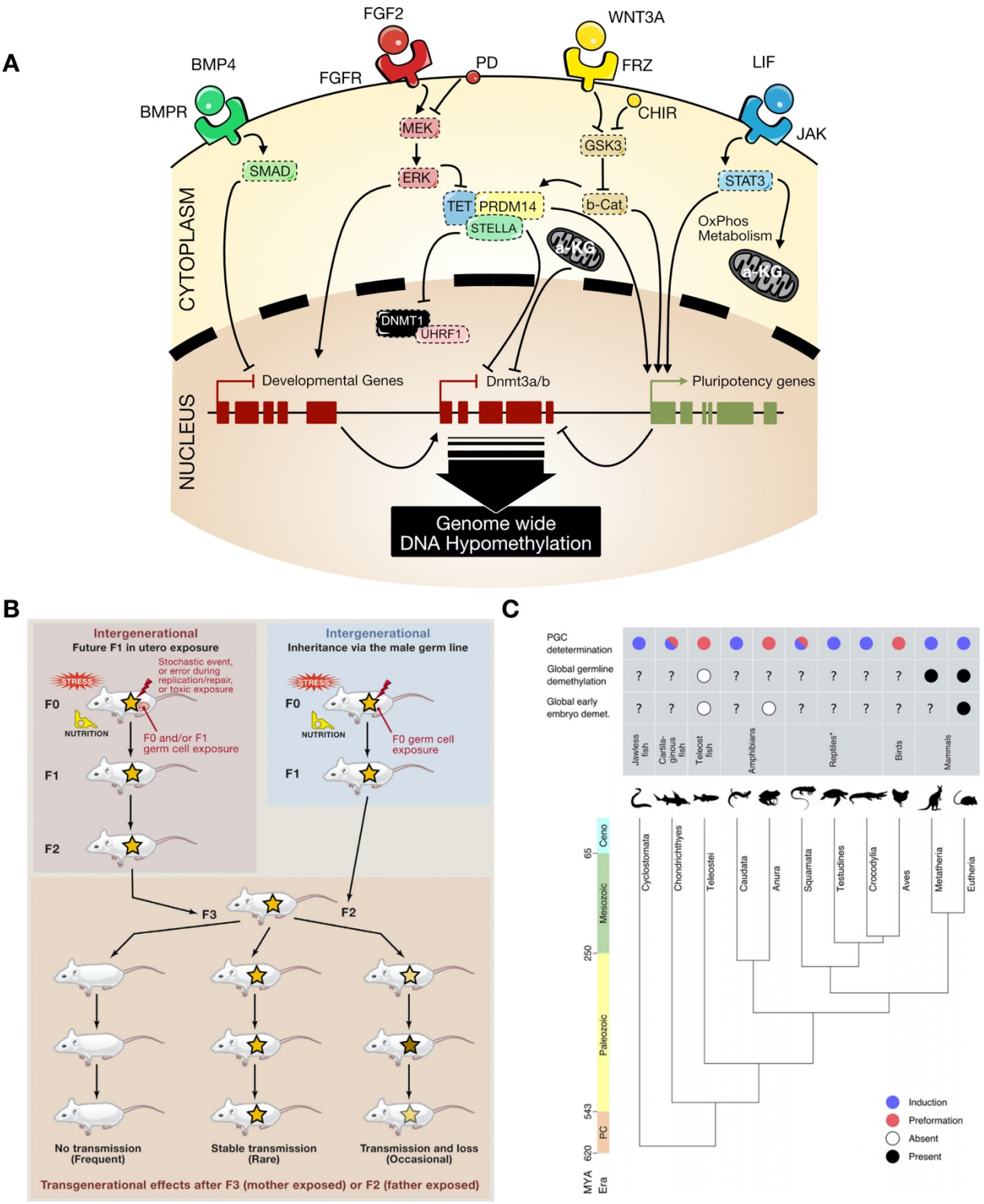
von Meyenn et al., 2016b). Similarly, *Dppa3* overexpression in hepatocellular carcinoma is also associated with DNA hypomethylation and up-regulation of pluripotency markers *(Yan et al., 2021)*. However, this mechanism is unlikely to happen in the pre-implantation embryo *in vivo*, as TET1 and TET2 are not expressed at this stage and that DPPA3 is already present, inherited from the oocyte store of proteins, and thus does not need to be activated by the joint action of PRDM14 and TET enzymes. Yet, in PGCs, similar mechanism of *Dppa3* activation by PRDM14 might exist to induce the first wave of passive DNA demethylation, since TET1 and TET2 are expressed in the PGCs from E7.5 to E12.5 *(Greenberg, 2021; Hackett et al., 2013)*. Finally, PRDM14 can repress *Dnmt3B* through the recruitment of PRC2 at its promoter and the deposition of repressive histone H3K27me3 marks *(Yamaji et al., 2013)*. Consistently, in metastable serum/LIF ESCs, clones with lower *Prdm14* levels have increased *Dnmt3B* and lower *Tcl1*, an inhibitor of the FGF signaling pathway. In summary, PRDM14 is a central regulator of reprogramming and a driver of hypomethylation, by inducing DNA methylation loss both genome-wide and at specific targets through DNMT1 exclusion, DNMT3B down-regulation and recruitment of TET demethylases. However, while *Prdm14* loss is linked with downregulation of naïve genes and upregulation of developmental genes, its activity is not necessary for the establishment of pluripotency *per se* and for PGC specification: upon *Prdm14*-KO, ESCs still express OCT4, SOX2 and NANOG pluripotency genes alike WT ESCs, despite displaying a hypermethylated genome, notably around germline gene CGI promoters. Finally PGCs can still be specified albeit in lower number *(Yamaji et al., 2013, 2008)*. Finally, while DNA hypomethylation represents a "*carte blanche*" critical for resetting the epigenetic landscape for cell identity and avoid the transmission of potential epimutations to the next generation *(Carlini et al., 2022; Seisenberger et al., 2013b)*, some loci evade DNA methylation reprogramming in both the early embryo and in PGCs, with therefore a potential for transmitting their methylation status to the next generation.

Germline epigenetic inheritance is divided in two modes: inter- and transgenerational inheritance. The distinction between these two modes is made based whether the offspring has also been exposed to the original stress. If the inheritance is mediated trough the paternal germline, effects observed in the second

generation can be qualified as transgenerational, while the third generation is to be considered when this inheritance is mediated by the maternal germline **(Figure 15B)** *(Heard and Martienssen, 2014)*. On one hand, because of the two waves of epigenetic reprogramming that take place during early development, evidence for transgenerational inheritance is scarce in mammalian species. On the other hand, mechanisms for epigenetic inheritance exist in mammals, as exemplified by the imprinted genes and IAPs that resist DNA methylation erasure, as described in this Introduction, involving many factors like ZFP57, KAP1 and DPPA3/STELLA. Reprogramming in PGCs most probably occurs because at this stage, the germline is induced from epiblast cells that are epigenetically set to give rise to somatic tissues. Hence, organisms in which the germline is preformed are also more susceptible to epigenetic inheritance and environmental perturbations, as they are not subjected to reprogramming **(Figure 15C)** *(Ortega-Recalde and Hore, 2019)*. For example, in some fishes, DNA methylation influences a switch from genetically driven to environmentally driven sex determination *(Shao et al., 2014)*. This suggests that DNA methylation can act as a vector of epigenetic inheritance in animals, including at some specific loci in mammals. We have previously mentioned studies in which impaired DNA methylation profiles in the sperm can impact offspring development independently from imprinted gene regulation. The main caveat is that we miss evidence on how impaired gametic methylation profiles can evade epigenetic reprogramming after fertilization, in the pluripotent embryo. Notably, since DNA methylation is a “transmissive” epigenetic mark—meaning that the mark is replicated and maintained during cell divisions—it is critical to demonstrate that impaired methylation profiles are indeed transmitted and maintained in the offspring *(Fitz-James and Cavalli, 2022)*.

### 4.1.3 Human and mouse germline epigenetics: so similar, yet so different

The global dynamics of DNA methylation in early development are very similar between mouse and human embryos, with the ICM and PGC cells also representing stages with the lowest DNA methylation levels genome wide *(Gkountela et al., 2015; Zhu et al., 2018)*. Moreover, transcriptomic analysis of the human germline also revealed striking similarities with mice: nascent human PGCs also reactivate




**A.** Graphical representation of the signaling involved in the regulation of DNA methylation in naïve cells, including embryonic stem cells and PGCs. Briefly, stimulation of the WNT, BMP and JAK/STAT pathways promotes the activity of pluripotency genes, inhibits *Dnmt* expression and induces *Dppa3*, *Prdm14* and *Tet* gene activity that promotes DNA hypomethylation genome-wide. In contrast, FGF signals are inhibited, and BMP repress expression of developmental genes. **B.** Distinction between Intergenerational/Transgenerational epigenetic inheritance. In the maternal line, transgenerational effects are measurable in the third generation, while in the paternal line, transgenerational effects are measurable in the second generation. In mammals the epigenetic inheritance of transgenerational traits is rare beyond the third generation, probably due to the two waves of epigenetic reprogramming that occur in early mammalian development (Heard and Martienssen, 2014 Cell). **C.** Classification of species demonstrating either induction or preformation of the germline, and DNA hypomethylation waves in the germline and early embryos. Following the most parsimonious hypothesis, germline induction is probably the most ancient mode of germline specification and is usually associated with DNA methylation reprogramming (Oretga-Recalde *et al*, 2019 Essays in Biochemistry).

pluripotency genes, express similar early germline markers, and activate meiotic genes after DNA methylation has been erased genome wide *(F. Guo et al., 2015)*. Coincidently, many factors found in mouse pluripotent stem cells and germ cells are also found in their human counterparts *(Weinberger et al., 2016)*. Whether these conserved factors also retain similar mechanisms and functions between mice and humans remains unclear. Namely, SOX17 appears to be the first inducer of the PGC program in humans—promoting the activation of the pluripotency network, as well as *Blimp1* and *Tfap2c*—in opposition to mice where this factor is not essential for PGC specification *(Irie et al., 2015; Tang et al., 2022)*. To conclude on the differences between human and mouse germline specification, it is also noteworthy that *in vitro* specification of hPGCLCs revealed that the progenitors of the germline arise from intermediate cells that reactivate naïve pluripotency program before engaging into the germline fate *(Chen et al., 2019)*, which is in contrast with the dominant model of mouse PGCLC specification that requires a complete dissolution of the naïve network.

Moreover, notable differences exist between human and mouse DNA methylation regulation in the germ and pluripotent cells. First, not only mice express a specific *de novo* DNA methylation transferase DNMT3C during male spermatogenesis, which does not exist in human *(Barau et al., 2016)*, but they also rely on the activity of the DNMT3L co-factor for DNA methylation establishment in oocyte, while DNMT3L is not present during human oogenesis *(Okae et al., 2014)*. Moreover, while the final levels of methylation in the sperm is overall similar between mice and men (~75% in humans and ~80% in mice), human oocytes display a higher average of DNA methylation than mouse oocytes (~55% versus 40%, respectively) *(H. Guo et al., 2014; Kobayashi et al., 2012; Oakes et al., 2007)*. Higher levels of DNA methylation levels in human gametes could be linked either to a stronger activity of DNMTs compared to the mouse, or simply due to longer expression of DNMTs in humans, given that spermatogenesis and oogenesis last longer in humans than in mice (in humans *vs* mice: 76 days *vs* 40 days for spermatogenesis and 150 days *vs* 21 days for oogenesis) *(Griswold, 2016; Hanna et al., 2018; Hess and de Franca, 2009)*.

Another example on how DNA methylation differently impacts human and mouse development can be found in embryonic stem cells. If cultured murine ESCs tolerate the triple knock-out of *Dnmt1*, *Dnmt3A* and *Dnmt3B* *(Tsumura et al., 2006)*,

human ESCs do not survive the loss of DNMT1. In addition, loss of the DNMTs in primed human stem cells counter-intuitively increases transcriptional repression and cell-to-cell heterogeneity *(Liao et al., 2015; Tsankov et al., 2019)*. Moreover, a transient wave of *de novo* DNA methylation deposition on the paternal genome has been found to take place during human pre-implantation development, primarily targeting some classes of young TEs *(Amouroux et al., 2016; Zhu et al., 2018)*. Interestingly, in pre-implantation embryos, both human and mouse embryos rely on the same set of transcription factors to regulate early development, although with different downstream targets and temporality *(Blakeley et al., 2015)*.

Overall, this suggests that the mouse is an interesting model for the study of human development. Global levels and dynamics of DNA methylation are similar between mouse and human development, and many of the transcription factors are conserved. Yet, differences exist between the two species, and they might have profound impact on the way we can interpret epigenetic regulation from mice to humans. For example, around 5% of histones are retained in sperm in mice, but there is twice more in human sperm. This could provide more opportunities for epigenetic inheritance to happen in humans compared to mice, since these retained histones are found close to genes involved in development *(Miller et al., 2010)*.

## 4.2 Emergence of somatic lineages—all the rest that matters

Germ layers are fundamental and distinct embryonic structures that form early in development, from which somatic tissues and organs arise. During the 19$^{th}$ century, extensive descriptions of those fundamental layers were made. Most notably, in 1849, Thomas Huxley–who specialized in comparative anatomy–demonstrated the similarities between the endoderm and ectoderm primary germ layers among vertebrates. Then, he coined the term 'mesoderm' to describe a secondary germ layer that arises by inductive signals coming from the two primary layers. Later, English zoologist Edwin Lankester, classified organisms as homoblastica (like protozoan, originating from one single cell "layer"), diploblastica (like cnidarians) and triploblastica (like vertebrates). In mammals, the ectoderm gives rise to the nervous systems and epidermis, the endoderm is the precursor of the midgut and some inner organs, the mesoderm gives rise to muscle tissues, bones and blood. A debate exists

today among scientists, whether the formation of the neural crest cells in the neural tube could be considered as a fourth germ layer in craniates, making them "quadroblastic" organisms *(Hall, 2018, 2000)*. This is supported by the fact that the neural crest cells gives rise to many different and specific cell types, like melanocytes and many sensory neurons *(Erickson et al., 2022)*.

During embryonic development, gastrulation is the developmental stage in which the germ layers are specified. During this time, pluripotency is fading away while lineage-associated gene expression is strengthened. How epigenetics can influence germ layer formation and somatic fate adoption remains obscure, although this field of research is extensively studied. During this last part of my Introduction, I would like to make a brief description of the regulation of somatic lineage emergence, highlighting first how epigenetic regulation can prime germ layer formation. Then, I will cover how the germ layers influence each other, some constituting a "default" route in somatic development, and how well somatic lineages are insulated (or not) from the germline.

### 4.2.1 Transcriptional and epigenetic control of somatic lineage specification

The emergence of the three somatic germ layers is primarily induced by external signals during gastrulation. First, ectoderm formation is driven by FGF signaling *in vivo* and in cultured ESCs, while BMP signaling must be inhibited. This induction is conserved in most vertebrates from frogs to chicken *(Delaune et al., 2005; Kuroda et al., 2005; Tropepe et al., 2001)*. Then, mesoderm and endoderm induction is mediated by a combination of NODAL, WNT and BMP signals **(Figure 16A)** *(Arnold et al., 2008; Gouti et al., 2014; Kurek et al., 2015; Rivera-Pérez and Hadjantonakis, 2015)*. In fact, the strong similarities of the signaling pathways involved in the emergence of mesoderm and endoderm opened the possibility of the existence of mesendoderm bipotential progenitors. In support of this, it is thought that the mesoderm germ layer has emerged from the endoderm 40 million years ago at a time when ancient mesodermal genes were initially involved in ecto/endoderm specification, but were progressively co-opted to support mesoderm specification *(Martindale et al., 2004; Stainier, 2005)*. Hence, bipotential mesendoderm progenitors have been observed in

cultures of ESCs and in chick and mouse embryos where cells expressing *Foxa2* have been proposed as possible mesendoderm progenitor cells *(Kimura et al., 2006; Kubo et al., 2004)*. However, the co-expression of both mesoderm and endoderm markers in some cells does not necessarily translate into bipotentiality.

In regard of the gene network regulating somatic cell fate choices, germ layer specification first requires the dissolution of the pluripotency network, especially the extinction of *Nanog*, while *Oct4* and *Sox2* are differentially expressed in mesoderm and ectoderm, respectively. Indeed, initiation of somatic differentiation is linked to a 100-fold drop in NANOG level, while *Dnmt3A* gene expression is upregulated. Moreover, variation of OCT4 and SOX2 levels seems to be directly associated with ectoderm or mesoderm driving signals, and precedes cell fate specification *(Blassberg et al., 2022; Thomson et al., 2011)*. Then SOX1 appears to act as an early inducer of ectoderm fate in mouse ESCS, as its overexpression favors a neural fate **(Figure 16A)**. But *Sox1* expression is only transient and associated to a progenitor state, since knock-down of *Sox1* precipitates expression of genes associated with a more terminally differentiated state, like *Pax6* *(Suter et al., 2009; Q. L. Ying et al., 2003)*. In mice, *Foxa2* and *Brachury (T)* are found mostly expressed in the endoderm and mesoderm layers, respectively. Their expression actually happens as early as E6.0, in cells of the primitive streak, a structure that has not completely undergone germ layer induction *(Arnold and Robertson, 2009)*, highlighting their inductive role for these lineages. Many other factors specific to mesoderm and endoderm lineages can be mentioned, and show conservation in different species, like SNAIL and TWIST that drive mesodermal fate also in *Drosophila*, and GATA factors that are also involved in epithelial to mesenchymal transition (EMT) of the endoderm and mesoderm in zebrafish *(Nowotschin et al., 2019; Technau and Scholz, 2003)*. Recently, it has been shown that BRACHURY and EOMES are mediator of meso/endoderm fate specification: in response to differentiation signals, these two factors repress both pluripotency and neuroectoderm programs, and induce chromatin opening at regulatory elements of mesodermal and endodermal genes **(Figure 16A)** *(Tosic et al., 2019)*.

Indeed, chromatin changes appear to profoundly impact germ layer emergence. Active histone marks can be critical for differentiation. For example,

embryos deficient for H3K4 methyltransferases and H3K27 acetylases arrest their embryonic development, display aberrant *Hox* gene patterning and developmental defects *(Bledau et al., 2014; Glaser et al., 2006; Yao et al., 1998)*. Indeed, H3K4me3 and H3K27ac can be found at lineage-specific genes during priming of the embryo *(Xiang et al., 2019)*.

As we already mentioned, and in sharp contrast with pluripotent stem cells, repressive chromatin marks are essential to sustain somatic differentiation. The knock-out of DNMTs, H3K27 methyltransferase EZH2 and H3K9 methyltransferase G9A all lead to embryonic lethality peri-gastrulation *(Lei et al., 1996; O'Carroll et al., 2001; Okano et al., 1999; Tachibana et al., 2002)*. In primed stem cells, lineage decisions can be made through redistribution of repressive chromatin marks and resolving bivalent states. For example, in neural crest cell differentiation, activation of a specific set of developmental genes, previously kept silent in bivalent chromatin domains, can be achieved by target dissolution of H3K27me3 and maintenance of active H3K4me3 *(Minoux et al., 2017)*. Then, H3K9me3 levels are transiently increased in mesoderm and endoderm progenitors before being removed or conserved at specific loci, in a lineage-specific manner **(Figure 16B)**, and it has been observed that H3K27me3 and H3K9me3 can act synergistically to maintain repressive heterochromatin domain in differentiated cells *(Nicetto et al., 2019)*. Similarly, changes in both DNA methylation and H3K27me3 can concur in the regulation of lineage specific genes, as demonstrated during ESC commitment toward neural fate *(Mohn et al., 2008)*. DNA methylation in neural progenitors is acquired at promoters of genes involved in the regulation of other somatic lineages that were marked by H3K27me3 in ESCs, while repressive H3K27me3 domains are found at genes required for subsequent terminal neural differentiation. This repression by H3K27me3 on terminal differentiation genes remains only transient and can be overcome by appropriate differentiation signals. This suggests that during germ layer specification, after the exit of pluripotency and at the progenitor stage, repressive epigenetic marks can transiently repress genes involved in later, terminal commitment. This has also been observed to some extent in the germline, with the precocious expression of meiotic genes in PGCs in absence of *Dnmt1* *(Hargan-Calvopina et al., 2016)*.

### 4.2.2 Cracks in the lineage specification program

**Default routes in somatic differentiation programs?**

It has been shown on several occurrences and in different model organisms, that the neural fate constitutes a default route of differentiation during early lineage specification. For example, explant cellular cultures from dissociated Xenope embryos spontaneously differentiate toward neural tissues in absence of differentiation signals *(Grunz and Tacke, 1989; Sato and Sargent, 1989)*. This phenomenon also occurs in culture of mouse ESCs, which adopt a neural stem cell identity, above all other somatic lineages, when maintained in minimal culture conditions *(Smukler et al., 2006)*. At the center of this neural default route model lies the BMP signaling pathway. Indeed, it was shown that the absence of BMPs promotes brain development in the Xenope embryos *(Reversade et al., 2005)*, which is in accordance with the notion that inhibiting BMP signals favors neural development, like we described previously. Yet, targeting BMP2/4/7 or BMP signal transducers SMAD1/5 does not promote ectodermal fate specification in mouse or in chick embryos *(Muñoz-Sanjuán and Brivanlou, 2002)*. Hence, the simplistic view of the emergence of neural fate by BMP inhibition has been furthermore complexified with the realization that requirement for BMP signals could also vary across differentiation time, like in the chick where misexpression of BMP4 only affects the expression of late neural genes *(Linker and Stern, 2004)*.

Thus, we might now consider that other factors could play a role in supporting the neural route as a default pathway. Recently, chromatin factors have been proposed to participate in promoting neural fate commitment. Using multi-omic approaches in mouse embryos, the team of Wolf Reik discovered that regulatory elements (i.e., mostly enhancers) of the ectoderm display an open chromatin state very early in development, as soon as E4.5, in contrast to mesoderm and endoderm regulatory elements **(Figure 16C)**. Ectoderm enhancers gain DNA methylation and restrict their accessibility only when cells engage in endoderm and mesoderm fates, which promotes their decommissioning *(Argelaguet et al., 2019a)*. This might prime the epiblast cells to adopt the ectoderm as a default route in early embryonic development. Interestingly, differential epigenetic regulation in lineage specification

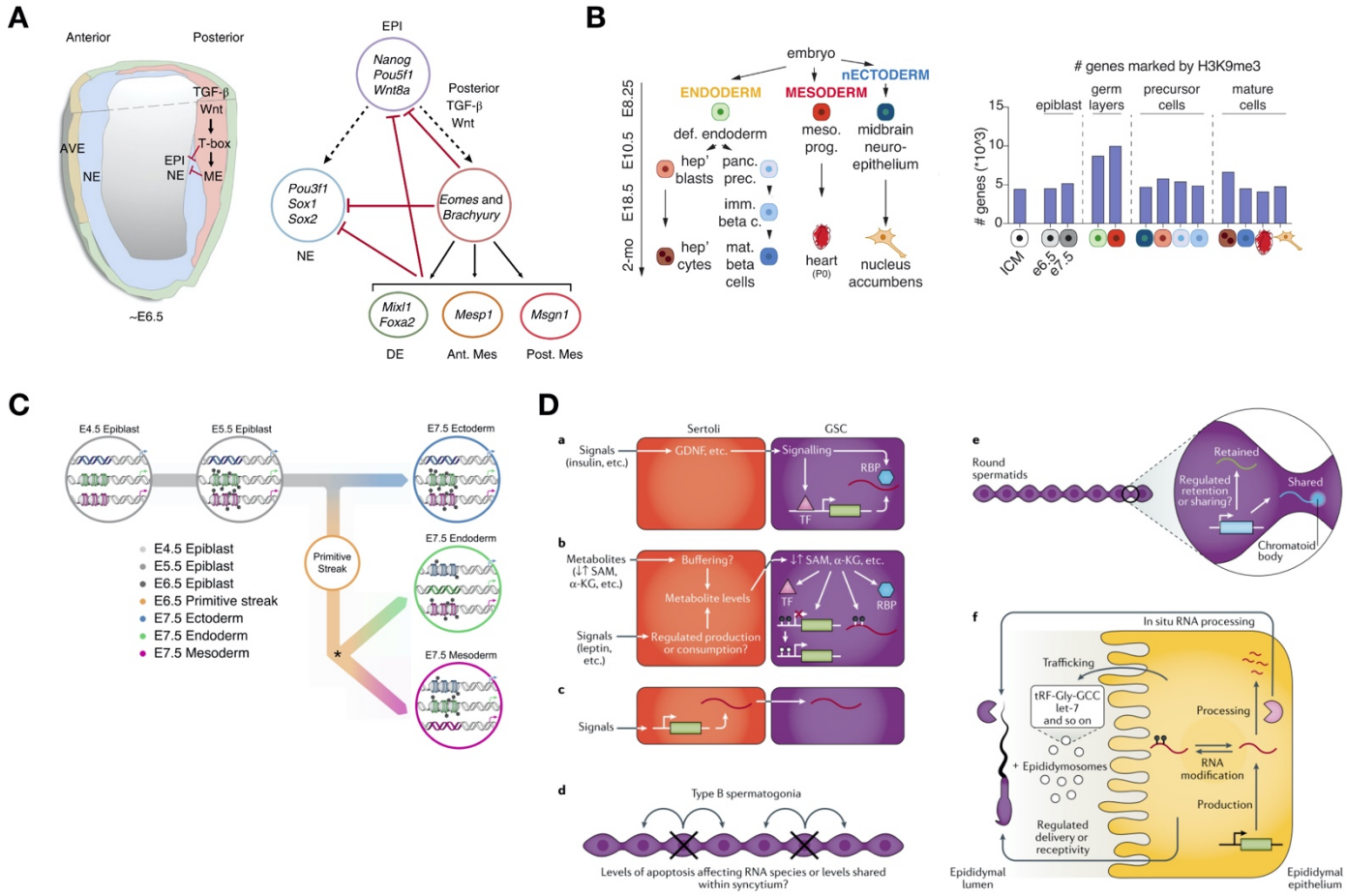



**Fig. 16 l Lineage specification and soma-to-germline communication for inheritance**

**A.** Graphical model of signaling pathways and gene networks regulating Neuroectoderm (NE), Meso/Endoderm (ME) germ layers in the E6.5 embryo. Briefly, WNT and TGF signals induce T-box factors *Eomes* and *Brachury*, that inhibits NE fate, and promote the formation of Distal endoderm (DE), and anterior/posterior Mesoderm (Tosic *et al*, 2019 Nat. Cell Biology). **B.** Graphical hierarchy of cell types during lineage specification in the gastrulating embryo. Compared to the ICM, and stem cells from the E6.5 and E7.5 embryos, endoderm and mesoderm progenitors demonstrate strong and transient enrichment in repressive H3K9me3, before acquiring more specialized somatic fates (Nicetto *et al*, 2019 Science). **C.** Graphical representation of the epigenetic priming of the neural fate. Ectoderm-associated enhancers demonstrate high accessibility and low DNA methylation in E4.5 to E5.5 embryos, while endoderm and mesoderm enhancers are only activated at the onset of gastrulation, in the primitive streak, with the removal of DNA methylation and acquisition of accessible chromatin in a lineage-specific manner (Argelaget *et al*, 2019 Nature). **D.** Schematic representation of possible soma-to-germline communication routes that could impact epigenetic inheritance. (**a, b, c**) First, neighboring somatic Sertoli cells could affect germline stem cells (GSCs) transcription factors (TF), epigenome and RNA stability through various signals like hormonal communication (glial cell- derived neurotrophic factor GDNF), RNAs or buffering for metabolites known to affect the epigenome (alpha-ketoglutarate a-KG, s-adenosyl methionine SAM). (**d, e**) Then spermatogonia and round spermatids, through cytoplasmic bridges, can translate apoptotic signals and RNAs via chromatoid bodies to other germ cells, propagating environmental cues through the syncytium. (**f**) Finally, as spermatozoa navigate through the epididymal epithelium, epididymosomes, and changes in the pH or ReDox mediated by somatic tissues, could also affect the molecular composition of the sperm and competence for fertilization (Conine and Rando 2021 Nature Reviews Genetics).

has been observed in other instances. Notably, the differentiation of human mesenchymal stromal stem cells (MSCs) into osteoblasts is driven by a set of enhancers that are already accessible in undifferentiated MSCs, while the formation of adipocytes requires a more profound rewiring of these enhancers found in MSCs *(Rauch et al., 2019)*. The default activation of the osteogenic route seems to be mediated by a specific set of transcription factors found at common active enhancers in MSCs and osteoblasts, like TEAD, ELK and HIF1 $\alpha$ . Similarly, in the endoderm layer, mid-gut endoderm cells already bear accessible enhancers that promote the differentiation into foregut cells, while intestine cells emergence depends first on the repression of foregut regulatory elements *(Banerjee et al., 2018)*. Finally, in hematopoietic lineages, disruption of the DNA methylation landscape through the deletion of TET2 or DNMT3A promotes the differentiation of hematopoietic stem cells either toward myelomonocytes or erythrocytes, respectively, notably by affecting the binding of DNA methylation-sensitive transcription factors *(Izzo et al., 2020)*.

Therefore, it is possible that asymmetric chromatin regulation of regulatory elements might bias progenitor cells to adopt some differentiation routes by default. This epigenetic priming mostly likely occurs with specific set of enhancers being more accessible than other at the progenitor stage, maybe regulated by a specific set of transcription factors. In this context, repressive epigenetic marks could play a central role in regulating lineage specification during differentiation. Indeed, such default routing could be resolved or mitigated by repressive epigenetic mechanisms like DNA methylation, or other histone marks, that would correct this priming, as it is the case for ectoderm enhancers in meso/endoderm lineages.

**Are the soma and germline really completely isolated from each other?**

The Weissman barrier postulates the complete insulation of the germline from the soma. This view still holds true in the sense that mutations affecting the somatic cells will not be transmitted to the next generation. Yet, more and more evidences now accumulate in favor of somatic tissues influencing the development of the germline, and possibly affecting the epigenome of the germ cells, and potentially, of the next generation.

First of all, it is important to highlight the critical function of reproductive somatic tissues in the support of the germline. In the case of the mammalian testis,

the somatic Sertoli cells provide critical signals that promote spermatogenesis, like GDNF (glial cell derived neurotrophic factor) and SCF (stem cell factor) signals, as well as nutrients for the germ cells *(Hai et al., 2014)* **(Figure 16E)**. This is best exemplified by the fact that *in vitro* specified PGCLCs can develop further in gametogenesis only with the support of somatic gonadal cells or by being transplanted in gonads *in vivo (Ishikura et al., 2021, 2016)*. Moreover, even after the completion of spermatogenesis, spermatozoa get capacitated in the seminiferous tubule as they get exposed to proteins, RNAs, and changes in pH and Red-Ox. These modifications of the environment are controlled by distinct segments of the somatic tissue composing the epididymis, and they affect the spermatozoa in their motility and ability to fertilize the future oocyte *(Gervasi and Visconti, 2017; Johnston et al., 2005; Sullivan et al., 2007)*. Similarly, in females, somatic cells of the granulosa that surround the oocyte extend protrusions that allows direct exchanges of proteins and metabolites less than 1kDa in size with the oocyte *(Coticchio et al., 2014)*. Hence, it has become evident that the somatic niche around the germline is critical for proper gametogenesis, and can direct germline development and capacitation by establishing direct communications with the future gametes.

The question whether this soma-to-germline communication could be translated into transmissible traits of epigenetic inheritance is under extensive scrutiny. One straight forward way for the somatic reproductive tissues to impact germline epigenetic inheritance would be through exchanges of RNAs or vesicles **(Figure 16E)**. In some studies, exposing sperm to epididymal tissues that underwent environmental insults, like ethanol or corticosterone exposure, affects sperm ncRNAs, and results in over-anxious and stress-related behavior in the offspring *(Chan et al., 2020; Rompala et al., 2020)*. Yet, evidence for a direct control of somatic tissues on the germline epigenome is still missing. Since somatic reproductive tissues also act as a buffer layer for the germline against varying levels of metabolites coming from non-reproductive somatic tissues, it is possible that if the somatic niche is subject to metabolic disfunctions, the germline could become a primary target for epigenetic impairments. While maternal disturbance of folate metabolism (an important source of methyl groups) was found to affect development of the offspring for several generations—causing tissue-specific DNA hypomethylation and epigenetic

instabilities—how this impaired metabolism in somatic cells mechanistically affects the germline is still unclear *(Padmanabhan et al., 2013)*. In addition, overstimulation of the germline by signals regulating DNA methylation, like BMP or JAK/STAT pathways, coming from neighboring somatic tissues, could also directly affect the germline DNA methylome **(Figure 16E)**. Although still very hypothetical, these potential mechanisms could become interesting new research avenues, shedding light onto how somatic cells could crack the Weismann barrier and affect germline epigenetic inheritance.

# RESULTS

# Impact of DNA methylation in cell fate decisions

My PhD project aimed at investigating the impact of DNA methylation on early cell fate decisions during mammalian development. I generated a mouse ESC line devoid of DNA methylation, by knocking out the three DNA methyltransferases that are active in the early mouse embryo (*Dnmt1*, *Dnmt3A* and *Dnmt3B*, TKO), and relied on *in vitro* differentiation protocols to recapitulate mouse early development in a petri dish.

First, I found that despite being DNA methylation-free, TKO cells still bear the capacity to adopt a neural progenitor identity, when subjected to neural differentiation. However, over long time in culture, TKO cells upregulate pluripotency genes and fail to consolidate the somatic neurogenic program. In contrast, I found that DNA methylation is completely dispensable for epiblast differentiation. The absence of DNA methylation does not impact the regulatory network of the pluripotency transitioning, allowing TKO EpiLCs to acquire functional features of epiblast cells *in vitro*, and contribute to the epiblast *in vivo* in chimeric embryos. Interestingly, during epiblast differentiation, the absence of DNA methylation seems to extend the window for germline competence.

By performing genome-wide chromatin profiling, I observed that the absence of DNA methylation is associated with the retention of high chromatin accessibility and active histone marks at putative enhancers specific to the neural and germline lineages. I found that these enhancers were targets of DNA methylation-sensitive transcription factors, whose binding motifs appeared more mobilized in TKO EpiLCs. Moreover, when differentiating TKO ESCs toward other lineages, I observed that they would preferentially adopt a neural or germline identity, suggesting that these two lineages constitute default routes in development when DNA methylation is absent.

The data were generated mostly by myself, with the assistance of Elena de la Mata, a Master student I supervised (who performed neural progenitor differentiation and transcriptomics), Aurélie Teissandier and Pierre Gestraud, bioinformatic engineers (they performed chromatin profiling analysis, and single-cell pre-processing, respectively). This work is close to completion and will be posted soon in BioRxiv while submitted to a peer-reviewed journal. It is presented to you with the tentative title:

# DNA methylation restricts default deterministic fates towards neural and germline lineages during early development

Mathieu SCHULZ-[1], Aurélie TEISSANDIER[1], Elena DE LA MATA[1], Mélanie ARMAND[1], Julian IRANZO[1], Fatima El MARJOU[1], Pierre GESTRAUD[2], Maxim V.C. GREENBERG[3], Sarah KINSTON[4], Marius WALTER[5], Berthold GÖTTGENS[4], Deborah BOURC'HIS[1]

1. Institut Curie, PSL Research University, INSERM U934, CNRS UMR3215, Paris, France
2. Institut Curie, PSL Research University, INSERM U900, MINES ParisTech, Paris, France
3. Institut Jacques Monod, Université de Paris, CNRS, Paris, FRANCE
4. Wellcome-MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK
5. Fred Hutchinson Cancer Center, Seattle, USA

Corresponding author: deborah.bourchis@curie.fr

## ABSTRACT

Dynamic DNA methylation changes accompany early mammalian development. Somatic DNA methylation patterns are acquired as embryonic stem cells (ESCs) transition from naïve to primed pluripotency during epiblast development. This transition precedes the emergence of the three somatic germ layers—which maintain high DNA methylation levels throughout life—but also of the germline, where the primordial germ cells (PGCs) undergoes genome-wide DNA demethylation right after their specification. DNA methylation is known to be dispensable for naïve pluripotency, but strictly required for development in the mouse. However, at which developmental transition DNA methylation becomes critical and whether all somatic and germline lineages equally depend on DNA methylation is not known. Using a cultured system of genetically engineered DNA methylation-free ESCs, we report here that DNA methylation is dispensable for exiting naïve pluripotency and priming. However, DNA methylation-free ESCs demonstrated skewed differentiation abilities in favor of both neural and germline fates. Based on chromatin profiling, we propose that DNA methylation warrants multi-lineage specification by decommissioning DNA methylation-sensitive NPC and PGCLC putative enhancers, tempering neural and germline fates as default routes for differentiation.

# INTRODUCTION

DNA methylation (5mC) is an epigenetic mark associated with repression of gene expression, providing control over cellular identities. Hence, DNA methylation is critical for development, as exemplified by the lethality observed in knock-out mice for DNA methyltransferases expressed in the early embryo (*Dnmt1*, *Dnmt3A* or *Dnmt3B*) (Li et al., 1992; Okano et al., 1999). DNA methylation is especially dynamic in embryonic pluripotent cells during the first days of development, coinciding with functional changes in their potency (Yu Zhang et al., 2018). Notably, hypomethylated embryonic pluripotent cells transition from a ground-state naïve pluripotency state to somatic primed pluripotency state as the epiblast is formed and somatic DNA methylation patterns are acquired genome-wide (Morgani et al., 2017; Nichols and Smith, 2009). During naive-to-primed transitioning, embryonic pluripotent cells go through a transient formative state at E5.5-6.5, which represents the high point for germline specification (Hayashi et al., 2011; Kinoshita and Smith, 2018; Smith, 2017). Moreover, following their emergence from the methylated epiblast (Lawson et al., 1999; Ohinata et al., 2005; Yamaji et al., 2008), primordial germ cells (PGCs) undergo genome-wide DNA methylation erasure as part of their specification program (Seisenberger et al., 2012).

Cultures of embryonic stem cells (ESCs) can recapitulate early development and the associated DNA methylation dynamics. These were instrumental to reveal that DNA methylation was dispensable for naive pluripotency, as *Dnmt1*, *Dnmt3A* and *Dnmt3B* triple-KO (*Dnmt*-TKO) ESCs can be maintained in naive-prone culture medium (Tsumura et al., 2006). While the dissolution of pluripotency genes is necessary for cell fate induction (Chambers et al., 2007; Hayashi et al., 2011; Mulas et al., 2017), DNA methylation is strongly acquired at regulatory elements of naive genes during epiblast differentiation (Bell et al., 2020; Kalkan et al., 2017), suggesting a possible role in consolidating their repression and allowing proper differentiation (Schmidt et al., 2012). However, there is contrasting evidence as to whether DNA methylation participates to the induction of cell identities (Bell et al., 2020; Rulands et al., 2018; Sakaue et al., 2010; Schmidt et al., 2012). DNA methylation is especially dynamic at enhancers and recent observations revealed heterogeneous DNA methylation levels at lineage-associated enhancers in naïve and primed stem cells *in*

*vitro*, which could regulate and consolidate cell fate decisions (Bleckwehl et al., 2021; Charlton et al., 2020; Choi et al., 2016). Notably, it has recently been shown that ectoderm enhancers are hypomethylated and accessible early in development, supporting the idea of a "default" neural route of development (Argelaguet et al., 2019a). Mechanistically, some transcription factors (TFs) are known to be DNA methylation-sensitive (Zhu et al., 2016), and could participate in promoting deterministic epigenetic profiles, priming the cells to adopt specific fates.

Whilst early development represents an ideal developmental window to study the interplay between DNA methylation and cell fate decisions, the impact of DNA methylation on lineage commitment remains poorly understood. Most notably, functional characterization of *Dnmt*-TKO ESCs during priming and early lineage induction is still missing, and their ability to adopt a germline fate is poorly investigated. Here, we showed that DNA methylation is dispensable for ESCs to adopt transcriptional and functional features associated with priming *in vitro* and *in vivo*, and is not required for inducing neural and germline lineages. More saliently, the absence of DNA methylation provides a temporally extended opportunity to give rise to the germline. We propose that DNA methylation and DNA-methylation sensitive-factors are required during priming at a subset of DNA methylation sensitive enhancers to restrict neural and germline as default routes of differentiation.

# RESULTS

**DNA methylation is dispensable for neural induction, but critical for consolidation of neurogenesis**

Various studies yielded contrasting results regarding the impact of the lack of DNA methylation on somatic differentiation (Sakaue et al., 2010; Schmidt et al., 2012), although using the same originally derived DNA methylation-free *Dnmt1*, *Dnmt3A* and *Dnmt3B* triple-KO (TKO) embryonic stem cell (ESC) line (Tsumura et al., 2006). To address this discrepancy—and to mitigate potential genetic and epigenetic drift from this older TKO cell line—we engineered a new *Dnmt*-TKO line from low passage ESCs (E14 background) by addressing CRISPR-Cas9 to the sequence coding the catalytic loop of *Dnmt1*, *Dnmt3A*, and *Dnmt3B* **(Extended Figure 1A)**. This TKO cell line demonstrated homozygous frame-shift deletions at the *Dnmt* genes, resulting in complete absence of DNMT proteins and of 5mC methylation **(Extended Figure 1B-D)**. As previously shown (Tsumura et al., 2006), E14 TKO ESCs also expressed pluripotency markers at normal levels under naïve pluripotency-sustaining conditions **(Extended Figure 1E)**.

To test the potential of TKO cells for the adoption of a somatic fate, we first performed undirected differentiation into the three somatic germ layers using 3D embryoid body (EB) formation. While WT EBs continuously grew during the eight days of differentiation (going from 80µm to 420µm in diameter) and formed cystic structures—a read-out of proper differentiation (Koike et al., 2007; Leaky et al., 1999)—TKO EBs demonstrated extended cell death around day 3 (D3), reaching a maximum of 125µm diameter size **(Figure 1A-B)**. This growth arrest could not be rescued by increasing initial cell density **(Extended Figure 1F)**. Therefore, DNA methylation appears critical for EB formation and in sustaining multi-differentiation into the three somatic germ layers.

However, we thought that this approach may bear confounding variables as we needed to assume that mutant cells are able to *i)* form aggregates, and *ii)* autonomously coordinate towards the three germ layers. Hence, we performed a directed neural progenitor cell (NPC) differentiation to probe the effect of DNA methylation in adopting a specific fate (Mulas et al., 2019; Q. L. Ying et al., 2003). During NPC differentiation, WT cells demonstrated exponential growth until D6, and

acquisition of neural-like morphological features like “axon-like protrusions” and neural rosettes **(Figure 1C-D)**. Unexpectedly, TKO cells survived and formed protrusions at D6, despite a lower growth rate that was associated with a drop in viability from 92% to 66% at D8 **(Figure 1C-E)**. To gain deeper insight into the identity of TKO NPCs, we performed bulk RNA sequencing at D0, D2, D4, D6 and D8 of NPC differentiation. Multivariate analysis revealed that cells were ordered first by time (PC1, 58.1%), then by genotype (PC2, 12.5%) due to the strong up-regulation of germline genes in absence of DNA methylation **(Figure 1F and Extended Figure 1G)**. Accordingly, between D4 and D8, gametogenesis-associated GO terms were significantly enriched in the TKO background **(Figure 1G)**. This is expected as DNA methylation is directly involved in the regulation of germline gene promoters in somatic cells (Borgel et al., 2010; Dahlet et al., 2021; Velasco et al., 2010). In addition, we found strong enrichment of GO terms associated with neurogenesis at D2 in DNA methylation-free cells, but most of these terms were lost at D8, as neurogenesis-associated genes became drastically down-regulated (pvalue cutoff=0.05) **(Figure 1G-H)**. It is noteworthy that both WT and TKO ESCs demonstrated synchronous exit of pluripotency, after having transiently up-regulated primed pluripotency genes at D2. This suggests that during directed somatic differentiation, ESCs may still undergo sequential naïve-to-primed pluripotency transitioning, before inducing neurogenesis. However, we found that the pluripotency network became abnormally reactivated in late TKO NPCs at D8 **(Figure 1I)**. While we cannot discriminate between true developmental reversion or negative selection of TKO NPCs, we show that DNA methylation is necessary for the consolidation of neural identity in the long term, by securing the repression of pluripotency program. In sum, in contrast to previous reports (Schmidt et al., 2012), the absence of DNA methylation did not transcriptionally affect pluripotency exit and somatic fate induction in our newly generated DNA methylation-free TKO ESCs.

**Naïve-to-primed pluripotency transition is disconnected from DNA methylation**

As DNA methylation-free TKO NPCs ultimately fail to consolidate pluripotency gene repression at D6 and D8, we questioned the role of DNA methylation in pluripotency transitioning. To this end, we performed Epiblast-like cell (EpiLC) differentiation,

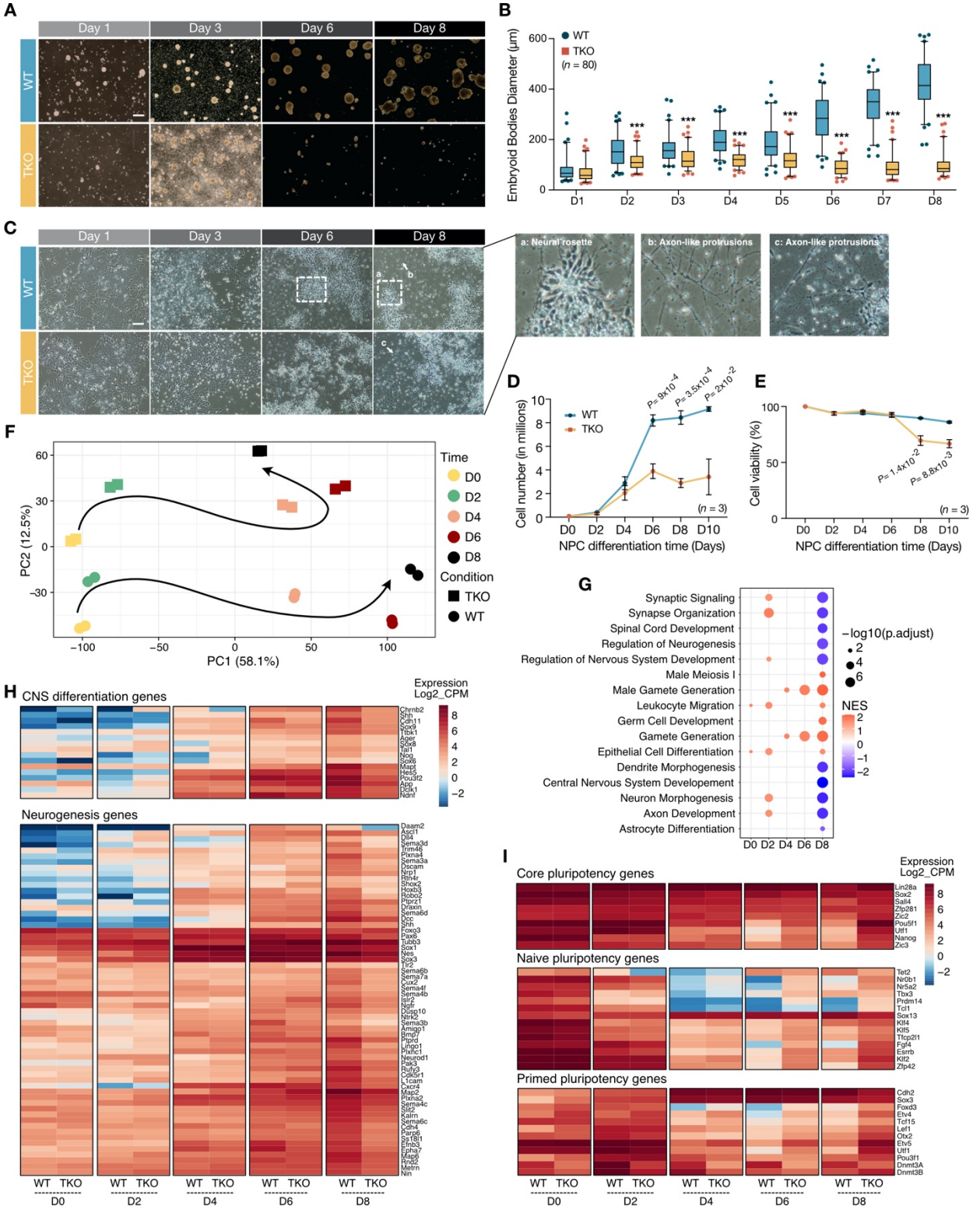



**Fig. 1 I DNA-methylation-free TKO ESCs show proper neural fate induction but fail to consolidate long-term repression of pluripotency. A,** Brightfield images of embryoid bodies (EBies) generated from WT (blue) and TKO (orange) ESCs cultured in serum for 8 days. Scale: 500$\mu$m. **B,** Boxplot representation of EBies diameter across differentiation. Boxes represent 5-95 confidence interval, *n*= 80 EBies counted/genotype/timepoint (two-tailed unpaired student t-test). **C,** Brightfield images of neural progenitor cells (NPCs). Scale: 200$\mu$m (NPCs) generated from WT and TKO ESCs cultured in N2B27 for 8 days. White arrows: protrusion; white dotted square: neural rosettes. **D, E,** Neural differentiation (D) Growth curve and (E) Trypan blue-based viability curve for WT and TKO cells. Data are mean ± SD in biological triplicate (two-tailed unpaired student t-test). **F,** Principal component (PCA) analysis from bulk RNA-seq of WT (circles) and TKO (squares) biological duplicates during NPC differentiation. Axes indicate two main components for variability, the black arrow represents hypothetical trajectory. **G,** Gene Ontology (GO) analysis showing set enrichment score (NES) and -log10(p.adj) in TKO over WT cells during NPC differentiation, pvalue cutoff=0.05. **H, I,** Heatmaps of (H) pluripotency and (I) neurogenesis markers showing normalized Log2 CPM counts during NPC differentiation. Expression level was obtained for each condition from the average between biological duplicates.

starting from WT and TKO ESCs cultured in 2i/LIF. The medium was also supplemented with Vitamin C, to promote hypomethylation in naïve ESCs (Blaschke et al., 2013; Gao et al., 2013). This *in vitro* model recapitulates genome-wide acquisition of *de novo* DNA methylation **(Figure 2A-B)** along naïve-to-primed pluripotency transitioning, with a transient "formative" phase reported to occur at around 40-48h of differentiation (Buecker et al., 2014; Morgani et al., 2017; Shirane et al., 2016; Smith, 2017). Upon EpiLC differentiation, WT and TKO cells displayed similar morphology and growth rate, and both underwent strong negative selection with apoptosis rate rising up to 65.8% and 75% at 96h for WT and TKO cells, respectively **(Figure 2C-D and Extended Figure 2A)**. Such massive epiblast apoptosis has been previously observed *in vitro* and *in vivo* and associated with negative selection against unfit primed EpiLCs (Bedzhov and Zernicka-Goetz, 2014; Sancho et al., 2013). Altogether, this suggests that despite being established shortly after EpiLC induction, DNA methylation is not necessary for EpiLC viability. To better capture the identity of TKO EpiLCs, we assessed their functional features. First, we performed a dual luciferase assay probing the switch of enhancers of *Oct4*—whose expression relies on a distal enhancer in naïve stem cells and a proximal enhancer in primed stem cells (Choi et al., 2016)—as a read-out for correct priming. This assay revealed proper activation of the *Oct4* proximal enhancer in both D4 WT and TKO EpiLCs, compared to D0 naïve ESCs **(Figure 2E)**. Next, we tested the ability of TKO EpiLCs to be reverted into naïve ESC colonies (rESCs) upon culture in 2i-based medium. The absence of DNA methylation did not provide higher reversion ability since both WT and TKO yielded similar and progressively decreasing numbers of reverted alkaline phosphatase-positive colonies, from around 2000 to 500 at D2 and D7 EpiLCs, respectively **(Figure 2F and Extended Figure 2B)**. Finally, we wanted to validate the ability of TKO ESCs to contribute to the epiblast *in vivo*. For this, we generated WT and TKO ESC lines expressing a constitutive H2B::tdTOMATO reporter to trace them after aggregation with host morulae **(Extended Figure 3A-C)**. We found integration rates in the blastocyst inner cell mass (ICM) to be high for both WT and TKO H2B::tdTOMATO ESCs (84.7% and 91.7%, respectively), which supports the dispensability of DNA methylation in naïve pluripotent stem cells **(Extended Figure 3D-E)** (Sakaue et al., 2010). After a week of development in pseudo-pregnant females,

we harvested chimeric embryos and checked cellular contribution by performing immunofluorescence against AP2y to discriminate extra-embryonic tissues. We observed that at E6.5, both WT and DNA methylation-free TKO reporter cells contributed the embryonic epiblast **(Figure 2G)**, which is consistent with our *in vitro* findings. Taken together, these results demonstrated that DNA methylation is dispensable for the acquisition of primed epiblastic features, both in vitro and in vivo.

