UNIVERSITAT DE BARCELONA

Mediator complex and transcriptional control of cellular identity and plasticity

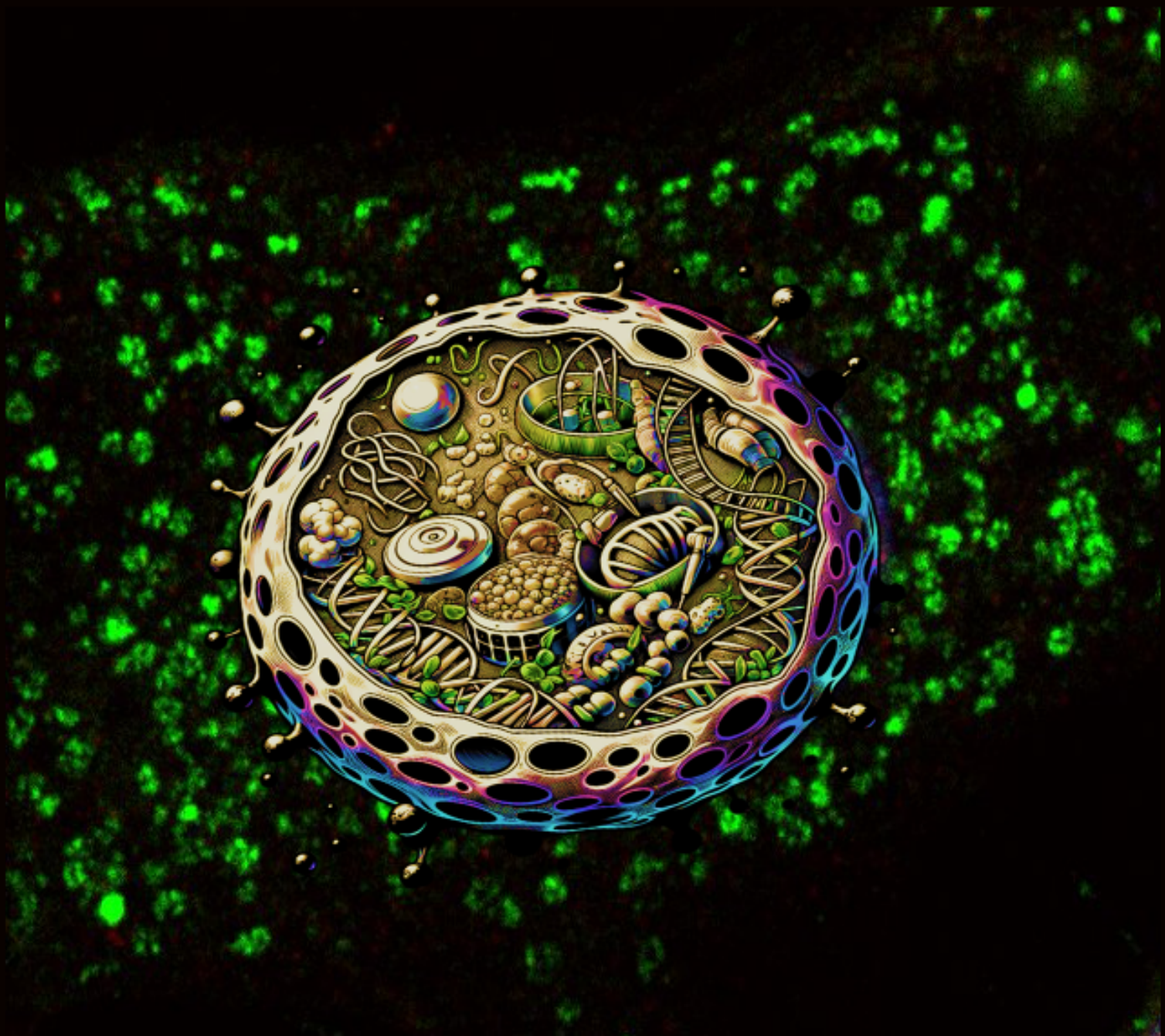
Isabel Calvo Serrano

UNIVERSITAT DE BARCELONA

DOCTORAL THESIS

Mediator Complex and Transcriptional Control of Cellular Identity and Plasticity

Isabel Calvo Serrano
Barcelona, 2023



MEDIATOR COMPLEX AND THE TRANSCIPTIONAL CONTROL OF CELLULAR IDENTITY AND PLASTICITY

DOCTORAL THESIS

Doctoral Programme in Biomedicine

Thesis submitted by **Isabel Calvo Serrano** to qualify for the degree of Doctor from the Universitat de Barcelona.

The work presented in this thesis has been carried out in the Cellular Plasticity and Disease laboratory at the Biomedical Research Institute of Barcelona (IRB) under the guidance and supervision of **Dr. Manuel Serrano Marugán**.

Barcelona, 2023

Dr. Manuel Serrano Marugán
Thesis Director

Dr. Manuel Palacín Prieto
Thesis Tutor

Isabel Calvo Serrano
PhD Student

“You cannot teach a man anything;
you can only help him discover it in himself.”

Galileo Galilei

“*Per aspera ad Astra”*

Séneca

A mis padres, a mi hermana y a mis tíos

Acknowledgements

The journey that has brought me to the completion of this thesis has been a great adventure that would not have been possible, nor would I have enjoyed it to the same extent, without all of you. I want to express my sincere gratitude to all the people who, in one way or another, have contributed to this beautiful process.

Manolo, I want to express my deepest gratitude to you, not only as my thesis supervisor but as a mentor and family figure -more than an uncle, a constant presence since my earliest memories. Your guidance, support, and passion for science have been instrumental throughout this research journey. You've not only shaped my academic pursuits but also inspired a profound love for molecular biology. Instilling in me a passion for the pursuit of knowledge. I am immensely thankful for your support, both personally and professionally.

To Cian, the first person who guided me from the basics of pipetting. Cian, you've been more than a mentor. At the end I got to understand your impactful lesson, 'you must suffer´. You've been a constant companion, sharing in both the celebrations and challenges of everyday life, from the early days in Madrid to the transition to Barcelona and now across the distance from Cambridge. Your passion and unwavering dedication to science have been a great source of inspiration. Your mentorship has taught me the importance of maintaining emotional equilibrium, not letting oneself be swayed excessively by highs or lows. I am profoundly grateful for the time and effort you've invested in me, Cian. Your impact on my journey has been immeasurable, and I carry your teachings with me as I continue to grow both personally and professionally.

To Manuel Palacin, my thesis advisor, for showing interest in the topics I investigated. Our conversations and discussions about my work, and your guidance has played an important role in my academic journey.

To my CNIO laboratory companions in Madrid and Barcelona: Miguel, Raquel, Noe, Dafni, Selim, and Gianluca, you were my first inspirations, enjoying good times together, even outside the lab, like our moments at the bowling. Thank you for the lasting memories and the supportive environment you provided. Dafni, you've been a true example of scientific perseverance and how things are done right in a laboratory. Thank you for always helping after each lab meeting and for taking an interest in my projects. Raquel, my CDK8 team partner! We've faced challenges together. You've helped me grasp the

necessary concepts at the beginning of my project and navigate both the experiments and the socials in the laboratory. Thanks for the good times and the support.

To my thesis colleagues Elena, Ines, Fernanda, Valentina and Vanessa: with whom I have shared invaluable moments, from joys to less happy times in the laboratory. I appreciate the celebrations and good times we have shared. Elena, thank you for being an example not only in the scientific realm but also for your constant joy and energy. You've been a companion in our adventures, from climbing in nature to enjoying beers on the terrace in front of the Sagrada Familia. Valentina, "Hacemolo," my Italianini, I carry your friendship with me forever. Your camaraderie during late nights of shared work, and your Christmas sweets. Not only are you one of the brightest individuals I know in the scientific realm, but your qualities as a person are truly priceless. May nothing separate what quarantine and COVID brought together! Vanessa, arguably one of the most hardworking and systematic individuals I know, and also my friend and of course my gym buddy (gymsis). We've shared everything from late-night hours working in the lab to express trips to Andorra thank you for being yourself.

Mar, Maribel, Fede, Jose-Alberto, Mate, Marta, Kathleen, Nayuta,: an incredible team of mentors and colleagues, I've learned so much from each of you, and we've truly had a great time together. Maribel, thank you for all your guidance with the animal facility work, for the shared moments and conversations about life. Mar, thank you for being our mother in the lab, for scolding and helping us when needed, and for paying attention to every detail. Fede, even though our time in the lab was short, I keep your friendship forever, your mark is indelible—one in a million! José Alberto, what luck to have met and learned from you. Thanks for accompanying me on ketogenic adventures, for being an example and a friend. Mate, probably one of the most unique and special minds I've encountered, always questioning everything, and prompting you to reconsider any knowledge in search of reality. Marta, one of the brightest individuals, a vibrant and restless mind in all aspects of life. I wish there were more people like you.

Laia, I want to express my deep gratitude for all the bioinformatics assistance you've provided. Your involvement and dedication in understanding the biological foundations have been essential. Your contribution has given purpose and depth to this study. Thank you for your invaluable help.

I also want to express my gratitude to all the people who have been crucial to the development of this thesis, even those outside the work environment. Adri Candelas, my friend from university, with whom this scientific adventure began, and with whom I have also shared good times in Barcelona. To Marc, Don Andrés and Godoy my Master´s colleges, my friends, who are always there, thanks Barcelona for letting me know you! To my flatmates when I arrived in Barcelona, Steven, Carmen, Manu, and of course, Laura and Gonzalo, you were my first real adult life companions! Thank you all for being part of this journey.

Also, to each and every one of my friends, the family you choose, those who have always been and will always be there since Madrid, to the new people I have met in Barcelona, and, in general, to all of you who provide me with energy, support in difficult times, and joy in the good ones. To the ones that I love and who love me. Thank you for being a vital part of my life.

Finally, I want to express my gratitude to my family, who have shaped the person I am today from the very beginning. They have taught me to be consistent and strong, to finish what I start, and to set high goals that keep me inspired and fulfilled in life.

To my mother, the person who knows me the most, whose hard work has shaped me and who has seen me through the process of writing and presenting my own thesis. She gives me a reality check when needed and supports and encourages me in everything I set out to do.

To my father, an example of mental strength and perseverance, with whom I have long conversations that help me grow as a person in all aspects. He supports and takes an interest in each of my passions, whatever they may be.

To my sister Sofia, my greatest inspiration, with the purest heart I know—a resilient and strong person, yet sensitive at the same time. You will achieve everything you set out to do, and I will always be by your side in good times and bad.

To Esteban, with whom I have spent countless hours not only learning mathematics but discovering sports, music, and literature. Thank you for inspiring me to take an interest in every single thing around me.

Summary

Our understanding of the transcriptional control of cellular identity is rapidly emerging. Mediator complex enriches at active enhancers, forms a protein bridge to target genes, and recruits RNA Polymerase II (RNA-Pol II). Thus, Mediator plays an essential role in transcriptional regulation. However, precisely how Mediator complex completes these critical steps remains unclear. In this study, we delve into the intricacies of Mediator interactions with newfound transcriptional regulators. For this, we establish and employ a proximity-ligation assay (PLA) to characterize protein interactions in the nucleus by immunofluorescence and quantitative imaging. We detect changes in the Mediator-Pol II interaction between different cell types. Moreover, we show that small molecule manipulation of the Mediator-Pol II interaction correlates with subsequent changes in cell identity. This suggests an important role for the Mediator-Pol II interaction in establishing cell identity transitions. We chose Mediator´s biggest subunit: MED1 (Mediator of RNA polymerase II transcription subunit 1) and performed immunoprecipitation followed by mass spectrometry (IP-MS) analysis of mouse ESCs with the aim of identifying its protein interactome. We observe a strong overlap between the proteins and biological functions of our mouse Mediator interactome and previously reported Mediator interactomes. Beyond the already know interactors. We present an updated interaction network of MED1 with novel proteins connected with Mediator. We further explored the particularly strong interaction with all the constitutive subunits of the Pyruvate Dehydrogenase complex (PDH). We provide evidence that the nuclear PDH is not randomly distributed in the nucleus, but instead tightly associates with Mediator and RNA-Pol II. Chromatin Immunoprecipitation combined with DNA Sequencing (ChIP-Seq) of the DLAT subunit of PDH complex confirms that the nuclear PDH complex is enriched at enhancers and super- enhancers, it coincides with highly acetylated regions of the DNA and is involved in biological processes that regulate cellular identity and chromatin maintenance and regulation. Experimental manipulation of the PDH complex reflects the functional significance of this nuclear complex in the transcriptional regulation of cell identity and viability independently of the mitochondrial PDH activity. Taken together, our data provide novel insights about how the Mediator-PDH axis controls cell identity.

Table of contents

Table of contents

Abbreviations

2i	MEK inhibitor + GSK3ß inhibitor
5-METH-THF	from 5-methyl-tetrahydrofolate
AHCY	Adenosylhomocysteinase
ACLY	ATP citrate synthase
ACSS1	Acetyl-CoA Synthetase 1
ACSS2	Acetyl-CoA Synthetase 2
ASCS	Adult Stem Cells
ATP	Adenosine Triphosphate
BRE	TFIIB recognition element
BHMT	Betaine-Momocysteine S-methyltransferase
BRD4	Bromodomain Containing 4
cDNA	Complementary DNA
CDK8	Cyclin dependent Kinase 8
CDK19	Cyclin Dependent Kinase 19
CDK8i	Cyclin dependent Kinase 8/19 Inhibitor
CBS	Cystathionine Beta-Synthase
ChIP-Seq	Chromatin Immunoprecipitation combined with DNA Sequencing
CRMs	Cis-Regulatory Modules
CTD	Carboxyl Terminal Domain
DHF	Dihydrofolate
DLAT	Dihydrolipoamide S-Acetyltransferase
DLD	Dihydrolipoamide Dehydrogenase
DNA	Deoxyribonucleic Acid
DPE	Downstream Promoter Element
EMT	Epithelial-to-Mesenchymal Transition
EPI	Epiblast
ES	Embryonic Stem
ESCs	Embryonic Stem Cells
FADH	Flavin Adenine Dinucleotide
FRET	Fluorescence Resonance Energy Transfer
FBS	Fetal Bovine Serum
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GFP	Green Fluorescent Protein

GSK3ß	Glycogen Synthase Kinase 3 Beta
GSK3ßi	Glycogen Synthase Kinase 3 Beta Inhibitor
GO	Gene Ontology
HATs	Histone acetyltransferases
HCys	Homocysteine
ICM	Inner Cell Mass
IDRs	Intrinsically Disordered Regions
Inr	Initiator
IP-MS	Immunoprecipitation followed by Mass Spectrometry
LncRNAs	long non-coding RNAs
LIF	Leukemia Inhibitory Factor
LLPS	Liquid-Liquid Phase Separation
MAPK	Mitogen-Activated Protein Kinase
MEF	Mouse Embryonic Fibroblasts
MEKi	MAPK/ERK Kinase Inhibitor
MED1	Mediator of RNA polymerase II transcription subunit 1
METH	Methamphetamine
Meth	Methionine
MAT2A	Methionine Adenosyltransferase 2A
MPC	Mitochondrial Pyruvate Carrier
MPC1	Mitochondrial Pyruvate Carrier 1
MPC2	Mitochondrial Pyruvate Carrier 2
MS	Methionine Synthase
MTs	Methyltransferases
mRNA	Messenger RNA
NADH	Nicotinamide Adenine Dinucleotide
NcRNAs	Non-Coding RNAs
NUDIX5	Nucleoside Diphosphate Linked Moiety X-type motif 5
PBS	Phosphate-Buffered Saline
PDHA1	Pyruvate Dehydrogenase E1 Subunit Alpha 1
PDHB	Pyruvate Dehydrogenase E1 Subunit Beta
PDH	Pyruvate Dehydrogenase
PE	Primitive Endoderm

PLA	Proximity Ligation Assay
PSCs	Pluripotent Stem Cells
PTMs	Post-Translational Modifications
qRT-PCR	Quantitative Real Time PCR
rRNA	Ribosomal RNA
RA	Retinoic Acid
RNA	Ribonucleic Acid
RNA Pol I	RNA Polymerase I
RNA Pol II	RNA Polymerase II
RNAi	RNA interference
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SEs	Super-Enhancers
shRNA	Short Hairpin RNA
shSCR	Scramble shRNA
siRNA	Small Interfering RNA
SMC1	Structural Maintenance of Chromosomes 1
snRNAs	Small Nuclear RNAs
STORM	Stochastic Optical Reconstruction Microscopy
TADs	Topologically Associated Domains
TE	Trophoectoderm
TEs	Typical Enhancers
THF	Tetrahydrofolate
TFIIB	Transcription Factor IIB
TFs	Transcription Factors
TSS	Transcription Start Site
Xa	Active X Chromosome
Xi	Inactive X Chromosome

Introduction

1. CELLULAR IDENTITY AND PLASTICITY

1.1. Cellular identity

Cellular identity is a fundamental characteristic of each cell. It determines cell phenotype, function, and interaction with other cells of the organism, underpinning the division of labour and coordination required in multi-cellular species.

While an adult and developed organism is composed of a rich variety of cells with many different phenotypes, them all have differentiated from successive divisions initiated from a single fertilized cell: the zygote. This means that they all share the same genetic background and, through development and differentiation, cells not only multiply in number, but diversify and specialize into a wide variety of distinct identities.

This process integrates complex signaling decisions to selectively activate the cell-type-specific transcriptional program, involving multiple regulatory layers that we are in the process of understanding.

It was in the late 50s, when Conrad Waddington proposed the "epigenetic landscape" (Waddington, 1957) illustrating this process of acquisition of a terminal cellular identity (differentiation) with a hill, where a totipotent (zygote and blastomers) cell starts rolling down while progressively acquiring a more and more differentiated identity (Figure 1).

It is important to emphasize that every cell has an identity, and toti- or pluripotent cells are not cells that "lack" an identity they need to gain, but in contrast, they are cells that express specific pluripotency markers and transcription factors. Their identity is to be progenitors for other tissues and cells (De Los Angeles et al., 2015).

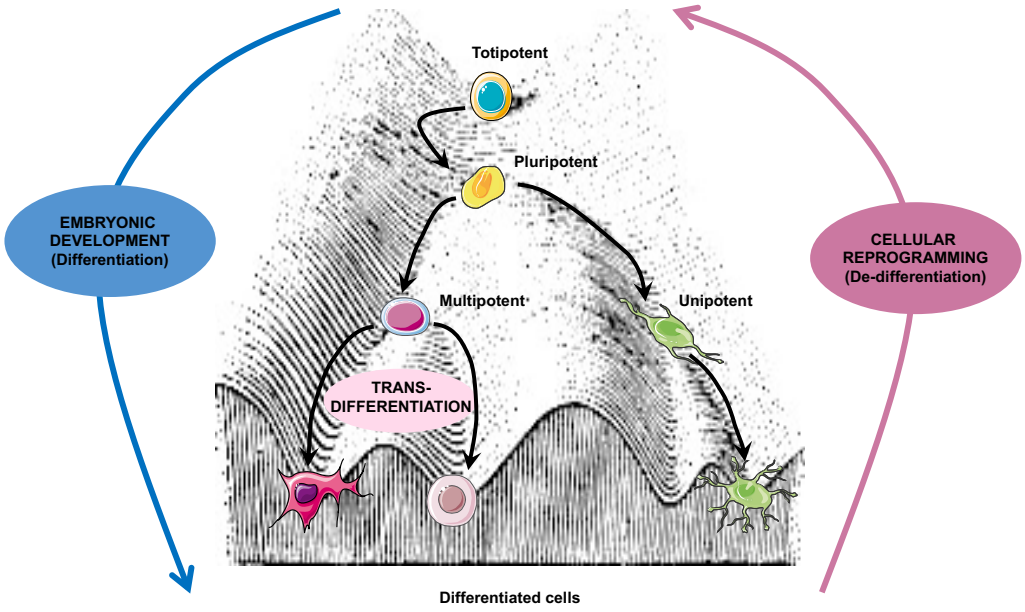


Figure 1 Schematic representation of an updated Waddington's epigenetic landscape model.

In the course of embryonic development, a totipotent cell undergoes a journey "rolling down" a hill with distinct segments on an incline. Depending on the specific segment it lands in, it will adopt distinct cell fates. Some cells have a high plasticity potential and are capable of trans-differentiating. In those cases, a cell specific to a particular tissue transforms directly into a related tissue-specific cell without passing through an enduring intermediate pluripotent state. In the context of cellular reprogramming, a specialized cell ascends from the base to the peak, thus regaining its pluripotent potential.

1.2. Embryonic Development

This process of differentiation operates embryonic development (Figure 2) initiated from a totipotent cell: the zygote, that starts its journey to develop a full organism through sequential divisions and cell fate decisions.

In mice, the first cell fate decision occurs at the eight-cell stage E2.5 (Embryonic day 2.5), also known as "Morula", at this stage, cells become pluripotent, they start gaining differentiation marks and loosing potency; two different identities emerge: the Inner Cell Mass (ICM) and the Trophoectoderm (TE).

One day after, at E3.5, the TE cells will again differentiate, giving raise to the extraembryonic placenta (Tanaka et al., 1998), and a second cell fate decision is made, and the ICM cells differentiate into two identities: the Primitive Endoderm (PE), which will give raise to the extraembryonic yolk sac (Kunath et al., 2005) and the Epiblast (EPI),

that will form the three germinal layers (endoderm, mesoderm and ectoderm) (G. R. Martin, 1981).

The stage E4.5 is determined by the movement of the PE cells to the cavity side before the implantation. When the implantation occurs, the anterior and posterior domains get stablished and the EPI structures in a cup-shaped epithelial that surrounds the luminal space and is covered by the PE (Shahbazi & Zernicka-Goetz, 2018). At this point, gastrulation starts, and EPI cells undergo Epithelial-to-Mesenchymal Transition (EMT) to develop the three germ layers.

Cells continue dividing and differentiating, giving raise to the adult organism, where cells are either multipotent Adult Stem Cells (ASCs) or (and mostly) unipotent and terminally differentiated.

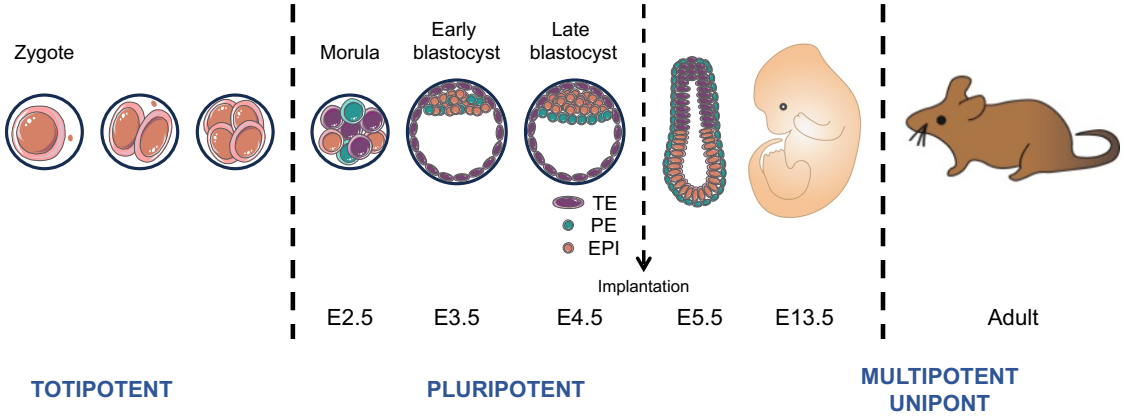


Figure 2. Overview of mouse embryonic development. *Schematic representation of the main differentiation events that occur during embryonic development in mice, from the zygote, until the adult organism.*

1.3. In vitro Pluripotent Stem Cells

The last years of research we have progressively provided us with powerful tools to study this plasticity of pluripotent cells, Embryonic Stem Cells (ESCs) represent a prototypical model of how cells change their identity by re-directing their transcriptional program, faithfully reflecting the changes that occur naturally in the mouse early embryo.

ES cells are Pluripotent Stem Cells (PSCs), this means that they have the ability to differentiate into any of the cell types that make up an adult organism (Hackett & Azim Surani, 2014). ES cells are obtained from the inner cell mass of developing blastocysts, whereupon they can divide indefinitely *in vitro*, while retaining their physiological cell identity. Interestingly, ES cells *in vitro* can exist in either of two different states that reflect two successive developmental stages in vivo and which are driven by discrete transcriptional activity: (i) naïve state and (ii) primed state (Figure 3).

1.3.1. Pluripotent "naïve" stem cells

Naïve stem cells recapitulate *in vitro* the state of undifferentiated pluripotency that exists transiently *in vivo* in the E4.5 pre-implantation epiblast of the early embryo (Boroviak et al., 2014; Marks et al., 2012).

