



Dynamic of nuclear changes occurring during the conversion between naïve (ESCs) and primed (EpiSCs) pluripotent cells

Matteo Tosolini

► To cite this version:

Matteo Tosolini. Dynamic of nuclear changes occurring during the conversion between naïve (ESCs) and primed (EpiSCs) pluripotent cells. Subcellular Processes [q-bio.SC]. Université Paris Saclay (COMUE), 2016. English. NNT : 2016SACLS511 . tel-01552175

HAL Id: tel-01552175

<https://theses.hal.science/tel-01552175>

Submitted on 1 Jul 2017

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.

NNT: 2016SACLS511

THESE DE DOCTORAT
DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE PARIS-SUD

ECOLE DOCTORALE N° 577
Structure et dynamique des systèmes vivants (SDSV)
Spécialité de doctorat: Sciences de la Vie et de la Santé

Par

Mr. Matteo Tosolini

*Dynamique de la réorganisation nucléaire accompagnant la conversion
entre deux états pluripotents: l'état naïf (ESCs) et amorcé (EpiSCs)*

Thèse présentée et soutenue à Jouy-en-Josas, le 12 Décembre 2016:

Composition du Jury :

Dr. Sébastien Bloyer	Professeur, Université Paris-Sud	Président
Dr. Véronique Azuara	Professeur, Imperial College London	Rapporteur
Dr. Antoine H.F.M. Peters	Professeur, University of Basel	Rapporteur
Dr. Pablo Navarro-Gil	Chargé de Recherche, Institut Pasteur	Examineur
Dr. Amélie Bonnet-Garnier	Chargé de Recherche, INRA	Co-encadrant de thèse
Dr. Alice Jouneau	Chargé de Recherche, INRA	Directeur de thèse

*“Per quanto lontano scorra l’acqua,
non dimentica mai la sua fonte.”*

*“Je n'ai pas peur de la route,
faudrait voir,
faut qu'on y goûte.”*

*“When you talk, you are only repeating what you already know.
But if you listen, you may learn something new.”*

ACKNOWLEDGMENTS

First of all I would like to thank the members of the jury for accepting to read and evaluate my PhD manuscript and thesis defense. Many thanks to Dr. Véronique Azuara and Dr. Antoine H.F.M. Peters the reviewers of my manuscript and to the PhD examiners Dr. Sébastien Bloyer, Dr. Pablo Navarro-Gil and Dr. Nathalie Beaujean.

All my gratitude to my PhD supervisor, Dr. Alice Jouneau, for her true commitment to science, her profound knowledge and expertise in the ESCs and EpiSCs world and her infinite passion in her work, all of which she passed on to me in these years.

The word “thanks” is not enough to express my appreciation to my co-supervisor Dr. Amélie Bonnet-Garnier for her unconditional support, for believing in me and pushing me to always do my best. She introduced me to the world of epigenetic and heterochromatin and it has been a real pleasure to work with her.

A very big thank to Dr. Nathalie Beaujean for her constant support, even from far away. I really appreciated.

Many thanks to all the members of the BDR-ER1 EPEE group: first of all Dr. Veronique Duranthon, you always had a good word to say to make me smile. Dr. Pierre Adenot I will never thank you enough for the help you gave me, for all the constructive discussions at the building 212 and all the time you invested in me. Martine Chebrout and Tiphaine Aguirre-Lavin thank you for the help you provided me and all the tips you gave me. The list is very long: thanks to Eugenie Canon, Nathalie Daniel, Dr. Sophie Calderari, Nathalie Peynot, Catherine Archilla, Magali Monnoye, Dr. Fabienne Nuttinck, Linda Maulny, Dr. Juliette Salvaing, Claire Boulesteix, Renaud Fleurot, Silvie Ruffini, Ludivine Laffont, Dr. Brigitte Le Guienne and Dr. Alline De Paula Reis for the good times at coffee breaks and at the canteen, for the discussions and all the support you gave me in these three years. Special thanks to Dr. Laurent Boulanger for having repaired all the things that I broke during my PhD, I am so sorry!

Vincent, you deserve a thanking paragraph all for yourself. I do not know how I can thank you for all your help in these years: with the cell culture, qRT-PCR, western-blot, driving back and forth home-lab an infinite number of times and driving to IKEA to buy furniture for the apartment. I have no idea how I would have made all this without you, thanks a lot.

Many thanks to the people of the Student Room - Office 55 past and present, in particular Anne-Clemence Veillard, Sophie Veniel, Violette Navia, Luc Maillet, Melanie Bernard-

Cacciarella, Anaïs Carvalho and Delphine Dube. Delphine I owe you a special thank as you supported me from the beginning till the end. It has been a pleasure to do the PhD "adventure" side by side with you, sharing the good and bad moments and helping each other.

Thanks to all the students of the BDR unit: Maxime, Jean-Philippe, Mouna, Clara, Audrey P, Sofiane, Audrey L, Sarah, Lessly and all the others for the good times we have spent together.

I am grateful to you, Polina, as you helped me since the first day I arrived in the lab. You forced me to switch from English to French and you reintroduced me to volleyball, my favourite sport, after many years of not playing. I will never forget all of our discussions about science and much more while going from the lab to the Vauboyen station and during the RER C trips to Massy.

Thank you, to everyone in the BDR unit who has somehow helped me during my PhD.

Many thanks to the guys of the Monday's volley at lunchtime: you made me start each week in a so positive and smiling way!

I apologize to Dr. Pierre Adenot and Luc Jouneau for not including the work we have done together in this thesis. We have all invested so much time in my "side" project and I am so grateful to you, but for evident reasons I did not have the time to write this part here. I am confident we will pursue this work and publish it next year.

I would like also to thank the PhD committee that has followed me during these three years, especially the president Dr. Claire Rougeulle and all the other members: Dr. Anne Gabory, Dr. Christian Muchardt and Dr. Anselme Perrier for the scientific help in leading my project.

I thank the collaborators for the technical and scientific help: Dr. Claire Francastel, Dr. Guillaume Velasco and Giacomo Grillo.

I must thank all the people outside the lab that helped me through my life in France all along my PhD studies. Many thanks to the Italian friends of the "Magistère Européen de Génétique", in particular Maddalena, Giacomo & Francesca and Giulia, still here in Paris, as well as Boris and Debora that are overseas. I would have never made it without you guys!

Thanks to my volleyball team in Massy. Guys you have been so important for my integration in France, thank you for teaching me so many French words (official and unofficial), and thank you for helping me distress after the hard days in the lab. Aleksandar, Xavier, Boris, Thibault, Johnnie, Simon, Sebastian, and all the others: I will never forget you!

Special thanks the "Cratere team" for the nice indoor and summer volleyball tournament we made in these three years and also to the people of the beach-volley in "Supelec".

It is more than four years that I am not living in Italy but you are still there to discuss, hang out and have a good time together every time I am back in Udine. I am so grateful for the long-standing friendship and support from my friends of "La Banda": Alessandro, Luca, Stefano, Martina e Gaia. I miss you guys and I thank you for everything.

Last but not least, I am thankful to my family. I love you mum, dad and sister. I know that these years have been very hard for you, as well as for me. I really appreciate your long-distance unconditional support. There are not enough words to thank you. I know you are proud of my life choices even if these have brought me far from you.

TABLE OF CONTENTS

Abbreviations	5
INTRODUCTION.....	9
1 EPIGENETICS AND GENOME ORGANIZATION	9
1.1 Organization of the eukaryotic genome	9
1.2 Epigenetics	11
1.2.1 DNA methylation	11
1.2.1.1 DNA methylation machinery.....	13
1.2.2 Histone modifications	16
1.2.2.1 Active histone modifications	19
1.2.2.1.1 H3K4me3 and H3K9ac and their enzymes	19
1.2.2.2 Repressive histone modifications	19
1.2.2.2.1 H3K9me3, H4K40me3 and their enzymes	19
1.2.2.2.2 H3K27me3, H2AK119ub and their enzymes.....	21
1.2.3 Histones variants.....	23
1.2.4 Euchromatin and Heterochromatin.....	24
1.2.4.1 Constitutive heterochromatin: the repetitive sequences of mammalian genome.....	25
1.2.5 Non-coding RNAs (ncRNAs).....	28
1.2.5.1 Satellite non-coding RNAs.....	29
2 PLURIPOTENCY	32
2.1 Totipotency and Pluripotency	32
2.2 Early mouse embryo development.....	32
2.3 In vitro pluripotency in mouse	35
2.3.1 The core pluripotency factors: OCT4, SOX2 and NANOG.....	36
2.3.2 Signaling pathways in pluripotency	38
2.3.3 The naïve state of mouse pluripotency	40
2.3.3.1 Naïve-metastable mESCs in serum/LIF	40
2.3.3.2 Ground-naïve mESCs in 2i/LIF.....	42
2.3.4 The primed state of mouse pluripotency: EpiSCs in ActivinA/FGF2.....	45
2.3.5 Comparison between mouse states of in vitro pluripotency.....	48

2.4	In vitro human pluripotency	51
2.4.1	The primed state of human pluripotency: conventional hESCs	51
2.4.2	The naïve state of human pluripotency: naïve hESCs	52
3	HETEROCHROMATIN ORGANIZATION IN MOUSE PLURIPOTENCY	55
3.1	Chromatin organization during early mouse embryonic development	55
3.1.1	Dynamic organization of constitutive heterochromatin during early mouse embryo development	55
3.1.2	DNA methylation dynamics of heterochromatin in early mouse embryo development	58
3.1.3	H3K9me3 dynamics in early mouse embryo development.....	60
3.1.4	H3K27me3 dynamics in early mouse embryo development.....	60
3.1.5	Satellite non-coding transcription in early mouse embryo development	61
3.2	Chromatin plasticity in mouse in vitro pluripotency.....	63
3.2.1	Chromatin bivalency in naïve mESCs.....	64
3.2.2	Heterochromatin organization in mouse in vitro pluripotency.....	65
3.2.3	Plasticity of heterochromatin in mouse naïve pluripotency	68
	OBJECTIVES OF THIS THESIS	75
	MATERIALS AND METHODS	79
1	Cell culture	79
2	Western-blot	79
3	Immunostaining.....	80
4	DNA-FISH	81
5	In situ Proximity Ligation Assay (PLA)	82
6	Three-dimensional structured image acquisition and analysis.....	82
7	Southern-blot	83
8	qRT-PCR.....	83
9	Bioinformatics analysis of ChIP-Seq datasets for satellite repeats	85
	RESULTS.....	89
1	Global epigenetic organization in the different types of mouse pluripotent stem cells	89
1.1	EUCHROMATIN IN MOUSE PLURIPOTENCY	89
1.2	BIVALENCY IN MOUSE PLURIPOTENCY	91
1.3	HETEROCHROMATIN IN MOUSE PLURIPOTENCY	93
1.3.1	Heterochromatin domains are characterized by different epigenetic histone marks depending on the pluripotent stem cell type.....	93

1.3.2	Pericentromeric heterochromatin is subjected to an epigenetic switch in 2i-ESCs compared to serum ones and EpiSCs.	98
1.3.3	DNA methylation state of repetitive sequences reflects the pluripotent stem cell type..	100
1.3.4	2i condition leads to decondensation of the well-organized structure of the chromocenter	103
1.3.5	The transcriptional state of major and minor satellites depends on pluripotent cell type.....	105
1.3.6	Absence of Suv39h1/2 induces different phenotypes depending on the pluripotent stem cell type	107
1.3.7	Absence of DNA methylation has limited effects on satellite transcription	112
1.3.8	Reduced levels of H3K27me3 do not up-regulate satellite transcription	114
2	Conversion from naive to primed pluripotency in mouse.....	118
DISCUSSION		129
1	Uncoupling epigenetic state of naïve ESCs with transcription regulation of major and minor satellites	129
2	The problem of major satellite repeats quantification.....	131
3	Cross-talk between H3K27me3 and 5-meC in mouse pluripotency	132
4	Hypothetical SUV39H-independent H3K9me3 deposition at PCH in 2i-ESCs	133
5	EpiSC: a pluripotent cell with a somatic epigenetic state	134
6	Major and minor satellites sequences respond to different epigenetic pathways	135
7	Genome stability in ESCs	137
8	Does the epigenetic status of in vitro pluripotency reflect in vivo pluripotency of mouse embryo?	137
9	In vitro conversion: why such an inefficient process compared to in vivo?	139
10	Bivalent Domains	141
11	Nomenclature ambiguity: ground, naive, primed in mouse and human pluripotency	142
12	Concluding remarks of this thesis	144
RESUME SUBSTANTIEL DE LA THESE EN FRANÇAIS.....		149
1	INTRODUCTION.....	149
2	OBJECTIFS	151
3	RESULTATS	152
3.1	L'hétérochromatine est caractérisée par différentes modifications d'histones selon le type de cellules pluripotentes.	152
3.2	Faible niveau de méthylation de l'ADN au PCH dans les ESCs en 2i	154

3.3	Le PCH est décondensé dans les ESCs en 2i mais transcriptionnellement réprimé.....	155
3.4	L'absence des SUV39H1/2 induit des phénotypes différents en fonction du type de cellule pluripotente.....	156
3.5	L'absence de méthylation de l'ADN augmente le dépôt d'H3K27me3 mais a un effet limité sur la transcription des séquences satellites.....	159
3.6	La réduction des niveaux d'H3K27me3 n'induit pas une sur-expression des séquences satellites.....	160
4	DISCUSSION	162
4.1	Dialogue entre H3K9me3, H3K27me3 et 5-meC dans la pluripotence chez la souris.....	162
4.2	EpiSC: une cellule pluripotente avec un état épigénétique somatique.....	163
4.3	Découplage de l'état épigénétique des ESCs avec régulation de la transcription des satellites.....	163
	REFERENCES.....	167
	APPENDIXES	189

Abbreviations

2i	Two inhibitors
5h-meC	5-hydroxymethylcytosine
5-meC	5-methylcytosine
AP	Alkaline phosphatase
BCA	Bicinchoninic acid
BMP	Bone morphogenic factor
BSA	Bovine serum albumin
CAF-1	Chromatin assembly factor 1
CDM	Chemically define medium
CENP-A	Centromere protein A
cEpiSC	Converted epiblast stem cell
CH	Centromeric heterochromatin
ChIP	Chromatin immunoprecipitation
CpG	CG dinucleotide
DAPI	4',6-diamidino-2-phénylindole
DMEM	Dulbecco's Modified Eagle Medium
dn	double null
DNMT	DNA methyl transferase
EB	Embryoid body
EGA	Embryonic genome activation
Epi	Epiblast
EpiLC	Epiblast stem cell-like
EpiSC	Epiblast stem cell
ERK	Extracellular signal-regulated kinase
ESC	Embryonic stem cell
ExE	Extra embryonic endoderm
EZH2	Enhancer of Zeste
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FISH	Fluorescence in situ hybridization
GSK-3	Glycogen synthase kinase 3
H2AK119ub	Ubiquitylation on the lysine 119 of the histone H2A
H3K27me3	Trimethylation on the lysine 27 of the histone H3

H3K4me3	Trimethylation on the lysine 4 of the histone H3
H3K9ac	Acetylation of the lysine 9 of the histone H3
H3K9me3	Trimethylation on the lysine 9 of the histone H3
HMT	Histone methyl transferase
HP1	Heterochromatin protein 1
ICM	Inner cell mass
iPSC	induced pluripotent stem cell
KDM	Lysine demethylase
LIF	Leukemia inhibitory factor
MAPK	Mitogen-activated protein kinase
MEF	Mouse embryonic fibroblast
miR	micro RNA
ncRNA	Non-coding RNA
NOD	Non-obese diabetic
NPB	Nucleolar precursor body
NPC	Neural progenitor cell
NuRD	Nucleosome Remodeling Deacetylase
PCH	Pericentromeric heterochromatin
PFA	Paraformaldehyde
PLA	Proximity ligation assay
PRC	Polycomb repressive complex
PrE	Primitive endoderm
PTGS	Post transcriptional gene silencing
qRT-PCR	Quantitative retro-transcribed polymerase chain reaction
RIPA	Radioimmunoprecipitation assay buffer
RNAi	RNA interference
SCID	Severe combined immunodeficiency
STAT	Signal Transducer and Activator of Transcription
SUV39H	Suppressor of variegation
TBS	Tris-borate saline
TE	Trophoectoderm
TET	Ten eleven translocation
TGF β	Transforming growth factor beta
TKO	Triple knock-out
WGCNA	Weighted Gene Co-expression Network Analysis

INTRODUCTION

INTRODUCTION

1 EPIGENETICS AND GENOME ORGANIZATION

1.1 Organization of the eukaryotic genome

Eukaryotic cells have to pack their immense genome made by a long linear DNA molecule into the confined volume of their nucleus. This problem is solved by wrapping the DNA around proteins called histones in order to form the nucleosome which is the structural unit of the chromatin (DNA-protein complex). One nucleosome is composed by two turns of double-helix DNA (146bp) around the core histone octamer made by two copies of heterodimers of histones H2A - H2B, and a tetramer of H3 and H4 (Figure 1) (Luger et al., 1997; Rhodes, 1997). Histones have a globular portion and a tail that is extruding from the nucleosomal core.

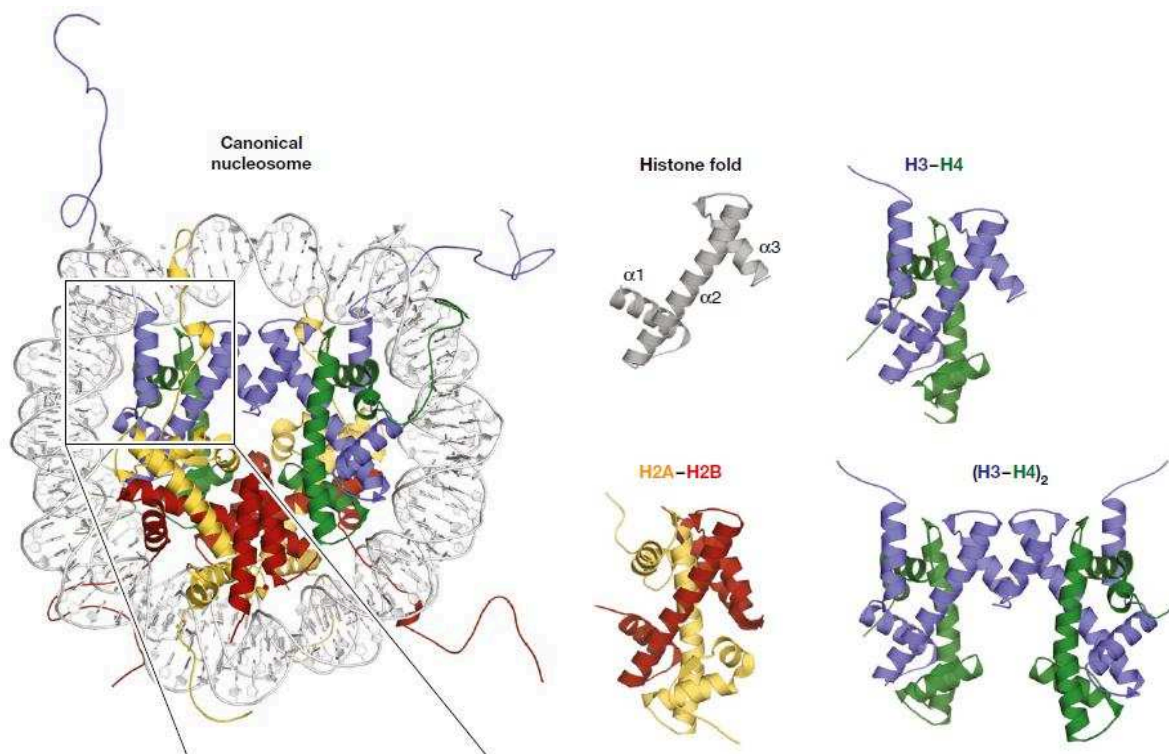


Figure 1: Nucleosome core and tails (From Mattioli et al., 2015)

Canonical nucleosome structure made by histone H3 (blue), H4 (green), H2A (yellow), H2B (red), and DNA (white). Tails of each histone are extruding from the nucleosome core.

Two successive nucleosomes are separated by 20-80bp of DNA linker (Routh et al., 2008) giving rise to the 10nm-fiber (bead on a string structure). The histone H1 is the only one that does not take part in the nucleosome core and its role is bringing together linkers DNA between nucleosomes creating the 30nm-fiber, which is the one found in interphase. The chromatin could be then further compacted till the maximum condensation of mitotic chromosomes (Figure 2) (Happel and Doenecke, 2009; Izzo et al., 2008).

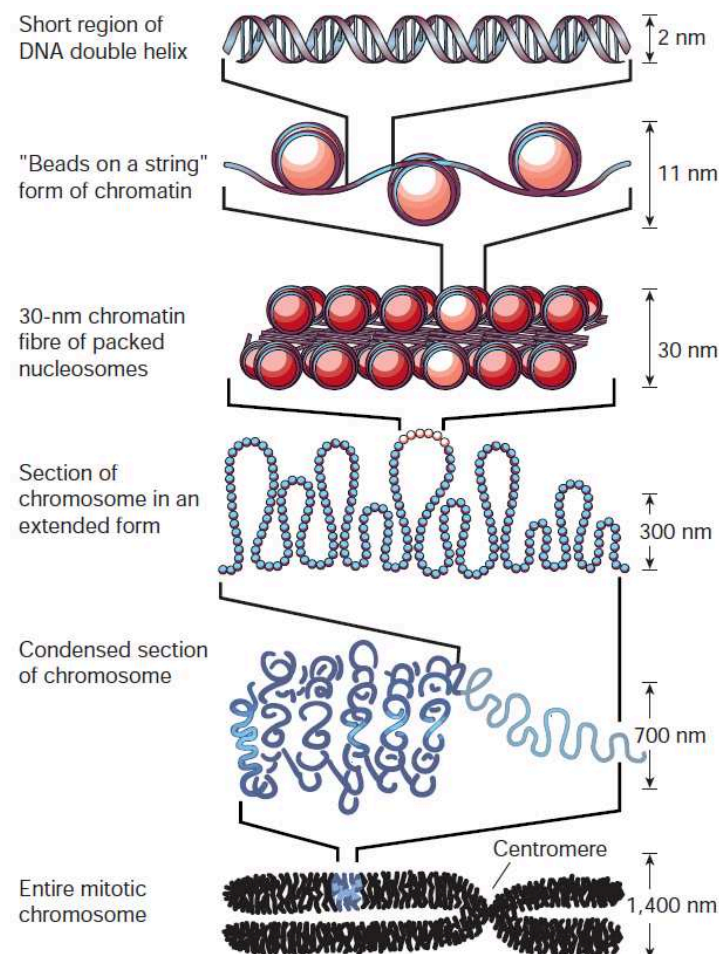


Figure 2: Organization of the eukaryotic genome (From Felsenfeld and Groudine, 2003)

In eukaryotes the double helix of DNA is wrapped on nucleosomes which are separated one another by a stretch of DNA linker. Successively the string of nucleosomes is folded into a fiber about 30 nm in diameter. These fibers are then further folded into higher-order structures till the hyper-condensation of a mitotic chromosome.

1.2 Epigenetics

The term Epigenetics has been coined by Conrad Waddington in 1942 to describe “processes by which genotype gives rise to phenotype”. In more recent times epigenetics has been defined as mitotically and/or meiotically heritable changes in gene functions that do not entail changes in DNA sequence and in the absence of the factor responsible for these changes (Wu Ct and Morris, 2001). However the “gene” concept in epigenetics must be extended also to non-coding sequences.

There are many mechanisms by which epigenetics can act, the first described is DNA methylation where a methyl-group is added and covalently bound to nucleotides principally on cytosines. Others are the covalent modification of histones, in particular on C- or N-tails, or the substitution of histone in toto (histone variants), both influencing the DNA indirectly as its helix is wrapped on the protein of the histone core. Finally a non-negligible portion of epigenetic inheritance and maintenance is mediated by RNAs and particularly non-coding RNAs (ncRNA) (Allis and Jenuwein, 2016).

1.2.1 DNA methylation

DNA methylation is an epigenetic mark that consists in the covalent binding of a methyl group on the carbon 5 of the pyrimidine ring of cytosine producing a 5-methylcytosine (5-meC) using as a donor S-adenosyl methionine (SAM) (Figure 3). This mark seems to be inherited through cell divisions via a mechanism that recognizes an hemimethylated palindrome CpG and induces the DNA methylation in the newly synthesized strand to obtain a fully methylated CpG (Bird, 2002). These CpGs are not randomly distributed into the genome but enriched in specific short regions (around 1kb) called CpG Islands found near the majority of vertebrate genes (Deaton and Bird, 2011). CpG sites are not the only one susceptible to DNA methylation, particularly in embryonic stem cells (ESCs) compared to somatic tissue there is a significant presence of CpA methylated sites and to a lesser extent CpT (Ramsahoye et al., 2000).

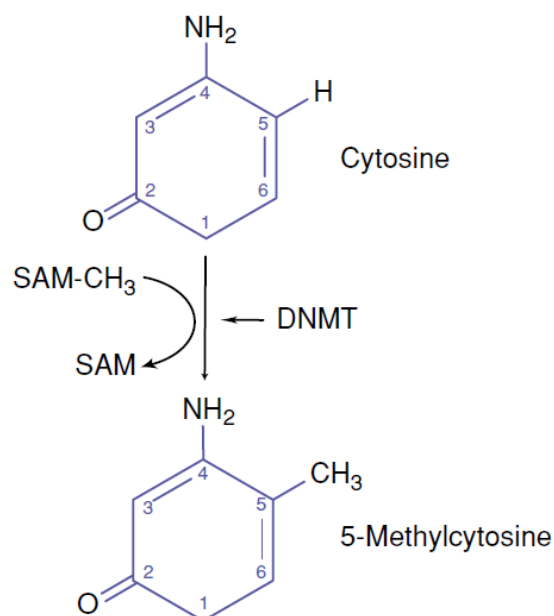


Figure 3: The mechanism of DNA methylation (From Strathdee and Brown, 2002)

5-Methylcytosine is produced by the action of the DNA methyltransferases (DNMT1, 3A or 3B), which catalyze the transfer of a methyl group (CH_3) from S-adenosylmethionine (SAM) to the carbon-5 position of cytosine.

DNA methylation has been historically associated with transcriptional repression but the position of the DNA methylation in the sequence of the transcriptional unit strongly influences the final effect (Jones, 2012). DNA methylation in close proximity of the transcription start site (TSS) or promoter region is known to repress the initiation of transcription (Kass et al., 1997), but when present in the gene body the effect is opposite and it could even stimulate the elongation of transcription (Jones, 1999). Moreover CpG methylation in the gene body could also influence the splicing (Jones, 2012). DNA methylation plays also key role in the stability of the genome allowing correct chromosomal segregation during mitosis when present on centromeric repeat sequences or suppressing the expression of transposable elements (Jones, 2012).

The majority of CpGs islands are unmethylated, however CpGs islands within promoters are methylated for stable repression of the corresponding gene. This is the case of genes subjected to imprinting, inactivated X chromosome- and germ-cell lineage-genes that undergo DNA methylation to suppress their inappropriate expression. DNA methylation seems to be the final lock of already silenced genes (Jones, 2012).

Methylated DNA is recognized by specific proteins with Methyl Binding Domains, such as MBD1 and MBD3 (Wade and Wolffe, 2001) and methyl-CpG binding protein 2 (MeCP2) (Fuks et al., 2003) that promote a compacted heterochromatic state and prevent from transcription factor bindings (Trojer and Reinberg, 2007).

1.2.1.1 DNA methylation machinery

Mammals have three different DNA methyltransferases (DNMTs) which catalyze the methylation (Okano et al., 1999) and one enzymatic inactive cofactor Dnmt3-like (Bourc'his et al., 2001) (Figure 4). DNMT3A and DNMT3B are the de novo DNMTs which are able to methylate the naked DNA in concert with the cofactor DNMT3L, thus establishing the DNA methylation pattern during development (Denis et al., 2011). Dnmt3a encodes for two isoforms DNMT3A1 and DNMT3A2 with the second being shorter at the N-terminal part than the first one and specifically interacting with DNMT3L at heterochromatin foci in ESCs (Nimura et al., 2006). The DNA methylation pattern is maintained through cell divisions by DNMT1 and it copies the methylation from hemimethylated DNA during replication (Cheng and Blumenthal, 2008).

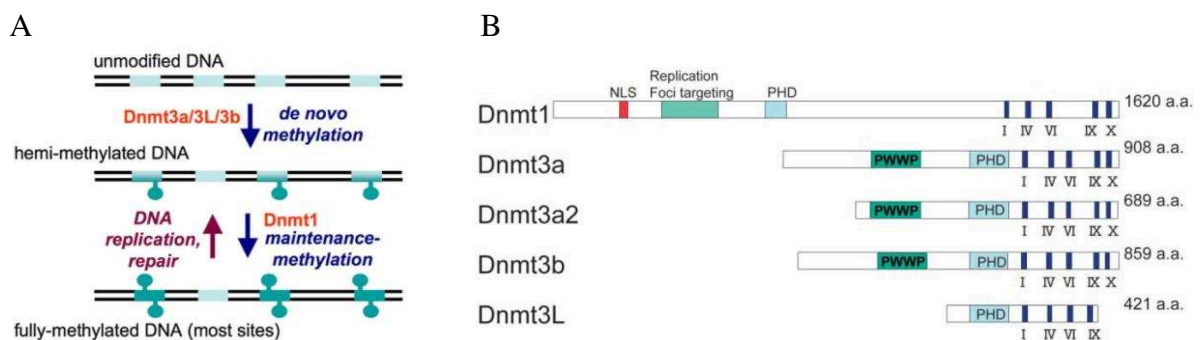


Figure 4: DNMTs, their functions and structures (From Cheng and Blumenthal, 2008 and Nimura et al., 2006)

A: de novo methylation is firstly established on unmodified CpG by DNMT3A and DNMT3B in concert with their cofactor DNMT3L. The maintenance role of DNMT1 is then to fully methylate the hemi-methylated sites and to maintain this state through DNA replication and DNA repair. B: schematic representation of the DNMT family in mouse. NLS (nuclear localization signal), PWWP (PWWP DNA- and protein-binding domain), PHD (the cysteine-rich PHD zinc-finger domain), I, IV, VI, IX and X correspond to conserved methyltransferase catalytic motifs. Of note DNMT3L lacks motif X.

To accomplish this process DNMT1 localizes in the replication fork and interacts with Proliferative Cell Nuclear Antigen (PCNA) and UHRF1 which binds hemimethylated DNA (Sharif et al., 2007). However the maintenance role of DNMT1 alone is not sufficient as

Dnmt3a^{-/-} and Dnmt3b^{-/-} embryonic stem cells (ESCs) gradually lose their DNA methylation pattern after several passages (Chen et al., 2003; Jackson et al., 2004). In the absence of Dnmt1, ESCs show reduced global level of 5-meC, but some sequences remained fully methylated such as major satellites (Arand et al., 2012). These observations showed that the separation in de novo and maintenance function of DNMTs is not as strict as it seems, as they can partially compensate each other. All the enzymatic DNMTs are strictly necessary for mammalian development (Li et al., 1992; Okano et al., 1999), as only Dnmt3l knock-out mice are viable even though males are sterile (Bourc'his et al., 2001). Indeed Dnmt1^{-/-} embryos die at E8.5-9.0 and showed only one third of 5-meC compared to wild-type condition (Li et al., 1992). Dnmt3b mutant embryo develop even further but after E9.5 they show multiple defects and do not go to term, while Dnmt3a^{-/-} develop to term but die one month after (Okano et al., 1999) (Figure 5).

DNA methylation is not an irreversible epigenetic mark and can be reverted through two principal ways: passive by replication and active via the action of specific enzymes. Demethylation could simply occur by the lack of maintenance after DNA replication inducing a dilution effect and a progressive loss of methylation after numerous rounds of cell division (Hill et al., 2014). However an active enzymatic mechanism of demethylation has been discovered in recent years thanks to the action of the Ten Eleven Translocation enzymes (TETs) which catalyze the oxidation of 5-methylcytosine (5-meC) into 5-hydroxymethylcytosine (5h-meC) (Tahiliani et al., 2009). 5h-meC could be then transformed in 5-formylcytosine (5-fC) also by TETs enzymes and subsequently in 5-carboxylcytosine (5-caC). 5-caC finally can enter in the Base-Excision Repair process (BER), removed by Thymine DNA Glycosidase and replaced by new unmodified cytosine (He et al., 2011; Ito et al., 2011). Function of 5h-meC is still controverted as it could be considered simply as a short-lived entity or as an epigenetic modification on its own. Its genome profiling shows a distinct distribution compared to 5-meC and it is associated to gene transcription (active promoters) as well as gene silencing (Polycomb-mediated) (Branco et al., 2012).

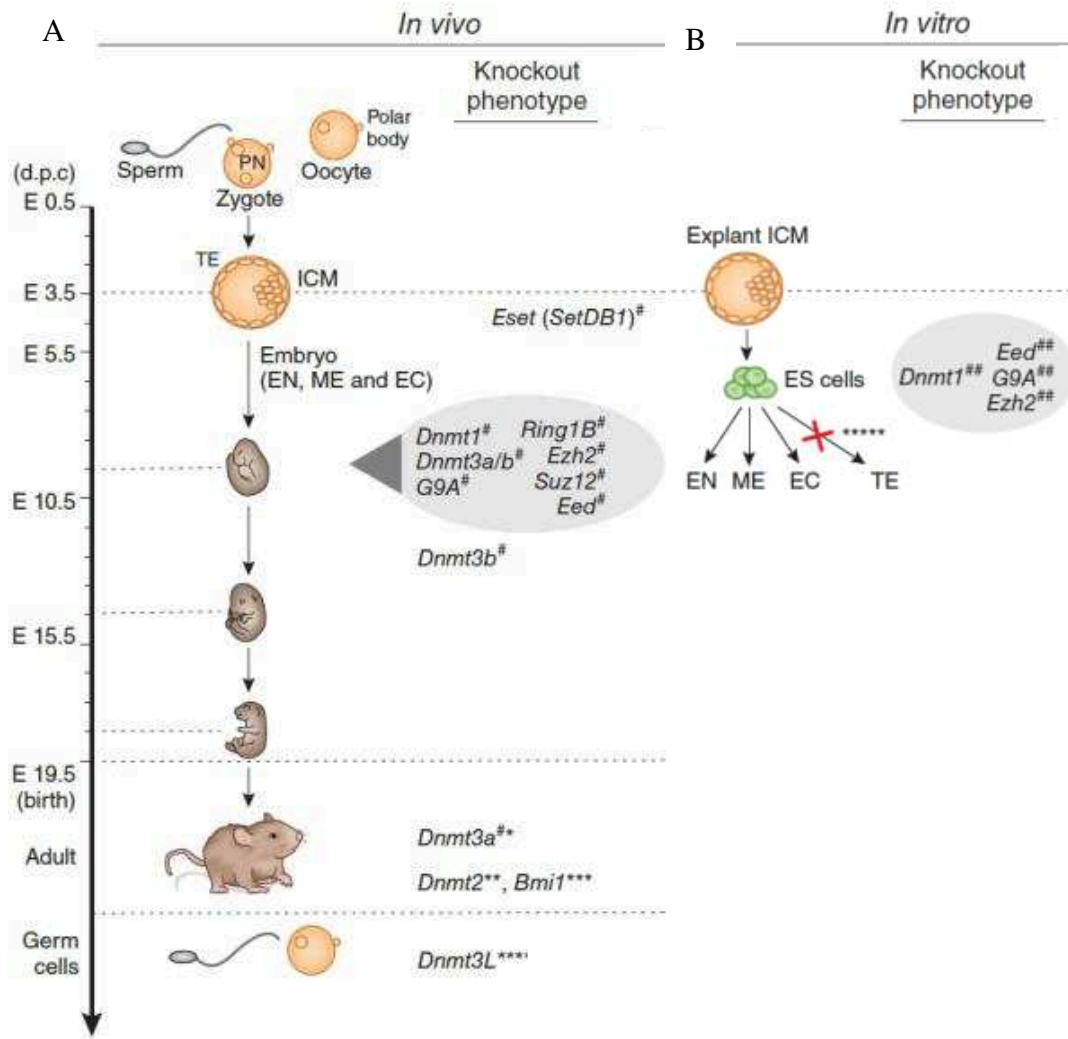


Figure 5: DNMTs and histone modifiers knock-out effects in mouse development and ESCs (From Meissner, 2010)

A: Stage of embryonic lethality of knock-out in vivo embryos for different epigenetics modifiers. B: in vitro knock-out ESCs for different epigenetics modifiers.

(lethal), ## (normal ESCs maintenance, but differentiation defects), * (Dnmt3a knockout mice die around 3 weeks postnatal and are smaller/runted), ** (No observed phenotype), *** (Mice are viable, but have hematopoietic and neural abnormalities), **** (Homozygous mice are sterile, offspring of homozygous female mice and heterozygous crosses show imprinting defects and die), ***** (Wild-type ES cells cannot differentiate into trophectodermal cells), d.p.c. (days post coitum), PN (pronuclei), EN (endoderm), ME (mesoderm), EC (ectoderm), TE (trophectoderm), ICM = inner cell mass.

1.2.2 Histone modifications

Another mechanism by which the epigenetics can act is through post-translational covalent modification of histone tails that are extruding from the core octamer. Many different post-translational modifications of histones exist, mainly methylation, acetylation, phosphorylation and ubiquitylation (Kouzarides, 2007) (Figure 6A). The chemical group attached, the number of these modifications, the position in the histone tail and the type of residues accepting the modification could lead to different effects (Jenuwein and Allis, 2001; Strahl and Allis, 2000). Histone modifications are reversible and dynamic marks under the control of chromatin "writers", that establish the modification catalyzing the deposition of the chemical group, and "erasers", which conversely remove the mark. The function and the "message" of an histone modification are then transmitted by chromatin "readers" that have specific protein binding motifs in their aminoacid sequence that recognize and bind to specific histone modifications (Figure 6B) (Allis and Jenuwein, 2016; Yun et al., 2011). Altogether the sequential combinations of one or more histone modifications, that are read by proteins influencing down-stream events, form the so called histone code or histone language (Strahl and Allis, 2000). "Readers" can have different function such as architectural proteins, chromatin remodelers and modifiers (Figure 7) (Yun et al., 2011).

Histone methylation could occur on arginine (R) and on lysine (K), but while arginines can only accept up to two methyl groups, lysines can be even tri-methylated (Greer and Shi, 2012). The function of histone methylation is strictly dependent on the position in the tails and could even have opposite effect, for example tri-methylation on lysine 4 of the histone H3 (H3K4me3) and H3K36me3 are active chromatin marks that promote the transcription, while H3K27me3, H3K9me3 or H4K20me3 are repressive marks leading to transcriptional silencing. Histone methylation is deposited by writers called histone methyl transferases (HMTs) and more specifically lysine methyl transferases (KMTs), while the erasers are histone lysine demethylases (KDM). The readers for histone methylation are proteins containing the chromodomain (CD) binding site (Sims et al., 2003). CD-proteins can recognize different type of methylated histone and generally recruits, in collaboration with RNAs and DNA-binding proteins, other proteins forming a larger complex (Tajul-Arifin et al., 2003).

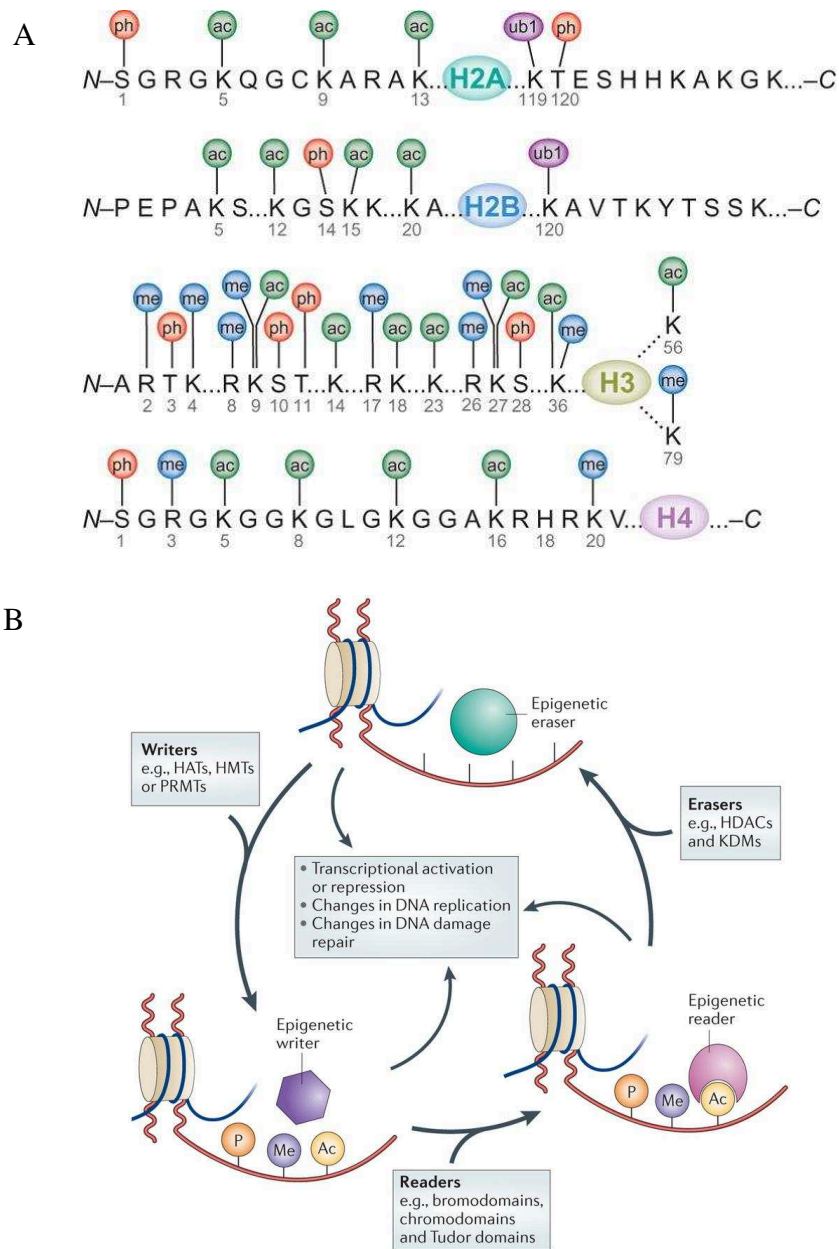


Figure 6: Histone code and chromatin modifiers: Writers, Erasers and Readers (From Advanced BioDesign and Falkenberg and Johnstone, 2014)

A: possible modification on the different residues of the tails of each histone. S (Serine), K (Lysine), R (Arginine), T (Threonine). Ph (Phosphorylation), ac (acetylation), ub1 (ubiquitylation), me (methylation). B: dynamics of epigenetic regulations: chromatin writers deposit the histone mark which is then recognized by chromatin readers or eventually reversed by chromatin erasers.

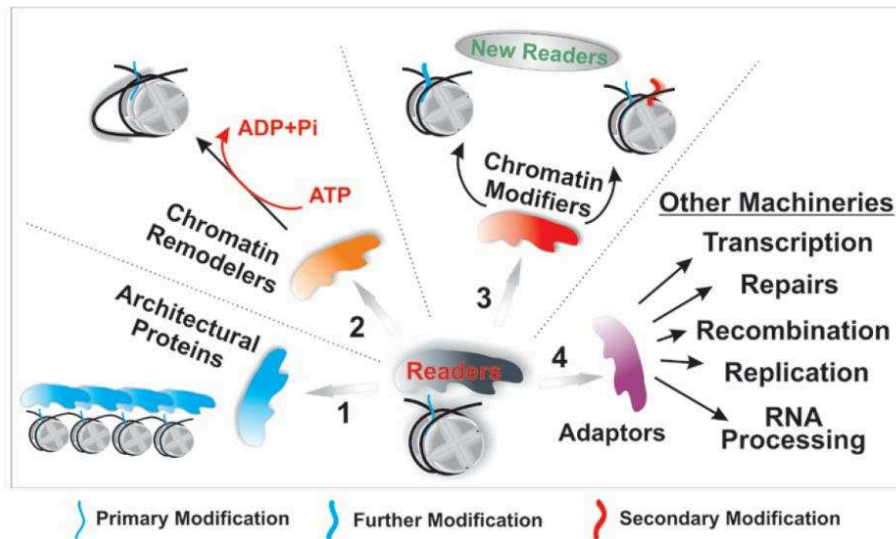


Figure 7: Chromatin readers functions (From Yun et al., 2011)

Chromatin readers can have different roles, acting as: architectural protein (inducing compaction for example), remodeling (increasing accessibility of nucleosomal DNA, energy dependant (use of ATP), chromatin modifiers (adding a new mark) or adaptors (recruiting other protein machineries such as transcription factors).

Histone acetylation occurs on lysine residues and it induces the neutralization of the positive charge of histone tail. This leads to decompaction of the nucleosomal structure by a less intimate interaction with the negatively charged DNAs. For that reason hyperacetylation of histones is generally associated with a more decondensed and open conformation of the chromatin which is more prone to active transcription of genes due to increased accessibility of the transcriptional machinery to the DNA. Conversely hypoacetylated regions correlate with compact and silent domains (Sterner and Berger, 2000). The writers in this case are the histone acetyltransferases (HATs) while the erasers are the histone deacetylases (HDACs) (Berndsen and Denu, 2008). Readers must contain a bromodomain to recognize histone acetylation (Kouzarides, 2007). Bromodomain proteins play a key role anchoring complexes to the acetylated chromatin in order for example to remodel chromatin and recruit transcription factors (Josling et al., 2012).

Histone ubiquitylation is a less common epigenetic mark that consists in the binding of the ubiquitin (a polypeptide of 76-amino acids) to lysines via the successive action of E1, E2 and E3-ligases. The most studied is the monoubiquitylation of the lysine 119 of the histone H2A (H2AK119ub) which is linked to gene silencing (Wang et al., 2004; Zhang, 2003).

1.2.2.1 Active histone modifications

1.2.2.1.1 H3K4me3 and H3K9ac and their enzymes

Trimethylation of lysine 4 on histone H3 (H3K4me3) is the hallmark of actively transcribed genes as it is specifically enriched at the 5' region of these genes (promoter), while H3K4me2 is more distributed along the gene body (Martin and Zhang, 2005). However H3K4me3 when present in association to H3K27me3 at the promoter level (bivalent domain) leads to the loading of the RNA Polymerase II on the gene promoter that stalled in a paused condition without elongating the transcription (Azuara et al., 2006; Bernstein et al., 2006). H3K4 is principally methylated by the KMT called mixed-lineage leukemia 1 (MLL1) which is a SET-domain protein belonging to the Tritorax complex (TrxG). Mll expression is necessary for the correct regulation of the Hox genes during embryo development, in fact Mll^{-/-} embryo died early during development (Schuettengruber et al., 2011; Terranova et al., 2006).

Another active histone modification is the acetylation of the lysine 9 on the histone H3 (H3K9ac) present at open regions of the genome that are sensitive to DNase I and enriched in active transcription factors (Hezroni et al., 2011). H3K9ac is enriched at the TSS region of active genes, highly correlating with H3K14ac, H3K27ac, H3K4me2 and H3K4me3 (Hezroni et al., 2011; Karmodiya et al., 2012). In mouse the main HATs responsible for H3K9ac is GCN5/PCAF as its deletion dramatically reduces the global level of H3K9ac with little or no effect on the level of H3K14ac or other histone H3 and H4 acetylations (Jin et al., 2011). On the other hand HDAC1 is the main enzyme responsible for deacetylation of histones and Hdac1^{-/-} leads to embryonic lethality before E10.5 with embryos presenting proliferation defects and retard in development (Lagger et al., 2002). Moreover mutant ESCs showed reduced proliferation rates that cannot be compensated by the expression of HDAC2 and HDAC3 (Lagger et al., 2002).

1.2.2.2 Repressive histone modifications

1.2.2.2.1 H3K9me3, H4K40me3 and their enzymes

Trimethylation of the lysine 9 of the histone H3 (H3K9me3) is found from fission yeast to human on repeat-rich centromeric, pericentromeric and telomeric regions (Peters et al., 2001), but also in block of tissue-specific genes (Becker et al., 2016). It is thought that H3K9me3 protects clusters of repetitive genes and non-coding repeats from illicit recombination,

suppressing as well their transcription. This histone modification generally prevents the binding of transcription factors and this is probably why H3K9me3-chromatin blocks are the last to be reprogrammed during induced pluripotent stem cell (iPSC) generation being the less accessible (Becker et al., 2016). All H3K9-KMTs have a SET-domain and can be divided into two groups: the first comprises G9a (Ehmt2) and GLP (Ehmt1) that are the KMTs necessary to catalyze H3K9me1 and H3K9me2, while the trimethylation is achieved by the second group composed by SET domain bifurcated 1 (SETDB1), SUV39H1 and SUV39H2 (Mozzetta et al., 2015). This repressive mark is known to induce heterochromatin compaction and spreading via the recruitment of heterochromatin proteins 1 (HP1), thanks to the chromoshadow domain (Bannister et al., 2001; Lachner et al., 2001). In mammals there are three HP1 proteins: α , β and γ (Jones et al., 2000). HP1 α and β seem to share the same function accumulating together over H3K9me3 (Figure 8), while HP1 γ has a diffuse genome localization (Dialynas et al., 2007). HP1 α and β can self-oligomerize and recruit other repressive machineries like DNMTs to depose DNA methylation and SUV4-20H2 that specifically catalyze H4K20me3, another repressive histone modification, inducing further more compaction and repression (Figure 8; Schotta et al., 2004; Wongtawan et al., 2011).

In mouse SUV39H KMTs are encoded by two gene loci *Suv39h1* and *Suv39h2* which are both expressed in embryogenesis. *Suv39h* double null (*Suv39h^{dn}*) condition impairs severely the viability of mice (which are growth retarded and infertile), inducing chromosomal instability and increased risk of tumorigenesis. However mice deficient for either *Suv39h1* or *Suv39h2* are fertile and showed normal viability showing redundant functions of the two enzymes (Peters et al., 2001). Interestingly *Suv4-20h2* *-/-* mice have no apparent defects and develop normally however *Suv4-20h^{dn}* display peri-natal lethality and are smaller (Schotta et al., 2008).

SETDB1 is described as the principal KMT responsible for H3K9me3 in the genome outside centromeric, pericentromeric and telomeric repeats (Schultz et al., 2002). However its loss in mice is associated with a substantial reduction of H3K9me3 also at pericentromeric heterochromatin (Mozzetta et al., 2015). It is not clear if this reduction is a result of SETDB1 mono-methyltransferase activity necessary for successive trimethylation by SUV39H1 and SUV39H2 or potentially direct trimethyltransferase activity of SETDB1 at pericentromeric repeats.

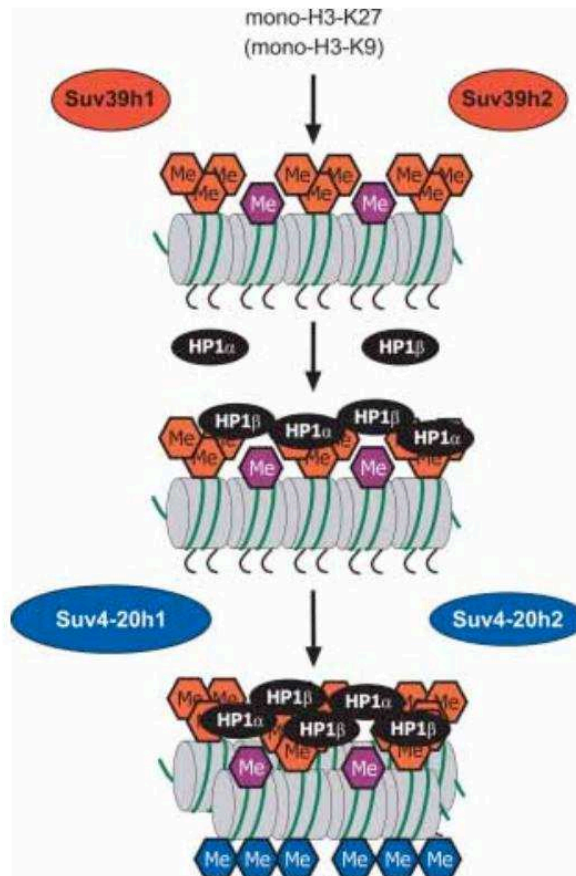


Figure 8: Conserved pathway SUV39H1/2-HP1-SUV4-20H2 (From Schotta et al., 2004)

Proposed model of sequential induction of H3K9me3 deposition by SUV39H1/2. H3K9me3 is then recognized by HP1 α and β inducing compaction and recruiting SUV4-20H2 that finally deposits H4K20me3 condensing furthermore the chromatin.

1.2.2.2.2 H3K27me3, H2AK119ub and their enzymes

Trimethylation of the lysine 27 on the histone H3 (H3K27me3) is another repressive histone mark present principally in tissue-specific gene regions and on the inactive X chromosome (Boyer et al., 2006; Margueron and Reinberg, 2011). H3K27me3 is not enriched in focal peak but on larger and broader regions in the genome, and these H3K27me3 blocks are negatively correlated with transcription such as the Hox cluster in differentiated cells (Pauler et al., 2008). Interestingly during iPSCs derivation H3K27me3-block are reprogrammed earlier than H3K9me3 ones and in general H3K27me3 and H3K9me3 blocks are largely exclusive (Becker et al., 2016).

Polycomb group proteins (PcGs) were originally identified as important regulators of developmentally related genes like the Hox cluster in *Drosophila*. In mammals there are two PcG called Polycomb repressive complex 1 and 2 (PRC1 and PRC2, Figure 9). Enhancer of

zeste homolog 2 (EZH2) is the catalytic subunit of PRC2, which contains a SET domain and it is thought to be the only KMT for di- and tri-methylation of H3K27 (Trojer and Reinberg, 2007). However in *Ezh2*^{-/-} ESCs a residual H3K27me3 is found in the genome and it is likely due to EZH1 that seems to partially complement the absence of EZH2 (Margueron et al., 2008; Shen et al., 2008). *Ezh2* transcription is up-regulated after fertilization and it is highly expressed all along the pre-implantation development. *Ezh2*-null condition is lethal at early stages of mouse development as these embryos die between pre- and post-implantation development (Figure 5) (Becker et al., 2016). *Ezh2*-null blastocysts have an impaired potential to outgrowth preventing the establishment of mutant ESCs (O'Carroll et al., 2001). The other PcG complex is PRC1 that mediates via, its catalytic subunit RING1B (E3-ubiquitin ligase), the ubiquitylation of lysine 119 of histone H2A (H2AK119ub) which is also a repressive mark. Mouse RING1B is coded by the *Rnf2* gene which when ablated causes an arrest during gastrulation with developmental defects occurring in both embryonic and extra-embryonic tissues (Voncken et al., 2003) (Figure 5).

PcG complexes are composed by different combination of subunits. The most common and studied PRC1 is called the canonical PRC1 (cPRC1) and it is recruited on PRC2 targets thanks to Cbx subunit that recognize H3K27me3 (Cao et al., 2002). However more recently it has been discovered that a variant or non-canonical PRC1 (vPRC1 or ncPRC1) complex could be firstly recruited at unmethylated CpG island by KDM2B and only in a second time PRC2 is enrolled via H2AK119ub recognition (Blackledge et al., 2014; He et al., 2013).

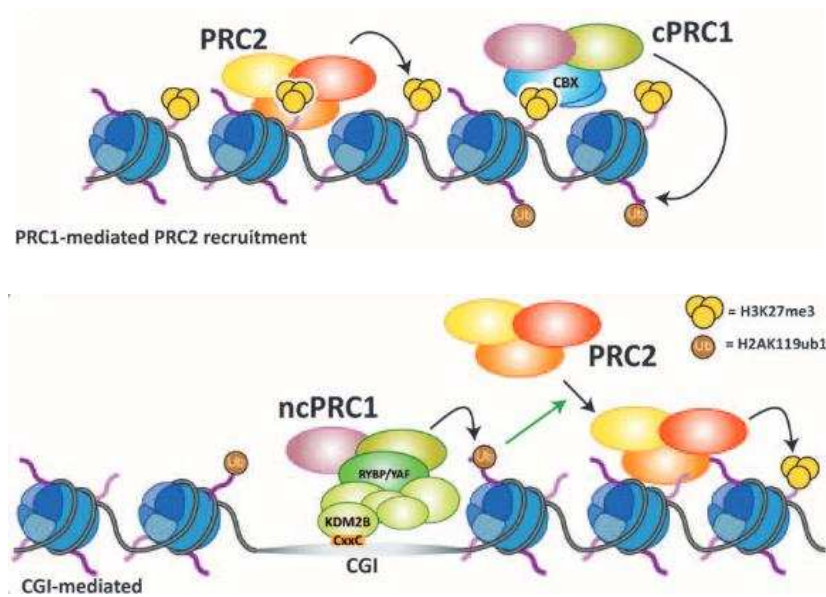


Figure 9: Mechanism of PcG recruitment to chromatin (From Aranda et al., 2015)

Upper panel: PRC2 complex deposits H3K27me3 that mediates the recruitment of PRC1 complex by interacting with Cbx (its chromatin “reader” subunit).

Lower panel: KDM2B binding to CpG could induce recruitment of non-canonical PRC1 complexes depositing H2AK119ub. This mark is then recognized by PRC2 complexes that will successively appose H3K27me3. Canonical PRC1 (cPRC1), non-canonical PRC1 (ncPRC1), CpG island (CGI).

1.2.3 Histones variants

The histone H3 variants, CENP-A (Centromeric protein A) or CenH3 is highly enriched at centromere, defining it and directing kinetochore assembly (Müller et al., 2014). Its functionality is very well conserved among Eukaryotes and it is the epigenetic determinant of centromeres. The regulation of CENP-A deposition by the histone chaperon HJURP (Holliday junction recognition protein) is crucial for fidelity of chromosome segregation and cell division (Rop et al., 2012).

H3.3 is another histone variant of H3 differing only in five residues and it is principally found at pericentromeric and telomeric region where it is specifically deposited by DAXX (Death domain-associated protein 6) (Mattioli et al., 2015; Santenard et al., 2010). However H3.3 is present also at transitionally active regions of the genome where it is deposited by another histone chaperone (HIRA) (Mattioli et al., 2015).

1.2.4 Euchromatin and Heterochromatin

Historically, inside an eukaryotic nucleus the chromatin can be divided into two compartments based on the degree of compaction: euchromatin and heterochromatin (Tamaru, 2010). Euchromatin is mainly made by active transcriptional gene-rich regions of the genome, marked by H3K4me3, H3K36me3 and hyperacetylated histones, replicating in early S phase, highly dispersed and diffused in interphase (not stained in electron microscopy) and compacted only during mitosis. On the other hand heterochromatin is composed by gene-poor regions (especially for constitutive heterochromatin) which is supposed to be not transcribed, hypoacetylated in histones, highly condensed and compacted (even in interphase), replicated in late S phase and heavily stained in electron microscopy (Heinz, 1928) (Allis and Jenuwein, 2016) (Figure 10).

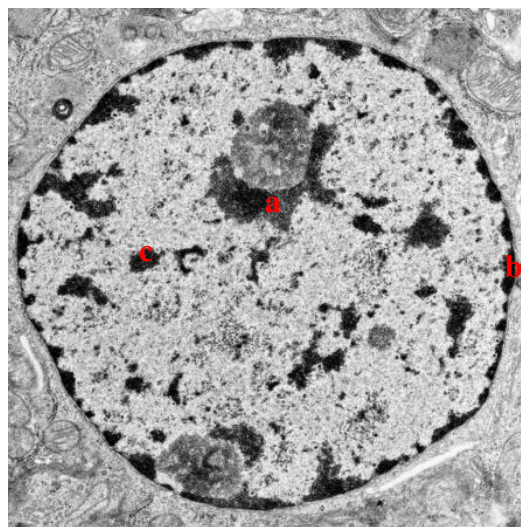


Figure 10: Eukaryotic cell under transmission electron microscopy (TEM).

From a TEM image of an eukaryotic nucleus it is possible to distinguish two different types of heterochromatin: heterochromatin blocks (black) around nucleoli (a), at the nuclear periphery (b) and dispersed into the nucleoplasm (c) and more diffuse heterochromatin (light gray).

The heterochromatin could be further divided in facultative and constitutive parts. Facultative heterochromatin can go from local gene to genomic regions (Hox gene cluster) up to entire chromosomes (X-chromosome in female mammalian cells). It is composed by regions of the genome that have the opportunity to adopt an open and dynamic or closed and compact conformation depending on space and time (Trojer and Reinberg, 2007). Facultative heterochromatin is silenced and generally marked by H3K27me3, but also by H2AK119ub, (Trojer and Reinberg, 2007). Conversely constitutive heterochromatin consists in large blocks

made by repetitive sequences (especially present near centromeres and telomeres) which maintain their characteristics on both homologous chromosomes (Dimitri et al., 2005). Compared to facultative, constitutive heterochromatin is generally marked by H3K9me3 and H4K20me3 and it is recognized by HP1 proteins (Trojer and Reinberg, 2007).

1.2.4.1 Constitutive heterochromatin: the repetitive sequences of mammalian genome

Only 4% of the mouse genome encodes for proteins while the majority of the DNA is made by repetitive sequences (44%) and non-coding sequences (52%) (Martens et al., 2005). The human genome content is similar to the mouse one with only 2% of protein coding sequences and the remaining (98%) is made up of transposable elements and tandem repeats (López-Flores and Garrido-Ramos, 2012). Now it is known that this “junk DNA” plays a role in the formation of specialized structure but, due to its repetitive characteristic, this DNA is also an issue for genome stability as more easily subjected to recombination, deletion or translocation than single-copy sequences (Jaco et al., 2008).

Mouse chromosomes are all acrocentric while human chromosomes can be divided in three groups: metacentric, sub-metacentric and acrocentric. However chromosomes of both species present around their centromeres the constitutive heterochromatin which is cytologically visible (Figure 11A) (Padilla-Nash et al., 2007). Constitutive heterochromatin consists principally in centric, pericentric and telomeric region. At the end of each linear chromosome there is the telomeric region composed from hundreds to thousand repeats of TTAGGG sequence, which is conserved between all the mammals, and coated with protecting proteins in order to prevent DNA damage and inappropriate recombinant events (Calado and Dumitriu, 2013). In addition to these large tandem arrays of repeats, the mouse genome contains single repetitive elements that are all along the chromosomes, such as DNA transposons that represent 1% of this interspersed repetitive element, while RNA transposons or retrotransposons represent about 25%. Retrotransposons include: LTR (Long terminal repeats) transposons, principally intracisternal A particle (IAP), non-LTR transposons or LINEs (Long Interspersed Nuclear Element) which represent the largest fraction with 19% of the genome and the Short Interspersed Nuclear Element (SINE) (Martens et al., 2005) (Figure 11B).

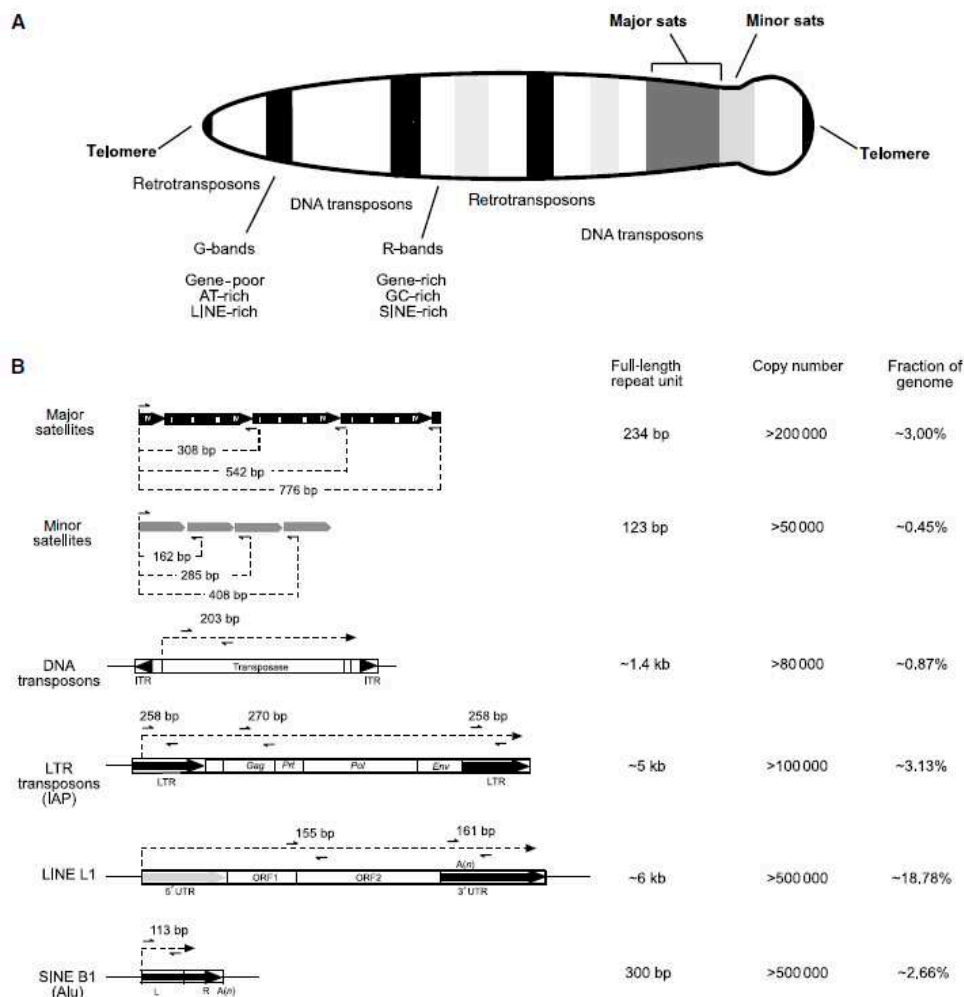


Figure 11: The different repetitive sequences in the mouse genome (From Martens et al., 2005)

(A) Schematic representation of a mitotic mouse chromosome illustrating the distribution of major and minor satellite repeats, respectively pericentromeric and centromeric heterochromatin and of the various interspersed repetitive elements. (B) Summary of repetitive elements in mouse with repeat organization, length, copy number and overall abundance in the mouse genome. Specific primers (black arrows) can be designed to generate PCR fragments from within one or more successive repeat of the repetitive elements.

In mouse the core of the centromere is made by tandem arrays of minor satellite repeats (123bp) in approximately 2000 copies (Martens et al., 2005). Minor satellites are heavily methylated at the DNA level, transcriptionally silenced and characterized by the H3 histone variant CENP-A (Scott, 2013). Juxtaposed to the centromere, the pericentromeric region is composed by major satellite repeats (234bp) which are A/T-rich sequences that are divided in four sub-repeats and present in more than 10000 copies. Major satellites are characterized by H3K9me3, HP1 and H4K20me3. Like minor satellites, they are characterized by an hypermethylation at the DNA level and a repressed transcriptional state (Lehnertz et al., 2003). Together centromeric and pericentromeric regions represent more than 3.5% of the mouse genome, with major satellites alone representing 3% (Martens et al., 2005). A

characteristic of mouse nuclei is the clusterization of pericentromeric and centromeric heterochromatin in chromocenters. These structures are formed by the coalescence of major satellites coming from different chromosomes with the corresponding minor satellites that are located in surrounding separate domains at the periphery (Guenatri et al., 2004) (Figure 12). These prominent heterochromatin domains being rich in A/T sequences (major satellites) are well discernible in the nucleus as 4',6-diamidino-2-phenylindole (DAPI)-dense foci (Dambacher et al., 2013).

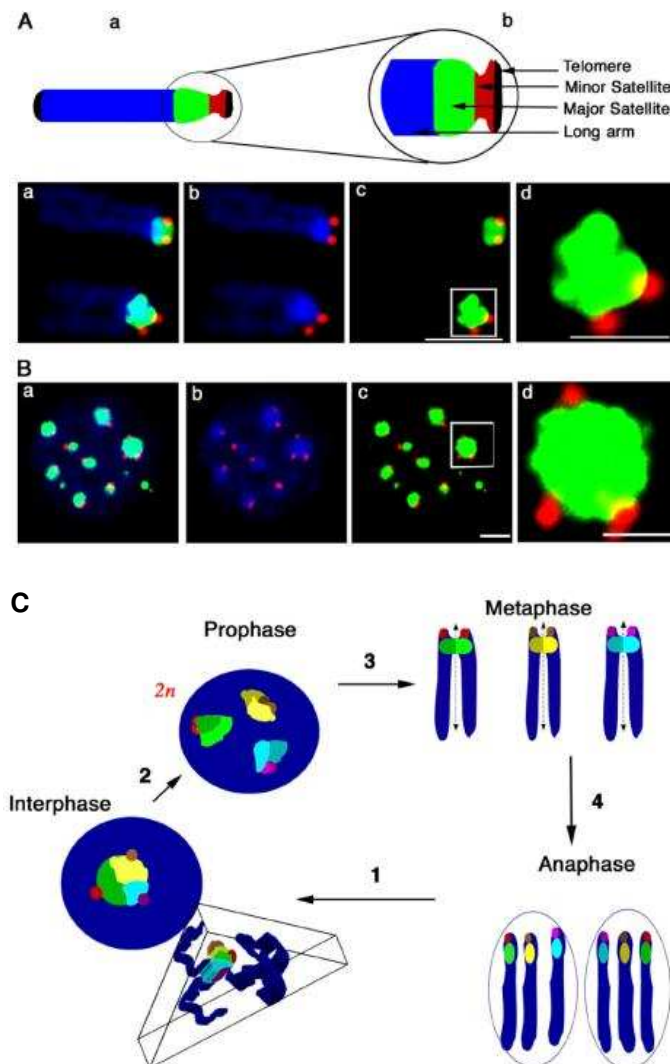


Figure 12: Mouse chromosomes organize together their pericentromeric region to form Chromocenters (From Guenatri et al., 2004)

(A) Schematic organization and DNA-FISH images of major and minor satellites along the mouse chromosome. DAPI DNA counterstaining (blue), majors satellites (green), minor satellites (red).

(B) DNA-FISH images for major and minor satellites revealing the chromocenter organization inside a mouse interphase nucleus. DAPI DNA counterstaining (blue), majors satellites (green), minor satellites (red).

(C) Schematic organization of mouse chromosome during cell-cycle. (1) In interphase major satellites from different chromosomes associate in clusters (chromocenters). (2) In prophase major satellites from different chromosomes dissociate. (3) In metaphase minor satellites from sister chromatids dissociate, whereas the major satellite sister chromatids still cohere. (4) In anaphase finally major satellites from sister chromatids separate.

In human these regions are also composed by satellite repeats, however the DNA sequence of these repeats is different. Human centromeric region is made by alpha satellite repeats of 171bp that are present in all chromosomes, while the closer pericentromeric region is composed by Satellite 2 and 3 principally with different size and composition between the chromosomes (Saksouk et al., 2015).

1.2.5 Non-coding RNAs (ncRNAs)

The discovery that mammalian genome is transcribed quite entirely even if only a small fraction codes for protein, introduces the concept of the existence of non-coding RNAs (“Dark matter” RNA) (Mattick, 2007). ncRNAs are divided in two groups based on their size: small and long non-coding RNAs.

The small non-coding RNAs are RNA species of less than 200nt that are in many cases associated with 5’ or 3’ regions but also in introns of protein-coding genes (Mattick and Makunin, 2006). The most studied are Dicer-dependent microRNA (miRNA) and small interfering RNA (siRNA), and the Dicer-independent PIWI-interacting RNA (Carthew and Sontheimer, 2009). These ncRNA are largely involved in post-transcriptional gene silencing (PTGS) (Agrawal et al., 2003).

The long non-coding RNAs (lncRNA) are defined as transcripts of more than 200nt that lack an open reading frame (Cao, 2014). LncRNAs are mainly transcribed by RNA polymerase II, they can undergo splicing and can also contain a poly-A tail. They can be developmentally-regulated and/or tissue-specific, being implicated in alternative splicing, modulation of protein activity, alternative protein localization, epigenetic regulation, transcriptional silencing. LncRNAs can act as signals, guides, decoy and scaffold (Sana et al., 2012).

While for many years the repetitive constitutive heterochromatin has been considered as a transcriptionally inactive domain due to its compacted organization, high DNA methylation content and presence of repressive histone mark like H3K9me3, it is now clear that in many cases these sequences are transcribed giving rise to satellite non-coding RNAs (Biscotti et al., 2015).

1.2.5.1 Satellite non-coding RNAs

Transcripts homologous to centromeric and pericentromeric repetitive sequences have been identified in several organism from yeast to human (Saksouk et al., 2015).

In yeast the RNA polymerase II transcribes satellites repeats and these ncRNA are involved in the heterochromatin formation, maintenance and silencing via the RNA interference machinery (RNAi). Paradoxically the heterochromatin needs to be transcriptionally active to maintain its inactive state by the recruitment of H3K9-KMTs (Biscotti et al., 2015).

In mouse the implication of the RNAi pathway in the heterochromatin is controverted (Kanellopoulou et al., 2005; Murchison et al., 2005) and still under debate with no real proof (Plohl et al., 2014). However transcription of ncRNA from both major and minor satellites is observed in physiological conditions as well as in pathological conditions (Figure 13) (Saksouk et al., 2015). Interestingly it has been shown that an RNA component is involved in the high structured three-dimensional chromatin at pericentromeric regions, as RNase treatment disrupts the H3K9me3-HP1 foci (Maison et al., 2002). Whether this RNA component is also made by pericentromeric satellite transcripts is still unknown.

Non-coding pericentromeric RNA are produced by RNA polymerase II in both orientation in mouse as well as in human: sense or forward (T-rich in mice) and antisense or reverse (A-rich) (Figure 13) (Saksouk et al., 2015) and can be also polyadenylated (Lehnertz et al., 2003).

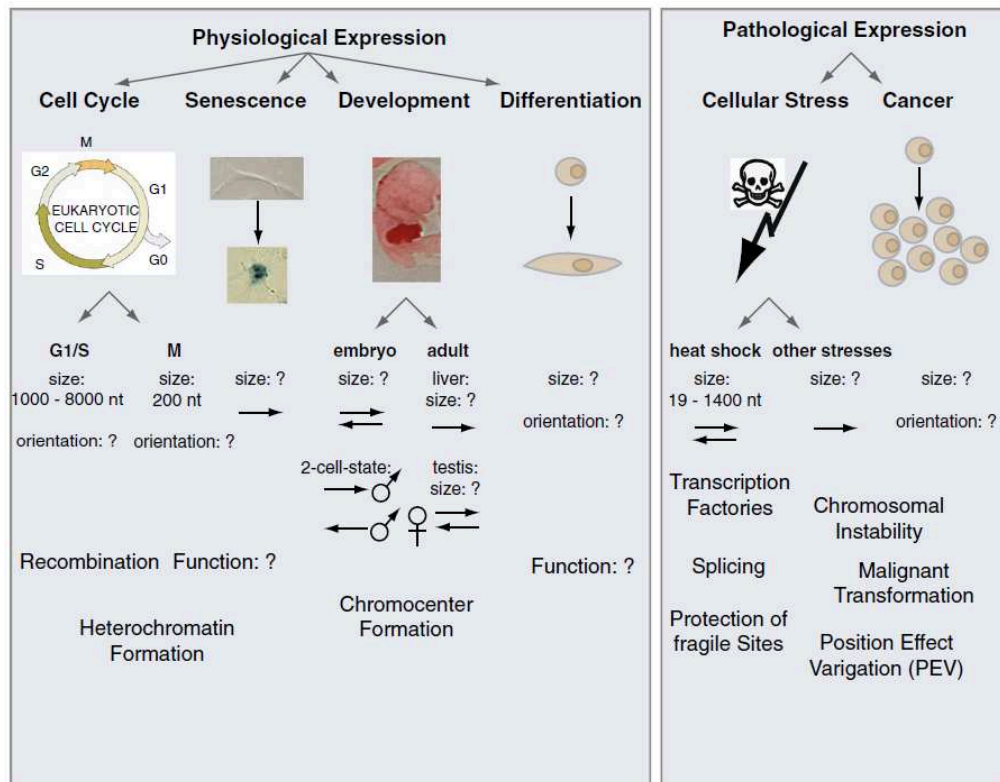


Figure 13: Pericentromeric satellite transcription in different context (From Saksouk et al., 2015)
Physiological expression of pericentromeric satellite repeats has been reported during cell cycle, senescence, development and differentiation. Pathological expression has been observed upon cellular stress and in cancer. The size, orientation and putative functions of the transcripts are indicated when known.

In the major satellite repeat sequence many binding sites for transcription factors were found (Figure 14). For example these repeats contain sites for PAX3 and PAX9 Paired box-transcription factors, which were found to be necessary to repress non-coding RNA transcription from these regions and also to help the recruitment of SUV39H enzymes needed for the deposition of H3K9me3 (Bulut-Karslioglu et al., 2012).

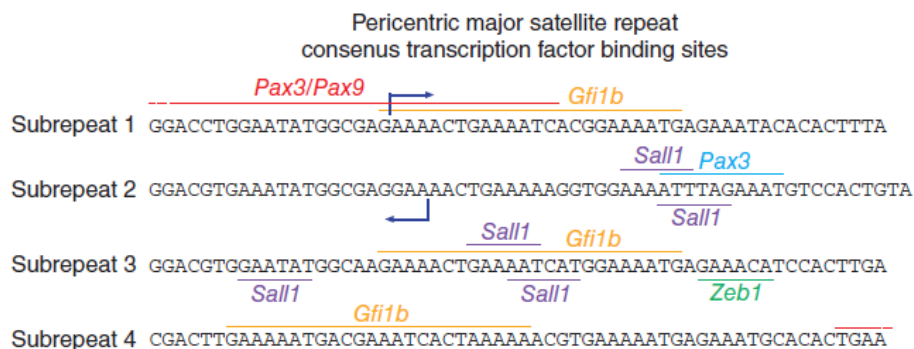


Figure 14: Transcriptional factor binding sites on major satellite repeat sequence (From Bulut-Karslioglu et al., 2012)
Consensus sequence of a full-length major satellite repeat. Transcription factor binding sites are highlighted above or below the DNA sequence according if their binding motif is respectively sense or antisense.

Different populations of satellite transcripts seem to be generated according to the cell cycle stage. First major satellite transcription is Cdk-dependent because cells do not transcribe if they are not proliferating or maybe in this condition there are extremely short-lived (Lu and Gilbert, 2007). During mitosis small RNA species (less than 200nt) are produced and their half-life is very short (less than one hour). Conversely a more abundant, large and heterogeneous population of transcripts (from 1kb to more than 8kb) is produced between late G1 and early S, and strongly down-regulated after mid-S phase when pericentric heterochromatin is replicated (Lu and Gilbert, 2007). Major satellites are found to be transcribed in adult mice only in highly proliferative tissues as liver and testes. In particular in the liver transcription occurs only in sense orientation while in testes it is antisense in immature germ cells and sense in the mature ones (Saksouk et al., 2015). During replicative senescence and aging, pericentromeric transcripts were detected especially from the sense orientation and concomitant to reduction of methylation levels and decondensation of constitutive heterochromatin (Figure 13) (Saksouk et al., 2015). Similar observations were made in several cancer cells with an associated genetic instability and chromosomal disorder (Frescas et al., 2008).

Minor satellites have also been shown to produce heterogeneous populations of transcripts and to be cell-cycle regulated. In in vitro cultured mouse cells, centromeric transcripts are present in two long forms of 2kb and 4 kb but also in a smaller form around 120nt. These small minor satellite transcripts accumulate with culture-time (Bouzinba-Segard et al., 2006). No centromeric transcripts were found at the range of size for siRNAs (22-30nt) suggesting no RNAi involvement in mouse. The 120nt population of minor satellites increased during stress condition, using a DNA demethylation agent or inducing apoptosis, leading to an impaired centromeric function during mitosis and promoting defects in chromosomal segregation (Bouzinba-Segard et al., 2006). In addition minor satellites are lowly expressed in G1 phase and increased with cell-cycle progression peaking at G2/M phases just prior the kinetochore assembly (Ferri et al., 2009). Furthermore centromeric transcripts were found to be an RNA component of the CENP-A associated complex in concert with Chromosomal Passenger Complex (CPC), Aurora B kinase and Survivin providing implications of minor satellite transcripts in chromosome segregation (Ferri et al., 2009).

2 PLURIPOTENCY

2.1 Totipotency and Pluripotency

Pluripotency is defined as the capacity of cells to self-renew (auto-maintain themselves) in vitro and to differentiate into the three embryonic lineages (mesoderm, endoderm and ectoderm) that will give rise to all the different tissues of an adult organism, but not to extra-embryonic tissues. Totipotency is the capability of a cell to give rise to a fertile adult individual, so generating also the extra-embryonic tissues (like placenta or yolk-sack). In mouse, totipotency is restricted to the zygote (1-cell stage) and to blastomeres of a 2-cell stage embryo (Condic, 2014). In the following pre-implantation embryo stages, cells are pluripotent up to the blastocyst stage where we assist to the first real differentiation process: the inner cell mass (ICM) and the trophectoderm (TE). At this stage only ICM cells are still pluripotent. Pluripotent cells are also present after the implantation in the epiblast up to E7.5 (Nichols and Smith, 2009; Osorno et al., 2012).

2.2 Early mouse embryo development

The early mouse embryo development goes from totipotency to pluripotency. It begins with fertilization, followed by dividing pre-implantation embryo stages that lead to the blastocyst formation (Figure 15A) which then will implant into the uterus starting the post-implantation development and gastrulation. The pre-implantation development starts in the oviduct where the oocyte is fertilized by the sperm giving rise to the zygote (fertilized egg) (Wang and Dey, 2006).

At the blastocyst stage E3.5 two types of cells can be observed: the pluripotent inner cell mass (ICM) and the trophectoderm (TE), distinguished by Oct4-Sox2-Nanog or Cdx2-Eomes-Gata3 expression, respectively (Chazaud et al., 2006). At this stage Oct4 repressed TE lineage and Cdx2 conversely ensures the repression of Oct4 and Nanog (Figure 15B). In E3.5 early blastocyst individual ICM cells show exclusive expression of epiblast genes (such as Nanog) or primitive endoderm genes (Gata6 and Gata4) in a “salt and pepper” manner (Chazaud et al., 2006).

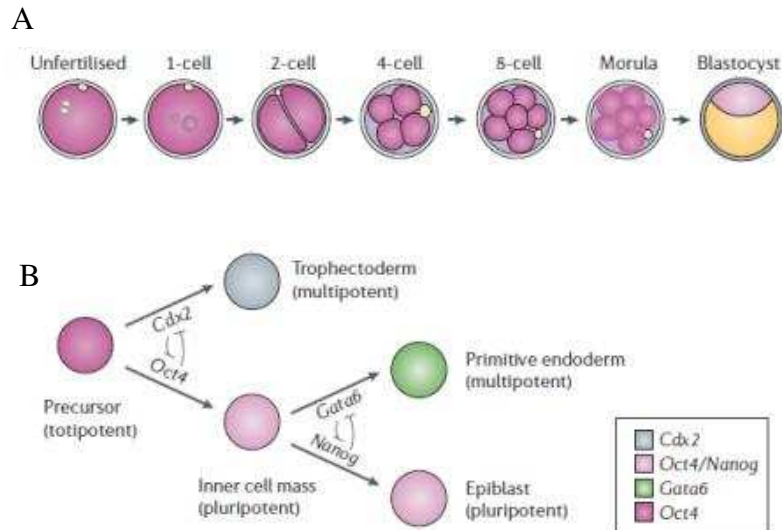


Figure 15: Pre-implantation development and first gene specifications (From Wang and Dey, 2006)

A: Schematic representation of mouse pre-implantation embryo development: once the egg is fertilized by the sperm successive cell divisions will give rise at E3.5 to the blastocyst.

B: Gene expression pattern governing blastocyst specification. Oct4 (dark pink) is expressed throughout the embryo before the late morula stage. Nanog (light pink) is specifically induced in the inside cells of late morula. Cdx2 (blue) is expressed in the outer layer of cells in late morula and is required for the repression of Oct4 and Nanog in the trophoblast of the blastocyst. Gata6 (green) is expressed in the primitive endoderm of the late blastocyst, where Oct4 and Nanog are repressed. Oct4 represses Cdx2 expression, which in turn represses Oct4 expression allowing segregation of the ICM and trophoblast of the blastocyst. Nanog and Gata6 antagonized each other segregating epiblast and primitive endoderm within the ICM.

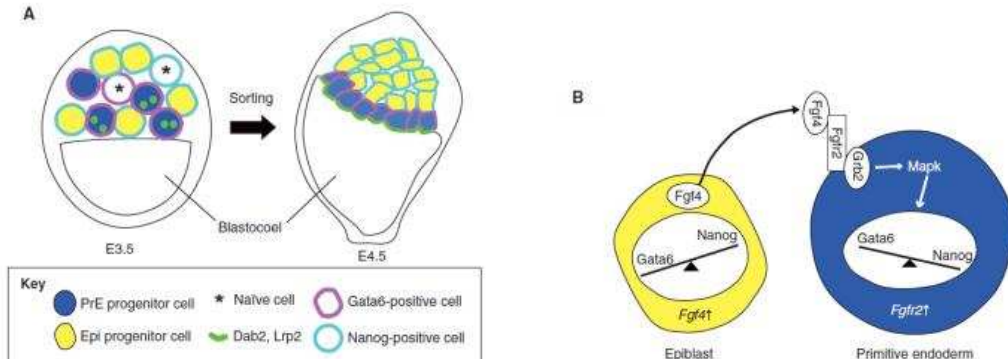


Figure 16: Epiblast vs. Primitive endoderm specification in early- to late-blastocyst stage (From Takaoka and Hamada, 2012)

(A) Primitive endoderm (blue) and epiblast (yellow) progenitors are randomly positioned in the inner cell mass (ICM) at E3.5. At this stage, some of the cells in the ICM are still naïve (asterisk) for Nanog or Gata6 expression. By E4.5, the primitive endoderm cells migrate at the surface of the blastocoel cavity, whereas the Epiblast cells are confined to the inner part of ICM. Dab2 and Lrp2 are localized to the apical surface of PrE cells. (B) FGF signaling guides primitive endoderm formation. Fgf4 secreted by epiblast progenitors interacts with Fgfr2 and thereby activates Grb2 and Mapk in primitive endoderm progenitors. Activated Mapk induces the expression of PrE-specific genes, such as Gata6. Fgf4 simultaneously represses Nanog expression, further promoting PrE fate while inhibiting Epi fate. Epiblast (Epi), primitive endoderm (PrE), adaptor protein disabled (Dab2), lipoprotein receptor-related protein 2 (Lrp2), fibroblast growth factor 4 (Fgf4), Fgf receptor 2 (Fgfr2), growth factor receptor bound protein 2 (Grb2), microtubule-associated protein kinase (Mapk).

FGF signaling is involved into primitive endoderm specification (Figure 16) (Artus and Chazaud, 2014). By E4.5, just prior to implantation, this mosaic pattern is resolved in a late blastocyst where ICM cells are sorted and relocated in epiblast cells (Epi) (Figure 16) which are pluripotent and will give rise to the fetus, and primitive endoderm (PrE) from which will predominantly originates the yolk sac. The PrE cells form a layer at the surface of the ICM facing the blastocoel cavity and are positive for Gata6. Conversely Epi cells are positive for Nanog and are surrounded by PrE cells and TE cells (Rossant and Tam, 2009).