Using bulk RNA-seq, we then tested whether the absence of DNA methylation impacted the pluripotency priming gene network. PCA highlighted transcriptional similarity between WT and TKO EpiLCs, especially at D0 and D2 **(Figure 2H and Extended Figure 2D)**. Indeed, when looking at the expression of major pluripotency markers across differentiation times, no striking difference could be detected between the two cell lines **(Figure 2I and Extended Figure 2C)**. More specifically, naïve pluripotency markers were properly repressed and primed markers normally induced in TKO EpiLCs. However, stronger divergence was observed at D4 and D7 EpiLCs, as TKO cells demonstrated significant up-regulation of 1406 and 1967 genes, respectively (|Log2FC|>1 and FDR<5%). Most notably, germline genes composed a significant part of up-regulated genes in late TKO EpiLCs, which was associated with an enrichment of GO terms linked to gametogenesis and reproductive processes **(Figure 3A-C)**. Corroborating the fact that germline genes acquire DNA methylation at their promoter during EpiLC differentiation, we indeed observed around 581 genes that are normally targeted by promoter DNA methylation to be up-regulated in TKO D7 EpiLCs, among which 103 are known germline genes (Mochizuki et al., 2021) **(Extended Figure 2E-F)**. Finally, some families of DNA methylation-repressed transposable elements like IAP, LINE-1 and MERVL (|Log2FC|>1) were also up-regulated in TKO EpiLCs compared to *WT* **(Extended Figure 2G)**. Overall, DNA methylation played no apparent role in the exit of naïve pluripotency and in the regulation of the primed pluripotency gene network. Taken together, these results indicate that DNA methylation is largely dispensable for priming and epiblast development, *in vitro* and *in vivo*. However, it remains unclear how critical DNA methylation is for the induction of the germline.

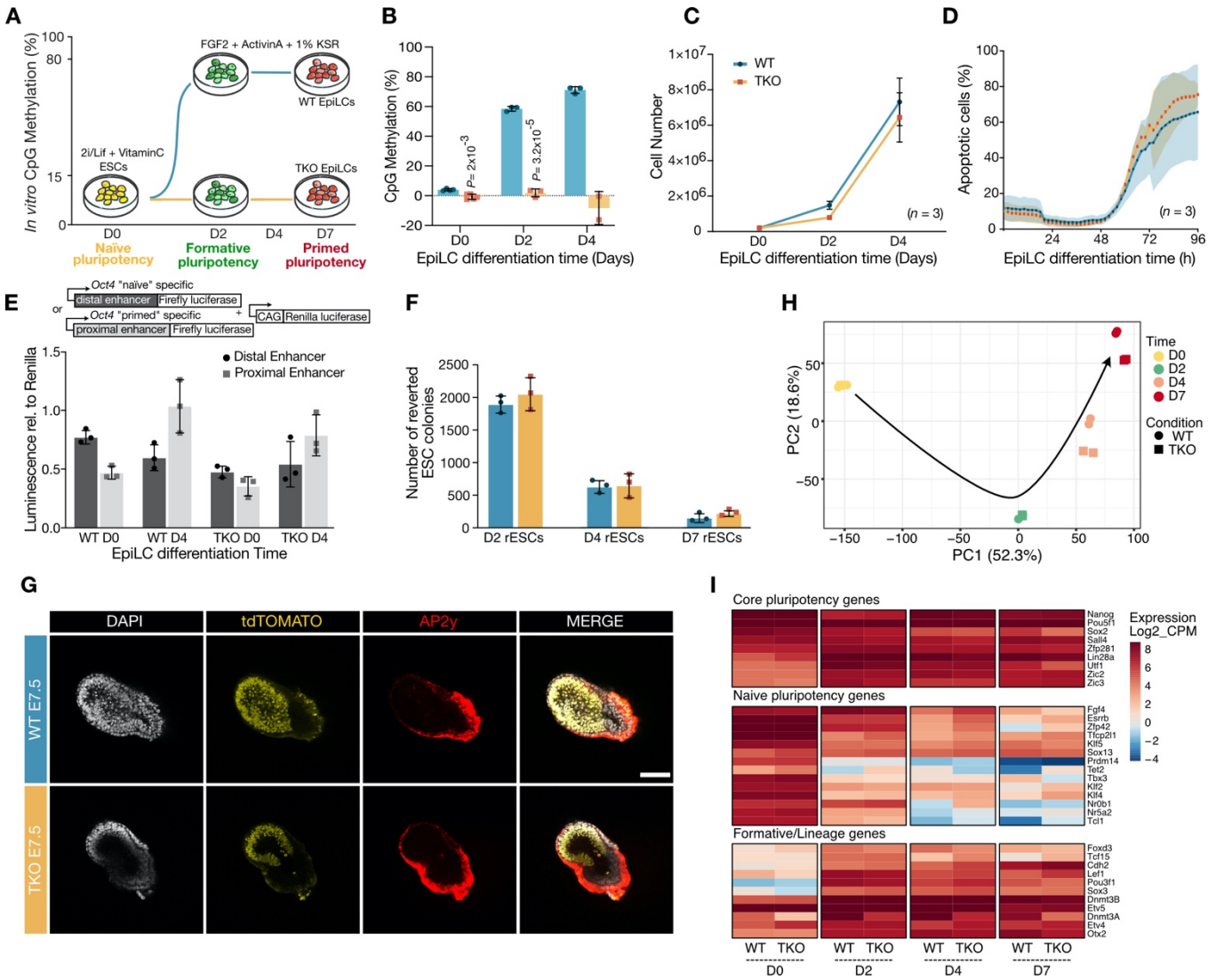



**Fig. 2 I TKO ESCs are competent for naïve-to-primed pluripotency transitioning. A,** Schematic representation of Epiblast-like cell (EpiLC) differentiation. Axes represent genomic methylation levels and key pluripotency states. **B,** Barplot representing CpG methylation measured by LUMA assay during WT and TKO EpiLC differentiation. Data shown are mean ± SD from biological replicates (*n*= 3 or 2) (two-tailed unpaired student t-test). **C,** Growth curve of WT (blue) and TKO (orange) EpiLC differentiation. Data shown are from biological triplicates with mean ± SD. **D,** Apoptosis rate measured by live staining of Caspase 3/7 on Incucyte system of WT and TKO cells. Data shown are from biological triplicates with mean ± SD. **E,** Barplot representing Luciferase activity from the distal (dark grey) and proximal (light grey) enhancer of *Oct4* in WT and TKO ESCs and D4 EpiLCs. Data shown are mean ± SD from biological triplicates with mean ± SD. **F,** Barplot displaying numbers of reverted ESC colonies by alkaline phosphatase staining of WT and TKO EpiLCs placed in 2i+VitC medium at different days of EpiLC differentiation. Data shown are from biological triplicates with mean ± SD. **G,** Confocal images of chimeric embryos at E7.5 obtained by morula aggregation with WT or TKO H2B::tdTOMATO-transgenic ESCs . Embryos were stained for DAPI (gray), and AP2Y (Red) to assess contribution to the epiblast, scale bar 100$\mu$m. **H,** Principal component analysis (PCA) from bulk RNA-seq of WT (circles) and TKO (squares) EpiLCs from D0 to D7 of differentiation in biological duplicates. Axes indicate two main components for variability, the black arrow represents hypothetical trajectory. **I,** Heatmaps of pluripotency markers showing normalized Log2 CPM counts during EpiLC differentiation. Expression is obtained for each condition from the average between biological duplicates.

**DNA methylation restricts the temporal window of germline induction**

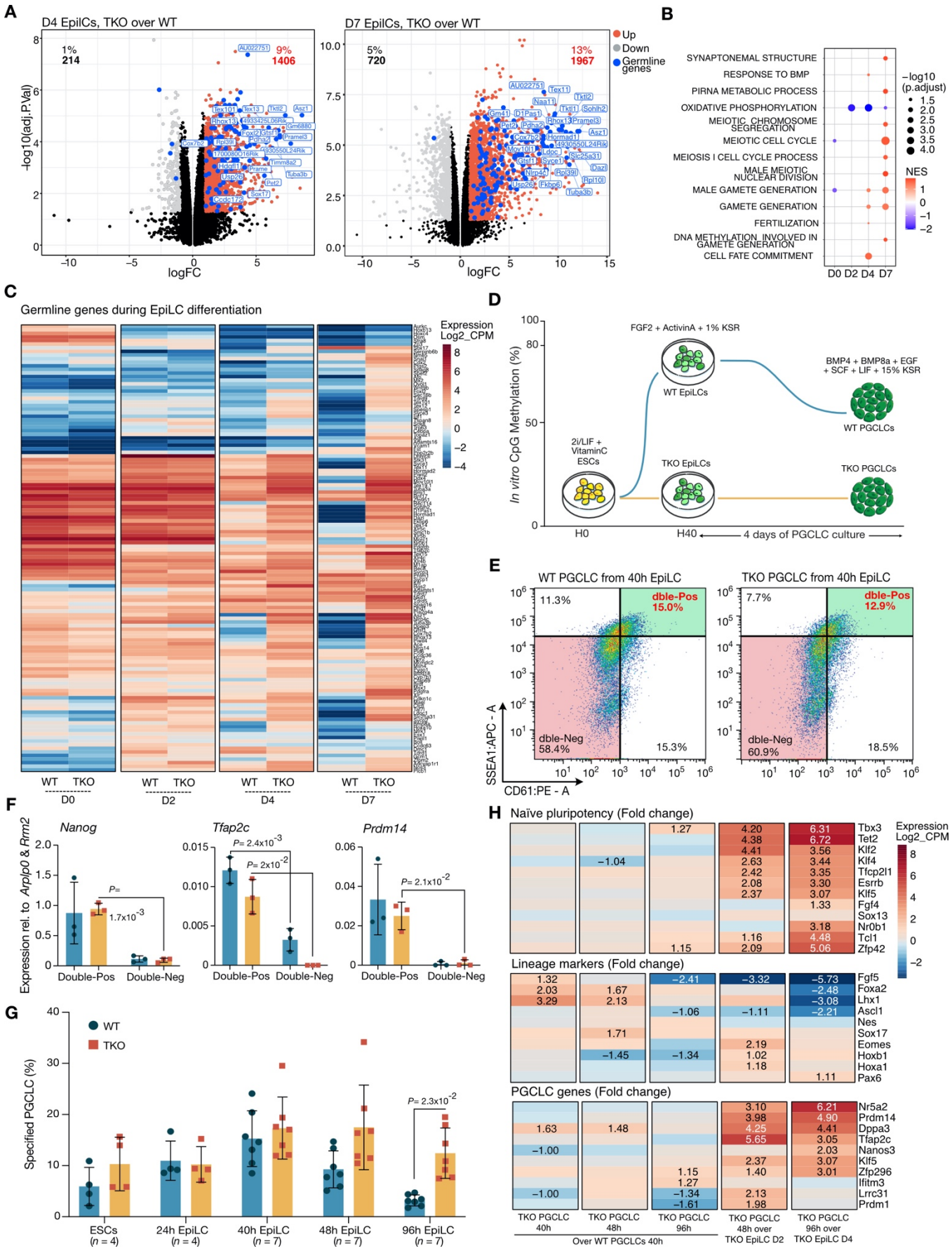
*As our* prior results highlighted that DNA methylation is mostly involved in silencing germline genes during EpiLC differentiation, we hypothesized that ESCs devoid of DNA methylation could be biased towards adopting a germline fate. This hypothesis is further reinforced by the fact that primordial germ cells (PGCs) acquire an hypomethylated genome as part of their specification program and re-express many genes that are typical of naïve stem cells (Bleckwehl and Rada-Iglesias, 2019; Seisenberger et al., 2012; Shirane et al., 2016; von Meyenn et al., 2016a). Accordingly, PGCs are sometimes referred to as « naïve-like », because of this retention of naïve pluripotency markers (Leitch and Smith, 2013).

We took advantage of PGC-like cell (PGCLC) differentiation to induce *in vitro* a germline fate in DNA methylation-free ESCs **(Figure 3D)**. As this protocol requires cells to form 3D aggregates, we first checked the ability of TKO cells to properly survive this protocol by inducing PGCLC fate after 40h EpiLC differentiation, a time point that was previously identified as optimal for germline induction (Gruhn and Günesdogan, 2021; Hayashi et al., 2011). In sharp contrast with our previous observations during embryoid body differentiation, TKO EpiLCs were able to form and grow as aggregates for at least 4 days, being undistinguishable from WT cells **(Extended Figure 4B)**. Although unexpected, we suppose this discrepancy might reflect different adhesion molecule requirement under somatic and germline-driving conditions (Leitch et al., 2014; Ohtsuka et al., 2012). Specified PGCLCs were analyzed by FACS using the SSEA1 and CD61 surface markers **(Figure 3E and Extended Figure 4A)**. Both WT and TKO cells gave rise to double-positive live PGCLCs, with slightly higher efficiency when DNA methylation is absent (14.85% for WT and 18.31% for TKO) **(Figure 3E-G)**. PGCLCs were sorted to perform RT-qPCR validation on known early PGC markers and compared with double-negative cell populations. Using RT-qPCR, we validated the expression of PGCLC markers (*Nanog*, *Tfap2c* and *Prdm14*) in double-positive compared to double-negative sorted cells, and similarly in WT and TKO **(Figure 3F and Extended Figure 4C)**. This indicates that the absence of DNA methylation does not compromise germline specification *in vitro*.

Next, we induced PGCLCs across different times of EpiLC differentiation, to test whether the competence timing of germline specification was conditioned by

DNA methylation during epiblast progression. As reported, naïve WT ESCs were poorly competent to give rise to PGCLCs (6% in average) (Mulas et al., 2017), and progressively gained germline specification capacity, reaching a maximum after 40h of EpiLC differentiation. Then, this ability quickly declined in later EpiLCs (48h and 96h) as WT cells progress toward somatic priming (3.1% of specified WT PGCLCs after 96h of EpiLC differentiation) **(Figure 3G)**. While TKO cells demonstrated similar rates of PGCLC induction in early EpiLCs, they noticeably maintained increased ability germline competence beyond 40h. More specifically, after 96h, TKO EpiLCs were four times more efficient for PGCLC specification than WT (12.4% of specified TKO PGCLCs) **(Figure 3G and Extended Figure 4D)**. Taken together, these results suggest that while 40h of EpiLC differentiation normally represents the tipping point for PGCLC specification *in vitro*, the absence of DNA methylation provides longer opportunities to adopt the germline fate.

We next investigated whether the lack of DNA methylation impacted the PGCLC transcriptional program, upon normal and late EpiLC induction. We observed that all WT and TKO PGCLCs were ordered closely together in the PCA analysis, demonstrating that WT and DNA methylation-free PGCLCs were globally similar transcriptomically, and distinct from EpiLCs **(Extended Figure 5A-B)**. Furthermore, TKO PGCLCs after 40h, 48h or even 96h of EpiLC differentiation showed very little variations in the expression of pluripotency and PGC markers compared to WT PGCLCs generated after 40h of EpiLC differentiation **(Figure 3H)**. Finally, comparing late induced TKO PGCLCs and their corresponding D2 and D4 EpiLC counterparts demonstrated expected differences. Notably, we identified 2905 down- and 2871 up-regulated genes between 48h TKO PGCLCs and D2 TKO EpiLCs, and 3506 down- and 3017 up-regulated genes between 96h TKO PGCLCs and D4 TKO EpiLCs (|Log2FC|>1 and FDR<5%) **(Extended Figure 5C)**. Among them, many early PGC markers and naïve pluripotency genes were up-regulated in late TKO PGCLCs, while early lineage-associated genes were mostly down-regulated **(Figure 3H)**. This further implies correct specification of PGCLCs from late TKO EpiLCs. Taken together, these results indicate that DNA methylation is dispensable for the transcriptional regulation of PGCLC specification. Moreover, its absence rather extends the time-window for germline fate adoption, while WT primed cells rapidly lose that ability.


A
D4 EpiLCs, TKO over WT
1%
214
9%
1406
-log10(adj.P.Val)
logFC
D7 EpiLCs, TKO over WT
5%
720
13%
1967
Up
Down
Germline genes
B
SYNAPTONEMAL STRUCTURE
RESPONSE TO BMP
PIRNA METABOLIC PROCESS
OXIDATIVE PHOSPHORYLATION
MEIOTIC CHROMOSOME SEGREGATION
MEIOTIC CELL CYCLE
MEIOSIS I CELL CYCLE PROCESS
MALE MEIOTIC NUCLEAR DIVISION
MALE GAMETE GENERATION
GAMETE GENERATION
FERTILIZATION
DNA METHYLATION INVOLVED IN GAMETE GENERATION
CELL FATE COMMITMENT
D0 D2 D4 D7
-log10 (p.adjust)
NES
C
Germline genes during EpiLC differentiation
Expression Log2_CPM
WT TKO
D0
D2
D4
D7
D
In vitro CpG Methylation (%)
FGF2 + ActivinA + 1% KSR
WT EpiLCs
BMP4 + BMP8a + EGF + SCF + LIF + 15% KSR
WT PGCLCs
2i/LIF + VitaminC ESCs
TKO EpiLCs
TKO PGCLCs
H0
H40
4 days of PGCLC culture
E
WT PGCLC from 40h EpiLC
TKO PGCLC from 40h EpiLC
11.3%
dble-Pos 15.0%
dble-Neg 58.4%
15.3%
7.7%
dble-Pos 12.9%
dble-Neg 60.9%
18.5%
SSEA1:APC - A
CD61:PE - A
F
Nanog
Tfap2c
Prdm14
Expression rel. to Arplp0 & Rrm2
P= 1.7x10-3
P= 2.4x10-3
P= 2x10-2
P= 2.1x10-2
Double-Pos
Double-Neg
G
WT
TKO
Specified PGCLC (%)
P= 2.3x10-2
ESCs (n = 4)
24h EpiLC (n = 4)
40h EpiLC (n = 7)
48h EpiLC (n = 7)
96h EpiLC (n = 7)
H
Naïve pluripotency (Fold change)
Tbx3
Tet2
Klf2
Klf4
Tfcp2l1
Esrrb
Klf5
Fgf4
Sox13
Nr0b1
Tcl1
Zfp42
Expression Log2_CPM
Lineage markers (Fold change)
Fgf5
Foxa2
Lhx1
Ascl1
Nes
Sox17
Eomes
Hoxb1
Hoxa1
Pax6
PGCLC genes (Fold change)
Nr5a2
Prdm14
Dppa3
Tfap2c
Nanos3
Klf5
Zfp296
Ifitm3
Lrrc31
Prdm1
TKO PGCLC 40h
TKO PGCLC 48h
TKO PGCLC 96h
TKO PGCLC 48h over TKO EpiLC D2
TKO PGCLC 96h over TKO EpiLC D4
Over WT PGCLCs 40h

◀**Fig. 3 | TKO EpiLCs up-regulate germline genes and demonstrate longer opportunities for germline contribution *in vitro*. A,** Volcano plots showing differential expression in TKO over WT EpiLCs in Log2FC versus -log10(adj. pvalues). Red: up-regulated genes; Grey: down-regulated genes; Green: germline genes. Thresholds were set up at FDR <5% and FC>1. Data obtained from biological duplicates. **B,** Gene Ontology (GO) analysis showing set enrichment score (NES) and -log10(p.adj) for germline-associated pathways in TKO over WT cells during EpiLC differentiation. Pvalue cutoff= 0.05. **C,** Heatmaps of germline markers showing normalized Log2 CPM counts. Expression was obtained for each condition from the average between biological triplicates. **D,** Schematic representation of PGC-like cell (PGCLC) differentiation, with axes representing genomic methylation levels and time. **E,** Representative FACS plot of SSEA1-Pos and CD61-Pos WT and TKO PGCLCs generated from 40h EpiLCs. Cell percentages are indicated in each quarter. **F,** Expression of PGC-related genes measured by RT-qPCR in WT (blue) and TKO (orange) PGCLCs. Data shown are mean ± SD from biological triplicates as ΔCT values normalized to *Arplp0* and *Rrm2* (two-tailed unpaired student t-test). **G,** Barplot showing time-window percentage of specified WT and TKO PGCLCs generated from ESCs and EpiLCs at different times. Data shown are mean ± SD from biological replicates (*n*=4 or 7) (two-tailed unpaired student t-test). **H,** Heatmaps of naïve pluripotency genes, lineage markers and PGCLC markers showing normalized Log2 CPM counts Fold change of TKO over WT PGCLCs and TKO PGCLCs over TKO EpiLCs. Numbers are indicated when log2FC>1. Expression was obtained for each condition from the average between biological triplicates.

## The absence of DNA methylation does not promote a "germline-competent" subpopulation

Previous studies have postulated that the retention of naïve pluripotency gene expression at the EpiLC stage could prime cells for germline specification *in vitro* (Aramaki et al., 2021; Kim et al., 2020), although this remains debated (Bleckwehl et al., 2021; Mulas et al., 2017). We therefore questioned whether a subpopulation of TKO ESCs could exit naivety with a delayed kinetics, explaining their extended temporal opportunity for germline specification. In particular, as DNA methylation may participate to the repression of some naïve pluripotency genes (Bell et al., 2020; Kalkan et al., 2017), we hypothesized that the exit from naïve pluripotency and/or transition towards priming could occur asynchronously in absence of DNA methylation. To identify subpopulations that we may have missed in our previous bulk approach, we performed single cell SMART-seq2 RNA-seq analysis at critical time-points of naïve pluripotency exit and formative/primed transitioning (D0, D1, D2 and D4 EpiLCs), for a total of 672 cells **(Extended Figure 6A)**.

Using uniform manifold approximation and projection (UMAP) representation, we observed that WT and TKO cells were primarily ordered by time point of EpiLC differentiation, and not by genotype **(Figure 4A-B)**. Clustering analysis revealed five clusters that mostly overlayed with each time point analyzed. Each cluster was delineated by first investigating the expression of known markers of naïve

pluripotency (like *Esrrb, Tfcp2l1*, *Prdm14*), transient (*Dnmt3L*, *Enox1*, *Dppa4*) and primed genes (*Otx2*, *Fgf5*, *Pou3f1*), and top specific markers for each cluster **(Figure 4C-D and Extended Figure 6B)**. Based on their expression profile, we characterized these five clusters as "Naïve", "Exit Naivety", "Transitioning", "Formative" and "Primed" **(Figure 4C)**. Interestingly, the "Primed" cluster demonstrated partition between WT and TKO EpiLCs, which we attributed to the up-regulation of germline genes in TKO cells **(Extended Figure 6C)**. However, according to our clustering analysis, this difference was not sufficient to categorize WT and TKO cells as two different subpopulations.

Previous work reported transcriptional heterogeneity as an indicator preceding cell fate decisions (Bleckwehl et al., 2021; Eldar and Elowitz, 2010; Mohammed et al., 2017; Raj and van Oudenaarden, 2008). However, we did not observe striking differences in overall transcription heterogeneity between WT and TKO EpiLCs at the single cell level **(Extended Figure 6D)**. When focusing on key PGCLC specifiers, we did not observe clusters with higher expression of PGCLC fate inductors (*Prdm1*, *Tfap2c*, *Lrrc31*, *Nr5a2*, *Nanos3*), inhibitors (*Otx2*, *Foxd3*), or markers of epiblast differentiation (*Oct4*, *Dnmt3L*, *Esrrb*, *Fgf5*) **(Figure 4E and Extended Figure 6E)**. Overall, these results indicate that the absence of DNA methylation does not promote the emergence of a 'germline-competent' sub-population with residual naïve gene activity or higher PGCLC gene expression.

Finally, we performed pseudotime analysis, to investigate differentiation trajectories and the dynamics of naïve-to-primed transition. However, we did not observe any diverging trajectory in TKO compared to WT cells **(Figure 4F)**. Both WT and TKO cells were ordered according to their differentiation time point or assigned cluster **(Figure 4F, and Extended Figure 7A-B)**. Moreover, when ordered according to their pseudotime position, no difference in the expression dynamics of major pluripotency and EpiLC differentiation genes could be observed between WT and TKO cells **(Figure 4G and Extended Figure 7C)**. Therefore, we concluded that the ability of DNA methylation-free TKO embryonic cells to commit to the germline does not originate from an asynchronous exit of naivety nor delay in priming.

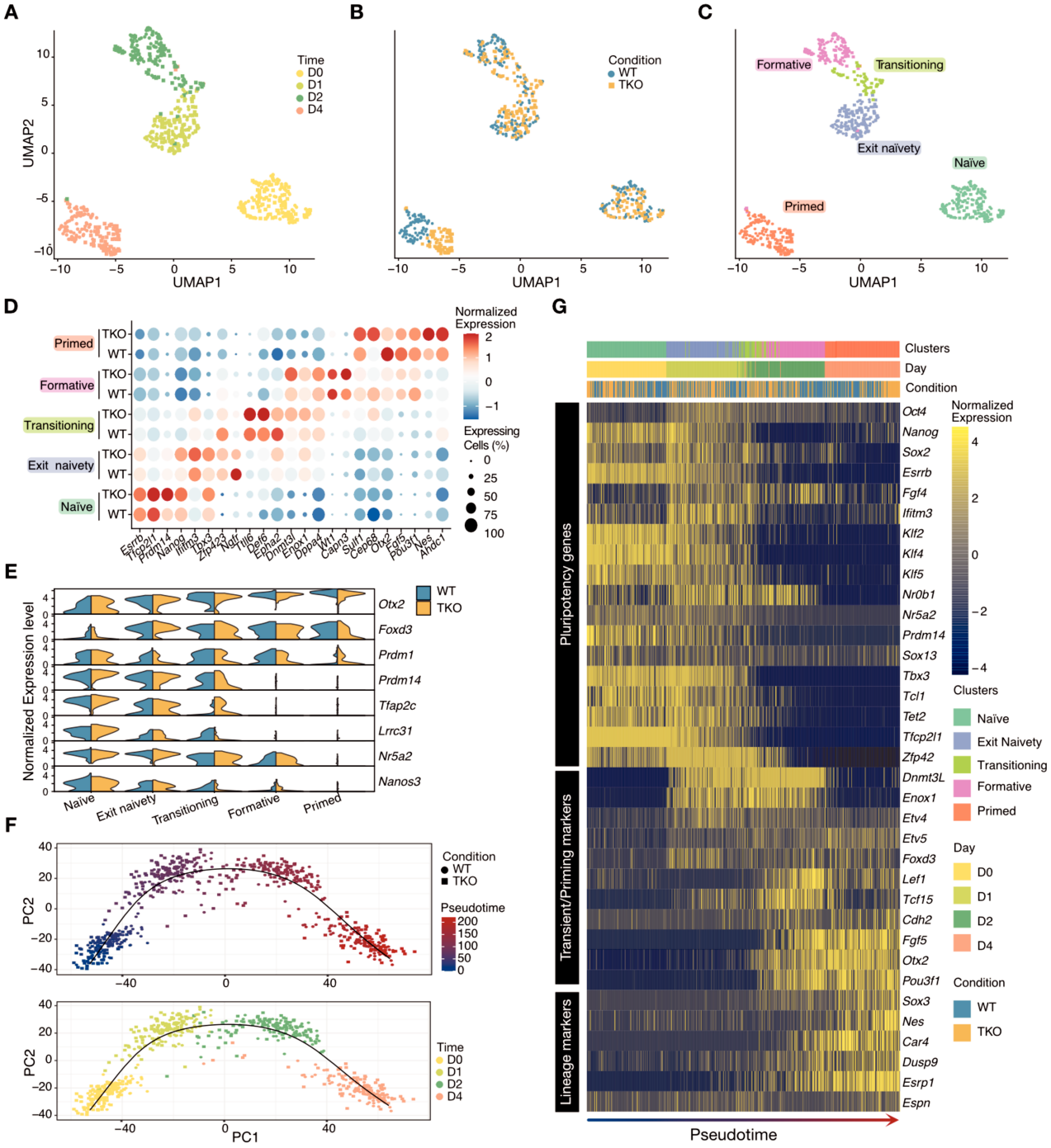



**Fig. 4 I Increased ability of TKO EpiLCs for germline specification does not originate from a "germline-competent" subpopulation. A, B, C,** UMAP dimensionality reduction of scRNA-seq data from WT and TKO EpiLCs highlighting (**A**) Timing of differentiation, (**B**) Condition (WT in blue, TKO in orange) and (**C**) unbiased cell clustering for a total *n*= 672 cells, from two biological duplicates. **D,** Dotplot of expression levels of key pluripotency and priming markers in each cluster detected for WT and TKO cells. Diameter of the dots represents the percentage of cell expressing the marker and colors represent average expression levels. **E,** Violin plot showing expression variation of PGC fate activators and repressors in WT and TKO EpiLCs in each of the identified clusters. **F,** PCA dimensionality reduction for pseudotime trajectory of WT and TKO EpiLCs from first ordered cell (dark blue) to last (dark red), then highlighted for timing of differentiation. **G,** Heatmap of normalized expression levels of different makers of pluripotency, priming and early lineage commitment in WT and TKO cells highlighted by their assigned clusters and days of EpiLC differentiation, and ordered according to their pseudotime coordinates.

**The absence of DNA methylation induces poised chromatin states for both neural and germline fates**

DNA methylation is not the only chromatin mark that is reshaped during naïve-to-primed pluripotency transition. Permissive histone marks such as H3K27ac are notably lost around pluripotency regulators, and H3K27me3 repressive marks are reorganized around the bivalent promoters of developmental genes (Buecker et al., 2014; Kurimoto et al., 2015; Shirane et al., 2016). Our rationale was that the interplay between histone marks and DNA methylation at regulatory elements—promoters and enhancers—might be critical for defining cellular identities (King et al., 2016). To investigate if the absence of DNA methylation could impact priming-associated chromatin program, prior to germline and neural fate induction, we performed Assay for Transposase-Accessible Chromatin (ATAC-seq) and Cleavage Under Targets and Release Using Nuclease (CUT&RUN) against H3K27ac and H3K27me3 in D0 ESCs, and in D2-D4 EpiLCs when TKO cells display extended abilities for germline differentiation **(Extended Figure 8A)**. Our aim was to investigate how the lack of DNA methylation may impact putative enhancers, marked by H3K27ac and ATAC peaks, and developmental genes, marked by H3K27me3.

When quantifying global enrichment for each mark, TKO cells demonstrated genome-wide retention of sharp ATAC and H3K27ac-enriched regions notably at D2 and D4 of EpiLC differentiation, while these tended to decline in WT. Concomitantly, H3K27me3 occurred in wider domains, which were transiently more enriched in WT EpiLCs at D2, while TKO cells demonstrated stronger enrichment at D0 and D4 of EpiLC differentiation **(Figure 5A-C)**. These patterns could relate to previous reports that increased H3K27me3 compensates DNA methylation loss in *Dnmt*-TKO pluripotent cells (Brinkman et al., 2012; Sachs et al., 2013; Walter et al., 2016). Taken together, these results suggest that TKO EpiLCs retain a relaxed chromatin landscape, both in terms of accessibility and active histone marks genome-wide. More precisely, using Differential Enrichment (DE) analysis, 11079 and 14525 regions demonstrated increased ATAC and H3K27ac enrichment, respectively, in D4 TKO EpiLCs compared to WT. This contrasted with earlier time points (D0 ESCs and D2 EpiLCs), where very few peaks showed differential enrichment for these features in TKO versus WT cells. Regarding H3K27me3, according to the trend observed by

global quantification, loss of H3K27me3-marked regions was mostly observed in TKO EpiLCs at D2 (*n*= 283) **(Figure 5D and Extended Figure 8B)**. Noticeably, nearly 70% of H3K27me3-depleted regions mapped to the promoters of genes associated with neurogenic pathways **(Figure 5E-F)**, and more than half of these regions (*n*= 194) subsequently gained an ATAC signal at D4 in TKO EpiLCs, such as the neural genes *Ascl1* and *Scg5* **(Extended Figure 8B-D)**. To investigate whether this peculiar H3K27me3 dynamics in D2 TKO EpiLCs was specific to neural genes, we looked at H3K27me3 enrichment at a previously published list of bivalent developmental genes, which are enriched in both H3K27me3 and H3K4me3 in EpiLCs (Kurimoto et al., 2015). We observed no significant changes in the enrichment of H3K27me3 at the promoter of these bivalent genes in D2 TKO EpiLCs **(Extended Figure 8C)**.

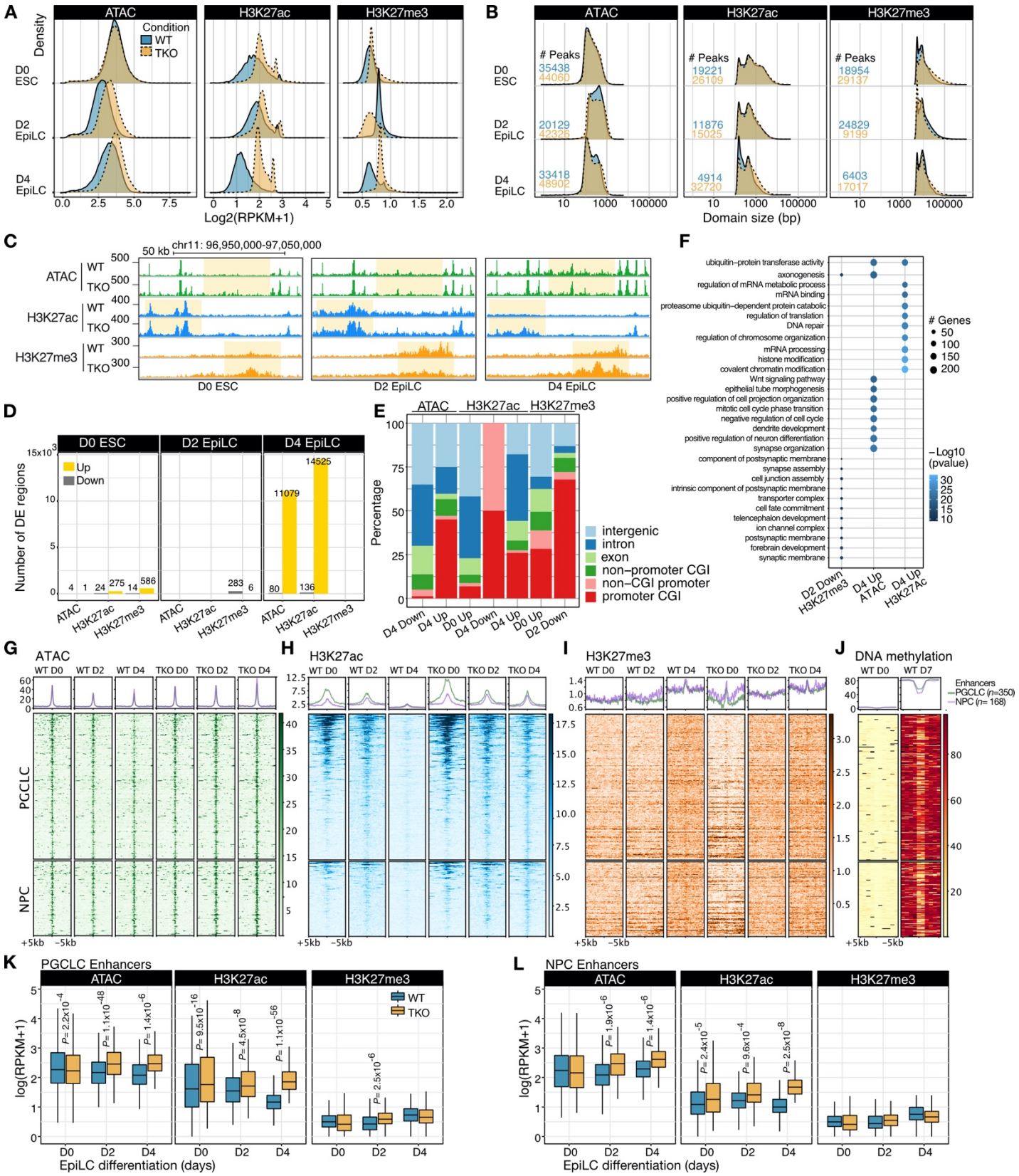
We next sought to investigate how these chromatin changes at promoters could impact gene expression. Notably, over 3000 regions showed dual ATAC and H3K27ac gain in D4 TKO EpiLCs, among which some were found at promoters of genes of interest, like the naïve pluripotency and PGC-associated genes *Zfp42* and *Tfcp2l1* **(Extended Figure 8E)**. However, these differential enrichments of chromatin marks at promoters appear to be only mildly associated with changes in the expression level of associated genes. Indeed, when looking at promoter regions (+/- 5kb from TSS) gaining H3K27ac or ATAC signal, only few were associated with differential gene expression in D4 TKO EpiLCs (265 and 527 genes, respectively). Similarly, regions loosing H3K27me3 signal only affected the expression of 29 genes in D2 TKO EpiLCs **(Extended Figure 9A-C)**. Consequently, whilst H3K27me3 is transiently lost in TKO EpiLCs around neural genes, we do not consider that these changes in promoter chromatin could directly impact cell fate decisions, as few changes in gene expression levels are observed. This is consistent with our previous observations during pluripotency priming, where the lack of DNA methylation also had little impact on transcriptional regulation.

Next, we specifically focused on NPC and PGCLC putative enhancers, addressing whether the absence of DNA methylation could impact their H3K27ac and ATAC enrichment. We defined these putative NPC (*n*=168) or PGCLC (*n*=350) enhancers as regions with an overlapped H3K27ac and ATAC signals at any time point of WT EpiLC differentiation, located in a 50kb window of known NPC- or

PGCLC-specific genes determined from our own NPC and PGCLC transcriptomic datasets (genes were declared specific of a cell type if they demonstrate FDR < 0.05 and Log2CPM>1 in comparison to other cell states, see methods). Interestingly, we found that in absence of DNA methylation, PGCLC- or NPC- associated enhancers globally retained H3K27ac and/or ATAC signal in D2 and D4 TKO EpiLCs. In contrast, these features tended to attenuate upon WT EpiLC differentiation as these enhancers acquired intermediate levels of DNA methylation (around 50-60% of mCpGs quantified at D7) **(Figure 5G-L)** (Greenberg et al., 2019). Intriguingly, no major change could be observed for H3K27me3 at these defined NPC and PGCLC regulatory elements.

Next, we looked at all regions gaining both ATAC and H3K27ac signals in TKO EpiLCs, defining a list of differentially enriched putative enhancers (DE enhancers). We specifically focused on D4 EpiLCs since at this time-point, lack of DNA methylation has the strongest effect on chromatin, and it also coincides with the highest bias for TKO EpiLCs to adopt a germline fate when compared to WT cells. We identified over 800 putative DE enhancers scattered in the genome of D4 TKO EpiLCs, among which we found 74 and 18 previously defined PGCLC and NPC putative enhancers, respectively. However, again, we could not associate these DE enhancers with increased expression of EpiLC-, PGCLC- or NPC-associated genes in the TKO background **(Extended Figure 9D-E)**. This further reinforces that chromatin changes are decoupled from transcription in TKO cells, as observed so far.

Overall, these results reveal that the lack of DNA methylation affects chromatin in transitioning EpiLCs by promoting *i)* global chromatin relaxation with increased accessibility, *ii)* retention or differential enrichment for H3K27ac/ATAC marks, notably around PGCLC and NPC putative enhancers, and *iii)* transient loss of H3K27me3 repressive marks at neural gene promoters. Although these changes are not impactful at the transcriptional level, we hypothesized that they might poise embryonic cells for promptly adopting neural or germline fates when exposed to appropriate signals in culture.


A
ATAC
H3K27ac
H3K27me3
Density
Condition
WT
TKO
D0 ESC
D2 EpiLC
D4 EpiLC
Log2(RPKM+1)
B
# Peaks
35438
44060
19221
26109
18954
29137
20129
42326
11876
15025
24829
9199
33418
48902
4914
32720
6403
17017
Domain size (bp)
C
chr11: 96,950,000-97,050,000
50 kb
D0 ESC
D2 EpiLC
D4 EpiLC
D
Number of DE regions
Up
Down
11079
14525
283
6
80
136
4
1
24
275
14
586
E
Percentage
intergenic
intron
exon
non–promoter CGI
non–CGI promoter
promoter CGI
D4 Down
D4 Up
D0 Up
D4 Down
D4 Up
D0 Up
D2 Down
F
ubiquitin–protein transferase activity
axonogenesis
regulation of mRNA metabolic process
mRNA binding
proteasome ubiquitin–dependent protein catabolic
regulation of translation
DNA repair
regulation of chromosome organization
mRNA processing
histone modification
covalent chromatin modification
Wnt signaling pathway
epithelial tube morphogenesis
positive regulation of cell projection organization
mitotic cell cycle phase transition
negative regulation of cell cycle
dendrite development
positive regulation of neuron differentiation
synapse organization
component of postsynaptic membrane
synapse assembly
cell junction assembly
intrinsic component of postsynaptic membrane
transporter complex
cell fate commitment
telencephalon development
ion channel complex
postsynaptic membrane
forebrain development
synaptic membrane
# Genes
–Log10 (pvalue)
D2 Down H3K27me3
D4 Up ATAC
D4 Up H3K27Ac
G
ATAC
WT D0
WT D2
WT D4
TKO D0
TKO D2
TKO D4
PGCLC
NPC
+5kb
–5kb
H
H3K27ac
I
H3K27me3
J
DNA methylation
WT D7
Enhancers
PGCLC (n=350)
NPC (n= 168)
K
PGCLC Enhancers
log(RPKM+1)
EpiLC differentiation (days)
WT
TKO
L
NPC Enhancers

▲ **Fig. 5 | TKO EpiLCs retain active chromatin marks at NPC and PGCLC putative enhancers. A, B,** Ridge plots displaying global (**A**) Density in Log2(RPKM+1) and (**B**) Domain size in bp, for ATAC-, H3K27ac- and H3K27me3-marked regions in WT (blue) and TKO (orange) EpiLCs at D0, D2 and D4 of differentiation. **C,** Genome tracks for ATAC, H3K27ac and H3K27me3 for WT and TKO EpiLCs at D0, D2 and D4. **D,** Differential Enrichment (DE) analysis of ATAC, H3K27ac and H3K27me3 peak enrichment in TKO over WT cells. Number of gained peaks ("up", yellow) and lost peaks ("down", grey") are shown at D0, D2 and D4 of EpiLC differentiation (FDR<0.05). **E,** Genomic annotation analysis of DE regions. **F,** GO analysis of genes in proximity to ATAC, H3K27ac or H3K27me3 DE peaks (< 3kb distance from TSS). Color gradient represents -log10(pvalue) with circle size showing the number of genes associated with each GO pathway. The top 30 enriched ontologies are shown (based on FDR) per gene set. **G, H, I,** Heatmaps and metaplot for (**G**) ATAC, (**H**) H3K27ac and (**I**) H3K27me3 , for PGCLC (purple, *n*=350)) and NPC (red, *n*=168) enhancers in WT and TKO cells during EpiLC differentiation. Enrichment in RPKM for 10kb windows centered around peaks are represented. **J,** Heatmap and metaplot for CpG methylation levels in WT and TKO cells at D0 and D7 of EpiLC differentiation at NPC (purple) and PGCLC (green) enhancers. Enrichment is displayed in RPKM for 1kKb windows centered around peaks. **K, L,** Boxplot showing enrichment in Log(RPKM+1) of ATAC, H3K27ac and H3K27me3 marks at (**K**) PGCLC enhancers and (**L**) NPC enhancers, for WT and TKO cells during EpiLC differentiation. Data shown are the median with upper and lower hinges corresponding to 75 and 25% quantile (two-tailed unpaired student t-test).

### NPC and PGCLC enhancers share similar network of DNA methylation-sensitive transcription factors

To gain insights into why NPC-PGCLC enhancers, as defined above, may gain H3K27ac and ATAC signals in TKO cells, we performed transcription factor (TF) binding motif analysis using HOMER (Heinz et al., 2010). Interestingly, we found that putative PGCLC and NPC enhancers displayed largely overlapping sets of binding motifs **(Figure 6A and Extended Figure 10A)**. In addition, similar binding motifs could also be found when considering all DE enhancers detected in TKO EpiLCs at D4, suggesting that all these elements might be co-regulated. Most importantly, we noticed that many of these binding motifs relate to TFs that are DNA methylation-sensitive or regulated by DNA methylation, like SP, KLF, CEBP, CTCF, ETS and BACH factors **(Figure 6A)** (Domcke et al., 2015; Dura et al., 2022; Héberlé and Bardet, 2019; Luo et al., 2020; Suzuki et al., 2017; Zhu et al., 2016). In addition, some motifs corresponded to TFs involved in regulation during naïve-to-primed pluripotency transitioning, most notably the pluripotency triad (OCT4, SOX2, NANOG), KLF, ZIC and ETS factors, as well as ZFP281 (Akagi et al., 2015; Buecker et al., 2014; Mayer et al., 2020; Yang et al., 2018). Motif analysis at NPC, PGCLC and DE enhancers also showed that some of the TF binding motifs that were identified contain CpG dinucleotides, supporting the idea of possible DNA methylation-based regulation **(Extended Figure 10B)**.

To identify which of these TFs might be mobilized in D4 TKO EpiLCs, we performed footprinting analysis from ATAC-seq data, using TOBIAS (Bentsen et al., 2020). We observed an increased binding score at NPC, PGCLC and DE enhancers, for motifs associated with DNA methylation-sensitive factors or lineage inductors, such as ASCL1 at NPC enhancers, and the pluripotency triad at PGCLC enhancers **(Figure 6B)**. This indicates possible increased or more frequent binding of these factors in absence of DNA methylation. Taken together, these results suggest that NPC, PGCLC and DE enhancers are probably regulated by a similar set of transcription factors sensitive to DNA methylation. The absence of DNA methylation might promote the binding of DNA methylation-sensitive factors, which in turn would favor the retention of permissive chromatin states at these enhancers during EpiLC differentiation.

**Neural and Germline fates are preferential differentiation routes in absence of DNA methylation**

Importantly, these characteristics were specific to NPC and PGCLC enhancers: no clear enrichment of ATAC and H3K27ac signal was found at enhancers previously associated with endoderm and mesoderm development (Argelaguet et al., 2019a) **(Figure 6C and Extended Figure 10C)**. Moreover, binding motif analysis suggested that putative meso/endoderm enhancers are likely regulated by a different set of TFs than NPC-PGCLC enhancers, with only FOXO factors demonstrating known DNA methylation sensitivity **(Extended Figure 10B)**. Moreover, when looking at GC content and DNA methylation dynamics using public WGBS datasets at D0, D3 and D7 of WT EpiLC differentiation (Greenberg et al., 2019), we found that meso/endoderm enhancers were globally GC-poor and rapidly acquired high DNA methylation levels upon EpiLC differentiation **(Extended Figure 10D)**. In contrast, NPC-PGCLC-DE enhancers were more GC-rich and reached lower DNA methylation levels. Intermediate levels of DNA methylation suggest that a proportion of WT EpiLCs might carry high methylation at these elements, while others remain hypomethylated, indicating that DNA methylation acquisition in a fraction of WT EpiLCs restricts the timing of cell fate acquisition at the population level. This is reminiscent of the neural default pathway hypothesis, as it was shown that ectoderm enhancers remain “open”

and relatively hypomethylated during epiblast development *in vivo*, compared to other lineages, like mesoderm and endoderm (Argelaguet et al., 2019a). Therefore, we propose that the germline could constitute also a default route during priming, as putative PGCLC enhancers demonstrate the same DNA methylation dynamics and TF network as NPC enhancers.

To functionally test if the neural and germline fates indeed constitute default differentiation routes, we first performed somatic differentiation towards *in vitro* meso/endoderm. We found that while TKO cells survived the protocol and repressed pluripotency markers—once again demonstrating the irrelevance of DNA methylation in the exit of pluripotency—they were unable to properly activate the meso/endoderm markers *Gata4* or *Foxa2*, but rather upregulated neural genes like *Sox1*, *Nestin* and *Ascl1* **(Figure 6D and Extended Figure 10E)**. Then, we repeated PGCLC induction from WT and TKO EpiLCs without adding BMPs and other pro-germline cytokines. After sorting with PGCLC surface markers, we found that TKO cells retained an inherent higher ability to give rise to the germline, compared to WT cells, despite a lower efficiency compared to normal conditions **(Figure 6E**). This corroborates that the germline and the neural lineage are indeed default routes of differentiation in absence of DNA methylation.

To sum up, we propose that DNA methylation instructs cell fate decisions, not via transcriptional control, but rather by tempering neural and germline fates as default routes of development. We suggest that the acquisition of DNA methylation is important for decommissioning a subset of NPC-PGCLC enhancers during priming. Indeed, in absence of DNA methylation, stem cells appear to be in an “hyper formative state” at later time points of EpiLC differentiation, with high chromatin accessibility, strong footprinting for DNA methylation-sensitive TFs at neural and germline enhancers, and priming of TKO EpiLCs to adopt theses fates when exposed to proper differentiation cues **(Figure 6F)**.