Naïve ES cell cultures are homogenous in cellular morphology, forming colonies that aggregate in a characteristic “dome” shape. Naïve cells uniformly express a set of markers that configurate the pluripotency transcriptional program (Ying et al., 2008) and they lack expression of differentiation genes. Their chromatin is in an open and highly accessible configuration (Leitch et al., 2013) and their developmental potential is unbiased.
From a metabolic point of view, Naïve embryonic stem cells are able to perform oxidative phosphorylation, glycolysis and fatty acid oxidation (Mathieu & Ruohola-Baker, 2017).

1.3.2. Pluripotent “primed” stem cells

Primed pluripotent stem cells *in vitro* represent a subsequent developmental stage corresponding to the post-implantation E6.5 epiblast (Nichols & Smith, 2009; Shahbazi & Zernicka-Goetz, 2018) They display heterogeneous cell morphologies, and start expressing early differentiation markers, and global epigenetic remodeling with higher levels of DNA methylation(Manor et al., 2015; Torres-Padilla & Chambers, 2014).
Primed ESCs rely almost exclusively on glycolysis to meet their bioenergetic demands (Martinez-Val et al., 2021; Mathieu & Ruohola-Baker, 2017).

While both identities can differentiate into the three germ layers and form teratomas, only naïve cells can contribute to the formation of chimeras (Nichols & Smith, 2009; Ying et al., 2008).

Supporting these differences in transcriptional program of expressed genes and developmental potential, several studies suggest specialized metabolic adaptions. From a metabolic point of view, Naïve embryonic stem cells are able to perform oxidative phosphorylation, glycolysis and fatty acid oxidation (Mathieu & Ruohola-Baker, 2017). (It would be good to expand this point, either here or further below in the Metabolism Section)

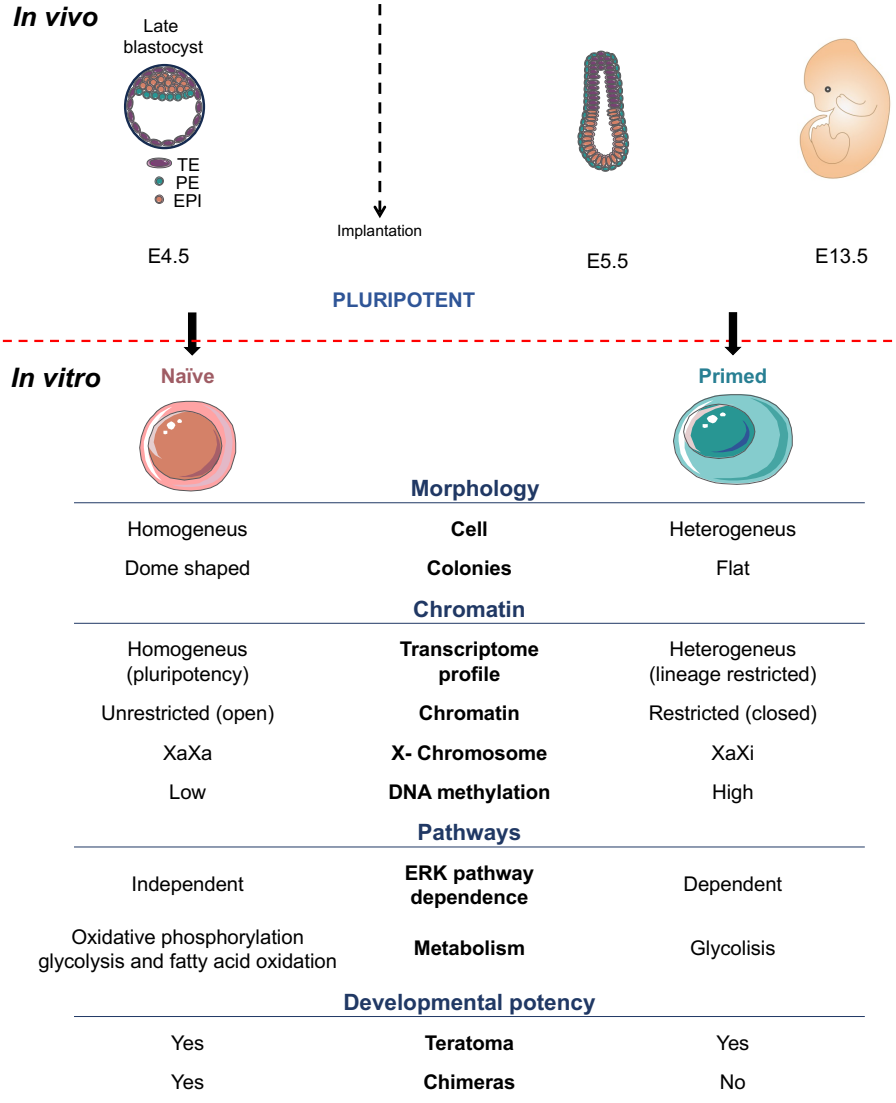


Figure 3. Naive and primed pluripotent ESCs states comparison. *A comparative of some of the main differential characteristics that can be found between pluripotent naïve and pimed cells.*

2. TRANSCRIPTION FOR CELL IDENTITY

2.1. General overview

The central dogma of molecular biology, proposed by Francis Crick in 1958, and revisited afterwards (Crick, 1970) serves as a foundational framework for understanding the flow of genetic information within living organisms.
Initially it described the flow of genetic information as a unidirectional process: DNA is transcribed into RNA, and RNA is translated into proteins. This concept emphasized the unidirectional transfer of genetic information.

However, our understanding of molecular biology continues to evolve, and the central dogma has been refined to accommodate new findings. This paradigmatic concept outlines the sequential processes of replication, transcription, and translation that govern the transmission of genetic instructions from DNA to RNA to proteins. (Figure 4.).

Replication, the initial phase of the central dogma, is a meticulously orchestrated process during cell division. Its primary objective is to ensure the faithful duplication of the genetic material. DNA, the double-stranded helical molecule housing the genetic code, undergoes replication, giving rise to identical copies. This meticulous duplication process safeguards the integrity and consistency of genetic information across generations of cells.

Transcription, the second component of the central dogma, is a crucial step in the conversion of genetic information from DNA to RNA. Transcription occurs in the cell nucleus, where the enzyme RNA polymerase reads the DNA template and synthesizes a complementary RNA strand. This newly formed RNA, known as messenger RNA (mRNA), carries the genetic instructions from the nucleus to the cytoplasm, serving as an intermediary messenger for subsequent protein synthesis. While the central dogma emphasized unidirectional flow, we now know that some RNA molecules can be reverse transcribed into DNA (Baltimore, 1970; Mizutani & Temin, 1970). This process, called

reverse transcription, is catalyzed by enzymes like reverse transcriptase, allowing the synthesis of complementary DNA (cDNA) from an RNA template.

Translation: As genetic information reaches the cytoplasm, the final stage of the central dogma, translation, takes place. Ribosomes, the cellular machinery responsible for protein synthesis, decode the mRNA sequence. Transfer RNA (tRNA) molecules ferry amino acids to the ribosome, where they are assembled into a polypeptide chain. This chain ultimately folds into a functional protein, embodying the genetic information initially encoded in the DNA. Initially, the focus was on protein-coding genes. However, the discovery of non-coding RNAs (ncRNAs), such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), highlighted the importance of RNA molecules beyond their role as templates for protein synthesis(Inouye & Delihast, 1988; Mizuno et al., 1984). These non-coding RNAs participate in various regulatory processes, influencing gene expression.

The original central dogma did not account for the impact of epigenetic modifications on gene regulation, unknown at the time. Epigenetic changes, such as DNA methylation and histone modifications, can alter chromatin structure and affect gene expression without changing the underlying DNA sequence (Mathieu & Ruohola-Baker, 2017). Also, RNA molecules were initially seen as intermediaries in the transfer of genetic information. However, discoveries like RNA interference (RNAi) revealed that certain RNAs play active roles in regulating gene expression by targeting and silencing specific mRNAs.

The coordination of replication, transcription, and translation is fundamental to cellular function. Known regulatory relationships and feedback loops are shown in Figure 4.

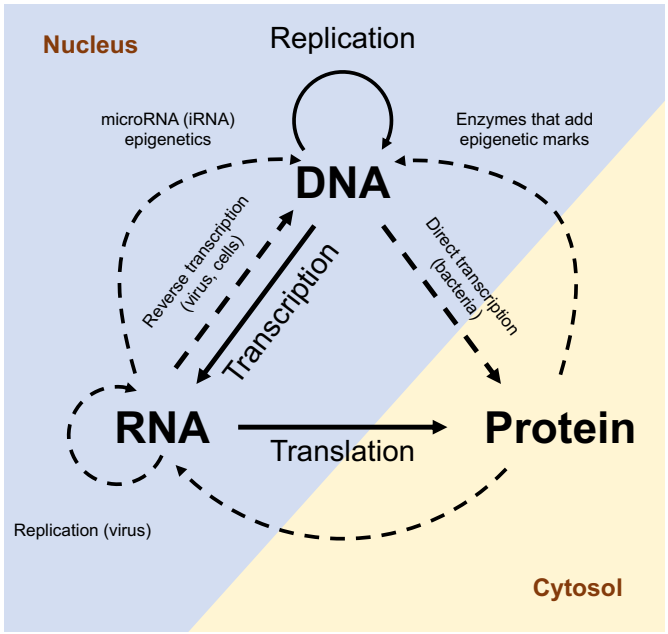


Figure 4. Schematic of molecular biology's central dogma. *Illustration of the central dogma with nucleus (blue) and cytoplasm (yellow) distinctions. Continuous arrows represent original processes, while non-continuous arrows denote new modifications. Key events: DNA replication, transcription, RNA processing (blue), translation, post-translational modifications (yellow), and emerging RNA-mediated processes.*

Eukaryotic cells employ three distinct RNA polymerases—RNA Polymerase I, II, and III—to transcribe different classes of RNA. RNA Polymerase I (RNA Pol I) is responsible for rRNA synthesis, RNA Polymerase II (RNA Pol II) transcribes mRNA and some small nuclear RNAs (snRNAs), while RNA Polymerase III (RNA Pol III) synthesizes tRNA and other small RNAs (Roeder & Rutter, 1969).

Transcription driven by RNA Pol II stands as the primary decission point for gene expression, wielding a pivotal role for efficient regulation of cellular processes. This critical step in the central dogma is the moment when genetic information encoded in DNA is transcribed into mRNA, effectively determining which genes are activated or repressed. The regulation of transcription driven by RNA Pol II allows cells to finely tune gene expression in response to internal and external cues.

In essence, transcription serves as the command center, dictating the blueprint for protein synthesis and, consequently, determining the functional characteristics of a cell. The regulation of gene expression through transcription is a finely tuned mechanism that

enables cells to adapt and respond to their environment, contributing to the diversity and complexity of biological systems. Understanding this central role of transcription provides valuable insights into the molecular orchestration of cellular activities and the maintenance of cellular homeostasis.

The concept of transcriptional control extends beyond the initiation of transcription. Post-transcriptional modifications and RNA processing further influence the fate of the mRNA transcript. These modifications can dictate the stability, localization, and translatability of the mRNA, ultimately impacting the abundance and functionality of the corresponding protein (Filtz et al., 2014).

2.2. Enhancer-driven transcription

As described above, RNA Pol II governs gene transcription. To do so, it relies on two primary classes of regulatory elements: promoters and enhancers.

Promoters reside proximal to the transcription start site (TSS) and play a central role in initiating transcription. They comprise a core promoter with essential motifs (TATA box, TFIIB recognition element (BRE), initiator (Inr) (Butler & Kadonaga., 2002) and the downstream promoter element (DPE)(Kutach & Kadonaga., 2000). Promoters facilitate the binding of RNA Pol II and general transcription factors. Adjacent cis-regulatory modules (CRMs) further contribute to their regulatory complexity, binding sequence-specific transcription factors to determine the specificity of transcriptional initiation and ensuring basal transcription (Berman et al., 2002).

Transcription factor proteins that bind to specific DNA sequences, either promoting or inhibiting the recruitment of RNA Pol II (Orphanides et al., 1996). This dynamic interplay ensures that only the necessary genes are transcribed under specific conditions (Becker et al., 2002).

Enhancers are DNA regions found distally located (Serfling et al., 1985). Their remarkable feature lies in their ability to activate transcription over long distances. Independent of orientation and distance, enhancers interact with promoters, influencing

the initiation of transcription. They are composed of clusters of recognition sites for transcription factors, enhancers respond to cellular cues and environmental signals, providing specificity to gene expression patterns.
Landmark studies have revealed the existence of two general classes of enhancer, divided by genomic size and strength of influence over cell identity (Hnisz et al., 2013). Typical enhancers (TEs) are small. In contrast, Super-enhancers are bigger. 12-25 Kb or larger. Consists of the amalgamation of multiple typical enhancers, which form constituents that coverge to more than the sum of their parts. They are defined by being exponentially larger than typical enhancers. Defined by the amount of enhancer and transcription proteins contained, in a system known as R.O.S.E Rank Order of Super-enhancers. Super-enhancers are of particular note as crucially, super-enhancers are the key drivers of cell-type-specific gene expression (Hnisz et al., 2013; Whyte et al., 2013).

The intricate interplay between promoters and enhancers defines enhancer-driven transcription, and it is crucial for the nuanced and context-dependent regulation of gene expression and cellular identity, contributing to the diversity and adaptability of cellular functions.

The concept of transcriptional control extends beyond the initiation of transcription. Post-transcriptional modifications and RNA processing further influence the fate of the mRNA transcript. These modifications can dictate the stability, localization, and translatability of the mRNA, ultimately impacting the abundance and functionality of the corresponding protein (Filtz et al., 2014).

2.3. Mediator Complex and Cellular Identity

Our understanding of the transcriptional control of cellular identity is rapidly emerging. Each cell type has a specific tightly regulated chromatin configuration, determined by the patterns of DNA packing, and their unique epigenetic and histones modifications.

Each cell type possesses a repertoire of active DNA enhancer regions, characterised by the presence of high concentrations of particular histone marks and transcription factors(Heinz et al., 2015; Hnisz et al., 2013; Whyte et al., 2013).

Active enhancer regions are recognized by the Mediator complex, which recruits RNA Pol II to the pre-initiation complex (PIC) of target genes. Thus, Mediator forms a protein bridge between active enhancers and the promoters of the genes that have to be transcribed (Malik & Roeder, 2016).

Moreover, the largest concentrations of Mediator and RNA Pol II co-exist s super-enhancers. These extensive enhancer networks act as hubs for the recruitment of various proteins involved in transcriptional regulation, making them particularly potent regulators of gene expression. Super enhancers have been implicated in driving the expression of genes critical for cell fate determination and are associated with the maintenance of cell identity in both normal development and disease states (Hnisz et al., 2013; Whyte et al., 2013).

These regions are also characterized by enrichment of bromodomain containing 4 (BRD4), chromatin modifiers such (Ep300 and Kdm1a: LSD1 complex), and chromatin remodelers (Chd7, Brg1: SWI-SNF complex; Chd4 NuRD: complex; Smc1a: Cohesin complex). Together these regulate the transcription of key cell identity genes and master transcription factors (Hnisz et al., 2013; Whyte et al., 2013). Thus, super-enhancers are central to the control of cellular identity.

The flexibility and specificity provided by enhancer- and super-enhancer- driven transcription plays a crucial role in the complexity and adaptability of eukaryotic gene regulation. This mechanism allows for the fine-tuning of gene expression patterns, ensuring that genes are activated or repressed in a context-dependent manner during development, differentiation, and in response to various stimuli.

Mediator complex consists of 30 subunits, contained in four main domains. The head, central and tail domains constitute the core of Mediator and their roles are mainly structural, participating in the binding to enhancers and in the recruitment of RNA Pol II to TSSs.

The fourth domain is the kinase module, containing Cyclin dependent Kinase 8 (CDK8) or its paralogue Cyclin dependent Kinase 19 (CDK19) and their catalytic binding partner

Cyclin C (Verger et al., 2019). CDK8 and CDK19 possess an identical kinase domain, and in each Mediator complex, the presence of CDK8 or CDK19 is mutually exclusive, such that only one may occupy Mediator at any given time (Fant & Taatjes, 2019). It acts as a negative regulator of RNA Pol II recruitment by Mediator (Galbraith et al., 2010; Hengartner et al., 1998).

In accordance with this, it has been observed that CDK8/19 phosphorylates multiple subunits of Mediator (Poss et al., 2016) and chemical inhibition of its kinase activity results in hyperactivation of enhancer function (Lynch et al., 2020; Pelish et al., 2015). In addition, CDK8/19 binds and dissociates from Mediator dynamically and, interestingly, the association of CDK8/19 to Mediator prevents the recruitment of RNA Pol II by steric hindrance (Clark et al., 2015; Jeronimo & Robert, 2017; Poss et al., 2016) (Figure 5). CDK8/19 may also regulate transcription through phosphorylation of the C-terminal regulatory domain of RNA Pol II (Hsin & Manley, 2012), regulators of chromatin, and transcription factors (Galbraith et al., 2010; Soutourina, 2018).

Lastly, in addition to phospho-regulation of Mediator, it was recently reported that Mediator and the proximal transcriptional machinery are heavily acetylated (Weinert et al., 2018). Many questions remain regarding how the 30 subunits of Mediator work together, receive post-translational inputs, and achieve gene-specific control of transcription.

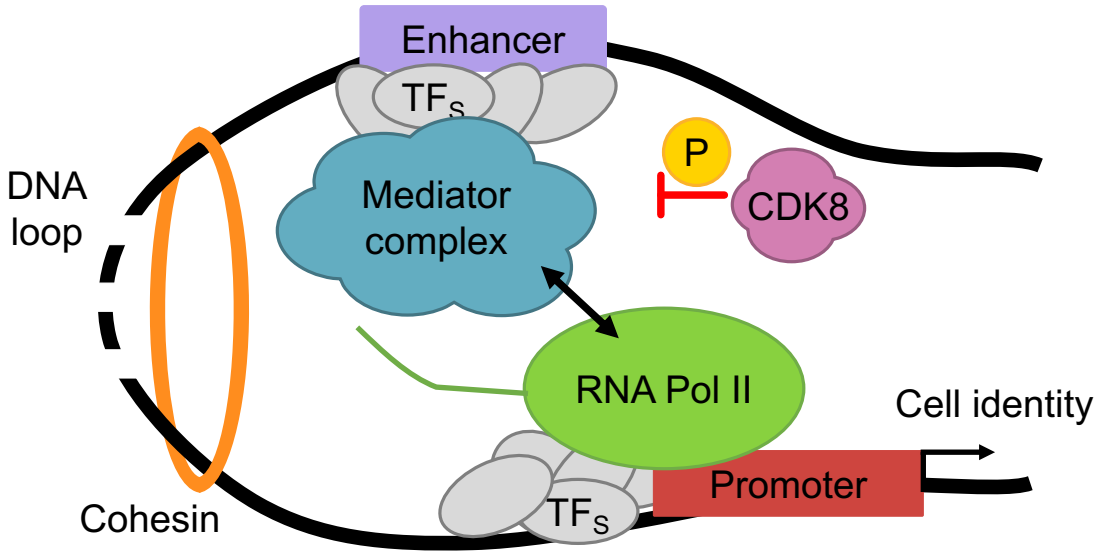


***Figure 5. Transcription and Mediator complex**. The role of Mediator complex and its CDK8/19 kinase module in the regulation of Pol II recruitment and transcription mediated by enhancers.*

2.4. Manipulating the transcriptional machinery to toggle ES cell identity

When mouse ESCs are cultured in standard conditions, we obtain an heterogenous culture where both naïve and primed identities are present (Canham et al., 2010), this two identities are interchangeable, in a constant fluctuation driven by Leukemia Inhibitory Factor (LIF)/STAT3 and FGF/ERK pathways (Hackett & Azim Surani, 2014; Kunath et al., 2007; Yeo et al., 2014). This phenomenon is driven by the relative strength of LIF or MAPK signaling (Chambers et al., 2007), stablishing a transcriptional antagonism between the pluripotency program and the differentiation programs that will give raise to the three germ layers. (Figure 6)

Remarkably, ES cells can be reprogrammed quickly, efficiently, and reversibly, between these two states of pluripotency (Hackett & Azim Surani, 2014). Specifically, the heterogeneity of the primed state is rapidly eliminated, and its transcriptional program is rapidly re-configured, by blocking the MEK/ERK and GSK3ß signaling pathways (Hackett & Azim Surani, 2014; Ying et al., 2008) (Ying et al., 2008; Hackett and Surani., 2014). This is achieved using two small chemical inhibitors, commonly called "2i", which

together reset primed ES cells back into the naïve state in a matter of days (Figure 6). Importantly, the inhibition of the MEK pathway has been identified as the key element in triggering these effects on PSCs (Ficz et al., 2013; Marks et al., 2012; Nichols & Smith, 2009; Ying et al., 2008). However, it remains unclear how mechanistically 2i can rapidly re-direct the transcriptional program.

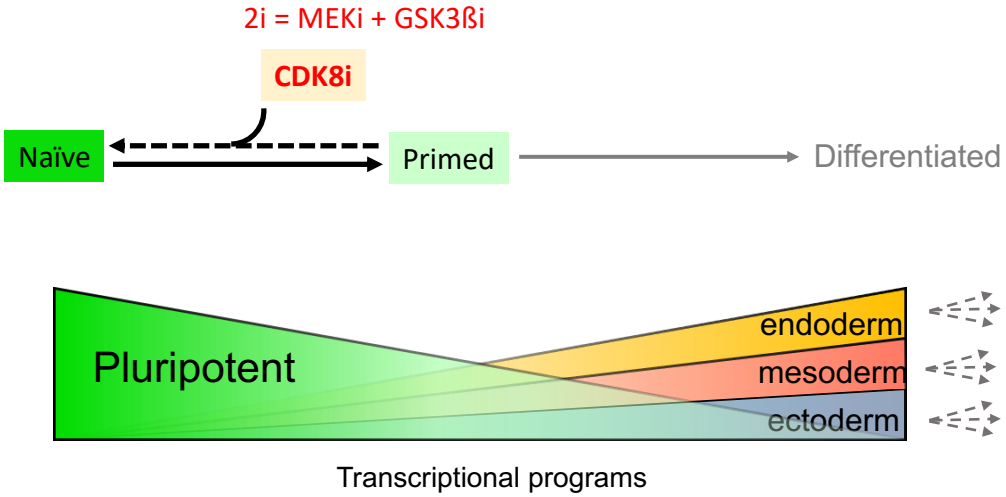


Figure 6. In vitro naïve pluripotent mESCs stabilization and transcriptional reinforcement by Mediator hyperactivation. *Stabilization and transcriptional reinforcement of in vitro naïve pluripotent mouse embryonic stem cells (mESCs) facilitated by Mediator hyperactivation. Strong pluripotent transcriptional activity is observed in naïve cells, contrasting with weakened pluripotent transcriptional programs in primed cells. The latter marks the initiation of new transcriptional programs associated with the differentiation process into the three distinct germ layers.*

Previously, we developed a method to inhibit the CDK8/19 subunit of Mediator based on a small chemical compound abbreviated as CDK8i (Lynch et al., 2020).

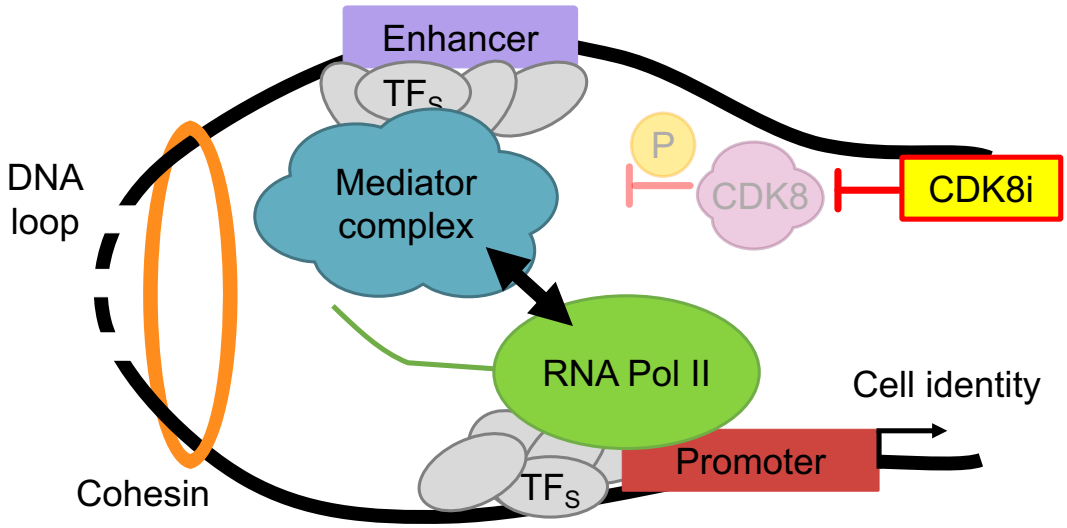


Figure 7. Mediator complex hyperactivation through CDK8 inhibition. *lustration depicting the enhanced activation of the Mediator complex through CDK8 inhibition. The removal of CDK8 negative modulation results in a strengthened and more potent interaction between RNA Polymerase II (Pol II) and the Mediator complex.*

Surprisingly, we found that CDK8i mimics the effect of 2i on ES cell identity and restores the naïve transcriptional program (Figure 7).