At this stage (E4.5) the blastocyst invades the uterine tissues and implants. Mural TE surrounding the blastocyst cavity will first makes contacts with maternal tissues and differentiates into primary TE giant cells which mediate the mother-embryo exchanges. Conversely polar TE surrounding the outer face of the ICM proliferates and differentiates in extra-embryonic ectoderm (ExE) and ectoplacental cone (EPC) that will build the proximal half of the egg cylinder and later on the placenta (Figure 17). After implantation at E5.0 there is a burst of cell proliferation and growth resulting in the expansion of embryonic as well as extra-embryonic lineages into the blastocyst cavity. From PrE derives the visceral endoderm cells that cover both Epi and ExE. The elongating egg cylinder grows, reorganizes and forms the primitive streak in order to initiate the gastrulation under the control of TGF β /Activin/Nodal pathway. Nodal play a key role in the maintenance of pluripotency of the epiblast and the posterior cell fate. Indeed in Nodal $-/-$ condition the epiblast is precociously differentiated into the neuronal fate (Camus et al., 2006). During peri-implantation stages the Epi reorganized itself from a compact ball non-polarized to a rosette-like organization, at the time of implantation, and finally to a cup-shaped polarized epithelium surrounding the pro-amniotic cavity in 24 hours (Figure 17) (Bedzhov et al., 2014). The anterior-posterior axis of the mouse embryo is established when distal visceral endoderm (DVE) will form the anterior visceral endoderm (AVE) on the future anterior side of the E6.5 embryo. Once established the AVE will secrete signal like Nodal antagonists (Lefty1) to the Epi for anterior specification (Perea-Gomez et al., 2002).

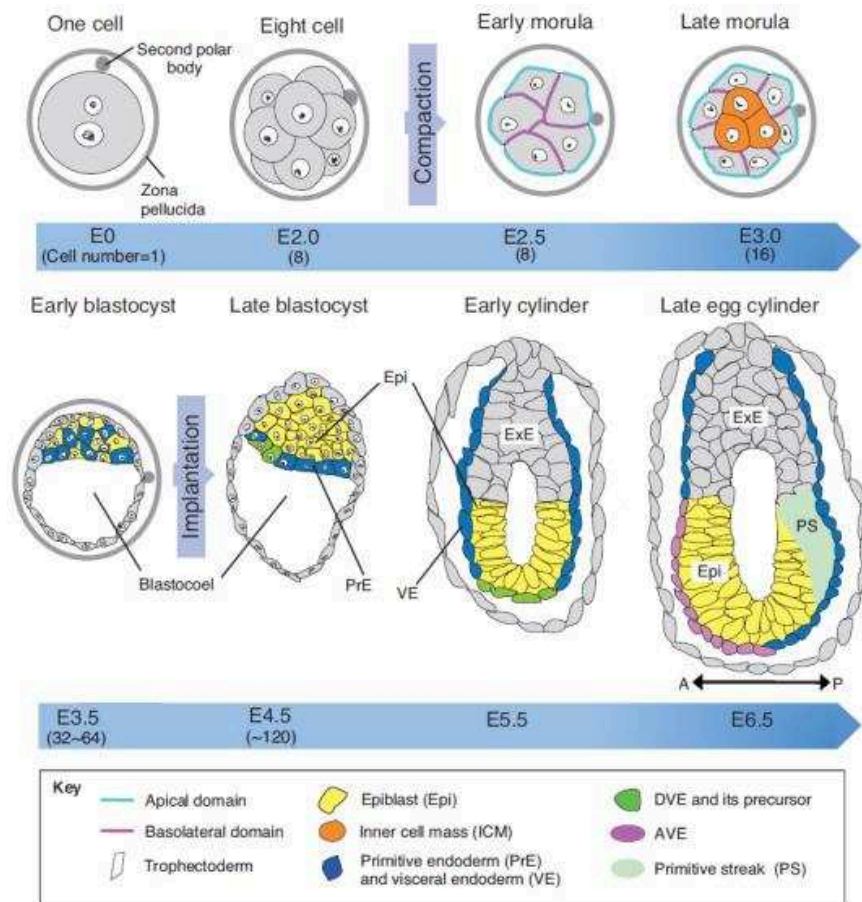


Figure 17: Early mouse development from 1-cell stage to gastrulation (From Takaoka and Hamada, 2012)
 Kinetics of morphological changes and cell-specification from zygote to late egg cylinder. The cell types in the embryos are coded with different colors: epiblast (Epi), primitive ectoderm (PrE), visceral endoderm (VE), extra-embryonic ectoderm (Exe), distal visceral endoderm (DVE), anterior visceral endoderm (AVE), primitive streak (PS).

2.3 In vitro pluripotency in mouse

The attractive potential of embryo-pluripotent stem cells has fascinated scientists who attempted derivation of these cells as a stable line in vitro. In 1981 two groups established for the first time an in vitro culture of pluripotent stem cells starting from 129 strain mouse blastocyst-embryos E3.5 (Evans and Kaufman, 1981; Martin, 1981). These cells have the capacity to differentiate into all the three primary germ layers and were called embryonic stem cells (ESCs) in order to distinguish them from embryonal carcinoma cells (ECCs), that were isolated from teratocarcinoma induced by transplantation of mouse embryos in an extra-uterine site of a host. Many years later in 2007 two other groups established another type of pluripotent stem cells this time using epiblast of the egg-cylinder of a post-implantation embryo E5.5 (Figure 18) (Brons et al., 2007; Tesar et al., 2007). These cells were called epiblast stem cells (EpiSCs) in order to distinguish from the ESCs and while they form

teratomas, as ESCs, they cannot form chimeras when injected in the ICM of a blastocyst. These two type of cells are clearly distinct at the molecular point of view (intracellular and extracellular signaling) and they have been defined as naive pluripotency for mouse ESCs and primed pluripotency for mouse EpiSCs (Figure 18) (Nichols and Smith, 2009). However these states share some characteristics like the expression of the transcription factors Oct4, Sox2 and Nanog which represent the molecular foundation of pluripotency.

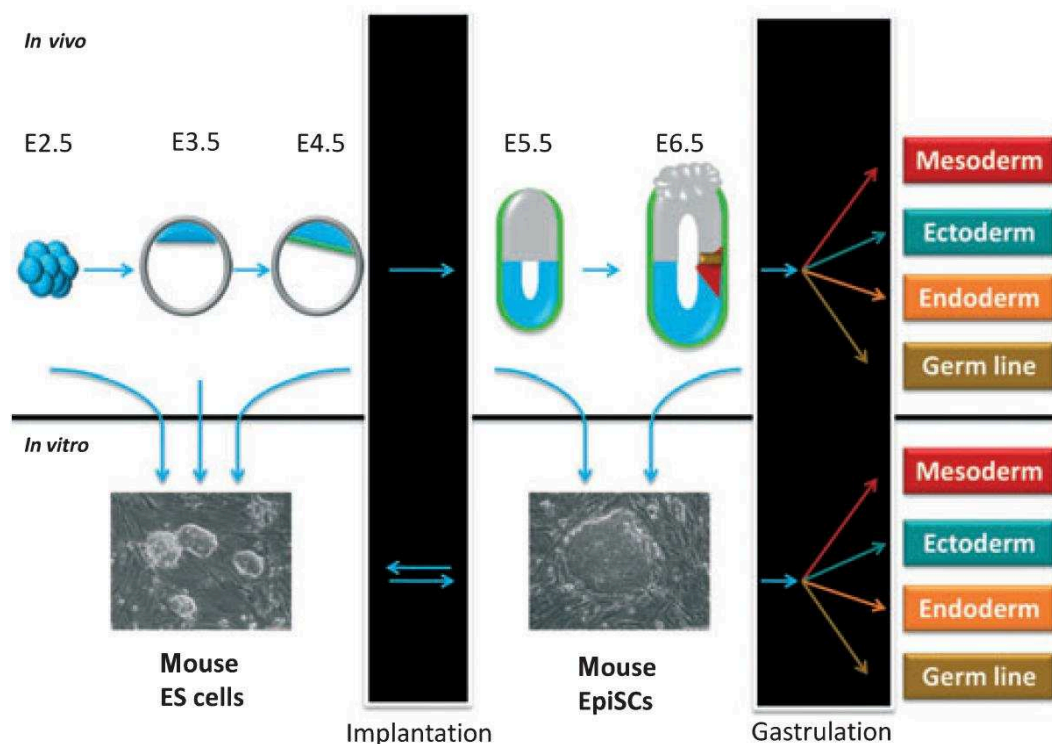


Figure 18: From in vivo to in vitro mouse pluripotency (From Chenoweth et al., 2010)

Pluripotent cells from a developmental window in the pre-implantation embryo grow as mouse embryonic stem cells (ESCs) and those from the post-implantation epiblast grow as epiblast stem cells (EpiSCs).

2.3.1 The core pluripotency factors: OCT4, SOX2 and NANOG

OCT4 is a POU domain-containing transcription factor which is encoded by Pou5f1 gene. It binds to the octamer DNA sequence motif ATGCAAAT and it controls the expression of many genes related to pluripotency in tandem with SOX2 (Loh et al., 2006). The role of OCT4 in pluripotency is crucial and it is one of the transcription factors (“Yamanaka factors”) that are needed to induce the pluripotent state from differentiated cells such as fibroblasts (Takahashi and Yamanaka, 2006). Oct4 is expressed all along the pre-implantation development (Figure 19) but at the blastocyst stage it becomes restricted to ICM cells while

TE cells are Oct4-negative. Oct4^{-/-} embryos can only survive up to the morula stage and failed to form ICM *in vivo*, or ESCs *in vitro*, because blastomeres fate is redirected towards the trophoblast lineage (Loh et al., 2006; Nichols et al., 1998). However the expression of Oct4 in ESCs must be tightly regulated because when expressed above the endogenous levels it induces differentiation in the extra-embryonic endoderm lineage (Niwa et al., 2000).

SOX2 is a member of the SOX (SRY-related HMG box) transcription factor family which has a highly conserved single DNA-binding domain HMG (High Mobility Group). It is expressed all along the pre-implantation embryo development (Figure 19) but also in the ICM, epiblast, extra-embryonic ectoderm and neural tissues (Avilion et al., 2003; Huang et al., 2015). Sox-null embryos die just after implantation and Sox2-null ESCs differentiate in the trophoblast-like cells as Oct4^{-/-} ESCs. It is likely that SOX2 maintains the pluripotency via the transcriptional activation of Oct4 (Masui et al., 2007).

NANOG is a Q50 homeodomain-containing protein with the amino acid glutamine at position 50 of the homeodomain that makes direct contact at the 5' of its preferred consensus sequence ATTA tetramer. NANOG regulates the expression of Pou5f1 and Sox2 sustaining pluripotency (Huang et al., 2015) and when over-expressed in ESCs it confers self-renewal independency to LIF/STAT3 signaling. Nanog-null embryos are able to give rise to pluripotent stem cells even if they rapidly differentiate into extra-embryonic endoderm lineage (Chambers et al., 2007). This bias can be partially explained by the well-known negative regulation of Nanog on Gata6 which induce primitive endoderm differentiation. Conversely to Oct4^{-/-} or Sox2^{-/-}, Nanog-null ESCs can be derived meaning that Nanog is not strictly required for the maintenance of self-renewal *in vitro*. Furthermore Nanog is not within the canonical “Yamanaka factors” necessary to reprogram somatic cells to iPSCs (Takahashi and Yamanaka, 2006). NANOG controls important factors of ESCs such as Esrrb, which can induce LIF independency when over-expressed in Nanog^{-/-} ESCs, rescuing the absence of Nanog (Festuccia et al., 2012).

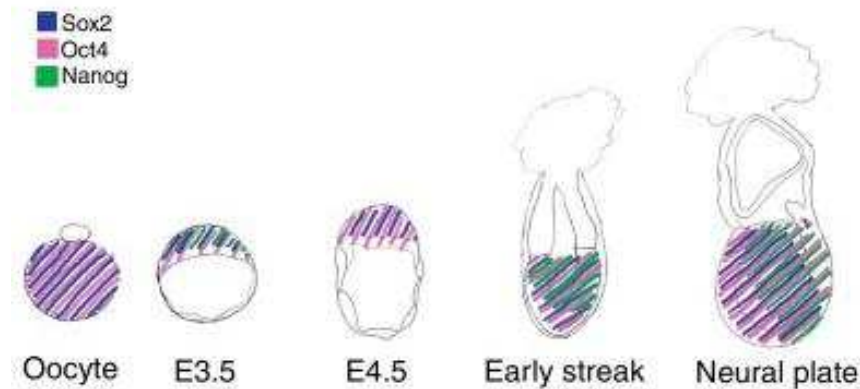


Figure 19: Transcript expression of the core pluripotency factor during mouse development (From Festuccia et al., 2013)

Oct4 and Sox2 mRNAs are maternally inherited in oocyte while Nanog mRNA is first detected in blastomeres of the 8-cell stage embryo. At E3.5 Nanog is expressed heterogeneously in the ICM and Nanog-negative cells will form the hypoblast or primitive endoderm at E4.5. Around implantation Nanog is down-regulated before being re-expressed in the posterior epiblast. Oct4 and Sox2 also become regionally expressed post-implantation, with Sox2 higher in the anterior epiblast and Oct4 becoming progressively posterior. Sox2 (blue), Oct4 (pink) and Nanog (green).

2.3.2 Signaling pathways in pluripotency

Different pathways are involved in the control of pluripotency (Figure 20). The LIF/STAT3 signaling is probably the principal pathway governing self-renewal of mESCs (Niwa et al., 1998). LIF is a cytokine of the IL-6 family that binds the LIF receptor (LIFR) and the co-receptor subunit glycoprotein 130 (gp130) which together form an heterodimer and activate the tyrosine kinase activity of Janus kinases (JAKs). The JAKs phosphorylate themselves becoming docking sites for the SH2 domain of cytoplasmic STAT3 (Signal Transducer and Activator of Transcription 3). Once STAT3s are recruited, they become phosphorylated and dimerize. The dimerization promotes in turn the translocation to the nucleus where they activate gene transcription (Huang et al., 2015). Constitutive STAT3 activation makes the self-renewal of mESCs independent to LIF. STAT3 activation prevents ESCs differentiation to mesoderm and endoderm lineages, but a stronger activation induces also ESCs differentiation towards the trophoblast lineage (Tai et al., 2014). In vivo studies reported a role of LIF/STAT3 during diapause when TE cells secrete LIF to sustain the ICM in the arrested embryo development and delayed implantation (Renfree and Shaw, 2000). However in normal development without diapause LIF signaling is not required, but STAT3 activation by IL-6 is essential in vivo for ICM maintenance as it binds to Oct4 and Nanog enhancers (Do et al., 2013).

TGF β /SMAD pathway also sustains mESCs self-renewal via Bone Morphogenic Protein (BMP) and in collaboration with LIF/STAT3 prevents neural differentiation (Huang et al., 2015). BMP binds to transmembrane receptor type I and II that, once dimerized, activates the serine/threonine kinase activity of type I receptors (ALKs), which phosphorylates the SMAD1 proteins. Phosphorylated SMAD forms complexes with co-SMAD (SMAD4) which translocated into the nucleus activating in turns transcription of target genes (Shi and Massagué, 2003). The same basic pathway is also activated by Activin/Nodal signaling which is fundamental to sustain pluripotency of mEpiSCs, but they depend specifically on SMAD2/3 (Greber et al., 2010; Sakaki-Yumoto et al., 2013).

FGF/ERK pathway is important in mESCs as well as mEpiSCs while inducing opposite effects. FGF ligand induces auto-phosphorylation on tyrosin residues of intracellular domains of FGF receptor (FGFR) (Huang et al., 2015). Fibroblast growth factor receptor substrate 2 (FRS2) and Grb2 are recruited, phosphorylated and activate RAS/MEK/ERK pathway which promotes self-renewal of mEpiSCs while it is a differentiation cue for mESCs (Kunath et al., 2007). To suppress differentiation and reach the ground-naïve state of pluripotency, mESCs are cultured with an inhibitor of FGF/ERK signaling: PD0325901 (PD) (Ying et al., 2008).

Wnt/ β -catenin pathway is also implicated in pluripotency as well as differentiation. In absence of Wnt (ligand), glycogen syntase kinase-3 (GSK3) forms the destruction complex which phosphorylates β -catenin. Once phosphorylated, β -catenin is ubiquitinated and degraded via the proteasome. However when Wnt is present it inhibits the assembly of the destruction complex, leading to accumulation of β -catenin that can inhibit T Cell Factor (TCF) (Sokol, 2011). TCF is a repressor of pluripotency genes, hence by using a GSK3 inhibitor (CHIR99021) it is possible to stabilize the naïve pluripotency network (Wray et al., 2011; Ying et al., 2008) and to inhibit the epiblast transition (Berge et al., 2011). In primed EpiSCs, however, activation of the Wnt pathway induces mesendoderm differentiation and the use of inhibitors of this pathway leads to the block of this spontaneous differentiation and homogeneous propagation of EpiSCs colonies (Sumi et al., 2013).

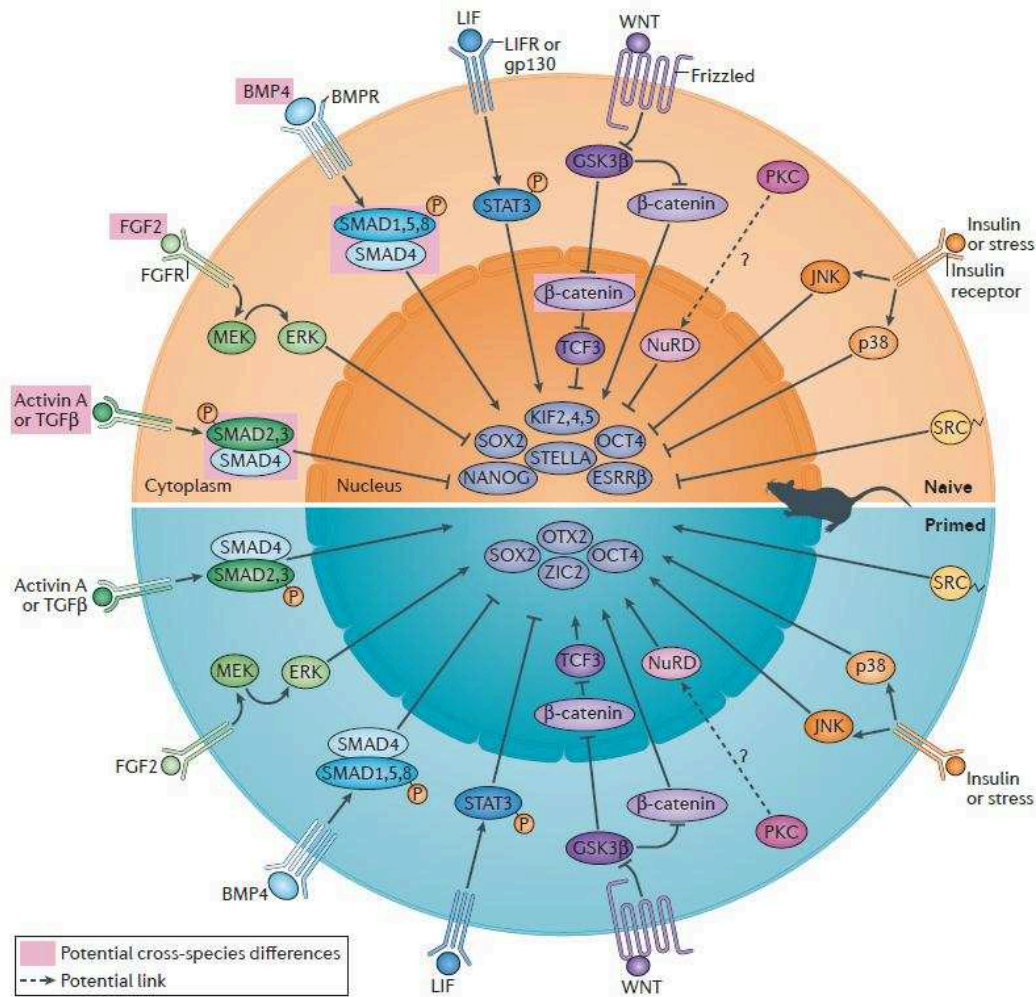


Figure 20: Signaling pathways in naive Vs. primed mouse pluripotency (From Weinberger et al., 2016). Principal signaling pathways that can positively or negatively regulate naive and primed murine pluripotent stem cells. The majority of the signaling pathways shown have opposing effects on the naive and primed pluripotent states.

2.3.3 The naïve state of mouse pluripotency

2.3.3.1 Naïve-metastable mESCs in serum/LIF

mESCs were initially maintained on gelatin-coated dishes in fetal bovine serum (FBS)-containing medium and in co-culture with mitotically inactivated feeder fibroblast (called simply “feeders”) or mouse embryonic fibroblast (MEFs) which produce trophic factors sustaining ESCs growth. Then it has been shown that “feeders” produce LIF and so the addition of this chemokine to the medium can replace feeders which release also many other unknown factors (Smith et al., 1988). Naïve-metastable mESCs can be cultivated also in BMP4/LIF instead of serum to inhibit differentiation (Ying et al., 2003). Standard serum/LIF

conditions and feeders permitted the establishment of ESCs principally from the 129 mouse strain, hence called "permissive" in contrast to other strains like NOD (Non-obese diabetic) which were "non-permissive" (refractory) to ESC derivation.

mESC in serum/LIF conditions grow in dome-like colonies (Figure 21) and can be passaged as single-cells with trypsin.

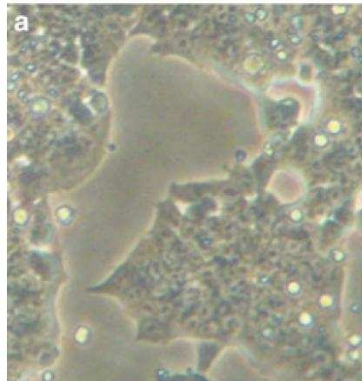


Figure 21: Morphology of serum-ESCs under the phase-contrast microscope (From Tosolini and Jouneau, 2016)

Dome-like colonies ESCs in serum/LIF condition (on gelatin coated plates).

Serum-ESCs appear to be subjected to uncontrolled multi-factorial perturbations, and thus are considered to be in a metastable condition, characterized in particular by transcription of lineage specification genes and heterogeneous expression of some pluripotency factors (Kalkan and Smith, 2014; Toyooka et al., 2008). Serum-ESCs are indeed transcriptionally hyperactive with also high transcription of repetitive elements (major and minor satellites, IAPs, LINEs, SINEs) but also transcripts of tissue-specific genes (Efroni et al., 2008). The metastability of serum-ESCs comes from cell population studies that showed the fluctuation of these cells between Nanog-high and Nanog-low states but also Rex1-high and Rex1-low states (Marks et al., 2012; Wray et al., 2010). The bases of this heterogeneity in serum-ESCs can be explained at least partially by the dynamics of gene expression. It is likely that the heterogeneity does not arise from stochastic fluctuations or noise in gene expression but from coexistence of multiple cellular states. Using single molecule RNA-FISH Singer and colleagues (2014) detected in serum-ESCs three types of gene expression distribution: Unimodal, Long-tailed and bimodal. Among bimodal genes were Nanog and Rex1. Cells remain into one of the two expression states for multiple cell cycles before transiting to the

other state in a dynamic way. Interestingly only Rex1-high population of ESCs can efficiently colonize the blastocyst and contribute to development while Rex1-low cells are more prone to differentiation, showing that naive-ESCs conventionally cultured in serum are a mosaic of subpopulations (Toyooka et al., 2008). Another recent study also highlighted the heterogeneity of serum-ESCs by single-cell analysis showing that it is possible to distinguish three cellular states: the naive population with high levels of Nanog, Sox2 and Klf2, the primed population expressing pluripotency markers (Nanog and Sox2) but also differentiation markers (Gata4, Gata6 and Lamb1) and PrE lineage cell showing high levels of Gata4 and Gata6 and low levels of pluripotency markers (Guo et al., 2016).

Naive metastable-ESCs are characterized by a peculiar cell cycle with a short G1 phase of less than 2h, lack of the pRB and p53 DNA damage checkpoints, high levels of CyclinE and A in complex with Cdk2. This short G1 phase seems to be regulated by LIF signaling, as LIF starvation reduce the progression of Rex1-high cells to the next phase (Coronado et al., 2013).

2.3.3.2 Ground-naive mESCs in 2i/LIF

After more than 30 years of ESCs cultivation in LIF with serum or BMP, Ying and colleagues (2008) suggested the use of serum-free medium that abrogated the differentiation-inducing signal with the addition of small molecule inhibitors. In particular they showed that inhibition of MAPK and GSK3 enabled self-renewal of mouse ESCs, suppressing residual differentiation (Ying et al., 2008). This spontaneous differentiation is due to the auto-inductive stimulation of the ERK 1/2 in MAPK pathway mediated by autocrine FGF4 which is not blocked by BMP4. It is likely that BMP4 blocks the ESCs differentiation commitment downstream the phosphorylation and activation of ERK. Initially a combination of three inhibitors (3i) was used: SU5402 inhibiting FGF receptor tyrosine kinase, PD184352 inhibiting ERK cascade and CHIR99021 inhibiting GSK3. Triple inhibition (3i) was applied in serum-free medium added with B27 and N2 supplements that contain insulin, transferrin and albumin to give cells the optimal condition. In these conditions, when the self-renewal was sustained by such inhibition of FGF, MAPK and GSK3 pathways, ESCs were defined as being in the “ground state” of pluripotency, as they were completely freed from external stimuli (being without LIF) (Weinberger et al., 2016). Then a MEK inhibitor PD0325901 (instead of FGF+ERK inhibitors) together with a GSK3 inhibitor CHIRON99021 (2i) were used and supplemented with LIF. Expansion of ESCs in this 2i/LIF medium was enhanced

compared to 3i-ESCs that have reduced proliferation (Weinberger et al., 2016). 2i/LIF condition permits ESCs derivation not only more efficiently from 129 strain but also from previously called “non-permissive” strains or refractory like the NOD one (Batlle-Morera et al., 2008; Nichols et al., 2009). 2i-ESCs can grow on gelatin coated plates forming ball-like structures. However to let them spread it is needed to pre-treat petri dish with laminin (Figure 22) (Tosolini and Jouneau, 2016a).

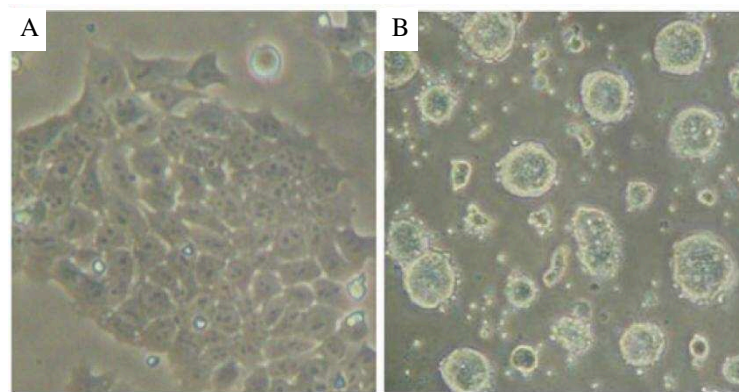


Figure 22: Morphology of 2i-ESCs under the phase-contrast microscope (From Tosolini and Jouneau, 2016a)

(a) ESCs in 2i/LIF condition on plates coated with laminin (spread and attached colonies). (b) ESCs in 2i/LIF condition on plates coated with gelatin (ball-like colonies).

ESCs in 2i/LIF condition formed colonies composed purely of undifferentiated cells compared to serum/LIF condition that induces a mixture of undifferentiated and some differentiated cells. Moreover 2i/LIF leads to an homogenization of expression levels of some pluripotency genes in particular *Nanog* and *Rex-1* that are high in all cells. The ground-naïve state of pluripotency stabilized by the double inhibition (2i) blocks the fluctuation of *Nanog* level to a high homogenous state (Wray et al., 2010). Interestingly once ESCs are cultivated in 2i/LIF medium bimodal genes become unimodal (reducing variability in the cell population) due to the increment of switching rate from *Nanog*-low to -high at the expenses of the transition *Nanog*-high to -low (Singer et al., 2014). However a more recent study addressed the question of heterogeneity and showed by single-cell RNA-seq that serum- and 2i-ESCs both show transcriptional heterogeneity. What differs between the two cell types are only the groups of genes that varies most: while serum-ESCs are more heterogeneous in term of expression of pluripotency genes, 2i-ESCs showed more prominent variability for genes related to cell cycle (Kolodziejczyk et al., 2015). Indeed, although Coronado and colleagues

(2013) showed that 2i-ESCs had a cell-cycle regulation and timing closely related to conventional serum-ESCs with a rapid doubling rate and short G1 phase, a more recent paper showed slower doubling time of ESCs in 2i compared to serum condition (Kolodziejczyk et al., 2015).

The first complete transcriptome comparison of ESCs cultured in serum/LIF and 2i/LIF has been made by Marks and colleagues (2012) where they showed that the core pluripotency factors and other stem cell maintenance genes were expressed at similar levels: Pou5f1, Nanog, Sox2, Esrrb, Klf2, Klf4 and Tbx3. The up-regulated genes in 2i-ESCs concerned metabolic processes and cell-cycle regulation. Tcf1 is also up-regulated and is implicated in self-renewal. On the contrary up-regulated genes in serum-ESCs are linked to developmental processes (especially ectoderm and mesoderm markers) such as Pax6, T (Brachyury), Runx1, Runx3, Sox18, Cdx4 and Tail1 which are undetectable or very lowly expressed in 2i-ESCs. Genes related to germline or endoderm lineages are expressed similarly in 2i- or serum-ESCs. These transcriptome profiles are inter-convertible after switching culture conditions (Marks et al., 2012).

Marks and colleagues have also observed differences at the epigenetic level. Although the global level of H3K27me3 is unchanged, 2i-ESCs present an important reduction of this mark compared to serum-ESCs on promoters of lowly-expressed genes (Marks et al., 2012). The loss of H3K27me3 in 2i- versus serum-ESCs is neither linked to reduced expression of PRC2 component nor increased demethylase expression. Interestingly, while ESCs in serum present around 3000 bivalent (presence of both H3K27me3 and H3K4me3) genes, after 2i adaptation this number drops to less than 1000 due to the reduced presence of H3K27me3 at the promoter level (Marks et al., 2012; Weiner et al., 2016). This epigenomic reorganization is only secondary to the early rewiring of Oct4, Sox2 and Nanog bindings in the genome which are already reconfigured after 24h of 2i induction (Galonska et al., 2015).

At the DNA methylation level, the 2i condition induces a hypomethylated state in ESCs (even more exacerbated in female than in male) which parallels the ICM cell state, while serum (male) ESCs were hypermethylated resembling more post-implantation epiblast E6.5 (Habibi et al, 2013). Compared to male serum-ESCs, 2i-ESCs loose DNA methylation over CpG islands of promoters, enhancers and bivalent loci. However there are some regions that maintain methylation, principally in retroviral sequences (ERVs and IAPs) and to a lesser extent satellite sequences. The hypomethylated state in 2i-ESCs seems to be related to the up-

regulation of *Prdm14* which leads to the down-regulation of DNMTs levels (Yamaji et al., 2013) paralleling the in vivo situation where between pre- and post-implantation epiblast these enzymes are 10-times more expressed (Habibi et al., 2013). More recently it has been shown that the main cause of ESCs DNA demethylation in 2i condition is the reduction of UHRF1 which is necessary for targeting the maintenance DNMT1 at the replication foci (von Meyenn et al., 2016). As for the transcriptome, the DNA methylome is also dynamic and can be reverted/adapted, switching ESCs medium from serum to 2i and vice versa (Habibi et al., 2013).

2.3.4 The primed state of mouse pluripotency: EpiSCs in ActivinA/FGF2

Following the observation that mouse ESCs (mESCs) do not resemble human ESCs (hESCs), two different groups tried to derive mouse pluripotent stem cells similar to human cells. They showed that starting from the late epiblast of a post-implantation mouse embryo at E5.5-6.5, just before gastrulation, it was possible to derive pluripotent stem cells that could be maintained under culture conditions for hESCs, using Activin/Nodal and FGF signaling. To distinguish this new cell type from mESCs they were called epiblast stem cells (EpiSCs) (Brons et al., 2007; Tesar et al., 2007). EpiSCs grow as large, flat, colonies (Figure 23), in contrast to rounded dome-like colonies of mESCs.

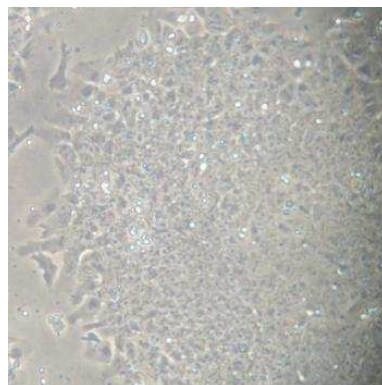


Figure 23: Morphology of EpiSCs under the phase-contrast microscope (Image from the laboratory)
EpiSCs colony in ActivinA/FGF2 condition on plates coated with fibronectin.

Passage in single-cell with trypsin induces death of EpiSCs so a less aggressive treatment with Collagenase is necessary to pass them as small clumps (Brons et al., 2007; Tosolini and Jouneau, 2016b). EpiSCs strictly depend on Activin/Nodal signaling to self-renew, while being independent of LIF. FGF signaling is not strictly required but it improves the overall quality of the culture reinforcing the Activin signaling (Brons et al., 2007; Greber et al., 2010). On the contrary BMP4 condition leads to differentiation of EpiSCs in mesoderm and endoderm.

EpiSCs express Oct4 (Pou5f1), Nanog, Sox2, Ssea-1 as ESCs, however at the molecular level some differences are detected. Even though Oct4 is a shared marker, it is regulated differently because in EpiSCs its expression is mediated by the proximal enhancer (as in post-implantation epiblast), while in ESCs Oct4 is transcribed using the distal enhancer (as ICM cells) (Tesar et al., 2007). The marked difference in Oct4 regulation between EpiSCs and ESCs has been further investigated by Ding and colleagues (2015) using a genome-wide RNAi screen. They have shown that even though the knockdown of many genes has similar effects on Oct4 expression in EpiSCs compared to ESCs, others however have divergent effects. In particular they conclude that Oct4 in EpiSCs is mostly under a complex repressive control, as the knockdown of many genes induces up-regulation of Oct4, rather than down-regulation. Such phenomenon is rarely seen in ESCs (Ding et al., 2015). Moreover EpiSCs are negative for alkaline phosphatase (AP) which is a marker of ESCs and present reduced (or even undetectable) levels for ICM-markers (Rex1, Stella, Esrrb, Grbx2, Tbx3) compared to ESCs (Brons et al., 2007). On the contrary EpiSCs express some late epiblast and early germ layers markers like Fgf5, Nodal, Otx2, Eomes, Foxa2, Gata6, T (Brachyury), Lefty2.

The fact that EpiSCs express the core pluripotency factors but also some early differentiation markers prompted the scientific community to define them as primed pluripotency (Nichols and Smith, 2009). Among the EpiSC-specific genes, Otx2 seems to play a key role in exiting ESCs from pluripotency while protecting EpiSCs from neural fate differentiation (Acampora et al., 2013).

More recent studies have shown that it is possible to derive different types of EpiSCs depending on culture conditions. For example a new type of EpiSCs can be established using FGF2 in combination with IWR1, an inhibitor of Wnt signaling, leading to more homogenous colonies with less spontaneous differentiation and corresponding more specifically to the posterior-proximal epiblast (Wu et al., 2015). Another type of primed EpiSC was obtained

without extracellular signaling only by supplementation with GSK3i and IWR1 which prevent differentiation therefore inducing a less primed state (Kim et al., 2013; Weinberger et al., 2016). However these two novel EpiSCs types have not extensively studied and will not be analyzed further in this thesis.

EpiSCs can make embryoid bodies (EB) and teratomas. However they were not able to integrate when injected into mouse embryos and do not incorporate into the ICM suggesting that pre-implantation embryo is not a compatible environment for EpiSCs (Rossant, 2008). This issue was resolved with the demonstration that EpiSCs were able to integrate a mouse embryo but only at the right developmental window corresponding to their in vivo timing: E5.5-6.5 (Huang et al., 2012).

EpiSCs have a longer G1 phase with 17.5% cell population in this phase, in contrast to 7.8% of naive ESCs (Coronado et al., 2013).

Epigenetically while female ESCs have two active X chromosomes, female mEpiSCs showed only one active X, in line with the random X-inactivation occurring during mouse embryo implantation (Nichols and Smith, 2009).

ESCs can be easily converted in vitro into EpiSCs by adapting them to Activin A and FGF2 culture condition (Buecker et al., 2014; Orkin and Hochedlinger, 2011; Schulz et al., 2014). However in vitro reversion of EpiSCs to ESCs is possible but very inefficient when just medium conditions are changed (Bao et al., 2009; Guo et al., 2009) eliciting the concept of an epigenetic barrier to overcome, such as the re-activation of the X chromosome in female cells or the DNA demethylation. To gain efficiency and bypass the epigenetic barrier it is necessary to over-express one or more naïve factors such as Klf4 (Guo et al., 2009) or to use small molecules inhibitors targeting LSD1, ALK5, MEK, FGFR and GSK3 (Zhou et al., 2010). The reversion from primed to naive pluripotency can be considered as a real de-differentiation process as it induces a reactivation of an X chromosome in female cells, a transcriptional profile completely comparable to ESCs, high capacity to generate blastocyst chimeras and germline transmission (Bao et al., 2009; Guo et al., 2009). More recently it has been shown that inhibition of the KMT MLL1, which induces a global down-regulation of H3K4me1, leads to a very efficient and rapid reversion of primed to naïve mouse pluripotency stressing once more the requirement of a strong epigenetic reorganization in order to “reprogram” EpiSCs into ESCs (Zhang et al., 2016).

EpiSCs showed strong differences from ESCs also at the miRNA level. Indeed one third of the miRNAs (miR) were differentially expressed between ESCs and EpiSCs: the cluster miR320/367 is the most abundant in EpiSCs (and hESCs) indicating a specificity of primed pluripotency. The majority of highly expressed miRs in mESCs are classified as ESC- or gonads-associated miRs, while those preponderant in EpiSC are somatic-type miRs (Jouneau et al., 2012).

Concerning DNA methylation our laboratory and other groups has previously demonstrated that at the global level and genomic distribution EpiSCs are similar to mESCs (Hackett et al., 2013; Veillard et al., 2014), however they showed a more pronounced DNA methylation at germline-related gene-promoters and on CpG islands of gene-promoters that were bivalent (copresence of H3K27me3 and H3K4me3) in ESCs. Globally EpiSCs are even more methylated than the *in vivo* epiblast at E6.5-7.5 (Veillard et al., 2014). A recent study has compared the EpiSCs derived from different stages of post-implantation development (from E5.5 to E8.0) and shows that once derived *in vitro*, the transcriptome of EpiSCs resembles each other regardless their developmental origin (Kojima et al., 2014). They also demonstrate by gene expression profiling and functional *in vivo* integration that EpiSCs display properties of anterior primitive streak.

2.3.5 Comparison between mouse states of *in vitro* pluripotency

Only in recent years some studies finally begin to analyze simultaneously the three different types of mouse *in vitro* pluripotent stem cells and also to compare them to the *in vivo* situation. A first comparison of *in vitro* to *in vivo* pluripotency showed that 2i-ESC share strong transcriptome similarities with early pre-implantation epiblast E4.5 while clustering far from ICM cells of E3.5 (Figure 24) (Boroviak et al., 2014). In contrast serum-ESCs and EpiSCs cluster separately and far away even from late epiblast E5.5 (Figure 24) (Boroviak et al., 2014). This study was based only on the expression of 96 genes enriched for pluripotency genes, lineage markers and signaling pathway genes. The 2i-ESCs transcriptional resemblance of pre-implantation epiblast E4.5 was further confirmed by RNA-seq data (Boroviak et al., 2015)

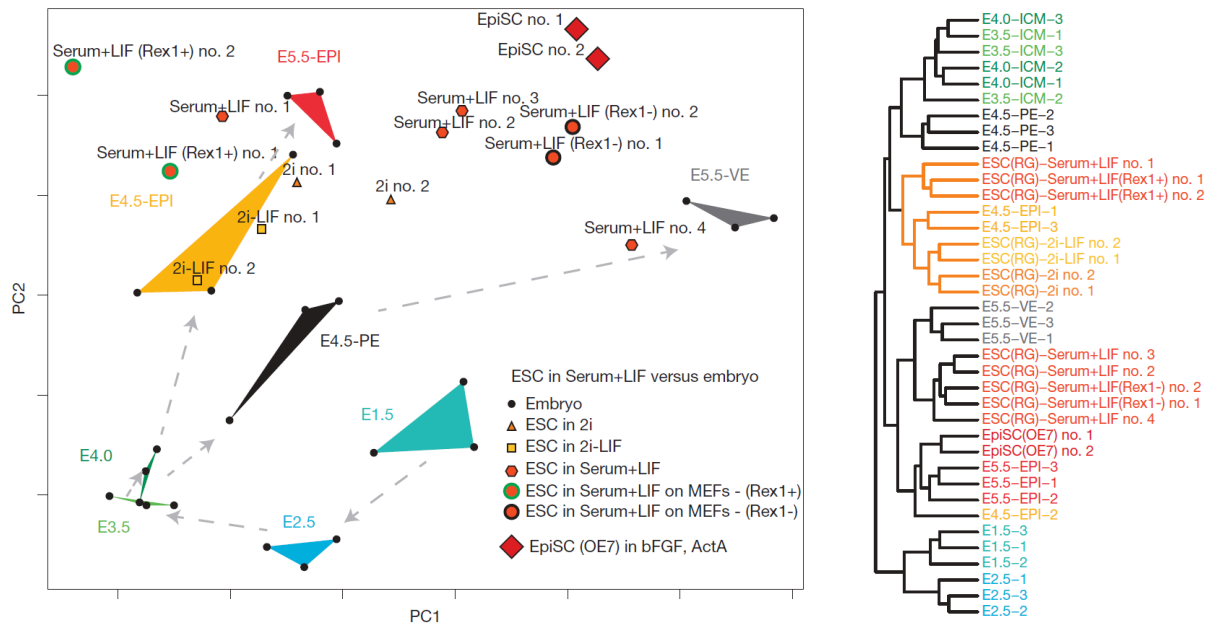


Figure 24: Correlation of gene expression of the different types of mouse pluripotent stem cells to the early embryo stages. (From Boroviak et al., 2014)

Principal Component Analysis (PCA) (Left part) and Hierarchical clustering (Right part) of 2i-ESCs (2i-LIF), serum-ESCs (Serum+LIF), EpiSCs and early embryonic stages for expression of lineage markers and pathway-associated genes assayed by qRT-PCR array.

A very recent review finally made a global comparison of the states of mouse in vitro pluripotency defining which features (molecular, epigenetic and signaling) are considered as naive or primed and attributing them to each cell type (Weinberger et al., 2016) (Figure 25). Serum-ESCs share the majority of naive-configuration properties with 2i-ESCs, however they present also some primed-features with EpiSCs such as the global DNA hypermethylation. Thus Weinberger et al., 2016 considers serum-ESCs as naive even though less naive than 2i-ESCs (Figure 25-26).
















































			Mouse		
Pluripotent cell property	Naive pluripotent cell 	Primed pluripotent cell 	2i, LIF	FBS, LIF	FGF2, Activin A or FGF2, TGFβ
MEK-ERK dependence	No	Yes			
Long-term dependence on FGF2 signalling	No	Yes			
Long-term dependence on TGFβ or Activin A signalling	No	Yes			
Dominant OCT4 enhancer	Distal	Proximal			
H3K27me3 on developmental regulators	Low	High			
Global DNA hypomethylation	Yes	No			
X chromosome inactivation	No	Yes			
Dependence on DNMT1, DICER, METTL3, MBD3	No	Yes			
Priming markers (OTX2, ZIC2)	↓	↑			
Pluripotency markers (NANOG, KLFs, ESRRB)	↑	↓			
Promotion of pluripotency maintenance by NANOG or PRDM14	Yes	No			
Metabolism	OxPhos, Glycolytic	Glycolytic			
Competence as initial starting cells for PGCLC induction	High	Low			
Capacity for colonization of host pre-implantation ICM and contribution to advanced embryonic chimeras	High	Low			
Hypomethylation of promoter and enhancer regions	Yes	No			

Figure 25: Schematic recapitulation of naive and primed features of the different states of mouse pluripotent stem cells (From Weinberger et al., 2016)

Pluripotency properties defined as naïve-specific (orange) or primed-specific (blue). 2i-ESCs (2i,LIF), serum-ESC (FBS,LIF) and EpiSCs (FGF2, Activin A).

In the same review Weinberger et al., 2016 try to find the best molecular and functional criteria that define the naive state in contrast to the primed state (Figure 25). They suggest that the source from where pluripotent stem cells are derived, meaning pre- or post-implantation embryo, cannot be used as a criteria, as the state is dictated by the derivation growth conditions. They also do not consider as a good criteria the fact that naive ESCs can make chimeras when injected into a blastocyst while primed EpiSCs do not however they can integrate a post-implantation epiblast ex-vivo conversely to ESCs (Huang et al., 2012; Nichols and Smith, 2009). They suggest that the molecular "major divider" naive Vs. primed pluripotency is the response to MEK inhibition: primed EpiSCs or hESCs differentiate in absence of MEK pathway while naive ESCs highly tolerate such inhibition and even

consolidate their pluripotency network, as happen in 2i-ESCs (Weinberger et al., 2016) (Figure 26).

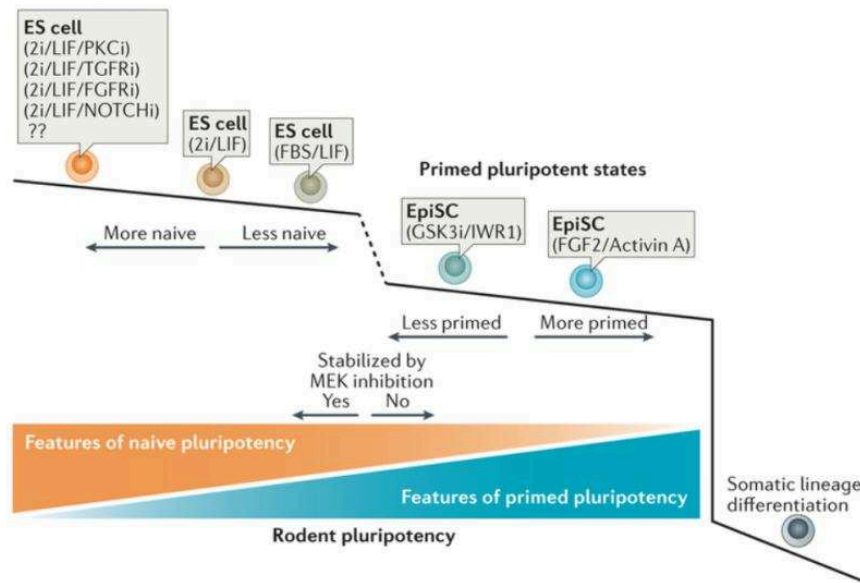


Figure 26: Schematic representation of the different states of mouse pluripotent stem cells (From Weinberger et al., 2016)

Model explaining the spectrum of naive and primed pluripotency. According to Weinberger et al., 2016 the major criterion discriminating is their ability to maintain and stabilize their pluripotent state upon blockade of MEK activity. Naive and primed pluripotent cells cultured in different conditions have different features and varying degrees of naivety or priming.

2.4 In vitro human pluripotency

2.4.1 The primed state of human pluripotency: conventional hESCs

Many years after the establishment of mouse ESCs line from a mouse blastocyst, Thomson and colleagues in 1998 derived for the first time human ESCs from a blastocyst. The hESCs that they obtained had a flat colony morphology, express stage-specific embryonic antigen (SSEA)-3 and -4, TRA-1-60 and -81 (Thomson et al., 1998). However, hESCs did not expressed SSEA-1, which on the contrary is expressed in mouse ICM cells and mESCs, giving first insights about the interspecies differences between human and mouse. Another striking difference with the mouse was that hESCs cannot be clonally expanded starting from a single-cell, but must be passaged in clumps as EpiSCs. Pluripotency of hESCs could not be assessed by chimera assay for obvious bioethical issues, but was proven by teratoma formation after injection in immunodeficient mice. Human ESCs were strictly dependent on

feeders but independent of LIF, conversely they are stimulated by TGF β /Activin A and FGF signaling (Thomson et al., 1998).

In addition primed hESCs like EpiSCs do not incorporate in mouse blastocyst. Another similarity between mEpiSCs and hESCs is that female cells already present one inactive chromosome X while female mESCs have both active X. In spite of all these differences hESCs (and mEpiSCs) retained the expression of OCT4, SOX2 and NANOG. In contrast to mESCs which relies on Oct4 distal enhancer (DE), mEpiSCs and hESCs depend on the proximal one (PE). Although sharing many properties, hESCs appear to differ from mEpiSCs as FGF2/ERK inhibition does not influence Nanog in mEpiSCs while rapidly down-regulates NANOG in hESCs. Moreover hESCs do not express FGF5 that is a key mEpiSC-marked but on the contrary expresses the ICM-marker REX1 (De Los Angeles et al., 2012). Interestingly hESCs stained positive for Alkaline Phosphates (AP) as mESCs (Thomson et al., 1998), while mEpiSCs do not. Altogether, hESCs are closer to the primed pluripotent state of mouse EpiSCs rather than to the naïve mESCs. Interestingly, the same characteristics as hESCs have been obtained by reprogramming human fibroblasts to induced pluripotent stem cells hiPSCs using the “Yamanaka factors”: OCT4, SOX2, KLF4 and c-MYC (Takahashi et al., 2007)

2.4.2 The naïve state of human pluripotency: naïve hESCs

Based on the parallel of the two states of pluripotency in mouse and on the discovery that small molecules inhibitors were able to stabilize naïve pluripotency of historically “non-permissive” mouse strains as well as rat species, many groups world-wide tried to find a way to obtain naïve human pluripotent stem cells (Hanna et al., 2010; Gafni et al., 2013; Chan et al., 2013; Ware et al., 2014; Valamehr et al., 2014; Theunissen et al., 2014; Takashima et al., 2014; Duggal et al., 2015; Chen et al., 2015).

The variety of conditions to obtain the naïve-hESCs is summarized in the (Table 1).

	Thompson et al.,1998 Conventional (primed)	Hanna et al., 2010 2i/LIF	Gafni et al., 2013 NHSM	Chan et al., 2013 3iL	Valamehr et al., 2014 FMM	Ware et al., 2014 2iF	Theunissen et al., 2014 5i/L/A	Takashima et al., 2014 / 2i/L+Go	Duggal et al., 2015	Chen et al., 2015
Inhibitors		MEKi	MEKi	MEKi	MEKi	MEKi	MEKi	MEKi	MEKi	MEKi
		GSK3i	GSK3i	GSK3i	GSK3i	GSK3i	GSK3i	GSK3i	GSK3i	GSK3i
		Forskolin (FK)							Forskolin (FK)	
			JNKi							
			P38i							
			PKCi					PKCi (Go6983)		
			ROCKi		ROCKi					
				BMPi						
						HDACi				
							BRAFi			
Growth factors	bFGF		bFGF	bFGF	bFGF	bFGF			bFGF	
	TGFb/ActivinA		TGFb	TGFb			ActivinA			
		hLIF	hLIF	hLIF	hLIF		hLIF	hLIF	hLIF	
	(MEF)		MEF	MEF			MEF	MEF	MEF	(MEF)
Medium Base	20% KSR ?	N2B27	20% KSR	mTESR1	20% KSR	20% KSR	N2B27	N2B27	20% KSR	20% KSR
Transient expression		KLF4						KLF4		STAT3
		KLF2								
								NANOG		
OCT4 enhancer	Proximal	Distal	Distal			Distal	Distal	Distal		Distal
Bivalent domains	high		low	low		low	low			low
DNA methylation	high		low					low	low	low
X chromosome	inactive	pre-inactive	pre-inactive	inactive	pre-inactive	pre-inactive	inactive	eroded	pre-inactive	partial-reactivation
Metabolism	glycolytic		mitochondrial			mitochondrial	mitochondrial	mitochondrial		mitochondrial
Contribution mouse chimera	no		yes				no	yes		

Table 1: Recap chart of the differences between primed-hESCs and the different attempts to obtain naive-hESCs. From Hanna et al., 2010; Gafni et al., 2013; Chan et al., 2013; Ware et al., 2014; Valamehr et al., 2014; Theunissen et al., 2014; Takashima et al., 2014; Duggal et al., 2015; Chen et al., 2015 and Weinberger et al., 2016.

Huang and colleagues (2014) compared all the different types of naive-hESCs described up to 2014 at the transcriptomic level and also made a comparison with mouse pluripotency and with early human embryo developmental stages. They performed a system biology approach called Weighted Gene Co-expression Network Analysis (WGCNA) and discovered that there are large transcriptome changes between each independently established naive hESCs. Comparison with mouse pluripotency showed that the majority of naive-hESCs do not significantly overlap with 2i/LIF mESCs with the exception of those obtained by Takashima et al., 2014 and Theunissen et al., 2014 (Figure 27) (Huang et al., 2014). Early human embryo comparison leads to similar results because naïve-hESCs generated by Takashima et al., 2014 and Theunissen et al., 2014 have the highest resemblance to the expression profile of human blastocyst (Vassena et al., 2011; Xie et al., 2010; Yan et al., 2013) (Figure 27) (Huang et al., 2014). Primed hESCs on the other hand significantly overlap with mEpiSCs and early passages ICM-outgrowths, suggesting a rapid adaptation to in vitro culture conditions. Interestingly they also share some gene modules (cell-cycle and mitosis) with pre-EGA

embryos from 1- to 4-cell stages (Figure 27) (Huang et al., 2014). The only consensus module shared by all naive-hESCs significantly overlaps with 8-cell and morula stages suggesting that naive-hESCs may have some cellular and metabolic features of post-EGA human blastomers (Figure 27) (Huang et al., 2014).

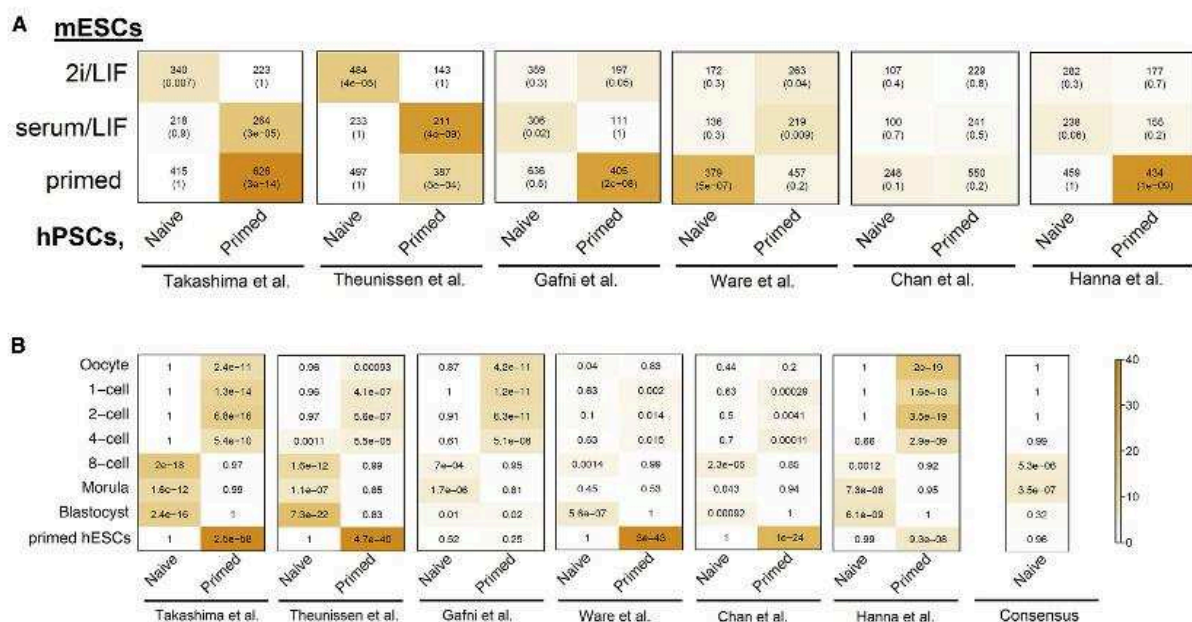


Figure 27: Transcriptional network comparison of primed-hESCs and the different attempts to obtain naive-hESCs with the different mouse pluripotent stem cells and with human pre-implantation embryos (From Huang et al., 2014)

(A) Heatmap showing the significance of gene network overlaps using weighted gene co-expression network analysis (WGCNA) between mouse 2i-ESCs (2i/LIF), serum-ESCs (serum/LIF) and EpiSCs (primed) modules (y axis) with naive and primed hESCs from six different methods (x axis). Number of intersecting genes and the p-value of the intersection are indicated. (B) Heatmap showing the significance of gene network overlaps using weighted gene co-expression network analysis (WGCNA) between human pre-implantation embryos (x axis) and six different methods for producing naive hESCs (y axis). The average p-value (geometric mean) from gene intersects of naive and primed modules with three separate human pre-implantation data sets (Vassena et al., 2011; Xie et al., 2010; Yan et al., 2013) is indicated. Color legend represents $-\log_{10}$ p-value based on the hypergeometric test.

To conclude it seems that there are multiple routes to induce naive human pluripotency: strong inhibitions of MEK in NHSM (naive human stem cell medium) (Gafni et al., 2013), inhibition of BMP signaling with 3iL (Chan et al., 2013), reversed toggle with HDACi (Ware et al., 2014), suppression of MAPK/b-RAF and SRC in 5i/L/A (Theunissen et al., 2014), inhibition of PKC (Takashima et al., 2014) and their combination which activates PI3K/AKT/mTOR (Duggal et al., 2015) and forced activation of STAT3 with 2iL (Chen et al., 2015).

3 HETEROCHROMATIN ORGANIZATION IN MOUSE PLURIPOTENCY

3.1 Chromatin organization during early mouse embryonic development

In mammals, epigenetic modifications are globally rearranged after the fertilization of the egg by sperm. These two gametes carry unique epigenetic signatures in terms of DNA methylation, H3K9me3 and H3K27me3, and a peculiar nuclear organization of pericentromeric sequences. The nuclear and epigenetic state of sperm and egg is reprogrammed by the oocyte's cytoplasm during the first cell-cycles of a developing embryo (Beaujean, 2014). These modifications are concomitant to the gradual loss of totipotency of the embryo and the acquisition of pluripotency of epiblast cells.

3.1.1 Dynamic organization of constitutive heterochromatin during early mouse embryo development

Just after fertilization, at the late 1-cell stage, pericentromeric regions are organized in rings or “shell”-like structures around nucleolar-precursor bodies (NPBs) with centromeric spots associated with the periphery of NPBs (Figure 28) (Aguirre-Lavin et al., 2012; Probst et al., 2010).

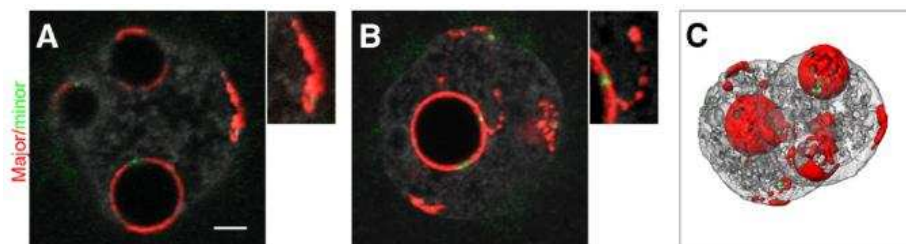


Figure 28: Organization of major and minor satellite sequences at the late 1-cell stage of mouse embryo (From Aguirre-Lavin et al., 2012).

(A, B) DNA-FISH for major (red) and minor (green) satellites repeats. In both pronuclei pericentromeric sequences organized mainly around nucleolar-precursor bodies (NPBs). (C) Three-dimensional reconstruction of the signals of both pronuclei (A, B). DAPI signal is shown in grey.

At the 2-cell stage when the activation of the embryonic genome occurs, the genome is reorganized and the major satellite repeats switched from a ring to a spherical patched structure (Figure 29) (Aguirre-Lavin et al., 2012; Probst et al., 2010). At the 4-cell stage classic chromocenter structures appeared as compact mass of pericentromeric sequences

surrounded by individual centromere. By the blastocyst stage the general heterochromatin nuclear organization is very similar to that of somatic cells (Aguirre-Lavin et al., 2012; Guenatri et al., 2004; Probst et al., 2010).

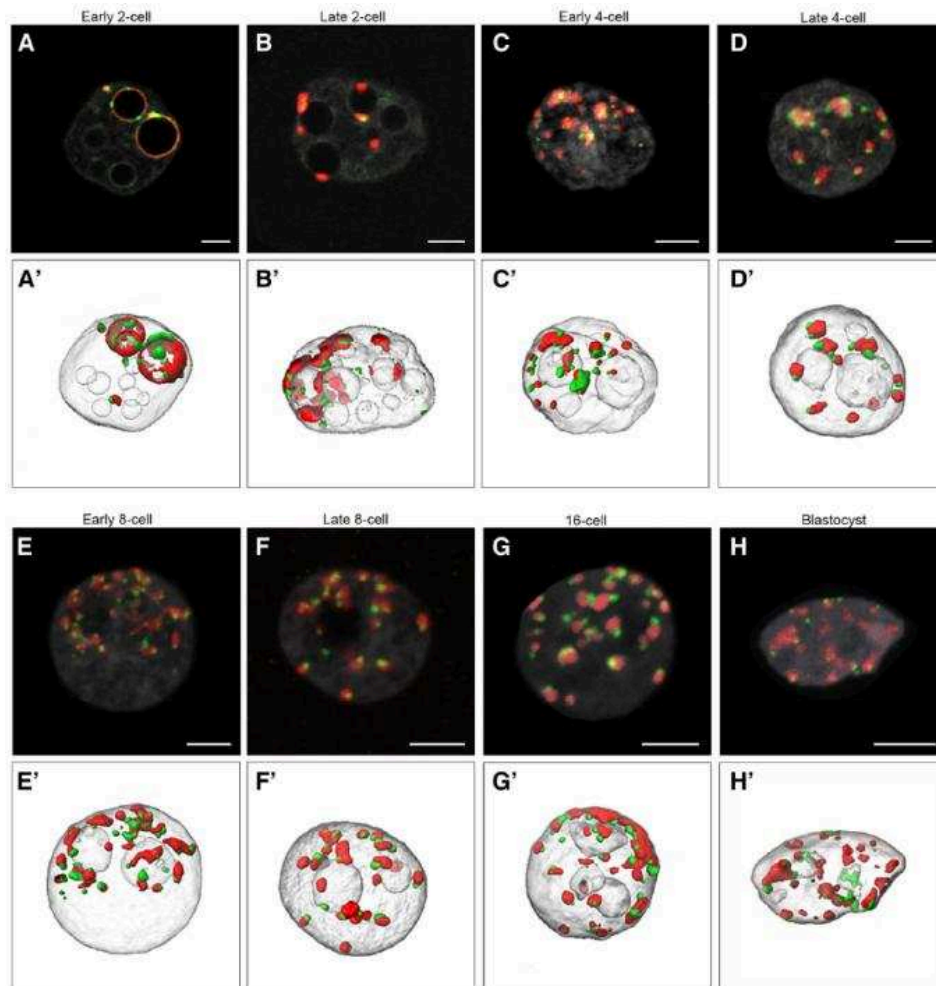


Figure 29: Organization of major and minor satellite sequences during early embryo development (From Aguirre-Lavin et al., 2012).

(A-H) DNA-FISH for major (red) and minor (green) satellites repeats of each pre-implantation developmental stage. Starting from the late 2-cell stage embryo, major and minor satellite migrates from the nucleolar-precursor bodies (NPBs) clustering together in chromocenters. (A'-H') Three-dimensional reconstruction corresponding to the upper image. DAPI signal is shown in grey.

Such three-dimensional heterochromatin reorganization during early mouse embryo development can be impaired by the loss of the histone chaperone CAF-1/p150. Chromatin Assembly Factor 1 (CAF-1) is a three subunit complex made by p150, p60 and p48, that promotes histone H3 and H4 deposition at newly synthesized DNA during replication (Kaufman et al., 1995). The p150 subunit guarantees the pool of replication-specific HP1 α at

replication sites in pericentromeric heterochromatin during mid- and late-S phase (Quivy et al., 2004). Mutation of this CAF-1 subunit arrests development between 8- and 16-cell stages. The nuclei of mutant embryos show diffused HP1 α staining and barely detectable DAPI-dense foci compared to wild-type, with some enrichment only at the nuclear periphery and around nucleoli (Houlard et al., 2006) (Figure 30). CAF-1 seems to be a key player of the reorganization of pericentromeric heterochromatin which is necessary for embryo survival. In the same way, a burst of non-coding major satellite transcripts occurring at the 2-cell stage (described later on in this thesis) is also crucial for heterochromatin reorganization, chromocenter formation and further development (Probst et al., 2010)

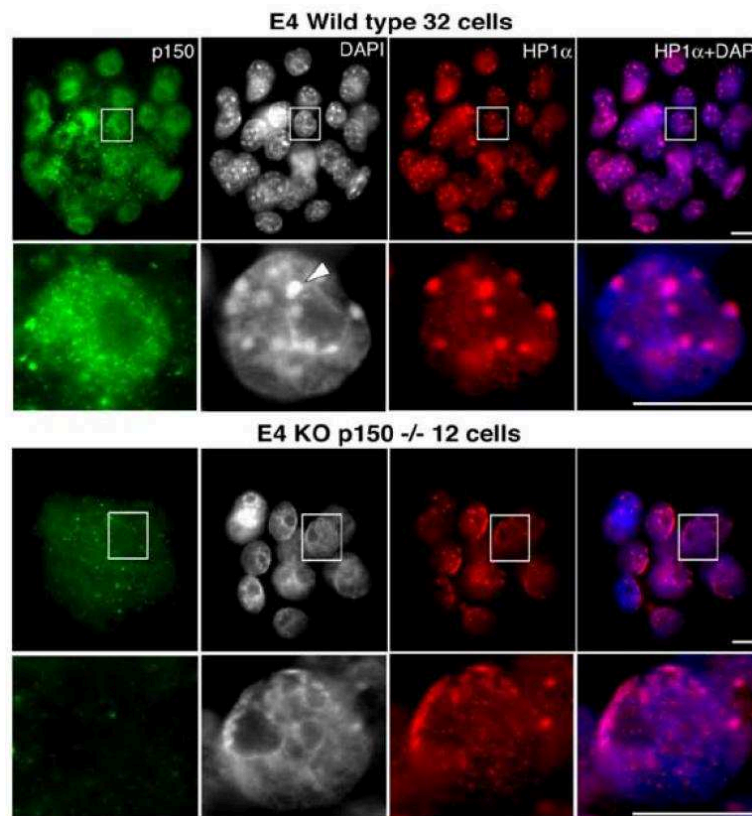


Figure 30: Comparison between DAPI and HP1 α distribution in wild-type and p150 $-/-$ mouse E4 embryos (From Houlard et al., 2006)

Immunostaining for CAF-1/p150 (green) and HP1 α (red) in E4 wild-type embryos (upper panel) and E4 p150 $-/-$ embryos, with DAPI counterstaining for DNA (grey/blue). HP1 α and DAPI foci-enrichment are lost in the knock-out (KO) condition.

3.1.2 DNA methylation dynamics of heterochromatin in early mouse embryo development

DNA methylation is highly reprogrammed during early embryo development with a massive global demethylation starting from hypermethylated gametes (Figure 31A). The methylation levels are at the lowest by the early blastocyst stage (Kohli and Zhang, 2013). Globally 5-meC staining on single blastomeres for each developmental stage reveals a reorganization of this epigenetic mark with a clear foci-enrichment that starts at the 8-cell stage (Li et al., 2016) (Figure 31B).

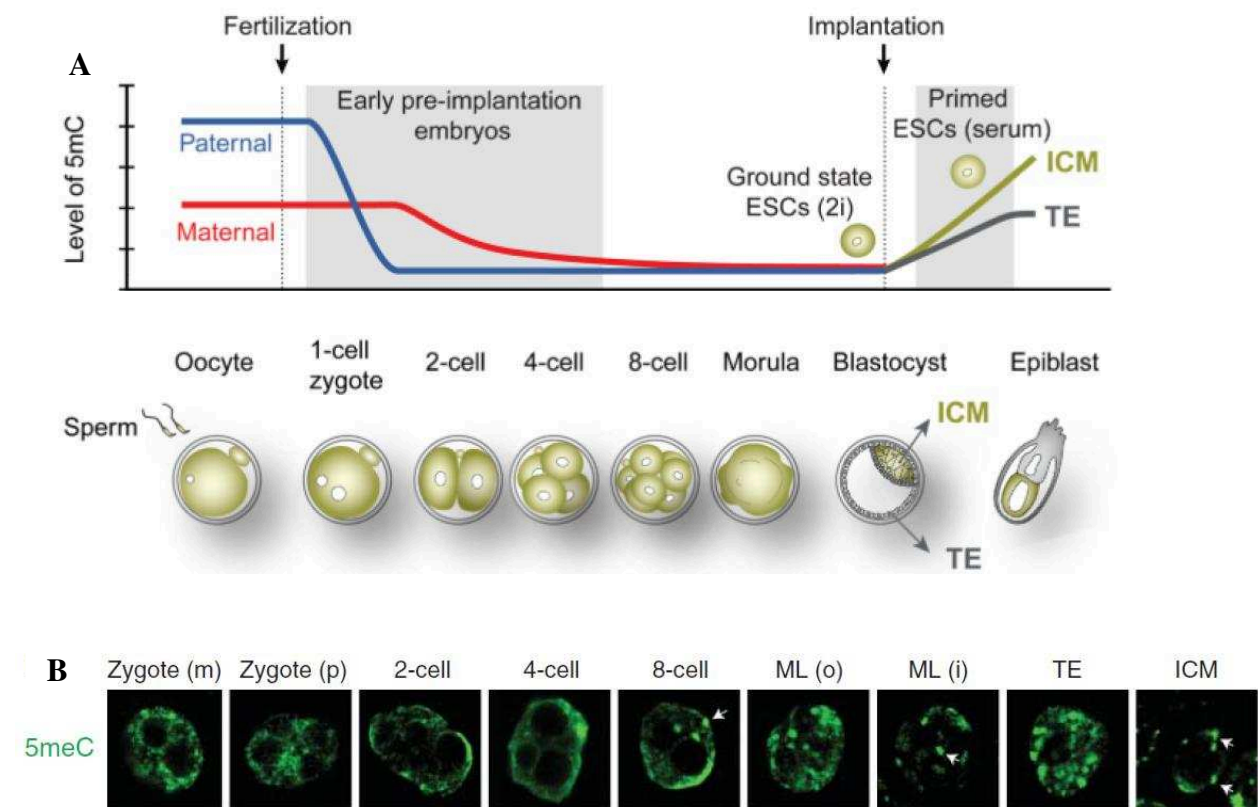


Figure 31: DNA methylation dynamics during early mouse embryo development (From Wu and Zhang, 2014 and Li et al., 2016)

(A) Schematic representation of the dynamics of 5-meC paternal and maternal during early embryo development. (B) Immunostaining of 5-meC (green) in individual maternal pronucleus (zygote (m)), paternal pronucleus (zygote (p)), and nuclei from 2-cell, 4-cell, 8-cell embryos, the outer (ML (o)) and inner (ML (i)) cells of morula, trophoectoderm (TE) and inner cell mass (ICM) of blastocysts. White arrows show representative examples of 5meC-intense staining foci.

Concerning major and minor satellites, they are hypermethylated in all adult tissues except for sperm and metaphase II-arrested egg that possess a unique hypomethylated state (Figure 32B) (Yamagata et al., 2007). This hypomethylated status persisted through all the pre-implantation development (Figure 32C) (Yamagata et al., 2007). Interestingly the paternal genome undergoes a rapid demethylation after fertilization (Salvaing et al., 2012) and methylated CpGs in major and minor satellites drop from 40/50% of epididymal sperm to 20% of zygote (Figure 32) (Yamagata et al., 2007). The de novo DNA methylation at pericentromeric and centromeric region occurs after the implantation of the blastocyst thus concomitant with the global genome-wide remethylation (Figure 32C) (Yamagata et al., 2007).

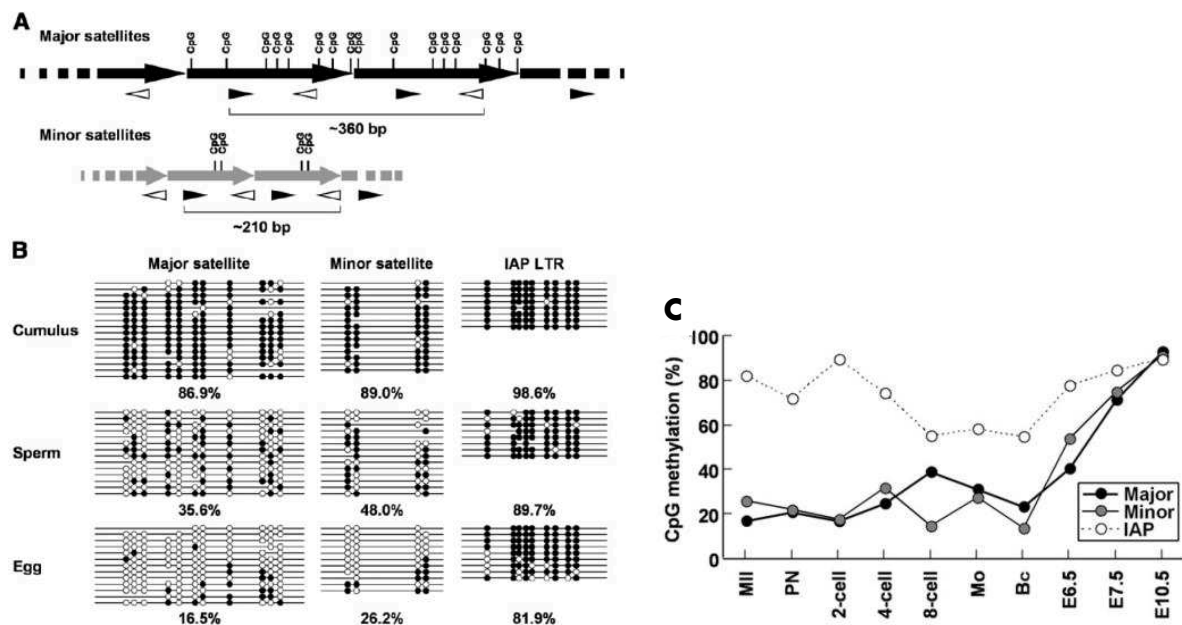


Figure 32: Dynamics of DNA methylation on major and minor satellite repeats from gametes to late mouse embryonic development (From Yamagata et al., 2007).

(A) CpG loci of the major and minor satellites. Closed and open arrowheads represent the oligonucleotide primers used for bisulfite sequencing. (B) Methylated patterns analyzed by bisulfite sequencing of major and minor satellites and IAP LTR in mouse cumulus cells, caudal epididymal sperm, and metaphase II-arrested eggs. Black and white dots indicate the methylated and unmethylated status of CpG sites, respectively. (C) CpG methylation kinetics analyzed by bisulfite sequencing of major satellites, minor satellites and intracisternal A particles (IAPs) during pre- and post-implantation mouse embryo development. Data are presented as % of methylated CpG sites per total CpG sites. Metaphase II-arrested egg (MII), pronuclear-stage embryo (PN), morula (Mo) and blastocyst (Bc).

3.1.3 H3K9me3 dynamics in early mouse embryo development

H3K9me3 strongly stain the female pronucleus at 1-cell stage with a particular intensity for NPBs periphery (pericentromeric heterochromatin), while being absent from the male one. This asymmetry between the two parental genomes is still visible at the 2-cell stage with only half of each blastomere nuclei marked by H3K9me3 and maintained till the 4-cell stage (Beaujean, 2014; Puschendorf et al., 2008). Globally all along the early mouse embryo development H3K9me3 is enriched in foci (Figure 33) and also in E5.5 epiblast cells (Rugg-Gunn et al., 2010).

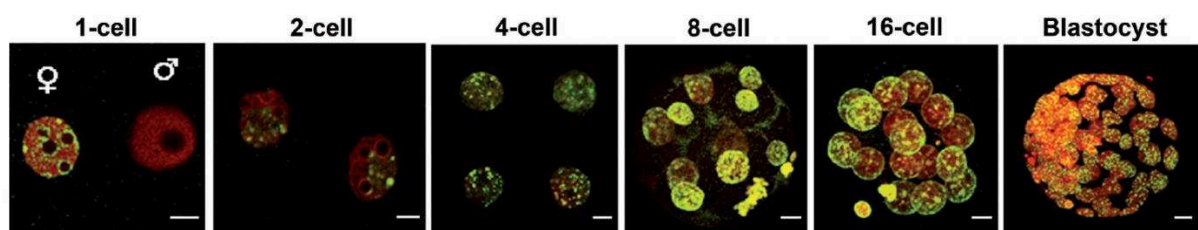


Figure 33: Subnuclear distribution of H3K9me3 during pre-implantation mouse embryo development (From Beaujean, 2014)

Immunostaining for H3K9me3 (green) and DNA (red) in pre-implantation embryos from one-cell to blastocyst stage. Only the maternal pronucleus is labeled at the 1-cell stage and at the 2-cell stage the staining is asymmetric.

3.1.4 H3K27me3 dynamics in early mouse embryo development

H3K27me3 is detected in both pronucleus at the 1-cell stage but interestingly it is enriched in pericentromeric heterochromatin (ring-structures) only in paternal pronucleus (Puschendorf et al., 2008; Santenard et al., 2010) (Figure 34A). This asymmetric distribution is resolved already at the 2-cell stage once the chromocenters start to be formed and H3K27me3 showed a foci-enrichment distribution pattern at least up to 8-cell stage (Puschendorf et al., 2008) (Figure 34B). The H3K27me3 subnuclear distribution particularly in ICM and post-implantation epiblast has not been extensively investigated except for the inactive X in female embryos.

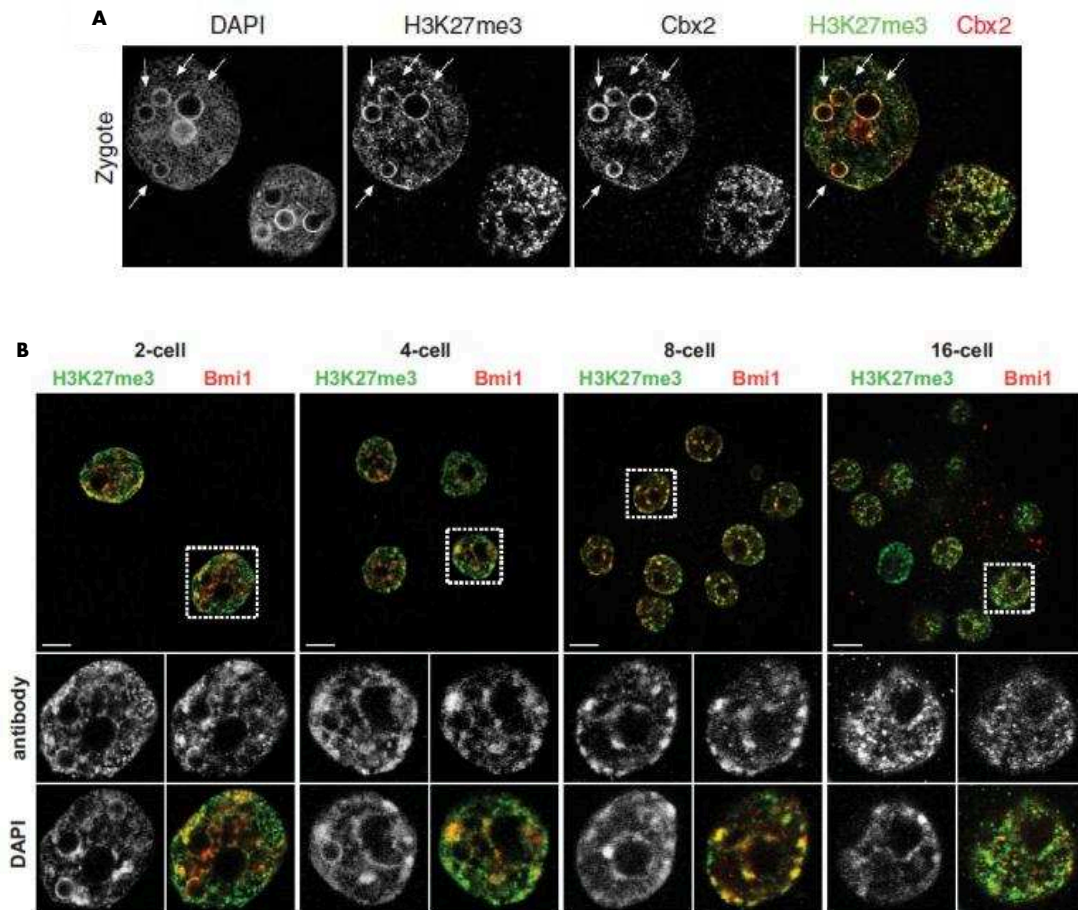


Figure 34: H3K27me3 nuclear distribution during early stages of mouse embryo development (From Puschendorf et al., 2008)

(A) Immunostaining for H3K27me3 (green) in late zygotes showing that only paternal constitutive heterochromatin is labeled (arrows). (B) Immunostaining for H3K27me3 (green) from 2-cell to 16-cell stage showing substantial colocalization with DAPI-dense foci till the 8-cell stage. At 16-cell stage H3K27me3 seems to be less enriched at constitutive heterochromatin.

3.1.5 Satellite non-coding transcription in early mouse embryo development

Knowing the great rearrangement of pericentromeric satellite repeats occurring during early embryo development, Probst and colleagues (2010) analyzed the transcription status of these repeats. They observed an initial transcription at the 1-cell stage followed by a burst of major satellite transcription at the 2-cell stage (concomitant to the reorganization of pericentromeric regions from rings around NPBs to compacted chromocenters) and finally down-regulated at the 8-cell stage (Figure 35) (Probst et al., 2010). At the 1-cell stage the transcription has only a paternal genome origin and interestingly RNA-FISH spots are found from DAPI-bright ring structure not colocalizing with H3K27me3-enriched ones (Probst et al., 2010). The transcription occurs from both senses with different dynamics: the forward ncRNAs

accumulate during S phase and peaks at early 2-cell stage and originate predominantly from paternal genome, while the reverse ones peak at the late 2-cell stage and are expressed widely from paternal as well as maternal embryonic genome (Probst et al., 2010). Knock-down of reverse major satellite transcripts is sufficient to impede nuclear reorganization of pericentromeric heterochromatin into chromocenters and development beyond the 2-cell stage (Casanova et al., 2013).

No similar study has been made for minor satellite non-coding transcripts till now.

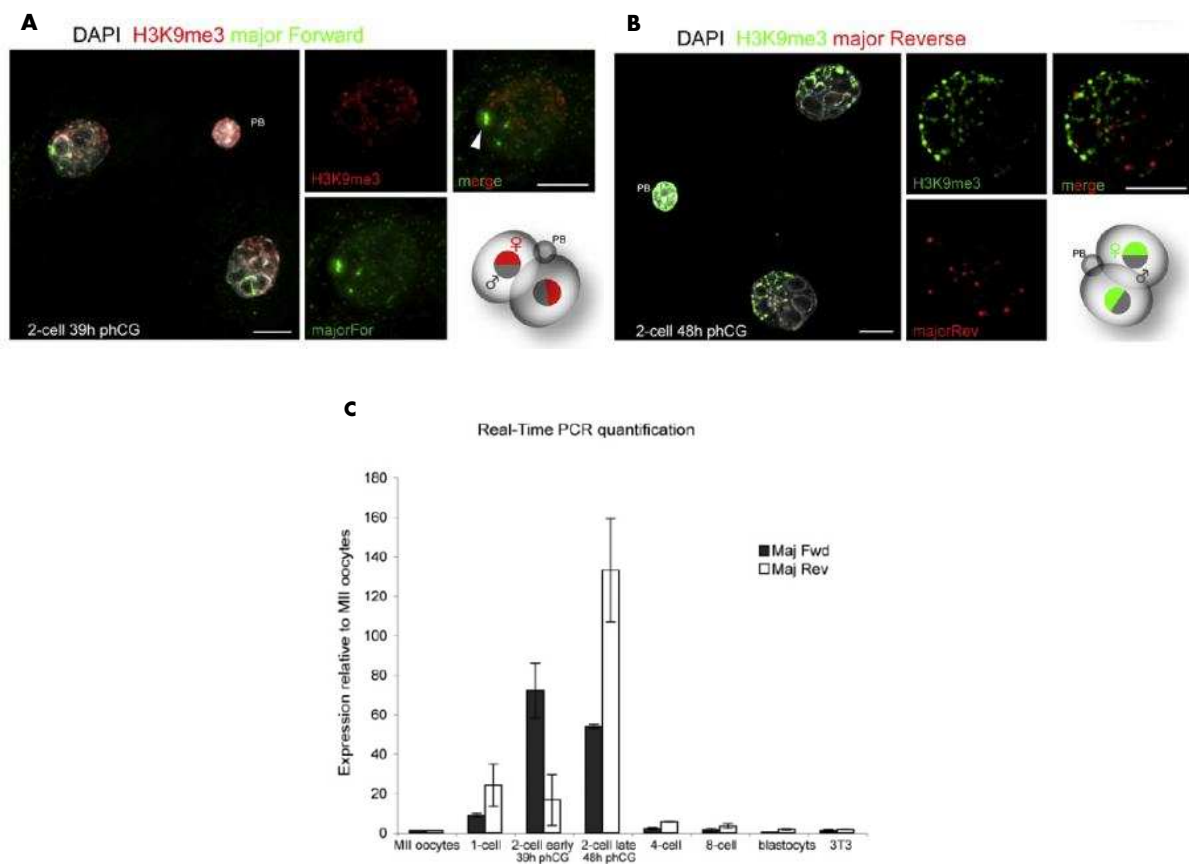


Figure 35: Major satellite repeat transcripts during pre-implantation mouse embryo development (From Probst et al., 2010).

(A) Immuno-RNA-FISH for H3K9me3 (red) and major forward satellites RNAs (green) at early two-cell embryos (39h phCG). Nascent transcripts have only paternal origin (not marked by H3K9me3). (B) Immuno-RNA-FISH for H3K9me3 (green) and major reverse satellites RNAs (red) at late two-cell embryos (48h phCG). Transcripts originate from both maternal and paternal genome. (A,B) DNA is counterstained with DAPI (gray). PB, polar body. (C) Reverse transcription with strand-specific primers followed by quantification of major transcripts using qPCR. Transcript levels relative to MII oocytes (set to 1). Major satellites transcription begins at late 1-cell stage, forward transcripts (Maj Fwd) peaks at early 2-cell stage while reverse one (Maj Rev) at late 2-cell stage. At 4-cell stage major satellites transcription is shut-down as blastocyst or 3T3 mouse fibroblasts.

Taken together at the blastocyst stage, when ESCs are derived, *in vivo* pluripotent ICM cells are characterized by a compacted organization in chromocenters of pericentromeric major satellites which are silenced, hypomethylated and strongly enriched by H3K9me3 but likely not by H3K27me3.