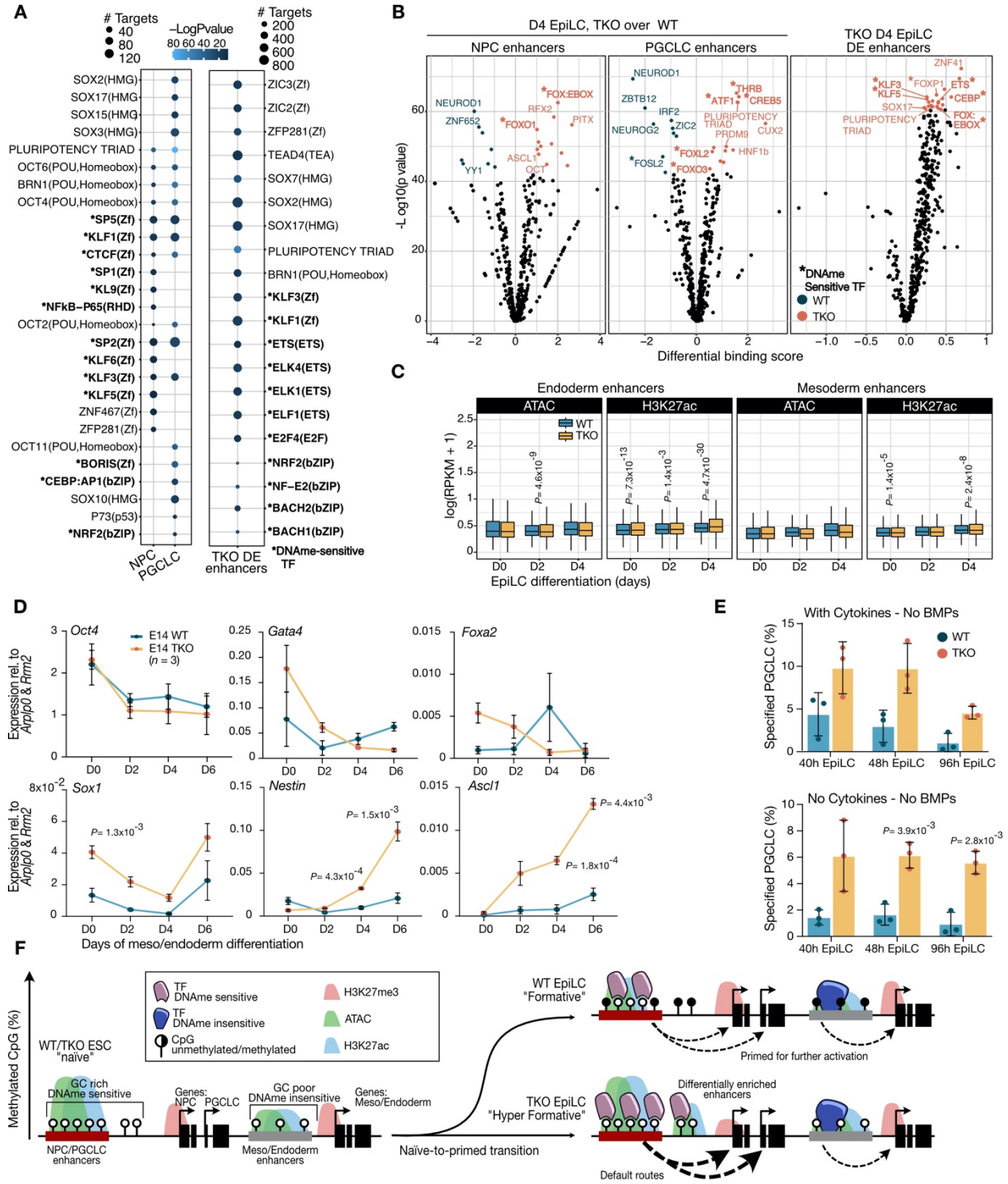



**Fig. 6 I DNA methylation-sensitive factors define NPC and PGCLC fates as default routes. A,** Bubble plot of HOMER analysis for top 20 TF binding motifs at PGCLC-, NPC- and Differentially Enriched enhancers. Factors that are DNA methylation-sensitive or regulated by DNA methylation are highlighted with (*). **B,** Footprint analysis of predicted TF occupancy at NPC-, PGCLC- and DE enhancers in TKO over WT at D4 EpiLC stage, using TOBIAS. **C,** Boxplot showing enrichment in Log(RPKM+1) of ATAC and H3K27ac marks at Endoderm (*n*=1957) and Mesoderm enhancers (*n*=1441) as defined in Argelaguet *et al.,* 2019, during WT and TKO EpiLC differentiation. Data shown are the median with upper and lower hinges corresponding to 75 and 25% quantile (two-tailed unpaired student t-test). **D,** Expression of pluripotency genes (*Oct4*), meso/endoderm (*Gata4*, *Foxa2*) and neural (*Sox1*, *Nestin*, *Ascl1*) markers measured by RT-qPCR in WT and TKO during mesoderm differentiation. Data shown are mean ± SD from biological triplicates. ΔCT values were normalized to *Arplp0* and *Rrm2* (two-tailed unpaired student t-test). **E,** Barplot showing percentage of specified WT and TKO PGCLCs generated from ESCs and EpiLCs at different times, in absence of BMPs with and without cytokines, data shown are mean ± SD from biological replicates (*n*=3). **F,** Hypothetical model for DNA methylation as a restrictive mark to attune cell fate decisions and repress neural and germline as default routes of differentiation in early primed pluripotent EpiLCs.

# DISCUSSION

Despite being recognized for its key role on development, when DNA methylation becomes critical and for which lineages specifically was still poorly understood. Using various differentiation protocols applied to a newly generated DNA methylation-free *Dnmt1; Dnmt3A; Dnmt3B*-TKO ESC line, our work provided an unprecedented resolution of the role of DNA methylation on key developmental bifurcations. First, it was known that DNA methylation plays little role in regulating gene expression in naïve pluripotency (Tsumura et al., 2006), but we demonstrated here that this also applies to primed pluripotency. In particular, while acquisition of DNA methylation coincides with naïve-to-primed transition, TKO ESCs were nonetheless apt at exiting naïve pluripotency and at acquiring functional features of the primed epiblast, *in vitro* and *in vivo*. Second, DNA methylation also appeared unnecessary for the induction of neural progenitors during somatic differentiation, although it becomes later necessary for consolidating the neural program. Moreover, while neural fate induction was possible, TKO cells were otherwise unable to properly commit to the meso/endoderm lineage and spuriously expressed neural markers. Finally, germline induction was also not impaired by the lack of DNA methylation, and TKO cells even showed extended and self-autonomous opportunities for germline specification. In some recent work, it was found that treating embryoid bodies with the demethylating agent 5-azacytidine promoted the adoption of PGC identity associated with a delay in the exit of naïve pluripotency (Kim et al., 2020). However, we did not find evidences suggesting that the extended time-window for germline specification in absence of DNA methylation was due to an improper repression of naïve pluripotency genes or even in a delayed priming phase in TKO EpiLCs.

By performing a step-wise chromatin profiling of H3K27me3, H3K27ac and ATAC signal, we propose that DNA methylation acquisition in the epiblast is critical for allowing multi-lineage decisions, by tempering neural and germline fates as default routes of differentiation. Indeed, we observed that in TKO EpiLC, NPC and PGCLC gene-associated putative enhancers maintain high level of chromatin accessibility and active H3K27ac. We also found that NPC and PGCLC enhancers may be targeted by the same network of DNA methylation-sensitive factors, suggesting possible co-regulation of the two lineages. Moreover, footprint analysis revealed high binding

score of these DNA-methylation sensitive TFs at neural and germline enhancers in TKO cells. While we also observed changes of chromatin marks at promoters, most notably a loss of H3K27me3 at neural genes, these changes are unlikely to promote neural fate in TKO cells, as they are only transient and poorly associate with changes in gene expression level.

The decoupling between gene silencing and repressive epigenetic in naïve stem cells is well studied. As of recently, it was shown that forced deposition of DNA methylation in 2i-cultured ESCs poorly impacted transcription (Gretarsson and Hackett, 2020; McLaughlin et al., 2019). Whilst it is widely admitted that repressive chromatin marks are critical for development, it is important to make the distinction between the relevance of these marks in primed pluripotent cells versus terminally differentiated cells. In fact, it appears that DNA methylation and other marks, like H3K27me3 and H3K27ac, do not directly impact gene regulation in primed cells, and in precursors of somatic and germline lineages (Huang et al., 2021; Riising et al., 2014; Sankar et al., 2022). This could be understandable especially for repressive chromatin marks that act as lockers of cell identities over time, as illustrated by the necessity of active DNA demethylation in somatic cell reprogramming (Pei et al., 2014). However, in the context of rapidly dividing stem cells, which need to remain plastic enough to respond to various differentiation cues, the reliance on repressive marks for gene regulation might be minimal. This was best illustrated in our chimera experiments, where we consistently observed strong contribution of TKO ESCs in the primed epiblast. While we did not observe extra-embryonic contribution, as previously reported in *Dnmt3A; Dnmt3B* double-KO chimerae (Kinoshita et al., 2021), we found that TKO chimerae were delayed and smaller than WT chimerae at E6.5. This supports that DNA methyltransferases may become necessary for epiblast development after priming (Dahlet et al., 2020; Lei et al., 1996; Okano et al., 1999).

It has been reported in various differentiation systems that accessible enhancers in progenitors cells could pre-determine cell routes in differentiation (Banerjee et al., 2018; Rauch et al., 2019). Moreover, DNA methylation at lineage-associated enhancers is very heterogenous due to the simultaneous activity of DNMTs and TET enzymes (Charlton et al., 2020; Ginno et al., 2020; Rulands et al., 2018),suggesting a potential role in instructing enhancer activity during differentiation.

More recently, is was shown that low DNA methylation levels at ectoderm enhancers during epiblast differentiation promoted the establishment of this germ layer as a default pathway in embryonic development (Argelaguet et al., 2019a). Similarly, around 3% of somatic lineage-specific enhancers were estimated to be regulated by methyl-sensitive transcription factors (Kreibich et al., 2022), some of which were found common at our NPC and PGCLC putative enhancers. Here we propose that, similarly to the ectoderm, the germline is also subject to such DNA methylation-based restriction, and that its absence skews cells to adopt a germline fate. A similar extended window for PGCLC specification has been described upon exposure of EpiLCs to alpha-ketoglutarate (aKG), which promotes a mitochondrial metabolism of oxidative phosphorylation (OxPhos) (Tischler et al., 2018). In addition, the LIF-STAT signaling pathway has been shown to regulate DNA methylation and to promote OxPhos in naïve cells (Betto et al., 2021). Hence, the question can be raised whether the absence of DNA methylation could promote OxPhos and germline fate acquisition. However, we did not observe misregulation of OxPhos-associated genes in TKO EpiLCs (data not shown), suggesting that the lack of DNA methylation might not impact the metabolism of pluripotent stem cells. In conclusion, while our study demonstrated the relatively minimal impact of DNA methylation on gene regulation in early development, it shows that DNA methylation is critical for instructing chromatin switches at gene regulatory elements, thereby tempering default routes of differentiation.

# MATERIAL AND METHODS

### Mouse embryonic stem cell lines

Mouse embryonic stem cells (ESCs, E14 background derived from the 1291*0la* strain) were used to generate a triple knockout (tKO) for the three DNA methyltransferase genes *Dnmt1*, *Dnmt3A* and *Dnmt3B*. The strategy for CRIPSR-Cas9 editing of the catalytic PCN/Q loop of the DNMT MTase domain was previously described in Domcke et al., 2015.

### ESC culture

ESCs were cultured in feeder-free conditions at 37°C in 5% $CO_2$ atmosphere. Mycoplasma-free status was assessed using the MycoplasmaCheck service from Eurofins.

*Serum-LIF culture medium:* ESCs were cultured on 0.1% gelatin-coated plates with serum-LIF medium (Glasgow medium (Sigma), 2mM L-glutamine (Gibco), 0.1mM MEM non-essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), 15% Fetal Bovine Serum (FBS), 0.1mM β-Mercaptoethanol and 1000 U/ml leukemia inhibitory factor (LIF, Chemicon)). Cells were counted on a Beckman Cell Counter Vi-Cell XR and passed with TrypLE (Trypsin replacement enzyme, Gibco) every 2-3 days.

*2i + Vitamin C culture medium:* To reduce DNA methylation levels, ESCs were grown for at least two weeks on fibronectin-coated plates (1/60 dilution in PBS) in 2i+VitC medium (N2B27 (50% Neurobasal medium and 50% DMEM/F12 (Gibco) complemented with 2mM L-glutamine, 0.1mM β-Mercaptoethanol, N2 supplement (Gibco) and B27 serum-free supplement (Gibco)), complemented with 1000U/ml LIF, 2i (3μM Gsk3 inhibitor and 1μM of MEK inhibitor), and vitamin C 100μg/ml final (Sigma)) and passed with accutase (GIBCO) every 2-3 days.

*Epiblast-like cell (EpiLC) differentiation:* To initiate EpiLC differentiation, 2i+VitC-grown ESCs were first split with accutase (Gibco) and plated on fibronectin-coated 6-well/plates at a density of $0.20 \times 10^6$ cells/well. The next day, the medium was replaced with AF medium: N2B27 medium supplemented with 12ng/ml Fgf2 (R&D) and 20ng/ml Activin A (R&D), 1% Knock-out Serum Medium (Gibco), to drive EpiLC differentiation. Cells were harvested at Day 2, 4, and 7 of differentiation, with an additional splitting step at D4 with accutase to reduce cell density.

*Neuronal progenitor cell (NPC) differentiation:* For monolayer NPC differentiation, serum-LIF-grown ESCs were split with accutase onto laminin-coated 6-well/plates (Sigma, 10μg/ml in PBS) at a density of $0.75 \times 10^4$ cells/well in plain N2B27 medium. Cells were cultured up to 10 days with daily medium change.

*Embryoid body differentiation:* Embryoid bodies were generated by dissociating 2i+VitC-grown ESCs with TrypLE and $3 \times 10^6$ cells were seeded in uncoated B10 plates in serum without LIF to allow the formation of cell aggregates. Medium was changed every 2 days.

*Meso-endoderm differentiation:* Meso-endoderm differentiation was adapted from Kinoshita et al., 2020. Briefly, serum-LIF-grown ESCs were split with accutase onto fibronectin-coated plates (Sigma, 10µg/ml in PBS) at a density of $1x10^4$ cells/ well in 6-well/plates in N2B27 basal medium supplemented with 10 ng/ml Activin A and 3 µM of GSK3 inhibitor CHIR99021. Cells were cultured for 6 days, and the medium was changed every two days.
*PGC-like cell differentiation:* Dissociated EpiLCs were resuspended in PGCLC basic medium: GMEM (GIBCO), KSR 15% (GIBCO), Non-Essential Amino Acid (GIBCO), Sodium Pyruvate 1mM (GIBCO), L-Glutamine 2mM (GIBCO), Penicillin-Streptomycin 100U/ml (GIBCO), β-Mercaptoethanol 0.5mM (GIBCO). Then EpiLCs were distributed into ultra-low adhesive U-bottom 96-well plates at a density of $2.5x10^3$ cells/well in 100µl of PGCLC medium complemented with cytokines for PGCLC induction (PGCLC basic medium, 0.5 µg/ml BMP4 (R&D system), 0.5 µg/ml BMP8 (R&D system), 0.1 µg/ml stem cell factor (SCF, R&D system), 50 ng/ml epidermal growth factor (EGF, R&D system), 1000 u/ml mouse LIF). After 4 days of culture, embryoid bodies were harvested and dissociated with 10mM EDTA in PBS+3%FBS for subsequent FACS analyses.

**Cell transfection and cell line isolation**

All vectors were generated using Gibson cloning from Gibson et al., 2009. For sgRNA cloning, pX330 (Addgene #42230) or pX459 (Addgene #62988) plasmids were digested with *BbsI* restriction enzyme (NEB) and annealed with double-stranded sgRNA oligos using QuickLigase (NEB). The H2B::tdTomato-containing piggyBac vector was a kind gift from Dr. Jamie Hackett (EMBL Rome). Transgenic ESC lines were generated by nucleofection using Amaxa 4D nucleofector (Lonza) using 3-5× $10^6$ cells with 1-3µg of vectors and sgRNA/Cas9 plasmids. Cells were then seeded in serial dilutions in 0.1% gelatin-coated B10 plates in serum-LIF. After two days of recovery, cells underwent puromycin (1µg/ml, Life Technologies) or hygromycin B (200µg/ml Sigma) selection for 2 or 4 days, respectively. After selection, the culture medium was changed for serum-LIF only and individual colonies were picked and screened by PCR genotyping.

**Live-imaging for apoptosis analysis**

To quantify apoptosis, cells were grown on a 96 well-plate and placed in the incubated chamber of an Incucyte Microscope (IncuCyte ZOOM, Sartorius). Then, EpiLC differentiation media was added and supplemented with 5µM of Caspase 3/7 Green dye (Sartorius) to detect for apoptosis in live cells. Media was changed every day, brightfield and green fluorescence images were taken every 2 hours for 96 hours. Counting of dead cells was made using IncuCyte ZOOM Software.

**FACS**

For FACS, cells were collected, dissociated with TrypLE, resuspended with PBS/FBS 3% supplemented with DAPI (1µg/ml) to discriminate dead cells. For PGCLC specification quantification, 2.5µL/$10^5$ PGCLCs were incubated with Anti-CD61-PE (Biolegend #104307) and Anti-SSEA1-eFluor660 (eBioscience #50-8813-42) antibodies in PBS/FBS 3% for 30 min on ice, in the dark, before being analyzed on the ZE5 Flow cytometer (Biorad). PGCLC sorting was performed on a Sony SH800. Quantification of tdTomato fluorescence in ESCs was also measured on a ZE5 (BioRad).

**Karyotyping**

ESCs were treated with KaryoMAX Colcemid (Gibco) in PBS for 2h at 37°C. After dissociation using TrypLE, cells were resuspended in 3ml of hypotonic solution (1 volume of FBS+3 volumes of H2O) for 10min at 37°C. After another centrifugation, cells were fixed with 3ml of Carnoy solution (Acetic acid diluted ¼ in 100% Ethanol) and incubated at room temperature (RT) for 30 min. Cells were then dropped from ~50cm height onto glass slides. Slides were air-dried and DNA was stained with DAPI (1µg/ml). Chromosomes were imaged with an upright widefield microscope (Leica) and around 100 cells were analyzed per cell line on Fiji for counting.

**Western Blot analysis**

Proteins were extracted by sonicating cell pellets after treatment with protease inhibitors (Aprotinin, Leupeptin, Pepstatin and PMSF) and BC250 buffer containing 1M TRIS pH8, 20% glycerol, and 0.5M EDTA. The mixture was centrifuged (10 min-9000g at 4°C). The supernatant was denatured at 95°C for 10min and equal quantities of protein were transferred on a 4%-12% Bis-Tris Gel (NuPage). After electrophoresis, the gel content was transferred on 0.45µm nitro-cellulose membranes (GE Healthcare). After blocking with 5% milk and PBS/NP40 (PBS+0.3% NP40), membranes were incubated overnight with antibodies diluted 1:5000 in 1% milk in PBS/NP40. After washes in PBS/NP40, membranes were incubated 2h at RT with HRP-conjugated secondary antibodies. Revealing was performed by 5 min incubation with LumiLight$^{Plus}$ (Roche) substrate on a Chemidoc MP Imaging system (BioRad). Antibodies are listed in Table 1.

**LUMA analysis**

Global DNA methylation levels were assessed using the Luminometric Methylation Assay (LUMA) (Karimi et al., 2006) from WT and TKO ESCs in serum-LIF medium, and WT ESCs cultured in 2i medium complemented with VitC. Briefly, 500ng of genomic DNA was digested with *MspI*+*EcoRI* and *HpaII*+*EcoRI* (NEB) in parallel reactions at 37°C for 4h, followed by heat inactivation at 65°C. *EcoRI* is included as an internal reference. Filling of the genome-wide protruding ends of the restriction digestions

were quantified in a pyrosequencing reaction (PyroMark Q48 autoprep, Qiagen, dispensation order: ACTCGA) and all results were analyzed with the associated software (Q48 Autoprep Software).

**Liquid chromatography-mass spectrometry (LC-MS)**

LC-MS-based quantification of methyl-cytosines was performed on 1 µg of DNA degraded to nucleosides with nuclease P1 (Roche), snake venom phosphodiesterase (Worthington) and alkaline phosphatase (Fermentas). An equal volume of isotopic standard mixture (15N3-C (Silantes), 2H3-5mC (TRC) and self-synthesized 15N3-5hmC, 15N3-5fC and 15N3-5caC were added to the nucleoside mixture and injected for LC-MS/MS analysis as described previously Schomacher et al., 2016. Quantitative analysis was performed on an Agilent 1290 Infinity Binary LC system (Agilent technologies) using a ReproSil 100 C18 column (Jasco) coupled to an Agilent 6490 triple quadrupole mass spectrometer (Agilent technologies).

**Dual Luciferase assay**

Dual luciferase assays were performed using Promega DLR kit (#E1910) according to the manufacturer's instructions. WT and TKO ESCs were cultured in 2i+VitC for transfection at D0, or cultured in AF medium for 2 days for EpiLC transfection. Cells were co-transfected with pGL3EV/pCAGRLuc (Empty control vector) and pGL3DE/pCAGRLuc (distal enhancer) or pGL3PE/pCAGRLuc (proximal enhancer). Briefly, 1.6µg of each plasmid was diluted in a mix of 500µl of OptiMEM media (GIBCO) and 4µl of Lipofectamine 2000 (Thermoscientific) per well. Culture medium was replaced with the transfection mix for half a day, and then completed with 500µl of AF medium. The next day, medium was changed for AF or 2i+VitC and cells were allowed to grow one more day before harvest and lysis using PLB buffer for luminescence analysis on CalrioStar (BMG Labtech).

**ESC Reversion assay**

ESCs were first differentiated into EpiLCs. At 2, 4 and 7 days of differentiation, cells were split in accutase, and plated in fibronectin-coated 6-well plates at the density $1x10^5$ cells/well in 2i+VitC medium. Colonies from D2, D4 or D7 EpiLCs were grown for 4, 7 and 10 days, respectively, before being stained for alkaline phosphatase activity. Briefly, cells were rinsed with PBS and fixed with ethanol 100% for 10 min at RT. After 3 PBS rinses, cells were incubated at 37°C for 10 min with FastBCIP/NBT (Sigma #B5655) dissolved in MiliQ water, and air-dried. Pictures were taken and analyzed on Fiji to count AP positive stained reverted ESC colonies.

**Chimerae**

*Mice:* Mice were hosted on a 12h light/12h dark cycle at 22 +/- 2°C ambient temperature and 40-70% humidity, with free access to food and water, in the pathogen-free Animal Care Facility of the Institut Curie (agreement C75-05-18). All

animal experiments were performed following European (Directive 2010/63/EU) legislation and were approved by the ethics committee of the Institut Curie CEEA-IC #118 and by French Ministry of Research (APAFiS#30920-2021040622279416-v1).
*Chimera aggregation:* Eight-week-old B6D2F1 (C57BL/6J × DBA2) females, were superovulated by intraperitoneal (i.p.) administration of 5 IU of Pregnant Mare Serum Gonadotropin (PMS013 Centravet) followed by an additional i.p. injection of 5 IU Human Chorion Gonadotropin (CH0003 Centravet) 48 hours later. Females were mated to a stud male of the same genetic background. Morula embryos were collected and washed in M2 medium (Sigma M7167) and cultured in Cleave medium (Cook K-RVCL) at 37 °C under 5% CO2. Meanwhile, low passage WT and TKO ESCs containing the H2B::tdTomato reporter were thawed and cultured in 2i medium. Cells were dissociated by accutase in small clumps of 5-10 cells. Right before aggregation, the zona pellucida of morulae was removed by Tyrode's acid treatment (Sigma T1788-100ML), before being washed with M2 medium. Each morula-stage embryo was placed into a depression well performed using an aggregation needle (BLS DN-09) on a plastic plate. A dissociated ESC clump was added close to the embryo to allow aggregation and development to the blastocyst stage. Integrity and proper integration of transgenic ESCs in the inner cell mass were assessed with a fluorescence inverted microscope (Zeiss Axio Vert A.1) before re-implantation into pseudo-pregnant foster mothers (NMRI). Embryos were harvested at E7.5 and E8.5 for whole-mount immunostaining.

**Immunofluorescence**

For whole-mount embryo immunostaining, E7.5 and E8.5 embryos were fixed in 4% PFA for 30 min at RT. Then, embryos were washed in PBT (PBS+0,1% Triton) and permeabilized for one hour at room temperature in PBS+0,5% triton. Blocking was performed by adding 0.2% BSA and 5% Donkey serum in PBT for 30 minutes. Primary antibodies diluted in blocking solution were incubated for 72 hours at 4°C with the embryos, then three washes of 15min at RT were performed. Similarly, secondary antibodies were diluted in blocking solution (Table 1) and incubated for 3 hours at RT. After the final washes, embryo mounting was made through increasing percentages of glycerol in increasing times as follow: 2.5% (5 minutes), 5% (5 minutes), 10% (10 minutes), 20% (15 minutes), 50% (15 minutes), and DTG (Glycerol, DABCO 2.5%, 50mM Tris pH8.6) (15 minutes). Images were obtained using a Zeiss LSM800 inverted confocal microscope

**RT-qPCR**

RNA extraction was performed on cell pellets using standard Trizol (Life Technologies)-chloroform protocol. RNA pellet was treated with DNase before clean up with RNeasy Mini kit (QIAGEN). To generate cDNA first-strand, a minimum of a 100ng of RNA was reverse transcribed using random priming with SuperScript III (Life Technologies). The design of the qPCR primers was made using Primer3 program

(Table 2). Amplification reactions were conducted on the Viia7 thermal cycling system (Applied Biosystems) using SYBR Green Reagent (Thermo Fisher Scientific). Expression levels were normalized to housekeeping genes using ΔΔCT method.

**Bulk RNA sequencing**

For NPC and EpiLC bulk RNA sequencing, total DNase-treated RNA was isolated as described above. RNA-seq libraries were prepared from a minimum of 200 ng of DNase-treated total RNAs with the TRuSeq Stranded mRNA protocol (Illumina). Between 59M and 130M reads were sequenced per sample in a 100bp paired-end format using a NovaSeq 6000 (Illumina) platform. For PGCLC bulk RNA sequencing, libraries were generated from 50ng of isolated of DNase-treated total RNA using SMARTer Stranded Total RNA-seq Kit for mammalian Pico Input. Between 38M and 45M reads were sequenced per sample.

**Single-cell RNA sequencing**

Single cells were index-sorted individually by FACS (SONY SH800) into wells of a 96-well PCR plate containing 2.3 µl of lysis buffer (200U/µL SUPERase-IN (LifeTechnologies), 10% Triton X100, RNAse free H2O). scRNA-seq SMARTseq2 was performed as previously described (Picelli et al., 2014). The Illumina Nextera XT DNA kit was used to prepare libraries that were sequenced on the Illumina HiSeq 4000 (paired-end 50bp). Samples from all cell lines were included in each sequencing lane, to control for technical lane effects. Only cells with over 7500 detected reads assigned to genes were kept in the analysis.

**Cut&Run**

The Cut&Run protocol was optimized from Skene and Henikoff, 2017. Concanavalin A beads (Polysciences) were activated in 1ml of Binding Buffer (20mM HEPES-KOH, 10mM KCl, 1mM $CaCl_2$, 10mM $MnCl_2$). Between 200 000 and 300 000 cells were aliquoted in 1ml of Wash Buffer (20mM HEPES-KOH, 150mM NaCl, 0.5mM Spermidine (Sigma) and 1X Complete EDTA-free (Roche) protease inhibitor cocktail) and 20µL of activated beads were added to the cells, followed by 5 min incubation at RT on a rotating wheel. Cells were collected on a magnetic rack; the supernatant was replaced by 400µL of primary antibody (Table 1) diluted to 1:200 in Wash Buffer complemented with 2mM EDTA and 0.02% Digitonin (Millipore). Cells were incubated for one hour at RT. Samples were then washed twice in Dig-Wash Buffer (Wash Buffer + 0.02% Digitonin) for 15 min at RT and incubated with 1:400 home-made pA-MNase in Dig-Wash for 15 min at RT. After two more washes, cells were resuspended in 150µl of Dig-Wash and cooled-down for five min at 0°C. For targeted digestion, $CaCl_2$ was added to the cells to a final concentration of 2mM followed by 30 min incubation on ice. The reaction was stopped by adding 150µl of 2X STOP Buffer (340mM NaCl, 20mM EDTA, 4mM EGTA, 0.02% Digitonin, 100ng/ml RNase1, 250ng/ml glycogen) and incubating the samples at 37°C for 20 min before centrifugation at 4°C 16000g

for 5 min. Samples were placed on a magnetic rack and supernatant was recovered in a low binding eppendorf tube. Following addition of 0.1% SDS and 0.17mg/ml Proteinase K, samples were incubated at 70°C for 30 min. Purified DNA was obtained by phenol/chloroform extraction and precipitated with 100% Ethanol incubation at -20°C for at least 20 min. The DNA pellet was washed in 85% ethanol, spun down and air-dried before being resuspended in 25µl of 1mM Tris-HCl pH8.0.

Library preparation was made according to the manufacturer's instructions (Accel NGS 2S DNA library kit, Swift Biosciences). For library amplification, 12 PCR cycles were applied as followed: 98°C for 10 sec, 60°C for 15 sec and 68°C for one minute. Library size and quantification was assessed on Agilent 4200 Tapestation. Cut&Run libraries were sequenced on NovaSeq (Illumina) using PE 100bp run, with biological duplicates for all samples. Between 25M and 75M reads were sequenced per sample.

**ATAC sequencing**

The ATAC protocol was optimized from Fursova et al., 2021. First, $1x10^7$ ESCs were mixed with $2.5x10^6$ *Drosophila* S2 cells, and resuspended and incubated in 1ml Lysis Buffer (50mM KCl, 10mM $MgSO_4.7H_20$, 5mM HEPES, 0.05% NP40) for 5 min at RT. After a spin at 4°C 1500g for 5 min, the supernatant was discarded and nuclei pellet was resuspended in ice-cold RS Buffer (10mM NaCl, 10mM Tris-HCl, 3mM $MgCl_2$). Nuclei were counted using Beckman Coulter ViCell XR cyto-counter, to ensure nuclei integrity. Tagmentation was performed on 50 000 nuclei per sample, with TDE1 Tn5 transposase (Nextera), at 37°C for 30 min. The reaction was stopped by the addition of PB Buffer and tagmented DNA was eluted using minElute columns from Reaction clean-up Qiagen Kit. We used 50ng of tagmented genomic DNA isolated from the same nuclei preparation to control for Tn5 transposition bias. Libraries were generated by PCR amplification (maximum of 12 cycles) as previously described in Buenrostro et al., 2013 using custom-made barcoded primers. Libraries were cleaned on minElute columns and quality was assessed using Agilent 4200 Tapestation. Sequencing was performed on NovaSeq platform (Illumina) using PE 100bp run. Between 40M and 91Mreads were sequenced per sample.

**Definition of PGCLC-NPC-EpiLC-specific genes**

To define PGCLC and NPC specific genes, differential gene expression was compared between NPCs at D8, PGCLCs after 40h of EpiLCs induction, and primed EpiLCs at D7. Genes were declared as NPC-specific or PGCLC-specific if FDR < 0.05 and log2(CPM) > 1 when compared to the other two states (comparisons were NPC vs PGCLC/EpiLCD7, and PGCLC vs NPC/EpiLCD7). To define EpiLC-specific genes, the same strategy and thresholds were applied, using EpiLC at D2 instead of EpiLC at D7 to capture transiently expressed "formative" epiblast genes (comparisons were EpiLC D2 vs PGCLC/NPC).

**Bulk RNA-seq analysis**

Paired-end read alignment was made using Kallisto pseudo-aligner (v0.46.2). Gene annotation was obtained from Ensembl (EnsDb.Mmusculus.v79). After EdgeR (v3.30.3) CPM normalization, differential gene expression analysis was made using Voom transformation from Limma R package (v3.44.3). P-Values were adjusted with the BHH correction, cutoff was set up at 0.05. Threshold for differentially expressed genes were set at the FDR<5% and |log2FC| >1.

**Single-cell RNA-seq analysis**

Paired-end reads were trimmed with trim Galore (v0.6.5) in order to remove adapter sequences and Ns nucleotides from both side of the read. Cleaned reads were aligned onto a concatenated genome (the Mouse reference genome mm10 + fasta file with the ERCC spike-in sequences) with STAR (v2.7.5a) reporting random alignment with at most 6 % of mismatches. ERCCs, genes and TE families-based quantification was performed with featureCounts (v1.5.1) using a concatenated file with repeatMasker, ERCC spike-in and Gencode vM25 annotations. Filtering of cells was performed with the following criteria: cells with less than 7500 genes with at least one read and cells with more than 2% of mitochondrial reads or more than 10% of ribosomal reads were filtered out. On the remaining cells, the raw counts were normalized with the SCTransform from Seurat (v4.0.1). PCA was performed on the 10000 most variable genes defined by the vst method. UMAP was performed on the 10 first principal components with 10 neighbors, a minimal distance of 0.1 and a spread of 1. Cells clusters were defined with the FindNeighbors and FindClusters functions from Seurat on 5 first principal components with 14 neighbors and a resolution of 0.4. Pseudotime was inferred with the R package slingshot (v1.8.0) on the 10 first principal components. Transcriptional noise analysis was conducted with the same procedure as in Mohammed et al., 2017.

**CUT&RUN analysis**

Paired-end reads were trimmed with trim Galore (v0.4.4) in order to remove adapter sequences and Ns nucleotides from both side of the read. Cleaned reads were aligned onto the Mouse reference genome mm10 with bowtie2 (v2.2.9) in end-to-end and very sensitive mode. Duplicated reads were removed using Picard (v2.6.0). Bigwig files for UCSC genome browser and heatmaps were created with deepTools (v2.5.3). Sliding windows approach was used to define enriched regions with csaw R package (v1.24.3). Window size of 150bp and 5 kb was used for H3K27ac and H3K27me3 samples with a spacing interval of 38bp and 1250bp respectively. Window-level counts were quantified using the *windowCounts* function, discarding blacklisted regions (ENCODE, accession ENCFF547MET), with mapping quality threshold of 20. For each genotype (WT and TKO), non-insteresting windows were filtered out using a global background enrichment (*filterWindowsGlobal* function) with large bins (750bp for H3K27ac and 25kb for H3K27me3 samples). Normalization factors were

calculated using the large bins and used for normalization of enriched windows to remove composition biases. Differential binding analysis was performed with csaw. Windows that are less than 100bp and 1kb (for H3K27ac and H3K27me3 samples respectively) apart were merged into regions. Regions with a FDR<5 % were declared as differentially bound.

**ATAC-seq analysis**

Paired-end reads were trimmed with trim Galore (v0.4.4) in order to remove adapter sequences and Ns nucleotides from both side of the read. Cleaned reads were aligned onto the Mouse reference genome mm10 with bowtie2 (v2.2.9) in end-to-end and very sensitive mode. Mitochondrial reads were removed from the rest of the analysis. Duplicated reads were removed using Picard (v2.6.0). Bigwig files for UCSC genome browser and heatmaps were created with deepTools (v2.5.3). Enriched regions were defined using MACS2 (v2.1.1) with a q-value threshold of 0.01, discarding the ones that overlap with blacklisted regions (ENCODE, accession ENCFF547MET). Overlapping regions between biological replicates were used for the following analyses. Region-level counts were quantified using the *windowCounts* function with mapping quality threshold of 20. Normalization factors were calculated using 10kb-long bins to remove composition biases. Differential binding analysis was performed with csaw. Regions that are less than 500bp apart were merged. Regions with a FDR<5 % were declared as differentially bound.

**Integration of CUT&RUN and ATAC-seq data**

Genes annotated in close proximity (less than 5kb to the TSS) to differentially bound regions were used for gene ontology enrichment analysis with the R clusterProfiler package (*enrichGO* function). Overlaps between differentially bound regions defined in different time points and epigenetics marks were performed using intervene (v0.6.4). Enhancers were defined as the co-occurrence of ATAC and H3K27ac-enriched regions that are distal from promoter regions (+/- 2kb around the TSS). PGCLC and NPC enhancers were defined as enhancers that are +/- 50kb from the TSS of PGC and NPC genes that were defined as previously explained. Quantification of NPC, PGCLC, endoderm, mesoderm enhancers and bivalent genes was performed with featureCounts (v1.5.1). Bivalent genes in EpiLCs and mesoderm/endoderm enhancers were originally defined in Argelaguet et al., 2019 and Kurimoto et al., 2015. Enrichment of known motifs and *de novo* motif finding were performed using HOMER (v4.11) (Heinz et al., 2010) with the default database. Transcription factor activity, also called footprinting, was assessed using the software TOBIAS (v.0.13.3) (Bentsen et al., 2020). Tn5 bias correction was first performed using the ATACorrect module. Then footprinting scores were calculated using the corrected cut sites with ScoreBigwig module. Differential binding score between conditions was estimated with BINDetect module with the HOMER motif database.

**Softwares**

All the statistical analyses and graphs were performed using GraphPad Prism (v9.4).
All microscopy image analysis were made possible using Fiji (ImageJ2 v2.3.0).
All transgene constructions were made possible with Geneious (v7.1)
All figures were made with Affinity Designer (v1.8.6).

# REFERENCES


Akagi, Kuure, Uranishi, Koide, Costantini, Yokota. 2015. ETS-related transcription factors Etv4 and Etv5 are involved in proliferation and induction of differentiationassociated genes in embryonic stem (ES) cells. *J Biol Chem* **290**:22460–22473. doi:10.1074/jbc.M115.675595

Aramaki, Kagiwada, Wu, Obridge, Adachi, Kutejova, Lickert, Hübner, Schöler. 2021. Residual pluripotency is required for inductive germ cell segregation. *EMBO Rep* 1–15. doi:10.15252/embr.202152553

Argelaguet, Clark, Mohammed, Stapel, Krueger, Kapourani, Imaz-Rosshandler, Lohoff, Xiang, Hanna, Smallwood, Ibarra-Soria, Buettner, Sanguinetti, Xie, Krueger, Göttgens, Rugg-Gunn, Kelsey, Dean, Nichols, Stegle, Marioni, Reik. 2019. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**:487–491. doi:10.1038/s41586-019-1825-8

Banerjee, Saxena, Kumar, Chen, Cavazza, Toke, O'Neill, Madha, Jadhav, Verzi, Shivdasani. 2018. Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development. *Genes Dev* **32**:1430–1442. doi:10.1101/gad.318832.118

Bedzhov, Zernicka-Goetz. 2014. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**:1032–1044. doi:10.1016/j.cell.2014.01.023

Bell, Curry, Megchelenbrink, Jouneau, Brochard, Tomaz, Mau, Atlasi, de Souza, Marks, Stunnenberg, Jouneau, Azuara. 2020. Dynamic CpG methylation delineates subregions within super-enhancers selectively decommissioned at the exit from naive pluripotency. *Nat Commun* **11**:1112. doi:10.1038/s41467-020-14916-7

Bentsen, Goymann, Schultheis, Klee, Petrova, Wiegandt, Fust, Preussner, Kuenne, Braun, Kim, Looso. 2020. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat Commun* **11**. doi:10.1038/s41467-020-18035-1

Betto, Diamante, Perrera, Audano, Rapelli, Lauria, Incarnato, Arboit, Pedretti, Rigoni, Guerineau, Touboul, Stirparo, Lohoff, Boroviak, Grumati, Soriano, Nichols, Mitro, Oliviero, Martello. 2021. Metabolic control of DNA methylation in naive pluripotent cells. *Nat Genet*. doi:10.1038/s41588-020-00770-2

Blaschke, Ebata, Karimi, Zepeda-Martínez, Goyal, Mahapatra, Tam, Laird, Hirst, Rao, Lorincz, Ramalho-Santos. 2013. VitaminC induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**:222–226. doi:10.1038/nature12362

Bleckwehl, Crispatzu, Schaaf, Respuela, Bartusel, Benson, Clark, Dorighi, Barral, Laugsch, van IJcken, Manzanares, Wysocka, Reik, Rada-Iglesias. 2021. Enhancer-associated H3K4 methylation safeguards in vitro germline competence. *Nat Commun* **12**:5771. doi:10.1038/s41467-021-26065-6

Bleckwehl, Rada-Iglesias. 2019. Transcriptional and epigenetic control of germline competence and specification. *Curr Opin Cell Biol* **61**:1–8. doi:10.1016/j.ceb.2019.05.006

Borgel, Guibert, Li, Chiba, Schübeler, Sasaki, Forné, Weber. 2010. Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* **42**:1093–1100. doi:10.1038/ng.708

Brinkman, Gu, Bartels, Zhang, Matarese, Simmer, Marks, Bock, Gnirke, Meissner, Stunnenberg. 2012. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res* **22**:1128–1138. doi:10.1101/gr.133728.111

Buecker, Srinivasan, Wu, Calo, Acampora, Faial, Simeone, Tan, Swigut, Wysocka. 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**:838–53. doi:10.1016/j.stem.2014.04.003

Buenrostro, Giresi, Zaba, Chang, Greenleaf. 2013. Transposition of native chromatin for fast and

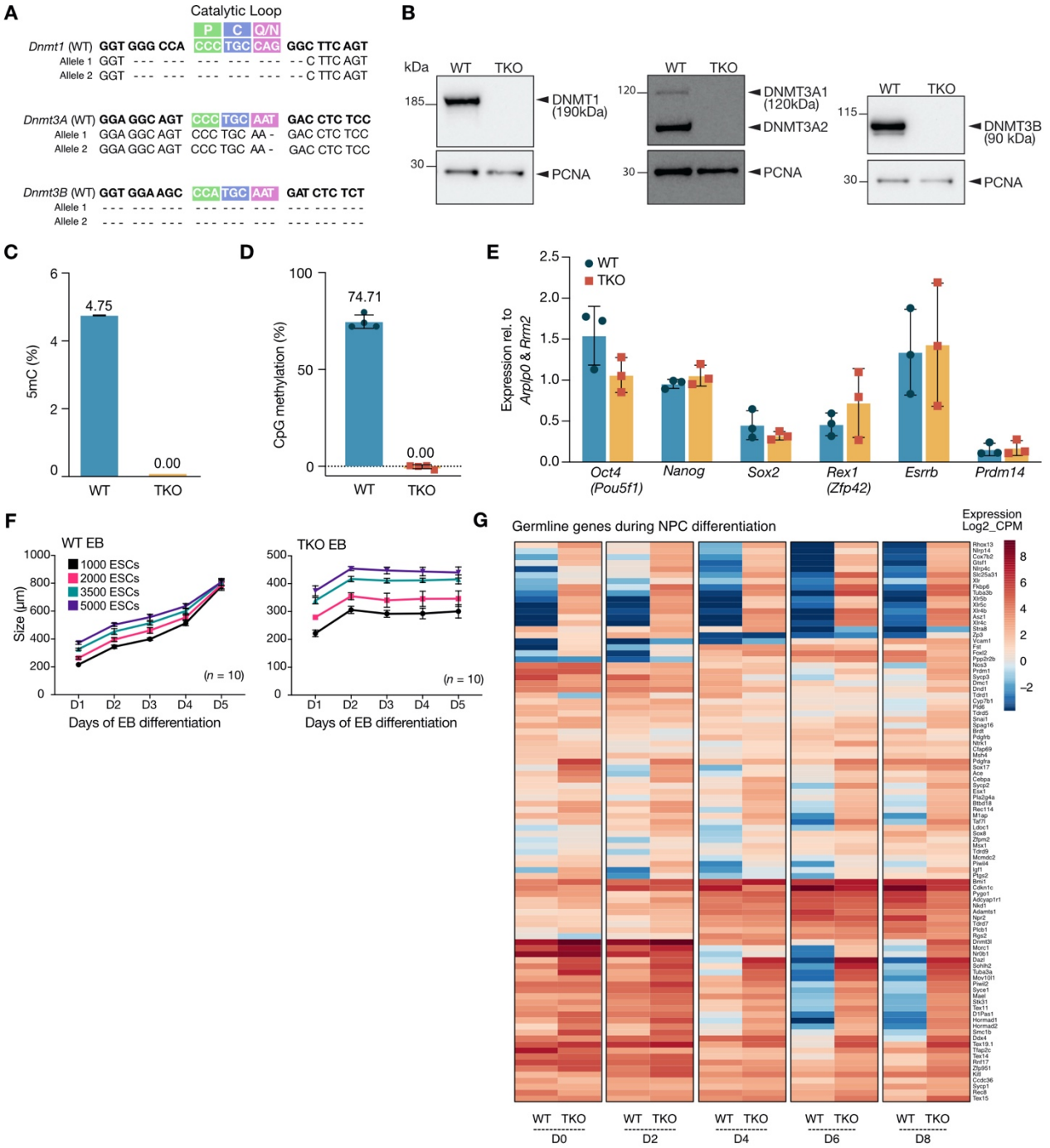
sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**:1213–1218. doi:10.1038/nmeth.2688
Chambers, Silva, Colby, Nichols, Nijmeijer, Robertson, Vrana, Jones, Grotewold, Smith. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature* **450**:1230–4. doi:10.1038/nature06403
Charlton, Jung, Mattei, Bailly, Liao, Martin, Giesselmann, Brändl, Stamenova, Müller, Kiskinis, Gnirke, Smith, Meissner. 2020. TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. *Nat Genet*. doi:10.1038/s41588-020-0639-9
Choi, Joo, Hong, Kim, Song, Lee, Wu, Schöler, Do. 2016. Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency. *Stem Cell Reports* **7**:911–926. doi:10.1016/j.stemcr.2016.09.012
Dahlet, Argüeso Lleida, Al Adhami, Dumas, Bender, Ngondo, Tanguy, Vallet, Auclair, Bardet, Weber. 2020. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. *Nat Commun* **11**:1–14. doi:10.1038/s41467-020-16919-w
Dahlet, Truss, Frede, Al Adhami, Bardet, Dumas, Vallet, Chicher, Hammann, Kottnik, Hansen, Luz, Alvarez, Auclair, Hecht, Robinson, Hagemeier, Weber. 2021. E2F6 initiates stable epigenetic silencing of germline genes during embryonic development. *Nat Commun* **12**:3582. doi:10.1038/s41467-021-23596-w
Domcke, Bardet, Adrian Ginno, Hartl, Burger, Schübeler. 2015. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* **528**:575–579. doi:10.1038/nature16462
Dura, Teissandier, Armand, Barau, Lapoujade, Fouchet, Bonneville, Schulz, Weber, Baudrin, Lameiras, Bourc'his. 2022. DNMT3A-dependent DNA methylation is required for spermatogonial stem cells to commit to spermatogenesis. *Nat Genet* **54**:469–480. doi:10.1038/s41588-022-01040-z
Eldar, Elowitz. 2010. Functional roles for noise in genetic circuits. *Nature* **467**:167–173. doi:10.1038/nature09326
Fursova, Turberfield, Blackledge, Findlater, Lastuvkova, Huseyin, Dobrinic, Klose. 2021. BAP1 constrains pervasive H2AK119ub1 to control the transcriptional potential of the genome. *Genes Dev* **35**:749–770. doi:10.1101/GAD.347005.120
Gao, Yang, Chen, Wang, Wu, Ai, Du, Liu, Shi, Wu, Guo, Zhang. 2013. Vitamin C facilitates pluripotent stem cell maintenance by promoting pluripotency gene transcription. *Biochimie* **95**:2107–2113. doi:10.1016/j.biochi.2013.08.001
Gibson, Young, Chuang, Venter, Hutchison, Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**:343–345. doi:10.1038/nmeth.1318
Ginno, Gaidatzis, Feldmann, Hoerner, Imanci, Burger, Zilbermann, Peters, Edenhofer, Smallwood, Krebs, Schübeler. 2020. A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nat Commun* **11**:2680. doi:10.1038/s41467-020-16354-x
Greenberg, Teissandier, Walter, Noordermeer, Bourc'his. 2019. Dynamic enhancer partitioning instructs activation of a growth-related gene during exit from naïve pluripotency. *Elife* **8**:1–20. doi:10.7554/eLife.44057
Gretarsson, Hackett. 2020. Dppa2 and Dppa4 counteract de novo methylation to establish a permissive epigenome for development. *Nat Struct Mol Biol* **27**:706–716. doi:10.1038/s41594-020-0445-1
Gruhn, Günesdogan. 2021. Generation of Primordial Germ Cell-like Cells on Small and Large Scales. pp. 75–89. doi:10.1007/978-1-0716-0958-3_6
Hayashi, Ohta, Kurimoto, Aramaki, Saitou. 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**:519–532. doi:10.1016/j.cell.2011.06.052
Héberlé, Bardet. 2019. Sensitivity of transcription factors to DNA methylation. *Essays Biochem* **63**:727–741. doi:10.1042/EBC20190033
Heinz, Benner, Spann, Bertolino, Lin, Laslo, Cheng, Murre, Singh, Glass. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**:576–589. doi:10.1016/j.molcel.2010.05.004
Huang, Wang, Vazquez-Ferrer, Theofel, Requena, Hanna, Kelsey, Hajkova. 2021. Sex-specific chromatin remodelling safeguards transcription in germ cells. *Nature* **600**. doi:10.1038/s41586-021-04208-5
Kalkan, Olova, Roode, Mulas, Lee, Nett, Marks, Walker, Stunnenberg, Lilley, Nichols, Reik, Bertone,

Smith. 2017. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**:1221–1234. doi:10.1242/dev.142711
Karimi, Johansson, Stach, Corcoran, Grandér, Schalling, Bakalkin, Lyko, Larsson, Ekström. 2006. LUMA (LUminometric Methylation Assay)—A high throughput method to the analysis of genomic DNA methylation. *Exp Cell Res* **312**:1989–1995. doi:10.1016/j.yexcr.2006.03.006
Kim, Wu, Rahme, Battaglia, Dixit, Gaskell, Chen, Pinello, Bernstein. 2020. Parallel Single-Cell RNA-Seq and Genetic Recording Reveals Lineage Decisions in Developing Embryoid Bodies. *Cell Rep* **33**:108222. doi:10.1016/j.celrep.2020.108222
King, Huang, Rubbi, Liu, Wang, Wang, Pellegrini, Fan. 2016. Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *Cell Rep* **17**:289–302. doi:10.1016/j.celrep.2016.08.083
Kinoshita, Barber, Mansfield, Cui, Spindlow, Stirparo, Dietmann, Nichols, Smith. 2020. Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* 2020.09.04.283218. doi:10.1016/j.stem.2020.11.005
Kinoshita, Li, Barber, Mansfield, Dietmann, Smith. 2021. Disabling de novo DNA methylation in embryonic stem cells allows an illegitimate fate trajectory. *Proc Natl Acad Sci U S A* **118**:1–11. doi:10.1073/pnas.2109475118
Kinoshita, Smith. 2018. Pluripotency Deconstructed. *Dev Growth Differ*. doi:10.1111/dgd.12419
Koike, Sakaki, Amano, Kurosawa. 2007. Characterization of embryoid bodies of mouse embryonic stem cells formed under various culture conditions and estimation of differentiation status of such bodies. *J Biosci Bioeng* **104**:294–299. doi:10.1263/jbb.104.294
Kreibich, Kleinendorst, Barzaghi, Kaspar, Krebs. 2022. Single molecule multi-omics reveals context-dependent regulation of enhancers by DNA methylation. *Biorxv* 1–23.
Kurimoto, Yabuta, Hayashi, Ohta, Kiyonari, Mitani, Moritoki, Kohri, Kimura, Yamamoto, Katou, Shirahige, Saitou. 2015. Quantitative dynamics of chromatin remodeling during germ cell specification from mouse embryonic stem cells. *Cell Stem Cell* **16**:517–532. doi:10.1016/j.stem.2015.03.002
Lawson, Dunn, Roelen, Zeinstra, Davis, Wright, Korving, Hogan. 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* **13**:424–436. doi:10.1101/gad.13.4.424
Leaky, Weixiong, Kuhnert, Stuhlmann. 1999. Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation. *J Exp Zool* **284**:67–81. doi:10.1002/(SICI)1097-010X(19990615)284:1<67::AID-JEZ10>3.0.CO;2-O
Lei, Oh, Okano, Jüttermann, Goss, Jaenisch, Li. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**:3195–205.
Leitch, Okamura, Durcova-Hills, Stewart, Gardner, Matsui, Papaioannou. 2014. On the fate of primordial germ cells injected into early mouse embryos. *Dev Biol* **385**:155–159. doi:10.1016/j.ydbio.2013.11.014
Leitch, Smith. 2013. The mammalian germline as a pluripotency cycle. *Development* **140**:2495–2501. doi:10.1242/dev.091603
Li, Bestor, Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**:915–926. doi:10.1016/0092-8674(92)90611-F
Luo, van der Veer, Sun, Bartoccetti, Boretto, Vankelecom, Khoueiry, Koh. 2020. Coordination of germ layer lineage choice by TET1 during primed pluripotency. *Genes Dev* 1–21. doi:10.1101/gad.329474.119
Mayer, Stadler, Rittirsch, Hess, Lukonin, Winzi, Smith, Buchholz, Betschinger. 2020. Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *EMBO J* **39**:1–22. doi:10.15252/embj.2019102591
McLaughlin, Flyamer, Thomson, Mjoseng, Shukla, Williamson, Grimes, Illingworth, Adams, Pennings, Meehan, Bickmore. 2019. DNA Methylation Directs Polycomb-Dependent 3D Genome Re-organization in Naive Pluripotency. *Cell Rep* **29**:1974-1985.e6. doi:10.1016/j.celrep.2019.10.031
Mochizuki, Sharif, Shirane, Uranishi, Bogutz, Janssen, Suzuki, Okuda, Koseki, Lorincz. 2021. Repression of germline genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing. *Nat Commun* **12**:1–8. doi:10.1038/s41467-021-27345-x
Mohammed, Hernando-Herraez, Savino, Scialdone, Macaulay, Mulas, Chandra, Voet, Dean, Nichols, Marioni, Reik. 2017. Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep* **20**:1215–1228. doi:10.1016/j.celrep.2017.07.009

