DOCTOR OF SCIENCES

Role of heme biosynthesis in the early stages of pluripotency

DETRAUX, Damien

Award date:
2021

Awarding institution:
University of Namur

University of Namur
Faculté des Sciences – Département de Biologie
Namur Research Institute for Life Sciences (NARILIS)
Unité de recherche en Biologie Cellulaire (URBC)
Rue de Bruxelles, 61, B-5000 Namur, Belgique

Role of heme biosynthesis in the early stages of pluripotency

Dissertation originale présentée par
Damien DETRAUX
en vue de l'obtention du grade de
Docteur en Sciences

Composition du jury :

Prof. Patsy Renard (Promoteur)
Unité de recherche en Biologie Cellulaire
Narilis, UNamur, Namur

Prof. Julie Mathieu
Department of comparative medicine
University of Washington, Seattle, USA

Prof. Thierry Arnould (Co-promoteur)
Unité de recherche en Biologie Cellulaire
Narilis, UNamur, Namur

Prof. Frederic Lluis Vinas
Department of Development and Regeneration
K.U. Leuven, Leuven, Belgium

Prof. Nicolas Gillet (Président)
Unité de recherche Vétérinaire Intégrée
Narilis, UNamur, Namur

Prof. Michel Jadot
Unité de recherche en physiologie moléculaire
Narilis, UNamur, Belgique

Avril 2021

Acknowledgments

Here we are! This section is the final stone for building this manuscript, and this manuscript is the pinnacle of my PhD. Along the way there are many people I would like to thank for the role they played, directly or indirectly to the completion of this important step.

First of all, I have to thank my promoter Patsy for the past 4 years. Even if this manuscript is submitted before the end of the four years of the thesis, I will gladly include the work we did together since the redaction of the first FRIA draft. I really believe that our personalities matched on several point and that it definitely made the path of the PhD easier. I guess our common optimism was due to lead in a good combination. I hope our recurrent impromptu meetings just to show the latest results because I'm excited about it were not too much, I just can't keep the good news for myself. Thank you for listening, for the discussions and of course for helping building the story together. Along with Patsy, I should also thank Thierry, as the co-promoter, for the multiple meetings and discussions together

Then, there is a particular acknowledgment for the members of the jury, and members of the committee meeting: First Pr. Nicolas Gillet for agreeing to be the president of the jury, then Profs. Julie Mathieu, Frederic Lluis Vinas and Michel Jadot. Thank you all for being part of this adventure, especially for taking time to read the manuscript and for your comments and the constructive discussions during the defense. I will add a special mention for Julie since you have been part of my scientific education since 2015 in the HRB lab: thank you for being by my side since that time.

Ultimately, what makes the day-to-day life in the lab enjoyable is the people you share it with. For that, I believe that the atmosphere in URBC and within the team has always been one of the major factors leading to the completion of this work. First, I would like to mention my colleagues of the 4127 office: Myriam, Camille and Alexis. Thank you for the trips to Coffee & More, the Badminton games, the games nights, the numerous talks in the office or even the Trips to Canada, Japan, the Moselle Valley or the Ardennes. In these, the "4133 & associates" was also a huge part, somehow it was my second office. Thank you to Julie, Marino, Céline and Lola, we shared so many moments together, unfortunately especially before Covid, and I hope we will still share many more in the future. Now that Sébastien is also an official 4133 member he should also be included in this, especially for the good times we share him, Marino and I (sometimes with Lola) at the personal benches in the lab. Our weekly discussions about the latest results were always useful.

Overall, I believe the atmosphere in general was a great factor that raised morale when it was more challenging. In this, the rest of the Dyso Team was fundamental, thank you for listening, sharing and helping whenever it was needed. More than just the Dyso Team, the whole lab was always fun to work with, but not only work, the activities we would do together, ALL the food or the Beer Hours!

Practically, this project would not have moved like it did without the technical team supporting the URBC. Thank you to them, but especially to Maude for her involvement for the

latest experiments, without you, or M. Dieu, I definitely wouldn't have been able to complete any of this.

The support received outside of the lab was also fundamental for the successful completion of the thesis. I believe I owe Minou, Fred, Chabwabwa and Julie many thanks for being present in the good and the bad times, fortunately for them mostly the good ones. Being able to escape mentally during our numerous parties, restaurants, trips, or wine and games was crucial. It is too bad we cannot properly celebrate this year… Along with them, thanks to my family for trying to understand the issues but also for believing in me all his time.

Finally, and oddly enough, I would like the world pandemic for closing all the restaurants and bar, giving me all the time need without distractions for the completion of this manuscript. It is over now, you can disappear!

Abstract

Using human embryonic stem cells (hESC) in regenerative medicine or in disease modeling requires a complete understanding of these cells. Developmentally speaking, two distinct states of ESC have been stabilized, a naïve pre-implantation stage and a post-implantation primed stage. Among the numerous changes observed during this transition, the mitochondria show a transition to maturity on a morphologic point of view but also a strong reduction in its metabolic oxidative phosphorylation activity, suggesting another role for the organelle. Performing a deep analysis of two recently published CRISPR-Cas9 KO functional screens, we identified the heme synthesis, a partly mitochondrial pathway, as critical for the naïve-to-primed transition *in vitro*. We show here an impairment of the exit of the naïve state upon blockade of the pathway by reducing the heme synthesis rate with succinylacetone (SA), a chemical inhibitor of the ALAD, the second enzyme of the pathway. Interestingly, ESC pushed to exit the naïve stage with SA fail to activate the two crucial signaling pathways for the transition: the MAPK and TGFβ. In parallel, heme synthesis inhibition promoted the acquisition of features found in the 2-cell embryo, as well as in a subpopulation of ESCs. We showed that this effect was heme-independent and due to the accumulation of succinate in the mitochondria, leaking to the other cellular compartments, since blocking the mitochondrial exit of succinate was able to prevent the acquisition of the 2C-like features. We propose that extra-mitochondrial succinate induces the succinylation of proteins, including regulators of the 2C-like-cells (2CLCs) such as nucleolin and TRIM28. Further experiments are required to pinpoint the exact mechanisms controlling both the exit of the naïve state and the acquisition of 2CLCs features. Overall, this study aims to unveil the mechanisms underlying the maintenance of pluripotent cells in early development.

Table of contents

ACKNOWLEDGMENTS....2

ABSTRACT....4

LIST OF ABBREVIATIONS....7

INTRODUCTION....11

I. Embryo development....11
A. Generalities....11
B. From zygote to blastocyst....11
C. From blastocyst to post-implantation differentiation....14

II. Stem cells....16
A. Potency....16
B. Classification....18
C. Embryonic stem cells....19
D. The naïve and primed pluripotent stem cells....21
1. Signaling dependency....22
2. Molecular features....24
3. Naïve-to-primed ESC transition....29
E. Alternative states....31
1. Formative pluripotency....31
2. Poised state....31
3. 2C-like cells....32
4. Diapause....34

III. Heme: Synthesis and function....34
A. The heme biosynthetic pathway....35
B. Heme as a prosthetic group....37
C. Heme trafficking....39

IV. Succinate, succinyl-CoA and protein succinylation....40
A. Metabolic functions....40
B. Succinate as a post-translational modification....41
C. Succinate-mediated regulation of cellular functions....43

OBJECTIVES....44

MATERIAL & METHODS....45
A. Cell culture....45
B. mESC treatments....45
C. RNA extraction and RT-qPCR....45
D. Western blot analyses....47
E. Puromycin-incorporation assay....48
F. Immunofluorescence....48
G. Protein digestion and Succinyl-lysine pull-down....48
H. Mass spectrometry....49

1. Mass spectrometry analyses....49
2. Database searching....50
3. Criteria for protein identification....50
I. Co-Immunoprecipitation....50
J. RNA sequencing....50
K. Data analysis....51

RESULTS....52

I. Identification of critical pathways in the naïve-to-primed ESC transition in human and mouse....52

II. Heme biosynthesis in the naïve-to-primed mESC transition....53
A. Heme synthesis inhibition....53
B. Validation of the requirement for heme synthesis....55
C. Investigation of the mechanisms for the transition defect....56
1. BACH1....56
2. ISR....56
3. Signaling pathways....58

III. Heme synthesis inhibition in naïve mESCs....59
A. Acquisition of a 2C-like phenotype....59
B. Heme-independent acquisition of the 2C features....61

DISCUSSION AND PERSPECTIVES....64
A. Identification of the developmental stage....64
B. Signaling pathways....64
1. MAPK....65
2. Activin A – SMADs....67
3. Other pathways....69
C. 2C reprogramming....69
D. General considerations....72
1. Human – Mouse differences....72
2. Generation of KO lines....73
3. *In vitro* models....73

CONCLUSIONS....75

BIBLIOGRAPHY....76

SUPPLEMENTARY INFORMATION....94

List of abbreviations

α-KGDH= oxoglutarate dehydrogenase complex
2C= 2-cell embryo
2CLCs= 2C-like cells
5mC= 5-methylcytosine
ABCB10= ATP-binding cassette protein B10
ACVR2B= Activin receptor type-2B
ADP= adenosine diphosphate
AIP= acute intermittent porphyria
AKT= protein kinase B
ALA= δ-aminolevulinic acid
ALAD= ALA dehydratase ALAD
ALAS= aminolevulinate synthase
ATF4= activating transcription factor 4
BACH1= BTB and CNC homolog 1
BM-MSCs= bone marrow-derived mesenchymal stem cells
BMP2= Bone morphogenic protein 2
BMP4= Bone morphogenic protein 4
BRAF= serine-threonine kinase B-Raf
BTdCPU= 1-(benzo[d][1,2,3]thiadiazol-6-yl)-3-(3,4-dichlorophenyl)urea
CAF-1= chromatin assembly factor 1
cAMP= cyclic adenosine monophosphate
CER1= Cerberus protein
cGMP= cyclic guanosine monophosphate
CPOX= coproporphyrinogen oxidase
CPT1= carnitine palmitoyltransferase 1
DE= distal enhence
DNA= deoxyribonucleic acid
DNMT3a= DNA methyltransferases 3a
DNMT3b= DNA methyltransferases 3b
Dox= Doxycyclin
DUB1= Ubiquitin carboxyl-terminal hydrolase
Dux= double homeobox
EB= embryoid bodies
EGLN1= Egl nine homolog 1
EGR-1= Early growth response 1
EIF2α= eukaryotic translation initiation factor 2α
ELK1= ETS domain-containing protein Elk-1
EMT= epithelial-to-mesenchymal transition
ENO1= enolase
EpiLC= epiblast-like stem cell
EpiSCs= epiblast stem cells = post-implantation mESCs
ER= endoplasmic reticulum
ERK= Extracellular signal-Regulated Kinases
ERV= endogenous retrovirus

ESC= embryonic stem cell
ESRRB= estrogen Related Receptor β
ETC= electron transport chain
FA= fatty acids
$FADH_2$= flavin adenine dinucleotide
FAH= fumarylacetoacetic hydrolase
FECH= Ferrochelatase
FGF 5= Fibroblast growth factor 5
FGF15= Fibroblast growth factor 2
FGF2= Fibroblast growth factor 2
FLCN= folliculin
FOXA2= Forkhead box protein A2
FRS2= Fibroblast Growth Factor Receptor Substrate 2
GAPDH= Glyceraldehyde 3-phosphate dehydrogenase
GATA1= GATA-binding factor 1
GATA3= GATA-binding factor 3
GDP= Guanosine diphosphate
GFP= Green fluorescent protein
GO= Gene ontology
GRB2= Growth factor receptor-bound protein 2
GSEA= gene set enrichment analysis
GSK3β= glycogen synthase kinase 3β
H3K27me3= Histone 3 lysine 27 tri-methyl mark
HAT1= histone acetyltransferase 1
HCCS= holocytochrome c synthase
HDAC= Histone deacetylase
HDM= histone demethylase
HERVK= Human endogenous retrovirus K
HIF-1α= Hypoxia-inducible factor 1-α
HRE= hypoxia-responsive element
HRI= Heme-regulated inhibitor
HRM= heme regulatory motif
ICM= inner cell mass
IDO1= indolamine2,3 dioxygenase
IL-10= interleukin 10
IL-8= interleukin 8
IMM= inner mitochondrial membrane
IMS= intermembrane space
iNOS= inducible nitric oxide synthase
iPSCs= induced pluripotent stem cells
ISR= Integrated stress response
ISY1= Pre-mRNA-splicing factor ISY1 homolog
IVF= in-vitro fertilization
JAK= Janus kinase
JmjD3= Jumonji domain-containing protein D3
JNK= c-Jun N-terminal kinase
KAP1= KRAB-associated protein-1

KAT2A= lysine acetyltransferase 2A
KEGG= Kyoto Encyclopedia of Genes and Genomes
KLF2= Kruppel Like Factor 2
KLF4= Kruppel Like Factor 4
KO= knock-out
LEF1= Lymphoid Enhancer Binding Factor 1
LIF= leukemia inhibitory factor
LIN28= lin-28 homolog
LINEs= long-interspaced nuclear elements
List of abbreviations
MAM= mitochondria-associated membrane
MAPK= mitogen-activated protein kinase
MARE= Maf recognition element
MEF= murine embryonic fibroblast
MEK= mitogen-activated protein kinase kinase
MFF= mitochondrial fission factor
MHC= major histocompatibility complex
MSC= mesenchymal stem cells
MTCH2= mitochondrial carrier homolog 2
mTOR= mammalian target of rapamycin
MuERV-L= murine endogenous retrovirus-like
MZT= maternal-to-zygotic transition
NAD= nicotinamide adenine dinucleotide
NADH= Nicotinamide adenine dinucleotide
Nanog= Nanog homeobox
NHSM= Naive Human Stem cell Medium
NMPP= N-methylprotoporphyrin IX
NNMT= Nicotinamide N-methyl transferase
NO= nitric oxide
NOS= nitric oxide synthases
Oct4= octamer-binding transcription factor 4
OTX2= Orthodenticle homeobox 2
OXPHOS= oxidative phosphorylation
PBG= porphobilinogen
PBGD= PBG deaminase
PBGS= porphobilinogen synthase
PCA= principal component analysis
PDK1= pyruvate dehydrogenase kinase 1
PE= prximal enhancer
PGC= primordial germ cell
PGC1α= peroxisome proliferator-activated coactivator 1α
PGRMC1= progesterone receptor membrane component 1
PGRMC2= progesterone receptor membrane component 2
PHD= prolyl-hydroxylase domain
PIAS4= small ubiquitin-like modifier (Sumo) E3 ligase protein inhibitor of activated STAT 4
PKC= Protein kinase C
PPOX= protoporphyrinogen oxidase

PRC2= polycomb repressive complex 2
PSC= Pluripotent stem cell
PTM= post-translational modification
pVHL= von Hippel-Lindau protein
RAF= Rapidly Accelerated Fibrosarcoma
RAR= retinoic acid receptor
RAS= proto-oncogene protein p21
REX1= Reduced Expression 1
RNS= reactive nitrogen species
ROCK= Rho-associated, coiled-coil containing protein kinase
ROS= reactive oxygen species
SA= succinylacetone
SAM= S-adenosyl methionine
SDH= succinate dehydrogenase complex
SINEs= short-interspaced nuclear elements
Sirt7=sirtuin 7
Sirt5= sirtuin 5
SMAD= Mothers against decapentaplegic homolog
SOS= Son of Sevenless
Sox2= sex-determining region Y-box 2
SRC= Proto-oncogene tyrosine-protein kinase
STAT3= signal transducer and activation of transcription 3
SUCL= succinate-CoA ligase
SUCNR1= succinate receptor 1
TBP= TATA-binding protein
TBX3= T-box 3
TBXT= T-box transcription factor T
TCA = tricarboxylic acid cycle
TCSTV1= 2-cell-stage, variable group, member 1
TE= trophectoderm
TET= Ten eleven translocation
TF= Transcription factor
TFCP2L1= Transcription factor CP2-like 1
TFE3= Transcription Factor Binding To IGHM Enhancer 3
TGF-β= Transforming growth factor
TSS= Transcription start site
UbQ= ubiquinone
UHRF1= Ubiquitin-like, with PHD and RING finger domains 1
uORF= upstream open reading frame
UROD= Uroporphyrinogen decarboxylase
UROS= URO3S= uroporphyrinogen III synthase
YAP1= yes-associated protein 1
ZGA= zygote genome activation
ZIC2 = Zic family member 2
ZO-1= zona occludens 1
Zscan4= Zinc finger and SCAN domain containing 4

Introduction

I. Embryo development

A. Generalities

It is always fascinating to try to understand the mechanisms that ultimately lead to the formation of a new being. This question of the origins always has been a major challenge and lead to the developmental biology field. In the past, most of the research was limited to the embryo collections such as the Carnegie collection, the Kyoto collection or the Madrid collection. These collections gather human embryo samples at different stages that allow morphological analysis, crucial to understand development (Yamada, Hill and Takakuwa, 2015). But while these collections allow the understanding of the later steps of embryo development, they however don't inform much on the earlier steps, from the zygote to the implantation of the blastocyst, steps that are at the center of this work. It is only with the rise of in-vitro fertilization (IVF) techniques that knowledge on these stages started to build. Historically, the work on IVF began in the middle of the 20th century with the pioneer work of R. Edwards, P, Steptoe and J. Purdy (Edwards, Bavister and Steptoe, 1969; Edwards, Steptoe and Purdy, 1970), that allowed the culture of *in-vitro* human eggs for a few days, which lead ultimately to the birth of the first IVF baby in 1978. Apart from the birth of IVF babies, these developments also allowed the characterization of the early developmental steps such as the zygote genome activation (detailed in the next section), the different steps of cleavage, compaction or even blastulation. Following these findings, embryonic stem cell lines from human or mouse were then stabilized for *in-vitro* culture, expanding even more our knowledge on early development (Evans and Kaufman, 1981; Martin, 1981; Thomson *et al.*, 1998).

As much as human development is captivating, this work will be focused on mouse development. In the embryology field, mice are attractive since they produce a relatively large litter (between 8 to 20) but also develop quickly (21 days). These features, combined to the ability to edit their genome, gave a serious advantage to this model, allowing to decipher the molecular mechanisms governing embryo development. Hence, since mice share the same tissues and organs as humans, understanding the mechanisms leading to the development of these organs is the key to someday be able to reconstruct these organs ex-vivo for a potential use in regenerative medicine. Finally, the use of mice tissues and samples is first preferred for obvious ethical barriers. In the next sections, we will discuss the major steps in embryo development, first from the zygote to the blastocyst, then from the blastocyst to the early differentiation steps.

B. From zygote to blastocyst

The formation of a zygote is the result of the fusion of two gametes: The egg as the female gamete and spermatozoa as male gamete. The union between those two is a really complicated process, which involves many molecular steps. The complexity of this association will not be detailed in this document but is nicely reviewed in (Trebichalská and Holubcová,

2020). The resulting zygote marks the entry into the pre-implantation period of embryonic development (**Fig. 1A**). From then on, the age of the embryo is reported depending on the number of days post-mating since the exact time of fertilization is unknown *in-vivo*. In the following section, we will thus refer as Embryonic (E) day to discriminate the developmental stage.

Soon after the fertilization, both of the germ cell genomes must be reorganized in order to become totipotent. Indeed, after their conjugation, there is a stage with little to no transcription, that instead relies almost entirely on the maternal transcripts already present in high number in the egg. These transcripts actually drive this dramatic reprograming event, as elegantly shown by the work of J. Gurdon, who showed that replacing the genome of the zygote by the nuclei of a somatic cell could reprogram this nuclei to give rise to a developing embryo (Gurdon, 1962). This nuclear transfer elegantly showed the plasticity of the genome when exposed to the right stimuli, but also showed that the maternal egg contained all the factors necessary for this reprogramming.

At this time in the zygote, there is thus a switch between the maternal mRNAs and proteins present in the egg, to the transcription and translation of zygote-encoded genes. This switch is called maternal-to-zygotic transition (MZT). It is during this transition that the zygote genome activation (ZGA) takes place (Schulz and Harrison, 2019). This is actually a two-step process, the first one proceeding during the S-phase of the 1 cell stage (the minor ZGA),

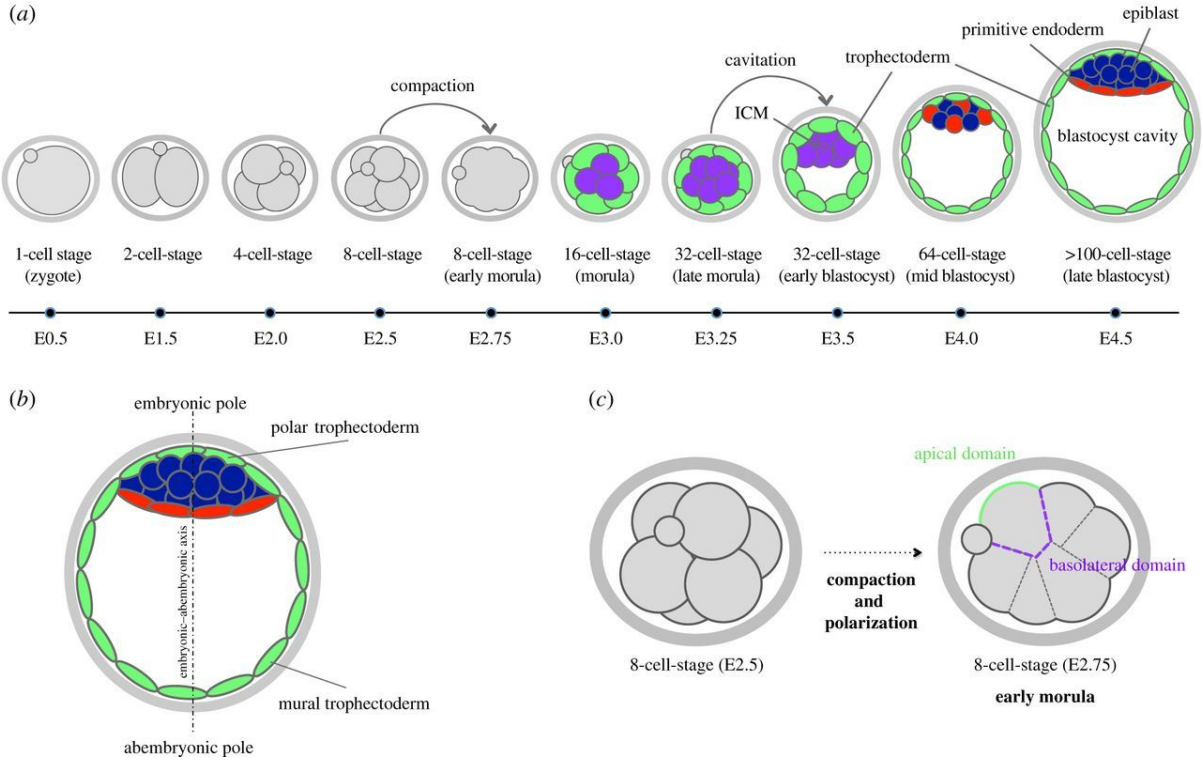


Figure 1 Schematic diagram of the pre-implantation mouse embryo development. *A) Temporal timeline of the pre-implantation embryo from the zygote (E0.5) to the mature blastocyst (E4.5). The first and second cell fate commitment are represented by the change of color of the cells. Green= trophectoderm lineage (TE), purple= Inner cell mass (ICM), red= primitive endoderm and blue= epiblast. B) Representation of the polarity of the blastocyst at E4.5 with the polar embryonic pole harboring the ICM and the abembryonic pole on the other side of the blastocoel cavity (also caller blastocyst cavity). C) Representation of the compaction step forming the morula, and representation on the tight junction created between the blastomeres, ultimately creating a cellular polarity giving rise to the first cell commitment: TE cells with an apical domain and ICM with only basolateral domains. From (Mihajlović and Bruce, 2017).*

followed by a major ZGA in the mid-to-late 2 cell stage. There is thus a time lag between fertilization and the ZGA. The causes for the delay in ZGA after fertilization are numerous. We can cite, among others, the need for accumulation of regulators of transcription, such as the TATA-binding protein (TBP), required for the initiation of transcription. Indeed, these TBP are translated from maternal mRNAs present in the egg and it would thus take some time to translate enough of the protein to induces a strong enough transcriptional activity (Veenstra, Destrée and Wolffe, 1999). One can also mention the replacement of protamines packing tightly the chromatin in the male gamete by maternally-encoded histones to allow an effective transcription of the zygote's chromatin (Braun, 2001; Joseph *et al.*, 2017), but also changes in histone post-translational modifications, reviewed in (Schulz and Harrison, 2019). All these processes take time to settle in and result in a delay of transcription of the zygote genome. In mammals, this ZGA coincides with the reactivation of some transposons or endogenous retroviruses but only temporarily and they are being re-silenced later. This is exemplified with the endogenous retroviruses MuERV-L (murine endogenous retrovirus-like) in mouse (Macfarlan *et al.*, 2012) or HERVK (Human endogenous retrovirus K) in humans (Grow *et al.*, 2015).

Specific activating transcription factors have also been identified as driving the initiation of the ZGA such as Zelda in drosophila (Harrison *et al.*, 2011) or the core pluripotency proteins Nanog (Nanog homeobox), Oct4 (Octamer-binding transcription factor 4) and SoxB1 (SRY (sex determining region Y)-box) in zebrafish (Lee *et al.*, 2013). However, while these factors are known to be involved in the transcriptional activity of pre-implantation stem cells in mouse and human, their implication in the ZGA is far from being understood. On the other hand, in common in humans and mice, several actors have been designated. First, the transcriptional activator YAP1 (yes-associated protein 1) is present in the cytoplasm of the zygote but is progressively translocated in the nuclei starting from the 2-cell stage and its genetic ablation results in a drastic decrease in the expression of ZGA genes (Abbassi *et al.*, 2016). Then, transcription factors such as the DUX protein (double homeobox; or *DUX4* in human) are major drivers of 2-cell stage phenotype. This gene has been linked to the activation of genes during ZGA as well as transposons and retrovirus such as HERVK (De Iaco *et al.*, 2017; Hendrickson *et al.*, 2017). Although its role is demonstrated, the mechanisms leading to its *in vivo* activation are not. Finally the *Zscan4* cluster of genes (Zinc finger and SCAN domain containing 4; *Zscan4a-f*) was shown to regulate events related to the telomeres maintenance or the chromatin decompaction (Falco *et al.*, 2007).

Following this ZGA, the embryo proceeds with the division cycles without affecting the total cytoplasmic volume until the late blastocyst stage (E4) (Kojima, Tam and Tam, 2014). Because of this, after the 4th division, the cells will begin to undergo compaction events, triggering the establishment of tight junctions between the blastomeres (Hyenne *et al.*, 2005; Fierro-González *et al.*, 2013), leading to the formation of the morula stage (E2.5-E2.75). At this stage, the tight junctions create a polarity of the cell, with a basal and apical site, depending on their presence or not (**Fig. 1C**). This will allocate the appropriate subsequent cell fate identity (Fleming *et al.*, 1989). These tight junctions and their associated proteins, such as the zona occludens 1 (ZO-1) proteins are necessary for the progression toward the next stages as the depletion of this protein impairs blastocyst formation (Wang *et al.*, 2008). This polarity will give rise to the first cell fate decision of the embryo: separation of the trophectoderm (TE) on the outside, with both an apical and a basolateral side with tight junctions, and the inner

cell mass (ICM) on the inside, with only basolateral sides with the surrounding cells. At E3.5 an osmotic gradient will be created by an active transport of Na^+ by the TE cells, triggering the transport of water across the developing TE layer, forming a fluid-filled cavity, the blastocoel (Manejwala, Cragoe and Schultz, 1989). This cavity is asymmetrical, leaving the ICM on one side of the blastocyst, creating two different sides of this embryo: embryonic (with the ICM) and abembryonic (**Fig. 1B**).

C. From blastocyst to post-implantation differentiation

As the blastocyst progress in development, the ICM segregates into two different lineages at E4.5: the monolayer of primitive endoderm, on the path of differentiation and found at the border with the blastocoel cavity, and the pluripotent epiblast, between the polar TE and the primitive endoderm (**Fig. 1B**). This is the second cell fate decision made by the developing embryo (Chazaud *et al.*, 2006). At this point, the blastocyst emerges from the zona pellucida, the glycoprotein layer surrounding it since the egg stage, in a step defined as the blastocyst hatching. Finally free, the blastocyst will then implant in the uterine wall and proceeds with the post-implantation development. This step is one of the most critical in the establishment of a successful pregnancy as it is estimated that up to 50% of fertilized eggs fail to implant (Wilcox *et al.*, 2020).

As the blastocyst nests in the uterus, the ICM grows toward the blastocoel cavity, forming a peri-implantation epiblast, later organized as an epithelium around a central lumen (the proamniotic cavity) and becoming the post-implantation epiblast at E6 (Bedzhov and

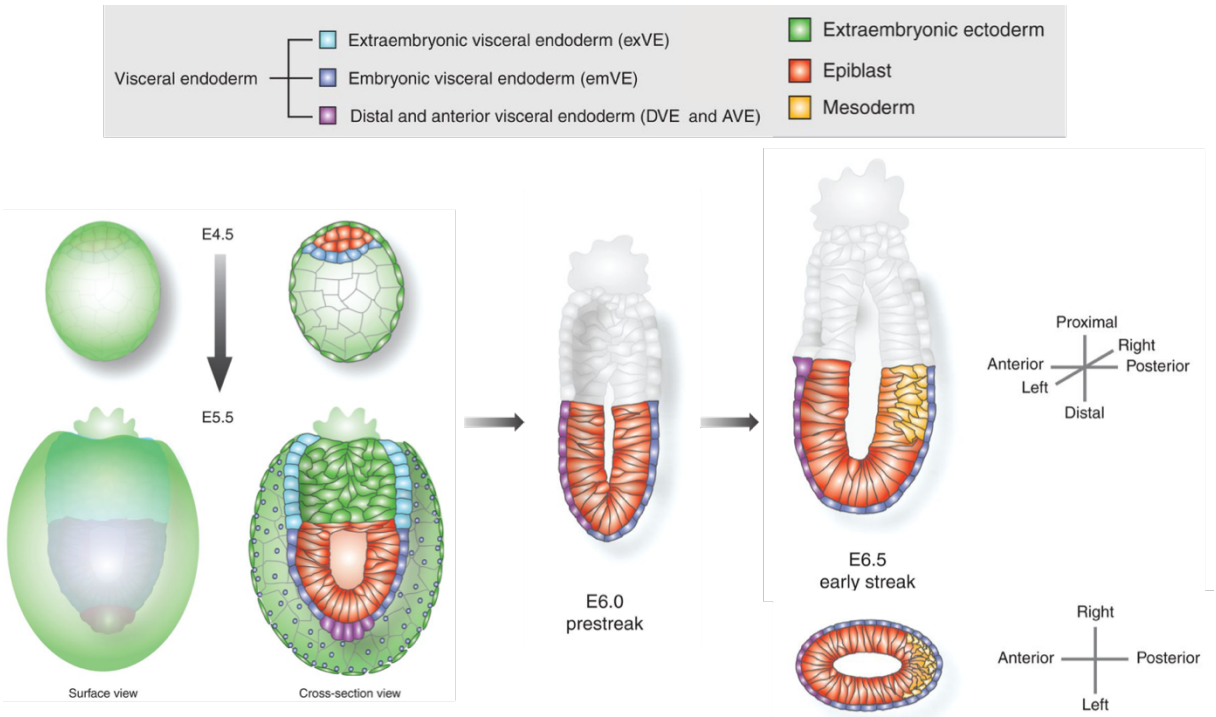


Figure 2 Schematic representation of the development of the post-implantation blastocyst. *The hatched blastocyst implants in the uterine wall at around E5. Following this step, the egg cylinder elongates, a second step of cavitation taking place in the middle of the epiblast cells, forming the proamniotic cavity. Around the epiblast, the visceral endoderm separates in three distinct regions: Extraembryonic visceral endoderm (exVE), the embryonic visceral endoderm (emVE) and the distal visceral endoderm (DVE), respectively from the most proximal to the most apical. At E6, the reorganization of the distal endoderm to the anterior region is concomitant to the initiation of the gastrulation. This alignment gives also rise to the mesoderm layer in the posterior region, between the epiblast and the emVE. Modified from (Rivera-Pérez and Hadjantonakis, 2014).*

Zernicka-Goetz, 2014) (**Fig. 2**). The specification of the three primary germ layers is then set in the posterior region of the epiblast: the primitive streak. In this region, at E6.5, dramatic changes are observed as the epiblast undergoes an epithelial-to-mesenchymal transition (EMT) reorganizing the junctional complexes, the cytoskeleton and the basement membrane (Nakaya *et al.*, 2008). This initiates the key step of gastrulation, with emergence of mesoderm and further specifications of germ layers. Indeed, the epiblast bordering the proamniotic cavity will progressively become the ectoderm layer, the primitive endoderm facing the blastocoel cavity, and the mesoderm in between the two (**Fig. 2**). This gastrulation event marks the point when the cells in the embryo start losing their pluripotent capacity. The different lineages continue their specification, with colonization of the anterior part by the

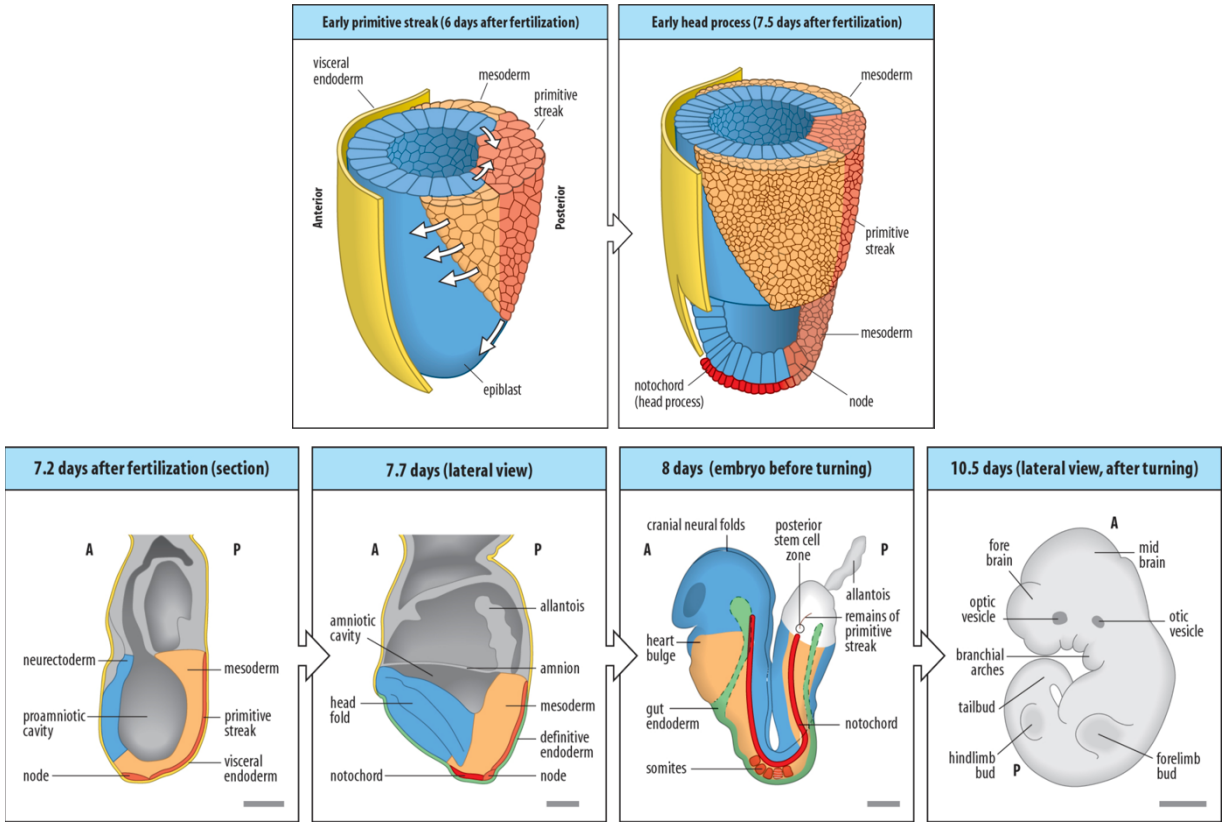


Figure 3 Schematic representation of the post-implantation embryo development. *During the later stages of gastrulation, the mesoderm layer (in orange) colonize the embryo leading to the heart bulge at E8. Meanwhile, the epiblast is specified into neurectoderm (blue), while the primitive streak expand to ultimately form the notochord (red). Around E8, the endoderm invaginates to form the gut endoderm. Finally, between E8 and E10 the embryo turns on itself to end up completely surrounded by the amniotic cavity. From (Wolpert, 2015)*

endoderm to form the gut endoderm, elongation of the primitive streak forming ultimately the notochord and the mesoderm giving rise to specialized regions such as the heart bulge in the anterior region (Wolpert, 2015) (**Fig. 3**).

At this point the cells have committed to a particular cell fate and will continue on the path of differentiation and commitment. The study of these differentiation mechanisms is crucial for the field of regenerative medicine as the development of stem cell-based therapies constitutes a promising field. But beyond the understanding of the differentiation process and organogenesis, understanding the cellular and molecular mechanisms involved in stem cell maintenance and specification are also imperative. This is why the following section will be focused on pluripotent stem cells.

II. Stem cells

As fully developed adult beings, we rely on two crucial processes to fully form and to maintain the homeostasis of this state: A correct and controlled development and a capacity for regeneration to balance cell death in organs. These two processes are held by stem cells, bearing the capacity for cellular differentiation but also for self-renewal. Because of these features, stem cells are also of growing interest in regenerative medicine with applications in cellular therapy or even organ replacement. Besides, to be able to harness their full potential, we also need to understand them in depth, deciphering the mechanisms governing them and regulating their potential.

The stem cell identity is pretty heterogenous and complex. To better understand them, they can be defined by different categories depending on two major characteristics: their potency or their tissue of origin.

A. Potency

As stem cells commit into a particular fate, they gradually lose potency such as explained in the embryo development. In 1957 Conrad Hal Waddington described the commitment of stem cells as a linear descent of a ball from a mountain (Waddington and Kacser, 1957) (**Fig. 4**). The cellular commitment is symbolized by the rolling down into a particular valley until reaching the most differentiated state at the bottom of the hill. The ball, or stem cell in this case, at the very top of the mountain would be thus considered as totipotent and as the ultimate stem cell potency state. These totipotent cells are represented by a very restricted type of cells: in mammals it is the zygote and the first blastomeres from the newly formed embryo. At this stage, all cellular fates are possible, from the three germ layers of the embryo to the extra-embryonic tissues such as the placenta. Very quickly, after a few rounds of divisions and at E2.5, the cells will have to enter one of the mountain's valley

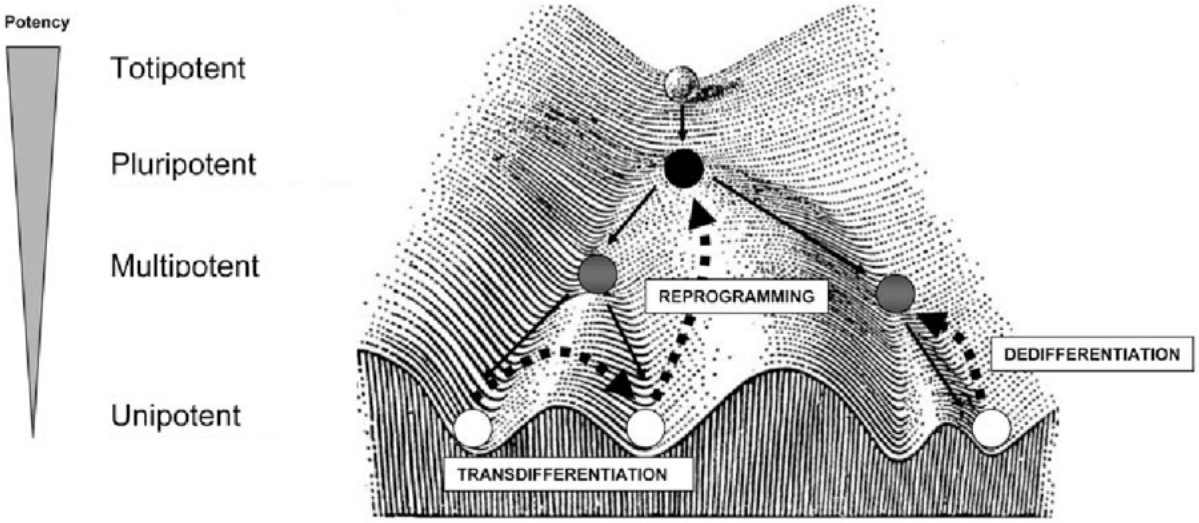


Figure 4 Representation of Waddington's landscape in the context of stem cell differentiation*. Totipotent are on top of the hill, with the potential to differentiate into any cell type. While going down the hill, stem cells lose potential, with first pluripotent cells leading to the embryo three germ layers and multipotent cells committed to one lineage. Finally, unipotent and differentiated cells lie at the bottom of the hill. Reprogramming or dedifferentiation events bring the cells back up, toward a pluri- or multipotent stage while transdifferentiation convert differentiated cells of one type to another.* (Eguizabal *et al.*, 2013)

to commit to one of the two paths: The TE leading to extraembryonic tissues or the ICM for embryonic tissues. Having segregated from the TE cells, the cells committed toward an embryonic fate are now pluripotent, holding only the ability to form any of the embryonic tissues of the three germ layers. *In vitro*, this stage would be represented by the embryonic stem cells in the ICM (ESCs) that will be further described in section II.C. With this high potential, both totipotent and pluripotent cells are especially of interest for cellular therapies since they can differentiate into any adult cell type. Further along the development, or the hill's descent, embryonic stem cells of the post-implantation epiblast will need to commit to one of the three embryonic germ layers: endoderm, mesoderm or ectoderm. Passed this point, the cell potential is restricted to one germ layer fate and will thus be called multipotency. The stem cells will finally become more and more specified as such as only one cell type is possible as potential fate, the cells became unipotent. It is typically the case for satellite cells or epidermal stem cells.

According to this classification, embryonic stem cells, as pluripotent, hold great promises for regenerative medicine. However, there are still many limitations to their use as therapies: their sourcing is obviously ethically complicated since it implies harvesting developing embryos. In addition, there are also some concerns about their immunogenicity since they can hardly be used as autogenic grafts. Furthermore, there are also safety concerns as it has been shown that their incomplete differentiation could give rise to teratomas in the site of injection (Gordeeva and Khaydukov, 2017). To circumvent these issues, a lot of efforts have been put in the use of adult stem cells. Indeed, it would be convenient to harvest highly potent stem cells from the body, differentiate them *ex-vivo* and then implant them as cell therapy. These cells can actually be found in the bone marrow, as hematopoietic stem cells or mesenchymal stem cells (MSCs) or in different tissues such as adipose tissue or dental pulp. As opposed to ESCs, these somatic MSCs are "only" multipotent, offering a restricted differentiation capacity and *in vitro* self-renewal. The advances on their role and use in the clinic is nicely reviewed in (Levy *et al.*, 2020).

Long considered as unidirectional, the evolution and the restriction of the cell potential has been challenged by many research groups. For decades, a lot of efforts have been put into cell reprogramming, the art of bringing back stemness capabilities to fully differentiated cells. These efforts culminated in the 2012 Physiology and Medicine Nobel prize awarded to J. Gurdon and S. Yamanaka. J. Gurdon, whose work was described in section I.B, showed the reprogramming of a somatic nuclei by the egg's factors present in the cytoplasm. This nuclear transfer method paved the way for further studies in the field. In 2006, S. Yamanaka and K. Takahashi identified a series of factors that, when overexpressed in somatic cells, were able to reprogram the cells into a pluripotent state. Successful reprogramming of dermal fibroblasts into pluripotent stem cells, then called induced pluripotent stem cells (iPSCs), was a major breakthrough (Takahashi and Yamanaka, 2006). With the exogenous expression of a few key transcription factors (OCT3/4, KLF4 (Kruppel-like factor 4), SOX2 and MYC (Proto-oncogene c-Myc); called the OKSM cocktail, or the Yamanaka factors), the fully specialized fibroblast went back up Waddington's hill to re-acquire pluripotent features and differentiation capacity. This discovery started a golden age in the stem cell field since these iPSCs could bypass all the caveats of ESCs, such as sourcing or immunogenicity.