Given the similarity in the effects of 2i and CDK8i on ES cell identity, the known role of CDK8/19 module within the Mediator complex, and the ability of Mediator to control RNA Pol II, we have hypothesized that 2i and CDK8i may affect the Mediator complex in a similar way, resulting in a recruitment and control of RNA Pol II, in a manner which reinforces the naïve PSC program.

Indeed, our previous data found that 2i and CDK8i appear to boost the activity of Mediator, such that more RNA Pol II is recruited to the promoters of target genes. Therefore, this enhanced Mediator-RNA Pol II interaction may explain the stabilization of the naïve state transcriptional program.

2.5. Transcriptional condensates. Mediator Complex-RNA Polymerase II interaction.

The Mediator-RNA Pol II interaction has been directly observed in living cells. Live-cell super-resolution microscopy has revealed that RNA Pol-II-mediated transcription takes place at nuclear condensates (Cho et al., 2018; Y. E. Guo et al., 2019a).

Core subunits of Mediator: Mediator of RNA polymerase II transcription subunit 1 (MED1) and RNA Pol II: RNA polymerase II subunit A (POLR2A) were fused to fluorescent-protein labels by genetic modification, and this led to the conclusion that multiple complexes of Mediator and RNA Pol II concentrate in close proximity as phase-separated transcriptional condensates, exhibiting properties of liquid, as a fast recovery of fluorescence after photobleaching and sensitivity to 1,6-hexanediol. Phase separated liquid transcriptional condensates are formed by the LLPS of RNA Pol II and various factors harboring intrinsically disordered regions (IDRs), as Mediator complex and transcription factors (TFs).

Moreover, the largest condensates of Mediator and RNA Pol II with phase-separation properties appear to co-exist at the largest and most powerful enhancers, known as super-enhancers (see above).

However, fusion to a fluorescent protein label may affect Mediator-RNA Pol II function. Furthermore, to observe the co-localization of major complexes in the transcriptional machinery, it is useful to consider their size, and the imaging resolution limits of microscopy. The Mediator complex is approximately 15-20 nm in diameter, while the RNA Pol II complex is ~10 nm across (Hahn, 2004; Tsai et al., 2017) (Figure 8).

By comparison, using standard confocal microscopy, the resolution limit is 200 nm (Kaufmann et al., 2014). More recently, sub-diffraction limit super-resolution approaches as stochastic optical reconstruction microscopy (STORM) can approach a resolution limit of ~10-30 nm (Rust et al., 2006), while fluorescence resonance energy transfer (FRET) also operates at a distance of <10 nm (Sekar & Periasamy, 2003). However, quantification of fluorescent protein co-localization using super- resolution microscopy or FRET requires significant expertise and optimization.

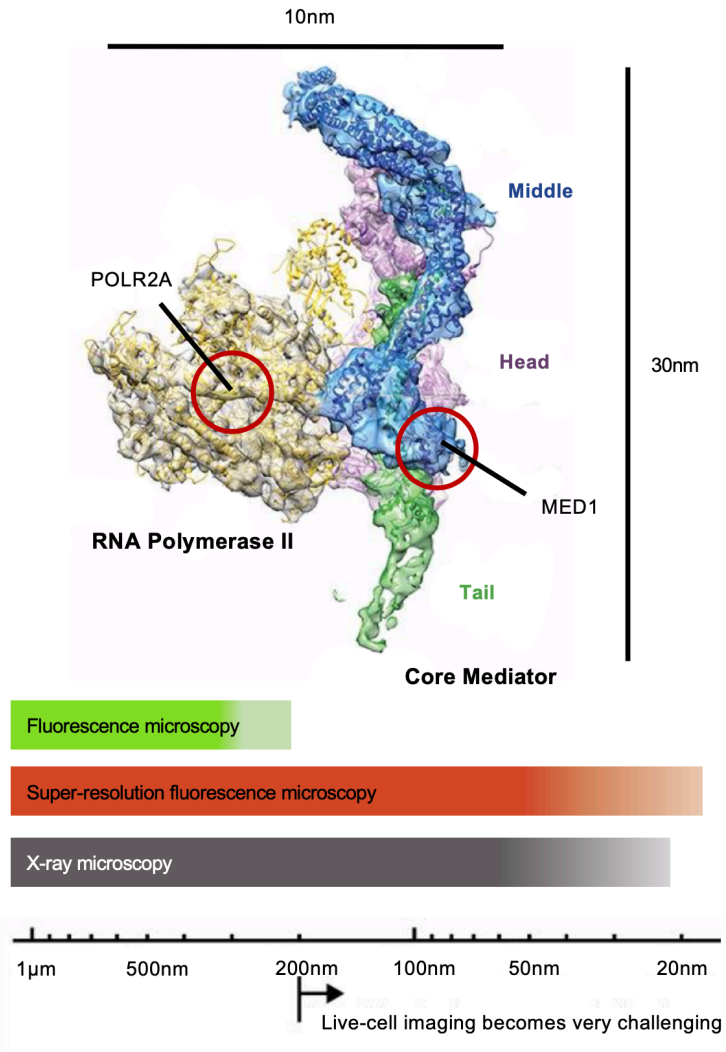


Figure 8. RNA Pol II and Mediator complex interacting. *Mediator head, tail and middle modules are shown. Also, subunits referred to in this study (MED1 and POLR2A are indicated). Below, resolution limits of microscopy. Figures adapted from Kaufmann et al., 2014 and Tsai et al., 2017*

2.6. Proximity Ligation Assay

Proximity Ligation Assay (PLA) (Jarvius et al., 2007) is a method that allows the detection of protein-protein interactions in situ using standard confocal fluorescence microscopy, yet it can demonstrate an interaction distance within less than 40 nm (Bagchi et al., 2015). Thus, the resolution conferred by PLA compares well with advanced imaging techniques like STORM (Rust et al., 2006). PLA relies on two antibody-conjugated oligonucleotides that can hybridize to two other "connector" oligonucleotides only if sufficiently proximal, that is, specifically within 40 nm due to the restriction of the oligo length (Figure 9). After hybridization, the connector oligonucleotides are ligated, forming a circular DNA molecule which can be amplified by the phi29 DNA

polymerase in a rolling circle amplification (RCA) mechanism primed by oligonucleotides on one of the antibodies. This results in a localized single-stranded DNA molecule that can be detected through fluorescently labeled complementary oligonucleotides.

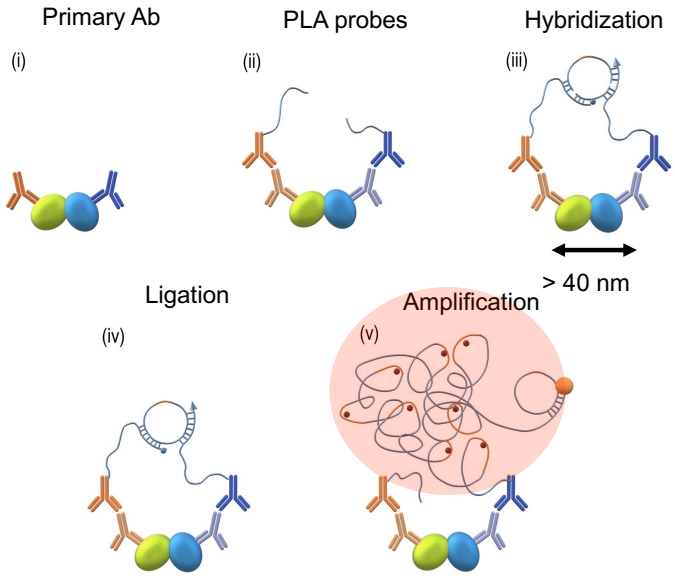


Figure 9. Schematics of a PLA reaction. *(i) primary antibody incubation (ii) PLA probe incubation, (iii) hybridization of connector nucleotides, (iv) ligation and (v) amplification*

2.7. Post-translational and epigenetic modifications

Finally, to have a complete overall picture of how transcription is regulated, we need to understand the role of post-translational modifications (PTMs) of histones and DNA. As they dictate chromatin architecture and therefore Mediator accessibility to chromatin and RNA Pol II. PTMs have been extensively studied since the early 1960s, and their understanding and identification is in constant growth.

Distinct histone PTMs are locally enriched in specific genomic locations and genes determining transcriptional activity (both positively and negatively) by relaxation or condensation of chromatin.

PTMs also include complex and non-direct mechanisms of regulation as recruitment of different binding proteins implicated in the regulation of gene expression. The most

extensively studied histone and DNA PTMs are acetylation and methylation; but many others are being characterized. They are classified in four main groups that include acylations, ubiquitin-like, and others (Methylation and Non-lysine PTMs) (Millán-Zambrano et al., 2022).

Depending on the modification and the target histone residue, the outcome of the modification in cell activity and transcription will be different, with local genomic activation or repression.

A correct pattern and distribution of the complex histone and DNA PTMs is carefully coordinated in the nucleus by histone modifiers, a process which is essential to carry out fundamental cellular processes, such as chromatin phase transitions, development, recombination, DNA repair, replication, and gene topology. Despite efforts to understand the involved mechanisms, many unknowns remain.

We will further elaborate on methylation and acetylation during this introduction, after having explained some key metabolic concepts that are key to understand this PTMs.

3. METABOLISM

3.1. General overview

Metabolism is the set of chemical reactions that occur within cells to fuel organisms with energy to sustain life. It is a complex process that encompasses many interconnected and tightly regulated cellular pathways, which together are responsible for energy production, the synthesis of biomolecules, and the signaling and regulation of cellular processes.

We can classify metabolic processes in two main categories: anabolic and catabolic processes.

(i). **Anabolic** processes include all the metabolic reactions that build complex molecules. Anabolism forms bonds between small molecules consuming energy.

(ii). On the other hand, **catabolic** reactions, are the ones that produce energy through the breakdown of bonds in large and complex molecules. The products of catabolic reactions are simpler, smaller molecules and energy. This energy will be used by anabolic processes in a continuous cycle.

The balance of intermediate metabolites and molecules are sensed by the cell and become signaling factors that regulate cellular activity (Schuster et al., 2000).

Within cells, a multitude of meticulously coordinated metabolic pathways operate in concert, akin to the intricate components of a clock, to orchestrate cellular function, maintain energy equilibrium, and facilitate the biosynthesis of molecules.
The exploration of each of these pathways and their interrelation is extensive.

3.2. Mitochondria and energy production

Mitochondria are the cellular organelles responsible for energy production (Nunnari & Suomalainen, 2012). Through cellular respiration, mitochondria generate adenosine triphosphate (ATP), the primary currency of cellular energy (Mitchell, 1961).

Among the myriad of metabolic routes, cellular respiration involves three main stages: glycolysis in the cytoplasm, the tricarboxylic acid (TCA) cycle in the mitochondrial matrix, and oxidative phosphorylation on the inner mitochondrial membrane (Figure 10).

Glycolysis: Glycolysis is the central pathway for the breakdown of glucose to produce energy in the form of ATP. It involves a series of enzymatic reactions that convert one molecule of glucose into two molecules of pyruvate. Along the way, ATP and NADH are generated. Glycolysis is a common pathway in both aerobic and anaerobic conditions.

Tricarboxylic Acid (TCA) Cycle: Also known as the Krebs cycle or citric acid cycle, the TCA cycle involves a series of biochemical reactions that complete the oxidation of acetyl-CoA derived from carbohydrates, fats, and proteins. It generates NADH and

$FADH_2$, which carry high-energy electrons to the electron transport chain for ATP production. The TCA cycle is a key hub connecting various metabolic pathways.

Oxidative Phosphorylation: Oxidative phosphorylation is the final stage of cellular respiration, where electrons from NADH and $FADH_2$, produced in glycolysis and the TCA cycle, move through a series of protein complexes in the electron transport chain. This movement creates a proton gradient across the inner mitochondrial membrane. ATP synthase utilizes the energy from this gradient to phosphorylate ADP to ATP, a process known as chemiosmosis.

These three pathways are interconnected and crucial for energy production in eukaryotic cells. Glycolysis breaks down glucose to pyruvate, the TCA cycle oxidizes the products of glycolysis to generate electron carriers, and oxidative phosphorylation uses these carriers to produce ATP, completing the process of cellular respiration (Spriet et al., 2020).

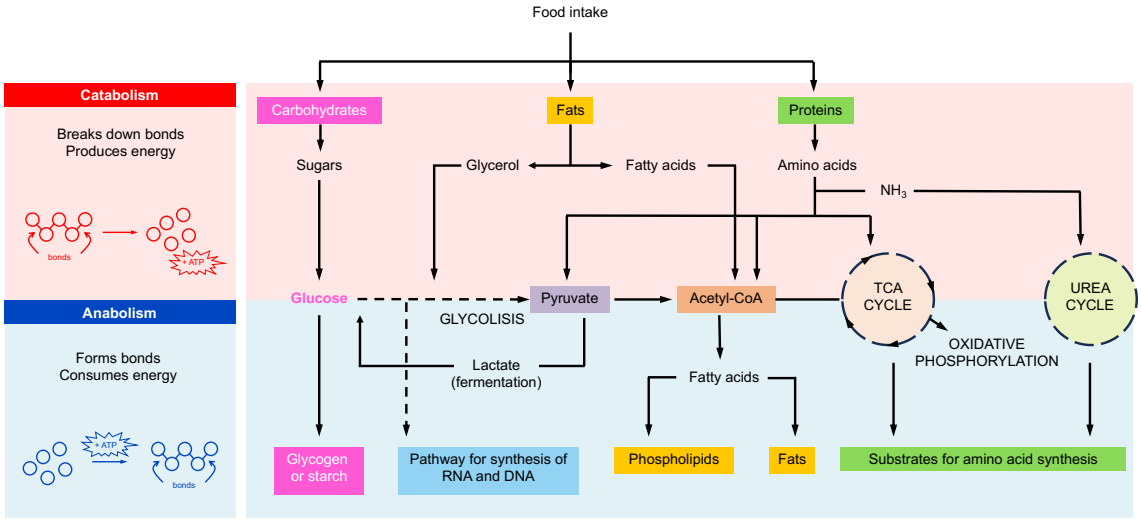


Figure 10. Overview of Interconnected Metabolic Pathways. *Schematic illustration of the interconnections of metabolic pathways with biomolecules serving as sources of energy. Catabolic processes are highlighted in red, representing energy-releasing pathways, while anabolic processes are depicted in blue, symbolizing energy-consuming pathways. This visual representation provides a clear overview of the dynamic balance between catabolism and anabolism in cellular metabolism.*

3.3. Metabolism: A source of signaling molecules for the control of transcription

There is a rapidly increasing amount of evidence pointing out the signaling role of metabolic intermediates, traditionally deemed important for biosynthetic purposes. These intermediates are now recognized as signaling molecules with pivotal functions in controlling chromatin modifications, DNA methylation, the hypoxic response, and immunity (Chandel, 2015).

For the interest of this thesis, we will focus into the metabolic processes that generate methyl and acetyl groups.

3.3.1. Production of methyl groups: One-carbon metabolism

Methyl groups, consist of a carbon atom bonded to three hydrogen atoms (CH_3), and are involved in numerous biological processes, including DNA methylation, nucleotide metabolism, maintenance of redox status and the synthesis of key molecules like neurotransmitters and phospholipids.

One-carbon (1C) metabolism is comprised by a set of interlinked metabolic pathways, that utilize a variety of nutrients including glucose, vitamins, and amino acids, encompassing the methionine and folate cycles, and the pathway of homocysteine remethylation and transsulfuration. The methionine cycle is directly linked to histone methylation through the generation of S-adenosylmethionine (SAM) (Ducker & Rabinowitz, 2017) (Figure 11).

Folate cycle: Derived from dietary sources, folate (B9) is a B-vitamin crucial for one-carbon metabolism. Folate is firstly converted to dihydrofolate (DHF) and then into its active form, tetrahydrofolate (THF). THF serves as a carrier of one-carbon units, which can be added to various molecules in subsequent reactions. These reactions include the conversion of serine to glycine, threonine also contributes as a precursor for the synthesis of glycine (Bailey & Gregory, 2018).

Methyonine cycle: Vitamin B12 serves as a cofactor for methionine synthase (MS), the enzyme responsible for transferring a methyl group from 5-methyl-tetrahydrofolate (5-methyl-THF) to homocysteine (hCys), generating methionine (MET). Methionine, an essential amino acid, is converted to SAM, through the action of methionine adenosyltransferase 2A (MAT2A). SAM donates its methyl group to various substrates, in different cellular reactions that include DNA methylation, histone methylation, and the methylation of lipids and neurotransmitters. During methylation reactions catalyzed by methyltransferases (MTs), SAM donates its methyl group, leading to the formation of S-adenosylhomocysteine (SAH). Adenosylhomocysteinase (AHCY) catalyzes the hydrolysis of SAH, producing homocysteine and adenosine. Homocysteine then can participate in subsequent reactions within the methionine cycle (Sanderson et al., 2019).

Homocysteine remethylation and transsulfuration: Homocysteine, generated from the demethylation of SAM, can follow two paths: it can be remethylated to methionine through betaine-homocysteine S-methyltransferase (BHMT), using methyl groups that come from THF or betaine; or alternatively, it can enter the transsulfuration pathway, forming cystathionine through the cystathionine beta-synthase (CBS) and finally leading to the synthesis of cysteine by the cystathionine gamma-lyase (CSE), cysteine is a precursor for antioxidant glutathione (Selhub, 1999).

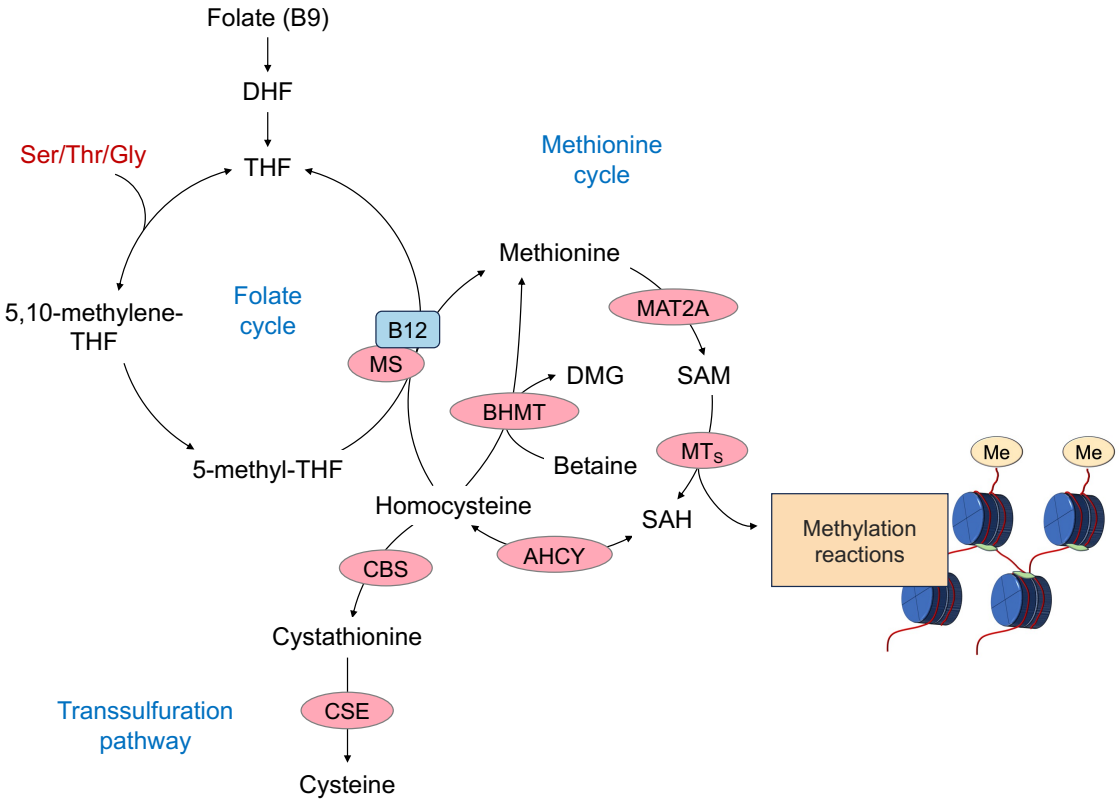


Figure 11. Overview of One-Carbon Metabolism. *concise summary of the mammalian folate and methionine cycles, collectively known as one-carbon (1C) metabolism, alongside the transsulfuration pathway. Crucial enzymes involved in these processes are highlighted in pink.. Additionally, it illustrates the connection to histone methylation, emphasizing the integrated nature of 1C metabolism with epigenetic regulation.*

3.3.2. Production of acetyl groups: Acetyl-CoA metabolism

Acetyl-CoA is a crucial metabolite, it occupies a central position at the crossroads of various metabolic pathways (Figure 11), playing a pivotal role in regulating cellular functions.

It is synthesized through diverse metabolic sources and pathways. The formation of acetyl-CoA is intricately linked to the catabolism of different macronutrients, reflecting the adaptability of cells to varying nutritional conditions. There are distinct metabolic sources and pathways contributing to the synthesis of acetyl-CoA. Here we highlight the main sources of Acetyl-CoA (Figure 12):

Glucose → **Pyruvate**: Glucose undergoes glycolysis to produce pyruvate. Pyruvate is the substrate that gets converted into acetyl-CoA through the pyruvate dehydrogenase

complex (PDH). This process can occur both in the mitochondria and in the nucleus (Sutendra et al., 2014).

Citrate: Citrate is a TCA cycle intermediate. It can be transported to the cytoplasm and to the nucleus, where the ATP citrate synthase (ACLY) enzyme cleaves it into acetyl-CoA (Sivanand et al., 2017).

Acetate: Acetate can be converted to acetyl-CoA in the mitochondria by the acetyl-CoA synthetase 1 (ACSS1);or in the cytosol and nucleus through the action of acetyl-CoA synthetase 2 (ACSS2) (X. Li et al., 2017).

Amino acids: Certain amino acids can be converted into acetyl-CoA through intermediary metabolic pathways (Shi & Tu, 2015).

Fatty acids: Fatty acids, derived from dietary fats or adipose tissue, are broken down through beta-oxidation in the mitochondria. Each round of beta-oxidation generates acetyl-CoA molecules, contributing to the pool of acetyl-CoA available for energy production or other metabolic processes (Shi & Tu, 2015).

Ketone bodies: During periods of fasting or low carbohydrate intake, ketone bodies (acetone, acetoacetate, and beta-hydroxybutyrate) are produced in the liver. These ketone bodies can be converted into acetyl-CoA in extrahepatic tissues, providing an alternative energy source, especially for the brain, heart, and skeletal muscles (Puchalska & Crawford, 2017).

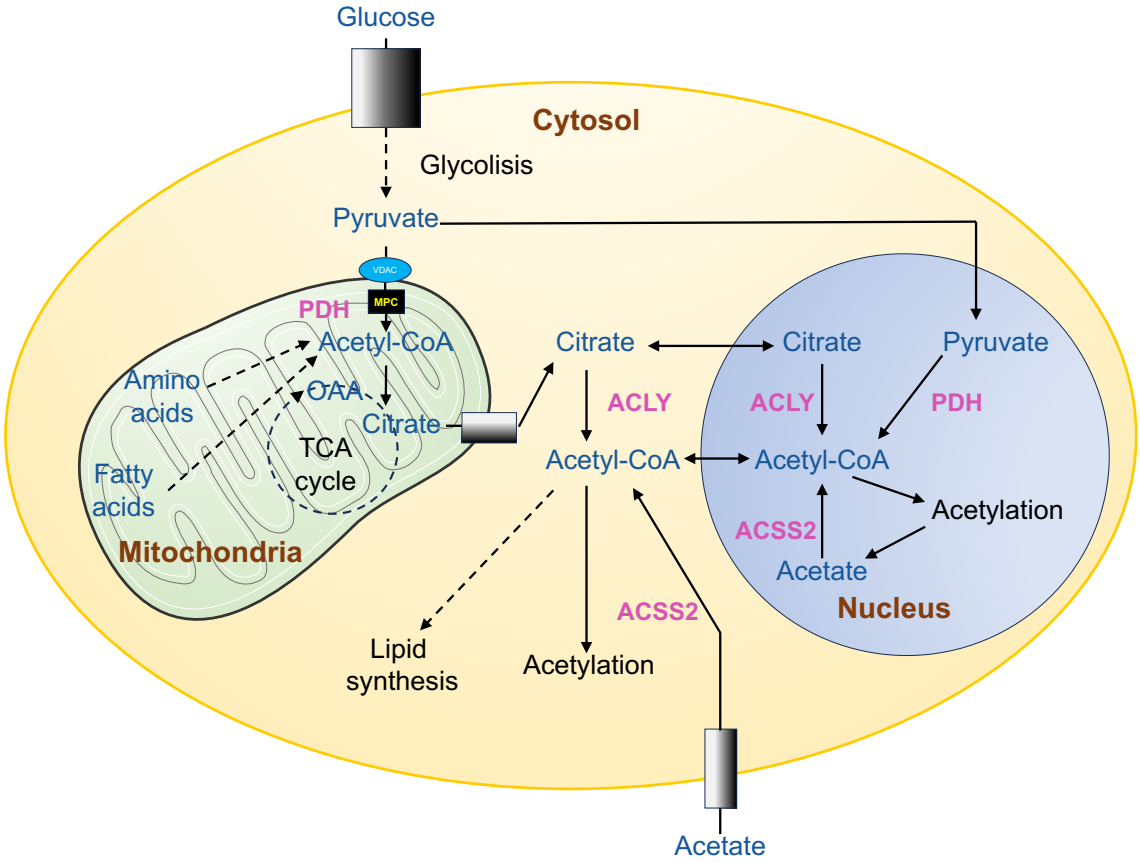


Figure 12. Acetyl CoA metabolic sources. *Scheme showing the main different sources and uses of Acetyl-CoA in the three compartments: Mitochondria, Cytosol and Nucleus. Implicated enzymes appear represented in pink and the different substrates and products in blue.*

Acetyl-CoA serves as a sentinel metabolite, providing cells with essential information about their metabolic state. This guidance is achieved through unique protein acetylation modifications that are intricately linked to the abundance of acetyl-CoA (Galdieri et al., 2014).