3.2 Chromatin plasticity in mouse *in vitro* pluripotency

Pluripotent stem cells and especially naïve ESCs have a special and unique epigenetic signature that reflects their broad developmental potential. Most of the epigenetics marks including DNA methylation are not required for survival and maintenance of their self-renewal (Meissner, 2010). However most of naïve ESCs lacking epigenetics marks (such as DNA methylation, H3K27me3 or H3K9me3) show impaired developmental and differentiation potential (Déjardin, 2015). Serum-ESCs are also characterized by hyperdynamic, loosely bound or soluble fraction of chromatin proteins. Such plasticity is necessary particularly during differentiation for the reshaping of the genome architectural organization especially the heterochromatin (Meshorer et al., 2006). In serum-ESCs, heterochromatin is organized in larger but fewer H3K9me3 domains, which become smaller, more abundant and hyper-condensed as cells differentiate (Figure 36) (Meshorer et al., 2006).

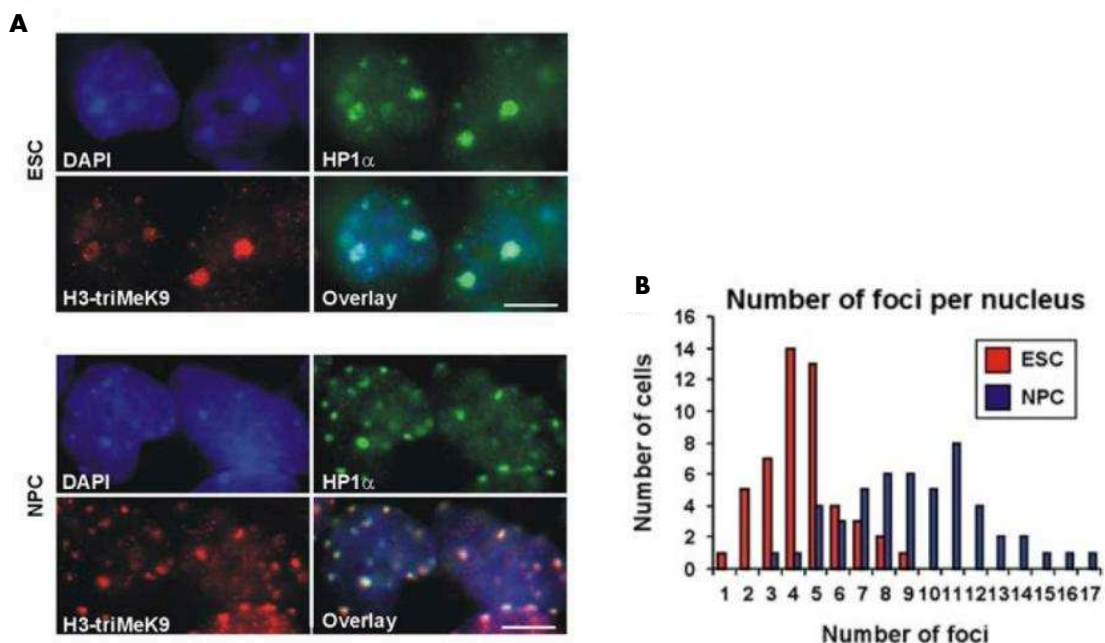


Figure 36: Constitutive heterochromatin organization in serum-ESCs Vs. NPCs (From Meshorer et al., 2006)

(A) Immunostaining for H3K9me3 (H3-triMeK9, red) and HP1 α (green) in serum-ESCs (top) or neural progenitor cells (NPCs, bottom), DNA counterstaining with DAPI staining (blue). (B) Distribution of heterochromatin foci number per nucleus in ESCs (red bars) or NPCs (blue bars).

In addition to the global decondensation, loose binding of chromatin proteins and enrichment in active histone modifications, serum-ESCs are characterized by a pervasive and elevated global transcription (Efroni et al., 2008) as total RNA and mRNA (which represent 5–10% of total RNA) levels, normalized to DNA content, are almost 2-fold higher in serum-ESCs compared to neural progenitor cells (NPCs). Interestingly 2i-ESCs have about half less RNA compared to serum-ESCs (Kolodziejczyk et al., 2015). Transcription of all repetitive non-coding elements including major and minor satellite repeats, is significantly higher in serum-ESCs than differentiated cells where they are normally silenced, as shown by qRT-PCR and RNA-FISH (Figure 37) (Efroni et al., 2008).

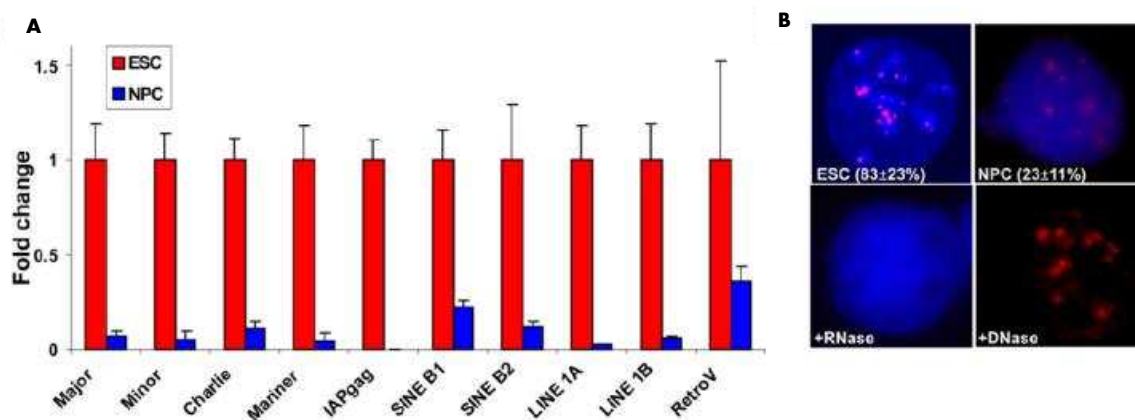


Figure 37: Pervasive transcription of serum-ESCs compared to NPCs (From Efroni et al., 2008). (A) qRT-PCR of repeat sequences, transposable and retroviral elements in serum-ESCs (red) and neural progenitor cells (NPCs, blue). (B) RNA-FISH for the major satellite transcripts (red) in serum-ESCs and NPCs. When serum-ESCs were pretreated with RNase A signal was abolished (+ RNase), while DNase I treatment retained the signal (DNase).

3.2.1 Chromatin bivalency in naïve mESCs

ESCs present a peculiar chromatin signature with the co-presence of an active histone mark H3K4me3 and a repressive histone mark H3K27me3 called bivalency (Azuara et al., 2006; Bernstein et al., 2006). These bivalent domains are largely enriched in serum-ESCs compared to differentiated cells and the genes harboring this combination of epigenetic marks encode for transcriptional factors that play roles in embryonic development and lineage specification (Bernstein et al., 2006). Bivalency is associated with silencing of these developmentally-related genes in serum-ESCs, preserving at the same time the potential to become rapidly activated upon differentiation. Interestingly these genes are also replicated in early S phase like active regions even though they are not expressed in serum-ESCs (Azuara et al., 2006). During initiation of differentiation bivalent genes are resolved keeping only one of the two

histone modifications (Bernstein et al., 2006). In PRC2-deficient ESCs, it has been observed a premature expression of these genes as they lose the repressive histone mark H3K27me3 (Azuara et al., 2006), further showing that lineage-specification genes are already primed to be transcribed in serum-ESCs.

CpG-rich promoters are marked by H3K4me3 in ESCs and 22% of them also present H3K27me3 and are lowly expressed (Mikkelsen et al., 2007). Interestingly bivalent domains were also found in human ESCs and largely overlap with the mouse ones (Voigt et al., 2013).

All these studies described in this paragraph have been made in serum-ESCs, however different studies have shown that after ESCs adaptation to 2i condition, the number of bivalent genes drastically drops (Marks et al., 2012; Weiner et al., 2016). No available data exist at the moment to evaluate if bivalency is maintained in EpiSCs or if it is lost as in differentiated cells.

3.2.2 Heterochromatin organization in mouse in vitro pluripotency

A biophysical study has shown that in serum-ESC, nuclei of Low-NANOG (LN) cells are more deformable and their chromatin is less condensed than the High-NANOG (HN) nuclei counterparts. Conversely to what one would expect in the more pluripotent ESCs chromatin is more compact while the ones that are primed for differentiation it is more dynamic and decondensed. In support of these data LN serum-ESCs present a distribution of H3K9me3/HP1 α more diffuse than in HN-ESCs (Chalut et al., 2012).

A molecular link between states of pluripotency and heterochromatin organization has been recently made by Novo and colleagues (2016) demonstrating the role of the transcription factor NANOG in the pericentromeric architecture. They have found that the absence of NANOG induces heterochromatin compaction in mutant ESCs but also in wild type-EpiSCs which lowly express Nanog (Novo et al., 2016). EpiSCs, as well as E5.5 epiblast, have been previously shown to have a more compacted heterochromatin compared to serum-ESCs and ICM E3.5 by electron microscopic imaging (Ahmed et al., 2010) (Figure 38A).

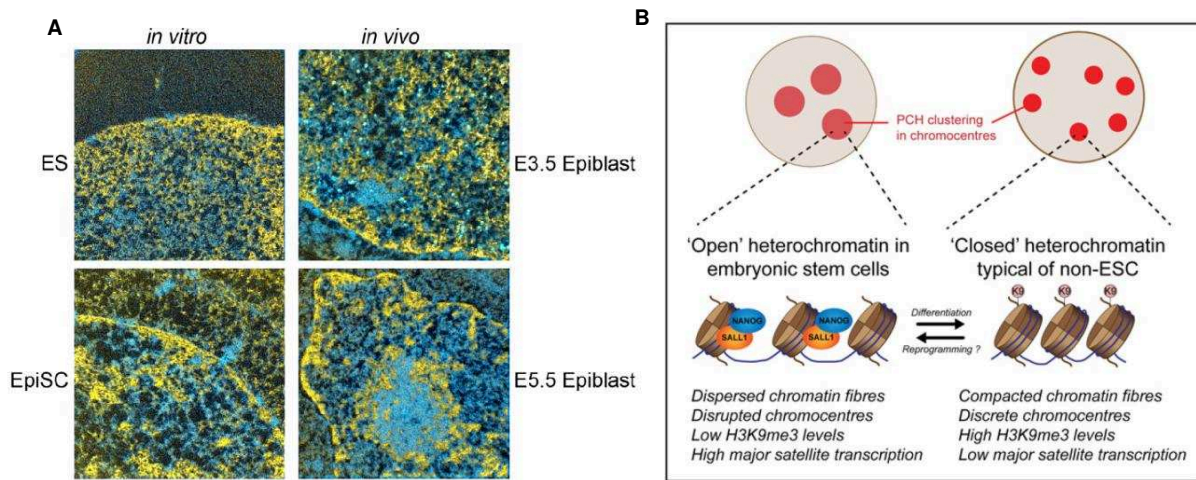


Figure 38: Heterochromatin organization of naive Vs. primed pluripotency (From Ahmed et al., 2010 and Novo et al., 2016).

(A) Electron spectroscopic imaging (ESI) images of *in vitro* pluripotent serum-ESCs and EpiSCs (left column) and *in vivo* E3.5 and E5.5 epiblast (right column). Chromatin (yellow) and protein and RNPs (shades of blue) are determined from nitrogen and phosphorus maps. Chromatin clusterizes more in EpiSC and E5.5 epiblast than serum-ESCs and E3.5 epiblast. (B) Schematic representation of heterochromatin organization in serum-ESCs (Open chromatin, low H3K9me3 and high transcription of major satellites) versus non-ESCs or EpiSCs (Compacted chromatin, high H3K9me3 and low transcription of major satellites).

Forced expression of NANOG in EpiSCs is sufficient to decompact heterochromatin, reducing H3K9me3 levels at major satellite repeats and increasing the ncRNA transcripts coming from these regions. In addition, within the heterogeneous populations of serum-ESCs, higher transcription of major satellites is found in Nanog-high cells compared to the low-fraction (Novo et al., 2016). The authors have proposed that NANOG is recruited to major satellites thanks to heterochromatin associated protein SALL1, as *Sall1*^{-/-} ESCs recapitulates the same Nanog^{-/-} phenotype (Novo et al., 2016).

As in embryos, the histone chaperone CAF-1 plays a critical role in three-dimensional organization of pericentric heterochromatin in serum-ESCs: when depleted for p150, the DAPI-dense foci are lost and only some bright staining at the periphery or around nucleoli are still observed (Figure 39A) (Houlard et al., 2006). In addition these ESCs lose H3K9me3, HP1 and H4K20me3 foci-enrichment, showing a diffused staining (Houlard et al., 2006). Absence of CAF-1 in serum-ESCs induces a complete decondensation of major satellites sequences and the consequent destructuration of chromocenters (Figure 39C), without however affecting the hypermethylated state of major satellites (Houlard et al., 2006). Interestingly when the same experiment was performed on differentiated cells (MEFs) no alteration of the heterochromatin organization was found (Figure 39B), suggesting that only

pluripotent stem cells need CAF-1 during replication of pericentromeric regions in order to maintain their compacted structure (Houlard et al., 2006).

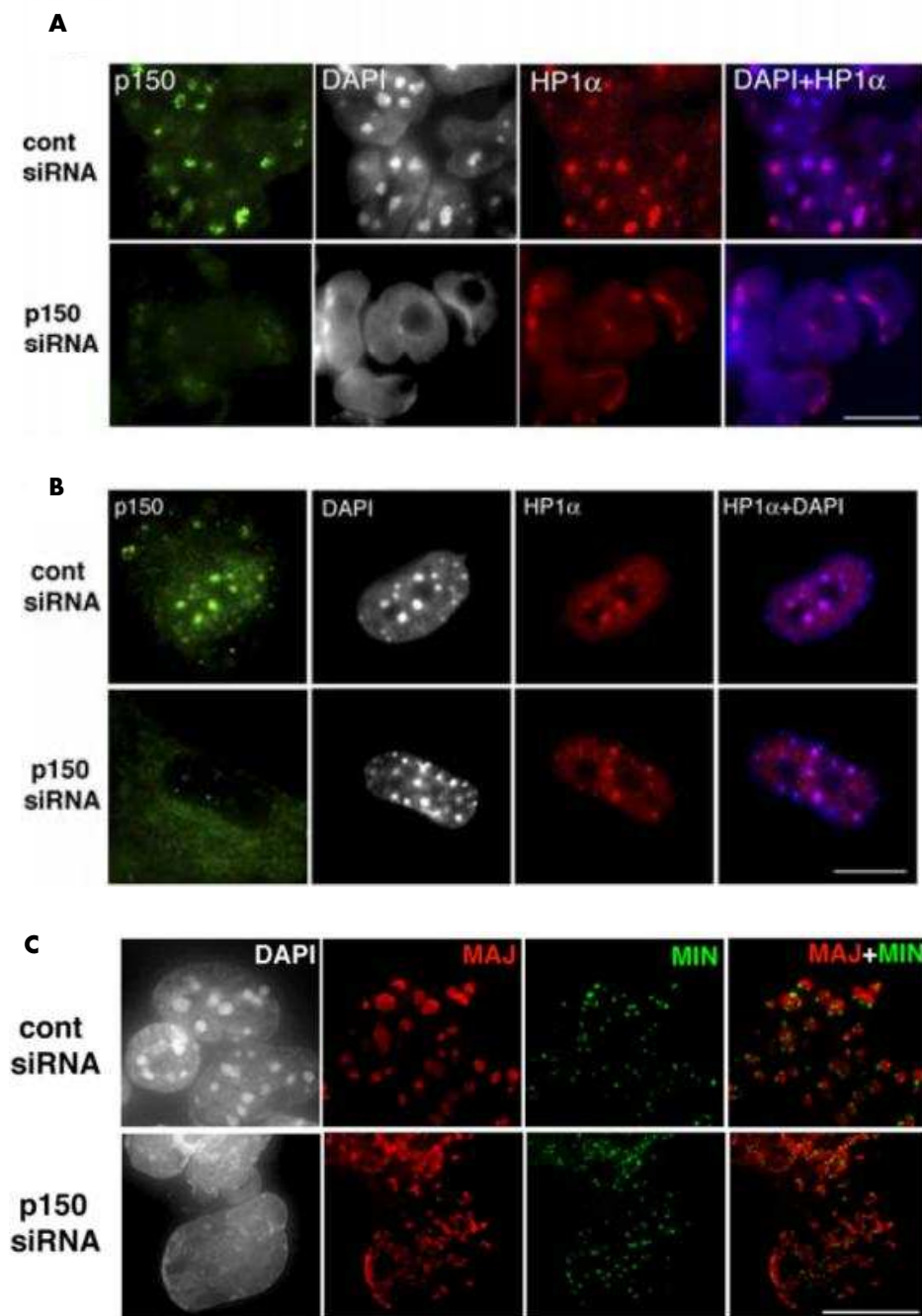


Figure 39: Effect of CAF-1/p150 down-regulation in serum-ESCs and MEFs (From Houlard et al., 2006)
 (A, B) Immunostaining for CAF-1/p150 (green) and HP1α (red) in serum-ESCs (A) or MEFs (B) transfected with control (cont) or CAF-1 (p150) siRNA. While down-regulation of CAF-1 in serum-ESCs disrupt the heterochromatin organization (loss of DAPI-dense and HP1α foci-enrichment), such phenomenon is not observed in MEFs. (C) DNA-FISH for major (red) and minor (green) satellites in serum-ESCs transfected with control (cont) and CAF-1 (p150) siRNA. Down-regulation of CAF-1 disorganizes chromocenters structures, pericentromeric sequences decompact, going particularly at the nuclear periphery.

3.2.3 Plasticity of heterochromatin in mouse naive pluripotency

Pericentromeric heterochromatin is normally characterized by high levels of H3K9me3 and DNA methylation, inducing a strong repressive environment. However many studies reveal that the composition of the constitutive heterochromatin can be highly plastic and dynamic in particular in the context of naive pluripotency (Déjardin, 2015).

Serum-ESCs depleted from SUV39H enzymes show impressive loss of H3K9me3 in particular at major and minor satellites (Figure 40A) (Cooper et al., 2014; Lehnertz et al., 2003; Martens et al., 2005; Peters et al., 2003; Saksouk et al., 2014). At the same time these ESCs acquired H3K27me3 enrichment in foci and specific deposition of this mark at major satellite but not minor satellites (Figure 40B) (Cooper et al., 2014; Lehnertz et al., 2003; Martens et al., 2005; Peters et al., 2003; Saksouk et al., 2014). Peters and colleagues (2003) have thus suggested a cross-talk between constitutive and facultative heterochromatin to rescue the repressed state when the structure of the chromatin was compromised (Figure 40C) (Peters et al., 2003). Moreover the absence of SUV39H enzymes induces a dramatic loss of DNA methylation specifically at major satellites (not on minor satellites) and a slight up-regulation of non-coding RNAs from these repeats (Lehnertz et al., 2003). DNMT1 and DNMT3A interact directly with SUV39H1 and HP1 β (Fuks et al., 2003) while the absence of SUV39H1/2 abrogates the foci staining for DNMT3B in serum-ESCs (Lehnertz et al., 2003). Moreover while DNMT3B is solely responsible for minor satellite DNA methylation, both DNMT3A and DNMT3B cooperate to deposit methyl-groups on major satellites in serum-ESCs (Okano et al., 1999).

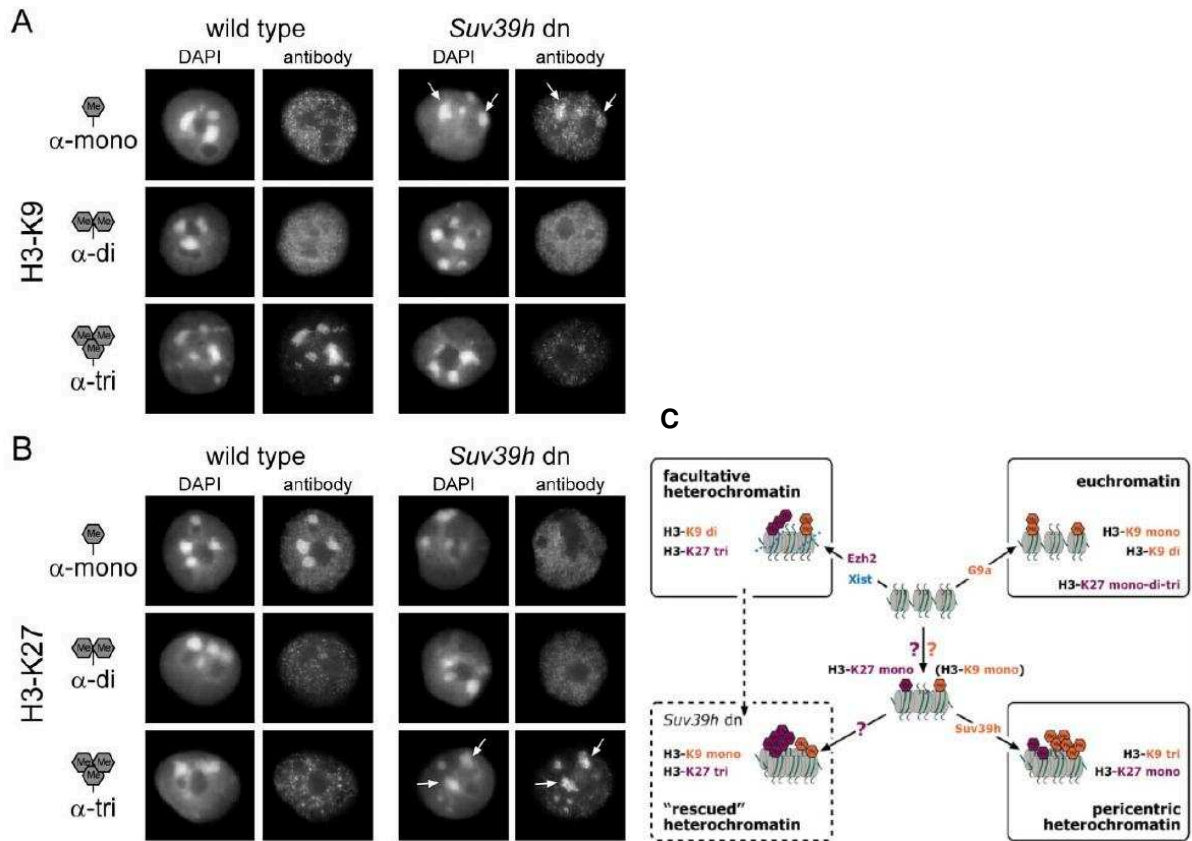


Figure 40: Suv39hdn condition on serum-ESCs heterochromatin (From Peters et al., 2003)

(A, B) Immunostaining for H3K9me1, me2, me3 (A) and H3K27me1, me2 and me3 (B) in serum-ESCs wild-type and *Suv39h*dn. The *Suv39h*dn condition leads to loss of H3K9me3 foci and enrichment in H3K9me1 foci while increasing H3K27me3 foci at the expense of H3K27me1. (C) Schematic representation of the heterochromatin that rescues the absence of H3K9me3 by H3K27me3.

While *Suv39h*dn serum-ESCs show a great restructuration of the epigenetic state of heterochromatin, the complete absence of 5-meC in *DnmtTKO* serum-ESCs induces a reduction of H3K9me3 and an enrichment of H3K27me3 at major satellites with a presence of both marks in heterochromatin foci at the subnuclear level (Cooper et al., 2014; Saksouk et al., 2014; Tsumura et al., 2006). Interestingly in *DnmtTKO* serum-ESCs, H3K27me3 and H3K9me3 are enriched in distinct subdomains of the chromocenters as showed by super resolution microscopy (Figure 41A) (Cooper et al., 2014). Moreover *DnmtTKO* serum-ESCs gain also the mark H2AK119ub at DAPI-dense foci likely via the recruitment of variant PRC1 by KDM6B which recognizes unmethylated CpG islands, thus driving in turns the recruitment of PRC2 for H3K27me3 deposition (Figure 41B) (Blackledge et al., 2014; Cooper et al., 2014).

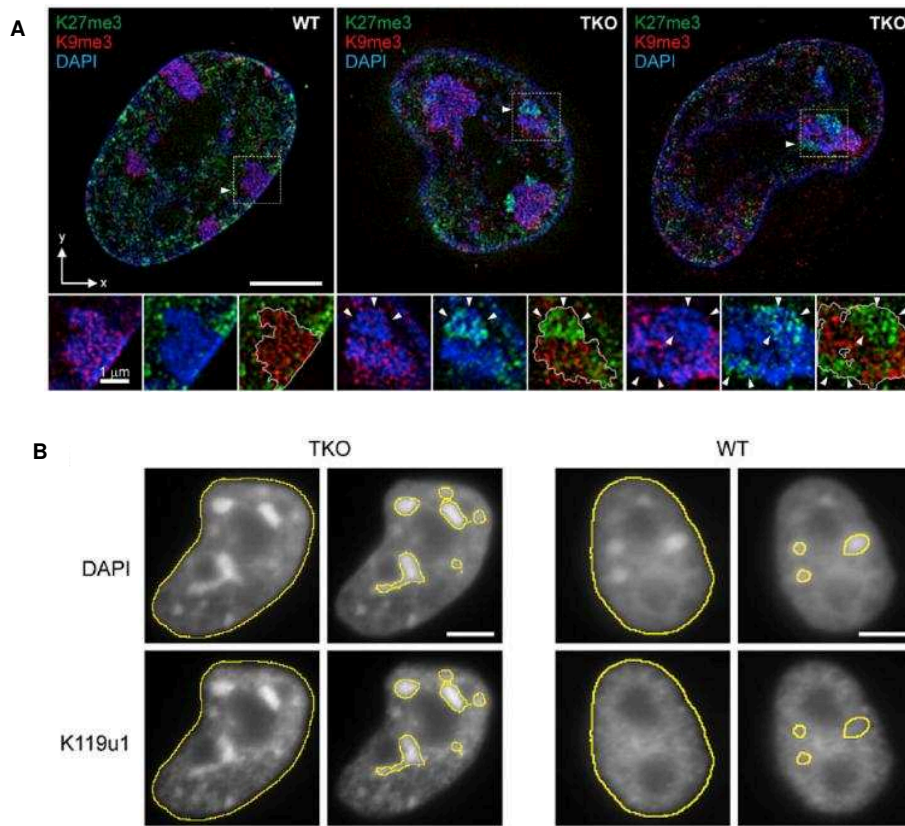


Figure 41: DnmtTKO condition on serum-ESCs heterochromatin (From Cooper et al., 2014).

(A) Three dimensional super-resolution structured illuminated microscopy (3D-SIM) images of immunostaining for H3K27me3 (green), H3K9me3 (red), and DAPI (blue) on serum-ESCs wild-type and DnmtTKO. Small panels show a single PCH region (delineated with a white line), and merges show overlap of each color. Arrowheads mark H3K27me3 staining within PCH in DnmtTKO serum-ESCs. (B) Immunostaining for H2AK119ub in serum-ESCs DnmtTKO and wild-type with DNA counterstaining with DAPI. Only in the mutant condition DAPI-dense foci are enriched in H2AK119ub.

Saksouk and colleagues (2014) have proposed another common molecular mechanism of PRC2 recruitment for H3K27me3 deposition at pericentromeric satellites in Suv39h^{dn} and DnmtTKO serum-ESCs via the relocalization of the methylation-sensitive DNA binding protein BEND3 to hypomethylated major satellites (Saksouk et al., 2014). BEND3 strongly binds unmethylated major satellite repeats, as shown also by treatment of wild-type serum-ESCs with 5-azacytine (a DNMTs inhibitor), and recruits the Nucleosomal Remodeling Deacetylase (NuRD) complex which in turns can recruit PRC2 (Saksouk et al., 2014).

To establish the general relationship between DNA methylation and H3K27me3, Hagarman and colleagues (2013) have studied the loss of either PRC2 complex (Eed subunit) or DNMTs in serum-ESCs. They have found that the absence of 5-meC induces a genome wide increment in H3K27me3 meaning that DNA methylation is directly and globally antagonizing the placement of H3K27me3 (Hagarman et al., 2013). Interestingly the genes that lost 5-meC

and gain H3K27me3 do not change their expression. Conversely in the reciprocal experiment the loss of H3K27me3 has only a modest effect on DNA methylation, with only 4% of genes changing their methylation status and with some becoming hyper- but other also hypomethylated (Hagarman et al., 2013).

Altogether pluripotent state of serum-ESCs is characterized by a dynamic and plastic constitutive heterochromatin compartment which is highly transcribed, organized into a more open state and governed by interconnected pathways that can rescue each other between H3K9me3, DNA methylation and H3K27me3 principally. However no such studies have been made till now on 2i-ESCs and EpiSCs, in order to have a complete and comparative view of mouse in vitro pluripotency.

OBJECTIVES

OBJECTIVES OF THIS THESIS

For more than 35 years mouse ESCs in serum-condition, so in a naive but metastable pluripotent state, have been extensively studied in terms of transcriptional profile, epigenetics, nuclear organization and proteomics. However less than 10 years ago a primed pluripotent stem cell type called EpiSC has been derived in mouse in order to have a mouse parallel of the conventional human ESCs. The characterization of these EpiSCs is still ongoing and far to be sufficient for a real comparison with either mESCs or hESCs. More recently the discovery of a chemically defined medium (2i) that inhibits differentiation of mESCs more efficiently than standard serum condition, introduced a new type of naive pluripotent stem cell here called ground-naive. Our knowledge on 2i-ESCs especially in comparison to either serum-ESCs or EpiSC is getting more and more in depth during these last years, particularly concerning the transcriptomics and epigenomics.

However exhaustive studies comparing directly and simultaneously 2i-ESCs, serum-ESCs and EpiSCs are still lacking. Such studies could be fundamental to really establish markers and peculiar features in the definition of each state. In addition extrapolation from comprehensive mouse studies could be extremely important also for the classification of human pluripotent states as many groups attempt to obtain a more naive human pluripotency close to mouse ESCs in 2i condition playing with inhibition of different signaling pathways.

This thesis presents a comprehensive and comparative study of the two kinds of naïve states and the primed state of mouse pluripotency concerning epigenetics features and transcriptional dynamics. The objectives of the studies conducted here are to fill some of the many gaps of knowledge between the different states of mouse pluripotency and to give some insight on what scientists should also focus in the case of human pluripotent cells.

The majority of the epigenomic studies conducted until now on the different states of pluripotency were concentrated on single-copy sequences of the genome as normally in ChIP-seq analysis only these sequences are taken in account. However in mouse the repetitive regions are the most represented and in particular pericentromeric and centromeric regions composed respectively by major and minor satellites correspond to 3.5% of the genome and can be analyzed separately starting from ChIP-seq datasets. Noteworthy these regions seem to be important for the stability of the genome and are normally reorganized during early development to acquire a compacted and silent structure in somatic tissues. Serum-ESCs are known to be dynamic even for major and minor satellites with a more decompacted condition

compared to differentiated cells and characterized by a pervasive transcription of these non-coding sequences as well. Whether this open chromatin and hyper-transcriptional state of the constitutive heterochromatin is maintained also in 2i-ESCs as well as in EpiSCs is going to be elucidated in this thesis. In addition the concept of an already established epigenetic barrier in EpiSCs compared to the open and dynamic state of ESCs prompted the idea to deeply study the conversion kinetics from naïve to primed state. Such studies aim to gain more insight on the molecular interconnection between transcriptional regulation and epigenetic state and vice versa allowing maybe the identification of key actors of this process. This thesis mainly focuses on the heterochromatin compartment but will also give some light on the euchromatin especially analyzing the acetylation of histones which is probably the major “physical force” of the chromatin decompaction. In between euchromatin and heterochromatin, chromatin bivalency is probably one of the most peculiar and interesting features of ESCs, especially when cultivated in serum condition. Whether this chromatin condition is maintained also in primed pluripotency will be also investigated in this thesis.

MATERIALS AND METHODS

MATERIALS AND METHODS

1 Cell culture

ESC lines (R1, WT01, Suv39hdn and DnmtTKO) were cultured in 2i or serum medium as described in Tosolini and Jouneau, 2016a. Briefly, 2i-ESCs were cultured on Laminin-coated dishes (Sigma) in Chemically Defined Medium (CDM) (Brons et al., 2007) supplemented with LIF (700 U/ml) (CellGS), PD0332552 (1 μ M) (AxonMedchem) and CHIR99201 (final 3 μ M) (AxonMedchem), while serum-ESCs were cultured on gelatine-coated dishes in DMEM supplemented with 15% serum (Thermo Fisher Scientific) and LIF (1000 U/ml) (CellGS).

In vitro conversion of ESC into cEpiSC was performed by switching ESCs from serum/LIF or 2i/Lif mediums to CDM with FGF2 (12 ng/ml) (CellGS) and ActivinA (20 ng/ml) (CellGS) as described in Tosolini and Jouneau, 2016b. These converted cells were used 3-5 passages after the conversion. EpiSC lines (FT129.1 and 9.73) were cultured as described in Brons et al., 2007.

EZH2 inhibition was performed by culturing ESC either in 2i or serum-containing medium supplemented with 1 μ M EPZ-6438 (AxonMedchem) for 72h (changing medium daily) or with Dimethyl Sulphoxide (DMSO) (Sigma) as control.

2 Western-blot

Cells were lysed for 30min on ice with RIPA buffer (150mM NaCl, 1% NP-40, 0.5% NaDeoxycholate, 0.1% SDS, 50mM Tris-HCl pH8.0) in presence of protease and phosphatase inhibitors (Pierce). Proteins were quantified using BCA assay (Pierce). 3 μ g of proteins were charged on pre-cast polyacrylamide gel 4-15% (Biorad) and the running has been performed for 1h at 100V. Transfer was then performed on Trans-Blot Turbo (Biorad) for 7min on a PVDF membrane (Hybond-P, GE Healthcare). After blocking in TBS-Tween 20 0.01% (TBS-T) with either 4% non-fatty milk or 5% BSA (Sigma), membranes were incubated overnight (O/N) at 4°C with primary antibodies. After washes in TBS-T, membranes were incubated with secondary antibodies for 1h and washed again before the revelation with ECL2 Western blotting substrate (Pierce). Chemiluminescent signals were captured on a Fuji camera LAS-1000plus and then analysed with ImageJ (imagej.nih.gov/ij). H3 total was used for normalization.

For sequential protein detection, membranes were stripped with 25mM Glycine and 1% SDS at pH 2 for 30min, followed by washes in TBS-T and blocking (milk or BSA) according to the new primary antibody. Antibodies used are listed in (Table 2).

3 Immunostaining

Cells were grown on coated glass-coverslips for 24h, then fixed with PFA 2% (EMS) for 20min, permeabilized with Triton X100 0.5% for 30min and blocked with BSA 2% (Sigma) for 1h. Primary antibody was incubated at 4°C O/N. After washes, the secondary antibody was incubated for 1h. Cells were then washed, post-fixed with PFA 2% (EMS) for 20min, incubated with 1/500 DAPI (Invitrogen) at 37°C for 15min and finally mounted on slide with VectaShield (Vector Laboratories). For 5-meC staining cells were fixed for 5min in methanol 100% and before the BSA blocking, a denaturation step was added with HCl 4N for 1h at 37°C to open the double-helix structure of DNA. For 5h-meC staining standard protocol fixation was used but as for 5-meC the denaturation step was added. Unless specified all the steps were done at room temperature. Antibodies used are listed in (Table 2).

	Antigen	Host species	Dilution for immunostaining	Dilution for western-blot	Reference
Primary	H3K9me3	rabbit	1/300	1/1000	Active Motif 39161
	HP1 β	mouse	1/200		Euromedex 1mod_1A9.A5
	H3K27me2/3	mouse	1/300	1/1000	Active Motif 39538
	H3K27me3	rabbit	1/300		Cell Signaling C36B11#9733
	H3K27me3	rabbit		1/1000	Millipore DAM07-774
	H2AK119ub	rabbit	1/200	1/1000	Cell Signaling 8240S
	H3K4me3	rabbit	1/300	1/1000	Abcam 8580
	H3K9ac	rabbit	1/100	1/1000	Abcam 4441
	EZH2	mouse	1/200	1/1000	NovocastraNCL-L-EZH
	SUV39H1	rabbit	1/100	1/1000	Cell Signaling D11B6 #8729
	RING1B	mouse	1/200	1/1000	Active Motif 39663
	CAF-1 p150	goat	1/50		Santa Cruz D-16 sc-10206
	5-meC	mouse	1/500		Eurogentech BY-MECY
	5h-meC	rabbit	1/500		Active Motif 39769
	NANOG	mouse	1/200		Cell Signaling D2A3#8822
	NANOG	rabbit		1/1000	Abcam ab80892
	DNMT3B	rabbit	1/200	1/500	Active Motif 39207
	DNMT3A	mouse		1/500	Active Motif 39206
	H3total	rabbit		1/20,000	Abcam 1791
Secondary	Anti-Rabbit-Cy3		1/200		Jackson ImmunoResearch
	Anti-Mouse-FITC		1/200		Jackson ImmunoResearch
	Anti-Goat-Cy3		1/200		Jackson ImmunoResearch
	Anti-mouse-HPO			1/5000	Jackson ImmunoResearch
	Anti-rabbit-HPO			1/5000	Jackson ImmunoResearch

Table 2: List of Antibodies used in immunostaining and western-blot with respective dilutions.

4 DNA-FISH

For FISH, cells grown on coated glass-coverslips for 24h were fixed with PFA 4% (EMS) for 15min. They were permeabilized with Triton X100 0.5% for 30min and treated with RNase A 200 μ g/mL (Sigma) for 30min at 37°C. After an equilibration step in the hybridization buffer (50% formamide, SSC 2X, Denhardt 1X, 40 mM NaH₂PO₄, 10% dextran sulfate) for 45min at 37°C, cells were denaturated in presence of probes at 75°C for 3 min and then incubated O/N at 37°C. The day after they were washed three times with SSC 2X pH5.8-50% formamide at 39°C and then three times with SSC 2X pH6.3 at 39°C. Cells were finally incubated with 1/500 DAPI (Invitrogen) at 37°C for 15min and then mounted with

VectaShield (Vector Laboratories). Unless specified all the steps were done at room temperature.

For the detection of major and minor satellites sequences, we used probes described in Aguirre-Lavin et al., (2012) prepared by PCR on mouse genomic DNA using the following primer pairs: 5'-CACTTTAGGACGTGAAATATGGCG-3' and 5'-CATATTCCAGGTCCTTCAGTGTGC-3' for major satellites and 5'-AAAACACATTCGTTGGAAACGCG-3' and 5'-ACTCATCTAATGTTCTACAGTG-3' for minor satellites. PCR products were labelled with Cy3 and Cy5, respectively, using a random labelling kit (Invitrogen).

5 In situ Proximity Ligation Assay (PLA)

To assess the co-presence of two different histone marks on the same nucleosome we used the Duolink in situ PLA (Olink, Bioscience) largely following the manufacturer protocol. Briefly, cells grown on coated glass-coverslips for 24h were fixed with PFA 4% (EMS) for 10min. They were permeabilized with Triton X100 0.5% for 30min and blocked with the Olink blocking solution for 30min at 37°C. Primary antibodies, H3K27me2me3 (Active Motif 39538, mouse antibody, 1/300) and H3K4me3 (Abcam 8580, rabbit antibody, 1/300) were incubated at 4°C O/N. After washes with Olink buffer A, the secondary antibodies Olink Plus (anti-rabbit) and Minus (anti-mouse) PLA probes was incubated for 1h at 37°C. Cells were then washed again with Olink buffer A and incubated for ligation with Olink Ligase and buffer for 30min at 37°C. Successively the amplification of the in situ PLA signal is performed incubating cells with Olink polymerase and buffer for 100min at 37°C. After washed with Olink buffer B, cells were finally incubated with 1/500 DAPI (Invitrogen) at 37°C for 15min and then mounted with VectaShield (Vector Laboratories). Unless specified all the steps were done at room temperature.

6 Three-dimensional structured image acquisition and analysis

Imaging was performed at the MIMA2 platform (<http://www6.jouy.inra.fr/mima2>) with an inverted ZEISS AxioObserver Z1 microscope equipped with an ApoTome slider, a Colibri light source, AxioCam MRm camera and driven by the Axiovision software 4.8.2. Observations were carried out using a 63X oil-immersion objective. Cells were scanned entirely using a z-distance of 0.24 µm between optical sections. Fluorescent wavelengths of

405, 488, 555, and 639 nm were used to excite DAPI, FITC, Cy3, and Cy5, respectively. Images were then analysed on ImageJ (imagej.nih.gov/ij) to perform linescan, merge of channels and z-projections.

Three-dimensional (3D) reconstructions of DNA-FISH signals were done with AMIRA software 3.1 after 3D nuclei segmentation using an unpublished Python script of the laboratory.

For analysis of in situ PLA images nuclei were automatically segmented in 3D using an unpublished Python script of the laboratory and subsequently the number of spots inside each nuclei were counted, adapting a previously published Python script (Pauloin et al., 2016).

7 Southern-blot

Southern blot on genomic DNA was performed as described in Thijssen et al. 2015. For major satellite analysis 200ng of genomic DNA were digested with HpyCH4IV (New England Biolabs) for 1h at 37°C. For minor satellites 500ng of gDNA were digested with HpaII (New England Biolabs) and 300ng with MspI (New England Biolabs), both O/N at 37°C. Digested samples were separated for 5h run on 1% agarose gel. Gels were then denaturated in a 1.5M NaCl and 0.5M NaOH solution for 20min and neutralized with 0.5M Tris-HCl pH 7.5 and 1.5M NaCl for 40min. Transfer was performed O/N on Hybond-N+ membranes (GE Healthcare) in SSC 20X. After ultraviolet cross-linking, membranes were pre-hybridized in SSC 6X, Denhardt 5X and 0.1% SDS for 1h at 42°C and hybridized with 32P-labelled probes for 2h at 42°C. Unless specified all the steps were done at room temperature. After membrane washing, signals were detected using FLA 7000 phosphorimager (Fuji). Images were then analyzed with ImageJ (imagej.nih.gov/ij) to perform linescan for major satellites and intensity ratio HpaII/MspI for the lower six bands of each lane for minor satellites.

Probe used: Major satellites 5' –CAC GTC CTA CAG TGG ACA TTT CTA AAT TTT CCA CCT TTT TCA GTT- 3' and minor satellites 5' –ACA TTC GTT GGA AAC GGG ATT TGT AGA ACA GTG TAT ATC AAT GAG TTA CAA TGA GAA ACA T- 3'.

8 qRT-PCR

Total RNA was extracted from cells using TRIzol (Ambion). 3µg of RNA were subjected to DNase treatment using Turbo DNA-free kit (Ambion). Retro-transcription of 500ng of DNase treated-RNA was performed using Random primers (Invitrogen) and Superscript III

(Invitrogen). In satellite transcripts quantification, for each sample, a negative control was included (no Superscript enzyme). Quantitative PCR was carried out in triplicates using SybrGreen mix (Applied Biosystem) on a StepOne Plus thermal cycler (Applied Biosystem). Data were normalized using the geometric mean of housekeeping genes *Sdha* and *Pbgd* using Qbase software (Biogazelle). Results were expressed as Calibrated Normalized Relative Quantities (CNRQ) and presented in accordance to Weissgerber et al., 2015 with the exception of the kinetics of conversion. The Primers are described in (Table 3).

Primers for qRT-PCR				
Gene	Primer Forward Sequence (5'-3')	Primer Revers Sequence (5'-3')	Reference or Primer Bank ID	Annealing
Major satellite	GACGACTTGAAAAATGACGAAATC	CATATTCCAGGTCCTTCAGTGTGC	Lehnertz et al., 2003	60°C
Minor satellite	GAACATATTAGATGAGTGAGTTAC	GTTCTACAAATCCCGTTTCCAAC	Ferri et al., 2009	60°C
<i>Sdha</i>	GGAACACTCCAAAAACAGACCT	CCACCACTGGGTATTGAGTAGAA		60°C
<i>Pbgd</i>	CCTGGCATACAGTTTGAAATCAT	TTTTTCCAGGGCGTTTTCT	Bernardo et al., 2011	60°C
<i>Ezh2</i>	AGTGACTTGGATTTTCCAGCAC	AATTCTGTTGTAAGGGCGACC		60°C
<i>Eed</i>	AAGAACCTGGAGGGAGGCG	TCCTGGTGCATTTGGCGTAT		60°C
<i>Kdm6b</i>	GCATCTATTTGGAGAGCAAACGAG	GGTACGGACCTCCACCGTA		60°C
<i>Suv39h1</i>	GCAGTGTGTGCTGTAAATCTTCT	ATACCCACGCCACTTAACCAG		60°C
<i>Suv39h2</i>	CTGCCCAGGATAGCATTGTTC	CAAGTCTCGGCTCCACATTTAC		60°C
<i>Dnmt3a1</i>	GAGGGAAGTGAAGACCCAC	CTGGAAGGTGAGTCTTGGCA	6681209a1	60°C
<i>Dnmt3a</i>	TGGAGCTGCAAGAGTGTCTG	GACGTCTGTGTAGTGGACGG		60°C
<i>Dnmt3b</i>	TCAGATGAGCAAGGTCAAGG	TGTACCAAAGCAAGGGGAAG		60°C
<i>Tet1</i>	GAGCCTGTTCTCGATGTGG	CAAACCCACCTGAGGCTGTT	Koh et al., 2011	60°C
<i>Tet2</i>	AACCTGGCTACTGTCATTGCTCCA	ATGTTCTGCTGGTCTCTGTGGGAA	Koh et al., 2011	60°C
<i>Rnf2 (Ring1b)</i>	GAGTTACAACGAACACCTCAGG	CAATCCGCGCAAAACCGATG		60°C
<i>Cbx7</i>	TGCGGAAGGGCAAAGTTGAAT	ACAAGGCGAGGGTCCAAGA		60°C
<i>Pou5f1 (Oct4)</i>	CAGCCAGACCACCATCTGTC	GTCTCCGATTTGCATATC	7305399a3	58°C
<i>Nanog</i>	CTTTCACCTATTAAGGTGCTTGC	TGGCATCGGTTTCATCATGGTAC	Hayashi and Surani, 2009	58°C
<i>Sox2</i>	GCGGAGTGGAACTTTTGTCC	CGGGAAGCGTGTACTTATCCTT		60°C
<i>Prdm14</i>	GCATCCTGGTTCCACAGAG	CTGCAGAACACGCCAAAGTG	Gillich et al., 2012	60°C
<i>Fgf5</i>	TGTGTCTCAGGGGATTGTAGG	AGCTGTTTTCTTGGAATCTCTCC	6753854a1	60 °C
<i>Brachyury (T)</i>	CCCAGAGACCCAGTTCATAG	ATTACATCTTTGTGGTCGTTTC		60 °C
<i>Otx2</i>	TATCTAAAGCAACCGCCTTACG	GCCCTAGTAAATGTCGTCCTCTC	158518427c1	60°C
<i>Esrrb</i>	ATGCGAGTACATGCTTAACGC	CATCCCCACTTTGAGGCATTT		60°C
<i>Klf4</i>	GCAGTCACAAGTCCCCTCTC	GACCTTCTTCCCCTCTTTGG	Jouneau et al., 2012	58°C

Table 3: List of qRT-PCR primers used with sequences, references and annealing.

9 Bioinformatics analysis of ChIP-Seq datasets for satellite repeats

To compare the enrichment in H3K27me3 mark over major and minor satellite repeats we used the following ChIP-seq datasets: GSM590115 (E14-serum) and GSM590116 (E14-2i) from Marks et al., 2012; GSM1725687 (EpiSC1), GSM1725686 (Input EpiSC1), GSM1725726 (Input 2i-ESC1), GSM1725727 (2i-ESC1), GSM1725730 (Input 2i-ESC2), GSM1725731 (2i-ESC2), GSM1725689 (Input EpiSC2), GSM1725690 (EpiSC2), from Zylicz et al., 2015.

To compare the enrichment in H3K9me3 mark over major and minor satellite repeats we used the following ChIP-seq datasets: GSM850406 (E14-serum) and GSM850407 (E14-2i) from Marks et al., 2012.

An “in silico” library was made up exclusively of major and minor fasta consensus sequences taken from Lehnertz et al., 2003. Each sequence was duplicated and juxtaposed in order to detect reads which may map at the junction between two consecutive repeats. For the repeat analysis of ChIP-seq profiles, mappings were performed with the bowtie2 aligner version 2.1.0 with default options (Langmead and Salzberg, 2012). Each read that mapped on major or minor satellite repeats was counted from the resulting output BAM file. Results were expressed as percentage of total number of reads that mapped on the whole mouse genome GRCm38.84.

RESULTS

RESULTS

1 Global epigenetic organization in the different types of mouse pluripotent stem cells

In this first part we characterized and compared the nuclear distribution as well as the global amount of different epigenetic marks of euchromatin and heterochromatin in mouse ESCs (both serum and 2i conditions) and EpiSCs. We mainly performed immunostainings and western-blot in order to identify the epigenetic landscape of our cells of interest. Successively we focused on the heterochromatin compartment as the histone modifications that changed their subnuclear distribution were hallmarks of heterochromatin.

1.1 EUCHROMATIN IN MOUSE PLURIPOTENCY

The most studied serum-ESCs are considered to have a plastic and open chromatin state, so we wondered whether there are any differences between the three types of pluripotent stem cells in terms of euchromatin histone marks. In particular we performed immunostaining to study the nuclear organization of two majors active histone marks H3K4me3 and H3K9ac, as well as western-blot on total cell extracts to evaluate their bulk levels.

H3K4me3 immunostaining reveals an identical diffused pattern in all pluripotent stem cells with some brighter spots (Figure 42A). Furthermore western-blot quantification showed no difference in terms of global H3K4me3 levels in 2i-ESCs, serum-ESCs and EpiSCs (Figure 42B).

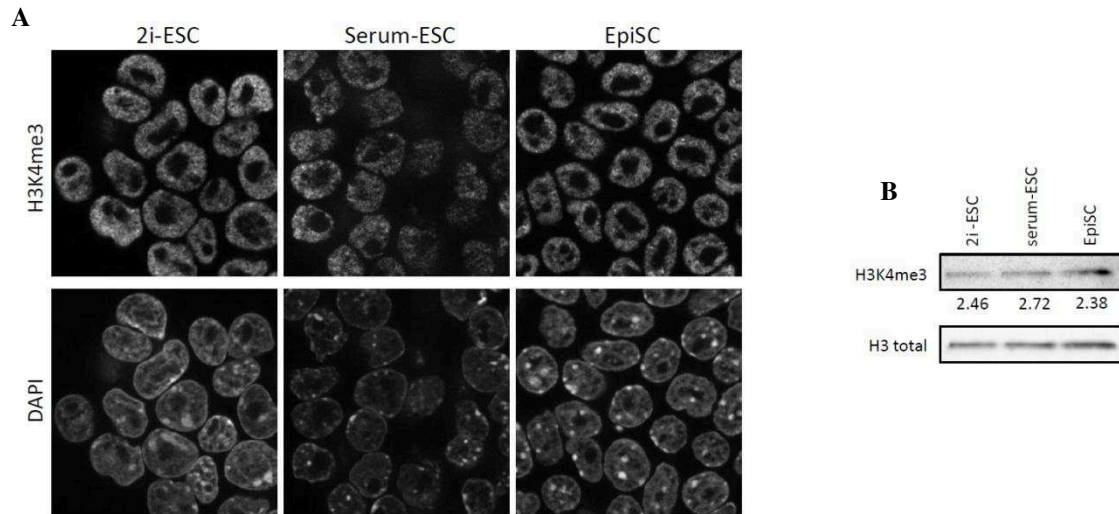


Figure 42: H3K4me3 landscape in the different types of mouse pluripotent stem cells

(A) Immunostaining images (single-plan) for H3K4me3 with DAPI DNA counterstaining. Scale bars represent 5 μ m. (B) Western-blot analysis for quantification of bulk levels of the repressive histone modification H3K4me3 related to the total level of H3. 2i- and serum-ESC: R1. EpiSC: FT129.1.

Immunostaining for H3K9ac showed a similar pattern in the three types of pluripotent stem cells with a diffused signal with bright spots (Figure 43A). Interestingly western-blot analysis revealed an important decrease of H3K9ac in EpiSCs compared to ESCs (either 2i- or serum-) suggesting a less open chromatin state (Figure 43B).

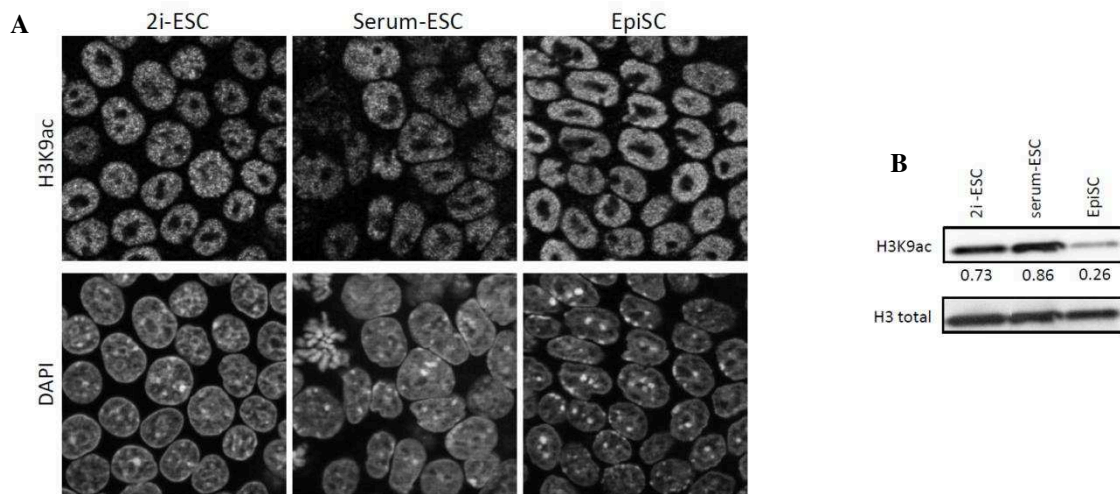


Figure 43: H3K9ac landscape in the different types of mouse pluripotent stem cells

(A) Immunostaining images (single-plan) for H3K9ac with DAPI DNA counterstaining. Scale bars represent 5 μ m. (B) Western-blot analysis for quantification of bulk levels of the repressive histone modification H3K9ac related to the total level of H3. 2i- and serum-ESC: R1. EpiSC: FT129.1.

Altogether euchromatin does not seem to change its global nuclear distribution in the three pluripotent states, at least concerning the two major active histone modifications H3K4me3 and H3K9ac analyzed here. However, while H3K4me3 seems to be equally enriched in the three different types of mouse pluripotent stem cells, H3K9ac is strongly reduced in primed EpiSCs compared to naïve ESCs (either 2i- or serum-). This finding goes along with the idea of an open, hyperacetylated chromatin typical of ESCs and lost in later stages, thus EpiSCs show an epigenetic landscape closer to that of somatic cells than to ESCs.

1.2 BIVALENCY IN MOUSE PLURIPOTENCY

In serum-ESCs the chromatin is enriched in bivalent domains, so characterized by the co-presence of active H3K4me3 and repressive H3K27me3 histone marks at the same locus and even on the same nucleosome (Azuara et al., 2006; Bernstein et al., 2006; Hattori et al., 2013; Marks et al., 2012; Weiner et al., 2016). In order to study the combination of histone modification at the single cell level, Hattori and colleagues applied the *in situ* Proximity Ligation Assay (PLA) technique which produces a fluorescent spot inside the cell when two proteins are in close proximity, defined as less than 30-40nm (Hattori et al., 2013). The number of spots inside a cell does not represent the absolute number of the co-occurrence of the two proteins due to the limited or variable efficiency of the method. Despite this limitation, comparison between different cell types can be done when processed at the same time.

We decided to apply this technique to study bivalency in our three cell types. We performed three independent experiments of *in situ* PLA using antibodies against H3K4me3 and H3K27me2/3 (Figure 44A). Nuclei were then segmented in three-dimensions and the number of spots in each nucleus was automatically counted (See Material and Methods).

In all the three experiments we found that EpiSCs showed approximately a 2-fold reduction of the PLA spots number compared to serum-ESCs (Figure 44B). This suggests that bivalent domains are lost going from naïve to primed pluripotency similarly to what happens in differentiating cells (Hattori et al., 2013). Unexpectedly we observed that 2i-ESCs presented twice more PLA spots than serum-ESCs (Figure 44B). This increment in H3K4me3-H3K27me2/3 co-presence within less than 30-40nm was consistent in all three experiments. Such result contrasts with the largely reduced number of bivalent domains evidenced by

ChIP-seq in 2i-ESCs compared to serum-ESCs (Marks et al., 2012) as well by co-ChIP (Weiner et al., 2016).

We preferred to analyse the three independent experiments separately rather than merging them, as the efficiency of the ligation and amplification steps could vary from one experiment to the other. Indeed while experiment 1 and 3 give very similar spot numbers, experiment 2 shows a 50% increment in spot number per cell in each cell type, suggesting a ligation and/or an amplification step more efficient compared to the other experiments (Figure 44B).

Note that for serum-ESCs our medians (10, 15 and 9) are close to the mean for wild-type ESCs (12) in the Hattori et al study, even though the antibodies we used were different from those in their study. This observation strength our results confirming that the global amount of bivalent domains are on average roughly the same in different types of serum-ESCs and that there is not a strong bias due to the use of different antibodies.

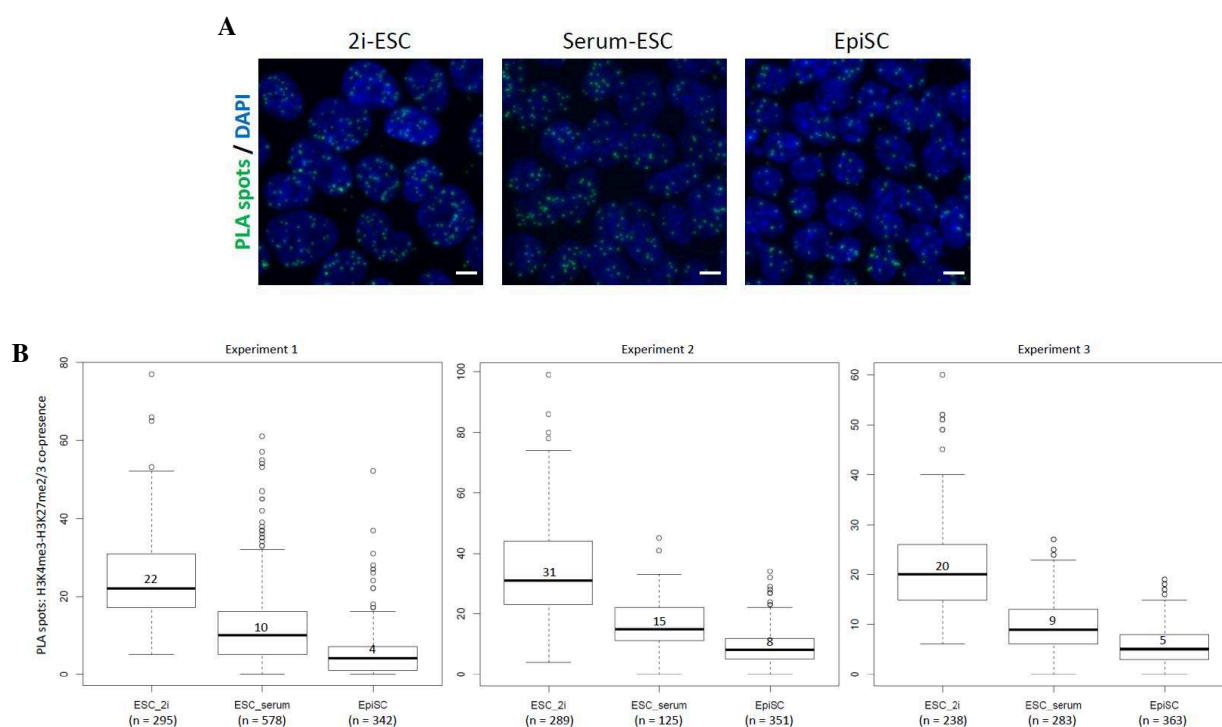


Figure 44: Bivalent domain revealed by in situ PLA in the different types of mouse pluripotent stem cells (A) in situ PLA images (single-plan) for H3K4me3-H3K27me2/3 (green spots) with DAPI DNA counterstaining (blue). Scale bars represent 5µm. (B) Counting of PLA spots per cells in 2i-ESCs, serum-ESCs and EpiSCs in the three independent experiments. Median of each cell population is indicated in the each box, while the n at the bottom represents the cell number per population. 2i- and serum-ESC: R1. EpiSC: FT129.1.

Chromatin bivalency was mainly assessed by H3K4me3 and H3K27me3 ChIP-seq analysis in ESCs. Here we present another protocol to compare the global number of bivalent domain

between the different states of mouse pluripotency using in situ proximity ligation assay. For the first time we showed that H3K4me3-H3K27me2/3 bivalent domains are strongly reduced in EpiSCs compared to ESCs. An important reduction in bivalency was already shown by Hattori et al., 2013 when comparing serum-ESCs to MEFs strengthening once more the hypothesis that EpiSCs have a somatic epigenome. In addition we found that in contrast to what previously shown (Marks et al., 2012; Weiner et al., 2016), 2i-ESCs seems to have an increased co-presence of H3K4me3-H3K27me2/3 compared to serum-ESCs.

1.3 HETEROCHROMATIN IN MOUSE PLURIPOTENCY

The majority of the results of this chapter are included into a publication that has been accepted under major revisions at the journal Scientific Reports.

1.3.1 Heterochromatin domains are characterized by different epigenetic histone marks depending on the pluripotent stem cell type

We studied more in-depth the constitutive heterochromatin regions (PCH/CH), so the chromocenters that are specifically marked by 4',6-diamidino-2-phenylindole (DAPI), showing a foci-enrichment staining. We assessed the percentage of cells presenting specific histone modification foci-enrichment and determined their colocalization with DAPI staining using linescan across nuclei. Our data show that the distribution of the hallmarks of heterochromatin H3K9me3 and H3K27me3 was different according to the cell type. In 2i-ESCs, only one third of the population display H3K9me3-enrichment at PCH/CH foci (Figure 45A). In such cells, H3K9me3-enriched foci were rare and small, and located close to the nuclear periphery (Figure 45B). Conversely all serum-ESCs and EpiSCs showed large and numerous H3K9me3 foci, perfectly co-localizing with DAPI-dense foci (Figure 45A-B). Quantification by western-blot revealed a slight decrease of the H3K9me3 global level in 2i-ESCs compared to serum-ESCs and EpiSCs, which showed no major changes between them (Figure 45C). HP1 β which is one of the typical proteins bound to constitutive heterochromatin also followed the same distribution patten of H3K9m3, as expected by the fact that thanks to its chromodomain it binds specifically to this histone mark (Figure 45D).

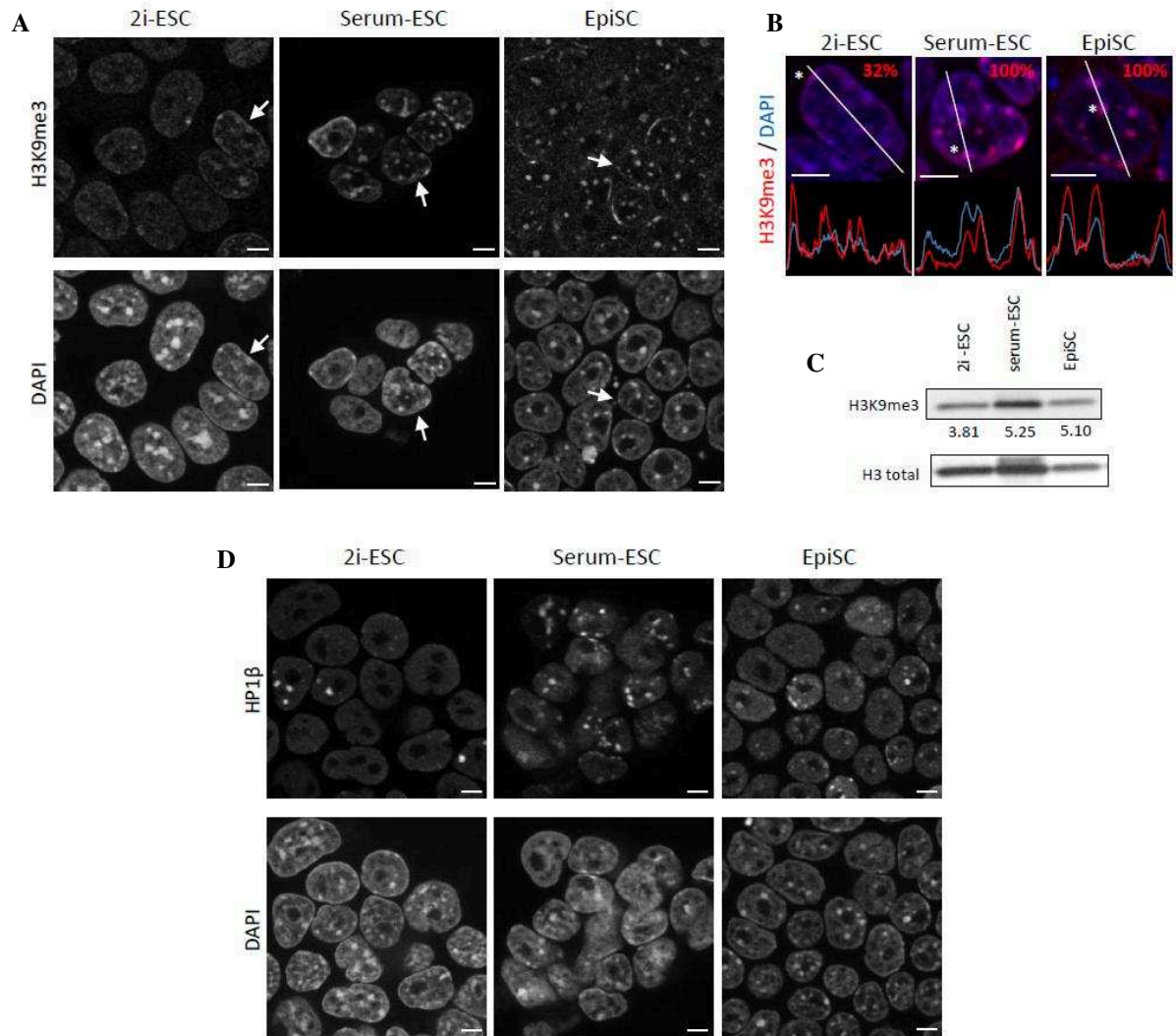


Figure 45: H3K9me3-HP1β landscape in the different types of mouse pluripotent stem cells

(A) Immunostaining images (single-plan) for H3K9me3 with DAPI DNA counterstaining. Scale bars represent 5μm. (B) Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5μm. (C) Western-blot analysis for quantification of bulk levels of the repressive histone modification H3K9me3 related to the total level of H3. (D) Immunostaining images (single-plan) for HP1β with DAPI DNA counterstaining. Scale bars represent 5μm. 2i- and serum-ESC: WT01 (A, B) and R1 (C, D). EpiSC: cWT01 (A, B) and 129.1 (C, D).

We then examined the expression of the enzymes responsible for H3K9me3 deposition and we found that both mRNA and protein levels of Suv39h1 increased considerably between 2i-ESC and EpiSC (Figure 46A). Accordingly, immunostaining revealed that SUV39H1 accumulated in foci in serum-ESCs in 31% of cells and in the vast majority (93%) of EpiSCs, while no enrichment was observed in 2i-ESCs (Figure 46B).

We conclude that H3K9me3, HP1β and SUV39H1 share the same pattern in terms of bulk-levels and subnuclear localization: lower levels and mostly dispersed organization in 2i-ESC, high-levels and strong foci-enrichment in serum-ESCs and EpiSCs.

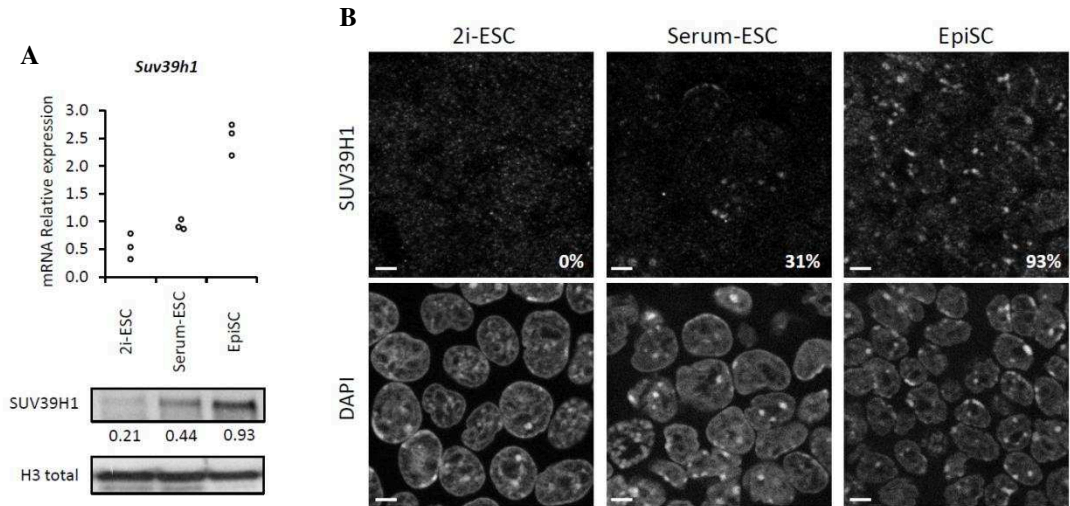


Figure 46: H3K9 trimethylase SUV39H1 in the different types of mouse pluripotent stem cells

(A) Upper part: Relative expression (CNRQ) of *Suv39h1* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent biological replicate. Lower part: Western-blot analysis for quantification of bulk levels of SUV39H1 related to total H3. (B) Immunostaining images (single-plan) for SUV39H1 with DAPI DNA counterstaining. % indicates the percentage of cells in the population displaying a foci-enrichment pattern. Scale bars represent 5 μ m. 2i- and serum-ESC: R1. EpiSC: FT129.1.

Regarding the “so called” facultative heterochromatin histone mark H3K27me3, we observed two other remarkable phenomena. First, immunostaining revealed an unexpected H3K27me3 nuclear distribution: the majority of 2i-ESCs (69%) presented a foci-enrichment pattern of H3K27me3, which co-localized with DAPI-bright regions (Figure 47A-B). In contrast serum-ESCs had, in 95% of the cases, a completely diffuse pattern. Second, all EpiSCs presented a very low H3K27me3 signal that correlated with a reduced bulk level of this histone mark, as revealed by western-blotting (Figure 47C).

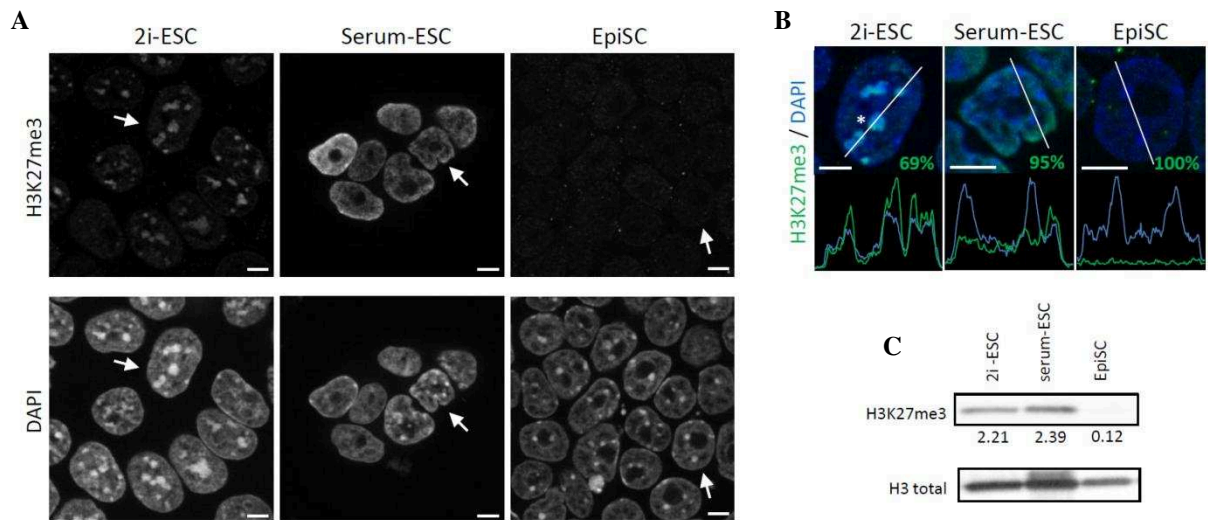


Figure 47: H3K27me3 landscape in the different types of mouse pluripotent stem cells

(A) Immunostaining images (single-plan) for H3K27me3 with DAPI DNA counterstaining. Scale bars represent 5 μ m. (B) Magnification on a single cell (arrow) with merge of signals: H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5 μ m. (C) Western-blot analysis for quantification of bulk levels of the repressive histone modification H3K27me3 related to the total level of H3. 2i- and serum-ESC: WT01 (A, B) and R1 (C). EpiSC: cWT01 (A, B) and 129.1 (C).

Conversely H2AK119ub, which is another mark of facultative heterochromatin, does not share the same pattern of H3K27me3 as it is diffused in the nucleoplasm presenting only some bright spots with no foci enrichment even in 2i-ESCs (Figure 48A), suggesting no PRC1 recruitment by H3K27me3. In addition bulk levels of H2AK119ub were substantially unchanged even between ESCs and EpiSCs (Figure 48B).

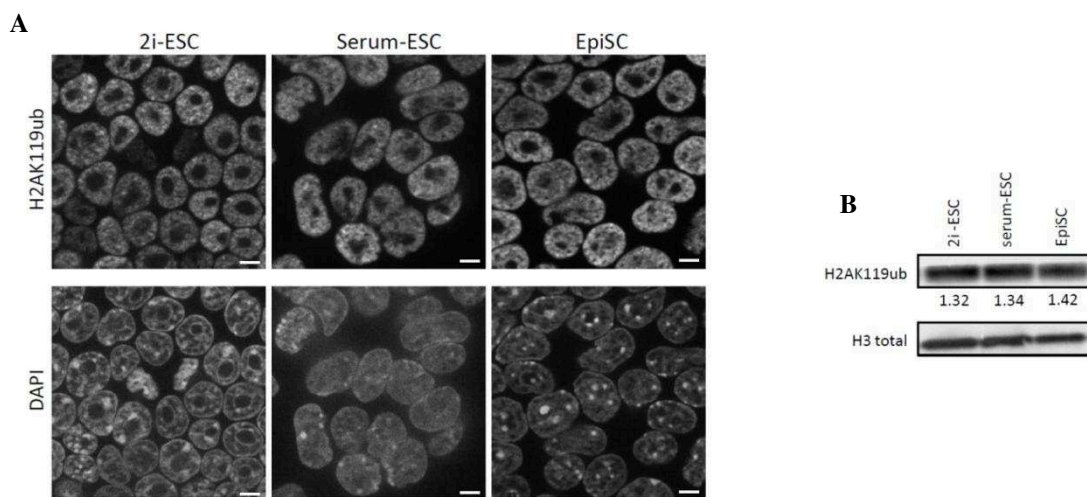


Figure 48: H2AK119ub landscape in the different types of mouse pluripotent stem cells

(A) Immunostaining images (single-plan) for H2AK119ub with DAPI DNA counterstaining. Scale bars represent 5 μ m. (B) Western-blot analysis for quantification of bulk levels of the repressive histone modification H2AK119ub related to the total level of H3. 2i- and serum-ESC: R1. EpiSC: FT129.1.

To complete the study, we investigated the expression of EZH2, the enzyme that within the PRC2 complex deposits H3K27me3. While qRT-PCR did not evidence any major difference in *Ezh2* transcript levels, western-blot revealed a strong reduction of the EZH2 protein amount in EpiSCs compared to both ESCs (Figure 49A). However in the three types of cells, EZH2 displayed the same diffuse signal with some bright spots but no foci-enrichment (Figure 49B). Moreover the expression of *Kdm6b* gene, which encodes for a specific H3K27me3 demethylase (KDM6B also called JMJD3), was increased in EpiSCs compared to ESCs (Figure 49C). Unfortunately, we did not find any antibody that was working on western-blot for the corresponding protein. Altogether, the decreased H3K27me3 level in EpiSCs could be explained by both a reduction of its histone methyltransferase and an increase of its histone demethylase.

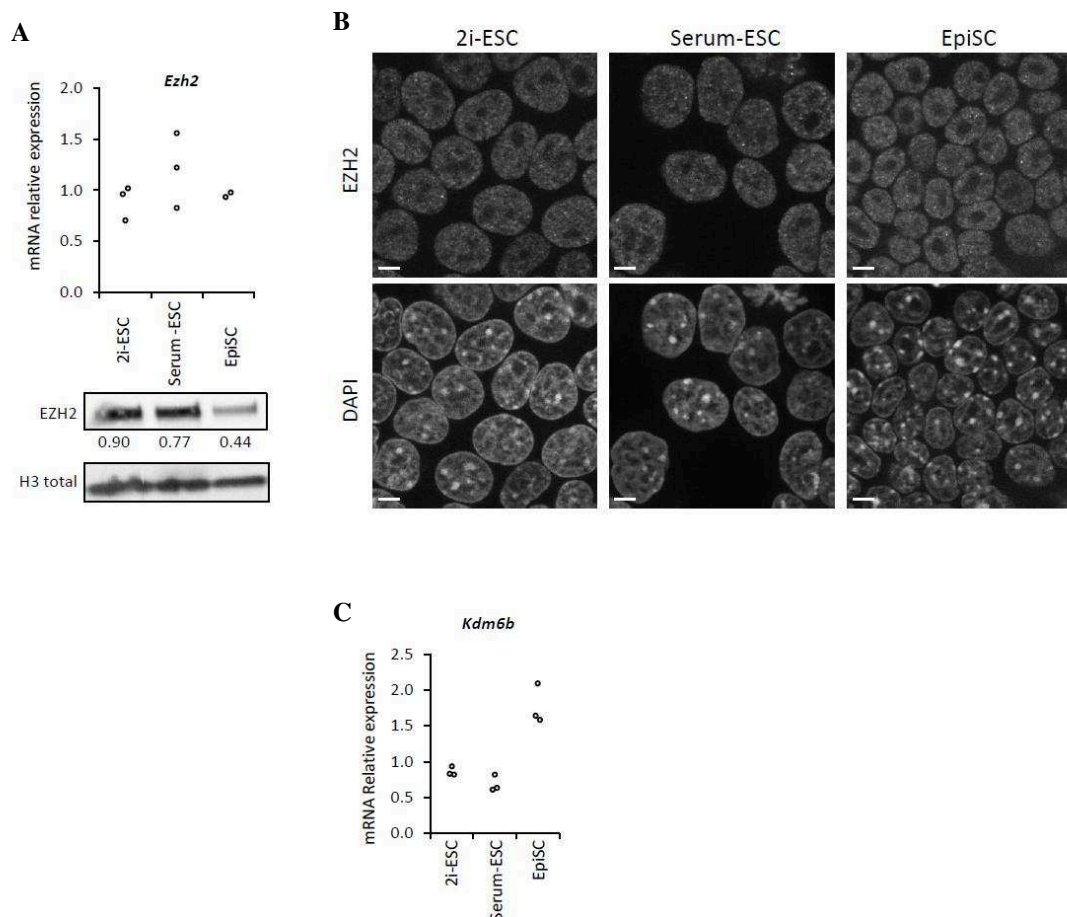


Figure 49: H3K27 methylases and demethylases in 2i-ESCs, serum-ESCs and EpiSCs

(A) Upper part: Relative expression (CNRQ) of *Ezh2* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Lower part: Western-blot analysis for quantification of bulk levels of EZH2 related to H3 total. (B) Right part: immunostaining images (single-plan) for EZH2 with DAPI DNA counterstaining. Scale bars represent 5μm. (C) Relative expression (CNRQ) of *Kdm6b* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. 2i- and serum-ESC: R1. EpiSC: FT129.1.

In addition, as Ezh2, Rnf2 does not show change at the transcriptional level (Figure 50A), but at the protein level RING1B, the enzyme responsible for H2AK119ub deposition seems to be slightly more expressed in serum-ESCs compared to 2i-ESCs, as shown by western-blot, and clearly drops down in EpiSC as EZH2 does (Figure 50B) showing a parallel between PcG proteins.

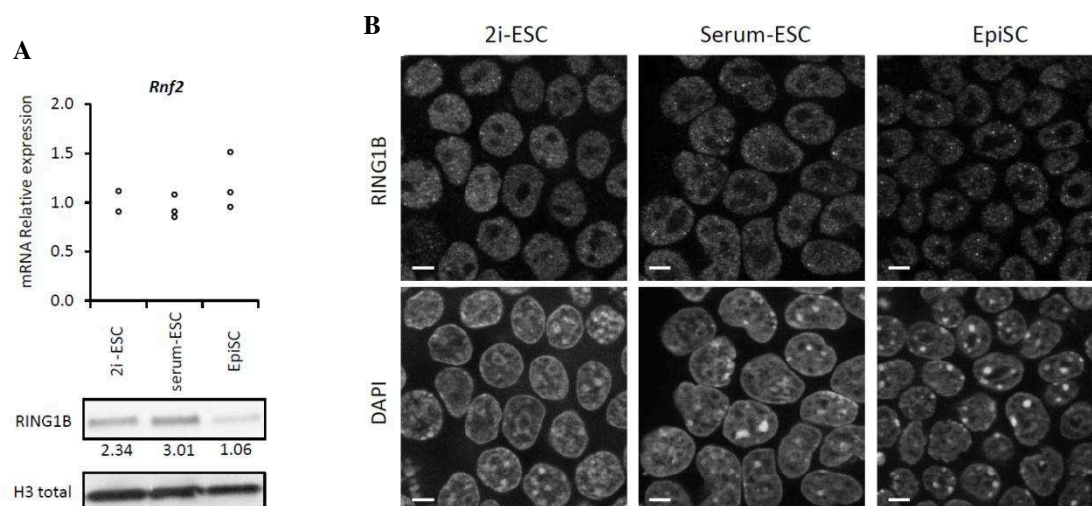


Figure 50: H2A ubiquitin ligase in the different types of mouse pluripotent stem cells.

(A) Upper part: Relative expression (CNRQ) of *Rnf2* (*Ring1b*) transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Lower part: Western-blot analysis for quantification of bulk levels of RING1B related to H3 total. (B) Right part: immunostaining images (single-plan) for RING1B with DAPI DNA counterstaining. Scale bars represent 5 μ m. 2i- and serum-ESC: R1. EpiSC: FT129.1.

1.3.2 Pericentromeric heterochromatin is subjected to an epigenetic switch in 2i-ESCs compared to serum ones and EpiSCs.

Knowing that in mouse somatic cells H3K9me3 is particularly enriched in pericentromeric and centromeric regions (PCH/CH) (Bouzinba-Segard et al., 2006; Martens et al., 2005; Peters et al., 2001), we wondered whether the 2i condition led to a switch in the epigenetic marks at these sequences. To test this hypothesis we analyzed cells in late anaphase and we found that H3K27me3 was clearly enriched specifically at the pericentromeric ends of mitotic chromosomes in 2i-ESCs, while H3K9me3 was present all along these chromosomes (Figure 51). Conversely, in serum-ESCs and EpiSCs the ends of mitotic chromosomes were strongly enriched in H3K9me3, while H3K27me3 presented a continuous staining in serum-ESCs or was quite undetectable in EpiSCs (Figure 51).

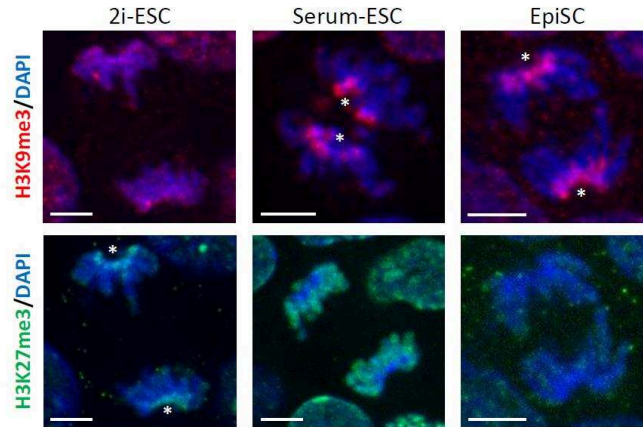


Figure 51: Enrichment of H3K9me3 or H3K27me3 at PCH/CH depending on the type of mouse pluripotent stem cell

Immunostaining images of anaphase chromosome plates for H3K9me3 (red) or H3K27me3 (green) with DAPI DNA counterstaining (blue). Stars indicate enrichment of the histone mark at the PCH/CH region. Scale bars represent 5µm. 2i- and serum-ESC: WT01. EpiSC: cWT01.

To further confirm our observation we used previously published H3K27me3 ChIP-seq datasets to quantify the reads mapping specifically on the major and minor satellite repeats. With the first dataset (Marks et al., 2012) we observed a 2-fold enrichment in 2i-ESC compared to serum-ESC for H3K27me3 at major satellites, while no difference was found for minor satellites sequences (Figure 52A). Another dataset allowed comparing 2i-ESCs and EpiSCs (Zylicz et al., 2015) and it shows that 2i-ESCs presents a more than 2-fold enrichment of H3K27me3 on major satellites compared to input (confirming Marks et al., 2012), but no enrichment at all in EpiSC (Figure 52B). Conversely minor satellites showed no H3K27me3 enrichment in any cell type confirming once more Marks et al., 2012 (Figure 52B). Interestingly by ChIP-seq analysis from the first dataset (Marks et al., 2012) no reduction of H3K9me3 was observed from either major or minor satellite when serum-ESCs were adapted into the 2i medium (Figure 52C), contrasting to what we visually see in mitotic chromosome ends. However in the second study (Zylicz et al., 2015) they did not perform ChIP-seq for H3K9me3 so we could not confirm the result of Marks et al., 2012 concerning 2i- and serum-ESCs neither compare those to EpiSCs for this histone mark.

Altogether our data indicate a considerable reorganization especially for PCH (major satellites) in the different types of pluripotent cells, with a likely progressive enrichment of H3K9me3 and a clear depletion of H3K27me3 from 2i to serum-conditions of ESCs and up to EpiSCs.