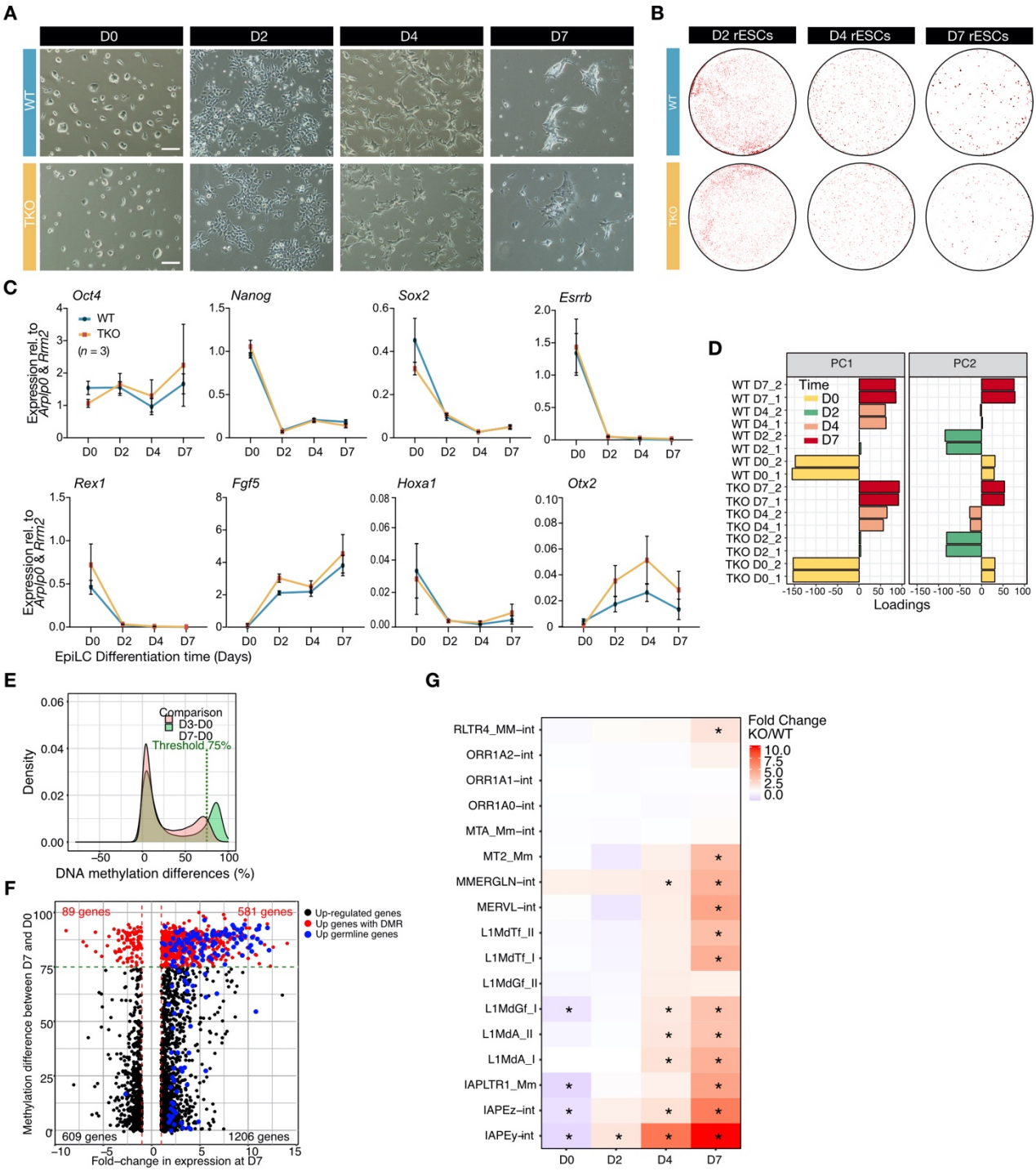
Morgani, Nichols, Hadjantonakis. 2017. The many faces of Pluripotency: In vitro adaptations of a continuum of in vivo states. *BMC Dev Biol* **17**:10–12. doi:10.1186/s12861-017-0150-4
Mulas, Kalkan, Smith. 2017. NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from the Ground State. *Stem Cell Reports* **9**:77–91. doi:10.1016/j.stemcr.2017.05.033
Mulas, Kalkan, von Meyenn, Leitch, Nichols, Smith. 2019. Defined conditions for propagation and manipulation of mouse embryonic stem cells. *Development* **146**:dev173146. doi:10.1242/dev.173146
Nichols, Smith. 2009. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**:487–492. doi:10.1016/j.stem.2009.05.015
Ohinata, Payer, O'Carroll, Ancelin, Ono, Sano, Barton, Obukhanych, Nussenzweig, Tarakhovsky, Saitou, Surani. 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**:207–213. doi:10.1038/nature03813
Ohtsuka, Nishikawa-Torikai, Niwa. 2012. E-Cadherin Promotes Incorporation of Mouse Epiblast Stem Cells into Normal Development. *PLoS One* **7**. doi:10.1371/journal.pone.0045220
Okano, Bell, Haber, Li. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**:247–257. doi:10.1016/S0092-8674(00)81656-6
Pei, Mao, Li, Li, Xu, Xu, Shi, Gao, Hu, Hu, Chen, Wu, Liu, Zhang, Zhang, Guo, Tang, Dai. 2014. Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. *Cell Stem Cell* **14**:512–522. doi:10.1016/j.stem.2014.01.001
Raj, van Oudenaarden. 2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135**:216–226. doi:10.1016/j.cell.2008.09.050
Rauch, Haakonsson, Madsen, Larsen, Forss, Madsen, Van Hauwaert, Wiwie, Jespersen, Tencerova, Nielsen, Larsen, Röttger, Baumbach, Scheele, Kassem, Mandrup. 2019. Osteogenesis depends on commissioning of a network of stem cell transcription factors that act as repressors of adipogenesis. *Nat Genet* **51**:716–727. doi:10.1038/s41588-019-0359-1
Riising, Comet, Leblanc, Wu, Johansen, Helin. 2014. Gene Silencing Triggers Polycomb Repressive Complex 2 Recruitment to CpG Islands Genome Wide. *Mol Cell* **55**:347–360. doi:10.1016/j.molcel.2014.06.005
Rulands, Lee, Clark, Angermueller, Smallwood, Krueger, Mohammed, Dean, Nichols, Rugg-Gunn, Kelsey, Stegle, Simons, Reik. 2018. Genome-Scale Oscillations in DNA Methylation during Exit from Pluripotency. *Cell Syst* **7**:63-76.e12. doi:10.1016/j.cels.2018.06.012
Sachs, Onodera, Blaschke, Ebata, Song, Ramalho-Santos. 2013. Bivalent Chromatin Marks Developmental Regulatory Genes in the Mouse Embryonic Germline InVivo. *Cell Rep* **3**:1777–1784. doi:10.1016/j.celrep.2013.04.032
Sakaue, Ohta, Kumaki, Oda, Sakaide, Matsuoka, Yamagiwa, Niwa, Wakayama, Okano. 2010. DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr Biol* **20**:1452–1457. doi:10.1016/j.cub.2010.06.050
Sancho, Di-Gregorio, George, Pozzi, Sánchez, Pernaute, Rodríguez. 2013. Competitive interactions eliminate unfit embryonic stem cells at the onset of differentiation. *Dev Cell* **26**:19–30. doi:10.1016/j.devcel.2013.06.012
Sankar, Mohammad, Sundaramurthy, Wang, Lerdrup, Tatar, Helin. 2022. Histone editing elucidates the functional roles of H3K27 methylation and acetylation in mammals. *Nat Genet*. doi:10.1038/s41588-022-01091-2
Schmidt, Bultmann, Meilinger, Zacher, Tresch, Maier, Peter, Martin, Leonhardt, Spada. 2012. Global DNA Hypomethylation Prevents Consolidation of Differentiation Programs and Allows Reversion to the Embryonic Stem Cell State. *PLoS One* **7**:e52629. doi:10.1371/journal.pone.0052629
Schomacher, Han, Musheev, Arab, Kienhöfer, Von Seggern, Niehrs. 2016. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nat Struct Mol Biol* **23**:116–124. doi:10.1038/nsmb.3151
Seisenberger, Andrews, Krueger, Arand, Walter, Santos, Popp, Thienpont, Dean, Reik. 2012. The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. *Mol Cell* **48**:849–862. doi:10.1016/j.molcel.2012.11.001
Shirane, Kurimoto, Yabuta, Yamaji, Satoh, Ito, Watanabe, Hayashi, Saitou, Sasaki. 2016. Global Landscape and Regulatory Principles of DNA Methylation Reprogramming for Germ Cell Specification by Mouse Pluripotent Stem Cells. *Dev Cell* **39**:87–103. doi:10.1016/j.devcel.2016.08.008
Skene, Henikoff. 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA

binding sites. *Elife* **6**:1–35. doi:10.7554/eLife.21856
Smith. 2017. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**:365–373. doi:10.1242/dev.142679
Suzuki, Maeda, Furuhata, Shimizu, Nishimura, Kishima, Suzuki. 2017. A screening system to identify transcription factors that induce binding site-directed DNA demethylation. *Epigenetics and Chromatin* **10**:1–14. doi:10.1186/s13072-017-0169-6
Tischler, Gruhn, Reid, Allgeyer, Buettner, Marr, Theis, Simons, Wernisch, Surani. 2018. Metabolic regulation of pluripotency and germ cell fate through α-ketoglutarate. *EMBO J* e99518. doi:10.15252/embj.201899518
Tsumura, Hayakawa, Kumaki, Takebayashi, Sakaue, Matsuoka, Shimotohno, Ishikawa, Li, Ueda, Nakayama, Okano. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to Cells* **11**:805–814. doi:10.1111/j.1365-2443.2006.00984.x
Velasco, Hube, Rollin, Neuillet, Philippe, Bouzinba-Segard, Galvani, Viegas-Pequignot, Francastel. 2010. Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc Natl Acad Sci* **107**:9281–9286. doi:10.1073/pnas.1000473107
von Meyenn, Berrens, Andrews, Santos, Collier, Krueger, Osorno, Dean, Rugg-Gunn, Reik. 2016. Comparative Principles of DNA Methylation Reprogramming during Human and Mouse In Vitro Primordial Germ Cell Specification. *Dev Cell* **39**:104–115. doi:10.1016/j.devcel.2016.09.015
Walter, Teissandier, Pérez-Palacios, Bourc'his. 2016. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife* **5**:1–30. doi:10.7554/eLife.11418
Yamaji, Seki, Kurimoto, Yabuta, Yuasa, Shigeta, Yamanaka, Ohinata, Saitou. 2008. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* **40**:1016–1022. doi:10.1038/ng.186
Yang, Andrabi, Biss, Baker, Iqbal, Sharrocks. 2018. ZIC3 Controls the Transition from Naïve to Primed Pluripotency. *Cell Rep* **27**:3215-3227.e6. doi:10.2139/ssrn.3305317
Ying, Stavridis, Griffiths, Li, Smith. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21**:183–186. doi:10.1038/nbt780
Zhang, Xiang, Yin, Du, Peng, Wang, Fidalgo, Xia, Li, Zhao, Zhang, Ma, Xu, Wang, Li, Xie. 2018. Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat Genet* **50**:96–105. doi:10.1038/s41588-017-0003-x
Zhu, Wang, Qian. 2016. Transcription factors as readers and effectors of DNA methylation. *Nat Rev Genet* **17**:551–65. doi:10.1038/nrg.2016.83

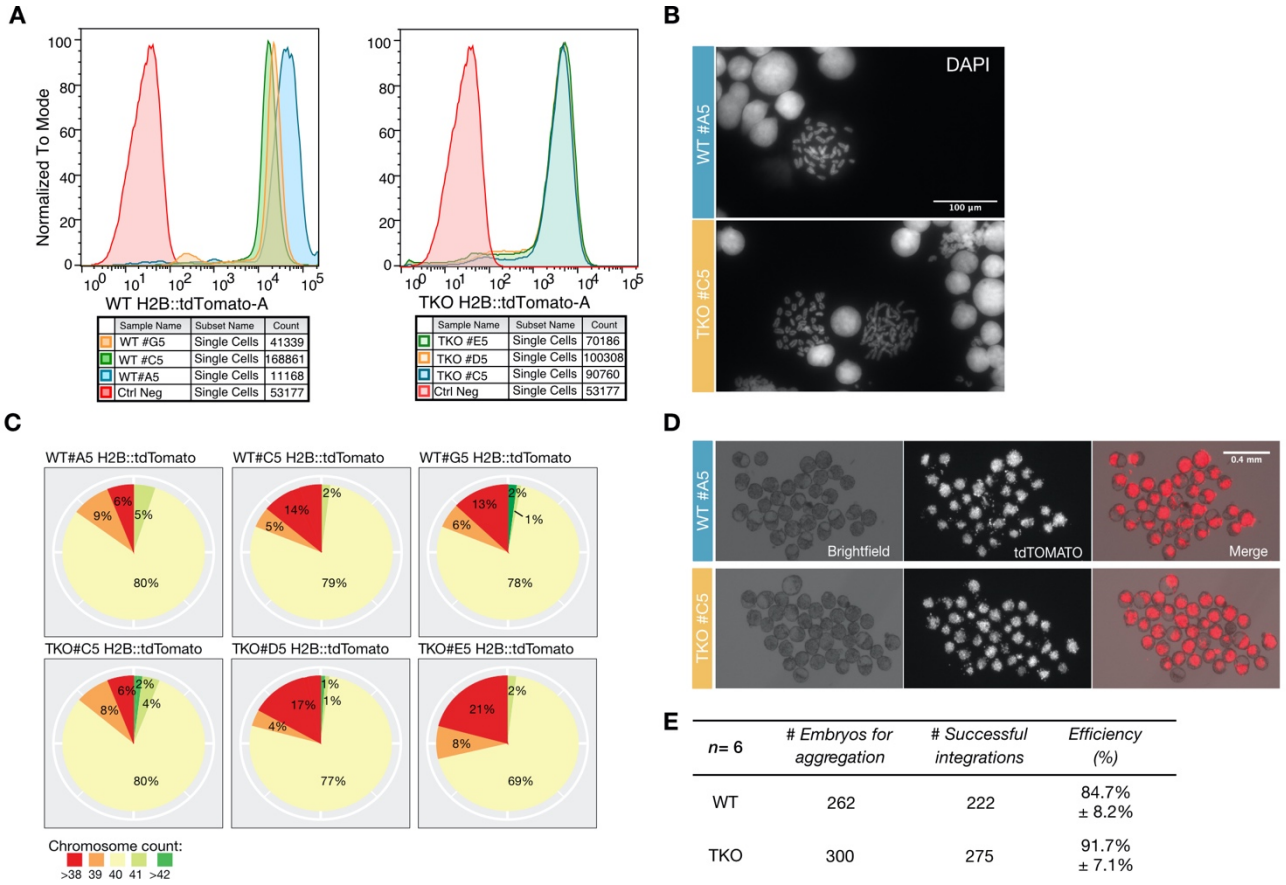
# EXTENDED FIGURES AND TABLES




**Extended Fig. 1 I DNA-methylation-free *Dnmt*-triple KO (TKO) ESCs: generation and NPC differentiation. A,** Sequencing results for the E14 *Dnmt*-triple-KO (TKO) ESC line showing allelic homozygous deletions of *Dnmt* genes after catalytic motif targeting by CRIPSR-Cas9. **B,** Immunoblotting showing DNMT1, DNMT3A and DNMT3B in WT and TKO ESCs grown in Serum/Lif). PCNA is used as a loading control. **C,** Barplot showing LC/MS quantification of methylated cytosines in serum/LIF-grown WT (blue) and TKO (orange) ESCs. Data shown are mean ± SD in technical duplicates. **D,** LUMA assay of genome-wide CpG methylation quantification of serum/LIF-grown WT and TKO ESCs. Data shown are *n*= 4 biological replicates with mean ± SD. **E,** Expression of pluripotency genes measured by RT-qPCR in 2i/LIF-grown WT and TKO ESCs. Data shown are mean ± SD from biological triplicates, ΔCT values are normalized to *Arplp0* and *Rrm2*. **F,** EBies diameter size curve showing the effect of increasing initial dilution on EBies growth during five-day long differentiation. Data shown are biological replicates (*n*=10) with mean ± SD. **G,** Heatmaps of germline markers showing normalized Log2 CPM counts during NPC differentiation. Expression is obtained for each condition from the average between biological duplicates.



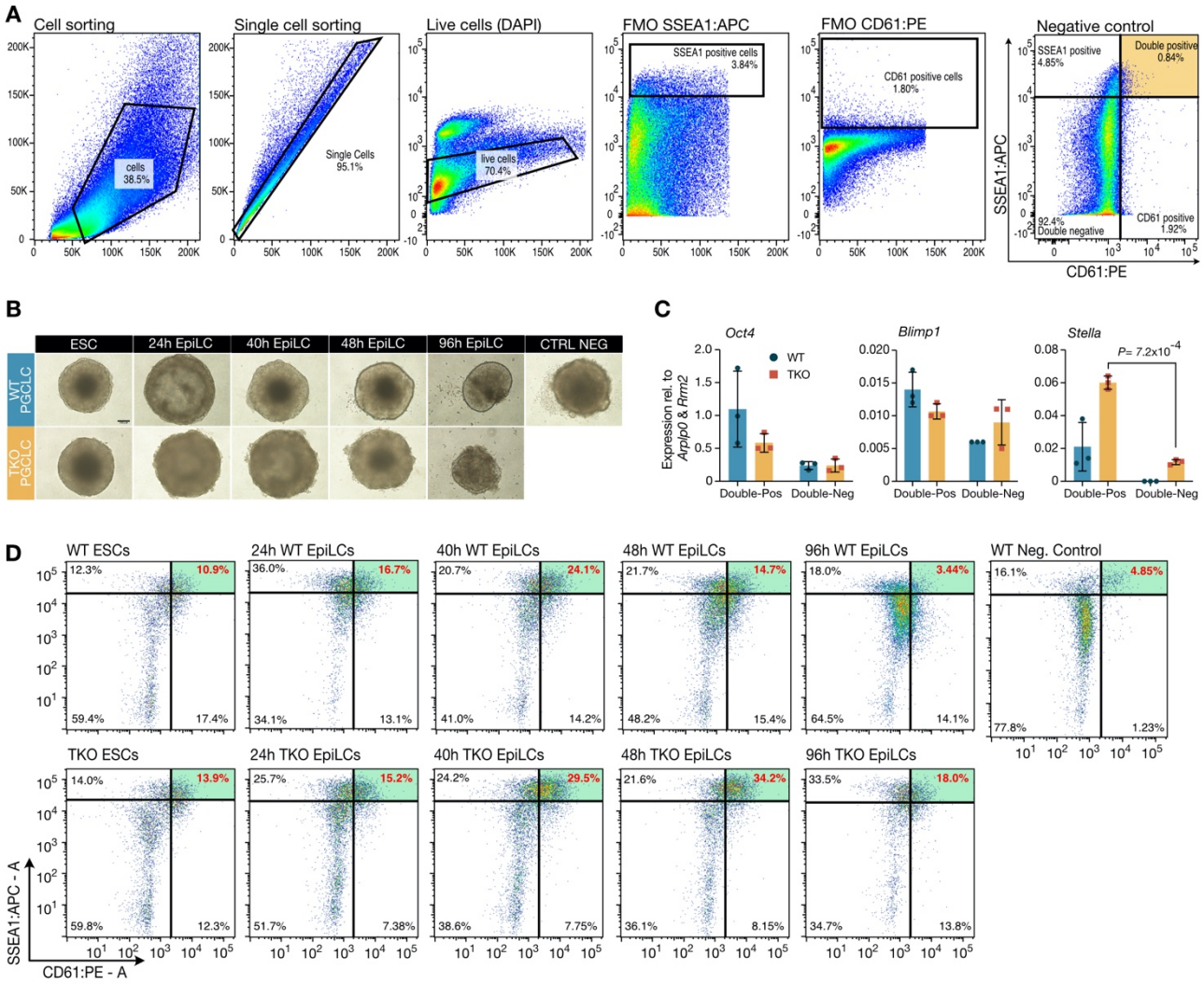

**Extended Fig. 2 | TKO ESCs can undergo EpiLC differentiation *in vitro*. A,** Representative images of WT (blue) and TKO (orange) cells during EpiLC differentiation. Scale: 200$\mu$m. **B,** Fiji layout picture of 6 well-plates of reverted ESC colonies derived from WT and TKO EpiLCs, after staining for alkaline-phosphatase activity and counted with FIJI. **C,** Expression of pluripotency and differentiation genes measured by RT-qPCR during WT and TKO EpiLC differentiation. Data shown are mean ± SD from biological triplicates. ΔCT values were normalized to *Arplp0* and *Rrm2*. **D,** Small multiple bar plot showing contribution of each WT and TKO sample to the variance of the main principal components during EpiLC differentiation. **E**, Density plot showing differences in DNA methylation observed between D0 and D7 of EpiLC differentiation in WT cells as measured by WGBS (Walter *et al.*, 2016 and Greenberg *et al.*, 2019). Threshold for regions defined as differentially methylated was arbitrary set-up at 75%. **F**, Volcano plot showing up-regulated genes in TKO compared to WT EpiLCs at D7. Red: up-regulated genes with a DMR promoter; Blue: germline genes. **G,** Heatmaps of transposable element expression showing normalized fold-change in TKO over WT cells during EpiLC differentiation. Black stars represent significance (FDR<1%, Log2FC>1). Expression is obtained for each condition from the average between biological duplicates.



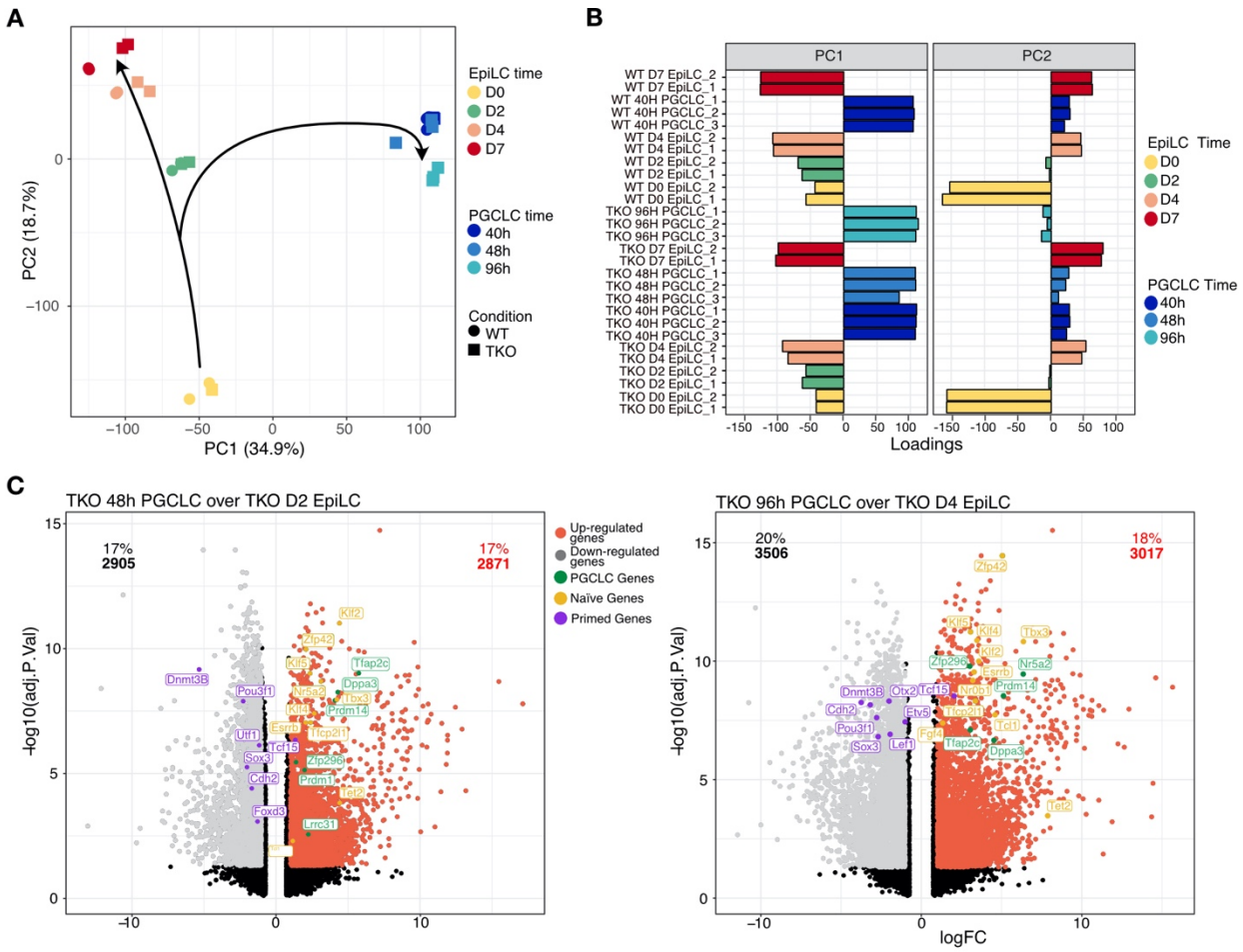

**E**

| *n= 6* | *# Embryos for aggregation* | *# Successful integrations* | *Efficiency (%)* |
|---|---|---|---|
| WT | 262 | 222 | 84.7% ± 8.2% |
| TKO | 300 | 275 | 91.7% ± 7.1% |

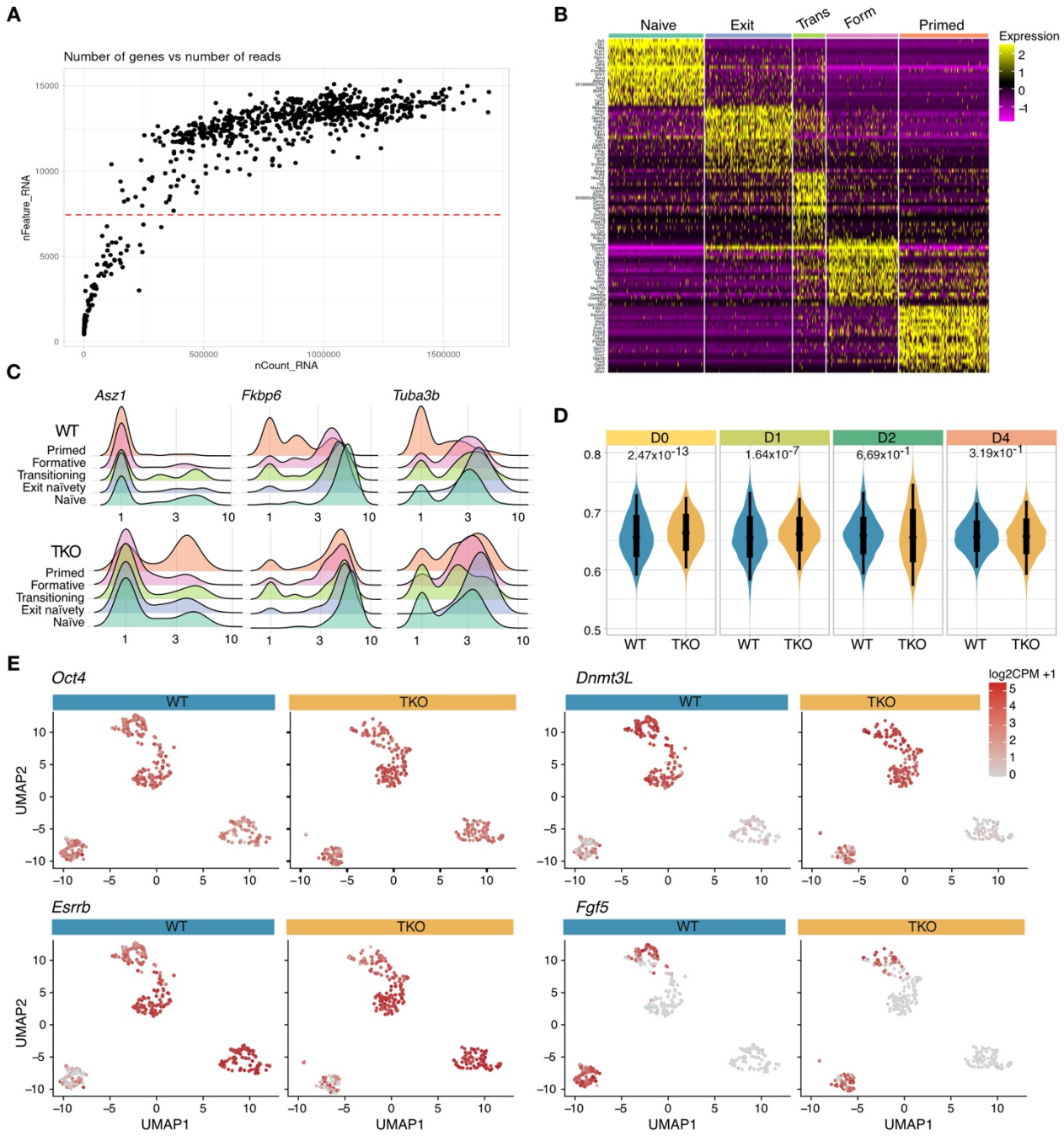
**Extended Fig. 3 I Generation of WT and TKO ESCs with a constitutive H2B::tdTomato reporter for chimera generation. A,** Flow cytometry analysis of H2B::tdTOMATO reporter expression in WT and TKO ESCs, and compared to untransfected negative control cells. **B,** Representative mitotic chromosome spreading of DAPI-stained WT and TKO ESCs bearing the H2B:tdTOMATO reporter. **C,** Counting of chromosome number in WT and TKO reporter cells. Scale: 100$\mu$m. Clones with the highest number of correct chromosome counts were chosen for chimera aggregation. **D,** Representative microscopy pictures after 24h of culture of aggregates of host-morula/ H2B::tdTOMATO ESCs. The merge reveals contribution of both WT and TKO ESCs to the blastocyst ICM. Scale: 400$\mu$m. **E,** Statistics for aggregation experiments. Five experiments of chimera generation were performed, totalizing 202 and 231 aggregates of host morulae with WT and TKO reporter ESC lines, respectively. Integration efficiency, measured by the number of correctly integrated reporter ESCs within embryos after 24h of culture, rose to around 84.7% with WT cells, and 91.7% with TKO cells.



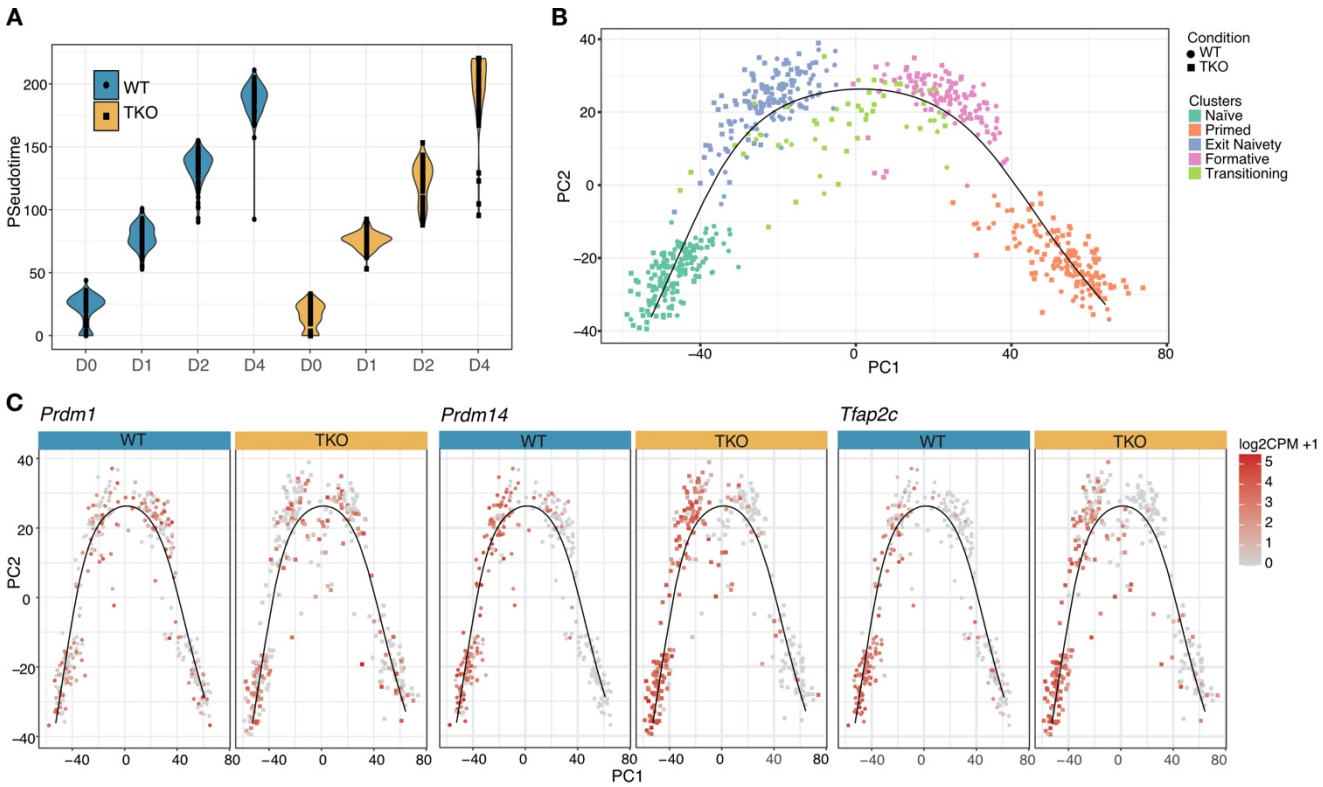

**Extended Fig. 4 I TKO EpiLCs can undergo germline differentiation *in vitro*. A,** After dissociation from aggregates grown under germline-inducing conditions, single cells were selected based on SSC/FSC. Then, single cells were identified by comparison between SSC-area/SSC-height. Next, live cells were selected based on DAPI integration. Finally, PGCLCs were identified for positive staining with SSEA1 and CD61 surface markers. Thresholds were set by fluorescence minus one (FMO) conditions and negative control, i.e. cells cultured without germline-inducing cytokines. **B,** Representative brightfield images of cultivated PGCLC aggregates at D4 generated from WT and TKO cells cultured in presence or absence of PGCLC-inducing cytokines. Scale: 250$\mu$m. **C,** Expression of PGCLC markers measured by RT-qPCR in WT (blue) and TKO (orange) PGCLCs. Data shown are mean ± SD from biological triplicates ΔCT values normalized to *Arplp0* and *Rrm2* (two-tailed unpaired student t-test). **D,** Representative FACS plot of SSEA1-Pos and CD61-Pos WT and TKO PGCLCs generated from ESCs and EpiLCs at different time points. Cell percentages are indicated in each quarter.



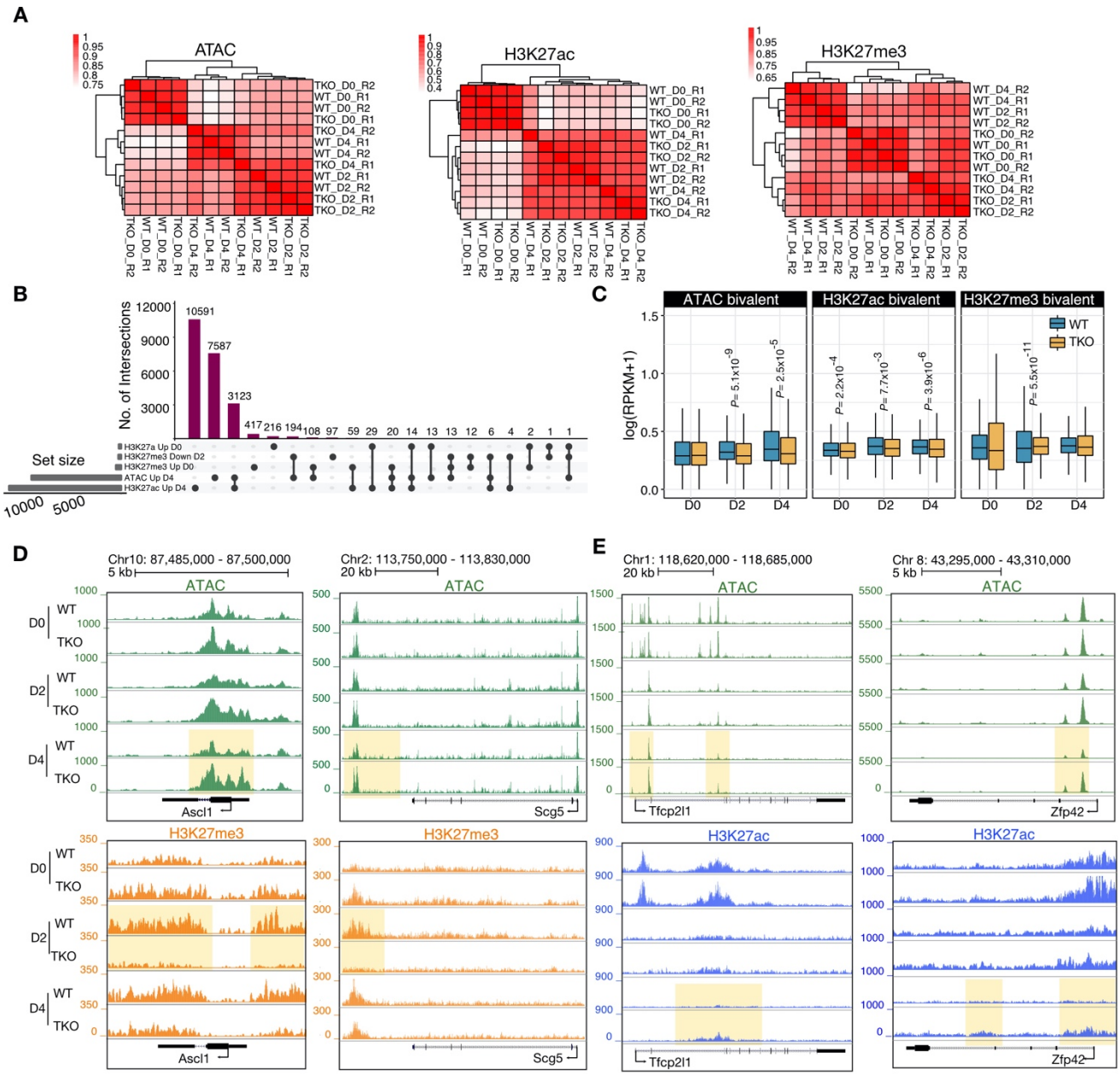

**Extended Fig. 5 | Bulk RNA-seq analysis of WT and TKO PGCLCs. A,** PCA from bulk RNA-seq of WT (circles) and TKO (squares) EpiLC and PGCLC differentiation in biological duplicates. Axes indicate variance, black arrows represent hypothetical trajectory. **B,** Small multiple bar plot showing contribution of each WT and TKO sample to the variance of the main principal components during EpiLC and PGCLC differentiation. **C,** Volcano plots showing differential gene expression in late TKO PGCLCs over TKO EpiLCs in Log2FC versus -log10(adj. pvalues). Red: Up-regulated genes; Grey: down-regulated genes ; Green: PGCLC genes; Yellow: naive pluripotency genes; Purple: primed pluripotency genes. Thresholds set up at FDR <5% and FC>1. Data shown were obtained from biological triplicates or duplicates.



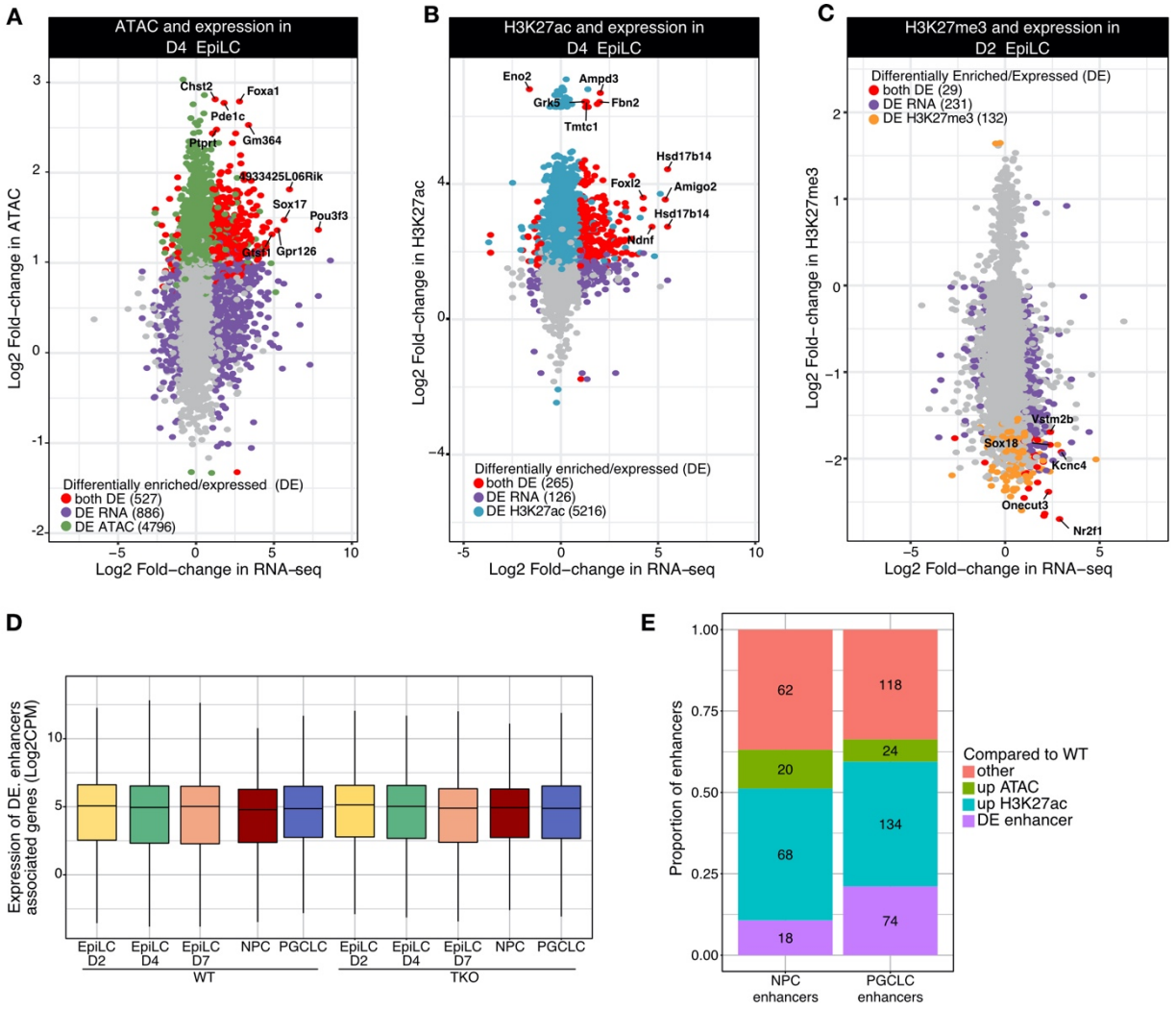

**Extended Fig. 6 I Cluster identification in scRNA-seq analysis of WT and TKO EpiLCs. A,** Dotplot displaying the number of genes over the number of reads per analyzed cell. Threshold was set up at a minimum of 7500 genes detected (red dashed line). **B,** Heatmap of top expressed markers identified in each cluster compared to all the others. Data shown is LogFC. **C,** Ridgeplot displaying germline gene expression (in Log scale) in each identified cluster for WT and TKO cells. **D,** Violin plot of WT (blue) and TKO (orange) transcriptional noise, calculated from the top 500 most variable genes as described in Mohammed *et al.* 2017. **E,** UMAP dimensionality reduction of scRNA-seq data from WT and TKO EpiLCs highlighting expression of key markers of EpiLC differentiation for a total *n*= 672 cells, from two biological duplicates.



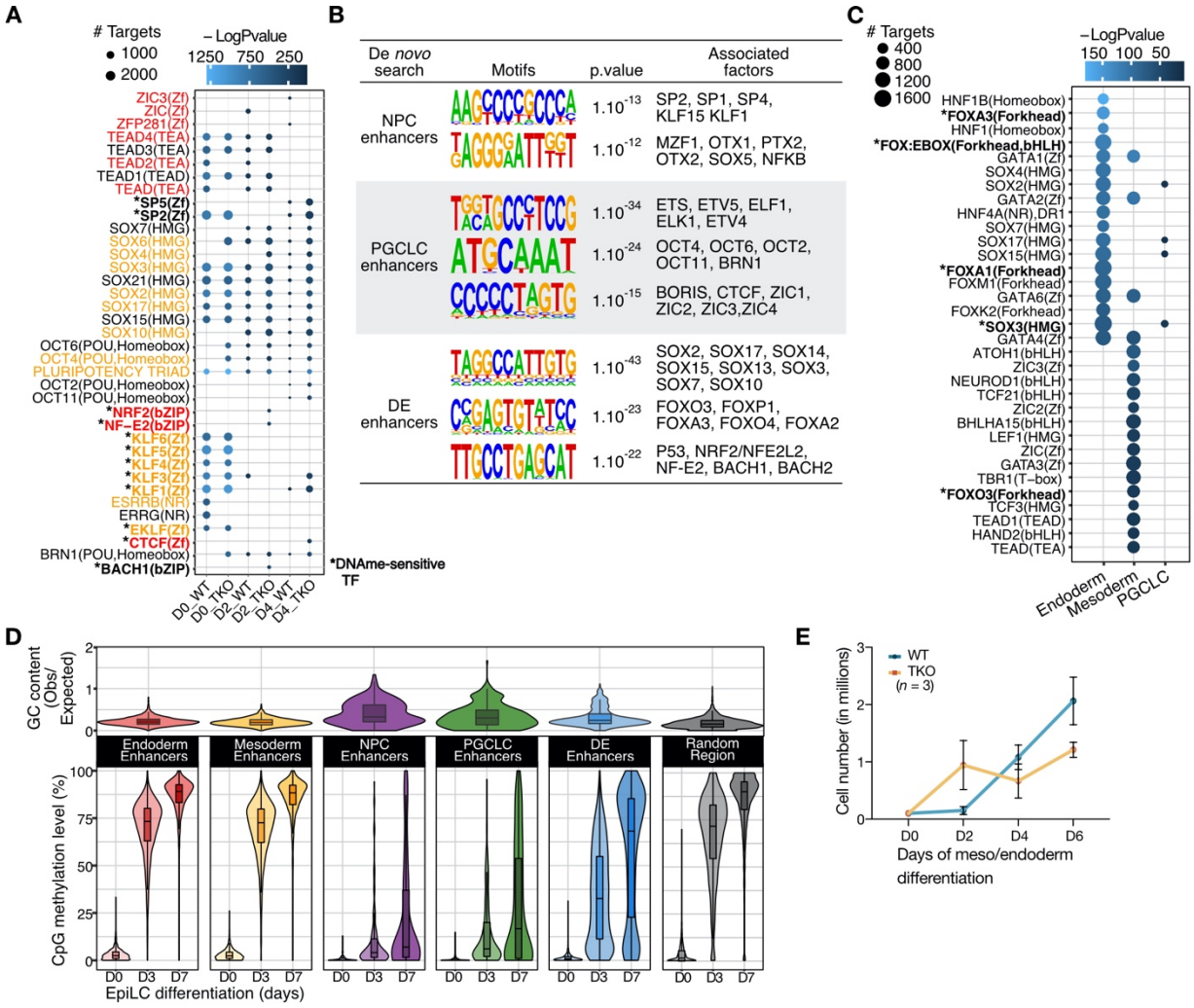

**Extended Fig. 7l TKO ESCs show proper timing of pluripotency transitioning. A,** Violin plot showing coordinates of WT (blue) and TKO (orange) cells along the pseudotime trajectory. **B,** PCA dimensionality reduction for pseudotime trajectory of WT and TKO EpiLCs highlighted for identified clusters during EpiLC differentiation. **C,** PCA dimensionality reduction for pseudotime trajectory of WT (blue) and TKO (orange) EpiLCs according to *Prdm1*, *Prdm14* or *Tfap2c* expression levels, expressed in Log2CPM+1.




**Extended Fig.8 | Characterization of ATAC, H3K27ac and H3K27me3 profiles during WT and TKO EpiLC differentiation. A,** Correlation heatmaps of ATAC, H3K27ac and H3K27me3 during WT and TKO EpiLC differentiation. Scale represents Pearson correlation with the relationships between the clusters represented by cluster dendrogram. **B,** Upset barplot showing overlaps between DE peaks in TKO over WT for ATAC, H3K27ac and H3K27me3. **C,** Boxplot showing enrichment in Log(RPKM+1) of ATAC, H3K27ac and H3K27me3 marks at bivalent regions (*n*=1568) as defined in Kurimoto *et al.,* 2015, for WT (Blue) and TKO (orange) cells during EpiLC differentiation. Data shown are the median with upper and lower hinges corresponding to 75 and 25% quantile (two-tailed unpaired student t-test). **D,** UCSC tracks for ATAC (green) and H3K27me3 (red) around 2 genes involved in neurogenesis (*Ascl1* and *Scg5)* showing decreased H3K27me3 enrichment at D2 and increased ATAC at D4 in TKO EpiLCs over WT. **E,** UCSC tracks for ATAC (green) and H3K27ac (blue) around 2 genes involved in PGC specification (*Tfcp2l1* and *Zfp42*) showing increased ATAC (green) and H3K27ac (blue) in D4 TKO over WT EpiLCs.