In a state of growth or nourishment, elevated nucleocytosolic acetyl-CoA levels serve as a signature, promoting its utilization for processes such as lipid synthesis and histone acetylation. Conversely, during survival or fasting states, acetyl-CoA is preferentially directed into the mitochondria, contributing to mitochondrial-dependent functions like ATP synthesis and ketone body formation. This adaptive metabolic shift highlights acetyl-CoA's versatility in responding to the dynamic energy needs of the cell.

The fluctuation of acetyl-CoA within distinct subcellular compartments facilitates substrate-level regulation of acetylation modifications. This process necessitates the activity of sirtuin deacetylases. These enzymes play a crucial role in catalyzing the removal of spontaneous acetylation modifications that may occur unintentionally.

Finally, histone acetylation is fueled by acetyl-CoA. Histone acetylation emerges as a central switch that controls DNA transcription, replication and repair; allowing interconversion between permissive and repressive chromatin structures and domains (Grunstein, 1997; Vogelauer et al., 2002).

3.4. Metabolic enzymes in the nucleus

The direct influence of metabolites on transcriptional decisions is emerging. Of note, enzymes traditionally associated with mitochondria have been recently detected in the nucleus, and they are turning out to be newly discovered major players in processes such as transcriptional regulation and chromatin organization:

In 2006, ACSS2 was firstly located in the nucleus of cells, and directly linked to chromatin regulation (Takahashi et al., 2006). Three years later, another metabolic enzyme: ACLY was found in the nucleus, working as a source of Acetyl-CoA for histone acetylation (Wellen et al., 2009). All of the pyruvate dehydrogenase complex enzymes where found in the nucleus, where they were found to have a critical role in histone acetylation (Sutendra et al., 2014). In 2016 the group of Dr. Beato described how chromatin remodeling energy can be derived from poly-ADP-ribose by nucleoside diphosphate linked moiety x-type motif 5 (NUDIX5) within the nucleus, bypassing the usual reliance on ATP from the cytoplasmic mitochondria. (Wright et al., 2016). A portion of folate pathway enzyme methylenetetrahydrofolate dehydrogenase, cyclohydrolase and formyltetrahydrofolate synthetase 1 (MTHFD1) was recently found to be present in the nucleus, where it is recruited to specific genomic locations through a direct interaction with BRD4, implying a direct involvement in the regulation of gene expression (Sdelci et al., 2019).

As nuclear metabolism is an emerging field of study, many unknowns regarding the regulation of these enzymes, the subnuclear localization and their working partners remain to be unveiled.

The direct readout is that nuclear metabolic enzymes and intermediates may drive, nuclear processes, of which RNA Pol II-driven transcription is perhaps the most relevant for the control of cell identity.

3.4.1. . The pyruvate dehydrogenase complex

The PDH complex is, , a crucial enzyme complex that plays a central role in cellular energy metabolism and in the generation of acetyl groups. It is abundant in the mitochondria, where it catalyzes the oxidative decarboxylation of pyruvate, to produce acetyl-CoA, a critical metabolic intermediate for several processes (Linn et al., 1969). The intact PDH complex of ~30-40 nm diameter has been reported to reside not only in the mitochondria, but also in the nucleus (Sutendra et al., 2014).

Importantly, the transcriptional machinery and histones are heavily acetylated, as mentioned above, to maintain the chromatin in an open or accessible state during active transcription(Clayton et al., 2006; Cramer, 2019; Hebbes et al., 1988; Weinert et al., 2018).

Global levels of histone acetylation respond to the abundance of acetyl-CoA, although the mechanisms of specificity in gene regulation by acetyl-CoA are unclear. Acetyl-CoA is produced in three main cellular compartments: the mitochondria, cytosol, and the nucleus through several direct and undirect pathways (Sivanand et al., 2017) (Figure 12).

Although acetyl-CoA is permeable through nuclear pores, recent metabolic evidence suggests that nuclear and cytosolic pools of acetyl-CoA may be, in large part, functionally distinct. This is explained by the obstruction of metabolite diffusion by macromolecular crowding. The nucleus is highly packed with high concentrations of proteins and DNA. Moreover, the use of exogenous acetate at physiological concentrations is inefficient for histone acetylation in cells with impaired acetyl-CoA production (Sivanand et al., 2017).

While the PDH complex is not the exclusive source of acetyl-CoA for chromatin acetylation, its role in the maintenance of a correct pattern of histone acetylation is critical (Chen et al., 2018; Nagaraj et al., 2017; Sutendra et al., 2014)

The pyruvate dehydrogenase Complex (Figure 13) is constituted by three catalytic enzymes: pyruvate dehydrogenase: E1 (PDHA1 and PDHB), dihydrolipoamide transacetylase: E2 (DLAT), and dihydrolipoamide dehydrogenase: E3 (DLD and PDHX), as well as the tethering protein, E3-binding protein (E3BP) (Behal et al., 1993).

DLAT, the catalytic core subunit of the PDH complex, is known to use a lipoyl side-chain in its enzymatic centre to catalyse a rate-limiting step in acetyl-group production.

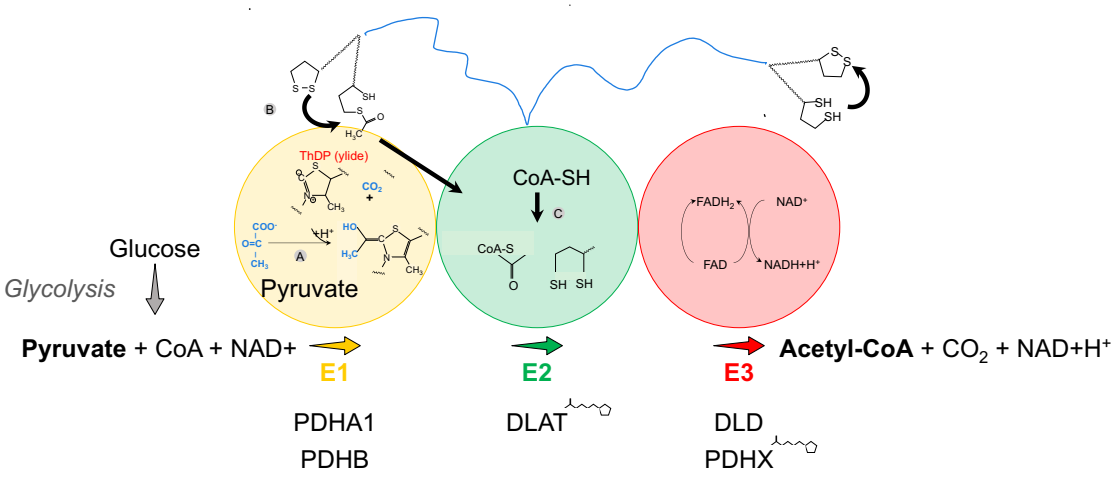


Figure 13. The Pyruvate Dehydrogenase Complex reaction. *PDH catalyzes multiple reactions in its three different subcomplexes (E1, E2 and E3) for the conversion of pyruvate into acetyl-CoA.*

Since histone acetylation marks sites of active RNA Pol II-transcription, the presence of PDH in the nucleus may provide a mechanism by which the abundance of the acetyl-CoA intermediate can directly modify the transcriptional program. However, nothing is known about the sub-nuclear distribution of the PDH or its nuclear partners.

4. ORGANIZATION AND METABOLIC CONTROL OF TRANSCRIPTION

4.1. Chromatin structure and histone marks

Chromatin is the complex structure formed by DNA, histone proteins, and other regulatory proteins in the nucleus of eukaryotic cells. It serves as the packaging material for genetic information, allowing the cell to efficiently organize, compact, and regulate access to its DNA (Kornberg, 1974).

The structural hierarchy of chromatin extends from the fundamental nucleosome level to higher-order conformations, ultimately forming chromosomes. Nucleosomes are comprised of DNA wound around histone octamers, and they serve as the basic repeating unit. Beyond nucleosomes, chromatin undergoes folding and compaction, leading to the formation of chromosomal territories. The spatial organization of these territories is orchestrated by factors, including DNA sequence, architectural proteins, and non-coding RNAs (Boyle et al., 2001).

Chromosomal organization governs the spatial arrangement of genetic material within the nucleus. This organization plays a pivotal role in regulating gene expression, DNA replication, and repair processes(Schneider & Grosschedl, 2007).

Epigenetic modifications, particularly histone marks and DNA methylation, contribute significantly to chromosomal organization. Histone tails undergo various post-translational modifications, such as methylation, acetylation, and phosphorylation, influencing chromatin structure and gene accessibility. DNA methylation, often associated with transcriptional repression, further refines chromatin domains. The intricate crosstalk between different epigenetic marks dynamically shapes the functional chromatin landscape, delineating regions into heterochromatin and euchromatin (Handy et al., 2011) (Figure 14).

Heterochromatin: It is the tightly compacted form of chromatin. It is transcriptionally repressive, characterized by condensed nucleosomes and enriched in specific histone modifications from which we can highlight a global increase in methylation and the role of non-coding RNAs. Heterochromatin plays a role in gene silencing, genome stability, and centromere function.

Euchromatin: Euchromatin represents a more open and transcriptionally active state of chromatin. It is associated with high levels of histone acetylation and actively transcribed genes. Due to the high levels of acetylation, it exhibits a more accessible nucleosome structure, facilitating the binding of regulatory proteins.

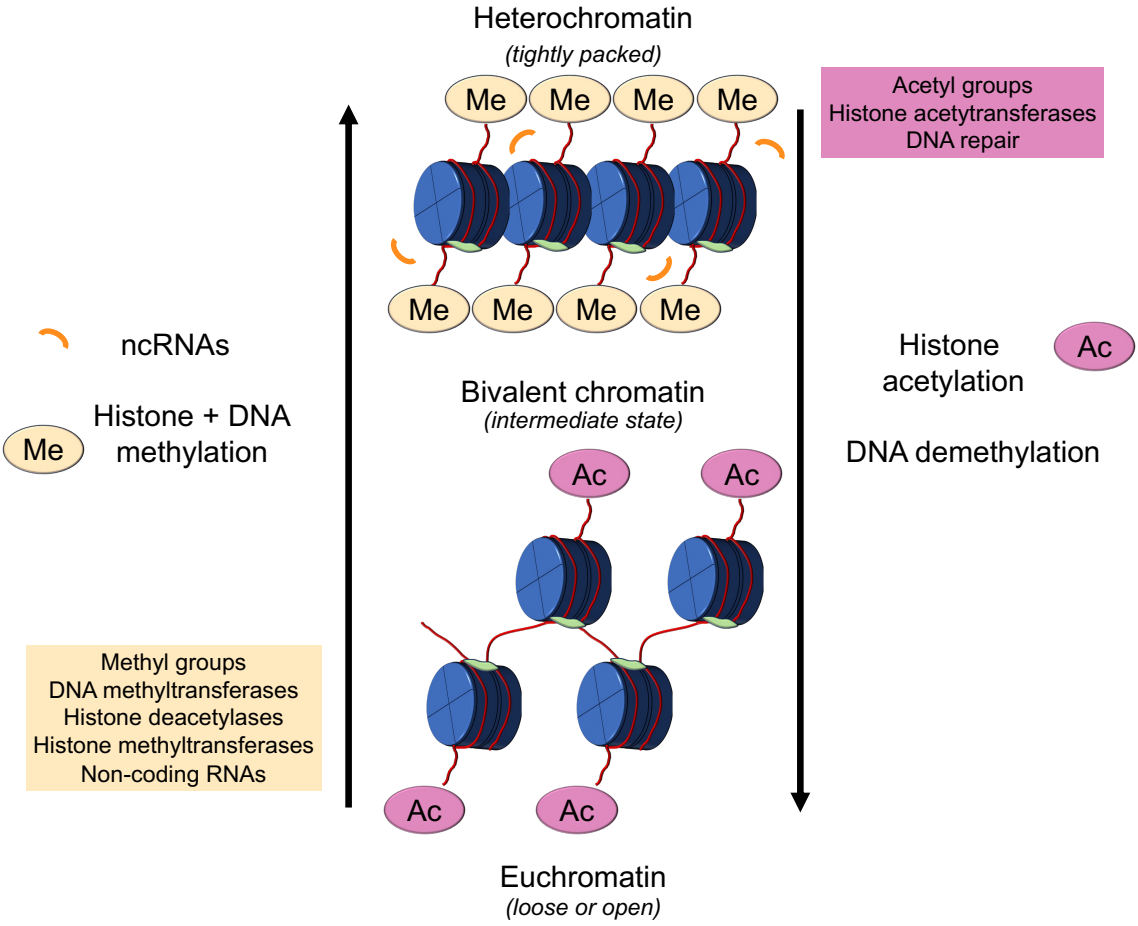


Figure 14. Chromatin States and Transitions. *This schematic illustrates the distinctive characteristics of heterochromatin, euchromatin, and the intermediate state of bivalent chromatin. Key features include chromatin compaction and methylation in heterochromatin, an open, acetylated and accessible structure in euchromatin, and a bivalent state during the transition between both. Additionally, the figure highlights main events mediating transitions between these states.*

Chromosomes are not randomly distributed within the nucleus; instead, they exhibit specific spatial arrangements. The positioning of chromosomes is influenced by factors such as gene density, transcriptional activity, and interactions with nuclear structures. Emerging techniques, including Hi-C and 3C-based methods, provide unprecedented insights into the spatial organization of chromosomes, revealing intricate patterns of interactions between genomic loci (Akgol Oksuz et al., 2021).

Topologically Associated Domains (TADs) represent genomic regions with enhanced internal interactions, delineating functionally distinct chromatin domains. TAD boundaries, often demarcated by architectural proteins, act as insulators, preventing inappropriate interactions between neighboring genomic regions. The disruption of TAD organization has been implicated in diseases, emphasizing the functional significance of these chromatin domains (Dekker & Heard, 2015).

Chromosomal organization is not static but exhibits dynamic changes in response to cellular cues and developmental stages. Chromatin loops, formed within TADs, facilitate long-range interactions between distant genomic elements, allowing enhancers and target genes to coincide, despite being distantly located. Transient chromatin interactions play a crucial role in orchestrating rapid responses to environmental stimuli, contributing to the adaptability and plasticity of the genome (Yu & Ren, 2017).

4.2. Phase separation and chromatin organization

As introduced before, liquid-liquid phase separation (LLPS) is a biophysical phenomenon demanding significant attention as it appears to underlie most processes in cell biology. LLPS, also referred to as phase-separation involves the formation of distinct liquid phases within cellular compartments. Up to 50% of human proteins are calculated to contain repetitive amino acid sequences with unstructured character. These are referred to as intrinsically disordered repeat regions or IDRs. Protein-protein interactions via their IDRs are weak and transient, permitting liquid-like freedom of movement, all of which resemble the interaction and qualities of water molecules, leading to the concept of protein-liquids within which proteins containing IDRs of similar chemical character, can collate within proximity to achieve their collective fuction. These protein-based liquids can be rapidly dissolved or nucleated by post-translational modifications. Thus, LLPS achieves a high degree of compartmentalization that is also rapid and highly regulatable, all without a membrane. Essentially, LLPS achieves the formation of membraneless organelles.

Well-studied examples of phase separated membranless organelles include P-granules (Brangwynne et al., 2009) and stress granules (Molliex et al., 2015) in the cytoplasm, and the nucleolus (Feric et al., 2016), histone locus bodies (Hur et al., 2020), and nuclear speckles (Lu et al., 2018), in the nucleus (Sabari et al., 2020). Heterochromatin formation

has also been proposed to be influenced by LLPS condensates of HP1(Larson et al., 2017).

Although these examples primarily involve proteins and RNA, comparable phase behavior is observed in complexes formed by proteins and DNA.

The orchestration of transcriptional events within the nucleus and specifically during transcription are of particular interest in the context of understanding mediator complex role in the control of cellular identity and transcription.

Biomolecular phase separation happens through multiple weak interactions between proteins and/or (ribo)nucleic acids, that separate a single homogeneous phase into multiple phases, similar to what happens between oil and water. Typically, proteins that phase separate include IDRs in their structure that act as biophysical addresses to localize interactor proteins, and they contribute an essential structural component to stabilize the LLPS condensates (Dignon et al., 2020). Overall, this suggests the minute compartments of nuclear organization, such that regulators are held in proximity to target regions.

Chromatin can undergo LLPS to form condensates, dynamic structures that influence the spatiotemporal organization of gene expression.

Super-resolution imaging has allowed to observe how Mediator and RNA polymerase II association happens through phase separation, forming transcriptional condensates (Cho et al., 2018).These transcriptional condensates, or "transcriptional hubs," are implicated in coordinating the assembly of transcriptional machinery and regulatory elements. Importantly, super-enhancers, which contain the highest amounts of Mediator complex and drive robust expression of genes with critical roles in cell identity, are phase separated, compartmentalizing, and concentrating the transcription apparatus. IDRs from Mediator complex other co-ativators, and transcription factors are key in this process (Sabari et al., 2018) (Figure 15).

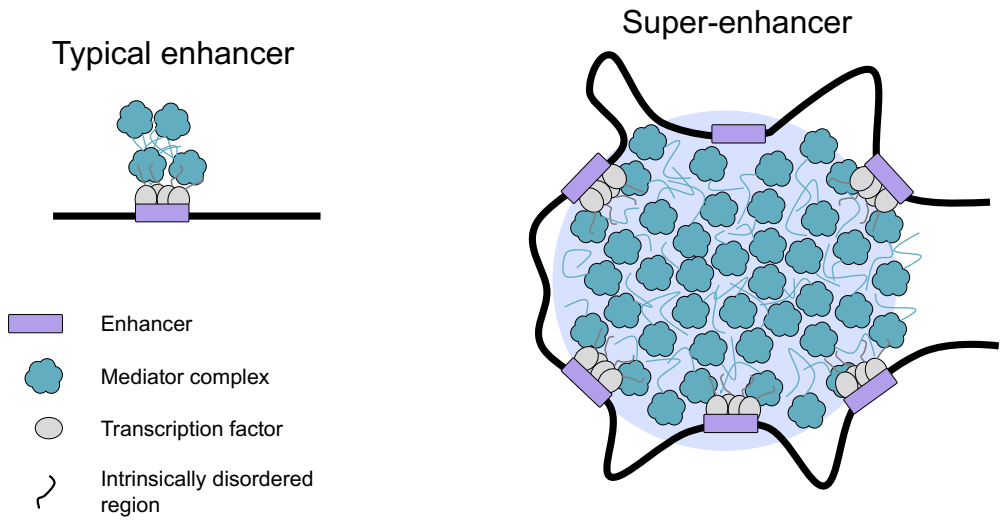


Figure 15. Phase Separation, Enhancers, and Super Enhancers. *Visual representation of phase separation, highlighting the distinctions between typical enhancers and super enhancers. It illustrates how super enhancers undergo liquid-liquid phase separation (LLPS), creating distinct condensates. The figure further emphasizes the high concentration of the mediator complex within super enhancer condensates. Intrinsically disordered regions of transcription factors and mediator complex are depicted, showcasing their interactions within the dynamic environment of super enhancers.*

The mechanisms implied in the regulation of transcriptional condensates are being studied, and modifications as phosphorylation or acetylation, that affect the charge of the proteins implicated have been shown to have a direct impact in the formation and dissolution of these condensates. For example, the phosphorylation of the carboxyl terminal domain (CTD) of RNA Pol II regulates a switch between the transcriptional initiation phase separated condensates (Figure 16), and elongation and splicing condensates (Y. E. Guo et al., 2019a). Histone acetylation driven by the multifunctional transcriptional co-activators CBP/p300 also regulates transcriptional phase separation by dissolving and forming new phase-separated droplets and therefore modulating the organization of nuclear chromatin subdomains (Gibson et al., 2019; Weinert et al., 2018)

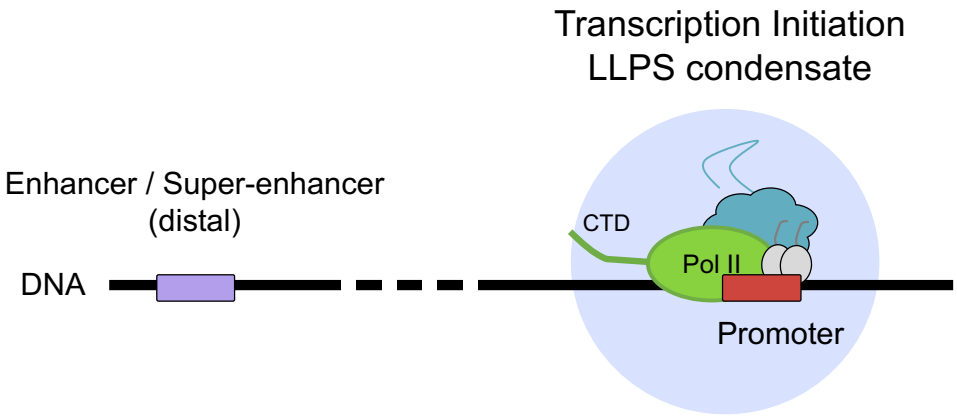


Figure 16. Transcription Initiation Phase-Separated Condensates. *transcription initiation, emphasizing the formation of phase-separated condensates at gene promoters. The illustration highlights the condensation of RNA Polymerase II (Pol II) and the Mediator complex at gene promoters, showcasing the localized concentration of key transcription initiation components.*

4.3. Metabolic control of transcription: A nexus of histone methylation and acetylation.

Two key histone epigenetic marks, methylation and acetylation, are of particular interest for this thesis, given the previously mentioned connections between chromatin organization, and metabolic pathways. Studies of histone acetylation and methylation will likely shed light on how metabolic cues shape the chromatin landscape and influence gene expression (Filtz et al., 2014; Gibney & Nolan, 2010; Grunstein, 1997; Vogelauer et al., 2002).

Above, we described the source of both acetyl and methyl groups. Next, we will better understand their presence and impact in chromatin.

Histone acetylation consists of the addition of acetyl groups to lysine residues on histone tails. It is catalyzed by histone acetyltransferases (HATs), generally correlates with transcriptional activation by neutralizing the positive charge on histones, leading to a more open chromatin structure. That facilitates the binding of transcriptional machinery and regulatory proteins (Marmorstein & Zhou, 2014).

Conversely, **histone methylation** is the process of addition of methyl groups to lysine or arginine residues on histone tails. Histone methyltransferases (HMTs) are the enzymes that add methyl groups. It can be associated with both activation and repression,

depending on the specific lysine or arginine residue modified and the degree of methylation (mono-, di-, or trimethylation) (Andrew J et al., 2002).

In the table below (Table 1), we provide a list of some of the best documented acetylation and methylation marks, with a very brief description of their role in transcription.