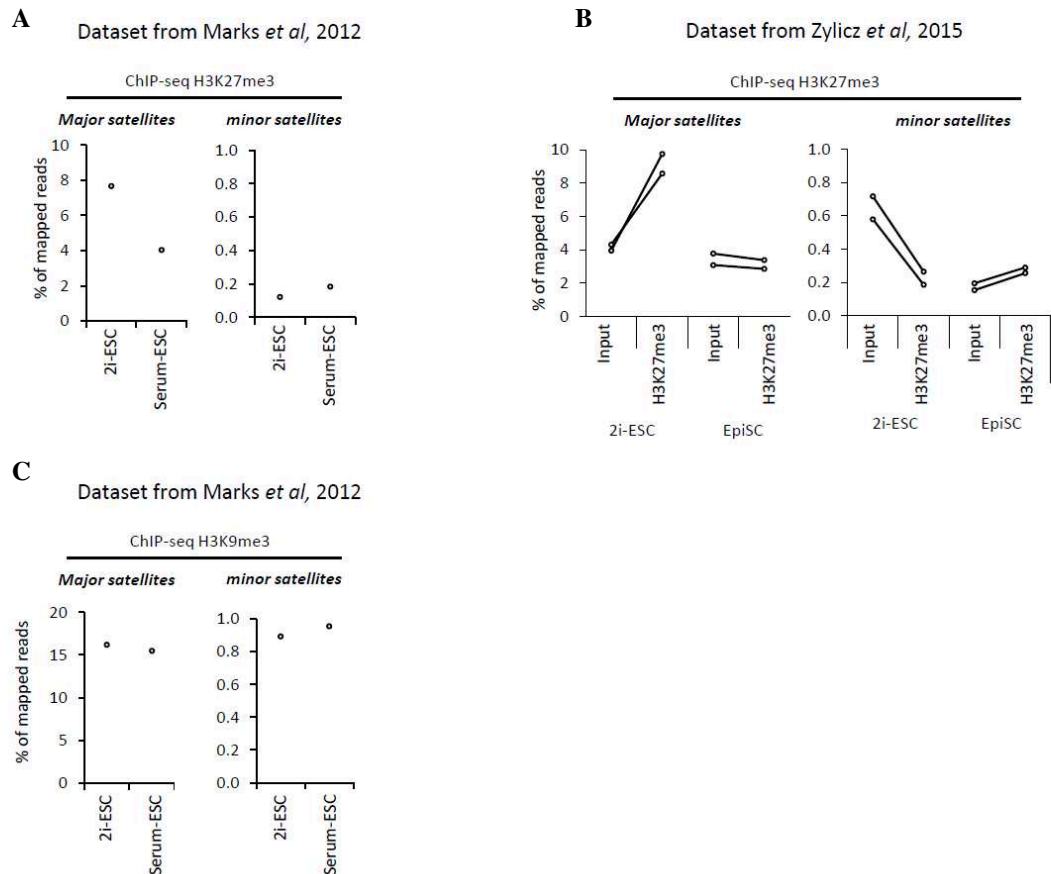


Figure 52 : H3K27me3 is specifically enriched at PCH (Major satellites repeats) only in 2i-ESCs

(A) Percentage of mapped reads on major and minor satellite repeats sequence using H3K27me3 ChIP-seq (Marks *et al.*, 2012). (B) Percentage of mapped reads on major and minor satellite repeat sequences using H3K27me3 ChIP-seq compared to input (Zylitz *et al.*, 2015). (C) Percentage of mapped reads on major and minor satellite repeats sequence using H3K9me3 ChIP-seq (Marks *et al.*, 2012). 2i and serum-ESC: E14 (Marks *et al.*, 2012) and GGO (Zylitz *et al.*, 2015). EpiSC: X6 (Zylitz *et al.*, 2015).

1.3.3 DNA methylation state of repetitive sequences reflects the pluripotent stem cell type

Another important actor of the heterochromatin compartment is DNA methylation. We have first investigated the organization of the 5-meC, but also 5h-meC, in our cells by immunostaining and then evaluated the DNA methylation levels on major and minor satellites by Southern blot on genomic DNA digested with methyl-sensitive enzymes. It should be noted that only a methanol treatment permitted a correctly 5-meC staining in EpiSCs, while a classic PFA-Triton treatment was insufficient to deproteinize the chromatin (Data not shown), suggesting that EpiSC's chromatin is more compacted than ESC's one. With this aggressive treatment we observed a similar staining pattern in 2i-ESCs, serum-ESCs and EpiSCs with 5-meC aggregating in clusters in all nuclei (Figure 53). 5h-meC however showed a nucleoplasmic diffused signal in every pluripotent stem cells but interestingly it is also

accumulated in foci in some 2i-ESCs only (Figure 53). However previous data have shown that 5h-meC is 10-times less abundant than 5-meC (Ficz et al., 2011; Habibi et al., 2013; Leitch et al., 2013) we considered that the contribution of 5h-meC should be minimal compared to 5-meC, so in the following parts we will discuss only about DNA methylation.

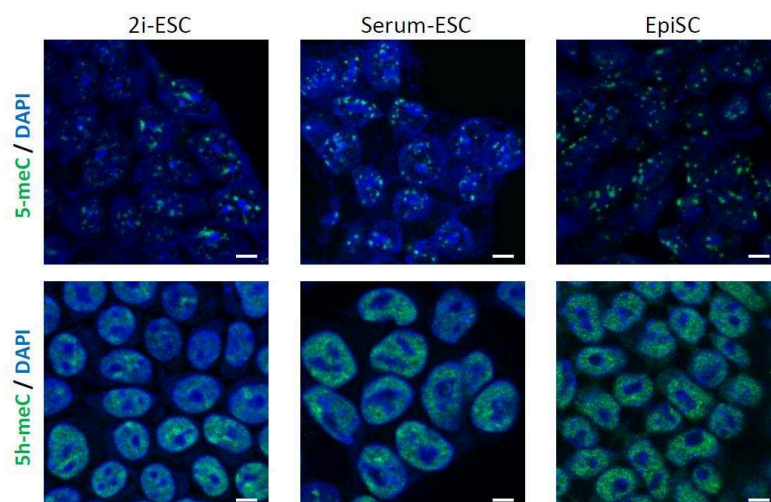


Figure 53: Global distribution of methylation and hydroxymethylation of DNA in the different types of pluripotent stem cells.

Upper part: Immunostaining images (single-plan) for 5-meC (green) with DAPI DNA counterstaining (blue). Scale bars represent 5 μ m. Lower part: Immunostaining images (single-plan) for 5h-meC (green) with DAPI DNA counterstaining (blue). Scale bars represent 5 μ m. 2i- and serum-ESC: R1. EpiSC: FT129.1.

Conversely to the 5-meC staining patterns that do not highlighted any differences, Southern-blot analysis revealed impressive changes of DNA methylation on major and minor satellites between the different types of pluripotent stem cells (Figure 54). Major satellites were partially demethylated in 2i-ESCs, as shown by the linescan profile (red line) which is intermediate between the fully demethylated DnmtTKO cells (purple line) and the hypermethylated fibroblasts (MEFs-black line) (Figure 54A). On the contrary these sequences in both serum-ESCs (blue line) and EpiSCs (green line) were as methylated as the MEFs (Figure 54A). A similar situation was observed for minor satellites that were partially demethylated in 2i-ESCs and as methylated as MEFs in EpiSCs (Figure 54B). Serum-ESCs showed for minor satellites an intermediate pattern between 2i-ESCs and EpiSCs (Figure 54B).

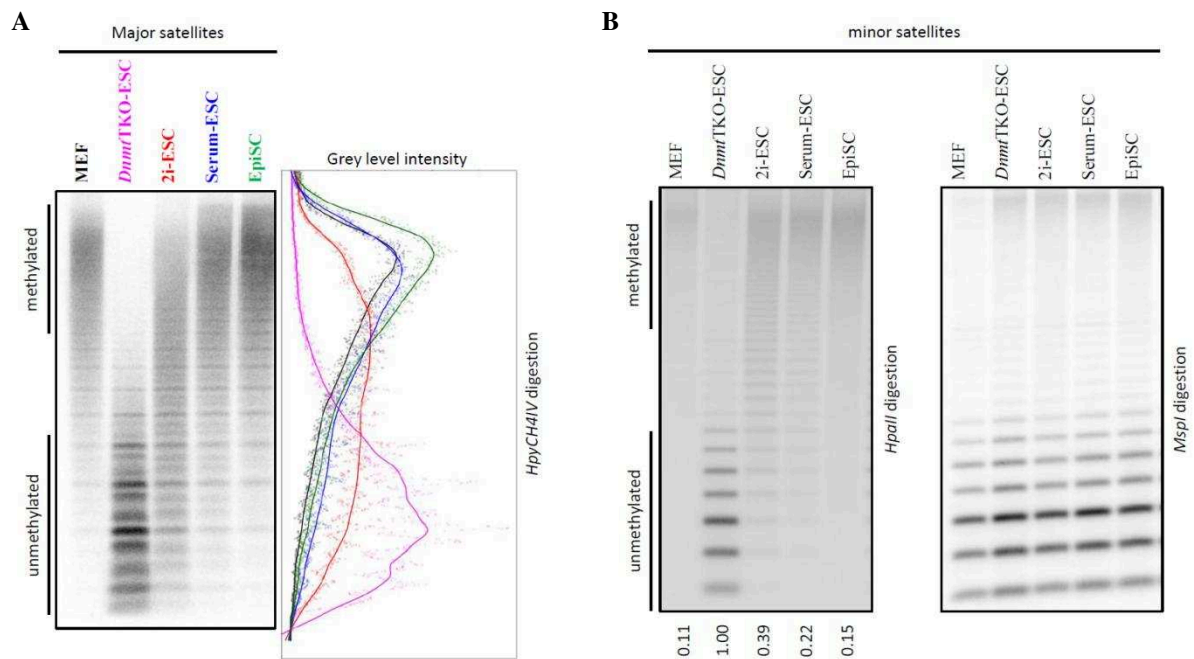


Figure 54: Methylation profile at major and minor satellites in 2i-ESCs, serum-ESCs and EpiSCs.

(A) Southern-blot analysis of gDNA digested with HpyCH4IV revealed with probe for major satellites. Linescan quantification for each lane: MEF (black), DnmtTKO (pink), 2i-ESC (red), serum-ESC (blue) and EpiSC (green). (B) Southern-blot analysis of gDNA digested with HpaII revealed with probe for minor satellites. Quantification related to southern-blot of gDNA digested with MspI. DnmtTKO-ESC is set to 1. 2i- and serum-ESC: R1. EpiSC: FT129.1.

We then assessed the expression level of the two main enzymes responsible for de novo methylation: DNMT3A and DNMT3B (Figure 55A-B). Dnmt3a expression level increased strongly between 2i-ESCs and serum-ESCs and even more in EpiSCs. In parallel the embryonic isoform of the DNMT3A protein was about ten times more abundant in serum-ESCs than in 2i-ESCs, while EpiSCs gained the additional somatic form as shown by western-blot analysis (Chen et al., 2002; Veillard et al., 2012) (Figure 55A). Unfortunately we did not find any antibody working in immunostaining for DNMT3A to assess its nuclear distribution. In parallel Dnmt3b was transcribed at low level in ESCs (2i- or serum-) and strongly increased in EpiSCs (Figure 55B). DNMT3B protein level showed a sequential increment going from 2i- to serum-ESCs and finally EpiSCs. Of particular interest is the subnuclear localization of DNMT3B which presents foci-enrichment in ESCs (2i- or serum-) and a diffuse nucleoplasmic signal in EpiSCs (Figure 55B).

In conclusion we observed a progressive methylation of major and minor satellites, accompanied by an increased expression of the de novo methyltransferases, from 2i-ESCs to EpiSCs.

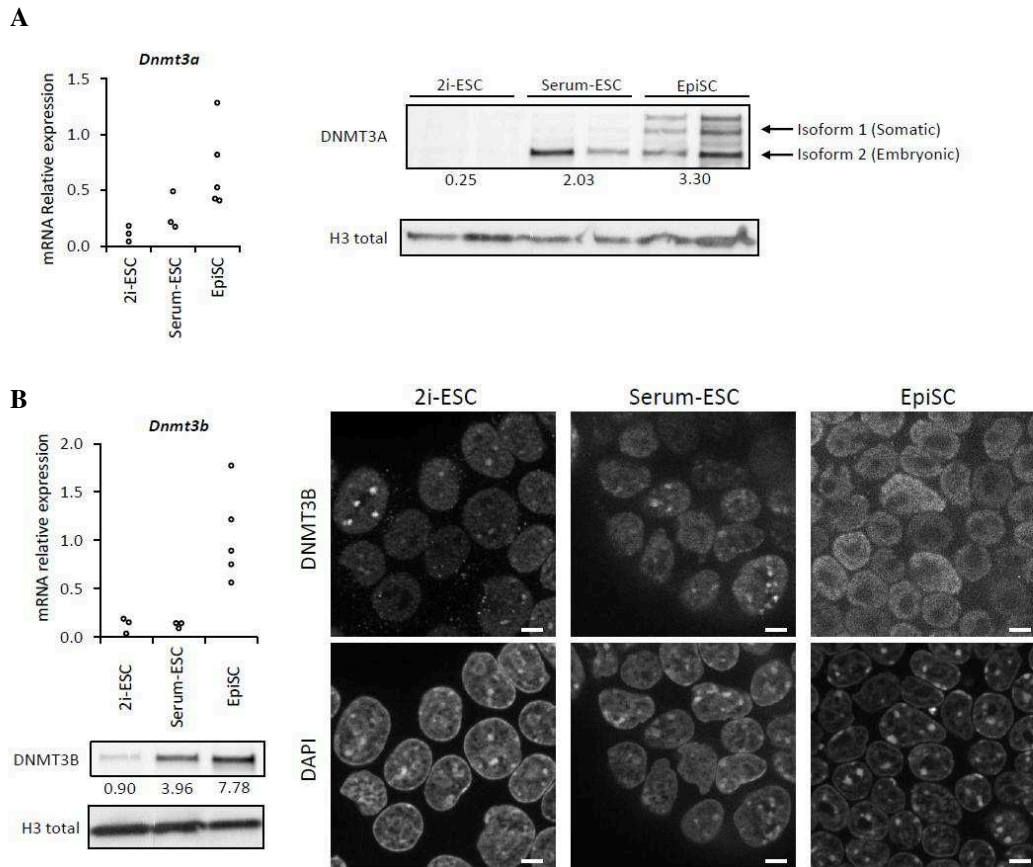


Figure 55: De novo DNA methylation machinery in 2i-ESCs, serum-ESCs and EpiSCs
 (A) Left part: Relative expression (CNRQ) of *Dnmt3a* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Right part: Western-blot analysis for quantification of bulk levels of DNMT3A related to H3 total. (B) Upper left part: Relative expression (CNRQ) of *Dnmt3b* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Lower left part: Western-blot analysis for quantification of bulk levels of DNMT3B related to H3 total. Right part: Immunostaining images (single-plan) for DNMT3B with DAPI DNA counterstaining. Scale bars represent 5 μ m. 2i- and serum-ESC: R1. EpiSC: FT129.1 and cR1.

1.3.4 2i condition leads to decondensation of the well-organized structure of the chromocenter

After having shown that PCH differed in their epigenetic marks, we investigated whether there were any differences also in the three-dimensional organization of these regions depending on the state of pluripotency. We performed DNA-FISH for major and minor satellites (Figure 56), followed by nuclei-segmentation and three-dimensional single-cell reconstruction with AMIRA 3.1 software (Figure 56). In EpiSCs and serum-ESCs major satellites were organized into round compact domains, surrounded by smaller dots of minor satellite domains, as classic chromocenters (Figure 56) (Guenatri et al., 2004).

By contrast, they formed decondensed domains of irregular size and shape in 2i-ESCs (blebs, half-rings around nucleolus or at the nuclear periphery), reflecting globally unstructured chromocenters (Figure 56).

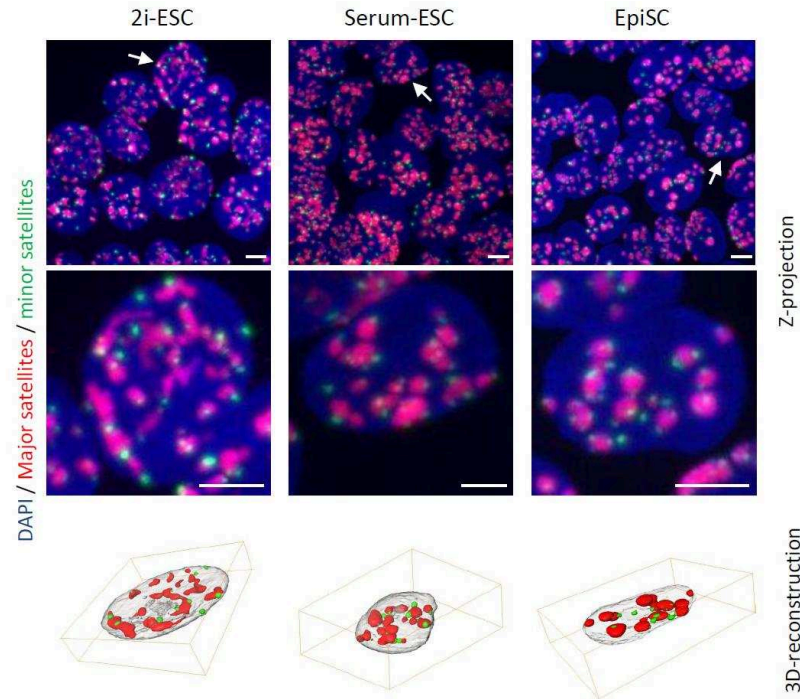


Figure 56: Three-dimensional organization of major and minor satellites in the different types of mouse pluripotent stem cells.

Three-dimensional DNA-FISH images (z-projection) for major (red) and minor (green) satellites with DAPI DNA counterstaining (blue). Magnification on a single cell (arrowed) and three-dimensional reconstruction of major and minor satellite signals using AMIRA 3.1 software. Scale bars represent 5µm. 2i- and serum-ESC: R1. EpiSC: FT129.1.

A similar phenomenon has been previously observed as a consequence of down-regulation of the histone chaperone CAF-1/p150 in ESCs (Houlard et al., 2006). By immunostaining we showed that only 25% of 2i-ESCs displayed almost one CAF1/p150 positive foci in their nucleus. This proportion increased in serum-ESCs (33%) and reached 42% in EpiSCs (Figure 57). This suggests that such decondensation of major satellites in 2i-ESCs could be the consequence of a reduction of CAF-1/p150 association to these sequences.

Taken together we observed that 2i-ESCs display highly decompacted major satellite organization in contrast to serum-ESCs and EpiSCs which present a well-structured chromocenter, close to the somatic one (particularly for EpiSCs).

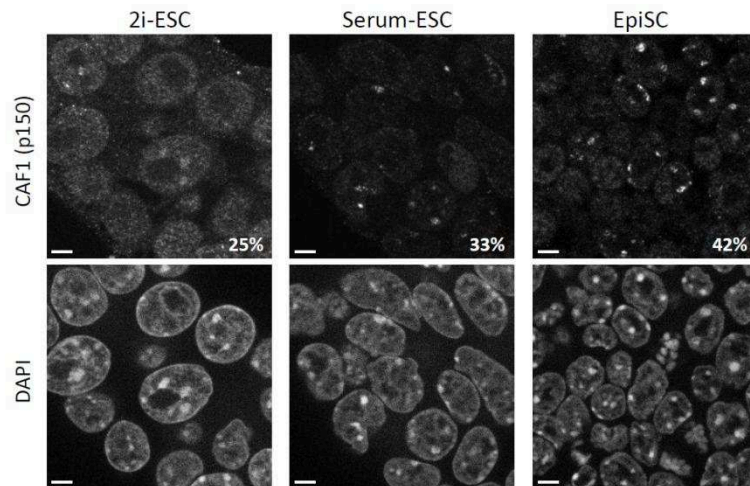


Figure 57: CAF-1, the histone chaperon of constitutive heterochromatin in mouse pluripotent stem cells.

Immunostaining images (single-plan) for CAF-1 p150 with DAPI DNA counterstaining. % indicates the percentage of cell in the population displaying a foci-enrichment pattern. Scale bars represent 5 μ m. 2i- and serum-ESC: R1. EpiSC: FT129.1.

1.3.5 The transcriptional state of major and minor satellites depends on pluripotent cell type

Next we wondered whether the different epigenetic states of centromeric and pericentromeric regions reflected different levels of transcription of these sequences. In EpiSCs, in which these sequences showed high DNA methylation levels, strong H3K9me3-enrichment and classic chromocenter structure, both major and minor satellites are very lowly expressed as in somatic cells (Figure 58) (Efroni et al., 2008; Lu and Gilbert, 2007; Probst et al., 2010). On the contrary, serum-ESCs in which we observed strong enrichment of H3K9me3 but slightly less DNA methylation, presented high (but variable) levels of both satellite transcripts (Figure 58). Interestingly, 2i-ESCs, which showed partial DNA hypomethylation and strong enrichment of H3K27me3 compared to H3K9me3 on these sequences, expressed major and minor satellites at very low level (Figure 58).

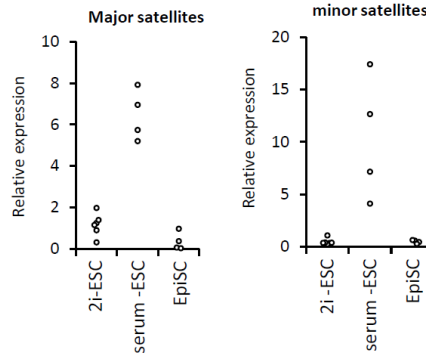


Figure 58: Transcription of major and minor satellites ncRNAs in the different types of pluripotent stem cells.

Relative expression (CNRQ) of major and minor satellite transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent biological replicate. 2i- and serum-ESC: R1 and WT01. EpiSC: FT129.1 and cWT01.

Based on the recent discovery that NANOG regulates the special PCH organization and transcription of these regions in serum-ESCs (Novo et al., 2016) we evaluated by western blot the NANOG global levels and by immunostaining its subnuclear localization. EpiSCs presented a very low level of NANOG compared to serum-ESCs and interestingly we found a reduced one in 2i-ESCs as well (Figure 59). By immunostaining we saw that 2i condition led to the expected homogeneous NANOG expression while serum condition induces the well-known NANOG heterogeneity with high- and low- NANOG expressing ESCs (Wray et al., 2010) (Figure 59).

In conclusion we observed a striking difference in major and minor satellites transcripts levels which are both highly expressed in serum-ESCs, strongly reduced in 2i-ESCs and even more in EpiSCs, in accordance to NANOG expression levels.

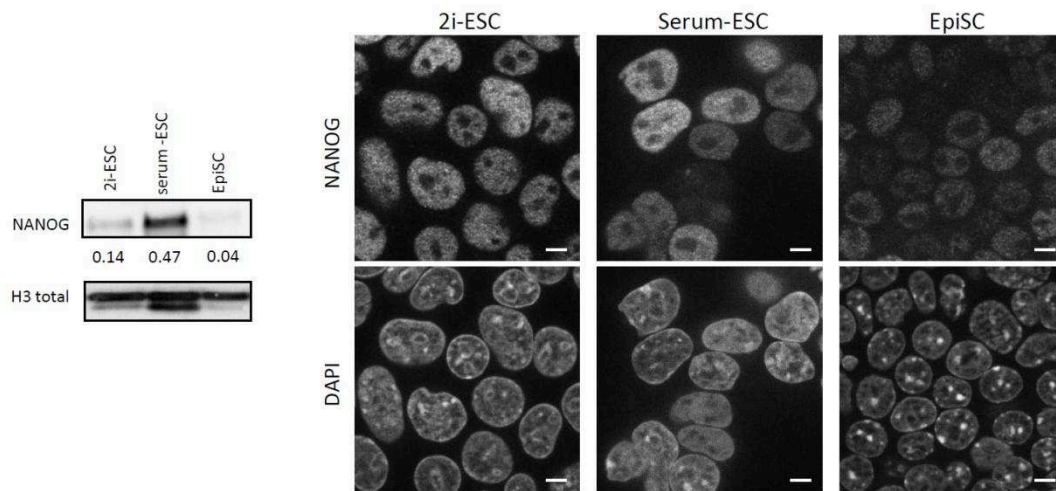


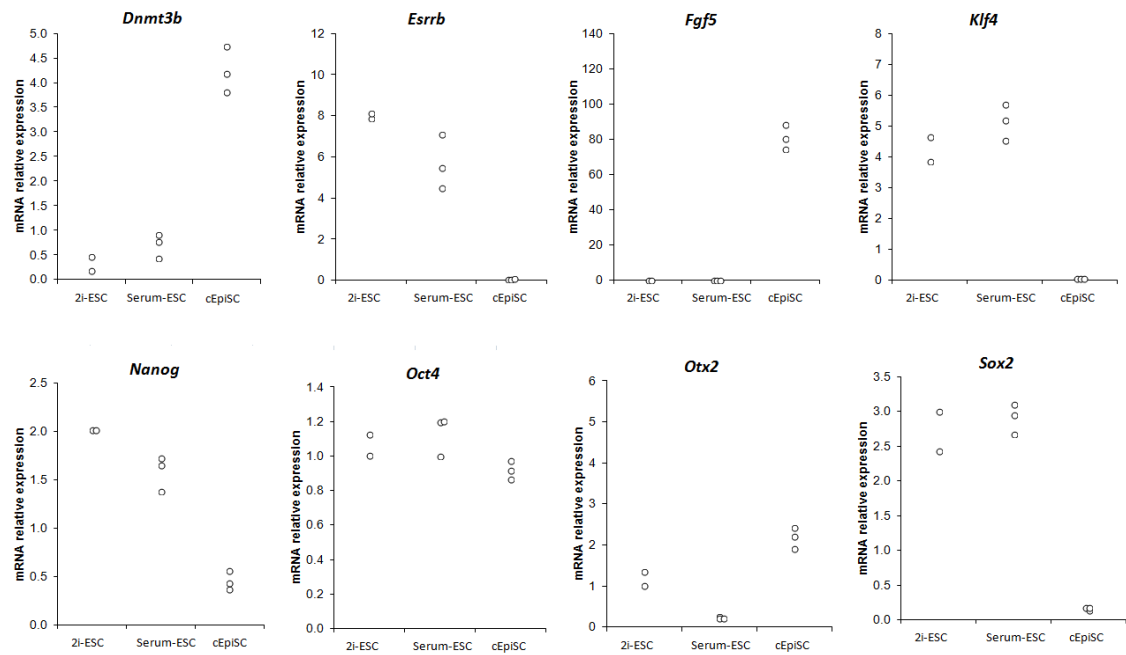
Figure 59: NANOG expression and distribution in mouse pluripotent stem cells.

Left part: Western-blot analysis for quantification of bulk levels of NANOG related to H3 total. Right part: immunostaining images (single-plan) for NANOG with DAPI DNA counterstaining. Scale bars represent 5 μm. 2i- and serum-ESC: R1. EpiSC: FT129.1.

1.3.6 Absence of Suv39h1/2 induces different phenotypes depending on the pluripotent stem cell type

To further decipher the different epigenetic pathways that control transcription of constitutive heterochromatin, we examined the PCH organization in Suv39hdn ESCs, which lack the HMTs responsible for H3K9 trimethylation specifically at pericentromeric regions (Peters et al., 2003). We adapted these cells in 2i culture conditions and converted them in vitro in EpiSCs (cEpiSCs). We next verified by qRT-PCR that the primed state of pluripotency is correctly established even in absence of H3K9me3 at the constitutive heterochromatin. We found no differences between wild-type and Suv39hdn concerning the transcription of naive- or primed-specific genes (Figure 60).

Wild-type



Suv39hdn

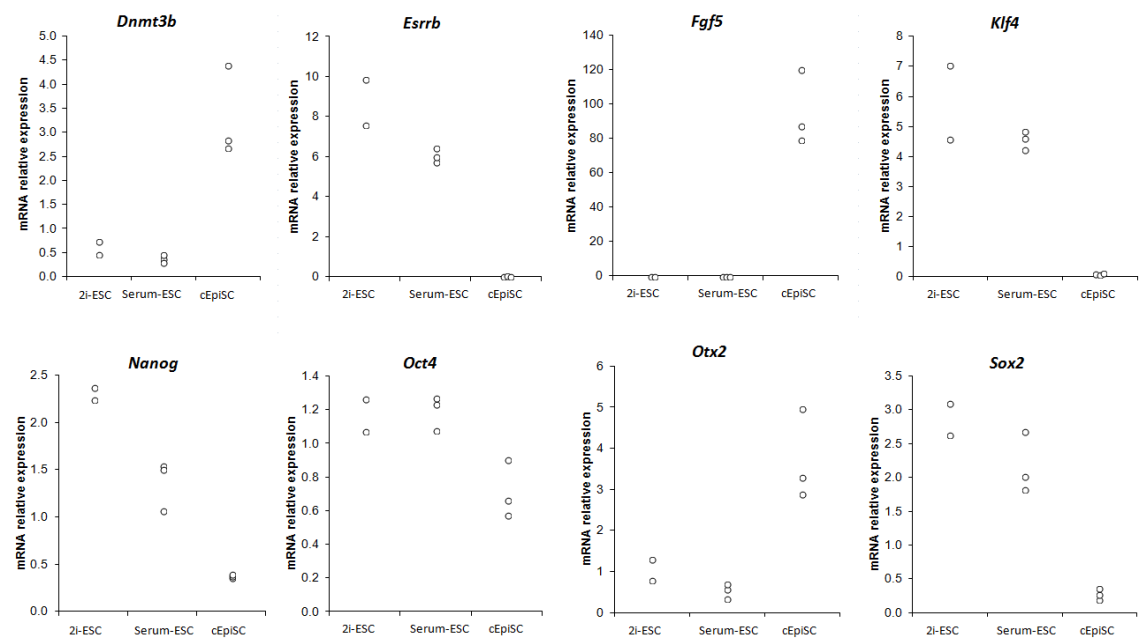


Figure 60: Validation of conversions from Wild-Type and Suv39hdn ESCs

Relative expression (CNRQ) of different pluripotent (Nanog, Esrrb, Klf4, Sox2, Oct4) and primed (Dnmt3b, Fgf5, Otx2) specific transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Three independent conversions were made. Wild-type: WT01.

We first investigated the subnuclear organization of H3K9me3 and H3K27me3 by immunostaining and their global levels by western-blot. As previously published (Cooper et al., 2014; Martens et al., 2005; Peters et al., 2003; Saksouk et al., 2014), we confirmed that in serum-ESCs, the absence of H3K9me3 at PCH induced deposition of H3K27me3 at these regions (Figure 61). Indeed we found that 60% of Suv39hdn serum-ESCs presented clusters of H3K27me3 staining colocalizing with DAPI-dense foci (Figure 61). Strikingly in both mutant and wild-type cEpiSCs, we observed the same low and diffuse pattern of H3K27me3, with no foci-enrichment (Figure 61). In addition Suv39hdn and wild-type cEpiSCs present the same decreased bulk level of H3K27me3 compared to ESCs (Figure 61).

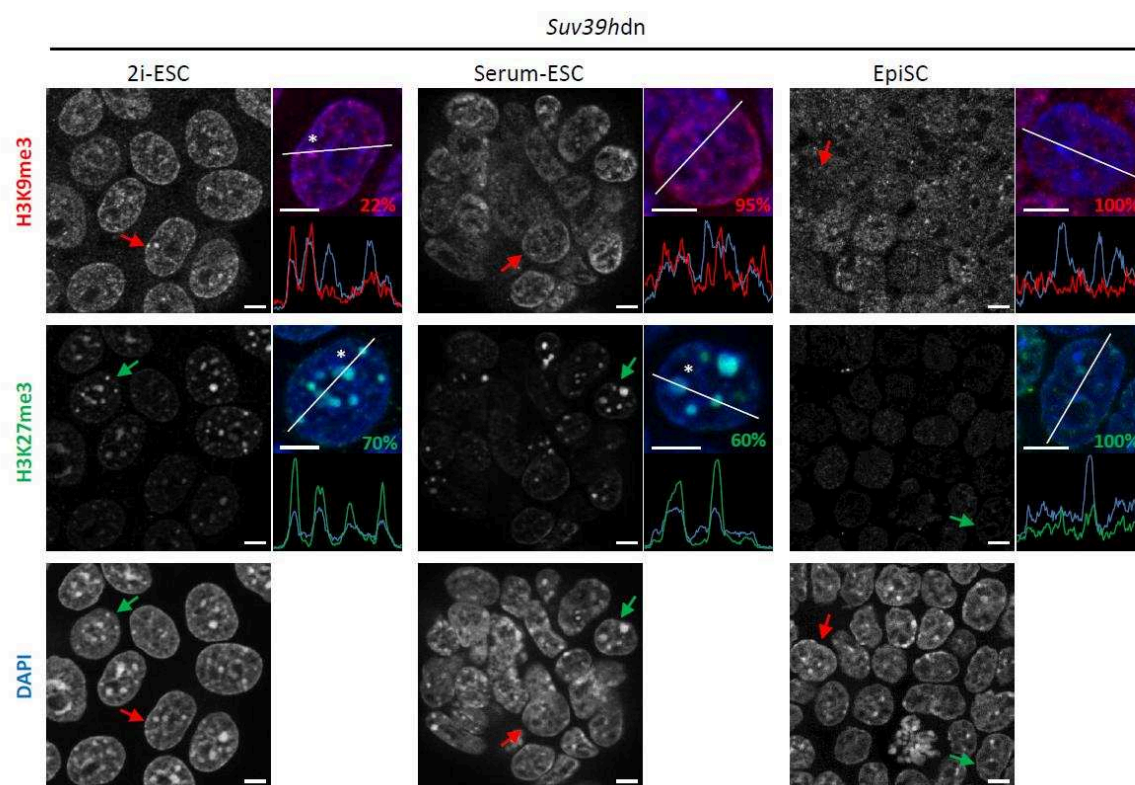


Figure 61: Contrasting organization of repressive histone marks due to Suv39hdn condition in 2i-ESCs, serum-ESCs and EpiSCs.

Immunostaining images (single-plan) for H3K9me3 and H3K27me3 with DAPI DNA counterstaining in Suv39hdn condition, to compare with wild-type condition (Fig. 1A). Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) or H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5µm.

H3K9me3 staining revealed mostly a diffuse pattern in serum-ESCs (with only 5% of cells presenting some foci-enrichment) and cEpiSCs (no foci at all) (Figure 61). The bulk level of H3K9me3, as expected, was reduced in the mutant condition (Figure 62). The 2i adapted

mutant ESCs showed no major differences in terms of epigenetic organization compared to the wild-type conditions (Figure 62). Notably, a small proportion (22%) of these mutant ESCs in 2i exhibited small H3K9me3-enriched foci similar to those seen in wild-type cells (Figure 62), indicating an H3K9me3-deposition at PCH independent of SUV39H1/2 enzymes. Such a phenomenon seems to be highly marginal in serum-ESCs, as only 5% of cells presented such foci in absence of SUV39H1/2.

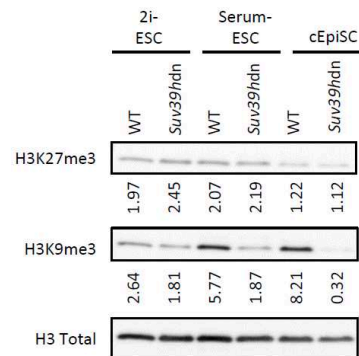


Figure 62: Global levels of repressive histones in wild-type and Suv39h1/2 condition in 2i-ESCs, serum-ESCs and EpiSCs.

Western-blot analysis for quantification of bulk levels of repressive histone modification H3K9me3 and H3K27me3 related to H3 total in wild-type and Suv39h1/2 condition. WT: WT01.

Next, we evaluated DNA methylation levels on satellite sequences by Southern-blot in these mutant cells and confirmed that the absence of Suv39h1/2 in serum-ESCs induced a reduction of DNA methylation compared to wild-type on major satellites (Figure 63A) (Note the shift of the blue dotted line compared to the continuous one) but not on minor satellites (Figure 63B) (Lehnertz et al., 2003). Hence, the 2i condition induces a phenotype that recapitulates the absence of Suv39h1/2 in naive ESCs: increased H3K27me3 and reduced DNA methylation at PCH. Interestingly no such decrease of DNA methylation on major satellites was observed in the mutant cEpiSCs (Figure 63A) (the green lines are quite superimposed). Therefore the absence of H3K9me3 does not interfere with the establishment of a hypermethylated state at major satellites in primed EpiSCs.

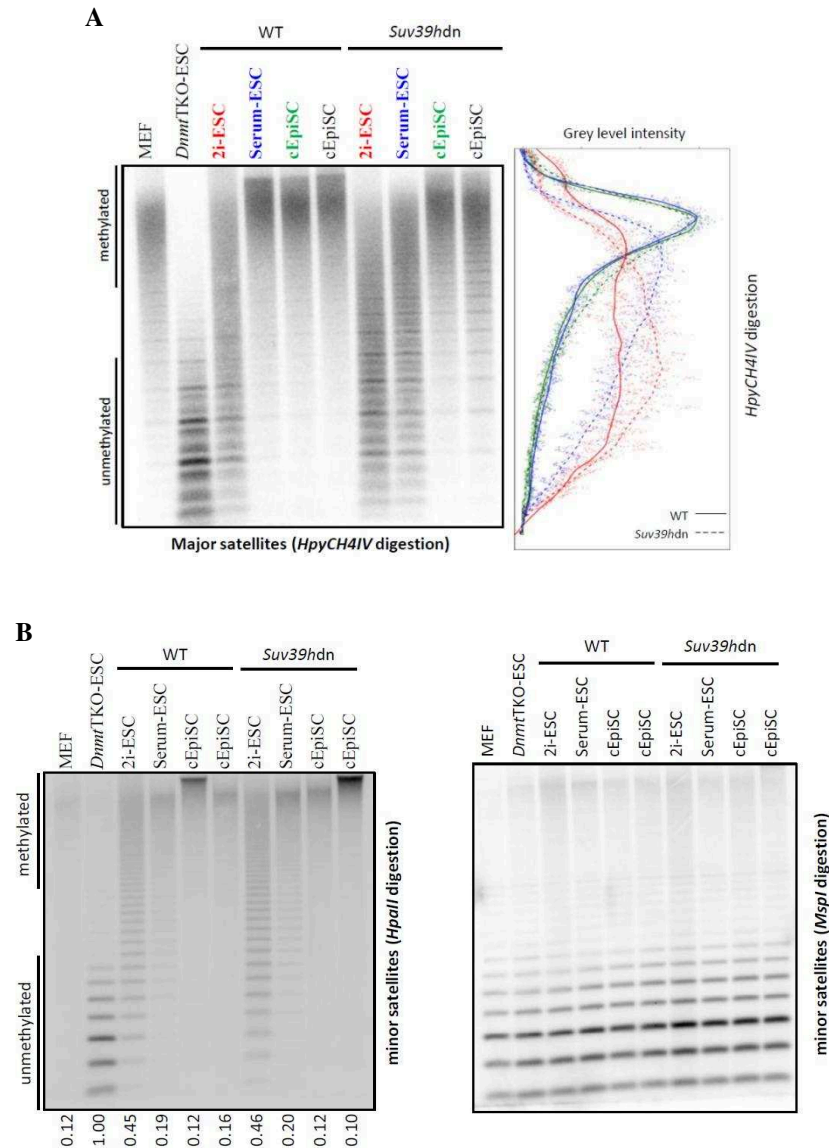


Figure 63: Methylation profile at major and minor satellites in wild-type and Suv39hdn conditions in 2i-ESCs, serum-ESCs and EpiSCs.

(A) Southern-blot analysis of gDNA digested with *HpyCH4IV* revealed with probe for major satellites in wild-type and Suv39hdn conditions. Linescan quantification for each lane: 2i-ESC (red), serum-ESC (blue) and EpiSC (green). Wild-type condition is represented with a continuous line, while Suv39hdn with a dotted line. (B) Southern-blot analysis of gDNA digested with *HpaII* revealed with probe for minor satellites in Wild-type and Suv39hdn conditions. Quantification related to southern-blot of gDNA digested with *MspI*. DnmtTKO-ESC set to 1. WT: WT01

In contrast to what would be expected from the loss of a repressive mark, transcription of major satellites was slightly decreased in Suv39hdn serum ESCs (Figure 64A), in contrast to what was previously shown by Lehnertz et al., 2003 using standard PCR (semi-quantitative PCR). In this same context minor satellite transcripts were 2-fold reduced in mutant compared to wild-type conditions (Figure 64A). This suggests that both the absence of repressive H3K9me3 and a low level of DNA methylation do not always implicate up-regulation of

transcription. Very interestingly in Suv39hdn cEpiSCs, transcription of major satellites was de-repressed, as we detected a similar level of transcripts as in the mutant serum-ESCs (Figure 64A). We verified that this up-regulation of major satellites in Suv39hdn-cEpiSCs was not due to a deregulation of NANOG (Novo et al., 2016). Indeed, NANOG bulk level does not vary whatever the conditions (Figure 64B). Unfortunately transcription of minor satellites was less conclusive, with variable levels of transcription within biological replicates (Figure 64A). In 2i condition we observed a tendency for up-regulation of major and minor satellite transcripts in Suv39hdn cells compared to wild-type (Figure 64A). Altogether, while in serum-ESCs the absence of H3K9me3 at PCH leads to reduced methylation level and accumulation of H3K27me3, in EpiSCs no such changes occur and the transcription of major satellites is not repressed anymore.

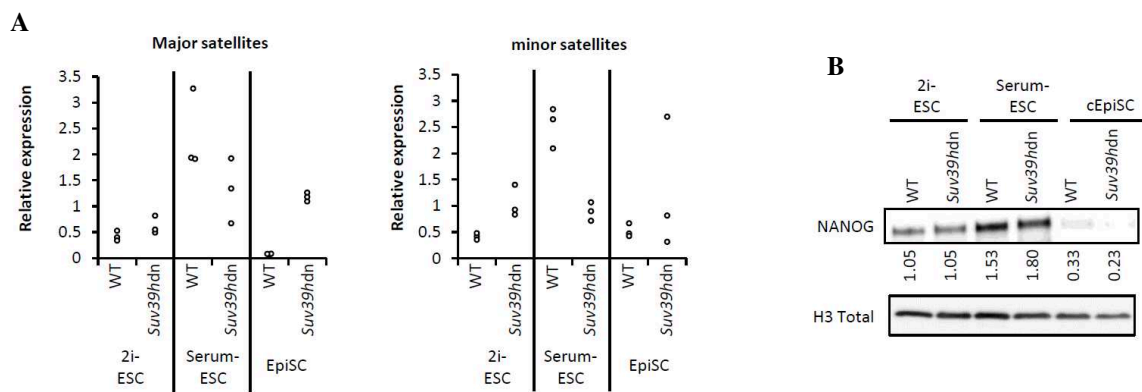


Figure 64: Transcription of major and minor satellites ncRNAs in wild-type and Suv39hdn conditions in the different types of mouse pluripotent stem cells with associated NANOG levels.

(A) Relative expression (CNRQ) of major and minor satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in Wild-type and Suv39hdn condition. Each point is an independent biological replicate. (B) Western-blot analysis for quantification of bulk levels of NANOG related to total H3 in wild-type and Suv39hdn condition. WT: WT01.

1.3.7 Absence of DNA methylation has limited effects on satellite transcription

To study the effects of DNA methylation in the epigenetic pathway involved in regulation of major and minor satellite organization and transcription, we used *Dnmt1*, 3a, and 3b triple knock-out ESCs (*DnmtTKO*). These cells do not have any methylated cytosine in their genome (Tsumura et al., 2006). We analyzed the H3K9me3 and H3K27me3 subnuclear organization by immunostaining (Figure 65). We were able to adapt these cells in 2i but not to convert them into cEpiSC because of cell death (apoptosis), in agreement with their low contribution to the development of the epiblast in *DnmtTKO* nuclear-transfer embryos

(Sakaue et al., 2010). DnmtTKO serum-ESCs gained strong enrichment of H3K27me3 at PCH/CH foci, as expected from previous studies (Figure 65) (Cooper et al., 2014; Saksouk et al., 2014). Indeed we found that most (84%) of mutant serum-ESCs present H3K27me3 enrichment at DAPI-dense foci (Figure 65). This enrichment was at the expense of H3K9me3 (Figure 65), as only 52% of DnmtTKO serum-ESCs contain H3K9me3 foci compared to the totality in wild-type condition (Figure 65). In 2i condition, the complete absence of DNA methylation led to an increased proportion of ESCs with H3K27me3 marked foci: 94% 2i-DnmtTKO compared to 69% in wild-type (Figure 65). The pattern of H3K9me3 remained very similar in 2i-DnmtTKO compared to wild-type, so small and rare foci (Figure 65), but the proportion of 2i-ESCs with H3K9me3-foci rise up to 41% (Figure 65).

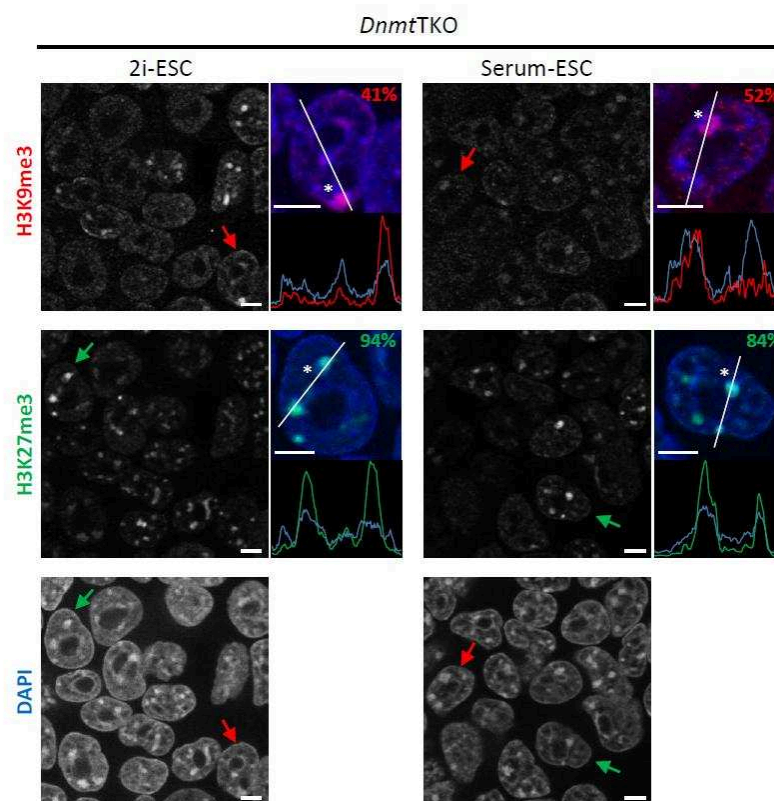


Figure 65: Repressive histone modifications rearrangement in consequence of DNA methylation absence in 2i- and serum-ESCs.

Immunostaining images (single-plan) for H3K9me3 and H3K27me3 with DAPI DNA counterstaining in DnmtTKO condition, to compare with wild-type condition (Fig. 1A). Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) or H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5 μm.

Transcription of satellites was globally unchanged, except in serum-DnmtTKO where minor transcripts were strongly down-regulated (Figure 66). In summary the absence of DNA methylation dramatically modify the epigenetic state of PCH/CH in serum-ESCs, inducing a strong enrichment of H3K27me3 and a reduction of H3K9me3-enriched foci, and also causes a repression of minor satellite transcription in the same cells.

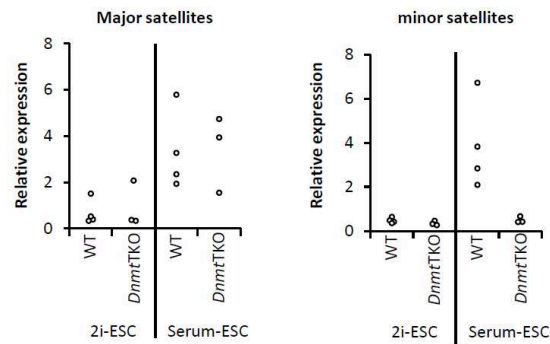


Figure 66: Transcriptional consequences of DNA methylation absence on major and minor satellites in 2i- and serum-ESCs.

Relative expression (CNRQ) of major and minor satellites noncoding transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in Wild-type and DnmtTKO condition. Each point is an independent biological replicate. WT: R1 and WT01.

1.3.8 Reduced levels of H3K27me3 do not up-regulate satellite transcription

In order to study the role of H3K27me3 in the regulation of PCH/CH transcription status, we used an inhibitor of the HMT activity of EZH2. We chose EPZ-6438 (further call for simplicity EPZ), a selective EZH2 inhibitor with a strong potential use in cancer therapy (Knutson et al., 2014). The treatment of ESCs with EPZ for 72 hours led to an impressive reduction of the bulk levels of H3K27me3 (at least 70% in each cell types tested), with no major changes in H3K9me3 levels (Figure 67A). Immunostaining of 2i-ESCs treated with EPZ also confirmed the loss of H3K27me3 foci in the vast majority of cells with no changes in H3K9me3 organization compared to DMSO-treated control cells (Figure 67B). In addition the hypomethylated status of satellite sequences in 2i-ESCs treated cells was not affected (Figure 67C). We then analyzed the transcription of major and minor transcripts. Unexpectedly, the loss of H3K27me3 in 2i-ESCs did not lead to an up-regulation of satellite transcription. On the contrary we observed a slight reduction of major satellite transcripts and no appreciable changes for minor satellite transcripts (Figure 67D).

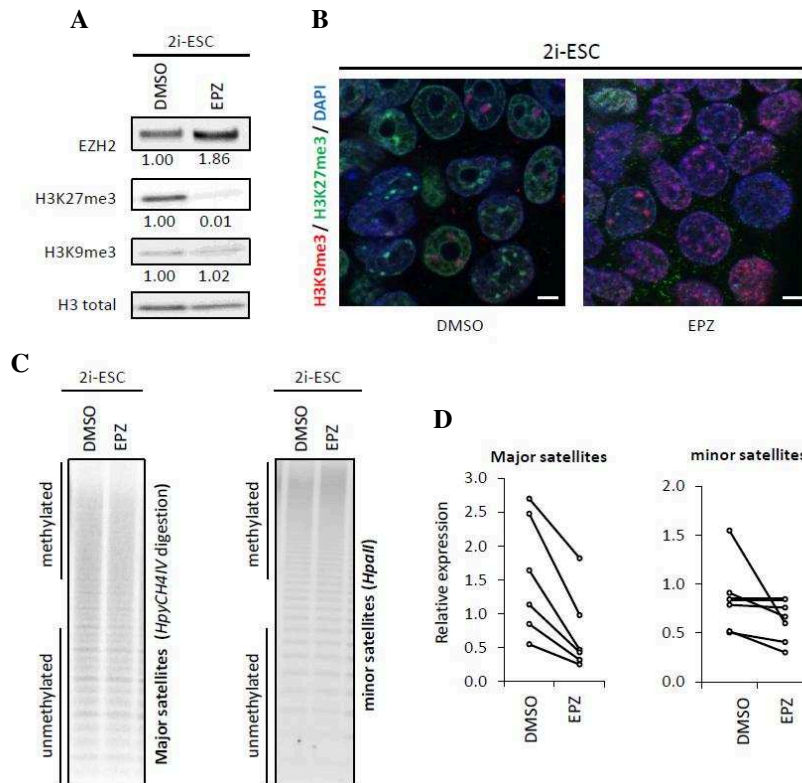


Figure 67: Reduced levels of H3K27me3 do not induce up-regulation of satellite repeats in 2i-ESCs.

(A) Western-blot analysis for quantification of bulk levels of the HMT enzyme EZH2 and the repressive histone modification H3K9me3 and H3K27me3 related to total H3 in 2i-ESC treated with DMSO (Control) or EPZ (EZH2 inhibition). (B) Immunostaining images (single-plan) for H3K9me3 (red) and H3K27me3 (green) with DAPI DNA counterstaining (blue) in 2i-ESC treated with DMSO or EPZ. Scale bars represent 5 μ m. (C) Southern-blot analysis of gDNA digested with HpyCH4IV or HpaII revealed with probe for major or minor satellites respectively in 2i-ESCs DMSO or EPZ-treated. (D) Relative expression (CNRQ) of major satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in 2i-ESC treated with DMSO or EPZ. Each point is an independent biological replicate. 2i-ESC: R1.

When serum-ESCs were treated with the EZH2 inhibitor, although H3K27me3 is not as highly enriched at PCH as in 2i condition, we also observed a reduction of major as well as minor satellite transcripts (Figure 68A). We performed the same inhibition in mutant cells that present a similar enrichment of H3K27me3 at PCH (namely *Suv39hdn* serum-ESCs and both *DnmTKO* ESCs). In all cases, we never observed up-regulation of satellite transcription, but in contrast a slight reduction of major satellites transcripts, at least in serum-ESCs (Figure 68B, C, D). To exclude a possible effect of EPZ treatment on pluripotency of ESC we verified by western-blot that the global levels of NANOG were not altered (Figure 68E). Altogether, treatment with EZH2 inhibitor strikingly does not induce an up-regulation of satellite transcripts even in absence of the other repressive marks (such as H3K9me3 or DNA methylation) at the same sequences.

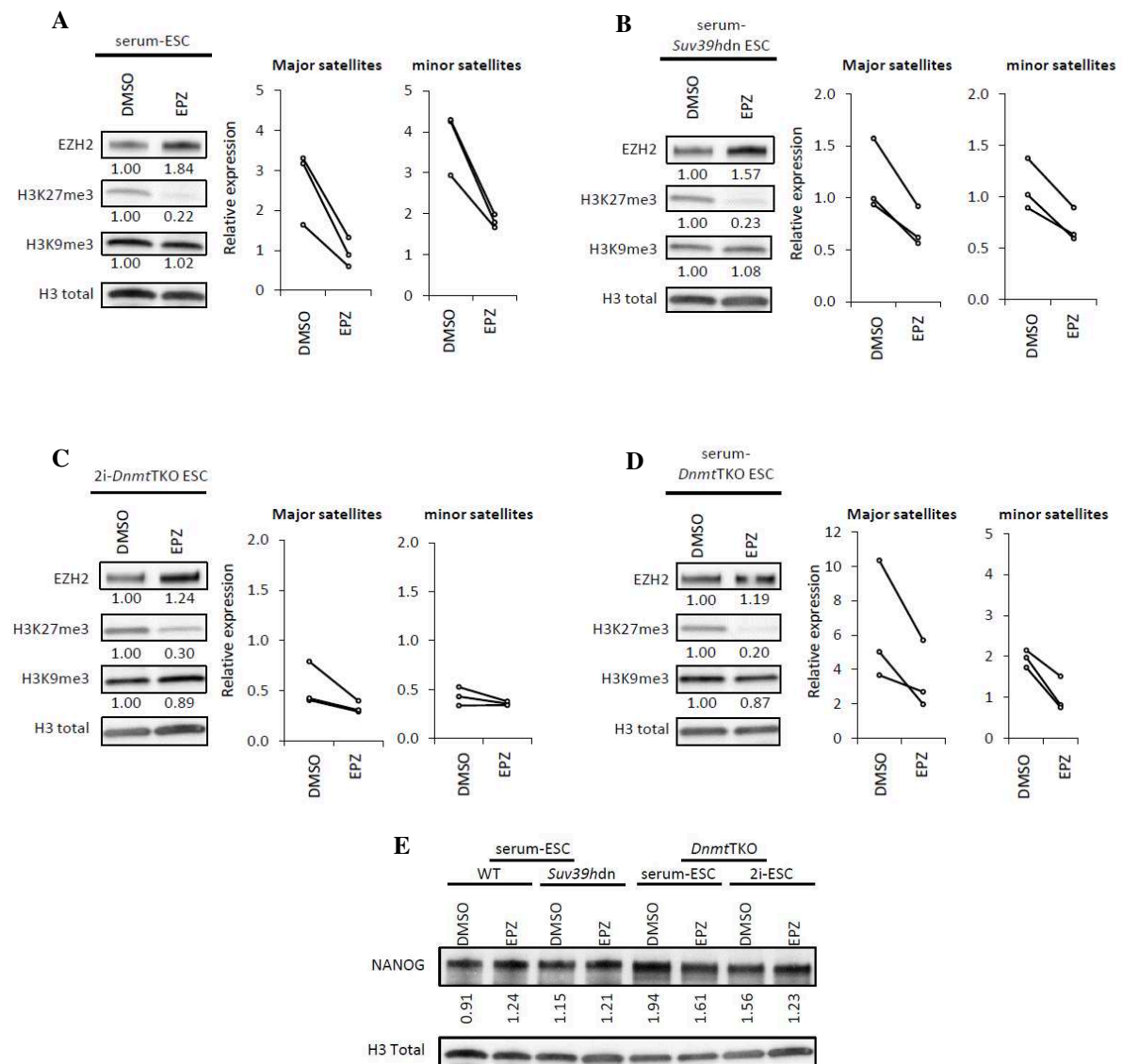


Figure 68: Reduced levels of H3K27me3 do not induce up-regulation of satellite repeats in wild-type as well as mutant conditions.

(A, B, C, D) For each condition: Wild-type serum-ESC (A), Suv39hdn serum-ESC (B), DnmtTKO 2i-ESC (C) and DnmtTKO serum-ESC (D). Left part: Western-blot analysis for quantification of bulk levels of the HMT enzyme EZH2 and the repressive histone modification H3K9me3 and H3K27me3 related to H3 total. Right part: Relative expression (CNRQ) of major and minor satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes after treated with DMSO (Control) or EPZ (EZH2 inhibition). Each point is an independent biological replicate. (E) Quantification by Western-blot analysis of NANOG bulk levels related to H3 total after DMSO (Control) and EPZ treatment in serum-ESC (Wild-type and Suv39hdn) and in DnmtTKO (2i- and serum-ESC). Serum-ESC (WT): WT01.

In conclusion, we showed that the interplay between different repressive marks at PCH in particular is modulated according to the pluripotency states and culture conditions (Figure 69). We showed that serum-ESCs present a repressive state at the centromeric and pericentromeric regions with high level of H3K9me3, 5-meC and compacted chromocenters. Paradoxically in these cells transcription of major and minor satellites is variable but globally

elevated (Figure 69). However in EpiSCs such chromatin state correlate with low levels of both satellites transcripts (Figure 69). On the other hand, 2i-ESCs present a strong enrichment in H3K27me3 compared to H3K9me3, reduced levels of 5-meC and deconstructed chromocenters but unexpectedly transcription of major and minor satellites is low (Figure 69). Reduced major satellite transcripts in 2i-ESCs and EpiSCs correlate with reduced levels of NANOG compared to serum-ESCs (Figure 69). Interestingly we noticed that ESCs in 2i condition recapitulated the phenotype of *Suv39hdn*-ESCs at least concerning major satellite: reduced H3K9me3, reduced 5-meC and enrichment in H3K27me (Figure 69). However *Suv39hdn* mutant condition revealed the strict dependence of EpiSCs to H3K9me3 in order to fully repress major satellites transcription (Figure 69). The complete absence of 5-meC is not compatible with viability of primed pluripotent stem cells. Conversely this condition in naive ESCs induces an epigenetic switch with an enrichment in H3K27me3 at the expenses of H3K9me3 at major satellites repeats but strikingly does not affect the transcription activity at these sequences (Figure 69). Finally using an EZH2 inhibitor we showed that, in wild-type as well as mutant conditions, H3K27me3 seems not to play a key role in repression of major and minor satellites transcription.

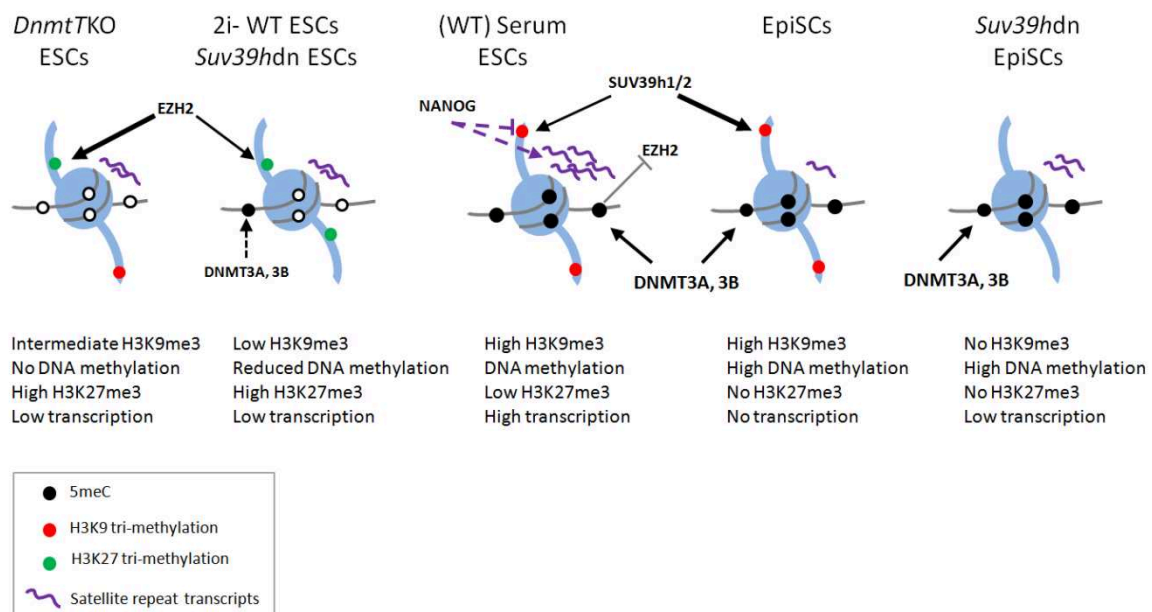


Figure 69 : Model of the epigenetic organization at PCH in the different pluripotent cells and in mutants' conditions

Schematic drawings recapitulating the organization and transcription status at PCH based on our findings and published data (Cooper et al., 2014; Fuks et al., 2003; Lehnertz et al., 2003; Saksouk et al., 2014)

2 Conversion from naïve to primed pluripotency in mouse

ESCs can be converted in vitro into EpiSCs (Figure 70) in about two weeks mimicking the in vivo development of ICM cells into post-implantation epiblast that occurs in two days. Based on this fact we wondered about the in vitro kinetics of the transcriptional changes in the conversion process. Another unsolved question was also whether the kinetic of conversion into primed EpiSCs state was different starting from 2i-ESCs or serum-ESCs. In particular we focused our interest on genes connected to naïve and primed pluripotency, genes coding for the epigenetic modifiers and satellites non-coding RNAs.

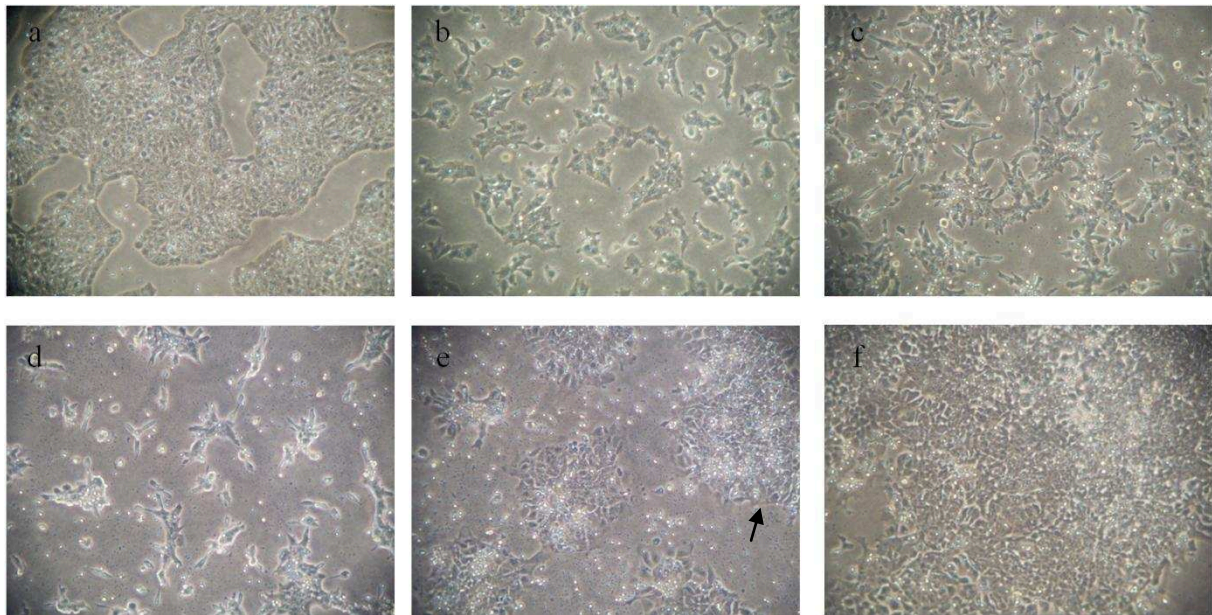


Figure 70: Morphological changes of ESCs during conversion to EpiSCs under the phase contrast microscope (Images from Tosolini and Jouneau, 2016).

(a) ESCs in CDM/2i/LIF. (b) ESCs at day 1 of conversion. (c) Cells at day 2 of conversion. (d) Cells at day 3 of conversion (high mortality rate). (e) Cells at day 5 of conversion: first appearance of EpiSC-like colonies before the first collagenase treatment (arrow). (f) Cells at day 7 of conversion (after the first passage): real cEpiSCs colonies. ESC: R1.

On a morphological point of view, 5 to 7 days are needed to obtain flat EpiSC colonies, but two weeks are usually necessary to stabilize the phenotype and reduce differentiation. The laboratory has previously shown that by day 3 considerable changes have already occurred at the transcriptional level (Veillard et al., 2014). To get more insight on this process we decided to focus on the first three days of conversion in order to deeply investigate rapid transcriptional changes.

We performed three independent conversions starting either from 2i-ESCs or serum-ESCs and took samples at day 1, 2 and 3 and then after one (day 7) and two weeks (established cEpiSCs), as well as from the starting ESCs population (day 0). We first performed qRT-PCR on these samples for naïve and primed pluripotency genes.

Interestingly we observed that during conversion from 2i-ESCs the majority of naïve pluripotency genes *Nanog*, *Klf4*, *Prdm14* and *Esrrb* were already strongly down-regulated at day 1, close to their level found at day 3 (Figure 71). During conversion from serum-ESCs the transcriptional down-regulation was more gradual for the same genes compared to 2i-ESC (Figure 71). *Sox2* expression seems to be gradually down-regulated during both types of conversion, while *Pou5f1* (*Oct4*) expression was only slightly reduced in cEpiSC, confirming that *Oct4* is present even in primed pluripotency (Figure 71).

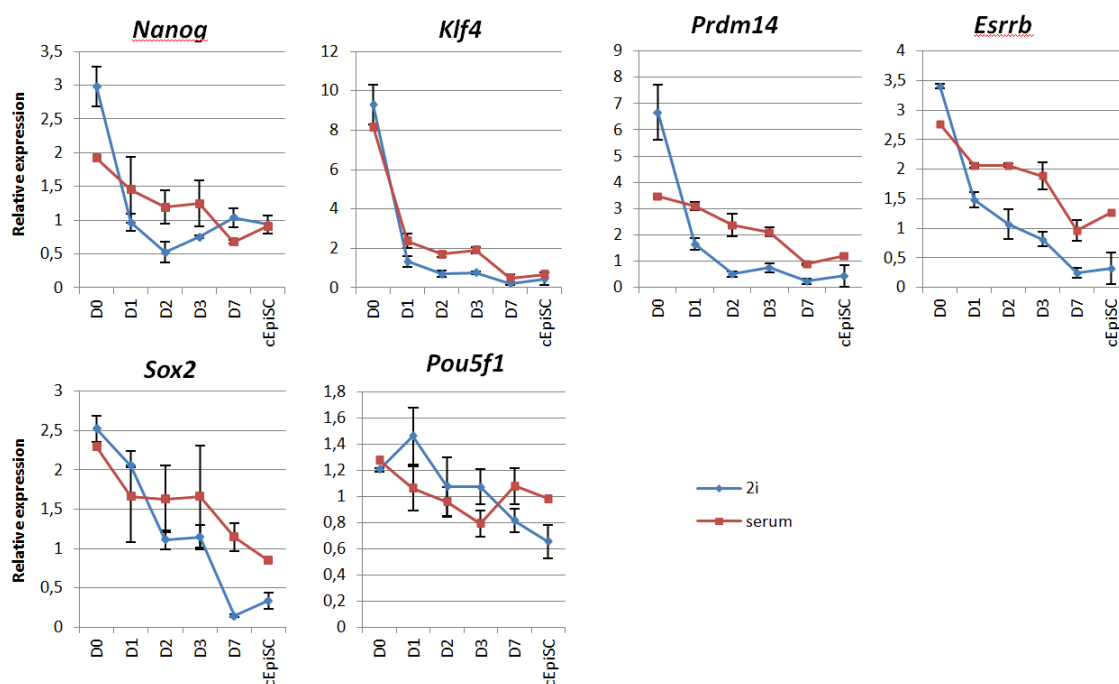


Figure 71: Kinetics of transcription of pluripotency genes during conversion of ESCs into cEpiSCs.

Relative expression (CNRQ) of *Nanog*, *Klf4*, *Prdm14*, *Esrrb*, *Sox2* and *Pou5f1* (*Oct4*) transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Blue line corresponds to the kinetics of 2i-ESC conversion, and the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). D0: R1.

The primed pluripotent specific genes showed different kinetics in between them. *Otx2* showed a burst of transcription between day 2 and day 3 of conversion, higher in 2i-ESCs compared to serum-ESCs, and it is then slightly down-regulated in cEpiSCs (Figure 72). *Fgf5*

showed a slow and gradual up-regulation along the conversion but its transcription is induced more rapidly when starting from 2i-ESCs than serum-ESCs (Figure 72). Finally Brachyury (also known as T), being an early differentiation marker, increases drastically only later on during conversion, after day 3 or day 7 starting from 2i-ESCs or serum-ESCs, respectively (Figure 72).

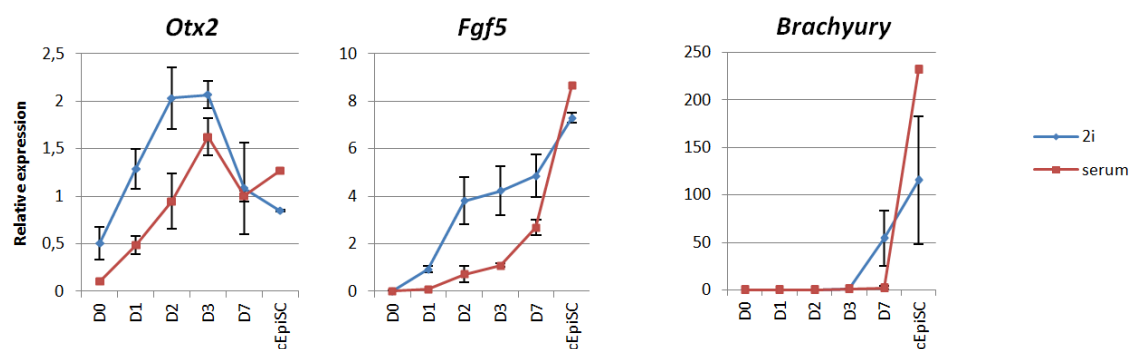


Figure 72: Kinetics of transcription of primed-associate genes during conversion of ESCs into cEpiSCs.

Relative expression (CNRQ) of *Otx2*, *Fgf5* and Brachyury (T) transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Blue line corresponds to the kinetics of 2i-ESC conversion, and the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). D0: R1.

As we have shown that DNA methylation machinery is more expressed in EpiSCs than ESCs, we examined the dynamics of the de novo Dnmts expression during conversion. Using isoform-specific primers we showed by qRT-PCR that the somatic long form of *Dnmt3a* (called here *Dnmt3a1*) is gradually up-regulated during conversion from ESCs to EpiSCs (Figure 73A). *Dnmt3a* (both isoforms) and *Dnmt3b* showed interestingly the same pattern of expression during conversion. Starting from 2i-ESCs *Dnmt3a* and *Dnmt3b* both showed a transient burst of transcription with a peak at day 2, while from serum the up-regulation was more gradual (Figure 73A). Moreover using western-blot, we confirmed that the protein level of DNMT3B parallels the dynamics of its transcripts with a peak at day 2 (Figure 73B). After this burst DNMT3B is gradually down-regulated but remains higher in cEpiSCs than in 2i-ESCs (day 0).

We then examined the kinetic of transcription of genes coding for the DNA demethylation enzymes. Interestingly, *Tet1* expression remains stable up to day 3, especially in serum-ESC conversion, while *Tet2* was rapidly down-regulated already at day 1 (Figure 73C). Finally, *Tet3* was not analyzed as it is expressed only in oocyte and zygote (Iqbal et al., 2011).

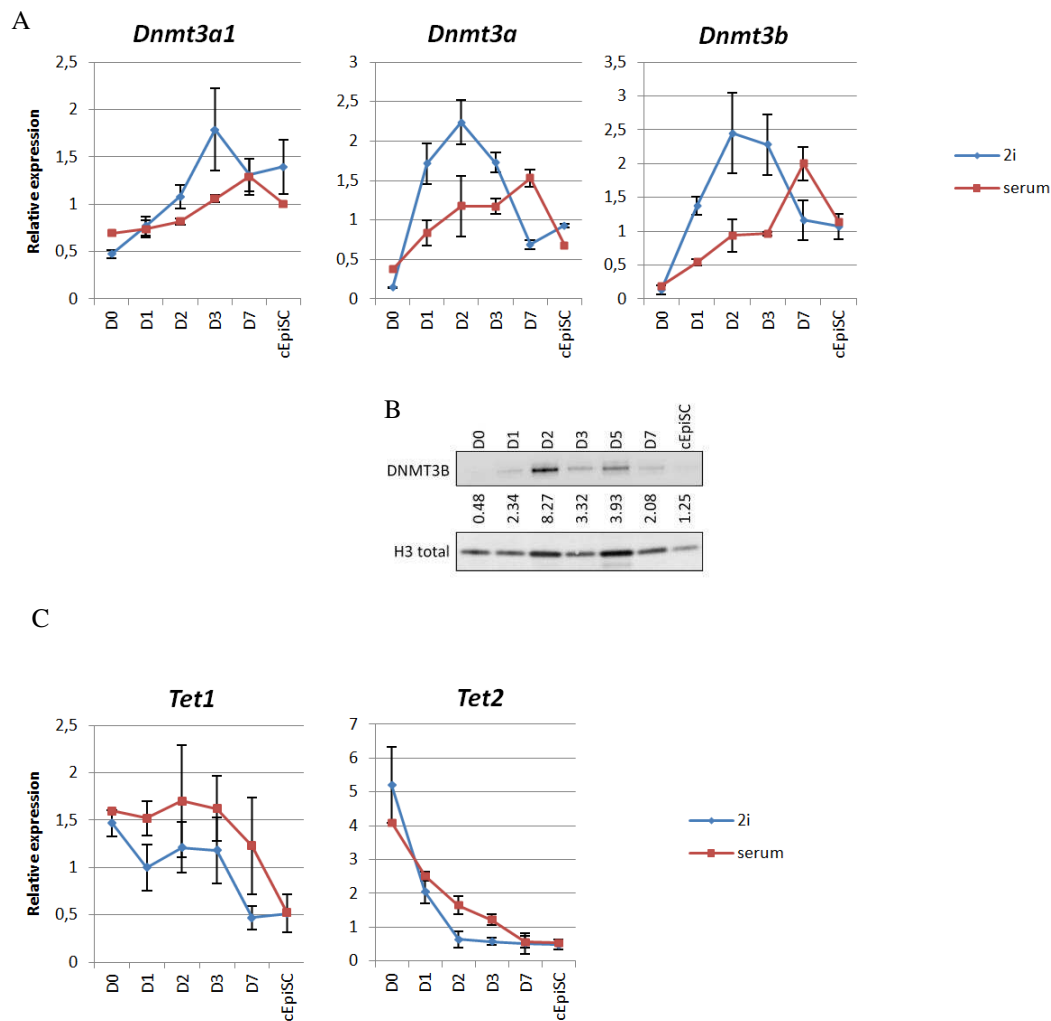


Figure 73: Kinetics of transcription of DNA methylation and demethylation machinery genes during conversion of ESCs into cEpiSCs.

(A) Relative expression (CNRQ) of de novo Dnmts transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Blue line corresponds to the kinetics of 2i-ESC conversion, while the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). (B) Western-blot analysis for quantification of bulk levels of DNMT3B related to H3 total during conversion from 2i-ESCs. (C) Relative expression (CNRQ) of *Tet*s transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Blue line corresponds to the kinetics of 2i-ESC conversion, while the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). D0: R1.

We have shown that the heterochromatin compartment is highly remodeled between ESCs and EpiSCs. We therefore decided to analyze also the kinetics of expression of genes related to constitutive and facultative heterochromatin during conversion.

The genes coding for H3K9me3 KMTs specific for constitutive heterochromatin (*Suv39h1/2*) show progressive up-regulation during conversion of both 2i-ESCs and serum-ESCs and in particular the increase in *Suv39h1* transcription is stronger than the one of *Suv39h2* (Figure 74A).

However transcription of genes coding for Polycomb repressive complex subunits (PRCs) responsible for the facultative heterochromatin compartment does not show such up-regulation. Regarding the PRC2, transcription of the catalytic subunit Ezh2 is stable through the conversion while one of the cofactor subunit Eed is only slightly down-regulated (Figure 74B). Furthermore we assessed also EZH2 protein levels during conversion by western-blot and interestingly transcript and protein do not follow the same kinetic. EZH2 protein level remains stable up to day 2 and then it is down-regulated. EZH2 remains low in cEpiSCs, as we already showed for EpiSCs (Figure 74C). Concerning PRC1, the Rnf2 gene coding for RING1B is stably expressed during conversion from serum-ESCs, while it seems to be down-regulated starting from 2i-ESCs however the variability is so huge that we cannot really conclude (Figure 74D). Cbx7, a gene coding for a cofactor subunit of PRC1 is slightly upregulated during the first three days of conversion while coming back to ESCs level (day 0) in cEpiSC (Figure 7D).

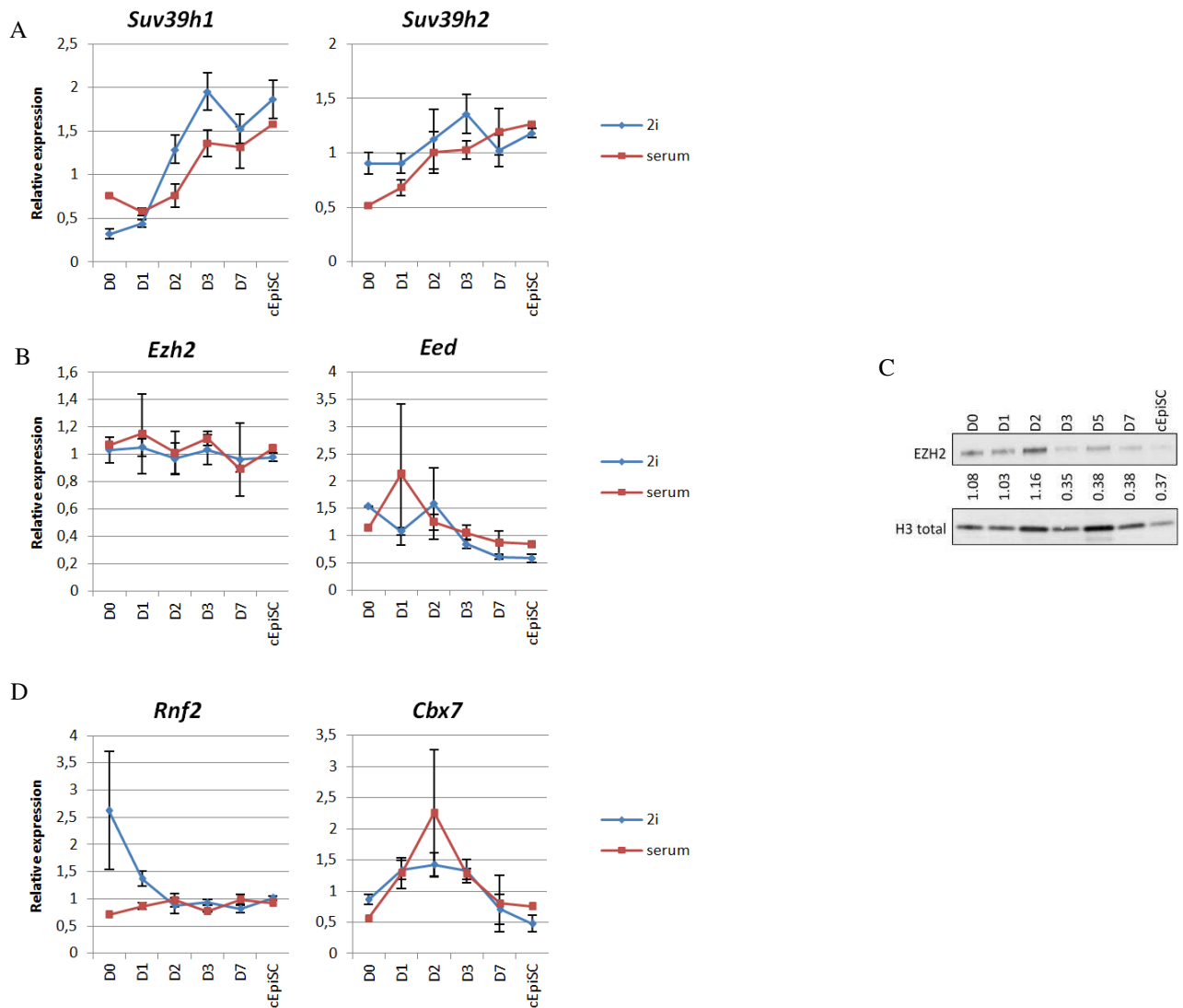


Figure 74: Kinetics of transcription of epigenetic modifiers genes during conversion of ESCs into cEpiSCs. (A, B) Relative expression (CNRQ) of (A) PCH specific H3K9 KMTs Suv39h, (B) PRC2 subunits Ezh2 and Eed, transcripts by qRT-PCR analysis normalized to Sdha and Pbgd housekeeping genes. Blue line corresponds 2i-ESC conversion, and the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). (C) Western-blot analysis for quantification of bulk levels of EZH2 related to H3 total during conversion from 2i-ESCs. (D) Relative expression (CNRQ) of PRC1 subunits Rnf2 (Ring1b) and Cbx7 transcripts by qRT-PCR analysis normalized to Sdha and Pbgd housekeeping genes. Blue line corresponds 2i-ESCs conversion, and the red one to serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). D0: R1.

As we showed that 2i-ESCs present a decompacted organization of major satellites and that during embryo development a transcriptional burst of major satellites is necessary to induce the chromocenter formation and compaction (Probst et al., 2010), we hypothesized that the same phenomenon should happen during conversion especially when starting from 2i-ESCs. Indeed, we have shown that major satellites sequences were highly relaxed in 2i-ESCs and

became clustered into somatic-like chromocenter in EpiSCs. To test this hypothesis we performed qRT-PCR for major and minor satellites transcriptional kinetics during conversion to cEpiSCs starting from 2i-ESCs as well as serum-ESCs.

Our data show that during conversion of 2i-ESCs to EpiSCs non-coding major satellite transcripts which are already lowly expressed, are even further down-regulated up to the cEpiSC stage. Minor satellite transcripts are slightly up-regulated during the first three days of conversion but came back in cEpiSCs to day 0 low levels (Figure 75). Conversion from serum-ESCs show a strong decrease of the high transcriptional levels of both major and minor non-coding satellites transcript as early as day 1 (Figure 75).

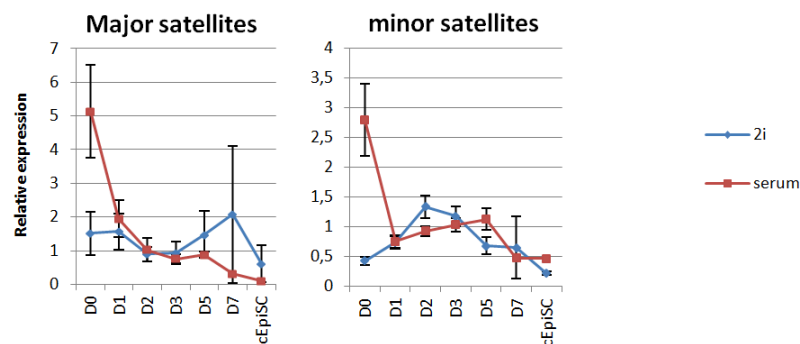


Figure 75: Kinetics of transcription of satellite repeats non-coding RNAs during conversion of ESCs into cEpiSCs.

Relative expression (CNRQ) of major and minor satellite transcripts by qRT-PCR analysis normalized to *Sdh*a and *Pb*gd housekeeping genes. Blue line corresponds to the kinetics of 2i-ESCs conversion, and the red one of serum-ESCs. Each point corresponds to the mean of the three independent conversions and the error bar indicates the Standard Error Mean (SEM). D0: R1.

Altogether these results show that even though a week is necessary to obtain morphologically EpiSC-like colonies, at the transcriptional level, naïve pluripotency genes are mainly down-regulated at day 1 of conversion. Meanwhile primed-associate genes were up-regulated with different kinetics: within the first days (*Otx2*) or gradually all along the conversion (*Fgf5*). A lineage associated gene such as *Brachyury* is up-regulated only later. Interestingly starting from 2i-ESCs the down-regulation of naïve pluripotent genes was more rapid and important than from serum-ESCs. In addition the up-regulation of primed pluripotency-associated genes was also earlier and stronger starting from 2i-ESCs compared to serum-ESCs.

Concerning DNA methylation machinery gene, they are up-regulated during conversion, and especially in 2i-ESC conversion, show a transient burst of expression at day 2. Such increase in expression during conversion is reminiscent of the DNA methylation wave that occurs

concomitant to implantation in embryo development. Conversely genes coding for enzymes of the DNA demethylation machinery are down-regulated either late (Tet1) or early (Tet2).

Globally the genes governing the machinery for H3K9 trimethylation at constitutive heterochromatin are clearly up-regulated through the conversion, confirming the idea that EpiSCs present a more compact and close chromatin than ESCs. Interestingly genes coding for PRCs subunits do not seem to be strongly affected during the process of priming from naïve pluripotency.

Finally major satellite non-coding transcripts are not up-regulated during conversion from 2i-ESC to cEpiSCs and minor satellites showed only a slightly up-regulation, suggesting that in mouse ESCs there is no need to increase the transcription of major satellite sequences to induce the compaction of heterochromatin as in early stage of embryonic development (Probst et al., 2010). Interestingly the characteristic high levels of major and minor satellites non-coding transcripts of serum-ESCs are rapidly and strongly down-regulated, so at day 1 of conversion they have reached the same low levels as found in 2i-ESCs.

DISCUSSION

DISCUSSION

1 Uncoupling epigenetic state of naïve ESCs with transcription regulation of major and minor satellites

We showed that serum-ESCs present a repressive state at the constitutive heterochromatin with local accumulations of H3K9me3, HP1 β , 5-meC and compacted chromocenters. Paradoxically transcription of major and minor satellites is elevated. On the other hand 2i-ESCs present at the same regions a strong enrichment in H3K27me3 compared to H3K9me3, reduced levels of 5-meC and deconstructed chromocenters. Unexpectedly transcription at major and minor satellites is low.

In a somatic environment the presence of repressive epigenetics marks such as H3K9me3, H3K27me3 and DNA methylation leads to the formation of a compacted domain. Such condition is normally inaccessible to transcription factors and thereby transcriptionally inactive (Reviewed in: Tessarz and Kouzarides, 2014).

Therefore, such uncoupling of the epigenetic state of centromeric and pericentromeric heterochromatin (CH/PCH) and its transcriptional status leaves open the nature of the driver of satellite transcription in serum-ESCs and its repression in 2i-ESCs. First, we can speculate that satellites transcription is a consequence of the in vitro culture condition. In particular it could be due to unknown factors present in the serum and not found in the chemically defined medium used in 2i condition. Alternatively, inhibition of MAPK and GSK3 pathways may induce the repression of satellite transcription. Very preliminary data of the laboratory show that the two inhibitors (PD and CH of 2i medium) have a synergic negative effect on satellite transcription when added to ESCs cultured in serum condition. It would be very interesting to perform the reverse experiment, adding different percentages of serum to ESCs cultured in 2i medium and evaluating if there is up-regulation of satellite repeats ncRNAs. Moreover both major and minor satellites are strongly down-regulated already at day 1 of conversion from serum-ESCs and, also in this case, such decrease can be due either to the drastic withdrawal of serum and switch to chemically define medium (CDM) or to the presence of Activin A and FGF2 that may activate pathways repressing rapidly satellites transcription.