**Extended Fig. 9 | Uncoupling between chromatin changes and expression changes in TKO EpiLCs. A, B, and C,** Scatterplots showing Differential Enrichment in log2 Fold-change in (**A**) ATAC at D4, (**B**) H3K27ac at D4, (**C**) H3K27me3 at D2 vs Differential Expression in log2 Fold-change in RNA-seq. Peaks were annotated to genes at close proximity (<3kb from TSS). Colored dots depict the status of expression and/or enrichment. The numbers between brackets represent the number of genes marked by (**A**) ATAC in green, (**B**) H3K27ac in blue at D4, and (**C**) H3K27me3 in orange at D2 EpiLCs. For DE genes with DE peaks in close proximity (red dots), the top 10 differentially expressed genes are displayed. **D,** Expression of genes in the vicinity of Differentially Enriched enhancers (within 50kb) in WT and TKO EpiLCs, NPCs and PGCLCs (40h). **E,** Barplot showing the proportion and absolute number of putative enhancers in D4 TKO EpiLCs gaining ATAC or H3K27ac, or being differentially enriched in both marks when compared to D4 WT EpiLCs.




**Extended Fig. 10 | Transcription factors and differential dynamics of DNA methylation prime TKO cells for neural and germline fate. A,** Bubble plot of HOMER analysis for top 20 enriched TF binding motifs at all enhancers in TKO and WT cells during EpiLC differentiation for each time points. Factors that are DNA methylation-sensitive or regulated by DNA methylation are highlighted with (*). Factors are color-coded for their association with ESC (Yellow) and EpiLC (Red) states, as defined in Yang *et al.*, 2018. **B,** Table for *de* novo motif analysis in HOMER at NPC-, PGCLC- and DE Enhancers. Groups of motifs and associated factors that were identified in previous binding motifs analysis are represented. **C,** Bubble plot of HOMER analysis for top 20 enriched TF binding motifs at endoderm and mesoderm enhancers. **D,** DNA methylation analysis at enhancers. Top panel: global GC content as a ratio of observed/expected GC nucleotides. Lower panel: dynamic of mCpGs in WT cells during EpiLC differentiation at enhancers determined using public WGBS datasets (Greenberg *et al.,* 2019). **E,** Growth curve during mesoderm differentiation of WT and TKO cells. Data are mean ± SD in biological triplicates (two-tailed unpaired student t-test).

## Table 1. Antibody list.

| Target/Conjugate | Reference | Features | Application and dilution |
|---|---|---|---|
| SSEA1-APC | Thermo 50-8813-42 | Mouse Monoclonal | PGCLC Sorting<br>Undiluted |
| CD61-PE | Biolegend 04307 | Hamster Monoclonal | PGCLC Sorting<br>1/10 |
| TFAP2C | Santa-Cruz sc12762 | Mouse Monoclonal | PGC staining<br>1/200 |
| Anti-mouse CY5 | Jackson ImmunoResearch<br>715175151 | Donkey Polyclonal | PGC staining<br>1/100 |
| Anti-Rabbit IgG (H+L)<br>Alexa Fluor 488 | Invitrogen A21206 | Donkey Polyclonal | PGC staining<br>1/200 |
| H3K27ac | Active motif 39133 | Rabbit Polyclonal | Cut&Run<br>1/200 |
| H3K27me3 | CST 9733S | Rabbit Monoclonal | Cut&Run<br>1/200 |
| IgG | Sigma I5006 | From Rabbit serum | Cut&Run<br>1/200 |
| pMNase | Home made | | Cut&Run<br>1/400 |
| DNMT1 | CST 5032S | Rabbit Monoclonal | Western blot<br>1/1000 |
| DNMT3A | Abcam ab188470 | Rabbit Monoclonal | Western blot<br>1/2000 |
| DNMT3B | Abcam ab122932 | Rabbit Polyclonal | Western blot<br>1/1000 |
| PCNA | DAKO M0879 | Mouse Monoclonal | Western blot<br>1/2000 |
| Anti-Mouse IgG (H+L),<br>HRP | Thermo Fisher Scientific G-21040 | Goat Polyclonal | Western blot<br>1/10000 |
| Anti-Rabbit IgG (H+L),<br>HRP | Thermo Fisher Scientific G-21234 | Goat Polyclonal | Western blot<br>1/10000 |

## Table 2. Oligos et Primers

| Targets | Forward | Reverse | Application |
|---|---|---|---|
| *Oct4 (Pou5f1)* | GTTCAGCCAGACCACCATCT | ACTCCACCTCACACGGTTCT | qPCR Pluripotency |
| *Nanog* | TACCTCAGCCTCCAGCAGAT | CCACCGCTTGCACTTCAT | qPCR Pluripotency |
| *Sox2* | CACAGATGCAACCGATGCA | GGTGCCCTGCTGCGAGTA | qPCR Pluripotency |
| *Esrrb* | CAGGCAAGGATGACAGACG | GAGACAGCACGAAGGACTGC | qPCR Pluripotency |
| *Rex1 (Zfp42)* | ATCGCTGTGGGCATTAGAGA | CGATCCTGCTTTCTTCTGTGT | qPCR Pluripotency |
| *Fgf5* | GGGATTGTAGGAATACGAGGAGTT | TGGCACTTGCATGGAGTTT | qPCR Priming |
| *Otx2* | CCACTTCGGGTATGGACTTG | GTCCTCTCCCTTCGCTGTTT | qPCR Priming |
| *Hoxa1* | GTGACTAGTCTTCTGCATGTCG | TCTGCTCTGGACCACATCACTC | qPCR Mesoderm |
| *Foxa2* | CCCTACGCCAACATGAACTCG | GTTCTGCCGGTAGAAAGGGA | qPCR Mesoderm |
| *Gata4* | GAAAACGGAAGCCCAAGAACC | TGCTGTGCCCATAGTGAGATGAC | qPCR Mesoderm |
| *Lhx1* | ACAAGTTCGTTTGTAAAGAGGATT | TGGCACTTTCAGAGTCCTTGG | qPCR Mesoderm |
| *T (Brachury)* | TCCCGAGACCCAGTTCATAG | TTCTTTGGCATCAAGGAAGG | qPCR Mesoderm/ |
| *Sox1* | GCCTTCATGGTGTGGTCCC | GGTGTTCCTTCATGTGCAGC | qPCR Neural |
| *Nestin* | CTCCTGTGACAGCCTTTCTGAAG | AGGATAGGGAGCCTCAGACATAGG | qPCR Neural |
| *Ascl1* | CCAACTACTCCAACGACTTGAACTC | TCCTGCCATCCTGCTTCCAAAG | qPCR Neural |
| *Tfap2c (Ap2y)* | ATCAAGATCGGACACCCAAC | ATGGCGATTAGAGCCTCCTT | qPCR Germline |
| *Prdm14* | ACAGCCAAGCAATTTGCACTAC | TTACCTGGCATTTTCATTGCTC | qPCR Germline |
| *Prdm1 (Blimp1)* | GAGGATCTGACCCGAATCAA | CATGGAGGTCACATCGACAC | qPCR Germline |
| *Stella* | AGGCTCGAAGGAAATGAGTTTG | TCCTAATTCTTCCCGATTTTCG | qPCR Germline |
| *Rrm2* | CCGAGCTGGAAAGTAAAGCG | ATGGGAAAGACAACGAAGCG | qPCR Normalisation |
| *Rplp0* | TCCAGAGGCACCATTGAAATT | TCGCTGGCTCCCACCTT | qPCR Normalisation |
| *Dnmt1* | ATCTCCCGAGTAGGCCTGAG | CCACGAGAGGCCATTTCTGT | Genotyping TKO |
| *Dnmt3a* | GCTGTTCAGTGCTCCACCT | CACCTCTTTGAACAAGAGAACTTTCT | Genotyping TKO |
| *Dnmt3b* | AAGCATGAAGGCCAGATCAAA | AAGCCATTCTCGGTGCCAA | Genotyping TKO |
| *sgRNA Dnmt1* | CACCGTGTGGTGGGCCACCCTGCCA | CACCGTGTGGTGGGCCACCCTGCCA | CRISPR.Cas9 KO Dnmt1 |
| *sgRNA Dnmt3a* | CACCGACAATGGAGAGGTCATTGC | AAACGCAATGACCTCTCCATTGTC | CRISPR.Cas9 KO Dnmt3a |
| *sgRNA Dnmt3b* | CACCGCGTTAGAGAGATCATTGCAT | AAACATGCAATGATCTCTCTAACGC | CRISPR.Cas9 KO Dnmt3b |

# DISCUSSION

# 1. General Summary

Robin Holliday, in 1975, was among the first scientists to postulate that "**certain base modifications could [...] control the activity of adjacent structural genes [...] and operate developmental clocks which turn genes on or off.**" *(Holliday and Pugh, 1975)*. The study of the interplay between DNA methylation and development has since become a very active field of research for over 50 years. DNA methylation has proven to be an extremely stable and heritable mark, and its repressive role on gene expression has been well-established. Early on, DNA methylation appeared as a strong candidate for epigenetic regulation, with implications spanning from regulating cell identities during development, to non-genetic heredity and emergence of pathologies *(Greenberg and Bourc'his, 2019)*. But it is only with the advent of molecular biology and *in vitro* models of cell culture that the exploration of the function of DNA methylation in development was truly unleashed. Indeed, while *in vivo* studies unambiguously demonstrated that DNA methylation was a developmental necessity, it is only with the derivation of *in vitro* embryonic stem cells (ESCs) devoid of DNA methylation that the relationships between DNA methylation and embryonic stem cell identity started to be uncovered *(Lei et al., 1996; Okano et al., 1999; Sakaue et al., 2010; Tsumura et al., 2006)*. Since then other groups have used TKO cells to uncover regulation of DNA methylation on transcription factors activity *(Domcke et al., 2015; Grand et al., 2021)*, on histone marks *(Brinkman et al., 2012; Karimi et al., 2011)*, regulation of transposable elements or germline genes *(Velasco et al., 2010; Walter et al., 2016)*, early differentiation of embryonic stem cells *(Schmidt et al., 2012)*, and so on… Yet, as I started my PhD work, some key questions remained to be answered.

First, while it was shown that DNA methylation is densely acquired during epiblast differentiation at regulatory elements of naïve pluripotency genes, the exact role of DNA methylation in regulating the exit of naïve pluripotency and priming of stem cells remained nebulous. Then, despite a crucial role in repressing germline genes in somatic cells and in preventing precocious gametogenesis in early germ cells, the function of DNA methylation in the specification of primordial germ cells (PGCs) remained largely unexplored. Finally, contrasting results had been previously observed in somatic differentiation of *Dnmt*-TKO ESCs, thus we sought to shed some

new light on the potentialities of DNA methylation-free ESCs to undergo early somatic differentiation. To this end, I generated our own *Dnmt*-TKO cell line using CRISPR-Cas9 editing, and made extensive use of differentiation protocols, functional assays and genome-wide profiling to characterize the ability of TKO cells to exit pluripotency, be primed and induce progenitors of somatic and germline lineages. During this study, I only worked with male stem cells. While mammalian female ESCs can sustain the same differentiation protocols, they display a delayed kinetics when compared to male ESCs, as the regulation of X chromosome inactivation impacts the exit of naïve pluripotency and the differentiation of the epiblast towards the germline *(Schulz et al., 2014; Severino et al., 2022)*. In my opinion, while *in vitro* differentiation protocols have their limitations and drawbacks—that I will cover later in this discussion—they also present some significant advantages. Most notably, they recapitulate the main dynamics of major chromatin marks and allow for proper gene expression compared to *in vivo* pluripotent embryonic cells (Hayashi et al., 2011; Shirane et al., 2016; Q. L. Ying et al., 2003), and they also provide enough material for a wide-range of downstream applications, overcoming material limitation faced by *in vivo* models at these developmental stages.

In this chapter, I will discuss some aspects of the main conclusions that I drew from my study, and will try to open some new perspectives. Briefly, I showed that, alike naïve stem cells, epiblast stem cells do not rely on DNA methylation to support their identity and functional properties. Therefore, I am proposing to extend the notion of dispensability of DNA methylation in stem cells from the naïve to the primed state. Then, I found that DNA methylation is also dispensable for the induction of somatic neural progenitor cells (NPCs), as gene misregulation only became evident after several days of culture during NPC differentiation, with the reactivation of pluripotency genes in late TKO NPCs. Moreover, it appeared that TKO cells have longer opportunities to give rise to the germline without any apparent impact on transcription, thus showing that DNA methylation is not critical for the regulation of the transcriptome in early differentiating stem cells. Finally, I observed that the absence of DNA methylation primes neural and germline fates as default routes, likely via the action of DNA methylation-sensitive transcription factors that correlates with a “relaxed” chromatin state that we found at putative enhancers of these lineages.

# 2. DNA methylation in stem cells: a one-way relationship

## 2.1 DNA methylation is dispensable for stem cell identities

### 2.1.1 Redundancy of repressive epigenetic marks in naivety

Mouse naïve embryonic stem cells were known to tolerate combinatorial loss of *Dnmt1*, *Dnmt3a* and *Dnmt3b* *(Tsumura et al., 2006)*. This effect was also visible to us, as our own E14 TKO cell line could be maintained in naïve conditions (in serum/LIF or 2i-based media) without changes in morphology, growth rate or expression level of main pluripotency markers. This dispensability of DNA methylation in naïve stem cells can also be extended to other epigenetic marks, like repressive H3K27me3 histone modifications *(Chamberlain et al., 2008)*. However, additive knock-outs of DNA methylation, H3K9me3 and H3K27me3 induce cellular death after six to eight of cultures, even in in 2i media *(Walter et al., 2016)*. This suggests that, while repressive epigenetic marks taken one at the time are largely unnecessary to support naïve identity, these marks may display some redundancy. The main view is that embryonic epigenetic reprogramming—ending up with a reduced levels of repressive DNA methylation at the blastocyst stage—is a by-product of the need to erase in the embryo the parentally inherited gametic patterns of DNA methylation. As a consequence, as hypomethylated stem cells become exposed to new threats like the reactivation of TEs, this might have promoted the development of some redundant and compensating strategies through other repressive chromatin marks. However, since repressive epigenetic marks are also thought to be pivotal in stabilizing cellular identity, their relevance for embryonic stem cells is not intuitive: these cells are dividing fast and need to remain plastic and highly-responsive to signaling cues. Instead, the control of naïve pluripotent identity might fall in other hands than repressive epigenetic marks. High levels of chromatin accessibility, activity of transcription factor networks, and active histone marks may play a larger influence. In line with this, we saw like others, that ESCs have high chromatin accessibility and also display more diffuse heterochromatin domains *(Meshorer et al., 2006; Schlesinger and Meshorer, 2019)*.

### 2.1.2 DNA methylation is not required for priming

I could further show that DNA methylation is also dispensable for exiting naïve pluripotency and for transitioning toward a primed pluripotency state. This was in contrast to the general view *(Geula et al., 2015; Weinberger et al., 2016)*. I validated this observation first at the transcriptomic level, where I observed no striking changes in gene expression compared to WT cells, with the exception of germline genes. Then I demonstrated that in absence of DNA methylation, Epiblast-like cells properly activated the proximal enhancer of *Oct4 and* were not prone to developmental reversion. These findings, in our opinion, extend the window of dispensability of DNA methylation to primed stem cells. Similarly, it has been formerly argued that repressive H3K27me3 marks could also be dispensable for gene repression in early differentiation and in PGCLC specification *(Huang et al., 2021; Riising et al., 2014)*. However, it might be necessary to point that we performed EpiLC differentiation for a maximum of seven days only, which might represent an early state of primed pluripotency, in opposition to stable epiblast-like primed colonies that may need few rounds of passaging and longer culture time to be stably established *(Brons et al., 2007; Morgani et al., 2017)*. Yet, I also showed that DNA methylation-free stem cells can contribute to the epiblast *in vivo*, in chimeric embryos.

### 2.1.3 Different *Dnmt* Knock-outs, different chimerae

The contribution of our E14 *Dnmt*-TKO cells to the epiblast of chimeric embryos appeared at first puzzling to us. Indeed, it has long been shown that *Dnmt1*-KO, *Dnmt3B*-KO or double *Dnmt3A; Dnmt3B*-KO mice poorly develop embryonic lineages beyond E9.0-10.0 *(Dahlet et al., 2020; Lei et al., 1996; Okano et al., 1999)*. Accordingly, I observed a reduction in the size of the TKO chimerae at E7.5, supporting growth retardation of embryonic tissues in absence of DNA methylation, but yet, TKO embryonic cells were abundantly present in the epiblast. In particular, former studies concluded that the lack of DNA methylation is only compatible with contribution to extra-embryonic fate. This was observed in nuclear-transfer chimeras using the original J1 *Dnmt*-TKO cell line (*Sakaue et al., 2010)*, and in a recent aggregation assay with double *Dnmt3a*; *Dnmt3b*-KO E14 ESCs *(Kinoshita et al., 2021)*. Both studies

showed that the *Ascl2* imprinted gene—encoding a trophoblast transcription factor—was upregulated during early differentiation of these DNA methylation-defective cell lines, and most importantly, that *Ascl2* knock-out restored proper epiblast contribution in chimerae *(Kinoshita et al., 2021)* **(Figure 17A- 17B)**. However, this *Ascl2* upregulation did not appear to originate from loss of imprinting, as DNA methylation is not involved in the regulation of this imprinted gene. In contrast, we did not notice striking differences in *Ascl2* expression in our own E14 *Dnmt*-TKO cell line during epiblast differentiation compared with WT cells, which may explain the high epiblastic contribution that we observed **(Figure 17C)**. How our TKO cells properly regulate *Ascl2* expression remains to be explained. But it is worth noticing that both the genetic background and the knock-out strategy can greatly influence the ability of a mutant to differentiate, or not. As an example, it was observed that ESCs with a catalytically inactive DNMT1 protein showed higher propensity to differentiate into the primitive endoderm, which was not observed in complete *Dnmt1*-KO cells *(Ito et al., 2022)*. Moreover, cells with knock-down of *Dnmt3A* and *Dnmt3B*, demonstrate delayed exit of naïve pluripotency and epiblast differentiation, associated with loss of DNA methylation at the *Nanog* promoter *(Li et al., 2017)*. However, while the *Dnmt3A*; *Dnmt3B*-DKO E14 ESCs from the study of Dr. Kinoshita also targeted the catalytic domain of DNMTs, the J1 TKO ESCs generated by Dr. Okano were made by targeted disruption of *Dnmt* genes using vectors carrying antibiotic resistance genes. Hence it could be interesting to see if knocking out *Dnmt* genes by CRISPR-Cas9 at catalytic domains in the J1 ESC background could modify the contribution of J1 ESCs from extra-embryonic to embryonic lineages, thus pointing toward a background or KO strategy-dependent origin for differential contribution of ESCs in chimeras. Moreover, we could also generate an additional *Dnmt1*-KO in the *Dnmt3A*; *Dnmt3B*-DKO E14 ESCs from Dr. Kinoshita's study, and check if this could restore *Ascl2* expression, thus pointing at a mutant-type origin for this chimeric contribution discrepancy.

## 2.2 Uncoupling transcription and epigenetic reprogramming

### 2.2.1 DNA methylation, critical for silencing naivety?

A recurrent observation I made throughout this work is the consistent ability of TKO ESCs to properly silence naïve pluripotency genes during EpiLC, NPC or meso/endoderm differentiation, despite the absence of DNA methylation. It is at odd with previous observations that naïve pluripotency genes acquire DNA methylation at their enhancers and promoters during priming *(Bell et al., 2020; Kalkan et al., 2017)*. This had led to the prevailing view that DNA methylation is an important repressor of pluripotency genes during differentiation. However, while this applied to a subset of DNA methylation-sensitive genes like *Rex1*/*Zfp42*, *Esrrb*, *Nr0b1*, or *Nanog*, other core pluripotency factors, like *Sox2* and *Oct4*, or priming genes like *Zic2*, *Zic3* and *Zfp281*, do not show strong correlation between their transcription status and DNA methylation level. Hence, it is unlikely that promoter DNA methylation could really impact transcription and regulation of priming *(Bogdanović et al., 2011; Kalkan et al., 2017)*. It is interesting to note that *Rex1* expression is a classic marker for monitoring the exit from naïve pluripotency, but its expression is anticorrelated with DNA methylation acquisition at its promoter *(Kalkan et al., 2017; Wray et al., 2011)*. Since, we have shown that naïve pluripotency exit does not rely on DNA methylation, maybe using another naïve marker, not targeted by DNA methylation, like *Nr5a2*, could be an interesting alternative to follow the dynamics of pluripotency transitioning.

Finally, single-cell RNA-seq (using SMARTseq2) showed no TKO-specific cluster of cells, unequivocally demonstrating homogenous and synchronous regulation of gene expression between WT and TKO cells at the single-cell level during epiblast differentiation. This data also argue against the emergence of PGCLCs from a population of EpiLCs with residual expression of naïve genes, as it was recently proposed *(Aramaki et al., 2021; Kim et al., 2020)*. This is true in both WT and TKO EpiLCs, furthermore disconnecting DNA methylation from a transcriptional control of PGC specification. Yet, it is in contrast to human PGCs, which have been shown to emerge from a population of progenitors that reactivate the naïve pluripotency program before specializing into germ cells *(Chen et al., 2019)*. Why mouse PGCLCs need a complete dissolution of pluripotency before specification is still an open

question. Maybe it is wrong of us to think of naïve pluripotent stem cells as the starting point of embryonic lineages development in mice. In fact, naïve pluripotency needs to be overcome to enable differentiation in general, as exemplified by the necessity to down-regulate *Nanog* to allow somatic and germline differentiation *(Di Giovannantonio et al., 2021; Trott and Martinez Arias, 2013)*, which we also observed in all of our differentiation protocols **(Figure 17D)**. In contrast, in humans, stem cells in the blastocyst more closely resemble mouse primed pluripotent stem cells. The accelerated preimplantation development in mice may be the reason for the persistence of naïve stem cells in the blastocyst *(Messmer et al., 2019)* but this state might just represent the final stage of "blank" reprogramming. Finally, priming of these naïve cells may truly allow them to acquire pluripotency, functionally speaking, as postulated by the formative pluripotency model of Austin Smith *(Smith, 2017)*. Since DNA methylation might be required for the repression of naïve pluripotency and PGCs reactivate the pluripotency program, our initial hypothesis was that TKO cells could take a developmental 'short-cut', and be more prone to give rise to the germline by incomplete loss of pluripotency gene expression. While TKO ESCs do in fact show extended ability to give rise to PGCLCs, this effect seems to be independent from a role of DNA methylation on gene transcription.

### 2.2.2 DNA methylation at naïve regulatory elements, critical for priming?

As *de novo* DNA methylation occurs simultaneously with naïve-to-primed pluripotency transition in the epiblast, it is legitimate to question the role of DNA methylation in this transition. We found that the lack of DNA methylation changes the distribution of H3K27ac and H3K27me3, as well as global chromatin accessibility in EpiLCs. More specifically, we observed retention of H3K27ac and ATAC signal in putative enhancers associated with NPC and PGCLC genes. While I will discuss later how these local, specific enrichments are achieved, I will first focus on the relationships between DNA methylation and these marks, and why these changes may also not impact transcription.

The globally accepted "follower model" is that DNA methylation is directed and deposited according to histone modifications, in particular H3K4me3 and H3K36me3 that are sensed by specialized domains in *de novo* DNMT3 proteins. However, the

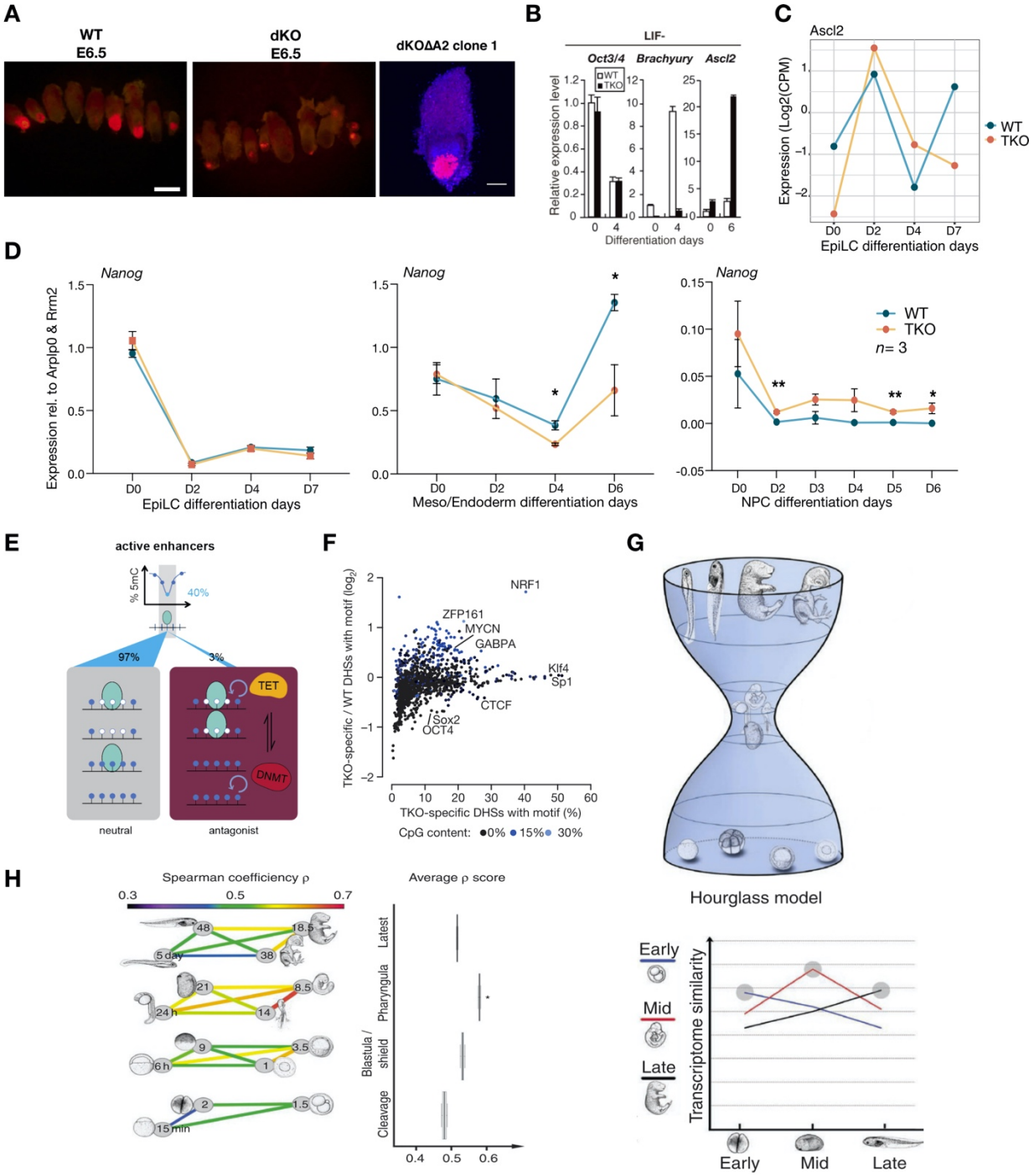



**Fig. 17 I Chimerae, DNA methylation-sensitive enhancers and phylotypic stage**

**A.** Chimerae made with *Dnmt3A*/*Dnmt3B*-DKO ESCs, at E6.5, with rescued epiblastic contribution when *Ascl2* gene is knocked-out (Kinoshita *et al*, PNAS 2021). **B,** Barplot showing *Oct4*, *T*, and *Ascl2* expression in *Dnmt*-TKO ESCs undergoing general differentiation after LIF withdrawal, assessed by RT-qPCR. (M. Sakaue *et al*, Current Biology 2010) **C.** Expression of *Ascl2* in our RNAseq data for EpiLC differentiation in E14 WT and TKO cells (*n*=2). **D.** Downregulation of *Nanog* in early days of NPC, Meso/endoderm and EpiLC differentiation, as a read-out of exit of naïve pluripotency before lineage commitment. **E.** Graphical representation of the control of DNA methylation at sensitive enhancers. Briefly, a subset of enhancers demonstrate differential binding of transcription factors (TFs) depending of methylation status, while the majority of enhancer remain mostly insensitive (Kreibich *et al*, BioRxv 2022). **F.** Volcano plot showing TFs binding occurrence in accessible regions (DHS) in TKO cells compared to all accessible regions in WT, colored according to CpG content in binding motifs (Domcke *et al,* Nature 2015). **G.** Graphical representation of the 'Hourglass" model, representing the bottleneck of high transcriptomic similarity in different vertebrate species during early embryonic development. **H.** Spearman correlation score (ρ) from the comparison of ortholog gene expression profile at different embryonic development times, in different vertebrate species. High ρ score is associated with higher transcriptomic similarity between these species at this specific stage (Irie *et al,* Nature Communications 2011).

hierarchy may be different at promoters and enhancers for histone marks H3K27ac and H3K27me3 for which DNA methylation is rather instructive. Moreover, this effect seem to be highly dynamic and reversible, as reintroduction of DNMTs allows re-establishing WT chromatin profiles *(King et al., 2016)*. On one side DNA methylation could limit H3K27ac profiles by regulating the binding of DNAme-sensitive transcription factors (TFs), and on the other side DNA methylation could impact the deposition of H3K27me3 depending on the CpG density and genomic context. While we saw thousands of regions with differential gain of H3K27ac and ATAC signal in TKO EpiLCs, only a few hundreds of genes were subsequently upregulated. While it has been previously shown that accumulation of H3K27ac does not necessarily correlate with changes in transcription during epiblast differentiation *(Zylicz et al., 2015)*, it remains puzzling to us why so few genes are affected. This decoupling could be due to weak enhancer/promoter interactions, as it was shown that naïve cells demonstrate more open chromatin landscape with loss of local chromatin interactions *(McLaughlin et al., 2019; Schlesinger and Meshorer, 2019)*. Promoter/enhancer interactions are more likely to happen when they are found in the same topologically associated domain (TAD) *(Zuin et al., 2022)*. Because we see higher chromatin accessibility in absence of DNA methylation in TKO EpiLCs, it is possible that this relaxed chromatin state reduces the definition of TADs and thus, the probability of promoter/enhancer interactions, hence minimizing the impact of local changes in histone modifications on gene expression.

# 3. A head start in cell fate decisions for DNA methylation

## 3.1 DNA methylation, an instructor of formative pluripotency

### 3.1.1 The absence of DNA methylation promotes "*hyper-formativity*"?

We propose that while DNA methylation does not regulate the transcriptional program associated with priming, its absence promotes the adoption of an "hyper-formative" state, rendering the cells more sensitive to specific differentiation cues. There are indeed a few defining characteristics that could support this notion *(Morgani et al., 2017; Smith, 2017)*. First of all, the competence to give rise to the germline is considered a defining feature of formative stem cells, and while we clearly observed that WT ESCs lose this ability after 40h of epiblast differentiation, the TKO cells demonstrate an extended time-window for germline specification. Of note, we never managed to locate specialized TKO PGCs in our chimeras by whole-mount immunofluorescence, because of low number of PGCs specialized at E6.5, low yield and retarded growth in aggregated chimeras, and suboptimal antibodies against markers of the nascent germline. Despite this, when performing the dual-luciferase assay, I found normal *Oct4* enhancer switch in TKO ESCs vs EpiLCs, but this trend was less pronounced in absence of DNA methylation, suggesting that a large proportion of TKO EpiLCs might still carry an active *Oct4* distal enhancer or both active enhancers. This goes along with the proposition that the *Oct4* enhancer switch is not binary *in vivo*, and that simultaneous mobilization of both enhancers is to be expected in intermediate pluripotent states *(Han et al., 2010)*. Finally, fully primed stem cells normally lose the active chromatin signature reflective of the naïve state, with strong increase in H3K9me3, H3K27me3 at bivalent regions, or reduced chromatin accessibility, for example (Bernard et al., 2022; Nicetto et al., 2019; C. Wang et al., 2018). In contrast, I observed that TKO EpiLCs seem to retain a relaxed chromatin landscape, possibly rendering them more sensitive to differentiation cues. It would be interesting to functionally assert the hyper-formativity of our TKO EpiLCs by testing their contribution in chimerae, as it is expected that fully primed stem cells could have poorer contribution to chimerae than formative—or early-primed—TKO EpiLCs.

### 3.1.2 How DNA methylation and *in vitro* models affect formativity?

Using RNA-seq, we also noticed that transient priming occurs during NPC differentiation of TKO cells, around two days of differentiation. This supports the notion that formativity is acquired when differentiating stem cells toward somatic lineages, even in absence of DNA methylation. Hence, it is unlikely that DNA methylation, or other repressive marks, play a direct role in the regulation of formativity, as its establishment is made in full in primed EpiLCs or even in lineage progenitors *(Minoux et al., 2017; Nicetto et al., 2019)*. But they could contribute more indirectly by prepping formative cells to be properly addressed to specific lineages later in development. Most notably, some studies argued that the progressive establishment of epigenetic marks, despite not directly controlling gene expression, could contribute to transcriptional heterogeneity and thus, facilitate cell fate commitment *(Bleckwehl et al., 2021; Eldar and Elowitz, 2010; Mohammed et al., 2017)*. However, when we looked in our TKO EpiLCs, we could not detect changes in transcriptional heterogeneity, excluding this mechanism as a way for DNA methylation to regulate formative pluripotency.

One limitation that we could raise for *in vitro* models, is that they tend to demonstrate a “slowed-down” development. Indeed, two days are required to reach a formative state during EpiLC differentiation, and then at least a week to obtain stable cultures of primed stem cells. Instead, naïve-to-primed transitioning occurs in 2-3 days *in vivo*. Hence, *in vitro* models tend to artificially stretch development as a sequence of well-defined, separate, events. In his definition of formative pluripotency features, Austin Smith argued for a clear separation between transcriptional programs, with complete absence of naïve genes in formative cells, and low induction of early post-implantation markers. However, it has been shown that naïve genes, like *Klf4*, and post-implantation genes, like *Otx2* and *Oct6*, can be found co-expressed *in vivo* in the transition from naïve-to-primed pluripotency, at E5.0 *(Neagu et al., 2020)*.

## 3.2 DNA methylation harmonizes differentiation routes

### 3.2.1 DNA methylation decommissions neural and germline-associated enhancers

We have shown that the lack of DNA methylation allows for priming, and promotes some kind of "hyper-formativity" in stem cells, without major impact on transcription, but rather by promoting the retention of active chromatin signatures. This peculiar epigenetic landscape in TKO EpiLCs is associated with extended opportunities to give rise to the germline, and proper induction of a somatic neural fate. We could reconcile these two observations by observing that H3K27ac and ATAC signal gain was especially strong at neural and germline putative enhancers in TKO EpiLCs. This is in contrast with enhancers of other somatic lineages, like mesoderm and endoderm, that kept normal levels of these marks compared to WT EpiLCs. Further characterization of these putative elements revealed that neural and germline enhancers are more GC-rich, and less methylated during EpiLC differentiation than other lineage enhancers, which is in line with previous observations (Argelaguet et al., 2019a). Moreover, transcription factors binding motifs and footprinting analysis revealed that neural and germline enhancers share a large overlapping network of TF binding sites, among them many are known to be sensitive to DNA methylation.

It was recently shown that for a subset of DNA methylation-sensitive lineage enhancers, DNA methylation turnover is central for regulating TF binding in pluripotent and in early differentiated cells *(Charlton et al., 2020; Kreibich et al., 2022)*. Therefore, it is tempting to speculate that DNA methylation establishment at these sensitive putative enhancers during epiblast differentiation is critical to temper the induction of specific lineages, avoiding their strong activation **(Figure 17E-17F)**. Moreover, it was shown that DNA methylation also affects histone acetylation levels *(Cusack et al., 2020; Hon et al., 2014)*. Hence, the absence of DNA methylation and subsequent TF binding might promote local chromatin accessibility and active histone marks retention, resulting in the priming of neural and germline putative enhancers for rapid activation upon differentiation. Then, beyond DNA methylation-sensitive factors, we also noticed strong mobilization at the binding site for the pluripotency factors, like OCT4, in TKO EpiLCs at these DNA methylation putative enhancers. These proteins

are well-studied pioneer transcription factors that can promote chromatin accessibility. For example, OCT4 protein, even when present at low level, has been associated with chromatin opening *(Xiong et al., 2022)*.

### 3.2.2 DNA methylation counteracts neural and germline fates as default routes

The absence of DNA methylation does not provide permissiveness for differentiation into all lineages. In fact, when we performed meso/endoderm somatic differentiation, we observed that TKO cells properly repressed pluripotency genes, but mis-regulated meso/endoderm markers while they up-regulating neural genes. Similarly, in absence of pro-germline differentiation cytokines, TKO cells still yielded the capacity to generate PGCLCs (yet, to a lesser extent than in normal conditions). This supports that neural and germline fates constitute default routes of differentiation in absence of DNA methylation, when exposed to differentiation cues, in regard to other fates.

We propose that the apparent inability of TKO cells to efficiently undergo meso/endoderm differentiation is probably due to the "over-induction" of the neural network, impairing correct harmonization of the different cell fates. This could resolve the puzzling observation we made when we observed strong cellular lethality for TKO cells in embryoid bodies grown in somatic culture conditions, but not when these aggregates are formed in germline culture conditions. Indeed, in somatic embryoid bodies, the cells have to drive themselves in their differentiation toward the three somatic germ layers *(Koike et al., 2007)*. If the lack of DNA methylation affects the ability of the cells to correctly adopt mesoderm of endoderm identities due to a strong epigenetic determinism promoting neural fate, this could result in decreased cellular viability. In contrast, when we derive PGCLCs *in vitro*, we promote the differentiation toward one specific lineage, furthermore one with strong epigenetic determinism, explaining their greater survivability.

In fact, they are other occurrences of epigenetic determinisms in development, promoting default routes and requiring partial epigenetic reprogramming in progenitor cells to facilitate the adoption of one fate compared to another one *(Banerjee et al., 2018; Izzo et al., 2020; Rauch et al., 2019)*. Most notably, it has been proposed that the ectoderm lineage—and neural identity—constitutes a default pathway of

differentiation in early ESCs (Argelaguet et al., 2019a; Muñoz-Sanjuán and Brivanlou, 2002). Our observations during differentiation protocols strongly supports this model, and suggests that DNA methylation restricts neural identity and harmonizes cell fate decisions against epigenetic predeterminism. Additionally, we propose that the germline lineage constitute another default route of development, as PGCLC enhancers showed similar sensitivity to DNA methylation than neural enhancers and seem to be targeted by the same set of DNA methylation-sensitive TFs.

### 3.2.3 Why neural and germline identities constitute default routes in development?

Why such asymmetries exist in cell fate decisions remains unanswered. Besides, since default routing exists across diverse differentiation paths during embryogenesis and adulthood, it will be probably hard to find a unifying model explaining all processes underlying default routing. However, in the examples of which I have knowledge, chromatin marks, like DNA methylation, are central for attenuating default routing, with their removal often skewing differentiation in one direction *(Izzo et al., 2020; Kim et al., 2020)*. Yet, I would like to attempt to propose one possible explanation as to why some fates (especially the neural fate) need such tempering, in the context of the early embryo.

Among vertebrates, there is a highly conserved phylotypic stage of early development, around the pharyngula stage **(Figure 17G-17H)**—the exact stage is still debated—where various ectoderm-derived structures emerge (like neural tube, neural crest cells, olfactory pit) *(Irie and Kuratani, 2011)*. We can imagine that such conservation promotes high selective pressure on upstream development, in which any impairment could have drastic impact in the subsequent developmental stages. Therefore, we can imagine that embryonic stem cells may be evolutionary constrained to give rise to neural and ectoderm-derived structures. In this optic, epigenetic marks could be put in place to mitigate this tendency. Indeed, it was shown in a recent preprint that DNA methylation, and DNA methylation-sensitive TFs (notably KLF, ETS, FOXO factors) have a highly conserved function of stabilizers of tissue-specific cell identities among vertebrates *(Klughammer et al., 2022)*. Why the lack of DNA methylation also favors the germline is more difficult to explain. In my opinion, the fact

the many TFs are involved in the regulation of both neural and germline fates could provide a partial explanation, in which the epigenetic control over germline enhancers is a secondary effect to the ectoderm/neural specificity. It is interesting to note that there is a striking similarity between the proteome and transcriptome of the brain and testes tissues in mice and humans *(Matos et al., 2021)*. While on an evolutionary angle, such brain-testes axis could be explained by the central role played by these tissues on the speciation process *(Wilda et al., 2000)*, the question about why some TFs commonly regulate NPC and PGCLC enhancers warrants more investigation.

The role of enhancers methylation in the regulation of lineage routes has been observed in other species. For example, in zebrafish embryos, some enhancers regulating expression of genes involved in adult tissue formation get hypermethylated. This "dememorization" process is thought central to temporally regulate the activity of these enhancers and avoid their precocious activation during embryogenesis *(Wu et al., 2021)*. This is remnant of the function we propose for DNA methylation at PGCLC/NPC enhancers that get progressively methylated in embryogenesis to timely regulate their activation.

Then, other hypothesis were proposed to explain how default developmental routes could emerge. One model simply propose that different cis-regulating strategies could be in control of different lineage allocation. While one lineage could depend on differential availability of some specific cis-regulatory elements (CREs) determinant in its induction, other lineages would rather rely on differential binding strategies where binding of specific factors could be determinant in the lineage specification *(Delas et al., 2022)*. As a consequence, while the former strategy require deep chromatin remodeling for proper lineage allocation, the differential binding strategy only require different gene regulatory network (GRN) configurations. Finally, some argued that default routes also could emerge from differential attractivity of different GRNs. Indeed, while GRNs are central in defining new cell identities and cell types, they play also a central role in preventing cellular trans-differentiation or reprogramming. Hence, one could imagine that some GRN act like "higher attractors" and are more efficient in prevent cell fate changes. But this "Higher attractor theory" is still debated in the field *(Sparta et al., 2022)*.

# 4. Final Perspectives and Conclusions

## 4.1 Limitations and future perspectives of the study

### 4.1.1 Be aware of your genetic background and culture conditions

When I started this project, I found intriguing that two research groups came to different conclusions about the role of DNA methylation upon differentiation while using similar protocols and the same J1 *Dnmt*-TKO ESC line (129S4/SvJae strain). Briefly, one team showed that J1 TKO cells could not properly survive EBies differentiation, but nevertheless normally repressed pluripotency genes *(Sakaue et al., 2010)*, while the other observed long-term culture of J1 TKO EBies, but impaired repression of pluripotency genes and somatic markers *(Schmidt et al., 2012)*. This discrepancy prompted us to generate our own *Dnmt*-TKO cell line from an ESC line with lower passage to avoid any possible genetic and epigenetic drifts. We also used an ESC line with a slightly different genetic background (E14, 129/Ola strain), which is widely used in the ESC community. At the end, the results I obtained with our E14 TKO cells are more in line with the findings of Dr. Sakaue *(Sakaue et al., 2010)*: poor viability of TKO EBies, with nonetheless a proper exit from pluripotency.

Early on in the project, I tested how would the J1 TKO ESCs survive epiblast-like differentiation in our hands. Interestingly, during epiblast differentiation, the J1 TKO cells grew less that WT cells, but did not demonstrate stronger apoptosis. However, contrary to the early passage TKO E14 EpiLCs, J1 TKO EpiLCs had impaired activation of the proximal enhancers of *Oct4*, and higher reversion abilities toward a naïve state, suggesting improper or delayed priming. This was associated with uncomplete repression of naïve pluripotency genes. Finally, I tested a directed neural differentiation protocol that required a transient neurosphere growth phase **(Figure 18A-F)**. Our E14 TKO cells could not survive this aggregated phase, which was in contrast to the J1 TKO which never managed to activate neural genes and to repress pluripotency markers, in line with the results obtained by Dr. Schmidt *(Schmidt et al., 2012)*.

How such differences can arise between different ESC lines is difficult to assess. It was shown in some studies that genetic background can influence the regulation of naïve pluripotency and progression during epiblast differentiation *(Byers et al., 2022; Ortmann et al., 2020)*. Moreover, time spent in culture is also a source of stress for embryonic stem cells, leading to chromosomal abnormalities and affecting the ability of the cells to properly differentiate. Strikingly, it has been described that both E14 and J1 ESC lines are especially susceptible to aneuploidy *(Choi et al., 2017; Gaztelumendi and Nogués, 2014)*. Despite this, when we karyotyped our E14 cells, I did not see high proportion of aneuploid cells. Furthermore, functional analyses on primed E14 TKO cells at a higher passage did not demonstrate strong differences with low passage E14 TKO cells in regards to *Oct4* enhancer switch and reversion propensity (data not shown). Overall, this suggests that culture conditions and background can greatly influence the conclusions that can be drawn from cultured ESCs. Yet, the fact that we can still find convergent interpretations between our experiments on E14 TKO cells, and the findings from Dr. Sakaue with J1 TKO cells, supports our conclusions, beyond strain and technical variations.

### 4.1.2 Adapting our strategy

When I found that TKO cells presented longer opportunities to give rise to the germline, I initially thought that this could be due to subtle variations in the expression of some genes associated with the germline fate, in absence of DNA methylation. Hence I turned to a SMARTseq2 single cell RNA-seq approach, which can detect genes with low expression levels, but only allows processing of a limited number of cells, mostly because of operation costs *(Picelli et al., 2014)*. However, later on, our epigenome profiling revealed that the absence of DNA methylation primed the cells to adopt neural or germline fates, by allowing stronger licensing of a subset of putative enhancers for these cell types. This suggests that normally, in the WT background, only a few cells likely carry low methylation levels at these enhancers, and are being skewed toward these fates during EpiLC differentiation. If we had anticipated that these events might be rarer than we initially expected, we could have adopted a more suitable strategy to capture specific gene expression patterns only present in a few

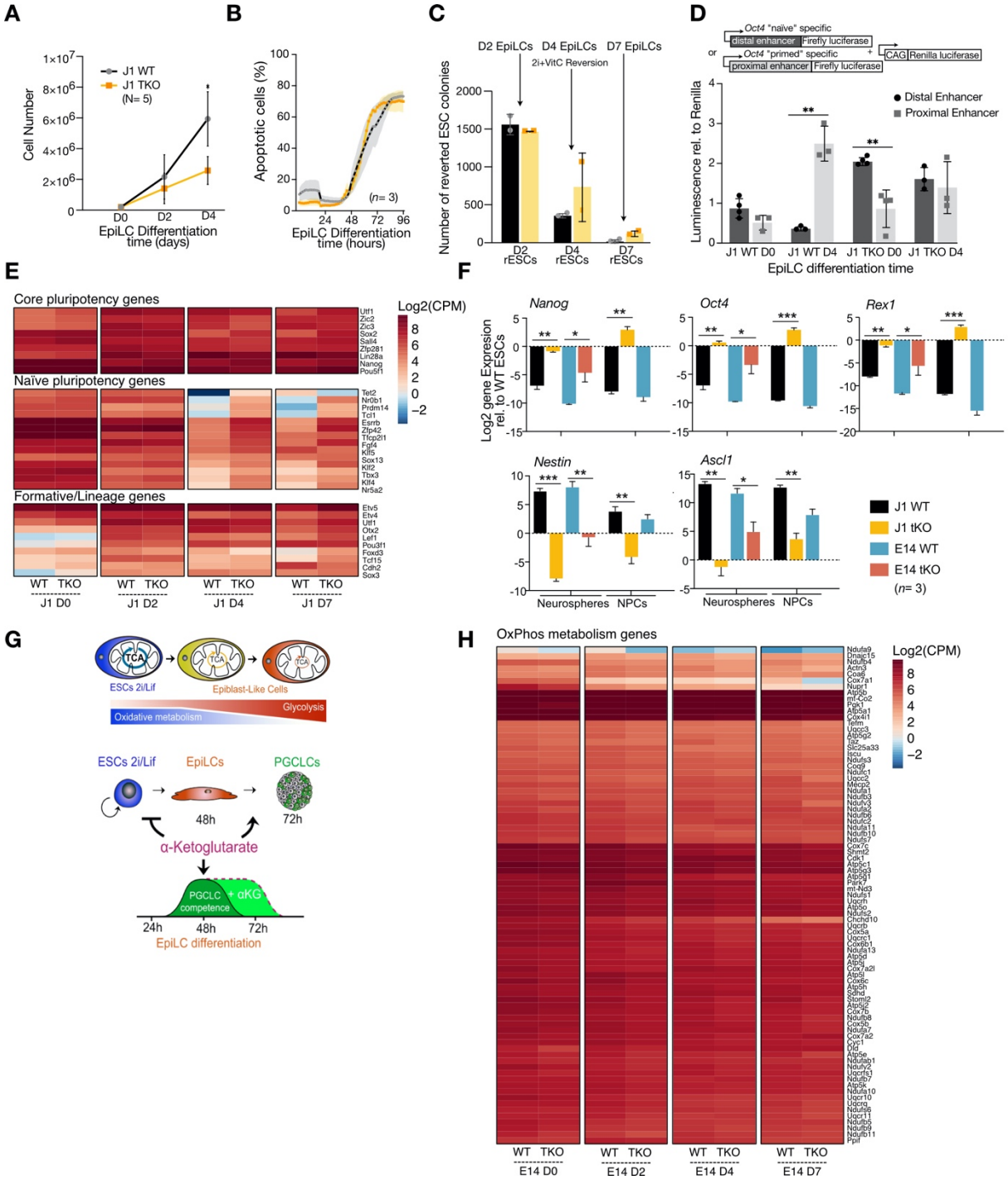



**Fig. 18 I ESCs of the J1 background diverge in their abilities for differentiation and the role of metabolism**

**A.** Growth curve of J1 WT (Black) and J1 TKO (Yellow) cells during EpiLC differentiation, data shown are from biological triplicates with mean ± SD. **B.** Apoptosis rate measured by live staining of Caspase 3/7 on Incucyte system of J1 WT and J1 TKO cells, data shown are from biological triplicates with mean ± SD. **C.** Barplot displaying number of counted reverted ESC colonies after alkaline phosphatase staining from J1 WT and J1 TKO EpiLCs at different days of EpiLC differentiation, data shown are from biological triplicates with mean ± SD. **D.** Barplot representing Luciferase activity from distal (dark grey) and proximal (light grey) enhancers of *Oct4* in J1 WT and TKO ESCs and D4 EpiLCs, data shown are mean ± SD from biological triplicate (two-tailed unpaired student t-test, *p<0.05, **p<0.005, ***p<0.0005). **E.** Heatmaps of pluripotency-associated markers showing normalized Log2 CPM counts during J1 WT and TKO EpiLC differentiation, expression is obtained for each condition from the average between biological duplicates. **F.** Barplot representing Log2 gene expression over WT ESCs, in E14 and J1 backgrounds for WT and TKO cells during neural differentiation, data shown are mean ± SD from biological triplicate with mean ± SD (two-tailed unpaired student t-test, *p<0.05, **p<0.005, ***p<0.0005). **G.** Graphical representation of metabolic switch during naïve-to-primed transitioning. Briefly, EpiLCs lose OxPhos metabolism as Glycolysis is established, but exposing EpiLCs to OxPhos metabolite aKG extends OxPhos and promotes germline fate (Tischler *et al*, Embo Reports 2018). **H.** Heatmaps of OxPhos metabolism-associated genes showing normalized Log2 CPM counts during E14 WT and TKO EpiLC differentiation, expression is obtained for each condition from the average between biological duplicates.

WT cells. Namely, we could have turned instead to Chromium 10X approach which, whilst demonstrating lower coverage than SMART, would have allowed to process more cells, increasing our chance to capture EpiLCs with skewed fates towards a neural or the germline fate *(Guo et al., 2019)*.

Moreover, other histone marks have been shown to define the priming of enhancers for specific cell fate decisions in epiblast differentiation. In particular, a subset of PGCLC-associated enhancers transiently accumulate MLL3/4-dependent H3K4me1 during EpiLC differentiation: this mark highlights germline competence and disappear from these enhancers upon further priming *(Bleckwehl et al., 2021)*. Hence it could be interesting to check if the absence of DNA methylation could affect the decommissioning of H3K4me1 at PGCLC enhancers, thus extending the time-window of germline competence during epiblast differentiation.