B. Classification

The description of stem cells based on their potential for differentiation is usually the most informative. Aside from that, stem cells can also be defined with regards to their tissue of origin.

Totipotent stem cells, as described in the previous section, are only found in a tight window of time and space: at the 1- and 2-cell stages of the embryo. There, the blastomeres still haven't committed to any particular fate, capabilities that will disappear soon after as the cells divide (see section I.B.). In the case of pluripotent stem cells two subdivisions are made: on one hand it encompass cells found in the ICM of the embryo (from E2.5 to E6.5-7) that are thus called embryonic stem cells (ESCs) and that will account for the "naturally" occurring pluripotent stem cells, and on the other hand, the iPSCs developed by Yamanaka (Takahashi and Yamanaka, 2006), that would be the reprogrammed, artificial ones. These two populations retain the ability to form the three germ layers.

Adult/somatic tissues also contain stem cells, although these are at best multipotent. The MSCs represent a wide variety of those and are initially derived from the mesoderm layer. They are found in a wide variety of tissues such as adipose tissues or the dental pulp, but initial studies focused on bone marrow-derived stem cells (BM-MSCs). The general definition of MSCs involves the expression of a subset of surface markers (such as CD73, CD90 or CD103; clusters of differentiation) and lack the expression of others (HLA-DR, CD14, CD19 or CD45) but especially the ability to differentiate into fat, bone and cartilage (Wobma and Satwani, 2021). Importantly, there is a lot of interest in those cells as they lack of expression of the major histocompatibility complex (MHC) thus avoiding the immune system recognition. This capacity, coupled to their immunosuppressive action makes them attractive for regenerative medicine. Indeed, they have been shown to secrete a variety of anti-inflammatory or immunosuppressive proteins such as kynurenine through the expression of indolamine2,3 dioxygenase (IDO1) enzyme, or cytokines such as interleukins 8 or 10 (IL-8 and -10) (Pittenger *et al.*, 2019). These properties explained that MSCs are harnessed, and used as therapies, for several pathologies linked with inflammation such as osteoarthritis or myocardial infraction or even amyotrophic lateral sclerosis. However, the hype that followed the positive results of preclinical studies is fading as most of the current clinical studies fail to replicate them robustly enough (Levy *et al.*, 2020).

For the control of tissue homeostasis, resident populations of stem cells are also present to balance the natural cell death or damages from injuries. The cells, such as satellite cells in muscle or intestinal stem cells, will divide and replace the lost cells to maintain the organ function (Ferraro, Lo Celso and Scadden, 2010). These cells are somewhat rare and are often considered as unipotent: Intestinal stem cells will divide asymmetrically to provide a daughter stem cell and a daughter cell that will replenish the intestinal crypt epithelium. While this process of division and differentiation is constant in this population, since the crypts are renewed frequently (in mouse the small intestine renews every 5 days (Barker, De Wetering and Clevers, 2008)), some adult stem cell populations can remain dormant for a while if not presented with the right stimuli, such as muscle stem cells that will be activated for the production of new myofibrils in response to injury of weight-bearing exercise (Schultz, 1996).

C. Embryonic stem cells

The successful isolation of cells from the mouse pre-implantation blastocyst by Evans, Kaufman and Martin (Evans and Kaufman, 1981; Martin, 1981) in 1981 enabled scientists to investigate the molecular mechanisms that are the root of pluripotency. Before then, researchers eventually used cells from a teratocarcinoma, a form of malignancy from the germ line cells. From this tumor, they were able to isolate embryonal carcinoma cells, much resembling the ESCs from the developing embryo (Martin and Evans, 1974). However, as their name indicates, these cells are malignant and their use was thus less significant than the normal ESCs. At the time, the trick to obtain a successful isolation and growth in culture of ESCs was the use of a layer of mitotically inactive murine embryonic fibroblasts (termed feeder layer; MEFs). Interestingly, not all the mouse strains would be permissive to the isolation of ESCs, the 129 strain would be pretty efficient at *in vitro* stabilization whereas the CBA strain would require adaptations to the original protocol (Batlle-Morera, Smith and Nichols, 2008; Nichols and Smith, 2011). The aim to the isolation and culture of ESCs *in vitro* lies in the goal to be able to edit their genome to generate transgenic animals for a better understanding of the physiology.

The key cytokine that enables growth of ESCs is the leukemia inhibitory factor (LIF), activating the Janus kinase (JAK) – signal transducer and activation of transcription 3 (STAT3) pathway (Williams *et al.*, 1988). By itself however, this molecule still doesn't allow the stabilization of ESC lines from non-permissive strains such as CBA. This was achieved by an active repression of the ERK (Extracellular signal-Regulated Kinases) pathway, shown responsible for the early differentiation of ESCs (Burdon *et al.*, 1999). Ultimately, *in vitro* maintenance of ESCs from the pre-implantation blastocyst was broadly achieved adding a final inhibitor for the glycogen synthase kinase 3β (GSK3β), promoting protein biosynthesis and Wnt activation (Ying *et al.*, 2008). This culture condition with the two inhibitors and LIF is thus called 2iL. More recently, ESCs have also been isolated from more advanced, post-implantation epiblasts (Brons *et al.*, 2007; Tesar *et al.*, 2007). Compared to the 2iL ESCs from the pre-implantation blastocyst, these post-implantation ESCs require FGF2 (Fibroblast growth factor 2) and activin A to sustain their growth.

These ESCs retain their pluripotent features *in vitro*. Their phenotype, as described earlier, is characterized by the ability to form the three germ layers of the embryo and has to be associated with a capacity for self-renewal. To date, several assays have been developed to truly assess the pluripotent capacities of cells grown *in vitro*. First, the cells have to be able to differentiate upon withdrawal of cytokines cocktails maintaining them in the pluripotent state (the 2iL or FGF-activin cocktails), with or without the addition of specific differentiating molecules. This differentiation process can also be performed in a 3D setup, giving rise to embryoid bodies (EB), that recapitulate some aspects of early embryogenesis (Zeevaert *et al.*, 2020). Since there isn't any direct assessment of functionality of the product of this differentiation, this would be considered as a weak proof of pluripotency. The formation of teratomas, by injecting the cells subcutaneously in a mouse followed by the histological analysis of the resulting tumor would bring this functionality to the pluripotency assessment. However, this assay has caveats as incompletely reprogrammed cells could also form teratomas (Chan *et al.*, 2009). Instead, the formation of chimeras can overcome these issues. If true pluripotent cells are injected in a pre-implantation blastocyst, they should be able to

reenter development and contribute to the three germ layers of the chimeric embryo. This is further highlighted if the generated chimeras are able to give functional offspring, showing the capacity for germline transmission. But since the resident stem cells of the blastocyst could compensate for some defects in the tested cells, the complementation of a tetraploid blastocyst would be needed to certify the pluripotency. Indeed, the tetraploid cells cannot develop normally (Nagy *et al.*, 1993) so if a complemented tetraploid blastocyst manages to develop into a functional embryo, then it proves the true pluripotent features of the tested cells (De Los Angeles, 2019).

With this classification method on the developmental capacities of pluripotent stem cells, the exploration of the molecular mechanisms for the maintenance of these features is now possible. At the core of the pluripotency phenotype lies a group of three transcription factor: octamer-binding transcription factor 4 (Oct4), Nanog homeobox (Nanog) and SRY (sex determining region Y)-box 2 (Sox2). The coordinated action of these factors maintains the stemness phenotype in pluripotent cells, and it is thus unsurprising to find them as the factors used for cellular reprogramming (Takahashi and Yamanaka, 2006; Yu *et al.*, 2007). Their control over pluripotency is such that knockout (KO) of any of the three is embryo lethal, through a loss of self-renewal and pluripotency features (Ivanova *et al.*, 2006; Masui *et al.*, 2007). Since the isolation of embryonic stem cells from blastocysts (Evans and Kaufman, 1981), these factors have been under scrutiny for their role in the maintenance of pluripotency and it appears that their tight regulation is essential to maintain balance between self-renewal, stemness and differentiation.

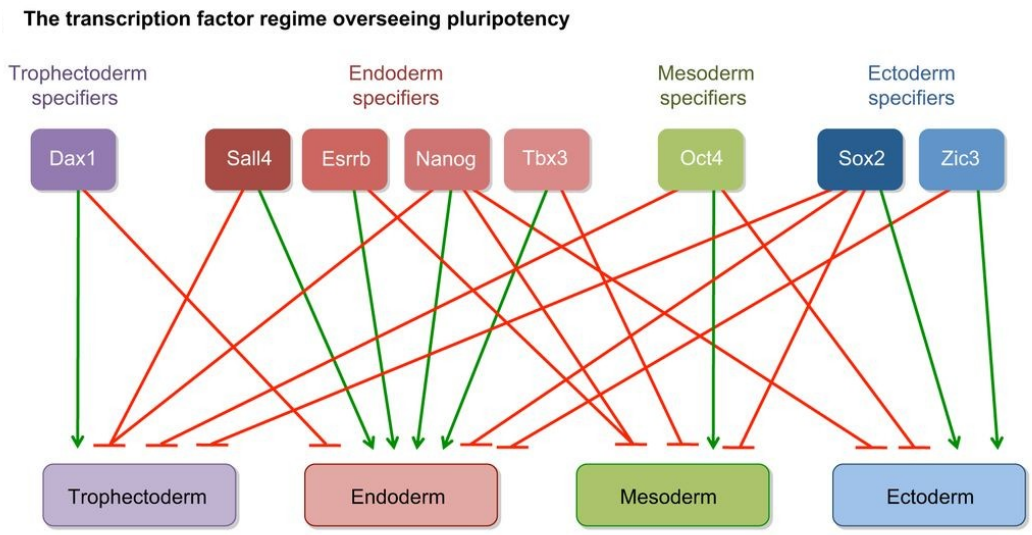


Figure 5 Regulation of lineage specification by key transcription factors. *The maintenance of pluripotency and the initiation of differentiation is tightly regulated by a titration of the various transcription factors (TFs). In a pluripotent state, most TFs balance each other for the repression and/or induction of the specific lineages. When this balance is lost and one TF takes over the others, the specification into a particular lineage is triggered. A green arrow reflects an activating function while a red arrow represents an inhibition. Dax1= DSS-AHC Critical Region On The X Chromosome Protein 1; Sall4= Spalt Like Transcription Factor 4; Esrrb= Estrogen Related Receptor Beta; Nanog= Nanog homeobox; Tbx3= T-Box Transcription Factor 3; Oct4= POU Class 5 Homeobox 1; Sox2= SRY-Box Transcription Factor 2; Zic3= Zic Family Member 3. From* (Loh, Lim and Ang, 2015)

As opposed to the two other transcription factors (TFs), Nanog is the only one that seems dispensable for maintenance of the stemness phenotype *in vitro,* while being indispensable *in vivo* (Chambers *et al.*, 2007). Depletion of Oct4 or Sox2 in ESCs leads to differentiation. Outside of their core regulation of pluripotency, these factors can have different outcomes on pluripotency. Oct4, uniformly expressed in all pluripotent cells, promotes mesoderm differentiation and represses ectoderm specification, while its downregulation results in a shift to the trophectoderm fate (Niwa, Miyazaki and Smith, 2000). On the other hand, Sox2 balances Oct4 as it represses mesoderm and promotes ectoderm fates (**Fig. 5**) (Thomson *et al.*, 2011). These two TF, along with Nanog, also work cooperatively and with positive feedback autoregulatory loops: Nanog regulates the expression of both Oct4 and Sox2, and the Oct4-Sox2 complex activates the transcription of Nanog (Boyer *et al.*, 2005). This complex can also bind the respective promoters, mostly activating their transcriptions, indeed it has been proposed that Oct4 could bind to the Nanog promoter and repress it (**Fig. 6**) (Loh *et al.*, 2006; Kalmar *et al.*, 2009). It is not surprising that many of their DNA-binding site are overlapping and also correlate with enhancers associated with terms such as "Stemness" or "Differentiation" (Boyer *et al.*, 2005; Loh *et al.*, 2006; Young, 2011). Interactions between these factors in regulation of stemness or differentiation is further explored in recent reviews by M. Li and J. C. Izpisua Belmonte (Li and Belmonte, 2017; Li and Izpisua Belmonte, 2018).

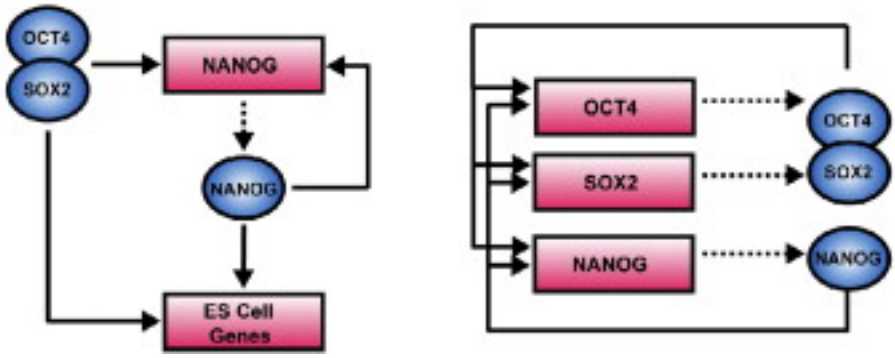


Figure 6 Autoregulatory loops between the three core transcription factors in pluripotency. *Cooperation of the three core TFs maintains the pluripotency state of ESCs, acting together to maintain their own expression. The transcription factors are represented by the blue circles and the promoter regions by the pink rectangles. A dotted line shows the product of the activation of a promoter while a solid arrow represents the binding to a particular promoter.* (Boyer *et al.*, 2005)

D. The naïve and primed pluripotent stem cells

Even though the core pluripotency network is conserved in pluripotent stem cells, many differences also separate different populations. As described in the previous section, ESCs can be derived from either the pre- and the post-implantation epiblasts. While both of them are able to self-renew, differentiate into the three germ layers and form teratomas when injected in mouse, respecting the fundamental principles of pluripotency, they do not match in their ability to form chimeras. Indeed, while pre-implantation ESCs successfully form chimeras when injected in pre-implantation blastocysts, post-implantation ESCs (also called EpiSCs) rarely do, except when injected in developmentally matched blastocysts (Brons *et al.*, 2007; Tesar *et al.*, 2007). This major developmental difference prompted the use of two specific names respectively for pre- or post-implantation derived ESCs: naïve ESCs and primed

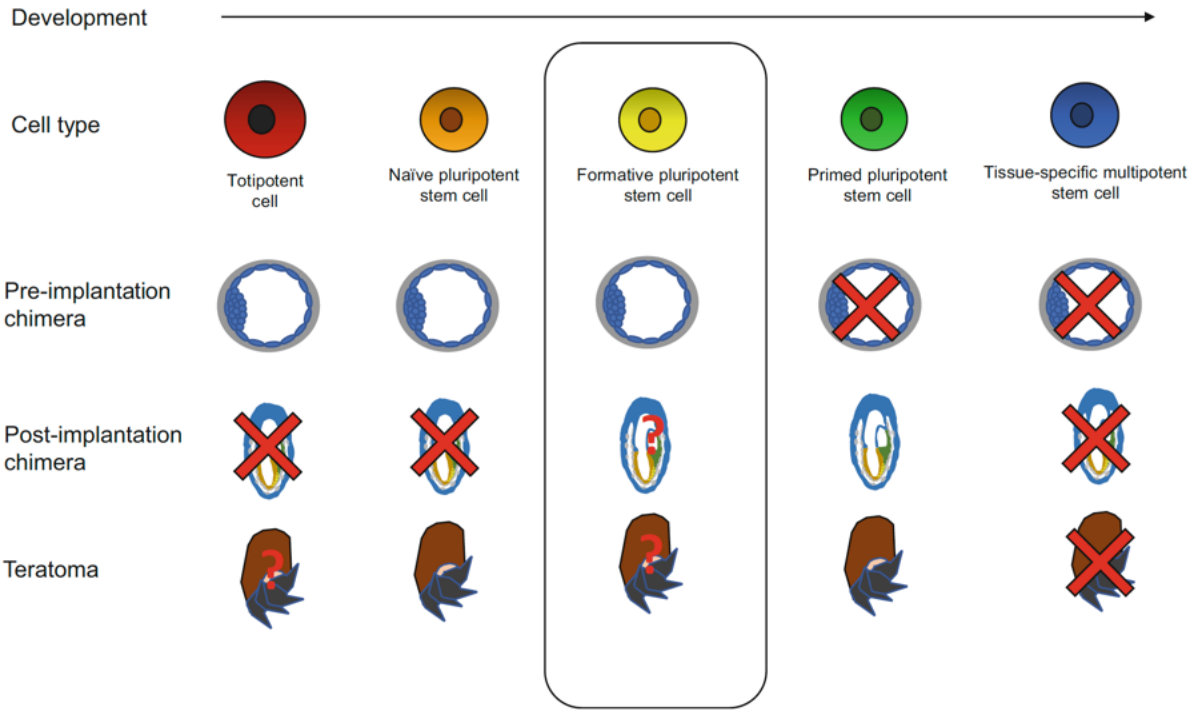


Figure 7 Schematic representation of the developmental potential of the different stages of stem cells. *Depending on the stem cell stage, different outcomes are expected for stem cells in the in vivo assays for pluripotency. While teratoma formation after subcutaneous injection is relatively easy, it lacks assessment of functionality. On the other hand, chimera formation by blastocyst injection represents better the pluripotency of embryos. However, in the case of primed PSCs, only a stage-matched blastocyst allows the development of a chimera.* (De Los Angeles, 2019)

ESCs (**Fig. 7**) (Nichols and Smith, 2009). Interestingly, the isolated human ESC (hESCs) from pre-implantation blastocysts share many similarities with the primed mESCs, such as the morphology or the signaling dependency (FGF2 and activin A), as detailed hereafter (Thomson *et al.*, 1998). Retrospectively, even if these cells were isolated from the pre-implantation blastocyst, they were actually stabilized in the primed state. It took about 15 years to stabilize hESCs in the naïve state, thanks to modulation of the growth culture conditions (Gafni *et al.*, 2013a; Takashima *et al.*, 2014; Theunissen *et al.*, 2014; Ware *et al.*, 2014). At the molecular level, numerous differences exist between the two states, as described in the following paragraphs.

1. Signaling dependency

Over time, the culture conditions of ESCs from either mouse or human changed, for either a better preservation from spontaneous differentiation or for better self-renewal capacities.

a) mESC

The pioneering experiments in 1981 were able to stabilize the first mESCs by using a feeder layer of MEFs and serum products (Evans and Kaufman, 1981; Martin, 1981). The addition of LIF to the growth media a few years later further enhanced the stability of mESCs and allowed their growth without the need of MEFs (Williams *et al.*, 1988). Together, this culture condition is now referenced as the Serum/LIF (SL) medium and is able to maintain

$Oct4^+/Nanog^+$ cells in culture for many passages. The pluripotency of the mESC was shown by the ability to form chimeras in pre-implantation blastocysts.

However, subsequent efforts were made in order to switch to serum-free conditions, the serum being highly variable from batch to batch. This was achieved using BMP2 or BMP4 (Bone morphogenic proteins 2/4) in conjunction with LIF (Ying *et al.*, 2003). Used alone, BMP4 induces differentiation into the mesoderm lineage (Johansson and Wiles, 1995) but together with LIF it enables feeder-free and serum-free mESC growth (Ying *et al.*, 2003). Then, since reports showed that activation of ERK1/2 signaling was responsible for hasty differentiation, inhibitors of the pathway, mostly via MEK (mitogen-activated protein kinase kinase) inhibition, were introduced (Burdon *et al.*, 1999). Today, mESCs are considered at a "ground state" when grown in a 2i/LIF (2iL) medium, using GSK3β inhibition in addition to the MEK inhibition. GSK3β inhibition was shown to maintain self-renewal of mESCs and hESCs through Wnt activation (Sato *et al.*, 2004). The name ground state comes from the fact that mESCs grown in those conditions are less heterogenous and better resist to spontaneous differentiation than the SL counterparts (Ying *et al.*, 2008). To date, most of the research is performed with these two main growth conditions, SL and 2iL, but combinations of the serum, the LIF and the inhibitors have been used at some point.

Primed mESCs, or EpiSCs, require completely different signals for *in vitro* maintenance. Instead of a repression of the MAPK pathway through MEK inhibition, they require FGF2 signals for its activation, coupled with Transforming growth factor β (TGFβ) and/or activin A pathway activation (Brons *et al.*, 2007; Tesar *et al.*, 2007). Alternative growth conditions have also been described for EpiSC maintenance, for example using the GSK3β inhibitor and a tankyrase inhibitor or FGF2 with the tankyrase inhibitor but these conditions have been shown to retain higher levels of naïve markers expression (Kim *et al.*, 2013). The inhibition of tankyrase leads to a stabilization of AXIN and a subsequent reduction in Wnt signaling by promoting the β-Catenin destruction complex (Huang *et al.*, 2009). To date, the FGF2 and activin A are still the canonical signals required for derivation of primed mESCs (Kinoshita *et al.*, 2020).

b) hESC

For years, the isolation of pre-implantation epiblast stem cells from human blastocyst was only possible in EpiSC-like conditions (activin A and FGF2) since LIF wasn't able to support their growth (Thomson *et al.*, 1998). Sometimes thought to be due to inter-species differences, the quest for the true naïve hESC state didn't stop. All within a year, multiple teams and subsequent protocols have been described to reach this naïve stage (Chan *et al.*, 2013; Gafni *et al.*, 2013a; Takashima *et al.*, 2014; Theunissen *et al.*, 2014; Ware *et al.*, 2014). The strategies used by those research groups are different and some even use the expression of transgenes. Indeed, the overexpression of either Oct4 and KLF4 or KLF2 and KLF4 was able to revert primed hESCs to the naïve state if cultured in 2iL conditions (Hanna *et al.*, 2010). However, finding transgene-free naïve hESCs would shed light on the mechanisms governing these pluripotent cells in human, allowing a comparison with rodents. It is a few years later that such cocktails were developed. First, Gafni and coworkers used a cocktail of FGF, LIF and TGF-β, with inhibitors of MEK, GSK3β, p38 and JNK (c-Jun N-terminal kinase) (Gafni *et al.*, 2013a). With this cocktail of factors, named Naive Human Stem cell Medium (NHSM), they

were able to bring hESCs to a state that allows chimera formation in mouse blastocysts for the first time. In parallel, others have used an almost as complex cytokine and small molecule cocktail as in the team of R. Jaenisch, using LIF and activin A together with inhibitors of MEK, GSK3β, BRAF (serine-threonine kinase B-Raf), SRC (Proto-oncogene tyrosine-protein kinase) and ROCK (Rho-associated, coiled-coil containing protein kinase); a cocktail named 5iLA (Theunissen *et al.*, 2014). Occasionally, FGF2 was included for a 5iLAF formulation. This transgene-free cocktail was able to revert primed ESC into a naïve state (state described in section II.D.2) and also allowed the isolation of ESCs directly from blastocysts. Alternatively a pan-PKC (protein kinase C) inhibitor (targeting PKCα, PKCβ, PKCγ and PKCδ) or a pan-HDAC (Histone deacetylase) inhibitor were used (Takashima *et al.*, 2014; Ware *et al.*, 2014).

2. Molecular features

The characterization of the different stages and/or culture conditions of ESCs relies on the analysis of molecular markers such as gene expression, transcription factor localization and/or epigenetic landscape. Since these cell stages were first uncovered in mice, we will focus here on those. These differences are compiled in **Figure 8**.

a) Transcription factor circuitry

Even though all ESCs express the pluripotent marker Oct4, one of the most crucial difference between naïve and primed cells is the regulation of expression of this transcription factor. The promoter of the *Pou5f1* gene, coding for Oct4, contains two distinct enhancer-binding regions, one being proximal (PE) and the other one located further away from the transcription start site (TSS) called distal enhancer (DE). It has been shown that while naïve mESCs use preferentially the DE, primed ones use the PE (Tesar *et al.*, 2007; Han *et al.*, 2010; Choi *et al.*, 2016). Hanna and his team used this specificity to validate the naïve state of their hESCs using luciferase constructs (Hanna *et al.*, 2010).

The transcription factor circuitry found in naïve and primed cells also differ drastically. In mouse, a robust expression of *Klf2/4*, *Tfcp2l1* (Transcription factor CP2-like 1), *Esrrb* (estrogen Related Receptor β) or even *Tbx3* (T-box 3) is observed and their downregulation is required for proper transition to the primed stage (Dunn *et al.*, 2014). The expression of these factors is reminiscent of the pathways required for stabilization of the naïve state: *Esrrb* and *Tbx3* are expressed downstream of Wnt activation (Waghray *et al.*, 2015; Huang *et al.*, 2018) while *Klf4* and *Tfcp21* are regulated by the LIF-STAT3 axis (Chen *et al.*, 2015). At the other end of the spectrum, EpiSCs cells express higher levels of *Otx2* (Orthodenticle homeobox 2), *Zic2* (Zic family member 2), *Fgf-5* and *-15* (fibroblast growth factor 5/15) (**Fig. 8**) (De Los Angeles, 2019; Kinoshita *et al.*, 2020).

The subcellular localization of TFE3 (Transcription Factor Binding To IGHM Enhancer 3) transcription factor is also shown to be regulated between the two stages: while observed almost exclusively in the nucleus in naïve cells, TFE3 is kept in the cytoplasm of primed cells (Betschinger *et al.*, 2013; Mathieu *et al.*, 2019). Again, this reinforce the naïve-specific circuitry as many of the genes targeted by TFE3 are involved in the activation of the Wnt pathway, critical for the maintenance of the naïve state (Xu *et al.*, 2016; Mathieu *et al.*, 2019). In turn,

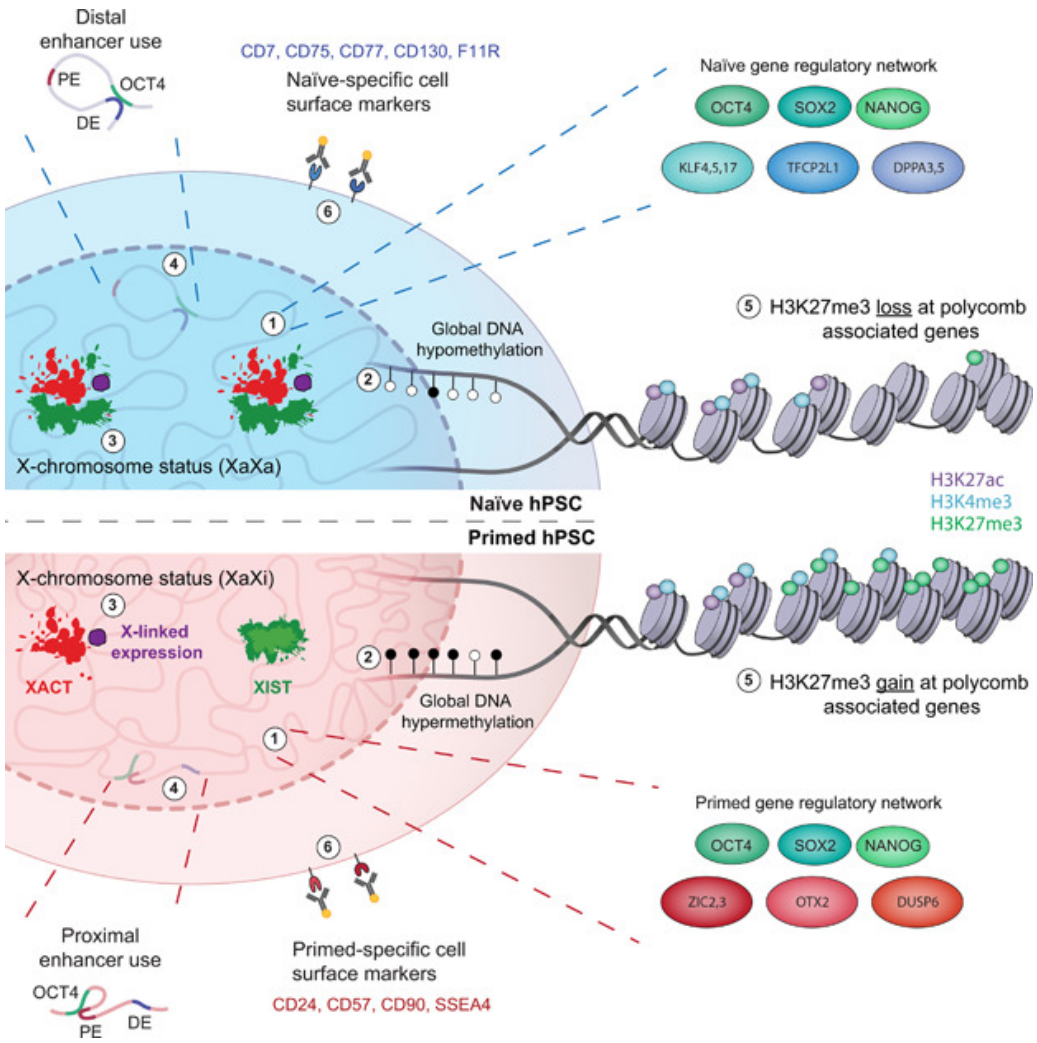


Figure 8 Hallmarks of naïve and primed hESC features. *1) While both states express the core pluripotency factors (OCT4, SOX2, NANOG), naïve cells express specific markers (KLF-4;-5;-17, TFCP2L1 or DDPA3 and 5) while primed ESCs express a another subset of genes as well (ZIC2 and 3, OTX2 or DUSP6). (2) Naïve cells show a global hypomethylated DNA profile compared to primed and (3) two active X chromosomes in female. (4) Two enhancers for OCT4 expression have been identified and naïve ESCs preferentially use the distal enhancer (DE) while primed cells use the proximal one (PE). (5) A net increase of repressive H3K27me3 marks at polycomb associated genes is observed in primed cells. (6) Specific surface marker pattern characterizes naïve and primed hESCs.* (Collier and Rugg-Gunn, 2018)

the regulation of TFE3 subcellular localization is known to be regulated by the activity of the mTOR complex (mammalian target of rapamycin) (Martina *et al.*, 2014; Mathieu *et al.*, 2019).

b) Epigenetic regulation

Another important feature in the distinction between naïve and primed ESCs is the activation state of the second X chromosome in female ESCs. Indeed, in order to compensate for the presence of two X chromosomes, there is a random inactivation of one of the chromosomes, to restore the correct dosage of X-linked gene expression (Disteche, 2016). This inactivation occurs early in the development, so that naïve cells still carry two active X (XaXa), while primed present one inactive chromosome (XaXi) (**Fig. 8**) (Silva *et al.*, 2009).

Aside from the X chromosome pattern of activation, the epigenetic landscape of naïve and primed ESCs is also dramatically differently organized. On one hand, the naïve ESCs show

a global DNA hypomethylation status compared to EpiSCs (Hackett *et al.*, 2013). This difference was also highlighted when comparing cells grown in 2iL or in SL conditions, the later presenting a higher methylation state, although still lower than EpiSCs. This underlines the more permissive state of the SL growth condition, with cells presenting an heterogenous pattern of gene expression (Kumar *et al.*, 2014). This difference in methylation patterns is complex but includes two main features : a lower expression of the DNA methylation machinery (DNA methyltransferases 3a and 3b; DNMT3a/b) but also a defective recruitment machinery, with decreased protein levels of UHRF1 (Ubiquitin-like, with PHD and RING finger domains 1) in naïve cells (Ficz *et al.*, 2013; von Meyenn *et al.*, 2016). Differences in the parental imprinting of gene methylation is also observed between the two stages: it was shown that the reprogramming of primed ESCs to naïve or the long term culture of naïve cells induce a loss of methylation in imprinted regions leading sometimes to biallelic gene expression, that is not reacquired when transitioned back to the primed stage (Pastor *et al.*, 2016; Liu *et al.*, 2017).

On the other hand, dramatic changes in the landscape of histones modifications have also been noted. These changes have been extensively reviewed in (Gö Kbuget and Blelloch, 2019). An example of epigenetic marks that is modified during the progression of development is the trimethylation of the lysine 27 of the histone 3 (H3K27me3), increasing dramatically in primed cells (**Fig. 8**) (Marks *et al.*, 2012; Sperber *et al.*, 2015). This repressive mark is shown to regulate for example the expression of Wnt pathway genes such as *WNT5A/5B/8A, AXIN2, LEF1 (*Lymphoid Enhancer Binding Factor 1*)* (Sperber *et al.*, 2015). This is in accordance with the known role of Wnt pathway in the maintenance of the naïve state (Sperber *et al.*, 2015; Xu *et al.*, 2016). Interestingly, it has been shown that the presence of the polycomb repressive complex 2 (PCR2), the complex actively methylating the H3K27, is not required in ground state mESCs (Chamberlain, Yee and Magnuson, 2008). This has been also proved for the human naïve cells, offering another criteria for the characterization of naïve state (Moody *et al.*, 2017). However, although PRC2 is dispensable for the maintenance of the pluripotent stage, PCR2-deficient cells present a compromised capacity for further differentiation highlighting the need for epigenetic rewiring in the differentiation processes (Landeira *et al.*, 2010; Walker *et al.*, 2010).

c) Metabolic activity

Finally, if we consider glycolytic activity, oxidative phosphorylation (OXPHOS) usage and fatty-acid β-oxidation, naïve and primed ESCs present a completely different profile. These differences are summarized in **figure 9** and reviewed over the years (Wu, Ocampo and Belmonte, 2016; Mathieu and Ruohola-Baker, 2017; Tsogtbaatar *et al.*, 2020).

The naïve-to-primed transition is marked by a dramatic switch between a naïve bivalent (glycolytic and oxidative) energetic state to a primed almost exclusively glycolytic state (Zhou *et al.*, 2012; Sperber *et al.*, 2015). This metabolic difference can be the driver of the transition since the activation of OXPHOS enhances the reprogramming of primed murine PSC into naïve state, while inhibition of the OXPHOS activity through the activity of Lin28 (lin-28 homolog A) or HIF1α (Hypoxia-inducible factor 1-α; described below) pushes naïve cells forward (Sperber *et al.*, 2015; Zhang *et al.*, 2016; Sone *et al.*, 2017). As the metabolic hub of the cell, the mitochondrion is a key player in this transition and, interestingly, its morphology

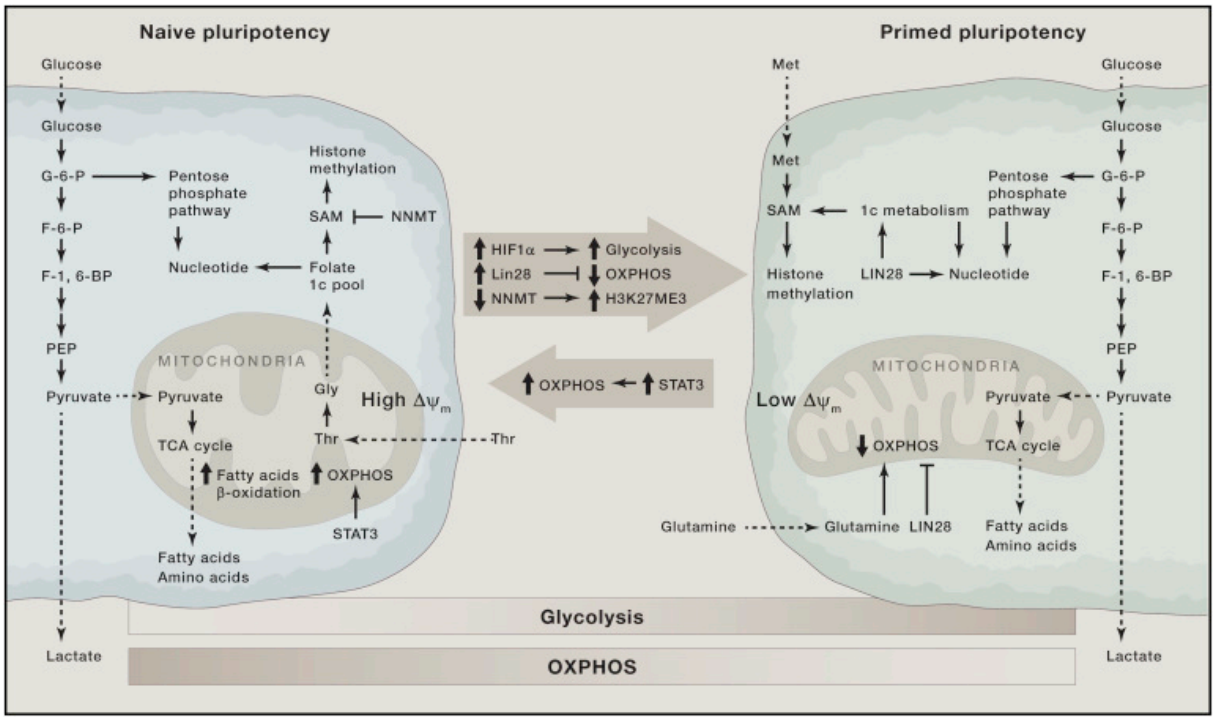


Figure 9 Distinct metabolic features of naïve and primed ESCs. *Compared to primed ESCs, naïve ESCs display a higher mitochondrial oxidative metabolic activity than primed cells, highlighted by a higher mitochondrial membrane potential (Δψm), increase in OXPHOS activity and fatty acid β-oxidation. On the other hand, NNMT reduces the SAM pool available for histone and DNA methylation. Drivers of the metabolic remodeling such as HIF1α or Lin28 are able to precipitate the switch to the primed state, along with the loss of NNMT unlocking SA for epigenetic remodeling. Increasing the OXPHOS activity in primed cells, through STAT3 for example, is able to reprogram primed ESC into naïve.* (Wu, Ocampo and Belmonte, 2016)

changes dramatically during this conversion. Albeit using their mitochondria and their electron transport chain (ETC) complexes at a higher rate, naïve cells do not possess mature mitochondria. It is only during the transition that this organelle goes from a round shaped with sparse and irregular cristae to more elongated mitochondria with well-defined transverse cristae. All these changes were described as one of the major hallmarks of the implantation of the embryo in human (Sperber *et al.*, 2015), mice (Zhou *et al.*, 2012) or even dogs (Tobias *et al.*, 2018). This morphological and metabolic remodeling occurring during the naïve-to-primed transition is reminiscent to the remodeling of the cell's metabolism observed during stem cell differentiation, during which the cell re-acquires an OXPHOS-based metabolism or conversely switches to a glycolytic based metabolism during reprogramming (reviewed in (Zhang *et al.*, 2012; Wanet *et al.*, 2015)). The same metabolic switch is also observed in cancer cells, an effect described as the Warburg effect (Warburg, Wind and Negelein, 1927).

Crucial to the metabolic switch observed when transitioning to the primed stage is the activation of the HIF1α response. Together with the beta subunit (HIF1β; *ARNT*), HIF1α forms a heterodimer functioning as a transcription factor for genes with a hypoxia-responsive element (HRE) in the promoter region. The regulation of the activity of the pathway is dependent on the O_2 concentration (**Fig. 10**). Indeed, in normoxic conditions, prolyl-hydroxylase domain (PHDs) proteins catalyze the hydroxylation of HIF1α on prolines 402 and 564, in a process dependent on 2-oxoglutarate and O_2. These post-translational modifications are then recognized by von Hippel-Lindau (pVHL) protein, which is part of the E3 ubiquitin ligase complex. This triggers the poly-ubiquitination of the HIF1α protein and its further degradation by the proteasome. HIF1α and 1β are constitutively expressed and thus

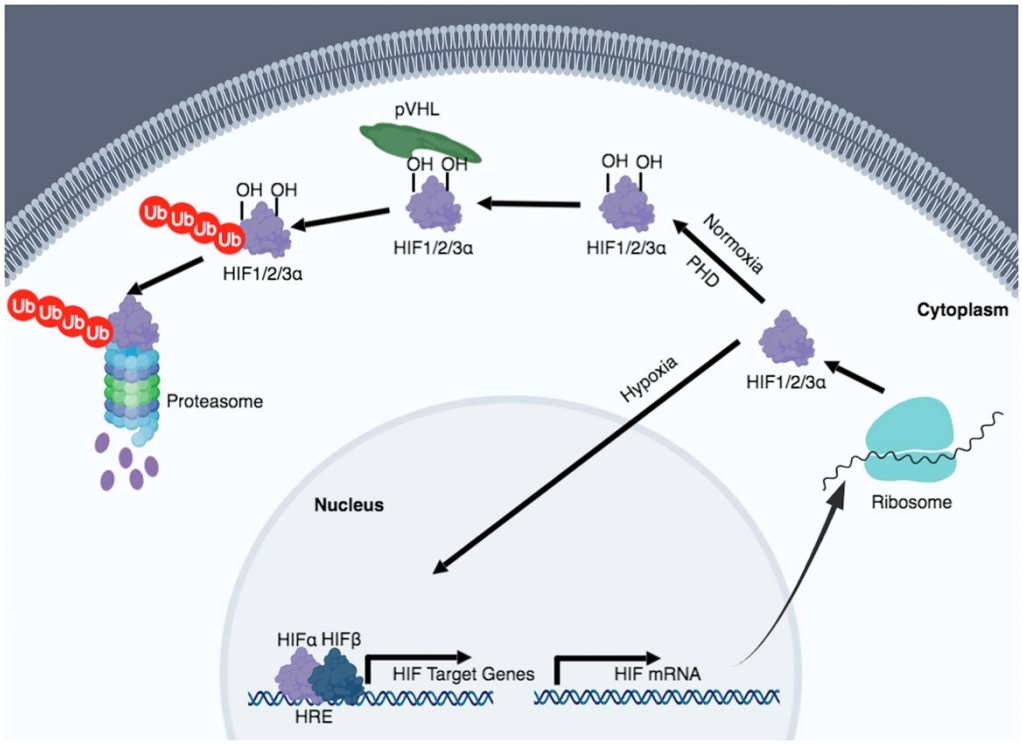


***Figure 10 Schematic representation of the cellular response to hypoxia through the hypoxia-inducible factors (HIFs)**. In normoxia, the HIFα subunits are recognized by prolyl-hydroxylase domain proteins (PHD) and hydroxylated thanks to O_2. These post-translational modifications are then recognized by the von Hippel-Lindau protein (pVHL) and targeted to the proteasomal degradation by a poly-ubiquitination. In hypoxia, the inactivation of PHDs allows the translocation of the HIFα subunits to the nucleus where HIFα acts as transcription factor as a heterodimer with the HIFβ subunit (also called ARNT)* (Lee *et al.*, 2019)

continuously degraded under a high pO_2. When the concentration in oxygen drops, a condition called hypoxia, the activity of the PHDs drops as well, promoting the stabilization of HIF1α that further translocates into the nucleus and exert its activity after dimerization with HIF1β. This mode of activation will trigger the cellular response to hypoxia. Indeed, HIF1α transcriptional activity increases the expression of the pyruvate dehydrogenase (PDH) kinase (PDK1), a repressor of the PDH complex activity, thus efficiently reducing the entry of pyruvate in the TCA cycle and thus reducing the activity of the oxygen-consuming ETC (Kim *et al.*, 2006). In parallel, HIF1α also triggers the transcription of genes involved in glycolysis, efficiently regulating the metabolic switch observed in the transition (Zhou *et al.*, 2012; Sperber *et al.*, 2015). By efficiently switching the metabolism from bivalent in naïve to mostly glycolytic in primed, HIF1α stabilization pushes thus ESCs toward the primed stage (Zhou *et al.*, 2012). On the other hand, HIF1α knock-out naïve ESCs fail to properly transition (Sperber *et al.*, 2015). Interestingly, this metabolic switch is also observed during the reprogramming of somatic cells, with bivalent metabolism, to iPSCs, mostly glycolytic. In a similar manner, HIF1α has also been shown to control the switch and enhance the reprogramming efficiency if overexpressed (Mathieu *et al.*, 2014).