Saksouk and colleagues (2014) used the PiCh (Proteomics of isolated Chromatin fragment) method to assess which are the proteins associated to major satellites (apart from histones) in mouse ESCs (in serum conditions). In addition to SUV39Hs and DNMTs enzymes, they

found several factors dedicated to five main pathways: Chromatin organization, DNA replication and cell cycle, RNA processing, DNA damage and Lamina. So pericentromeric heterochromatin seems to be a complex system of proteins and between them transcription factors play a role in organization and regulation of these regions. We thus wonder whether transcription factors can explain the transcriptional state of these sequences.

Very recently, Novo and co-workers (2016) have shown that NANOG is recruited to heterochromatin via SALL1, favors transcription of major satellites and relaxes structure of the heterochromatin in serum-ESCs. Indeed, they observed that down-regulation of Nanog (in serum-ESCs Nanog^{+/+} and ESCs Nanog^{-/-}) induces compaction of heterochromatin, increase of H3K9me3 level at pericentromeres and decrease of major satellite transcription. Strikingly we have observed that although NANOG is expressed homogeneously within the population of ESCs cultured in 2i, the global protein level is reduced compared to ESCs in serum medium (Figure 59). This, together with a more diffuse pattern of H3K9me3-HP1 β , correlates with a reduced expression of major and minor transcripts in 2i-ESCs. But conversely, we have also observed that 2i-ESCs have a disrupted organization of constitutive heterochromatin with major satellites being decompacted and relaxed compared to serum-ESCs, which we would have expected to be a chromatin state more prone to active transcription. We speculate that 2i condition induces a complex phenotype on heterochromatin of naïve ESCs that could not solely be explained by reduced NANOG levels.

NANOG is not the only factor that is found to bind to major and minor satellites, indeed PAX3 and PAX9 also bind to major satellites having a synergic functions to safeguard silencing of these sequences at least in MEFs (Bulut-Karslioglu et al., 2012). The way of action of these factors is not known but probably they can compete for the binding of the same region with other proteins or they can mask some binding motif, impeding the recruitment of the PolII complex. However Pax3 and Pax9, according to RNA-seq data of the laboratory, are lowly expressed in both naïve ESCs (2i or serum condition) and primed EpiSCs where satellite repeats seem to be completely silenced, so the differences of expression that we have observed cannot be explained with these factors. In addition another transcription factor SNAIL1 has been found to associate with pericentromeric region and regulate transcription of major satellite in MEFs and during Epithelial-to-Mesenchymal transition (EMT) (Millanes-Romero et al., 2013). Absence of Snail1 in MEFs leads to up-regulation of major satellite transcripts while a peak of SNAIL1 during EMT transiently inhibits transcription from pericentromeric regions (Millanes-Romero et al., 2013). As SNAIL1 is more expressed in 2i-ESCs (thanks to GSK-3 inhibition) compared to serum-ESCs

(Lin et al., 2014), we hypothesize that major satellites are more expressed in serum condition due to a reduced presence of SNAIL1 at these sequences.

Moreover, an RNA component seems to be involved in the highly structured three-dimensional chromatin at pericentromeric regions as RNase treatment on living cells lead to the disappearance of HP1-enrichment foci (Maison et al., 2002; Probst et al., 2010). Whether these RNAs come directly from satellite transcription is not known so far. On the contrary in yeast the link between satellites transcription and constitutive heterochromatin organization is well established, and it involves the RNA interference pathway that in turns is necessary for H3K9me3 deposition (Saksouk et al., 2015). In mammals there are only little (Kanellopoulou et al., 2005) or no evidence (Murchison et al., 2005) for RNAi machinery involvement in heterochromatin formation and maintenance (reviewed in: Biscotti et al., 2015; Saksouk et al., 2015).

Our study aimed to answer whether the epigenetic state of satellite repeats in naïve ESCs influences transcriptional activity of these sequences but it would be interesting as well to ask the opposite question: would satellite non-coding transcripts influence the epigenetic state of centromere and pericentromere in pluripotency? The cellular model we describe in this thesis seems to be perfect to answer such question as we have different epigenetic and transcription states for these regions. It would be very interesting to study whether up-regulation of major and minor satellite ncRNAs in 2i-ESCs and EpiSCs will have different consequences on these two cell types. In particular with such experiments we could assess the role of satellite transcripts on histone marks and DNA methylation deposition, chromocenter status, genome and chromosome stability as well as self-renewal property. Moreover the down-regulation of satellite ncRNAs in serum-ESCs would also help to infer about the function of these transcripts on the epigenetic state as well as on the pluripotent state.

2 The problem of major satellite repeats quantification

Major satellite transcription was found to be either slightly up-regulated (Lehnertz et al., 2003) or clearly up-regulated (Martens et al., 2005) in Suv39hdn-ESCs compared to wild-type. However our data do not confirm these previous results as we found major satellite transcripts to be slightly down-regulated in mutant serum-ESCs compared to wild-type. The first hypothesis to explain this difference is a technical one as we used the same primers of Lehnertz et al., and Martens et al., but while they assessed the transcription by semi-quantitative PCR, we used qPCR (Syber green technology). However this discrepancy cannot

influence only some types of samples but all the samples and with the same bias. The major consideration that we propose to explain such divergence is the effect of time of culture and confluence of cells. Indeed the culture effect has already been shown for minor satellites (Bouzinba-Segard et al., 2006), while major satellites were so far only linked to cell-cycle (Lu et al., 2008). The cell-cycle of cells in culture is likely to be influenced by the rhythm of trypsinization, re-plating and confluence. Moreover preliminary results of the laboratory suggest that density/confluence of ESCs is inversely correlated with the transcription of major satellites in serum-condition. It would be very interesting to study also the effect of time of culture on major satellite transcription to see whether there is an accumulation effect. Since we detected such differences we tried to standardize as much as we could our culture conditions of serum-ESCs, passaging the different cell lines (Wild-type and mutant) the same day with the same ratio or cell number to avoid any bias or at least to introduce the same bias in both conditions.

3 Cross-talk between H3K27me3 and 5-meC in mouse pluripotency

The chromatin of ESCs has long been considered as very plastic and dynamic compared to that in somatic cells (Ahmed et al., 2010; Koh et al., 2010; Meissner, 2010; Meshorer et al., 2006). One of the features of this plastic chromatin is the cross-talk between DNA methylation and repressive histone modifications H3K27me3 in particular. Indeed we observed that a reduction of 5-meC is replaced to some extent by H3K27me3 in ESCs in three different contexts. First after 2i adaptation of ESCs, which leads to a global 5-meC depletion (Habibi et al., 2013), we observe H3K27me3-foci enrichment at the expense of H3K9me3 foci. Second the Suv39hdn condition induces, even in hypermethylated serum-ESCs, a partial demethylation of major satellites and also an H3K27me3-foci enrichment, as previously observed in several studies (Cooper et al., 2014; Martens et al., 2005; Peters et al., 2003; Saksouk et al., 2014). Finally in the total absence of 5-meC (DnmtTKO), serum-ESCs present strong enrichment in H3K27me3 foci compared to wild-type cells (also shown in: Saksouk et al., 2014, Cooper et al., 2014). Along with these findings it was shown that DNA methylation globally antagonizes H3K27me3 deposition, indeed in DnmtTKO or 5-azacytidine treated ESCs, there is a redistribution of this mark, being now enriched in demethylated CpG sites, at the expense of the canonical PRC2 target sites (Cooper et al., 2014; Hagarman et al., 2013). However it is interesting to notice that while H3K27me3 replaces DNA methylation when this mark is abrogated, in the inverse situation when there is no H3K27me3, DNA methylation

does not compensate its absence, meaning that these two repressive marks are not interchangeable (Hagarman et al., 2013). In naïve ESCs in 2i condition, it has also been shown that many bivalent promoters (characterized by the co-occurrence of H3K4me3 and H3K27me3) lose the trimethylation on lysine 27 of histone H3 (Marks et al., 2012). Therefore, while this mark is reduced at unique sequences, we now show that it is redeployed at PCH in 2i-ESCs, making these cells an interesting, non-pathological model of the cross-talk between PRC2 regulation of gene expression and PHC epigenetic state (Déjardin, 2015).

An additional evidence of this H3K27me3-DNA methylation cross-talk has been made by Walter and colleagues (2016) in ESCs adapted from serum to 2i medium also containing VitamineC. This switch induces on various transposons families a loss of 5-meC, a reduction in H3K9me2 levels and an increment in H3K27me3 levels, while H3K9me3 stays globally constant (Walter et al., 2016). On an epigenetic point of view 2i condition of ESCs seems to convey the same status at transposons and pericentromeric satellites. However, at the transcriptional level transposons are mainly repressed in both serum- and 2i-ESCs conditions, while satellite repeats are in a repressed state only in 2i-ESCs. This means that the same epigenetic context even in the same cell type has different consequences depending on the nature of the sequences.

4 Hypothetical SUV39H-independent H3K9me3 deposition at PCH in 2i-ESCs

In 2i-ESCs, H3K9me3 was enriched in small foci present in part of the ESC population. These foci do not always correspond to DAPI-dense regions and remain remarkably similar in the absence of both Suv39h enzymes. In addition, SUV39H1 was not detectable after immunostaining of 2i-ESCs while accumulating at DAPI-dense foci in serum-ESCs and even more in EpiSCs. We speculate that these H3K9me3 small foci may depend on other HMT enzymes, such as SETDB1. Indeed it is known that in absence of SETDB1 there is a reduction of H3K9me3 at the euchromatin but also at PCH, meaning that SETDB1 is likely responsible for a portion of H3K9me3 even at PCH where SUV39H1/2 play the major HMT role (Mozzetta et al., 2015; Schultz et al., 2002).

To test such hypothesis we could simply adapt Setdb1^{-/-} ESCs in 2i condition and evaluate the H3K9me3 foci-enrichment compared to wild-type. A completely loss of these foci in mutated 2i-ESCs compared to wild type will prove the action of SETDB1 as KMT at PCH.

5 EpiSC: a pluripotent cell with a somatic epigenetic state

We showed that EpiSCs are closer to somatic cells than ESCs as they present high levels of H3K9me3 and 5-meC, enrichment of CAF-1, SUV39H1 and HP1 β at DAPI-dense foci, and compacted chromocenters. Conversely to serum-ESC, the repressive epigenetic state of EpiSCs at major and minor satellites correlates with a transcriptionally inactive compartment. It is possible that the degree of H3K9me3 enrichment and DNA methylation in serum-ESCs is still below the threshold that would prevent accessibility of chromatin for transcription, as it does in EpiSCs. But we cannot exclude the importance of other factors that are also reduced in EpiSCs compared to serum-ESCs (such as NANOG). Novo et al (2016) indeed have demonstrated that the over-expression of NANOG in EpiSCs induces the up-regulation of major satellite transcripts and decompaction of H3K9me3 foci. From our findings it appears clearly that albeit pluripotent, EpiSCs present an epigenetic state closer to somatic cells than naïve ESCs. Indeed the absence of Suv39h1/2 in primed EpiSC induces neither DNA-demethylation nor H3K27me3-enrichment at major satellite loci, while this happens in naïve ESCs. Our results show that in primed EpiSCs, conversely to naïve ESCs, DNA methylation at major satellite is independent of the pathway SUV39Hs-H3K9me3-HP1 (Schotta et al., 2004) thus impeding to H3K27me3 to spread in Suv39hdn-EpiSCs. However the absence of H3K9me3 at PCH induces de-repression and up-regulation of major satellites transcripts levels in EpiSCs, thus suggesting that DNA methylation by itself is not sufficient to get a complete repression of these sequences. Hence, in a hypermethylated condition, H3K9me3 plays a key role in silencing major satellites at least in primed EpiSCs. Note that DNA methylation is strictly necessary in EpiSCs as our attempts to convert DnmtTKO naïve ESCs into primed EpiSCs have been all unsuccessful, likely due to abundant apoptosis. When we tried to treat EpiSCs with an inhibitor of DNMTs (5-aza-cytidine), even at low concentration that has no effect on ESCs, we induced strong apoptosis and EpiSCs colony morphology was drastically impaired (unpublished data of the laboratory).

In addition a further evidence of the loss of a plastic and open chromatin in EpiSCs compared to ESCs is brought by the fact that EpiSCs have globally reduced their levels of histone acetylation, at least in terms of H3K9ac. It would be interestingly to investigate whether the reduced H3K9ac level in EpiSCs compared to ESCs is due to increased expression and/or activity of HDACs and/or reduced for HATs. Moreover western-blot analysis of another histone acetylation such as H3K14ac would enlighten the deacetylation process in primed EpiSCs and whether it is specific to H3K9ac or more general.

Another suggestion concerning the less plastic chromatin of EpiSCs is revealed by the reduction of bivalency that we demonstrated by PLA technology when compared to serum-ESCs. To strengthen our result we will use already published datasets of H3K4me3 and H3K27me3 ChIP-seq performed on EpiSCs by other groups. We will analyze if the promoters of genes that were bivalent in serum-ESCs (Marks et al., 2012) are still enriched in both histone marks or resolved in either active or repressive state, as no one have never look at them in primed mouse pluripotency (EpiSCs).

6 Major and minor satellites sequences respond to different epigenetic pathways

Even if pericentromeric and centromeric regions are located close to one another in each chromosome, they seem to be regulated differently and/or independently. Indeed we showed that in serum-ESCs major satellite sequences are as methylated as in EpiSCs and MEFs. Minor satellite sequences display an intermediate level of methylation between the hypomethylated 2i-ESCs and the fully methylated EpiSCs and MEFs. It has been previously shown that major and minor satellites sequences are methylated by different pathways: minor satellites sequences are mainly methylated by DNMT3B, while on major satellites sequences both DNMT3A and DNMTB are active and compensate each other. Only the double knock-out of these enzymes results in a hypomethylation of major satellite sequences (Okano et al., 1999).

In addition in Suv39hdn serum-ESCs major but not minor satellite sequences become demethylated at the DNA level and enriched for H3K27me3 (Lehnertz et al., 2003; Martens et al., 2005; Saksouk et al., 2014). These data suggest that DNMTs are recruited on minor satellite sequences via an independent pathway of SUV39H1 (Fuks et al., 2003). Indeed the association of SUV39H1 and HP1 with a DNMT has been observed only for DNMT3A and DNMT1 (Fuks et al., 2003) and not for DNMT3B which is the unique enzyme responsible for DNA methylation on minor satellite sequences (Okano et al., 1999). Thus it seems that DNA methylation on minor satellite sequences is completely independent of H3K9me3-associated pathway in ESCs as well as in EpiSCs. On the other hand major satellites DNA methylation strictly depends on H3K9me3 in ESCs while becoming independent on this pathway in EpiSCs. This result correlates with the fact that DNMT3B is enriched in foci in naïve ESCs

while in EpiSCs it is completely diffuse into the nucleoplasm suggesting no specific binding to PCH (Figure 55B).

Moreover in Suv39h^{dn} cEpiSCs a strong up-regulation of major satellites transcripts levels was revealed while minor satellites sequences did not show such clear effect, suggesting once more the independent regulation of major and minor satellites. Third in complete absence of DNA methylation only minor satellites and not major satellites sequences showed a strong reduction in their transcriptional levels in serum-ESCs.

All these differences could be related to many properties: first, the DNA sequences of the two satellites are different, in terms of nucleotides (different content of CpG for example) and size (234bp for major while 123bp for minor). Then the numbers of repetition within the regions are not comparable: 200,000 major satellites and “only” 50,000 minor satellites (Martens et al., 2005). Finally while the structural function of major satellite repeats is still unknown, minor satellite repeats play a key role in the segregation of chromosomes during mitosis.

The implications of minor satellite region and also the transcripts arising from these sequences on genome stability are clearly established molecularly (Bouzinba-Segard et al., 2006; Ferri et al., 2009). Some data also suggest that major satellite sequences seem to be also important for genome organization and genome stability but the mechanisms involved in this process are mainly still unknown. Firstly, the chromocenter organization of PCH during embryo development need a burst of major satellite transcript to occur and when this up-regulation is impaired the embryos do not develop further (Probst et al., 2010). However this burst of major satellite transcription seems not to be a general mechanism necessary to compact chromocenter as in our system when we convert 2i-ESCs (decondensed pericentromeres) into EpiSCs (compact chromocenters) no such transient up-regulation of major satellite ncRNAs is detected. Even though we could not exclude that the major satellites up-regulation is maybe too rapid compared to our window of time-points. Secondly, major satellites has been suggested to be a specific loading site for cohesion on heterochromatin via a pathway governed by SUV4-20H2 and when this protein is absent cohesion recruitment is reduced in MEFs leading to mitotic defects and genome instability (Hahn et al., 2013). Moreover the Suv39h^{dn} condition in MEFs induces chromosomal abnormalities but whether this problem is due to the absence of H3K9me3 on minor or major satellites or both is still to be investigated.

7 Genome stability in ESCs

Knowing that perturbing the epigenetic/transcriptional state or three dimensional structure of CH/PCH in differentiated cells lead to a genomic instability (Ferri et al., 2009; Peters et al., 2001), we wondered why this is not the case for ESCs. We have shown that 2i-ESCs present decompacted major satellite domains, likely with a loss of H3K9me3. This is unexpectedly not correlated with a de-repression of transcription from these regions and we did not observed chromosomal instability at least in terms of polyploidy, presence of chromosomal bridges or aberrant mitosis. It is already known that Suv39hdn condition in mice induces chromosomal instability and perturbed chromosome segregation, indeed MEFs are more prone to aneuploidy (Peters et al., 2001). However Suv39hdn serum-ESCs present a normal karyotype (García-Cao et al., 2004). Based on these results we were less surprised that 2i-ESCs did not show abnormalities as they recapitulated the epigenetic environment of Suv39hdn serum-ESCs. This fact strongly suggest that in naive pluripotency there are mechanisms, unknown for the moment, to protect cells from chromosomal instability that occurs in absence of H3K9me3 at CH/PCH (as in Suv39hdn or 2i conditions) or when these regions are highly transcribed (serum-ESCs). High transcription of minor satellites is associated with mitotic instability in differentiated cells (Bouzinba-Segard et al., 2006) but not in pluripotency as serum-ESCs transcribe at high levels minor satellite without inducing such instability.

As EpiSCs resemble less to ESCs than to differentiated cells concerning their epigenetic state, it would be interesting to test whether up-regulation of major and/or minor satellites sequences will induce some genome instability in these cells. For example we could propose to deeply study Suv39hdn EpiSCs, which lack H3K9me3 at PCH and over-express major satellite ncRNA, investigating their karyotype or the eventual presence of abnormal mitotic plates or chromosome bridges.

8 Does the epigenetic status of in vitro pluripotency reflect in vivo pluripotency of mouse embryo?

ESCs cultured in 2i condition are known to be closer to cells of the early epiblast in E4.5 blastocyst than serum-ESCs at least at the transcriptional level (Boroviak et al., 2014, 2015). Conversely on the epigenetic point of view not much is known for now. We showed that 2i-ESCs are characterized by DNA hypomethylation and strong H3K27me3-foci enrichment at the expenses of H3K9me3 at PCH. For the moment it is still under debate and further

investigations are needed to better characterize the H3K27me3 subnuclear pattern in mouse pre-implantation development. It seems that such H3K27me3-foci enrichment pattern is not found in the ICM but is more typical of earlier developmental stages from the paternal pronucleus of 1-cell till 16-cell and morula stage embryos (Unpublished data of the laboratory; Puschendorf et al., 2008; Santenard and Torres-Padilla, 2009). Moreover pericentromeric sequences are marked by H3K9me3 as early as the 1-cell stage (maternal pronucleus) and remains so until the blastocyst stage even in the ICM and in further developmental stages (Beaujean, 2014; Santos et al., 2005). We showed that serum-ESCs and EpiSCs are enriched in H3K9me3 at PCH as pre- and post-implantation embryos from which they are derived.

Concerning DNA methylation levels, *in vivo* and *in vitro* pluripotency seems to behave similarly because ICM of early blastocyst has low levels of 5-meC, like 2i-ESCs (Ficz et al., 2013; Habibi et al., 2013). Hypermethylated serum-ESCs are transcriptionally closer to blastocyst outgrowth (Huang et al., 2014). However they share some similarities also with the epiblast at E5.5 (Boroviak et al., 2014) that is also hypermethylated as after implantation there is a wave of DNA methylation. Finally EpiSCs are even more methylated than the epiblast at E6-7.5 (Borgel et al., 2010; Smith et al., 2012; Veillard et al., 2012).

We showed that a striking difference of 2i-ESCs compared to serum-ESCs and EpiSCs is the decompaction of major satellites with the disruption of the chromocenter structure. At the blastocyst stage, ICM cells showed already a compacted chromocenter as in serum-ESCs and EpiSCs. A similar decondensation of major satellites is found only at 1-cell stage of mouse embryo development up to the 2-cell stage when the chromocenter formation occurs (Aguirre-Lavin et al., 2012; Probst et al., 2010). Thus concerning PCH organization 2i-ESCs are similar to early 2-cell stage embryos and not ICM cells from where they are derived.

Transcription of major satellites was reported in embryos only from the late 1-cell stage to the late 2-cell stage, prior to chromocenter formation. It seems to be completely down-regulated already at the 8-cell stage and its level remains low up to blastocyst stage as in somatic cells (Probst et al., 2010).

In conclusion even if 2i-ESCs highly resemble pre-implantation epiblast at the transcriptomic and DNA methylation levels, concerning the epigenetic of the PCH they clearly diverged from the *in vivo* situation in terms of histone modification, chromocenter organization and satellite transcription resembling earliest embryo stages. Serum-ESCs on the other hand seems to diverge transcriptionally and at the DNA methylation level from ICM cells however sharing with their *in vivo* source histone modification pattern and chromocenter

compaction. Finally EpiSCs seems to parallel late post-implantation mouse embryo epigenetically as well as transcriptionally.

To really push the comparison further we need to investigate specifically the transcription of satellites repeats at the blastocyst stage, distinguishing ICM cells and trophoblast cells, but also in the next stages. Moreover immunostaining should be made at the stages E3.5, 4.5, 5.5 and 6.5 for H3K9me3 and H3K27me3 in order to compare specifically the subnuclear organization in foci-enrichment as we made for the different type of pluripotent stem cells.

9 In vitro conversion: why such an inefficient process compared to in vivo?

The conversion is the in vitro process that push naive ESCs to become primed EpiSCs, mimicking the ICM cells that develop into epiblast cells in vivo after implantation of the blastocyst. However the timings of in vitro and in vivo processes are not comparable. Indeed while in vivo ICM cells became epiblast in 48 hours, in vitro ESCs need quite two weeks to stabilize their primed state of EpiSCs. We tried to investigate it molecularly in particular from the first day of the conversion in order to find the possible causes of this delayed process. Previous studies from the laboratory have shown that at day 3 of conversion many genes have already adapted their transcriptional state to the EpiSCs one although the primed state was not yet fully established (Veillard et al., 2014). Our results showed that particularly naive pluripotency genes are already shut-down after 24 hours of culture in EpiSCs medium condition suggesting that the delay in reaching the primed state is likely not due to a retention or a delayed down-regulation of the naive transcriptional state. However the transcription factors typical of primed pluripotency are gradually up-regulated, starting from 24 hours. Early differentiation genes are present only from day 7 or even later. Taken together it could be that the delay in the obtention of stable EpiSCs is partially explained by a slow establishment of the primed pluripotent transcriptional circuitry rather than a retained expression of naive factors.

Based on the fact that EpiSCs seems to have an epigenetic state closer to differentiated cells than ESCs, we can argue that the building of the epigenetic barrier that distinguishes ESCs from EpiSCs is time-dependent. Indeed we found that genes coding for H3K9 KMTs such as Suv39h1 and Suv39h2 are gradually up-regulated during conversion which suggests a slow mechanism. However concerning DNA methylation we would have expected a delay in the establishment of the hypermethylated EpiSCs status when we are starting from 2i-ESCs

that are hypomethylated compared to serum-ESCs. But this seems not to be the case as 2i-ESCs showed a transcriptional burst, at the beginning of conversion (from day 1 to 3), of *de novo* Dnmt. Transcriptional burst was also translated into temporary increased protein level at least for DNMT3B (Figure 73A, B) that probably leads to a rapid DNA methylation, flattening the differences between 2i- and serum-ESCs. Indeed we did not find any delay in conversion of 2i-ESCs compared to serum-ESCs even though we showed that these two states have a completely different heterochromatin epigenetic state.

In conclusion a slow up-regulation of primed associated genes and specific heterochromatic KMTs, could probably explain, at least partially, the delay of the *in vitro* process of conversion of ESCs into EpiSCs compared to *in vivo* establishment of primed state. Many other genes and proteins should be analyzed to get a complete view in order to find other molecular candidates that can explain this delay.

One hypothesis is that the *in vitro* environment for EpiSCs is sufficient and enough efficient to sustain the primed state of mouse pluripotency but is much less efficient to induce the conversion from naive pluripotency. For sure the *in vivo* environment is much more specialized than the *in vitro* one with increased number of soluble factors and more importantly with the physical constraints that the implanting embryo needs and that are completely absent on a petri dish (Bedzhov and Zernicka-Goetz, 2014).

Performing kinetics studies during conversion, we have discovered the discordance between mRNA *Ezh2* and protein levels (Figure 74B, C). In the conversion from either 2i- or serum-ESCs, while transcripts levels stays globally constant, the EZH2 protein level drops down at day 3 and remains low in cEpiSCs. Such phenomenon may be explained by a post-transcriptional regulation like the one mediated by micro RNAs (miRNAs) called post-transcriptional gene silencing (PTGS). To test this hypothesis we searched for putative miRNAs targeting the 3'UTR of *Ezh2* transcript and we crossed these results with the miRNAs strongly up-regulated in EpiSCs versus serum-ESCs (Jouneau et al., 2012). We found several matches in particular for miR-367 and Let-7e that may target *Ezh2* and that are specifically up-regulated in the primed state of pluripotency. A direct approach to study the relevance of these miRNAs on EZH2 expression will be to use antagomiR technology (oligomers with a reverse and complementary sequence of the miRNA) to test their presumed function. However using an indirect approach we used *Dgcr8* null-ESCs (kind gift of Constance Ciaudo, ETH Zurich) in which the miRNA pathway is aborted at the pre-miRNA stage in the nucleus. We converted them into cEpiSCs and we observed the same decrease in EZH2 protein levels as in wild-type condition (Unpublished data of the laboratory). This

suggests that the incoherence between mRNA levels and protein levels of EZH2 is not likely due to a miRNA mediated action. We cannot however exclude other pathways of post-translational modifications such as the proteasome targeted degradation mediated by ubiquitin (Schrader et al., 2009) to explain the incoherence between mRNA and protein levels of EZH2 in primed EpiSC.

10 Bivalent Domains

Bivalent domains are regions of chromatin characterized by the co-presence of H3K4me3 and H3K27me3. The simple way to study bivalent domain is to make independent H3K4me3 and H3K27me3-ChIP and then correlate the two to find where there is co-presence of these two histone marks. The problem of this approach is due to the possible effect of heterogeneity of the population that could introduce a bias. To have a more accurate answer it is necessary to perform sequential-ChIP starting with H3K27me3 and successively H3K4me3 and/or vice versa. In this way there is no bias due to the population heterogeneity. Recently a co-ChIP method has been published where Weiner and colleagues (2016) coupled immunoprecipitation and barcoding with DNA adaptors specific to each histone modification. Thus they performed independent ChIP on the same chromatin.

ChIP-seq data have the advantage of giving an absolute quantification, whereas in vitro proximity ligation assay (PLA) permits single-cell analysis in a simplistic way but with only a relative quantification. PLA technique allows the production of a fluorescent signal when two proteins (or in our case two histone modifications) are closer than 30-40nm. It is likely that PLA signals can originate from two or even three successive nucleosomes but also from non-successive nucleosomes brought in proximity by the three-dimensional conformation of the chromatin. In both cases we do not quantify a real bivalent domain defined as the co-presence of H3K4me3 and H3K27me3 on the same nucleosome. ChIP-seq studies do not usually assess bivalency on satellite repeats as they only include uniquely mapped sequences. PLA technique on the other hand can detect proximity of two histone marks even at highly repetitive sequences.

Our comparison between 2i- and serum-ESCs using PLA technique is not in agreement with published ChIP-seq analysis. Indeed previous studies using different kinds of ChIP-seq analysis showed a strong reduction in bivalent promoter genes when ESCs are switched from serum to 2i condition (Marks et al., 2012, Weiner et al., 2016). However with in vivo PLA we

observed an inverse phenomenon with increased bivalency in 2i- compared to serum-ESCs. We made different hypothesis to explain such discrepancy.

First, this can be due to the fact that we used an antibody recognizing both di- and three-methylated H3K27. Indeed to apply PLA technique we had to use antibodies made in mouse and rabbit (mouse for anti-H3K27me2/3 and rabbit for H3K4me3, respectively) and we did not have the opportunity to test whether using an antibody highly specific for H3K27me3 would give similar results as we did not find H3K27me3 antibody made in mouse that works in our hands for immunostaining detection.

Second, as mentioned above, PLA technique can reveal a positive signal even when H3K4me3 and H3K27me2/3 are on non-successive nucleosomes increasing the number of signals. Moreover we observed that major satellite sequences are H3K27me3-enriched in 2i-ESCs but also decompacted through the nucleus when compared to serum-ESCs. Together with the fact that these sequences represent 3% of the genome, the chance to get a PLA signal when in proximity with H3K4me3 enriched nucleosome is increased.

Third we do not know whether there is an enrichment of H3K4me3 on satellites sequences but we cannot completely exclude its presence. Indeed, the possible bivalency on repetitive sequences has never been taken in account previously as it needs a specific analysis on ChIP-seq results different from the standard one on single copy sequences.

To address this problem, we will first make ChIP-qPCR for H3K4me3 on major satellites to assess its eventual enrichment particularly in 2i-ESCs. Secondly we will perform PLA assay to evaluate H3K4me3-H3K27me3 vicinity in wild-type and Suv39hdn serum-ESCs to test whether even in serum condition the increment of H3K27me3 at PCH induces an increase in bivalent loci.

11 Nomenclature ambiguity: ground, naive, primed in mouse and human pluripotency

Historically the first pluripotent stem cells in human were derived from a blastocyst, which induced Thomson and colleagues in 1998 to defined them as human embryonic stem cells (hESCs) to be consistent with Evans, Kaufman and Martin in 1981 for mouse ones. However the deep characterization of hESCs revealed strong differences with mESCs. The more recent discovery of a second state of pluripotency in mouse by Brons and Tesar and colleagues in 2007 clearly showed that hESCs are not as naïve as mESCs but strikingly resemble to primed epiblast stem cells (EpiSCs) which derive from a post-implantation mouse epiblast. So the

first ambiguity in pluripotent stem cells emerged because cells in primed pluripotency state are called EpiSCs in mouse and ESCs in human.

The discovery of serum-free culture condition called 3i by Ying and colleagues in 2008 that makes mouse ESCs completely independent of extracellular signaling led to the definition of the ground state of pluripotency. However, very rapidly the 3i condition was modified to 2i/LIF to improve the cell proliferation. In such medium cells reacquired the dependency to LIF signaling thus losing the connotation of a true “ground” state (Weinberger et al., 2016). Mouse ESCs in 2i/LIF became widely used as they are closer to pre-implantation epiblast cells at the transcriptome and DNA methylation level (Wray et al., 2010; Marks et al., 2012; Habibi et al., 2013). However they have been referred with different names inducing confusion and ambiguity such as naïve, ground-naïve or even ground state (Joshi et al., 2015; Marks et al., 2012). This last designation is inappropriate because only 3i-ESCs (Ying et al., 2008) are really in a ground pluripotency as they are freed from external signals.

The nomenclature controversy goes on as initially “ground” state of pluripotency has been coined to distinguish 3i- (and improperly 2i-) ESCs from serum-ESCs that are considered in a naïve pluripotent state. However serum-ESCs are in a metastable state, constituted of different sub-populations of cells: some are really naïve and others more primed or engaged towards differentiation, likely due to the mixture of factors present in the serum that do not inhibit all the differentiation pathways (Guo et al., 2016; Singer et al., 2014; Toyooka et al., 2008). Knowing this heterogeneity and metastability, serum-ESCs have been also (improperly) referred as being in a primed state of pluripotency (Joshi et al., 2015; von Meyenn et al., 2016) but the true mouse primed pluripotent cells are the EpiSCs (Nichols and Smith, 2009). Moreover the switch of ESCs from 2i- to serum-containing medium has sometimes been called “priming” inducing even more confusion. To our point of view, the priming transition is done when ESCs are cultured in defined medium containing Activin and FGF2 that drive them towards primed EpiSCs.

As in this thesis we studied all the three type of mouse pluripotent stem cells, we tried to make as clear as possible the differences between them in accordance with a recent review on the subject (Weinberger et al., 2016). In particular we state that in mouse there are only two states of pluripotency and that it is the different in vitro culture conditions that can stabilized the circuitry of pluripotency within different “flavors” or types of pluripotent stem cells among the naïve or primed state.

Mouse ESCs are in a naïve state of pluripotency. 2i-ESCs are the more stable and naïve, closer to the ground state. While serum-ESCs are metastable and their instability pushes them

probably closer to the border of primed pluripotency. Conversely mouse EpiSC in ActivinA and FGF2 are in a primed state of pluripotency but alternative culture conditions can drive them closer to the barrier of the naïve state (Figure 26).

Concerning the human species the conventional hESCs, as already told, are in a primed pluripotent state (sharing indeed many features with mEpiSC) but recently many groups tried to define culture conditions to bypass the barrier and stabilize them in a naïve state (Hanna et al., 2010; Gafni et al., 2013; Chan et al., 2013; Ware et al., 2014; Valamehr et al., 2014; Theunissen et al., 2014; Takashima et al., 2014; Duggal et al., 2015; Chen et al., 2015).

The ambiguity in nomenclature that has been created is now inducing confusion when comparisons of the different states of pluripotency are made in the mouse or between mouse and human species. For example, reduction of the number of bivalent domains is considered as an indication of the transition from primed to naïve state for hESCs. However no one has yet compared the number of bivalent domains in primed EpiSCs compared to naïve ESCs. Bivalency in mouse has only been studied between ESCs cultured in 2i or serum conditions (Marks et al., 2012; Weiner et al., 2016). In this thesis, we now provide new data that strongly suggest that mouse primed EpiSCs display less bivalent domains compared to naïve serum-ESCs, as a considerable reduction of H3K4me3/H3K27me3 proximity “spots” is observed using PLA technology.

12 Concluding remarks of this thesis

With this thesis we showed the differences between the three types of mouse pluripotent stem cells, principally in terms of the epigenetic state, organization and transcription of satellites repeats at pericentromeric and centromeric regions. In particular we showed that, depending on the culture condition, naïve ESCs could present compacted, H3K9me3-rich major satellites that are highly transcribed (in serum/LIF) or decompacted, mostly H3K27me3-rich major satellites which are lowly transcribed (in 2i/LIF). We thus propose the H3K27me3-foci enrichment staining as a new marker to define hypomethylated naïve-ground 2i-ESCs. Moreover EpiSCs which correspond to mouse primed pluripotency showed a less plastic heterochromatin and an epigenome landscape closer to somatic cells than naïve ESCs.

In the last years many groups have defined a panel of different culture conditions to revert conventional primed-hESCs into naïve-hESC. Different assays between the studies have been used by the authors to really prove the “naivety” of their hESCs cultured in their own combination of inhibitors and factors, as it is not known which are the key features of human

naïve pluripotency. Making the parallel with the mouse we therefore propose to study the state of pericentromeric and centromeric heterochromatin (PCH/CH) in human pluripotent cells. Indeed also in human it is known that a stable naïve pluripotency is accompanied by a general hypomethylated state that should concern also PCH and CH. These regions have not been extensively studied in hESCs even in primed one or during pre-implantation embryo development (van de Werken et al., 2014). It will be very interesting to study epigenetic state, organization and transcription of pericentromeric and centromeric satellite sequences in human pluripotency, as it has been done extensively in mouse. It should be noticed that in human no all the chromosomes have the same structure (acrocentric, metacentric and sub-metacentric). Moreover the sequences of the repetitive satellites in the pericentromeric region are chromosome-specific (principally made by satellite II and III). Finally human chromosomes do not clusterize their PCH into chromocenters, which are a peculiar feature of mouse cells. Such human-specific features of constitutive heterochromatin may have induced different types of regulation compared to mouse however the specific pericentromeric pathway of H3K9me3 deposition by SUV39H1/2 is conserved also in human. Thus we speculate that PRC2-recruitment pathway at hypomethylated pericentromeric satellites could be also conserved in human.

**RESUME
SUBSTANTIEL
DE LA THESE EN
FRANÇAIS**

RESUME SUBSTANTIEL DE LA THESE EN FRANÇAIS

1 INTRODUCTION

La pluripotence est définie comme la capacité des cellules à pouvoir se différencier en tissus appartenant aux trois feuillets embryonnaires. Chez la souris, les cellules souches embryonnaire dérivées in vitro ont permis de distinguer l'état naïf et amorcé de la pluripotence (Nichols and Smith, 2009). Les cellules souches embryonnaires (ESCs) colonisent la masse cellulaire interne (ICM) lorsqu'elles sont injectés dans un blastocyste et permettent la génération de chimères à terme. En revanche, les cellules souches épiblastiques (EpiSCs) ne colonisent que l'épiblaste post-implantation et leur contribution au développement ultérieur reste à démontrer (Huang et al., 2012). Les deux états sont maintenus in vitro en utilisant des voies de signalisation spécifiques, principalement Lif/Stat3 (Facteur d'inhibition de la leucemie/Transducteur de signal et activateur de transcription) pour les cellules naïves (ESCs), et FGF et Activine pour les cellules amorcées (EpiSCs) (Brons et al., 2007). Les ESCs sont classiquement maintenues dans un milieu contenant du sérum mais peuvent aussi être cultivées dans un milieu sans sérum avec des inhibiteurs de deux voies de différenciation, la voie MAPK/ERK (protéine kinases activées par les mitogènes/kinases régulées par des signaux extracellulaires) et la voie GSK3 (Glycogène synthase kinase 3) (Ying et al., 2008). Dans ce milieu 2i/Lif (plus tard dénommé "2i" uniquement), les cellules acquièrent un état encore plus naïf, avec une répression plus efficace des marqueurs de la différenciation et une expression plus homogène des gènes de la pluripotence (Marks et al., 2012). En revanche, le transcriptome des EpiSCs reflète leur nature amorcée, car ils expriment déjà de nombreux marqueurs de différenciation précoce tandis que certains gènes de pluripotence sont réprimés (Brons et al., 2007). Les ESCs cultivées dans du sérum/Lif présentent un transcriptome intermédiaire, avec des niveaux hétérogènes de marqueurs de pluripotence et une faible, mais détectable, expression des gènes de différenciation (Marks et al., 2012). Les ESCs sont considérées comme ayant une plus grande ouverture de la chromatine et une organisation plastique du noyau par rapport aux cellules différenciées (Ahmed et al., 2010; Koh et al., 2010; Meissner, 2010; Meshorer et al., 2006). En effet, leur épigénome est rapidement et de manière réversible modifié en fonction du milieu de culture, 2i ou sérum (Habibi et al., 2013; Marks et al., 2012). Plus précisément, les ESCs en 2i

présentent une réduction significative du dépôt d'H3K27me3 au niveau des promoteurs ainsi qu'un ADN globalement hypométhylé (Habibi et al., 2013; Marks et al., 2012). Les ESCs peuvent être converties en EpiSCs (cEpiSCs) *in vitro* lorsqu'elles sont exposées à la signalisation FGF et Activine au lieu de Lif avec du sérum ou des inhibiteurs (2i) (Guo et al., 2009). En revanche, la réversion des EpiSCs en cellules naïves est un processus long et inefficace, mettant en évidence l'existence d'une barrière épigénétique (Bao et al., 2009). Bien que peu d'analyses approfondies aient été menées, les données disponibles indiquent que lors de la conversion vers l'état amorcé, de nombreux promoteurs deviennent hyperméthylés au niveau de l'ADN conjointement avec une réorganisation substantielle de l'activité des enhancers, par rapport aux ESCs (Factor et al., 2014; Veillard et al., 2012). Les études mentionnées ci-dessus suggèrent que l'organisation de l'épigénome serait caractéristique de chaque type de cellule pluripotente. Cette comparaison n'a toutefois pas encore été réalisée au niveau de l'hétérochromatine constitutive. Ce compartiment est composé de séquences d'ADN répétées situées aux télomères, centromères et au niveau des régions péri-centromériques (Biscotti et al., 2015). Le contrôle adéquat de ces régions est crucial pour la stabilité chromosomique (Ferreira et al., 2015). En dehors des séquences télomériques, il existe deux types de séquences répétées chez la souris: les séquences satellites majeurs et les séquences satellites mineurs qui forment respectivement l'hétérochromatine péri-centromérique (PCH) et centromérique (CH) (Guenatri et al., 2004). Les satellites majeurs consistent en une répartition de plus de 200.000 fois d'une séquence de 234bp et représente environ 3% du génome de la souris, tandis que l'unité de répétition des satellites mineurs est une séquence 123pb répétée au moins 50.000 fois (Martens et al., 2005).

Dans les cellules somatiques, PCH et CH provenant de chromosomes différents vont s'agréger pour former des "chromocentres", qui sont colorés densément avec du DAPI, en raison de leur teneur élevée en AT, et sont généralement enrichis pour la marque épigénétique répressive H3K9me3 déposée par SUV39H1/2 (Peters et al., 2003). La méthylation de l'ADN est une autre caractéristique de l'hétérochromatine constitutive qui coexiste avec H3K9me3 (Déjardin, 2015). La 5-méthylcytosine (5-meC) est une marque répressive *de novo* déposée spécifiquement par DNMT3A/B, et maintenue au cours de la réplication par DNMT1 (Okano et al., 1999). Un tel état épigénétique (méthylation de l'ADN et triméthylation de l'histone H3) n'est pas favorable à la transcription; par conséquent, en cellules somatiques, les répétitions satellites sont réprimés (Lu and Gilbert, 2007). Toutefois, lors la sénescence cellulaire, dans certains cancers et au début du développement, l'activation de la transcription de ces séquences a été observée, le plus souvent en corrélation avec une réduction de la méthylation

de l'ADN (Revue dans Saksouk et al., 2015). Dans les ESCs en sérum, bien que les séquences satellites soient enrichies en H3K9me3 et 5-meC (Okano et al., 1999; Peters et al., 2003), la transcription est plus élevée par rapport à celle de cellules différenciées telles que les progéniteurs neuronaux (Efroni et al., 2008). Si l'une ou l'autre marque de l'hétérochromatine constitutive (H3K9me3 ou 5mC) est manquante, comme dans le cas des ESCs invalidées pour les gènes *Suv39h* ou *Dnmt*, il est observé aux PCH un enrichissement dans la marque épigénétique caractéristique de l'hétérochromatine facultative: H3K27me3 (Cooper et al., 2014; Lehnertz et al., 2003; Peters et al., 2003).

2 OBJECTIFS

Cette thèse présente une étude comparative principalement des caractéristiques épigénétiques et de la dynamique de transcription de l'hétérochromatine constitutive de trois types de cellules pluripotentes chez la souris. Les objectifs des études menées ici sont d'apporter des connaissances nouvelles sur les différents états de pluripotence en lien avec les différentes conditions de culture *in vitro* chez la souris. En effet, la majorité des études épi-génomiques (analyses de données de séquençage après immunoprécipitation de la chromatine - ChIP-Seq) réalisées jusqu'à présent sur les différents états de pluripotence étaient concentrées sur les séquences codantes du génome (généralement en copie unique). Cependant, chez la souris, les régions répétées sont très représentées, en particulier les régions péri-centromériques et centromériques composées respectivement de séquences satellites majeurs et mineurs correspondent à 3,5% du génome. La chromatine des ESCs en sérum a été décrite comme très dynamique, notamment les séquences satellites présentent une organisation plus décompactée par rapport celles observées en cellules différenciées, favorisant ainsi la transcription de ces séquences non-codantes. Il a été montré dans des travaux précédent, que le taux d'ADN méthylé (5-meC) était grandement réduit et que la marque H3K27me3 était redistribuée lors de l'adaptation des ESCs en milieu 2i. Toutefois la distribution de ces marques au niveau de PCH n'a pas encore été étudiée dans ce contexte. De plus, comme nous supposons l'existence d'une barrière épigénétique entre les ESCs et les EpiSCs, nous étudierons aussi l'organisation, la présence des marques épigénétiques et l'état transcriptionnel de PCH dans les EpiSCs. La question de savoir si cette chromatine ouverte et l'état hypertranscriptionnel de l'hétérochromatine constitutive est maintenue également dans les ESCs en 2i ainsi que dans les EpiSCs va être étudiée dans cette thèse.

3 RESULTATS

3.1 L'hétérochromatine est caractérisée par différentes modifications d'histones selon le type de cellules pluripotentes.

Pour caractériser et comparer la distribution nucléaire des marques d'histones H3K9me3 et H3K27me3 au niveau des PCH/CH dans les ESCs (en sérum ou 2i) et EpiSCs, nous avons effectué des immunomarquages pour ces deux marques épigénétiques ainsi que la coloration DAPI pour détecter les chromocentres. Nous avons évalué le pourcentage de cellules présentant une accumulation en « foyers » de H3K9me3 ou H3K27me3. Nous avons aussi déterminé leur colocalisation avec les foyers DAPI en utilisant la fonction linscan de ImageJ qui permet de tracer le profil des intensités de fluorescence au travers du noyau. Nos données montrent que la distribution des deux marques est différente selon le type de cellule. Dans les ESCs en 2i, seulement un tiers de la population présente ces foyers de H3K9me3 (Figure 1A). Dans ces cellules, ces foyers sont rares et petits, et situés à proximité de la périphérie du noyau (voir le grossissement du noyau unique sur la Figure 1A). A l'inverse toutes les ESCs en sérum et les EpiSCs présentent de nombreux foyers d'H3K9me3, qui co-localisent parfaitement avec les zones riches en DAPI (Figure 1A).

En ce qui concerne H3K27me3, nous avons observé une distribution nucléaire distincte dans chacun des trois types cellulaires. En effet, dans les ESCs en 2i, la majorité des cellules (69%) présentent un enrichissement en H3K27me3 aux zones riches en DAPI (voir le grossissement du noyau unique sur la Figure 3A). La répartition au sein du noyau suggère que H3K27me3 se substitue à H3K9me3 dans ces régions. En revanche, H3K27me3 est diffus dans la plupart des noyaux des ESCs en sérum (95%), comme on s'y attend pour une marque d'hétérochromatine facultative (Figure 1A). Dans les EpiSCs, le signal H3K27me3 est aussi diffus mais surtout très faible (Figure 1A).

Pour explorer davantage ces différences, nous avons évalué le niveau global des deux marques d'histone dans les extraits cellulaires. La quantification par western-blot indique une légère baisse du niveau global de H3K9me3 dans les ESCs en 2i par rapport aux ESCs en sérum et aux EpiSCs, et surtout une perte très prononcée de H3K27me3 dans les EpiSCs par rapport aux ESCs (en sérum ou 2i) (Figure 1B). Par conséquent, dans les EpiSCs, H3K27me3 a été réduit non seulement au PCH, mais aussi ailleurs dans le génome.

Nos données suggèrent donc que H3K27me3 est redistribuée à partir de l'hétérochromatine facultative vers l'hétérochromatine constitutive lorsque les ESCs sont transférés dans le

milieu 2i, tout en étant perdu dans les EpiSCs. A l'inverse, H3K9me3 est légèrement diminuée dans les ESCs en 2i et peu enrichi dans les PCH/CH, et au contraire très enrichie dans ces régions dans les ESCs en sérum et les EpiSCs.

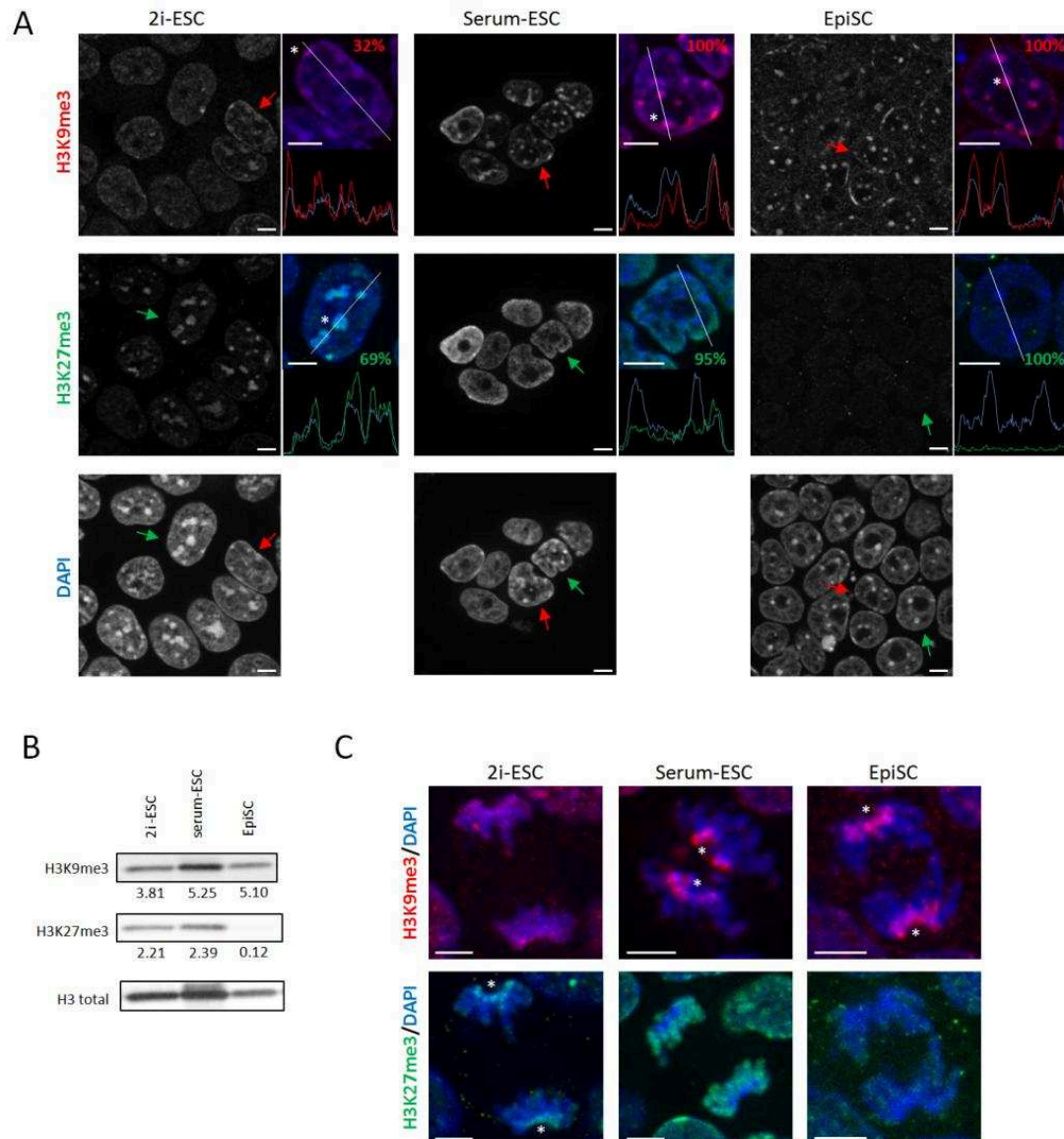


Figure 76 : Marques épigénétique de l'hétérochromatine dans les trois types de cellules pluripotentes chez la souris. (A) Immunomarquage pour H3K9me3 et H3K27me3 avec controcoulage de l'ADN avec du DAPI. Grossissement sur une seule cellule (flèche) avec fusion de signaux: H3K9me3 (rouge) ou H3K27me3 (vert) avec DAPI (bleu). Analyse « linescan » des enrichissements (mis en évidence avec l'étoile). % Indique le pourcentage de cellules dans la population affichant la même organisation. Les barres d'échelle représentent 5 μ m. (B) Analyse par western-blot pour la quantification des niveaux globaux des modifications d'histones H3K9me3 et H3K27me3 par rapport au niveau total de H3. (C) Immunomarquages sur des figures de mitoses pour H3K9me3 (rouge) ou H3K27me3 (vert) avec contre-coloration de l'ADN avec du DAPI (bleu). Les étoiles indiquent l'enrichissement de la marque histone aux PCH / CH. Les barres d'échelle représentent 5 μ m.

Pour renforcer ces observations, nous avons examiné les chromosomes des cellules en mitose. Chez la souris, les chromosomes sont acrocentriques, et donc les PCH/CH sont localisés à une

extrémité. H3K27me3 est effectivement clairement enrichi au PCH/CH dans les ESCs en 2i, tandis que H3K9me3 est présent tout au long des chromosomes (Figure 1C). A l'inverse, dans les ESCs en sérum et dans les EpiSCs les extrémités des chromosomes mitotiques sont fortement enrichies en H3K9me3, tandis que H3K27me3 est présent le long des chromosomes des ESCs en sérum et indétectable dans les EpiSCs (Figure 1C).

3.2 Faible niveau de méthylation de l'ADN au PCH dans les ESCs en 2i

Nous avons également analysé les niveaux de méthylation de l'ADN par digestion avec des enzymes sensibles à la méthylation, suivies par un Southern-blot révélant spécifiquement les séquences satellites majeurs. Nous avons observé des changements importants de méthylation de l'ADN dans ces séquences entre les différents types des cellules pluripotentes (Figure 2A).

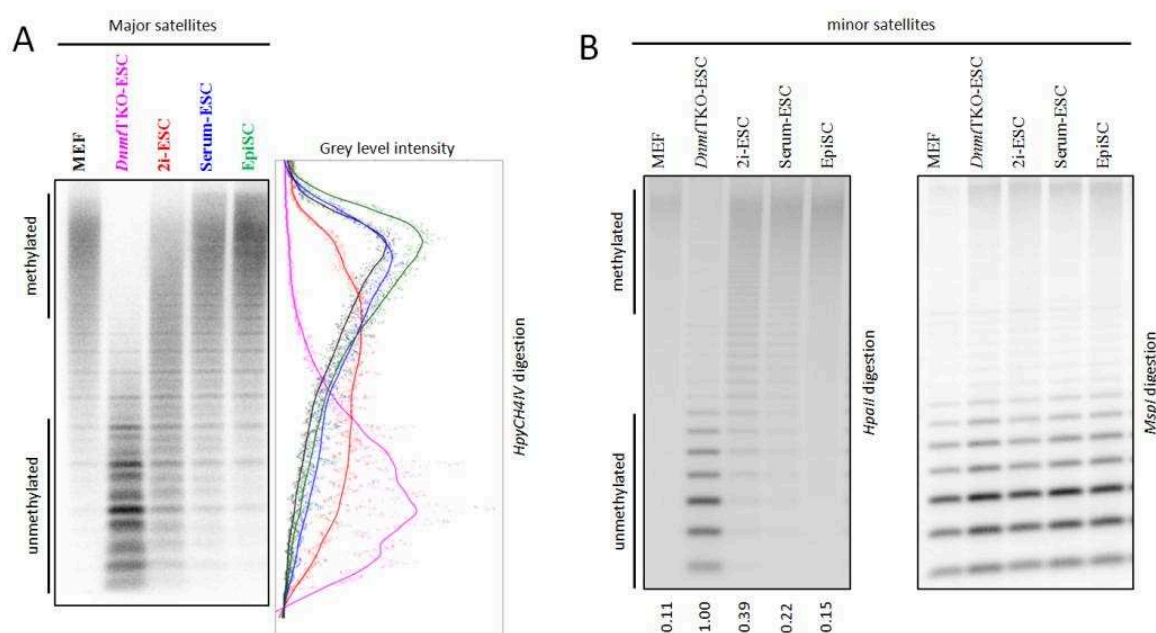


Figure 77 : Profil de méthylation des séquences satellites majeurs et mineurs dans les trois types de cellules pluripotentes.

(A) Analyse par southern-blot de l'ADN génomique digéré avec HpyCH4IV et révélé avec la sonde pour les satellites majeurs. Quantification « linescan » pour chaque piste: MEF (noir), DnmtTKO (rose), 2i-ESC (rouge), sérum ESCs (bleu) et EpiSCs (vert). (B) Analyse par southern-blot de l'ADN génomique digéré avec HpaII et révélé avec une sonde pour les satellites mineurs. Quantification concernant le southern-blot de l'ADN génomique digéré avec MspI. DnmtTKO-ESC est fixé sur 1.

Les satellites majeurs sont partiellement déméthylées dans les ESCs en 2i (comme le montre le profil rouge), le niveau de méthylation étant intermédiaire entre des cellules entièrement

déméthylées (mutants invalidés pour les trois enzymes de méthylation de l'ADN = DnmtTKO (ligne violette)) et des fibroblastes hyperméthylés (MEFs - ligne noire) (Figure 2A). Au contraire, ces séquences sont tout aussi méthylées dans les ESCs en sérum (ligne bleue) et dans les EpiSCs (ligne verte) que dans les MEFs. Une situation similaire a été observée pour les satellites mineurs qui sont partiellement déméthylées dans les ESCs en 2i et hyperméthylées dans les ESCs en sérum et dans les EpiSCs (Figure 2B).

3.3 Le PCH est décondensé dans les ESCs en 2i mais transcriptionnellement réprimé

Pour évaluer l'organisation spatiale des chromocentres en fonction de l'état de pluripotence, nous avons réalisé une DNA-FISH (Hybridation in situ en fluorescence) en utilisant des sondes spécifiques pour les séquences satellites majeures et mineures (Figure 3A), suivie d'une segmentation des noyaux et une reconstruction tridimensionnelle des cellules avec le logiciel AMIRA 3.1 (Figure 3A).

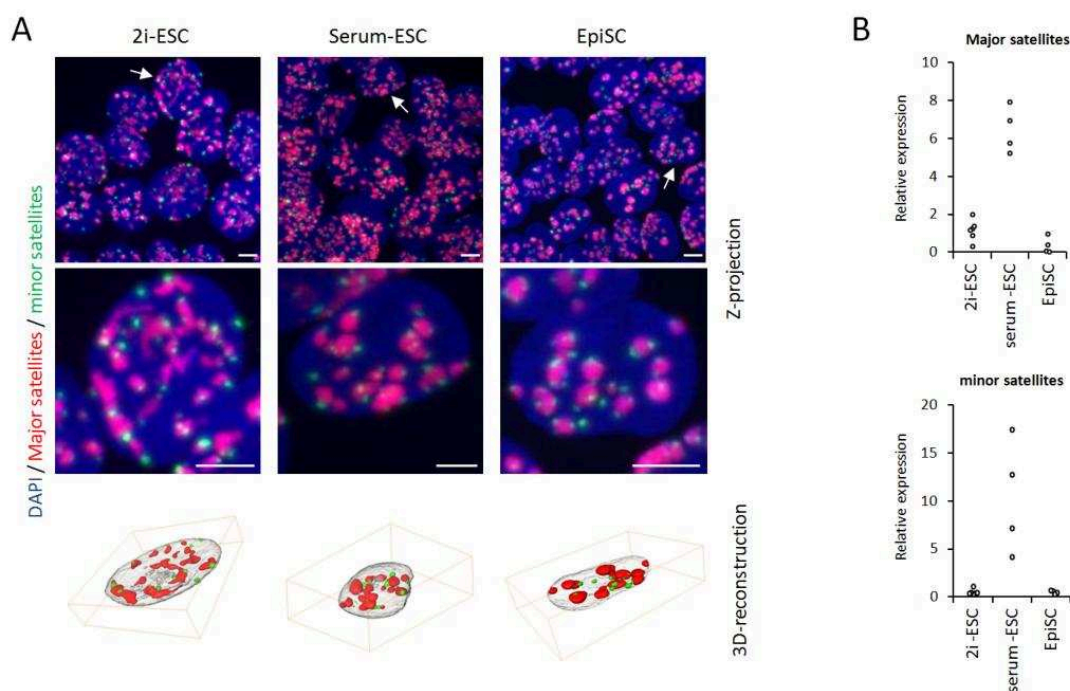


Figure 78 : Organisation tridimensionnelle et transcription des séquences satellites majeurs et mineurs dans les trois types de cellules pluripotentes.

(A) DNA-FISH pour les satellites majeurs (rouge) et mineurs (vert) avec contre-coloration de l'ADN en DAPI (bleu). Grossissement sur une seule cellule (fléchée) et reconstruction tridimensionnelle des signaux des satellites majeurs et mineurs en utilisant le logiciel AMIRA 3.1. Les barres d'échelle représentent 5 μ m. (B) Expression relative des transcrits des séquences majeures et mineures par analyse qRT-PCR normalisés par les gènes de référence *Sdha* et *Pbgd*. Chaque point est un réplicat biologique indépendant.

Dans les EpiSCs et les ESCs en sérum les séquences satellites majeures sont organisées en domaines compacts ronds, entourés par de petits points formés des satellites mineurs, comme des chromocentres classiques que l'on trouve dans les cellules somatiques (Guenatri et al., 2004). En revanche, dans les ESCs en 2i les satellites majeurs forment des domaines plus lâches, de forme irrégulière et allongée, souvent en demi-anneaux autour des nucléoles ou à la périphérie nucléaire, indiquant des chromocentres peu structurés. Une telle décompaction devrait créer un environnement transcriptionnellement permissif. Nous avons donc évalué le taux de transcrits associés aux séquences satellites par qRT-PCR dans les ESCs et les EpiSCs (Figure 3B). Paradoxalement, les ESCs en 2i ont un niveau de transcrits des séquences satellites majeurs inférieur à celui des ESCs en sérum. Dans les EpiSCs, l'expression est encore plus faible que dans les ESCs en 2i. La même tendance est observée pour les satellites mineurs. En conclusion, malgré des modes différents de répression et de compaction, les ESCs en 2i et EpiSCs répriment toutes deux la transcription à partir des PCH/CH. De leur côté, les ESCs en sérum sont largement permissive pour la transcription de ces séquences, malgré la présence des marques épigénétiques répressives H3K9me3 et 5-meC.

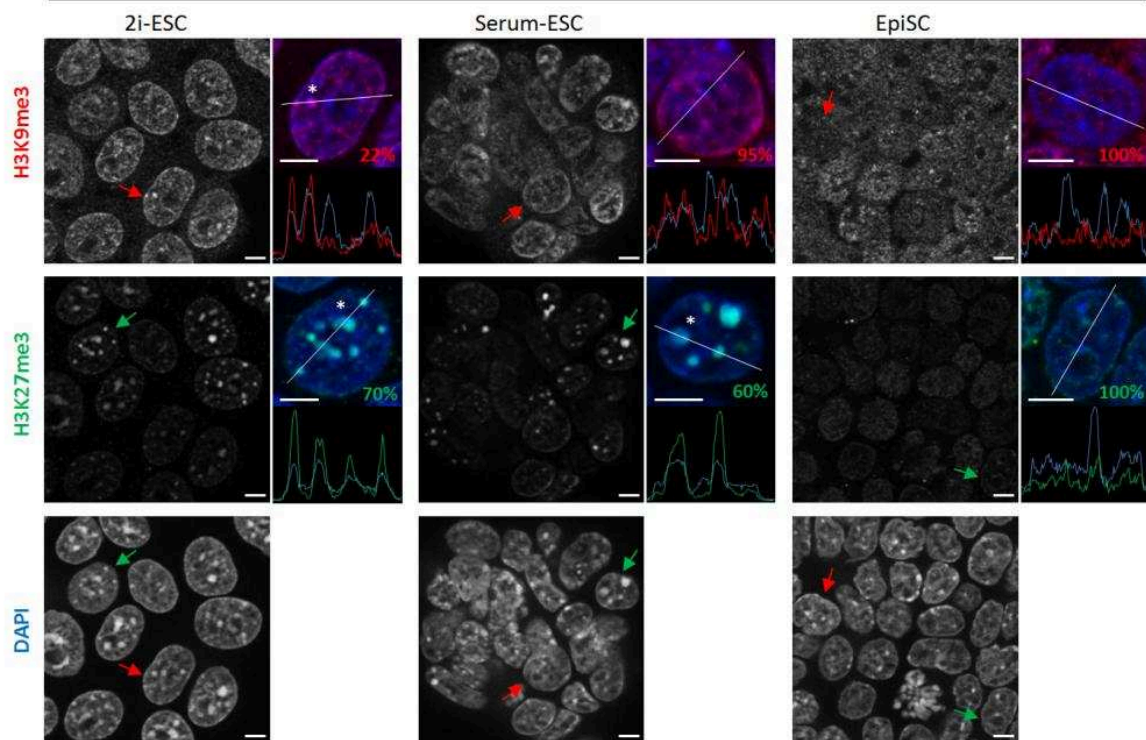
3.4 L'absence des SUV39H1/2 induit des phénotypes différents en fonction du type de cellule pluripotente.

Pour étudier les interactions entre les marques répressives au PCH dans les différents types des cellules pluripotentes, nous avons examiné leur distribution lors de l'absence d'H3K9me3. Pour cela, nous avons utilisé les ESCs mutantes Suv39hdn, dans lesquelles les deux enzymes Suv39h ont été invalidées (Peters et al., 2003). Nous avons adapté ces ESCs dans les conditions de culture en 2i et nous les avons converties in vitro en EpiSCs (cEpiSCs). Dans les ESCs mutantes en 2i, aucune modification évidente d'H3K9me3 n'a pu être observée par rapport aux cellules de type sauvage. Cependant, une proportion importante des cellules (22%) présentait encore les mêmes petits foyers d'H3K9me3 que dans les cellules contrôles, ce qui indique que ce dépôt d'H3K9me3 est indépendant des enzymes SUV39H1/2 (Figure 4A à comparer avec Figure 1A). Par ailleurs, le même enrichissement de H3K27me3 au PCH et la même hypométhylation de l'ADN aux séquences satellites majeures ont été observés (Figure 4A et 4B). Le niveau de transcription des séquences satellite majeurs est globalement inchangé dans les cellules mutantes par rapport au type sauvage. Donc, dans les ESCs en 2i, l'organisation de PCH est largement indépendante du dépôt de H3K9me3 par SUV39H1/2. Dans les ESCs en sérum, H3K9me3 est principalement diffus, avec 95% des

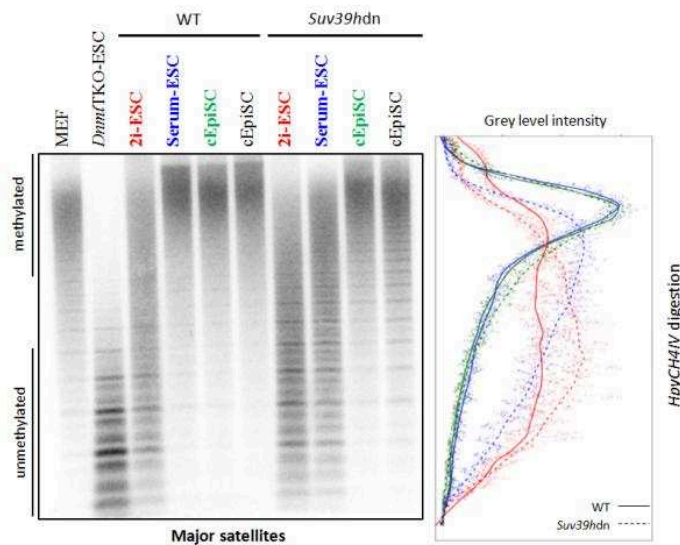
cellules ne présentant pas d'enrichissement en foyers (Figure 4A). Il avait déjà été démontré que H3K27me3 pouvait se substituer à H3K9me3 au PCH dans les ESCs mutantes en sérum (Peters et al., 2003; Martens et al., 2005). Nous avons confirmé ce phénotype, en effet 60% des ESCs Suv39hdn en sérum présentent des enrichissements de H3K27me3 aux régions riches en DAPI (Figure 4A). Ensuite, nous avons montré que l'absence de SUV39H1/2 dans les ESCs en sérum induit une réduction de la méthylation de l'ADN par rapport au type sauvage sur les satellites majeurs (Figure 4B), confirmant les travaux antérieurs (Lehnertz et al., 2003). La transcription des satellites majeurs est légèrement diminuée dans les ESCs Suv39hdn en sérum malgré la perte des marques répressives (Figure 4C). Par conséquent, l'absence de SUV39H1/2 dans les ESCs en sérum induit un phénotype qui récapitule les conditions des ESCs sauvages en 2i: augmentation H3K27me3 et réduction de la méthylation de l'ADN, ainsi que la réduction de la transcription à partir du PCH (mais pas aussi faible que dans ESC en 2i).

Dans les cEpiSCs mutantes, nous avons observé une situation différente. Ces cellules ont complètement perdu les foyers d'H3K9me3, mais contrairement aux ESCs, elles ne gagnent pas d'enrichissement en H3K27me3 (Figure 4A). Par conséquent, dans les cEpiSCs, H3K27me3 ne se substitue pas à H3K9me3 au PCH. En outre, les satellites majeurs sont hyperméthylés dans les cellules mutantes comme dans les sauvages (Figure 4B). De manière intéressante, dans les cEpiSCs Suv39hdn, la transcription des satellites majeurs est déréprimée, jusqu'au niveau de celle observée dans les ESC mutantes en sérum (Figure 4C). Pour conclure, dans les ESCs en sérum l'absence d'H3K9me3 au PCH entraîne une réduction de la méthylation de l'ADN et l'accumulation de H3K27me3, en revanche dans les EpiSCs mutants il n'y a pas d'enrichissement d'H3K27me3, la méthylation ADN est restaurée, mais cela n'empêche pas la de-répression de la transcription des satellites majeurs.

A

*Suv39h*dn

B



C

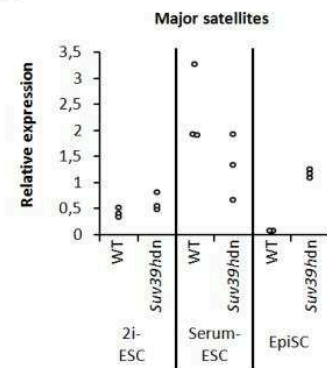


Figure 79 : Effets contrastés de la condition de Suv39hdn dans les trois types des cellules pluripotentes.

(A) Immunomarquage pour H3K9me3 et H3K27me3 avec contre-coloration ADN avec DAPI dans la condition Suv39hdn, à comparer avec la condition sauvage (Figure 3A). Grossissement sur une seule cellule (flèche) avec fusion de signaux: H3K9me3 (rouge) ou H3K27me3 (vert) avec DAPI (bleu). Analyse « linescan » présentant les enrichissements (mis en évidence avec l'étoile). % Indique le pourcentage de cellules dans la population affichant la même organisation. Les barres d'échelle représentent 5 μm. (B) Analyse par southern-blot de l'ADN génomique digéré avec HpyCH4IV révélé avec sonde pour les satellites majeurs dans les cellules type sauvage et Suv39hdn. Quantification « linescan » pour chaque piste: 2i-ESC (rouge), sérum ESC (bleu) et EpiSC (vert). Le type sauvage est représentée par une ligne continue, tandis que la condition Suv39hdn avec une ligne pointillée. (C) Expression relative des satellites majeurs par l'analyse qRT-PCR normalisée par les gènes de référence Sdhα et Pbgd dans le type sauvage et l'état Suv39hdn. Chaque point est un réplicat biologique indépendant.

3.5 L'absence de méthylation de l'ADN augmente le dépôt d'H3K27me3 mais a un effet limité sur la transcription des séquences satellites.

Pour étudier le rôle de la méthylation de l'ADN dans la régulation de l'organisation des satellites majeurs et leur transcription, nous avons utilisé des ESCs invalidées pour Dnmt1, 3a et 3b (DnmtTKO). Ces cellules n'ont pas de cytosine méthylée dans leur génome (Tsumura et al., 2006) et notamment au PCH (Figure 2A). Nous avons pu adapter ces cellules en milieu 2i, mais pas les convertir en cEpiSCs : toutes les cellules mouraient par apoptose pendant la conversion ce qui est en accord avec leur faible contribution au développement de l'épiblaste in vivo (Sakaue et al., 2010). Les ESCs DnmtTKO en sérum montrent un fort enrichissement en H3K27me3 aux PCH (Figure 5A), comme montré dans des études antérieures (Cooper et al., 2014; Saksouk et al., 2014).

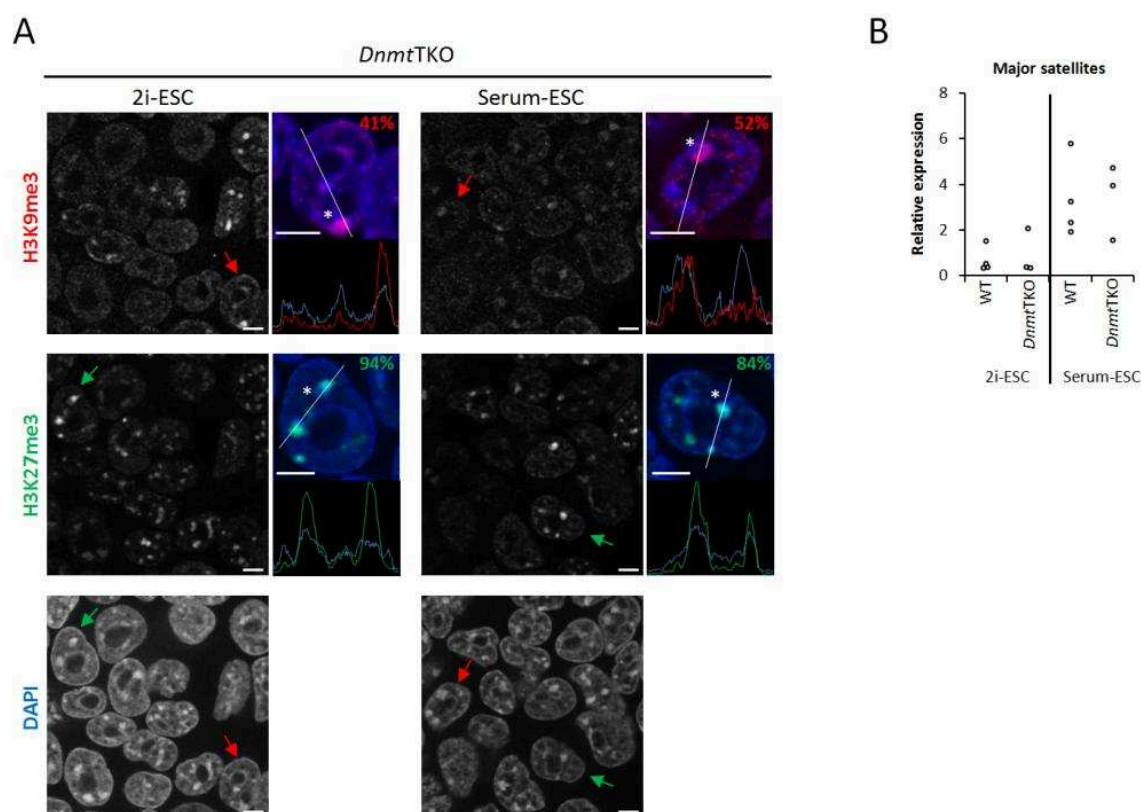


Figure 80 : Conséquences épigénétiques et transcriptionnelles de l'absence de la méthylation de l'ADN sur les satellites majeurs dans les ESCs en 2i ou sérum.

(A) Immunomarquages pour H3K9me3 et H3K27me3 avec contre-coloration de l'ADN en DAPI dans la condition DnmtTKO, à comparer avec la condition sauvage (Figure 3A). Grossissement sur une seule cellule (flèche) avec fusion de signaux: H3K9me3 (rouge) ou H3K27me3 (vert) avec DAPI (bleu). Analyse « linescan » présentant les enrichissements (mis en évidence avec l'étoile). % Indique le pourcentage de cellules dans la population affichant le même schéma. Les barres d'échelle représentent 5 μ m. (B) Expression relative des satellites majeurs par l'analyse qRT-PCR normalisé avec les gènes de référence *Sdha* et *Pbgd* dans les cellules de type sauvage ou DnmtTKO. Chaque point est un réplicat biologique indépendant.

En effet, nous avons constaté que la plupart (84%) des ESCs mutantes en sérum présentent un enrichissement d'H3K27me3 dans les régions riches en DAPI. Cet enrichissement se fait au détriment de celui d'H3K9me3 (Figure 5A) car 100% des ESCs sauvages en sérum contient des foyers d'H3K9me3 contre seulement 52% des cellules mutantes (Figure 5A à comparer avec Figure 1A). Dans le milieu 2i, l'absence totale de méthylation de l'ADN conduit à une proportion accrue de cellules présentant des foyers de H3K27me3 (94% 2i-DnmtTKO vs 69% dans le type sauvage, Figure 5A and Figure 1A). En revanche la distribution d'H3K9me3 dans les ESCs en milieu 2i est très similaire dans les DnmtTKO par rapport au type sauvage (Figure 5A and Figure 1A). En outre la transcription des satellites majeurs ne change pratiquement pas (Figure 5B). En conclusion, l'absence de méthylation de l'ADN modifie considérablement l'état épigénétique du PCH dans les ESCs en sérum, induisant un fort enrichissement en H3K27me3 et une réduction des foyers d'H3K9me3, mais sans effet clair sur la transcription des satellites majeurs.

3.6 La réduction des niveaux d'H3K27me3 n'induit pas une sur-expression des séquences satellites

Afin d'étudier le rôle de H3K27me3 dans la régulation de l'état transcriptionnel du PCH, nous avons utilisé un inhibiteur de l'activité méthyl-transférase de EZH2. Nous avons choisi EPZ-6438 (appelle par simplicité EPZ) qui est un inhibiteur sélectif d'EZH2 (Knutson et al., 2014). Le traitement des ESCs avec EPZ pendant 72 heures conduit à une réduction remarquable du niveau globale d'H3K27me3 (au moins 70% dans chaque type cellulaire testé), sans changement dans les niveaux de H3K9me3 (Figure 6A-E). L'immunomarquage des ESCs en 2i traités 72h avec EPZ a également confirmé la perte des foyers H3K27me3 dans la majorité des cellules sans changements majeurs dans l'organisation de H3K9me3 par rapport aux cellules témoins traités avec du DMSO (Figure 6A – panneau de droite). De plus, l'état hypométhylé des séquences satellites dans les ESCs en 2i traitées n'est pas modifié, ce qui montre qu'il n'y a pas de remplacement H3K27me3 par la méthylation de l'ADN au PCH (données non présentées). Nous avons ensuite analysé la transcription des satellites majeurs. Contre toute attente, la perte de H3K27me3 dans les ESCs en 2i n'a pas conduit à une augmentation de la transcription mais, au contraire, à une légère réduction (Figure 6A). Quand les ESCs en sérum ont été traitées avec l'inhibiteur de EZH2, bien que H3K27me3 ne soit pas aussi fortement enrichi au PCH comme dans la condition 2i, nous avons également observé une réduction des transcrits des séquences satellites (Figure 6B). Nous avons effectué

la même inhibition dans les cellules mutantes qui présentent un enrichissement similaire de H3K27me3 au PCH, à savoir les ESCs Suv39hdn en sérum et les ESCs DnmtTKO. Dans tous les cas, nous avons observé une légère réduction de la transcription des satellites majeurs (Figure 6C-E). En conclusion, la réduction d'H3K27me3 au PCH n'induit pas une surexpression des transcrits des satellites majeurs, même en absence des autres marques répressives (comme H3K9me3 ou méthylation de l'ADN).

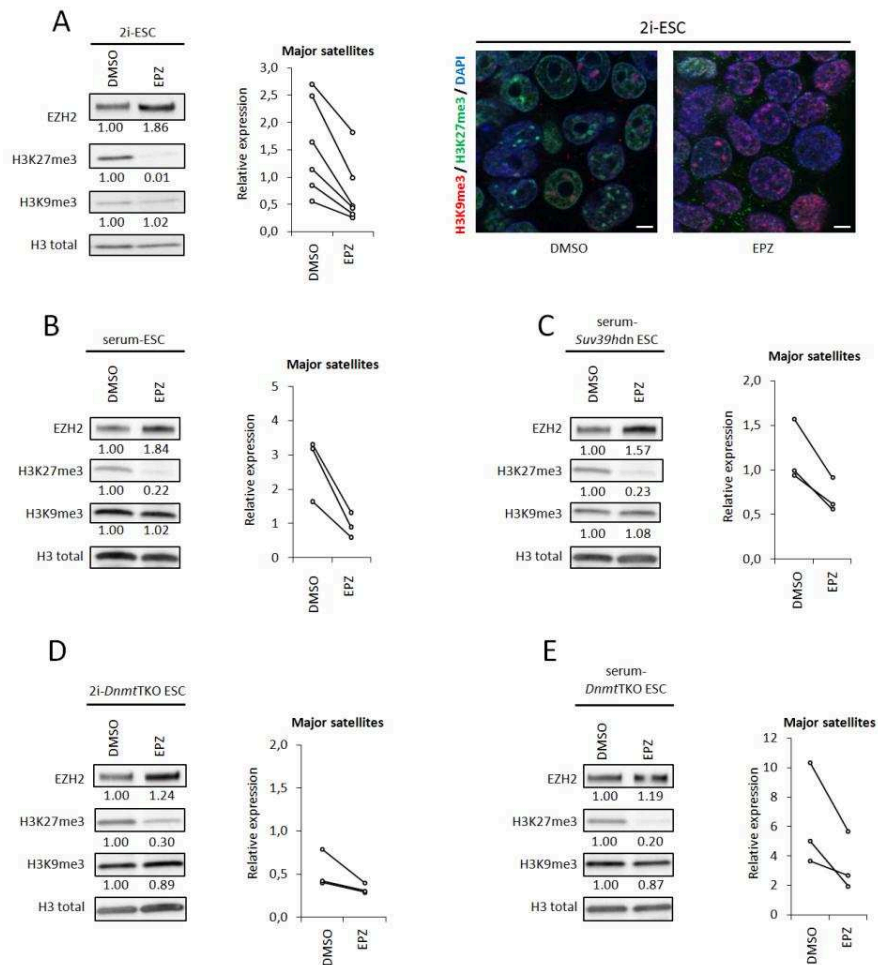


Figure 81 : La réduction des niveaux de H3K27me3 n'induisent pas une sur-expression des séquences satellites majeurs.

(A) Partie gauche: analyse par western-blot pour la quantification des niveaux global de l'enzyme HMT EZH2 et les modifications des histones H3K9me3 et H3K27me3 par rapport a H3 total dans les ESCs en 2i traités avec du DMSO (témoin) ou EPZ (Inhibition EZH2). Partie centrale: Expression relative des transcrits des satellites majeurs par analyse qRT-PCR normalisés par des gènes de référence *Sdha* et *Pbgd* dans les ESCs en 2i traités avec du DMSO ou EPZ. Chaque point est un réplicat biologique indépendant. Partie droite: immunomarquage pour H3K9me3 (rouge) et H3K27me3 (vert) avec contre-coloration de l'ADN en DAPI (bleu) dans les ESCs en 2i traités avec du DMSO ou EPZ. Les barres d'échelle représentent 5 μ m. (B, C, D, E) Pour chaque condition: ESCs sérum de type sauvage (B), ESCs sérum Suv39hdn (C), ESCs en 2i DnmtTKO (D) et ESC en sérum DnmtTKO (E). Partie gauche: analyse par western-blot pour la quantification de niveau global de l'enzyme HMT EZH2 et les modifications d'histones H3K9me3 et H3K27me3 par rapport à H3 totale. Partie droite: Expression relative des transcrits satellites par analyse qRT-PCR normalisés par les gènes de ménage *Sdha* et *Pbgd* après traitement avec du DMSO (témoin) ou EPZ (Inhibition de EZH2). Chaque point est un replicat biologique indépendant.

4 DISCUSSION

4.1 Dialogue entre H3K9me3, H3K27me3 et 5-meC dans la pluripotence chez la souris

Nous montrons que l'interaction entre les différentes marques répressives au PCH est modulée en fonction de l'état de pluripotence et des conditions de culture. Nous avons observé que lors d'une réduction de la méthylation de l'ADN, H3K27me3 s'enrichit au PCH dans trois contextes différents: (i) après adaptation des ESCs au milieu 2i, qui conduit à une déméthylation générale au niveau de l'ADN (Habibi et al., 2013); (ii) dans la condition Suv39hdn, qui induit dans les ESCs en sérum une déméthylation partielle d'ADN des satellites majeurs, comme précédemment observé (Cooper et al., 2014; Martens et al., 2005; Peters et al., 2003; Saksouk et al., 2014); (iii) en l'absence totale de 5-meC (DnmtTKO) dans les ESCs en sérum (également montré dans Saksouk et al., 2014). Ailleurs dans le génome, la méthylation de l'ADN antagonise également le dépôt d'H3K27me3, comme le montre les ESCs DnmtTKO ou dans les ESCs traité avec 5-azacytidine, où cette marque est redistribuée vers des sites CpG déméthylés, au détriment des sites canoniques cibles de PRC2 (Cooper et al., 2014; Hagarman et al., 2013). Lors de la mise des ESCs dans le milieu 2i, il a été démontré qu'il y avait une perte considérable de H3K27me3 à des promoteurs de gènes (Marks et al., 2012). Par conséquent, alors que cette marque est réduite à des séquences uniques, nous montrons maintenant qu'elle est redéployée au PCH dans les ESCs en 2i, ce qui fait des ESCs un modèle intéressant pour étudier le dialogue entre la régulation PRC de l'expression génique et l'état épigénétique du PCH. Il a été montré précédemment que, en absence de méthylation de l'ADN, H3K9me3 et H3K27me3 coexistent au sein de l'hétérochromatine constitutive, mais dans des sub-domaines différents, ou dans des séquences qui ne se chevauchent pas (Cooper et al., 2014). Nous confirmons cette observation et montrons également que H3K27me3 remplace H3K9me3 dans un contexte de réduction de la méthylation de l'ADN. En effet, dans les ESCs DnmtTKO en sérum, la moitié de la population présente uniquement des enrichissements en H3K27me3, et dans des cellules qui gardent des foyers H3K9me3, ceux-ci ne chevauchent pas ceux de H3K27me3. Dans les ESCs en 2i, les clusters d'H3K9me3 ne sont pas dépendants de SUV39H1 et persistent également en l'absence d'enzymes Suv39h. Nous pensons qu'ils peuvent dépendre d'autres histone méthyltransférases (KMT), telles que SETDB1. Les données antérieures ont en effet

mis en évidence une réduction de H3K9me3 non seulement à l'euchromatine mais aussi au PCH en l'absence de SETDB1 (Mozzetta et al., 2015; Schultz et al., 2002).

4.2 EpiSC: une cellule pluripotente avec un état épigénétique somatique

Nous avons montré que, comme les cellules somatiques, les EpiSCs présentent des niveaux élevés de H3K9me3, de méthylation de l'ADN et de SUV39H1 aux régions riches en DAPI, et des chromocentres compactés, en accord avec des séquences transcriptionnellement inactives. En outre, l'expression de Nanog est réduite dans les EpiSCs et dans une étude très récente, Novo et al 2016 ont démontré que sa surexpression dans les EpiSCs induit la régulation positive de la transcription des satellites majeurs et la décompaction des clusters d'H3K9me3. Il a été montré que les DNMTs dans les ESCs sont recrutés au PCH suite à l'enrichissement d'H3K9me3, probablement grâce à l'interaction avec SUV39H/HP1 (Lehnertz et al., 2003; Fuks et al., 2003). Nos données montrent que dans les EpiSCs une telle interaction n'est pas nécessaire, car dans les cEpiSCs Suv39hdn la méthylation de l'ADN au niveau de séquences du PCH est rétabli. Le statut de de-répression des satellites majeurs en Suv39hdn cEpiSCs indiquent que H3K9me3 joue un rôle clé dans la répression de ces séquences et que la méthylation élevée de l'ADN ne suffit pas à maintenir un état répressif au PCH dans les EpiSCs.