### 4.1.3 Metabolism, a possible new avenue?

Little is known about the control of metabolism in the nascent germline. However, some recent studies have started to uncover the regulation of metabolism during early development. Notably, metabolites have been shown to be play critical roles in the regulation of chromatin marks in the embryos, like the methyl-donor SAM (S adenosyl methionine) and L-2-hydroxyglutarate (L-2-HG) *(Haws et al., 2020; Zhao et al., 2021)*. In particular, a metabolic transition occurs between naïve and primed stem cells—with the former relying mostly on Oxidative Phosphorylation (OxPhos), the latter on Glycolysis—and transition seems to be partly regulated by many pluripotency and early differentiation genes *(Stirparo et al., 2021; Tsogtbaatar et al., 2020)*. In line of this work, it was shown that exposing EpiLCs to alpha-Ketoglutarate (aKG), a byproduct of OxPhos metabolism and LIF/STAT3 signaling, led to prolonged germline competence, in a very similar phenotype to our own findings in cells devoid of DNA methylation *(Tischler et al., 2018)*. The authors claimed that aKG in EpiLCs promotes the expression of naïve markers, like *Tfcp2l1* and *Esrrb*, and represses epiblast genes notably *Dnmt3B*, hence promoting germline competence **(Figure 18G)**. However, we cannot see how the lack of DNA methylation in TKO EpiLCs could impact metabolism, since we did not observe any majors shifts in OxPhos-associated gene expression in TKO EpiLCs **(Figure 18H)**. Moreover, it appeared from some other recently published

work that metabolism regulates DNA methylation in stem cells, but it is less clear if the reverse is true *(Betto et al., 2021)*. This suggests that metabolism, and the OxPhos/aKG axis in particular, could constitute a new actor of the regulation of naïve-to-primed pluripotency transition and cell fate decisions.

## 4.2 Final perspective and personal appreciation

The main conclusions drawn from my work have wide implications in the field of developmental biology and stem cell biology. First of all, I propose to extend the dispensability of DNA methylation to early primed epiblast stem cells, suggesting that DNA methylation might act as a guide and consolidator of cellular identities rather than a direct controller of the epiblast differentiation. Indeed, I found strong disconnection between DNA methylation and gene transcription, yet strong implication for DNA methylation in shaping the chromatin landscape of epiblast cells. Finally, the lack of DNA methylation promotes neural and germline fate acquisition, supporting default models of differentiation, and highlighting a central role for epigenetic marks in the regulation of such default pathways. This could have potential implications in the way we envision the impact of impaired DNA methylation in cancer cells. Indeed, many cancer cells are affected in their methylation profiles, which is associated with germline gene reactivation *(Simpson et al., 2005; Van Tongelen et al., 2017)*. While we can draw a direct link between DNA methylation and its repressive role at the promoters of meiotic genes, my work also supports the notion that the lack of DNA methylation could have additional effects, notably by affecting other chromatin marks at distal regulatory elements/enhancers of germline genes, skewing cancer cells towards adopting a germline-like identity.

Finally, this four year long research project was at the same time, the hardest and the most exciting project I have ever undertaken. For this thesis, I had the opportunity to explore biology at so many different levels: from single cells to entire cell populations, from *in vitro* to *in vivo*, from molecular variations to functional and phenotypical characterizations. All of this work implied many failed experiments, negative results, patience, endurance and many optimizations, to get to my final results. But I could not have hoped for a better (default?) route to learn how to become a scientist.

# BIBLIOGRAPHY

Aapola, Shibuya, Scott, Ollila, Vihinen, Heino, Shintani, Kawasaki, Minoshima, Krohn, Antonarakis, Shimizu, Kudoh, Peterson. 2000. Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* **65**:293–298. doi:10.1006/geno.2000.6168

Acampora, Di Giovannantonio, Simeone. 2013. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. *Development* **140**:43–55. doi:10.1242/dev.085290

Aguilera-Castrejon, Oldak, Shani, Ghanem, Itzkovich, Slomovich, Tarazi, Bayerl, Chugaeva, Ayyash, Ashouokhi, Sheban, Livnat, Lasman, Viukov, Zerbib, Addadi, Rais, Cheng, Stelzer, Keren-Shaul, Shlomo, Massarwa, Novershtern, Maza, Hanna. 2021. Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature* **593**:119–124. doi:10.1038/s41586-021-03416-3

Akagi, Kuure, Uranishi, Koide, Costantini, Yokota. 2015. ETS-related transcription factors Etv4 and Etv5 are involved in proliferation and induction of differentiationassociated genes in embryonic stem (ES) cells. *J Biol Chem* **290**:22460–22473. doi:10.1074/jbc.M115.675595

Alajem, Roth, Ratgauzer, Bavli, Motzik, Lahav, Peled, Ram. 2021. DNA methylation patterns expose variations in enhancer-chromatin modifications during embryonic stem cell differentiation. *PLoS Genet* **17**:1–24. doi:10.1371/journal.pgen.1009498

Alarcon. 2010. Cell polarity regulator PARD6B is essential for trophectoderm formation in the preimplantation mouse embryo. *Biol Reprod* **83**:347–358. doi:10.1095/biolreprod.110.084400

Alda-Catalinas, Bredikhin, Hernando-Herraez, Santos, Kubinyecz, Eckersley-Maslin, Stegle, Reik. 2020. A Single-Cell Transcriptomics CRISPR-Activation Screen Identifies Epigenetic Regulators of the Zygotic Genome Activation Program. *Cell Syst* **11**:25-41.e9. doi:10.1016/j.cels.2020.06.004

Allen. 2015. Compelled by the Diagram: Thinking through C. H. Waddington's Epigenetic Landscape. *Contemp Hist Presence Vis Cult* **4**:119–142. doi:10.5195/contemp.2015.143

Allfrey, Faulkner, Mirsky. 1964. Acetylation and Methylation of Histones and Their Possible Role in the. *Proc Natl Acad Sci United States* **51**:786–794. doi:10.1073/pnas.51.5.786

Allis, Jenuwein. 2016. The molecular hallmarks of epigenetic control. *Nat Rev Genet* **17**:487–500. doi:10.1038/nrg.2016.59

Almeida, Vernaz, Putman, Miska. 2022. Taming transposable elements in vertebrates: from epigenetic silencing to domestication. *Trends Genet*. doi:10.1016/j.tig.2022.02.009

Amouroux, Nashun, Shirane, Nakagawa, Hill, D'Souza, Nakayama, Matsuda, Turp, Ndjetehe, Encheva, Kudo, Koseki, Sasaki, Hajkova. 2016. De novo DNA methylation drives 5hmC accumulation in mouse zygotes. *Nat Cell Biol* **18**:225–233. doi:10.1038/ncb3296

Anani, Bhat, Honma-Yamanaka, Krawchuk, Yamanaka. 2014. Initiation of Hippo signaling is linked to polarity rather than to cell position in the pre-implantation mouse embryo. *Dev* **141**:2813–2824. doi:10.1242/dev.107276

Angermueller, Clark, Lee, Macaulay, Teng, Hu, Krueger, Smallwood, Ponting, Voet, Kelsey, Stegle, Reik. 2016. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* **13**:229–232. doi:10.1038/nmeth.3728

Aoki, Worrad, Schultz. 1997. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev Biol* **181**:296–307. doi:10.1006/dbio.1996.8466

Aoto, Saitoh, Ichimura, Niwa, Nakao. 2006. Nuclear and chromatin reorganization in the MHC-Oct3/4 locus at developmental phases of embryonic stem cell differentiation. *Dev Biol* **298**:354–367. doi:10.1016/j.ydbio.2006.04.450

Appleby, Bredenoord. 2018. Should the 14-day rule for embryo research become the 28-day rule? *EMBO Mol Med* **10**:7–10. doi:10.15252/emmm.201809437

Aramaki, Kagiwada, Wu, Obridge, Adachi, Kutejova, Lickert, Hübner, Schöler. 2021. Residual pluripotency is required for inductive germ cell segregation. *EMBO Rep* 1–15. doi:10.15252/embr.202152553

Argelaguet, Clark, Mohammed, Stapel, Krueger, Kapourani, Imaz-Rosshandler, Lohoff, Xiang, Hanna, Smallwood, Ibarra-Soria, Buettner, Sanguinetti, Xie, Krueger, Göttgens, Rugg-Gunn, Kelsey, Dean, Nichols, Stegle, Marioni, Reik. 2019a. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**:487–491. doi:10.1038/s41586-019-1825-8

Argelaguet, Mohammed, Clark, Stapel, Krueger, Kapourani, Xiang, Hanna, Smallwood, Ibarra-soria, Buettner, Sanguinetti, Krueger, Xie, Rugg-gunn, Kelsey, Dean, Nichols, Stegle, Marioni, Reik. 2019b. Single cell multi-omics profiling reveals a hierarchical epigenetic landscape during mammalian germ layer specification. *bioRxiv*. doi:10.1101/519207

Arnold, Hofmann, Bikoff, Robertson. 2008. Pivotal roles for eomesodermin during axis formation,

epithelium-to-mesenchyme transition and endoderm specification in the mouse. *Development* **135**:501–511. doi:10.1242/dev.014357
Arnold, Robertson. 2009. Making a commitment: Cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol* **10**:91–103. doi:10.1038/nrm2618
Artus, Hadjantonakis. 2011. Generation of Chimeras by Aggregation of Embryonic Stem Cells with Diploid or Tetraploid Mouse EmbryosMethods. pp. 37–56. doi:10.1007/978-1-60761-974-1_3
Atlasi, Megchelenbrink, Peng, Habibi, Joshi, Wang, Wang, Logie, Poser, Marks, Stunnenberg. 2019. Epigenetic modulation of a hardwired 3D chromatin landscape in two naive states of pluripotency. *Nat Cell Biol* **21**:568–578. doi:10.1038/s41556-019-0310-9
Aubin, Davy, Soriano. 2004. In vivo convergence of BMP and MAPK signaling pathways: Impact of differential Smad1 phosphorylation on development and homeostasis. *Genes Dev* **18**:1482–1494. doi:10.1101/gad.1202604
Auclair, Guibert, Bender, Weber. 2014. Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. *Genome Biol* **15**:545. doi:10.1186/s13059-014-0545-5
Azuara, Perry, Sauer, Spivakov, Jørgensen, John, Gouti, Casanova, Warnes, Merkenschlager, Fisher. 2006. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8**:532–538. doi:10.1038/ncb1403
Banerjee, Saxena, Kumar, Chen, Cavazza, Toke, O'Neill, Madha, Jadhav, Verzi, Shivdasani. 2018. Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development. *Genes Dev* **32**:1430–1442. doi:10.1101/gad.318832.118
Bao, Tang, Li, Hayashi, Gillich, Lao, Surani. 2009. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**:1292–1295. doi:10.1038/nature08534
Bao, Tang, Wu, Kim, Li, Li, Kobayashi, Lee, Chen, Wei, Li, Dietmann, Tang, Li, Surani. 2018. Derivation of hypermethylated pluripotent embryonic stem cells with high potency. *Cell Res* **28**:22–34. doi:10.1038/cr.2017.134
Barau, Teissandier, Zamudio, Roy, Nalesso, Hérault, Guillou, Bourc'his. 2016. The novel DNA methyltransferase DNMT3C protects male germ cells from transposon activity. *Science* **2**:(In Press). doi:10.1126/science.aah5143
Barral, Pozo, Ducrot, Papadopoulos, Sauzet, Oldfield, Cavalli, Déjardin. 2022. SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Mol Cell* 1–17. doi:10.1016/j.molcel.2021.12.037
Bartolomei, Ferguson-Smith. 2011. Mammalian genomic imprinting. *Cold Spring Harb Perspect Biol* **3**:1–17. doi:10.1101/cshperspect.a002592
Barton, Surani, Norris. 1984. Role of paternal and maternal genomes in mouse development. *Nature* **311**:374–376. doi:10.1038/311374a0
Baubec, Colombo, Wirbelauer, Schmidt, Burger, Krebs, Akalin, Schübeler. 2015. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**:243–247. doi:10.1038/nature14176
Baubec, Ivánek, Lienert, Schübeler. 2013. Methylation-dependent and -independent genomic targeting principles of the mbd protein family. *Cell* **153**:480–492. doi:10.1016/j.cell.2013.03.011
Beccari, Moris, Girgin, Turner, Baillie-Johnson, Cossy, Lutolf, Duboule, Arias. 2018. Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* 1. doi:10.1038/s41586-018-0578-0
Bedzhov, Zernicka-Goetz. 2014. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* **156**:1032–1044. doi:10.1016/j.cell.2014.01.023
Bell, Curry, Megchelenbrink, Jouneau, Brochard, Tomaz, Mau, Atlasi, de Souza, Marks, Stunnenberg, Jouneau, Azuara. 2020. Dynamic CpG methylation delineates subregions within super-enhancers selectively decommissioned at the exit from naive pluripotency. *Nat Commun* **11**:1112. doi:10.1038/s41467-020-14916-7
Bellacosa, Drohat. 2015. Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites. *DNA Repair (Amst)* **32**:33–42. doi:10.1016/j.dnarep.2015.04.011
Bentsen, Goymann, Schultheis, Klee, Petrova, Wiegandt, Fust, Preussner, Kuenne, Braun, Kim, Looso. 2020. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat Commun* **11**. doi:10.1038/s41467-020-18035-1
Bernard, Dubois, Heurtier, Fischer, Gonzalez, Chervova, Tachtsidi, Gil, Owens, Bates, Vandormael-Pournin, Silva, Ulitsky, Cohen-Tannoudji, Navarro. 2022. OCT4 activates a Suv39h1 -repressive antisense lncRNA to couple histone H3 Lysine 9 methylation to pluripotency. *Nucleic Acids Res*

1–13. doi:10.1093/nar/gkac550
Bernstein, Mikkelsen, Xie, Kamal, Huebert, Cuff, Fry, Meissner, Wernig, Plath, Jaenisch, Wagschal, Feil, Schreiber, Lander. 2006. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**:315–326. doi:10.1016/j.cell.2006.02.041
Bestor, Laudano, Mattaliano, Ingram. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* **203**:971–983. doi:10.1016/0022-2836(88)90122-2
Betto, Diamante, Perrera, Audano, Rapelli, Lauria, Incarnato, Arboit, Pedretti, Rigoni, Guerineau, Touboul, Stirparo, Lohoff, Boroviak, Grumati, Soriano, Nichols, Mitro, Oliviero, Martello. 2021. Metabolic control of DNA methylation in naive pluripotent cells. *Nat Genet*. doi:10.1038/s41588-020-00770-2
Bewick, Hofmeister, Powers, Mondo, Grigoriev, James, Stajich, Schmitz. 2019. Diversity of cytosine methylation across the fungal tree of life. *Nat Ecol Evol* **3**:479–490. doi:10.1038/s41559-019-0810-9
Bewick, Vogel, Moore, Schmitz. 2017. Evolution of DNA methylation across insects. *Mol Biol Evol* **34**:654–665. doi:10.1093/molbev/msw264
Bierkamp, Luxey, Metchat, Audouard, Dumollard, Christians. 2010. Lack of maternal Heat Shock Factor 1 results in multiple cellular and developmental defects, including mitochondrial damage and altered redox homeostasis, and leads to reduced survival of mammalian oocytes and embryos. *Dev Biol* **339**:338–353. doi:10.1016/j.ydbio.2009.12.037
Biessmann, Mason, Ferry, d'Hulst, Valgeirsdottir, Traverse, Pardue. 1990. Addition of telomere-associated HeT DNA sequences "heals" broken chromosome ends in Drosophila. *Cell* **61**:663–673. doi:10.1016/0092-8674(90)90478-W
Bird. 2007. Perceptions of epigenetics. *Nature* **447**:396–398. doi:10.1038/nature05913
Bird. 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* **8**:1499–1504. doi:10.1093/nar/8.7.1499
Blackledge, Farcas, Kondo, King, McGouran, Hanssen, Ito, Cooper, Kondo, Koseki, Ishikura, Long, Sheahan, Brockdorff, Kessler, Koseki, Klose. 2014. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**:1445–1459. doi:10.1016/j.cell.2014.05.004
Blackshaw, Rodger. 2021. Why we should not extend the 14-day rule. *J Med Ethics* **47**:712–714. doi:10.1136/medethics-2021-107317
Blakeley, Fogarty, Valle, Wamaitha, Hu, Elder, Snell, Christie, Robson, Niakan, Blakeley, Fogarty, Valle, Wamaitha, Hu, Elder, Snell, Christie, Robson, Niakan. 2015. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Dev* **142**:3613. doi:10.1242/dev.131235
Blaschke, Ebata, Karimi, Zepeda-Martínez, Goyal, Mahapatra, Tam, Laird, Hirst, Rao, Lorincz, Ramalho-Santos. 2013. VitaminC induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**:222–226. doi:10.1038/nature12362
Blassberg, Patel, Watson, Gouti, Metzis, Delás, Briscoe. 2022. Sox2 levels regulate the chromatin occupancy of WNT mediators in epiblast progenitors responsible for vertebrate body formation. *Nat Cell Biol* **24**:633–644. doi:10.1038/s41556-022-00910-2
Bleckwehl, Crispatzu, Schaaf, Respuela, Bartusel, Benson, Clark, Dorighi, Barral, Laugsch, van IJcken, Manzanares, Wysocka, Reik, Rada-Iglesias. 2021. Enhancer-associated H3K4 methylation safeguards in vitro germline competence. *Nat Commun* **12**:5771. doi:10.1038/s41467-021-26065-6
Bleckwehl, Rada-Iglesias. 2019. Transcriptional and epigenetic control of germline competence and specification. *Curr Opin Cell Biol* **61**:1–8. doi:10.1016/j.ceb.2019.05.006
Bledau, Schmidt, Neumann, Hill, Ciotta, Gupta, Torres, Fu, Kranz, Stewart, Anastassiadis. 2014. The H3K4 methyltransferase Setd1a is first required at the epiblast stage, whereas Setd1b becomes essential after gastrulation. *Dev* **141**:1022–1035. doi:10.1242/dev.098152
Bogdanović, Long, Van Heeringen, Brinkman, Gómez-Skarmeta, Stunnenberg, Jones, Veenstra. 2011. Temporal uncoupling of the DNA methylome and transcriptional repression during embryogenesis. *Genome Res* **21**:1313–1327. doi:10.1101/gr.114843.110
Bogutz, Brind'Amour, Kobayashi, Jensen, Nakabayashi, Imai, Lorincz, Lefebvre. 2019. Evolution of imprinting via lineage-specific insertion of retroviral promoters. *Nat Commun* **10**:1–14. doi:10.1038/s41467-019-13662-9
Bone, Lavender, Richman, Palmer, Turner, Kuroda. 1994. Acetylated histone H4 on the male X

chromosome is associated with dosage compensation in Drosophila. *Genes Dev* **8**:96–104. doi:10.1101/gad.8.1.96
Bonev, Mendelson Cohen, Szabo, Fritsch, Papadopoulos, Lubling, Xu, Lv, Hugnot, Tanay, Cavalli. 2017. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**:557-572.e24. doi:10.1016/j.cell.2017.09.043
Borgel, Guibert, Li, Chiba, Schübeler, Sasaki, Forné, Weber. 2010. Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* **42**:1093–1100. doi:10.1038/ng.708
Bostick, Kim, Estève, Clark, Pradhan, Jacobsen. 2007. UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science (80- )* **317**:1760–1764. doi:10.1126/science.1147939
Boulard, Edwards, Bestor. 2015. FBXL10 protects Polycomb-bound genes from hypermethylation. *Nat Genet* **47**:479–485. doi:10.1038/ng.3272
Bourc'his, Xu, Lin, Bollman, Bestor. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science (80- )* **294**:2536–2539. doi:10.1126/science.1065848
Bourque, Leong, Vega, Chen, Yen, Srinivasan, Chew, Ruan, Wei, Huck, Liu. 2008. Evolution of the mammalian transcription factor binding repertoire via transposable elements. *Genome Res* **18**:1752–1762. doi:10.1101/gr.080663.108
Braunstein, Sobel, Allis, Turner, Broach. 1996. Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. *Mol Cell Biol* **16**:4349–4356. doi:10.1128/mcb.16.8.4349
Brind'Amour, Kobayashi, Richard Albert, Shirane, Sakashita, Kamio, Bogutz, Koike, Karimi, Lefebvre, Kono, Lorincz. 2018. LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. *Nat Commun* **9**. doi:10.1038/s41467-018-05841-x
Brinkman, Gu, Bartels, Zhang, Matarese, Simmer, Marks, Bock, Gnirke, Meissner, Stunnenberg. 2012. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res* **22**:1128–1138. doi:10.1101/gr.133728.111
Brockdorff, Ashworth, Kay, Cooper, Smith, McCabe, Norris, Penny, Patel, Rastan. 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351**:329–331. doi:10.1038/351329a0
Brons, Smithers, Trotter, Rugg-Gunn, Sun, Chuva De Sousa Lopes, Howlett, Clarkson, Ahrlund-Richter, Pedersen, Vallier. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**:191–195. doi:10.1038/nature05950
Brook, Gardner. 1997. The origin and efficient derivation of embryonic stem cells in the mouse. *Proc Natl Acad Sci U S A* **94**:5709–5712. doi:10.1073/pnas.94.11.5709
Brown. 1966. Heterochromatin. *Science (80- )* **151**:417–425. doi:10.1126/science.151.3709.417
Brown C, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF. 1991. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature* **349**:38–44.
Brown, Di Cerbo, Feldmann, Ahn, Ito, Blackledge, Nakayama, McClellan, Dimitrova, Turberfield, Long, King, Kriaucionis, Schermelleh, Kutateladze, Koseki, Klose. 2017. The SET1 Complex Selects Actively Transcribed Target Genes via Multivalent Interaction with CpG Island Chromatin. *Cell Rep* **20**:2313–2327. doi:10.1016/j.celrep.2017.08.030
Brykczynska, Hisano, Erkek, Ramos, Oakeley, Roloff, Beisel, Schübeler, Stadler, Peters. 2010. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* **17**:679–687. doi:10.1038/nsmb.1821
Buecker, Srinivasan, Wu, Calo, Acampora, Faial, Simeone, Tan, Swigut, Wysocka. 2014. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**:838–53. doi:10.1016/j.stem.2014.04.003
Buenrostro, Giresi, Zaba, Chang, Greenleaf. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**:1213–1218. doi:10.1038/nmeth.2688
Byers, Spruce, Fortin, Hartig, Czechanski, Munger, Reinholdt, Skelly, Baker. 2022. Genetic control of the pluripotency epigenome determines differentiation bias in mouse embryonic stem cells. *EMBO J* **41**:1–19. doi:10.15252/embj.2021109445
Caldwell, Liu, Prasasya, Wang, DeNizio, Leu, Amoh, Krapp, Lan, Shields, Bonasio, Lengner, Kohli, Bartolomei. 2021. Functionally distinct roles for TET-oxidized 5-methylcytosine bases in somatic reprogramming to pluripotency. *Mol Cell* **81**:859-869.e8. doi:10.1016/j.molcel.2020.11.045
Campbell, McWhir, Ritchie, Wilmut. 1996. Sheep cloned by nuclear transfer from a cultured cell line.

*Nature* **380**:64–66. doi:10.1038/380064a0
Carlini, Policarpi, Hackett. 2022. Epigenetic inheritance is gated by naïve pluripotency and Dppa2. *EMBO J* 2021.05.11.443595. doi:10.15252/embj.2021108677
Carone, Fauquier, Habib, Shea, Hart, Li, Bock, Li, Gu, Zamore, Meissner, Weng, Hofmann, Friedman, Rando. 2010. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* **143**:1084–1096. doi:10.1016/j.cell.2010.12.008
Carrel, Burrows. 1911. Cultivation of tissues in vitro and its technique. *J Exp Med* **13**:387–396. doi:10.1084/jem.13.3.387
Casser, Israel, Witten, Schulte, Schlatt, Nordhoff, Boiani. 2017. Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos. *Sci Rep* **7**:1–15. doi:10.1038/s41598-017-08266-6
Catania, Dumesic, Pimentel, Nasif, Stoddard, Burke, Diedrich, Cook, Shea, Geinger, Lintner, Yates, Hajkova, Narlikar, Cuomo, Pritchard, Madhani. 2020. Evolutionary Persistence of DNA Methylation for Millions of Years after Ancient Loss of a De Novo Methyltransferase. *Cell* 149385. doi:10.1016/j.cell.2019.12.012
Cattanach, Kirk. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* **315**:496–498. doi:10.1038/315496a0
Cermola, D'Aniello, Tatè, De Cesare, Martinez-Arias, Minchiotti, Patriarca. 2021. Gastruloid Development Competence Discriminates Different States of Pluripotency. *Stem Cell Reports*. doi:10.1016/j.stemcr.2020.12.013
Chamberlain, Yee, Magnuson. 2008. Polycomb Repressive Complex 2 Is Dispensable for Maintenance of Embryonic Stem Cell Pluripotency. *Stem Cells* **26**:1496–1505. doi:10.1634/stemcells.2008-0102
Chambers, Colby, Robertson, Nichols, Lee, Tweedie, Smith. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**:643–655. doi:10.1016/S0092-8674(03)00392-1
Chambers, Silva, Colby, Nichols, Nijmeijer, Robertson, Vrana, Jones, Grotewold, Smith. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature* **450**:1230–4. doi:10.1038/nature06403
Chan, Morgan, Adrian Leu, Shetty, Cisse, Nugent, Morrison, Jašarević, Huang, Kanyuch, Rodgers, Bhanu, Berger, Garcia, Ament, Kane, Neill Epperson, Bale. 2020. Reproductive tract extracellular vesicles are sufficient to transmit intergenerational stress and program neurodevelopment. *Nat Commun* **11**:1–13. doi:10.1038/s41467-020-15305-w
Chang, Hemberg, Barahona, Ingber, Huang. 2008. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**:544–547. doi:10.1038/nature06965
Charlet, Duymich, Lay, Mundbjerg, Dalsgaard Sørensen, Liang, Jones. 2016. Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers. *Mol Cell* **62**:422–431. doi:10.1016/j.molcel.2016.03.033
Charlton, Jung, Mattei, Bailly, Liao, Martin, Giesselmann, Brändl, Stamenova, Müller, Kiskinis, Gnirke, Smith, Meissner. 2020. TETs compete with DNMT3 activity in pluripotent cells at thousands of methylated somatic enhancers. *Nat Genet*. doi:10.1038/s41588-020-0639-9
Chazaud, Yamanaka, Pawson, Rossant. 2006. Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Dev Cell* **10**:615–624. doi:10.1016/j.devcel.2006.02.020
Chen, Mann, Hsieh, Riggs, Chédin. 2005. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *J Cell Biochem* **95**:902–917. doi:10.1002/jcb.20447
Chen, Shi, Tao, Zernicka-Goetz. 2018. Tracing the origin of heterogeneity and symmetry breaking in the early mammalian embryo. *Nat Commun* **9**:1–11. doi:10.1038/s41467-018-04155-2
Chen, Sun, Hou, Kim, Faith, Aslanyan, Tao, Zheng, Fu, Liu, Kellis, Clark. 2019. Human Primordial Germ Cells Are Specified from Lineage-Primed Progenitors. *Cell Rep* **29**:4568-4582.e5. doi:10.1016/j.celrep.2019.11.083
Chen, Tsujimoto, Li. 2004. The PWWP Domain of Dnmt3a and Dnmt3b Is Required for Directing DNA Methylation to the Major Satellite Repeats at Pericentric Heterochromatin. *Mol Cell Biol* **24**:9048–9058. doi:10.1128/mcb.24.20.9048-9058.2004
Chen, Yan, Cao, Li, Zhang, Shi, Feng, Peng, Zhang, Zhang, Qian, Duan, Zhai, Zhou. 2016. Sperm tsRNAs contribute to intergenerational inheritance of an acquired metabolic disorder. *Science (80- )* **351**:397–400. doi:10.1126/science.aad7977

Chénais. 2013. Transposable elements and human cancer: A causal relationship? *Biochim Biophys Acta - Rev Cancer* **1835**:28–35. doi:10.1016/j.bbcan.2012.09.001

Chew, Loh, Zhang, Chen, Tam, Yeap, Li, Ang, Lim, Robson, Ng. 2005. Reciprocal Transcriptional Regulation of Pou5f1 and Sox2 via the Oct4/Sox2 Complex in Embryonic Stem Cells . *Mol Cell Biol* **25**:6031–6046. doi:10.1128/mcb.25.14.6031-6046.2005

Choi, Huebner, Clement, Walsh, Savol, Lin, Gu, Di Stefano, Brumbaugh, Kim, Sharif, Rose, Mohammad, Odajima, Charron, Shioda, Gnirke, Gygi, Koseki, Sadreyev, Xiao, Meissner, Hochedlinger. 2017. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**:219–223. doi:10.1038/nature23274

Choi, Joo, Hong, Kim, Song, Lee, Wu, Schöler, Do. 2016. Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency. *Stem Cell Reports* **7**:911–926. doi:10.1016/j.stemcr.2016.09.012

Clark, Argelaguet, Kapourani, Stubbs, Lee, Alda-Catalinas, Krueger, Sanguinetti, Kelsey, Marioni, Stegle, Reik. 2018. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* **9**:781. doi:10.1038/s41467-018-03149-4

Cooper, Dienstbier, Hassan, Schermelleh, Sharif, Blackledge, DeMarco, Elderkin, Koseki, Klose, Heger, Brockdorff. 2014. Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep* **7**:1456–1470. doi:10.1016/j.celrep.2014.04.012

Cota, Sánchez-Romero, Hernández, Pucciarelli, García-Del Portillo, Casadesús. 2015. Epigenetic Control of Salmonella enterica O-Antigen Chain Length: A Tradeoff between Virulence and Bacteriophage Resistance. *PLoS Genet* **11**. doi:10.1371/journal.pgen.1005667

Coticchio, Dal Canto, Renzini, Guglielmo, Brambillasca, Turchi, Novara, Fadini. 2014. Oocyte maturation: Gamete-somatic cells interactions, meiotic resumption, cytoskeletal dynamics and cytoplasmic reorganization. *Hum Reprod Update* **21**:427–454. doi:10.1093/humupd/dmv011

Coucouvanis, Martin. 1995. Signals for death and survival: A two-step mechanism for cavitation in the vertebrate embryo. *Cell* **83**:279–287. doi:10.1016/0092-8674(95)90169-8

Cusack, King, Spingardi, Kessler, Klose, Kriaucionis. 2020. Distinct contributions of DNA methylation and histone acetylation to the genomic occupancy of transcription factors. *Genome Res* 1–14. doi:10.1101/gr.257576.119

Cuypers, Dumetz, Meysman, Laukens, De Muylder, Dujardin, Domagalska. 2020. The absence of c-5 dna methylation in leishmania donovani allows dna enrichment from complex samples. *Microorganisms* **8**:1–18. doi:10.3390/microorganisms8081252

Dahl, Jung, Aanes, Greggains, Manaf, Lerdrup, Li, Kuan, Li, Lee, Preissl, Jermstad, Haugen, Suganthan, Bjørås, Hansen, Dalen, Fedorcsak, Ren, Klungland. 2016. Broad histone H3K4me3 domains in mouse oocytes modulate maternal-to-zygotic transition. *Nature* **537**:548–552. doi:10.1038/nature19360

Dahlet, Argüeso Lleida, Al Adhami, Dumas, Bender, Ngondo, Tanguy, Vallet, Auclair, Bardet, Weber. 2020. Genome-wide analysis in the mouse embryo reveals the importance of DNA methylation for transcription integrity. *Nat Commun* **11**:1–14. doi:10.1038/s41467-020-16919-w

Dahlet, Truss, Frede, Al Adhami, Bardet, Dumas, Vallet, Chicher, Hammann, Kottnik, Hansen, Luz, Alvarez, Auclair, Hecht, Robinson, Hagemeier, Weber. 2021. E2F6 initiates stable epigenetic silencing of germline genes during embryonic development. *Nat Commun* **12**:3582. doi:10.1038/s41467-021-23596-w

Dai, Wang, Yang, Chen, Zhu, Sun, Ge, Wang, Chapman, Tang, Sun, Xu. 2016. TET-mediated DNA demethylation controls gastrulation by regulating Lefty-Nodal signalling. *Nature* **538**:528–532. doi:10.1038/nature20095

Dawlaty, Breiling, Le, Barrasa, Raddatz, Gao, Powell, Cheng, Faull, Lyko, Jaenisch. 2014. Loss of tet enzymes compromises proper differentiation of embryonic stem cells. *Dev Cell* **29**:102–111. doi:10.1016/j.devcel.2014.03.003

Dawlaty, Ganz, Powell, Hu, Markoulaki, Cheng, Gao, Kim, Choi, Page, Jaenisch. 2011. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**:166–175. doi:10.1016/j.stem.2011.07.010

De Los Angeles, Ferrari, Xi, Fujiwara, Benvenisty, Deng, Hochedlinger, Jaenisch, Lee, Leitch, Lensch, Lujan, Pei, Rossant, Wernig, Park, Daley. 2015. Hallmarks of pluripotency. *Nature* **525**:469–478. doi:10.1038/nature15515

De Mendoza, Bonnet, Vargas-Landin, Ji, Hong, Yang, Li, Hori, Pflueger, Buckberry, Ohta, Rosic, Lesage, Lin, Lister. 2018. Recurrent acquisition of cytosine methyltransferases into eukaryotic

retrotransposons. *Nat Commun* **9**:1–11. doi:10.1038/s41467-018-03724-9
de Mendoza, Lister, Bogdanovic. 2020. Evolution of DNA Methylome Diversity in Eukaryotes. *J Mol Biol* **432**:1687–1705. doi:10.1016/j.jmb.2019.11.003
Deathridge, Antolović, Parsons, Chubb. 2019. Live imaging of erk signalling dynamics in differentiating mouse embryonic stem cells. *Dev* **146**. doi:10.1242/dev.172940
Delas, Kalaitzis, Fawzi, Demuth, Zhang, Stuart, Costantini, Ivanovitch, Tanaka, Briscoe. 2022. Developmental cell fate choice employs two distinct cis regulatory strategies. *Biorxv*. doi:10.1101/2022.06.06.494792
Delaune, Lemaire, Kodjabachian. 2005. Neural induction in Xenopus requires early FGF signalling in addition to BMP inhibition. *Development* **132**:299–310. doi:10.1242/dev.01582
Dhaliwal, Miri, Davidson, Tamim El Jarkass, Mitchell. 2018. KLF4 Nuclear Export Requires ERK Activation and Initiates Exit from Naive Pluripotency. *Stem Cell Reports* **10**:1308–1323. doi:10.1016/j.stemcr.2018.02.007
Dhayalan, Rajavelu, Rathert, Tamas, Jurkowska, Ragozin, Jeltsch. 2010. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J Biol Chem* **285**:26114–26120. doi:10.1074/jbc.M109.089433
Di Giovannantonio, Acampora, Omodei, Nigro, Barba, Barbieri, Chambers, Simeone. 2021. Direct repression of Nanog and Oct4 by OTX2 modulates the contribution of epiblast-derived cells to germline and somatic lineage. *Development* **148**. doi:10.1242/dev.199166
Dias, Ressler. 2014. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* **17**:89–96. doi:10.1038/nn.3594
Dixon, Pan, Yang, Rosen, Jashari, Verma, Pulecio, Caspi, Lee, Stransky, Glezer, Liu, Rivas, Kumar, Lan, Torregroza, He, Sidoli, Evans, Elemento, Huangfu. 2021. QSER1 protects DNA methylation valleys from de novo methylation. *Science (80- )* **372**. doi:10.1126/science.abd0875
Domcke, Bardet, Adrian Ginno, Hartl, Burger, Schübeler. 2015. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* **528**:575–579. doi:10.1038/nature16462
Dong, Maksakova, Mohn, Leung, Appanah, Lee, Yang, Lam, Mager, Schübeler, Tachibana, Shinkai, Lorincz. 2008. DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27**:2691–2701. doi:10.1038/emboj.2008.193
Duffié, Ajjan, Greenberg, Zamudio, Escamilla, Iranzo, Okamoto, Barbaux, Fauque. 2014. The Gpr1 / Zdbf2 locus provides new paradigms for transient and dynamic genomic imprinting in mammals. *Genes Dev* **28**:463–478. doi:10.1101/gad.232058.113.4
Dumesic, Stoddard, Catania, Narlikar, Madhani. 2020. ATP Hydrolysis by the SNF2 Domain of Dnmt5 Is Coupled to Both Specific Recognition and Modification of Hemimethylated DNA. *Mol Cell* **79**:127-139.e4. doi:10.1016/j.molcel.2020.04.029
Dunican, Cruickshanks, Suzuki, Semple, Davey, Arceci, Greally, Adams, Meehan. 2013. Lsh regulates LTR retrotransposon repression independently of Dnmt3b function. *Genome Biol* **14**:1–23. doi:10.1186/gb-2013-14-12-r146
Dunn, Martello, Yordanov, Emmott, Smith. 2014. Defining an essential transcription factor program for naive pluripotency. *Science (80- )* **344**:1156–1160. doi:10.1126/science.1248882
Dupressoir, Marceau, Vernochet, Benit, Kanellopoulos, Sapin, Heidmann. 2005. Syncytin-A and syncytin-B, two fusogenic placenta-specific murine envelope genes of retroviral origin conserved in Muridae. *Proc Natl Acad Sci* **102**:725–730. doi:10.1073/pnas.0406509102
Dura, Teissandier, Armand, Barau, Bonneville, Weber, Baudrin, Lameiras, Bourc'his. 2021. DNMT3A-dependent DNA methylation is required 2 for spermatogonial stem cells to commit to spermatogenesis. *bioRxiv* 2021.04.19.440465. doi:10.1101/2021.04.19.440465
Dura, Teissandier, Armand, Barau, Lapoujade, Fouchet, Bonneville, Schulz, Weber, Baudrin, Lameiras, Bourc'his. 2022. DNMT3A-dependent DNA methylation is required for spermatogonial stem cells to commit to spermatogenesis. *Nat Genet* **54**:469–480. doi:10.1038/s41588-022-01040-z
Duval, Vaslin, Barata, Frarma, Contremoulins, Baudin, Nedelec, Ribes. 2019. Bmp4 patterns smad activity and generates stereotyped cell fate organization in spinal organoids. *Dev* **146**. doi:10.1242/dev.175430
Eckersley-Maslin, Alda-Catalinas, Blotenburg, Kreibich, Krueger, Reik. 2019. Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program. *Genes Dev* **33**:194–208. doi:10.1101/gad.321174.118
Eckersley-Maslin, Svensson, Krueger, Stubbs, Giehr, Krueger, Miragaia, Kyriakopoulos, Berrens, Milagre, Walter, Teichmann, Reik. 2016. MERVL/Zscan4 Network Activation Results in Transient

Genome-wide DNA Demethylation of mESCs. *Cell Rep* **17**:179–192. doi:10.1016/j.celrep.2016.08.087
Eldar, Elowitz. 2010. Functional roles for noise in genetic circuits. *Nature* **467**:167–173. doi:10.1038/nature09326
En, Beard, Jaenisch. 1993. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature* **366**:362–365.
Endoh, Endo, Shinga, Hayashi, Farcas, Ma, Ito, Sharif, Endoh, Onaga, Nakayama, Ishikura, Masui, Kessler, Suda, Ohara, Okuda, Klose, Koseki. 2017. PCGF6-PRC1 suppresses premature differentiation of mouse embryonic stem cells by regulating germ cell-related genes. *Elife* **6**:1–26. doi:10.7554/eLife.21064
Epsztejn-Litman, Feldman, Abu-Remaileh, Shufaro, Gerson, Ueda, Deplus, Fuks, Shinkai, Cedar, Bergman. 2008. De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* **15**:1176–1183. doi:10.1038/nsmb.1476
Erickson, Kameneva, Adameyko. 2022. The transcriptional portraits of the neural crest at the individual cell level. *Semin Cell Dev Biol*. doi:10.1016/j.semcdb.2022.02.017
Erives. 2017. Phylogenetic analysis of the core histone doublet and DNA topo II genes of Marseilleviridae: Evidence of proto-eukaryotic provenance. *Epigenetics and Chromatin* **10**:1–11. doi:10.1186/s13072-017-0162-0
Falco, Lee, Stanghellini, Bassey, Hamatani, Ko. 2007. Zscan4: A novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells. *Dev Biol* **307**:539–550. doi:10.1016/j.ydbio.2007.05.003
Fang, Munera, Friedman, Mandlik, Chao, Banerjee, Feng, Losic, Mahajan, Jabado, Deikus, Clark, Luong, Murray, Davis, Keren-Paz, Chess, Roberts, Korlach, Turner, Kumar, Waldor, Schadt. 2012. Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using single-molecule real-time sequencing. *Nat Biotechnol* **30**:1232–1239. doi:10.1038/nbt.2432
Fatemi, Hermann, Pradhan, Jeltsch. 2001. The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol* **309**:1189–1199. doi:10.1006/jmbi.2001.4709
Festuccia, Gonzalez, Navarro. 2017. The Epigenetic Paradox of Pluripotent ES Cells. *J Mol Biol* **429**:1476–1503. doi:10.1016/j.jmb.2016.12.009
Festuccia, Owens, Chervova, Dubois, Navarro. 2021. The combined action of Esrrb and Nr5a2 is essential for murine naïve pluripotency., Development. doi:10.1242/dev.199604
Ficz, Hore, Santos, Lee, Dean, Arand, Krueger, Oxley, Paul, Walter, Cook, Andrews, Branco, Reik. 2013. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**:351–359. doi:10.1016/j.stem.2013.06.004
Fioravanti, Mautone, Rovere, Rotili, Mai. 2020. Targeting histone acetylation/deacetylation in parasites: an update (2017–2020). *Curr Opin Chem Biol* **57**:65–74. doi:10.1016/j.cbpa.2020.05.008
Fitz-James, Cavalli. 2022. Molecular mechanisms of transgenerational epigenetic inheritance. *Nat Rev Genet*. doi:10.1038/s41576-021-00438-5
Fu, Wu, Djekidel, Zhang. 2019. Myc and Dnmt1 impede the pluripotent to totipotent state transition in embryonic stem cells. *Nat Cell Biol* **21**:835–844. doi:10.1038/s41556-019-0343-0
Fu, Zhang, Zhang. 2020. Epigenetic regulation of mouse preimplantation embryo development. *Curr Opin Genet Dev* **64**:13–20. doi:10.1016/j.gde.2020.05.015
Fursova, Turberfield, Blackledge, Findlater, Lastuvkova, Huseyin, Dobrinic, Klose. 2021. BAP1 constrains pervasive H2AK119ub1 to control the transcriptional potential of the genome. *Genes Dev* **35**:749–770. doi:10.1101/GAD.347005.120
Gao, Yang, Chen, Wang, Wu, Ai, Du, Liu, Shi, Wu, Guo, Zhang. 2013. Vitamin C facilitates pluripotent stem cell maintenance by promoting pluripotency gene transcription. *Biochimie* **95**:2107–2113. doi:10.1016/j.biochi.2013.08.001
Gapp, Jawaid, Sarkies, Bohacek, Pelczar, Prados, Farinelli, Miska, Mansuy. 2014. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. *Nat Neurosci* **17**:667–669. doi:10.1038/nn.3695
Gardiner-Garden, Frommer. 1987. CpG Islands in vertebrate genomes. *J Mol Biol* **196**:261–282. doi:10.1016/0022-2836(87)90689-9
Gaztelumendi, Nogués. 2014. Chromosome Instability in mouse Embryonic Stem Cells. *Sci Rep* **4**:1–8. doi:10.1038/srep05324
Gendrel, Apedaile, Coker, Termanis, Zvetkova, Godwin, Tang, Huntley, Montana, Taylor, Giannoulatou,

Heard, Stancheva, Brockdorff. 2012. Smchd1-Dependent and -Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome. *Dev Cell* **23**:265–279. doi:10.1016/j.devcel.2012.06.011

Georgel, Horowitz-Scherer, Adkins, Woodcock, Wade, Hansen. 2003. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J Biol Chem* **278**:32181–32188. doi:10.1074/jbc.M305308200

Gervasi, Visconti. 2017. Molecular changes and signaling events occurring in spermatozoa during epididymal maturation. *Andrology* **5**:204–218. doi:10.1111/andr.12320

Geula, Moshitch-Moshkovitz, Dominissini, Mansour, Kol, Salmon-Divon, Hershkovitz, Peer, Mor, Manor, Ben-Haim, Eyal, Yunger, Pinto, Jaitin, Viukov, Rais, Krupalnik, Chomsky, Zerbib, Maza, Rechavi, Massarwa, Hanna, Amit, Levanon, Amariglio, Stern-Ginossar, Novershtern, Rechavi, Hanna. 2015. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science (80- )* **347**:1002–1006. doi:10.1126/science.1261417

Gibson, Young, Chuang, Venter, Hutchison, Smith. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**:343–345. doi:10.1038/nmeth.1318

Gillich, Bao, Grabole, Hayashi, Trotter, Pasque, Magnúsdóttir, Surani. 2012. Epiblast stem cell-based system reveals reprogramming synergy of germline factors. *Cell Stem Cell* **10**:425–439. doi:10.1016/j.stem.2012.01.020

Ginno, Gaidatzis, Feldmann, Hoerner, Imanci, Burger, Zilbermann, Peters, Edenhofer, Smallwood, Krebs, Schübeler. 2020. A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity. *Nat Commun* **11**:2680. doi:10.1038/s41467-020-16354-x

Girgin, Broguiere, Mattolini, Lutolf. 2021. Gastruloids generated without exogenous Wnt activation develop anterior neural tissues. *Stem Cell Reports* **16**:1–13. doi:10.1016/j.stemcr.2021.03.017

Gkountela, Zhang, Shafiq, Liao, Hargan-Calvopiña, Chen, Clark. 2015. DNA demethylation dynamics in the human prenatal germline. *Cell* **161**:1425–1436. doi:10.1016/j.cell.2015.05.012

Glaser, Schaft, Lubitz, Vintersten, van der Hoeven, Tuftteland, Aasland, Anastassiadis, Ang, Stewart. 2006. Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**:1423–1432. doi:10.1242/dev.02302

Gökbuget, Blelloch. 2019. Epigenetic control of transcriptional regulation in pluripotency and early differentiation. *Dev* **146**. doi:10.1242/dev.164772

Goll, Bestor. 2005. Eukaryotic Cytosine Methyltransferases. *Annu Rev Biochem* **74**:481–514. doi:10.1146/annurev.biochem.74.010904.153721

Goll, Kirpekar, Maggert, Yoder, Hsieh, Zhang, Golic, Jacobsen, Bestor. 2006. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science (80- )* **311**:395–398. doi:10.1126/science.1120976

Gouti, Tsakiridis, Wymeersch, Huang, Kleinjung, Wilson, Briscoe. 2014. In vitro generation of neuromesodermal progenitors reveals distinct roles for wnt signalling in the specification of spinal cord and paraxial mesoderm identity. *PLoS Biol* **12**. doi:10.1371/journal.pbio.1001937

Grabole, Tischler, Hackett, Kim, Tang, Leitch, Magnúsdóttir, Surani. 2013. Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep* **14**:629–637. doi:10.1038/embor.2013.67

Grand, Burger, Gräwe, Michael, Isbel, Hess, Hoerner, Iesmantavicius, Durdu, Pregnolato, Krebs, Smallwood, Thomä, Vermeulen, Schübeler. 2021. BANP opens chromatin and activates CpG-island-regulated genes. *Nature*. doi:10.1038/s41586-021-03689-8

Greenberg. 2021. Get Out and Stay Out: New Insights Into DNA Methylation Reprogramming in Mammals. *Front Cell Dev Biol* **8**:1–8. doi:10.3389/fcell.2020.629068

Greenberg, Bourc'his. 2019. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol*. doi:10.1038/s41580-019-0159-6

Greenberg, Glaser, Borsos, Marjou, Walter, Teissandier, Bourc'his. 2017. Transient transcription in the early embryo sets an epigenetic state that programs postnatal growth. *Nat Genet* **49**:110–118. doi:10.1038/ng.3718

Greenberg, Teissandier, Walter, Noordermeer, Bourc'his. 2019. Dynamic enhancer partitioning instructs activation of a growth-related gene during exit from naïve pluripotency. *Elife* **8**:1–20. doi:10.7554/eLife.44057

Gretarsson, Hackett. 2020. Dppa2 and Dppa4 counteract de novo methylation to establish a permissive epigenome for development. *Nat Struct Mol Biol* **27**:706–716. doi:10.1038/s41594-020-0445-1

Griswold. 2016. Spermatogenesis: The commitment to Meiosis. *Physiol Rev* **96**:1–17. doi:10.1152/physrev.00013.2015
Gruhn, Günesdogan. 2021. Generation of Primordial Germ Cell-like Cells on Small and Large Scales. pp. 75–89. doi:10.1007/978-1-0716-0958-3_6
Grunz, Tacke. 1989. Neural differentiation of Xenopus laevis ectoderm takes place after disaggregation and delayed reaggregation without inducer. *Cell Differ Dev* **28**:211–217. doi:10.1016/0922-3371(89)90006-3
Gu, Guo, Yang, Wu, Xu, Liu, Xie, Shi, He, Jin, Iqbal, Shi, Deng, Szabó, Pfeifer, Li, Xu. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**:606–612. doi:10.1038/nature10443
Gu, Hao, Woo, Huang, Guo, Lin, Guzman, Tovy, Rosas, Jeong, Zhou, Deneen, Huang, Li, Goodell. 2022. The disordered N-terminal domain of DNMT3A recognizes H2AK119ub and is required for postnatal development. *Nat Genet* **54**. doi:10.1038/s41588-022-01063-6
Guallar, Bi, Pardavila, Huang, Saenz, Shi, Zhou, Faiola, Ding, Haruehanroengra, Yang, Li, Sanchez-Priego, Saunders, Pan, Valdes, Kelley, Blanco, Chen, Wang, Sheng, Xu, Fidalgo, Shen, Wang. 2017. RNA-dependent chromatin targeting of TET2 for endogenous retrovirus control in pluripotent stem cells. *Nat Genet* **Accepted**. doi:10.1038/s41588-018-0060-9
Guo, Huss, Tong, Wang, Li Sun, Clarke, Robson. 2010. Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Dev Cell* **18**:675–685. doi:10.1016/j.devcel.2010.02.012
Guo F, Li, Liang, Li, Zhu, Guo, Wu, Wen, Gu, Hu, Walsh, Li, Tang, Xu. 2014. Active and Passive Demethylation of Male and Female Pronuclear DNA in the Mammalian Zygote. *Cell Stem Cell* **15**:447–459. doi:10.1016/j.stem.2014.08.003
Guo, Lin, Wang, Gao, Cao, Mai, Wu, Kuang, Liu, Yang, Chu, Song, Li, Liu, Wu, Liu, Wang, Pan, Hutchins, Liu, Pei, Chen. 2019. Resolving Cell Fate Decisions during Somatic Cell Reprogramming by Single-Cell RNA-Seq. *Mol Cell* **73**:815-829.e7. doi:10.1016/j.molcel.2019.01.042
Guo X, Wang, Li, Ding, Xiao, Yin, He, Shi, Dong, Li, Tian, Wang, Cong, Xu. 2015. Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. *Nature* **517**:640–644. doi:10.1038/nature13899
Guo F, Yan, Guo, Li, Hu, Zhao, Yong, Hu, Wang, Wei, Wang, Li, Yan, Zhi, Zhang, Jin, Zhang, Hou, Zhu, Li, Zhang, Liu, Ren, Zhu, Wen, Gao, Tang, Qiao. 2015. The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell* **161**:1437–1452. doi:10.1016/j.cell.2015.05.015
Guo H, Zhu, Yan, Li, Hu, Lian, Yan, Ren, Lin, Li, Jin, Shi, Liu, Wang, Wang, Wei, Li, Guo, Wu, Fan, Yong, Wen, Xie, Tang, Qiao. 2014. The DNA methylation landscape of human early embryos. *Nature* **511**:606–610. doi:10.1038/nature13544
Haberlandt. 1902. Kulturversuche mit Isolierten Pflanzenzellen.
Habibi, Brinkman, Arand, Kroeze, Kerstens, Matarese, Lepikhov, Gut, Brun-Heath, Hubner, Benedetti, Altucci, Jansen, Walter, Gut, Marks, Stunnenberg. 2013. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**:360–369. doi:10.1016/j.stem.2013.06.002
Hackett Jamie A., Dietmann, Murakami, Down, Leitch, Surani. 2013. Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Reports* **1**:518–531. doi:10.1016/j.stemcr.2013.11.010
Hackett, Kobayashi, Dietmann, Surani. 2017. Activation of Lineage Regulators and Transposable Elements across a Pluripotent Spectrum. *Stem Cell Reports* **8**:1645–1658. doi:10.1016/j.stemcr.2017.05.014
Hackett J. A., Sengupta, Zylicz, Murakami, Lee, Down, Surani. 2013. Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine. *Science (80- )* **339**:448–452. doi:10.1126/science.1229277
Hackett, Surani. 2014. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**:416–430. doi:10.1016/j.stem.2014.09.015
Hagarman, Motley, Kristjansdottir, Soloway. 2013. Coordinate Regulation of DNA Methylation and H3K27me3 in Mouse Embryonic Stem Cells. *PLoS One* **8**. doi:10.1371/journal.pone.0053880
Haggerty, Kretzmer, Riemenschneider, Kumar, Mattei, Bailly, Gottfreund, Giesselmann, Weigert, Brändl, Giehr, Buschow, Galonska, von Meyenn, Pappalardi, McCabe, Wittler, Giesecke-Thiel, Mielke, Meierhofer, Timmermann, Müller, Walter, Meissner. 2021. Dnmt1 has de novo activity targeted to transposable elements. *Nat Struct Mol Biol*. doi:10.1038/s41594-021-00603-8