In addition to OXPHOS, the mitochondria participate to other metabolic pathways such as the TCA, the fatty acid beta-oxidation, the heme synthesis, the degradation and synthesis of several amino acids, the folate pathway, etc. Several of these non-OXPHOS pathways are also involved during the naïve-to-primed transition. For instance, the beta oxidation of fatty acids is strongly reduced during the naïve-to-primed transition of human and murine cells, partly due to CPT-1 (carnitine palmitoyltransferase 1) inhibition-mediated blockage of fatty

acid import to the mitochondrial matrix, by a CPT1A downregulation by micro RNAs (miRNAs) and repressive chromatin marks, as shown with the oxygen consumption rate when the cells are presented with palmitate (C16:0) (Sperber *et al.*, 2015). These major differences in metabolic function were only shown with either glucose or palmitate (Sperber *et al.*, 2015) and so far primed cells do not show an increase of mitochondrial respiration when presented with these substrates. The energy source of these cells could however be non-classical substrates such as amino acids catabolism or other types of fatty acids. For example, exogenous glutamine was shown to support a robust ESC self-renewal and stemness maintenance in both mESCs and hESCs (Carey *et al.*, 2015; Marsboom *et al.*, 2016; Tohyama *et al.*, 2016). Furthermore, the mitochondrial function is also essential to provide the metabolites such as α-ketoglutarate (Carey *et al.*, 2015; TeSlaa *et al.*, 2016) or even S-adenosyl methionine (SAM) (Sperber *et al.*, 2015) that are necessary for epigenetics modifying enzymes (reviewed in (Matilainen, Quirós and Auwerx, 2017)) also playing a key role during the transition. These critical mitochondrial functions, independent from the OXPHOS activity, might be the reason why a mitochondrial maturation is observed during the conversion to the primed stage.

Interestingly, this metabolic switch seems to be intertwined with the changes in the epigenetic landscape. This is exemplified by the crosstalk between the NNMT enzyme (Nicotinamide N-methyl transferase) and the repressive methyl marks on histones. NNMT, whose expression is regulated by the LIF-STAT3 pathway, is highly expressed in naïve cells. This enzyme consumes SAM, a precursor for histone and DNA methyl modifications, acting thus as a "methyl-sink" and reducing the abundance of SAM available for these epigenetic modifications (Sperber *et al.*, 2015). This allows for example the maintenance of an activated Wnt pathway, necessary for naïve cell maintenance. During the transition to the primed stage, the LIF-STAT3 signal is lost, reducing the NNMT expression, then allowing the deposition of these repressive marks. Among the genes targeted by these modifications is EGLN1 (Egl nine homolog 1), a PHD protein involved in HIF1α degradation. The reduced expression of EGLN1 favors HIF1 transcriptional activity, hence efficiently repressing the oxidative metabolic activity.

3. Naïve-to-primed ESC transition

As much as the molecular features of naïve and primed ESCs have been extensively studied, the requirements for the transition from one stage to the other have not been yet fully elucidated and so far, a few studies have highlighted some mechanisms.

The first examples worth mentioning are focused on the metabolic switch described in the previous section. HIF1α, a master regulator in this metabolic rewiring, is crucial in the establishment of this transition and its forced expression is sufficient to induce the transition to the primed stage (Zhou *et al.*, 2012). Conversely, genetic ablation of HIF1α prevents this transition (Sperber *et al.*, 2015). In the same way, the RNA-binding proteins LIN28A and LIN28B are able to push naïve cells to the primed stage or to enhance the reprogramming of fibroblasts, in both cases the effect was mediated by a control of the metabolic switch (Zhang *et al.*, 2016).

Then, as the metabolic hub of the cell, the mitochondria also has its role in the transition, although being seemingly inactive in terms of cellular respiration in primed ESCs. The team of J. Hanna showed that the mitochondrial fusion can be one of the driving forces leading to the conversion of pre-implantation mESC to their post-implantation counterpart. By invalidating MTCH2 (mitochondrial carrier homolog 2), a regulator of the mitochondrial fusion, they show a reduced glutamine utilization that they correlate with impaired acetylated histone epigenetics marks, both of these contributing to the transition from naïve to primed state. This phenotype was similar to the overexpression of MFN2 (mitofusin 2), forcing the mitochondrial elongation and promoting the naïve-to-primed transition (Bahat *et al.*, 2018). Accordingly, a more recent study was able to show that an excess of fission, through the overexpression of the mitochondrial fission factor Mff, leads to a reduced differentiation capacity of pluripotent stem cells (Zhong *et al.*, 2019). Taken together these two recent studies emphasize the role of mitochondrial dynamics as one of the drivers during embryonic development, but also differentiation.

To unveil the regulators of this transition, the major strategy used is the genetic screening. A recent detailed review gathered the different strategies used to date, for the naïve-to-primed transition, but also to identify new regulators of pluripotency as well (Li, Rosen and Huangfu, 2020). Since the focus of this work is the naïve-to-primed transition, only screens focusing on this process will be detailed.

In mESC, the strategy used rely on the use of cells with a *Rex1-GFP* construct. *Rex1* (Reduced Expression 1; also known as ZFP-42) is a gene whose expression is restricted to the naïve pluripotent state. Upon the initiation of transition, its expression is rapidly downregulated, allowing to follow the progression of the exit of the naïve state in a high-throughput fashion (Chambers *et al.*, 2007). Indeed, the whole genome CRISPR guide library allowed to identify the genes responsible for the maintenance of GFP expression thus preventing the exit of the naïve state (Li *et al.*, 2018).

In human, a recent paper exploited the metabolic differences between naïve and primed cells to perform the functional screen (Mathieu *et al.*, 2019). Indeed, as described earlier, primed cells rely on SAM, comparatively to naïve cells, to regulate their epigenetic landscape, mainly through an increase in repressive histone marks (Sperber *et al.*, 2015; Mathieu *et al.*, 2019). As a consequence, primed cells exhibit an increased sensitivity to drugs such as methotrexate and acetaldehyde, both inhibiting the methionine synthase, thereby depleting the pool of SAM available. With a combination of these two molecules, and a CRISPR-Cas9 library (GecKo), only cells acquiring a mutation preventing them to transition to the primed stage would be spared upon selection. By comparing the enriched sgRNAs after selection to those before, a list of genes required for the transition, or the exit of the naïve state, was established (Mathieu *et al.*, 2019). Interestingly, these two papers highlighted the mTOR pathway as critical, an idea reinforced by the efficient reprogramming of primed hESCs into naïve with transient mTOR inhibition (Hu *et al.*, 2020). As a direct regulator of mTORC activity, folliculin (FLCN) was also found necessary for the transition in both mESCs and hESCs (Betschinger *et al.*, 2013; Mathieu *et al.*, 2019). The authors of the studies show that FLCN is able to regulates the subcellular localization of TFE3, whose regulation is a hallmark of the transition.

E. Alternative states

For years, this distinction between naïve pre-implantation and primed post-implantation epiblast cells governed the identification of ESCs. However, the segregation of ESCs in two distinct categories is simplistic. The same way that development proceeds, ESCs are actually found in more of a continuum with a progression from one stage to the other, with the existence of intermediate stages as first proposed by Austin Smith (Smith, 2017).

1. Formative pluripotency

In view of the developmental timeline between naïve and primed, the formative pluripotent state was proposed as a stage related to the peri-implantation epiblast (Smith, 2017). Indeed, when epiblast cells are derived from the embryo using the conventional FGF2 and activin A culture condition, the cultured cells always converge to a mid-gastrula-like stage (about E7). One of the features of these cells, is that they are refractory to give rise to primordial germ cells (PGCs) whereas naïve cells transitioned for one or two days into a transient epiblast-like stem cell (EpiLC) can (Hayashi *et al.*, 2011; Nakaki *et al.*, 2013). This capacity for PGC induction is a feature of the early post-implantation epiblast *in vivo*, suggesting that there could be a transient "formative" stage arising during the transition to EpiSC. Very recently, and after long efforts, a formative state was finally stabilized *in-vitro* for mouse and human ESCs (Kinoshita *et al.*, 2020). This stage was defined in mouse as the population of ESCs that, when pushed to EpiSC, first loses the expression of the naïve marker REX1 as assessed by the extinction of the REX1-GFP signal (Smith, 2017). Along with REX1, these cells have also lost the expression of the other naïve markers (*Tfcp2l1, Esrrb, Klfs*,…), have acquired the expression of the early post-implantation markers *Otx2* and *Fgf5*, but do not display the expression of the more definitive primed markers such as *Foxa2* (Forkhead box protein A2) or *Cer1* (Cerberus protein) (Smith, 2017). In this population, capable of PGC induction, the drop in oxygen consumption is already observed and the DNA methylation levels are intermediate to the EpiSCs (Kinoshita *et al.*, 2020). The key for their *in-vitro* stabilization was a combination of low concentration in activin A and Wnt, together with retinoic acid receptor (RAR) inhibition, efficiently stabilizing a formative epiblast stem cell population for many passages. The intermediate phenotype observed is also emphasized by the ability of formative cells to contribute to chimeras when injected to pre-implantation blastocyst (**Fig. 7**).

2. Poised state

Closer to the naïve stage, a particular "poised" stage was also stabilized (P. Du *et al.*, 2018). Compared to the formative state, the population retains the expression of the naïve markers but fails to upregulate the primed ones. To date, this stage has never been stabilized with a non-transgenic approach. Indeed, to allow the poised cell stabilization, the authors relied on the overexpression of ISY1 (Pre-mRNA-splicing factor ISY1 homolog) a regulator of microRNA biogenesis necessary for the cleavage of pri-miRNA into progenitor pro-miRNAs. This protein is transiently increased in expression during the early stages of the naïve-to-primed ESC transition, just before the loss of *Nanog* expression. The poised stage would thus represent an intermediate between naïve and formative cells. They further showed that this intermediate is required for proper transition to the primed stage as cells with ISY1 depletion

fail to transition. The role of ISY1 in this process was attributed to a subset of miRNAs. Interestingly, the capacity for these cells to form chimeras in pre-implantation blastocysts highlights an early post-implantation phenotype but, unlike formative ESCs, poised cells have yet to acquire the increase in DNA methylation (P. Du *et al.*, 2018).

3. 2C-like cells

The culture and maintenance of naïve cells in either 2iL ("ground") or in SL conditions is reported to present a certain level of cellular heterogeneity (Ying *et al.*, 2008; Canham *et al.*, 2010; Kumar *et al.*, 2014). This is especially the case in SL, where some cells have been shown to oscillate between a state close to the epiblast while others resemble the primitive endoderm lineage. This concept of heterogenous population is further supported by a subpopulation of cells that re-acquire features of the 2-cell embryo (2C). As a reminder, the 2 cell embryo is the stage when the genome of the zygote is activated (ZGA), through complex mechanisms of DNA decompaction, epigenetic remodeling and maternal transcripts destruction (reviewed in (Schulz and Harrison, 2019)). These decompaction events are also accompanied with the expression of a wide variety of retrotransposons, like endogenous retroviruses (ERVs; MuERV-L), long- or short-interspaced nuclear elements (LINEs and SINEs) (**Fig. 11**) (Macfarlan *et al.*, 2012; Percharde *et al.*, 2018; He *et al.*, 2019). The expression of these transposable elements is concomitant to the expression of a subset of 2C-specific genes such as the *Zscan4* family, *Tcstv1* (2-cell-stage, variable group, member 1), Dux, Dub1 (Ubiquitin carboxyl-terminal hydrolase) or *Eif1α* (Eukaryotic translation initiation factor 1A) (Macfarlan *et al.*, 2012; De Iaco *et al.*, 2017; Percharde *et al.*, 2018). This subpopulation of mESCs that expresses these factors has been named 2C-like cells (2CLCs) and was shown to represent a very low percentage of the cell population (<1%) (Macfarlan *et al.*, 2012;

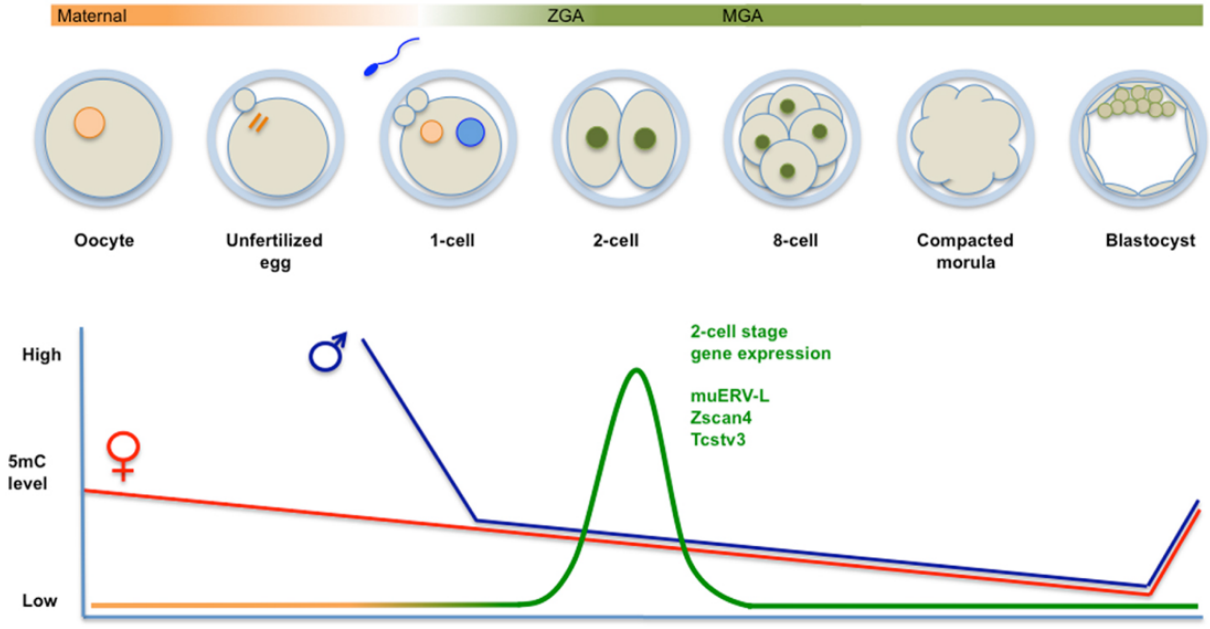


Figure 11 Schematic representation of the expression of retrovirus or 2C-like specific genes in relation with the embryo developmental stage in mouse. *In the bottom panel, representation of the time course of the female and male genome methylation levels, decreasing over time until the blastocyst implantation. Meanwhile, at the 2-cell stage, the expression of specific genes as well as some transposable elements is observed. ZGA= zygote genome activation; MGA= mid pre-implantation gene activation; muERV-L= murine endogenous retrovirus-like; Zscan4= Zinc finger and SCAN domain containing 4; TCSTV3= Two cell stage, variable group member 3.* (Schoorlemmer *et al.*, 2014)

Eckersley-Maslin *et al.*, 2019; Rodriguez-Terrones *et al.*, 2020). Interestingly, these cells are able to colonize the extraembryonic tissues when implanted in chimeras, demonstrating their "totipotent-like" capacity (Macfarlan *et al.*, 2012).

The apparition and maintenance of this 2CLC population is under intensive research and currently only partly understood. First, these cells are in a state allowing the telomere elongation, for example with a role of the *Zscan4* gene family and the expression of *Tcstv1* (Dan *et al.*, 2017). Then, crucial to this 2CLC population, is the transcription factor DUX. This 2C stage-specific TF is one of the earliest transcripts activated during the ZGA and its transcriptional activity triggers, among others, the transcription of *Zscan4.* Since it is a TF, many pathways identified for the acquisition of the 2CLC phenotype involve its activation (De Iaco *et al.*, 2017; Percharde *et al.*, 2018; Eckersley-Maslin *et al.*, 2019). The pathways triggering the acquisition of 2CLCs markers include, among others, the multifunctional protein KAP1 (also known as TRIM28), the chromatin assembly factor 1 (CAF-1) or PIAS4 (small ubiquitin-like modifier (Sumo) E3 ligase protein inhibitor of activated STAT 4) (Ishiuchi *et al.*, 2015; Percharde *et al.*, 2018; Yan *et al.*, 2019). KAP-1 and CAF-1 are negative regulators of *Dux* expression so their genetic ablation leads to a Dux-mediated 2CLCs phenotype acquisition. On

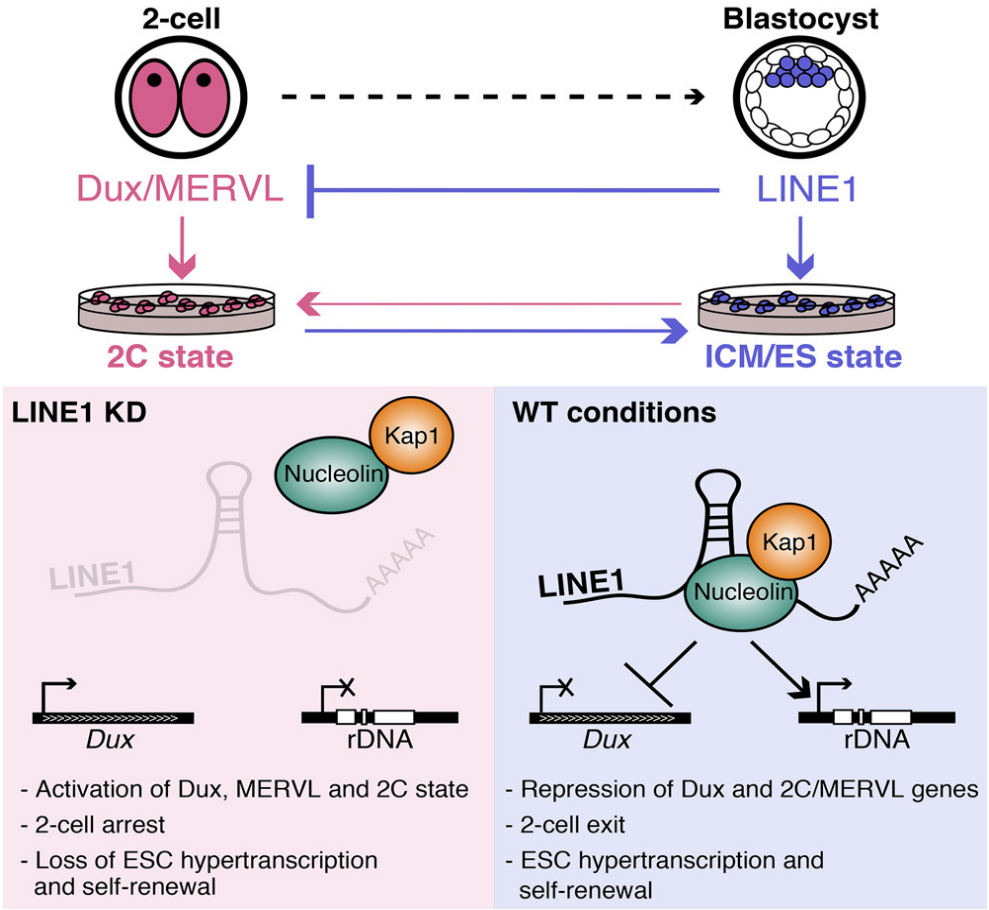


Figure 12 Proposed mechanism for the complex Nucleolin-KAP1-LINE1 in the regulation of Dux and the 2CLC population. *As proposed by* (Percharde *et al.*, 2018)*, the association of the protein complex Nucleolin-KAP1 with the retrotransposon LINE1 represses the expression of Dux, the master transcription factor of the 2CLCs. Upon dissociation of the complex, through either loss of any of the three components, the repression on Dux is relieved. The complex also regulates positively the expression of the ribosomal DNA locus, inducing a decrease in translational activity in 2CLC.* (Percharde *et al.*, 2018)

the other hand, *Dux* KO mESCs fail to acquire the 2CLC phenotype, alone or in a KAP-1 or CAF-1 KO background (Ishiuchi *et al.*, 2015; Percharde *et al.*, 2018). The activity of KAP-1 also involves other actors such as the LINE-1 transposable element and the Nucleolin protein. Assembled in a trimeric complex, these actors actively repress the expression of Dux (**Fig. 12**) (Percharde *et al.*, 2018).

Altogether, this highlights the heterogeneity and the plasticity of embryonic stem cells in culture, with naïve cells, even in the ground state, oscillating between different levels of pluripotency and either their transition to EpiSCs or the EpiSCs themselves showing also various degrees of post-implantation epiblasts.

4. Diapause

In some mammals, the developmental process occurring between the pre-implantation to the post-implantation stage can be paused in a state called diapause. This arrested stage can last for weeks and is reversible: the progression of development will resume normally leading to the implantation and embryo development. It is considered as a suspended stage, waiting for better conditions to restart, maximizing the chances of pregnancy success. Diapause is observed in the case of nutrient deprivation for example. *In vivo,* the hatched blastocyst will sit close to the uterine wall. In this state, the cellular metabolism is rewired, with a decrease of mitochondrial activity, increased glycolysis and lipolysis. These pathways are mostly controlled by a reduction in mTOR activity and activation of AMPK (AMP-activated protein kinase) (Bulut-Karslioglu *et al.*, 2016; Ehnes *et al.*, 2020; Hussein *et al.*, 2020). This paused state can be induced in mice through estrogen deprivation or ovariectomy (Yoshinaga and Adams, 1966; MacLean Hunter and Evans, 1999) or in mESCs through mTOR inhibition (Bulut-Karslioglu *et al.*, 2016; Hussein *et al.*, 2020).

III. Heme: Synthesis and function

A way to understand which genes or pathways are important for proper development and especially for proper embryo implantation is to assess the embryonic lethality of specific gene-knockout embryos. By doing so, it has been shown that homozygous mutations in the *Urod* gene (Uroporphyrinogen decarboxylase) were lethal shortly after the implantation step (before E7.5) (Phillips *et al.*, 2001). It is unexpected to observe this phenotype since this enzyme is part of the heme biosynthetic pathway, that has no obvious immediate role in the progression of pluripotency. So far, this pathway has mostly been investigated in hematopoietic lineages or in hepatic function, since those two categories of cells require high amounts of heme for proper functioning. However, this is to underestimate the importance of heme, as this molecule is crucial for a large variety of biological processes thanks to the intrinsic properties of heme to function both as an electron carrier and a catalyst for redox reactions. In the two following sections we will first summarize the heme biosynthesis pathway, before describing the roles of the heme molecule.

A. The heme biosynthetic pathway

The synthesis of the tetrapyrrolic heme molecule is a multistep process occurring between the mitochondria and the cytosol compartments (**Fig. 13**). The first enzyme, the aminolevulinate synthase (ALAS) catalyzes the condensation of glycine and succinyl-CoA in the mitochondrial matrix to form δ-aminolevulinic acid (ALA). ALAS exists under two isoforms, ALAS1 being ubiquitously expressed while ALAS2 is erythroid-specific (Ajioka, Phillips and Kushner, 2006). Both isoforms contain a heme-binding motif and binding of heme to ALAS blocks the import of the enzyme into mitochondria (Munakata *et al.*, 2004). This negative feedback of the final product on the first enzyme of the biosynthesis pathway is a prime example of post-translational regulation mechanism. While PGC1-α (peroxisome proliferator-activated coactivator 1α) drives the transcription of *Alas1*, the promoter sequence of *Alas2* is responsive to the erythroid-specific GATA1 transcription factor (Surinya, Cox and May, 1997).

The ALA generated is then shuttled to the cytosol through the ATP-binding cassette protein B10 (ABCB10), although the involvement of this protein is somewhat controversial (Seguin *et al.*, 2017). Once in the cytosol, two ALA molecules are condensed in monopyrrole

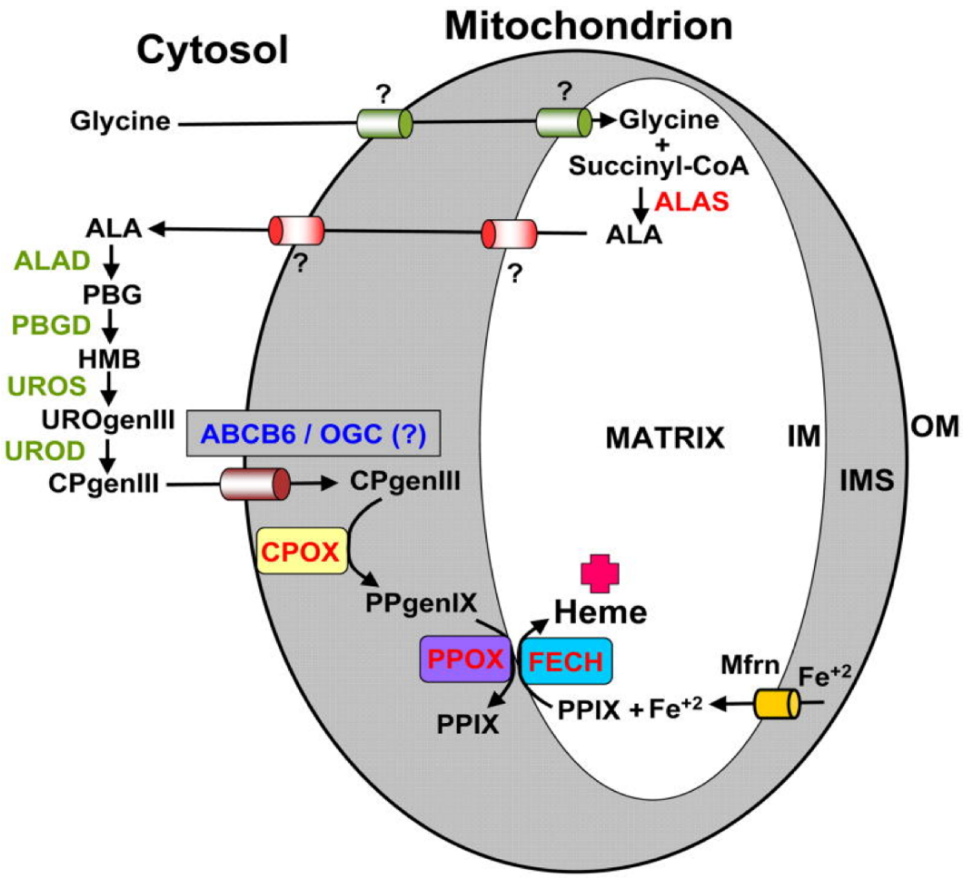


Figure 13 Diagram of the heme biosynthetic pathway and its relative subcellular localization as detailed in section III.A. *Mitochondrial enzymes are in red and cytosolic enzymes in green. OM= outer mitochondrial membrane; IMS= intermembrane space; IM= inner mitochondrial membrane; ALA= γ-aminolevulinic acid; ALAS= ALA synthase; ALAD= ALA dehydrogenase; PBG= porphobilinogen; PBGD= PBG deaminase; HMB= hydroxymethylbilane; UROS= uroporphyrinogen II synthase; UROgenIII= uroporphyrinogen III; UROD= URO decarboxylase; CPgenIII= coproporphyrinogen III; CPOX= coproporphyrinogen III oxidase; PPgenIX= protoporphyrinogen IX; PPOX= protoporphyrinogen IX oxidase; PPIX= protoporphyrinogen IX; FECH= ferrochelatase; MFRN= mitoferrin; ABCB6= ATP-binding cassette B6; OGC= 2-oxoglutarate/malate carrier protein* (Severance and Hamza, 2009)

porphobilinogen (PBG), a reaction catalyzed by the ALA dehydratase ALAD (also called porphobilinogen synthase, PBGS). This enzyme is octameric, actually a tetramer of homodimers, and uses zinc ions for the catalysis. Replacement of these Zn^{2+} ions by Pb^{2+} ions, resulting in loss of enzymatic activity, is one of the hallmarks of lead poisoning, leading to anemia and neuronal toxicity due to ALA accumulation (Scinicariello *et al.*, 2007). ALAD is also inhibited by succinylacetone, a metabolite formed in pathological conditions of hepatorenal tyrosinemia (Sassa and Kappas, 1983). Four of these PBG will be then condensed by the PBG deaminase (PBGD, also called uroporphyrinogen I synthase) as the third step of this pathway, to generate hydroxymethylbilane (HMB) prior to being taken up by the uroporphyrinogen III synthase (URO3S or UROS), catalyzing the closure of the porphyrinogen cycle. The formation of this uroporphyrinogen III marks the end of the ring formation steps of the pathway, the next steps serving as successive modifiers of the lateral chains.

Indeed, next is the elimination of the four carboxylic groups of the acetate side chains, by the uroporphyrinogen decarboxylase (UROD). The resulting coproporphyrinogen III is transported back into the mitochondria, in the intermembrane space (IMS), via the ABCB6 transporter, where the coproporphyrinogen oxidase (CPOX) decarboxylates the propionate groups of the pyrrole rings, generating protoporphyrinogen IX. At the outer surface of the inner mitochondrial membrane (IMM) lies the penultimate enzyme of the pathway, the protoporphyrinogen oxidase (PPOX), oxidizing the protoporphyrinogen IX into protoporphyrin IX, while translocating the product in the matrix. Fully functional heme (or heme b) is finally

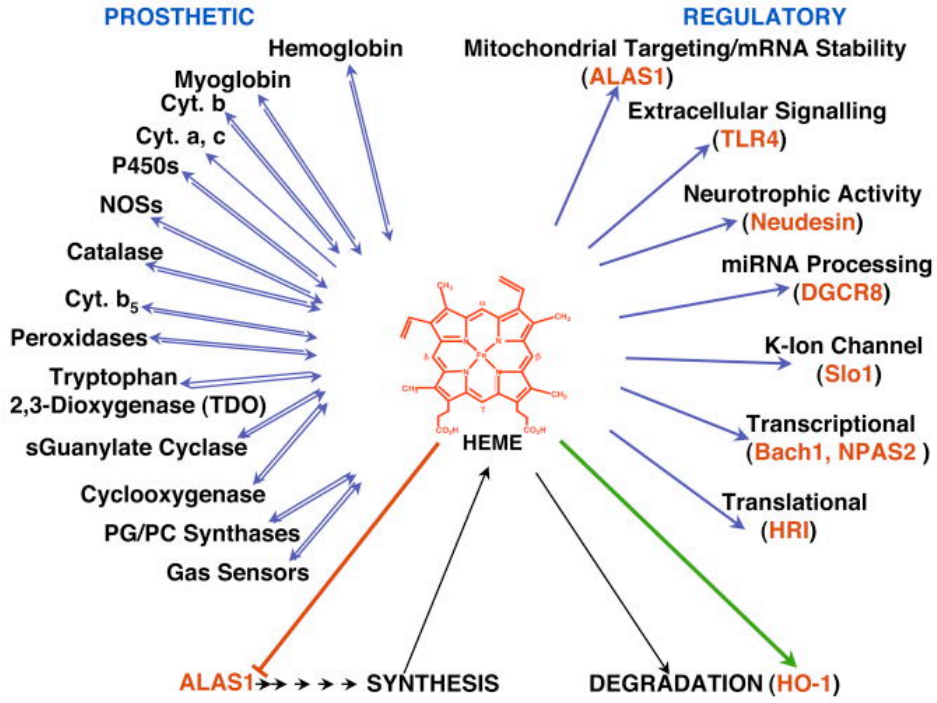


Figure 14 Schematic representation of the role of heme in cellular processes. *There is a tight control of heme abundance trough a negative feedback loop between heme and ALAS1 (red arrow), thereby reducing its synthesis rate, while exerting a positive feedback on its degradation enzyme Heme oxygenase-1 (HO-1) (green arrow). On the left, heme serves as a prosthetic group for the enzymatic activity of a wide variety of enzymes involved in ROS detoxification, xenobiotics detoxification or signaling molecule production. On the right, heme is used as a regulating molecule, modifying the DNA affinity of transcription factors (BACH1), or changing the conformation of the protein for a modification in enzymatic capacities. From (Correia, Sinclair et De Matteis, 2010).*

formed by addition of Fe^{2+} to protoporphyrin IX by the ferrochelatase on the matrix side of the IMM (FECH) (Ajioka, Phillips and Kushner, 2006).

B. Heme as a prosthetic group

The final product of this biosynthetic pathway is a heme b, which is the most used heme subtype in mammalian cells (Swenson *et al.*, 2020). Thanks to its properties of electron carrier and catalyst for redox reactions in a wide variety of proteins and enzymes, summarized in **Fig. 14** and described in the next paragraphs, heme is an important cofactor for many enzymes. The prime examples of heme utilization are the hemoglobin proteins, used in large amount in red blood cells to carry molecular oxygen throughout the body, and the cytochromes P450 family of xenobiotic detoxifying enzymes. However, these are especially present and relevant in highly specialized tissues and cell types, so they will not be detailed here. However, heme also contributes to various ubiquitous cellular functions. The electron carrier capacities are of great use in the ETC of the mitochondria, used for energy production. Indeed, 3 out of the 4 ETC complexes have at least one heme prosthetic group and its absence causes defect in their assembly, highlighting a structural role in addition to the biochemical one (Kim *et al.*, 2012). The heme groups are not all of the b-type of heme but from its derivatives.

Starting from heme b, different enzymes catalyze the formation of a, o or c-type of heme molecules (**Fig. 15**). All the heme molecules are free but the heme c, which is the only one covalently bound through a thioether linkage to its hemoprotein, such as the cytochrome c and the cytochrome c_1 subunit of the complex III. In mammals, this action is performed by HCCS (holocytochrome c synthase), reviewed in (Babbitt *et al.*, 2015). Heme o, an intermediate in the synthesis of heme a, is generated by the action of a protoheme IX farnesyltransferase (Cox10), in the inner side of the IMM. In mammals the unique identified role for this heme metabolite is to be an intermediate in the synthesis of heme a, as opposed to some bacteria such as *Escherichia coli* using heme o in cytochrome o-containing bo3 oxidase (Puustinen and Wikstrom, 1991). The Cox15 enzyme, located on the IMS side of the IMM, uses

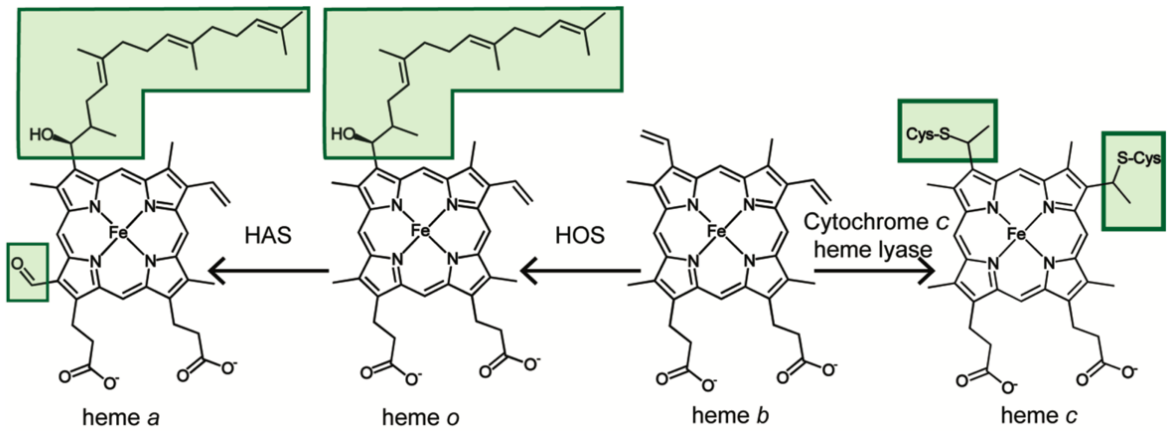


Figure 15 Schematic representation of the formation of the derivatives of heme b. *In mammalian cells, three major derivatives of heme b (or protoheme) are found. When covalently bound to its heme protein through thioether bounds with the cysteine residues to its hemoprotein, heme is of the c-type. This process is performed by the cytochrome c heme lyase also called the holocytochrome c synthase (*HCCS*). The synthesis of heme o is an intermediate to the synthesis of heme a. First vinyl groups are added at the C2 position of the porphyrin ring by the heme o synthase (HOS; COX10) then an aldehyde group is added by the heme a synthase (HAS; COX15).* (Swenson *et al.*, 2020)

the heme o generated by Cox10 to generate heme a. This cofactor is only used by the cytochrome c oxidase subunit 1 (COX1 in mouse and human) (Mogi, Saiki and Anraku, 1994). So far, the transport of heme o across the IMM is not understood (Swenson *et al.*, 2020).

An important role of the heme moiety is also found in the detoxification of reactive oxygen species (ROS), via catalase or peroxidases. Catalases and peroxidases are enzymes able to detoxify one of the most abundant ROS in mammalian cells, hydrogen peroxide (H_2O_2). This function is crucial to the maintenance of cell homeostasis: on one side the ROS entities have to be tightly balanced to avoid deleterious effects such as DNA or lipid oxidation but on the other side they are required for the proper activity of signaling pathways (Gough and Cotter, 2011; Duvigneau, Esterbauer and Kozlov, 2019). The presence of heme in catalase is fundamental for its activity since the iron ion will serve as an electron acceptor during the catalysis of the reaction. Ultimately, two H_2O_2 molecules will be reduced to dioxygen and 2 molecules of water. The balance between the beneficial and deleterious effects of ROS is decisive in stemness maintenance and in differentiation (reviewed in (Nugud, Sandeep and El-Serafi, 2018)).

Along with these ROS detoxifying enzymes, heme is also involved in the activity of the nitric oxide synthases (NOS), converting L-arginine to L-citrulline and nitric oxide (NO). This production of NO is known to play a role in the maintenance of pluripotency (Beltran-Povea *et al.*, 2015). Indeed, low μM of NO have been shown to protect mESCs from death or differentiation following LIF withdrawal, through a blocked caspase 3 activation, an upregulation of anti-apoptotic genes such as *Bcl-2*, and a downregulation of differentiation genes such as *Brachyury* or *Gata4* (GATA Binding Protein 4) (Tejedo *et al.*, 2010). However, this positive role of NO on stemness maintenance is tightly balanced since an increase in its concentration (low mM) induces a decrease of OCT4 and NANOG abundance, a p53-dependent effect. NO indeed increases p53 stability and induces p53 activation through post-translational modifications (Mora-Castilla *et al.*, 2010).

NO, the smallest signaling molecule in the cell, is in turn able to regulate different pathways through for example the formation of cGMP (cyclic guanosine monophosphate) or the formation of reactive nitrogen species (RNS) (Förstermann and Sessa, 2012). The nitric oxide (NO) produced in the cytoplasm is also known to stabilize HIF1α, shown to be important for the metabolic switch happening during the naïve-to-primed transition (Zhou *et al.*, 2012). This effect of NO on HIF1α is probably dose-dependent, but NO has been reported to induce the S-nitrosylation of the cysteine 533 of the HIF-1α protein, decreasing the action of the oxygen-dependent PHD proteins (Li *et al.*, 2007).

The action of heme on cellular processes goes beyond its role of electron carrier to also serve as cofactor involved in the stability of heme-sensing proteins. This is nicely illustrated by the transcriptional regulator BACH1 (BTB and CNC homolog 1). BACH1 is a transcriptional regulator known to bind MARE sequences (Maf recognition element) and inhibits the transcription of the downstream genes. The binding of heme to the multiple heme regulatory motifs (HRMs) induces a decrease in the transcriptional activity of multiple fronts: first, binding of heme reduces the DNA-binding affinity of BACH1 (Ogawa *et al.*, 2001) and then, it also triggers the translocation of the protein out of the nucleus while inducing its degradation by the proteasomal machinery (Suzuki *et al.*, 2004; Zenke-Kawasaki *et al.*, 2007). The genes

regulated by BACH1 comprise genes involved in heme degradation, redox regulation or even the cell cycle. Interestingly, in the context of the ESC transition, TFE3 was also shown to be a direct target of this TF (Warnatz *et al.*, 2011).

Another crucial heme-sensing protein is the Heme-regulated inhibitor (HRI) protein involved in the Integrated Stress Response (ISR). This kinase was originally described in the context of erythropoiesis: in erythroblasts, since the production of hemoglobin increases dramatically and that the lack of heme in globin chains induces their toxic aggregation, HRI acts as a safeguard reducing protein (and thus globin) synthesis under heme-deprivation conditions. In a basal situation, HRI forms an inactive dimer by stabilization with two heme molecules (Chen, 2007). When the concentrations of heme are low, the HRI dimer is activated by autophosphorylation. As a kinase, HRI is then able to phosphorylate its target, the eukaryotic translation initiation factor 2α (EIF2α). In turn, p-EIF2α has two major consequences: a reduction of the CAP-dependent protein translation and a concomitant enhancement of the translation of mRNAs with upstream open reading frames (uORFs) in their 5'UTR region. The activating transcription factor 4 (ATF4) is an example of protein with an uORF, that will in turn activate the transcription of genes involved in the stress response such as antioxidants genes promoting cellular recovery from the stress or apoptosis-inducing genes (Pakos-Zebrucka *et al.*, 2016).

C. Heme trafficking

After its production in the mitochondrial matrix, heme must be transported across the cell to reach a wide variety of hemoproteins. This mechanism must be tightly regulated as free heme is a toxic compound. As much as the biosynthesis of heme and its regulation are well characterized, the regulation of its transport across the cell is still relatively unknown. Only a few proteins have been demonstrated to act as heme chaperones among which the progesterone receptor membrane component 1 and 2 (PGRMC1/2) or even the glyceraldehyde 3-phosphate dehydrogenase (GAPDH), although the mechanisms underlying this transport have not been fully elucidated. PGRMC1 was shown to interact with FECH, directly binding the produced heme, and delivering to hemoproteins such as the P450 cytochromes (Hughes *et al.*, 2007), while its homolog PGRMC2 was proved to deliver heme to the nucleus in adipocytes (Galmozzi *et al.*, 2019). On the other hand, a recent study showed that GAPDH was able to bind heme, through conserved histidine residues, and deliver it to cytosolic proteins such as the inducible NOS (iNOS, NOS2) (Hannibal *et al.*, 2012; Sweeny *et al.*, 2018).

The transport of heme could also be a direct mechanism, without involving molecular chaperones. Indeed, the mitochondria-associated membranes (MAMs) could mediate the direct transport of heme from the mitochondria to the endoplasmic reticulum (ER). While this has never been demonstrated experimentally, a few heme-associated proteins (FECH, CPOX or PGRMC1) have been identified in proteomic analysis of MAMs (Poston *et al.*, 2011). The hydrophobic nature of heme would also support this idea.

Unlike most of the metabolites in the TCA that can contribute to other metabolic pathways, porphyrins fate is restricted to heme biosynthesis. Also, since the heme synthesis pathway is somewhat inefficient, precursors do not always become heme, with some excess

porphyrins and side products thus degraded or excreted in case of imbalance (Atamna, 2004). It is estimated that 2nmol/day are excreted in rats (Bowers *et al.*, 1992) and this estimation is multiplied by 100 to 1000 times for humans (Daniell *et al.*, 1997). This changes the stoichiometry in the consumption of precursors metabolites, glycine and succinyl-CoA, and throws the balance to higher than the 8 moles of each for 1 mole of heme. Together, this would mean that the heme biosynthesis is funneling down succinyl-CoA from the TCA. In the case of porphyria, the imbalance in heme synthesis observed was linked to a change in the succinyl-CoA availability (Homedan *et al.*, 2014).