4.3 Découplage de l'état épigénétique des ESCs avec régulation de la transcription des satellites

Nous avons montré que les ESCs en sérum présentent un état répressif à l'hétérochromatine constitutive avec un haut niveau de H3K9me3, méthylation de l'ADN et des chromocentres compactés. Paradoxalement la transcription des séquences satellites majeurs et mineurs est élevée (mais variable) à l'échelle globale. D'autre part, les ESCs en 2i présentent un fort enrichissement en H3K27me3 par rapport à H3K9me3, des niveaux réduits de 5-meC et des chromocentres déstructurés (voir le schéma récapitulatif de la Figure 9). De façon inattendue la transcription des satellites majeurs et mineurs est faible. En outre, alors que la répression de la transcription est dépendante de H3K9me3 dans les EpiSCs, nous montrons que ni la méthylation de l'ADN, ni H3K9me3 ne peuvent clairement moduler la transcription des satellites majeurs dans les ESCs. Par conséquent, nos données laissent ouverte la question de la nature des régulateurs (activateurs et répresseurs) de la transcription des séquences

satellites dans les ESCs. La transcription dans les ESC en sérum peut être régulée par des facteurs inconnus présents dans le sérum et non trouvés dans le milieu défini chimiquement utilisé dans la condition 2i. D'autre part, les inhibiteurs des voies GSK3 et MAPK peuvent peut-être eux-mêmes induire la répression de la transcription des séquences satellites.

En conclusion, cette étude apporte de nouvelles connaissances sur l'organisation de l'hétérochromatine dans les différents types des cellules pluripotentes. La marque épigénétique H3K27me3 est particulièrement dynamique et permet de distinguer l'état le plus naïf (comme les ESCs dans le milieu 2i) des autres cellules pluripotentes. Cette modification d'histone semble transiter depuis les séquences uniques (codantes et non-codantes) vers le PCH lors du transfert des ESCs à partir du milieu sérum vers le 2i, et montre une perte globale des EpiSCs.

REFERENCES

REFERENCES

- Acampora, D., Di Giovannantonio, L.G., and Simeone, A. (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.
- Agrawal, N., Dasaradhi, P.V.N., Mohmmmed, A., Malhotra, P., Bhatnagar, R.K., and Mukherjee, S.K. (2003). RNA Interference: Biology, Mechanism, and Applications. *Microbiol. Mol. Biol. Rev.* 67, 657–685.
- Aguirre-Lavin, T., Adenot, P., Bonnet-Garnier, A., Lehmann, G., Fleurot, R., Boulesteix, C., Debey, P., and Beaujean, N. (2012). 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.* 12, 1.
- Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. *PLoS ONE* 5, e10531.
- Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487–500.
- Arand, J., Spieler, D., Karius, T., Branco, M.R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., et al. (2012). In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases. *PLoS Genet* 8, e1002750.
- Aranda, S., Mas, G., and Croce, L.D. (2015). Regulation of gene transcription by Polycomb proteins. *Sci. Adv.* 1, e1500737.
- Artus, J., and Chazaud, C. (2014). A close look at the mammalian blastocyst: epiblast and primitive endoderm formation. *Cell. Mol. Life Sci.* 71, 3327–3338.
- Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* 17, 126–140.
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M., et al. (2006). Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* 8, 532–538.
- Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120–124.
- Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 461, 1292–1295.

- Batlle-Morera, L., Smith, A., and Nichols, J. (2008). Parameters influencing derivation of embryonic stem cells from murine embryos. *Genesis* 46, 758–767.
- Beaujean, N. (2014). Histone post-translational modifications in preimplantation mouse embryos and their role in nuclear architecture. *Mol. Reprod. Dev.* 81, 100–112.
- Becker, J.S., Nicetto, D., and Zaret, K.S. (2016). H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet.* 32, 29–41.
- Bedzhov, I., and Zernicka-Goetz, M. (2014). Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* 156, 1032–1044.
- Bedzhov, I., Graham, S.J.L., Leung, C.Y., and Zernicka-Goetz, M. (2014). Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. *Phil Trans R Soc B* 369, 20130538.
- Berge, D. ten, Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell Biol.* 13, 1070–1075.
- Bernardo, A.S., Faial, T., Gardner, L., Niakan, K.K., Ortmann, D., Senner, C.E., Callery, E.M., Trotter, M.W., Hemberger, M., Smith, J.C., et al. (2011). BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* 9, 144–155.
- Berndsen, C.E., and Denu, J.M. (2008). Catalysis and Substrate Selection by Histone/Protein Lysine Acetyltransferases. *Curr. Opin. Struct. Biol.* 18, 682–689.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.
- Biscotti, M.A., Canapa, A., Forconi, M., Olmo, E., and Barucca, M. (2015). Transcription of tandemly repetitive DNA: functional roles. *Chromosome Res.* 23, 463–477.
- Blackledge, N.P., Farcas, A.M., Kondo, T., King, H.W., McGouran, J.F., Hanssen, L.L.P., Ito, S., Cooper, S., Kondo, K., Koseki, Y., et al. (2014). Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. *Cell* 157, 1445–1459.
- Borgel, J., Guibert, S., Li, Y., Chiba, H., Schubeler, D., Sasaki, H., Forne, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* 42, 1093–1100.
- Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* 16, 516–528.

- Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., and Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* 35, 366–382.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536–2539.
- Bouzinba-Segard, H., Guais, A., and Francastel, C. (2006). Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci.* 103, 8709–8714.
- Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., et al. (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353.
- Branco, M.R., Ficz, G., and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.* 13, 7–13.
- Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195.
- Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of Enhancer Patterns in Transition from Naive to Primed Pluripotency. *Cell Stem Cell* 14, 838–853.
- Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., de la Rosa-Velazquez, I.A., van de Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M., et al. (2012). A transcription factor-based mechanism for mouse heterochromatin formation. *Nat. Struct. Mol. Biol.* 19, 1023–1030.
- Calado, R.T., and Dumitriu, B. (2013). Telomere dynamics in mice and humans. *Semin. Hematol.* 50, 165–174.
- Camus, A., Perea-Gomez, A., Moreau, A., and Collignon, J. (2006). Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev. Biol.* 295, 743–755.
- Cao, J. (2014). The functional role of long non-coding RNAs and epigenetics. *Biol. Proced. Online* 16, 11.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039–1043.
- Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655.
- Casanova, M., Pasternak, M., El Marjou, F., Le Baccon, P., Probst, A.V., and Almouzni, G. (2013). Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. *Cell Rep.* 4, 1156–1167.

- Chalut, K.J., Höpfler, M., Lautenschläger, F., Boyde, L., Chan, C.J., Ekpenyong, A., Martinez-Arias, A., and Guck, J. (2012). Chromatin Decondensation and Nuclear Softening Accompany Nanog Downregulation in Embryonic Stem Cells. *Biophys. J.* 103, 2060–2070.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450, 1230–1234.
- Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K.A.U., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S., and Ng, H.-H. (2013). Induction of a Human Pluripotent State with Distinct Regulatory Circuitry that Resembles Preimplantation Epiblast. *Cell Stem Cell* 13, 663–675.
- Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell* 10, 615–624.
- Chen, H., Aksoy, I., Gonnot, F., Osteil, P., Aubry, M., Hamela, C., Rognard, C., Hochard, A., Voisin, S., Fontaine, E., et al. (2015). Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat. Commun.* 6, 7095.
- Chen, T., Ueda, Y., Xie, S., and Li, E. (2002). A Novel Dnmt3a Isoform Produced from an Alternative Promoter Localizes to Euchromatin and Its Expression Correlates with Active de Novo Methylation. *J. Biol. Chem.* 277, 38746–38754.
- Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol. Cell. Biol.* 23, 5594–5605.
- Cheng, X., and Blumenthal, R.M. (2008). Mammalian DNA Methyltransferases: A Structural Perspective. *Struct. Lond. Engl.* 1993 16, 341–350.
- Chenoweth, J.G., McKay, R.D.G., and Tesar, P.J. (2010). Epiblast stem cells contribute new insight into pluripotency and gastrulation. *Dev. Growth Differ.* 52, 293–301.
- Condic, M.L. (2014). Totipotency: What It Is and What It Is Not. *Stem Cells Dev.* 23, 796–812.
- Cooper, S., Dienstbier, M., Hassan, R., Schermelleh, L., Sharif, J., Blackledge, N.P., De Marco, V., Elderkin, S., Koseki, H., Klose, R., et al. (2014). Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep.* 7, 1456–1470.
- Coronado, D., Godet, M., Bourillot, P.-Y., Tapponnier, Y., Bernat, A., Petit, M., Afanassieff, M., Markossian, S., Malashicheva, A., Iacone, R., et al. (2013). A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* 10, 118–131.
- Dambacher, S., Hahn, M., and Schotta, G. (2013). The compact view on heterochromatin. *Cell Cycle Georget. Tex* 12, 2925–2926.
- De Los Angeles, A., Loh, Y.-H., Tesar, P.J., and Daley, G.Q. (2012). Accessing naïve human pluripotency. *Curr. Opin. Genet. Dev.* 22, 272–282.

- Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes Dev* 25, 1010–1022.
- Déjardin, J. (2015). Switching between Epigenetic States at Pericentromeric Heterochromatin. *Trends Genet.* 31, 661–672.
- Denis, H., Ndlovu, 'Matladi N, and Fuks, F. (2011). Regulation of mammalian DNA methyltransferases: a route to new mechanisms. *EMBO Rep.* 12, 647–656.
- Dialynas, G.K., Terjung, S., Brown, J.P., Aucott, R.L., Baron-Luhr, B., Singh, P.B., and Georgatos, S.D. (2007). Plasticity of HP1 proteins in mammalian cells. *J. Cell Sci.* 120, 3415–3424.
- Dimitri, P., Corradini, N., Rossi, F., and Vernì, F. (2005). The paradox of functional heterochromatin. *BioEssays News Rev. Mol. Cell. Dev. Biol.* 27, 29–41.
- Ding, L., Paszkowski-Rogacz, M., Winzi, M., Chakraborty, D., Theis, M., Singh, S., Ciotta, G., Poser, I., Roguev, A., Chu, W.K., et al. (2015). Systems Analyses Reveal Shared and Diverse Attributes of Oct4 Regulation in Pluripotent Cells. *Cell Syst.* 1, 141–151.
- Do, D.V., Ueda, J., Messerschmidt, D.M., Lorthongpanich, C., Zhou, Y., Feng, B., Guo, G., Lin, P.J., Hossain, M.Z., Zhang, W., et al. (2013). A genetic and developmental pathway from STAT3 to the OCT4-NANOG circuit is essential for maintenance of ICM lineages in vivo. *Genes Dev.* 27, 1378–1390.
- Duggal, G., Warriar, S., Ghimire, S., Broekaert, D., Van der Jeught, M., Lierman, S., Deroo, T., Peelman, L., Van Soom, A., Cornelissen, R., et al. (2015). Alternative Routes to Induce Naïve Pluripotency in Human Embryonic Stem Cells. *Stem Cells Dayt. Ohio* 33, 2686–2698.
- Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D.G., Buetow, K.H., et al. (2008). Global Transcription in Pluripotent Embryonic Stem Cells. *Cell Stem Cell* 2, 437–447.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292.
- Falkenberg, K.J., and Johnstone, R.W. (2014). Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* 13, 673–691.
- Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. *Nature* 421, 448–453.
- Ferri, F., Bouzinba-Segard, H., Velasco, G., Hubé, F., and Francastel, C. (2009). Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.* 37, 5071–5080.
- Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* 11, 477–490.
- Festuccia, N., Osorno, R., Wilson, V., and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. *Curr. Opin. Genet. Dev.* 23, 504–511.

- Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473, 398–402.
- Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.-L., Walter, J., et al. (2013). FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency. *Cell Stem Cell* 13, 351–359.
- Frescas, D., Guardavaccaro, D., Kuchay, S.M., Kato, H., Poleshko, A., Basrur, V., Elenitoba-Johnson, K.S., Katz, R.A., and Pagano, M. (2008). KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state. *Cell Cycle* 7, 3539–3547.
- Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* 31, 2305–2312.
- Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504, 282–286.
- Galonska, C., Ziller, M.J., Karnik, R., and Meissner, A. (2015). Ground State Conditions Induce Rapid Reorganization of Core Pluripotency Factor Binding before Global Epigenetic Reprogramming. *Cell Stem Cell* 17, 462–470.
- García-Cao, M., O’Sullivan, R., Peters, A.H.F.M., Jenuwein, T., and Blasco, M.A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nat. Genet.* 36, 94–99.
- Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M.W.B., Pasque, V., Magnúsdóttir, E., and Surani, M.A. (2012). Epiblast Stem Cell-Based System Reveals Reprogramming Synergy of Germline Factors. *Cell Stem Cell* 10, 425–439.
- Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sternecker, J., Tesar, P., et al. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* 6, 215–226.
- Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* 13, 343–357.
- Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* 166, 493–505.
- Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136, 1063–1069.
- Guo, G., Pinello, L., Han, X., Lai, S., Shen, L., Lin, T.-W., Zou, K., Yuan, G.-C., and Orkin, S.H. (2016). Serum-Based Culture Conditions Provoke Gene Expression Variability in Mouse Embryonic Stem Cells as Revealed by Single-Cell Analysis. *Cell Rep.* 14, 956–965.

- Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H.D., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells. *Cell Stem Cell* 13, 360–369.
- Hackett, J.A., Dietmann, S., Murakami, K., Down, T.A., Leitch, H.G., and Surani, M.A. (2013). Synergistic Mechanisms of DNA Demethylation during Transition to Ground-State Pluripotency. *Stem Cell Rep.* 1, 518–531.
- Hagarman, J.A., Motley, M.P., Kristjansdottir, K., and Soloway, P.D. (2013). Coordinate Regulation of DNA Methylation and H3K27me3 in Mouse Embryonic Stem Cells. *PLOS ONE* 8, e53880.
- Hahn, M., Dambacher, S., Dulev, S., Kuznetsova, A.Y., Eck, S., Wörz, S., Sadic, D., Schulte, M., Mallm, J.-P., Mäyser, A., et al. (2013). Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev.* 27, 859–872.
- Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl. Acad. Sci.* 107, 9222–9227.
- Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: Contribution to chromatin structure and function. *Gene* 431, 1–12.
- Hattori, N., Niwa, T., Kimura, K., Helin, K., and Ushijima, T. (2013). Visualization of multivalent histone modification in a single cell reveals highly concerted epigenetic changes on differentiation of embryonic stem cells. *Nucleic Acids Res.* 41, 7231–7239.
- Hayashi, K., and Surani, M.A. (2009). Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro. *Development* 136, 3549–3556.
- He, J., Shen, L., Wan, M., Taranova, O., Wu, H., and Zhang, Y. (2013). Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat. Cell Biol.* 15, 373–384.
- He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333, 1303–1307.
- Hezroni, H., Sailaja, B.S., and Meshorer, E. (2011). Pluripotency-related, Valproic Acid (VPA)-induced Genome-wide Histone H3 Lysine 9 (H3K9) Acetylation Patterns in Embryonic Stem Cells. *J. Biol. Chem.* 286, 35977–35988.
- Hill, P.W.S., Amouroux, R., and Hajkova, P. (2014). DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. *Genomics* 104, 324–333.
- Houlard, M., Berlivet, S., Probst, A.V., Quivy, J.-P., Héry, P., Almouzni, G., and Gérard, M. (2006). CAF-1 Is Essential for Heterochromatin Organization in Pluripotent Embryonic Cells. *PLoS Genet.* 2, e181.

Huang, G., Ye, S., Zhou, X., Liu, D., and Ying, Q.-L. (2015). Molecular basis of embryonic stem cell self-renewal: from signaling pathways to pluripotency network. *Cell. Mol. Life Sci. CMLS* 72, 1741–1757.

Huang, K., Maruyama, T., and Fan, G. (2014). The Naive State of Human Pluripotent Stem Cells: A Synthesis of Stem Cell and Preimplantation Embryo Transcriptome Analyses. *Cell Stem Cell* 15, 410–415.

Huang, Y., Osorno, R., Tsakiridis, A., and Wilson, V. (2012). In Vivo Differentiation Potential of Epiblast Stem Cells Revealed by Chimeric Embryo Formation. *Cell Rep.* 2, 1571–1578.

Iqbal, K., Jin, S.G., Pfeifer, G.P., and Szabo, P.E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U A* 108, 3642–3647.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 333, 1300–1303.

Izzo, A., Kamieniarz, K., and Schneider, R. (2008). The histone H1 family: specific members, specific functions? *Biol. Chem.* 389, 333–343.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol. Cell. Biol.* 24, 8862–8871.

Jaco, I., Canela, A., Vera, E., and Blasco, M.A. (2008). Centromere mitotic recombination in mammalian cells. *J. Cell Biol.* 181, 885–892.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* 293, 1074–1080.

Jin, Q., Yu, L.-R., Wang, L., Zhang, Z., Kasper, L.H., Lee, J.-E., Wang, C., Brindle, P.K., Dent, S.Y.R., and Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J.* 30, 249–262.

Jones, P.A. (1999). The DNA methylation paradox. *Trends Genet. TIG* 15, 34–37.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13, 484–492.

Jones, D.O., Cowell, I.G., and Singh, P.B. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *BioEssays* 22, 124–137.

Joshi, O., Wang, S.-Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. *Cell Stem Cell* 17, 748–757.

Josling, G.A., Selvarajah, S.A., Petter, M., and Duffy, M.F. (2012). The Role of Bromodomain Proteins in Regulating Gene Expression. *Genes* 3, 320–343.

Jouneau, A., Ciaudo, C., Sismeiro, O., Brochard, V., Jouneau, L., Vandormael-Pournin, S., Coppee, J.Y., Zhou, Q., Heard, E., Antoniewski, C., et al. (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. *RNA* 18, 253–264.

Kalkan, T., and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. *Phil Trans R Soc B* 369, 20130540.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 19, 489–501.

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13, 424.

Kass, S.U., Pruss, D., and Wolffe, A.P. (1997). How does DNA methylation repress transcription? *Trends Genet. TIG* 13, 444–449.

Kaufman, P.D., Kobayashi, R., Kessler, N., and Stillman, B. (1995). The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. *Cell* 81, 1105–1114.

Kim, H., Wu, J., Ye, S., Tai, C.-I., Zhou, X., Yan, H., Li, P., Pera, M., and Ying, Q.-L. (2013). Modulation of β -catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. *Nat. Commun.* 4.

Knutson, S.K., Kawano, S., Minoshima, Y., Warholic, N.M., Huang, K.-C., Xiao, Y., Kadowaki, T., Uesugi, M., Kuznetsov, G., Kumar, N., et al. (2014). Selective Inhibition of EZH2 by EPZ-6438 Leads to Potent Antitumor Activity in EZH2-Mutant Non-Hodgkin Lymphoma. *Mol. Cancer Ther.* 13, 842–854.

Koh, F.M., Sachs, M., Guzman-Ayala, M., and Ramalho-Santos, M. (2010). Parallel gateways to pluripotency: open chromatin in stem cells and development. *Curr. Opin. Genet. Dev.* 20, 492–499.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., et al. (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8, 200–213.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502, 472–479.

Kojima, Y., Kaufman-Francis, K., Studdert, J.B., Steiner, K.A., Power, M.D., Loebel, D.A.F., Jones, V., Hor, A., de Alencastro, G., Logan, G.J., et al. (2014). The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak. *Cell Stem Cell* 14, 107–120.

Kolodziejczyk, A.A., Kim, J.K., Tsang, J.C.H., Ilicic, T., Henriksson, J., Natarajan, K.N., Tuck, A.C., Gao, X., Bühler, M., Liu, P., et al. (2015). Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell Stem Cell* 17, 471–485.

- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.
- Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (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.
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
- Lagger, G., O’Carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R., Jenuwein, T., et al. (2002). Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *EMBO J.* 21, 2672–2681.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol. CB* 13, 1192–1200.
- Leitch, H.G., McEwen, K.R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J.G., Smith, A., et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* 20, 311–316.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.
- Li, Y., Seah, M.K.Y., and O’Neill, C. (2016). Mapping global changes in nuclear cytosine base modifications in the early mouse embryo. *Reprod. Camb. Engl.* 151, 83–95.
- Lin, Y., Li, X.-Y., Willis, A.L., Liu, C., Chen, G., and Weiss, S.J. (2014). Snail1-dependent control of embryonic stem cell pluripotency and lineage commitment. *Nat. Commun.* 5, 3070.
- Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* 38, 431–440.
- López-Flores, I., and Garrido-Ramos, M.A. (2012). The Repetitive DNA Content of Eukaryotic Genomes. In *Genome Dynamics*, M.A. Garrido-Ramos, ed. (Basel: S. KARGER AG), pp. 1–28.
- Lu, J., and Gilbert, D.M. (2007). Proliferation-dependent and cell cycle-regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* 179, 411–421.
- Lu, L.Y., Wood, J.L., Minter-Dykhouse, K., Ye, L., Saunders, T.L., Yu, X., and Chen, J. (2008). Polo-like kinase 1 is essential for early embryonic development and tumor suppression. *Mol Cell Biol* 28, 6870–6876.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.

- Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* 30, 329–334.
- Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343–349.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 Maintain Repressive Chromatin through Different Mechanisms. *Mol. Cell* 32, 503–518.
- Marks, H., Kalkan, T., Menafrá, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis Stewart, A., Smith, A., et al. (2012). The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. *Cell* 149, 590–604.
- Martens, J.H., O’Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* 24, 800–812.
- Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* 78.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838–849.
- Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., et al. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.* 9, 625–635.
- Mattick, J.S. (2007). A new paradigm for developmental biology. *J. Exp. Biol.* 210, 1526–1547.
- Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. *Hum. Mol. Genet.* 15, R17–R29.
- Mattioli, F., D’Arcy, S., and Luger, K. (2015). The right place at the right time: chaperoning core histone variants. *EMBO Rep.* 16, 1454–1466.
- Meissner, A. (2010). Epigenetic modifications in pluripotent and differentiated cells. *Nat. Biotechnol.* 28, 1079–1088.
- Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. *Dev. Cell* 10, 105–116.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.-K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553–560.
- Millanes-Romero, A., Herranz, N., Perrera, V., Iturbide, A., Loubat-Casanovas, J., Gil, J., Jenuwein, T., García de Herreros, A., and Peiró, S. (2013). Regulation of heterochromatin

transcription by Snail1/LOXL2 during epithelial-to-mesenchymal transition. *Mol. Cell* 52, 746–757.

Mozzetta, C., Boyarchuk, E., Pontis, J., and Ait-Si-Ali, S. (2015). Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat. Rev. Mol. Cell Biol.* 16, 499–513.

Müller, S., Montes de Oca, R., Lacoste, N., Dingli, F., Loew, D., and Almouzni, G. (2014). Phosphorylation and DNA binding of HJURP determine its centromeric recruitment and function in CenH3(CENP-A) loading. *Cell Rep.* 8, 190–203.

Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12135–12140.

Nichols, J., and Smith, A. (2009). Naive and Primed Pluripotent States. *Cell Stem Cell* 4, 487–492.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391.

Nichols, J., Jones, K., Phillips, J.M., Newland, S.A., Roode, M., Mansfield, W., Smith, A., and Cooke, A. (2009). Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. *Nat. Med.* 15, 814–818.

Nimura, K., Ishida, C., Koriyama, H., Hata, K., Yamanaka, S., Li, E., Ura, K., and Kaneda, Y. (2006). Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. *Genes Cells* 11, 1225–1237.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 12, 2048–2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* 24, 372–376.

Novo, C.L., Tang, C., Ahmed, K., Djuric, U., Fussner, E., Mullin, N.P., Morgan, N.P., Hayre, J., Sienerth, A.R., Elderkin, S., et al. (2016). The pluripotency factor Nanog regulates pericentromeric heterochromatin organization in mouse embryonic stem cells. *Genes Dev.*

O’Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The Polycomb-Group Gene Ezh2 Is Required for Early Mouse Development. *Mol. Cell. Biol.* 21, 4330–4336.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.

Orkin, S.H., and Hochedlinger, K. (2011). Chromatin Connections to Pluripotency and Cellular Reprogramming. *Cell* 145, 835–850.

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development* 139, 2288–2298.

Padilla-Nash, H.M., Barenboim-Stapleton, L., Difilippantonio, M.J., and Ried, T. (2007). Spectral karyotyping analysis of human and mouse chromosomes. *Nat. Protoc.* 1, 3129–3142.

Pauler, F.M., Sloane, M.A., Huang, R., Regha, K., Koerner, M.V., Tamir, I., Sommer, A., Aszodi, A., Jenuwein, T., and Barlow, D.P. (2008). 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.

Pauloin, A., Adenot, P., Hue-Beauvais, C., and Chanat, E. (2016). The perilipin-2 (adipophilin) coat of cytosolic lipid droplets is regulated by an Arf1-dependent mechanism in HC11 mouse mammary epithelial cells. *Cell Biol. Int.* 40, 143–155.

Perea-Gomez, A., Vella, F.D., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H., et al. (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Dev Cell* 3, 745–756.

Peters, A.H., Kubicek, S., Mechtler, K., O’Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y., et al. (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* 12, 1577–1589.

Peters, A.H.F.M., O’Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schöfer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A., et al. (2001). Loss of the Suv39h Histone Methyltransferases Impairs Mammalian Heterochromatin and Genome Stability. *Cell* 107, 323–337.

Plohl, M., Meštrović, N., and Mravinac, B. (2014). Centromere identity from the DNA point of view. *Chromosoma* 123, 313–325.

Probst, A.V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P., and Almouzni, G. (2010). A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. *Dev. Cell* 19, 625–638.

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A.P., Koseki, H., Orkin, S.H., et al. (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* 40, 411–420.

Quivy, J.-P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y., and Almouzni, G. (2004). A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J.* 23, 3516–3526.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci.* 97, 5237–5242.

Renfree, M.B., and Shaw, G. (2000). Diapause. *Annu. Rev. Physiol.* 62, 353–375.

- Rhodes, D. (1997). Chromatin structure. The nucleosome core all wrapped up. *Nature* 389, 231, 233.
- Rop, V.D., Padeganeh, A., and Maddox, P.S. (2012). CENP-A: the key player behind centromere identity, propagation, and kinetochore assembly. *Chromosoma* 121, 527–538.
- Rossant, J. (2008). Stem Cells and Early Lineage Development. *Cell* 132, 527–531.
- Rossant, J., and Tam, P.P.L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Dev. Camb. Engl.* 136, 701–713.
- Routh, A., Sandin, S., and Rhodes, D. (2008). Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci.* 105, 8872–8877.
- Rugg-Gunn, P.J., Cox, B.J., Ralston, A., and Rossant, J. (2010). Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc Natl Acad Sci U A* 107, 10783–10790.
- Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N., and Derynck, R. (2013). Smad2 Is Essential for Maintenance of the Human and Mouse Primed Pluripotent Stem Cell State. *J. Biol. Chem.* 288, 18546–18560.
- Sakaue, M., Ohta, H., Kumaki, Y., Oda, M., Sakaide, Y., Matsuoka, C., Yamagiwa, A., Niwa, H., Wakayama, T., and Okano, M. (2010). DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr Biol* 20, 1452–1457.
- Saksouk, N., Barth, T.K., Ziegler-Birling, C., Olova, N., Nowak, A., Rey, E., Mateos-Langerak, J., Urbach, S., Reik, W., Torres-Padilla, M.-E., et al. (2014). Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. *Mol. Cell* 56, 580–594.
- Saksouk, N., Simboeck, E., and Déjardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* 8, 3.
- Salvaing, J., Aguirre-Lavin, T., Boulesteix, C., Lehmann, G., Debey, P., and Beaujean, N. (2012). 5-Methylcytosine and 5-hydroxymethylcytosine spatiotemporal profiles in the mouse zygote. *PloS One* 7, e38156.
- Sana, J., Faltejskova, P., Svoboda, M., and Slaby, O. (2012). Novel classes of non-coding RNAs and cancer. *J. Transl. Med.* 10, 103.
- Santenard, A., and Torres-Padilla, M.E. (2009). Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. *Epigenetics* 4, 80–84.
- Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J., and Torres-Padilla, M.-E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat. Cell Biol.* 12, 853–862.
- Santos, F., Peters, A.H., Otte, A.P., Reik, W., and Dean, W. (2005). Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* 280, 225–236.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* 18, 1251–1262.

Schotta, G., Sengupta, R., Kubicek, S., Malin, S., Kauer, M., Callén, E., Celeste, A., Pagani, M., Opravil, S., De La Rosa-Velazquez, I.A., et al. (2008). A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* 22, 2048–2061.

Schrader, E.K., Harstad, K.G., and Matouschek, A. (2009). Targeting proteins for degradation. *Nat. Chem. Biol.* 5, 815–822.

Schuettengruber, B., Martinez, A.-M., Iovino, N., and Cavalli, G. (2011). Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* 12, 799–814.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 16, 919–932.

Schulz, E.G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., Borensztein, M., Saitou, M., Blüthgen, N., and Heard, E. (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.

Scott, K.C. (2013). Transcription and ncRNAs: at the cent(rome)re of kinetochore assembly and maintenance. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 21, 643–651.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450, 908–912.

Shen, X., Liu, Y., Hsu, Y.-J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.-C., and Orkin, S.H. (2008). EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. *Mol. Cell* 32, 491–502.

Shi, Y., and Massagué, J. (2003). Mechanisms of TGF- β Signaling from Cell Membrane to the Nucleus. *Cell* 113, 685–700.

Sims, R.J., Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet. TIG* 19, 629–639.

Singer, Z.S., Yong, J., Tischler, J., Hackett, J.A., Altinok, A., Surani, M.A., Cai, L., and Elowitz, M.B. (2014). Dynamic Heterogeneity and DNA Methylation in Embryonic Stem Cells. *Mol. Cell* 55, 319–331.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336, 688–690.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484, 339–344.

Sokol, S.Y. (2011). Maintaining embryonic stem cell pluripotency with Wnt signaling. *Development* 138, 4341–4350.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. *Microbiol. Mol. Biol. Rev. MMBR* 64, 435–459.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45.

Strathdee, G., and Brown, R. (2002). Aberrant DNA methylation in cancer: potential clinical interventions. *Expert Rev. Mol. Med.* 4, 1–17.

Sumi, T., Oki, S., Kitajima, K., and Meno, C. (2013). Epiblast Ground State Is Controlled by Canonical Wnt/ β -Catenin Signaling in the Postimplantation Mouse Embryo and Epiblast Stem Cells. *PLoS ONE* 8, e63378.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935.

Tai, C.-I., Schulze, E.N., and Ying, Q.-L. (2014). Stat3 signaling regulates embryonic stem cell fate in a dose-dependent manner. *Biol. Open* 3, 958–965.

Tajul-Arifin, K., Teasdale, R., Ravasi, T., Hume, D.A., and Mattick, J.S. (2003). Identification and Analysis of Chromodomain-Containing Proteins Encoded in the Mouse Transcriptome. *Genome Res.* 13, 1416–1429.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* 131, 861–872.

Takaoka, K., and Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. *Dev. Camb. Engl.* 139, 3–14.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficuz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting Transcription Factor Control Circuitry toward Ground-State Pluripotency in Human. *Cell* 158, 1254–1269.

Tamaru, H. (2010). Confining euchromatin/heterochromatin territory: jumonji crosses the line. *Genes Dev.* 24, 1465–1478.

Terranova, R., Agherbi, H., Boned, A., Meresse, S., and Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6629–6634.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196–199.

Tessarz, P., and Kouzarides, T. (2014). Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* 15, 703–708.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell* 15, 471–487.

Thijssen, P.E., Ito, Y., Grillo, G., Wang, J., Velasco, G., Nitta, H., Unoki, M., Yoshihara, M., Suyama, M., Sun, Y., et al. (2015). Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat. Commun.* 6, 7870.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145–1147.

Tosolini, M., and Jouneau, A. (2016a). Acquiring Ground State Pluripotency: Switching Mouse Embryonic Stem Cells from Serum/LIF Medium to 2i/LIF Medium. *Methods Mol. Biol. Clifton NJ* 1341, 41–48.

Tosolini, M., and Jouneau, A. (2016b). From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. *Methods Mol. Biol. Clifton NJ* 1341, 209–216.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* 135, 909–918.

Trojer, P., and Reinberg, D. (2007). Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell* 28, 1–13.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R., et al. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* 11, 805–814.

Valamehr, B., Robinson, M., Abujarour, R., Rezner, B., Vranceanu, F., Le, T., Medcalf, A., Lee, T.T., Fitch, M., Robbins, D., et al. (2014). Platform for Induction and Maintenance of Transgene-free hiPSCs Resembling Ground State Pluripotent Stem Cells. *Stem Cell Rep.* 2, 366–381.

Vassena, R., Boué, S., González-Roca, E., Aran, B., Auer, H., Veiga, A., and Belmonte, J.C.I. (2011). Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. *Dev. Camb. Engl.* 138, 3699–3709.

- Veillard, A.-C., Maruotti, J., and Jouneau, A. (2012). Reprogramming and pluripotency of epiblast stem cells. In *Stem Cells and Cancer Stem Cells*, M.A. Hayat, ed. (Springer Science+Business Media), pp. 133–146.
- Veillard, A.C., Marks, H., Bernardo, A.S., Jouneau, L., Laloe, D., Boulanger, L., Kaan, A., Brochard, V., Tosolini, M., Pedersen, R., et al. (2014). Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. *Stem Cells Dev.* 23, 2014–2029.
- Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters. *Genes Dev.* 27, 1318–1338.
- Voncken, J.W., Roelen, B.A.J., Roefs, M., Vries, S. de, Verhoeven, E., Marino, S., Deschamps, J., and Lohuizen, M. van (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proc. Natl. Acad. Sci.* 100, 2468–2473.
- von Meyenn, F., Iurlaro, M., Habibi, E., Liu, N.Q., Salehzadeh-Yazdi, A., Santos, F., Petrini, E., Milagre, I., Yu, M., Xie, Z., et al. (2016). Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* 0.
- Wade, P.A., and Wolffe, A.P. (2001). ReCoGnizing methylated DNA. *Nat. Struct. Mol. Biol.* 8, 575–577.
- Walter, M., Teissandier, A., Pérez-Palacios, R., and Bourc’his, D. (2016). An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* 5, e11418.
- Wang, H., and Dey, S.K. (2006). Roadmap to embryo implantation: clues from mouse models. *Nat. Rev. Genet.* 7, 185–199.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873–878.
- Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., et al. (2014). Derivation of naive human embryonic stem cells. *Proc. Natl. Acad. Sci.* 111, 4484–4489.
- Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* 17, 155–169.
- Weiner, A., Lara-Astiaso, D., Krupalnik, V., Gafni, O., David, E., Winter, D.R., Hanna, J.H., and Amit, I. (2016). Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. *Nat. Biotechnol.* 34, 953–961.
- Weissgerber, T.L., Milic, N.M., Winham, S.J., and Garovic, V.D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol.* 13, e1002128.
- van de Werken, C., van der Heijden, G.W., Eleveld, C., Teeuwssen, M., Albert, M., Baarends, W.M., Laven, J.S.E., Peters, A.H.F.M., and Baart, E.B. (2014). Paternal heterochromatin formation in human embryos is H3K9/HP1 directed and primed by sperm-derived histone modifications. *Nat. Commun.* 5, 5868.

- Wongtawan, T., Taylor, J.E., Lawson, K.A., Wilmut, I., and Pennings, S. (2011). Histone H4K20me3 and HP1 α are late heterochromatin markers in development, but present in undifferentiated embryonic stem cells. *J Cell Sci* 124, 1878–1890.
- Wray, J., Kalkan, T., and Smith, A.G. (2010). The ground state of pluripotency. *Biochem. Soc. Trans.* 38, 1027–1032.
- Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (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.
- Wu, J., Okamura, D., Li, M., Suzuki, K., Luo, C., Ma, L., He, Y., Li, Z., Benner, C., Tamura, I., et al. (2015). An alternative pluripotent state confers interspecies chimaeric competency. *Nature* 521, 316–321.
- Wu Ct, null, and Morris, J.R. (2001). Genes, genetics, and epigenetics: a correspondence. *Science* 293, 1103–1105.
- Xie, D., Chen, C.-C., Ptaszek, L.M., Xiao, S., Cao, X., Fang, F., Ng, H.H., Lewin, H.A., Cowan, C., and Zhong, S. (2010). Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome Res.* 20, 804–815.
- Yamagata, K., Yamazaki, T., Miki, H., Ogonuki, N., Inoue, K., Ogura, A., and Baba, T. (2007). Centromeric DNA hypomethylation as an epigenetic signature discriminates between germ and somatic cell lineages. *Dev. Biol.* 312, 419–426.
- Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. *Cell Stem Cell* 12, 368–382.
- Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1131–1139.
- Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115, 281–292.
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519–523.
- Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. *Cell Res.* 21, 564–578.
- Zhang, Y. (2003). Transcriptional regulation by histone ubiquitination and deubiquitination. *Genes Dev.* 17, 2733–2740.

Zhang, H., Gayen, S., Xiong, J., Zhou, B., Shanmugam, A.K., Sun, Y., Karatas, H., Liu, L., Rao, R.C., Wang, S., et al. (2016). MLL1 Inhibition Reprograms Epiblast Stem Cells to Naive Pluripotency. *Cell Stem Cell*.

Zhou, H., Li, W., Zhu, S., Joo, J.Y., Do, J.T., Xiong, W., Kim, J.B., Zhang, K., Scholer, H.R., and Ding, S. (2010). Conversion of Mouse Epiblast Stem Cells to an Earlier Pluripotency State by Small Molecules. *J. Biol. Chem.* 285, 29676–29680.

Zylicz, J.J., Dietmann, S., Günesdogan, U., Hackett, J.A., Cougot, D., Lee, C., and Surani, M.A. (2015). Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *eLife* 4, e09571.

APPENDIXES

APPENDIXES

Paper 1: Veillard, A.C., Marks, H., Bernardo, A.S., Jouneau, L., Laloe, D., Boulanger, L., Kaan, A., Brochard, V., **Tosolini, M.**, Pedersen, R., et al. (2014). Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. *Stem Cells Dev.* 23, 2014–2029.

Paper 2: Tosolini, M., and Jouneau, A. (2016). Acquiring Ground State Pluripotency: Switching Mouse Embryonic Stem Cells from Serum/LIF Medium to 2i/LIF Medium. *Methods Mol. Biol.* Clifton NJ 1341, 41–48.

Paper 3: Tosolini, M., and Jouneau, A. (2016). From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. *Methods Mol. Biol.* Clifton NJ 1341, 209–216.

Paper 4: Under revision in Scientific Reports. **Matteo Tosolini**, Vincent Brochard, Pierre Adenot, Martine Chebrou, Giacomo Grillo, Violette Navia, François Piumi, Nathalie Beaujean, Claire Francastel, Amelie Bonnet-Garnier, and Alice Jouneau Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells

Stable Methylation at Promoters Distinguishes Epiblast Stem Cells from Embryonic Stem Cells and the In Vivo Epiblasts

Anne-Clémence Veillard,¹ Hendrik Marks,² Andreia Sofia Bernardo,^{3,4} Luc Jouneau,¹ Denis Laloë,⁵ Laurent Boulanger,¹ Anita Kaan,² Vincent Brochard,¹ Matteo Tosolini,¹ Roger Pedersen,⁴ Henk Stunnenberg,² and Alice Jouneau¹

Embryonic Stem Cells (ESCs) and Epiblast Stem Cells (EpiSCs) are the *in vitro* representatives of naïve and primed pluripotency, respectively. It is currently unclear how their epigenomes underpin the phenotypic and molecular characteristics of these distinct pluripotent states. Here, we performed a genome-wide comparison of DNA methylation between ESCs and EpiSCs by MethylCap-Seq. We observe that promoters are preferential targets for methylation in EpiSC compared to ESCs, in particular high CpG island promoters. This is in line with upregulation of the *de novo* methyltransferases Dnmt3a1 and Dnmt3b in EpiSC, and downregulation of the demethylases Tet1 and Tet2. Remarkably, the observed DNA methylation signature is specific to EpiSCs and differs from that of their *in vivo* counterpart, the postimplantation epiblast. Using a subset of promoters that are differentially methylated, we show that DNA methylation is established within a few days during *in vitro* outgrowth of the epiblast, and also occurs when ESCs are converted to EpiSCs *in vitro*. Once established, this methylation is stable, as ES-like cells obtained by *in vitro* reversion of EpiSCs display an epigenetic memory that only extensive passaging and sub-cloning are able to almost completely erase.

Introduction

TWO KINDS OF PLURIPOTENT STEM CELLS can be captured *ex vivo* from the mouse embryo: embryonic stem cells (ESCs) are derived from the inner cell mass (ICM) of the blastocyst, whereas epiblast stem cells (EpiSCs) are isolated from the late epiblast of postimplantation embryos. Although both express the core triad of transcription factors Oct4/Sox2/Nanog, some other pluripotency factors identified in ESCs are absent in EpiSCs, such as *Esrrb*, *Klf4*, *Rex1/Zfp42*, and *Dppa3/Stella* [1,2]. Moreover, transcriptome comparisons have indicated that EpiSCs may be closer to the postimplantation epiblast, whereas ESCs share more characteristics with ICM cells [3]. Indeed, EpiSCs express late epiblast markers such as *Nodal*, *Fgf5*, *Brachyury (T)*, or *Cer1*, which are low or absent in ESCs [4]. These contrasting signatures suggest these cell types represent two different states of pluripotency, naïve for ESCs and primed for EpiSCs [5]. In contrast to ESCs, EpiSCs are unable to form chimeras fol-

lowing injection into blastocysts [3,6]. However, they can contribute, at least to some extent, to embryo development if injected in the postimplantation epiblast [7].

Different signaling pathways control self-renewal of these pluripotent states: ESCs require LIF and BMP4, while EpiSCs are dependent on FGF2 and Activin/Nodal [3,8]. By switching between the appropriate culture conditions for each cell type, ESCs can be readily converted into EpiSCs, whereas EpiSCs can also be reverted into naïve ES-like cells *in vitro* albeit with low efficiencies [9–11]. Such interconversion abilities provide new avenues to study the relationships between the two states of pluripotency and have elicited the notion of an “epigenetic barrier” separating ESCs from EpiSCs, since reprogramming EpiSCs into naïve ESCs is an inefficient and long process. In particular, *de novo* DNA methylation, which takes place during epiblast development [12,13], may be a constituent of this barrier as it is in somatic cell reprogramming [14]. *De novo* DNA methylation is catalyzed by the *de novo* methyltransferase 3

¹INRA, UMR1198 Biologie du Développement et Reproduction, Jouy-en-Josas, France.

²Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences (RIMLS), Radboud University, Nijmegen, The Netherlands.

³Systems Biology Division, MRC National Institute for Medical Research, Mill Hill, London, United Kingdom.

⁴Laboratory for Regenerative Medicine, Department of Surgery, University of Cambridge, Cambridge, United Kingdom.

⁵INRA, UMR GABI, Jouy-en-Josas, France.

(Dnmt3) enzymes [15]. Three have been isolated in mammals: Dnmt3a and Dnmt3b are catalytically active, while Dnmt3l is a cofactor for both. In the epiblast, Dnmt3b is the first to be expressed at E3.5 while Dnmt3a expression starts 1 day later [16–18]. Dnmt3l is expressed transiently in the epiblast, between E4.5 and E6.5 [16,18]. The ICM of E3.5 blastocyst is globally hypomethylated, while a massive wave of de novo methylation occurs between the epiblast stages E3.5 and E6.5 [13]. Interestingly, the methylome of ESC is shaped by their culture conditions: ESCs cultured in the presence of the two kinase inhibitors, inhibiting Gsk3 and FGF signaling, respectively, are globally hypomethylated and resemble ICM cells. ESCs cultured using serum conditions display a methylation profile more similar to that of the early postimplantation epiblast [12,19,20]. In contrast, little is known about the methylome of EpiSCs. Quantitatively, the global level of 5-methylcytosine in EpiSCs was found to be similar to that of ESC cultured in serum [20,21]. A recent study by Senner et al. [21] showed that four types of stem cells derived from the mouse embryo, either extra-embryonic or embryonic, contained a unique DNA methylation signature. However, an in-depth comparison of the methylome of ESCs and EpiSCs is currently lacking.

Here, we compare the patterns of DNA methylation in EpiSCs and ESCs and observed a clear bias toward promoter-associated hypermethylation in EpiSCs. By following the kinetics of methylation during ESC to EpiSC conversion, we show that de novo methylation seems to occur very rapidly and concomitant to the molecular switch. Conversely, reversion of EpiSC into ES-like cells shows that the reprogramming of methylation at the promoters is very slow and incomplete, suggesting the persistence of an epigenetic memory. Finally, a comparison of EpiSCs and late epiblast cells reveals that the *in vitro* and “*in embryo*” cells show a remarkably different promoter methylation profile.

Materials and Methods

Preparation of samples for sequencing

MethylCap. Genomic DNA was sonicated to generate 300 bp fragments on average. MethylCap was performed using the IP-STAR robot (Diagenode) as described before [22]. In short, 1 µg DNA was incubated with paramagnetic beads coated with the MBD domain of MeCP2 fused to GST. After washing with 200, 400, and 500 mM NaCl, the bound methylated DNA was eluted in two fractions using 600 and 800 mM NaCl, respectively. Twenty nanogram of DNA eluates was prepared for sequencing.

Double-stranded cDNA synthesis. Total RNA was isolated with TRIzol (Invitrogen) according to the manufacturer's recommendations. One hundred microgram total RNA was subjected to two rounds of poly(A) selection (Oligotex mRNA Mini Kit; QIAGEN), followed by DNaseI treatment (QIAGEN). About 100–200 ng mRNA was fragmented by hydrolysis (5× fragmentation buffer: 200 mM Tris acetate, pH8.2, 500 mM potassium acetate, and 150 mM magnesium acetate) at 94°C for 90 s and purified (RNAeasy Minelute Kit; QIAGEN). cDNA was synthesized using 5 µg random hexamers by Superscript III Reverse Transcriptase (Invitrogen). Double-strand cDNA synthesis was performed in second strand buffer (Invitrogen) according to the manu-

facturer's recommendations and purified (Minelute Reaction Cleanup Kit; QIAGEN).

Sequencing

DNA or cDNA samples were prepared for sequencing by end repair of 20 ng DNA as measured by Qubit (Invitrogen). Adaptors were ligated to DNA fragments, followed by size selection (~300 bp) and 14 cycles of PCR amplification. Integrity of DNA libraries was confirmed by running the products on a Bioanalyzer (BioRad). Cluster generation and sequencing (36 bp) was performed with the Illumina Genome Analyzer IIx (GAIIx) (MethylCap-seq) or HiSeq (RNA-seq) platform according to standard Illumina protocols. Initial data processing, base calling, and alignment to the mouse reference genome was performed using the Illumina Analysis Pipeline allowing one mismatch. Only tags aligning to one position on the genome were considered for further analysis. For RNA-seq, further analysis was performed with the 36 bp aligned sequence. For MethylCap-seq, the uniquely mapped sequence reads were directionally extended to 300 bp, the estimated median length of the original DNA library. If multiple tags were mapped on the same genomic position, only one was included for further analysis. Mapped reads from the initial 600 and 800 mM NaCl eluate libraries were combined. For both RNA-seq and MethylCap-seq data were converted to Browser Extensible Data files for downstream analysis. To compensate for differences in sequencing depth and mapping efficiency among samples, the total number of unique reads of each sample was uniformly equalized relative to the sample with the lowest number of sequence reads, allowing quantitative comparisons. Wiggle (WIG) files for viewing the data in the UCSC Genome Browser were generated from the normalized files. All sequencing analyses were conducted based on the *Mus musculus* NCBI m37 genome assembly (MM9; assembly July 2007). Supplementary Table S1 summarizes the sequencing output.

MethylCap-seq analyses

Peak calling. Data analysis was performed using in-house generated scripts written in LINUX shell, Perl, and R. Enriched regions (peaks) were called on the basis of a Poisson distribution of overlapping sequence reads within a dynamic window. A false discovery rate (FDR) was calculated relative to the total covered sequence, and peaks with an FDR of $\leq 1 \times 10^{-6}$ were selected. All peaks from the four samples (three EpiSCs and one ESC) were merged. Per sample, the number of normalized sequence reads overlapping each region (peak) of interest was calculated, which is referred to as read density and used as a measure of DNA methylation.

Annotation. Methylated peaks were annotated according to their localization in the genome, as intron, exon, promoter (−900 to +400 bp around the gene start), or intergenic based on the Ensembl release 66 (MM9 assembly). As peaks often span different genomic regions, we used the summit genomic coordinate of each peak as representative. Due to the presence of overlapping transcripts in the mouse genome, about 5% of the peaks were annotated to multiple genes. Therefore, peaks were annotated with (at least one) gene name and sequence type (exon, intron, and promoter). Nonassigned peaks were considered as intergenic.

Gene Ontology and KEGG analysis was performed using DAVID (<http://david.abcc.ncifcrf.gov>) [23,24]. Further data analysis was performed using IPA (Ingenuity® Systems, www.ingenuity.com).

RNA-seq analysis

To obtain RNA-seq gene expression values (RPKM), we used Genomatix (www.genomatix.de).

Statistical analysis of MethylCap seq data

Read counts for all peaks and transcripts were normalized using the normalization procedure used in the Bioconductor package “DESeq” [25]. We used these normalized counts to perform a hierarchical clustering analysis of methylome samples using the distance function 1-*c*, where *c* is the correlation coefficient, and the Ward linkage method.

To study the relation between DNA methylation and expression, we built a reduced dataset where methylation peaks corresponding to promoters were combined with read counts for the corresponding transcripts. The following thresholds were used for including genes: methylation read density >0 for either the ESC sample or for at least 2 EpiSC samples.

After log2 transformation of normalized counts, we used the R package “flexclust” [26] to group methylation and expression profiles in clusters. Only clusters having at least 10 profiles and a Pearson correlation coefficient greater than 0.9 (with the center profile of the cluster; radius—that is maximum distance to the cluster center profile = 0.1), were included.

Cell lines and culture conditions

Derivation of the 129S2 ESC line was performed as previously described [27]. The rESCs (129S2 rESCs) and ESCs (129S2 ESCs, 129B6 ESCs, and R1) were grown on irradiated mouse embryonic fibroblasts in medium containing DMEM with either 15% serum or (for 129S2 ESCs) 20% KSR (Invitrogen), 0.1 mM β-mercaptoethanol and 1,000 U/mL LIF (ESGRO; Millipore) and plated feeder-free on gelatin-coated dishes for two passages before collection. EpiSCs (129S2 EpiSCs, EpiSC1, 2, and 3 described in Maruotti, 2010 and 129B6 EpiSCs) and cEpiSCs (129S2 cEpiSCs) were grown in serum-free medium (CDM) with FGF2 (12 ng/mL; R&D) and Activin A (20 ng/mL; R&D) on serum-coated dishes as previously described [3]. Conversion of 129S2 ESCs was performed as previously described [9,28]. In summary, the ESCs were trypsinized and 1.5×10^6 – 3×10^6 cells were seeded in 35 mm, serum-coated dishes in CDM+FGF2 and Activin. At day 4, cells were detached with collagenase-II (Sigma) and replated without dilution. This first passage promotes the appearance of flat colonies with typical EpiSC morphology. Converted cells were then cultured for several passages before harvesting. For reversion, EpiSCs were passaged onto mouse irradiated feeders in the presence of ESC medium as described. After 7 days, cells were trypsinized and passaged as ESCs for at least five passages.

Epiblast dissection

E6.5 and E7 epiblasts were dissected from CDI mouse embryos in Flushing and Handling Medium (FHM). The

embryonic region was cut out from extra-embryonic tissue and incubated for 10 min in FHM containing 0.1% Trypsin (Type II, Sigma) and 2.5% Pancreatin (Sigma). The epiblast was then isolated using glass needles and either snap-frozen or plated in four-well plate in CDM supplemented with Activin and Fgf2 for EpiSC derivation as described [29].

Real-time PCR analysis

Total RNA was extracted and reverse-transcribed with Superscript III (Invitrogen). Real-Time PCR were carried out using SybrGreen mix (Qiagen) on a Step One Plus thermal cycler (Applied Biosystem) and repeated thrice on independent experiments and/or cell lines. Data were normalized using the geometric mean of *Hprt* and *Pbgd* using Qbase software (Biogazelle). Primers used are listed in Supplementary Table S2.

Western-blot analysis

Cellular samples were lysed in 3×Laemli-SDS buffer. The polypeptides were separated through 4%–12% Bis-Tris Gel NuPage electrophoresis and transferred onto a polyvinylidene difluoride membrane (Hybond-P PVDF; Amersham). After blocking with 1/1,000 Tween 20-PBS (PBS-T) containing 4% (w/v) nonfat dried milk, the membranes were incubated with primary antibodies overnight at 4°C. Antibodies used were as follows: mouse monoclonal anti-Dnmt3b (Abcam; 1/1,500 dilution) or anti-Dnmt3a (Active Motif; 1/1,000 dilution). The membranes were washed thrice with PBS-T, incubated with a peroxidase-conjugated anti-mouse antibody and washed again. Peroxidase activity was measured using the ECL-Plus Western Blot detection system (Amersham) and a LAS 1000 camera (Fuji). The membranes were incubated with an anti-Actin antibody for loading control. Band intensities were quantified using ImageJ software (<http://imagej.nih.gov/ij/index.html>) and normalized using Actin.

DNA methylation analysis by bisulfite sequencing

Genomic DNA was purified using DNA extraction kit (Promega) according to the manufacturer’s instructions. For epiblast DNA, pools of seven (E6.5) or three (E7) epiblasts were digested by proteinase K in lysis saline buffer and DNA was extracted using NaCl/EtOH precipitation.

Bisulfite conversion was performed as previously described [30] on 1 µg of genomic DNA for the cells, or all DNA obtained from epiblasts. Regions of interest were then amplified by PCR using the KAPA HiFi HotStart Uracil+ mix (Clinisciences) using primers listed in Supplementary Table S2. The PCR program was as follows: 5 min at 95°C followed by 40 cycles of 20 s at 98°C, 30 s at 60°C, and 15 s at 72°C, with a final extension of 5 min at 72°C. PCR products were directly sequenced or subcloned into pGEMTeasy vector (Promega). Clones were amplified by PCR using the Platinum taq polymerase (Invitrogen) with 5% DMSO at 95°C 15 min followed by 35 cycles of 30 s at 95°C and 3 min at 64°C with a final extension 10 min at 64°C. PCR products of the expected size and quantity were sequenced and analyzed using BiQ Analyzer software [31].

Processing of publicly available datasets for methylation and expression in the epiblast

For comparisons with epiblast methylation, we used profiles of GEO series GSE22831 [12]. Oligo sequences of NimbleGen Mouse Promoter Array (GPL9485) were mapped to the MM9 genome assembly using bowtie2 with a maximum of three mismatches. Sequences mapping on multiple loci in the genome were discarded. For the methylation profiles of the three epiblast (E6.5) replicates within GSE22831 we computed the average log-ratios per NimbleGen probe. For each promoter regions covered by our Methylcap dataset and by the NimbleGen probes (only regions containing ≥ 5 probes were included), we computed the number of probes, the percentage of probes having an average log-ratio larger than 0.5 and the average log-ratios of all probes within the promoter. The classification of promoters was according to Borgel et al. [12]: low or no methylation ($\log_2\text{ratio} < 0.3$), or highly methylated ($\log_2\text{ratio} > 0.4$).

To determine expression levels for genes showing hypermethylated promoters in EpiSCs, we used profiles within GEO series GSE4622 [32], that is, the microarray data for epiblast at prestreak (two replicates) and mid-streak (three replicates) stages (same stages as used for EpiSC derivation, bisulfite sequencing, and RT-qPCR in this study). A gene was considered to be expressed in one sample in case of a detection P -value < 0.05 , and considered to be expressed in the epiblast if detected in at least four out of five samples.

Accession numbers

The GEO accession number for the MethylCap-seq and RNA-seq data for EpiSCs reported in this article is GSE47793. The GEO accession numbers for the previously generated MethylCap-seq and RNA-seq profiles of ESCs are GSE31343 and GSE23943, respectively [33,34]. Raw sequencing data of MeDIP-Seq from Senner et al. [21] were downloaded from the EBI European Nucleotide Archive (ENA) accession number PRJEB4263.

Results

DNA methylation profiles of ESCs and EpiSCs

To investigate the DNA methylome of ESCs and EpiSCs at a genome-wide scale, we applied MethylCap-sequencing. This method involves capture of methylated DNA using the MBD domain of MeCP2, followed by parallel sequencing of the captured DNA. In comparison to other genome-wide DNA methylation profiling methods, MethylCap-seq stands out for its robustness, sensitivity, and costeffectiveness [22]. MethylCap-seq was performed on three EpiSC lines (EpiSC1 and 3, male; EpiSC2, female) derived from fertilized B6D2F1 embryos and characterized in [29]. We compared these new profiles with an ESC line (male, E14Tg2a) [33]. Overall a total of 90,474 methylated regions were identified, with a median length of 2,034 bp. These regions are distributed across all chromosomes (Supplementary Fig. S1A). Furthermore, the read density distribution plots of all enriched regions, representative for the level of methylation at the individual loci, in either ESCs or EpiSCs were well overlaid (Supplementary Fig. S1B). This suggests that,

overall, DNA methylation is similar in ESCs and EpiSCs at a global level in accordance with Senner et al. [21].

Methylated regions were annotated according to their genomic localization: intergenic, exonic, intronic, and promoter regions. The partition of the methylated regions into these categories was the same in both pluripotent cell types, with an overrepresentation of intragenic methylation compared with the genomic background distribution within the nonrepetitive portion of the genome (Supplementary Fig. S1C). Despite these similarities between ESC and EpiSC methylation, hierarchical clustering clearly shows that the ESC methylome is distinct from that of EpiSCs (Fig. 1A). To get insight into this distinction, we plotted the read density distribution of the regions according to their genomic localization and observed that the read distribution of regions annotated as promoters was shifted toward higher read densities (ie, higher methylation) in EpiSCs compared with ESCs (Fig. 1B). Classifying promoters according to their CpG density [35], high-, intermediate-, and low-CpG content promoters (HCPs, ICPs, and LCPs, respectively), the shift revealed that the higher methylation in EpiSCs is mainly present in HCPs (Fig. 1C), and not so much in ICPs and LCPs.

Relation of promoter methylation to gene expression

To understand the functional importance of methylation differences in the promoters, we first examined the statistic correlation between DNA methylation at promoters and gene expression in ESCs and EpiSCs. We generated transcriptome profiles of the same three EpiSC lines by RNA-seq and used previously generated RNA-Seq data for the E14 ESC line used in this study [34]. We determined the correlation coefficients between promoter methylation and expression of the corresponding genes (Fig. 2A), followed by a principal component analysis performed on the matrix of these correlations (Supplementary Fig. S2A). Overall, there is a limited variability within EpiSC lines, exhibiting high positive correlations between methylation peaks (0.61–0.85). It is noteworthy that these correlation coefficients are lower between the male lines (EpiSC1 and EpiSC3) and the female one (EpiSC2), reflecting the presence of an inactivated and highly methylated X-chromosome in the female. In addition, although gene expression of the different EpiSC and ESC lines is closely related (between 0.9 and 0.99 correlation), lower correlations were observed between ESC and EpiSC methylation (0.31–0.57). Lastly, an anti-correlation between gene expression and methylation peaks was observed (between -0.08 and -0.35).

To gain more insight in the anti-correlation between gene expression and DNA methylation we performed quality threshold clustering (QT-Clust; Supplementary Fig. S2B), summarized in Fig. 2B. Fifty-eight percent of the genes present in the combined datasets followed pattern 1, that is, showed both high promoter methylation and low expression. Conversely, 29% of the genes followed the opposite pattern (pattern 2), characterized by low promoter methylation and high expression. An example of a gene following pattern 1 is shown on Fig. 2C. However, a few genes seem to escape from methylation-induced repression, as illustrated with *Car4* (Supplementary Fig. S2C). Interestingly, pattern 3

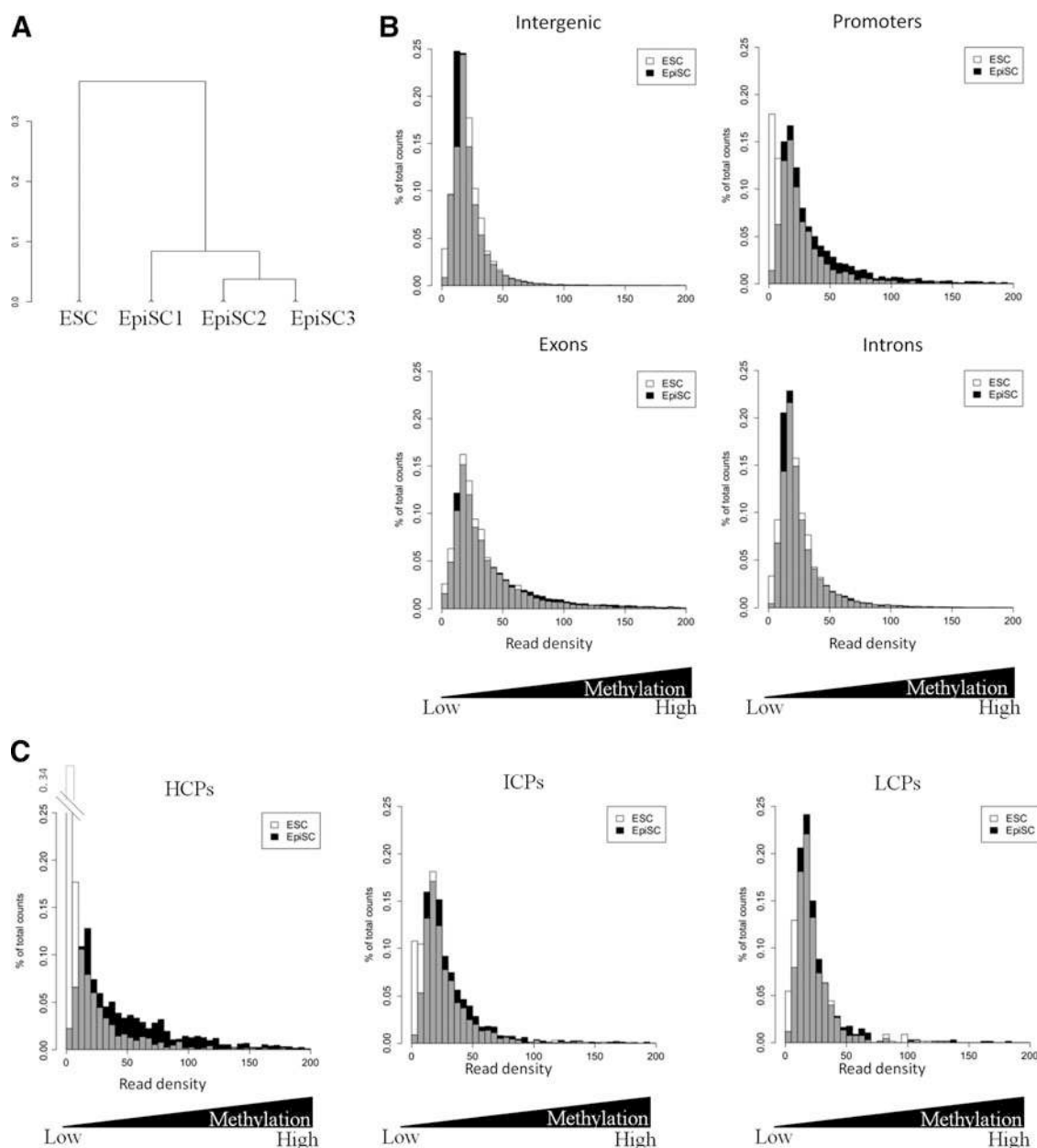


FIG. 1. Global analysis of DNA methylation in embryonic stem cell (ESC) and epiblast stem cells (EpiSC) lines. **(A)** Hierarchical clustering of methylation profiles, based on a Pearson's correlation distance matrix. **(B, C)** Distribution of read density for ESCs and EpiSCs (mean of the three cell lines) in each genomic category **(B)** and in promoters annotated as high (HCP), intermediate (ICP), or low (LCP) CpG content **(C)**. Bars for EpiSC are in white, bars for ESC in black, the overlay in gray.

characterized by high methylation only in the female EpiSC line (EpiSC2) contained mostly genes located on chromosome X (90%), in agreement with the presence of an inactivated X in this female line. This specificity of the EpiSC2 line probably explains the weaker anti-correlation (-0.08) compared with the others (-0.23 to -0.35).

Functional characterization of methylated promoters

We then investigated the biological functions and pathways associated with genes having methylated promoters, and therefore likely to be repressed in ESCs or EpiSCs. We

selected genes containing highly methylated promoters in EpiSCs or ESCs, that is, those with a read density of at least 20, which corresponds to the median of the read density distribution (Fig. 1C and Supplementary Table S3), resulting in 1,528 genes for ESCs and 2,151 for EpiSCs. Gene ontology analysis using DAVID showed only two terms associated with ESC specific methylation: transmembrane transport and translation (Fig. 3A). Terms showing up only in EpiSCs concerned response to endogenous or extracellular stimuli. Most terms (14/21) were common to both cell types, such as germ cell development and reproductive function, ion transport, and cell adhesion. In addition, two

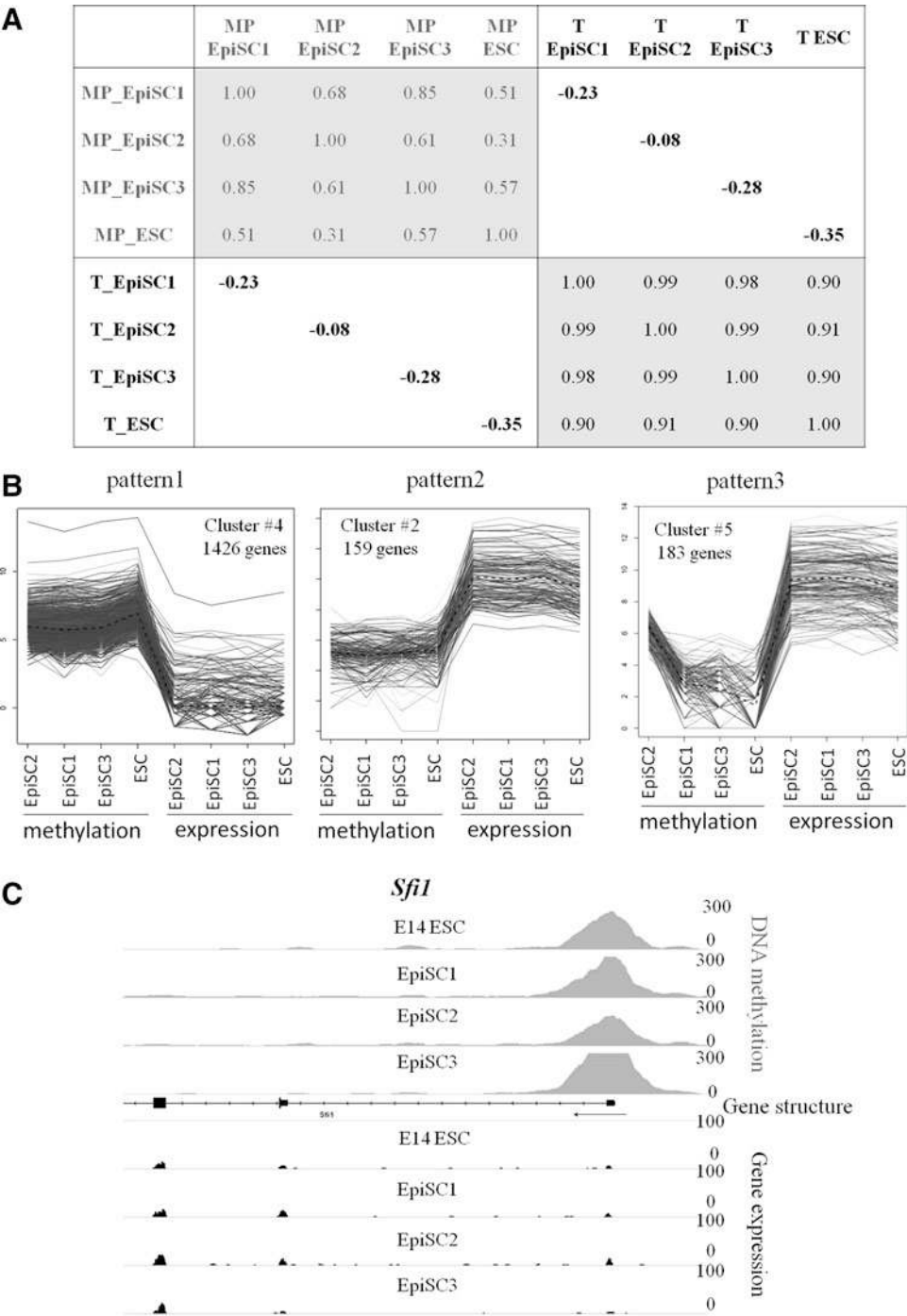


FIG. 2. Relationships between promoter methylation and gene expression. **(A)** Correlation coefficients between DNA methylation on promoters (MP) and expression data (T). **(B)** Three main clusters (Qt-clust) illustrating the relation between promoter methylation and gene expression. **(C)** Example of a gene following pattern 1.

Kegg pathways were common to ESCs and EpiSCs: neuroactive ligand-receptor interaction and ribosome. Interestingly, there were always more genes with methylated promoters in EpiSCs belonging to these common terms than for ESCs. This suggests that the repression of certain biological processes initiated in ESCs was amplified in EpiSCs. To further document this trend, we selected genes belonging to two well-represented categories, “sexual reproduction” and “gamete generation” and compared the levels of both DNA methylation and gene expression. Most genes belonging to these categories displayed much lower expression, while higher DNA methylation, in EpiSCs as

compared with ESCs (Fig. 3B). It has been reported that several germ cell markers are downregulated in EpiSCs compared with ESCs including *Piwi1* and *Nr0b1* [6]. We now show that both genes are hypermethylated in EpiSCs compared with ESCs (Fig. 3C). Among genes showing hypermethylated promoters in EpiSCs, *Dppa3* (*Stella*) and *Zfp42* (*Rex1*) have been shown to be methylated at their promoters and their expression repressed in EpiSCs, in contrast to ESCs ([9,36], and Supplementary Fig. S3B for *Zfp42*). Interestingly, also *Tbx3*, another “naïve” gene important for ESC maintenance [37], is specifically methylated in EpiSC, with concordant lower

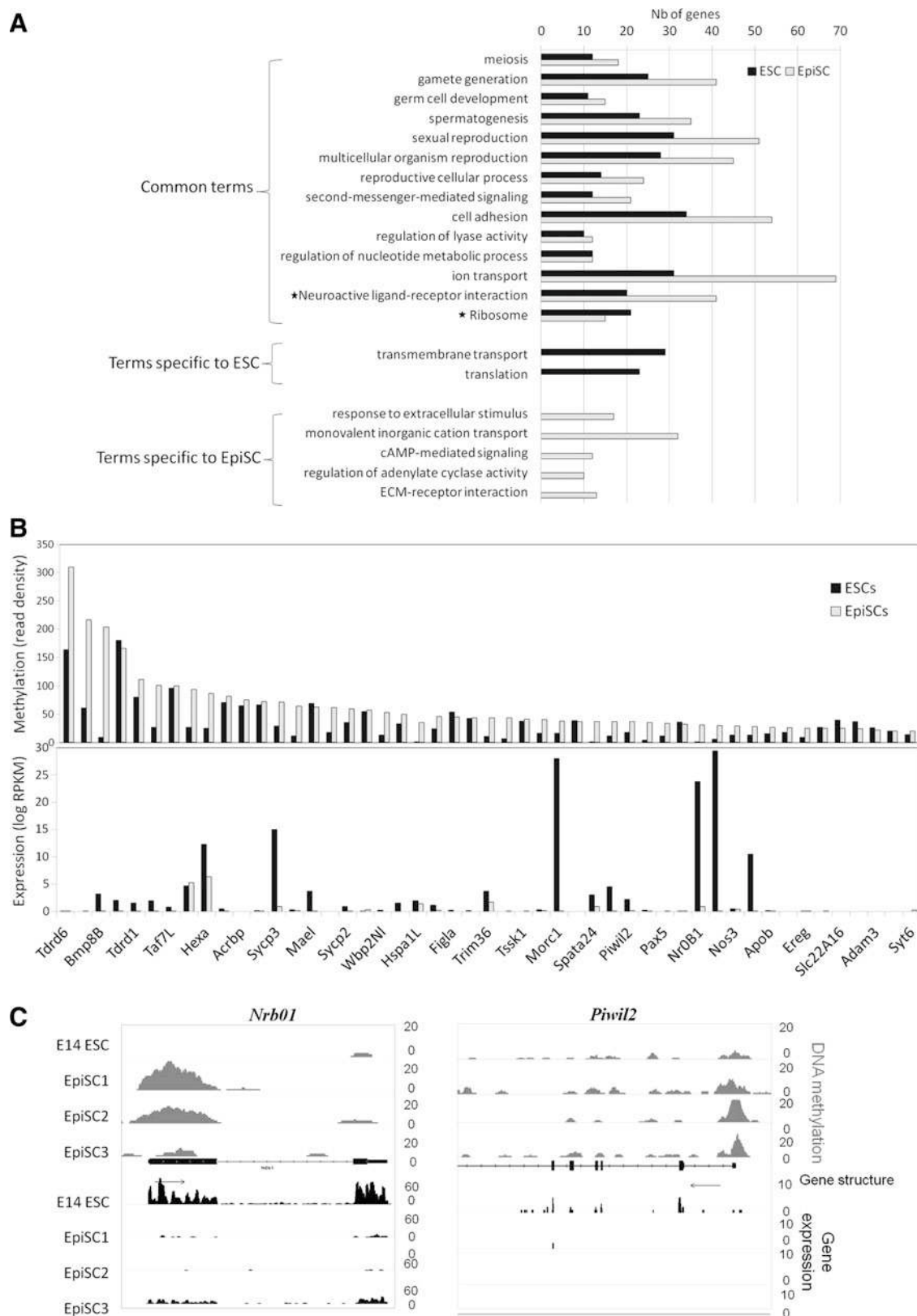


FIG. 3. Functional annotation of methylated promoters. **(A)** GO terms associated with promoters methylated (read density > 20) in EpiSCs and in ESCs, according to DAVID analysis (P -value $\leq 1\%$). The * indicate KEGG pathways. **(B)** Plots comparing methylation and expression of the 49 genes associated with the union of “gamete generation” and “sexual reproduction” GO terms. **(C)** IGV browser view of two germline and ESC-specific genes, *Nr0b1* and *Pwili2*, showing the de novo methylation and loss of gene expression in EpiSCs compared to ESCs.

expression in EpiSCs as compared with ESCs (Supplementary Fig. S3C). The zygotic promoter of *Dnmt3l*, which controls its expression during preimplantation stages, was also methylated in EpiSCs (Supplementary Fig. S3D), as is the case in the postimplantation epiblast [16]. Expression of these genes are quickly downregulated during conversion of ESCs to EpiSCs [38]; (see below). Together, our data suggest that DNA methylation at promoters in EpiSCs contributes to the regulation of genes that are differentially expressed in naïve versus primed pluripotent cells.

Identification and characteristics of differentially methylated promoters

To identify differentially methylated regions (DMRs) between ESCs and EpiSCs, we selected regions displaying a read density ratio of at least three between the two cell types with a minimum of 20 reads for the category with the highest count. This analysis was performed on the whole dataset and revealed 1,226 hypermethylated regions specific for ESCs and twice more (2,852) for EpiSCs. Among these DMRs, 724 (25%) are annotated as promoters specifically hypermethylated in EpiSCs, and only 58 (5%) are promoters that are hypermethylated specifically in ESCs, again illustrating that promoters tend to become hypermethylated in EpiSCs (Fig. 4A). Promoters hypermethylated in EpiSCs were found associated with molecular transport, metabolism, signaling, and nervous system development (Supplementary Fig. S3A). The same analysis performed on the promoters that are hypermethylated in ESCs did not yield significant results because of the low number of genes.

When CpG density was taken into account, the proportion of each category was identical for the small set of hypermethylated promoters in ESCs, compared to that of all methylated promoters (Fig. 4B). By contrast, only ICPs and HCPs were represented among promoters hypermethylated in EpiSCs, the latter being the most abundant and clearly overrepresented when compared to the population of all methylated promoters (78% compared to 35%).

Using published ChIP-seq data on histone modifications in ESCs [39,40], we determined that half of the hypermethylated promoters in EpiSCs were associated with bivalent domains (H3K4me3/H3K27me3) in ESCs (Fig. 4C), while the remaining were mainly associated with H3K4me3. Hence, these (bivalent) genes are likely stably silenced

in EpiSCs by deposition of dense methylation at their promoters.

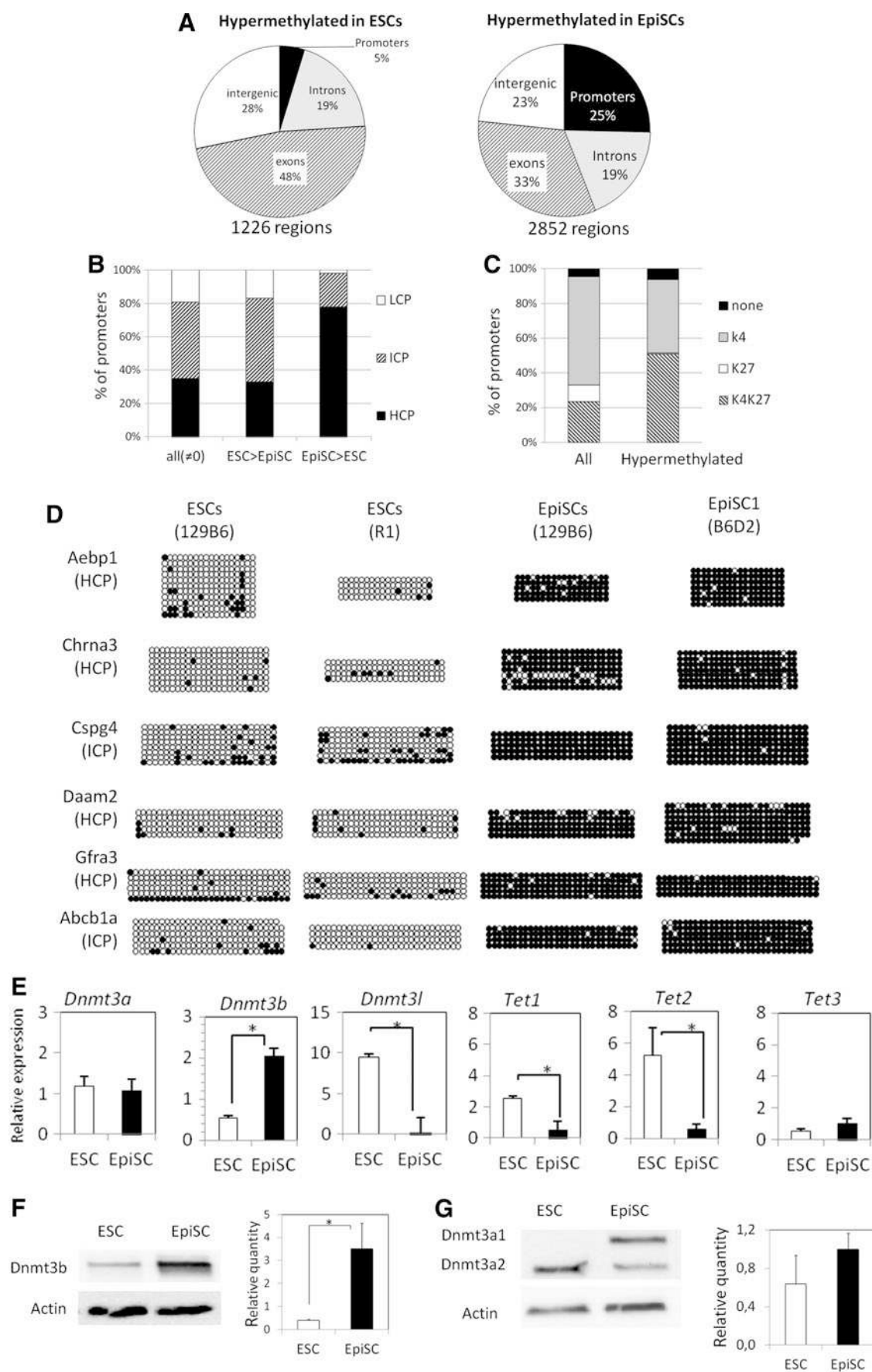
To validate the current MethylCap-seq, six DMRs in promoters were selected and assessed by bisulfite sequencing. The six genes were chosen among those with hypermethylated promoter and no expression in EpiSCs, and associated with bivalent promoters in ESCs (Supplementary Fig. S4). Of these, *Abcb1a* is a membrane transporter whereas the others are involved in development [41]. *Chrna3*, *Cspg4*, *Daam2*, and *Gfra3* play roles in nervous system development [42–46], while *Aebp1* is required for smooth muscle formation [47–50]. We verified the low level of DNA methylation in two different ESC lines, which contrasted with the high DNA methylation level (90%–98% methylated CpGs) in two EpiSC lines (Fig. 4D). These results highly correlated with the MethylCap-seq data.

To better understand the basis of the methylation difference between ESCs and EpiSCs, we assessed the RNA expression of the de novo methyltransferases *Dnmt3a*, *3b*, and *3l*, and the *Tet* enzymes that are involved in active demethylation (Fig. 4E). *Dnmt3a* is expressed at a similar level in the two cell types, whereas *Dnmt3b* expression is about four-fold higher in EpiSCs, while *Dnmt3l* is only expressed in ESCs. Expression of both *Tet1* and *Tet2* is lower in EpiSCs, while *Tet3* expression is low in both cell types. The protein level of both DNMT3A and 3B was evaluated by western-blotting (Fig. 4F, G). In good correlation with transcript level, DNMT3B is greater than eightfold higher in EpiSCs compared with ESCs. As expected, the total quantity of DNMT3A was similar in both cell types, but distributed over two different isoforms: the lower band of ~75 kDa corresponding to DNMT3A isoform 2 (DNMT3A2, [51]) is higher in ESCs, whereas the higher band (~100 kDa; DNMT3A1) appeared only in EpiSCs and accounted for about half of the total quantity of DNMT3A in these cells. In conclusion, compared with ESCs, EpiSCs have more abundant DNMT3B and DNMT3A1, whereas *Dnmt3l*, *Tet1*, and *Tet2* are much lower.

Dynamics and role of DNA methylation changes during conversion of ESCs into cEpiSCs

EpiSCs can be obtained directly in vitro from ESCs by applying EpiSC culture conditions to the ESCs (cEpiSCs, [9,10]. In cEpiSCs harvested 13–15 passages after conversion, the level of expression of the *Dnmt3* and *Tet* genes is very similar to that of embryo-derived EpiSCs

FIG. 4. EpiSCs tend to contain hypermethylated promoters. (A) Pie charts showing the classification of hypermethylated regions in ESCs (read density ratio ESC/EpiSC > 3) and hypermethylated in EpiSCs (read density ratio EpiSC/ESC > 3). The number of DMRs is indicated below the pies. ESC and EpiSC distribution differ significantly (Chi-squared test, P -value < 10^{-53}). (B) Classification of all methylated (read density $\neq 0$) and differentially methylated promoters as HCP, ICP, and LCP. Hypermethylated promoters in EpiSCs are significantly enriched in HCP (Chi-squared test, P -value < 10^{-76}). (C) Classification of all promoters (read density in EpiSC $\neq 0$) and differentially methylated promoters in EpiSC (ratio EpiSC/ESC > 3) according to their association in ESC with H3K4me3 (gray), H3K27me3 (white), bivalent (H3K4me3 + H3K27me3, hatched) or neither mark (black). Hypermethylated promoters are significantly different from all promoters (Chi-squared test, P -value < 10^{-22}). (D) Validation of differential methylation between EpiSCs and ESCs. The class of each promoter according to their CpG content is indicated. Circles represent CpG nucleotides either methylated (closed) or unmethylated (open). (E) Gene expression of DNA methylation modifying enzymes in ESCs and EpiSCs determined by RT-qPCR. Error bars represent SEM of three different cell lines. (F, G) Western blots showing the protein level of *Dnmt3a* and *Dnmt3b* in ESCs and EpiSCs. The average quantity relative to Actin is shown on the right. Error bars represent SEM of two (ESCs) to three (EpiSCs) different cell lines. *in E, F: $P < 0.05$, Mann–Whitney U test.



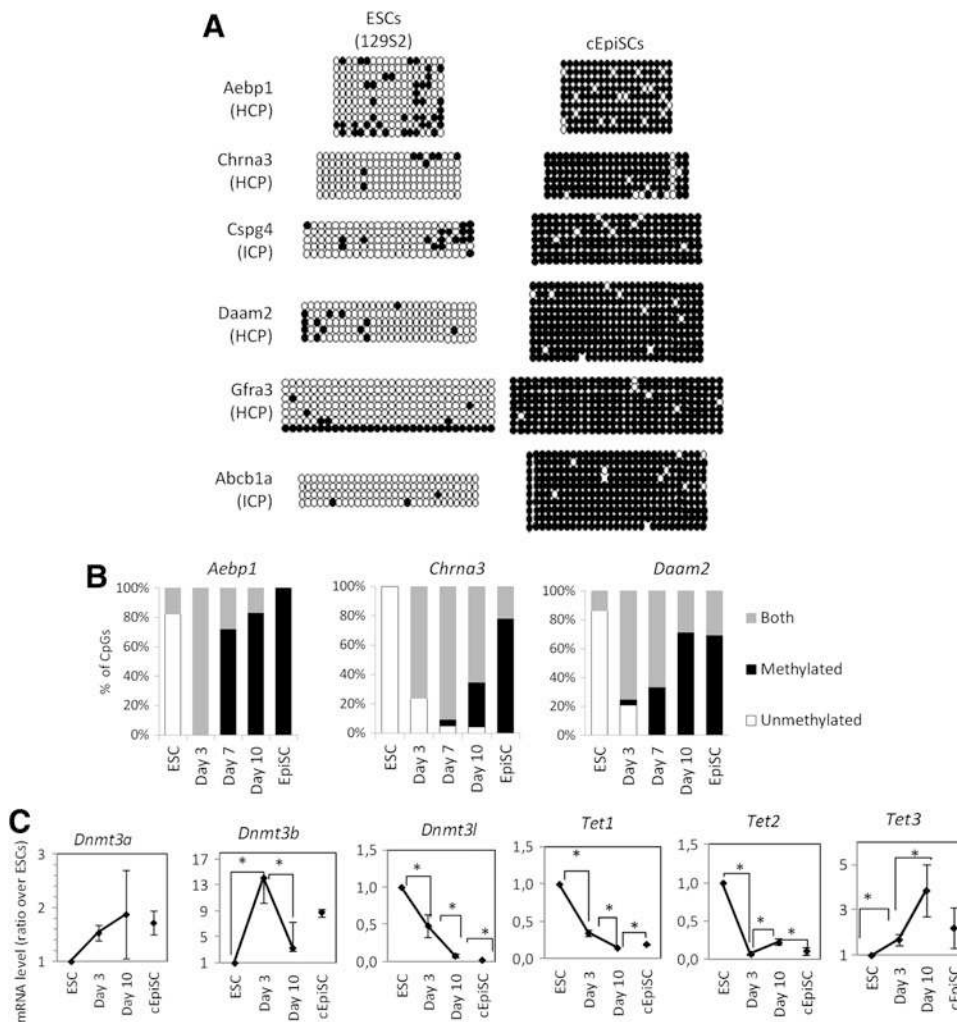


FIG. 5. Characterization of ESC conversion into cEpiSC. **(A)** Changes of methylation in the six promoters assessed by bisulfite-sequencing during conversion of ESCs into cEpiSCs. **(B)** Changes in the CpG level of methylation in the promoter of *Aebp1*, *Chrna3*, and *Daam2* during the conversion process. Genomic DNA after bisulfite conversion was directly sequenced. Each CpG has been classified according to the presence of a C (methylated), a T (unmethylated), or a polymorphism meaning heterogeneity between the two forms in the cell population (both). **(C)** Changes in expression of DNA methylation modifying enzymes, during the conversion of ESCs determined by RT-qPCR. Bars represent SEM of three independent experiments. * $P < 0.05$, Mann-Whitney U test.