Hai, Hou, Liu, Liu, Yang, Li, He. 2014. The roles and regulation of Sertoli cells in fate determinations of spermatogonial stem cells and spermatogenesis. *Semin Cell Dev Biol* **29**:66–75. doi:10.1016/j.semcdb.2014.04.007

Hall. 2018. Germ layers, the neural crest and emergent organization in development and evolution. *Genesis* **56**:1–9. doi:10.1002/dvg.23103

Hall. 2000. The neural crest as a fourth germ layer and vertebrates as quadroblastic not triploblastic. *Evol Dev* **2**:3–5. doi:10.1046/j.1525-142X.2000.00032.x

Hamatani, Carter, Sharov, Ko. 2004. Dynamics of global gene expression changes during mouse preimplantation development. *Dev Cell* **6**:117–131. doi:10.1016/S1534-5807(03)00373-3

Hamilton, Mosesson, Monteiro, Emdal, Knudsen, Francavilla, Barkai, Olsen, Brickman. 2019. Dynamic lineage priming is driven via direct enhancer regulation by ERK. *Nature* **575**:355–360. doi:10.1038/s41586-019-1732-z

Hammoud, Nix, Zhang, Purwar, Carrell, Cairns. 2009. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**:473–478. doi:10.1038/nature08162

Han, Ren, Li, Li, Ge, Wang, Miao, Guo, Moley, Wang, Shu. 2018. Embryonic defects induced by maternal obesity in mice derive from Stella insufficiency in oocytes. *Nat Genet*. doi:10.1038/s41588-018-0055-6

Han, Tapia, Joo, Greber, Araúzo-Bravo, Bernemann, Ko, Wu, Stehling, Do, Schöler. 2010. Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. *Cell* **143**:617–627. doi:10.1016/j.cell.2010.10.015

Hancock, Wamaitha, Peretz, Clark. 2021. Mammalian primordial germ cell specification. *Development* **148**:dev189217. doi:10.1242/dev.189217

Hanna, Demond, Kelsey. 2018. Epigenetic regulation in development: is the mouse a good model for the human? *Hum Reprod Update* 1–21. doi:10.1093/humupd/dmy021

Hansen. 2002. Conformational dynamics of the chromatin fiber in solution: Determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* **31**:361–392. doi:10.1146/annurev.biophys.31.101101.140858

Hargan-Calvopina, Taylor, Cook, Hu, Lee, Yen, Chiang, Chen, Clark. 2016. Stage-Specific Demethylation in Primordial Germ Cells Safeguards against Precocious Differentiation. *Dev Cell* **39**:75–86. doi:10.1016/j.devcel.2016.07.019

Hashimoto, Liu, Upadhyay, Chang, Howerton, Vertino, Zhang, Cheng. 2012. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* **40**:4841–4849. doi:10.1093/nar/gks155

Haston, Tung, Reijo Pera. 2009. Dazl functions in maintenance of pluripotency and genetic and epigenetic programs of differentiation in mouse primordial germ cells in vivo and in vitro. *PLoS One* **4**. doi:10.1371/journal.pone.0005654

Hawkins, Hon, Lee, Ngo, Lister, Pelizzola, Edsall, Kuan, Luu, Klugman, Antosiewicz-Bourget, Ye, Espinoza, Agarwahl, Shen, Ruotti, Wang, Stewart, Thomson, Ecker, Ren. 2010. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**:479–491. doi:10.1016/j.stem.2010.03.018

Haws, Yu, Ye, Wille, Nguyen, Krautkramer, Tomasiewicz, Yang, Miller, Liu, Igarashi, Sridharan, Tu, Cryns, Lamming, Denu. 2020. Methyl-Metabolite Depletion Elicits Adaptive Responses to Support Heterochromatin Stability and Epigenetic Persistence. *Mol Cell* 1–14. doi:10.1016/j.molcel.2020.03.004

Hayashi, Ohta, Kurimoto, Aramaki, Saitou. 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**:519–532. doi:10.1016/j.cell.2011.06.052

He, Li, Li, Liu, Wang, Tang, Ding, Jia, Chen, Li, Sun, Li, Dai, Song, Zhang, He, Xu. 2011. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science (80- )* **333**:1303–1307. doi:10.1126/science.1210944

Heard, Martienssen. 2014. Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell* **157**:95–109. doi:10.1016/j.cell.2014.02.045

Héberlé, Bardet. 2019. Sensitivity of transcription factors to DNA methylation. *Essays Biochem* **63**:727–741. doi:10.1042/EBC20190033

Heinz, Benner, Spann, Bertolino, Lin, Laslo, Cheng, Murre, Singh, Glass. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell* **38**:576–589. doi:10.1016/j.molcel.2010.05.004

Hess, de Franca. 2009. Spermatogenesis and Cycle of the Seminiferous Epithelium. pp. 1–15. doi:10.1007/978-0-387-09597-4_1

Hillman, Sherman, Graham. 1972. The effect of spatial arrangement on cell determination during mouse development. *J Embryol Exp Morphol* **28**:263–278. doi:10.1242/dev.28.2.263
Holliday. 1994. Epigenetics: An overview. *Dev Genet* **15**:453–457. doi:10.1002/dvg.1020150602
Holliday, Pugh. 1975. DNA modification mechanisms and gene activity during development. *Science (80- )* **187**:226–232. doi:10.1126/science.1111098
Hon, Song, Du, Jin, Selvaraj, Lee, Yen, Ye, Mao, Wang, Kuan, Edsall, Zhao, Xu, He, Ren. 2014. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell* **56**:286–297. doi:10.1016/j.molcel.2014.08.026
Hong, Schroth, Matthews, Yau, Bradbury. 1993. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J Biol Chem* **268**:305–314. doi:10.1016/s0021-9258(18)54150-8
Howell, Bestor, Ding, Latham, Mertineit, Trasler, Chaillet. 2001. Genomic Imprinting Disrupted by a Maternal Effect Mutation in the Dnmt1 Gene. *Cell* **104**:829–838. doi:10.1016/S0092-8674(01)00280-X
Hu, Gao, Cao, Morgan, Mas, Smith, Volk, Bartom, Crispino, Di Croce, Shilatifard. 2017. Not All H3K4 Methylations Are Created Equal: Mll2/COMPASS Dependency in Primordial Germ Cell Specification. *Mol Cell* **65**:460-475.e6. doi:10.1016/j.molcel.2017.01.013
Huang, Wang, Vazquez-Ferrer, Theofel, Requena, Hanna, Kelsey, Hajkova. 2021. Sex-specific chromatin remodelling safeguards transcription in germ cells. *Nature* **600**. doi:10.1038/s41586-021-04208-5
Huff, Zilberman. 2014. Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. *Cell* **156**:1286–1297. doi:10.1016/j.cell.2014.01.029
Hughes, Kelley, Klose. 2020. Understanding the interplay between CpG island-associated gene promoters and H3K4 methylation. *Biochim Biophys Acta - Gene Regul Mech* **1863**:194567. doi:10.1016/j.bbagrm.2020.194567
Hutnick, Huang, Loo, Ma, Fan. 2010. Repression of retrotransposal elements in mouse embryonic stem cells is primarily mediated by a DNA methylation-independent mechanism. *J Biol Chem* **285**:21082–21091. doi:10.1074/jbc.M110.125674
Hyun, Wilkerson, Johnston. 2016. Embryology policy: Revisit the 14-day rule. *Nature*. doi:10.1038/533169a
Illingworth, Gruenewald-Schneider, Webb, Kerr, James, Turner, Smith, Harrison, Andrews, Bird. 2010. Orphan CpG Islands Identify numerous conserved promoters in the mammalian genome. *PLoS Genet* **6**. doi:10.1371/journal.pgen.1001134
Inoue, Jiang, Lu, Suzuki, Zhang. 2017. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature* **547**:419–424. doi:10.1038/nature23262
International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**:860–921. doi:10.1038/35057062
Irie, Kuratani. 2011. Comparative transcriptome analysis reveals vertebrate phylotypic period during organogenesis. *Nat Commun* **2**. doi:10.1038/ncomms1248
Irie, Weinberger, Tang, Kobayashi, Viukov, Manor, Dietmann, Hanna, Surani. 2015. SOX17 is a critical specifier of human primordial germ cell fate. *Cell* **160**:253–268. doi:10.1016/j.cell.2014.12.013
Ishikura, Ohta, Sato, Murase, Yabuta, Kojima, Yamashiro, Nakamura, Yamamoto, Ogawa, Saitou. 2021. In vitro reconstitution of the whole male germ-cell development from mouse pluripotent stem cells. *Cell Stem Cell* 1–13. doi:10.1016/j.stem.2021.08.005
Ishikura, Yabuta, Ohta, Hayashi, Nakamura, Okamoto, Yamamoto, Kurimoto, Shirane, Sasaki, Saitou. 2016. In Vitro Derivation and Propagation of Spermatogonial Stem Cell Activity from Mouse Pluripotent Stem Cells. *Cell Rep* **17**:2789–2804. doi:10.1016/j.celrep.2016.11.026
Ito, Kubiura-Ichimaru, Murakami, Bogutz, Lefebvre, Suetake, Tajima, Tada. 2022. DNMT1 regulates the timing of DNA methylation by DNMT3 in an enzymatic activity-dependent manner in mouse embryonic stem cells. *PLoS One* **17**:1–21. doi:10.1371/journal.pone.0262277
Ito, Shen, Dai, Wu, Collins, Swenberg, He, Zhang. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science (80- )* **333**:1300–1303. doi:10.1126/science.1210597
Iyer, Abhiman, Aravind. 2011. Natural history of eukaryotic DNA methylation systems, 1st ed, Progress in Molecular Biology and Translational Science. Elsevier Inc. doi:10.1016/B978-0-12-387685-0.00002-0
Izzo, Lee, Poran, Chaligne, Gaiti, Gross, Murali, Deochand, Ang, Jones, Nam, Kim, Kothen-Hill,

Schulman, Ki, Lhoumaud, Skok, Viny, Levine, Kenigsberg, Abdel-Wahab, Landau. 2020. DNA methylation disruption reshapes the hematopoietic differentiation landscape. *Nat Genet*. doi:10.1038/s41588-020-0595-4

Jachowicz, Bing, Pontabry, Bošković, Rando, Torres-Padilla. 2017. LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat Genet* **49**:1502–1510. doi:10.1038/ng.3945

Jackson-Grusby, Beard, Possemato, Tudor, Fambrough, Csankovszki, Dausman, Lee, Wilson, Lander, Jaenisch. 2001. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat Genet* **27**:31–39. doi:10.1038/83730

Jacob, Monod. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**:318–356. doi:10.1016/S0022-2836(61)80072-7

Jamai, Imoberdorf, Strubin. 2007. Continuous Histone H2B and Transcription-Dependent Histone H3 Exchange in Yeast Cells outside of Replication. *Mol Cell* **25**:345–355. doi:10.1016/j.molcel.2007.01.019

Janssen, Lorincz. 2021. Interplay between chromatin marks in development and disease. *Nat Rev Genet* **0123456789**:1–17. doi:10.1038/s41576-021-00416-x

Jedrusik, Cox, Wicher, Glover, Zernicka-Goetz. 2015. Maternal-zygotic knockout reveals a critical role of Cdx2 in the morula to blastocyst transition. *Dev Biol* **398**:147–152. doi:10.1016/j.ydbio.2014.12.004

Jenkins, Aston, Pflueger, Cairns, Carrell. 2014. Age-Associated Sperm DNA Methylation Alterations: Possible Implications in Offspring Disease Susceptibility. *PLoS Genet* **10**. doi:10.1371/journal.pgen.1004458

Jenuwein, Allis. 2001. Translating the histone code. *Science (80- )* **293**:1074–1080. doi:10.1126/science.1063127

Jeppesen, Turner. 1993. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**:281–289. doi:10.1016/0092-8674(93)90419-Q

Jia, Jurkowska, Zhang, Jeltsch, Cheng. 2007. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449**:248–251. doi:10.1038/nature06146

John B. Gurdon. 1962. The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles. *J Embryol Exp Morphol* **10**.

Johnson, Ziomek. 1981. The foundation of two distinct cell lineages within the mouse morula. *Cell* **24**:71–80. doi:10.1016/0092-8674(81)90502-X

Johnston, Jelinsky, Bang, Dicandeloro, Wilson, Kopf, Turner. 2005. The mouse epididymal transcriptome: Transcriptional profiling of segmental gene expression in the epididymis. *Biol Reprod* **73**:404–413. doi:10.1095/biolreprod.105.039719

Jourdin. 2008. Gottlieb Haberlandt (1854-1945) et la culture in vitro de tissus et de cellules végétales . *Bull d'histoire d'épistémologie des Sci la vie* **Volume 15**:197. doi:10.3917/bhesv.152.0197

Kalkan, Bornelöv, Mulas, Diamanti, Lohoff, Ralser, Middelkamp, Lombard, Nichols, Smith. 2019. Complementary Activity of ETV5, RBPJ, and TCF3 Drives Formative Transition from Naive Pluripotency. *Cell Stem Cell* **24**:785-801.e7. doi:10.1016/j.stem.2019.03.017

Kalkan, Olova, Roode, Mulas, Lee, Nett, Marks, Walker, Stunnenberg, Lilley, Nichols, Reik, Bertone, Smith. 2017. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**:1221–1234. doi:10.1242/dev.142711

Kane, Kimmel. 1993. The zebrafish midblastula transition. *Development* **119**:447–456. doi:10.1242/dev.119.2.447

Kang, Garg, Hadjantonakis. 2017. Lineage Establishment and Progression within the Inner Cell Mass of the Mouse Blastocyst Requires FGFR1 and FGFR2. *Dev Cell* **41**:496-510.e5. doi:10.1016/j.devcel.2017.05.003

Kang, Piliszek, Artus, Hadjantonakis. 2013. FGF4 is required for lineage restriction and salt-and-pepper distribution of primitive endoderm factors but not their initial expression in the mouse. *Dev* **140**:267–279. doi:10.1242/dev.084996

Kapitonov, Jurka. 2005. RAG1 core and V(D)J recombination signal sequences were derived from Transib transposons. *PLoS Biol* **3**:0998–1011. doi:10.1371/journal.pbio.0030181

Karimi, Goyal, Maksakova, Bilenky, Leung, Tang, Shinkai, Mager, Jones, Hirst, Lorincz. 2011. DNA Methylation and SETDB1/H3K9me3 Regulate Predominantly Distinct Sets of Genes, Retroelements, and Chimeric Transcripts in mESCs. *Cell Stem Cell* **8**:676–687.

doi:10.1016/j.stem.2011.04.004
Karimi, Johansson, Stach, Corcoran, Grandér, Schalling, Bakalkin, Lyko, Larsson, Ekström. 2006. LUMA (LUminometric Methylation Assay)—A high throughput method to the analysis of genomic DNA methylation. *Exp Cell Res* **312**:1989–1995. doi:10.1016/j.yexcr.2006.03.006
Katayama, Ellersieck, Roberts. 2010. Development of monozygotic twin mouse embryos from the time of blastomere separation at the two-cell stage to blastocyst. *Biol Reprod* **82**:1237–1247. doi:10.1095/biolreprod.109.082982
Kigami, Minami, Takayama, Imai. 2003. MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol Reprod* **68**:651–654. doi:10.1095/biolreprod.102.007906
Kim, Wu, Rahme, Battaglia, Dixit, Gaskell, Chen, Pinello, Bernstein. 2020. Parallel Single-Cell RNA-Seq and Genetic Recording Reveals Lineage Decisions in Developing Embryoid Bodies. *Cell Rep* **33**:108222. doi:10.1016/j.celrep.2020.108222
Kimura, Yasugi, Stern, Fukuda. 2006. Fate and plasticity of the endoderm in the early chick embryo. *Dev Biol* **289**:283–295. doi:10.1016/j.ydbio.2005.09.009
King, Huang, Rubbi, Liu, Wang, Wang, Pellegrini, Fan. 2016. Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *Cell Rep* **17**:289–302. doi:10.1016/j.celrep.2016.08.083
Kinoshita, Barber, Mansfield, Cui, Spindlow, Stirparo, Dietmann, Nichols, Smith. 2020. Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency. *Cell Stem Cell* 2020.09.04.283218. doi:10.1016/j.stem.2020.11.005
Kinoshita, Li, Barber, Mansfield, Dietmann, Smith. 2021. Disabling de novo DNA methylation in embryonic stem cells allows an illegitimate fate trajectory. *Proc Natl Acad Sci U S A* **118**:1–11. doi:10.1073/pnas.2109475118
Kinoshita, Smith. 2018. Pluripotency Deconstructed. *Dev Growth Differ*. doi:10.1111/dgd.12419
Klughammer, Romanovskaia, Nemc, Posautz, Seid, Linda, Keinath, Sebastian, Ramos, Kosack, Evankow, Kirchberger, Ergüner, Datlinger, Fortelny, Schmidl, Farlik, Skjærven, Bergthaler, Liedvogel, Thaller, Pamela, Hermann, Distel, Distel, Kübber-heiss, Bock. 2022. Comparative analysis of genome-scale , base-resolution DNA methylation profiles across 580 animal species. *Biorxv*.
Kobayakawa, Kubota. 1981. Temporal pattern of cleavage and the onset of gastrulation in amphibian embryos developed from eggs with the reduced cytoplasm. *J Embryol Exp Morphol* **Vol. 62**:83–94. doi:10.1242/dev.62.1.83
Kobayashi, Sakurai, Imai, Takahashi, Fukuda, Yayoi, Sato, Nakabayashi, Hata, Sotomaru, Suzuki, Kono. 2012. Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish Oocyte-specific heritable marks. *PLoS Genet* **8**. doi:10.1371/journal.pgen.1002440
Koike, Sakaki, Amano, Kurosawa. 2007. Characterization of embryoid bodies of mouse embryonic stem cells formed under various culture conditions and estimation of differentiation status of such bodies. *J Biosci Bioeng* **104**:294–299. doi:10.1263/jbb.104.294
Kong, Banaszynski, Geng, Zhang, Zhang, Zhang, O'Neill, Yan, Liu, Shido, Palermo, David Allis, Rafii, Rosenwaks, Wen. 2018. Histone variant H3.3–mediated chromatin remodeling is essential for paternal genome activation in mouse preimplantation embryos. *J Biol Chem* **293**:3829–3838. doi:10.1074/jbc.RA117.001150
Korotkevich, Niwayama, Courtois, Friese, Berger, Buchholz, Hiiragi. 2017. The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse Embryo. *Dev Cell* **40**:235-247.e7. doi:10.1016/j.devcel.2017.01.006
Kossel. 1884. Ueber einen peptonartigen Bestandtheil des Zellkerns. *bchm* **8**:511–515. doi:10.1515/bchm1.1884.8.6.511
Kreibich, Kleinendorst, Barzaghi, Kaspar, Krebs. 2022. Single molecule multi-omics reveals context-dependent regulation of enhancers by DNA methylation. *Biorxv* 1–23.
Krietenstein, Rando. 2020. Mesoscale organization of the chromatin fiber. *Curr Opin Genet Dev* **61**:32–36. doi:10.1016/j.gde.2020.02.022
Krishnakumar, Chen, Pantovich, Danial, Parchem, Labosky, Blelloch. 2016. FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and Repressing Enhancer Activity. *Cell Stem Cell* **18**:104–117. doi:10.1016/j.stem.2015.10.003
Krogan, Kim, Tong, Golshani, Cagney, Canadien, Richards, Beattie, Emili, Boone, Buratowski, Greenblatt, Shilatifard. 2003. Methylation of Histone H3 by Set2 in Saccharomyces cerevisiae Is Linked to Transcriptional Elongation by RNA Polymerase II Downloaded from http://mcb.asm.org/ on September 18 , 2012 by IMPERIAL COLLEGE LONDON Methylation of Histone H3 by Set2 in

Saccharom **23**:4207–4218. doi:10.1128/MCB.23.12.4207
Krupa, Mazur, Szczepańska, Filimonow, Maleszewski, Suwińska. 2014. Allocation of inner cells to epiblast vs primitive endoderm in the mouse embryo is biased but not determined by the round of asymmetric divisions (8→16- and 16→32-cells). *Dev Biol* **385**:136–148. doi:10.1016/j.ydbio.2013.09.008
Ku, Koche, Rheinbay, Mendenhall, Endoh, Mikkelsen, Presser, Nusbaum, Xie, Chi, Adli, Kasif, Ptaszek, Cowan, Lander, Koseki, Bernstein. 2008. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**. doi:10.1371/journal.pgen.1000242
Kubinyecz, Santos, Drage, Reik, Eckersley-Maslin. 2021. Maternal Dppa2 and Dppa4 are dispensable for zygotic genome activation but important for offspring survival. *Development* **148**. doi:10.1242/dev.200191
Kubo, Shinozaki, Shannon, Kouskoff, Kennedy, Woo, Fehling, Keller. 2004. Development of definitive endoderm from embryonic stem cells in culture. *Development* **131**:1651–1662. doi:10.1242/dev.01044
Kumar, Cinghu, Oldfield, Yang, Jothi. 2021. Decoding the function of bivalent chromatin in development and cancer. *Genome Res* gr.275736.121. doi:10.1101/gr.275736.121
Kunath, Saba-El-Leil, Almousailleakh, Wray, Meloche, Smith. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**:2895–2902. doi:10.1242/dev.02880
Kurek, Neagu, Tastemel, Tüysüz, Lehmann, Van De Werken, Philipsen, Van Der Linden, Maas, Van Ijcken, Drukker, Ten Berge. 2015. Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports* **4**:114–128. doi:10.1016/j.stemcr.2014.11.007
Kurimoto, Yabuta, Hayashi, Ohta, Kiyonari, Mitani, Moritoki, Kohri, Kimura, Yamamoto, Katou, Shirahige, Saitou. 2015. Quantitative dynamics of chromatin remodeling during germ cell specification from mouse embryonic stem cells. *Cell Stem Cell* **16**:517–532. doi:10.1016/j.stem.2015.03.002
Kurimoto, Yamaji, Seki, Saitou. 2008. Specification of the germ cell lineage in mice: A process orchestrated by the PR-domain proteins, Blimp1 and Prdm14. *Cell Cycle* **7**:3514–3518. doi:10.4161/cc.7.22.6979
Kuroda, Fuentealba, Ikeda, Reversade, De Robertis. 2005. Default neural induction: Neuralization of dissociated Xenopus cells is mediated by Ras/MAPK activation. *Genes Dev* **19**:1022–1027. doi:10.1101/gad.1306605
Lackner, Sehlke, Garmhausen, Giuseppe Stirparo, Huth, Titz-Teixeira, van der Lelij, Ramesmayer, Thomas, Ralser, Santini, Galimberti, Sarov, Stewart, Smith, Beyer, Leeb. 2021. Cooperative genetic networks drive embryonic stem cell transition from naïve to formative pluripotency. *EMBO J* e105776. doi:10.15252/embj.2020105776
Lambrot, Chan, Shao, Aarabi, Kwan, Bourque, Moskovtsev, Librach, Trasler, Dumeaux, Kimmins. 2021. Whole-genome sequencing of H3K4me3 and DNA methylation in human sperm reveals regions of overlap linked to fertility and development. *Cell Rep* **36**:109418. doi:10.1016/j.celrep.2021.109418
Lambrot, Xu, Saint-Phar, Chountalos, Cohen, Paquet, Suderman, Hallett, Kimmins. 2013. Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. *Nat Commun* **4**. doi:10.1038/ncomms3889
Larson, Elnatan, Keenen, Trnka, Johnston, Burlingame, Agard, Redding, Narlikar. 2017. Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature* **547**:236–240. doi:10.1038/nature22822
Law, Jacobsen. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. doi:10.1038/nrg2719
Lawson, Dunn, Roelen, Zeinstra, Davis, Wright, Korving, Hogan. 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev* **13**:424–436. doi:10.1101/gad.13.4.424
Leaky, Weixiong, Kuhnert, Stuhlmann. 1999. Use of developmental marker genes to define temporal and spatial patterns of differentiation during embryoid body formation. *J Exp Zool* **284**:67–81. doi:10.1002/(SICI)1097-010X(19990615)284:1<67::AID-JEZ10>3.0.CO;2-O
Lee, Bonneau, Takacs, Bazzini, Divito, Fleming, Giraldez. 2013. Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature* **503**:360–364. doi:10.1038/nature12632

Lee, Hayes, Pruss, Wolffe. 1993. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**:73–84. doi:10.1016/0092-8674(93)90051-Q

Lee, Iskow, Yang, Gokcumen, Haseley, Luquette, Lohr, Harris, Ding, Wilson, Wheeler, Gibbs, Kucherlapati, Lee, Kharchenko, Park. 2012. Landscape of Somatic Retrotransposition in Human Cancers. *Science (80- )* **337**:967–971. doi:10.1126/science.1222077

Lei, Oh, Okano, Jüttermann, Goss, Jaenisch, Li. 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122**:3195–205.

Leitch, Okamura, Durcova-Hills, Stewart, Gardner, Matsui, Papaioannou. 2014. On the fate of primordial germ cells injected into early mouse embryos. *Dev Biol* **385**:155–159. doi:10.1016/j.ydbio.2013.11.014

Leitch, Smith. 2013. The mammalian germline as a pluripotency cycle. *Development* **140**:2495–2501. doi:10.1242/dev.091603

Lewis. 1978. A Gene Complex Controlling Segmentation in DrosophilaNature. Dordrecht: Springer Netherlands. pp. 229–242. doi:10.1007/978-1-4020-6345-9_10

Li, Amaral, Cheung, Bergmann, Kinoshita, Kalkan, Ralser, Robson, Vonmeyenn, Paramor, Yang, Chen, Nichols, Spector, Kouzarides, He, Smith. 2017. A lncRNA fine tunes the dynamics of a cell state transition involving lin28, let-7 and de novo DNA methylation. *Elife* **6**:1–24. doi:10.7554/eLife.23468

Li, Bestor, Jaenisch. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**:915–926. doi:10.1016/0092-8674(92)90611-F

Li, Cai, Cai, Wang, Zhang, Petersen, Yang, Xu. 2011. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* **118**:4509–4518. doi:10.1182/blood-2010-12-325241

Li, Ito, Zhou, Youngson, Zuo, Leder, Ferguson-Smith. 2008. A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. *Dev Cell* **15**:547–557. doi:10.1016/j.devcel.2008.08.014

Liao, Karnik, Gu, Ziller, Clement, Tsankov, Akopian, Gifford, Donaghey, Galonska, Pop, Reyon, Tsai, Mallard, Joung, Rinn, Gnirke, Meissner. 2015. Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet* **47**:469–478. doi:10.1038/ng.3258

Libby, Joy, Elder, Bulger, Krakora, Gaylord, Mendoza-Camacho, Butts, McDevitt. 2021. Axial elongation of caudalized human organoids mimics aspects of neural tube development. *Dev* **148**. doi:10.1242/dev.198275

Lienert, Mohn, Tiwari, Baubec, Roloff, Gaidatzis, Stadler, Schübeler. 2011. Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells. *PLoS Genet* **7**. doi:10.1371/journal.pgen.1002090

Ling, Harkness, Schultz, Fisher-Adams, Grunstein. 1996. Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: Redundant and position-independent functions in assembly but not in gene regulation. *Genes Dev* **10**:686–699. doi:10.1101/gad.10.6.686

Linker, Stern. 2004. Neural induction requires BMP inhibition only as a late step, and involves signals other than FGF and Wnt antagonists. *Development* **131**:5671–5681. doi:10.1242/dev.01445

Liu, Gao, Li, Zhao, Zhang, Li, Koseki, Wong. 2013. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nat Commun* **4**. doi:10.1038/ncomms2562

Liu, Wang, Liu, Li, Li, Kou, Chen, Zhao, Gao, Wang, Zhang, Gao, Gao. 2016. Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature* **537**:558–562. doi:10.1038/nature19362

Lu, Zhang, Wang, Wang, Wang, Xu, Xiang, Chen, Kong, Xia, Lin, Ma, Liu, Wang, Ni, Li, Guo, Xie. 2021. Evolutionary epigenomic analyses in mammalian early embryos reveal species-specific innovations and conserved principles of imprinting. *Sci Adv* **7**. doi:10.1126/sciadv.abi6178

Luger, Dechassa, Tremethick. 2012. New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nat Rev Mol Cell Biol* **13**:436–447. doi:10.1038/nrm3382

Luger, Mäder, Richmond, Sargent, Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**:251–260. doi:10.1038/38444

Luger, Richmond. 1998. The histone tails of the nucleosome. *Curr Opin Genet Dev* **8**:140–146. doi:10.1016/S0959-437X(98)80134-2

Luo, van der Veer, Sun, Bartoccetti, Boretto, Vankelecom, Khoueiry, Koh. 2020. Coordination of germ layer lineage choice by TET1 during primed pluripotency. *Genes Dev* 1–21.

doi:10.1101/gad.329474.119
Lyko. 2018. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. *Nat Rev Genet*. doi:10.1038/nrg.2017.80
Lynch, Smith, De Gobbi, Flenley, Hughes, Vernimmen, Ayyub, Sharpe, Sloane-Stanley, Sutherland, Meek, Burdon, Gibbons, Garrick, Higgs. 2012. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *EMBO J* **31**:317–329. doi:10.1038/emboj.2011.399
Lyon. 1961. Gene Action in the X-chromosom (Mus musculus L.). *Nature* **190**:372–373.
Maatouk, Kellam, Mann, Lei, En, Bartolomei, Resnick. 2006. DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133**:3411–3418. doi:10.1242/dev.02500
Macfarlan, Gifford, Driscoll, Lettieri, Rowe, Bonanomi, Firth, Singer, Trono, Pfaff. 2012. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**:57–63. doi:10.1038/nature11244
Maenohara, Unoki, Toh, Ohishi, Sharif, Koseki, Sasaki. 2017. Role of UHRF1 in de novo DNA methylation in oocytes and maintenance methylation in preimplantation embryos. *PLoS Genet* **13**. doi:10.1371/journal.pgen.1007042
Maeshima, Imai, Tamura, Nozaki. 2014. Chromatin as dynamic 10-nm fibers. *Chromosoma* **123**:225–237. doi:10.1007/s00412-014-0460-2
Malaguti, Migueles, Blin, Lin, Lowell. 2019. Id1 Stabilizes Epiblast Identity by Sensing Delays in Nodal Activation and Adjusting the Timing of Differentiation. *Dev Cell* **50**:1–16. doi:10.1016/j.devcel.2019.05.032
Malik, Henikoff. 2003. Phylogenomics of the nucleosome. *Nat Struct Biol* **10**:882–891. doi:10.1038/nsb996
Margueron, Li, Sarma, Blais, Zavadil, Woodcock, Dynlacht, Reinberg. 2008. Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol Cell* **32**:503–518. doi:10.1016/j.molcel.2008.11.004
Marinus, Casadesus. 2009. Roles of DNA adenine methylation in host-pathogen interactions: Mismatch repair, transcriptional regulation, and more. *FEMS Microbiol Rev* **33**:488–503. doi:10.1111/j.1574-6976.2008.00159.x
Martin, Evans. 1975. Differentiation of clonal teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc Natl Acad Sci U S A* **72**:1441–1445. doi:10.1073/pnas.72.4.1441
Martindale, Pang, Finnerty. 2004. Investigating the origins of triplosblasty: “Mesodermal” gene expression in a diploblastic animal, the sea anemone Nematostella vectensis (phylum, Cnidaria; class, Anthozoa). *Development* **131**:2463–2474. doi:10.1242/dev.01119
Masui, Wang. 1998. Cell cycle transition Xenopus laevis. *Biol Cell* **90**:537–548.
Matos, Publicover, Castro, Esteves, Fardilha. 2021. Brain and testis: More alike than previously thought? *Open Biol* **11**. doi:10.1098/rsob.200322
Matsui, Leung, Miyashita, Maksakova, Miyachi, Kimura, Tachibana, Lorincz, Shinkai. 2010. Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**:927–931. doi:10.1038/nature08858
Matthews, Wagner, Warmflash. 2021. Stem cell-based models of embryos: The need for improved naming conventions. *Stem Cell Reports* **16**:1014–1020. doi:10.1016/j.stemcr.2021.02.018
Maunakea, Chepelev, Cui, Zhao. 2013. Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res* **23**:1256–1269. doi:10.1038/cr.2013.110
Mayer, Stadler, Rittirsch, Hess, Lukonin, Winzi, Smith, Buchholz, Betschinger. 2020. Zfp281 orchestrates interconversion of pluripotent states by engaging Ehmt1 and Zic2. *EMBO J* **39**:1–22. doi:10.15252/embj.2019102591
Mazid, Ward, Luo, Liu, Li, Lai, Wu, Li, Jia, Jiang, Liu, Fu, Yang, Ibañez, Lai, Wei, An, Guo, Yuan, Deng, Wang, Liu, Gao, Wang, Zaman, Qin, Wu, Maxwell, Xu, Liu, Li, Esteban. 2022. Rolling back human pluripotent stem cells to an eight-cell embryo-like stage. *Nature* **605**:315–324. doi:10.1038/s41586-022-04625-0
McLaughlin, Flyamer, Thomson, Mjoseng, Shukla, Williamson, Grimes, Illingworth, Adams, Pennings, Meehan, Bickmore. 2019. DNA Methylation Directs Polycomb-Dependent 3D Genome Re-organization in Naive Pluripotency. *Cell Rep* **29**:1974-1985.e6. doi:10.1016/j.celrep.2019.10.031
Meissner, Mikkelsen, Gu, Wernig, Hanna, Sivachenko, Zhang, Bernstein, Nusbaum, Jaffe, Gnirke, Jaenisch, Lander. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated

cells. *Nature* **454**:766–770. doi:10.1038/nature07107
Meshorer, Yellajoshula, George, Scambler, Brown, Misteli. 2006. Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell* **10**:105–116. doi:10.1016/j.devcel.2005.10.017
Messerschmidt, de Vries, Ito, Solter, Ferguson-Smith, Knowles. 2012. Trim28 Is Required for Epigenetic Stability During Mouse Oocyte to Embryo Transition. *Science (80- )* **335**:1499–1502. doi:10.1126/science.1216154
Messerschmidt, Knowles, Solter. 2014. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev* **28**:812–828. doi:10.1101/gad.234294.113
Messmer, von Meyenn, Savino, Santos, Mohammed, Lun, Marioni, Reik. 2019. Transcriptional Heterogeneity in Naive and Primed Human Pluripotent Stem Cells at Single-Cell Resolution. *Cell Rep* **26**:815-824.e4. doi:10.1016/j.celrep.2018.12.099
Mihajlovic, Thamodaran, Bruce. 2015. The first two cell-fate decisions of preimplantation mouse embryo development are not functionally independent. *Sci Rep* **5**:1–16. doi:10.1038/srep15034
Miller, Brinkworth, Iles. 2010. Paternal DNA packaging in spermatozoa: More than the sum of its parts? DNA, histones, protamines and epigenetics. *Reproduction* **139**:287–301. doi:10.1530/REP-09-0281
Milne, Kim, Wang, Stadler, Basrur, Whitcomb, Wang, Ruthenburg, Elenitoba-Johnson, Roeder, Allis. 2010. Multiple Interactions Recruit MLL1 and MLL1 Fusion Proteins to the HOXA9 Locus in Leukemogenesis. *Mol Cell* **38**:853–863. doi:10.1016/j.molcel.2010.05.011
Minoux, Holwerda, Vitobello, Kitazawa, Kohler, Stadler, Rijli. 2017. Gene bivalency at Polycomb domains regulates cranial neural crest positional identity. *Science (80- )* **355**. doi:10.1126/science.aal2913
Mira-Bontenbal, Tan, Gontan, Goossens, Boers, Boers, Dupont, van Royen, IJcken, French, Bedalov, Gribnau. 2022. Genetic and epigenetic determinants of reactivation of Mecp2 and the inactive X chromosome in neural stem cells. *Stem Cell Reports* **5**:988–995. doi:10.1016/j.stemcr.2022.01.008
Mochizuki, Hayashi, Sekinaka, Otsuka, Ito-Matsuoka, Kobayashi, Oki, Takehara, Kono, Osumi, Matsui. 2018. Repression of Somatic Genes by Selective Recruitment of HDAC3 by BLIMP1 Is Essential for Mouse Primordial Germ Cell Fate Determination. *Cell Rep* **24**:2682-2693.e6. doi:10.1016/j.celrep.2018.07.108
Mochizuki, Sharif, Shirane, Uranishi, Bogutz, Janssen, Suzuki, Okuda, Koseki, Lorincz. 2021. Repression of germline genes by PRC1.6 and SETDB1 in the early embryo precedes DNA methylation-mediated silencing. *Nat Commun* **12**:1–8. doi:10.1038/s41467-021-27345-x
Mohammed, Hernando-Herraez, Savino, Scialdone, Macaulay, Mulas, Chandra, Voet, Dean, Nichols, Marioni, Reik. 2017. Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep* **20**:1215–1228. doi:10.1016/j.celrep.2017.07.009
Mohn, Weber, Rebhan, Roloff, Richter, Stadler, Bibel, Schübeler. 2008. Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Mol Cell* **30**:755–766. doi:10.1016/j.molcel.2008.05.007
Molaro, Malik, Bourc'his, Satta. 2020. Dynamic Evolution of de Novo DNA Methyltransferases in Rodent and Primate Genomes. *Mol Biol Evol* **37**:1882–1892. doi:10.1093/molbev/msaa044
Morgani, Nichols, Hadjantonakis. 2017. The many faces of Pluripotency: In vitro adaptations of a continuum of in vivo states. *BMC Dev Biol* **17**:10–12. doi:10.1186/s12861-017-0150-4
Moris, Anlas, van den Brink, Alemany, Schröder, Ghimire, Balayo, van Oudenaarden, Martinez Arias. 2020. An in vitro model of early anteroposterior organization during human development. *Nature* **582**:410–415. doi:10.1038/s41586-020-2383-9
Morris, Graham, Jedrusik, Zernicka-Goetz. 2013. The differential response to Fgf signalling in cells internalized at different times influences lineage segregation in preimplantation mouse embryos. *Open Biol* **3**. doi:10.1098/rsob.130104
Motosugi, Bauer, Polanski, Solter, Hiiragi. 2005. Polarity of the mouse embryo is established at blastocyst and is not prepatterned. *Genes Dev* **19**:1081–1092. doi:10.1101/gad.1304805
Mouse genome sequencing consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**:520–562. doi:10.1038/nature01262
Mulas, Kalkan, Smith. 2017. NODAL Secures Pluripotency upon Embryonic Stem Cell Progression from

the Ground State. *Stem Cell Reports* **9**:77–91. doi:10.1016/j.stemcr.2017.05.033
Mulas, Kalkan, von Meyenn, Leitch, Nichols, Smith. 2019. Defined conditions for propagation and manipulation of mouse embryonic stem cells. *Development* **146**:dev173146. doi:10.1242/dev.173146
Mulholland, Nishiyama, Ryan, Nakamura, Yiğit, Glück, Trummer, Qin, Bartoschek, Traube, Parsa, Ugur, Modic, Acharya, Stolz, Ziegenhain, Wierer, Enard, Carell, Lamb, Takeda, Nakanashi, Bultmann, Leonhardt. 2020a. Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals. *Nat Commun* **11**:5972. doi:10.1038/s41467-020-19603-1
Mulholland, Traube, Ugur, Parsa, Eckl, Schönung, Modic, Bartoschek, Stolz, Ryan, Carell, Leonhardt, Bultmann. 2020b. Distinct and stage-specific contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. *Sci Rep* **10**:1–12. doi:10.1038/s41598-020-68600-3
Muñoz-Sanjuán, Brivanlou. 2002. Neural induction, the default model and embryonic stem cells. *Nat Rev Neurosci* **3**:271–280. doi:10.1038/nrn786
Nakaki, Hayashi, Ohta, Kurimoto, Yabuta, Saitou. 2013. Induction of mouse germ-cell fate by transcription factors in vitro. *Nature* **501**:222–226. doi:10.1038/nature12417
Nakaki, Saitou. 2014. PRDM14: A unique regulator for pluripotency and epigenetic reprogramming. *Trends Biochem Sci* **39**:289–298. doi:10.1016/j.tibs.2014.04.003
Nakamura, Arai, Umehara, Masuhara, Kimura, Taniguchi, Sekimoto, Ikawa, Yoneda, Okabe, Tanaka, Shiota, Nakano. 2007. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9**:64–71. doi:10.1038/ncb1519
Nakamura, Liu, Nakashima, Umehara, Inoue, Matoba, Tachibana, Ogura, Shinkai, Nakano. 2012. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**:415–419. doi:10.1038/nature11093
Neagu, van Genderen, Escudero, Verwegen, Kurek, Lehmann, Stel, Dirks, van Mierlo, Maas, Eleveld, Ge, den Dekker, Brouwer, van IJcken, Modic, Drukker, Jansen, Rivron, Baart, Marks, ten Berge. 2020. In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states. *Nat Cell Biol* **22**:534–545. doi:10.1038/s41556-020-0508-x
Neri, Incarnato, Krepelova, Rapelli, Pagnani, Zecchina, Parlato, Oliviero. 2013a. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol* **14**. doi:10.1186/gb-2013-14-8-r91
Neri, Krepelova, Incarnato, Maldotti, Parlato, Galvagni, Matarese, Stunnenberg, Oliviero. 2013b. Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. *Cell* **155**:121–34. doi:10.1016/j.cell.2013.08.056
Neri, Rapelli, Krepelova, Incarnato, Parlato, Basile, Maldotti, Anselmi, Oliviero. 2017. Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**:72–77. doi:10.1038/nature21373
Ng, Zhang, Hendrich, Johnson, Turner, Erdjument-Bromage, Tempst, Reinberg, Bird. 1999. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* **23**:58–61. doi:10.1038/12659
Nicetto, Donahue, Jain, Peng, Sidoli, Sheng, Montavon, Becker, Grindheim, Blahnik, Garcia, Tan, Bonasio, Jenuwein, Zaret. 2019. H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification. *Science (80- )* **363**:294–297. doi:10.1126/science.aau0583
Nicetto, Zaret. 2019. Role of H3K9me3 heterochromatin in cell identity establishment and maintenance. *Curr Opin Genet Dev* **55**:1–10. doi:10.1016/j.gde.2019.04.013
Nicholls, Page. 2021. Germ cell determination and the developmental origin of germ cell tumors. *Development* **148**. doi:10.1242/dev.198150
Nichols, Smith. 2009. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**:487–492. doi:10.1016/j.stem.2009.05.015
Nicolas Rivron, Martin Pera, Janet Rossant, Alfonso Martinez Arias, Magdalena Zernicka-Goetz, Jianping Fu, Susanne van den Brink, Annelien Bredenoord, Wybo Dondorp, Guido de Wert, Insoo Hyun, Megan Munsie & Rosario Isasi. 2018. Debate ethics of embryo models from stem cells. *Nature* **564**:183–185.
Niederhuth, Schmitz. 2017. Putting DNA methylation in context: from genomes to gene expression in plants. *Biochim Biophys Acta - Gene Regul Mech* **1860**:149–156. doi:10.1016/j.bbagrm.2016.08.009

Nielsen. 2002. Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1alpha. *EMBO J* **21**:5797–5806. doi:10.1093/emboj/cdf560

Nielsen, Nietlispach, Mott, Callaghan, Bannister, Kouzarides, Murzin, Murzina, Laue. 2002. Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**:103–107. doi:10.1038/nature722

Nishioka, Yamamoto, Kiyonari, Sato, Sawada, Ota, Nakao, Sasaki. 2008. Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mech Dev* **125**:270–283. doi:10.1016/j.mod.2007.11.002

Nishiyama, Mulholland, Bultmann, Kori, Endo, Saeki, Qin, Trummer, Chiba, Yokoyama, Kumamoto, Kawakami, Hojo, Nagae, Aburatani, Tanaka, Arita, Leonhardt, Nakanishi. 2020. Two distinct modes of DNMT1 recruitment ensure stable maintenance DNA methylation. *Nat Commun* **11**:1222. doi:10.1038/s41467-020-15006-4

Noh, Wang, Kim, Wenderski, Fang, Li, Dewell, Hughes, Melnick, Patel, Li, Allis. 2015. Engineering of a Histone-Recognition Domain in Dnmt3a Alters the Epigenetic Landscape and Phenotypic Features of Mouse ESCs. *Mol Cell* **59**:89–103. doi:10.1016/j.molcel.2015.05.017

Norton, Imai, Yau, Bradbury. 1989. Histone acetylation reduces nucleosome core particle linking number change. *Cell* **57**:449–457. doi:10.1016/0092-8674(89)90920-3

Nowotschin, Hadjantonakis, Campbell. 2019. The endoderm: A divergent cell lineage with many commonalities. *Dev* **146**:1–12. doi:10.1242/dev.150920

O'Carroll, Erhardt, Pagani, Barton, Surani, Jenuwein. 2001. The polycomb-group gene Ezh2 is required for early mouse development. *Mol Cell Biol* **21**:4330–6. doi:10.1128/MCB.21.13.4330-4336.2001

Oakes, La Salle, Smiraglia, Robaire, Trasler. 2007. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev Biol* **307**:368–379. doi:10.1016/j.ydbio.2007.05.002

Ogushi, Yamagata, Obuse, Furuta, Wakayama, Matzuk, Saitou. 2017. Reconstitution of the oocyte nucleolus in mice through a single nucleolar protein, NPM2. *J Cell Sci* **130**:2416–2429. doi:10.1242/jcs.195875

Ohinata, Payer, O'Carroll, Ancelin, Ono, Sano, Barton, Obukhanych, Nussenzweig, Tarakhovsky, Saitou, Surani. 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**:207–213. doi:10.1038/nature03813

Ohno. 1967. Sex chromosome and sex-linked gens. *Springer*.

Ohtsuka, Nishikawa-Torikai, Niwa. 2012. E-Cadherin Promotes Incorporation of Mouse Epiblast Stem Cells into Normal Development. *PLoS One* **7**. doi:10.1371/journal.pone.0045220

Okae, Chiba, Hiura, Hamada, Sato, Utsunomiya, Kikuchi, Yoshida, Tanaka, Suyama, Arima. 2014. Genome-Wide Analysis of DNA Methylation Dynamics during Early Human Development. *PLoS Genet* **10**:1–12. doi:10.1371/journal.pgen.1004868

Okano, Bell, Haber, Li. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**:247–257. doi:10.1016/S0092-8674(00)81656-6

Okano, Xie, Li. 1998. Cloning and characterization of a family of novel mammalian DNA ( cytosine-5 ) methyltransferases Non-invasive sexing of preimplantation stage mammalian embryos. *Nat Am Inc* **19**:219–220.