IV. Succinate, succinyl-CoA and protein succinylation

A. Metabolic functions

As metabolite of the TCA cycle, succinate is one of the crucial metabolites of the cell. In this cycle, succinyl-CoA, along with CO_2 and NADH, is the product of the oxoglutarate dehydrogenase complex (OGDC), also known as the alpha-ketoglutarate dehydrogenase complex (α-KGDH). Then, succinyl-CoA is taken up by the succinate-CoA ligase SUCL, composed of a heterodimer of an invariant α subunit (SUCLG1) and a β subunit, either SUCLA2 or SUCLG2. Depending on the association with SUCLA2 or SUCLG2, the byproduct will be the substrate level phosphorylation of an ADP or a GDP along with the release of a succinate molecule. Finally, the uptake of succinate by the succinate dehydrogenase complex (SDH) is at the crossroads of the TCA and the ETC since the SDH is the complex II of the ETC. The reaction catalyzed by SDH will consume the succinate, forming fumarate and $FADH_2$ (Flavin adenine dinucleotide). In fact, the SDH complex is composed of 4 different subunits (SDHA-D):

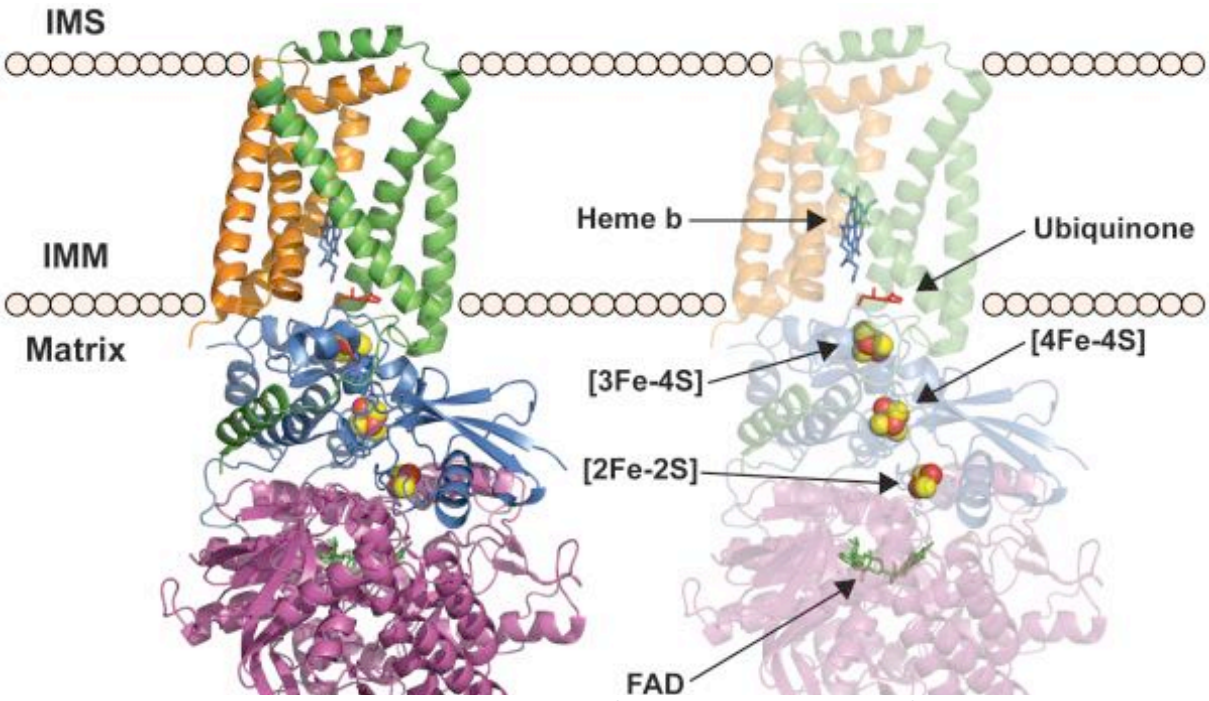


Figure 16 Structure of the succinate dehydrogenase complex (PDB accession number: 1ZOY). *The oxidation of succinate at the surface of SDHA reduces the FAD (flavin adenine dinucleotide) into $FADH_2$. The electrons are then transferred onto the successive Fe-S complexes in SDHB before ultimately reaching the ubiquinone. So far, the heme b group at the interface between SDHC and SDHD serves for the formation and stabilization of the complex. Purple= SDHA; blue= SDHB, green= SDHC; orange= SDHD.* (Van Vranken *et al.*, 2015).

SDHA shelters the FAD prosthetic group, while SDHB harbors 3 different Fe-S clusters. Both SDHA and SDHB are matrix-facing while SDHC and SDHD are embedded in the IMM (**Fig. 16**). Of note, the heme b group of the complex II is located in between these two subunits and is thought to be involved in the stabilization of the complex, more than a role in the electron transfer. The oxidation of succinate proceeds on SDHA, reducing the FAD. The $FADH_2$ then transfers its electrons one by one to the Fe-S centers of SDHB, ultimately handing them over to the ubiquinone (UbQ) (Sun *et al.*, 2005). There is thus no net flux of proton across the IMM, as opposed to complexes I and III. Succinyl-CoA is also the entry point in the TCA from anaplerotic reactions from branched-chain amino acids or even from propionate, reinforcing the role of this metabolite as a key node in mitochondrial metabolism.

B. Succinate as a post-translational modification

Apart from its role in metabolism, succinyl-CoA can also be used for post-translational protein (PTM) modification. For 50 years, utilization of chemical succinylation in the food industry is known to affect the behavior of proteins, changing the foaming capacity, swelling power or the solubility of the compounds, for example (Arueya and Oyewale, 2015). 10 years ago, the succinylation of lysine residues has been shown to also occur naturally *in vivo* (Zhang *et al.*, 2011). The impact of such a modification, compared for example to acetylation, is more drastic. Indeed, the two negative charges of the succinyl moiety switch the net charge of the otherwise positive lysine residue to a net negative one. Other than a charge modification, the succinylation of a lysine imply a relatively big structural change by adding a mass of 100 Daltons (Zhang *et al.*, 2011). In the same way than protein acetylation, protein succinylation is a direct link between the mitochondrial metabolic activity (TCA and ETC) to protein functions, even in remote cellular compartments.

So far, the mechanisms leading to protein succinylation are not fully understood. A controversial question resides in the mechanism of lysine succinylation, that could be enzymatic or non-enzymatic (Sreedhar, Wiese and Hitosugi, 2020). While there is no consensus yet, we could postulate that the two processes probably coexist. On one hand, succinyl-CoA has indeed been shown to *in vitro* succinylate bovine serum albumin in a dose-dependent manner (Colak *et al.*, 2013). Accordingly, since mitochondria hosts high concentrations of the metabolite, it has been postulated that the process of protein succinylation in that organelle could be mostly non-enzymatic.

On the other hand, several enzymes have been shown to catalyze the deposition of this modification. This is the case for the lysine acetyltransferase 2A (KAT2A) shown to mediate the histone 3 succinylation of the lysine K79 (H3K79Succ) in the context of cancer cell proliferation. In this setting, KAT2A is partnering with α-KGDH to locally produce the succinyl-CoA needed in the nucleus (Wang *et al.*, 2017). At the outer surface of mitochondria, CPT-1A, classically mediating the mitochondrial import of long chain fatty acids, was also attributed a succinyltransferase activity both *in vivo* and *in vitro* (Kurmi *et al.*, 2018). Interestingly, through mutation of the CPT catalytic site, they showed that the succinyltransferase activity was maintained, using thus a distinct catalytic domain. About 100 proteins were identified as targets for succinylation by CPT1A, among which the glycolytic enzyme enolase (ENO1), whose activity was reduced when succinylated, leading to an increase in cell proliferation under glutamine deprivation (Kurmi *et al.*, 2018). Finally, very recently a third enzyme was reported to be involved in protein succinylation: the histone acetyltransferase 1 (HAT1). As the name

indicates, HAT1 was first described as an acetyltransferase enzyme for histone and non-histone proteins (Nagarajan *et al.*, 2013; Sadler *et al.*, 2015). In the context of succinylation, HAT1 was shown to mediate histone (H3K112Succ) and non-histone succinylation (Phosphoglycerate mutase; PGAM). These two modifications were linked to the *in vitro* and *in vivo* cancer progression (Yang *et al.*, 2021).

The other side of the coin of protein succinylation is the desuccinylation. Whereas the process of succinylation is only recently investigated, the enzymes removing the various PTMs are pretty well identified and belong to the sirtuin family of enzymes. These enzymes are nicotinamide adenine dinucleotide (NAD)-dependent protein lysine deacylases. In particular, two members of the family have been identified as desuccinylases: Sirt5 and Sirt7. Sirt5 is mainly a mitochondrial enzyme matching with the observation of a predominance of protein succinylation in this organelle, although it has been recently observed in the cytosol (Du *et al.*, 2011; Y. Du *et al.*, 2018). Sirt7 has also been observed to desuccinylate proteins in the nucleus (Tong *et al.*, 2017). According to their roles, loss of the Sirt5 and Sirt7 induces a phenotype of hypersuccinylation of proteins (Rardin *et al.*, 2013; Li *et al.*, 2016).

The functional consequences of protein succinylation depends highly on the context. For histones, the succinylation of a lysine, changing the positive charge to a negative one, induces a decompaction of the chromatin since the DNA is negatively charged (Smestad *et al.*, 2018). In this study, the authors induced and hypersuccinylation by mutating the SDH complex, resulting in the accumulation of succinate that was able to exit mitochondria, by yet unidentified transporters. The succinylation of histones was thus a result of a defective TCA metabolism, linking once more the metabolic state of cells to their epigenetic landscape. Functionally, this chromatin decompaction was also linked to a defect in the DNA repair process, similarly observed in Sirt7 KO cells (Li *et al.*, 2016; Smestad *et al.*, 2018). In the context of metabolic signaling, the hypersuccinylation phenotype observed after Sirt5 KO leads to a rewiring of fatty acid metabolism from mitochondrial to peroxisomal. This confers an advantage for kidney proximal tubules during acute kidney injuries, reducing the mitochondrial ROS production in response to fatty acids, displacing the major part of β-oxidation in the peroxisome. Most of the enzymes for the fatty acid oxidation in both compartments were found succinylated but the resulting activity was opposite, activation in the peroxisome and inhibitory in the mitochondria (Chiba *et al.*, 2019).

The field investigating the protein succinylation is only booming for the past 4-5 years. Over this period, a few studies have tackled the task of identifying the proteins targeted by this PTM. Overall, and depending on the context such as *Sirt5*$^{-/-}$ or *SDH*$^{-/-}$ cells, up to 2000 lysine sites were succinylated, accounting for about 500 proteins (Park *et al.*, 2013; Rardin *et al.*, 2013; Smestad *et al.*, 2018; Guo *et al.*, 2020; Gut *et al.*, 2020). Most of the identified proteins are metabolic enzymes, among the most identified gene ontologies (GO) are the TCA cycle, branched-chain amino acid degradation, pyruvate metabolism or the fatty acid β-oxidation. However, even though these metabolic pathways are centered on the mitochondria, all the cellular compartments have been shown to be affected by protein succinylation, from mitochondria, to cytosol and nucleus (Park *et al.*, 2013; Smestad *et al.*, 2018). Interestingly, these high throughput studies have also revealed that many of the succinylated lysines are also targets for other PTMs such as acetylation (Park *et al.*, 2013;

Rardin *et al.*, 2013). This crosstalk between the PTMs could represent another way of regulating their deposition.

C. Succinate-mediated regulation of cellular functions

The impact of succinyl-CoA or succinate on the epigenetic regulation goes beyond the PTMs of histones. Succinate is indeed a known inhibitor of dioxygenases such as the histone demethylases (HDM) and the TET (Ten eleven translocation) family of DNA demethylases (Laukka *et al.*, 2016; Tretter, Patocs and Chinopoulos, 2016). For example, loss of SDH activity, leading to the accumulation of succinate throughout the cell, was shown to increase the methylation of the H3K27, due to a decrease in the activity of the HDM Jmjd3 (Jumonji domain-containing protein D3) (Cervera *et al.*, 2009). The 2012 Xiao paper confirmed these results and further showed a reduction in the hydroxylation of the 5-methylcytosine (5mC) due to reduced TET activity (Xiao *et al.*, 2012). These results were then also reported in the clinic as a "hypermethylator" phenotype of patient samples with SDH deficiency (Killian *et al.*, 2013)

The inhibitory action of succinate on dioxygenases (like HDM and TET) also includes the inhibition of the PHD enzymes involved in the degradation of HIF1α or HIF2α. By inhibiting PHD, succinate thus participates in the stabilization of the HIFα subunits, with subsequent activation of the transcription of the target genes. This phenomenon is thus called pseudohypoxia (Selak, Durán and Gottlieb, 2006; Kluckova and Tennant, 2018).

Finally, succinate is known to have paracrine functions on the cellular microenvironment, through the activation of its target orphan G-protein coupled receptor (GPR91). The recognition of succinate by GPR91, later renamed SUCNR1 (succinate receptor 1) induces a wide variety of signals such as 1) a drop of cAMP concentration related to a decrease in the adenylate cyclase activity (Högberg *et al.*, 2011), 2) the activation of the MAPK pathway leading to the phosphorylation of the ERK1/2 kinases (Gilissen *et al.*, 2015) and 3) regulation of the intracellular Ca^{2+} concentration by activation of PLCβ (phospholipase beta) (Sundström *et al.*, 2013). This succinate-SUCNR1 response is involved in the detection of local stresses in the liver (Correa *et al.*, 2007), in the heart (Aguiar *et al.*, 2014) or even in the inflammatory response (Tannahill *et al.*, 2013; Mills and O'Neill, 2014), stresses that manifest by the release of intracellular succinate.

Objectives

As much as the naïve and primed stages have been respectively described in details, a lot of information is still missing on the processes regulating the transition between the two stages. Among the strategies undertaken to answer this question, a few teams relied on the booming CRISPR-Cas9 technology to screen for required genes in the transition. This led for example to the identification of folliculin as crucial for the transition, regulating the Wnt and mTOR pathways (Mathieu *et al.*, 2019), mTOR being also highlighted by a separate group (Li *et al.*, 2018). By taking advantage of the published datasets from these high-throughput studies, this project aims to highlight new mechanisms regulating the progression from the naïve to the primed stages.

To address the question of the transition, we compared the genes found required for this naïve-to-primed transition in these two studies. Since they have been performed on mESCs and hESCs respectively, we aim for the identification of a pathway that would be common to both species. Intriguingly, the heme biosynthesis pathway appears in both screens, highlighting its importance in the process. We will follow up by deciphering the mechanisms involved. For this purpose, we will start by using the *in vitro* culture of mESCs as described before: grown in 2iL to represent the naïve state, and pushed to the primed stage by a cocktail of FGF2 and activin A. Using chemical inhibitors of the pathway will help us to first confirm the dependency of ESCs on this pathway to properly transition, and second to investigate the underlying mechanisms, among which the ISR activation, protein synthesis inhibition and the modulation of signaling pathways.

Together, this project aims at a better characterization of the two pluripotent states, modelling *in vitro* the implantation step during the establishment of a pregnancy *in vivo*, this implantation being one of the most critical events for the embryo development.

Material & Methods

A. Cell culture

mESCs (ES-E14TG2a) were cultured in N2B27 medium consisting in 1:1 mixture of DMEM/F12 (Gibco, 31330-038) and Neurobasal Medium (Gibco, 21103-049) supplemented with 1x N-2 Supplement (Gibco, 17502-048), 1x B-27 Supplement (Gibco, 17504-044), 1/100 penicillin-streptomycin (Gibco, 15140-122), 1x MEM nonessential amino acids (NEAA) (Gibco, 11140-035), 1x GlutaMAX (Gibco, 35050-038), 1x sodium-pyruvate (Gibco, 11360-039) and 0.1 mM β-mercaptoethanol (Gibco, 31350-010). Naïve mESCs were maintained on 0.2 % gelatin (Sigma, G1393)-coated plates at a density of 50 000 cells/cm^2 and in N2B27 medium complemented with 10^3 U/ml of mLIF (ESGRO, ESG1107), 3 µM of GSK3 inhibitor (CHIR99021) (Peprotech, 2520691) and 1 µM of MEK inhibitor (PD0325901) (referred to as 2iL) (SelleckChem, S1036). Cells were passaged every 2-3 days using accutase (Stemcell Technologies, #07920). Cells were then collected by centrifugation at 1200 rpm for 3 minutes and counted before seeding. The transition to EpiSC was obtained by transferring naïve mESCs on 15 µg/ml fibronectin (Gibco, 33010-018)-coated plates at a density of 30 000 cells/cm^2 and by supplementing the N2B27 medium with 12 ng/ml of bFGF (Peprotech, 100-18B) and 20 ng/ml of activin (Peprotech, 120-14P). Coating proteins were incubated 1h before seeding. mESCs were conserved at 37°C, 5 % CO_2 in a humidified incubator.

B. mESC treatments

Heme inhibitors are used at concentration of 0.5 mM for succinylacetone (SA) (Sigma, D1415) and 10 µM for NMPP (Cayman Chemical, 20846). Hemin (Sigma, 51280) is used at a concentration of 10 µM in 0,1N NaOH. Diethyl butylmalonate (BM) (Sigma, 112038) is used at a concentration of 1 mM The BTdCPU (1-(benzo[d][1,2,3]thiadiazol-6-yl)-3-(3,4-dichlorophenyl)urea) (Millipore, 324892) was used at 2 µM.

C. RNA extraction and RT-qPCR

RNA was extracted after 2 days of culture with the ReliaPrep™ RNA Tissue Miniprep System (Promega, Z6111) following manufacturer's protocol for non-fibrous tissue by adding RNA lysis buffer on pelleted cells. RNA concentrations were quantified with the Nanophotometer N60 (Implen). Reverse transcription (RT) was performed with the GoScript™ Reverse Transcriptase kit Random Primers (Promega, A2801) to convert 1 µg of RNA into cDNA. Briefly, RNA was mixed with RNAse-free water to obtain 1 µg of RNA in 12 µL and heated 5 minutes at 70°C. Then, 8 µL of RT mix (4 µL random primers buffer, 2 µL enzyme, 2 µL RNAse free water) was added and the reaction was performed in a thermocycler (5 minutes at 25°C, 60 minutes at 20°C and 15 minutes at 70°C).

The qPCR was performed on the ViiA 7 Real-Time PCR System (ThermoFisher) with 10 ng of cDNA per reaction, SYBR Green GoTaq qPCR Master Mix (Promega, A6002) and primers listed in the **table number 1** at a final concentration of 300 nM. Altogether, 2 µL of cDNA (5 ng/µL), 1 µL of forward primer (6 µM), 1µL of reverse primer (6 µM), 10 µL of Master Mix and

6 µL of RNAse free water were added in each well. Relative expression was calculated using the $2^{-\Delta Ct}$ method with GAPDH as an endogenous control.

Table 1: List of primers used in qPCR

Gene	**Sequences (5' → 3')**
DNMT3A	F: CTGCTGTGGAATACCCTGTTAG R: CTTTCTACCTGCTGCCATACTC
ESRRB	F: GCACCTGGGCTCTAGTTGC R: TACAGTCCTCGTAGCTCTTGC
FGF5	F: GGGATTGTAGGAATACGAGGAGTT R: CCAGAAGAATGGACGGTTGT
GAPDH	F: CATGGCCTTCCGTGTTCCT R: CCTGCTTCACCACCTTCTTG
KLF2	F: CTAAAGGCGCATCTGCGTA R: TAGTGGCGGGTAAGCTCGT
KLF4	F: CCAGCAAGTCAGCTTGTGAA R: GGGCATGTTCAAGTTGGATT
OCT4	F: CACGAGTGGAAAGCAACTCA R: AGATGGTGGTCTGGCTGAAC
OTX2	F: TATCTAAAGCAACCGCCTTACG R: AAGTCCATACCCGAAGTGGTC
REX1	F: CCCTCGACAGACTGACCCTAA R: TCGGGGCTAATCTCACTTTCAT
TFCP2L1	F: GCTGGAGAATCGGAAGCTAGG R: AAAACGACACGGATGATGCTC
ZIC2	F: CAAGGTCCGGGTGCTTACC R: ATTAAAGGGAGGCCCCGAATA
TBX3	F: CTCCATTCCAGTTTGGTCAA R: CAACAGCAGCCTGGTTACAC
OCT6	F: TTTCTCAAGTGTCCCAAGCC R: ACCACCTCCTTCTCCAGTTG
DNMT3B	F: GGCAAGGACGACGTTTTGTG R: GTTGGACACGTCCGTGTAGTGAG
DUX	F: AAAGGAAGAGCATGTGCCAGC R: GCAGTAAGCTGTCCTGGGAAC
ZFP352	F: AAGTCCCACATCTGAAGAAACAC R: GGGTATGAGGATTCACCCACA
TCSTV1	F: TGAACCCTGATGCCTGCTAAGACT R: AGATGGCTGCAAAGACACAACTGC
ZSCAN4C	F: CCGGAGAAAGCAGTGAGGTGGA R: CGAAAATGCTAACAGTTGAT
MuERV-L	F: CCCATCATGAGCTGGGTACT R: CGTGCAGATCCATCAGTAAA
DUB1	F: GCAGGCCAACCTCAAACAG R: CGCAGGGCTCTCCTAAATCTT
KRT18	F: GAGAAGATTTCAGTCTCAACGA R: CGATCTTACGGGTAGTTGTC
HAND1	F: ACAAACTGAAACCTTCAAGAGG R: TTCATGTTGGAGAGGCTCC

EOMES	F: CCAAGACTCAGACCTTCAC R: TTAGCTGGGTGATATCCGT
ELF5	F: TCAGACAGCCTGTGATTCC R: GAATTGGAGCCATTCCCAG
ID1	F: AACTCGGAGTCTGAAGTCG R: GACACAAGATGCGATCGTC

D. Western blot analyses

Pellets of cells were lysed by adding protein lysis buffer (Tris-HCl pH 7.5 (20 mM), NaCl (150 mM), Glycerol 15 %, SDS 2 %, 25x protease inhibitor cocktail (PIC, cOmplete protease inhibitor cocktail, Roche 11697498001), 25x phosphatase inhibitor buffer (PIB, composed of 25 mM of Na_3VO_3, 250 mM 4-nitrophenylphosphate, 250 mM β-glycerophosphate and 125 mM NaF), TRITON X-100 1 %, SuperNuclease (Sino Biologicals, 25U/10µL) and by pipetting up and down. The protein concentration was determined by Pierce protein assay (ThermoFisher, 22660). Samples were mixed with Laemmli buffer (SDS, β-mercaptoethanol, Bromophenol blue) and heated for 5 min at 95°C before loading. 10 µg of proteins were loaded on SDS-containing 10 or 12% polyacrylamide gels. At the end of migration, proteins were transferred to PVDF membranes (IPFL00010) by liquid transfer. Membranes were then blocked in LI-COR Intercept blocking buffer PBS for 1h at RT and incubated overnight at 4°C with the primary antibodies. Membranes were washed 5 minutes 3 times with PBS + 0.1 % Tween-20 (PBST) before and after 1h-incubation with the secondary antibodies at RT. Detections were performed and quantified with Odyssey LI-COR scanner. Primary antibodies and secondary antibodies used are listed in the **table 2** below. Primary and secondary antibodies were both diluted in Licor PBST. GAPDH was used as loading control. Primary and secondary antibodies for GAPDH detection were incubated 30 minutes.

Table 2: List of antibodies used in Western blot (WB) and/or Immunofluorescence (IF).

Antibody	**Reference**	**Species**	**Source**	**Dilution**
Anti-Catalase	Abcam #16731	Rabbit	WB	1/1000
Anti-GAPDH	Sigma #G8795	Mouse	WB	1/10 000
Anti-H3	CellSignaling # 4499	Rabbit	WB	1/2000
Anti-KLF4	R&D #AF3158	Goat	IF	1/100
Anti-Puromycin	MerckMillipore MABE343	Mouse	WB	1/1000
Anti-iNOS	Abcam #ab15323	Rabbit	WB	1/1000
Anti-Oct3/4	Santa-Cruz #sc-5279	Mouse	WB IF	1/1000 1/100
Anti-OTX2	R&D #AF1979	Goat	WB	1/1000
Anti-OXPHOS rodent cocktail	Abcam #ab110413	Mouse	WB	1/1000
Anti-Pan-succinyl-lysine	PTM Bio #PTM-401	Mouse	IF	1/300
Anti-TFE3	Sigma #HPA023881	Rabbit	IF	1/300
Anti-ZSCAN4	Millipore #AB4340	Rabbit	IF WB	1/300 1/1000

Anti-MUERVL-GAG	Novus #NBP2-66963	Rabbit	IF	1/100
Anti-ERK1/2	CellSignaling #9102	Rabbit	WB	1/1000
Anti-pERK1/2	CellSignaling #9101	Rabbit	WB	1/1000
Anti-SMAD2/3	CellSignaling #8685	Rabbit	IF	1/100
IRDye 800CW Anti-Goat	LI-COR Biosciences 926-32214	Donkey	WB	1/10 000
IRDye 800CW anti-Mouse	LI-COR Biosciences 926-32210	Goat	WB	1/10 000
IRDye 800CW anti-Rabbit	LI-COR Biosciences 926-32211	Goat	WB	1/10 000
IRDye 680RD anti-Mouse	LI-COR Biosciences 926-68070	Goat	WB	1/10 000
IRDye 680RD anti-Rabbit	LI-COR Biosciences 926-68071	Goat	WB	1/10 000
Anti-Goat Alexa Fluor 488	ThermoFisher #A11055	Donkey	IF	1/1000
Anti-Mouse Alexa Fluor 568	ThermoFisher #A11004	Goat	IF	1/1000
Anti-Rabbit Alexa Fluor 488	ThermoFisher #A11008	Goat	IF	1/1000

E. Puromycin-incorporation assay

Puromycin (Sigma, P8833) was added to the culture media for 10 minutes directly after 2 days of culture to obtain a final concentration of 10 µg/mL. The medium was discarded and cells washed twice with PBS. Proteins were extracted by adding protein lysis buffer directly in wells as described above. Puromycin-treated samples were then analyzed by western blot with anti-puromycin antibody (MerckMillipore, MABE343) and anti-histone 3 antibody (CellSignaling, #4499) as loading control.

F. Immunofluorescence

Cells were seeded on coated glass cover slips 2 days before fixation with 4 % paraformaldehyde (Sigma, 30525-89-4) for 15 minutes. Coating was performed for 1h with 15 µg/ml of fibronectin or with 3.5 µg/cm^2 Cell-Tak (VWR, 734-1081) diluted in sodium bicarbonate 0.1 M. Cells were permeabilized and blocked for 30 minutes incubation in blocking buffer (PBS, 0.1 % TRITON, 1 % BSA). Immunostaining was performed by an O/N incubation of cover slips at 4°C on 30 µL drops containing primary antibody diluted in blocking buffer. After three washes of 5 minutes in blocking buffer, the coverslips were incubated 1h at RT with 30 µL drops containing secondary antibody and DAPI (Sigma, 10 236 276 001) diluted 1:1000 in blocking buffer, in the dark. Cover slips were mounted with Mowiol after three new 5 minutes washes in blocking buffer. Analyses were performed with a Leica TCS SP5 confocal microscope (Leica microsystems).

G. Protein digestion and Succinyl-lysine pull-down

Cells were washed with PBS before protein lysis with IP buffer (1% IGPAL CA630, 150 mM NaCl, 50 mM Tris pH 7.4, 1 mM EDTA, 50 mM Nicotinamide, 25 mM sodium butyrate, cOmplete minitabs (1 for 10 ml)). Lysates were gently agitated for 30 min prior to centrifugation 5 min at 16000 g at 4 °C. The supernatants were incubated with 10% TCA in acetone for 4h at -20 °C to precipitate the proteins. Precipitates were collected by centrifugation at 16 000 x g for 15 min at 4 °C, washed 3 times with 80% acetone to remove trace of TCA and finally acetone was removed by air drying for 5 min. To ensure resolubilization, NaOH 0.2M was added for 2 min without disturbing the pellet before adding Rapigest 0.2% in NH_4HCO_3 50 mM. The Protein solution was agitated for 20 min at RT on a Thermomixer at 1400 RPM and then sonicated 3 times during 10 sec. Non-solubilized proteins were removed by centrifugation 5 min at 16000 g at 4 °C. The protein concentration was measured by Pierce.

1.2 mg of proteins were digested overnight at 37 °C using trypsin at an enzyme-to-protein ratio of 1:50 (w/w). Afterward, proteins were reduced by incubation with 1,4-Dithiothreitol (DTT) 10 mM 45 min at 37 °C, followed by alkylation with iodoacetamide (IAA) 40 mM for 45 min at 37°C in the dark. Samples were digested a second time using trypsin at 37 °C for 4hr at a ratio of 1:100 (w/w). After a 10 min at 16000 g at 4°C, the supernatant volume was almost completely dried by speed vac. Purified peptides were reconstituted in NETN buffer (0.5% IGPAL CA630, 100 mM NaCl, 50 mM Tris pH 7.4 and 1 mM EDTA). Agarose beads already coupled to anti-pan-succinyl lysine antibody (PTM-402) were washed three times with PBS and then added to the peptides for O/N incubation with gentle agitation. Unbound peptides were removed by 4 washes with NETN buffer then 2 washes with ultra-pure water. Finally, peptides were eluted by acidification with 1 % trifluoroacetic acid (TFA) and cleaned up with the pierce C18 SpinTips according to the manufacture's protocol.

H. Mass spectrometry

The digest was analyzed using nano-LC-ESI-MS/MS tims TOF Pro (Bruker, Billerica, MA, USA) coupled with an UHPLC nanoElute (Bruker).

1. Mass spectrometry analyses

Peptides were separated by nanoUHPLC (nanoElute, Bruker) on a 75 µm ID, 25 cm C18 column with integrated CaptiveSpray insert (Aurora, ionopticks, Melbourne) at a flow rate of 400 nl/min, at 50°C. LC mobile phase A was water with 0.1% formic acid (v/v) and B was ACN with formic acid 0.1% (v/v). Samples were loaded directly on the analytical column at a constant pressure of 800 bar. The digest (1 µl) was injected, and the organic content of the mobile phase was increased linearly from 2% B to 15 % in 60 min, from 15 % B to 25% in 30 min, from 25% B to 37 % in 10 min and from 37% B to 95% in 5 min. Data acquisition on the tims TOF Pro was performed using Hystar 5.1 and tims Control 2.0. Tims TOF Pro data were acquired using 160 ms TIMS accumulation time, mobility (1/K0) range from 0.7 to 1.4 Vs/cm². Mass-spectrometric analyses were carried out using the parallel accumulation serial fragmentation (PASEF) (Meier *et al.*, 2018) acquisition method. One MS spectra followed by six PASEF MSMS spectra per total cycle of 1.16 s. Two injections per sample were done.

2. Database searching

Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Data analysis (Bruker) version 5.3. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.7.0). Mascot was set up to search the Uniprot-MusIsoform_191212 database (December 2019 version, 97858 entries) assuming the digestion enzyme trypsin. Mascot was searched with a fragment ion mass tolerance of 0.050 Da and a parent ion tolerance of 15 PPM. Carbamidomethyl of cysteine was specified in Mascot as fixed modifications. Oxidation of methionine, acetyl of the n-terminus and succinyl of lysine and the n-terminus were specified in Mascot as variable modifications.

3. Criteria for protein identification

Scaffold (version Scaffold_4.10.0, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established with a probability above 96.0% and an FDR below 1.0% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be determined with a probability above 5.0%, an FDR below 1.0% and with at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii *et al.*, 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. A semi-quantitative analysis was performed based on the spectral counting method and normalized on the total spectra. A pValue was calculated using a Fisher exact test.

I. Co-Immunoprecipitation

In order to confirm the interaction between Nucleolin and KAP-1, the protein complex was co-immunoprecipitated. Cells were washed with PBS before protein lysis with IP buffer. Lysates were gently agitated for 30 min prior to centrifugation for 5 min at 16000 g at 4 °C. Protein concentration of the supernatant was determined by a Pierce assay and 400 µg of proteins were incubated O/N at 4°C to anti-nucleolin pre-coated Dynabeads protein G, following the manufacturer's protocol. Unbound proteins were removed by 3 washes with NETN, then 2 washes with ETN (100 mM NaCl, 50 mM tris pH 7,4, 1 mM EDTA) and a final wash with ultrapure water. Bound protein complexes were eluted by adding WB 1x loading buffer. Beads and proteins were loaded directly onto an acrylamide gel.

J. RNA sequencing

Sequence libraries were prepared with the Lexogen QuantSeq 3' mRNA-Seq library prep kit according to the manufacturer protocol. Samples were indexed to allow for multiplexing. Library quality and size range were assessed using a Bioanalyzer (Agilent Technologies) with the DNA 1000 kit (Agilent Technologies, California, USA). Libraries were subsequently sequenced on an Illumina HiSeq4000 instrument. Single-end reads of 50 bp length were produced with a minimum of 1M reads per sample.

Quality control of raw reads was performed with FastQC v0.11.7, available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc. Adapters were filtered with ea-utils fastq-mcf v1.05 (Erik Aronesty (2011), ea-utils: "Command-line tools for processing biological sequencing data"; https://github.com/ExpressionAnalysis/ea-utils). Splice-aware alignment was performed with HiSAT2 against the mouse reference genome mm10. Reads mapping to multiple loci in the reference genome were discarded. Resulting BAM files were handled with Samtools v1.5 (Li *et al.*, 2009). Quantification of reads per gene was performed with HT-seq Count v2.7.14. Count-based differential expression analysis was done with R-based Bioconductor package DESeq2. Reported p-values were adjusted for multiple testing with the Benjamini-Hochberg procedure, which controls false discovery rate (FDR).

K. Data analysis

TMM normalized rLog transformed counts were used for Principal Component analysis using R package PCATools. Gene set enrichment analysis (GSEA) was made on gene list ranked on Log2FC using R package ClusterProfiler (Yu *et al.*, 2012). For genes with FC>2 in MUERVL::Tomato^{+} list from (Macfarlan *et al.*, 2012), Z score was calculated from TMM-rLog transformed counts and plotted as heatmap using R package Heatmap.plus. Analysis was made using statistical programming language R.

Results

I. Identification of critical pathways in the naïve-to-primed ESC transition in human and mouse

To identify common genes involved in the exit from the naïve pluripotent state in mouse and human, we compared two studies using a CRISPR/Cas9 whole genome knock-out screen during this transition from naïve to primed.

First, in human, a recent paper relied on the metabolic differences between naïve and primed cells to perform the functional screen (Mathieu *et al.*, 2019). Indeed, primed cells are more dependent on SAM, comparatively to naïve cells, to regulate their epigenetic landscape, mainly through an increase in repressive histone marks (Sperber *et al.*, 2015; Mathieu *et al.*, 2019). As a consequence, primed cells exhibit an increased sensitivity to drugs such as methotrexate and acetaldehyde, both inhibiting the methionine synthase, thereby depleting the pool of available SAM. With a combination of these two molecules, and a CRISPR-Cas9 library (GecKo), only cells acquiring a mutation preventing them to transition to the primed stage would be spared upon selection. By comparing the enriched sgRNAs after selection to those before, a list of required genes for the transition or the exit of the naïve state was established. In that paper, the role of folliculin (FLCN) in the regulation of the nuclear localization of the transcription factor TFE3 and its involvement with mTORC1/2 activity was highlighted (Mathieu *et al.*, 2019).

In mouse, the strategy used is slightly different. Instead of using a metabolic approach to discriminate naïve and primed cells, the authors relied on the use of a REX1-GFP construct. REX1 (Reduced Expression 1; also known as ZFP-42) is a gene whose expression is restricted to the naïve pluripotent state. Upon the initiation of transition, its expression is rapidly downregulated, allowing the monitoring of the progression of the naïve state exit (Chambers *et al.*, 2007). Therefore, the whole genome CRISPR guide library allowed to identify the genes responsible for the maintenance of GFP expression. In this setting, Li and coworkers removed the 2i and LIF cocktail maintaining the mESCs in the naïve state. After two days of withdrawal, mESCs should have lost the GFP expression. They thus analyzed which of the targeted genes prevented this loss (Li *et al.*, 2018).

We thus compared on one side the genes enriched after the negative selection for the human screen (log2FC>1) (187 hits) and the list of genes allowing the maintenance of GFP+ cells in the mouse screen (563 hits) to identify common genes involved in the exit of the naïve state of mESCs and hESCs. The significant hits were submitted to functional annotation tools such as DAVID. For the human screen, false positive hits for apoptosis were removed. Indeed, since a negative selection is applied to induce the death of primed hESCs, cells that acquired a resistance to cell death by mutating genes involved in apoptosis for example would be spared. **Figure 17** displays the top gene ontologies (GO) for the biological processes. As highlighted in red, the heme biosynthesis pathway is shown important in both screens. Of the different enzymes in this pathway, 7 out of 8 enzymes (ALAD, PBGD, UROS, UROD, CPOX, PPOX and FECH) came out as positive hits in the CRISPR screen during the naïve-to-primed mESC

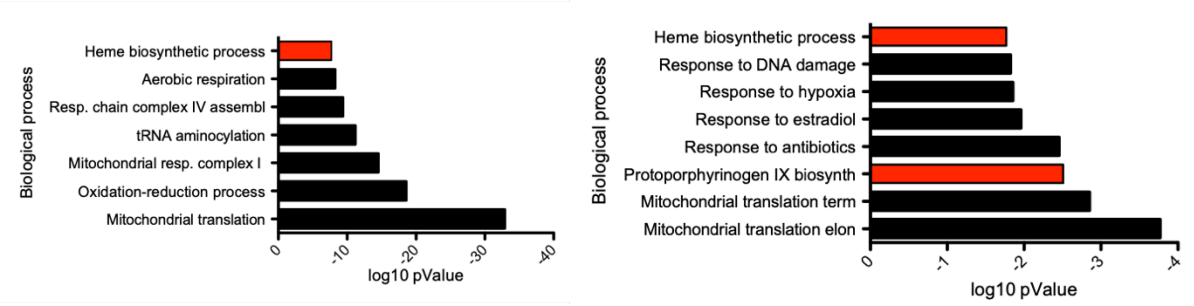


Figure 17 *Biological processes GO enrichment from two independent CRISPR-Cas9 screens for the exit of the naïve state exit, in mouse (left panel) (Li et al, 2018) and in human (right panel) (Mathieu et al, 2019).*

transition. In hESCs, only the 4 cytosolic enzymes were highlighted. Together, the results stress the importance of this pathway for the embryo's implantation. To our knowledge, the importance of this pathway in the maintenance of non-hematopoietic stem cells has never been shown.

II. Heme biosynthesis in the naïve-to-primed mESC transition

A. Heme synthesis inhibition

To study the transition between the two states of ESCs, we decided to first work on mESCs as their growth and maintenance *in vitro* is simpler than hESCs. In order to understand the requirements and the roles of heme biosynthesis in this model, interference with the pathway was needed. To do so, we used a naturally occurring inhibitor of ALAD, the second enzyme of the heme synthesis pathway, succinylacetone (SA), as described in the Introduction in section III.A. In patients with hereditary tyrosinemia, caused by a fumarylacetoacetic hydrolase (FAH) deficiency, the final enzyme in the tyrosine degradation pathway, there is a toxic build-up of SA, leading to functional consequences resembling the ones of acute intermittent porphyria (AIP) (Lindblad, Lindstedt and Steen, 1977; Sassa and Kappas, 1983). This accumulation of SA was shown to potently inhibit ALAD *in vitro*, replacing ALA in the binding pocket of the enzyme and acting as a suicide substrate, explaining why the symptoms of both diseases were comparable (Sassa and Kappas, 1983).

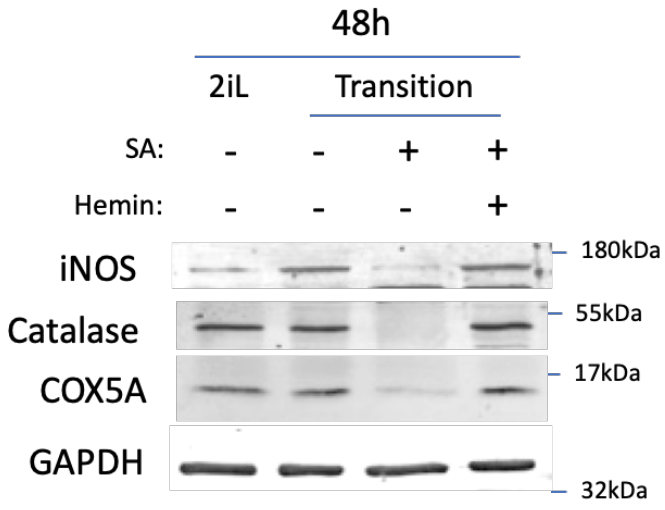


Figure 18 *Western blot analysis of the abundance of three hemoproteins (iNOS, Catalase, COX5A) after 48h of SA treatment, relative to GAPDH as a loading control for cells in naïve conditions (2iL), in transition to the Epi stage (EPI Ctl) with SA 0.5mM as heme synthesis inhibitor and 10 µM hemin supplementation. Representative blot of N=2 biological replicates.*

Prior its use in the experimental conditions, we first wanted to assess the activity of SA on the total heme content of mESCs. However, due to the low amounts of heme present in these cells, as

compared to erythrocytes, the assessment of this concentration with fluorescent approaches was not possible (Sinclair, Gorman and Jacobs, 1999). Indeed, to reach the limit of sensitivity of the test, a large amount of material was needed, which in return quenched the fluorescence readings (data not shown). To bypass this issue, we postulated that in the absence of heme production, the stability of hemoproteins would decrease. We thus chose to assess the protein abundance of three different hemoproteins: iNOS, the NO-producing enzyme; the catalase, detoxifying H2O2, and COX5A, a member of the complex IV of the ETC. As shown in figure 16, a 48h inhibition of heme biosynthesis decreased dramatically the abundance of each hemoprotein, an effect that was indeed caused by heme deprivation as supplementation with hemin rescued it (**Fig. 18**).

In addition, since the ETC relies heavily on the heme groups in multiple proteins, we decided to verify the integrity of the different complexes, using a cocktail of OXPHOS antibodies. This cocktail allows to monitor one protein in each of the complexes (I-V), giving an overview of their formation. Overall, this reveals, as expected, a drastic decrease in the abundance of these 5 proteins of the ETC, suggesting a global disruption of the OXPHOS (**Fig. 19**). It is worth noting that MTCO1 (or COX1) is a hemoprotein itself. This global disruption was rescued by the addition of exogenous hemin.

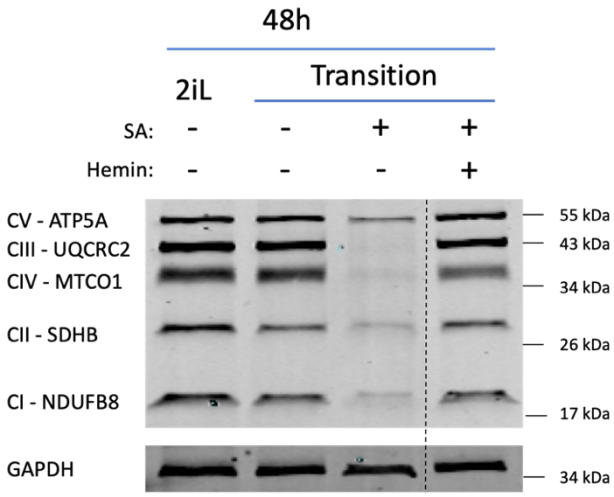


Figure 19 *Western blot analysis of the abundance members of the ETC complexes after 48h of SA treatment, relative to GAPDH as a loading control for cells in naïve conditions (2iL), in transition to the Epi stage (EPI Ctl) with SA 0.5mM as heme synthesis inhibitor and 10 μM hemin supplementation. Representative blot of N=2 biological replicates*

To further confirm these results, we also tested another inhibitor of the heme biosynthetic pathway: the N-methylprotoporphyrin (NMPP). This molecule is a known inhibitor of the last enzyme of the pathway, FECH (Jacobs *et al.*, 1998). Combining the results obtained with both inhibitors on the biology of ESCs would solidify any conclusions made. However, the maximal concentration of NMPP that mESCs tolerated (10 μM) was not able to reduce the abundance of hemoproteins as much as SA (**Fig. 20**). This difference could be due to the fact that NMPP is an analog of protoporphyrin IX and only competes for FECH activity, as opposed to the suicide activity of SA. Indeed, reports have shown that NMPP inhibition on FECH still allows the production of 40% of maximum levels of hemin (Jacobs *et al.*, 1998; Atamna *et al.*, 2002). For this reason, we decided to focus the rest of this work on heme synthesis inhibition with 0.5 mM of SA and a rescue with 10 μM of Hemin.