	Mark	Outcome	Role
Acetylation Marks	H3K9Ac	Activation	Neutralizes positive charge, opens chromatin,marks active promoters, facilitates binding of transcriptional activators.
	H3K122Ac	Activation	Open chromatin, marks active gene promoters and also a subset of active enhancers
	H3K18Ac	Activation	Linked to open chromatin and active gene expression. Enhances accessibility of chromatin.
	H3K27Ac	Activation.	Mark for active enhancers and actively transcribed genes. Marks regions of high transcriptional activity.
	H4K5Ac	Activation	Contributes to open chromatin structure. Supports binding of transcriptional activators.
	H4K8Ac	Activation	Contributes to open chromatin structure. Supports binding of transcriptional activators.
	H4K16Ac	Activation	Generally associated with transcriptional activation. Involved in preventing chromatin compaction.
Methylation Marks	H3K4me1	Bivalent	Repressed but inducible enhancers
	H3K4me3	Activation	Mark for active enhancers
	H3K9me3	Repression	Primarily associated with transcriptional repression. Establishes repressive chromatin, promotes heterochromatin formation.
	H3K27me3	Repression	Strongly associated with transcriptional repression. A hallmark of polycomb-repressed chromatin. Contributes to facultative heterochromatin formation.
	H3K36me	Activation	Associated with transcriptional activation and elongation. Marks actively transcribed regions. Facilitates transcriptional elongation.
	H3K79me	Activation	Primarily associated with transcriptional activation. Found in actively transcribed genes. Transcriptional regulation and DNA repair.
	H4K20me	Bivalent	Implicated in both transcriptional activation and repression, depending on the degree of methylation.

Table 1. Summary of Key Acetylation and Methylation Histone Marks. *This table provides a comprehensive overview of histone modifications relevant to understanding this thesis. Acetylation marks are highlighted in pink, representing their role in transcriptional activation and chromatin accessibility. Methylation marks are shown in yellow, indicating their involvement in diverse regulatory functions, such as gene activation or repression, depending on the specific lysine residue. The brief descriptions offer insights into the functional roles of each modification in shaping chromatin structure and gene expression.*

It is important to note that the functional outcomes of each histone modifications can vary depending on the context of specific genomic regions and the presence of other modifications. Additionally, crosstalk between different histone marks adds a layer of

complexity to the regulation of chromatin structure and gene expression (Suganuma & Workman, 2008).

The nucleus, once considered a place for the storing and processing of genetic information, is now recognized as a dynamic hub where metabolic and epigenetic pathways converge. As we mentioned above, more and more metabolic enzymes have been found in the nucleus during the last years.

Traditionally associated with cellular energy production in mitochondria and the cytoplasm, more recently these metabiolic enzymes also appear to operate within the nuclear envelope. This challenges conventional roles and expands our view of metabolic-epigenetic connections.

These discoveries prompt crucial questions as for example, where precisely within the nucleus do these enzymes reside, and how do they interact with and regulate chromatin sub-regions.

Objectives

The primary objective of this thesis is to enhance our comprehension of the Mediator's complex role in transcriptional control of cellular identity and plasticity. Additionally, this project seeks to uncover potential novel regulators influencing cell identity. To accomplish these overarching goals, our specific aims are as follows:

1. **Examine Changes in Mediator-Pol II Interaction:**
   1.1. Develop and apply novel visualization methods to efficiently capture the Mediator-Pol II interaction dynamics.
   1.2. Investigate alterations in the interaction dynamics between the Mediator complex and RNA Polymerase II (Pol II) during transitions in cellular identity.

2. **Identify Novel Mediator Interactors:**
   2.1. Uncover and characterize previously unrecognized interactors of the Mediator complex that may play pivotal roles in regulating cellular identity.

3. **Explore Mediator Interaction with Nuclear PDH Complex and Manipulate their Interaction:**
   3.1. Delve into the intricacies of the interaction between the Mediator complex and the nuclear Pyruvate Dehydrogenase (PDH) complex.
   3.2. Explore the potential functional activity of PDH in the nucleus.
   3.3. Investigate PDH chromatin relation through ChIP-Seq analysis.
   3.4. Experimentally manipulate the interaction between the Mediator complex and the nuclear PDH complex to gain deeper insights into the functional significance of this complex in the transcriptional regulation of cell identity.

In summary, we explore new aspects of the Mediator complex interactome, including RNA Pol II and the control of cell identity. Furthermore, we reveal a new interaction with the PDH complex within enhancers inside the nucleus and study its regulation.

Materials and methods

Cell culture

Mouse ES cells: E14Tg2a.4 (wild-type parental, 129/Ola background) were from BayGenomics/MMRRC resource, University of California; Wild-type ES cells were derived at the Transgenic Mouse Unit of CNIO from E4.5 C57BL6 blastocysts, or mixed background C57BL6/129 blastocysts. All ES cells were cultured on gelatin-coated plates in "Serum/LIF": 15%FBS (Gibco) in DMEM (high glucose) basal media with LIF (1000 Units/mL), non-essential amino acids, glutamax and ß-mercaptoethanol plus antibiotics (penicillin/streptomycin 100U/mL). (i) Primed media: no extra drugs were added to the culture conditions mentioned above; (ii) "2i" treated ES cells: "2i" drug cocktail, 1 µM Mek-inhibitor (PD0325901, Axon Medchem, #1408) plus 3 µM GSK3ß- inhibitor (CHIR 99021, Axon Medchem #1386) as described (Ying et al., 2008); (iii) CDK8/19i treated cells: 1µM CDK8/19i (developed and validated *in vitro* and *in vivo* by the Experimental Therapy Unit from the CNIO).

Primary mouse embryo fibroblasts (wild-type, MEFs, passage 2) were obtained at E13.5 from pure inbred C57BL6 background mice as described previously (Palmero and Serrano., 2001) and were maintained in DMEM medium with 10% FBS (Gibco) with antibiotics (penicillin/streptomycin 100 U/mL).

Human melanoma SKMEL-103 and mouse melanoma B16 cells were obtained from the ATCC. Both cell lines were maintained in standard DMEM medium supplemented with 10% heat inactivated fetal bovine serum (FBS) (Gibco) and 1% antibiotics (penicillin/streptomycin 100 U/mL; Gibco).

Galactose media: galactose media was prepared using as basal media no glucose DMEM medium (Life Technologies #11966025) suplemmented with 10 mM galactose (Sigma) 10% knockout serum replacement (KSR, Invitrogen) and 1mM pyruvate (Agilent).

mESC differentiation with LIF removal and retinoic acid. LIF was first removed for 24 h by culturing in LIF-free differentiation medium (as described for serum/ LIF medium, except LIF was omitted). Next, retinoic acid was added (10 µM) from 24 h to 72 h, followed by LIF-free differentiation medium alone for 5-9 days as indicated in the figure legends.

All cultures were routinely tested for mycoplasma.

Nuclear/cytoplasmic and Cytoplasmic/mitochondrial fractionation. .

Nuclear and cytoplasmic fractionation was performed with Thermo Scientific NE-PER Nuclear and Cytoplasmic Extraction Kit (Following the protocol provided in the commercial kit). Cytoplasmic and mitochondrial fractions were prepared using the Mitochondria Isolation Kit (89874, Thermo Fisher Scientific). Fresh protease inhibitors (Roche #11873580001) were added in the required steps.

Immunoblot

Cell pellets were lysed in 50mM Tris-Hcl pH 7.4, 0.5% NP-40, 250mM NaCl, 5mM EDTA, with freshly added 1mM Sodium Orthovanadate, 1mM NaF, protease inhibitors (#87785, Thermofisher) and phosphatase inhibitors (#4906837001, Roche). 15-25mg of proteins were run on NuPage 4-12% gradient Bis-Tris gels in NuPAGE™ MES SDS Running Buffer (20X) and wet-transferred to 0.2mM nitrocellulose membranes (#10600001, GE Healthcare). Blots were blocked in Intercept (TBS) Blocking Buffer (#927-60001, LI-COR) for 30 minutes and then incubated with the primary antibodies indicated in (table 2) over-night. The following day the membranes were incubated with the following secondary antibodies: IRDye 800 CW anti-rabbit (1:5000, #926-32211, Licor Odissey) and IRDye 800 CW anti-mouse (1:5000, #926-32210, Licor Odissey).

Cell immunofluorescence

PSCs and MEFs were grown on chamber slides using the same protocols as for the rest of the experiments. In addition, to aid cell attachment to the slides, slides were pre-treated for 30 minutes with fibronectin 5µg/mL (R&D Systems #1918-FN-02M) diluted in PBS. Cells were fixed 24 hours after plating on chamber slides with 4% paraformaldehyde for 3 minutes at room temperature, washed with PBS and permeabilized with PBS containing 0.1% Triton X100 for 1 hour. Cells were blocked in 10% FBS/1xPBS for 1h and incubated with antibodies (for a list of the antibodies used, see **Antibodies Table**) at 1:200 to 1:1000 in PBS-4%BSA, for 3 hours at 37oC or overnight at 4oC, washed with PBS and further incubated with secondary anti-rabbit or anti mouse antibodies conjugated with Alexa-

488, Alexa-555 and/or Alexa-647 (1:500 in PBS-4%BSA). Nuclei were counter-stained with DAPI.

Proximity Ligation Assay (PLA)

PSCs and MEFs were grown on chamber slides using the same protocols as for the rest of the experiments. In addition, to aid cell attachment to the slides, slides were pre-treated for 30 minutes with fibronectin 5µg/mL (R&D Systems #1918-FN-02M) diluted in PBS. Cells were fixed 24 hours after plating on chamber slides with 4% paraformaldehyde for 3 minutes at room temperature, washed with PBS and permeabilized with PBS containing 0.1% Triton X100 for 1 hour. Cells were blocked in Duolink® Blocking Solution in a previously heated humidity chamber for 60 minutes at 37 °C and incubated with antibodies (for a list of the antibodies used, see **Antibodies Table**) diluted in Duolink® Antibody Diluent at 1:80 to 1:100 for 1 hour at 37oC in a humidity chamber. After antibody incubation, Duolink® PLA probes PLUS and MINUS diluted in Duolink® Antibody Diluent at 1:3 were incubated for 1 hour at 37oC in a humidity chamber. Probes were ligated adding 1x Duolink Ligation Buffer with ligase for 30 minutes at 37oC in a humidity chamber. Finally, 1x Duolink® Amplification Buffer with polymerase was added and the slides were kept amplifying for 50-100 minutes in a humidity chamber at 37oC. Amplification time was optimized for each combination of antibodies and the same amplification time was always used between different treatments. Nuclei were counter stained with the Duolink® In Situ Mounting Media which contains DAPI.

siRNA transfection

All the siRNAs were purchased from siTOOLs. siRNAs have been used at 3nM final concentration. SK-Mel-103, E14Tg2a.4, and V6.4 (Wild type mESCs) were transfected with siRNAs for 5 days before performing viability assays. Lipofectamine reagent RNAiMAX (Cat. 13778075) was used at 2mL/mL diluted in Opti-MEM (Cat. 31985062) to perform the transfection.

After 5 days from the transfection, cell viability assays were performed (Crystal Violet Assay).

Cell viability assay

Cell viability assay were performed after siRNAs transfections (5 days) in proliferating, SK-Mel-103 and E14Tg2a.4 background mES cells in control media; in galactose media or in control media supplemented with 5 mM acetate (#S2889) using Crystal Violet (#C6158, Merck).

For crystal violet staining, cells were first washed once with PBS and then fixed stained through the crystal violet solution (6.179mM crystal violet in 20% methanol solution in H2O) for 30 minutes. The solution was then removed from the plates and cells were washed in H2O at least 5 times and dry under a chemical hood overnight. Images were taken using an HP scanner. For quantification, the staining was lysed using lysing solution (0.1M sodium citrate, 50% ethanol at pH 4.2). Absorbance was measured using a SYNERGY HTX Absorbance microplate reader (Agilent Technologies) at 570nM.

Chromatin Immunoprecipitation (ChIP) sample preparation

mESCs were cultured in standard 2i mESC media, were fixed with 1% (v/v) PFA (Fisher Sceintific #50980487) then, chromatin was prepared by using the iDeal ChIP-seq kit for Transcription Factors (Diagenode # C01010055), then sonicated using a Diagenode BioRuptor Pico (Diagenode #B01060010) for 12 cycles (30' on, 30' off) at 4°C. Lysates were clarified for 10 minutes at 8000 x g, 1% input samples were reserved, and supernatant was used for immunoprecipitation with Diagenode Protein A-coated Magnetic beads ChIP-seq grade (Diagenode #C03010020-660) and DLAT mAb (MyBioSource # MBS9404684) with 0.1% BSA (Sigma #10735094001). The following day, cells were washed once with the iDeal ChIP-seq kit for Transcription Factors (Diagenode # C01010055) buffer and eluted in 1% SDS, 100mM NaHCO3 buffer. Cross-links were reversed with RNaseA (ThermoFisher #EN0531), proteinase K (Merck #3115879001) and sodium chloride (Sigma #71376), and chromatin fragments were purified using QIAquick PCR purification kit (Qiagen #28104).

ChIP sequencing

The concentration of the DNA samples (inputs and IPs) was quantified with Qubit dsDNA HS kit and fragment size distribution was assessed on a TapeStation4200 using the D5000 HS assay (Agilent). Libraries for ChIP-Seq were prepared at IRB Barcelona Functional Genomics Core Facility. Briefly, dual-indexed DNA libraries were generated from 2.66 – 10.26 ng of DNA samples using the NEBNext Ultra II DNA Library Prep kit for Illumina (New England Biolabs). 14 and 15 cycles of PCR amplification were applied to all libraries of projects ICalvoJan23_ChIPSeq, ICalvoMay23_ChIPSeq, respectively.
The final libraries were quantified using the Qubit dsDNA HS assay (Invitrogen) and quality controlled with the Bioanalyzer 2100 DNA HS assay (Agilent). Two equimolar pools were prepared, containing the libraries of each project and sequenced using the PE150 strategy on a NovaSeq 6000 (Illumina). A minimum of 44 million of reads were obtained for all samples.

ChIP sequencing analysis

Adapters and low-quality bases (<Q20) were removed from DLAT and MED15 ChIP-Seq reads with Cutadapt (v4.4) (Martin., 2011) and TrimGalore (v0.6.10) (https://github.com/FelixKrueger/TrimGalore). Trimmed reads were mapped to the mouse reference genome mm10 with Bowtie2 (v2.5.1) (Langmead and Salzberg., 2012) using default paired-end parameters. PCR duplicates were removed with Picard Tools MarkDuplicates (v1.97). BigWig files were created using DeepTools bamCoverage (v3.0.2) (Ramírez et al., 2014) with the following parameters: --normalizeUsingRPKM –ignoreForNormalization chrX chrY –samFlagInclude 64 –extendReads --binSize 20 --smoothLength 40. Reads overlapping problematic, blacklisted regions (mm10-blacklist.v2.bed) (Amemiya et al., 2019) were excluded from the computation of coverage with the option --blackListFileName. Peak calling was performed using MACS2 callpeak (v2.2.9.1) (Zhang et al., 2008) (-q 0.05) on deduplicated bam files, using a ChIP-Seq for IgG as control. After discarding peaks overlapping blacklisted regions, peaks laying less than 5kb apart were merged.

The following public datasets were downloaded for comparison against the DLAT ChIP-Seq: Pol II ChIP-Seq in V6.5 mESc (GSE178848; (Sun et al., 2021)),), MED1 ChIP-Seq in V6.5 mESc (GSE178848; (Sun et al., 2021)), MED4 ChIP-Seq in V6.5 mESc (GSE178848(F. Sun et al., 2021)), H3K9Ac ChIP-Seq in V6.5 mESc (GSE178848; (Sun

et al., 2021), H4K5Ac ChIP-Seq in E14 mESc (GSE158736; (Radzisheuskaya et al., 2021)), H4K8Ac ChIP-Seq in E14 mESc (GSE158736; (Radzisheuskaya et al., 2021), H3K27Ac ChIP-Seq in V6.5 mESc (GSE200120; (Zhao et al., 2021)), H3K4me1 ChIP-Seq in V6.5 mESc (GSE200120; (Zhao et al., 2021)), H4K16Ac ChIP-Seq in E14 mESc (GSE43102; (Taylor et al., 2013)), H3K122Ac ChIP-Seq in 129S4/SvJae J1 mESc (GSE141525; (Zhang et al., 2020)), and H3K18Ac ChIP-Seq in 129S4/SvJae J1 mESc (GSE141525; (Zhang et al., 2020)). Processing of these experiments was done as described above for the in-house DLAT and MED15 datasets.

Annotation of DLAT, MED1, and H3K27Ac peaks over genomic features was done using the R libraries 'ChIPseeker' (Wang et al., 2022) and 'TxDb.Mmusculus.UCSC.mm10.knownGene'. DLAT target genes were defined as having a DLAT peak in their promoter (TSS +/- 3kb). DLAT target genes were then tested for enrichment in Gene Ontology (BP: Biological Process) terms (Ashburner et al., 2000; Gene Ontology Consortium et al., 2023).

Average plots of signal over enhancers were generated with Deeptools computeMatrix (v3.0.2) (Ramírez et al., 2014) and in-house scripts.

ChromHMM (Ernst et al., 2012) was used to understand how DLAT associates with Mediator and Pol II. The genome was analyzed at 200-bp intervals and a model of 8 states was chosen to characterize it according to the relative enrichment for DLAT, Pol II (Sun et al., 2021) MED1 (Sun et al., 2021), MED4 (Sun et al., 2021), H3K27Ac (Zhao et al., 2021), and H3K4me1 (Zhao et al., 2021).

Image acquisition

All immunofluorescence and PLA images were captured using a Leica SP5 confocal microscope. Once the settings were decided for each experiment, all the acquired images shared the same laser power settings for each channel independently of the cell type or treatment.

CHIP-qPCR

Quantitative real-time PCR (RT-qPCR) was performed using GoTaq PCR Master Mix (Promega, #A6002) and specific primers listed below (Table 3). The reaction was performed in a QuantStudio 6 Flex thermocycler (Applied Biosystems).-
All the reactions were performed in triplicates.

Quantification and statistical analysis

The number of PLA interactions were quantified using an image analysis script developed in collaboration with the Advanced Microscopy Unit from the IRB. This script was designed to count the number of foci detected inside the nucleus of each cell. This script not only allows foci quantification, but also, it gives information about the size and signal intensity of each focus. Importantly, it discriminates between background and signal, and includes a flexible threshold for spot size, thereby providing an unbiased count of foci per nucleus. The script has been trained and tested on multiple confocal immunofluorescent images followed by manual inspection, one nucleus at a time, to confirm accuracy. It performs an automated counting of foci per nucleus on max intensity projections of several z-slices through each nucleus. In this way, foci throughout the full height of each nucleus were included (not simply a single random section through part of the nucleus). This script uses a mask adjusted by DAPI that discards the cytoplasm and keeps the nuclei. A second mask counts the dots which are above a threshold of intensity (to discard possible noise and false positives). Quantitative data acquired with the image analysis script specifically designed for PLA quantification, was presented as mean +/- SD and significance was assessed by the two-tailed Student's t test; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.001$.

Table of antibodies

Antibodies used in this study		
Target	Company	Code
MED1	Cell Signaling Technology	#51613
POLR2A NTD DTLY	Cell Signaling Technology	#14958

POLR2A CTD 4H8	Cell Signaling Technology	#2629
POLR2A CTD 8WG16	Abcam	#ab817
PDHA1	Abcam	#ab110330
DBT	Novus	NBP1-89522
DLA T	MyBioSource	MBS9404684
TOM20	Santa Cruz Biotechnology	sc-11415 (FL-145)
LAMIN A/C	Santa Cruz Biotechnology	sc-6215 (N-18)
GAPDH	Sigma	G8795
SMC1	Bethyl Laboratories	# A300-055A

Table 2. List of the antibodies used in this study. *Targets, commercial brands and code appear as indicated.*

Table of primers

Primers used in this study		
Primer	Forward sequence	Reverse sequence
Random	TGCACTCCCGTCTCTGCTCT	TGCCTCCCTCACTGGAGCTT
Nanog Promoter	ACAATGTCCATGGTGGACCC	ACCCTACCCACCCCCTATTC
Peak 21	ATTGCTGGTCGGAGCGTAAA	ACGTCAGTCATGTCACGTCC
Peak 85	CCGCATGACGTCTGTACTCA	TAGGCTCCTGTATTTCCGGT
Peak 125	CCGATCTTCGGGGCAAAAAT	ACTCCAAAAACCGCAACCCT

Table 3. List of the primers used in this study. *Target regions and sequences appear as indicated.*

Results

1. MEDIATOR COMPLEX AND RNA POL II INTERACTION

1.1. Mediator Complex and RNA Polymerase II interaction

Mediator and RNA Pol II are large complexes, and to our knowledge, their proximity and direct interaction has not been studied using Proximity Ligation Assay (PLA). In order to directly observe and quantify the Mediator-RNA Pol II interaction in the nuclei of PSCs using PLA, we first needed to decide which protein subunits t subunits could successfully be used as surrogates to track these complexes, to identify specific antibodies for these subunits, and to confirm that these antibodies could also function in a PLA assay. Mediator consists of 30 subunits, while RNA Pol II consists of 12 subunits (Hahn, 2004; Verger et al., 2019). Previously, the localization of Mediator was tracked using one of its largest subunits, MED1, while the localization of RNA Pol II was tracked by its largest subunit, POLR2A (Cho et al., 2018; Y. E. Guo et al., 2019a; Sabari et al., 2018). Therefore, we tested antibodies for MED1 and POLR2A.

For MED1, we identified an antibody which stained one band of the expected size by Western blotting (Figure 17), and which detected a signal that was highly enriched in the nucleus as shown by nuclear/cytoplasmic fractionation.

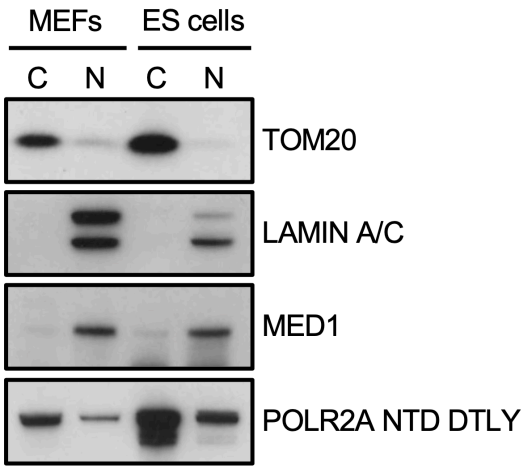


Figure 17. *Western Blot analysis of the indicated targets in both nuclear and cytoplasmic fractions of ES cells and MEFs. To confirm successful fractionation, we used TOM20 as a marker for the cytoplasmic fraction and LAMIN A/C as a marker for the nuclear fraction.*

We were also able to locate it in the nucleus by immunofluorescence (Figure 18). These data were consistent with the known size of MED1, and the expected location of the Mediator complex. Taken together, the data indicate that this antibody is specific for MED1 and could be used to track the position of the Mediator complex in the nucleus as previously shown (Sabari et al., 2018).

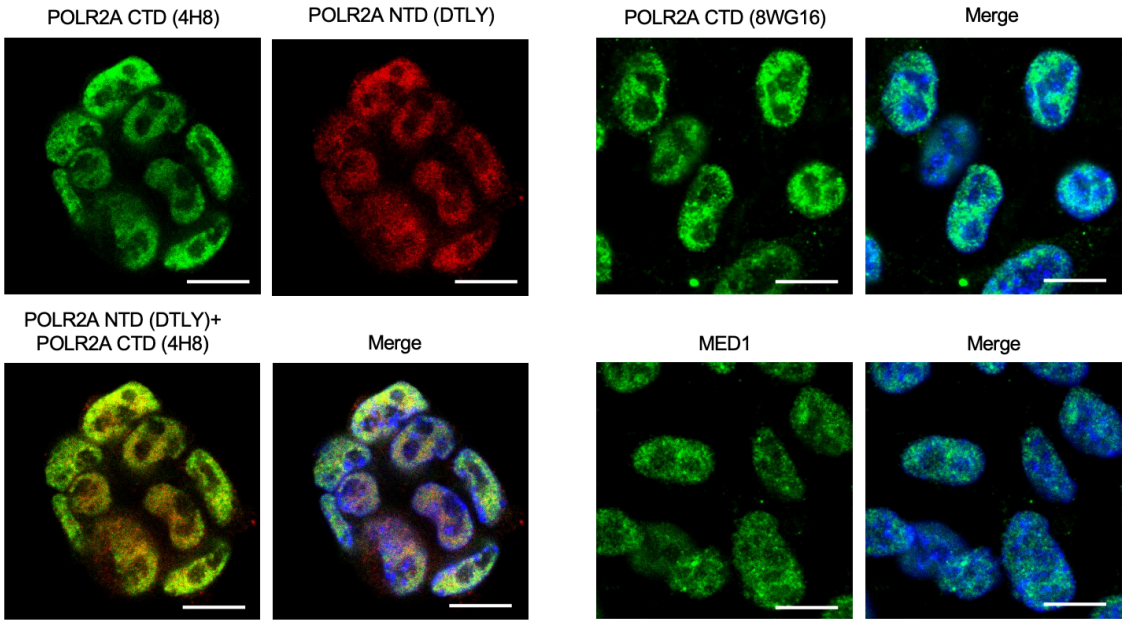


Figure 18. *Immunofluorescence for the indicated markers in mouse ES cells. Scale bar represents 10 µm.*

For POLR2A, we tested three antibodies which have been used to monitor Pol II for many years. Two antibodies were specific for POLR2A CTD: 4H8 and 8WG16. It is important to mention that POLR2A CTD 4H8 antibody specifically binds to POLR2A Ser5P CTD (phospho-RNA Pol II, engaged in active transcription), while POLR2A 8WG16 antibody binds to POLR2A CTD independently of its phosphorylation status (identifying all forms: total-Pol II). In addition, a third antibody specific for POLR2A amino terminal domain (NTD): POLR2A NTD DTLY was tested (also marking the location of total RNA Pol II). While all three antibodies against POLR2A showed the expected staining pattern in the nuclear fraction (Figure 17), we selected the two targeting POLR2A CTD (4H8 and 8WG16) as these had a cleaner and stronger signal, for further application in the PLA assay.