Okashita, Kumaki, Ebi, Nishi, Okamoto, Nakayama, Hashimoto, Nakamura, Sugasawa, Kojima, Takada, Okano, Seki. 2014. PRDM14 promotes active DNA demethylation through the Ten-eleven translocation (TET)-mediated base excision repair pathway in embryonic stem cells. *Development* **141**:269–280. doi:10.1242/dev.099622

Okashita, Suwa, Nishimura, Sakashita, Kadota, Nagamatsu, Kawaguchi, Kashida, Nakajima, Tachibana, Seki. 2016. PRDM14 Drives OCT3/4 Recruitment via Active Demethylation in the Transition from Primed to Naive Pluripotency. *Stem Cell Reports* **7**:1072–1086. doi:10.1016/j.stemcr.2016.10.007

Olmsted, Paluh. 2021. Co-development of central and peripheral neurons with trunk mesendoderm in human elongating multi-lineage organized gastruloids. *Nat Commun* **12**:1–19. doi:10.1038/s41467-021-23294-7

Ooi, Qiu, Bernstein, Li, Jia, Yang, Erdjument-Bromage, Tempst, Lin, Allis, Cheng, Bestor. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**:714–717. doi:10.1038/nature05987

Ortega-Recalde, Hore. 2019. DNA methylation in the vertebrate germline: Balancing memory and

erasure. *Essays Biochem* **63**:649–661. doi:10.1042/EBC20190038
Ortmann, Brown, Czechanski, Aydin, Muraro, Huang, Tomaz, Osnato, Canu, Wesley, Skelly, Stegle, Choi, Churchill, Baker, Rugg-Gunn, Munger, Reinholdt, Vallier. 2020. Naive Pluripotent Stem Cells Exhibit Phenotypic Variability that Is Driven by Genetic Variation. *Cell Stem Cell* **27**:470-481.e6. doi:10.1016/j.stem.2020.07.019
Otani, Nankumo, Arita, Inamoto, Ariyoshi, Shirakawa. 2009. Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. *EMBO Rep* **10**:1235–1241. doi:10.1038/embor.2009.218
Pachano, Sánchez-Gaya, Ealo, Mariner-Faulí, Bleckwehl, Asenjo, Respuela, Cruz-Molina, Muñoz-San Martín, Haro, van IJcken, Landeira, Rada-Iglesias. 2021. Orphan CpG islands amplify poised enhancer regulatory activity and determine target gene responsiveness. *Nat Genet*. doi:10.1038/s41588-021-00888-x
Padmanabhan, Jia, Geary-Joo, Wu, Ferguson-Smith, Fung, Bieda, Snyder, Gravel, Cross, Watson. 2013. Mutation in Folate Metabolism Causes Epigenetic Instability and Transgenerational Effects on Development. *Cell* **155**:81–93. doi:10.1016/j.cell.2013.09.002
Parry, Rulands, Reik. 2020. Active turnover of DNA methylation during cell fate decisions. *Nat Rev Genet*. doi:10.1038/s41576-020-00287-8
Pauler, Sloane, Huang, Regha, Koerner, Tamir, Sommer, Aszodi, Jenuwein, Barlow. 2009. H3K27me3 forms BLOCs over silent genes and intergenic regions and specifies a histone banding pattern on a mouse autosomal chromosome. *Genome Res* **19**:221–233. doi:10.1101/gr.080861.108
Peaston, Evsikov, Graber, de Vries, Holbrook, Solter, Knowles. 2004. Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev Cell* **7**:597–606. doi:10.1016/j.devcel.2004.09.004
Peat, Dean, Clark, Krueger, Smallwood, Ficz, Kim, Marioni, Hore, Reik. 2014. Genome-wide Bisulfite Sequencing in Zygotes Identifies Demethylation Targets and Maps the Contribution of TET3 Oxidation. *Cell Rep* **9**:1990–2000. doi:10.1016/j.celrep.2014.11.034
Pei, Mao, Li, Li, Xu, Xu, Shi, Gao, Hu, Hu, Chen, Wu, Liu, Zhang, Zhang, Guo, Tang, Dai. 2014. Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. *Cell Stem Cell* **14**:512–522. doi:10.1016/j.stem.2014.01.001
Petell, Alabdi, He, San Miguel, Rose, Gowher. 2016. An epigenetic switch regulates de novo DNA methylation at a subset of pluripotency gene enhancers during embryonic stem cell differentiation. *Nucleic Acids Res* **44**:7605–7617. doi:10.1093/nar/gkw426
Picelli, Faridani, Björklund, Winberg, Sagasser, Sandberg. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**:171–181. doi:10.1038/nprot.2014.006
Pieters, van Roy. 2014. Role of cell-cell adhesion complexes in embryonic stem cell biology. *J Cell Sci* **127**:2603–2613. doi:10.1242/jcs.146720
Piotrowska-Nitsche, Perea-Gomez, Haraguchi, Zernicka-Goetz. 2005. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**:479–490. doi:10.1242/dev.01602
Pisco, Fouquier d'Hérouël, Huang. 2016. Conceptual Confusion: The case of Epigenetics. *bioRxiv* 053009. doi:10.1101/053009
Plusa, Frankenberg, Chalmers, Hadjantonakis, Moore, Papalopulu, Papaioannou, Glover, Zernicka-Goetz. 2005. Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. *J Cell Sci* **118**:505–515. doi:10.1242/jcs.01666
Plusa, Piliszek, Frankenberg, Artus, Hadjantonakis. 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135**:3081–3091. doi:10.1242/dev.021519
Pohlers, Truss, Frede, Scholz, Strehle, Kuban, Hoffmann, Morkel, Birchmeier, Hagemeier. 2005. A role for E2F6 in the restriction of male-germ-cell-specific gene expression. *Curr Biol* **15**:1051–1057. doi:10.1016/j.cub.2005.04.060
Potok, Nix, Parnell, Cairns. 2013. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**:759–772. doi:10.1016/j.cell.2013.04.030
Pratt, Chakraborty, Surani. 1981. Molecular and morphological differentiation of the mouse blastocyst after manipulations of compaction with cytochalasin D. *Cell* **26**:279–292. doi:10.1016/0092-8674(81)90310-X
Proudhon, Duffié, Ajjan, Cowley, Iranzo, Carbajosa, Saadeh, Holland, Oakey, Rakyan, Schulz, Bourc'his. 2012. Protection against De Novo Methylation Is Instrumental in Maintaining Parent-of-Origin Methylation Inherited from the Gametes. *Mol Cell* **47**:909–920.

doi:10.1016/j.molcel.2012.07.010
Quenneville, Verde, Corsinotti, Kapopoulou, Jakobsson, Offner, Baglivo, Pedone, Grimaldi, Riccio, Trono. 2011. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* **44**:361–372. doi:10.1016/j.molcel.2011.08.032
Raj, van Oudenaarden. 2008. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135**:216–226. doi:10.1016/j.cell.2008.09.050
Rauch, Haakonsson, Madsen, Larsen, Forss, Madsen, Van Hauwaert, Wiwie, Jespersen, Tencerova, Nielsen, Larsen, Röttger, Baumbach, Scheele, Kassem, Mandrup. 2019. Osteogenesis depends on commissioning of a network of stem cell transcription factors that act as repressors of adipogenesis. *Nat Genet* **51**:716–727. doi:10.1038/s41588-019-0359-1
Reddington, Perricone, Nestor, Reichmann, Youngson, Suzuki, Reinhardt, Dunican, Prendergast, Mjoseng, Ramsahoye, Whitelaw, Greally, Adams, Bickmore, Meehan. 2013. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol* **14**. doi:10.1186/gb-2013-14-3-r25
Respuela, Nikolić, Tan, Frommolt, Zhao, Wysocka, Rada-Iglesias. 2016. Foxd3 Promotes Exit from Naive Pluripotency through Enhancer Decommissioning and Inhibits Germline Specification. *Cell Stem Cell* **18**:118–33. doi:10.1016/j.stem.2015.09.010
Reversade, Kuroda, Lee, Mays, De Robertis. 2005. Depletion of Bmp2, Bmp4, Bmp7 and Spemann organizer signals induces massive brain formation in Xenopus embryos. *Development* **132**:3381–3392. doi:10.1242/dev.01901
Riising, Comet, Leblanc, Wu, Johansen, Helin. 2014. Gene Silencing Triggers Polycomb Repressive Complex 2 Recruitment to CpG Islands Genome Wide. *Mol Cell* **55**:347–360. doi:10.1016/j.molcel.2014.06.005
Riveiro, Brickman. 2020. From pluripotency to totipotency: an experimentalist's guide to cellular potency. *Development* **147**. doi:10.1242/dev.189845
Rivera-Pérez, Hadjantonakis. 2015. The dynamics of morphogenesis in the early mouse embryo. *Cold Spring Harb Perspect Biol* **7**:1–18. doi:10.1101/cshperspect.a015867
Rodda, Chew, Lim, Loh, Wang, Ng, Robson. 2005. Transcriptional Regulation of Nanog by OCT4 and SOX2. *J Biol Chem* **280**:24731–24737. doi:10.1074/jbc.M502573200
Rodriguez-Terrones, Gaume, Ishiuchi, Weiss, Kopp, Kruse, Penning, Vaquerizas, Brino, Torres-Padilla. 2018. A molecular roadmap for the emergence of early-embryonic-like cells in culture. *Nat Genet* **50**:106–119. doi:10.1038/s41588-017-0016-5
Rompala, Ferguson, Homanics. 2020. Coincubation of sperm with epididymal extracellular vesicle preparations from chronic intermittent ethanol-treated mice is sufficient to impart anxiety-like and ethanol-induced behaviors to adult progeny. *Alcohol* **87**:111–120. doi:10.1016/j.alcohol.2020.05.001
Rothbart, Krajewski, Nady, Tempel, Xue, Badeaux, Barsyte-Lovejoy, Martinez, Bedford, Fuchs, Arrowsmith, Strahl. 2012. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* **19**:1155–1160. doi:10.1038/nsmb.2391
Rowe, Jakobsson, Mesnard, Rougemont, Reynard, Aktas, Maillard, Layard-Liesching, Verp, Marquis, Spitz, Constam, Trono. 2010. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**:237–240. doi:10.1038/nature08674
Rulands, Lee, Clark, Angermueller, Smallwood, Krueger, Mohammed, Dean, Nichols, Rugg-Gunn, Kelsey, Stegle, Simons, Reik. 2018. Genome-Scale Oscillations in DNA Methylation during Exit from Pluripotency. *Cell Syst* **7**:63-76.e12. doi:10.1016/j.cels.2018.06.012
Sachs, Onodera, Blaschke, Ebata, Song, Ramalho-Santos. 2013. Bivalent Chromatin Marks Developmental Regulatory Genes in the Mouse Embryonic Germline InVivo. *Cell Rep* **3**:1777–1784. doi:10.1016/j.celrep.2013.04.032
Sado, Fenner, Tan, Tam, Shioda, Li. 2000. X inactivation in the mouse embryo deficient for Dnmt1: Distinct effect of hypomethylation on imprinted and random X inactivation. *Dev Biol* **225**:294–303. doi:10.1006/dbio.2000.9823
Saitou, Barton, Surani, Ginsburg, Snow, McLaren. 2002. A molecular programme for the specification of germ cell fate in mice: Primordial germ cells in the mouse embryo during gastrulation. *Nature* **418**:293-300.
Saitou, Yamaji. 2010. Germ cell specification in mice: Signaling, transcription regulation, and epigenetic consequences. *Reproduction* **139**:931–942. doi:10.1530/REP-10-0043
Sakaue, Ohta, Kumaki, Oda, Sakaide, Matsuoka, Yamagiwa, Niwa, Wakayama, Okano. 2010. DNA

methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr Biol* **20**:1452–1457. doi:10.1016/j.cub.2010.06.050

Sánchez-Romero, Casadesús. 2020. The bacterial epigenome. *Nat Rev Microbiol* **18**:7–20. doi:10.1038/s41579-019-0286-2

Sancho, Di-Gregorio, George, Pozzi, Sánchez, Pernaute, Rodríguez. 2013. Competitive interactions eliminate unfit embryonic stem cells at the onset of differentiation. *Dev Cell* **26**:19–30. doi:10.1016/j.devcel.2013.06.012

Sankar, Mohammad, Sundaramurthy, Wang, Lerdrup, Tatar, Helin. 2022. Histone editing elucidates the functional roles of H3K27 methylation and acetylation in mammals. *Nat Genet*. doi:10.1038/s41588-022-01091-2

Sato, Sargent. 1989. Development of neural inducing capacity in dissociated Xenopus embryos. *Dev Biol* **134**:263–266. doi:10.1016/0012-1606(89)90096-1

Schlesinger, Meshorer. 2019. Open Chromatin, Epigenetic Plasticity, and Nuclear Organization in Pluripotency. *Dev Cell* **48**:135–150. doi:10.1016/j.devcel.2019.01.003

Schmidt, Bultmann, Meilinger, Zacher, Tresch, Maier, Peter, Martin, Leonhardt, Spada. 2012. Global DNA Hypomethylation Prevents Consolidation of Differentiation Programs and Allows Reversion to the Embryonic Stem Cell State. *PLoS One* **7**:e52629. doi:10.1371/journal.pone.0052629

Schmitz, Lewis, Goll. 2019. DNA Methylation: Shared and Divergent Features across Eukaryotes. *Trends Genet* **xx**. doi:10.1016/j.tig.2019.07.007

Schomacher, Han, Musheev, Arab, Kienhöfer, Von Seggern, Niehrs. 2016. Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. *Nat Struct Mol Biol* **23**:116–124. doi:10.1038/nsmb.3151

Schorderet, Gartler. 1992. Analysis of CpG suppression in methylated and nonmethylated species. *Proc Natl Acad Sci U S A* **89**:957–961. doi:10.1073/pnas.89.3.957

Schroeder, Jayashankar, Douglas, Thirkill, York, Dickinson, Williams, Samollow, Ross, Bannasch, Douglas, LaSalle. 2015. Early Developmental and Evolutionary Origins of Gene Body DNA Methylation Patterns in Mammalian Placentas. *PLoS Genet* **11**:1–20. doi:10.1371/journal.pgen.1005442

Schulz, Meisig, Nakamura, Okamoto, Sieber, Picard, Borensztein, Saitou, Blüthgen, Heard. 2014. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. *Cell Stem Cell* **14**:203–216. doi:10.1016/j.stem.2013.11.022

Schulz, Proudhon, Bestor, Woodfine, Lin, Lin, Prissette, Oakey, Bourc'his. 2010. The parental non-equivalence of imprinting control regions during mammalian development and evolution. *PLoS Genet* **6**. doi:10.1371/journal.pgen.1001214

Seisenberger, Andrews, Krueger, Arand, Walter, Santos, Popp, Thienpont, Dean, Reik. 2012. The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. *Mol Cell* **48**:849–862. doi:10.1016/j.molcel.2012.11.001

Seisenberger, Peat, Hore, Santos, Dean, Reik. 2013a. Reprogramming DNA methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Philos Trans R Soc B Biol Sci* **368**:1–11. doi:10.1098/rstb.2011.0330

Seisenberger, Peat, Reik. 2013b. Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr Opin Cell Biol* **25**:281–288. doi:10.1016/j.ceb.2013.02.013

Sendžikaitė, Hanna, Stewart-Morgan, Ivanova, Kelsey. 2019. A DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation in mice. *Nat Commun* **10**. doi:10.1038/s41467-019-09713-w

Senft, Bikoff, Robertson, Costello. 2019. Genetic dissection of Nodal and Bmp signalling requirements during primordial germ cell development in mouse. *Nat Commun* **10**:1089. doi:10.1038/s41467-019-09052-w

Senft, Macfarlan. 2021. Transposable elements shape the evolution of mammalian development. *Nat Rev Genet* **0123456789**. doi:10.1038/s41576-021-00385-1

Severino, Bauer, Mattimoe, Arecco, Cozzuto, Lorden, Hamada, Nosaka, Nagaoka, Audergon, Tarruell, Heyn, Hayashi, Saitou, Payer. 2022. Controlled X-chromosome dynamics defines meiotic potential of female mouse in vitro germ cells. *EMBO J*. doi:10.15252/embj.2021109457

Sevilla, Papatsenko, Mazloom, Xu, Vasileva, Unwin, LeRoy, Chen, Garrett-Bakelman, Lee, Trinite, Webb, Wang, Su, Gingold, Melnick, Garcia, Whetton, MacArthur, Ma'ayan, Lemischka. 2021. An Esrrb and Nanog Cell Fate Regulatory Module Controlled by Feed Forward Loop Interactions. *Front Cell Dev Biol* **9**. doi:10.3389/fcell.2021.630067

Shao, Li, Chen, Zhang, Lian, Hu, Sun, Jin, Liu, Wang, Zhao, Jin, Liang, Li, Zheng, Zhang, Wang, Zhang. 2014. Epigenetic modification and inheritance in sexual reversal of fish. *Genome Res* **24**:604–615. doi:10.1101/gr.162172.113

Sharif, Muto, Takebayashi, Suetake, Iwamatsu, Endo, Shinga, Mizutani-Koseki, Toyoda, Okamura, Tajima, Mitsuya, Okano, Koseki. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**:908–912. doi:10.1038/nature06397

Sharma, Conine, Shea, Boskovic, Derr, Bing, Belleannee, Kucukural, Serra, Sun, Song, Carone, Ricci, Li, Fauquier, Moore, Sullivan, Mello, Garber, Rando. 2016. Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. *Science (80- )* **351**:391–396. doi:10.1126/science.aad6780

Shen, Inoue, He, Liu, Lu, Zhang. 2014. Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes. *Cell Stem Cell* **15**:459–470. doi:10.1016/j.stem.2014.09.002

Shen, Yang, Li, Zhang, Peng, Wang, Chang, Ong, Du. 2021. Mouse totipotent stem cells captured and maintained through spliceosomal repression. *Cell* 1–17. doi:10.1016/j.cell.2021.04.020

Shi, Chen, Li, Zheng, Zhang, Qiao, Tang, Tao, Zhou, Duan. 2015. Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell rna-seq. *Dev* **142**:3468–3477. doi:10.1242/dev.123950

Shi, Strogantsev, Takahashi, Kazachenka, Lorincz, Hemberger, Ferguson-Smith. 2019. ZFP57 regulation of transposable elements and gene expression within and beyond imprinted domains. *Epigenetics and Chromatin* **12**:1–13. doi:10.1186/s13072-019-0295-4

Shirane, Kurimoto, Yabuta, Yamaji, Satoh, Ito, Watanabe, Hayashi, Saitou, Sasaki. 2016. Global Landscape and Regulatory Principles of DNA Methylation Reprogramming for Germ Cell Specification by Mouse Pluripotent Stem Cells. *Dev Cell* **39**:87–103. doi:10.1016/j.devcel.2016.08.008

Shirane, Miura, Ito, Lorincz. 2020. NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts Polycomb-associated silencing. *Nat Genet* **52**:1088–1098. doi:10.1038/s41588-020-0689-z

Shirane, Toh, Kobayashi, Miura, Chiba, Ito, Kono, Sasaki. 2013. Mouse Oocyte Methylomes at Base Resolution Reveal Genome-Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases. *PLoS Genet* **9**. doi:10.1371/journal.pgen.1003439

Shukla, Kavak, Gregory, Imashimizu, Shutinoski, Kashlev, Oberdoerffer, Sandberg, Oberdoerffer. 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**:74–79. doi:10.1038/nature10442

Siklenka, Erkek, Godmann, Lambrot, McGraw, Lafleur, Cohen, Xia, Suderman, Hallett, Trasler, Peters, Kimmins. 2015. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science (80- )* **350**. doi:10.1126/science.aab2006

Silva, Nichols, Theunissen, Guo, van Oosten, Barrandon, Wray, Yamanaka, Chambers, Smith. 2009. Nanog is the gateway to the pluripotent ground state. *Cell* **138**:722–37. doi:10.1016/j.cell.2009.07.039

Sim, Kim, Nayfeh, Yun, Kim, Park, Kim, Kim. 2017. 2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic Mechanisms. *Stem Cell Reports* **8**:1312–1328. doi:10.1016/j.stemcr.2017.04.001

Simpson, Caballero, Jungbluth, Chen, Old. 2005. Cancer/testis antigens, gametogenesis and cancer. *Nat Rev Cancer* **5**:615–625. doi:10.1038/nrc1669

Singer, Yong, Tischler, Hackett, Altinok, Surani, Cai, Elowitz. 2014. Dynamic Heterogeneity and DNA Methylation in Embryonic Stem Cells. *Mol Cell* **55**:319–331. doi:10.1016/j.molcel.2014.06.029

Skene, Henikoff. 2017. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**:1–35. doi:10.7554/eLife.21856

Smith. 2017. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**:365–373. doi:10.1242/dev.142679

Smith, Shi, Gu, Donaghey, Clement, Cacchiarelli, Gnirke, Michor, Meissner. 2017. Epigenetic restriction of extraembryonic lineages mirrors the somatic transition to cancer. *Nature* **549**:543–547. doi:10.1038/nature23891

Smukler, Runciman, Xu, Van Der Kooy. 2006. Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J Cell Biol* **172**:79–90. doi:10.1083/jcb.200508085

Song, Rechkoblit, Bestor, Patel. 2011. Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science (80- )* **331**:1036–1040.

doi:10.1126/science.1195380
Song, van den Berg, Markoulaki, Soldner, Dall'Agnese, Henninger, Drotar, Rosenau, Cohen, Young, Semrau, Stelzer, Jaenisch. 2019. Dynamic Enhancer DNA Methylation as Basis for Transcriptional and Cellular Heterogeneity of ESCs. *Mol Cell* **75**:905-920.e6. doi:10.1016/j.molcel.2019.06.045
Sparta, Hamilton, Hughes, Deeds. 2022. A lack of distinct cell identities in single-cell measurements: revisiting Waddington ' s landscape. *bioRxiv*.
Stadler, Murr, Burger, Ivanek, Lienert, Schöler, Nimwegen, Wirbelauer, Oakeley, Gaidatzis, Tiwari, Schübeler. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**:490–495. doi:10.1038/nature10716
Stainier. 2005. No organ left behind: Tales of gut development and evolution. *Science (80- )* **307**:1902–1904. doi:10.1126/science.1108709
Stavridis, Simon Lunn, Collins, Storey. 2007. A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. *Development* **134**:2889–2894. doi:10.1242/dev.02858
Stevens. 1958. Studies on transplantable testicular teratomas of strain 129 mice. *J Natl Cancer Inst* **20**:1257–1275. doi:10.1093/jnci/20.6.1257
Stirparo, Kurowski, Yanagida, Bates, Strawbridge, Hladkou, Stuart, Boroviak, Silva, Nichols. 2021. OCT4 induces embryonic pluripotency via STAT3 signaling and metabolic mechanisms. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/PNAS.2008890118
Strom, Emelyanov, Mir, Fyodorov, Darzacq, Karpen. 2017. Phase separation drives heterochromatin domain formation. *Nature* **547**:241–245. doi:10.1038/nature22989
Struhl. 1999. Fundamentally Different Logic of Gene Regulation in Eukaryotes and Prokaryotes. *Cell* **98**:1–4. doi:10.1016/S0092-8674(00)80599-1
Suetake, Shinozaki, Miyagawa, Takeshima, Tajima. 2004. DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J Biol Chem* **279**:27816–27823. doi:10.1074/jbc.M400181200
Sullivan, Frenette, Girouard. 2007. Epididymosomes are involved in the acquisition of new sperm proteins during epididymal transit. *Asian J Androl* **9**:483–491. doi:10.1111/j.1745-7262.2007.00281.x
Sun, Wei, Wu, Hu, Wang, Wang, Zhang, Chen, Huang, Chen. 2005. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J Biol Chem* **280**:35261–35271. doi:10.1074/jbc.M504012200
Sundaram, Cheng, Ma, Li, Xing, Edge, Snyder, Wang. 2014. Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome Res* **24**:1963–1976. doi:10.1101/gr.168872.113
Surani, Barton, Norris. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308**:548–550. doi:10.1038/308548a0
Suter, Tirefort, Julien, Krause. 2009. A Sox1 to Pax6 Switch Drives Neuroectoderm to Radial Glia Progression During Differentiation of Mouse Embryonic Stem Cells. *Stem Cells* **27**:49–58. doi:10.1634/stemcells.2008-0319
Suzuki, Maeda, Furuhata, Shimizu, Nishimura, Kishima, Suzuki. 2017. A screening system to identify transcription factors that induce binding site-directed DNA demethylation. *Epigenetics and Chromatin* **10**:1–14. doi:10.1186/s13072-017-0169-6
Tachibana, Sugimoto, Nozaki, Ueda, Ohta, Ohki, Fukuda, Takeda, Niida, Kato, Shinkai. 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**:1779–1791. doi:10.1101/gad.989402
Takahashi, Coluccio, Thorball, Planet, Shi, Offner, Turelli, Imbeault, Ferguson-Smith, Trono. 2019. ZNF445 is a primary regulator of genomic imprinting. *Genes Dev* **33**:49–54. doi:10.1101/gad.320069.118
Talbert, Meers, Henikoff. 2019. Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat Rev Genet* **20**:283–297. doi:10.1038/s41576-019-0105-7
Tam, Zhou. 1996. The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo. *Dev Biol* **178**:124–132. doi:10.1006/dbio.1996.0203
Tanaka, Nagamatsu, Tokitake, Kasa, Tam, Matsui. 2004. Regulation of expression of mouse interferon-induced transmembrane protein like gene-3, Ifitm3 (mil-1, fragilis), in germ cells. *Dev Dyn* **230**:651–659. doi:10.1002/dvdy.20085
Tanay, O'Donnell, Damelin, Bestor. 2007. Hyperconserved CpG domains underlie Polycomb-binding

sites. *Proc Natl Acad Sci U S A* **104**:5521–5526. doi:10.1073/pnas.0609746104
Tang, Castillo-Venzor, Gruhn, Kobayashi, Penfold, Morgan, Sun, Irie, Surani. 2022. Sequential enhancer state remodelling defines human germline competence and specification. *Nat Cell Biol*. doi:10.1038/s41556-022-00878-z
Tarkowski, Wróblewska. 1967. Development of blastomeres of mouse eggs isolated at the 4- and 8-cell stage. *J Embryol Exp Morphol* **18**:155–180. doi:10.1242/dev.18.1.155
Technau, Scholz. 2003. Origin and evolution of endoderm and mesoderm. *Int J Dev Biol* **47**:531–539. doi:10.1387/ijdb.14756329
ten Berge, Koole, Fuerer, Fish, Eroglu, Nusse. 2008. Wnt Signaling Mediates Self-Organization and Axis Formation in Embryoid Bodies. *Cell Stem Cell* **3**:508–518. doi:10.1016/j.stem.2008.09.013
Tesar, Chenoweth, Brook, Davies, Evans, Mack, Gardner, McKay. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**:196–199. doi:10.1038/nature05972
Thomson, Liu, Zou, Smith, Meissner, Ramanathan. 2011. Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **145**:875–889. doi:10.1016/j.cell.2011.05.017
Tischler, Gruhn, Reid, Allgeyer, Buettner, Marr, Theis, Simons, Wernisch, Surani. 2018. Metabolic regulation of pluripotency and germ cell fate through α-ketoglutarate. *EMBO J* e99518. doi:10.15252/embj.201899518
Todd, Deniz, Taylor, Branco. 2019. Functional evaluation of transposable elements as enhancers in mouse embryonic and trophoblast stem cells. *Elife* **8**:1–20. doi:10.7554/eLife.44344
Tosic, Kim, Pavlovic, Schröder, Mersiowsky, Barg, Hofherr, Probst, Köttgen, Hein, Arnold. 2019. Eomes and Brachyury control pluripotency exit and germ-layer segregation by changing the chromatin state. *Nat Cell Biol* **21**. doi:10.1038/s41556-019-0423-1
Toyooka. 2020. Trophoblast lineage specification in the mammalian preimplantation embryo. *Reprod Med Biol* **19**:209–221. doi:10.1002/rmb2.12333
Tremethick. 2007. Higher-Order Structures of Chromatin: The Elusive 30 nm Fiber. *Cell* **128**:651–654. doi:10.1016/j.cell.2007.02.008
Trojer, Reinberg. 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell* **28**:1–13. doi:10.1016/j.molcel.2007.09.011
Tropepe, Hitoshi, Sirard, Mak, Rossant, Van Der Kooy. 2001. Direct neural fate specification from embryonic stem cells: A primitive mammalian neural stem cell stage acquired through a default mechanism. *Neuron* **30**:65–78. doi:10.1016/S0896-6273(01)00263-X
Trott, Martinez Arias. 2013. Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. *Biol Open* **2**:1049–1056. doi:10.1242/bio.20135934
Tsankov, Wadsworth, Akopian, Charlton, Allon, Arczewska, Mead, Drake, Smith, Mikkelsen, Shalek, Meissner. 2019. Loss of DNA methyltransferase activity in primed human ES cells triggers increased cell-cell variability and transcriptional repression. *Development* **146**:dev174722. doi:10.1242/dev.174722
Tsogtbaatar, Landin, Minter-Dykhouse, Folmes. 2020. Energy Metabolism Regulates Stem Cell Pluripotency. *Front Cell Dev Biol* **8**:1–16. doi:10.3389/fcell.2020.00087
Tsumura, Hayakawa, Kumaki, Takebayashi, Sakaue, Matsuoka, Shimotohno, Ishikawa, Li, Ueda, Nakayama, Okano. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to Cells* **11**:805–814. doi:10.1111/j.1365-2443.2006.00984.x
Turco, Furia, Dietze, Diaz, Ronzoni, Sciullo, Simeone, Constam, Faretta, Lanfrancone. 2012. Cellular heterogeneity during embryonic stem cell differentiation to epiblast stem cells is revealed by the ShcD/RaLP adaptor protein. *Stem Cells* **30**:2423–2436. doi:10.1002/stem.1217
Ueda, Okano, Williams, Chen, Georgopoulos, Li. 2006. Roles for Dnmt3b in mammalian development: A mouse model for the ICF syndrome. *Development* **133**:1183–1192. doi:10.1242/dev.02293
Van Den Brink, Baillie-Johnson, Balayo, Hadjantonakis, Nowotschin, Turner, Arias. 2014. Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Dev* **141**:4231–4242. doi:10.1242/dev.113001
van den Brink, van Oudenaarden. 2021. 3D gastruloids: a novel frontier in stem cell-based in vitro modeling of mammalian gastrulation. *Trends Cell Biol* **xx**:1–13. doi:10.1016/j.tcb.2021.06.007
Van Der Heijden, Dieker, Derijck, Muller, Berden, Braat, Van Der Vlag, De Boer. 2005. Asymmetry in Histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* **122**:1008–1022. doi:10.1016/j.mod.2005.04.009

Van Tongelen, Loriot, De Smet. 2017. Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. *Cancer Lett* **396**:130–137. doi:10.1016/j.canlet.2017.03.029

Veenvliet, Bolondi, Kretzmer, Haut, Scholze-Wittler, Schifferl, Koch, Guignard, Kumar, Pustet, Heimann, Buschow, Wittler, Timmermann, Meissner, Herrmann. 2020. Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science (80- )* **370**:eaba4937. doi:10.1126/science.aba4937

Velasco, Hube, Rollin, Neuillet, Philippe, Bouzinba-Segard, Galvani, Viegas-Pequignot, Francastel. 2010. Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc Natl Acad Sci* **107**:9281–9286. doi:10.1073/pnas.1000473107

Verma, Pan, Doré, Shukla, Li, Pelham-Webb, Teijeiro, González, Krivtsov, Chang, Papapetrou, He, Elemento, Huangfu. 2018. TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells. *Nat Genet* **50**:83–95. doi:10.1038/s41588-017-0002-y

Vilkaitis, Suetake, Klimašauskas, Tajima. 2005. Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J Biol Chem* **280**:64–72. doi:10.1074/jbc.M411126200

von Meyenn, Berrens, Andrews, Santos, Collier, Krueger, Osorno, Dean, Rugg-Gunn, Reik. 2016a. Comparative Principles of DNA Methylation Reprogramming during Human and Mouse In Vitro Primordial Germ Cell Specification. *Dev Cell* **39**:104–115. doi:10.1016/j.devcel.2016.09.015

von Meyenn, Iurlaro, Habibi, Liu, Salehzadeh-Yazdi, Santos, Petrini, Milagre, Yu, Xie, Kroeze, Nesterova, Jansen, Xie, He, Reik, Stunnenberg. 2016b. Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol Cell* **62**:848–861. doi:10.1016/j.molcel.2016.04.025

Walter, Teissandier, Pérez-Palacios, Bourc'his. 2016. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife* **5**:1–30. doi:10.7554/eLife.11418

Wang C, Liu, Gao, Yang, Li, Liu, Chen, Kou, Zhao, Chen, Wang, Le, Wang, Duan, Zhang, Gao. 2018. Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat Cell Biol* **20**:620–631. doi:10.1038/s41556-018-0093-4

Wang, Liu, Tang, Archambeault, Li, Song, Tang, He, Matzuk, Wang. 2013. GASZ promotes germ cell derivation from embryonic stem cells. *Stem Cell Res* **11**:845–860. doi:10.1016/j.scr.2013.05.012

Wang, Mager, Chen, Schneider, Cross, Nagy, Magnuson. 2001. Imprinted X inactivation maintained by a mouse Polycomb group gene. *Nat Genet* **28**:371–375. doi:10.1038/ng574

Wang J, Wang, Feng, Wang, Li, Li, Liu, Jiao, Huang, Shi, Zhou, Chen, Liu, Li, Zhou. 2018. Asymmetric Expression of LincGET Biases Cell Fate in Two-Cell Mouse Embryos. *Cell* **175**:1887-1901.e18. doi:10.1016/j.cell.2018.11.039

Wang, Yu, Ming, Xia, Xu, Zhang, Zhang, Li, Huang, Xie, Zhu, Xie. 2020. Imprecise DNMT1 activity coupled with neighbor-guided correction enables robust yet flexible epigenetic inheritance. *Nat Genet*. doi:10.1038/s41588-020-0661-y

Wang, Zhang, Duan, Gao, Zhu, Lu, Yang, Zhang, Li, Ci, Li, Zhou, Aluru, Tang, He, Huang, Liu. 2014. Programming and inheritance of parental DNA methylomes in mammals. *Cell* **157**:979–991. doi:10.1016/j.cell.2014.04.017

Watson, Crick, Franklin. 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**:737–738. doi:10.1038/171737a0

Weber, Hellmann, Stadler, Ramos, Pääbo, Rebhan, Schübeler. 2007. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* **39**:457–466. doi:10.1038/ng1990

Weinberg, Papillon-Cavanagh, Chen, Yue, Chen, Rajagopalan, Horth, McGuire, Xu, Nikbakht, Lemiesz, Marchione, Marunde, Meiners, Cheek, Keogh, Bareke, Djedid, Harutyunyan, Jabado, Garcia, Li, Allis, Majewski, Lu. 2019. The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. *Nature* **573**:281–286. doi:10.1038/s41586-019-1534-3

Weinberg, Rosenbaum, Chen, Barrows, Horth, Marunde, Popova, Gillespie, Keogh, Lu, Majewski, Allis. 2021. Two competing mechanisms of DNMT3A recruitment regulate the dynamics of de novo DNA methylation at PRC1-targeted CpG islands. *Nat Genet*. doi:10.1038/s41588-021-00856-5

Weinberger, Ayyash, Novershtern, Hanna. 2016. Dynamic stem cell states: Naive to primed pluripotency in rodents and humans. *Nat Rev Mol Cell Biol* **17**:155–169. doi:10.1038/nrm.2015.28

Weismann. 1893. The Germ-Plasm: A Theory of Heredity.

Wen, Wu, Shinkai, Irizarry, Feinberg. 2009. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* **41**:246–250. doi:10.1038/ng.297

White. 1939. Potentially unlimited growth of excised plant callus in an artificial nutrient. *Am J Bot*.
Wicklow, Blij, Frum, Hirate, Lang, Sasaki, Ralston. 2014. HIPPO Pathway Members Restrict SOX2 to the Inner Cell Mass Where It Promotes ICM Fates in the Mouse Blastocyst. *PLoS Genet* **10**. doi:10.1371/journal.pgen.1004618
Wilda, Bächner, Zechner, Kehrer-Sawatzki, Vogel, Hameister. 2000. Do the constraints of human speciation cause expression of the same set of genes in brain, testis, and placenta? *Cytogenet Cell Genet* **91**:300–302. doi:10.1159/000056861
Williams, Hilton, Pease, Willson, Stewart, Gearing, Wagner, Metcalf, Nicola, Gough. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **336**:684–687. doi:10.1038/336684a0
Winther. 2001. August Weismann on Germ-Plasm Variation. *J Hist Biol Vol* 1–3. doi:10.1023/A:1012950826540
Wray, Kalkan, Gomez-Lopez, Eckardt, Cook, Kemler, Smith. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat Cell Biol* **13**:838–845. doi:10.1038/ncb2267
Wu, Dean. 2020. Maternal factors regulating preimplantation development in mice. *Curr Top Dev Biol* **140**:317–340. doi:10.1016/bs.ctdb.2019.10.006
Wu, Han, Gong, Sebastiano, Gentile, Singhal, Adachi, Fischedick, Ortmeier, Sinn, Radstaak, Tomilin, Schöler. 2013. Establishment of totipotency does not depend on Oct4A. *Nat Cell Biol* **15**:1089–1097. doi:10.1038/ncb2816
Wu B, Li, Li, Gao, Chen, Wei, Yang, Zhang, Li, Li, Wang, Surani, Li, Tang, Bao. 2020. Activin A and BMP4 Signaling Expands Potency of Mouse Embryonic Stem Cells in Serum-Free Media. *Stem Cell Reports* **14**:241–255. doi:10.1016/j.stemcr.2020.01.004
Wu K, Liu, Wang, He, Xu, Chen, Kuang, Liu, Guo, Li, Shi, Shen, Wang, Zhang, Wang, Pei, Chen. 2020. SETDB1-Mediated Cell Fate Transition between 2C-Like and Pluripotent States. *Cell Rep* **30**:25-36.e6. doi:10.1016/j.celrep.2019.12.010
Wu, Zhang. 2017. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* **18**:517–534. doi:10.1038/nrg.2017.33
Wu, Zhang, Xie. 2021. Methylome inheritance and enhancer dememorization reset an epigenetic gate safeguarding embryonic programs. *bioRxiv*. doi:10.1101/2021.10.10.463758
Wulansari, Sulistio, Darsono, Kim, Lee. 2021. LIF maintains mouse embryonic stem cells pluripotency by modulating TET1 and JMJD2 activity in a JAK2-dependent manner. *Stem Cells* 1–11. doi:10.1002/stem.3345
Xia, Xu, Yu, Yao, Xu, Ma, Zhang, Liu, Li, Lin, Chen, Li, Wang, Shi, Shi, Zhang, Song, Jin, Hu, Bu, Wang, Na, Xie, Sun. 2019. Resetting histone modifications during human parental-to-zygotic transition. *Science (80- )* **365**:353–360. doi:10.1126/science.aaw5118
Xiang, Tao, DiRusso, Hsu, Zhang, Xue, Pontis, Trono, Liu, Clark. 2022. Human reproduction is regulated by retrotransposons derived from ancient Hominidae-specific viral infections. *Nat Commun* **13**:463. doi:10.1038/s41467-022-28105-1
Xiang, Zhang, Xu, Zhou, Liu, Du, Zhang, Zhang, Wang, Gayen, Liu, Wang, Li, Wang, Kalantry, Li, Xie. 2019. Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. *Nat Genet*. doi:10.1038/s41588-019-0545-1
Xiong, Tolen, Choi, Velychko, Caizzi, Velychko, Adachi, MacCarthy, Lidschreiber, Cramer, Schöler. 2022. Oct4 differentially regulates chromatin opening and enhancer transcription in pluripotent stem cells. *Elife* **11**:1–28. doi:10.7554/eLife.71533
Xu, Li, Rong, Zhao, Wang, Dai, Chen, Liu, Gu, Liu, Guo, Shen, Wu, Lan. 2020. DNMT3A reads and connects histone H3K36me2 to DNA methylation. *Protein Cell* **11**:150–154. doi:10.1007/s13238-019-00672-y
Yaari, Katz, Domb, Harris, Zemach, Ohad. 2019. RdDM-independent de novo and heterochromatin DNA methylation by plant CMT and DNMT3 orthologs. *Nat Commun* **10**:1–10. doi:10.1038/s41467-019-09496-0
Yabuta, Kurimoto, Ohinata, Seki, Saitou. 2006. Gene Expression Dynamics During Germline Specification in Mice Identified by Quantitative Single-Cell Gene Expression Profiling1. *Biol Reprod* **75**:705–716. doi:10.1095/biolreprod.106.053686
Yagi, Kabata, Tanaka, Ukai, Ohta, Nakabayashi, Shimizu, Hata, Meissner, Yamamoto, Yamada. 2020. Identification of distinct loci for de novo DNA methylation by DNMT3A and DNMT3B during mammalian development. *Nat Commun* **11**. doi:10.1038/s41467-020-16989-w
Yamaguchi, Hong, Liu, Shen, Inoue, Diep, Zhang, Zhang. 2012. Tet1 controls meiosis by regulating

meiotic gene expression. *Nature* **492**:443–447. doi:10.1038/nature11709
Yamaji, Seki, Kurimoto, Yabuta, Yuasa, Shigeta, Yamanaka, Ohinata, Saitou. 2008. Critical function of Prdm14 for the establishment of the germ cell lineage in mice. *Nat Genet* **40**:1016–1022. doi:10.1038/ng.186
Yamaji, Ueda, Hayashi, Ohta, Yabuta, Kurimoto, Nakato, Yamada, Shirahige, Saitou. 2013. PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* **12**:368–382. doi:10.1016/j.stem.2012.12.012
Yamanaka, Lanner, Rossant. 2010. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**:715–724. doi:10.1242/dev.043471
Yan, Zhang, Fang, Liu, Wong, Gong, Liu, Yu, Liu, Jiang, Wang, Wei, Jia, Li, Sun, Tang, Zhou, Yuan, Li, Ma, Guan. 2021. PGC7 promotes tumor oncogenic dedifferentiation through remodeling DNA methylation pattern for key developmental transcription factors. *Cell Death Differ*. doi:10.1038/s41418-020-00726-3
Yang, Andrabi, Biss, Baker, Iqbal, Sharrocks. 2018. ZIC3 Controls the Transition from Naïve to Primed Pluripotency. *Cell Rep* **27**:3215-3227.e6. doi:10.2139/ssrn.3305317
Yang, Kalkan, Morissroe, Marks, Stunnenberg, Smith, Sharrocks. 2014. Otx2 and Oct4 Drive Early Enhancer Activation during Embryonic Stem Cell Transition from Naive Pluripotency. *Cell Rep* **7**:1968–1981. doi:10.1016/j.celrep.2014.05.037
Yang, Yu, Yu, Liang, Hu, Luo, Izsvák, Sun, Wang. 2022. Chemical-induced chromatin remodeling reprograms mouse ESCs to totipotent-like stem cells. *Cell Stem Cell* 1–19. doi:10.1016/j.stem.2022.01.010
Yao, Oh, Fuchs, Zhou, Ch'ng, Newsome, Bronson, Li, Livingston, Eckner. 1998. Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93**:361–372. doi:10.1016/S0092-8674(00)81165-4
Ye, Li, Tong, Ying. 2013. Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *EMBO J* **32**:2548–2560. doi:10.1038/emboj.2013.175
Yearim, Gelfman, Shayevitch, Melcer, Glaich, Mallm, Nissim-Rafinia, Cohen, Rippe, Meshorer, Ast. 2015. HP1 Is Involved in Regulating the Global Impact of DNA Methylation on Alternative Splicing. *Cell Rep* **10**:1122–1134. doi:10.1016/j.celrep.2015.01.038
Ying, Liu, Marble, Lawson, Zhao. 2000. Requirement of Bmp8b for the generation of primordial germ cells in the mouse. *Mol Endocrinol* **14**:1053–1063. doi:10.1210/mend.14.7.0479
Ying Q-L, Nichols, Chambers, Smith. 2003. BMP Induction of Id Proteins Suppresses Differentiation and Sustains Embryonic Stem Cell Self-Renewal in Collaboration with STAT3. *Cell* **115**:281–292. doi:10.1016/S0092-8674(03)00847-X
Ying QL, Stavridis, Griffiths, Li, Smith. 2003. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21**:183–186. doi:10.1038/nbt780
Ying, Wray, Nichols, Batlle-Morera, Doble, Woodgett, Cohen, Smith. 2008. The ground state of embryonic stem cell self-renewal. *Nature* **453**:519–523. doi:10.1038/nature06968
Ying, Zhao. 2001. Cooperation of endoderm-derived BMP2 and extraembryonic ectoderm-derived BMP4 in primordial germ cell generation in the mouse. *Dev Biol* **232**:484–492. doi:10.1006/dbio.2001.0173
Yokomine, Hata, Tsudzuki, Sasaki. 2006. Evolution of the vertebrate DNMT3 gene family: A possible link between existence of DNMT3L and genomic imprinting. *Cytogenet Genome Res* **113**:75–80. doi:10.1159/000090817
Yu, Ji, Dang, Sha, Yuan, Zhou, Yan, Qiao, Tang, Fan. 2016. Oocyte-expressed yes-associated protein is a key activator of the early zygotic genome in mouse. *Cell Res* **26**:275–287. doi:10.1038/cr.2016.20
Yu, Wei, Duan, Schmitz, Sakurai, Wang, Wang, Zhao, Hon, Wu. 2021. Blastocyst-like structures generated from human pluripotent stem cells, Nature. doi:10.1038/s41586-021-03356-y
Zalzman, Falco, Sharova, Nishiyama, Thomas, Lee, Stagg, Hoang, Yang, Indig, Wersto, Ko. 2010. Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* **464**:858–863. doi:10.1038/nature08882
Zemach, McDaniel, Silva, Zilberman. 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science (80- )* **328**:916–919. doi:10.1126/science.1186366
Zhang, Chen, Yin, Zhang, Racowsky, Zhang. 2019. Maternal-biased H3K27me3 correlates with paternal-specific gene expression in the human morula. *Genes Dev* **33**:382–387. doi:10.1101/gad.323105.118
Zhang, Gal-yam, Egger, Iniguez, Holster, Lin, Liang, Jones, Tanay, Sssi. 2008. Frequent switching of

Polycomb repressive marks and DNA hypermethylation in the PC3 prostate cancer cell line. *Proc Natl Acad Sci USA* **105**:12979–12984.

Zhang, Jurkowska, Soeroes, Rajavelu, Dhayalan, Bock, Rathert, Brandt, Reinhardt, Fischle, Jeltsch. 2010. Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic Acids Res* **38**:4246–4253. doi:10.1093/nar/gkq147

Zhang, Lai, Krupalnik, Guo, Guo, Zhou, Xu, Yu, Liu, Jiang, Li, Abdul, Ma, Li, Fu, Lv, Jiang, Tariq, Kanwal, Liu, Xu, Zhang, Huang, Wang, Chen, Babarinde, Luo, Wang, Zhou, Ward, He, Ibañez, Li, Zhou, Yuan, Feng, Arumugam, Di Vicino, Bao, Wu, Schambach, Wang, Sun, Gao, Qin, Hutchins, Doble, Hartmann, Cosma, Qin, Xu, Chen, Volpe, Chen, Hanna, Esteban. 2020. β-Catenin safeguards the ground state of mousepluripotency by strengthening the robustness of the transcriptional apparatus. *Sci Adv* **6**:eaba1593. doi:10.1126/sciadv.aba1593

Zhang, Lang, Zhu. 2018. Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol* **19**:489–506. doi:10.1038/s41580-018-0016-z

Zhang, Termanis, Özkan, Bao, Culley, de Lima Alves, Rappsilber, Ramsahoye, Stancheva. 2016. G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells. *Cell Rep* **15**:77–85. doi:10.1016/j.celrep.2016.03.007

Zhang Yu, Xiang, Yin, Du, Peng, Wang, Fidalgo, Xia, Li, Zhao, Zhang, Ma, Xu, Wang, Li, Xie. 2018. Dynamic epigenomic landscapes during early lineage specification in mouse embryos. *Nat Genet* **50**:96–105. doi:10.1038/s41588-017-0003-x

Zhang, Zhang, Acampora, Vojtek, Yuan, Simeone, Chambers. 2018. OTX2 restricts entry to the mouse germline. *Nature* 1. doi:10.1038/s41586-018-0581-5

Zhang Yunfang, Zhang, Shi, Tuorto, Li, Liu, Liebers, Zhang, Qu, Qian, Pahima, Liu, Yan, Cao, Lei, Cao, Peng, Liu, Wang, Zheng, Woolsey, Quilici, Zhai, Li, Zhou, Yan, Lyko, Zhang, Zhou, Duan, Chen. 2018. Dnmt2 mediates intergenerational transmission of paternally acquired metabolic disorders through sperm small non-coding RNAs. *Nat Cell Biol* **20**:535–540. doi:10.1038/s41556-018-0087-2

Zhao, Han, Chew, Liu, Chiu, Choo, Orlov, Sung, Shahab, Kuznetsov, Bourque, Oh, Ruan, Ng, Wei. 2007. Whole-Genome Mapping of Histone H3 Lys4 and 27 Trimethylations Reveals Distinct Genomic Compartments in Human Embryonic Stem Cells. *Cell Stem Cell* **1**:286–298. doi:10.1016/j.stem.2007.08.004

Zhao, Yao, Yu, Zhang, Xu, Chen, Sun, Zhu, Zhang, Qian, Ji, Pan, Zhang, Chen, Correia, Weiskittel, Lin, Zhao, Chandrasekaran, Fu, Zhang, Fan, Xie, Li, Hu, Zhang. 2021. Metabolic remodelling during early mouse embryo development. *Nat Metab*. doi:10.1038/s42255-021-00464-x

Zheng, Huang, Zhang, Xiang, Du, Xu, Li, Wang, Ma, Peng, Xu, Xie. 2016. Resetting Epigenetic Memory by Reprogramming of Histone Modifications in Mammals. *Mol Cell* **63**:1066–1079. doi:10.1016/j.molcel.2016.08.032

Zhu, Guo, Ren, Hou, Dong, Li, Lian, Fan, Hu, Gao, Wang, Wei, Liu, Yan, Ren, Yuan, Yuan, Yan, Wen, Yan, Qiao, Tang. 2018. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat Genet*. doi:10.1038/s41588-017-0007-6

Zhu, Wang, Qian. 2016. Transcription factors as readers and effectors of DNA methylation. *Nat Rev Genet* **17**:551–65. doi:10.1038/nrg.2016.83

Zuin, Roth, Zhan, Cramard, Redolfi, Piskadlo, Mach, Kryzhanovska, Tihanyi, Kohler, Eder, Leemans, van Steensel, Meister, Smallwood, Giorgetti. 2022. Nonlinear control of transcription through enhancer–promoter interactions. *Nature* **604**. doi:10.1038/s41586-022-04570-y

Zylicz, Dietmann, Günesdogan, Hackett, Cougot, Lee, Surani. 2015. Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *Elife* **4**:1–25. doi:10.7554/eLife.09571

## RÉSUMÉ

La méthylation de l'ADN est une marque épigénétique associée à la répression des gènes, dont le rôle central dans le développement est illustré par la létalité embryonnaire observée chez les souris sans méthylation de l'ADN. Durant la formation de l'épiblaste, les cellules souches embryonnaire murines (CSE) transitent d'un génome hypométhylé à hyperméthylé, alors qu'elles quittent la pluripotence naïve et s'engagent dans la pluripotence amorcée vers les lignages somatiques. En parallèle, les cellules primordiales germinales (CPG) vont, au contraire, perdre leurs profils de méthylation de l'ADN. Alors que le développement embryonnaire précoce représente une fenêtre idéale pour l'étude des liens entre dynamique de la méthylation de l'ADN et prise d'identité cellulaire, comment la méthylation impacte la transition de pluripotence, et l'adoption des différents lignages embryonnaires, notamment germinal, reste mal compris.

Au cours de ma thèse, j'ai utilisé des CSE dépourvues de méthylation de l'ADN (*Dnmt*-TKO) et des protocoles de différenciations *in vitro*, pour étudier le rôle de la méthylation dans la prise d'identité cellulaire. J'ai pu observer que si la méthylation semble dispensable pour la transition de pluripotence et l'induction du lignage somatique neurale, son absence étend la fenêtre de compétence à adopter un destin germinal. En se basant sur des analyses de cartographies chromatiniennes, je propose que la méthylation contrôle la temporalité des prises d'identité cellulaire neurale et germinale, en désengageant des éléments régulateurs de ces identités, évitant ainsi que ces deux lignées ne deviennent des voies de différentiation par défaut au cours du développement embryonnaire précoce.

## MOTS CLÉS

Méthylation de l'ADN, Pluripotence, Lignée germinale, Développement, Régulation de l'expression des gènes

## ABSTRACT

DNA methylation is an epigenetic mark associated with gene repression, playing a central role in development as demonstrated by the early lethality of mouse embryos lacking DNA methylation. During epiblast formation, embryonic stem cells (ESCs) transition from low to high DNA methylation levels, as they exit naïve pluripotency and get primed for somatic lineage commitment. In parallel, emerging primordial germ cells (PGCs) will undergo DNA methylation erasure. Whilst early development represents an ideal context to study the interplay between DNA methylation and cell fates, how DNA methylation impacts priming and early lineage induction, notably germline specification, remains unclear.

Here, we took advantage of DNA methylation-free ESCs (*Dnmt*-TKO) and *in vitro* differentiation to probe the role of DNA methylation in cell fate commitment. We found that DNA methylation is dispensable for priming and for neural induction, as TKO cells acquired functional primed features and could be engaged into neurogenesis. Strikingly, DNA methylation appeared unnecessary for adopting the germline fate, and its absence extended temporal competency for PGC specification. Based on our chromatin profiling results, we propose that DNA methylation allows multi-lineage diversification by decommissioning DNA methylation-sensitive neural and germline enhancers during epiblast formation, therefore tempering these fates as default differentiation routes.

## KEYWORDS

DNA methylation, Pluripotency, Germline, Development, Gene expression