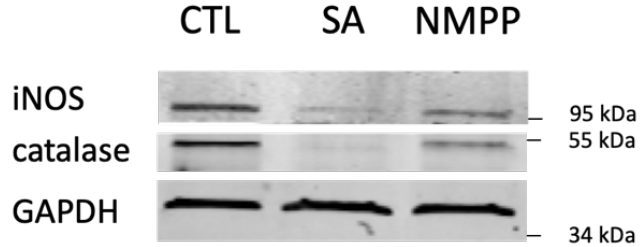


Figure 20 *Western blot analysis of the abundance the hemoproteins (iNOS and Catalase) after 48h of SA (0.5 mM) or NMPP (10 μM) treatment in 2iL mESCs, relative to GAPDH as a loading control. N=1*

B. Validation of the requirement for heme synthesis

Using succinylacetone (SA) to inhibit heme synthesis and 10 µM of hemin for its rescue, it is possible now to investigate the role of this pathway on the naïve-to-primed transition. As shown in figure **21A-B**, when mESCs are pushed for 48h to exit the naïve stage while blocking heme synthesis, the acquisition of the primed markers *Fgf5, Fgf15, Otx2, Oct6, Dnmt3a* and *Zic2* is prevented. On the other hand, the loss of naïve markers (*Esrrb, Tfcp2l1, Klf4, Tbx3*) is partially prevented. This was confirmed at the protein level by a decrease in the abundance of OTX2 and DNMT3A by western blot analysis and the increase in abundance of KLF4 by immunofluorescence, when treated with SA during the transition. Furthermore, the subcellular localization of TFE3, mainly nuclear in naïve and only cytosolic in primed (Betschinger *et al.*, 2013; Mathieu *et al.*, 2019), remains nuclear in the presence of SA (**Fig. 21C**). Hemin supplementation restores the gene expression, the protein abundance and the subcellular localization of TFE3 to levels similar to those without SA, highlighting the specific

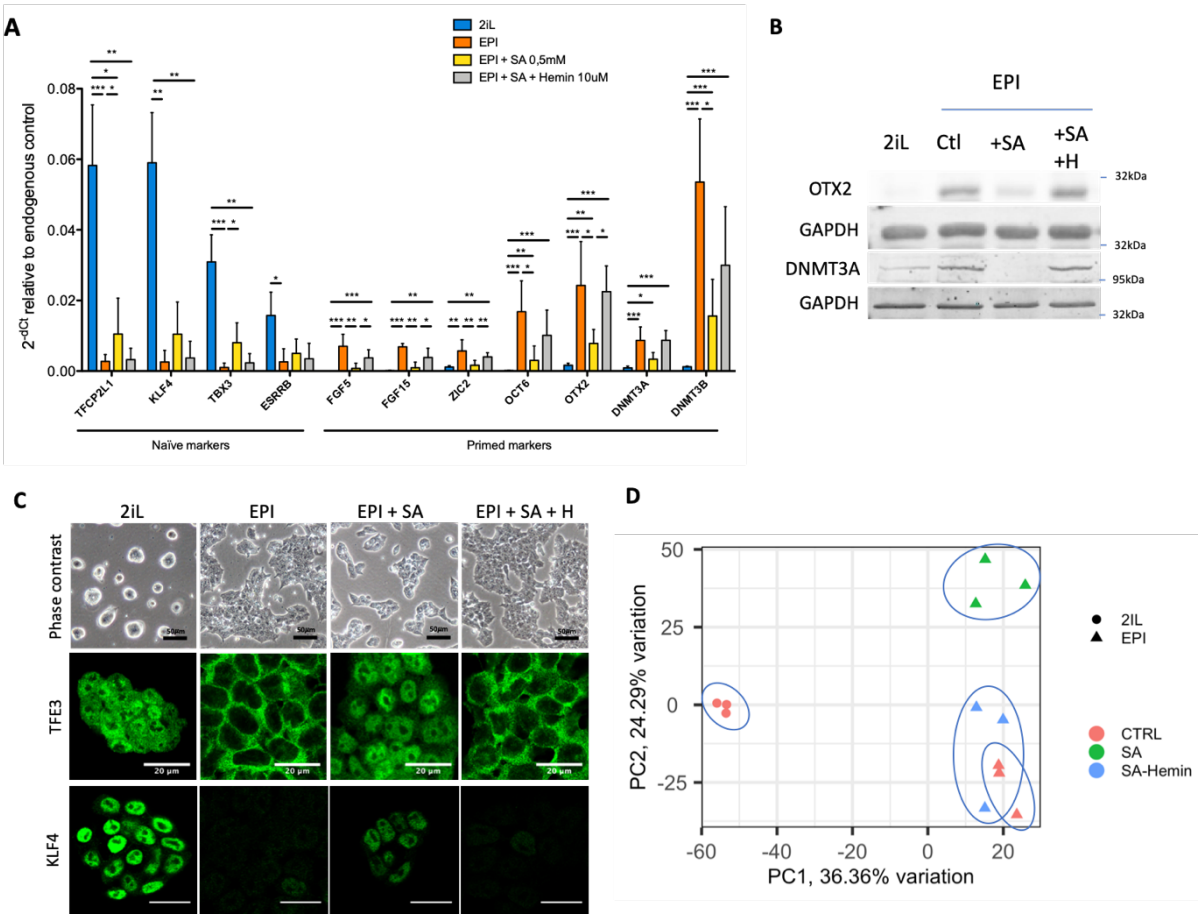


Figure 21 Heme synthesis inhibition impairs the exit of mESCs from the naïve state; effect mediated by heme. A) *Relative expression of naïve and primed markers of mESCs assessed by RT-qPCR relative to GAPDH expression (TFCP2L1, transcription factor CP2-like 1; ESSRβ, estrogen-related receptor β; KLF2/4, Kruppel-like factor 2/4; TBX3, T-Box Transcription Factor 3; FGF5/15, fibroblast growth factor-5/15; ZIC2, zic family member 2; OTX2, homeobox protein 2). S.D **p < 0.01, ***p < 0.001. ANOVA-1. n=4 independent biological replicates.* ***B)*** *Western blot analysis of the protein abundance of OTX2 and DNMT3a relative to GAPDH as a loading control for cells in naïve conditions (2iL), in transition for 2 days to the Epi stage (EPI Ctl) with 0.5 mM SA as heme synthesis inhibitor and 10 µM hemin (H) supplementation. Representative blot of n=3 biological replicates.* ***C)*** *Phase contrast micrographs of cells in naïve (2iL), primed (EPI) with treatment with heme synthesis inhibitor (SA) and hemin (H). Scale bar=50µm. Confocal micrographs of mESCs in naïve stage or in transition for TFE3 (Transcription Factor Binding To IGHM Enhancer 3) and KLF4. Scale bar =20µm. TFE3 n=3 and KLF4 n=2 biological replicates* ***D)*** *Principal component analysis (PCA) of the normalized RNAseq data transcripts.*

effect of heme in this blockage (**Fig 21A-C**). Separation of the samples with a principal component analysis (PCA) based on the normalized gene expression from RNA sequencing also reveals the segregation of the cells treated with SA (EPI + SA) from either the controls (EPI) or the cells rescued with hemin (EPI +SA +H) (**Fig. 21D**). Overall, this confirmed the screen results by showing that inhibition of heme biosynthesis impairs the naïve-to-primed mESC transition.

C. Investigation of the mechanisms for the transition defect

1. BACH1

To investigate the role of heme in this defect of transition, we first investigated the activation of the heme-sensing protein BACH1. This protein acts as a transcription factor whose action depends on the heme levels. Indeed, BACH1 possesses multiple HRE that, when bound to heme, reduces its affinity to DNA and also triggers its export from the nucleus as well as its proteasomal degradation (Ogawa *et al.*, 2001; Suzuki *et al.*, 2004; Warnatz *et al.*, 2011). Immunostaining of BACH1 during the transition doesn't reveal a change in the localization of BACH1 with SA (**Fig. 22**). The only effect observed is a decrease in its abundance when the culture media is supplemented with heme, an effect that is in accordance with the literature (Zenke-Kawasaki *et al.*, 2007). The effect of SA on the transition is thus unlikely to involve BACH1.

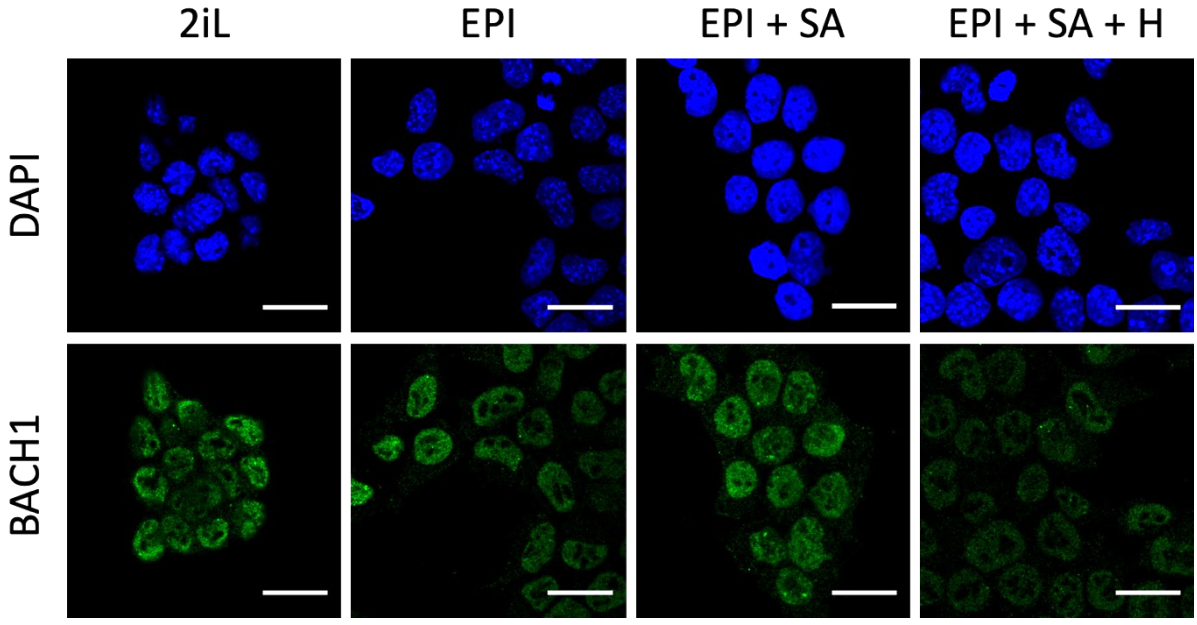


Figure 22 Heme synthesis inhibition does not modify the nuclear abundance of BACH1. *Confocal images of mESCs cells in naïve (2iL), primed (EPI) with treatment with heme synthesis inhibitor (SA) and hemin (H) rescue with immunostaining of BACH1 (BTB Domain And CNC Homolog 1) (green). Scale bar = 20 μm*

2. ISR

Heme deprivation is also a signal that induces the activation of the ISR, through the HRI kinase (see **section III.B**.). Activation of HRI induces the phosphorylation of EIF2α and a subsequent decrease in cap-dependent protein translation. To evaluate the activation of this

response, we thus measured the levels of global protein translation by using a puromycin-incorporation assay followed by western blot analysis of puromycin-labelled polypeptides. As expected from an ISR activation, blockade of heme synthesis significantly reduces the translational activity of mESCs, an effect that is rescued upon addition of hemin (**Fig. 21A-B**). Since this effect on translation is potent and could play a role in the transition defects observed in the presence of SA, we next evaluated whether the selective activation of HRI could provoke a transition blockade. To this end, we used BTdCPU (1-

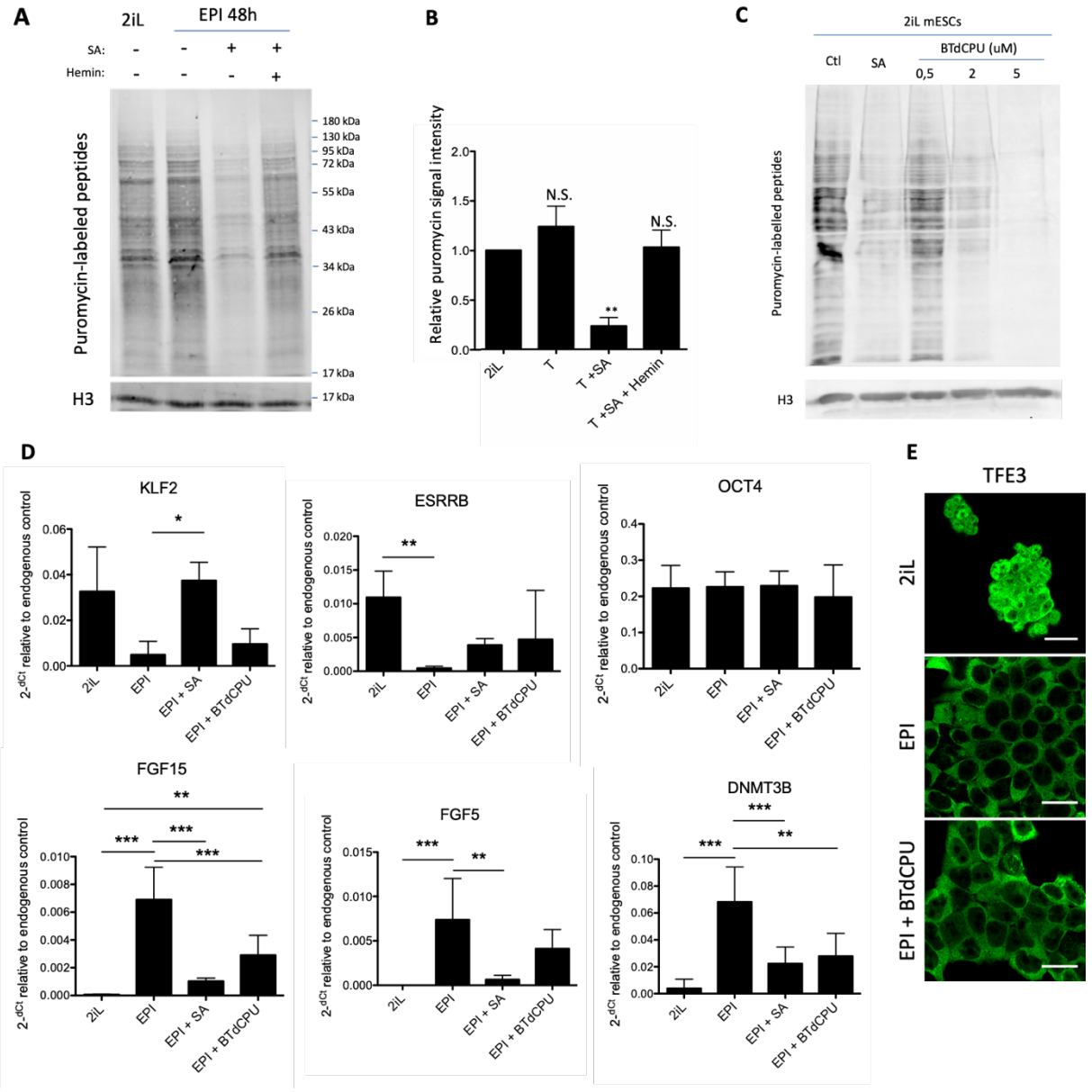


Figure 23 Activation of the ISR is not responsible for the transition defects under heme synthesis inhibition. A) *Representative blot of the puromycin-incorporation assay of mESCs in naïve (2iL), in transition for 48h (EPI) with treatment with 0.5 mM heme synthesis inhibitor (SA) and 10 µM hemin (H), relative to H3 loading control.* ***B)*** *Relative quantification of the puromycin signal from 5 independent experiments. N.S p>0.05; ** p<0.01, ANOVA1.* ***C)*** *Western blot image of the puromycin-incorporation assay of mESCs in naïve (2iL) treated with 0.5 mM SA or increasing conditions of BTdCPU for 48h, relative to the abundance of H3 loading control.* ***D)*** *Relative expression of naïve and primed markers of mESCs assessed by RT-qPCR relative to Gapdh expression. Oct4 serves as stemness marker control. (OCT4, octamer-binding transcription factor 4; Klf2, Kruppel-like factor 2; Fgf5/15, fibroblast growth factor-5/15; Dnmt3b; DNA methyltransferase 3 beta; Esrrb, estrogen related receptor beta). Results represented as means +/- S.D. *p<0.05, **p < 0.01, ***p < 0.001. ANOVA1. n=3 independent biological replicates.* ***E)*** *Confocal micrographs of TFE3 (Transcription Factor Binding to IGHM Enhancer 3) immunostaining of mESCs in naïve stage (2iL), in transition for 48h (EPI) with or without 2 µM BTdCPU. Scale bar =20 µm. n=2 independent biological replicates*

(benzo[d][1,2,3]thiadiazol-6-yl)-3-(3,4-dichlorophenyl)urea), a chemical activator of HRI (Chen *et al.*, 2011), thereby activating the same branch of the ISR pathway than SA. We selected a working concentration of 2 µM to mimic the transition inhibition to the same extent as SA (**Fig. 23C**). However, pushing mESCs in the transition in the presence of BTdCPU wasn't able to recapitulate the effect of SA on the gene expression of naïve (*Klf2*) or primed (*Fgf5, Fgf15*) markers or on the subcellular localization of TFE3 (**Fig. 23D-E**), although a trend is observed for some markers such as *Dnmt3b* or *Esrrb* (**Fig. 23D**).

3. Signaling pathways

To identify the mechanisms involved in the failure to properly transition, we performed a gene set enrichment analysis (GSEA) with the KEGG pathways (Kyoto Encyclopedia of Genes and Genomes) between EPI and EPI + SA RNAseq data. Interestingly, many crucial signaling pathways in development are shown downregulated by SA (**Fig. 22A**). We thus focused our attention on the pathways directly involved in the transition that are triggered by the combination of FGF2 and activin A in the growth media. Strikingly, on the one hand, the MAPK-ERK1/2 pathway (downstream of FGF2) isn't activated in Epi+SA cells, as shown by the absence

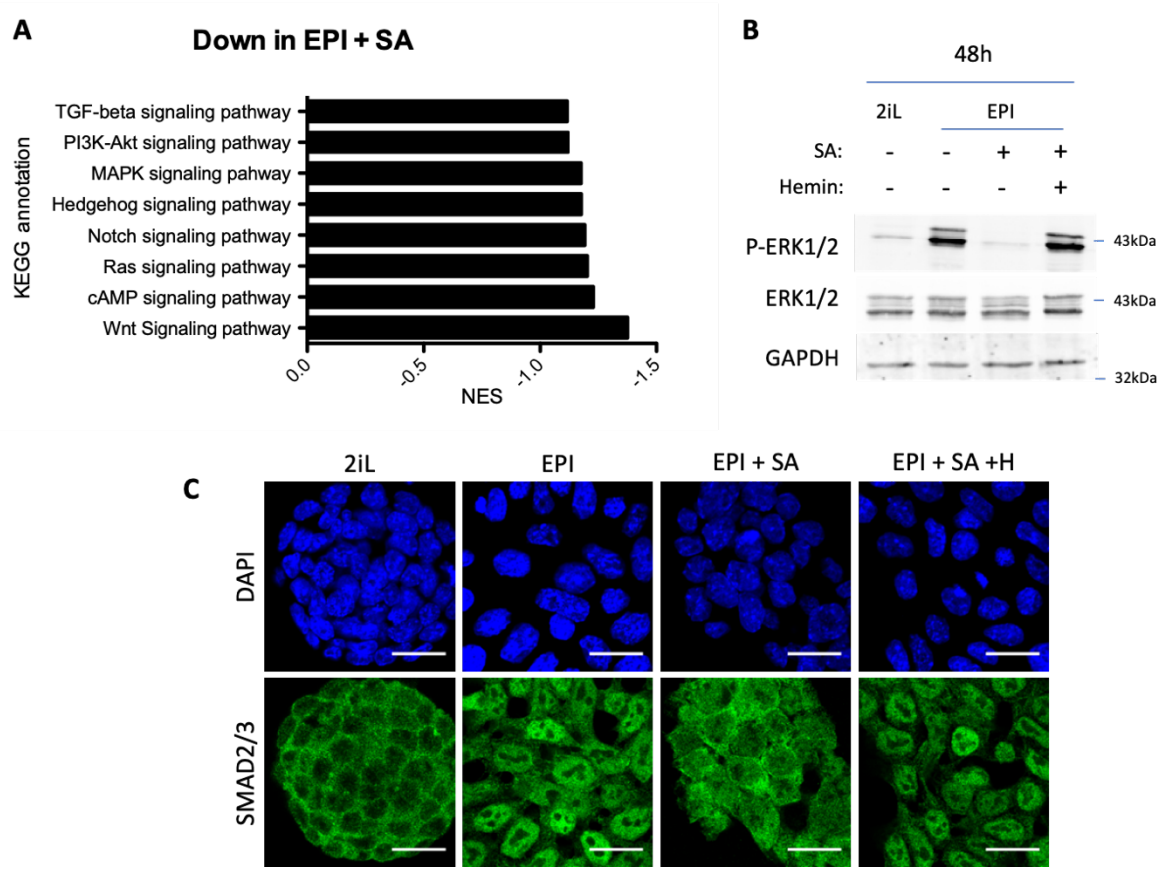


Figure 24 SA prevents the activation of the MAPK and Activin A-SMAD pathways during the mESC transition. A) *GSEA performed on RNAseq data were analyzed for gene ontology. KEGG pathways annotated down regulated in EPI+SA versus EPI ctl are represented as normalized enrichment scores (NES).* ***B)*** *Western blot analysis of the protein abundance of ERK1/2 and phospho-ERK1/2 (Thr202/Tyr204) relative to GAPDH as a loading control for cells in naïve conditions (2iL), in transition for 2 days to the Epi stage (EPI) with 0.5 mM SA as heme synthesis inhibitor and 10 µM hemin (H) supplementation. Representative blot of n=3 biological replicates.* ***C)*** *Confocal analysis of the immunostaining of SMAD2/3 (green) in cells in naïve conditions (2iL), in transition to the Epi stage (EPI) with SA as heme synthesis inhibitor and hemin (H) supplementation, representative of 2 independent experiments.*

of phosphorylation of ERK1/2 (**Fig. 24B**). On the other hand, the activin A-SMAD pathway activation is also compromised as shown by the difference in nuclear localization of SMAD2/3 compared to the EPI cells (**Fig. 24B-C**). We thus conclude that a failure to activate these key signaling pathways leads to a failure to properly transition to the primed stage, despite the presence of their respective ligands in the media.

III. Heme synthesis inhibition in naïve mESCs

A. Acquisition of a 2C-like phenotype

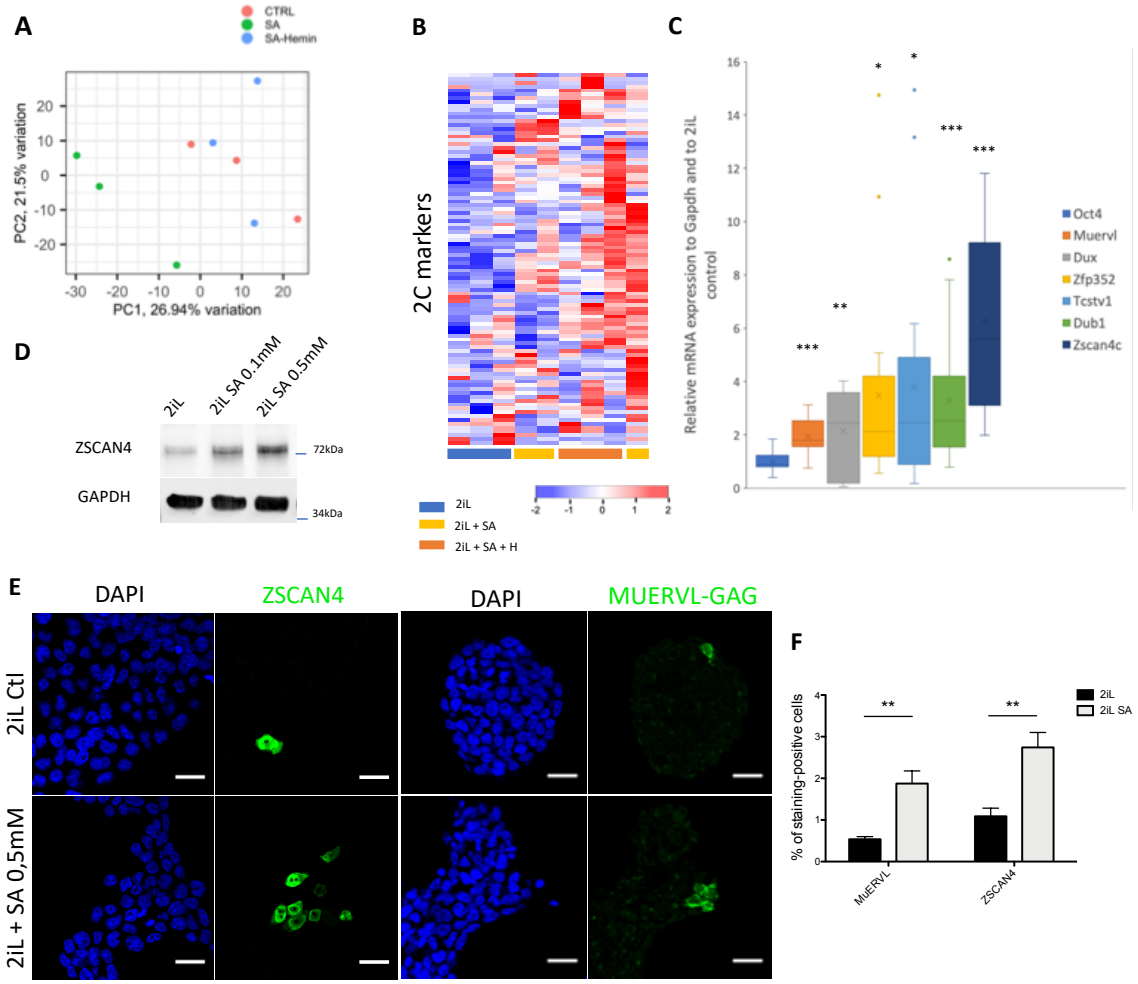


Figure 25 ***Heme synthesis inhibition pushes mESCs toward a 2C-like stage. A)*** *Principal component analysis (PCA) of the normalized RNAseq data transcripts of naïve mESCs (2iL) treated for 48h with 10 µM Hemin and/or 0.5 mM SA.* ***B)*** *Heatmap showing expression changes of 142 2C markers as defined in (Macfarlan et al. (2012), upon 48h heme synthesis inhibition (2iL + SA) and rescue with Hemin (2iL +SA +H).* ***C)*** *Relative expression of 2C gene markers of mESCs assessed by RT-qPCR relative to Gapdh expression and to 2iL naïve control. Oct4= octamer-binding transcription factor 4, Muervl= murine endogenous retrovirus-like, Dux= double homeobox, Zfp352= Zinc-finger protein 352, Tcstv1= 2-cell-stage variable group member 1, Dub1= Ubiquitin Specific Peptidase 36, Zscan4c= Zinc Finger And SCAN Domain Containing 4, isoform c. n=* ***D)*** *Western blot analysis of ZSCAN4 protein abundance relative to GAPDH as a loading control in naïve mESCs (2iL) treated for 2 days with SA at 0.1 mM or 0.5 mM. Representative image of 3 independent biological replicates.* ***E)*** *Immunostaining of ZSCAN4 (green) in naïve (2iL control) mESCs and treated with SA at 0,5mM. DAPI is used as a nuclear counterstain. Scale bar= 20um.* ***F)*** *Percentage of MUERVL- or ZSCAN4-positive cells in the whole population of naïve (2iL) mESCs or naïve treated with SA (2iL SA), counted from confocal micrographs as in (E) with 10 images per conditions for at least 1000 cells per condition. n=4 independent biological replicates. Results expressed as mean +/- S.D. ** p < 0.01 ; T-Tests.*

In parallel to this defect in the exit from the naïve state, we noticed that treatment of naïve 2iL cells with SA also modifies the gene expression as 2iL+SA samples cluster away from 2iL control cells in the PCA performed on RNAseq data (**Fig. 25A**). Our attention was drawn on markers reported to be expressed in the 2-cell embryo. Indeed, a small portion of the mESC population actually expresses a gene signature reminiscent of the 2C stage, cells thus called 2C-like cells (Macfarlan *et al.*, 2012). We aimed to confirm the expression of this subset of genes, first identified in 2012 (Macfarlan *et al.*, 2012). **Figure 25B** shows a heatmap of the expression of 2iL, 2iL +SA and 2iL +SA +Hemin of the 142 genes identified upregulated in this 2C-like population. A difference is clearly observed between the control samples versus the two other groups, showing an upregulation of this subset of genes. This was further confirmed by RT-qPCR for a selection of the most common markers (**Fig. 25C**) and by western blot analysis of ZSCAN4 abundance (**Fig. 25D**). Since the reports of this 2C-like population indicate a heterogeneity at the population level (Macfarlan *et al.*, 2012; Eckersley-Maslin *et al.*, 2016) we quantified the fraction of ZSCAN4^{+} and MUERVL-GAG^{+} cell population by confocal microscopy. As shown in **Fig. 25E and F**, treatment with SA increases the fraction of 2C-like cells from 2 to 3 folds.

Since this 2C-like population is reminiscent of the totipotency state of stem cells, we tested whether treatment with SA would increase the ability of mESCs to differentiate into trophoblast stem cells. To this end, we first plated mESCs for two days in the regular 2iL media with or without SA. The medium was then switched to the trophoblast differentiation media

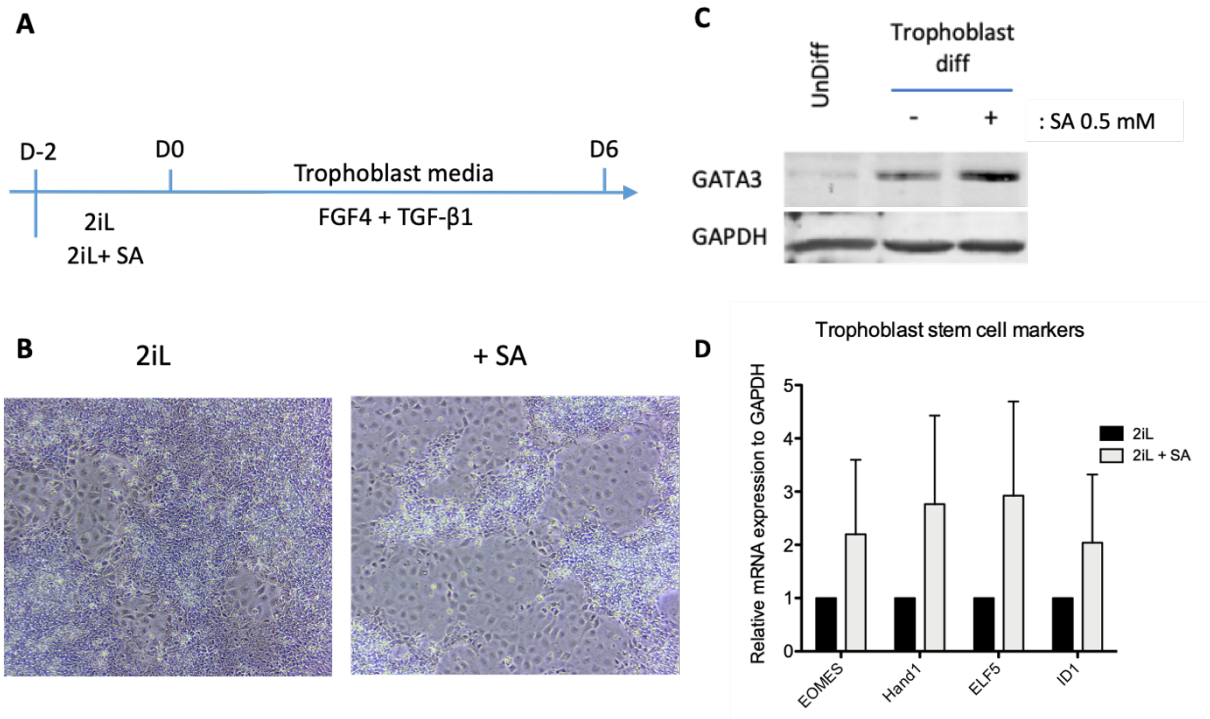


Figure 26 Pre-treatment with SA promotes a more efficient differentiation to trophoblast stem cells. A) *Experimental diagram of the differentiation process of mESCs to trophoblast stem cells. Cells are first plated in naïve media with or without SA for two days before switching to a trophoblast differentiation medium (FGF4+TGFβ1) for 6 days.* ***B)*** *Phase contrast micrographs of cells after the differentiation process, pre-incubated or not with 0.5 mM SA.* ***C)*** *Western blot analysis of the protein abundance of GATA3 (GATA Binding Protein 3) relative to GAPDH loading control, in undifferentiated mESCs (UnDiff) or after 6 days of trophoblast induction with or without heme synthesis inhibition (SA). Representative blot of 2independent experiments.* ***D)*** *Relative expression of trophoblast stem cell markers of mESCs differentiated for 6 days with (2iL +SA) or without (2iL) pre-treatment with the heme synthesis inhibitor succinylacetone, assessed by RT-qPCR relative to Gapdh expression. n=3 biological replicates. N.S.*

based on a combination of FGF4 and TGF-β1, as reported in (Kubaczka *et al.*, 2014) (**Fig. 26A**). Of note, this set up was selected as cell mortality was clearly visible for SA treatments longer than 2 days. After 6 days of differentiation, three features were analyzed: cellular morphology, gene expression and protein abundance. At the end of the trophoblast induction, we observe an increase in the abundance of cells with a different morphology, showing enlargement of the cell surface, reminiscent of the giant cells derived from the trophoblast (**Fig 26B**) (Tanaka *et al.*, 1998; Ullah *et al.*, 2008). This difference in phenotype is also observed by the increase in expression of the trophoblast lineage markers *Eomes, Hand1, Id1* or *Elf5* (although not significant) and in the abundance of GATA3 (**Fig. 26C-D**). Altogether this suggests that a 48h treatment of SA is able to expand the potential of mESCs toward the extraembryonic lineage.

B. Heme-independent acquisition of the 2C features

Interestingly, and as opposed to the naïve-to-primed setup, the observed phenotype seems independent of heme as hemin supplementation does not rescue it (**Fig. 25B and 27A**). Previous reports have shown that heme biosynthesis consumes a lot of the glycine and succinyl-CoA precursors in the mitochondria, acting as some sort of "sink" (Atamna, 2004). We thus hypothesized that heme synthesis inhibition would increase the abundance of succinyl-CoA in mitochondria, that could then exit the organelle and accumulate throughout the cell. Furthermore, another report showed that knock-down (KD) of SDHB could indeed increase protein succinylation throughout the cell, even impacting succinyl-histone modifications, demonstrating the ability of a buildup of succinate to exit mitochondria (Smestad *et al.*, 2018). We thus measured the abundance of succinylated proteins in naïve cells treated or not with SA using a pan-succinyl lysine antibody (**Fig. 27B**). In basal conditions, the bulk of succinyl-lysine modifications is located in the mitochondria, as expected. However, we observe a dramatic increase of succinylated proteins in all cellular compartments when heme synthesis is blocked, an effect that is not (or very limitedly) rescued upon hemin supplementation.

Since this increase in protein succinylation involves the exit of succinate through the mitochondrial membrane, we postulated that blocking the exit of this metabolite from mitochondria would prevent the acquisition of widespread succinyl-lysine post-translational modifications and impair the acquisition of the 2C-like cells markers, if this phenotype is associated with increased succinate concentration. This inhibition was achieved using diethyl butylmalonate (BM), an inhibitor of SLC25A10, the succinate transporter in the IMM (Mills *et al.*, 2018). As hypothesized, addition of BM in addition to SA prevents the acquisition of protein succinylation, to a certain extent (**Fig 27B**). This decrease in global lysine succinylation

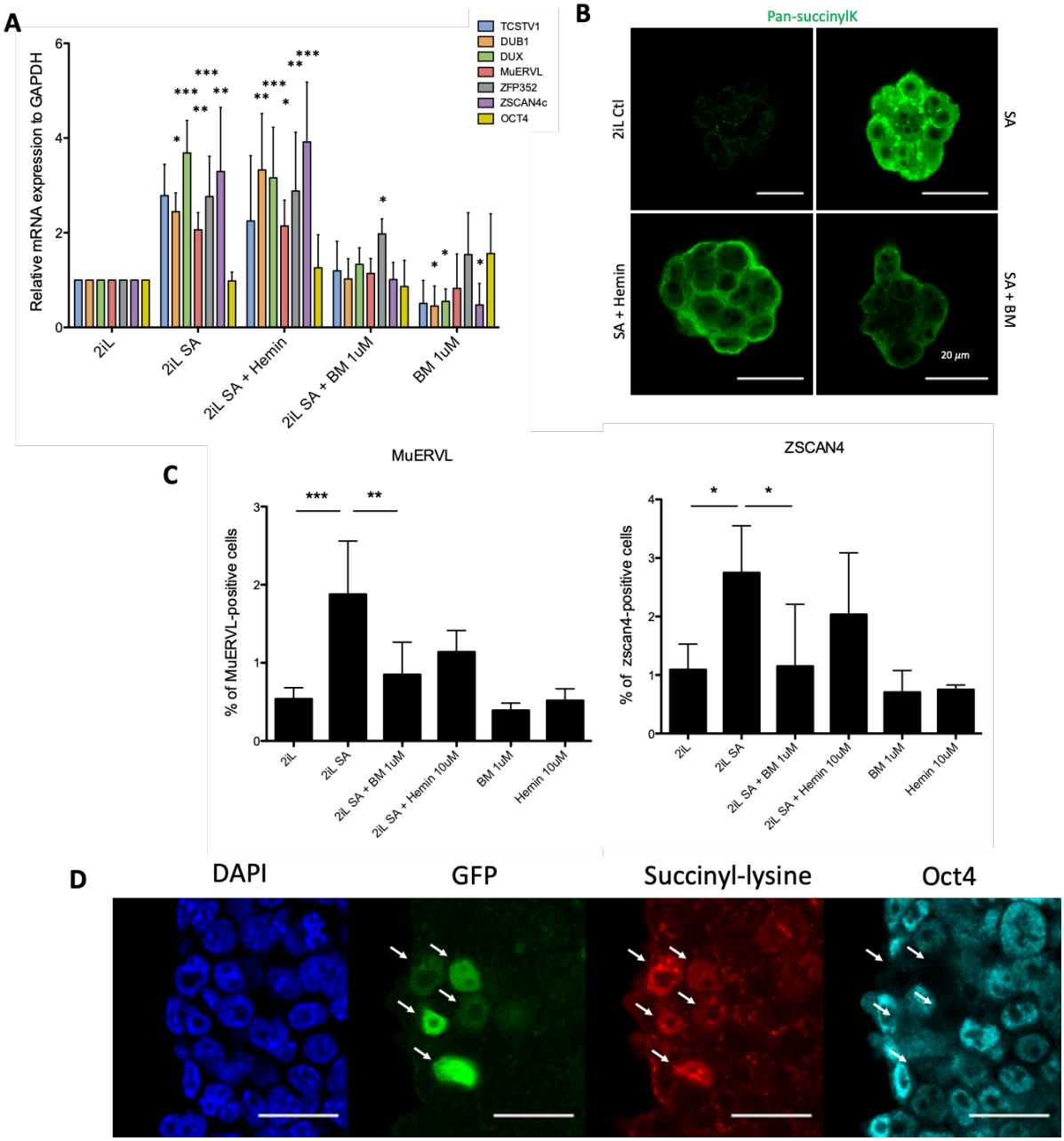


Figure 27 mESC "2C-like" reprograming by SA is due to defective protein succinylation. A) *Relative expression of 2C markers of mESCs assessed by RT-qPCR relative to GAPDH expression and to 2iL naïve control, in mESCs treated with 0.5 mM SA (2iL SA), with or without 10 µM Hemin (2iL SA + Hemin), 1 µM diethyl butylmalonate (2iL SA + BM 1µM) or BM alone (BM 1µM). S.D * p < 0.05, **p < 0.01, ***p < 0.001. ANOVA-1. n=4 independent biological replicates.* ***B)*** *Immunostaining of succinylated lysines (green) in mESCs treated with SA, with or without 10µM Hemin (SA + Hemin), 1uM diethyl butylmalonate (SA + BM 1µM). Representative image of n=3 independent experiments. Scale bar= 20µm.* ***C)*** *Percentage of MUERVL- or ZSCAN4-positive cells in the whole population of naïve (2iL) mESCs or naïve treated with SA (2iL SA) with or without 10 µM hemin (2iL SA + hemin 10 µM), diethyl butylmalonate (2iL SA + BM 1 µM). N=4 independent biological replicates. Results expressed as mean +/- S.D. * p < 0.05; ANOVA-1.* ***D)*** *Immunostaining of succinyl-lysines (Red) and Oct4 (cyan) of TBG4 cells (ES-E14TG2a mESCs with a 2C-GFP reporter construct). Representative image of n=3 independent experiments. Scale bar= 20µm.*

is correlated to a rescue of both the increase in 2C markers and the proportion of ZSCAN4 or MUERVL-positive cells in the population (**Fig. 27A-C**). Together, this shows that an accumulation of succinate in naïve mESCs induces a 2C-like phenotype, and provokes an increase in protein succinylation in all cell compartments. In order to identify the endogenous 2CLC and simultaneously observe the levels of protein succinylation in the mESC population, we took advantage of a reporter cell line for this 2C-state, via stable insertion of a construct

containing an EGFP-coding gene under control of the MERVL long terminal repeat (2C:::EGFP) previously described (Ishiuchi *et al.*, 2015; Rodriguez-Terrones *et al.*, 2020). This would allow to combine the endogenous fluorescence with the immunostaining of the succinyllysines. These observations showed an increase in protein succinylation specifically in the population of 2C-like cells (**Fig 27D**).

Since we showed that the increase in succinate exit from mitochondria triggers an increase in the reprogramming of mESCs to a 2C-like state, probably by increasing the protein succinylation, we decided to identify the protein differentially succinylated after SA treatment. To this end we immunoprecipitated the succinylated peptides in control 2iL or SA treated mESCs after trypsin digestion, followed by mass spectrometry analysis, as previously described (Guo *et al.*, 2020). This experiment, performed in biological duplicates, identified 426 and 482 proteins in the first replicate and 1064 and 1099 in the second, for 2iL and 2iL+SA respectively. These identifications and quantifications can be found in Supplementary table 1. Out of the proteins identified and overrepresented in the SA-treated condition, our attention was drawn on the enrichment of both nucleolin and TRIM28-succinylated peptides (**Fig. 28A**). These two proteins have been shown to be associated with a LINE1 (retrotransposon long interspersed element 1) transposable element in the repression of the expression of Dux, a master regulator of the 2C-like stage (**Fig. 12**). We postulate that an increase in succinylation of these proteins could thus regulate the association of the complex, in turn relieving the inhibition of Dux. We thus performed a co-immunoprecipitation (Co-IP) of nucleolin and TRIM28. Surprisingly, treatment of mESCs with SA increases the binding of TRIM28 on nucleolin, as opposed to what was expected (**Fig 28B**).