We next tested the selected antibodies for our target proteins (MED1 and POLR2A), to test their ability to track their parent complex (Mediator or RNA Pol II, respectively) by PLA assay. Although Mediator and RNA Pol II are known to interact directly (Cho et al., 2018; Y. E. Guo et al., 2019b), it was not known if MED1 and POLR2A would lie within the known limit of the PLA assay of 40 nm (Bagchi et al., 2015), or if these antibodies would function in the PLA conditions. Nevertheless, we were able to detect MED1-POLR2A co-localization signals by PLA with both 4H8 and 8WG16 antibodies against the phosphorylated or total CTD of POLR2A respectively (Figure 19). This was detected in euchromatic regions of the nucleus (DAPI-low), as expected. Importantly, we also performed negative controls for the PLA assay. As previously noted in the introduction, the Pol II complex is ~10 nm in diameter, the Mediator complex is ~15-20 nm in diameter (Hahn, 2004; Tsai et al., 2017; Verger et al., 2019), and the interaction-distance limit of the PLA assay is ~40 nm (Bagchi et al., 2015). Therefore, taken together, these data provide strong evidence that MED1 and POLR2A lie in close proximity if not direct contact in nuclei, and that it is possible to use the PLA assay to selectively observe Mediator-RNA Pol II complexes, via coincidence of their MED1-POLR2A sub-units within the PLA proximity limit of ~40 nm.

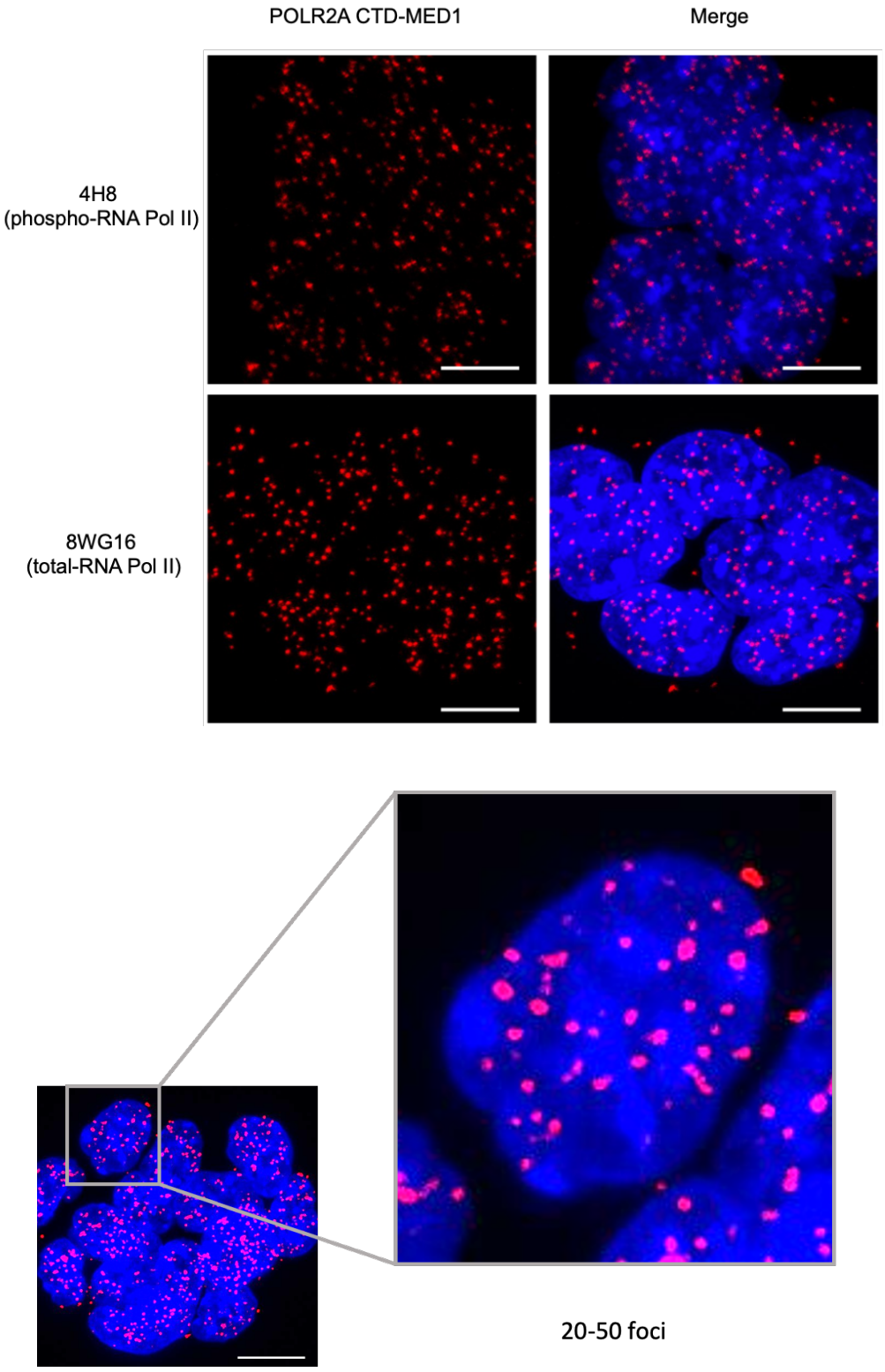


Figure 19. *Proximity Ligation Assay (PLA) in ES cells. Top: MED1- POLR2A CTD 4H8 (phospho-Pol II). Bottom: MED1- POLR2A CTD 8WG16 (total- Pol II). Scale bar represents 10 µm.Below: zoom image showing the number of foci per nucleus.*

Mediator and RNA Pol II interact at thousands of enhancer-promoter loops at any one moment in one nucleus. However crucially, not all enhancers are equivalent in size. Super-enhancers are the largest ~1% of enhancers, and they contain the largest concentration of Mediator and RNA Pol II, in order to drive very high expression of the master transcription factors and cell identity genes (Hnisz et al., 2013; Whyte et al., 2013). In agreement with this well-established model, a recent confocal study using fluorescently-labelled MED1 and POLR2A in live ES cells, reported the first visual observation of the Mediator-RNA Pol II interaction, with a relatively small number of large and stable foci (10-50 per nucleus) (Figure 19). In fact, the authors found that each large Mediator-RNA Pol II focus was co-localized with a super-enhancer, hence their observation of 10-50 foci per nucleus at any one moment in living cells (Cho et al., 2018; Guo et al., 2019). In our experiment, we found that using the PLA assay, we also detected 10-50 foci per nucleus (Figure 19). Based on this, we speculate that our PLA assay is capturing a similar set of the largest Mediator-RNA Pol II interactions, which most likely represent the super-enhancers active at the moment of cell fixation.

1.1.1. Mediator Complex and RNA Polymerase II interaction in naïve vs primed mESCs

When active, Mediator recruits RNA Pol II to promoters (Malik & Roeder, 2016). Therefore, we hypothesized that the number of Mediator-RNA Pol II interaction foci detected per nucleus could be used to infer the level of activity of Mediator.

We next quantified the number of interactions per nuclei, in order to detect possible differences between the number, or magnitude, of interactions in ES cells in different states of cellular identity. We did a first experiment regarding our second objective in which we compared primed ES cells versus 2i-induced naïve ES cells. We were able to quantify the number of interactions per nucleus between RNA Pol II and the Mediator complex, and to compare the two ES cell identity states. Interestingly, we found a significant increase of the number of foci/nuclei in the 2i-naïve state compared to primed state ES cells (Figure 20). Firstly, these results indicate that PLA can detect and quantify changes in Mediator and RNA Pol II interactions. Secondly, the data suggests that ES cells in the naïve state may display significantly more Mediator-RNA Pol II interactions

than cells in the primed state. This conclusion is in agreement with a reported decrease in RNA Pol II-Mediator interactions observed using fluorescently labelled proteins, during ES cell differentiation from the 2i-naïve state passing through to the primed state (Cho et al., 2018).

We then aimed to assess whether small molecule CDK8/19 inhibition (CDK8i) can also promote the naïve state by a similar mechanism to 2i, that is, by changing the activity of Mediator (Lynch et al., 2020). Therefore, we next quantified the Mediator-RNA Pol II interaction in ES cells in their primed and naïve states (where 2i treatment instills the naïve identity) and compared the effect of CDK8i treatment, using PLA as described above.

Given that the POLR2A CTD regions undergoes complex patterns of phosphorylation during Pol II translocation along genes (Hsin & Manley, 2012), we chose an antibody for the PLA assay which would detect the POLR2A CTD independently of its phosphorylation status. Thus, we decided to use the POLR2A CTD 8WG16 (total-RNA Pol II) antibody in the subsequent experiments.

By combining PLA with the POLR2A CTD 8WG16 (total-RNA Pol II) antibody and image quantification as above, we investigated the number of Mediator-RNA Pol II interactions per nucleus in ES cells under CDK8i treatment. We again detected a significant increase in the number of MED1-POLR2A proximity foci/nucleus compared to primed ES cells Lastly, we did not find significant differences between 2i-naïve ES cells and CDK8i treated ES cells regarding the frequency of MED1-POLR2A proximity foci (Figure 20). This data suggested that the parent complexes Mediator and Pol II interact with higher frequency in ES cells in the naïve state (induced with 2i), and in CDK8i treated ES cells, compared to primed state ES cells.

In summary, PLA results suggest that both 2i and CDK8i promote the association of Mediator and RNA Pol II, confirming our initial hypothesis that 2i and CDK8i may influence the core transcriptional machinery in a similar manner.

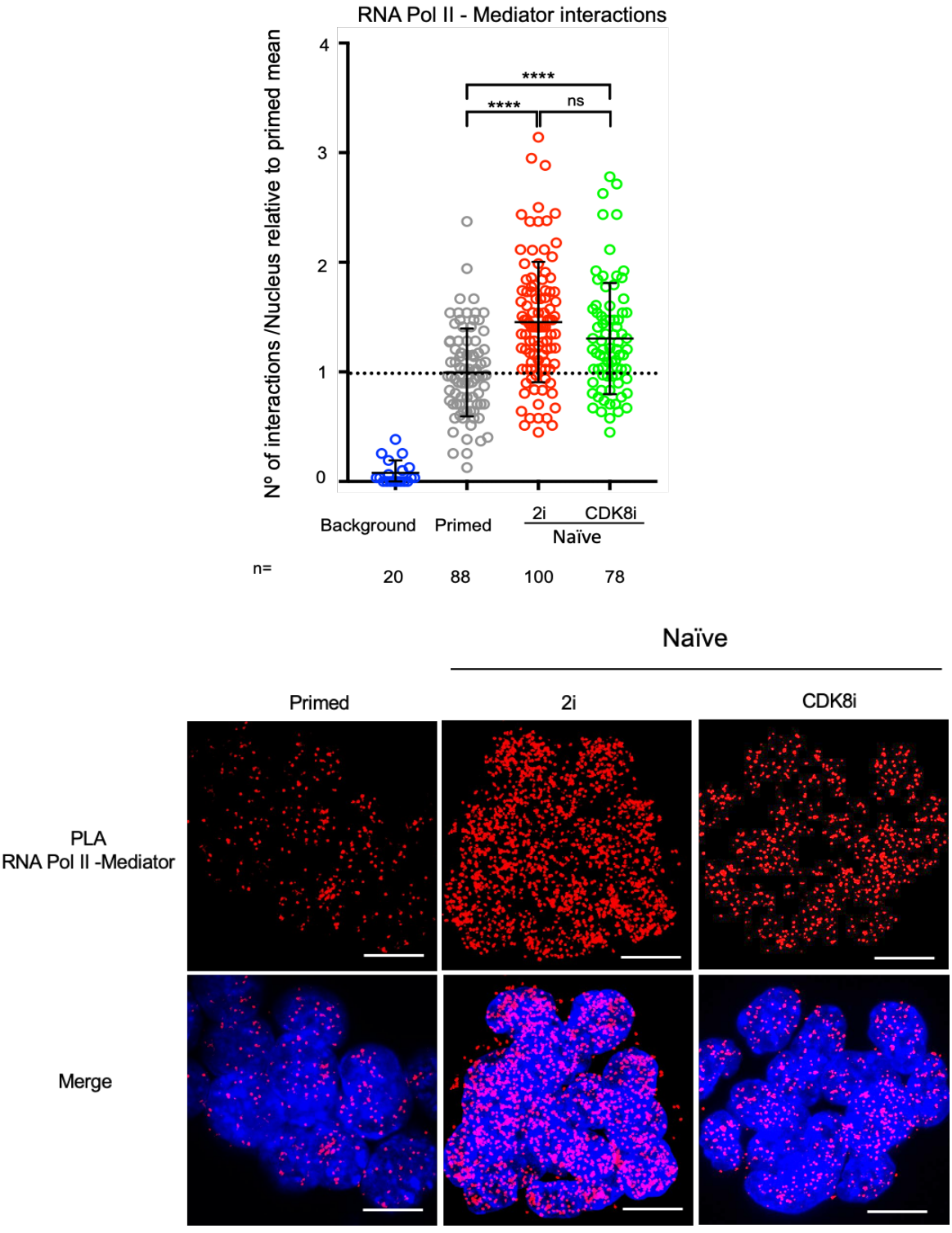


Figure 20. *Above: Quantification of the number of MED1-POLR2A CTD 8WG16 (total-RNA Pol II) interactions detected by PLA in mouse in primed condition (n=88 nuclei), 2i-naïve cells (n=100 nuclei), CDK8/19i-naïve ES cells (n=78 nuclei) and background/negative control (n=20 nuclei) as indicated. Below: representative microscopy images of the assay.*

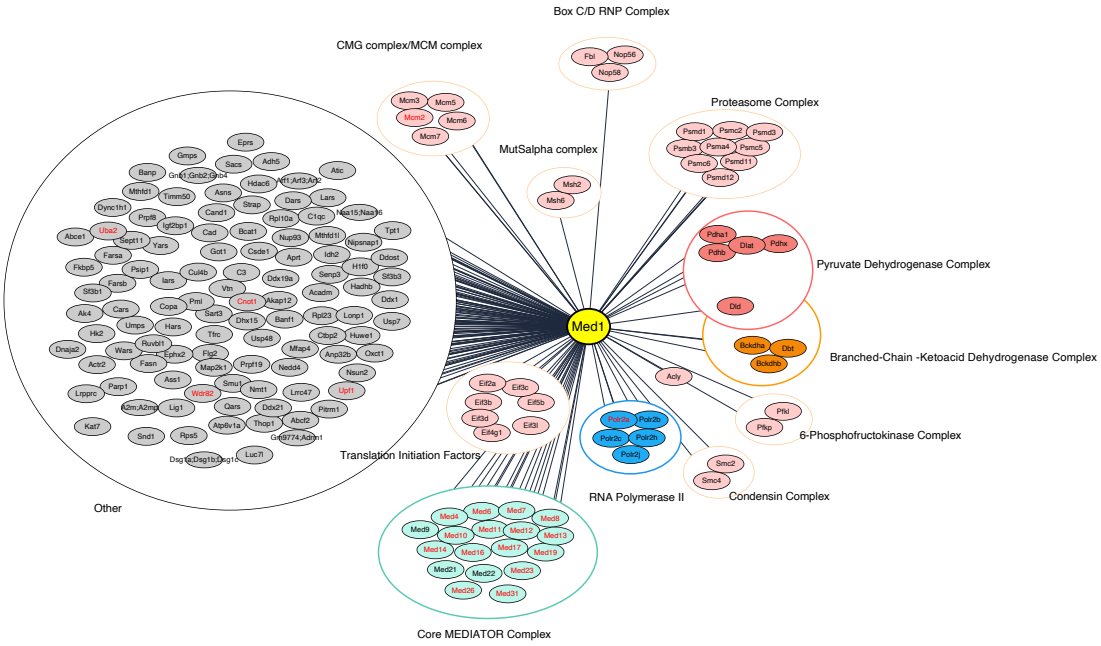
2. MEDIATOR COMPLEX INTERACTORS

2.1. Mediator Complex MED1 interactome: IP-MS identification of novel Mediator putative partners

As a central hub in the transcriptional machinery, the Mediator complex interacts with many critical factors and other complexes. To gain an overview of Mediator's interactome, we immunoprecipitated the Mediator complex from nuclear chromatin lysates and performed mass spectrometry to identify all accessory proteins.

With this aim, we choose Mediator´s biggest protein subunit: MED1 and performed Immunoprecipitation followed by Mass Spectrometry (IP-MS) analysis with the aim of further understanding Mediator dynamics.

Our MED1 IP-MS performed in mESCs showed high peptide counts for MED1.
the highest peptide counts for the target MED1, suggesting a successful immunoprecipitation. We performed our MED1 IP-MS in mESCs grown in control or primed state cells, versus 2i-naïve and CDK8i-naïve states, and we discarded reported common contaminants from the analysis. Also, we only considered as specific MED1 interactors, proteins that were represented with FC<2 compared to our IgG control. We observed a strong overlap in many of the proteins and biological functions of the Mediator interactome with transcription as expected, but also between yeast and our mouse cells (Quevedo et al., 2019; Uthe et al., 2017a). We short-listed previously reported core and conserved mediator interactors, which appear written in red in table 2. As expected Mediator interacted with many elements of the transcriptional process in the nucleus, including RNA Pol II itself, cohesin (Smc2/4) which Mediator associates with to form enhancer-promoter bridges, and 5 subunits of the MCM helicase complex (MCM2/3/5/6/7), which is typically associated with DNA unwinding during replication, but which has also been reported as part of the core RNA Pol II transcriptional machinery (Yankulov et al., 1999), where it assists in separating transcriptional elongation and replication (Liu et al., 2021). Altogether, these data further suggested a successful immunoprecipitation of the Mediator complex and its protein interactome.



GO Component	Gene name
Core MEDIATOR complex	Med4; Med14; Med26; Med21; Med9; Med31; Med12; Med13; Med7; Med10; Med17; Med22; Med11; Med19; Med8; Med23; Med16; Med6
Core RNA Pol II	Polr2a; Polr2b; Polr2c; Polr2h; Polr2j
Pyruvate Dehydrogenase complex	Dlat; Pdhb; Pdha1; Pdhx; Dld
Branched-chain α-ketoacid dehydrogenase complex	Dbt; Bckdha; Bckdhb
6-phosphofructokinase complex	Pfkp; Pfkl
Phenylalanine-tRNA ligase complex	Farsa; Farsb
CMG complex/MCM complex	Mcm3; Mcm5; Mcm7; Mcm2; Mcm6
MutSalpha complex	Msh6; Msh2
Box C/D RNP complex	Fbl; Nop56; Nop58
Eukaryotic translation initiation factor	Eif3c; Eif3b; Eif3l; Eif2a; Eif3d; Eif5b; Eif4g1
Proteasome complex	Psmc5; Psmd11; Psmc2; Psmc6; Psmd12; Psmd3; Psma4; Psmd1; Psmb3
Condensin complex	Smc2; Smc4
Other	Mthfd1; Rpl10a; Rps5; Asns; Ddx19a; Ddx21; Fasn; Eprs; Dars; Ass1; Nedd4; Got1; Igf2bp1; Nsun2; Tpt1; Hk2; Mfap4; Banp; Parp1; Atic; Psip1; Cad; Rpl23; Idh2; Usp7; Lars; Acly; Prpf19; Cand1; Dhx15; Snd1; Ddx1; Abcf2; Tfrc; Ruvbl1; Acadm; Qars; C3; Sf3b1; Adh5; Gnb1; Gnb2; Gnb4; Iars; Akap12; Dync1h1; Lig1; C1qc; Abce1; Arf1; Arf3; Arf2; Dnaja2; Anp32b; Hadhb; Banf1; Sf3b3; Nmt1; Prpf8; Ctbp2; Naa15; Naa16; Sart3; Gmps; Umps; Nup93; Ddost; Ak4; Uba2; Aprt; Thop1; Luc7l; Sacs; Atp6v1a; Copa; Upf1; Cars; Lrpprc; Cul4b; Smu1; Csde1; Map2k1; Senp3; Dsg1a; Dsg1b; Dsg1c; Pml; Strap; Gm9774; Adrm1; Sept11; Nipsnap1; Actr2; Lrrc47; A2m; A2mp; Bcat1; Mthfd1l; Yars; Ephx2; Wars; Lonp1; Kat7; Pitrm1; Oxct1; H1f0; Fkbp5; Vtn; Cnot1; Huwe1; Flg2; Hars; Hdac6; Wdr82; Usp48; Timm50

Figure 21. *Top: Mediator complex interactome derived from Med1 IP-MS.Bottom: List of names from the top enriched proteins FC <-2 in our interactome. Highlighted in red appear the previously reported interactors.*

Having validated that our MED1 IP-MS successfully represented in terms of peptides detected the previous knowledge about Mediator interactors network, we next focused our attention on some of the novel/unexpected interactors detected.

Regarding novel interactors, we noted a particularly strong interaction with several subunits of the PDH complex and the Branched Chain Ketoacid Dehydrogenase (BCKDH) (highlighted in pink in Table 2). This led us to elaborate an updated interaction network of Mediator with novel and potentially important proteins connected with Mediator complex (Figure 21).

In summary, the IP-MS of the Mediator complex and its interactome revealed the relative extent to which Mediator associates with accessory proteins in the enhancer-driven control of the transcriptional program. Overall, the Mediator interactome contained a large number of interactors and highly associated with proteins related to enhancers, chromatin, and transcriptional activity, consistent with a central role for Mediator in this process.

In addition, this overview suggested a number of surprising aspects to Mediator function. It is notable that while elements of the RNA Polymerase II complex were identified, their relative extent of detection within the Mediator interactome suggests that while the Mediator-RNA Pol II interaction is critical, Mediator may interact transiently with RNA Pol II, and also spend significant time associated with other factors. In particular, the high detection of elements of the PDH complex was striking and unexpected, leading us to further investigate this novel feature of Mediator occupancy.

3. MEDIATOR COMPLEX CONEXION TO THE PYRUVATE DEHYDROGENASE COMPLEX

3.1. Nuclear pyruvate dehydrogenase complex

From the list of novel Mediator interacting candidates, the presence and abundance in peptide count of different subunits of the Pyruvate Dehydrogenase Complex (PDH) specially called our attention (Figure 21).
Interestlingly, the PDH complex shares it´s E3 subunit (DLD) with another dehydrogenase: the BCKDH complex. The rest of BCKDH subunits were also detected by Mediator IP-MS.

Given the growing interest in the nuclear role of PDH in the maintenance of chromatin architecture and acetylation we decided to put our efforts in validating and further understanding the interplay between MED1 and PDH.
We first checked by microscopy the subcellular localization of two subunits from the PDH complex: PDHA1 and DLAT.

For E1 component Pyruvate Dehydrogenase E1 Subunit Alpha 1(PDHA1), we obtained a clean staining, showing that while the majority of PDHA1 was found in the cytoplasm of our samples, within mitochondria and co-localizing with the mitochondrial marker TOM20, there was some PDHA1 in the nuclear compartment of both Mouse Embryonic Fibroblasts (MEFs) and ES cells. We noticed that the staining of PDHA1 was similar to the MED1 pattern, observing a coincidence in the location of MED1 (which is nuclear as expected) and PDHA1. This nuclear PDHA1 was more abundant in ES cells compared to MEFs (Figure 23).