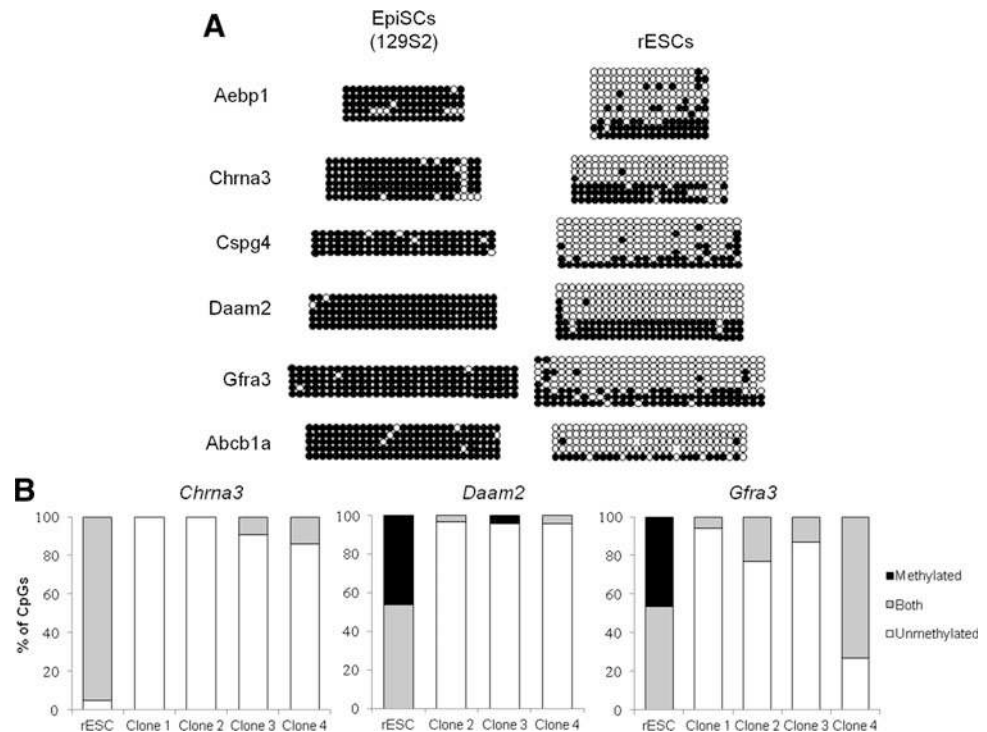
(Supplementary Fig. S5A). In line with this, the pattern of DNA methylation of embryo-derived EpiSCs was also correctly apposed in these cEpiSC: a very similar, dense methylation was observed in both EpiSCs and cEpiSCs (88%–98%) for the six promoters described above (Fig. 5A; see Fig. 4D for comparison). During conversion, colonies with an EpiSC-like morphology first appeared at day 6 (Supplementary Fig. S5B), while changes in gene expression levels of ESC markers such as *Klf4* and *Dppa3*, or the upregulation of the EpiSC marker *Fgf5*, occur as early as at day 3 (Supplementary Fig. S5C). We therefore asked when the de novo methylation occurred during the conversion. To this end, bisulfite sequencing was performed on *Aebp1* and *Chrna3* promoters at day 3, 7, and 10 of conversion (Fig. 5B). In parallel, we also examined the dynamics of expression of the *Dnmt3* and the *Tet* genes (Fig. 5C). At day 3, sequence polymorphism was present at most CpG loci (Fig. 5B), indicating that de novo methylation had already started, in accordance with the early upregulation of *Dnmt3b* and the downregulation of *Tet1* and *Tet2*. *Dnmt3l*, on the other hand, was quickly downregulated and, intriguingly, *Tet3* was transiently upregulated during conversion, although remaining at low level. Altogether, our results show that DNA hypermethylation at promoters occurs early in the transition from ESCs to EpiSCs.

Reprogramming of DNA methylation during reversion of EpiSCs to ESCs

As DNA methylation is considered to be a stable epigenetic mark, we next asked whether reprogramming of EpiSCs toward ESCs would efficiently reverse methylation at promoters. The DNA methylation status at the six promoters as mentioned above was analyzed in rESCs harvested 11–13 passages after the start of reversion by transferring the EpiSCs onto feeders in LIF-containing medium [9]. Although DNA methylation was largely reduced, all six genes displayed a higher level of DNA methylation at their promoters in the rESCs as compared with the embryo-derived ESCs (Fig. 6A). Remarkably, we observed heterogeneity between clones, representing different alleles in the population: some were highly methylated while others were unmethylated as is the case for *Daam2*. This is in contrast with embryo-derived ESCs in which the DNA methylation level at each allele was quite similar (see Fig. 4D for comparison). We verified that the level of expression of *Dnmt3s* and *Tets* was correctly reprogrammed in rESCs compared with embryo-derived ESCs (Supplementary Fig. S6).

Bisulfite sequencing of individual alleles after cloning does not allow distinguishing between allelic heterogeneity

FIG. 6. Characterization of EpiSC reversion into rESC. (A) Changes of methylation in the six promoters assessed by bisulfite-sequencing during reversion of EpiSCs into rESCs. (B) CpG level of methylation in the promoter of *Chrna3*, *Daam2*, and *Gfra3* after clonal expansion of rESCs. As in Fig. 5B, each CpG is classified as methylated, unmethylated, or polymorphic (both) in the cell population.



within cells or among the cell population, as linkage information between the different fragments is lost. Therefore, we grew four rESC clones originating from single rESCs and determined the methylation status of the three promoters showing the highest methylation heterogeneity (*Chrna3*, *Daam2*, and *Gfra3*; Fig. 6B). Surprisingly, clones were now mostly demethylated, with the notable exception of *Gfra3* for clone 4, which still exhibited some methylated CpGs. These results indicate that DNA methylation is stable and resistant to reprogramming although further passages and severe selection by sub-cloning is able to almost, but not totally, erase this epigenetic memory.

Comparison of methylated promoters in EpiSC compared to the epiblast

It has been reported that the DNA methylation signature at promoters in ESCs is closer to the early postimplantation epiblast than to the ICM that they are derived from [12]. We now asked how promoter methylation in EpiSCs compared with that of their in vivo counterpart. We isolated epiblasts from early (E6.5) or late (E7) gastrulating embryos and performed bisulfite sequencing on the six promoters that were strongly methylated in EpiSCs. The DNA methylation level in epiblasts was very low at the two stages and even lower than the level of methylation observed in ESCs for *Aebp1*, *Cspg4*, and *Gfra3* (Fig. 7A and Supplementary Fig. S7).

To perform this analysis at a global scale, we compared our data with the publicly available dataset of promoter methylation on E6.5 epiblasts obtained by MedIP arrays [12] after selection of promoters common to the two study (2,610 promoters; Supplementary Table S4 and Fig. 7B). The majority of promoters (77%, 1,861/2,416) that were

methylated in EpiSCs (at least 10 reads) had a low level of methylation in epiblast ($\log_2\text{ratio} < 0.3$). Conversely, very few promoters with low level of methylation in EpiSCs were methylated in the epiblast (8%, 15/194). Methylation at these promoters is therefore unique to EpiSCs, and does not recapitulate the methylation status of the epiblast stage they are derived from. The expression level of *Dnmt3s* and *Tets* in epiblast and EpiSCs could not explain this difference in methylation deposition, as these enzymes were similarly expressed, except for *Dnmt3b*, which is even higher expressed in the epiblast (Fig. 7C).

To get insight into the kinetic of this methylation process during EpiSC derivation, day 6.5 epiblasts were explanted in culture and outgrowths collected at different time points. Bisulfite conversion followed by direct sequencing was performed on three representative promoters, *Aebp1*, *Chrna3*, and *Daam2* (Fig. 7D). Completely methylated CpGs started to appear as early as day 2 for *Aebp1* and before day 9 for *Chrna3*. For *Aebp1*, it reached the level of established EpiSCs within 9 days. Such kinetic is not in favor of a slow deposition of methylation along passages in culture but rather suggests that the removal of the epiblast from its in vivo environment and/or the culture conditions may have released constraints that prevent promoter hypermethylation within the embryo.

Lastly, to gain insight into the functional consequence of the differential DNA methylation, we asked whether the difference in promoter DNA methylation between EpiSCs and epiblast cells translated into differences in expression of the corresponding genes. Using available microarray data on expression in the epiblast [32], we determined the expression status (expressed or not expressed) of genes in the epiblast with contain hypermethylated promoters in EpiSCs (reads ≥ 20) (Fig. 7E and Supplementary Table S4). Most

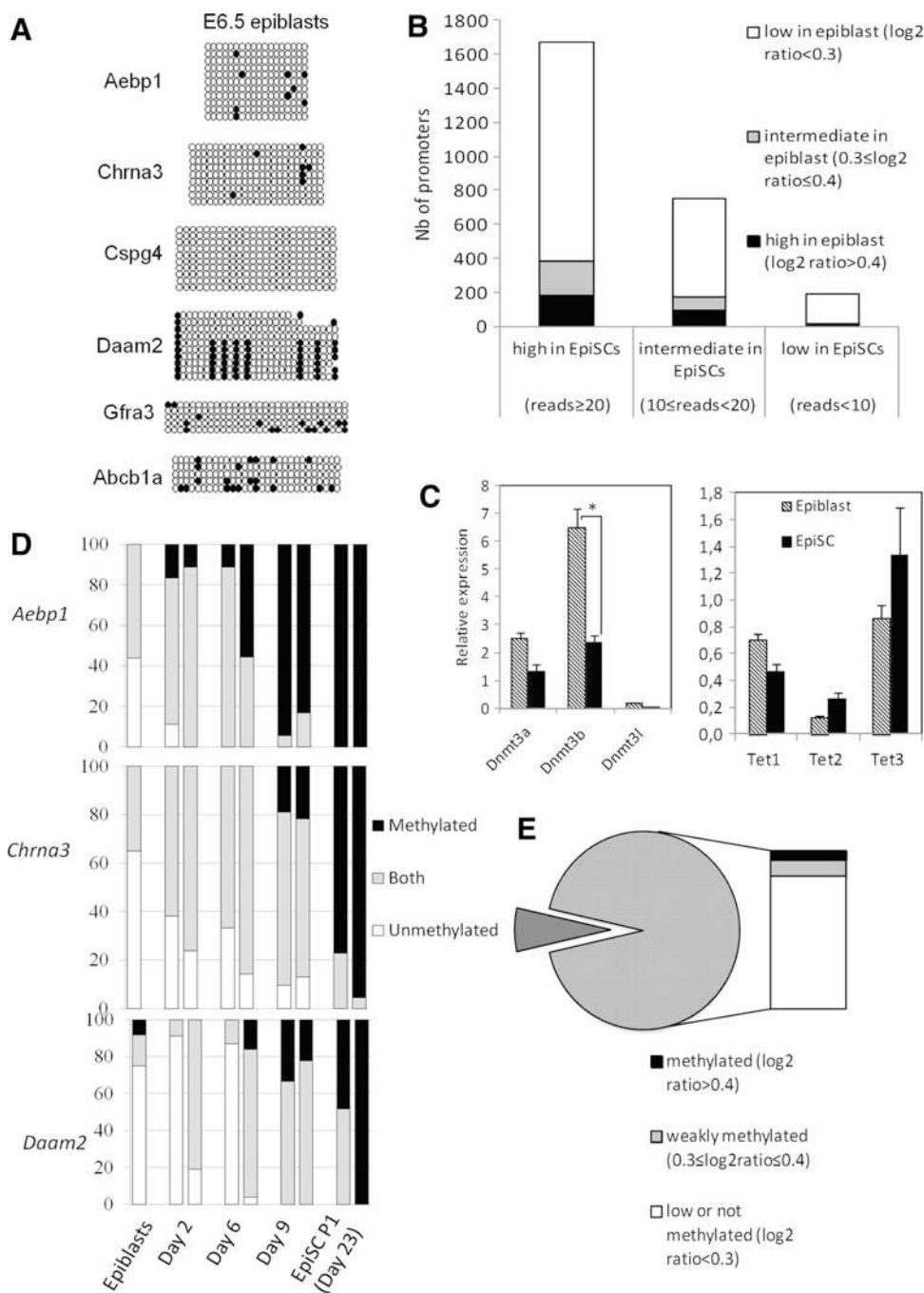


FIG. 7. Differential methylation in embryo-derived pluripotent cells and in the epiblast. **(A)** Methylation status in the epiblast (E6.5). **(B)** Comparison of methylation in promoters common to Borgel et al. [12] and our study. Promoters were classified according to their mean methylation status in EpiSCs and each category further separated according to their methylation values in E6.5 epiblasts. **(C)** Gene expression of DNA methylation modifying enzymes in E6.5 epiblast and EpiSCs determined by RT-qPCR. Bars represent SEM. $*P < 0.05$, Mann-Whitney U test. **(D)** Changes in the CpG level of methylation in the promoter of *Aebp1*, *Chrna3*, and *Daam2* during derivation of EpiSCs. Genomic DNA after bisulfite conversion was directly sequenced and the status of the CpGs was indicated as in Figs. 5 and 6. **(E)** Expression status in the epiblast of genes that contain methylated promoters in EpiSCs.

genes (93%, 554/593) were not expressed in the epiblast, although being largely unmethylated. This suggests that for a large set of genes that are repressed in both the epiblast and their in vitro counterparts (the EpiSCs), the epigenetic mechanism of repression is different.

Discussion

We have compared the DNA methylome of EpiSCs to that of ESCs and to their tissue of origin, the post-implantation epiblast. ESCs and EpiSCs have similar methylation levels and a similar distribution of methylation within the different genomic regions. However, several

features distinguish the two cell types, in particular the fact that there are significantly more regions specifically methylated in EpiSCs as compared to ESCs, a large part of those being hypermethylated HCP promoters. These promoters are mostly associated with either a bivalent signature (H3K4me3 and H3K27me3) or an active H3K4me3/me3 mark in the ESCs [39,40]. Furthermore, our study suggests that the promoter methylation pattern is quickly established during the in vitro conversion of ESCs into EpiSCs.

To further validate our findings on an independent dataset, we re-examined the MeDIP-seq data generated by Senner et al. [21]. This analysis yielded the same results as with our dataset: a larger set of hypermethylated, high-CpG

The comparison of our dataset of methylated promoters in EpiSCs with the dataset on epiblast cells [12] revealed that many methylated promoters in EpiSCs are poorly methyl-

In conclusion, our study shows that EpiSCs have a specific DNA methylome signature, in particular at promoters. It differs from both ESCs and from the epiblast they originate from, and cannot be easily erased by reprogramming EpiSCs to a more naïve ESC-like state through modulation of the culture conditions. The rapid molecular and epigenetic changes during the first days of ESC-to-EpiSC conversion make it an interesting system to further study the role of DNA methylation in the transition from naïve to a primed state of pluripotency.

We thank E.M. Janssen-Megens and Y. Tan for technical support with sequencing. Femke Simmer and Arjen Brinkman helped with the analysis of MethylCap-seq data. We are grateful to H  l  ne Jammes and Sylvie Ruffini for technical help in bisulfite sequencing and western-blotting, respectively. Many thanks to Gabriel Brons for the derivation of 129S2 EpiSC. The research leading to these results has received funding from PluriSys (FP7/2009: 223485) and by ANR Programme Investissements d’Avenir REVIVE. ACV and MT are recipients of a PhD fellowship from DIM STEM Pole and DIM Biotherapies Ile-de-France. HM is supported by a grant from The Netherlands Organization for Scientific

Research (NWO-VIDI 864.12.007). ASB was part sponsored by the British Heart Foundation (FS/12/37/29516).

Author Disclosure Statement

No financial competing interest exists.

References

- Hutchins AP, SH Choo, TK Mistri, M Rahmani, CT Woon, CK Ng, R Jauch and P Robson. (2013). Co-motif discovery identifies an Esrrb-Sox2-DNA ternary complex as a mediator of transcriptional differences between mouse embryonic and epiblast stem cells. *Stem Cells* 31:269–281.
- Nichols J and A Smith. (2012). Pluripotency in the embryo and in culture. *Cold Spring Harb Perspect Biol* 4:a008128.
- Brons IGM, LE Smithers, MWB Trotter, P Rugg-Gunn, B Sun, SM Chuva de Sousa Lopes, SK Howlett, A Clarkson, L Ahrlund-Richter, RA Pedersen and L Vallier. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448:191–195.
- Chenoweth JG, RDG McKay and PJ Tesar. (2010). Epiblast stem cells contribute new insight into pluripotency and gastrulation. *Dev Growth Differ* 52:293–301.
- Nichols J and A Smith. (2009). Naive and primed pluripotent states. *Cell Stem Cell* 4:487–492.
- Tesar PJ, JG Chenoweth, FA Brook, TJ Davies, EP Evans, DL Mack, RL Gardner and RDG McKay. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196–199.
- Huang Y, R Osorno, A Tsakiridis and V Wilson. (2012). In Vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Reprog* 2:1571–1578.
- Ying QL, J Nichols, I Chambers and A Smith. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115:281–292.
- Bao S, F Tang, X Li, K Hayashi, A Gillich, K Lao and MA Surani. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 461:1292–1295.
- Guo G, J Yang, J Nichols, JS Hall, I Eyres, W Mansfield and A Smith. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136:1063–1069.
- Zhou HY, WL Li, SY Zhu, JY Joo, JT Do, W Xiong, JB Kim, K Zhang, HR Scholer and S Ding. (2010). Conversion of mouse epiblast stem cells to an earlier pluripotency state by small molecules. *J Biol Chem* 285:29676–29680.
- Borgel J, S Guibert, Y Li, H Chiba, D Schubeler, H Sasaki, T Forne and M Weber. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* 42:1093–1100.
- Smith ZD, MM Chan, TS Mikkelsen, H Gu, A Gnirke, A Regev and A Meissner. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* 484:339–344.
- Mikkelsen TS, J Hanna, X Zhang, M Ku, M Wernig, P Schorderet, BE Bernstein, R Jaenisch, ES Lander and A Meissner. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454:49–55.
- Okano M, DW Bell, DA Haber and E Li. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99:247–257.
- Guenatri M, R Duffie, J Iranzo, P Fauque and D Bourc'his. (2013). Plasticity in Dnmt3L-dependent and -independent modes of de novo methylation in the developing mouse embryo. *Development* 140:562–572.
- Hirasawa R and H Sasaki. (2009). Dynamic transition of Dnmt3b expression in mouse pre- and early post-implantation embryos. *Gene Exp Patterns* 9:27–30.
- Hu Y-G, R Hirasawa, J-L Hu, K Hata, C-L Li, Y Jin, T Chen, E Li, M Rigolet, et al. (2008). Regulation of DNA methylation activity through Dnmt3L promoter methylation by Dnmt3 enzymes in embryonic development. *Hum Mol Gen* 17:2654–2664.
- Leitch HG, KR McEwen, A Turp, V Encheva, T Carroll, N Grabole, W Mansfield, B Nashun, JG Knezovich, et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol* 20:311–316.
- Habibi E, AB Brinkman, J Arand, LI Kroeze, HH Kerstens, F Matarese, K Lepikhov, M Gut, I Brun-Heath, et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* 13:360–369.
- Senner CE, F Krueger, D Oxley, S Andrews and M Hemberger. (2012). DNA methylation profiles define stem cell identity and reveal a tight embryonic-extraembryonic lineage boundary. *Stem Cells* 30:2732–2745.
- Bock C, EM Tomazou, AB Brinkman, F Muller, F Simmer, H Gu, N Jager, A Gnirke, HG Stunnenberg and A Meissner. (2010). Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol* 28:1106–1114.
- Huang DW, BT Sherman and RA Lempicki. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols* 4:44–57.
- Huang DW, BT Sherman and RA Lempicki. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucl Acids Res* 37:1–13.
- Anders S and W Huber. (2010). Differential expression analysis for sequence count data. *Genome Biol* 11:R106.
- Leisch F. (2006). A toolbox for -centroids cluster analysis. *Comput Stat Data Anal* 51:526–544.
- Bryja V, S Bonilla, L Cajanek, CL Parish, CM Schwartz, Y Luo, MS Rao and E Arenas. (2006). An efficient method for the derivation of mouse embryonic stem cells. *Stem Cells* 24:844–849.
- Jouneau A, C Ciaudo, O Sismeiro, V Brochard, L Jouneau, S Vandormael-Pournin, JY Coppee, Q Zhou, E Heard, C Antoniewski and M Cohen-Tannoudji. (2012). Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. *RNA* 18:253–264.
- Maruotti J, XP Dai, V Brochard, L Jouneau, J Liu, A Bonnet-Garnier, H Jammes, L Vallier, IG Brons, et al. (2010). Nuclear transfer-derived epiblast stem cells are transcriptionally and epigenetically distinguishable from their fertilized-derived counterparts. *Stem Cells* 28:743–752.
- Dupont JM, J Tost, H Jammes and IG Gut. (2004). De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem* 333:119–127.
- Bock C, S Reither, T Mikeska, M Paulsen, J Walter and T Lengauer. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* 21:4067–4068.

32. Kojima Y, K Kaufman-Francis, JB Studdert, KA Steiner, MD Power, DA Loebel, V Jones, A Hor, G de Alencastro, et al. (2014). The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak. *Cell Stem Cell* 14:107–120.
33. Matarese F, E Carrillo-de Santa Pau and HG Stunnenberg. (2011). 5-Hydroxymethylcytosine: a new kid on the epigenetic block? *Mol Syst Biol* 7:562.
34. Marks H, T Kalkan, R Menafrá, S Denissov, K Jones, H Hofemeister, J Nichols, A Kranz, A Francis Stewart, A Smith and HG Stunnenberg. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149:590–604.
35. Weber M, I Hellmann, MB Stadler, L Ramos, S Paabo, M Rebhan and D Schubeler. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39:457–466.
36. Hayashi K, SMCdS Lopes, F Tang and MA Surani. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* 3:391–401.
37. Niwa H, J Miyazaki and AG Smith. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24:372–376.
38. Hayashi K, H Ohta, K Kurimoto, S Aramaki and M Saitou. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146:519–532.
39. Meissner A, TS Mikkelsen, H Gu, M Wernig, J Hanna, A Sivachenko, X Zhang, BE Bernstein, C Nusbaum, et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454:766–770.
40. Mikkelsen TS, M Ku, DB Jaffe, B Issac, E Lieberman, G Giannoukos, P Alvarez, W Brockman, TK Kim, et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448:553–560.
41. Sawicki WT, M Kujawa, E Jankowska-Steifer, ET Mystkowska, A Hyc and C Kowalewski. (2006). Temporal/spatial expression and efflux activity of ABC transporter, P-glycoprotein/Abcb1 isoforms and Bcrp/Abcg2 during early murine development. *Gene Expr Patterns* 6:738–746.
42. Atluri P, MW Fleck, Q Shen, SJ Mah, D Stadfelt, W Barnes, SK Goderie, S Temple and AS Schneider. (2001). Functional nicotinic acetylcholine receptor expression in stem and progenitor cells of the early embryonic mouse cerebral cortex. *Dev Biol* 240:143–156.
43. Kida Y, T Shiraishi and T Ogura. (2004). Identification of chick and mouse Daam1 and Daam2 genes and their expression patterns in the central nervous system. *Brain Res Dev Brain Res* 153:143–150.
44. Nakaya MA, R Habas, K Biris, WC Dunty, Jr., Y Kato, X He and TP Yamaguchi. (2004). Identification and comparative expression analyses of Daam genes in mouse and *Xenopus*. *Gene Expr Patterns* 5:97–105.
45. Nishiyama A, XH Lin, N Giese, CH Heldin and WB Stallcup. (1996). Co-localization of NG2 proteoglycan and PDGF α -receptor on O2A progenitor cells in the developing rat brain. *J Neurosci Res* 43:299–314.
46. Widenfalk J, A Tomac, E Lindqvist, B Hoffer and L Olson. (1998). GFR α -3, a protein related to GFR α -1, is expressed in developing peripheral neurons and ensheathing cells. *Eur J Neurosci* 10:1508–1517.
47. Ith B, J Wei, SF Yet, MA Perrella and MD Layne. (2005). Aortic carboxypeptidase-like protein is expressed in collagen-rich tissues during mouse embryonic development. *Gene Expr Patterns* 5:533–537.
48. Layne MD, SF Yet, K Maemura, CM Hsieh, M Bernfield, MA Perrella and ME Lee. (2001). Impaired abdominal wall development and deficient wound healing in mice lacking aortic carboxypeptidase-like protein. *Mol Cell Biol* 21:5256–5261.
49. Kucharova K and WB Stallcup. (2010). The NG2 proteoglycan promotes oligodendrocyte progenitor proliferation and developmental myelination. *Neuroscience* 166:185–194.
50. Li J, C Klein, C Liang, R Rauch, K Kawamura and AJ Hsueh. (2009). Autocrine regulation of early embryonic development by the artemin-GFRA3 (GDNF family receptor- α 3) signaling system in mice. *FEBS Lett* 583:2479–2485.
51. Chen T, Y Ueda, S Xie and E Li. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *J Biol Chem* 277:38746–38754.
52. Hackett Jamie A, S Dietmann, K Murakami, TA Down, HG Leitch and MA Surani. (2013). Synergistic mechanisms of DNA demethylation during transition to ground-state pluripotency. *Stem Cell Rep* 1:518–531.
53. Velasco G, F Hube, J Rollin, D Neuillet, C Philippe, H Bouzinba-Segard, A Galvani, E Viegas-Pequignot and C Francastel. (2010). Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc Natl Acad Sci U S A* 107:9281–9286.
54. Williams K, J Christensen and K Helin. (2012). DNA methylation: TET proteins—guardians of CpG islands? *EMBO Rep* 13:28–35.
55. Wu H, AC D'Alessio, S Ito, Z Wang, K Cui, K Zhao, YE Sun and Y Zhang. (2011). Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* 25:679–684.
56. Neri F, A Krepelova, D Incarnato, M Maldotti, C Parlato, F Galvagni, F Matarese, HG Stunnenberg and S Oliviero. (2013). Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA Methylation at gene bodies in ESCs. *Cell* 155:121–134.
57. Kotini AG, A Mpakali and T Agalioti. (2011). Dnmt3a1 upregulates transcription of distinct genes and targets chromosomal gene clusters for epigenetic silencing in mouse embryonic stem cells. *Mol Cell Biol* 31:1577–1592.
58. Chin MH, MJ Mason, W Xie, S Volinia, M Singer, C Peterson, G Ambartsumyan, O Aimiwu, L Richter, et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5:111–123.
59. Buecker C, H-H Chen, JM Polo, L Daheron, L Bu, TS Barakat, P Okwieka, A Porter, J Gribnau, K Hochedlinger and N Geijsen. (2010). A Murine ESC-like state facilitates transgenesis and homologous recombination in human pluripotent stem cells. *Cell Stem Cell* 6:535–546.
60. Kim K, A Doi, B Wen, K Ng, R Zhao, P Cahan, J Kim, MJ Aryee, H Ji, et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* 467:285–290.
61. Rais Y, A Zviran, S Geula, O Gafni, E Chomsky, S Viukov, AA Mansour, I Caspi, V Krupalnik, et al. (2013). Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* 502:65–70.
62. Greber B, G Wu, C Bernemann, JY Joo, DW Han, K Ko, N Tapia, D Sabour, J Sterneckert, P Tesar and HR Scholer.

- (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* 6:215–226.
63. Vallier L, S Mendjan, S Brown, Z Chng, A Teo, LE Smithers, MW Trotter, CH Cho, A Martinez, et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* 136:1339–1349.
64. Ficiz G, TA Hore, F Santos, HJ Lee, W Dean, J Arand, F Krueger, D Oxley, Y-L Paul, et al. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* 13:351–359.
65. Grabole N, J Tischler, JA Hackett, S Kim, F Tang, HG Leitch, E Magnusdottir and MA Surani. (2013). Prdm14 promotes germline fate and naive pluripotency by repressing FGF signalling and DNA methylation. *EMBO Rep* 14:629–637.
66. Corson LB, Y Yamanaka, KM Lai and J Rossant. (2003). Spatial and temporal patterns of ERK signaling during mouse embryogenesis. *Development* 130:4527–4537.

Address correspondence to:

Dr. Alice Jouneau

INRA

UMR1198 Biologie du Développement et Reproduction

Jouy-en-Josas F-78350

France

E-mail: alice.jouneau@jouy.inra.fr

Received for publication December 24, 2013

Accepted after revision April 15, 2014

Prepublished on Liebert Instant Online April 16, 2014

Acquiring Ground State Pluripotency: Switching Mouse Embryonic Stem Cells from Serum/LIF Medium to 2i/LIF Medium

Matteo Tosolini and Alice Jouneau

Abstract

Mouse embryonic stem cells (ESCs) derive from the inner cell mass (ICM) of a blastocyst. These cells are pluripotent and thus able to generate both somatic and germinal lineages. It is possible to maintain ESCs in different pluripotent states depending on the in vitro culture conditions. Classically, ESCs are cultured in the presence of serum and LIF, which sustain the naive state of pluripotency but in this metastable state cells exhibit a large degree of heterogeneity. In the last few years, it has been discovered that when ESCs are cultured in a chemically defined medium (without serum), in the presence of LIF and with the addition of two small molecules (in particular the inhibitors of MAPK and Gsk-3 pathways), they reach a ground state of pluripotency where cells are more homogeneous and more “ICM-like.” In this protocol, we describe how we culture mouse ESCs and the way we switch them from naive to ground state.

Keywords: Mouse embryonic stem cells (ESCs), Serum, 2i, Chemically defined medium (CDM), Laminin

1 Introduction

Mouse embryonic stem cells (ESCs) derive from the inner cell mass (ICM) of an E3.5 embryo at the blastocyst stage (1). These cells are in a pluripotent state that means they are characterized by the ability to differentiate and generate both somatic and germ lineages, indeed when injected into a blastocyst, these cells are able to generate chimeric embryos. In addition, when injected subcutaneously into a mouse, they give rise to teratomas (2). Historically, mouse ESCs were maintained in the pluripotent state by culturing them in fetal bovine serum (FBS) containing medium, on a layer of mitotically inactivated fibroblasts called “feeder” cells, which provides trophic factors for the growth of ESCs (3). In particular, “feeder” cells were shown to produce Leukemia inhibitory factor (LIF) and so the addition of this interleukin into the medium could replace feeders (4). LIF is an activator of the transcriptional factor Stat3, which inhibits ESCs differentiation and promotes self-renewal (5). ESCs cultured in serum/LIF conditions

are defined to be naive pluripotent cells in order to distinguish them from Epiblast stem cells (EpiSCs), which are in a primed pluripotent state, more prone to differentiation (6). EpiSCs derive from late epiblast, a postimplantation mouse embryo at E5.5 and in vitro they required Activin A and Fibroblast Growth Factor 2 (FGF2) instead of LIF and serum to sustain pluripotency (7, 8). Analysis of ESCs cultured in serum conditions showed a strong heterogeneity in the population of cells, even in the expression of pluripotency factors, due to the uncontrolled and multifactorial stimulation by all the extracellular signals present in the serum. The identification of two small molecules that could substitute for serum to sustain ESCs culture was a turning point in the field. In particular, it is possible to culture ESCs without serum with the addition of PD0325901, an inhibitor of mitogen-activated protein kinase (MAPK), and CHIR99021, an inhibitor of glycogen synthase kinase-3 (Gsk3) (9), and in presence of LIF. This new defined medium called “2i” leads to obtain ESCs in a new state: ground pluripotency. ESCs in 2i are more homogeneous and show lower expression of lineage-associated genes and less DNA methylation, so they are closer to ICM-like cells (10, 11). Concerning mouse pluripotency three distinct states can be distinguished: ground, naive, and primed, which correspond to ESCs in 2i conditions, ESCs in serum condition and EpiSCs, respectively. These states are reversible and interconvertible. One of the advantages of using a completely defined medium is that the same basal medium can be used to convert ESC to EpiSC, with the only change being the added factors, in this specific case from 2i/LIF to Activin A/FGF2.

2 Materials

1. Conical centrifuge tubes 15 and 50 mL, sterile.
2. Graduated plastic pipettes (sterile, single package) of 2, 5, 10, and 25 mL.
3. Glass Pasteur pipettes sterilized in an aluminium container using a dry oven (4 h at 180 °C).
4. Plastic sterile petri dishes for cell culture of 35 and 60 mm of diameter.
5. Cryotube vials of 1.8 mL.
6. Freezing container for tube of 1.8 mL (rate of cooling $-1\text{ }^{\circ}\text{C}/\text{min}$).
7. Low temperature freezer ($-80\text{ }^{\circ}\text{C}$).
8. Liquid nitrogen container ($-196\text{ }^{\circ}\text{C}$).
9. Pipets P1000, P200, P20, P10, and sterile plastic tips.

10. 20 mL syringes.
11. Syringe membrane filters, 0.22 μm , in PES (Polyethersulfone).
12. Water bath.
13. Centrifuge (for 15 mL plastic tubes).
14. Incubator at 37 °C in a humid atmosphere with 5 % CO_2 .
15. Vertical laminar flow hood.
16. Aspiration system.
17. H_2O Milli-Q produced with a resistivity of 18.2 $\text{M}\Omega\text{ cm}$ at 25 °C and sterilized.
18. Dimethyl sulfoxide (DMSO) $\geq 99\%$.
19. Dulbecco's phosphate-buffered saline (DPBS) 1 \times sterile without Ca^{2+} and Mg^{2+} .
20. Trypsin-EDTA (0.25 %). Aliquots of 10 mL stored at -20°C .
21. Protease-free BSA (Bovine serum albumin). Powder stored at $+4^\circ\text{C}$.
22. 2-Mercaptoethanol (50 mM). Aliquots of 1 mL stored at $+4^\circ\text{C}$.
23. Transferrin. Resuspended in H_2O Milli-Q at the final concentration of 30 mg/mL. Aliquots stored at -20°C .
24. Recombinant Insulin. Resuspended in H_2O Milli-Q at the final concentration of 10 $\mu\text{g/mL}$. Aliquots stored at -20°C .
25. Leukemia Inhibitory Factor (LIF). Resuspend in PBS/BSA 0.1 % to the final concentration of 10 $\mu\text{g/mL}$ (10^6 U/mL). Aliquots stored at -20°C .
26. Laminin 1 mg/mL. Aliquots of 10 μL stored at -20°C .
27. CHIR99021. Resuspended in DMSO to a final concentration of 10 mM. Aliquots stored at -20°C .
28. PD0325901. Resuspended in DMSO to final concentration of 10 mM. Aliquots stored at -20°C .
29. Gelatine Type A from porcine skin. Resuspended at 0.2 % in H_2O Milli-Q and sterilized. Stored at $+4^\circ\text{C}$.
30. Ham's F-12 Nutrient Mix 1 \times , supplemented with 2 mM of L-glutamine. Stored at $+4^\circ\text{C}$.
31. Iscove's Modified Dulbecco's Medium (IMDM) 1 \times , supplemented with 2 mM of L-glutamine. Stored at $+4^\circ\text{C}$.
32. Dulbecco's Modified Eagle Medium (DMEM) 1 \times , supplemented with 2 mM of L-glutamine. Stored at $+4^\circ\text{C}$.
33. FBS (Fetal bovine serum) tested for ESC culture. Stock stored at -80°C , while aliquots of 50 mL at -20°C .
34. FBS (Fetal bovine serum). Stock stored at -80°C , while aliquots of 50 mL at -20°C .

35. Chemically Defined (CD) Lipid Concentrate. Aliquots of 10 mL stored at +4 °C.
36. 1-Thioglycerol ≥ 97 %. Aliquots of 50 μ L stored at +4 °C.

3 Methods

All the cell culture work is performed under sterile condition: manipulation of cells and preparation of solutions are done under a vertical laminar flow hood. ESCs are cultured at 37 °C in a humid atmosphere with 5 % of CO₂.

3.1 Serum-Containing Medium

1. For ESC culture, the serum-containing medium is prepared with 85 % DMEM 1 \times (supplemented with 2 mM of L-glutamine), 15 % FBS tested for ESC culture (*see Note 1*), 0.1 mM of 2-mercaptoethanol, and 800 U/mL of LIF. Serum/LIF medium can be kept for 1 month at +4 °C.
2. Inactivating medium: the serum-containing medium to inactivate the Trypsin is prepared with 90 % DMEM 1 \times (supplemented with 2 mM of L-glutamine), 10 % FBS, and 0.1 mM of 2-mercaptoethanol.

3.2 Chemically Defined Medium (CDM)

1. CDM is prepared with 50 % IMDM 1 \times (Supplemented with 2 mM of L-glutamine), 50 % Ham's F-12 Nutrient Mix 1 \times (Supplemented with 2 mM of L-glutamine), 5 mg/mL BSA (*see Note 2*), 1 % CD lipid concentrate, 450 μ M of 1-thioglycerol, 7 μ g/mL recombinant insulin, and 15 μ g/mL transferrin. The CDM is then sterilized by filtering with 0.22 μ m PES membrane filter. CDM can be kept for 1 month at +4 °C.
2. The final ESC 2i culture medium is prepared by adding 3 μ M of CHIR99021, 1 μ M of PD0325901, and 700 U/mL of LIF to CDM. CDM/2i/LIF can be kept for 1 month at +4 °C.

3.3 ESC Cultured in Serum-Containing Medium

1. Incubate dishes with Gelatin 0.2 % (1 mL for dishes of 35 mm) for at least 1 h at 37 °C.
2. Pre-warm ESC culture serum-containing medium, Trypsin-EDTA (0.25 %), and inactivating medium in the water bath at 37 °C.
3. Aspirate gelatin from the dishes with a sterile glass Pasteur pipette.
4. Replace the gelatin by serum-containing medium for ESC (1.5 mL for dishes of 35 mm) and put the dishes in the incubator for equilibration.
5. Aspirate the old medium from the dish with cells at confluence.

6. Quickly wash cells with DPBS kept at room temperature (2 mL for dishes of 35 mm).
7. Add Trypsin-EDTA (0.25 %) to cells (1 mL for dishes of 35 mm) and incubate for 2 min at 37 °C to detach cells from the dishes.
8. Add on top the same volume (1 mL for dishes of 35 mm) of inactivating medium and completely dissociate the cells by pipetting several times with a P1000 pipet.
9. Transfer the cells into a 15-mL plastic tube and centrifuge 5 min at $200 \times g$ at room temperature.
10. Aspirate the supernatant and resuspend thoroughly the visible cell pellet with fresh ESC culture serum-containing medium by pipetting with P1000 pipet.
11. Finally plate the cells in the new dish (*see Note 3*).

3.4 ESC Cultured in CDM/2i Medium

1. Incubate dishes with Laminin (*see Note 4*) diluted directly and freshly in DPBS (1 mL for dishes of 35 mm) at the final concentration of 10 µg/mL for at least 1 h at 37 °C (*see Note 5*).
2. Pre-warm CDM, CDM/2i/LIF, Trypsin-EDTA (0.25 %), and the inactivating medium in the water bath at 37 °C.
3. Remove Laminin from the dishes by aspiration.
4. Replace Laminin by CDM/2i/LIF medium (1.5 mL for dishes of 35 mm) and put the dishes in the incubator for equilibration.
5. Aspirate the old medium from the dish with cells at confluence (ESC in CDM/2i/LIF or ESC in serum/LIF with or without feeder cells, *see Note 6*).
6. Add directly, without DPBS washing step, Trypsin-EDTA (0.25 %) to cells (1 mL for dishes of 35 mm) and incubate for 3–4 min at 37 °C to detach cells from the dishes (*see Note 7*).
7. Add on top the same amount (1 mL for dishes of 35 mm) of inactivating medium and dissociate the cells by pipetting with P1000 pipette.
8. Transfer the cells into a 15-mL plastic tube and centrifuge 5 min at $200 \times g$.
9. Aspirate the supernatant and resuspend the visible cell pellet with fresh CDM by pipetting with P1000 pipette in order to wash cells from serum.
10. Centrifuge another time for 5 min at $200 \times g$.
11. Aspirate the supernatant and resuspend thoroughly the visible cell pellet with fresh CDM/2i/LIF medium by pipetting with P1000 pipette.
12. Finally plate the cells in the new dish (*see Notes 8 and 9*).

3.5 Freezing and Thawing ESC

1. Dissociate the cells with Trypsin and prepare a cell pellet as in Sect. 3.3 (steps 5–9) or Sect. 3.4 (steps 5–8).
2. Remove the supernatant by aspiration and resuspend the visible cell pellet by pipetting with P1000 pipette 1 mL of freezing medium freshly made with 70 % of ESC culture serum-containing medium, 20 % FBS, and 10 % of DMSO for ESC in serum conditions or 60 % of CDM, 30 % FBS, and 10 % of DMSO for ESC in 2i condition.
3. Transfer the cells in freezing medium into a cryotube vial of 1.8 mL and this one into a Freezing container, which is then put for a couple of days at -80°C .
4. Finally transfer the frozen vials into a liquid nitrogen container (-196°C) for long-term storage.
5. Pre-warm CDM, CDM/2i/LIF medium, serum/LIF medium, and inactivating medium in the water bath at 37°C .
6. Thaw the vials with cells in the water bath at 37°C .
7. Add the 1 mL of cells in freezing medium on top of 4 mL of serum-containing medium for ESC in serum conditions, or CDM for ESC in 2i condition, in a 15-mL plastic tube and centrifuge for 5 min at $200 \times g$.
8. Aspirate the supernatant and resuspend the visible cell pellet of with fresh serum/LIF medium or CDM/2i/LIF medium (according to ESC culture conditions) by pipetting with P1000 pipette.
9. Finally plate the cells in the new gelatin- or laminin-coated dish accordingly (*see Note 10*).

4 Notes

1. It is important to test different types of FBS in order to find a batch that leads to optimum cell growth and maintenance of pluripotency of ESCs.
2. It is also necessary to test different batches of BSA for ESC culture to check for optimal growth and absence of differentiation.
3. ESCs in serum-containing medium are usually passaged every 2 days with a dilution of 1/6, if there is a lot of mortality, the medium is changed daily. To let the cells grow for 3 days, dilute them 1/12.
4. ESCs in 2i condition could also be cultured on gelatin-coated dishes, but they will grow as ball-like colonies, sometimes loosely attached and this becomes an issue when performing

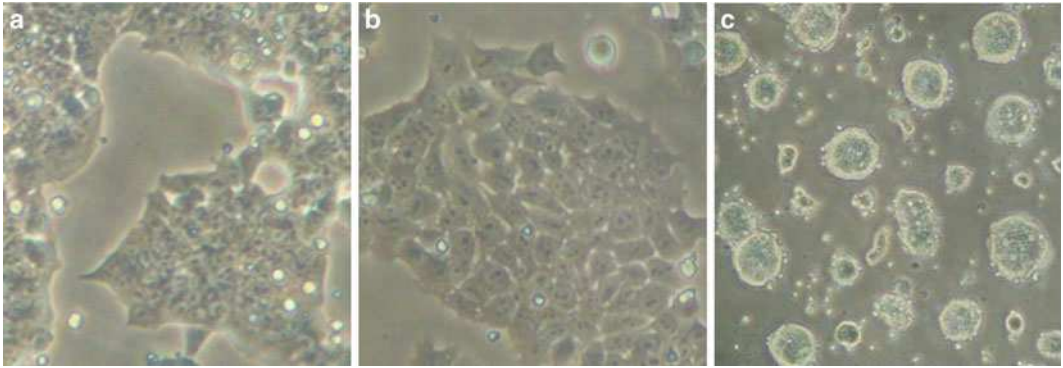


Fig. 1 Morphology of ESCs under the phase contrast microscope. **(a)** ESCs in serum/LIF condition. **(b)** ESCs in CDM/2i/LIF condition on plates coated with laminin (spread and attached colonies). **(c)** ESCs in CDM/2i/LIF condition on plates coated with gelatin (ball-like colonies)

immunofluorescence, for example. On the other hand, Laminin coating leads to full attachment and spreading of these cells, which is convenient for further manipulations (Fig. 1).

5. The aliquot of Laminin should be thawed gently at +4 °C and then diluted directly in DPBS in the dish that needs to be coated. It is not possible to use two times the same diluted Laminin to make the coating of a second dish, as not enough Laminin remains after the first incubation.
6. ESCs in serum condition can also be cultured on a layer of feeder cells. In this case, it is necessary to get rid of feeder cells in order to convert ESC into the ground state. To accomplish this, the pre-plating step is essential: after passing the ESC with feeders, cells are plated twice on noncoated cell culture plates for 20 min before the final plating on laminin-coated dishes. Because ESCs need many hours to attach, during the pre-plating steps, only the feeder cells will have time to attach.
7. When performing the switching, so starting from ESCs in serum (with or without feeders), it is necessary to perform the washing with DPBS and the trypsin treatment should be shorter about 2 min. Otherwise, during normal passage of ESCs that are already in 2i medium, no washing step is required.
8. ESCs in 2i condition are usually passaged every 3 or 4 days with a dilution 1/6 or 1/8 and the medium is changed every 2 days.
9. In first two or three passages after the switching of ESCs from serum/LIF condition to 2i/LIF condition, cells usually appear not completely attached to dished showing some flat colonies and some ball-like colonies typical of standard culture of ESCs in N2B27/2i/LIF on gelatin-coated dishes

(*see* Fig. 1). After 2 weeks of culture of ESCs under 2i/LIF condition, cells are fully attached and we could consider that they have reached the “ground state” (11).

10. Also after freezing and thawing, ESCs in 2i may not be completely attached and spread during the first passage.

Acknowledgments

We thank the ANR Programme Investissements d’Avenir REVIVE for their support. MT is a recipient of a PhD fellowship from DIM Biotherapies Ile-de-France.

References

1. Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. (Translated from eng). *Nature* 292(5819):154–156
2. Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. (Translated from eng). *Proc Natl Acad Sci U S A* 78(12):7634–7638
3. Brook FA, Gardner RL (1997) The origin and efficient derivation of embryonic stem cells in the mouse. (Translated from eng). *Proc Natl Acad Sci U S A* 94(11):5709–5712
4. Smith AG, Heath JK, Donaldson DD et al (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. (Translated from eng). *Nature* 336(6200):688–690
5. Niwa H, Burdon T, Chambers I et al (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. (Translated from eng). *Genes Dev* 12(13):2048–2060
6. Nichols J, Smith A (2009) Naive and primed pluripotent states. *Cell Stem Cell* 4(6):487–492
7. Brons IGM, Smithers L, Trotter M et al (2007) Derivation of pluripotent Epiblast Stem Cells from mouse and rat embryos. *Nature* 448:191–195
8. Tesar PJ, Chenoweth JG, Brook FA et al (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196–199
9. Wray J, Kalkan T, Smith AG (2010) The ground state of pluripotency. (Translated from eng). *Biochem Soc Trans* 38(4):1027–1032
10. Marks H, Kalkan T, Menafrá R et al (2012) The transcriptional and epigenomic foundations of ground state pluripotency. (Translated from eng). *Cell* 149(3):590–604
11. Habibi E, Brinkman AB, Arand J et al (2013) Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. (Translated from eng). *Cell Stem Cell* 13(3):360–369

From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells

Matteo Tosolini and Alice Jouneau

Abstract

Mouse embryonic stem cells (ESCs) derive from the inner cell mass (ICM) of a blastocyst at E3.5 while mouse epiblast stem cells (EpiSCs) derive from the late epiblast of a post-implantation embryo at E5.5–E7.5. Both cells are able to differentiate into derivatives of the three germs layers but only ESCs are able to produce chimeras when they are introduced into a blastocyst. To support the naive state of pluripotency, ESC culture requires Leukemia inhibitory factor (Lif) and either serum or inhibitors of Erk and Gsk3 pathways (2i) while the primed pluripotency of EpiSCs is maintained using Activin A and Fibroblast Growth Factor 2 (FGF2). It is possible to obtain EpiSCs in vitro starting from ESCs but also to induce ESCs starting from EpiSCs even if this second process is very difficult and inefficient. In this protocol we describe how we perform the process of conversion from ESCs to EpiSCs.

Keywords: Mouse embryonic stem cells (ESCs), Epiblast stem cells (EpiSCs), Conversion, Collagenase, Chemically defined medium (CDM)

1 Introduction

From a mouse embryo at different stages it is possible to capture two types of pluripotent cells: mouse embryonic stem cells (ESCs) from the inner cell mass (ICM) of a blastocyst at E3.5 and epiblast stem cells (EpiSCs) from the late epiblast layer of a post-implantation embryo at E5.5–E7.5 (1, 2). Both types of cells are pluripotent as they are able to differentiate into derivatives of all three germ layers either in vitro or in vivo through teratoma, but only ESCs are able to produce chimeras when they are injected into a blastocyst (3). ESCs and EpiSCs share the expression of the core pluripotency factors: Oct4, Nanog, Sox2, but ESCs express some naive pluripotent factor like Rex1 and Klf4 which are absent in EpiSCs, while EpiSCs express some epiblast specific genes such as Fgf5 and Otx2 which are already markers of differentiation absent in ESCs. A strong epigenetic difference between naive and primed state of pluripotency is that female ESCs have two active X chromosome while female EpiSCs present already one inactive X chromosome. To be maintained in vitro, ESCs required serum and Leukemia inhibitory factor (Lif), an activator of the transcriptional factor Stat3 which inhibits

differentiation and promotes self-renewal of naive pluripotency (4). On the other hand, EpiSCs do not respond to Lif while they required Activin A and Fibroblast Growth Factor 2 (FGF2) to sustain primed pluripotency in vitro (2, 5). EpiSCs share with ESCs a large nuclear-to-cytoplasmic ratio and prominent nucleoli, but their morphology is more two-dimensional and epithelial. In addition EpiSCs do not survive efficiently as isolated single cells but they need to be passaged in clumps, all these characteristics of primed pluripotency in mouse are shared with human embryonic stem cells (hESCs) (6). A recent study of our laboratory has shown that the two states of mouse pluripotency differ also in terms of DNA methylation, in particular EpiSCs present a higher proportion of methylation at a subset of gene promoters compared to ESCs and some of these genes are specific to the naive pluripotency (7). It is possible to convert in vitro ESCs into EpiSCs only by changing the culture conditions (8). Interestingly the in vivo transition from naive to primed pluripotency takes only 2 days while in vitro the conversion process needs about a week to obtain stable colonies morphologically similar to EpiSC (Fig. 1). However, a recent study has shown that within 2 days of conversion, the cells extinguish the expression of naive pluripotent genes, while adopting transiently the identity of an early epiblast (9). Moreover it is possible in vitro to induce the opposite process performing a reversion from EpiSCs to ESCs, also by switching the culture conditions, but this process is

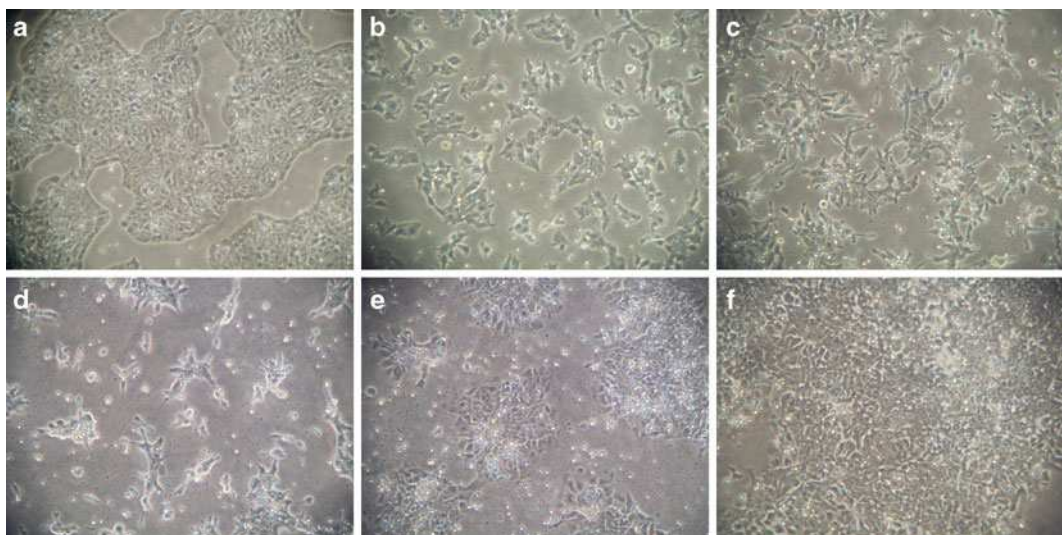


Fig. 1 Morphology of ESCs during conversion to EpiSCs under the phase contrast microscope. (a) ESCs in CDM/2i/LIF. (b) Converted ESCs (cESCs) at day 1 of conversion. (c) cESCs at day 2 of conversion. (d) cESCs at day 3 of conversion (high rate of mortality). (e) cESCs at day 5 of conversion: first appearance of EpiSC-like colonies before the collagenase treatment. (f) cESCs at day 7 of conversion (after the first passage): big EpiSC colonies

inefficient and long, suggesting the presence of an epigenetic barrier as during the reprogramming from somatic to pluripotent cells (10). Conversion from ESC to EpiSC is a useful model to study the molecular and epigenetic events that trigger the transition from the naive to the primed pluripotency (9, 11–13). In the last few years the discovery of two small molecules that inhibit the MAPK and Gsk-3 pathways and could substitute serum in ESCs culture makes a great change in the field (14). With this new completely defined (2i) culture condition, ESCs displayed a state of pluripotency closer to that of the ICM (15, 16). One of the advantages of using a chemically defined medium (CDM) is that the same basal medium can be used to convert ESC to EpiSC, with the only change being the added factors, in this specific case from 2i/LIF to Activin A/FGF2.

2 Materials

2.1 Materials

1. Conical centrifuge tubes 15 and 50 mL, sterile.
2. Graduated plastic pipettes (sterile, single package) of 2, 5, 10, and 25 mL.
3. Glass Pasteur pipettes sterilized in an aluminum container using a dry oven (4 h at 180 °C).
4. Plastic sterile Petri dishes for cell culture of 35 and 60 mm of diameter (*see Note 1*).
5. Pipettes P1000, P200, P20, P10 and sterile plastic tips.
6. 20 mL syringes.
7. Syringe membrane filters, 0.22 µm, in PES (Polyethersulfone).
8. Water bath.
9. Centrifuge (for 15 mL plastic tubes).
10. Incubator at 37 °C in a humid atmosphere with 5 % CO₂.
11. Vertical laminar flow hood.
12. Aspiration system.
13. Phase-contrast microscope and Neubauer chamber or other cell counting method.
14. H₂O Milli-Q produced with a resistivity of 18.2 MΩ cm at 25 °C and sterilized.
15. Dimethyl sulfoxide (DMSO) ≥99 %.
16. Dulbecco's Phosphate-Buffered Saline (DPBS) 1× sterile without Ca²⁺ and Mg²⁺.
17. Trypsin–EDTA (0.25 %). Aliquots of 10 mL stored at –20 °C.

18. Collagenase from *Clostridium histolyticum*. Powder stored at -20°C .
19. Protease-free BSA (Bovine serum albumin). Powder stored at $+4^{\circ}\text{C}$.
20. 2-mercaptoethanol (50 mM). Aliquots of 1 mL stored at $+4^{\circ}\text{C}$.
21. Transferrin. Resuspended in H_2O Milli-Q at the final concentration of 30 mg/mL. Aliquots stored at -20°C .
22. Recombinant Insulin. Resuspended in H_2O Milli-Q at the final concentration of 10 $\mu\text{g}/\text{mL}$. Aliquots stored at -20°C .
23. Recombinant Fibroblast Growth Factor-basic (FGF2). Resuspended in DPBS with 0.1 % BSA and 1 mM DTT at the final concentration of 12 $\mu\text{g}/\text{mL}$. Aliquots of 20 μL stored at -20°C .
24. Activin A. Resuspended in DPBS with 0.1 % BSA at the final concentration of 20 $\mu\text{g}/\text{mL}$. Aliquots of 20 μL stored at -20°C .
25. Ham's F-12 Nutrient Mix $1\times$, supplemented with 2 mM of L-glutamine. Stored at $+4^{\circ}\text{C}$.
26. Iscove's Modified Dulbecco's Medium (IMDM) $1\times$, supplemented with 2 mM of L-glutamine. Stored at $+4^{\circ}\text{C}$.
27. Dulbecco's Modified Eagle Medium (DMEM) $1\times$, supplemented with 2 mM of L-glutamine. Stored at $+4^{\circ}\text{C}$.
28. FBS (Fetal bovine serum). Stock stored at -80°C , while aliquots of 50 mL at -20°C .
29. Chemically Defined (CD) Lipid Concentrate. Aliquots of 10 mL stored at $+4^{\circ}\text{C}$.
30. 1-Thioglycerol $\geq 97\%$. Aliquots of 50 μL stored at $+4^{\circ}\text{C}$.

2.2 Collagenase II Solution

Collagenase II solution is prepared with: 50 % IMDM $1\times$ (supplemented with 2 mM of L-glutamine), 50 % Ham's F-12 Nutrient Mix $1\times$ (supplemented with 2 mM of L-glutamine), 3.5 mg/mL of collagenase from *Clostridium histolyticum* (see **Note 2**). The solution is then sterilized by filtering with 0.22 μm PES membrane filter. Collagenase II solution can be kept for 1 month at $+4^{\circ}\text{C}$.

2.3 Chemically Defined Medium (CDM)

CDM is prepared with: 50 % IMDM $1\times$ (supplemented with 2 mM of L-glutamine), 50 % Ham's F-12 Nutrient Mix $1\times$ (supplemented with 2 mM of L-glutamine), 5 mg/mL BSA (see **Note 3**), 1 % CD lipid concentrate, 450 μM of 1-thioglycerol, 7 $\mu\text{g}/\text{mL}$ recombinant insulin, 15 $\mu\text{g}/\text{mL}$ transferrin. The CDM is then sterilized by filtering with 0.22 μm PES membrane filter. CDM can be kept for 1 month at $+4^{\circ}\text{C}$.

2.4 Serum-Containing Medium

This medium will be used to inactivate the trypsin and as a source of extracellular matrix to coat the dishes for converting cells. It contains a basal medium such as DMEM, IMDM or F12 supplemented with 10 % FBS. It can be kept for 2 months at 4 °C.

3 Methods

All the cell culture work is performed under sterile condition: manipulation of cells and preparation of solutions are done under a vertical laminar flow hood. ESCs and EpiSCs are cultured at 37 °C in a humid atmosphere with 5 % of CO₂.

3.1 Conversion from ESCs to EpiSCs

1. Incubate 35 mm dishes with 1 ml of serum-containing medium for at least 1 h at 37 °C (*see Note 4*).
2. Pre-warm Trypsin–EDTA (0.25 %), serum-containing medium and CDM in the water bath at 37 °C.
3. Aspirate serum-containing medium from the dishes with a sterile glass Pasteur pipette.
4. Perform a DPBS wash (1 mL for dishes of 35 mm).
5. Replaced DPBS by CDM (1.5 mL for dishes of 35 mm) supplemented freshly with 12 µg/mL of FGF2 and 20 µg/mL of Activin A and put the dishes in the incubator for equilibration.
6. Aspirate the old medium from the dishes with ESCs (*see Note 5*).
7. Quickly wash cells with DPBS kept at room temperature (*see Note 6*).
8. Add Trypsin–EDTA (0.25 %) on ESCs and incubate at 37 °C for 2 or 3 min according to their condition, serum or 2i respectively, to detach cells from the dishes.
9. Add on top the same volume of serum-containing medium and completely dissociate the cells by pipetting several times with a P1000 pipet.
10. Take a little amount of solution with ESCs singularized to count them.
11. Transfer the cells into a 15 mL plastic tube and centrifuge 5 min at 200 g at room temperature.
12. Aspirate the supernatant and thoroughly resuspend the visible cell pellet with 1 mL of CDM by pipetting with P1000 pipette and centrifuge another time for 5 min at 200 g at room temperature.
13. Aspirate the supernatant and resuspend ESCs in the appropriate amount of CDM according to the number of cells counted.
14. Finally plate 1.5 millions of ESCs in the new dishes coated with serum or fibronectin and containing CDM supplemented with FGF2 and Activin A.

15. Every day replace the old medium with fresh CDM supplemented with the factors (*see* **Note 7**).
16. At the fifth day of conversion there is normally the appearance of first EpiSCs-like colonies (already quite flat or more ball-like ones) and they need to be passaged in order to help them to spread (Fig. 1). If the colonies are still small, wait one more day before passaging.

3.2 Passaging of EpiSCs Colonies

1. Prepare the new plates as in 3.1 (**steps 1, 3–5**).
2. Pre-warm collagenase II solution and CDM in the water bath at 37 °C.
3. Aspirate the old medium from the dishes.
4. Wash cells with DPBS kept at room temperature.
5. Add collagenase II solution on the EpiSC-like cells (400 µL for dishes of 35 mm) and incubate for 30 s at room temperature (*see* **Note 8**).
6. Wash with DPBS.
7. Flush the EpiSC-like colonies using 1 mL of CDM and a P1000 pipette in order to detach them in clumps without pipetting and singularizing cells.
8. Transfer the cells into a 15 mL plastic tube (*see* **Note 9**) and centrifuge 5 min at 200 g at room temperature.
9. Aspirate the supernatant and resuspend very gently the cells in CDM just with one or two pipetting of P1000 pipette in order to keep EpiSCs colonies in clumps and not single cells.
10. Transfer EpiSC colonies in the newly prepared dishes (*see* **Note 10**).

4 Notes

1. Test different types of plastic petri dishes for EpiSCs culture to check for optimal growth and absence of differentiation along passages.
2. To help dissolution of collagenase from *Clostridium histolyticum* leave the solution few minutes in the water bath at 37 °C before filtering it. The warm solution is naturally cloudy.
3. It is necessary to test different batches of BSA for EpiSCs culture to check for optimal growth and absence of differentiation.
4. The serum-containing medium can be replaced by fibronectin: use a dilution of 20 µg/ml in DPBS and incubate the dishes at least 20 mn at 37 °C.
5. The conversion can be performed starting from ESCs either cultured in 2i/Lif or in serum/Lif conditions. However,

conversion seems to induce less mortality when starting from ESCs cultured in 2i/Lif.

6. This step is necessary only if the ESCs are in serum condition.
7. During conversion there is a high rate of cell mortality so in this case it is better to wash with DPBS before changing the medium in order to eliminate the maximum of dead cells.
8. When the collagenase II solution is fresh-made, it is very active; so even 20 s will be sufficient to detach EpiSCs colonies from dishes.
9. Repeat 3.2 **step** 7 more than once using another 1 mL of CDM in order to detach the maximum of EpiSCs colonies trying not to break them too much.
10. For the first passaging at the fifth day of conversion the suspension is not diluted as there is still a lot of mortality. For the next passages dilution will be 1/2 then 1/3 every 2–3 days approximately.

Acknowledgments

We thank the ANR Programme Investissements d'Avenir REVIVE for their support. MT is a recipient of a PhD fellowship from DIM Biotherapies Ile-de-France.

References

1. Evans M, Kaufman MH (1981) Establishment in culture of pluripotent cells from mouse embryos. *Nature* 292:154–156
2. Brons IGM, Smithers LE, Trotter MWB et al (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448 (7150):191–195
3. Rossant J (2008) Stem cells and early lineage development. (Translated from eng). *Cell* 132 (4):527–531
4. Niwa H, Burdon T, Chambers I et al (1998) Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev* 12(13):2048–2060
5. Tesar PJ, Chenoweth JG, Brook FA et al (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448:196–199
6. Nichols J, Smith A (2012) Pluripotency in the embryo and in culture. (Translated from eng). *Cold Spring Harb Perspect Biol* 4(8):a008128
7. Veillard AC, Marks H, Bernardo AS et al (2014) Stable methylation at promoters distinguishes epiblast stem cells from embryonic stem cells and the in vivo epiblasts. (Translated from eng). *Stem Cells Dev* 23(17):2014–2029
8. Guo G, Yang J, Nichols J et al (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* 136:1063–1069
9. Buecker C, Srinivasan R, Wu Z et al (2014) Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* 14(6):838–853
10. Bao S, Tang F, Li X et al (2009) Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* 461:1292–5
11. Factor Daniel C, Corradin O, Zentner Gabriel E et al (2014) Epigenomic comparison reveals activation of “seed” enhancers during transition from naive to primed pluripotency. *Cell Stem Cell* 14(6):854–863
12. Hayashi K, Ohta H, Kurimoto K et al (2011) Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. (Translated from eng). *Cell* 146 (4):519–532

13. Schulz EG, Meisig J, Nakamura T et al (2014) The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. (Translated from eng). *Cell Stem Cell* 14 (2):203–216
14. Wray J, Kalkan T, Smith AG (2010) The ground state of pluripotency. (Translated from eng). *Biochem Soc Trans* 38(4):1027–1032
15. Marks H, Kalkan T, Menafrá R et al (2012) The transcriptional and epigenomic foundations of ground state pluripotency. (Translated from eng). *Cell* 149(3):590–604
16. Habibi E, Brinkman AB, Arand J et al (2013) Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. (Translated from eng). *Cell Stem Cell* 13(3):360–369

Contrasting epigenetic states of heterochromatin in the different types of mouse pluripotent stem cells

Tosolini Matteo¹, Brochard Vincent¹, Adenot Pierre¹, Chebrout Martine¹, Grillo Giacomo², Navia Violette¹, Piumi François¹, Beaujean Nathalie^{1,3}, Francastel Claire², Bonnet-Garnier Amélie^{1*} and Jouneau Alice^{1*}

1 UMR BDR, INRA, ENVA, Université Paris Saclay, 78350, Jouy en Josas, France

2 UMR7216 Epigenetics and cell fate, Université Paris Diderot Paris 7, 75013, Paris, France

3 present address : Université de Lyon 1, Inserm, Stem Cell and Brain Research Institute U1208, 69500, Bron, France

* corresponding authors: amelie.bonnet-garnier@jouy.inra.fr; alice.jouneau@jouy.inra.fr

ABSTRACT

Mouse embryonic stem cells (ESCs) and epiblast stem cells (EpiSCs) represent naïve and primed pluripotency states, respectively and are maintained *in vitro* using specific signalling pathways. Furthermore, ESCs cultured in serum-free medium with two kinase inhibitors (2i medium) are described as being the most naïve pluripotent state. Although several studies suggest that different epigenome organizations characterize each pluripotent state, no such comparison has yet been made concerning pericentromeric heterochromatin (PCH). Here we present a comparative study of the epigenetic and transcriptional state of PCH sequences in the distinctive naïve and primed pluripotency states. We show that the pattern of H3K27me3 is highly dynamic and discriminate the most naïve 2i-ESCs from the others. Whereas transcription is high in serum-ESCs, it is lower in 2i ESCs and even more repressed in the primed EpiSCs. Removal of either DNA methylation or H3K9me3 in ESCs leads to enhanced deposition of H3K27me3 but few changes in satellite transcription. By contrast, in EpiSCs, removal of H3K9me3 does not prevent DNA methylation but lift transcriptional repression in EpiSCs. Altogether our study reveals that PCH in mouse pluripotent stem cells display distinct features according to the pluripotency state and culture conditions.

INTRODUCTION

Pluripotency is defined by the ability to generate cells belonging to the three embryonic lineages. Cells derived *in vitro* in the Mouse have allowed distinguishing a naïve from a primed state of pluripotency ¹. These two states functionally differ by their ability to produce chimeras. Naïve embryonic stem cells (ESCs) colonize the inner cell mass (ICM) when injected into blastocyst and eventually allow generation of full-term chimeras. By contrast, primed epiblast stem cells (EpiSCs) only colonize the post-implantation epiblast and their contribution to later development remains to be demonstrated ². These two states are maintained *in vitro* using specific signalling pathways, mainly Lif/Stat3 for the naïve cells and FGF and Activin for the primed cells ^{3,4}. ESCs are classically maintained in serum-containing medium but can also be cultured in serum-free medium with Lif and inhibitors of two differentiation pathways, the MAPK/ERK (mitogen-activated protein kinases / Extracellular signal-regulated kinases) and the GSK3 (Glycogen synthase kinase 3) pathways ⁵. In this 2i/Lif medium (later referred as “2i” only), cells acquire an even more naïve state, with efficient repression of lineage commitment markers and homogenous expression of pluripotency genes ^{6,7}. By contrast, the EpiSC transcriptome reflects their primed nature, as they already express many lineage commitment markers while some pluripotency genes are down-regulated ^{3,8}. ESCs cultured in serum/Lif (serum-ESCs) display somehow an intermediate transcriptome, with heterogeneous levels of pluripotency markers and low but detectable expression of differentiation genes ^{6,7}.

ESCs are considered to have a more open and plastic chromatin than differentiated cells ^{9–12}. Indeed, the epigenome is rapidly and reversibly modified according to culture medium, 2i or serum ^{7,13,14}. Specifically, 2i-ESCs exhibit significantly reduced H3K27me3 deposition at promoters and a globally DNA hypomethylated state of the genome ^{7,13}. ESCs can be converted into EpiSCs (cEpiSCs) *in vitro* when exposed to FGF and Activin signalling instead of Lif and serum or inhibitors ¹⁵. By contrast, the reversion of EpiSCs into naïve cells is a long and inefficient process, eliciting the notion of epigenetic barrier ¹⁶. Although fewer in-depth analyses have been reported, available data indicate that during conversion to the primed state, many promoters become DNA hypermethylated together with substantial reorganization of enhancer landscapes, relative to ESCs ^{17,18}.

Studies mentioned above suggest that the epigenome organization characterizes each pluripotent cell type. Such comparison has not yet been made concerning the constitutive heterochromatin compartment. This compartment is composed of repeated DNA sequences located at telomeres, centromeres and pericentromeric regions ¹⁹. The proper control of these regions is crucial for chromosomal stability ²⁰. Beside telomeric sequences, there are two types of repeats in mouse: major and minor satellites at the pericentromeric (PCH) and centromeric (CH) heterochromatin, respectively ²¹. The major satellite consists of a 234bp sequence repeated over 200,000 times, representing about 3% of the mouse genome, while the minor satellite unit is a 123pb sequence

repeated at least 50,000 times ²². In somatic cells, both PCH and CH from different chromosomes aggregate in clusters (chromocenters), which densely stain for DAPI, due to their high AT-rich content, and are typically enriched in repressive SUV39H1/2-mediated H3K9me3 marker ²³. DNA methylation is another hallmark of constitutive heterochromatin that coexists with H3K9me3 ²⁴. 5-methylcytosine (5-meC) is a repressive mark specifically *de novo* deposited by DNMT3A/B, while being maintained throughout cell division by DNMT1 ^{25,26}. Such repressive epigenetic state is not favourable to transcription; hence in somatic cells satellite repeats are poorly expressed ²⁷. However, in cellular senescence, in some cancers and during early development, transcriptional activation has been observed, usually coincident with reduced DNA methylation at these sequences (reviewed in ²⁸). In serum ESCs, in line with the concept of an open chromatin, and although major satellites are also enriched in H3K9me3 and 5-meC ^{23,25}, transcription is higher compared to differentiated cells such as neural progenitors ²⁹. If one or the other constitutive heterochromatin mark is missing, as for *Suv39h* or *Dnmts* knock-out ESCs, PCH harbours the typical facultative heterochromatin mark H3K27me3 ^{23,30,31}. Both 5-meC and H3K27me3 are largely reduced and redistributed upon switching to 2i medium, yet the PCH organization has not been yet studied in 2i-ESCs. In addition, whether the organization of PCH and the transcription state are conserved beyond the epigenetic barrier in the primed EpiSCs is not known.

To address this question, we present a comprehensive and comparative study of the epigenetic and transcriptional state of PCH sequences in the distinctive naïve and primed pluripotency states. We show that PCH in naïve 2i-ESCs harbours an unusual constitution, with high enrichment of H3K27me3, at the expense of both H3K9me3 and DNA methylation. Despite a relaxed global structure, the satellite repeats are transcribed at a much lower level than in serum-ESCs. By contrast, PCH in primed EpiSCs harbours somatic like features, with dense deposition of H3K9me3 and 5-meC and repressed transcription of satellite repeats. Noteworthy, in EpiSCs, deposition of DNA methylation at PCH is independent of H3K9me3. At last we also show that transcription of major satellites is largely independent on the presence of DNA methylation and H3K9me3 in ESCs whereas it depends on H3K9me3 in EpiSCs. Altogether our study reveals that PCH in mouse pluripotent stem cells display distinct features according to the pluripotency state and culture condition and has important implications in the definition of the naïve state.

RESULTS

Heterochromatin domains are characterized by different epigenetic histone marks depending on the pluripotent state.

To characterize and compare the nuclear distribution of H3K9me3 and H3K27me3 histone marks at PCH/CH regions in ESCs (serum and 2i conditions) and EpiSCs, we performed immunostaining of both marks along with DAPI staining to detect chromocenters. We assessed the percentage of cell with H3K9me3 or H3K27me3 enrichment at these foci and determined their colocalization with DAPI staining using plot profile across nuclei. Our data show that the distribution of the two marks was different according to the cell type. In 2i-ESCs, only one third of the population display H3K9me3-enriched PCH/CH foci (Fig. 1A). In such cells, H3K9me3-enriched foci were rare and small, and located close to the nuclear periphery (see the single nucleus magnification in Fig. 1A). Conversely all serum-ESCs and EpiSCs showed numerous H3K9me3 foci, perfectly co-localizing with DAPI-dense foci (Fig. 1A).

Regarding H3K27me3, we also observed very distinctive nuclear distribution between ESCs and EpiSCs. Indeed, in 2i-ESCs, the majority of cells (69%) presented H3K27me3 enrichment at DAPI-dense PCH/CH foci (Fig. 1A, see the single nucleus magnification). The distribution within the nucleus suggests that H3K27me3 substitutes for H3K9me3 at DAPI-dense foci. By contrast, the pattern of H3K27me3 was diffuse in most (95%) serum-ESCs, as expected for a facultative heterochromatin mark (Fig. 1A). In the primed EpiSCs, H3K27me3 signal was also diffuse and very low (Fig. 1A).

To further explore these differences, we assessed the bulk level of the two histone marks in cell extracts from the three cell types. Quantification by western-blot indicated a slight decrease of the global level of H3K9me3 in 2i-ESCs compared to serum-ESCs and EpiSCs, and a much more pronounced loss of H3K27me3 in EpiSCs (Fig. 1B). Hence, in the primed EpiSCs, H3K27me3 was reduced not only at PCH but also elsewhere in the genome.

We then examined the expression of Suv39h enzymes and found that, both mRNA and protein levels of Suv39h1 increased considerably between 2i-ESC and serum-ESCs or EpiSCs (Supplementary Fig. 1A). Accordingly, immunostaining revealed that SUV39H1 was diffuse in 2i-ESCs while accumulating in foci in 31% of the serum and in the vast majority (93%) of EpiSCs (Supplementary Fig. 1A). We then investigated the distribution and expression of EZH2, the enzyme of the PRC2 complex that deposits H3K27me3. While transcription remains constant between cells types, the protein level of EZH2 was strongly reduced in EpiSCs relative to ESCs (Supplementary Fig. 1B). Yet, EZH2 distribution was unchanged, as we observed by immunostaining the same diffuse signal with some bright and tiny spots (Supplementary Fig. 1B).

Altogether our data suggests that H3K27me3 is redistributed from facultative to constitutive heterochromatin when ESCs are transferred from serum to 2i medium, while being lost in EpiSCs. Conversely, H3K9me3 shows lower levels and mostly dispersed organization in 2i-ESCs, and higher-levels with strong enrichment at DAPI-dense foci in serum-ESCs and EpiSCs.

To strengthen these observations, we examined mitotic cells, as in the mouse acrocentric chromosomes, the PCH/CH domains are localized at one end. H3K27me3 was indeed clearly enriched at the PCH/CH in 2i-ESCs, while H3K9me3 was present all along the chromosomes (Fig. 1C). Conversely, in serum-ESCs and EpiSCs the ends of mitotic chromosomes were strongly enriched in H3K9me3, while H3K27me3 presented a continuous staining in serum-ESCs or was undetectable in EpiSCs (Fig. 1C).

Finally, we used previously published H3K27me3 ChIP-seq datasets to quantify the reads mapping specifically on the major and minor satellite repeats. With the first dataset ⁷ we observed a 2-fold enrichment in 2i-ESC compared to serum-ESC for H3K27me3 at major satellites, while no difference was found for minor satellite (Supplementary Fig. 1C). Another dataset that allowed comparing 2i-ESCs and EpiSCs ³², showed that 2i-ESCs presented a more than 2-fold enrichment of H3K27me3 on major satellites compared to input, but no enrichment at all in EpiSCs (Supplementary Fig. 1D). Conversely minor satellites showed no H3K27me3 enrichment in any cell type and therefore are not concerned by the epigenetic switch.

Low level of DNA methylation at PCH in 2i-ESCs correlates with low level of *Dnmt3s* enzymes

To further dissect the organization of the PCH in the different pluripotent cells, we examined by Southern-blot the DNA methylation level at major satellite sequences. We revealed impressive changes at these sequences between the different states of pluripotency (Fig. 2A). Major satellites were partially demethylated in 2i-ESCs, as shown by the linescan profile (red line) which is intermediate between fully demethylated *Dnmt*TKO cells (purple line) and hypermethylated fibroblasts (MEFs - black line) (Fig. 2A). On the contrary, these sequences in both serum-ESCs (blue line) and EpiSCs (green line) were as methylated as the MEFs. A similar situation was observed for minor satellites that were partially demethylated in 2i-ESCs and hypermethylated in serum-ESCs and EpiSCs (Fig. 2B).

We then assessed the expression level of the two main enzymes responsible for *de novo* methylation: DNMT3A and DNMT3B (Supplementary Fig. 2). *Dnmt3a* expression level increased strongly between 2i-ESCs and serum-ESCs and even more in EpiSCs. In parallel the embryonic isoform of the DNMT3A protein was about ten times more abundant in serum-ESCs than in 2i-ESCs, while EpiSCs gained the additional somatic form ^{18,33} (Supplementary Fig. 2A). *Dnmt3b* was transcribed at low level in ESCs (2i- or serum-) and this level strongly increased in EpiSCs (Supplementary Fig. 2B).

DNMT3B protein level showed a sequential increment going from 2i- to serum-ESCs and finally EpiSCs.

In conclusion we observed a progressive methylation of satellites, accompanied by an increased expression of the de novo methyltransferases, from 2i-ESCs to EpiSCs.

PCH is decondensed in 2i ESCs but transcriptionally repressed

To assess the spatial organization of chromocenters depending on the state of pluripotency, we performed DNA-FISH using specific probes for major and minor satellites sequences (Fig. 3A), followed by nucleus segmentation and 3D single-cell reconstruction with AMIRA 3.1 software (Fig. 3A). In EpiSCs and serum-ESCs major satellites sequences were organized into round compact domains, surrounded by smaller dots of minor satellite domains, as classic chromocenters in somatic cells²¹. By contrast, major satellites formed decondensed domains of irregular size and shape in 2i-ESCs (blebs, half-rings around nucleolus or at the nuclear periphery), reflecting globally unstructured chromocenters. A similar decondensation has been previously observed as a consequence of down-regulation of the histone chaperone CAF-1/p150 in ESCs³⁴. The presence of this protein was assessed by immunostaining (Supplementary Fig. 3) and we observed a progressive increase in the proportion of cells displaying CAF1/p150 positive foci from 2i-ESCs (25 %) to serum-ESCs (33%) and to EpiSCs (42%). This suggests a role for this CAF-1 complex in the decondensation of major satellites in 2i-ESCs.

Such a decompacted PCH is expected to create a transcriptionally permissive environment. We therefore evaluated the level of expression of the associated transcripts by qRT-PCR on ESCs and EpiSCs (Fig. 3B). Surprisingly, 2i-ESCs expressed major satellites at lower level than serum-ESCs. In primed EpiSCs, expression was even lower than in 2i-ESCs. The same trend was observed for minor satellites.

In conclusion, 2i-ESCs and EpiSCs similarly silence their PCH/CH regions despite different mode of repression and level of compaction. On the other hand, serum-ESCs are largely permissive for satellite transcription, despite prominent deposition of H3K9me3 and 5mC at these loci.

Absence of Suv39h1/2 induces different phenotypes depending on the pluripotent cell state

To decipher the cross-talk between the three repressive marks in the different pluripotent states at PCH, we examined their distribution upon removal of H3K9me3. For that we used the *Suv39h*dn mutant model, in which both Suv39h enzymes were knocked-out²³. We adapted these ESCs in 2i culture conditions and converted them *in vitro* in EpiSCs (cEpiSCs, supplementary Fig. 4).

In 2i adapted mutant ESCs, no obvious modification in the pattern of H3K9me3 could be observed, relative to wild-type cells. However, a significant proportion of the cells (22%) still

exhibited small H3K9me3 enriched foci, indicating an H3K9me3-deposition at PCH independent of SUV39H1/2 enzymes (Fig.4A, compare with Fig.1A). Moreover, the same enrichment of H3K27me3 at PCH and a very similar, hypomethylated DNA methylation profile, were observed (Fig. 4A and B). Accordingly, the transcription level of major satellite sequences was overall unchanged in mutant and wild-type cells. Hence, in the 2i condition, the organization of PCH is largely independent of the deposition of H3K9me3 by SUV39H1/2.

In serum-ESCs H3K9me3 staining revealed a diffuse pattern (with 95% of cells presenting no foci-enrichment), as expected (Fig. 4A). It was previously shown that H3K27me3 could substitute for H3K9me3 at PCH foci in the mutant serum-ESCs ^{22,23}. Here we confirmed this phenotype, as 60% of *Suv39h*dn serum-ESCs presented H3K27me3 at the DAPI-dense foci (Fig. 4A). Next, we showed that the absence of *Suv39h1/2* in serum-ESCs induced a reduction of the methylation compared to wild-type on major satellites (Fig. 4B – Note the shift of the blue dotted line compared to the continuous one), confirming previous findings ³¹. In spite of the loss of a repressive mark, transcription of major satellites was slightly decreased in *Suv39h*dn serum ESCs (Fig. 4C). Hence, the absence of *Suv39h1/2* in serum-ESCs induces a phenotype that recapitulates the 2i culture condition: increased H3K27me3 and reduced DNA methylation, as well as lower transcription at PCH (although not as low as in 2i-ESCs)..

In the primed EpiSCs, we observed a contrasting situation. Mutant cells have completely lost H3K9me3 PCH foci, but in contrast to ESCs, they do not gain any H3K27me3 foci (Fig. 4A). Hence, in cEpiSCs, H3K27me3 does not substitute for H3K9me3 at PCH foci and their global level remains accordingly low (data not shown). In addition, major satellites were similarly methylated in wild-type and mutant cells (Fig. 4B – the green lines are quite superimposed). Very interestingly in *Suv39h*dn cEpiSCs, transcription of major satellites was de-repressed, up to the level of that seen in the mutant serum-ESCs (Fig. 4C). Altogether, while in serum-ESCs the absence of H3K9me3 at PCH leads to reduced DNA methylation and accumulation of H3K27me3, in mutant EpiSCs DNA methylation is restored but this does not prevent de-repression of major satellite transcription.

Absence of DNA methylation increases deposition of H3K27me3 but has limited effects on satellite transcription

To study the effects of DNA methylation in the epigenetic pathway involved in regulation of major satellite organization and transcription, we used *Dnmt1*, *3a*, and *3b* triple knock-out ESCs (*Dnmt*TKO). These cells do not have any methylated cytosine in their genome ³⁵ and notably at PCH (see Fig. 2A). We were able to adapt these cells in 2i but not to convert them into cEpiSC because of cell death (apoptosis), in agreement with their low contribution to the development of the epiblast when transferred in embryos ³⁶. *Dnmt*TKO serum-ESCs showed strong enrichment of H3K27me3 at PCH foci

(Fig. 5A), as expected from previous studies^{30,37}. Indeed we found that most (84%) mutant serum-ESCs presented H3K27me3 enrichment at DAPI-dense foci. This enrichment was at the expense of H3K9me3 (Fig. 5A), as 100% of wild-type serum-ESCs contain H3K9me3 foci and only 52% of mutant cells (Fig. 5A and Fig. 1A). In 2i condition, the complete absence of DNA methylation led to an increased proportion of cells with H3K27me3 marked foci (94% 2i-*Dnmt*TKO vs 69% in wild-type, see Fig. 5 and Fig. 1A). The pattern of H3K9me3 remained very similar in 2i-*Dnmt*TKO compared to wild-type, meaning small and rare foci (Fig. 5A and Fig. 1A). Transcription of major satellites was almost unchanged (Fig. 5B). In summary the absence of DNA methylation dramatically modifies the epigenetic state of PCH in serum-ESCs, inducing a strong enrichment of H3K27me3 and a reduction of H3K9me3-enriched foci, but no clear effect on major satellite transcription.

Reduced levels of H3K27me3 do not up-regulate satellite transcription

In order to study the role of H3K27me3 in the regulation of PCH transcription status, we used an inhibitor of the methyl-transferase activity of EZH2. We chose EPZ-6438 (further call for simplicity EPZ), a selective EZH2 inhibitor³⁸. The treatment of ESCs with EPZ for 72 hours led to an impressive reduction of the bulk levels of H3K27me3 (at least 70% in each cell types tested), with no changes in H3K9me3 levels (Fig. 6A-E). Immunostaining of 2i-ESCs treated with EPZ also confirmed a loss of H3K27me3 foci in the vast majority of cells with no major changes in H3K9me3 organization compared to DMSO-treated control cells (Fig. 6A-right panel). In addition the hypomethylated status of satellite sequences in 2i-ESCs treated cells was not affected, showing that there was no replacement of H3K27me3 by DNA methylation at PCH (Data not shown). We then analysed the transcription of major satellites. Unexpectedly, the loss of H3K27me3 in 2i-ESCs did not lead to an up-regulation of transcription, but on the contrary, to a slight reduction (Fig. 6A). When serum-ESCs were treated with the EZH2 inhibitor, although H3K27me3 is not as highly enriched at PCH as in 2i condition, we also observed a reduction of major satellite transcripts (Fig. 6B). We also performed the same inhibition in mutant cells that present a similar enrichment of H3K27me3 at PCH (namely *Suv39*hdn serum-ESCs and both *Dnmt*TKO ESCs). In all cases, we also observed a slight reduction of major satellite transcription (Fig. 6C-E). To exclude a possible effect of EPZ treatment on pluripotency of ESC, we verified that the expression level of pluripotency genes was not altered (Data not shown). Altogether, the reduction of H3K27me3 at PCH does not induce an up-regulation of major satellite transcripts even in absence of the other repressive marks (such as H3K9me3 or DNA methylation) at the same sequences.