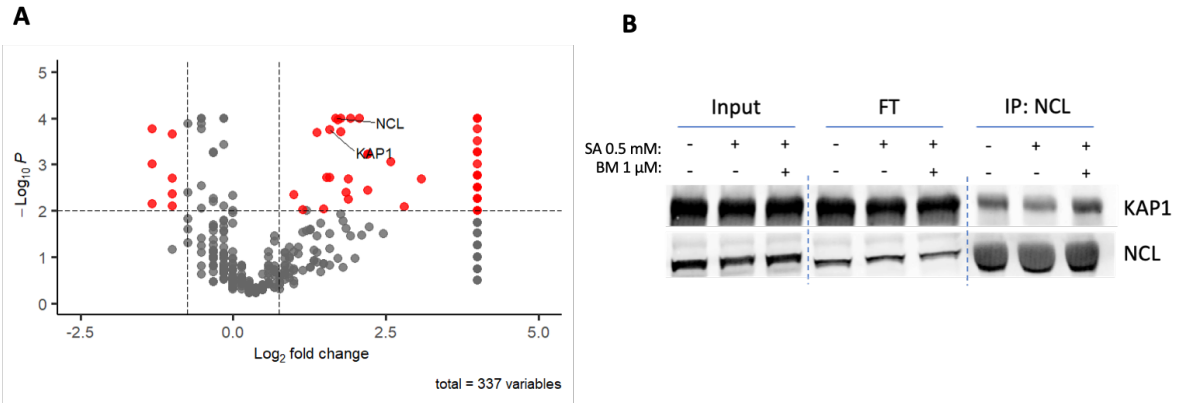


Figure 28 ***Succinylation of the Nucleolin-TRIM28 complex might be responsible for the 2C-like reprogramming. A)*** *Volcano plot of succinylated proteins identified in mass spectrometry after immunoprecipitation of succinylated peptides. NCL= nucleolin, KAP1=TRIM28= Tripartite motif containing 28. n=2 biological replicates.* ***B)*** *Western blot analysis of KAP1 after co-immunoprecipitation of Nucleolin in naïve mESCs (2iL) with or without heme synthesis inhibitor (SA) 0.5mM and butylmalonate (BM) 1µM. FT= flow-through fraction, IP= pull-down fraction, 10 µg were loaded on the gel. n=3.*

Discussion and perspectives

In the end, this project is made of two distinct part. We first investigated the heme synthesis pathway and its inhibition in the context of the naïve-to-primed ESC transition. Since we have not succeeded in deciphering the mechanisms in place in the context of this heme dependency, many leads remain to explain the phenotype. Second, the heme synthesis inhibition was also able to stimulate the apparition of 2CLCs in the population of naïve cells, an effect that was surprisingly heme-independent. Because the two processes are either heme-dependent in the transition or heme-independent in the 2CLCs, they will be discussed separately before a more global discussion at the end.

A. Identification of the developmental stage

While this project reports interesting finding about the maintenance of ESCs and the progression of development between pre- and post-implantation blastocysts, many questions remain unanswered. First, we demonstrated that the exit from the naïve 2iL state of mESC under heme synthesis inhibition is impaired, shown by gene and protein expression, transcription factor localization and whole cell RNA sequencing. However, while the SA-induced transition blockade is clear, the identity of the EPI + SA cells remains unclear. We could compare the gene expression of the 2iL, EPI and EPI + SA cells grown in this study with published RNA sequencing data for in-vivo epiblast stem cells at different time points (E4.5, E5.5, E6.5) (Nakamura *et al.*, 2016) or with stabilized cultures of EpiSC or formative cells (Kinoshita *et al.*, 2020). Using these data sets might help to more precisely identify some sort of temporal localization of the SA-treated cells in the developing blastocyst and in the pluripotency continuum. It is also essential to consider the experimental timing used in this study: 2 days of transition. This short timing is usually considered to reside in an early post-implantation stage called epiblast-like cells (EpiLCs), having left the naïve stage and on the path to the true EpiSC stage. Developmentally, this EpiLC stage resembles the stabilized formative stage. This consideration is important to identify the developmental point where SA actually acts as a blocking signal. Now that the formative stage has been stabilized, we could use their "formative-only" set of genes to measure the state of this transition. They report expression of genes such as Fbxo41, Pou2f2 or Slc45a1 only in the formative population and if the SA treatment doesn't prevent their acquisition, we could postulate a blockade in a formative state.

B. Signaling pathways

Regardless of the developmental stage reached by the EPI + SA cells, the mechanisms leading to this incapacity to properly transition haven't been exactly pinpointed yet. We postulate here that the inability to properly activate both the MAPK and the TGFβ pathways in response to FGF2 and activin A (**Fig. 24**), respectively, might be the cause of the observed phenotype especially since the majority of the genes separating the PC2 in **Fig. 21** are related to those pathways. Such effects from the blockage of these pathways have been previously described in the literature (Eiselleova *et al.*, 2009; Huang *et al.*, 2009; Lanner and Rossant, 2010; Hamilton and Brickman, 2014; Senft *et al.*, 2018; Lee, Park and Jung, 2019). Although

interesting, this hypothesis requires further testing. Technically, we could transition to the primed stage in the presence of inhibitors of both pathways and compare the RNAseq signature with the EPI+SA condition. If the hypothesis is correct, the dual MAPK-SMAD inhibition during the transition should present a phenotype similar to SA. For the inhibition of the MAPK, a MEK inhibitor such as the PD0325901, already used in the 2iL culture, could be added to the transition medium as well. The TGFβ/activin A pathway inhibition could be achieved with an activin A receptor inhibitor such as SB431542 or even with a SMAD3 phosphorylation inhibitor SIS3 (Jinnin, Ihn and Tamaki, 2006).

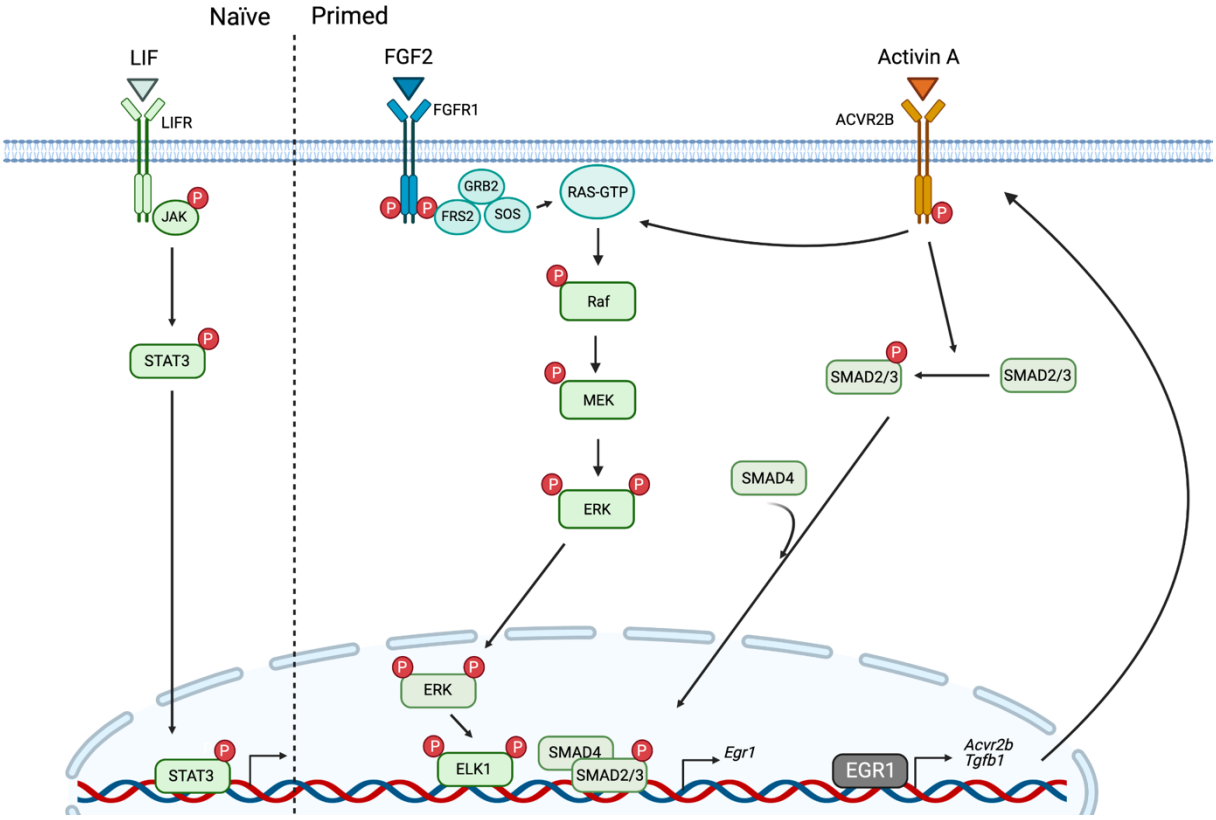
Unexpectedly, preliminary results obtained with the MEK inhibitor alone during the transition suggest that MEK inhibition does not impair the transition nor mimics the Epi+SA phenotype. This could be due to the fact that a combination of the inhibition of both pathways would be necessary to achieve the transition inhibition. Interestingly, the inhibition of SMAD signaling through a $SMAD2^{-/-}$/$SMAD3^{-/-}$ KO (dKO) mESC line shows that, in a setup comparable to the transition used in this study, ESCs cannot transition to the EpiLC state without proper SMAD signaling (Senft *et al.*, 2018). Excitingly, the top genes most downregulated in the Epi dKO versus the Epi WT are the same than in Epi SA vs Epi from this study: *Lefty1, Lefty2, Pitx2,* all part of the Nodal signaling. This further encourages us to investigate the SMAD2/3 pathway in the context of heme inhibition and naïve-to-primed transition. A rigorous comparison of the published RNAseq data from the SMAD study (Senft *et al.*, 2018) with our own data set would maybe highlight similarities between Epi dKO and Epi SA gene expression signature.

1. MAPK

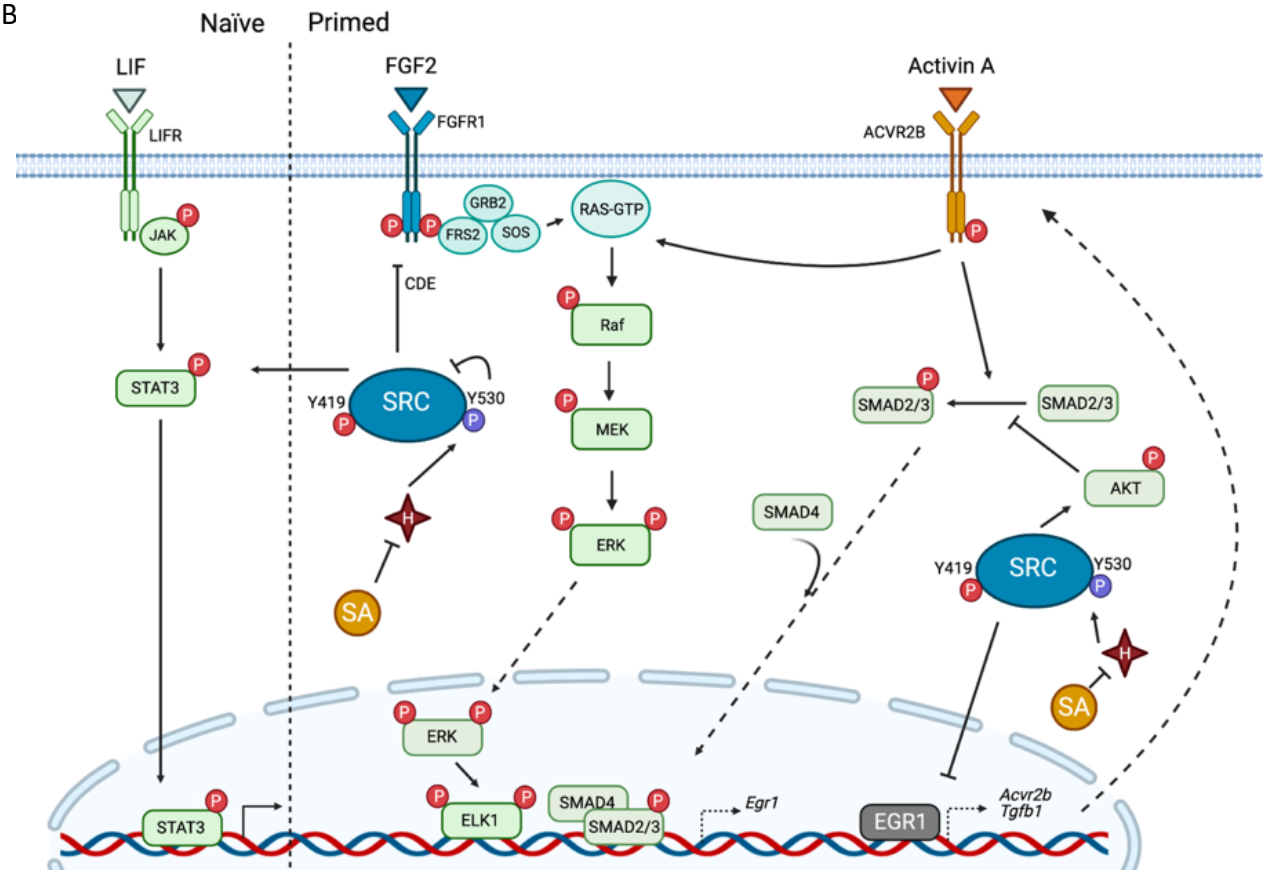
Identifying the pathway playing the key role in this transition impairment would somewhat narrow down the search for the missing link between heme and the observed phenotype. We propose a hypothetic model of the role of heme on signaling pathways in the transition in **figure 29**. Most of the research performed to understand the role of heme in cellular biology focus on either the BACH1 transcriptional activity or on the HRI/ISR activation, both hypotheses that we mostly ruled out. To our knowledge, only a single paper highlights a defective MAPK activation under heme deprivation. The team of Y. Zhu and coworkers (Zhu *et al.*, 2002) showed that PC12 neuronal cells treated with SA lose the ability to respond to nerve growth factor (NGF), leading to a defect in ERK1/2 phosphorylation. However, they did not demonstrate the involved mechanism. Since NGF signals activate the MAPK pathway through the TrkA (Tropomyosin Receptor Kinase A) receptor, we could hypothesize that the effect of heme deprivation on the MAPK pathway would act through a common target between the NGF-TrkA-ERK axis and the FGF2-FGFR2-ERK axis. Such proteins could be FRS2 (Fibroblast Growth Factor Receptor Substrate 2), GRB2 (Growth factor receptor-bound protein 2), SOS (Son of Sevenless), RAF (Rapidly Accelerated Fibrosarcoma), RAS (proto-oncogene protein p21) or even MEK. Since the activation of these proteins is signaled by a phosphorylation cascade, monitoring the phosphorylation levels of these intermediates in the pathway could inform at which level heme deprivation does play a role.

Signaling pathways are extremely interconnected and these connections are also dependent on the cell type and growth conditions. Connected to the MAPK pathway is also the SRC (Proto-oncogene tyrosine-protein kinase Src (sarcoma)) pathway. Activated by integrins, tyrosine kinase receptors or GPCRs, SRC can phosphorylate RAS, triggering the

A



B



(See legend of figure 29 on the next page)

cascade leading to ERK1/2 phosphorylation (Scapoli *et al.*, 2004; Lopez *et al.*, 2012).

Figure 29 Proposed hypothetical mechanism of SA-mediated interference with cell signaling crucial for the naïve-to-primed ESC transition*. A) Signaling pathways activated in naïve or primed ESCs with their respective cytokines LIF and FGF2/Activin A. LIF signaling in naïve cells induces the phosphorylation of JAK, then of STAT3 leading to its nuclear translocation and transcriptional activity. In primed cells FGF2 binds to the FGFR1, inducing its phosphorylation activating the MAPK phosphorylation cascade: FRS2, GRB2, SOS, RAS, RAF, MEK, and ERK. Phosphorylated ERK1/2 translocates into the nucleus activating transcription factors such as ELK1. Activin A, binding to the activin receptor ACVR2B, activates the SMAD2/3 proteins by phosphorylation, inducing the recruitment of SMAD4 ad their translocation to the nucleus, activating the transcription of target genes. ELK1 and SMAD2/3 can activate the transcription of Egr1, a transcription factor mediating the expression of TGFβ pathway members Acvr2b and Tgfb1 in a positive feedback loop. B) Proposed mechanism for the impact of SA on the different pathways. Binding of heme on SRC induces its phosphorylation on the 530 tyrosine residue, inhibiting the protein activity by a change in conformation. In absence of heme, by SA treatment, SRC is thus in an active conformation with a phosphorylated 419 tyrosine residue. This activation has been shown to phosphorylate STAT3 and induce the clathrin-dependent FGFR1 endocytosis (CDE) reducing the MAPK activation. Activated SRC is also known to phosphorylate AKT on the tyrosine 473 residue, in turn inhibiting the phosphorylation of SMAD2/3. Dashed arrows represent modifications in SA-treated Epi cells that have been shown in this work. LIF= leukemia inhibitory factor, LIFR= LIF receptor, JAK= Janus kinase, STAT3= Signal transducer and activator of transcription 3, FGF2= fibroblast growth factor 2, FGFR1= FGF receptor 1, MAPK= mitogen activated protein kinases, FRS2= Fibroblast Growth Factor Receptor Substrate 2, GRB2= Growth factor receptor-bound protein 2, SOS= Son of sevenless, RAS= proto-oncogene protein p21, RAF= Rapidly Accelerated Fibrosarcoma, MEK= Mitogen-activated protein kinase kinase, ERK= extracellular signal-regulated kinases, ELK1= ETS domain-containing protein Elk-1, ACVR2B= Activin receptor type-2B, SMAD= mothers against decapentaplegic, AKT= protein kinase B, SRC= Proto-oncogene tyrosine-protein kinase SRC, SA= succinylacetone, H= heme, P= phosphorylation residue (activating in red, inhibitory in purple).*

Interestingly, SRC has been shown to bind heme, an interaction that modulates the activation of the kinase (Yao *et al.*, 2010). Indeed, heme-binding CP motifs (cysteine-proline), also found in other heme-responsive proteins such as BACH1 or HRI, have been found on SRC. Direct binding of heme to these CP motifs induces phosphorylation of the Tyr530 residue, inducing a change in the protein conformation, inhibiting its activity (Sen and Johnson, 2011). Furthermore, the expression of the kinase mediating this phosphorylation, Csk (C-Src Kinase), is downregulated almost 4 folds in Epi+SA compared to Epi mESCs. While the phosphorylation of this residue should be assessed in SA-treated mESCs, we could postulate that heme deprivation would reduce this inhibitory phosphorylation, thereby activating SRC. Interestingly, the activation of SRC has been shown to trigger the endocytosis of FGF receptors in a clathrin-mediated mechanism (Sandilands *et al.*, 2007; Auciello *et al.*, 2013) leading to its targeting to the lysosome for degradation (Haugsten *et al.*, 2005; Porębska *et al.*, 2018). The clathrin dependent endocytosis (CDE) has been implicated in the maintenance of the pluripotent phenotype (Narayana *et al.*, 2019) and in the exit from the naïve state by regulating the intracellular ERK activity (De Belly *et al.*, 2021). Together, a reduction in heme content could activate SRC by decreasing the phosphorylation of an inhibitory residue, in turn able to promote FGFR internalization and to reduce the response to FGF2.

With the context of the naïve and primed ESCs and their respective activated pathways in mind, we could also link the transition defect to the action of SRC on a sustained activation of the STAT3 transcription factor (Turkson *et al.*, 1998; Garcia *et al.*, 2001). Since in naïve cells, LIF activates STAT3 through the JAK kinase, the hypothetical SA-mediated activation of SRC could maintain STAT3 activity even in absence of LIF, thereby impairing the transition. This could easily be assessed by western blotting of the phosphorylated form of STAT3.

2. Activin A – SMADs

On the other hand, the defect of activation of the activin A/SMAD pathway is also intriguing. To our knowledge, heme is not reported to exert a direct action on intermediates of the pathway. However, by digging into the RNAseq data, the expression of one of the

receptors for activin, ACVR2B, is strongly downregulated (3 folds) in SA-treated Epi cells. The protein abundance of this receptor would need to be monitored to explore the possibility of a decrease in the response to activin A due to a decrease of abundance of the receptor. Exploring the transcription factors driving the expression of *Acvr2b* shows that the promoter of this gene in controlled by several proteins, among which EGR-1 (Early growth response 1) that caught our interest. Indeed, heme has been shown to stimulate the transcription of *Egr1* notably through ERK1/2 activation (Hasan and Schafer, 2008; Gotoh *et al.*, 2011). This would explain the 4-fold decrease of *Egr1* transcript level in SA-treated cells observed in RNAseq. Further confirming a potential role of this transcription factor in the transition defect observed with SA, the top genes contributing to the PC2 separation in **Fig. 21** are either dependent of the transcriptional activity of EGR-1 (targeting *Lef1, Fgf5, Pou3f1* or *Wt1*) or SMADs (*Lefty1/2, Ina, Vrtn* or *Bcl11a*). Furthermore, SMAD3 and EGR-1 have been shown to physically interact, regulating the expression of several target genes (Fortin and Bernard, 2010) and EGR-1 could also promote the activation of the TGFβ pathway by the regulation of transcription of *Tgfb1* (Baron *et al.*, 2006). This TGF-β1 is one of the critical component required for the growth of primed ESCs (Tesar *et al.*, 2007; Gafni *et al.*, 2013a). Since EGR-1 expression is also positively regulated by the MAPK-ERK1/2 pathway (Hasan and Schafer, 2008), the effects of heme on the MAPK cascade could also, in turn, reduce EGR-1 expression followed by a decrease in the activin A response. SRC inhibition has also been shown to induce *Egr1* expression, further linking this kinase to the response observed in SA-treated ESCs since the potential activation of SRC by a reduction in heme concentration would thus reduce *Egr1* expression (Jones *et al.*, 2009). Finally, SRC has been shown to indirectly modulate the phosphorylation of SMADs, and their subsequent activation and translocation, by phosphorylating the AKT kinase (Ser473), in turn inhibiting the phosphorylation of SMAD2/3 (Kathiriya *et al.*, 2014).

Since the literature highlights a key role of SRC on the regulation of the critical pathways MAPK, SMAD and STAT3 during the transition, it would be crucial to monitor its activation, through immunoblotting of its phosphorylated Tyr416, in Epi + SA conditions. If shown activated, SRC inhibition with WH-4-023 might very well rescue the phenotype induced by SA.

Altogether these hypotheses pinpoint SRC as a heme-binding protein in the regulation of MAPK, through the clathrin-dependent endocytosis of the FGFR, aberrant STAT3 activation in the transition and/or inhibition of the activin A-SMAD pathway through the Akt kinase. There is also a possible link with SRC and EGR-1, both regulating SMAD2/3 activation, among others. EGR-1 is also shown to be controlled by both ERK1/2 and SMAD3 (Fortin and Bernard, 2010; Gregg and Fraizer, 2011; Hartney *et al.*, 2011). This makes the SRC kinase the next priority in understanding the SA-induced transition defect. So far there is no reports on the involvement of EGR-1 in the transition to the primed stage. This transcription factor was however among the top transcription factors controlling the differentially expressed genes in FLCN KO hESCs pushed for 7 days to the primed stage compared to the WT, a conditions shown to resist the conversion to the primed stage (Mathieu *et al.*, 2019). SRC however, has been investigated in the naïve-to-primed transition before. It is for example the target of a small molecule used on primed cells to reprogram them into the naïve state, mostly for its role on the inhibition of the MAPK pathway (Theunissen *et al.*, 2014).

3. Other pathways

Although a role for heme in the regulation of these signaling pathways crucial for the transition is appealing, other pathways should not be ruled out. For example, despite numerous reports of BACH-1 regulation by heme (Ogawa *et al.*, 2001; Suzuki *et al.*, 2004; Zhang *et al.*, 2018), we did not find any significant difference in the nuclear localization of this transcription factor in SA-treated cells (**Fig. 22**). This could either be a true effect but it could also be an artifact due to the antibody. Indeed, while attempting to validate by western blot analysis the results observed by immunofluorescence, multiple bands were detected, suggesting that the antibody, present in the lab for 10 years now and whose production has been interrupted since then, has lost sensitivity. Validating the results obtained with another validated antibody, both in IF and in WB would strengthen our claim.

Then, we couldn't fully implicate the heme-HRI-EIF2α axis in the effect of SA on the transition to the primed stage. While the activation of the pathway is clear (**Fig. 23**), its selective activation with BTdCPU is not bringing about the same response than SA, despite having a similar effect on protein translation. We cannot rule out that this pathway contributes to the transition blockade, in cooperation with another effector. To confirm that the activation of HRI is indeed not sufficient to recapitulate the transition defect induced by heme, a KO of *Eif2ak1* (HRI) would be interesting. Indeed, if HRI KO mESCs are pushed in the transition while treated with SA, they should still be blocked in the same way as WT cells. If the role of HRI activation in the process is a key step, then that would alleviate the transition defect and allow the conversion to the primed stage.

C. 2C reprogramming

In the second part of this work, we have shown the heme synthesis inhibition in naïve 2iL mESCs was increasing the number of cells cycling in the 2C-like state. As opposed to the defect in the naïve-to-primed transition, this effect was not rescued by the exogenous addition of hemin but, instead, was linked to an accumulation of succinate in the mitochondria and its leakage out of the organelle since pharmacological inhibition of SLC25A10, a succinate transporter, rescued that effect. The diethyl butylmalonate used to block this channel has however other roles in the cellular functions such as a decrease in gluconeogenesis or the inhibition of the aminoacylase I (Williamson, Anderson and Browning, 1970; Röhm, 1989). To bypass these potential off targets, genetic ablation of the *Slc25a10* gene would allow us to answer with precision if the mitochondrial exit of succinate is the cause of the increase in 2CLCs.

We then showed that the increase in extra-mitochondrial succinate concentration increased the succinylation of proteins. Among the differentially succinylated proteins, we suspect nucleolin and TRIM28 to mediate the increase of 2CLCs. Indeed, a recent report showed that these two proteins, together with LINE1 (**Fig. 12**) act together to repress the expression of *Dux*, a master transcriptional regulator of the 2C transcriptional program (Percharde *et al.*, 2018). We are currently pursuing the identification of the role of these two proteins in our model. First, we hypothesized that their succinylation would impair their complex assembly. Nucleolin was found succinylated on 10 residues in the control conditions but when the cells are treated with SA, a total of 30 residues are modified. TRIM28 was only

succinylated on 5 residues in basal conditions, that increased to 12 when treated with SA. However, since the complete structure of NCL hasn't be fully defined yet and thus the contact region between NCL and TRIM28 is not known either, it is currently not possible to correctly map if their binding would be impaired by the addition of the succinyl groups on lysine residues. We thus performed a Co-IP of these two proteins to assess their interaction. Surprisingly, the preliminary results suggest that the binding of NCL and TRIM28 is increased for about 2 folds. Regardless, since we actually observed an increase in the expression of Dux after treatment with SA, the role of succinylation in the process could affect other aspects of this interaction such as the interaction with LINE1.

The study reporting the role of this complex on the expression of *Dux* showed that the binding of the two proteins to the LINE1 retrotransposon was necessary to repress the expression of *Dux*. We could thus postulate that if the effect of succinylation on the complex is not due to the interaction between the two protein components, it might be due to the disruption of the NCL-LINE1 binding. This hypothesis is supported by the fact that the succinylated residues on NCL after SA treatment are located in the RNA binding region of the protein. To test this hypothesis, we could perform an RNA-immunoprecipitation to monitor the presence of LINE1 in the complex. In addition, because of the negative charge of the succinyl residue, and the negative charges of DNA, the actual binding of the complex to the DNA regulatory regions could also be compromised. This DNA binding activity can be measured with a DNA pull down assay followed by protein identification by mass spectrometry. We are thus planning to probe the regulatory regions of Dux identified in the original study (Percharde *et al.*, 2018) to assess the DNA binding activity of NCL.

The mechanism of regulation of Dux expression by the whole complex is so far unknown. We could postulate that, thanks to its DNA-binding domain (Dickinson and Kohwi-Shigematsu, 1995; Samuel *et al.*, 2008), NCL is the protein mediating the binding of the complex to the *Dux* promoter region, while TRIM28 would mediate the repressive action through the recruitment of the NuRD (nucleosome remodeling and deacetylase) complex or SETDB1 (SET Domain Bifurcated Histone Lysine Methyltransferase 1) (Iyengar and Farnham, 2011; Cheng, 2014). These two actors are known to mediate the formation of heterochromatin leading to repression of gene expression (Iyengar and Farnham, 2011). The recruitment of the NuRD complex and SETDB1 is mediated by the SUMOylation of three lysine residues in the C-terminal region of the protein (K554, K779 and K804) (Ivanov *et al.*, 2007). Interestingly, among these crucial residues, the K779 is specifically succinylated in SA-treated cells. We could thus postulate that the competition of succinylation and SUMOylation for this residue would impair the ability of TRIM28 to recruit the repressive complexes.

The dramatic effects of SA on the global protein succinylation observed via the immunostaining of succinylated lysines is also intriguing. One could postulate that the intracellular metabolism of SA itself could lead to the production of extra-mitochondrial succinate or to the production of a reactive SA-derived metabolite, leading to protein succinylation. The intracellular fate of SA is not well described but clinical studies of patients with hereditary tyrosinemia, leading to SA production and heme synthesis deficiency, report high concentration of the metabolite in urine, suggesting excretion of this toxic intermediate (Christensen *et al.*, 1981). However, *in vitro* studies also show a potential for SA oxidation by peroxynitrites, leading to the formation of either acetate or succinate depending on the

presence of O_2 (Royer *et al.*, 2004). However, if the production of these metabolites following SA degradation was actually leading to extramitochondrial protein succinylation, we shouldn't have observed a rescue of the phenotype by SLC25A10 inhibition. Alternatively, one can still hypothesize that such degradation would occur inside the mitochondrial matrix to produce succinate or that SA derived-product would exit mitochondria through the same path. So far this cannot be excluded but could be assessed by using another inhibitor of the heme synthesis pathway (Shetty and Corson, 2020), or even genetic ablation of one of the enzyme of the synthesis pathway.

Other than a role in protein modification, succinate has also other cellular effects that could explain the apparition of the 2CLCs. As described in the introduction, an increase in succinate concentration is known to inhibit the activity of dioxygenase enzymes (Laukka *et al.*, 2016; Tretter, Patocs and Chinopoulos, 2016). Among this family of enzymes, the TET DNA-demethylase enzymes would be particularly interesting. Indeed, it has been shown that the loss of TET enzymes is able to promote the acquisition of 2CLC in the mESC population (Lu *et al.*, 2014; Schüle *et al.*, 2019). In accordance to our experimental conditions, the treatment with SA increases the concentration of succinate outside of mitochondria. In turn, this could reduce the activity of TET enzymes, leading to the observed phenotype. It would be interesting to investigate if this effect takes place in the presence of SA by first measuring the increase in succinate in the cells and follow with the measure of activity of the TET enzymes, using a fluorometric assay for example (Shen and Zhang, 2012). If the enzymatic activity is actually down, monitoring the DNA methylation levels with bisulfite sequencing would be required to verify the hypothesis. On the other hand, the increase in succinate concentration would also inhibit the PHD enzymes responsible for HIF1α degradation. This way, the increase of succinate observed in SA-treated cells would thus increase the stability of HIF1α and its transcriptional activity. However, this would be in disagreement with the work of Macfarlan in 2012 showing that culture of mESCs in hypoxia actually reduces drastically the number of 2C::Tomato+ cells in the population (Macfarlan *et al.*, 2012).

To further validate that the increase of succinate concentration is responsible for the increase in 2CLCs, it would be interesting to raise succinate levels by an alternative way than heme synthesis inhibition. To this end, and as reported by the team of James Maher (Smestad *et al.*, 2018), we could inhibit the SDH complex activity, in turn leading to the accumulation of its substrate, succinate. This could be achieved either by the generation of SDH KO mESCs, or by chemically inhibiting the SDH complex with Atpenin A5, for example (Miyadera *et al.*, 2003). If the proposed model is correct, the loss of SDH activity should result in succinate accumulation and exit from the mitochondria, leading to an increase in 2CLCs. These effects would be rescued by butylmalonate. Since it also seems that the increase of protein succinylation occurs specifically in 2CLCs in culture (**Fig. 27**), it would be interesting to verify if this is also observed *in vivo* by immunostaining of mouse embryos from various stages (from 2C to blastocyst).

The original work on 2CLCs showed that this population of cells is highly dynamic, cycling in and out of this state during cell culture (Macfarlan *et al.*, 2012). Regardless of the process that we implicate in the phenotype, it would be interesting to monitor at which levels it plays out: it could increase the number of cells transitioning to the 2C-stage without changing the kinetics or it could slow down the 2C-to-mESC reversion, accumulating 2CLC in

the population. To answer this, we could take advantage of the 2C::GFP mESCs line to follow in live imaging the acquisition of GFP+ cells over time.

D. General considerations

In the big picture, this project aims to understand what regulates the implantation of blastocyst in the uterus. Indeed, it is necessary to understand this process as the implantation is the cause for the failure to establish a successful pregnancy for about 50% of the fertilized eggs (Wilcox *et al.*, 2020). While direct studies on human tissues and samples is extremely limited due to obvious ethical reasons, we propose here to bypass this in two ways: first, we replace the use of *in vivo* sampling and experimentation by an *in vitro* model consisting in the transition of embryonic stem cells from the naïve pre-implantation stage to the primed post-implantation stage. While this definitely doesn't recapitulate the whole embryo, with its surrounding tissues or the microenvironment, studying the modifications that are taking place in the ESCs, that reconstitute the whole individual, allow to catch a glimpse of the mechanisms that take place *in vivo.* Second, we decided to tackle the question of the requirements for this implantation step by using a mouse model. While using hESCs would have been ideal to study this process, we instead switched to mESCs since they share a lot of similarities like the metabolic switch or the reorganization of the epigenetic landscape.

1. Human – Mouse differences

Differences are however found in the signaling requirements for naïve stage maintenance and in the expression of some genes during the transition. For this reason, one of the first perspectives of this study, regardless of the highlighted mechanisms, will be to translate the mESCs results on hESCs. Despite having handled hESCs previously, setting up their culture in URBC happened to be more challenging than originally thought. Instead of wasting precious months on the resolution of these multiple challenges, we switched to the mESC culture that is way simpler. As described in the introduction, the stabilization of a naïve hESC stage was only achieved recently (Gafni *et al.*, 2013b; Takashima *et al.*, 2014; Theunissen *et al.*, 2014; Ware *et al.*, 2014) and so far none of the described protocols received unanimous support from the scientific community as "THE" naïve protocol. The most recognized protocol to date might be the one from R. Jaenisch group at the MIT, using the 5iLA cocktail. It would thus be interesting to test whether heme synthesis inhibition during the exit from the 5iLA state, by incubation with FGF2 or FGF2 and activin A, would also block the transition to a primed stage. The transition of hESCs is slower than in mESCs so we would need to study later time points than 48h in mESCs. Typically, 4 to 7 days are used to reach a stage primed enough and that would actually represent the duration of the transition performed in the CRISPR screen from which we based this study (Mathieu *et al.*, 2019).

Similarly to the transition in mESCs that was performed in this study, we can measure the transition of hESCs according to the gene expression and the concomitant protein abundance, but also with the TFE3 subcellular localization. In humans, however, the subset of genes that would be monitored slightly differs. Naïve hESCs express genes such as *TFCP2L1, KLF4, KLF17, DNMT3L* or *NNMT* while primed cells express *IDO1, OTX2* or *SOX11* (Sperber *et al.*, 2015; Mathieu *et al.*, 2019; Kinoshita *et al.*, 2020). Some primed hESC culture have also displayed the expression of early differentiation markers such as *FOXA2* or *TBXT* (T-box

transcription factor T; Brachyury) (Kinoshita *et al.*, 2020). Highlighting a mechanism for the naïve-to-primed ESC transition is both human and mouse would definitely stress the importance of the pathway in the process.

2. Generation of KO lines

So far, we demonstrated that the inhibition of heme synthesis in mESCs impairs the transition, by using a chemical inhibitor of the pathway. Ideally, the generation of a mESC (and/or hESC) cell line KO for one of the enzymes of the pathway would strengthen the results obtained. To this end, we plan to generate an ALAD KO mESC cell line with the CRISPR-Cas9 technique. Targeting the same enzyme that is the target of SA would confirm our results more elegantly than with an inhibitor. To date, the most efficient way of generating this KO would consist in a nucleofection of a CAS9-gRNA complex into the cells with the AMAXA nucleofector technology. We have previously shown that naïve cells are receptive to this technique (Moody *et al.*, 2017; Mathieu *et al.*, 2019). However, the risk is that a deletion of the gene would be lethal over a few generations seeing the drastic effects of heme synthesis inhibition on either protein translation or on the ETC. To bypass this, I suggest either the construction of a cell line expressing a GFP-tagged ALAD protein under the control of an inducible cassette, doxycycline (Dox) responsive. Knowing the tendency of PSC to silence exogenous DNA sequences, this cassette has to be inserted in the AAVS1 (Adeno-Associated Virus Integration Site 1) safe-harbor site in human (Smith *et al.*, 2008) or in the Rosa26 locus in mouse (Soriano, 1999). By then targeting the host ALAD gene with CRISPR-Cas9 while inducing the expression of the cassette with Dox, we would achieve an ALAD KO while maintaining the expression of the chimeric ALAD-GFP to maintain the production of heme. Of course, the ALAD-GFP construct inserted would have to be codon-optimized to be protected from the Cas9 recognition. While this approach offers the capacity to maintain ALAD KO cells under Dox supplementation, removal of Dox to stop the production of the recombinant protein might take a while to affect the cellular functions, depending on the protein stability. To our knowledge, the ALAD protein half-life is not known. The only available data shows that the half-life of the *Alad* mRNA in *Chlamydomonas reinhardtii*, a green algae, is about 51h, which would presumably offer a long stability in mammals (Matters and Beale, 1995). Since hemin supplementation was shown to salvage the heme synthesis inhibition, to bypass the potential bottlenecks of protein stability and inducible constructs and to simplify the process, we could supplement the media of ALAD KO cells with hemin to maintain sufficient heme intracellular levels. This way, removal of heme from the media would correspond to an induction of an ALAD KO.

3. *In vitro* models

While the use of this naïve-to-primed ESC transition indeed recapitulates the implantation step of the embryo, it dramatically simplifies it. Overall, it has been possible to develop embryos *in vitro* to recapitulate these early developmental stages, reviewed in (Shahbazi, Siggia and Zernicka-Goetz, 2019). Even very recently, the team of J. Hanna elegantly showed the development of the blastocyst even past the gastrulation stage, allowing the dissection of early events leading to the organogenesis (Aguilera-Castrejon *et al.*, 2021). Alternatively, the generation of blastocyst-like structures by self-organisation of ESCs render possible the study of these early developmental stages (Deglincerti *et al.*, 2016; Shahbazi *et al.*, 2016). Using such models in the context of heme synthesis inhibition would allow a

resolution in term of whole embryo, to dissect the role of heme in the developmental processes.

Conclusions

This research project was first supposed to interrogate the actors required for a proper implantation of embryos in the uterus, represented by the naïve-to-primed transition *in vitro.* This developmental step is one of the most crucial in the establishment of a successful pregnancy and thus understanding the molecular details governing it is fundamental. In this context of pluripotency, it was intriguing to highlight the heme biosynthetic pathway as crucial for this step. To our knowledge, it is the first report presenting a role for this pathway in embryonic stem cells. After pursuing the investigation for the role of this pathway on the cellular biology of ESCs, it seems the generation of heme serves several aspects of cell signaling that are crucial for the progression of the embryo development.

According to the results we presented in this work, heme seems to be more of a passenger than a driver in this process, a passenger required for the establishment of proper cellular signals, such as the MAPK or the TGFβ pathway. An appropriate response to signals activating these pathways is indeed crucial for the development of the embryo, as the cell fate is guided by the presence of such factors, *in vivo and in vitro*. This is in line with the embryonic lethality associated with the loss of most of the genes in the heme synthesis pathway. We showed the failure of activation of both MAPK and TGFβ pathways upon heme synthesis inhibition during the naïve-to-primed transition. We hypothesize that this is due to aberrant activation of the SRC kinase, activating STAT3, blocking the FGFR signaling through an induction of its endocytosis, and blocking the SMAD activation through the AKT kinase. The transcription factor EGR1 could also be involved in the phenotype, since it is directly activated by ERK and SMADs. Further work is however required to confirm these hypotheses.

Interestingly, we also showed that SA-induced heme synthesis inhibition can also have heme-independent effects, unrelated to eventual off targets of the molecule, but instead due to an accumulation of the upstream metabolite succinyl-CoA. We showed that this metabolite can then exit the mitochondria and directly impact the potency of stem cells, as shown by the increase in 2CLCs, a subpopulation of cells reminiscent of the totipotency state. We haven't pinpointed the exact mechanism leading to the succinate effect in this process but elements in the results point toward the modulation of gene expression due to the succinylation of proteins such as nucleolin and TRIM28, known to regulate the expression of *Dux*, the master transcriptional regulator in 2C embryos. A putative role for succinate as a repressor of the TET demethylases enzymes, also known to regulate the acquisition of the 2CLC population, could also contribute to this phenomenon, although this hypothesis was not explored in this work. Together this draw attention on the tight regulation between metabolism and gene expression and/or epigenetic regulation.

Bibliography

Abbassi, L. *et al.* (2016) 'Multiple mechanisms cooperate to constitutively exclude the transcriptional co-activator YAP from the nucleus during murine oogenesis', *Biology of Reproduction*. Society for the Study of Reproduction, 94(5). doi: 10.1095/biolreprod.115.137968.

Aguiar, C. J. *et al.* (2014) 'Succinate causes pathological cardiomyocyte hypertrophy through GPR91 activation', *Cell Communication and Signaling*. BioMed Central Ltd., 12(1). doi: 10.1186/s12964-014-0078-2.

Aguilera-Castrejon, A. *et al.* (2021) 'Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis', *Nature*. Nature Research, pp. 1–6. doi: 10.1038/s41586-021-03416-3.

Ajioka, R. S., Phillips, J. D. and Kushner, J. P. (2006) 'Biosynthesis of heme in mammals', *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*. Elsevier, 1763(7), pp. 723–736. doi: 10.1016/J.BBAMCR.2006.05.005.

Arueya, G. L. and Oyewale, T. M. (2015) 'Effect of varying degrees of succinylation on the functional and morphological properties of starch from acha (Digitaria exilis Kippis Stapf)', *Food Chemistry*. Elsevier Ltd, 177, pp. 258–266. doi: 10.1016/j.foodchem.2015.01.019.

Atamna, H. *et al.* (2002) 'Heme deficiency may be a factor in the mitochondrial and neuronal decay of aging', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 99(23), pp. 14807–14812. doi: 10.1073/pnas.192585799.

Atamna, H. (2004) 'Heme, iron, and the mitochondrial decay of ageing', *Ageing Research Reviews*. Elsevier, pp. 303–318. doi: 10.1016/j.arr.2004.02.002.

Auciello, G. *et al.* (2013) 'Regulation of fibroblast growth factor receptor signalling and trafficking by Src and Eps8', *Journal of Cell Science*. Company of Biologists, 126(2), pp. 613–624. doi: 10.1242/jcs.116228.

Babbitt, S. E. *et al.* (2015) 'Mitochondrial cytochrome c biogenesis: No longer an enigma', *Trends in Biochemical Sciences*. Elsevier Ltd, pp. 446–455. doi: 10.1016/j.tibs.2015.05.006.

Bahat, A. *et al.* (2018) 'MTCH2-mediated mitochondrial fusion drives exit from naïve pluripotency in embryonic stem cells', *Nature Communications*. Nature Publishing Group, 9(1), p. 5132. doi: 10.1038/s41467-018-07519-w.

Barker, N., De Wetering, M. Van and Clevers, H. (2008) 'The intestinal stem cell', *Genes and Development*. Cold Spring Harbor Laboratory Press, pp. 1856–1864. doi: 10.1101/gad.1674008.

Baron, V. *et al.* (2006) 'The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGFb1, PTEN, p53, and fibronectin', *Cancer Gene Therapy*, 13, pp. 115–124. doi: 10.1038/sj.cgt.7700896.

Batlle-Morera, L., Smith, A. and Nichols, J. (2008) 'Parameters influencing derivation of embryonic stem cells from murine embryos', *Genesis*. Genesis, 46(12), pp. 758–767. doi: 10.1002/dvg.20442.