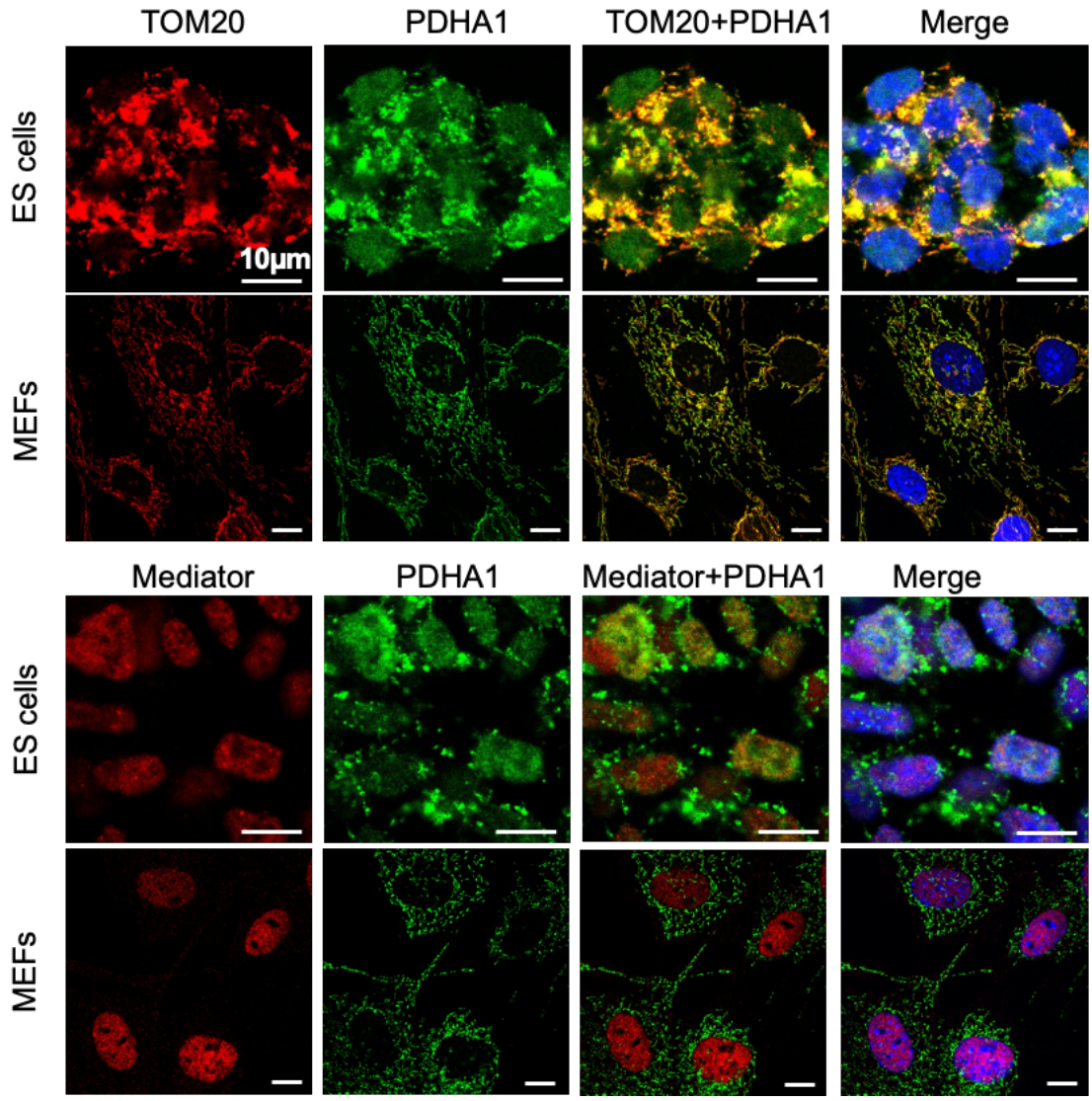


Figure 22. *Immunofluorescences for the indicated markers in mouse ES cells and MEFs. Scale bar represents 10 μm.*

For E2 component Dihydrolipoamide S-Acetyltransferase (DLAT) our imaging results again revealed a mostly cytoplasmic localization, consistent with its expected location in mitochondria, but also some presence of the enzyme in the nucleus. Again, while the majority of DLAT protein was detected in the cytoplasm with a mitochondrial distribution pattern, we also found DLAT in the nucleus of both ES cells and MEFs (Figure 23).

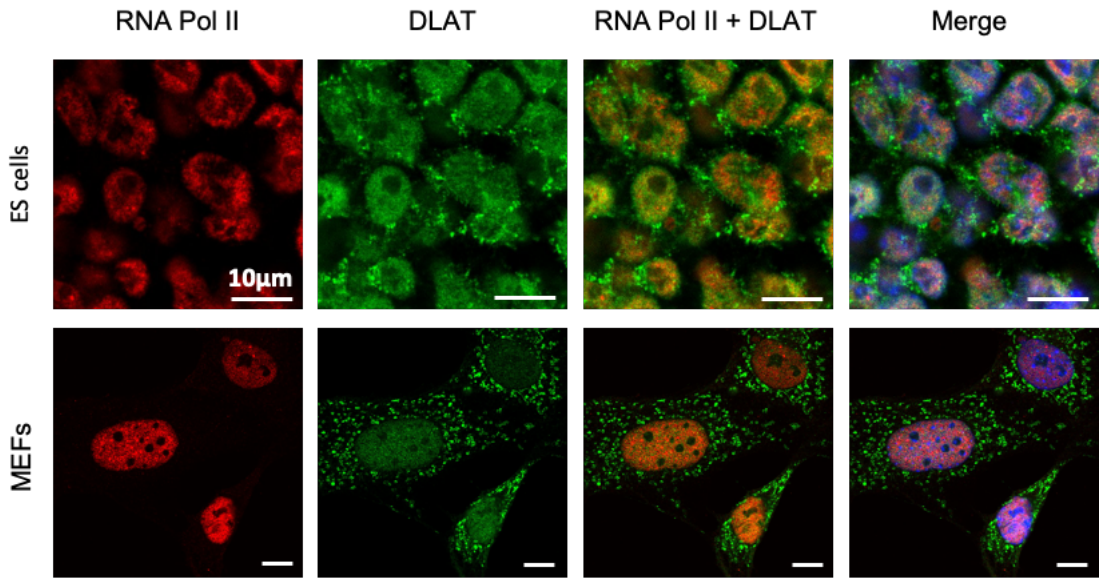


Figure 23. *Immunofluorescences for the indicated markers in mouse ES cells and MEFs. Scale bar represents 10 µm.*

We confirmed the immunostainings by western blotting the nuclear and cytosolic fractions of the samples. We were able to detect that the bands for each protein were detected at the expected size and the patterns of subcellular distribution were the same as the ones observed by microscopy (Figure 24).

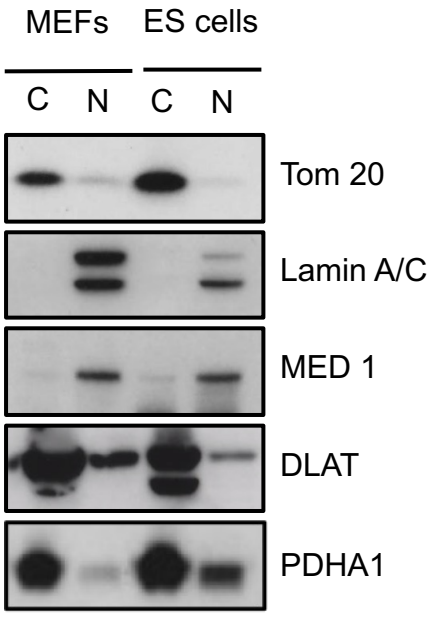


Figure 24. *Western Blot analysis of the indicated targets in both nuclear and cytoplasmic fractions of ES cells and MEFs. To confirm successful fractionation, we used TOM20 as a marker for the cytoplasmic fraction and LAMIN A/C as a marker for the nuclear fraction.*

The DLAT western blot from the fractionated samples showed that while nuclear DLAT had one band, the cytoplasmic fraction (including both mitochondrial and cytoplasmic compartments) had two migration fronts (Figure 24).

We explored in the literature and found that the lipoylated and non-lipoylated forms of a protein migrate at different rates (Tsvetkov et al., 2022), where the lipoylated form is bigger/slower in migration speed. In order to be functional and generate Acetyl-CoA from Pyruvate, PDH needs it´s DLAT subunit to be lipoylated. It is well known and reported that mitochondrial PDH is active and lipoylated, playing an essential role in the TCA cycle. In contrast, nothing is known about the lipoylation status of nuclear PDH.

We therefore decided to further study DLAT state in the 3 main cellular compartments: nucleus mitochondria and cytoplasm. Subcellular fractionation showed that while in the total cell extract there were two forms of DLAT protein, the nuclear and the mitochondrial compartments only had the big/slow version of DLAT (Figure 25).

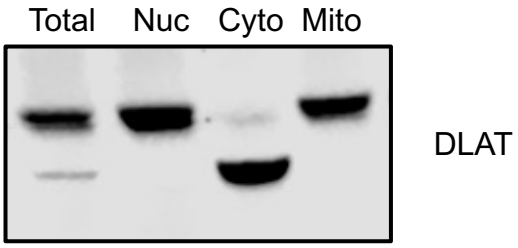


***Figure 25**. Western Blot analysis of DLAT in Total, nuclear, cytoplasmic, and mitochondrial fractions of SKMEL cells*

To confirm that this band corresponded to the lipoylated/active version of DLAT, we performed a specific staining against Lipoic acid, that detects any lipoylated protein in the sample.

To the date, only four mammal multimeric metabolic enzymes have been reported to be lipoylated (Rowland et al., 2018), where the proteins that get lipoylated are: GCSH (19 kDa), DLST (49 kDa), DBT (53 kDa), PDHX (54 kDa) and DLAT (70 kDa). Given that the staining for lipoic acid that we see in the fractionation is above PDHA1 (48 kDa) and in the expected size of the slow band of DLAT, the results pointed to an active form of nuclear DLAT (Figure 26).

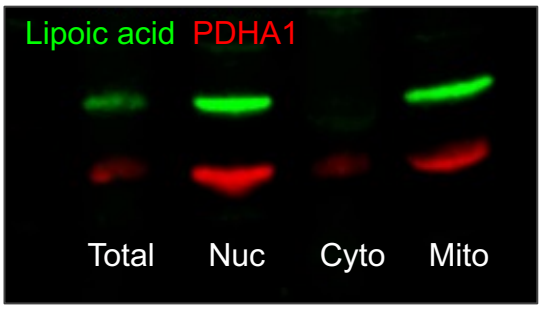


***Figure 26**. Western Blot analysis of the indicated targets in Total, nuclear, cytoplasmic, and mitochondrial fractions of SKMEL cells*

To confirm this result, we used Small Interfering RNA (siRNA) technology to generate cells that did not express PDHA1 and DLAT subunits of PDH and repeated the lipoic acid western blots (Figure 27). Confirming that nuclear DLAT is lipoylated and after it´s knockdown, the band of lipoic acid disappears.

We could also see that interfering with the mitochondrial pyruvate carrier (MPC) did not affect the expression of either PDHA1 or DLAT, nor did it affect the lipoylation pattern.

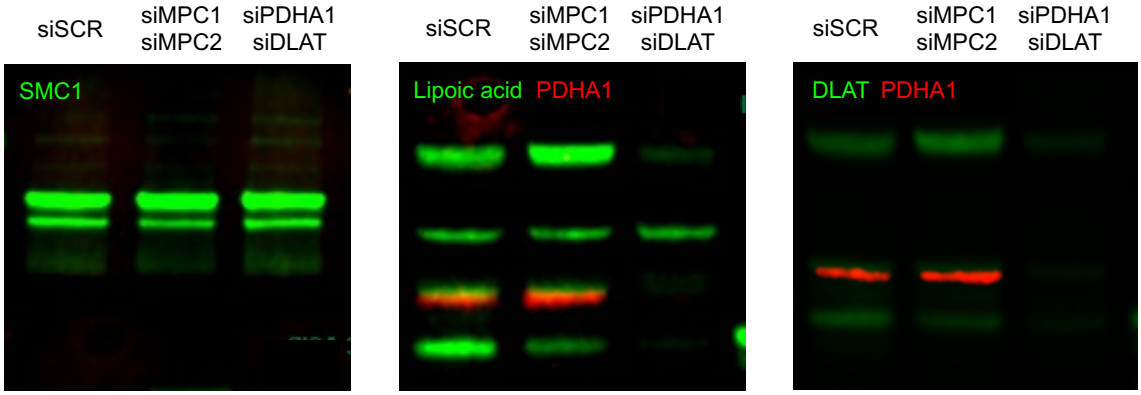


***Figure 27 .** Western Blot analysis of the indicated targets in Total, nuclear, cytoplasmic, and mitochondrial fractions of mES cells in control (siSCR), depleted MPC (siMPC1 + siMPC2) or depleted PDH complex (siPDHA1 + siDLAT) as appear in the figure.*

Altogether, this data strongly suggests the presence of an active form of the PDH in the nucleus.

3.2. The nuclear pyruvate dehydrogenase complex as part of core transcriptional machinery

Having confirmed the presence of an active nuclear PDH, we wanted to confirm that, as predicted from Mediator IP-MS, the nuclear form of PDH was tightly related to the transcriptional machinery. To assess this question, we performed PLA.

As described before in this Thesis, because of the inherent distance limit of PLA (Figure 8), this assay can detect proximity between proteins that are located in a range of 40 nm inside the studied cell or tissue.
Interestingly, we were also capable of confirming that, as predictable from our immunostainings (Figure 23), a high abundance of the DLAT subunit from the PDH complex tightly interacted with POLR2A subunit of RNA Pol II (Figure 28).

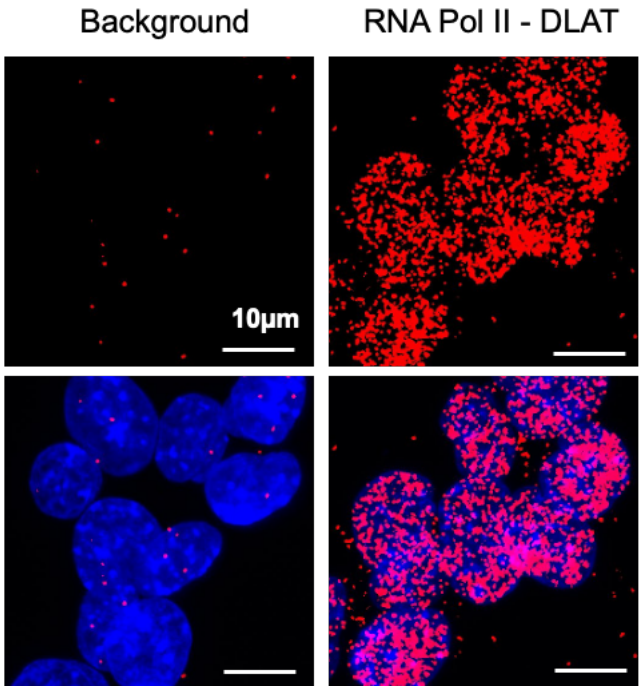


Figure 28. *POLR2A CTD 8WG16 (total-RNA Pol II)- DLAT Proximity Ligation Assay (PLA), in mouse ES cells. Scale bar represents 10 µm.*

In conclusion, we demonstrated that the nuclear distribution of the PDH complex is not random, but tightly associated with the core transcriptional machinery, where it exists in an active form (lipoylated-DLAT in the PDH complex).

3.3. Pyruvate dehydrogenase complex association with chromatin: DLAT ChIP-seq

In order to more deeply understand the localization of the PDH enzymatic complex with specific chromatin regions, we performed Chromatin Immunoprecipitation combined with DNA Sequencing (ChIP-Seq). We focused on ChIP of the DLAT subunit as a surrogate of the remainder of the PDH complex given the central structural position of DLAT, and its critical catalytic role in the PDH complex for the formation of acetyl-groups for histone acetylation (Figure 12), together with the practical advantage of a good antibody for DLAT which we had validated above. In addition, we chose to perform DLAT ChIPseq in mouse ES cells given their physiological relevance of this cell -type, and the availability of extensive datasets for chromatin-associated proteins for mouse ES cells.

ChIP for DLAT was performed under homeostatic/basal ES cell culture conditions. Following standard ChIP alignment and peak calling, we identified a total of 32,053 peaks in mouse ES cells. As a first check, we validated randomly chosen DLAT ChIPseq peaks, a peak located at Nanog promoter and a random region of the chromatin by ChIP-qPCR. Importantly, we could see that silencing DLAT subunit from PDH (siPDH) in ES cells resulted in the disappearance of DLAT ChIP-signal, confirming the ChIPseq dataset by a second approach (Figure 29).

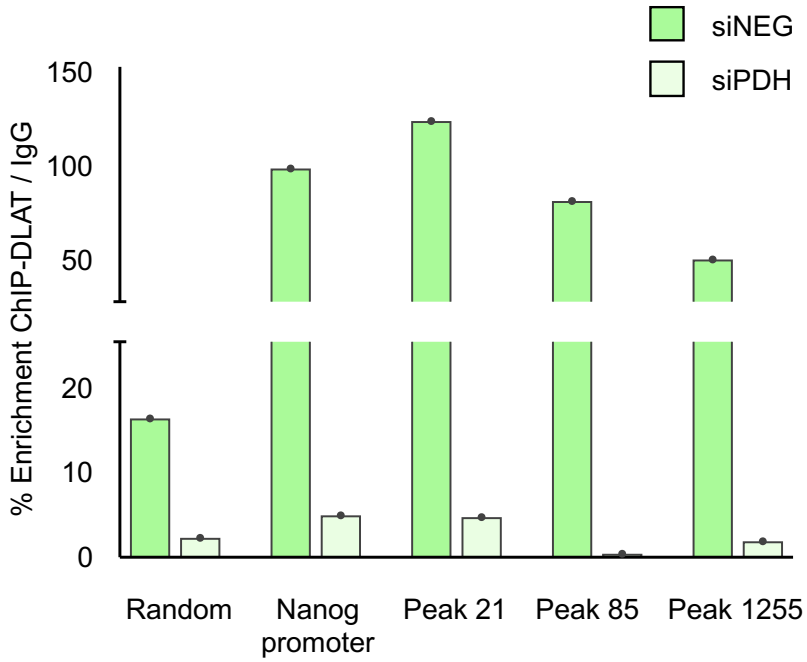


Figure 29. *Quantitative real-time PCR for the confirmation of ChIP-seq peaks % enrichment of PCR products from DLAT ChIP compared to IgG ChIP are shown as bar graphs. Dark green corresponds to control siNEG mES cells and light green corresponds to mES cells with depleted PDH (siDLAT).*

ChromHMM analysis uses published datasets of transcriptional activity, histone marks, and chromatin accessibility, to interpret and divide the genome into functional units (Bogu et al., 2016; Pintacuda et al., 2017). For a general insight into the role of DLAT on chromatin, DLAT-associated chromatin loci were annotated and the genomic distribution of DLAT-binding was assessed according to ChromHMM-defined regions. DLAT-chromatin loci were observed within euchromatic regions of actively transcribed chromatin together with more distal intergenic regions (Figure 30). Importantly, this observed genomic pattern of DLAT/PDH complex was consistent with the previously reported localization of the Mediator complex at enhancers and promoters (Hnisz et al., 2013; Whyte et al., 2013).

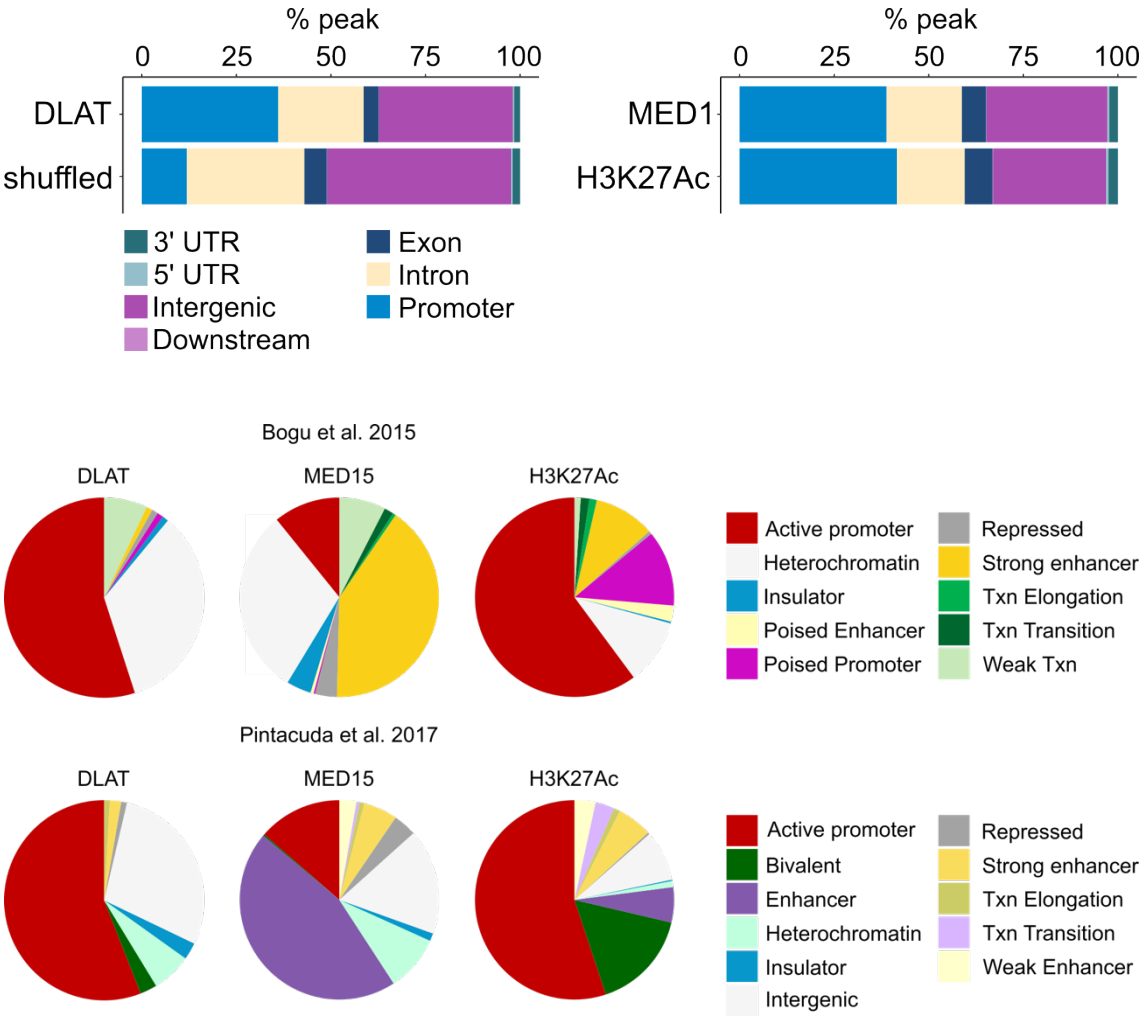


***Figure 30**. Top left: Left graph: Proportion of DLAT peaks overlapping genomic features. Top right: For comparison, annotation of MED1* (L. Sun et al., 2021)*) and H3K27Ac* (Zhao et al., 2021) *peaks over genomic features. Bottom: Pie charts showing the proportion of peaks for each ChIP-Seq that overlap with ChromHMM states in mESc defined by* (Bogu et al., 2016; Pintacuda et al., 2017).*In both cases, DLAT peaks accumulate at transcriptionally active promoters. In comparison, MED15 peaks mostly accumulate at enhancer regions.*

To understand how DLAT associates with Mediator and Pol II throughout the genome, and whether there are 'flavours' of Mediator that do not include DLAT, we used ChromHMM to bin the genome into 8 states according to their enrichment for DLAT, Pol II (L. Sun et al., 2021) MED1 (L. Sun et al., 2021), MED4 (L. Sun et al., 2021), H3K27Ac (Zhao et al., 2021), and H3K4me1(Zhao et al., 2021). The choice of 8 states is the result of

selecting the number of genomic partitions that is most informative. Overall, we can conclude that the strongest binding of DLAT to the chromatin occurs in positions bound

by Mediator. DLAT mostly interacts with Pol II in the presence of Mediator. And there are no genomic fragments bound by Mediator where DLAT is not present. (Figure 31).

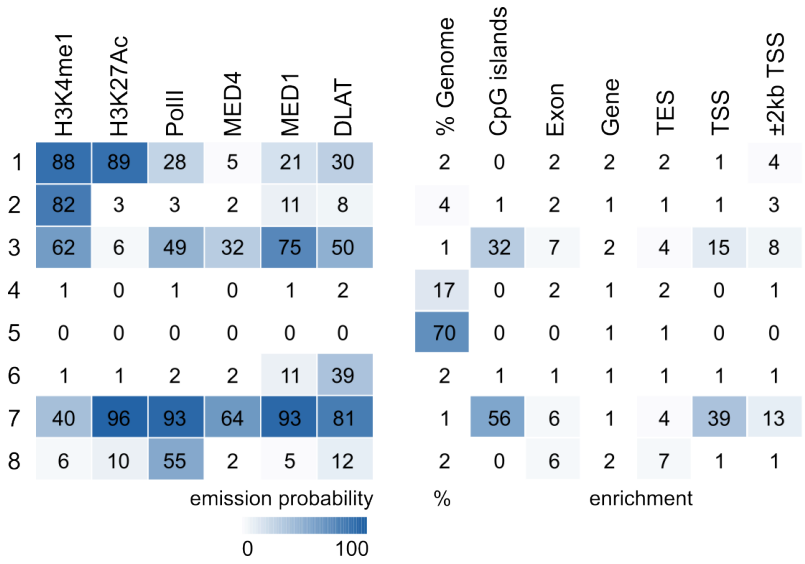


***Figure 31**. ChromHMM 8 states of the genome according to their enrichment for DLAT, Pol II (Sun et al., 2021) MED1 (Sun et al., 2021), MED4 (Sun et al., 2021), H3K27Ac (Zhao et al., 2021), and H3K4me1(Zhao et al., 2021).*

More broadly, we observed significant genomic coincidence of DLAT with regions of histone acetylation marks, all of which are associated with enhancers and active transcription, specifically H3K9ac, H3K27ac, H3K18ac, H3K122ac, H4K5ac, H4K8ac, and H4K16ac (Figure 32). Of note, DLAT associated more closely with enhancers compared to promoters and transcription start sites (Figure 32).