DISCUSSION

Cross-talk between H3K9me3, H3K27me3 and 5-meC in mouse pluripotency

Here we show that the interplay between different repressive marks at PCH is modulated according to the pluripotency state and culture condition (Fig. 7). We have observed that upon reduction of 5-meC, H3K27me3 becomes enriched at PCH in ESCs, in three different contexts: (i) after 2i adaptation of ESCs, which leads to a general demethylation at the DNA level ¹³ ; (ii) in the *Suv39h*dn condition, which induces in serum-ESCs partial DNA demethylation of major satellites, as previously observed ^{22,23,30,37} ; (iii) in the total absence of 5-meC (*DnmtTKO*) in serum-ESCs (also shown in ³⁷). Elsewhere in the genome, DNA methylation also antagonizes H3K27me3 deposition, as shown in *DnmtTKO* or 5-aza treated ESCs, where this mark is redistributed towards demethylated CpG sites, at the expense of PRC2 canonical target sites ^{30,39}. Conversely, DNA methylation cannot replace H3K27me3, as in 2i-ESCs treated with the EZH2 inhibitor we did not evidence any increase of 5meC at major satellites. This is in agreement with the relatively unchanged pattern of DNA methylation at promoters in PRC2 mutant ESCs ³⁹. Upon switching ESCs into 2i medium, it was shown that there was a considerable loss of H3K27me3 at gene promoters ⁷. Therefore, while this mark is reduced at unique sequences, we now show that it is redeployed at PCH in 2i-ESCs, making these cells an interesting model of the cross-talk between PRC regulation of gene expression and PCH epigenetic state ²⁴.

It was shown previously that in absence of DNA methylation, H3K9me3 and H3K27me3 coexist within the constitutive heterochromatin, but either in different domains, or in non-overlapping sequences ^{30,40}. We confirm this observation and also show that H3K27me3 takes over H3K9me3 in a context of reduced DNA methylation. Indeed, in *DnmtTKO* serum ESCs, half of the population display only H3K27me3-enriched foci, and in cells that keep H3K9me3 foci, these do not overlap H3K27me3.

In 2i-ESCs, H3K9me3 enriched foci do not accumulate SUV39H1 and also persist in absence of *Suv39h* enzymes. We speculate that they may depend on other histone methyltransferase (KMT) enzymes, such as SETDB1. Previous data have indeed evidenced a reduction of H3K9me3 not only at euchromatin but also at PCH in absence of SETDB1 ^{41,42}.

EpiSC: a pluripotent cell with a somatic epigenetic state

We showed that, like somatic cells, EpiSCs present high levels of H3K9me3, 5-meC, and SUV39H1 at DAPI-dense foci, and compacted chromocenters, in agreement with transcriptionally inactive sequences. In addition, Nanog expression is reduced in EpiSCs and in a very recent study, Novo et al

⁴³ have demonstrated that its over-expression in EpiSCs induces the up-regulation of major satellite transcripts and decompaction of H3K9me3 foci.

It has been shown that DNMTs are recruited at H3K9me3-enriched PCH foci in ESCs, probably through interaction with SUV39H/HP1 ^{31,44}. Our data show that in the primed EpiSCs such interaction is not necessary, as *Suv39hdn* converted cells regain DNA methylation at PCH sequences. The de-repressed status of major satellites in *Suv39hdn* cEpiSCs indicate that H3K9me3 plays a key role in silencing PCH sequences in these cells and that high DNA methylation and low NANOG level are not sufficient to maintain a repressive state at PCH in EpiSCs.

Uncoupling epigenetic state of ESCs with transcription regulation of satellites

We showed that serum-ESCs present a repressive state at the constitutive heterochromatin with high level of H3K9me3, 5-meC and compacted chromocenters. Paradoxically transcription of major and minor satellites is variable but globally elevated. Such a repressive environment is expected to be unfavourable to transcription (reviewed in ⁴⁵), as it is indeed the case in EpiSCs. On the other hand, 2i-ESCs present a strong enrichment in H3K27me3 compared to H3K9me3, reduced levels of 5-meC and deconstructed chromocenters (see the recap scheme of Fig. 7). Unexpectedly transcription at major and minor satellites is low. In addition, while repression of transcription is dependent on H3K9me3 in EpiSCs, we show that neither DNA methylation nor H3K9me3 can clearly modulate major satellite transcription in ESCs. Hence, our data leave open the question of the nature of the regulators (activators and repressors) of satellite transcription in ESCs. Transcription in serum-ESCs may be regulated by unknown factors present in the serum and not found in the chemically defined medium used in 2i condition. On the other hand, the inhibitors of MAPK and GSK3 pathways may themselves induce the repression of satellite transcription.

As mentioned above, Nanog activates major satellite transcription and reduces deposition of H3K9me3 at these sequences ⁴³. In 2i-ESCs, PCH are indeed more decondensed than in serum-ESCs, but on the other hand, transcription is strongly reduced compared to serum-ESCs. We checked the level of NANOG in cell extracts by western-blotting and showed that its level was unexpectedly reduced in 2i-ESCs compared to serum-ESCs, although not to the low level seen in EpiSCs (Supplementary Fig. 5). Hence in 2i-ESCs, reduced accumulation of H3K9me3 at PCH may not be directly linked to NANOG level.

Does the epigenetic status of in vitro pluripotency reflect in vivo pluripotency of mouse embryo?

ESCs cultured in 2i condition are transcriptionally closer to the E4.5 early epiblast than those cultured in serum ^{8,46,47}. The primed EpiSCs have characteristics of a late-gastrula stage epiblast ⁴⁸. Conversely on the epigenetic point of view not much is known for now, except for DNA methylation.

2i-ESCs display low level of 5-meC, as in the E3.5 ICM cells^{13,49}. After this stage, DNA methylation starts to increase in the epiblast cells and at E6.5, is similar to that in the serum-ESCs¹³. Finally EpiSCs seem to be even more methylated than the epiblast at E6.0-7.5^{18,32,50}.

Concerning histone modifications, a comprehensive survey during epiblast development is still lacking. However, available data suggest that the enrichment of PCH in H3K27me3 and the relative depletion in H3K9me3 as observed in 2i-ESCs do not recapitulate the situation *in vivo* in early epiblast cells. First, it seems that the H3K27me3-enriched-PCH pattern is not found in the ICM but is more typical of earlier developmental stages from the paternal pronucleus of 1-cell till 16-cell and morula stage embryos (Unpublished data of the lab and^{51,52}). Second, it is already known that pericentromeric sequences are marked by H3K9me3 as early as the 1-cell stage (maternal pronucleus) and remains so at blastocyst stage even in the ICM^{53,54}. At last, we showed that another striking difference of 2i-ESCs compared to serum-ESCs and EpiSCs is the decompaction of major satellites with the disruption of the chromocenter. At the blastocyst stage the ICM cells showed already a compacted chromocenter, while a more decondensed state was only found up to the late 2-cell stage when chromocenter formation occurs⁵⁵. The 2-cell stage is also characterized by a transient burst of major satellite transcripts that seems to be required for chromocenters formation and subsequent development⁵⁶. Further investigations are needed to really establish the transcription at the blastocyst stage distinguishing ICM cells and trophoblast cells for instance.

In conclusion, the present study brings new insights into the organization of the heterochromatin in naïve and primed pluripotent cells. The pattern of H3K27me3 is particularly dynamic and can easily distinguish the most naïve state (as ESCs in 2i medium) from the other pluripotent cells. Whereas this histone modification seems to shuttle from the coding genome to the PCH when cells transit from serum to 2i medium, it shows global loss in EpiSCs. Further investigation will be necessary to decipher how these different behaviors are regulated.

MATERIALS AND METHODS

Cell culture

ESC lines (R1, WT01, *Suv39hdn* and *DnmtTKO*) were cultured in 2i or serum medium as described⁵⁷. Briefly, 2i-ESCs were cultured on Laminin in Chemically Defined Medium (CDM)³ supplemented with LIF (700 U/ml), PD0332552 (1 μ M) and CHIR99201 (final 3 μ M), while serum-ESCs were cultured on gelatin in DMEM supplemented with 15% serum and LIF (1000 U/ml). *In vitro* conversion of ESC into cEpiSC was performed by switching ESCs from serum/LIF medium to CDM+FGF2 (12 ng/ml) and ActivinA (20 ng/ml) as described in⁵⁸. These converted cells were used 3-5 passages after the conversion. EpiSC lines (FT129.1 and 9.73) were cultured as described in³. EZH2 inhibition was performed by culturing ESC either in 2i or serum-containing medium supplemented with 1 μ M EPZ-6438 (AxonMedchem) for 72h (changing medium daily) or with DMSO as control.

Immunostaining

Cells were grown on coated glass-coverslips for 24h, then fixed with PFA 2% (EMS) for 20min, permeabilized with Triton X100 0.05% for 30min and blocked with BSA 2% for 1h. Primary antibody was incubated at 4°C O/N. After washes, the secondary antibody was incubated for 1h. Cells were then washed, post-fixed with PFA 2% (EMS) for 20min, incubated with 1/500 DAPI (Invitrogen) at 37°C for 15min and finally mounted on slide with VectaShield (Vector Laboratories). Antibodies used are described in Supplementary Table S1.

DNA-FISH

For the detection of major and minor satellites, we used probes⁵⁵ prepared by PCR on mouse genomic DNA using the following primer pairs: 5'-CATATTCCAGGTCCTTCAGTGTGC-3' and 5'-CACTTTAGGACGTGAAATATGGCG-3' (major), and 5'-ACTCATCTAATGTTCTACAGTG-3' and 5'-AAAACACATTCGTTGGAAACGCG-3' (minor). PCR products were labeled with Cy3 and Cy5, respectively, using random priming (Invitrogen). For FISH, cells grown on coated glass-coverslips for 24h were fixed with PFA 4% (EMS) for 15min. They were permeabilized with Triton X100 0.05% for 30min and treated with RNase A 200 μ g/mL (Sigma) for 30min at 37°C. After an equilibration step in the hybridization buffer (50% formamide, SSC 2X, Denhardt 1X, 40 mM NaH₂PO₄, 10% dextran sulfate) for 45min at 37°C, cells were denatured in presence of probes at 75°C for 3 min and then incubated O/N at 37°C. The day after they were washed three times with SSC2X pH5.8-50% formamide at 39°C and then three times with SSC2X pH6.3. Cells were finally incubated with 1/500 DAPI (Invitrogen) at 37°C for 15min and then mounted with VectaShield (Vector Laboratories).

3D-structured image acquisition and analysis

Imaging was performed at the MIMA2 platform (<http://www6.jouy.inra.fr/mima2>) with an inverted ZEISS AxioObserver Z1 microscope equipped with an ApoTome slider, a Colibri light source, Axiocam MRm camera and driven by the Axiovision software 4.8.2. Observations were carried out using a 63X oil-immersion objective. Cells were scanned entirely using a z-distance of 0.24 μm between optical sections. Fluorescent wavelengths of 405, 488, 555, and 639 nm were used to excite DAPI, FITC, Cy3, and Cy5, respectively. Images were then analysed on ImageJ (imagej.nih.gov/ij) to perform linescan, merge of channels and z-projections. 3D reconstructions of signals of DNA-FISH were done with AMIRA software 3.1 after 3D nuclei segmentation using an unpublished Python script of the laboratory.

Western blot

Cells were lysed for 30min on ice into RIPA buffer (150mM NaCl, 1% NP-40, 0.5% NaDeoxycholate, 0.1% SDS, 50mM Tris-HCl pH8.0) in presence of protease and phosphatase inhibitors (Pierce). Proteins were quantified using BCA assay (Pierce). 3 μg of proteins were charged on pre-cast polyacrylamide gel 4-15% (Biorad) for 1h run at 100V. Transfer was then performed on Trans-Blot Turbo (Biorad) for 7min on a PVDF membrane (Hybond-P, GE Healthcare). After blocking in TBS-Tween 20 0.01% (TBS-T) with either 4% non-fatty milk or 5% BSA, membranes were incubated O/N at 4°C with primary antibodies. After washes in TBS-T, membranes were incubated with secondary antibodies for 1h and washed again before the revelation with ECL2 Western blotting substrate (Pierce). Chemiluminescent signals were captured on a Fuji camera LAS-1000plus and then analysed with ImageJ (imagej.nih.gov/ij). H3 was used for normalization. For sequential protein detection, membranes were stripped with 25mM Glycine and 1% SDS at pH 2 for 30min, followed by washes in TBS-T and blocking (milk or BSA) according to the new primary antibody. Antibodies are described in Supplementary Table S1.

DNA methylation analysis of satellite repeats using Southern blot

Southern blot on genomic DNA was performed as described in Thijssen et al.⁵⁹. For major satellite analysis 200ng of genomic DNA were digested with *HpyCH4IV* (New England Biolabs) for 1h at 37°C, while for minor satellites 500ng of gDNA were digested with *HpaII* (New England Biolabs) and 300ng with *MspI* (New England Biolabs), both O/N at 37°C. Digested samples were separated for 5h on 1% agarose gel. Gels were then denaturated in a 1.5M NaCl and 0.5M NaOH solution for 20 min and neutralized with 0.5M Tris-HCl pH 7.5 and 1.5M NaCl for 40min. Transfer was performed O/N on Hybond-N+ membranes (GE Healthcare) in SSC 20X. After ultraviolet crosslinking, membranes were

pre-hybridized in SSC 6X, Denhardt 5X and 0.1% SDS for 1h at 42°C and hybridized with 32P-labelled probes for 2h at 42°C. After membrane washing, signals were detected using FLA 7000 phosphorimager (Fuji). Images were then analyzed with ImageJ (imagej.nih.gov/ij) to perform linescan for major satellites and intensity ratio *HpaII/MspI* for the lower six bands of each lane for minor satellites. Probe used: Major satellites 5' –CAC GTC CTA CAG TGG ACA TTT CTA AAT TTT CCA CCT TTT TCA GTT- 3' and minor satellites 5' –ACA TTC GTT GGA AAC GGG ATT TGT AGA ACA GTG TAT ATC AAT GAG TTA CAA TGA GAA ACA T- 3'.

qRT-PCR

Total RNA was extracted from cells using TRIzol (Ambion). 3µg of RNA were subjected to DNase treatment using Turbo DNA-free kit (Ambion). Retrotranscription of 500ng of DNase treated-RNA was performed using Random primers (Invitrogen) and Superscript III (Invitrogen). For each sample, a negative control was included (no Superscript enzyme). Quantitative PCR was carried out in triplicates using SybrGreen mix (Applied Biosystem) on a StepOne Plus thermal cycler (Applied Biosystem). Data were normalized using the geometric mean of *Sdha* and *Pbgd* using Qbase software (Biogazelle). Results were presented according to Weissgerber et al ⁶⁰. The Primers are described in Supplementary Table S2.

Bioinformatic analysis of ChIP-Seq datasets for satellite repeats

To compare the enrichment in H3K27me3 marks over major and minor satellite repeats we used the following ChIP-seq datasets: GSM590115 (E14-serum) and GSM590116 (E14-2i) from Marks et al. ⁷; GSM1725687 (EpiSC1), GSM1725686 (Input EpiSC1), GSM1725726 (Input 2i-ESC1), GSM1725727 (2i-ESC1), GSM1725730 (Input 2i-ESC2), GSM1725731 (2i-ESC2), GSM1725689 (Input EpiSC2), GSM1725690 (EpiSC2), from Zylitz et al ³².

An “*in silico*” library was made up exclusively of major and minor fasta consensus sequences ³¹. Each sequence was duplicated and juxtaposed in order to detect reads which may map at the junction between 2 consecutive repeats. For the repeat analysis of ChIP-seq profiles, mappings were performed with the bowtie2 aligner version 2.1.0 with default options ⁶².

Each read that mapped on major or minor satellite repeats was counted from the resulting output BAM file. Results were expressed as percentage of total number of reads that mapped on the whole mouse genome GRCm38.84.

REFERENCES

1. Nichols, J. & Smith, A. Naive and Primed Pluripotent States. *Cell Stem Cell* **4**, 487–492 (2009).
2. Huang, Y., Osorno, R., Tsakiridis, A. & Wilson, V. In Vivo Differentiation Potential of Epiblast Stem Cells Revealed by Chimeric Embryo Formation. *Cell Rep.* **2**, 1571–1578 (2012).
3. Brons, I. G. M. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
4. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
5. Ying, Q.-L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
6. Kolodziejczyk, A. A. *et al.* Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation. *Cell Stem Cell* **17**, 471–485 (2015).
7. Marks, H. *et al.* The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. *Cell* **149**, 590–604 (2012).
8. Chen, G. *et al.* Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. *Genome Res.* gr.201954.115 (2016). doi:10.1101/gr.201954.115
9. Ahmed, K. *et al.* Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. *PLoS ONE* **5**, e10531 (2010).
10. Koh, F. M., Sachs, M., Guzman-Ayala, M. & Ramalho-Santos, M. Parallel gateways to pluripotency: open chromatin in stem cells and development. *Curr. Opin. Genet. Dev.* **20**, 492–499 (2010).
11. Meissner, A. Epigenetic modifications in pluripotent and differentiated cells. *Nat. Biotechnol.* **28**, 1079–1088 (2010).
12. Meshorer, E. *et al.* Hyperdynamic Plasticity of Chromatin Proteins in Pluripotent Embryonic Stem Cells. *Dev. Cell* **10**, 105–116 (2006).

13. Habibi, E. *et al.* Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells. *Cell Stem Cell* **13**, 360–369 (2013).
14. Joshi, O. *et al.* Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. *Cell Stem Cell* **17**, 748–757 (2015).
15. Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).
16. Bao, S. *et al.* Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**, 1292–1295 (2009).
17. Factor, D. C. *et al.* Epigenomic Comparison Reveals Activation of ‘Seed’ Enhancers during Transition from Naive to Primed Pluripotency. *Cell Stem Cell* **14**, 854–863 (2014).
18. Veillard, A.-C., Maruotti, J. & Jouneau, A. in *Stem cells and Cancer stem cells* (ed. Hayat, M. A.) **8**, 133–146 (Springer Science+Business Media, 2012).
19. Biscotti, M. A., Canapa, A., Forconi, M., Olmo, E. & Barucca, M. Transcription of tandemly repetitive DNA: functional roles. *Chromosome Res.* **23**, 463–477 (2015).
20. Ferreira, D. *et al.* Satellite non-coding RNAs: the emerging players in cells, cellular pathways and cancer. *Chromosome Res.* **23**, 479–493 (2015).
21. Guenatri, M., Bailly, D., Maison, C. & Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **166**, 493–505 (2004).
22. Martens, J. H. *et al.* The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* **24**, 800–812 (2005).
23. Peters, A. H. *et al.* Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* **12**, 1577–1589 (2003).
24. Déjardin, J. Switching between Epigenetic States at Pericentromeric Heterochromatin. *Trends Genet.* **31**, 661–672 (2015).
25. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–57 (1999).

26. von Meyenn, F. *et al.* Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol. Cell* **0**, (2016).
27. Lu, J. & Gilbert, D. M. Proliferation-dependent and cell cycle-regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* **179**, 411–421 (2007).
28. Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015).
29. Efroni, S. *et al.* Global Transcription in Pluripotent Embryonic Stem Cells. *Cell Stem Cell* **2**, 437–447 (2008).
30. Cooper, S. *et al.* Targeting Polycomb to Pericentric Heterochromatin in Embryonic Stem Cells Reveals a Role for H2AK119u1 in PRC2 Recruitment. *Cell Rep.* **7**, 1456–1470 (2014).
31. Lehnertz, B. *et al.* Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192–1200 (2003).
32. Zyllicz, J. J. *et al.* Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *eLife* **4**, e09571 (2015).
33. Chen, T., Ueda, Y., Xie, S. & Li, E. A Novel Dnmt3a Isoform Produced from an Alternative Promoter Localizes to Euchromatin and Its Expression Correlates with Active de Novo Methylation. *J. Biol. Chem.* **277**, 38746–38754 (2002).
34. Houliard, M. *et al.* CAF-1 Is Essential for Heterochromatin Organization in Pluripotent Embryonic Cells. *PLoS Genet.* **2**, e181 (2006).
35. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
36. Sakaue, M. *et al.* DNA methylation is dispensable for the growth and survival of the extraembryonic lineages. *Curr Biol* **20**, 1452–7 (2010).
37. Saksouk, N. *et al.* Redundant Mechanisms to Form Silent Chromatin at Pericentromeric Regions Rely on BEND3 and DNA Methylation. *Mol. Cell* **56**, 580–594 (2014).

38. Knutson, S. K. *et al.* Selective Inhibition of EZH2 by EPZ-6438 Leads to Potent Antitumor Activity in EZH2-Mutant Non-Hodgkin Lymphoma. *Mol. Cancer Ther.* **13**, 842–854 (2014).
39. Hagarman, J. A., Motley, M. P., Kristjansdottir, K. & Soloway, P. D. Coordinate Regulation of DNA Methylation and H3K27me3 in Mouse Embryonic Stem Cells. *PLOS ONE* **8**, e53880 (2013).
40. Walter, M., Teissandier, A., Pérez-Palacios, R. & Bourc'his, D. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* **5**, e11418 (2016).
41. Mozzetta, C., Boyarchuk, E., Pontis, J. & Ait-Si-Ali, S. Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat. Rev. Mol. Cell Biol.* **16**, 499–513 (2015).
42. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919–932 (2002).
43. Novo, C. L. *et al.* The pluripotency factor Nanog regulates pericentromeric heterochromatin organization in mouse embryonic stem cells. *Genes Dev.* (2016). doi:10.1101/gad.275685.115
44. Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* **31**, 2305–2312 (2003).
45. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15**, 703–708 (2014).
46. Boroviak, T., Loos, R., Bertone, P., Smith, A. & Nichols, J. The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat. Cell Biol.* **16**, 516–28 (2014).
47. Boroviak, T. *et al.* Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
48. Kojima, Y. *et al.* The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak. *Cell Stem Cell* **14**, 107–120 (2014).

49. Ficiz, G. *et al.* FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
50. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
51. Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40**, 411–420 (2008).
52. Santenard, A. & Torres-Padilla, M. E. Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. *Epigenetics* **4**, 80–4 (2009).
53. Beaujean, N. Histone post-translational modifications in preimplantation mouse embryos and their role in nuclear architecture. *Mol. Reprod. Dev.* **81**, 100–112 (2014).
54. Santos, F., Peters, A. H., Otte, A. P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev. Biol.* **280**, 225–236 (2005).
55. Aguirre-Lavin, T. *et al.* 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.* **12**, 1 (2012).
56. Probst, A. V. *et al.* A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. *Dev. Cell* **19**, 625–638 (2010).
57. Tosolini, M. & Jouneau, A. Acquiring Ground State Pluripotency: Switching Mouse Embryonic Stem Cells from Serum/LIF Medium to 2i/LIF Medium. *Methods Mol. Biol. Clifton NJ* **1341**, 41–48 (2016).
58. Tosolini, M. & Jouneau, A. From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. *Methods Mol. Biol. Clifton NJ* **1341**, 209–216 (2016).
59. Thijssen, P. E. *et al.* Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat. Commun.* **6**, 7870 (2015).
60. Weissgerber, T. L., Milic, N. M., Winham, S. J. & Garovic, V. D. Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol.* **13**, e1002128 (2015).

61. Ferri, F., Bouzinba-Segard, H., Velasco, G., Hubé, F. & Francastel, C. Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.* **37**, 5071–5080 (2009).
62. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

ACKNOWLEDGMENTS

We are grateful to Dr Masaki Okano for the gift of the *Dnmt* TKO ESCs, and to Dr Antoine Peters for the WT and *Suv39hdn* ESC lines. We thank Tiphaine Aguirre-Lavin for technical advice on DNA-FISH and AMIRA 3.1 software and Luc Jouneau for helping with R. Thanks to Dr Guillaume Velasco for discussions. This work was supported by an ANR Programme Investissements d’Avenir REVIVE (ANR-10-LABX73). MT is a recipient of a PhD fellowship from DIM Biotherapies Ile-de-France.

AUTHOR CONTRIBUTIONS

MT, AJ and ABG designed and analysed the experiments and wrote the manuscript. MT performed most of the experiments, with some help from VB, MC and GG. PA did some image analysis and FP analysed the ChIPseq datasets. All authors reviewed the manuscript.

FIGURE LEGEND

Fig. 1: Heterochromatin landscape in the different state of mouse pluripotency

(A) Immunostaining images (single-plan) for H3K9me3 and H3K27me3 with DAPI DNA counterstaining. Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) or H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5µm. (B) Western-blot analysis for quantification of bulk levels of repressive histone modifications H3K9me3 and H3K27me3 related to the total level of H3. (C) Immunostaining images of anaphase chromosome plates for H3K9me3 (red) or H3K27me3 (green) with DAPI DNA counterstaining (blue). Stars indicate enrichment of the histone mark at the PCH/CH region. Scale bars represent 5µm.

Fig. 2: Methylation profile at major and minor satellites in 2i-ESCs, serum-ESCs and EpiSCs.

(A) Southern-blot analysis of gDNA digested with *HpyCH4IV* revealed with probe for major satellites. Linescan quantification for each lane: MEF (black), *DnmtTKO* (pink), 2i-ESC (red), serum-ESC (blue) and EpiSC (green). (B) Southern-blot analysis of gDNA digested with *HpaII* revealed with probe for minor satellites. Quantification related to southern-blot of gDNA digested with *MspI*. *DnmtTKO*-ESC is set to 1.

Fig. 3: 3D-organization and transcription of major and minor satellites in the different mouse pluripotency states.

(A) 3D DNA-FISH images (z-projection) for major (red) and minor (green) satellites with DAPI DNA counterstaining (blue). Magnification on a single cell (arrowed) and 3D-reconstruction of major and minor satellite signals using AMIRA 3.1 software. Scale bars represent 5µm. (B) Relative expression (CNRQ) of major and minor satellite transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent biological replicate.

Fig. 4: Contrasting effect of *Suv39h*dn condition in 2i-ESCs, serum-ESCs and EpiSCs.

(A) Immunostaining images (single-plan) for H3K9me3 and H3K27me3 with DAPI DNA counterstaining in *Suv39h*dn condition, to compare with wild-type condition (Fig. 1A). Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) or H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the

percentage of cells in the population displaying the same pattern. Scale bars represent 5µm. (B) Southern-blot analysis of gDNA digested with *HpyCH4IV* revealed with probe for major satellites in wild-type and *Suv39hdn* conditions. Linescan quantification for each lane: 2i-ESC (red), serum-ESC (blue) and EpiSC (green). Wild-type condition is represented with a continuous line, while *Suv39hdn* with a dotted line. (C) Relative expression (CNRQ) of major satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in Wild-type and *Suv39hdn* condition. Each point is an independent biological replicate.

Fig. 5: Epigenetic and transcriptional consequences of DNA methylation absence on major satellites in 2i- and serum-ESCs.

(A) Immunostaining images (single-plan) for H3K9me3 and H3K27me3 with DAPI DNA counterstaining in *DnmtTKO* condition, to compare with wild-type condition (Fig. 1A). Magnification on a single cell (arrow) with merge of signals: H3K9me3 (red) or H3K27me3 (green) with DAPI (blue). Linescan analysis showing peaks of foci-enrichment (highlighted with the star). % indicates the percentage of cells in the population displaying the same pattern. Scale bars represent 5µm. (B) Relative expression (CNRQ) of major satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in Wild-type and *DnmtTKO* condition. Each point is an independent biological replicate.

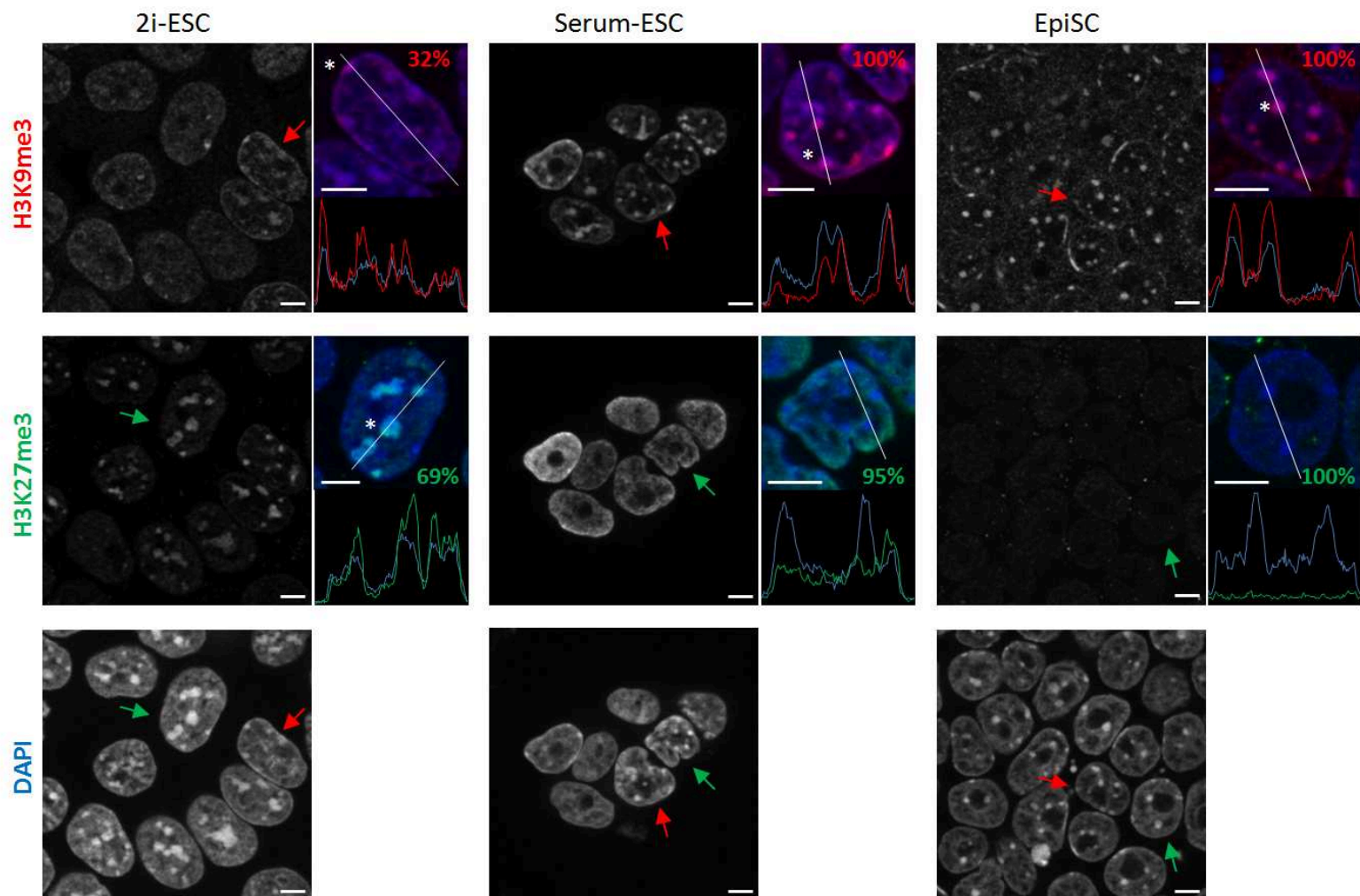
Fig. 6: Reduced levels of H3K27me3 do not induce up-regulation of major satellites.

(A) Left part: Western-blot analysis for quantification of bulk levels of the HMT enzyme EZH2 and the repressive histone modification H3K9me3 and H3K27me3 related to total H3 in 2i-ESC treated with DMSO (Control) or EPZ (EZH2 inhibition). Middle part: Relative expression (CNRQ) of major satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes in 2i-ESC treated with DMSO or EPZ. Each point is an independent biological replicate. Right part: immunostaining images (single-plan) for H3K9me3 (red) and H3K27me3 (green) with DAPI DNA counterstaining (blue) in 2i-ESC treated with DMSO or EPZ. Scale bars represent 5µm. (B, C, D, E) For each condition: Wild-type serum-ESC (B), *Suv39hdn* serum-ESC (C), *DnmtTKO* 2i-ESC (D) and *DnmtTKO* serum-ESC (E). Left part: Western-blot analysis for quantification of bulk levels of the HMT enzyme EZH2 and the repressive histone modification H3K9me3 and H3K27me3 related to H3 total. Right part: Relative expression (CNRQ) of major satellites transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes after treated with DMSO (Control) or EPZ (EZH2 inhibition). Each point is an independent biological replicate.

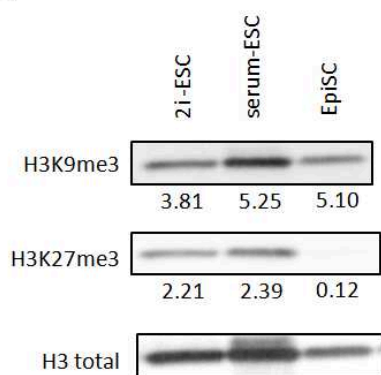
Fig. 7: Model of the epigenetic organization at PCH in the different pluripotent cells and in mutants

Schematic drawings recapitulating the organization and transcription status at PCH based on our findings and published data^{30,31,44}

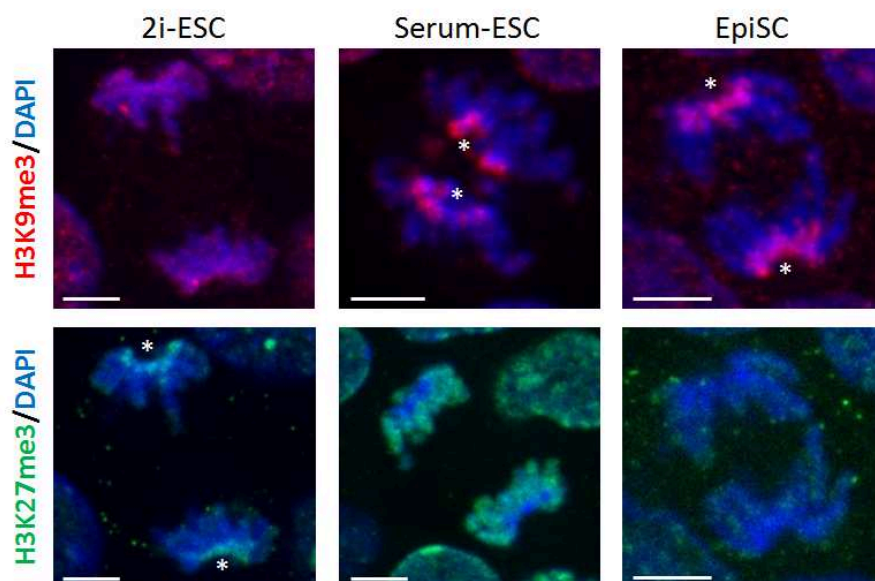
A



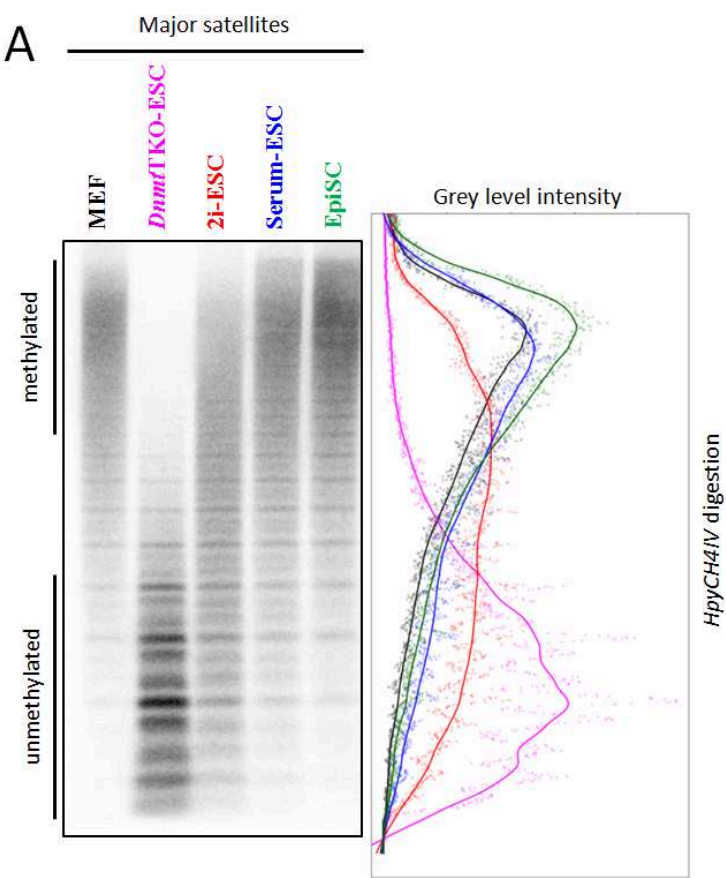
B



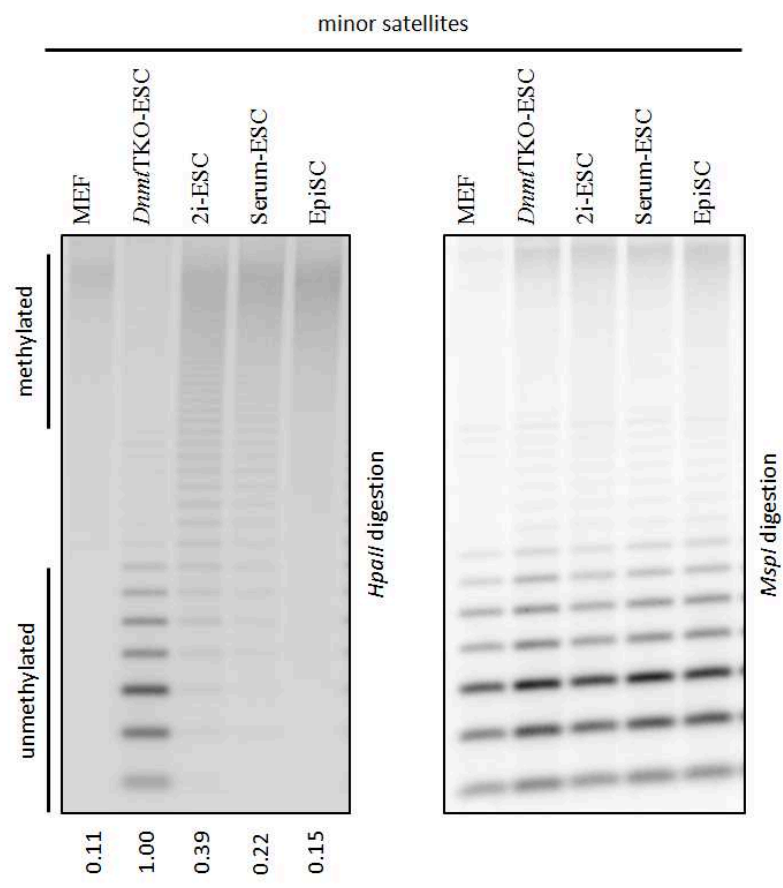
C



A



B



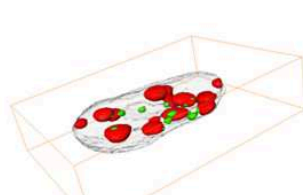
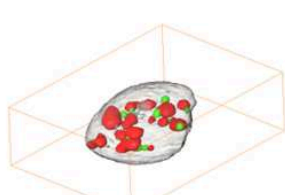
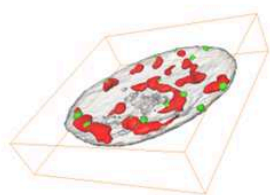
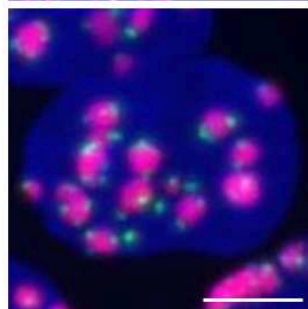
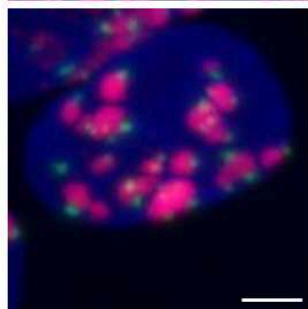
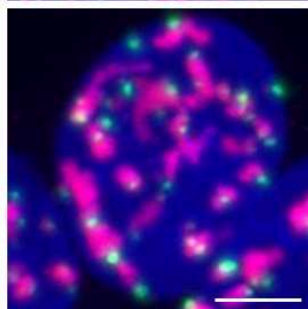
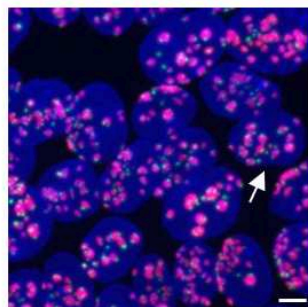
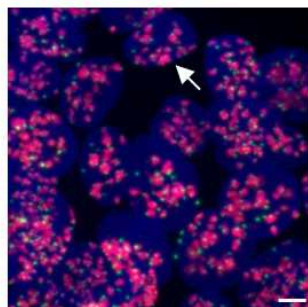
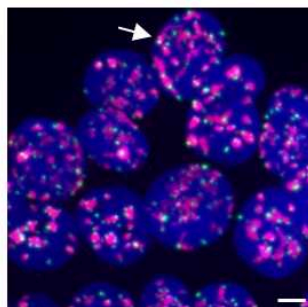
A

2i-ESC

Serum-ESC

EpiSC

DAPI / Major satellites / minor satellites

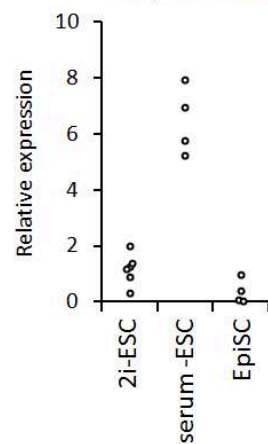


Z-projection

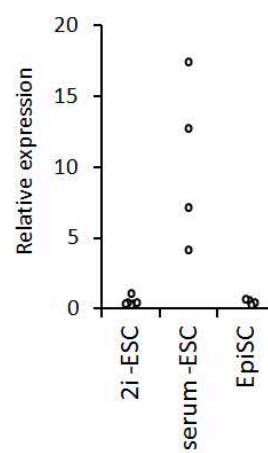
3D-reconstruction

B

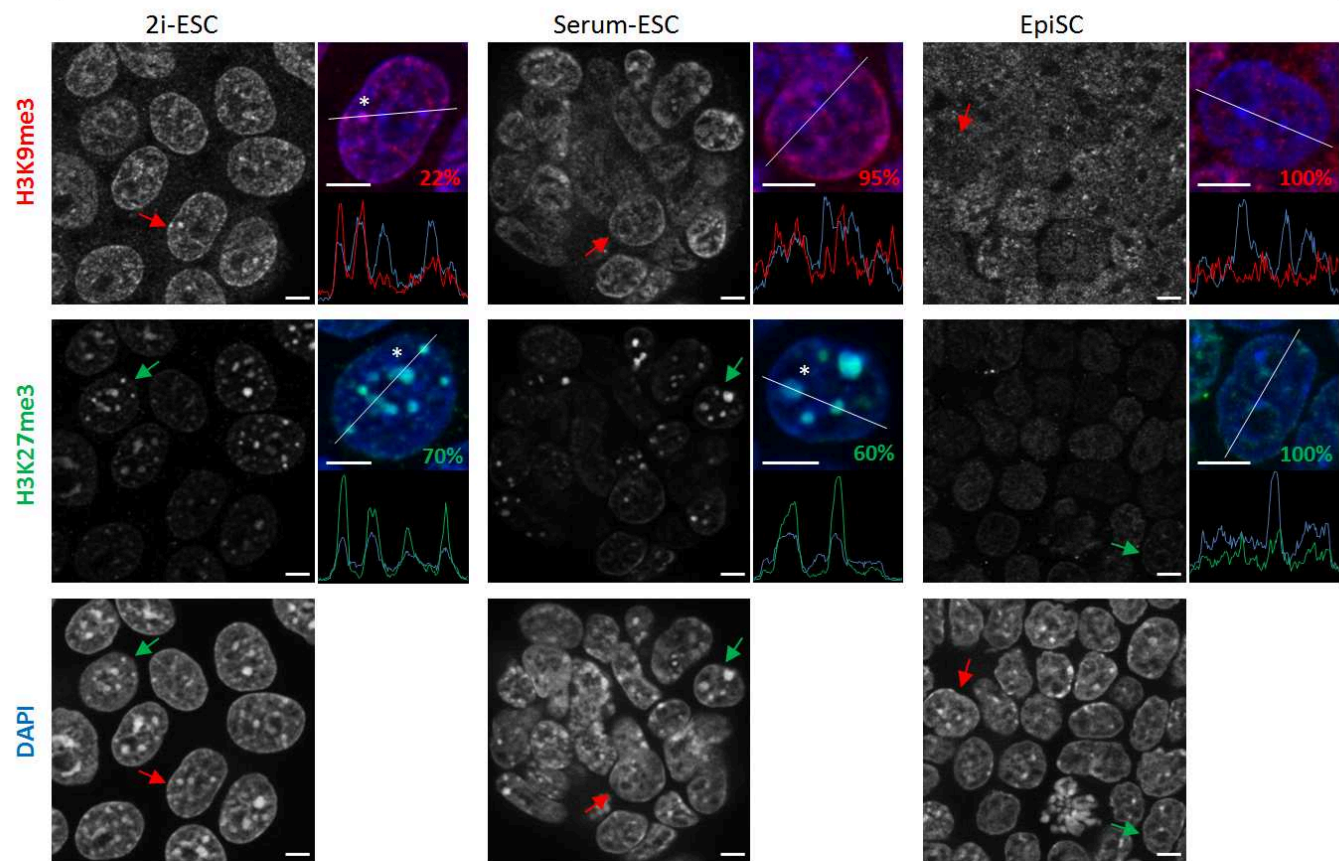
Major satellites



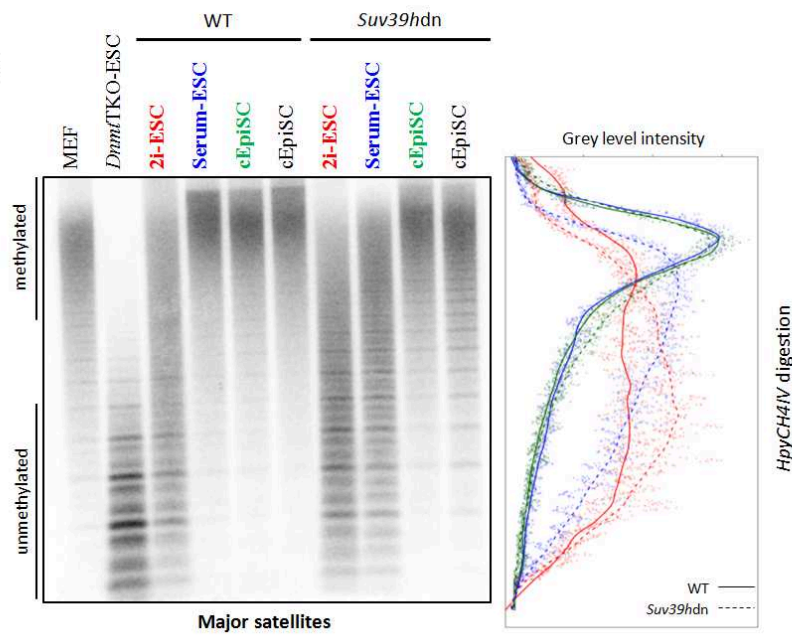
minor satellites



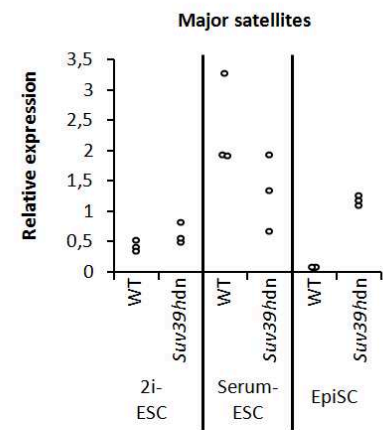
A

*Suv39h*dn

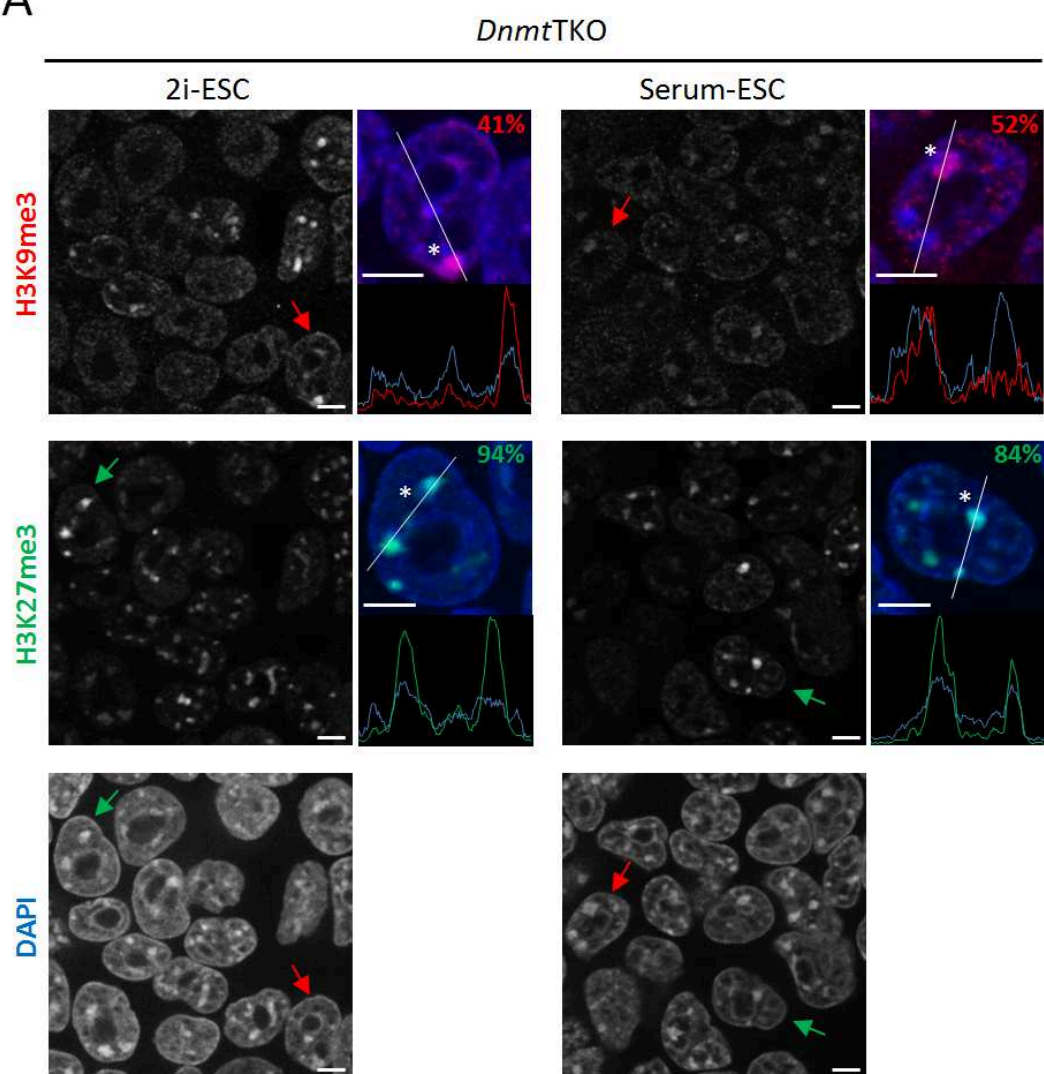
B



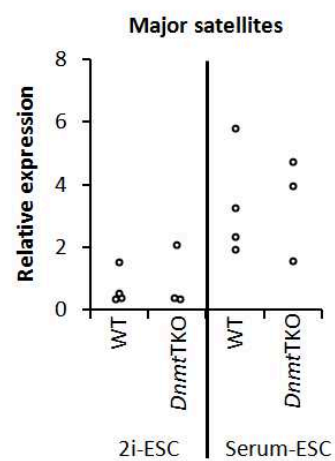
C



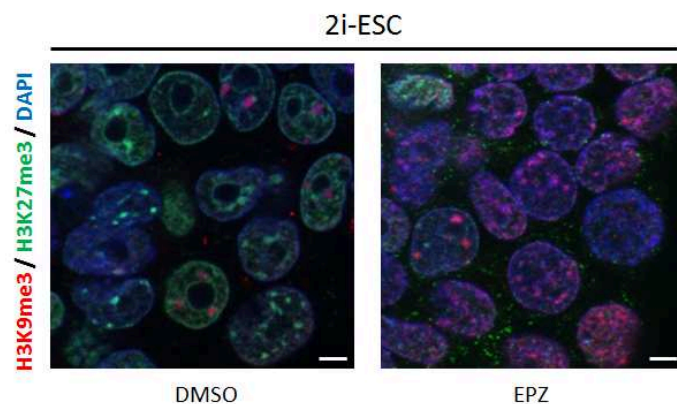
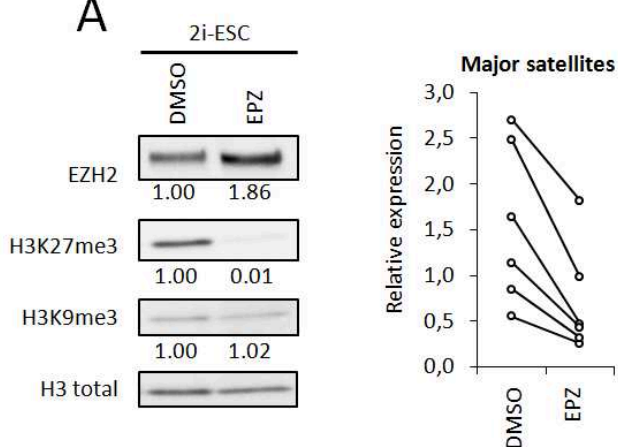
A



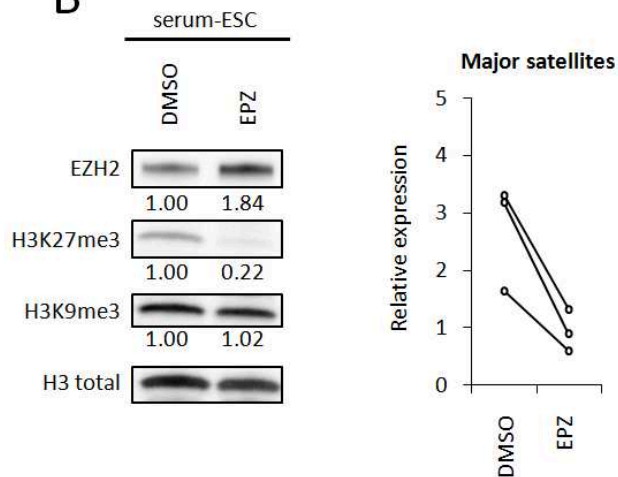
B



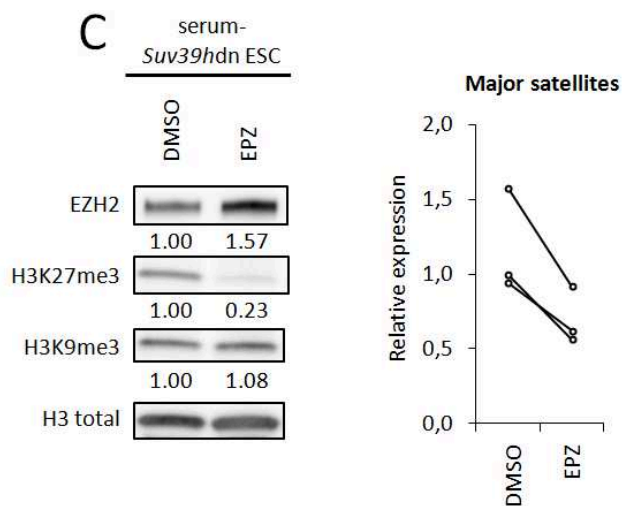
A



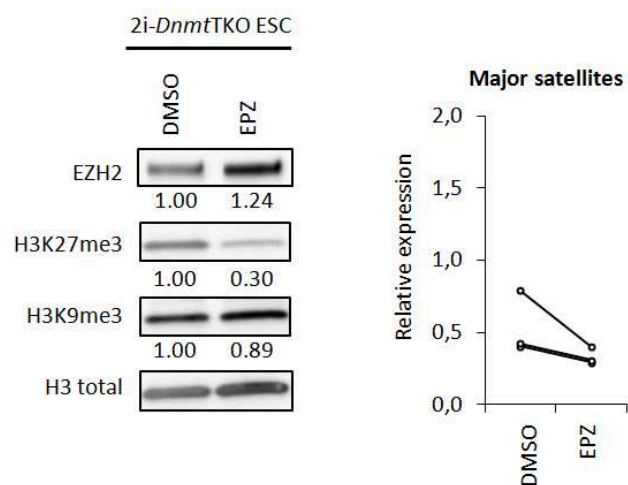
B



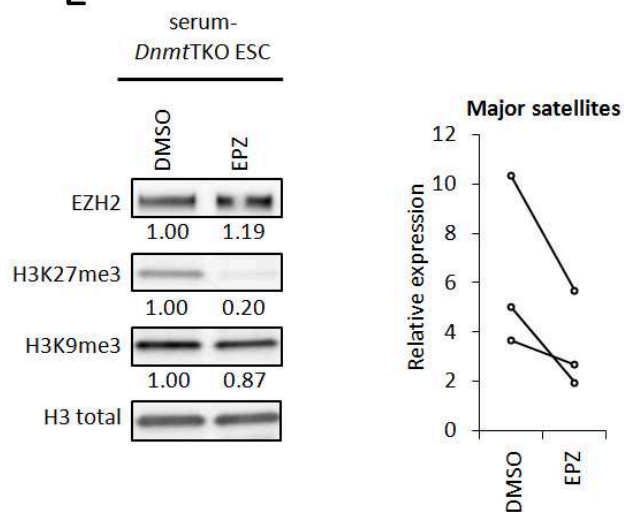
C



D



E



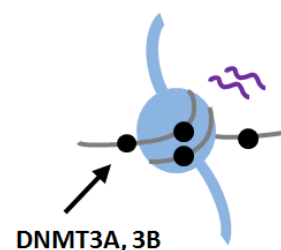
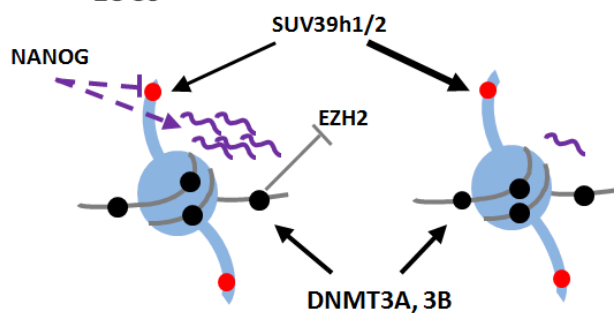
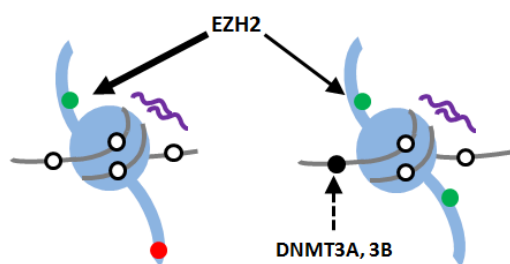
DnmtTKO
ESCs

2i- WT ESCs
Suv39hdn ESCs

(WT) Serum
ESCs

EpiSCs

Suv39hdn
EpiSCs



Intermediate H3K9me3
No DNA methylation
High H3K27me3
Low transcription

Low H3K9me3
Reduced DNA methylation
High H3K27me3
Low transcription

High H3K9me3
DNA methylation
Low H3K27me3
High transcription

High H3K9me3
High DNA methylation
No H3K27me3
No transcription

No H3K9me3
High DNA methylation
No H3K27me3
Low transcription

● 5meC

● H3K9 tri-methylation

● H3K27 tri-methylation

~ Satellite repeat transcripts

Supplementary Figure and Table Legends

Supplementary Fig. 1: Additional information on epigenetic modifiers and histone modification in the different states of mouse pluripotency

(A) Left part: Relative expression (CNRQ) of *Suv39h1* and *Suv39h2* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent biological replicate. Middle part: Western-blot analysis for quantification of bulk levels of SUV39H1 related to total H3. Right part: Immunostaining images (single-plan) for SUV39H1 with DAPI DNA counterstaining. % indicates the percentage of cells in the population displaying a foci-enrichment pattern. Scale bars represent 5µm.

(B) Left part: Relative expression (CNRQ) of *Ezh2* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Middle part: Western-blot analysis for quantification of bulk levels of EZH2 related to H3 total. Right part: immunostaining images (single-plan) for EZH2 with DAPI DNA counterstaining. Scale bars represent 5µm. (C) Relative expression (CNRQ) of *Kdm6b* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. (D) Percentage of mapped reads on major and minor satellite repeats sequence using H3K27me3 ChIP-seq ⁷(E) Percentage of mapped reads on major and minor satellite repeat sequences using H3K27me3 ChIP-seq compared to input ³²

Supplementary Fig. 2: De novo DNA methylation machinery in 2i-ESC, serum-ESC and EpiSC

(A) Left part: Relative expression (CNRQ) of *Dnmt3a* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Right part: Western-blot analysis for quantification of bulk levels of DNMT3A related to H3 total. (B) Left part: Relative expression (CNRQ) of *Dnmt3b* transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Each point is an independent sample. Right part: Western-blot analysis for quantification of bulk levels of DNMT3B related to H3 total.

Supplementary Fig. 3: CAF-1 a factor related to organization of heterochromatin in mouse pluripotent stem cells.

Immunostaining images (single-plan) for CAF-1 p150 with DAPI DNA counterstaining. % indicates the percentage of cell in the population displaying a foci-enrichment pattern. Scale bars represent 5µm.

Supplementary Fig. 4: validation of conversions from WT and *Suv39hdn* ESCs into cEpiSCs

Relative expression of different common pluripotency (Oct4, Sox2, Nanog), naïve-specific (Klf4, Esrrb) and epiblast specific (Dnmt3b, Fgf5, Otx2) transcripts by qRT-PCR analysis normalized to *Sdha* and *Pbgd* housekeeping genes. Three independent conversions were made.

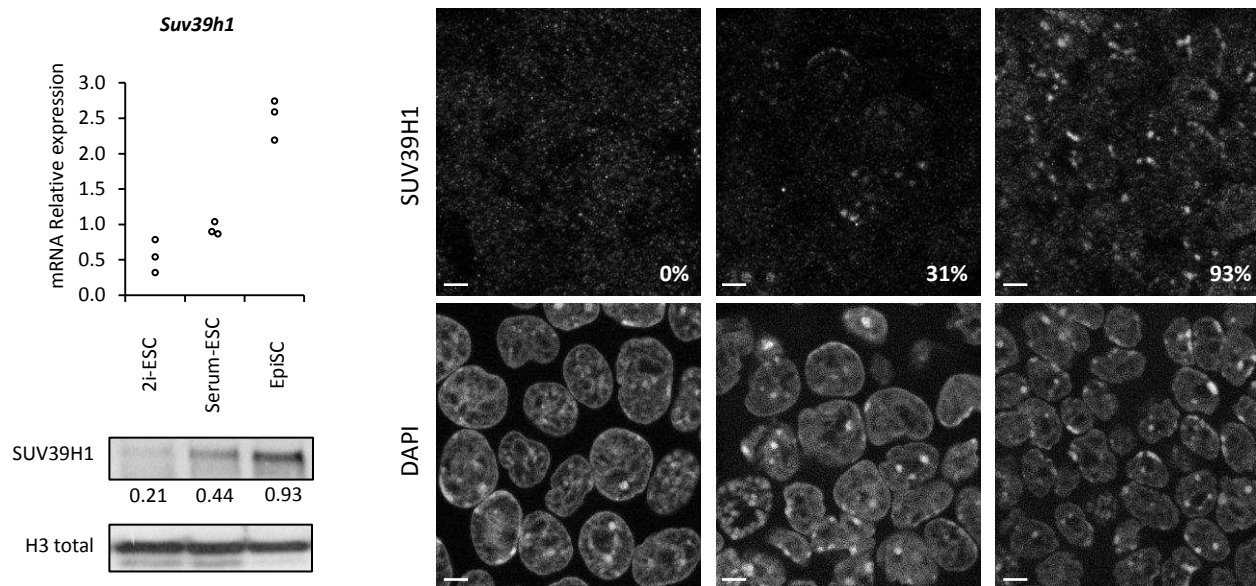
Supplementary Fig. 5: NANOG level is modulated according to the pluripotent state

Western-blot analysis for quantification of bulk levels of NANOG related to H3 total, in 2i and serum-ESCs and in EpiSCs.

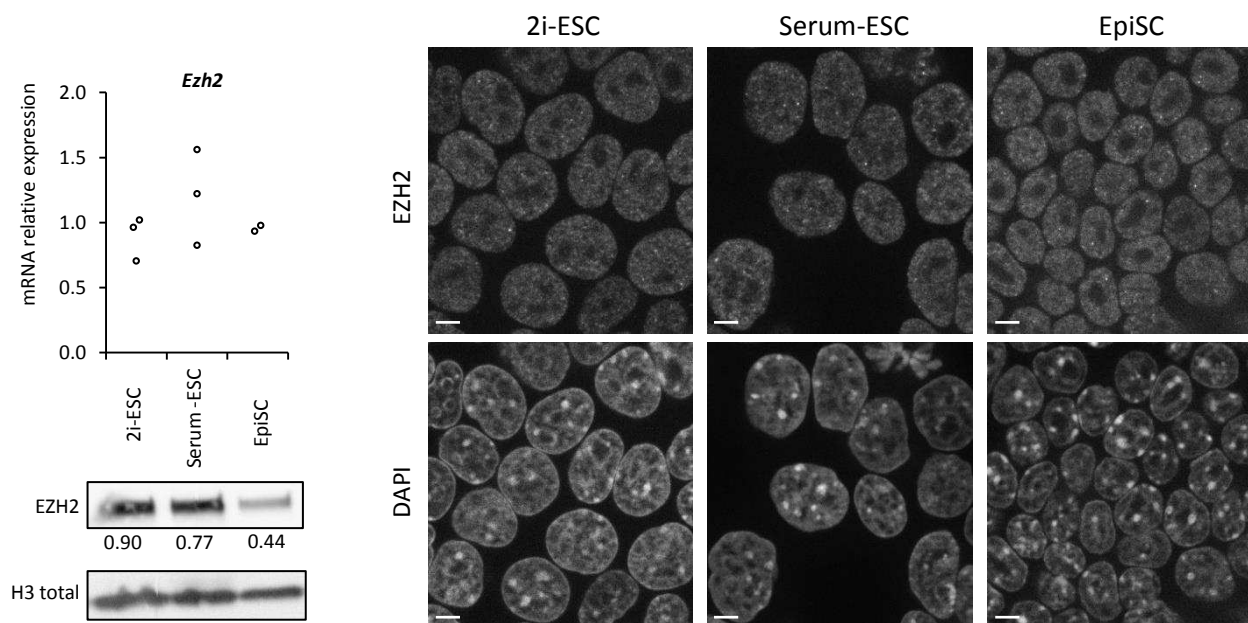
Supplementary Table S1: List of antibodies used for immunostaining and western-blot

Supplementary Table S2: list of primers used for qRT-PCR

A

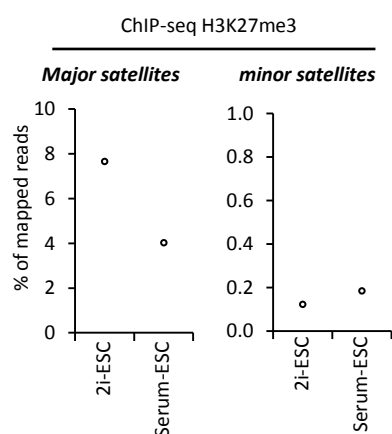


B



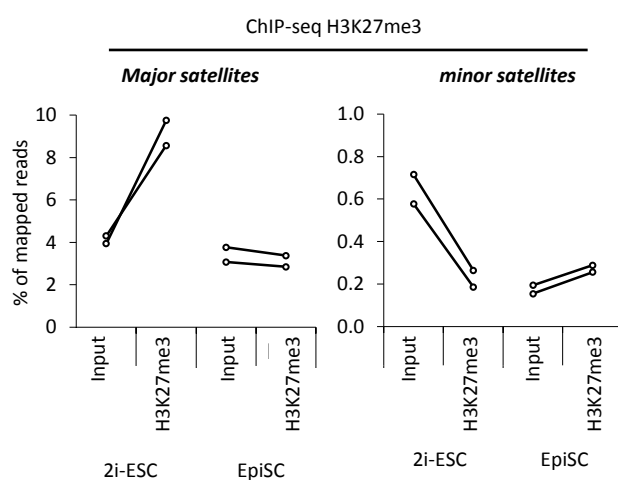
C

Dataset from Marks *et al*, 2012



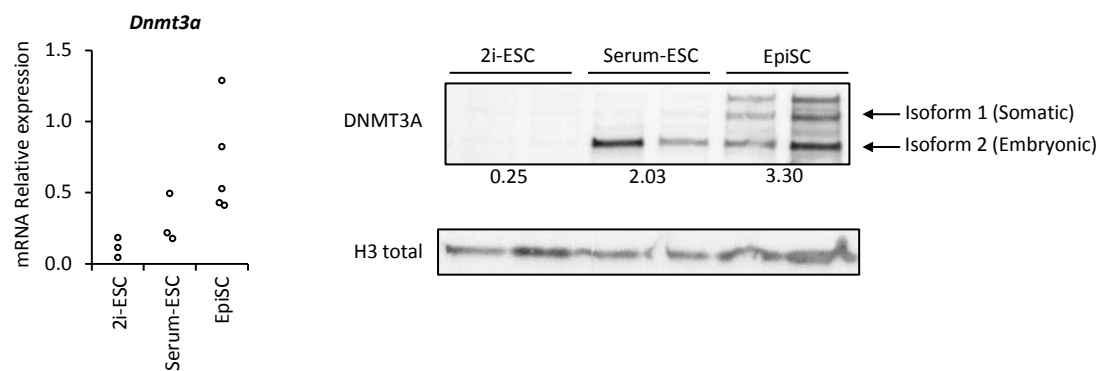
D

Dataset from Zylitz *et al*, 2015

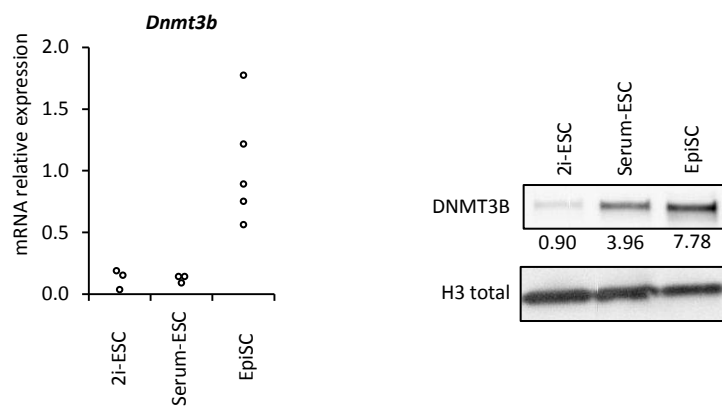


Supp Figure 1

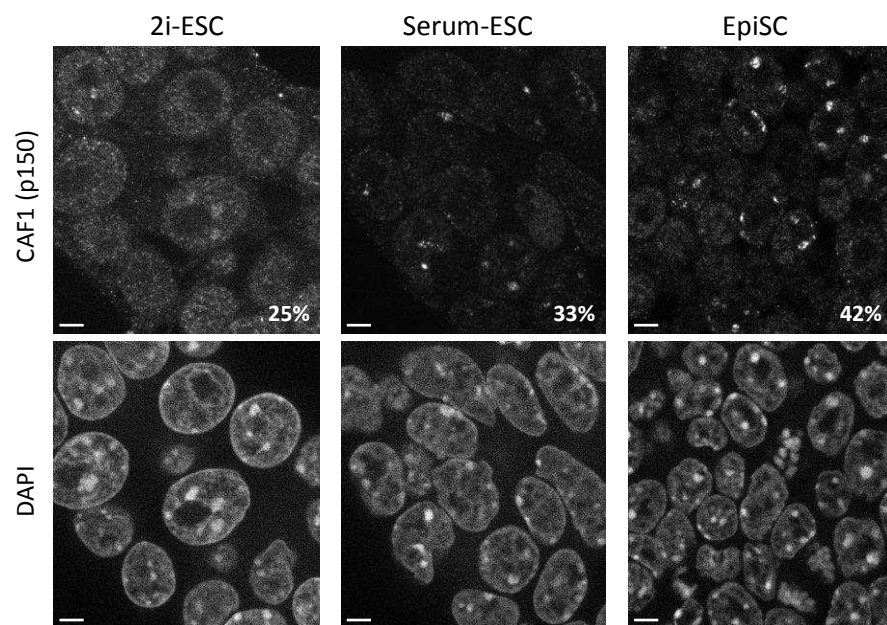
A



B

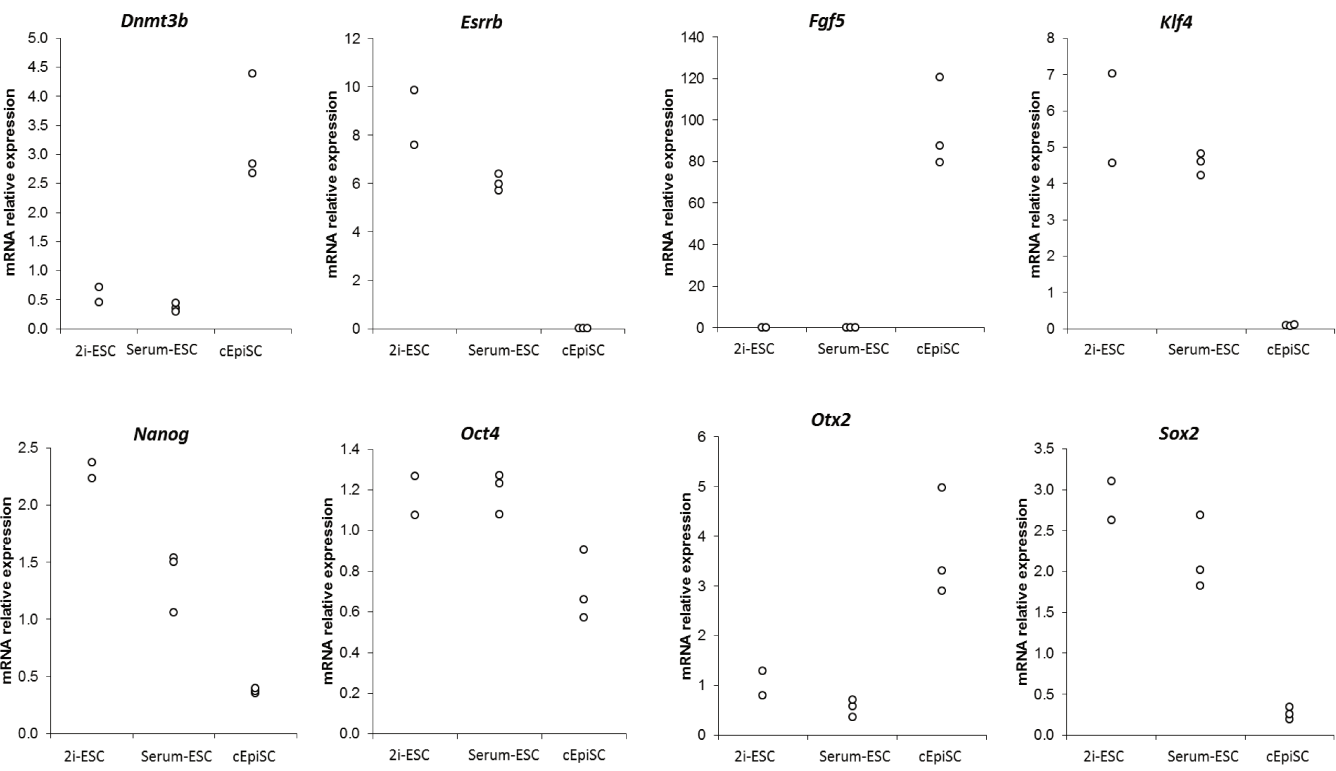


Supp Figure 2

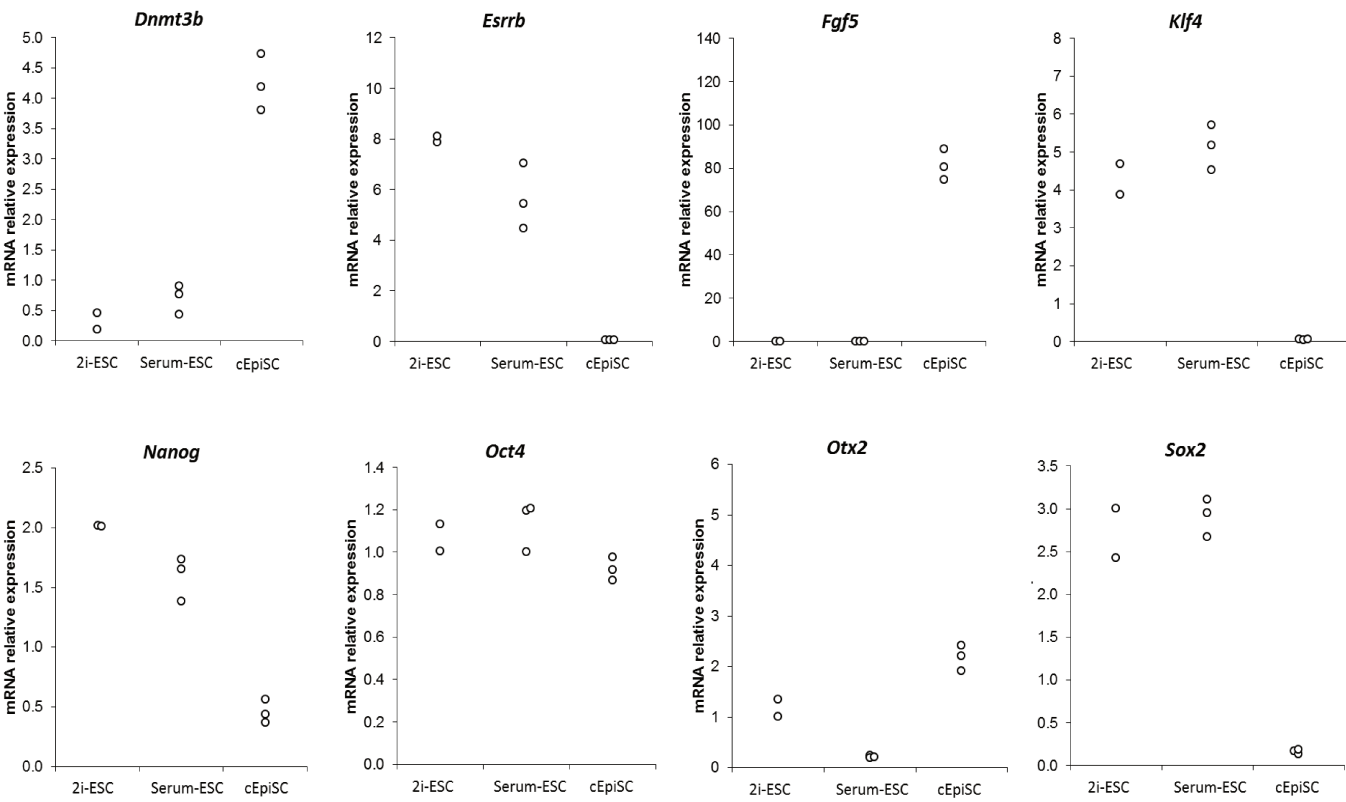


Supp Figure 3

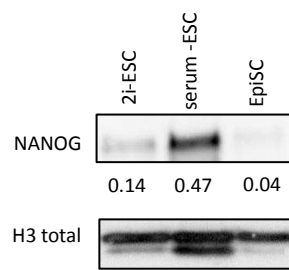
A Wild-type cells



B *Suv39*hdn cells



Supp Figure 4



Supp Figure 5

Antigen	dilution for immunostaining	dilution for western-blot	reference
H3K9me3	1/300	1/1000	Active Motif 39161
H3K27me2me3	1/300	1/1000	Active Motif 39538
H3K27me3	1/300		Cell Signaling C36B11#9733
H3K27me3		1/1000	Millipore DAM07-774
EZH2	1/200	1/1000	NovocastraNCL-L-EZH
SUV39H1	1/100	1/1000	Cell Signaling D11B6 #8729
CAF-1p150	1/50		Santa Cruz D-16 sc-10206
NANOG		1/1000	Abcam ab80892
DNMT3B		1/500	Active Motif 39207
DNMT3A		1/500	Active Motif 39206
H3total		1/20,000	Abcam 1791
Anti-Rabbit-Cy3	1/200		Jackson ImmunoResearch
Anti-Mouse-FITC	1/200		Jackson ImmunoResearch
Anti-Goat-Cy3	1/200		Jackson ImmunoResearch
Anti-mouse-HPO		1/5000	Jackson ImmunoResearch
anti-rabbit-HPO		1/5000	Jackson ImmunoResearch

Supp Table S1

Primers for qRT-PCR				
Gene	Primer Forward Sequence (5'-3')	Primer Revers Sequence (5'-3')	Reference or Primer Bank ID	Annealing
<i>Major satellite</i>	GACGACTTGAAAAATGACGAAATC	CATATTCCAGGTCCTTCAGTGTGC	1	60°C
<i>Minor satellite</i>	GAACATATTAGATGAGTGAGTTAC	GTTCTACAAATCCCCTTTCCAAC	2	60°C
<i>Sdha</i>	GGAACACTCCAAAAACAGACCT	CCACCACTGGGTATTGAGTAGAA		60°C
<i>Pbgd</i>	CCTGGCATACAGTTTGAAATCAT	TTTTTCAGGGCGTTTTCT	3	60°C
<i>Ezh2</i>	AGTGACTTGATTTCAGCAC	AATTCTGTTGTAAGGGCGACC		60°C
<i>Suv39h1</i>	GCAGTGTGTGCTGTAATCTTCT	ATACCCACGCCACTTAACCAG		60°C
<i>Dnmt3a</i>	GAGGGAAGTGAACCCAC	CTGGAAGGTGAGTCTTGGCA	6681209a1	60°C
<i>Dnmt3b</i>	TCAGATGAGCAAGGTCAAGG	TGTACCAAAGCAAGGGGAAG		60°C
<i>Oct4</i>	CAGCCAGACCACCATCTGTC	GTCTCCGATTTGCATATC	7305399a3	58°C
<i>Nanog</i>	CTTTCACCTATTAAGGTGCTTGC	TGGCATCGGTTTCATCATGGTAC	4	58°C
<i>Sox2</i>	GCGGAGTGGAACCTTTGTCC	CGGGAAGCGTGACTTATCCTT		60°C
<i>Fgf5</i>	TGTGTCTCAGGGGATTGTAGG	AGCTGTTTTCTTGAATCTCTCC	6753854a1	60 °C
<i>Otx2</i>	TATCTAAAGCAACCGCTTACG	GCCCTAGTAAATGTCGTCTCTC	158518427c1	60°C
<i>Esrrb</i>	ATGCGAGTACATGCTTAACGC	CATCCCCACTTTGAGGCATT		60°C
<i>Klf4</i>	GCAGTCACAAGTCCCCTCTC	GACCTTCTCCCCTCTTGG	5	58°C

- Lehnertz, B. et al. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* 13, 1192–1200 (2003).
- Ferri, F., Bouzinba-Segard, H., Velasco, G., Hubé, F. & Francstel, C. Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.* 37, 5071–5080 (2009).
- Bernardo, A. S. et al. BRACHYURY and CDX2 Mediate BMP-Induced Differentiation of Human and Mouse Pluripotent Stem Cells into Embryonic and Extraembryonic Lineages. *Cell Stem Cell* 9, 144–155 (2011).
- Hayashi, K. & Surani, M. A. Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro. *Development* 136, 3549–3556 (2009).
- Jouneau, A. et al. Naive and primed murine pluripotent stem cells have distinct miRNA expression profiles. *RNA* 18, 253–64 (2012).

Supp Table S2

Titre: Dynamique de la réorganisation nucléaire accompagnant la conversion entre deux états pluripotents: l'état naïf (ESCs) et amorcé (EpiSCs)

Mots clés : pluripotence, épigénétique, organisation du noyau, hétérochromatine

Résumé : Les cellules souches embryonnaires de souris (ESCs) et les cellules souches de l'épiblaste (EpiSCs) représentent, respectivement, les états naïf et amorcé de la pluripotence et sont maintenues in vitro par des voies de signalisation spécifiques. De plus les ESCs cultivées dans un milieu sans sérum avec deux inhibiteurs (2i) sont décrites comme étant les plus naïves. Plusieurs études ont suggéré que chaque type de cellules pluripotentes est caractérisé par une organisation différente de l'épigénome. Nous présentons ici une étude comparative de l'état épigénétique et transcriptionnel des séquences satellites répétées péricentromériques (PCH) entre les ESCs (2i et sérum) et EpiSCs. Nous montrons que H3K27me3 au PCH est très dynamique et peut discriminer les ESCs en 2i des autres cellules souches pluripotentes. Alors que la transcription des séquences satellites est élevée dans les ESCs en sérum, elle est plus faible dans

ESCs en 2i et encore plus réprimée dans les EpiSCs. La suppression de la méthylation de l'ADN ou d'H3K9me3 dans les ESCs conduit à un dépôt important de H3K27me3 au PCH, mais peu de changements transcriptionnels de ces séquences. En revanche, l'absence d'H3K9me3 dans les EpiSCs n'empêche pas la méthylation de l'ADN au PCH, mais induit la transcription de ces séquences. La conversion in vitro des ESCs en EpiSCs est plus longue que le passage des cellules de l'ICM en épiblaste in vivo. Cette inefficacité ne peut pas être expliquée par une mise en place retardée du nouveau réseau transcriptionnel. Pour conclure notre étude a révélé que les EpiSCs ont perdu de la plasticité par rapport au ESCs sur l'hétérochromatine ainsi que l'euchromatine, comme le montre la réduction des niveaux d'H3K9ac et des domaines bivalents, étant ainsi plus proche épigénétiquement de cellules somatiques que de la pluripotence naïve.

Title : Dynamic of nuclear changes occurring during the conversion between naïve (ESCs) and primed (EpiSCs) pluripotent cells

Keywords : pluripotency, epigenetics, nuclear organization, heterochromatin

Abstract: Mouse embryonic stem cells (ESCs) and epiblast stem cells (EpiSCs) represent naïve and primed pluripotency states, respectively and are maintained in vitro using specific signaling pathways. Furthermore, ESCs cultured in serum-free medium with two inhibitors (2i) are described as being the most naïve. Several studies have suggested that each pluripotent cell type is characterized by a different epigenome organization. Here we present a comparative study of the epigenetic and transcriptional state of pericentromeric (PCH) satellite repeats in ESCs (2i and serum ones) and EpiSCs. We show that the pattern of H3K27me3 at PCH is highly dynamic and discriminate 2i-ESCs from the other pluripotent stem cells. Whereas satellites transcription is high in serum-ESCs, it is lower in 2i-ESCs and even more repressed in EpiSCs.

Removal of either DNA methylation or H3K9me3 in ESCs leads to enhanced deposition of H3K27me3 but few changes in satellite transcription. By contrast, in EpiSCs removal of H3K9me3 does not prevent DNA methylation at PCH but de-represses the satellite transcription. In vitro conversion from naïve to primed pluripotency showed an important delay compared to the in vivo development of ICM cells into post-implantation epiblast. Such inefficiency cannot be explained by a delayed switch to the new transcriptional network. Altogether our study reveals that EpiSCs have lost the chromatin plasticity of ESCs on heterochromatin as well as euchromatin, as shown by the reduction of H3K9ac levels and bivalent domains, thus being closer to somatic cells in terms of epigenetics than naïve pluripotency.