Bedzhov, I. and Zernicka-Goetz, M. (2014) 'Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation', *Cell*. Elsevier, 156(5), pp. 1032–1044. doi: 10.1016/j.cell.2014.01.023.

De Belly, H. *et al.* (2021) 'Membrane Tension Gates ERK-Mediated Regulation of Pluripotent Cell Fate.', *Cell stem cell*. Cell Press, 28(2), pp. 273-284.e6. doi: 10.1016/j.stem.2020.10.018.

Beltran-Povea, A. *et al.* (2015) 'Role of nitric oxide in the maintenance of pluripotency and regulation of the hypoxia response in stem cells.', *World journal of stem cells*. Baishideng Publishing Group Inc, 7(3), pp. 605–17. doi: 10.4252/wjsc.v7.i3.605.
Betschinger, J. *et al.* (2013) 'Exit from Pluripotency Is Gated by Intracellular Redistribution of the bHLH Transcription Factor Tfe3', *Cell*, 153(2), pp. 335–347. doi: 10.1016/j.cell.2013.03.012.
Bowers, M. A. *et al.* (1992) 'Quantitative determination of porphyrins in rat and human urine and evaluation of urinary porphyrin profiles during mercury and lead exposures', *The Journal of Laboratory and Clinical Medicine*, 120(2), pp. 272–281. doi: 10.5555/uri:pii:0022214392901358.
Boyer, L. A. *et al.* (2005) 'Core transcriptional regulatory circuitry in human embryonic stem cells', *Cell*. NIH Public Access, 122(6), pp. 947–956. doi: 10.1016/j.cell.2005.08.020.
Braun, R. E. (2001) 'Packaging paternal chromosomes with protamine', *Nature Genetics*. Springer Nature, 28(1), pp. 10–12. doi: 10.1038/ng0501-10.
Brons, I. G. M. *et al.* (2007) 'Derivation of pluripotent epiblast stem cells from mammalian embryos', *Nature*, 448(7150), pp. 191–195. doi: 10.1038/nature05950.
Bulut-Karslioglu, A. *et al.* (2016) 'Inhibition of mTOR induces a paused pluripotent state', *Nature*. Nature Publishing Group, 540(7631), pp. 119–123. doi: 10.1038/nature20578.
Burdon, T. *et al.* (1999) 'Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells', *Developmental Biology*. Academic Press Inc., 210(1), pp. 30–43. doi: 10.1006/dbio.1999.9265.
Canham, M. A. *et al.* (2010) 'Functional Heterogeneity of Embryonic Stem Cells Revealed through Translational Amplification of an Early Endodermal Transcript', *PLoS Biology*. Edited by H. Hamada. Public Library of Science, 8(5), p. e1000379. doi: 10.1371/journal.pbio.1000379.
Carey, B. W. *et al.* (2015) 'Intracellular α-ketoglutarate maintains the pluripotency of embryonic stem cells', *Nature*, 518(7539), pp. 413–416. doi: 10.1038/nature13981.
Cervera, A. M. *et al.* (2009) 'Inhibition of succinate dehydrogenase dysregulates histone modification in mammalian cells', *Molecular Cancer*. Mol Cancer, 8, p. 89. doi: 10.1186/1476-4598-8-89.
Chamberlain, S. J., Yee, D. and Magnuson, T. (2008) 'Polycomb Repressive Complex 2 Is Dispensable for Maintenance of Embryonic Stem Cell Pluripotency', *Stem Cells*. Wiley, 26(6), pp. 1496–1505. doi: 10.1634/stemcells.2008-0102.
Chambers, I. *et al.* (2007) 'Nanog safeguards pluripotency and mediates germline development', *Nature*. Nature Publishing Group, 450(7173), pp. 1230–1234. doi: 10.1038/nature06403.
Chan, E. M. *et al.* (2009) 'Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells', *Nature Biotechnology*. Nature Publishing Group, 27(11), pp. 1033–1037. doi: 10.1038/nbt.1580.
Chan, Y.-S. *et al.* (2013) 'Induction of a Human Pluripotent State with Distinct Regulatory Circuitry that Resembles Preimplantation Epiblast', *Cell Stem Cell*. Cell Stem Cell, 13(6), pp. 663–675. doi: 10.1016/j.stem.2013.11.015.
Chazaud, C. *et al.* (2006) 'Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway', *Developmental Cell*. Dev Cell, 10(5), pp. 615–624. doi: 10.1016/j.devcel.2006.02.020.
Chen, H. *et al.* (2015) 'Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency', *Nature Communications*. Nature Publishing Group,

6. doi: 10.1038/ncomms8095.
Chen, J. J. (2007) 'Regulation of protein synthesis by the heme-regulated eIF2α kinase: Relevance to anemias', *Blood*. Blood, pp. 2693–2699. doi: 10.1182/blood-2006-08-041830.
Chen, T. *et al.* (2011) 'Chemical genetics identify eIF2α kinase heme-regulated inhibitor as an anticancer target', *Nature Chemical Biology*. Nature Publishing Group, 7(9), pp. 610–616. doi: 10.1038/nchembio.613.
Cheng, C.-T. (2014) 'KAPtain in charge of multiple missions: Emerging roles of KAP1', *World Journal of Biological Chemistry*. Baishideng Publishing Group Inc., 5(3), p. 308. doi: 10.4331/wjbc.v5.i3.308.
Chiba, T. *et al.* (2019) 'Sirtuin 5 regulates proximal tubule fatty acid oxidation to protect against AKI', *Journal of the American Society of Nephrology*. American Society of Nephrology, 30(12), pp. 2384–2398. doi: 10.1681/ASN.2019020163.
Choi, H. W. *et al.* (2016) 'Distinct Enhancer Activity of Oct4 in Naive and Primed Mouse Pluripotency', *Stem Cell Reports*. Cell Press, 7(5), pp. 911–926. doi: 10.1016/j.stemcr.2016.09.012.
Christensen, E. *et al.* (1981) 'Urinary excretion of succinylacetone and δ-aminolevulinic acid in patients with hereditary tyrosinemia', *Clinica Chimica Acta*. Clin Chim Acta, 116(3), pp. 331–341. doi: 10.1016/0009-8981(81)90052-8.
Colak, G. *et al.* (2013) 'Identification of lysine succinylation substrates and the succinylation regulatory enzyme CobB in escherichia coli', *Molecular and Cellular Proteomics*. Mol Cell Proteomics, 12(12), pp. 3509–3520. doi: 10.1074/mcp.M113.031567.
Collier, A. J. and Rugg-Gunn, P. J. (2018) 'Identifying Human Naïve Pluripotent Stem Cells − Evaluating State-Specific Reporter Lines and Cell-Surface Markers', *BioEssays*. John Wiley and Sons Inc., 40(5), p. 1700239. doi: 10.1002/bies.201700239.
Correa, P. R. A. V. *et al.* (2007) 'Succinate is a paracrine signal for liver damage', *Journal of Hepatology*. NIH Public Access, 47(2), pp. 262–269. doi: 10.1016/j.jhep.2007.03.016.
Dan, J. *et al.* (2017) 'Zscan4 Inhibits Maintenance DNA Methylation to Facilitate Telomere Elongation in Mouse Embryonic Stem Cells', *CellReports*, 20, pp. 1936–1949. doi: 10.1016/j.celrep.2017.07.070.
Daniell, W. E. *et al.* (1997) 'Environmental chemical exposures and disturbances of heme synthesis', *Environmental Health Perspectives*. Public Health Services, US Dept of Health and Human Services, pp. 37–53. doi: 10.1289/ehp.97105s137.
Deglincerti, A. *et al.* (2016) 'Self-organization of the in vitro attached human embryo', *Nature*. Nature Publishing Group, 533(7602), pp. 251–254. doi: 10.1038/nature17948.
Dickinson, L. A. and Kohwi-Shigematsu, T. (1995) 'Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential.', *Molecular and Cellular Biology*, 15(1), pp. 456–465. doi: 10.1128/MCB.15.1.456.
Disteche, C. M. (2016) 'Dosage compensation of the sex chromosomes and autosomes', *Seminars in Cell and Developmental Biology*. Academic Press, pp. 9–18. doi: 10.1016/j.semcdb.2016.04.013.
Du, J. *et al.* (2011) 'Sirt5 Is a NAD-Dependent Protein Lysine Demalonylase and Desuccinylase', *Science*. American Association for the Advancement of Science, 334(6057), pp. 806–809. doi: 10.1126/science.1207861.
Du, P. *et al.* (2018) 'An Intermediate Pluripotent State Controlled by MicroRNAs Is Required for the Naive-to-Primed Stem Cell Transition', *Cell Stem Cell*. Cell Press, 22(6), pp. 851-

864.e5. doi: 10.1016/j.stem.2018.04.021.
Du, Y. *et al.* (2018) 'Tissue distribution, subcellular localization, and enzymatic activity analysis of human SIRT5 isoforms', *Biochemical and Biophysical Research Communications*. Elsevier B.V., 503(2), pp. 763–769. doi: 10.1016/j.bbrc.2018.06.073.
Dunn, S. J. *et al.* (2014) 'Defining an essential transcription factor program for naïve pluripotency', *Science*. American Association for the Advancement of Science, 344(6188), pp. 1156–1160. doi: 10.1126/science.1248882.
Duvigneau, J. C., Esterbauer, H. and Kozlov, A. V. (2019) 'Role of heme oxygenase as a modulator of heme-mediated pathways', *Antioxidants*. MDPI AG, p. 475. doi: 10.3390/antiox8100475.
Eckersley-Maslin, M. *et al.* (2019) 'Dppa2 and Dppa4 directly regulate the Dux-driven zygotic transcriptional program', *Genes and Development*. Cold Spring Harbor Laboratory Press, 33(3–4), pp. 194–208. doi: 10.1101/gad.321174.118.
Eckersley-Maslin, M. A. *et al.* (2016) 'MERVL/Zscan4 Network Activation Results in Transient Genome-wide DNA Demethylation of mESCs', *Cell Reports*. Elsevier B.V., 17(1), pp. 179–192. doi: 10.1016/j.celrep.2016.08.087.
Edwards, R. G., Bavister, B. D. and Steptoe, P. C. (1969) 'Early stages of fertilization in vitro of human oocytes matured in vitro', *Nature*. Nature, 221(5181), pp. 632–635. doi: 10.1038/221632a0.
Edwards, R. G., Steptoe, P. C. and Purdy, J. M. (1970) 'Fertilization and cleavage in vitro of preovulator human oocytes', *Nature*, 227(5265), pp. 1307–1309. doi: 10.1038/2271307a0.
Eguizabal, C. *et al.* (2013) 'Dedifferentiation, transdifferentiation, and reprogramming: Future directions in regenerative medicine', *Seminars in Reproductive Medicine*. Semin Reprod Med, 31(1), pp. 82–94. doi: 10.1055/s-0032-1331802.
Ehnes, D. D. *et al.* (2020) 'Combinatorial metabolism drives the naive to primed pluripotent chromatin landscape', *Experimental Cell Research*. Elsevier Inc., 389(2), p. 111913. doi: 10.1016/j.yexcr.2020.111913.
Eiselleova, L. *et al.* (2009) 'A complex role for FGF-2 in self-renewal, survival, and adhesion of human embryonic stem cells', *Stem Cells*. Stem Cells, 27(8), pp. 1847–1857. doi: 10.1002/stem.128.
Evans, M. J. and Kaufman, M. H. (1981) 'Establishment in culture of pluripotential cells from mouse embryos', *Nature*. Nature Publishing Group, 292(5819), pp. 154–156. doi: 10.1038/292154a0.
Falco, G. *et al.* (2007) 'Zscan4: A novel gene expressed exclusively in late 2-cell embryos and embryonic stem cells', *Developmental Biology*. Academic Press Inc., 307(2), pp. 539–550. doi: 10.1016/j.ydbio.2007.05.003.
Ferraro, F., Lo Celso, C. and Scadden, D. (2010) 'Adult stem cells and their niches', *Advances in Experimental Medicine and Biology*. NIH Public Access, pp. 155–168. doi: 10.1007/978-1-4419-7037-4_11.
Ficz, G. *et al.* (2013) 'FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency', *Cell Stem Cell*. Cell Stem Cell, 13(3), pp. 351–359. doi: 10.1016/j.stem.2013.06.004.
Fierro-González, J. C. *et al.* (2013) 'Cadherin-dependent filopodia control preimplantation embryo compaction', *Nature Cell Biology*. Nature Publishing Group, 15(12), pp. 1424–1433. doi: 10.1038/ncb2875.
Fleming, T. P. *et al.* (1989) 'Development of tight junctions de novo in the mouse early

embryo: Control of assembly of the tight junction-specific protein, ZO-1', *Journal of Cell Biology*. J Cell Biol, 108(4), pp. 1407–1418. doi: 10.1083/jcb.108.4.1407.
Förstermann, U. and Sessa, W. C. (2012) 'Nitric oxide synthases: Regulation and function', *European Heart Journal*. Oxford University Press, p. 829. doi: 10.1093/eurheartj/ehr304.
Fortin, J. and Bernard, D. J. (2010) 'SMAD3 and EGR1 physically and functionally interact in promoter-specific fashion', *Cellular Signalling*. Cell Signal, 22(6), pp. 936–943. doi: 10.1016/j.cellsig.2010.01.019.
Gafni, O. *et al.* (2013a) 'Derivation of novel human ground state naive pluripotent stem cells', *Nature*, 504(7479), pp. 282–286. doi: 10.1038/nature12745.
Gafni, O. *et al.* (2013b) 'Derivation of novel human ground state naive pluripotent stem cells', *Nature*, 504(7479), pp. 282–286. doi: 10.1038/nature12745.
Galmozzi, A. *et al.* (2019) 'PGRMC2 is an intracellular haem chaperone critical for adipocyte function', *Nature*. Nature Research, 576(7785), pp. 138–142. doi: 10.1038/s41586-019-1774-2.
Garcia, R. *et al.* (2001) 'Constitutive activation of Stat3 by the Src and JAK tyrosine kinases participates in growth regulation of human breast carcinoma cells', *Oncogene*. Oncogene, 20(20), pp. 2499–2513. doi: 10.1038/sj.onc.1204349.
Gilissen, J. *et al.* (2015) 'Forskolin-free cAMP assay for Gi-coupled receptors', *Biochemical Pharmacology*. Elsevier Inc., 98(3), pp. 381–391. doi: 10.1016/j.bcp.2015.09.010.
Gö Kbuget, D. and Blelloch, R. (2019) 'Epigenetic control of transcriptional regulation in pluripotency and early differentiation'. doi: 10.1242/dev.164772.
Gordeeva, O. and Khaydukov, S. (2017) 'Tumorigenic and Differentiation Potentials of Embryonic Stem Cells Depend on TGF β Family Signaling: Lessons from Teratocarcinoma Cells Stimulated to Differentiate with Retinoic Acid', *Stem Cells International*. Hindawi Limited, 2017. doi: 10.1155/2017/7284872.
Gotoh, S. *et al.* (2011) 'Egr-1 regulates the transcriptional repression of mouse δ-aminolevulinic acid synthase 1 by heme', *Gene*. Elsevier, 472(1–2), pp. 28–36. doi: 10.1016/j.gene.2010.10.008.
Gough, D. R. and Cotter, T. G. (2011) 'Hydrogen peroxide: A Jekyll and Hyde signalling molecule', *Cell Death and Disease*. Nature Publishing Group, pp. e213–e213. doi: 10.1038/cddis.2011.96.
Gregg, J. and Fraizer, G. (2011) 'Transcriptional Regulation of EGR1 by EGF and the ERK Signaling Pathway in Prostate Cancer Cells', *Genes and Cancer*. Impact Journals, LLC, 2(9), pp. 900–909. doi: 10.1177/1947601911431885.
Grow, E. J. *et al.* (2015) 'Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells', *Nature*. Nature Publishing Group, 522(7555), pp. 221–246. doi: 10.1038/nature14308.
Guo, Z. *et al.* (2020) 'Systematic Proteome and Lysine Succinylome Analysis Reveals the Enhanced Cell Migration by Hyposuccinylation in Esophageal Squamous Cell Cancer', *Molecular & Cellular Proteomics*. Elsevier BV, p. mcp.RA120.002150. doi: 10.1074/mcp.ra120.002150.
Gurdon, J. B. (1962) 'The Developmental Capacity of Nuclei taken from Intestinal Epithelium Cells of Feeding Tadpoles', *Development*, 10(4).
Gut, P. *et al.* (2020) 'SUCLA2 mutations cause global protein succinylation contributing to the pathomechanism of a hereditary mitochondrial disease', *Nature Communications*. Nature Research, 11(1). doi: 10.1038/s41467-020-19743-4.
Hackett, J. A. *et al.* (2013) 'Synergistic mechanisms of DNA demethylation during transition

to ground-state pluripotency', *Stem Cell Reports*. Stem Cell Reports, 1(6), pp. 518–531. doi: 10.1016/j.stemcr.2013.11.010.
Hamilton, W. B. and Brickman, J. M. (2014) 'Erk Signaling Suppresses Embryonic Stem Cell Self-Renewal to Specify Endoderm', *Cell Reports*. Elsevier, 9(6), pp. 2056–2070. doi: 10.1016/j.celrep.2014.11.032.
Han, D. W. *et al.* (2010) 'Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages', *Cell*. Cell, 143(4), pp. 617–627. doi: 10.1016/j.cell.2010.10.015.
Hanna, J. *et al.* (2010) 'Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 107(20), pp. 9222–9227. doi: 10.1073/pnas.1004584107.
Hannibal, L. *et al.* (2012) 'Heme binding properties of glyceraldehyde-3-phosphate dehydrogenase', *Biochemistry*. American Chemical Society, 51(43), pp. 8514–8529. doi: 10.1021/bi300863a.
Harrison, M. M. *et al.* (2011) 'Zelda binding in the early Drosophila melanogaster embryo marks regions subsequently activated at the maternal-to-zygotic transition', *PLoS Genetics*. PLoS Genet, 7(10). doi: 10.1371/journal.pgen.1002266.
Hartney, T. *et al.* (2011) 'Xanthine oxidase-derived ROS upregulate Egr-1 via ERK1/2 in PA smooth muscle cells; model to test impact of extracellular ROS in chronic hypoxia', *PLoS ONE*. PLoS One, 6(11). doi: 10.1371/journal.pone.0027531.
Hasan, R. N. and Schafer, A. I. (2008) 'Hemin upregulates Egr-1 expression in vascular smooth muscle cells via reactive oxygen species ERK-1/2-Elk-1 and NF-κB', *Circulation Research*. Circ Res, 102(1), pp. 42–50. doi: 10.1161/CIRCRESAHA.107.155143.
Haugsten, E. M. *et al.* (2005) 'Different intracellular trafficking of FGF1 endocytosed by the four homologous FGF receptors', *Journal of Cell Science*. The Company of Biologists Ltd, 118(17), pp. 3869–3881. doi: 10.1242/jcs.02509.
Hayashi, K. *et al.* (2011) 'Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells', *Cell*. Cell Press, 146(4), pp. 519–532. doi: 10.1016/j.cell.2011.06.052.
He, J. *et al.* (2019) 'Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells', *Nature Communications*. Nature Publishing Group, 10(1), pp. 1–13. doi: 10.1038/s41467-018-08006-y.
Hendrickson, P. G. *et al.* (2017) 'Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons', *Nature Genetics*. Nature Publishing Group, 49(6), pp. 925–934. doi: 10.1038/ng.3844.
Högberg, C. *et al.* (2011) 'Succinate independently stimulates full platelet activation via cAMP and phosphoinositide 3-kinase-β signaling', *Journal of Thrombosis and Haemostasis*. J Thromb Haemost, 9(2), pp. 361–372. doi: 10.1111/j.1538-7836.2010.04158.x.
Homedan, C. *et al.* (2014) 'Acute intermittent porphyria causes hepatic mitochondrial energetic failure in a mouse model', *International Journal of Biochemistry and Cell Biology*, 51, pp. 93–101. doi: 10.1016/j.biocel.2014.03.032.
Hu, Z. *et al.* (2020) 'Transient inhibition of mTOR in human pluripotent stem cells enables robust formation of mouse-human chimeric embryos', *Science Advances*. American Association for the Advancement of Science, 6(20), pp. 298–311. doi: 10.1126/sciadv.aaz0298.

Huang, D. *et al.* (2018) 'LIF-Activated Jak signaling determines Esrrb expression during late-stage reprogramming', *Biology Open*. Company of Biologists Ltd, 7(1). doi: 10.1242/bio.029264.
Huang, S.-M. A. *et al.* (2009) 'Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling.', *Nature*. Nature Publishing Group, 461(7264), pp. 614–20. doi: 10.1038/nature08356.
Hughes, A. L. *et al.* (2007) 'Dap1/PGRMC1 Binds and Regulates Cytochrome P450 Enzymes', *Cell Metabolism*. Cell Metab, 5(2), pp. 143–149. doi: 10.1016/j.cmet.2006.12.009.
Hussein, A. M. *et al.* (2020) 'Metabolic Control over mTOR-Dependent Diapause-like State', *Developmental Cell*. Cell Press, 52(2), pp. 236-250.e7. doi: 10.1016/j.devcel.2019.12.018.
Hyenne, V. *et al.* (2005) 'Vezatin, a protein associated to adherens junctions, is required for mouse blastocyst morphogenesis', *Developmental Biology*. Academic Press Inc., 287(1), pp. 180–191. doi: 10.1016/j.ydbio.2005.09.004.
De Iaco, A. *et al.* (2017) 'DUX-family transcription factors regulate zygotic genome activation in placental mammals', *Nature Genetics*. Nature Publishing Group, 49(6), pp. 941–945. doi: 10.1038/ng.3858.
Ishiuchi, T. *et al.* (2015) 'Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly', *Nature Structural and Molecular Biology*. Nature Publishing Group, 22(9), pp. 662–671. doi: 10.1038/nsmb.3066.
Ivanov, A. V. *et al.* (2007) 'PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing', *Molecular Cell*. Mol Cell, 28(5), pp. 823–837. doi: 10.1016/j.molcel.2007.11.012.
Ivanova, N. *et al.* (2006) 'Dissecting self-renewal in stem cells with RNA interference', *Nature*. Nature Publishing Group, 442(7102), pp. 533–538. doi: 10.1038/nature04915.
Iyengar, S. and Farnham, P. J. (2011) 'KAP1 protein: An enigmatic master regulator of the genome', *Journal of Biological Chemistry*. J Biol Chem, pp. 26267–26276. doi: 10.1074/jbc.R111.252569.
Jacobs, J. M. *et al.* (1998) 'Formation of zinc protoporphyrin in cultured hepatocytes: Effects of ferrochelatase inhibition, iron chelation or lead', *Toxicology*. Toxicology, 125(2–3), pp. 95–105. doi: 10.1016/S0300-483X(97)00164-9.
Jinnin, M., Ihn, H. and Tamaki, K. (2006) 'Characterization of SIS3, a novel specific inhibitor of Smad3, and its effect on transforming growth factor-β1-induced extracellular matrix expression', *Molecular Pharmacology*. Mol Pharmacol, 69(2), pp. 597–607. doi: 10.1124/mol.105.017483.
Johansson, B. M. and Wiles, M. V (1995) 'Evidence for involvement of activin A and bone morphogenetic protein 4 in mammalian mesoderm and hematopoietic development.', *Molecular and Cellular Biology*. American Society for Microbiology, 15(1), pp. 141–151. doi: 10.1128/mcb.15.1.141.
Jones, J. E. *et al.* (2009) 'Src family kinase gene targets during myeloid differentiation: Identification of the EGR-1 gene as a direct target', *Leukemia*. Nature Publishing Group, 23(10), pp. 1933–1935. doi: 10.1038/leu.2009.118.
Joseph, S. R. *et al.* (2017) 'Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos', *eLife*. eLife Sciences Publications Ltd, 6. doi: 10.7554/eLife.23326.
Kalmar, T. *et al.* (2009) 'Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells', *PLoS Biology*. Edited by M. A. Goodell. Public Library of Science, 7(7), p. e1000149. doi: 10.1371/journal.pbio.1000149.

Kathiriya, J. J. *et al.* (2014) 'Presence and utility of intrinsically disordered regions in kinases', *Molecular BioSystems*. Royal Society of Chemistry, 10(11), pp. 2876–2888. doi: 10.1039/c4mb00224e.

Killian, J. K. *et al.* (2013) 'Succinate dehydrogenase mutation underlies global epigenomic divergence in gastrointestinal stromal tumor', *Cancer Discovery*. NIH Public Access, 3(6), pp. 648–657. doi: 10.1158/2159-8290.CD-13-0092.

Kim, H. *et al.* (2013) 'Modulation of β-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal', *Nature Communications*. Nature Publishing Group, 4. doi: 10.1038/ncomms3403.

Kim, H. J. *et al.* (2012) 'Structure, function, and assembly of heme centers in mitochondrial respiratory complexes', *Biochimica et Biophysica Acta - Molecular Cell Research*. NIH Public Access, pp. 1604–1616. doi: 10.1016/j.bbamcr.2012.04.008.

Kim, J. W. *et al.* (2006) 'HIF-1-mediated expression of pyruvate dehydrogenase kinase: A metabolic switch required for cellular adaptation to hypoxia', *Cell Metabolism*. Cell Metab, 3(3), pp. 177–185. doi: 10.1016/j.cmet.2006.02.002.

Kinoshita, M. *et al.* (2020) 'Capture of Mouse and Human Stem Cells with Features of Formative Pluripotency', *Cell Stem Cell*. Cell Press. doi: 10.1016/j.stem.2020.11.005.

Kluckova, K. and Tennant, D. A. (2018) 'Metabolic implications of hypoxia and pseudohypoxia in pheochromocytoma and paraganglioma', *Cell and Tissue Research*. Springer Verlag, pp. 367–378. doi: 10.1007/s00441-018-2801-6.

Kojima, Y., Tam, O. H. and Tam, P. P. L. (2014) 'Timing of developmental events in the early mouse embryo', *Seminars in Cell and Developmental Biology*. Elsevier Ltd, pp. 65–75. doi: 10.1016/j.semcdb.2014.06.010.

Kubaczka, C. *et al.* (2014) 'Derivation and maintenance of murine trophoblast stem cells under defined conditions', *Stem Cell Reports*. Cell Press, 2(2), pp. 232–242. doi: 10.1016/j.stemcr.2013.12.013.

Kumar, R. M. *et al.* (2014) 'Deconstructing transcriptional heterogeneity in pluripotent stem cells', *Nature*. Nature Publishing Group, 516(729), pp. 56–61. doi: 10.1038/nature13920.

Kurmi, K. *et al.* (2018) 'Carnitine Palmitoyltransferase 1A Has a Lysine Succinyltransferase Activity', *Cell Reports*. Elsevier B.V., 22(6), pp. 1365–1373. doi: 10.1016/j.celrep.2018.01.030.

Landeira, D. *et al.* (2010) 'Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators', *Nature Cell Biology*. Nat Cell Biol, 12(6), pp. 618–624. doi: 10.1038/ncb2065.

Lanner, F. and Rossant, J. (2010) 'The role of FGF/Erk signaling in pluripotent cells', *Development*. Company of Biologists Ltd, pp. 3351–3360. doi: 10.1242/dev.050146.

Laukka, T. *et al.* (2016) 'Fumarate and succinate regulate expression of hypoxia-inducible genes via TET enzymes', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 291(8), pp. 4256–4265. doi: 10.1074/jbc.M115.688762.

Lee, J., Park, Y. J. and Jung, H. (2019) 'Protein Kinases and Their Inhibitors in Pluripotent Stem Cell Fate Regulation', *Stem Cells International*. Hindawi Limited. doi: 10.1155/2019/1569740.

Lee, J. W. *et al.* (2019) 'Hypoxia signaling in human diseases and therapeutic targets', *Experimental and Molecular Medicine*. Nature Publishing Group, pp. 1–13. doi: 10.1038/s12276-019-0235-1.

Lee, M. T. *et al.* (2013) 'Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition', *Nature*. Nature, 503(7476), pp. 360–364. doi: 10.1038/nature12632.
Levy, O. *et al.* (2020) 'Shattering barriers toward clinically meaningful MSC therapies', *Science Advances*. American Association for the Advancement of Science, 6(30). doi: 10.1126/sciadv.aba6884.
Li, F. *et al.* (2007) 'Regulation of HIF-1α Stability through S-Nitrosylation', *Molecular Cell*. NIH Public Access, 26(1), pp. 63–74. doi: 10.1016/j.molcel.2007.02.024.
Li, H. *et al.* (2009) 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*. Oxford Academic, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.
Li, L. *et al.* (2016) 'SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability', *Nature Communications*. Nature Publishing Group, 7(1), p. 12235. doi: 10.1038/ncomms12235.
Li, M. *et al.* (2018) 'Genome-wide CRISPR-KO Screen Uncovers mTORC1-Mediated Gsk3 Regulation in Naive Pluripotency Maintenance and Dissolution', *Cell Reports*. Elsevier B.V., 24(2), pp. 489–502. doi: 10.1016/j.celrep.2018.06.027.
Li, M. and Belmonte, J. C. I. (2017) 'Ground rules of the pluripotency gene regulatory network', *Nature Reviews Genetics*. Nature Publishing Group, 18(3), pp. 180–191. doi: 10.1038/nrg.2016.156.
Li, M. and Izpisua Belmonte, J. C. (2018) 'Deconstructing the pluripotency gene regulatory network.', *Nature cell biology*. Nature Publishing Group, 20(4), pp. 382–392. doi: 10.1038/s41556-018-0067-6.
Li, Q. V., Rosen, B. P. and Huangfu, D. (2020) 'Decoding pluripotency: Genetic screens to interrogate the acquisition, maintenance, and exit of pluripotency', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. Wiley-Blackwell, 12(1). doi: 10.1002/wsbm.1464.
Lindblad, B., Lindstedt, S. and Steen, G. (1977) 'On the enzymic defects in hereditary tyrosinemia', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 74(10), pp. 4641–4645. doi: 10.1073/pnas.74.10.4641.
Liu, X. *et al.* (2017) 'Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming', *Nature Methods*. Nature Publishing Group, 14(11), pp. 1055–1062. doi: 10.1038/nmeth.4436.
Loh, K. M., Lim, B. and Ang, L. T. (2015) 'Ex uno plures: Molecular designs for embryonic pluripotency', *Physiological Reviews*. American Physiological Society, 95(1), pp. 245–295. doi: 10.1152/physrev.00001.2014.
Loh, Y. H. *et al.* (2006) 'The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells', *Nature Genetics*. Nature Publishing Group, 38(4), pp. 431–440. doi: 10.1038/ng1760.
Lopez, J. *et al.* (2012) 'Src tyrosine kinase inhibits apoptosis through the Erk1/2-dependent degradation of the death accelerator Bik', *Cell Death and Differentiation*. Nature Publishing Group, 19(9), pp. 1459–1469. doi: 10.1038/cdd.2012.21.
De Los Angeles, A. (2019) 'Frontiers of pluripotency', in *Methods in Molecular Biology*. Humana Press Inc., pp. 3–27. doi: 10.1007/978-1-4939-9524-0_1.
Lu, F. *et al.* (2014) 'Role of Tet proteins in enhancer activity and telomere elongation', *Genes and Development*. Cold Spring Harbor Laboratory Press, 28(19), pp. 2103–2119. doi: 10.1101/gad.248005.114.
Macfarlan, T. S. *et al.* (2012) 'Embryonic stem cell potency fluctuates with endogenous retrovirus activity', *Nature*. NIH Public Access, 487(7405), pp. 57–63. doi:

10.1038/nature11244.
MacLean Hunter, S. and Evans, M. (1999) 'Non-surgical method for the induction of delayed implantation and recovery of viable blastocysts in rats and mice by the use of tamoxifen and Depo-Provera', *Molecular Reproduction and Development*. John Wiley & Sons, Ltd, 52(1), pp. 29–32. doi: 10.1002/(SICI)1098-2795(199901)52:1<29::AID-MRD4>3.0.CO;2-2.
Manejwala, F. M., Cragoe, E. J. and Schultz, R. M. (1989) 'Blastocoel expansion in the preimplantation mouse embryo: Role of extracellular sodium and chloride and possible apical routes of their entry', *Developmental Biology*. Dev Biol, 133(1), pp. 210–220. doi: 10.1016/0012-1606(89)90312-6.
Marks, H. *et al.* (2012) 'The transcriptional and epigenomic foundations of ground state pluripotency', *Cell*. Elsevier, 149(3), pp. 590–604. doi: 10.1016/j.cell.2012.03.026.
Marsboom, G. *et al.* (2016) 'Glutamine Metabolism Regulates the Pluripotency Transcription Factor OCT4', *Cell Reports*. Elsevier B.V., 16(2), pp. 323–332. doi: 10.1016/j.celrep.2016.05.089.
Martin, G. R. (1981) 'Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.', *Proceedings of the National Academy of Sciences of the United States of America*, 78(12), pp. 7634–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6950406 (Accessed: 15 April 2019).
Martin, G. R. and Evans, M. J. (1974) 'The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture', *Cell*. Cell, 2(3), pp. 163–172. doi: 10.1016/0092-8674(74)90090-7.
Martina, J. A. *et al.* (2014) 'The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris', *Science Signaling*. American Association for the Advancement of Science, 7(309), pp. ra9–ra9. doi: 10.1126/scisignal.2004754.
Masui, S. *et al.* (2007) 'Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells', *Nature Cell Biology*. Nat Cell Biol, 9(6), pp. 625–635. doi: 10.1038/ncb1589.
Mathieu, J. *et al.* (2014) 'Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency', *Cell Stem Cell*. Cell Press, 14(5), pp. 592–605. doi: 10.1016/j.stem.2014.02.012.
Mathieu, J. *et al.* (2019) 'Folliculin regulates mTORC1/2 and WNT pathways in early human pluripotency', *Nature Communications*. Nature Publishing Group, 10(1), p. 632. doi: 10.1038/s41467-018-08020-0.
Mathieu, J. and Ruohola-Baker, H. (2017) 'Metabolic remodeling during the loss and acquisition of pluripotency', *Development*, 144(4), pp. 541–551. doi: 10.1242/dev.128389.
Matilainen, O., Quirós, P. M. and Auwerx, J. (2017) 'Mitochondria and Epigenetics – Crosstalk in Homeostasis and Stress', *Trends in Cell Biology*, 27(6), pp. 453–463. doi: 10.1016/j.tcb.2017.02.004.
Matters, G. L. and Beale, S. I. (1995) 'Blue-light-regulated expression of genes for two early steps of chlorophyll biosynthesis in Chlamydomonas reinhardtii', *Plant Physiology*. American Society of Plant Biologists, 109(2), pp. 471–479. doi: 10.1104/pp.109.2.471.
Meier, F. *et al.* (2018) 'Online parallel accumulation–serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer', *Molecular and Cellular Proteomics*. American Society for Biochemistry and Molecular Biology Inc., 17(12), pp. 2534–2545. doi: 10.1074/mcp.TIR118.000900.
von Meyenn, F. *et al.* (2016) 'Impairment of DNA Methylation Maintenance Is the Main

Cause of Global Demethylation in Naive Embryonic Stem Cells', *Molecular Cell*. Cell Press, 62(6), pp. 848–861. doi: 10.1016/j.molcel.2016.04.025.
Mills, E. L. *et al.* (2018) 'Accumulation of succinate controls activation of adipose tissue thermogenesis', *Nature*. Nature Publishing Group, 560(7716), pp. 102–106. doi: 10.1038/s41586-018-0353-2.
Mills, E. and O'Neill, L. A. J. (2014) 'Succinate: A metabolic signal in inflammation', *Trends in Cell Biology*. Elsevier Ltd, pp. 313–320. doi: 10.1016/j.tcb.2013.11.008.
Miyadera, H. *et al.* (2003) 'Atpenins, potent and specific inhibitors of mitochondrial complex II (succinate-ubiquinone oxidoreductase)', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 100(2), pp. 473–477. doi: 10.1073/pnas.0237315100.
Mogi, T., Saiki, K. and Anraku, Y. (1994) 'Biosynthesis and functional role of haem O and haem A', *Molecular Microbiology*. Mol Microbiol, 14(3), pp. 391–398. doi: 10.1111/j.1365-2958.1994.tb02174.x.
Moody, J. D. *et al.* (2017) 'First critical repressive H3K27me3 marks in embryonic stem cells identified using designed protein inhibitor.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 114(38), pp. 10125–10130. doi: 10.1073/pnas.1706907114.
Mora-Castilla, S. *et al.* (2010) 'Nitric oxide repression of Nanog promotes mouse embryonic stem cell differentiation', *Cell Death and Differentiation*. Cell Death Differ, 17(6), pp. 1025–1033. doi: 10.1038/cdd.2009.204.
Munakata, H. *et al.* (2004) 'Role of the heme regulatory motif in the heme-mediated inhibition of mitochondrial import of 5-aminolevulinate synthase', *Journal of Biochemistry*. J Biochem, 136(2), pp. 233–238. doi: 10.1093/jb/mvh112.
Nagarajan, P. *et al.* (2013) 'Histone Acetyl Transferase 1 Is Essential for Mammalian Development, Genome Stability, and the Processing of Newly Synthesized Histones H3 and H4', *PLoS Genetics*. PLoS Genet, 9(6). doi: 10.1371/journal.pgen.1003518.
Nagy, A. *et al.* (1993) 'Derivation of completely cell culture-derived mice from early-passage embryonic stem cells', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 90(18), pp. 8424–8428. doi: 10.1073/pnas.90.18.8424.
Nakaki, F. *et al.* (2013) 'Induction of mouse germ-cell fate by transcription factors in vitro', *Nature*. Nature Publishing Group, 501(7466), pp. 222–226. doi: 10.1038/nature12417.
Nakamura, T. *et al.* (2016) 'A developmental coordinate of pluripotency among mice, monkeys and humans', *Nature*. Nature Publishing Group, 537(7618), pp. 57–62. doi: 10.1038/nature19096.
Nakaya, Y. *et al.* (2008) 'RhoA and microtubule dynamics control cell-basement membrane interaction in EMT during gastrulation', *Nature Cell Biology*. Nat Cell Biol, 10(7), pp. 765–775. doi: 10.1038/ncb1739.
Narayana, Y. V. *et al.* (2019) 'Clathrin-Mediated Endocytosis Regulates a Balance between Opposing Signals to Maintain the Pluripotent State of Embryonic Stem Cells', *Stem Cell Reports*. Cell Press, 12(1), pp. 152–164. doi: 10.1016/j.stemcr.2018.11.018.
Nesvizhskii, A. I. *et al.* (2003) 'A statistical model for identifying proteins by tandem mass spectrometry', *Analytical Chemistry*. American Chemical Society , 75(17), pp. 4646–4658. doi: 10.1021/ac0341261.
Nichols, J. and Smith, A. (2009) 'Naive and Primed Pluripotent States', *Cell Stem Cell*. Cell Stem Cell, pp. 487–492. doi: 10.1016/j.stem.2009.05.015.

Nichols, J. and Smith, A. (2011) 'The origin and identity of embryonic stem cells', *Development*. Oxford University Press for The Company of Biologists Limited, 138(1), pp. 3–8. doi: 10.1242/dev.050831.

Niwa, H., Miyazaki, J. I. and Smith, A. G. (2000) 'Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells', *Nature Genetics*. Nat Genet, 24(4), pp. 372–376. doi: 10.1038/74199.

Nugud, A., Sandeep, D. and El-Serafi, A. T. (2018) 'Two faces of the coin: Minireview for dissecting the role of reactive oxygen species in stem cell potency and lineage commitment', *Journal of Advanced Research*. Elsevier B.V., pp. 73–79. doi: 10.1016/j.jare.2018.05.012.

Ogawa, K. *et al.* (2001) 'Heme mediates derepression of Maf recognition element through direct binding to transcription repressor Bach1', *EMBO Journal*. EMBO J, 20(11), pp. 2835–2843. doi: 10.1093/emboj/20.11.2835.

Pakos-Zebrucka, K. *et al.* (2016) 'The integrated stress response', *EMBO reports*. EMBO, 17(10), pp. 1374–1395. doi: 10.15252/embr.201642195.

Park, J. *et al.* (2013) 'SIRT5-Mediated Lysine Desuccinylation Impacts Diverse Metabolic Pathways', *Molecular Cell*. Cell Press, 50(6), pp. 919–930. doi: 10.1016/j.molcel.2013.06.001.

Pastor, W. A. *et al.* (2016) 'Naive Human Pluripotent Cells Feature a Methylation Landscape Devoid of Blastocyst or Germline Memory', *Cell Stem Cell*. Cell Press, 18(3), pp. 323–329. doi: 10.1016/j.stem.2016.01.019.

Percharde, M. *et al.* (2018) 'A LINE1-Nucleolin Partnership Regulates Early Development and ESC Identity', *Cell*. Cell Press, 174(2), pp. 391-405.e19. doi: 10.1016/j.cell.2018.05.043.

Phillips, J. D. *et al.* (2001) 'A mouse model of familial porphyria cutanea tarda', *Proceedings of the National Academy of Sciences*, 98(1), pp. 259–264. doi: 10.1073/pnas.011481398.

Pittenger, M. F. *et al.* (2019) 'Mesenchymal stem cell perspective: cell biology to clinical progress', *npj Regenerative Medicine*. Nature Research, pp. 1–15. doi: 10.1038/s41536-019-0083-6.

Porębska, N. *et al.* (2018) 'Targeting Cellular Trafficking of Fibroblast Growth Factor Receptors as a Strategy for Selective Cancer Treatment', *Journal of Clinical Medicine*. MDPI AG, 8(1), p. 7. doi: 10.3390/jcm8010007.

Poston, C. N. *et al.* (2011) 'Proteomic analysis of lipid raft-enriched membranes isolated from internal organelles', *Biochemical and Biophysical Research Communications*. Biochem Biophys Res Commun, 415(2), pp. 355–360. doi: 10.1016/j.bbrc.2011.10.072.

Puustinen, A. and Wikstrom, M. (1991) 'The heme groups of cytochrome o from Escherichia coli', *Proceedings of the National Academy of Sciences of the United States of America*. Proc Natl Acad Sci U S A, 88(14), pp. 6122–6126. doi: 10.1073/pnas.88.14.6122.

Rardin, M. J. *et al.* (2013) 'SIRT5 regulates the mitochondrial lysine succinylome and metabolic networks', *Cell Metabolism*. NIH Public Access, 18(6), pp. 920–933. doi: 10.1016/j.cmet.2013.11.013.

Rodriguez-Terrones, D. *et al.* (2020) 'A distinct metabolic state arises during the emergence of 2-cell-like cells', *EMBO reports*. EMBO, 21(1), p. e48354. doi: 10.15252/embr.201948354.

Röhm, K.-H. (1989) 'Butylmalonate is a transition state analogue for aminocylase I', *FEBS Letters*, 250(2), pp. 191–194. doi: 10.1016/0014-5793(89)80718-5.

Royer, L. O. *et al.* (2004) 'Succinylacetone oxidation by oxygen/peroxynitrite: a possible source of reactive intermediates in hereditary tyrosinemia type I.', *Chemical research in*

toxicology. American Chemical Society, 17(5), pp. 598–604. doi: 10.1021/tx0342520.
Sadler, A. J. *et al.* (2015) 'The acetyltransferase HAT1 moderates the NF-κB response by regulating the transcription factor PLZF', *Nature Communications*. Nature Publishing Group, 6. doi: 10.1038/ncomms7795.
Samuel, S. *et al.* (2008) 'Nucleolin binds specifically to an AP-1 DNA sequence and represses AP1-dependent transactivation of the matrix metalloproteinase-13 gene', *Molecular Carcinogenesis*. Mol Carcinog, 47(1), pp. 34–46. doi: 10.1002/mc.20358.
Sandilands, E. *et al.* (2007) 'Src kinase modulates the activation, transport and signalling dynamics of fibroblast growth factor receptors', *EMBO Reports*. European Molecular Biology Organization, 8(12), pp. 1162–1169. doi: 10.1038/sj.embor.7401097.
Sassa, S. and Kappas, A. (1983) 'Hereditary tyrosinemia and the heme biosynthetic pathway. Profound inhibition of δ-aminolevulinic acid dehydratase activity by succinylacetone', *Journal of Clinical Investigation*. J Clin Invest, 71(3), pp. 625–634. doi: 10.1172/JCI110809.
Sato, N. *et al.* (2004) 'Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor', *Nature Medicine*. Nat Med, 10(1), pp. 55–63. doi: 10.1038/nm979.
Scapoli, L. *et al.* (2004) 'Src-dependent ERK5 and Src/EGFR-dependent ERK1/2 activation is required for cell proliferation by asbestos', *Oncogene*. Nature Publishing Group, 23(3), pp. 805–813. doi: 10.1038/sj.onc.1207163.
Schoorlemmer, J. *et al.* (2014) 'Regulation of mouse retroelement MuERV-L/MERVL expression by REX1 and epigenetic control of stem cell potency', *Frontiers in Oncology*. Frontiers Research Foundation, 4 FEB, p. 14. doi: 10.3389/fonc.2014.00014.
Schüle, K. M. *et al.* (2019) 'GADD45 promotes locus-specific DNA demethylation and 2C cycling in embryonic stem cells', *Genes and Development*. Cold Spring Harbor Laboratory Press, 33(13–14), pp. 782–798. doi: 10.1101/gad.325696.119.
Schultz, E. (1996) 'Satellite cell proliferative compartments in growing skeletal muscles', *Developmental Biology*. Academic Press Inc., 175(1), pp. 84–94. doi: 10.1006/dbio.1996.0097.
Schulz, K. N. and Harrison, M. M. (2019) 'Mechanisms regulating zygotic genome activation', *Nature Reviews Genetics*. Nature Publishing Group, pp. 221–234. doi: 10.1038/s41576-018-0087-x.
Scinicariello, F. *et al.* (2007) 'Lead and δ-aminolevulinic acid dehydratase polymorphism: Where does it lead? A meta-analysis', *Environmental Health Perspectives*. National Institute of Environmental Health Sciences, 115(1), pp. 35–41. doi: 10.1289/ehp.9448.
Seguin, A. *et al.* (2017) 'Reductions in the mitochondrial ABC transporter Abcb10 affect the transcriptional profile of heme biosynthesis genes.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 292(39), pp. 16284–16299. doi: 10.1074/jbc.M117.797415.
Selak, M. A., Durán, R. V. and Gottlieb, E. (2006) 'Redox stress is not essential for the pseudo-hypoxic phenotype of succinate dehydrogenase deficient cells', *Biochimica et Biophysica Acta - Bioenergetics*. Biochim Biophys Acta, 1757(5–6), pp. 567–572. doi: 10.1016/j.bbabio.2006.05.015.
Sen, B. and Johnson, F. M. (2011) 'Regulation of Src Family Kinases in Human Cancers', *Journal of Signal Transduction*. Hindawi Limited, 2011, pp. 1–14. doi: 10.1155/2011/865819.
Senft, A. D. *et al.* (2018) 'Combinatorial Smad2/3 Activities Downstream of Nodal Signaling Maintain Embryonic/Extra-Embryonic Cell Identities during Lineage Priming', *Cell Reports*.

Elsevier B.V., 24(8), pp. 1977-1985.e7. doi: 10.1016/j.celrep.2018.07.077.
Severance, S. and Hamza, I. (2009) 'Trafficking of Heme and Porphyrins in Metazoa', *Chemical Reviews*, 109(10), pp. 4596–4616. doi: 10.1021/cr9001116.
Shahbazi, M. N. *et al.* (2016) 'Self-organization of the human embryo in the absence of maternal tissues', *Nature Cell Biology*. Nature Publishing Group, 18(6), pp. 700–708. doi: 10.1038/ncb3347.
Shahbazi, M. N., Siggia, E. D. and Zernicka-Goetz, M. (2019) 'Self-organization of stem cells into embryos: A window on early mammalian development', *Science*. American Association for the Advancement of Science, 364(6444), pp. 948–951. doi: 10.1126/science.aax0164.
Shen, L. and Zhang, Y. (2012) 'Enzymatic analysis of tet proteins: Key enzymes in the metabolism of DNA methylation', in *Methods in Enzymology*. Academic Press Inc., pp. 93–105. doi: 10.1016/B978-0-12-391940-3.00005-6.
Shetty, T. and Corson, T. W. (2020) 'Mitochondrial Heme Synthesis Enzymes as Therapeutic Targets in Vascular Diseases', *Frontiers in Pharmacology*. Frontiers Media S.A. doi: 10.3389/fphar.2020.01015.
Silva, J. *et al.* (2009) 'Nanog Is the Gateway to the Pluripotent Ground State', *Cell*. Cell, 138(4), pp. 722–737. doi: 10.1016/j.cell.2009.07.039.
Sinclair, P. R., Gorman, N. and Jacobs, J. M. (1999) 'Measurement of Heme Concentration', *Current Protocols in Toxicology*. Wiley, 00(1). doi: 10.1002/0471140856.tx0803s00.
Smestad, J. *et al.* (2018) 'Chromatin Succinylation Correlates with Active Gene Expression and Is Perturbed by Defective TCA Cycle Metabolism', *iScience*. Elsevier Inc., 2, pp. 63–75. doi: 10.1016/j.isci.2018.03.012.
Smith, A. (2017) 'Formative pluripotency: The executive phase in a developmental continuum', *Development (Cambridge)*. Company of Biologists Ltd, 144(3), pp. 365–373. doi: 10.1242/dev.142679.
Smith, J. R. *et al.* (2008) 'Robust, Persistent Transgene Expression in Human Embryonic Stem Cells Is Achieved with AAVS1-Targeted Integration', *Stem Cells*. Wiley, 26(2), pp. 496–504. doi: 10.1634/stemcells.2007-0039.
Sone, M. *et al.* (2017) 'Hybrid Cellular Metabolism Coordinated by Zic3 and Esrrb Synergistically Enhances Induction of Naive Pluripotency', *Cell Metabolism*, 25(5), pp. 1103-1117.e6. doi: 10.1016/j.cmet.2017.04.017.
Soriano, P. (1999) 'Generalized lacZ expression with the ROSA26 Cre reporter strain', *Nature Genetics*. Nature Publishing Group, 21(1), pp. 70–71. doi: 10.1038/5007.
Sperber, H. *et al.* (2015) 'The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition', *Nature Cell Biology*, 17(12), pp. 1523–1535. doi: 10.1038/ncb3264.
Sreedhar, A., Wiese, E. K. and Hitosugi, T. (2020) 'Enzymatic and metabolic regulation of lysine succinylation', *Genes & Diseases*. Chongqing University, 7(2), pp. 166–171. doi: 10.1016/j.gendis.2019.09.011.
Sun, F. *et al.* (2005) 'Crystal structure of mitochondrial respiratory membrane protein Complex II', *Cell*. Cell, 121(7), pp. 1043–1057. doi: 10.1016/j.cell.2005.05.025.
Sundström, L. *et al.* (2013) 'Succinate receptor GPR91, a Gαi coupled receptor that increases intracellular calcium concentrations through PLCβ', *FEBS Letters*. FEBS Lett, 587(15), pp. 2399–2404. doi: 10.1016/j.febslet.2013.05.067.
Surinya, K. H., Cox, T. C. and May, B. K. (1997) 'Transcriptional regulation of the human erythroid 5-aminolevulinate synthase gene. Identification of promoter elements and role

of regulatory proteins', *Journal of Biological Chemistry*. J Biol Chem, 272(42), pp. 26585–26594. doi: 10.1074/jbc.272.42.26585.
Suzuki, H. *et al.* (2004) 'Heme regulates gene expression by triggering Crm1-dependent nuclear export of Bach1', *EMBO Journal*. EMBO J, 23(13), pp. 2544–2553. doi: 10.1038/sj.emboj.7600248.
Sweeny, E. A. *et al.* (2018) 'Glyceraldehyde-3-phosphate dehydrogenase is a chaperone that allocates labile heme in cells', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology Inc., 293(37), pp. 14557–14568. doi: 10.1074/jbc.RA118.004169.
Swenson, S. A. *et al.* (2020) 'From Synthesis to Utilization: The Ins and Outs of Mitochondrial Heme', *Cells*. MDPI AG, 9(3), p. 579. doi: 10.3390/cells9030579.
Takahashi, K. and Yamanaka, S. (2006) 'Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors', *Cell*. Cell, 126(4), pp. 663–676. doi: 10.1016/j.cell.2006.07.024.
Takashima, Y. *et al.* (2014) 'Resetting transcription factor control circuitry toward ground-state pluripotency in human.', *Cell*. Elsevier, 158(6), pp. 1254–1269. doi: 10.1016/j.cell.2014.08.029.
Tanaka, S. *et al.* (1998) 'Promotion of trophoblast stem cell proliferation by FGF4.', *Science (New York, N.Y.)*. American Association for the Advancement of Science, 282(5396), pp. 2072–5. doi: 10.1126/science.282.5396.2072.
Tannahill, G. M. *et al.* (2013) 'Succinate is an inflammatory signal that induces IL-1β through HIF-1α', *Nature*. Nature, 496(7444), pp. 238–242. doi: 10.1038/nature11986.
Tejedo, J. R. *et al.* (2010) 'Low concentrations of nitric oxide delay the differentiation of embryonic stem cells and promote their survival', *Cell Death & Disease*. Nature Publishing Group, 1(10), pp. e80–e80. doi: 10.1038/cddis.2010.57.
Tesar, P. J. *et al.* (2007) 'New cell lines from mouse epiblast share defining features with human embryonic stem cells', *Nature*. Nature Publishing Group, 448(7150), pp. 196–199. doi: 10.1038/nature05972.
TeSlaa, T. *et al.* (2016) 'α-Ketoglutarate Accelerates the Initial Differentiation of Primed Human Pluripotent Stem Cells', *Cell metabolism*. NIH Public Access, 24(3), p. 485. doi: 10.1016/J.CMET.2016.07.002.
Theunissen, T. W. *et al.* (2014) 'Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency', *Cell Stem Cell*, 15(4), pp. 471–487. doi: 10.1016/j.stem.2014.07.002.
Thomson, J. A. *et al.* (1998) 'Embryonic stem cell lines derived from human blastocysts.', *Science*, 282(5391), pp. 1145–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9804556 (Accessed: 15 April 2019).
Thomson, M. *et al.* (2011) 'Pluripotency factors in embryonic stem cells regulate differentiation into germ layers', *Cell*. Cell, 145(6), pp. 875–889. doi: 10.1016/j.cell.2011.05.017.
Tobias, I. C. *et al.* (2018) 'Metabolic plasticity during transition to naïve-like pluripotency in canine embryo-derived stem cells', *Stem Cell Research*. Elsevier, 30, pp. 22–33. doi: 10.1016/J.SCR.2018.05.005.
Tohyama, S. *et al.* (2016) 'Glutamine Oxidation Is Indispensable for Survival of Human Pluripotent Stem Cells', *Cell Metabolism*. Cell Press, 23(4), pp. 663–674. doi: 10.1016/j.cmet.2016.03.001.
Tong, Z. *et al.* (2017) 'SIRT7 is an RNA-activated protein lysine deacylase', *ACS Chemical*

Biology. American Chemical Society, 12(1), pp. 300–310. doi: 10.1021/acschembio.6b00954.
Trebichalská, Z. and Holubcová, Z. (2020) 'Perfect date—the review of current research into molecular bases of mammalian fertilization', *Journal of Assisted Reproduction and Genetics*. Springer, pp. 243–256. doi: 10.1007/s10815-019-01679-4.
Tretter, L., Patocs, A. and Chinopoulos, C. (2016) 'Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis', *Biochimica et Biophysica Acta - Bioenergetics*. Elsevier B.V., 1857(8), pp. 1086–1101. doi: 10.1016/j.bbabio.2016.03.012.
Tsogtbaatar, E. *et al.* (2020) 'Energy Metabolism Regulates Stem Cell Pluripotency', *Frontiers in Cell and Developmental Biology*. Frontiers Media S.A., p. 87. doi: 10.3389/fcell.2020.00087.
Turkson, J. *et al.* (1998) 'Stat3 Activation by Src Induces Specific Gene Regulation and Is Required for Cell Transformation', *Molecular and Cellular Biology*. American Society for Microbiology, 18(5), pp. 2545–2552. doi: 10.1128/mcb.18.5.2545.
Ullah, Z. *et al.* (2008) 'Differentiation of trophoblast stem cells into giant cells is triggered by p57/Kip2 inhibition of CDK1 activity.', *Genes & development*. Cold Spring Harbor Laboratory Press, 22(21), pp. 3024–36. doi: 10.1101/gad.1718108.
Veenstra, G. J. C., Destrée, O. H. J. and Wolffe, A. P. (1999) 'Translation of Maternal TATA-Binding Protein mRNA Potentiates Basal but Not Activated Transcription in Xenopus Embryos at the Midblastula Transition', *Molecular and Cellular Biology*. American Society for Microbiology, 19(12), pp. 7972–7982. doi: 10.1128/mcb.19.12.7972.
Van Vranken, J. G. *et al.* (2015) 'Protein-mediated assembly of succinate dehydrogenase and its cofactors', *Critical Reviews in Biochemistry and Molecular Biology*. Informa Healthcare, pp. 168–180. doi: 10.3109/10409238.2014.990556.
Waddington, C. H. and Kacser, H. (1957) *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*. Edited by Allen and Unwin.
Waghray, A. *et al.* (2015) 'Tbx3 Controls Dppa3 Levels and Exit from Pluripotency toward Mesoderm', *Stem Cell Reports*. Cell Press, 5(1), pp. 97–110. doi: 10.1016/j.stemcr.2015.05.009.
Walker, E. *et al.* (2010) 'Polycomb-like 2 Associates with PRC2 and Regulates Transcriptional Networks during Mouse Embryonic Stem Cell Self-Renewal and Differentiation', *Cell Stem Cell*. Cell Stem Cell, 6(2), pp. 153–166. doi: 10.1016/j.stem.2009.12.014.
Wanet, A. *et al.* (2015) 'Connecting Mitochondria, Metabolism, and Stem Cell Fate', *Stem Cells and Development*, 24(17), pp. 1957–1971. doi: 10.1089/scd.2015.0117.
Wang, H. *et al.* (2008) 'Zonula occludens-1 (ZO-1) is involved in morula to blastocyst transformation in the mouse', *Developmental Biology*. Academic Press Inc., 318(1), pp. 112–125. doi: 10.1016/j.ydbio.2008.03.008.
Wang, Y. *et al.* (2017) 'KAT2A coupled with the α-KGDH complex acts as a histone H3 succinyltransferase', *Nature*. Nature Publishing Group, 552(7684), pp. 273–277. doi: 10.1038/nature25003.
Warburg, O., Wind, F. and Negelein, E. (1927) 'The metabolism of tumors in the body', *Journal of General Physiology*. The Rockefeller University Press, 8(6), pp. 519–530. doi: 10.1085/jgp.8.6.519.
Ware, C. B. *et al.* (2014) 'Derivation of naive human embryonic stem cells', *Proceedings of the National Academy of Sciences*, 111(12), pp. 4484–4489. doi: 10.1073/pnas.1319738111.
Warnatz, H. J. *et al.* (2011) 'The BTB and CNC homology 1 (BACH1) target genes are involved

in the oxidative stress response and in control of the cell cycle', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 286(26), pp. 23521–23532. doi: 10.1074/jbc.M111.220178.
Wilcox, A. J. *et al.* (2020) 'Preimplantation loss of fertilized human ova: estimating the unobservable', *Human reproduction (Oxford, England)*. NLM (Medline), 35(4), pp. 743–750. doi: 10.1093/humrep/deaa048.
Williams, R. L. *et al.* (1988) 'Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells', *Nature*. Nature, 336(6200), pp. 684–687. doi: 10.1038/336684a0.
Williamson, J. R., Anderson, J. and Browning, E. T. (1970) 'Inhibition of gluconeogenesis by butylmalonate in perfused rat liver.', *Journal of Biological Chemistry*. Elsevier, 245(7), pp. 1717–1726. doi: 10.1016/s0021-9258(19)77151-8.
Wobma, H. and Satwani, P. (2021) 'Mesenchymal Stromal Cells: Getting Ready for Clinical Primetime', *Transfusion and Apheresis Science*. Elsevier BV, p. 103058. doi: 10.1016/j.transci.2021.103058.
Wolpert, L. (2015) *Principles of Development*. Oxford University Press. Available at: https://books.google.be/books?id=hhlwAQAACAAJ.
Wu, J., Ocampo, A. and Belmonte, J. C. I. (2016) 'Cellular Metabolism and Induced Pluripotency', *Cell*. Cell Press, pp. 1371–1385. doi: 10.1016/j.cell.2016.08.008.
Xiao, M. *et al.* (2012) 'Inhibition of α-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors', *Genes and Development*. Genes Dev, 26(12), pp. 1326–1338. doi: 10.1101/gad.191056.112.
Xu, Z. *et al.* (2016) 'Wnt/β-catenin signaling promotes self-renewal and inhibits the primed state transition in naïve human embryonic stem cells', *Proceedings of the National Academy of Sciences*, 113(42), pp. E6382–E6390. doi: 10.1073/pnas.1613849113.
Yamada, S., Hill, M. and Takakuwa, T. (2015) 'Human Embryology', in *New Discoveries in Embryology*. InTech. doi: 10.5772/61453.
Yan, Y.-L. *et al.* (2019) 'DPPA2/4 and SUMO E3 ligase PIAS4 opposingly regulate zygotic transcriptional program', *PLOS Biology*. Edited by J. Wang. Public Library of Science, 17(6), p. e3000324. doi: 10.1371/journal.pbio.3000324.
Yang, G. *et al.* (2021) 'Histone acetyltransferase 1 is a succinyltransferase for histones and non-histones and promotes tumorigenesis', *EMBO reports*. EMBO, 22(2), p. e50967. doi: 10.15252/embr.202050967.
Yao, X. *et al.* (2010) 'Heme controls the regulation of protein tyrosine kinases Jak2 and Src', *Biochemical and Biophysical Research Communications*. Biochem Biophys Res Commun, 403(1), pp. 30–35. doi: 10.1016/j.bbrc.2010.10.101.
Ying, Q. L. *et al.* (2003) 'BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3', *Cell*. Cell Press, 115(3), pp. 281–292. doi: 10.1016/S0092-8674(03)00847-X.
Ying, Q. L. *et al.* (2008) 'The ground state of embryonic stem cell self-renewal', *Nature*. Nature Publishing Group, 453(7194), pp. 519–523. doi: 10.1038/nature06968.
Yoshinaga, K. and Adams, C. E. (1966) 'Delayed implantation in the spayed, progesterone treated adult mouse.', *Journal of reproduction and fertility*. Bioscientifica Ltd, 12(3), pp. 593–595. doi: 10.1530/jrf.0.0120593.
Young, R. A. (2011) 'Control of the embryonic stem cell state', *Cell*. Cell, pp. 940–954. doi: 10.1016/j.cell.2011.01.032.

Yu, G. *et al.* (2012) 'ClusterProfiler: An R package for comparing biological themes among gene clusters', *OMICS A Journal of Integrative Biology*. Mary Ann Liebert, Inc., 16(5), pp. 284–287. doi: 10.1089/omi.2011.0118.

Yu, J. *et al.* (2007) 'Induced pluripotent stem cell lines derived from human somatic cells', *Science*. Science, 318(5858), pp. 1917–1920. doi: 10.1126/science.1151526.

Zeevaert, K. *et al.* (2020) 'Cell Mechanics in Embryoid Bodies', *Cells*. NLM (Medline). doi: 10.3390/cells9102270.

Zenke-Kawasaki, Y. *et al.* (2007) 'Heme Induces Ubiquitination and Degradation of the Transcription Factor Bach1', *Molecular and Cellular Biology*. American Society for Microbiology, 27(19), pp. 6962–6971. doi: 10.1128/mcb.02415-06.

Zhang, J. *et al.* (2012) 'Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal.', *Cell stem cell*. NIH Public Access, 11(5), pp. 589–95. doi: 10.1016/j.stem.2012.10.005.

Zhang, J. *et al.* (2016) 'LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency', *Cell Stem Cell*. Cell Press, 19(1), pp. 66–80. doi: 10.1016/j.stem.2016.05.009.

Zhang, X. *et al.* (2018) 'Bach1: Function, regulation, and involvement in disease', *Oxidative Medicine and Cellular Longevity*. Hindawi Limited. doi: 10.1155/2018/1347969.

Zhang, Z. *et al.* (2011) 'Identification of lysine succinylation as a new post-translational modification', *Nature Chemical Biology*. Nature Publishing Group, 7(1), pp. 58–63. doi: 10.1038/nchembio.495.

Zhong, X. *et al.* (2019) 'Mitochondrial Dynamics Is Critical for the Full Pluripotency and Embryonic Developmental Potential of Pluripotent Stem Cells', *Cell Metabolism*, 29(4), pp. 979-992.e4. doi: 10.1016/j.cmet.2018.11.007.

Zhou, W. *et al.* (2012) 'HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition.', *The EMBO journal*. European Molecular Biology Organization, 31(9), pp. 2103–16. doi: 10.1038/emboj.2012.71.

Zhu, Y. *et al.* (2002) 'Heme deficiency interferes with the Ras-mitogen-activated protein kinase signaling pathway and expression of a subset of neuronal genes.', *Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research*. AACR, 13(9), pp. 431–9. Available at: http://cgd.aacrjournals.org/cgi/content/full/13/9/431 (Accessed: 20 August 2020).

Supplementary information

Supplementary table 1: List of succinylated peptide identified in mESC 2iL or 2iL + SA

Ddetraux210115-succi							Quantitative Value (Total Spectra)	Quantitative Value (Total Spectra)	Quantitative Value (Total Spectra)	Quantitative Value
							2IL	2IL	SA	SA
#	Identified Proteins (1189/1198)	Accession Number	Alternate ID	Molecular Weight	Fisher's Exact Test (p-value): *(p < 0,00305)	FC	2IL	2IL-210115	SA	SA-210115
1	Stress-70 protein, mitochondrial OS=Mus musculus OX=10090 GN=Hspa9 PE=1 SV=3	P38647	Hspa9	73 kDa	< 0,00010	0,9	106	200	80	181
2	60 kDa heat shock protein, mitochondrial OS=Mus musculus OX=10090 GN=Hspd1 PE=1 SV=1	P63038	Hspd1	61 kDa	< 0,00010	0,9	93	182	70	179
3	Cluster of Heat shock protein HSP 90-alpha OS=Mus musculus OX=10090 GN=Hsp90aa1 PE=1 SV=4 (P07901)	P07901 [3]	Hsp90aa1	85 kDa	0,036	1,6	34	110	57	167
4	Cluster of Heat shock protein HSP 90-beta OS=Mus musculus OX=10090 GN=Hsp90ab1 PE=1 SV=3 (P11499)	P11499 [3]	Hsp90ab1	83 kDa	0,046	1,6	35	93	55	144
5	Cluster of ADP/ATP translocase 2 OS=Mus musculus OX=10090 GN=Slc25a5 PE=1 SV=3 (P51881)	P51881 [6]	Slc25a5	33 kDa	0,49	1,3	34	92	33	130
6	Aconitate hydratase, mitochondrial OS=Mus musculus OX=10090 GN=Aco2 PE=1 SV=1	Q99KI0	Aco2	85 kDa	0,0047	1	50	121	38	126
7	ATP synthase subunit alpha, mitochondrial OS=Mus musculus OX=10090 GN=Atp5f1a PE=1 SV=1	Q03265	Atp5f1a	60 kDa	0,00037	0,9	53	140	50	123
8	Cluster of Heat shock cognate 71 kDa protein OS=Mus musculus OX=10090 GN=Hspa8 PE=1 SV=1 (P63017)	P63017 [13]	Hspa8	71 kDa	0,077	1,5	30	71	43	113
9	Cluster of Elongation factor 1-alpha OS=Mus musculus OX=10090 GN=Eef1a1 PE=2 SV=1 (Q3UA81)	Q3UA81 [2]	Eef1a1	50 kDa	0,16	1,5	33	70	41	110
10	Cluster of Glyceraldehyde-3-phosphate dehydrogenase OS=Mus musculus OX=10090 GN=GAPDH PE=2 SV=1 (D2KHZ9)	D2KHZ9 [4]	GAPDH	36 kDa	0,1	1,5	30	68	44	105
11	Isocitrate dehydrogenase [NADP], mitochondrial OS=Mus musculus OX=10090 GN=Idh2 PE=1 SV=3	P54071	Idh2	51 kDa	0,01	1	35	100	28	101
12	Aspartate aminotransferase, mitochondrial OS=Mus musculus OX=10090 GN=Got2 PE=1 SV=1	P05202	Got2	47 kDa	0,011	1	43	87	32	92
13	Elongation factor 2 OS=Mus musculus OX=10090 GN=Eef2 PE=1 SV=2	P58252 (+4)	Eef2	95 kDa	0,18	1,5	14	61	22	89
14	Cluster of Citrate synthase, mitochondrial OS=Mus musculus OX=10090 GN=Cs PE=1 SV=1 (Q9CZU6)	Q9CZU6 [2]	Cs	52 kDa	< 0,00010	0,7	31	126	27	88

15	Malate dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Mdh2 PE=1 SV=3	P08249	Mdh2	36 kDa	0,019	1	43	83	36	87
16	Nucleolin OS=Mus musculus OX=10090 GN=Ncl PE=1 SV=2	P09405 (+1)	Ncl	77 kDa	< 0,00010	3,2	17	33	75	85
17	Cluster of Fructose-bisphosphate aldolase A OS=Mus musculus OX=10090 GN=Aldoa PE=1 SV=2 (P05064)	P05064 [5]	Aldoa	39 kDa	0,24	1,4	24	54	30	82
18	Glutamate dehydrogenase 1, mitochondrial OS=Mus musculus OX=10090 GN=Glud1 PE=1 SV=1	P26443	Glud1	61 kDa	0,016	0,9	29	81	25	79
19	Heat shock protein 1 (Chaperonin 10) OS=Mus musculus OX=10090 GN=Hspe1 PE=1 SV=1	Q4KL76 (+1)	Hspe1	11 kDa	< 0,00010	0,7	42	113	30	72
20	Isoform 2 of Serine hydroxymethyltransferase, mitochondrial OS=Mus musculus OX=10090 GN=Shmt2	Q9CZN7-2	Shmt2	55 kDa	0,00055	0,8	35	88	29	72
21	Trifunctional enzyme subunit alpha, mitochondrial OS=Mus musculus OX=10090 GN=Hadha PE=1 SV=1	Q8BMS1	Hadha	83 kDa	0,00056	0,8	41	81	33	67
22	Cluster of Alpha-enolase OS=Mus musculus OX=10090 GN=Eno1 PE=1 SV=3 (P17182)	P17182 [5]	Eno1	47 kDa	0,15	1,5	16	46	31	64
23	Cluster of Heat shock 70 kDa protein 4 OS=Mus musculus OX=10090 GN=Hspa4 PE=1 SV=1 (Q3U2G2)	Q3U2G2 [4]	Hspa4	94 kDa	0,00021	2,6	8	24	23	61
24	ATP synthase subunit beta, mitochondrial OS=Mus musculus OX=10090 GN=Atp5f1b PE=1 SV=2	P56480	Atp5f1b	56 kDa	0,008	0,9	30	66	26	59
25	Cluster of MCG68069 OS=Mus musculus OX=10090 GN=Npm1 PE=1 SV=1 (Q5SQB7)	Q5SQB7 [3]	Npm1	33 kDa	0,0046	2	22	25	38	58
26	Presequence protease, mitochondrial OS=Mus musculus OX=10090 GN=Pitrm1 PE=1 SV=1	Q8K411 (+2)	Pitrm1	117 kDa	0,15	1,1	13	54	13	58
27	Leucine-rich PPR motif-containing protein, mitochondrial OS=Mus musculus OX=10090 GN=Lrpprc PE=1 SV=2	Q6PB66	Lrpprc	157 kDa	0,028	0,9	15	62	16	55
28	Cluster of Peptidyl-prolyl cis-trans isomerase A OS=Mus musculus OX=10090 GN=Ppia PE=1 SV=2 (P17742)	P17742 [4]	Ppia	18 kDa	0,45	1,2	21	47	29	55
29	Cluster of Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial OS=Mus musculus OX=10090 GN=Pdha1 PE=1 SV=1 (P35486)	P35486 [4]	Pdha1	43 kDa	0,027	0,9	24	55	20	53
30	HATPase_c domain-containing protein OS=Mus musculus OX=10090 GN=Trap1 PE=2 SV=1	Q3TK29 (+4)	Trap1	80 kDa	0,076	1	21	47	15	52
31	Cluster of Serine beta-lactamase-like protein LACTB, mitochondrial OS=Mus musculus OX=10090 GN=Lactb PE=1 SV=1 (A0A1L1SVF9)	A0A1L1SVF9 [4]	Lactb	39 kDa	0,0064	0,8	29	51	16	51
32	17beta-hydroxysteroid dehydrogenase type 10/short chain L-3-hydroxyacyl-CoA dehydrogenase OS=Mus musculus OX=10090 GN=Hsd17b10 PE=1 SV=1	Q99N15 (+1)	Hsd17b10	27 kDa	0,16	1,1	16	48	17	51

33	Stress-induced-phosphoprotein 1 OS=Mus musculus OX=10090 GN=Stip1 PE=1 SV=1	Q60864	Stip1	63 kDa	0,36	1,2	20	40	20	51
34	ATP synthase subunit d, mitochondrial OS=Mus musculus OX=10090 GN=Atp5pd PE=1 SV=3	Q9DCX2	Atp5pd	19 kDa	0,00013	0,7	21	68	12	50
35	Cluster of Actin, cytoplasmic 2 OS=Mus musculus OX=10090 GN=Actg1 PE=1 SV=1 (P63260)	P63260 [6]	Actg1	42 kDa	0,16	1,6	13	33	22	50
36	Isoform 2 of Cytosol aminopeptidase OS=Mus musculus OX=10090 GN=Lap3	Q9CPY7-2	Lap3	53 kDa	0,17	1,1	19	38	12	48
37	Cluster of 14-3-3 protein zeta/delta OS=Mus musculus OX=10090 GN=Ywhaz PE=1 SV=1 (P63101)	P63101 [11]	Ywhaz	28 kDa	0,4	1,4	5	38	11	48
38	60S ribosomal protein L29 OS=Mus musculus OX=10090 GN=Rpl29 PE=1 SV=2	P47915 (+1)	Rpl29	18 kDa	0,00011	3,3	2	16	13	47
39	Cluster of Tripartite motif-containing 28 OS=Mus musculus OX=10090 GN=Trim28 PE=2 SV=1 (Q5EBP9)	Q5EBP9 [3]	Trim28	89 kDa	0,00018	3	7	16	21	47
40	Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial OS=Mus musculus OX=10090 GN=Sdha PE=1 SV=1	Q8K2B3	Sdha	73 kDa	0,017	0,9	16	53	13	47
41	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Hadh PE=1 SV=2	Q61425	Hadh	34 kDa	0,33	1,2	21	34	17	47
42	60S ribosomal protein L7a OS=Mus musculus OX=10090 GN=Rpl7a PE=1 SV=2	P12970 (+2)	Rpl7a	30 kDa	< 0,00010	3,4	5	14	19	45
43	Isoform Mt-VDAC1 of Voltage-dependent anion-selective channel protein 1 OS=Mus musculus OX=10090 GN=Vdac1	Q60932-2	Vdac1	31 kDa	0,01	2,3	4	19	8	45
44	Cluster of T-complex protein 1 subunit alpha OS=Mus musculus OX=10090 GN=Tcp1 PE=1 SV=3 (P11983)	P11983 [3]	Tcp1	60 kDa	0,23	1,5	6	30	10	45
45	Dihydrolipoyl dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Dld PE=1 SV=2	O08749	Dld	54 kDa	0,022	0,9	21	47	16	44
46	Cluster of Histone H1.3 OS=Mus musculus OX=10090 GN=H1-3 PE=1 SV=2 (P43277)	P43277 [7]	H1-3	22 kDa	0,4	1,2	25	30	23	43
47	L-threonine 3-dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Tdh PE=1 SV=1	Q8K3F7	Tdh	41 kDa	0,15	1	13	38	10	42
48	Protein disulfide-isomerase A3 OS=Mus musculus OX=10090 GN=Pdia3 PE=1 SV=2	P27773	Pdia3	57 kDa	0,0002	3,4	2	14	14	40
49	Cluster of Uncharacterized protein OS=Mus musculus OX=10090 GN=Cct8 PE=2 SV=1 (Q8BVY8)	Q8BVY8 [3]	Cct8	60 kDa	0,061	1,9	5	22	11	40
50	Cluster of 40S ribosomal protein S25 OS=Mus musculus OX=10090 GN=Rps25 PE=1 SV=1 (P62852)	P62852 [3]	Rps25	14 kDa	0,43	1,4	11	28	13	40
51	Cluster of Isoform 2 of Protein SET OS=Mus musculus OX=10090 GN=Set (Q9EQU5-2)	Q9EQU5-2 [3]	Set	32 kDa	0,0097	2,2	9	18	21	39

52	Cluster of Uncharacterized protein OS=Mus musculus OX=10090 GN=Hnrnpu PE=2 SV=1 (Q3TVV6)	Q3TVV6 [5]	Hnrnpu	88 kDa	< 0,00010	25	0	2	11	38
53	Eukaryotic translation elongation factor 1 gamma OS=Mus musculus OX=10090 GN=Eef1g PE=1 SV=1	Q4FZK2 (+1)	Eef1g	50 kDa	0,08	1,8	4	26	16	38
54	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Acadm PE=1 SV=1	P45952 (+1)	Acadm	46 kDa	0,11	1	14	35	10	38
55	Cluster of Ubiquitin-40S ribosomal protein S27a OS=Mus musculus OX=10090 GN=Rps27a PE=1 SV=2 (P62983)	P62983 [5]	Rps27a	18 kDa	0,18	1	14	34	12	38
56	Aldehyde dehydrogenase, mitochondrial OS=Mus musculus OX=10090 GN=Aldh2 PE=1 SV=1	P47738 (+5)	Aldh2	57 kDa	0,53	1,3	10	28	11	38
57	Cluster of ATP synthase subunit O, mitochondrial OS=Mus musculus OX=10090 GN=Atp5po PE=1 SV=1 (Q9DB20)	Q9DB20 [3]	Atp5po	23 kDa	0,00013	0,6	19	54	10	37
58	ATP synthase subunit gamma OS=Mus musculus OX=10090 GN=Atp5c1 PE=1 SV=1	Q8C2Q8	Atp5c1	30 kDa	0,041	0,9	9	44	10	37
59	Adenylate kinase 4, mitochondrial OS=Mus musculus OX=10090 GN=Ak4 PE=1 SV=1	Q3U489 (+1)	Ak4	25 kDa	0,00017	0,7	28	57	23	36
60	Fumarate hydratase, mitochondrial OS=Mus musculus OX=10090 GN=Fh PE=1 SV=3	P97807 (+1)	Fh	54 kDa	0,02	0,8	17	41	13	36
61	Cluster of S1 motif domain-containing protein OS=Mus musculus OX=10090 GN=Pnpt1 PE=2 SV=1 (Q3TST0)	Q3TST0 [4]	Pnpt1	86 kDa	0,11	1	7	36	5	36
62	Histone H1.5 OS=Mus musculus OX=10090 GN=H1-5 PE=1 SV=2	P43276 (+1)	H1-5	23 kDa	0,21	1,6	13	18	13	36
63	Cluster of Cofilin-1 OS=Mus musculus OX=10090 GN=Cfl1 PE=1 SV=3 (P18760)	P18760 [4]	Cfl1	19 kDa	0,47	1,3	9	27	12	36
64	Histone H1.1 OS=Mus musculus OX=10090 GN=H1-1 PE=1 SV=2	P43275	H1-1	22 kDa	0,0019	2,9	5	11	12	35
65	Cluster of T-complex protein 1 subunit eta OS=Mus musculus OX=10090 GN=Cct7 PE=1 SV=1 (P80313)	P80313 [9]	Cct7	60 kDa	0,027	2,3	3	14	4	35
66	Cluster of Peptidylprolyl isomerase OS=Mus musculus OX=10090 GN=Fkbp3 PE=1 SV=1 (Q3UBU9)	Q3UBU9 [3]	Fkbp3	25 kDa	< 0,00010	4,2	5	7	16	34
67	Cluster of GTP-binding nuclear protein Ran OS=Mus musculus OX=10090 GN=Ran PE=1 SV=3 (P62827)	P62827 [4]	Ran	24 kDa	0,11	1,8	8	20	15	34
68	Cluster of Keratin 5 OS=Mus musculus OX=10090 GN=Krt5 PE=1 SV=2 (Q32P04)	Q32P04 [21]	Krt5	62 kDa	0,062	0,9	32	20	15	33
69	40S ribosomal protein S3a OS=Mus musculus OX=10090 GN=Rps3a PE=1 SV=3	P97351 (+1)	Rps3a	30 kDa	0,21	1,6	10	21	16	33
70	Acetyl-CoA acetyltransferase, mitochondrial OS=Mus musculus OX=10090 GN=Acat1 PE=1 SV=1	Q8QZT1	Acat1	45 kDa	0,32	1,1	14	25	11	33

71	Endoplasmic reticulum chaperone BiP OS=Mus musculus OX=10090 GN=Hspa5 PE=1 SV=3	P20029 (+2)	Hspa5	72 kDa	0,0019	3	3	12	13	32
72	Peroxiredoxin-1 OS=Mus musculus OX=10090 GN=Prdx1 PE=1 SV=1	P35700	Prdx1	22 kDa	0,29	1,5	9	21	13	32
73	Fatty acid synthase OS=Mus musculus OX=10090 GN=Fasn PE=1 SV=1	A0A0U1RNJ1 (+1)	Fasn	272 kDa	0,11	1,9	4	16	6	31
74	Electron transfer flavoprotein subunit alpha, mitochondrial OS=Mus musculus OX=10090 GN=Etfa PE=1 SV=2	Q99LC5	Etfa	35 kDa	0,2	1	12	26	8	31
75	Ornithine aminotransferase, mitochondrial OS=Mus musculus OX=10090 GN=Oat PE=1 SV=1	P29758 (+3)	Oat	48 kDa	0,51	1,2	3	25	4	31
76	Peptidyl-prolyl cis-trans isomerase B OS=Mus musculus OX=10090 GN=Ppib PE=1 SV=2	P24369 (+1)	Ppib	24 kDa	0,029	2,2	7	12	12	30
77	Endoplasmin OS=Mus musculus OX=10090 GN=Hsp90b1 PE=1 SV=2	P08113 (+3)	Hsp90b1	92 kDa	0,061	2,1	3	15	7	30
78	Cluster of Uncharacterized protein (Fragment) OS=Mus musculus OX=10090 PE=2 SV=1 (Q8BNM0)	Q8BNM0 [4]		21 kDa	0,19	1,6	13	19	21	30
79	Cluster of Tubulin alpha-1A chain OS=Mus musculus OX=10090 GN=Tuba1a PE=1 SV=1 (P68369)	P68369 [5]	Tuba1a	50 kDa	0,42	1,4	7	25	14	30
80	Cluster of T-complex protein 1 subunit zeta OS=Mus musculus OX=10090 GN=Cct6a PE=1 SV=3 (P80317)	P80317 [6]	Cct6a	58 kDa	0,43	1,4	5	21	6	30
81	T-complex protein 1 subunit epsilon OS=Mus musculus OX=10090 GN=Cct5 PE=1 SV=1	P80316	Cct5	60 kDa	0,0003	4,5	0	8	7	29
82	Superoxide dismutase [Mn], mitochondrial OS=Mus musculus OX=10090 GN=Sod2 PE=1 SV=3	P09671 (+1)	Sod2	25 kDa	0,027	0,8	16	35	14	29
83	Cluster of Succinate--CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial OS=Mus musculus OX=10090 GN=Suclg1 PE=1 SV=4 (Q9WUM5)	Q9WUM5 [2]	Suclg1	36 kDa	0,2	1	14	24	10	29
84	Cluster of Eukaryotic translation initiation factor 5A-1 OS=Mus musculus OX=10090 GN=Eif5a PE=1 SV=2 (P63242)	P63242 [3]	Eif5a	17 kDa	0,35	1,1	9	26	11	29
85	Glycine--tRNA ligase OS=Mus musculus OX=10090 GN=Gars1 PE=1 SV=1	Q9CZD3	Gars1	82 kDa	0,49	1,2	9	24	12	29
86	Transaldolase OS=Mus musculus OX=10090 GN=Taldo1 PE=1 SV=1	A0A1B0GR11 (+1)	Taldo1	42 kDa	0,00027	4,9	0	7	6	28
87	Cluster of Heterogeneous nuclear ribonucleoprotein A1 OS=Mus musculus OX=10090 GN=Hnrnpa1 PE=1 SV=2 (P49312)	P49312 [6]	Hnrnpa1	34 kDa	0,026	2,4	3	11	6	28
88	Pyruvate kinase PKM OS=Mus musculus OX=10090 GN=Pkm PE=1 SV=4	P52480	Pkm	58 kDa	0,19	1,6	9	18	16	28

89	Aminomethyltransferase OS=Mus musculus OX=10090 GN=Amt PE=1 SV=1	A2RSW6 (+1)	Amt	44 kDa	0,27	1,1	8	28	11	28
90	Isocitrate dehydrogenase [NAD] subunit, mitochondrial OS=Mus musculus OX=10090 GN=Idh3b PE=1 SV=1	Q91VA7	Idh3b	42 kDa	0,34	1,1	6	21	2	28
91	Isocitrate dehydrogenase [NAD] subunit, mitochondrial OS=Mus musculus OX=10090 GN=Idh3a PE=1 SV=1	A0A1L1STE6 (+1)	Idh3a	42 kDa	0,0055	0,8	18	38	15	27
92	Cluster of Ribosomal protein L4 OS=Mus musculus OX=10090 GN=Rpl4 PE=1 SV=1 (Q564E8)	Q564E8 [3]	Rpl4	47 kDa	0,0094	2,8	2	11	9	27
93	Cluster of Enoyl-CoA delta isomerase 2, mitochondrial OS=Mus musculus OX=10090 GN=Eci2 PE=1 SV=1 (Q3TCD4)	Q3TCD4 [4]	Eci2	42 kDa	0,14	1	14	24	10	27
94	Cluster of Acyl-coenzyme A thioesterase 9, mitochondrial OS=Mus musculus OX=10090 GN=Acot9 PE=1 SV=1 (Q9R0X4)	Q9R0X4 [3]	Acot9	51 kDa	0,4	1,2	4	22	3	27
95	Uncharacterized protein OS=Mus musculus OX=10090 GN=Slc25a3 PE=2 SV=1	Q3THU8 (+1)	Slc25a3	40 kDa	0,42	1,2	7	25	11	27
96	Cluster of HABP4_PAI-RBP1 domain-containing protein OS=Mus musculus OX=10090 GN=Serbp1 PE=2 SV=1 (Q3UJK2)	Q3UJK2 [5]	Serbp1	43 kDa	< 0,00010	3,8	6	7	24	26
97	GrpE protein homolog 1, mitochondrial OS=Mus musculus OX=10090 GN=Grpel1 PE=1 SV=1	Q99LP6	Grpel1	24 kDa	0,0029	0,7	15	37	10	26
98	High mobility group protein B2 OS=Mus musculus OX=10090 GN=Hmgb2 PE=1 SV=3	P30681 (+3)	Hmgb2	24 kDa	0,08	1,8	14	12	22	26
99	Cluster of Uncharacterized protein OS=Mus musculus OX=10090 GN=Rpl3 PE=2 SV=1 (Q3T9U9)	Q3T9U9 [4]	Rpl3	46 kDa	0,1	1,9	3	14	7	26
100	Prohibitin OS=Mus musculus OX=10090 GN=Phb PE=1 SV=1	P67778	Phb	30 kDa	0,24	1,1	8	27	11	26