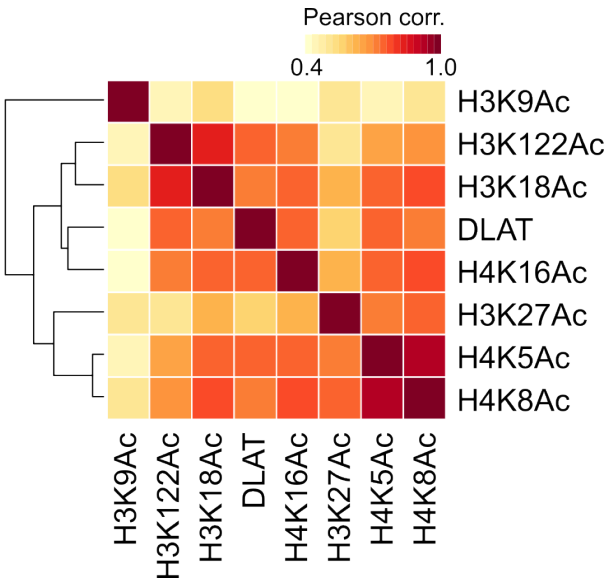


***Figure 32**. Pearson correlation heatmap of ChIP signal intensity across the genome for the indicated histone post-translational acetylation modifications with each other and versus DLAT. DLAT enriches with several forms of acetylated histones in chromatin.*

We observed that 97.83% of superenhancers (SE) and 44.47% of typical enhancers (TE) contain a DLAT peaks (Figure 33). To determine if this degree of overlap is above what is expected by chance, we shuffled the lists of SE and TE 10 times and computed the overlap with DLAT peaks (represented in grey) this data allowed us to determine that the overlap of DLAT peaks at enhancers and super enhancers is not random but in contrast specific.

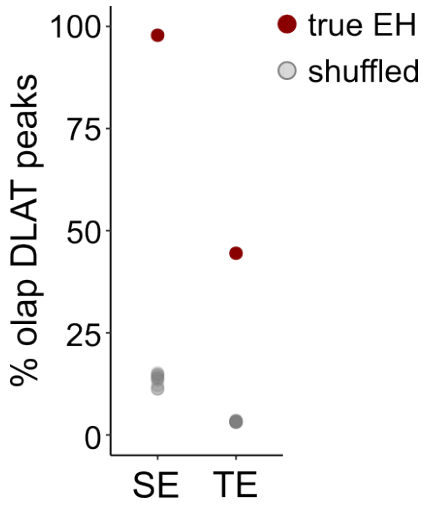


***Figure 33.** Intersection of DLAT peaks with typical enhancers (n = 9,981)* (Hnisz et al., 2013) *and superenhancers (n = 231)* (Whyte et al., 2013).

Moreover, among enhancer loci in mouse ES cells, DLAT associated to a greater extent with the largest enhancers (Super-Enhancers; SEs) versus smaller "typical" enhancers (TEs) (Figures 33-35).

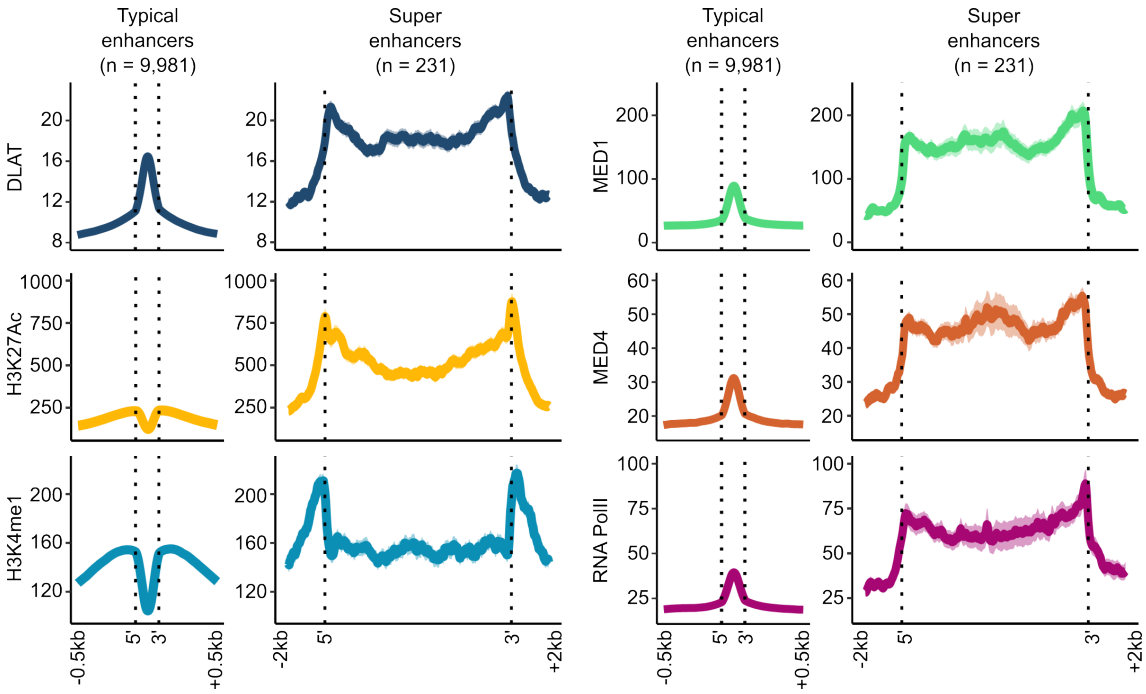


Figure 34. *Metagene average ChIP signal intensity at TEs or SEs for DLAT/PDH complex versus the indicated histone post-translational modifications that mark enhancers (H3K27ac, H3K4me1) and Mediator subunits (MED1, MED4). DLAT/PDH complex enrichment coincides closely with enhancer loci, with the largest enrichments observed within the largest 1% of enhancers -the SEs.*

SEs are known to contain up to 40% of Mediator complex in mouse ES cells, and they drive the expression of the master regulators and transcriptional factors for cell identity (Whyte et al., Hnisz et al., 2013, Cell). Similarly, we observed a majority of DLAT signal within SEs, further suggesting that PDH chromatin localisation mimics the pattern of Mediator at enhancers. In agreement, association-analysis of DLAT-bound chromatin regions and putative cis-regulated target genes revealed a strong enrichment for the control of pluripotency and cell proliferation, two key features of ES cell identity (Figure 35).

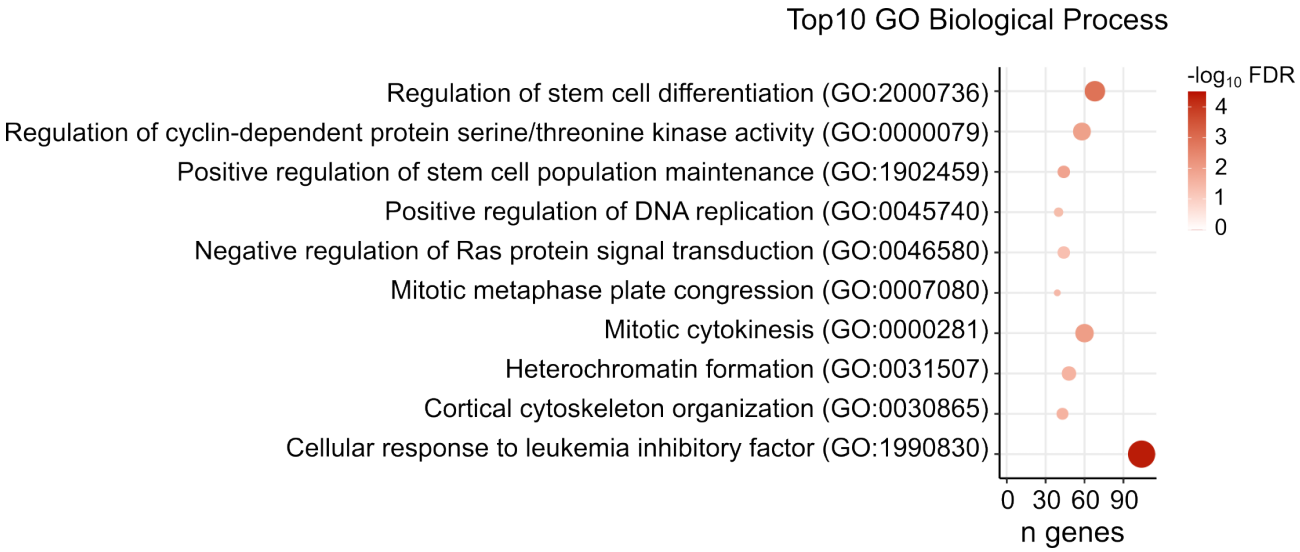


***Figure 35**. Bubble plot depicting the statistical significance of the top 10 most significantly enriched GO terms among DLAT targets, where the size of the points reflects the number of genes in each term, and the statistical significance is encoded by the colour of the points.*

Taken together, ChIPseq of the key subunit of the PDH complex, DLAT, further supports the specific enrichment of the PDH complex within enhancers, and strongly suggests the proximal production of acetyl groups by the PDH complex within enhancers and transcription units, which are major sites of acetyl-group consumption by histone acetylation.

3.4. Pyruvate dehydrogenase complex manipulation

Given the results obtained in our DLAT ChIP-seq experiment we decided to start a preliminary exploration of the effects caused by manipulating the PDH complex by different means in different cell types and models *in vitro*.

Our first approach was to knockdown DLAT, the core catalytic subunit from the PDH with Short Hairpin RNA (shRNA) technology. We performed the knockdown in two different mES cell lines, trying different shRNAs. In subsequent experiments, we chose shDLAT #1 as it knocked down DLAT most efficientlyin the two mESC lines that we tested (Figure 36).

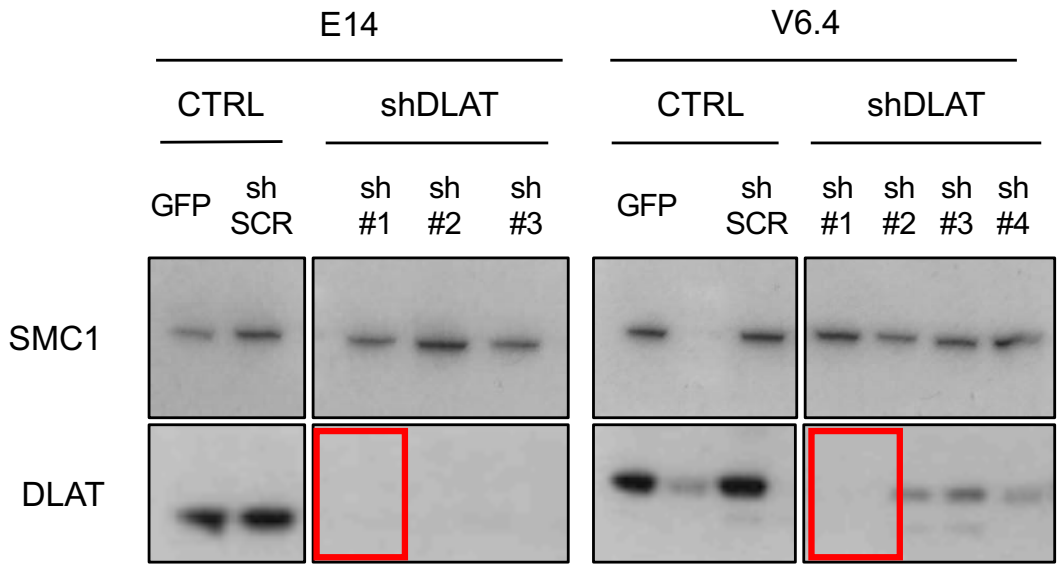


***Figure 36.**. Western Blot analysis of the DLAT knockdown by shRNA in E14 and V6.4 mES cell lines. SMC1 has been used as a housekeeper and sh for Green Fluorescent Protein (shGFP) has been used as a control.*

Upon shRNA knockdown of DLAT, mESCs grown in 2i-naïve state culture conditions did not show great changes, and cells could be maintained in culture. We challenged the system and induced differentiation. We did so by removing LIF and "2i" cocktail from the media and adding retinoic acid. In this culture conditions, ES cells start differentiating in an undirected manner, dramatically rewiring their transcriptional programs and establishing new identities.

After 7 days in differentiating media, we could see that while control Scramble shRNA (shSCR) cells were successfully differentiating, cells with shDLAT- knockdown, not only failed to differentiate, but could not survive the differentiation stimulus (Figure 37).

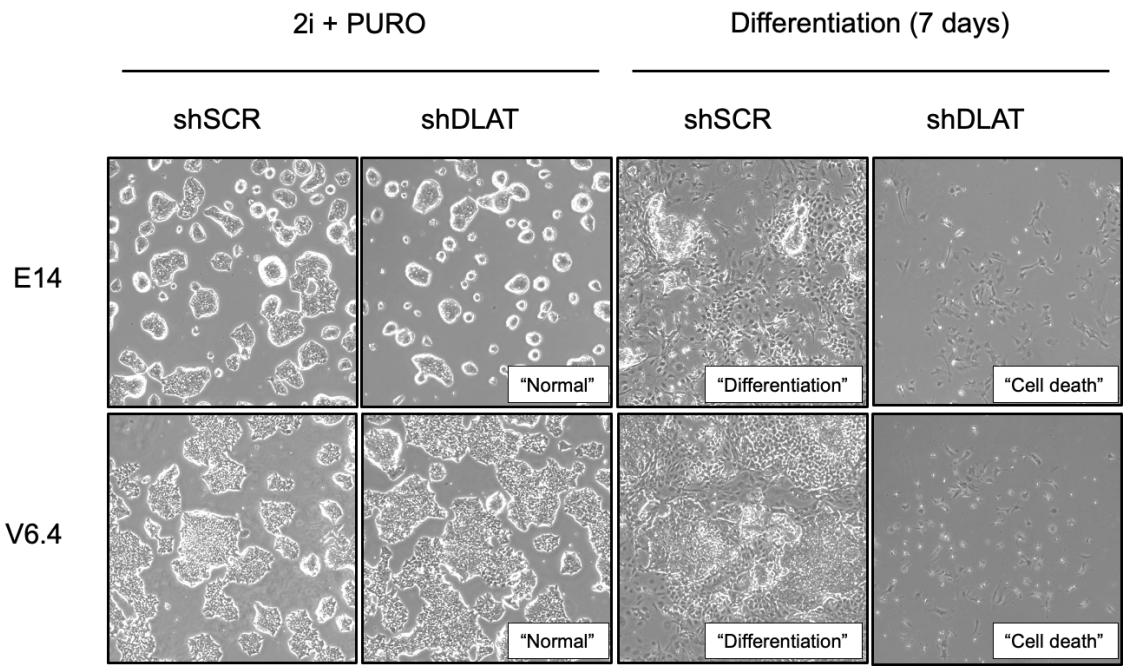


***Figure 37.**. Brightfield microscopy analysis of theeffects of DLAT knockdown by shRNA in E14 and V6.4 mES cell lines. Scramble sh (shSCR) was used as control. Cells were grown as indicated: left: 2i-naïve culture media; right differentiation (no LIF + retinoic acid) media for 7 days.*

Chromatin acetylation is important during periods of homeostasis, however duing cell identity transitons, when chromatin architecture is remodeled and new regions of the chromatin need to be acetylated and opened to stablish new active SE and TE enhancers that will stablish a new identity, we speculate that the efficient supply of Acetyl-CoA proximal to its site of consumption (histone acetylation) may be particularly critical.

As our knockdown depleted the whole pool of cellular PDH, we wanted to explore whether the observed effects were due to the nuclear PDH linked to the transcriptional machinery, the mitochondrial PDH linked to energy production or the combination of both.

To enter to the inner membrane of the mitochondria, pyruvate goes through a specific transport channel, named the mitochondrial pyruvate carrier (Figure 38), there are two proteins that behave as mitochondrial pyruvate carrier: MPC1 and MPC2. These two proteins are the only known manner of pyruvate entering through the mitochondrial inner membrane, and there are two well validated (Linden & Corbet, 2018; Zhong et al., 2015)inhibitors of the MPC: 7ACC2 and UK5099.

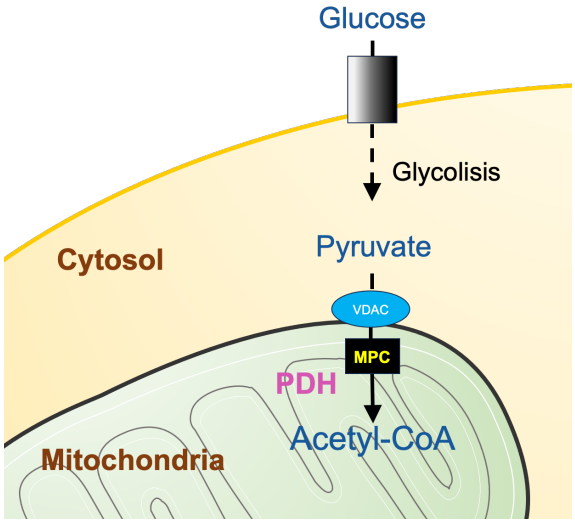


Figure 38. *Schematics of metabolic pathway related to the entrance of Pyruvate to the mitochondria fot acetyl Co-A production by the Pyruvate Dehydrogenase complex. Mitochondrial Pyruvate Carrier (MPC) inhibitors 7ACC2 and UK5099 are shown in red.*

By using those two inhibitors, we were able to inhibit the mitochondrial PDH activity by restricting the availability of its substrate: pyruvate. We repeated the previous experiment with +/- DLAT shRNA-knockdown, and induced mESC differentiation, but in this occasion, we included +/- MPC inhibitors.

We observed that when we only inhibit mitochondrial (and not global) PDH activity, cells are able to successfully differentiate when we remove LIF and add retinoic acid to the culture medium (Figure 39).

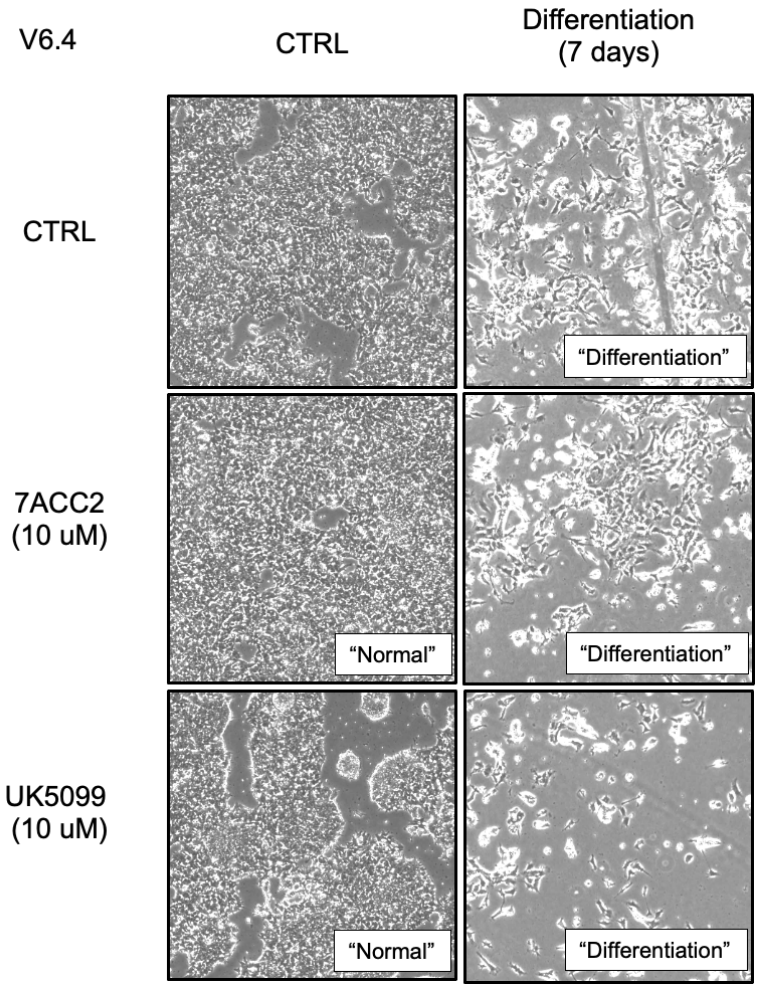


Figure 39. *Brightfield microscopy analysis in mES cells of the effects of Mitochondrial Pyruvate Carrier (MPC) inhibition by either 7ACC2 or UK5099. Cells were grown as indicated: left: control S/L mES cell media; right differentiation (no LIF + retinoic acid) media for 7 days.*

We then tested the effects of silencing DLAT in a different cell type: we used siRNA technology to silence DLAT in B16 melanoma cells. We again used the two previously mentioned MPC inhibitors: MPC:7ACC2 and UK5099. Under the described conditions, we observed that the effects of DLAT silencing in B16 melanoma cell viability were strong. After 5 days of siDLAT knockdown, B16 cells were mostly dead. In contrast, when we added either 7ACC2 or UK5099 inhibitors without siDLAT knockdown, B16 melanoma cells did not show any defects in either proliferation or viability (Figure 40).

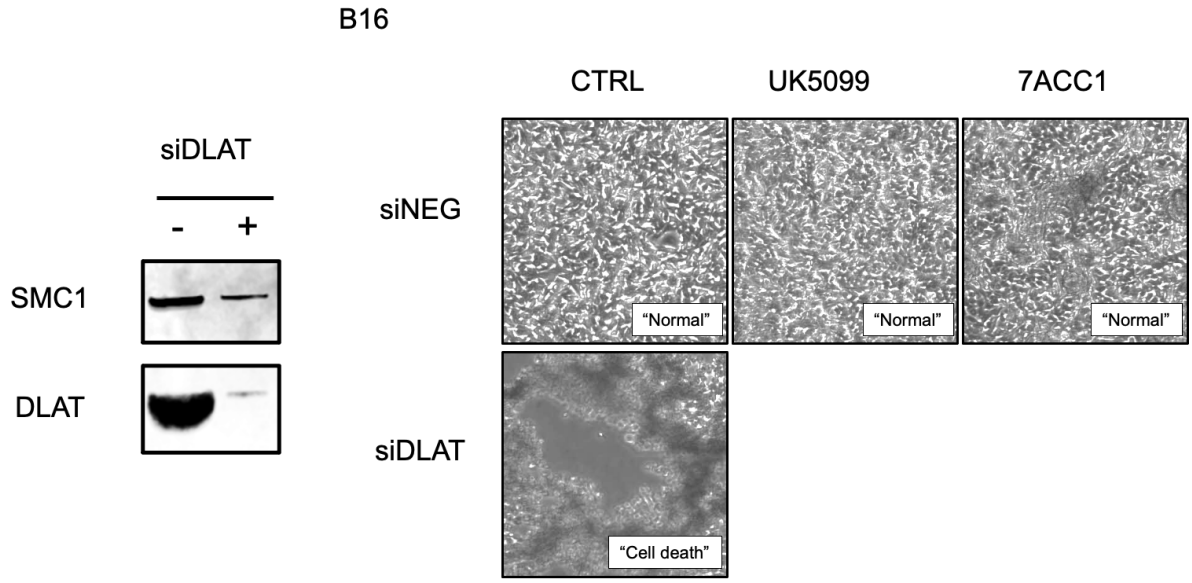


Figure 40. *Left: Western Blot analysis of the DLAT depletion by siRNA in mouse B16 melanoma cells. SMC1 has been used as a housekeeper and si Scramble (-) has been used as a control.Right: Brightfield microscopy analysis inmouse B16 melanoma cells of the effects of siRNA depletion of DLAT vs control si Scramble (siNEG) in combination with Mitochondrial Pyruvate Carrier (MPC) inhibition by either 7ACC2 or UK5099 as indicated cells were grown for 5 days in the indicated conditions. .*

We then tried this siRNA technology in mouse ES cells, but now, we silenced not only one (DLAT), but two subunits from the PDH complex (siPDH=siPDHA1+siDLAT). We were able to observe that the combinatorial effect of the double knockdown was greater, and by this method we could see the effects in cell viability without the need of challenging the system with differentiation. While at day 3 of siPDH infection cells did not show any phenotype, at day 6 of infection the culture of mESCs that lacked PDH was not viable. We used the silencing technology as an alternative way to genetically impair the entrance of pyruvate to the mitochondria and therefore the mitochondrial PDH. To do so we combined siMPC1 and siMPC2 (siMPC) again, limiting the mitochondrial PDH activity did not affect cell viability (Figure 41).

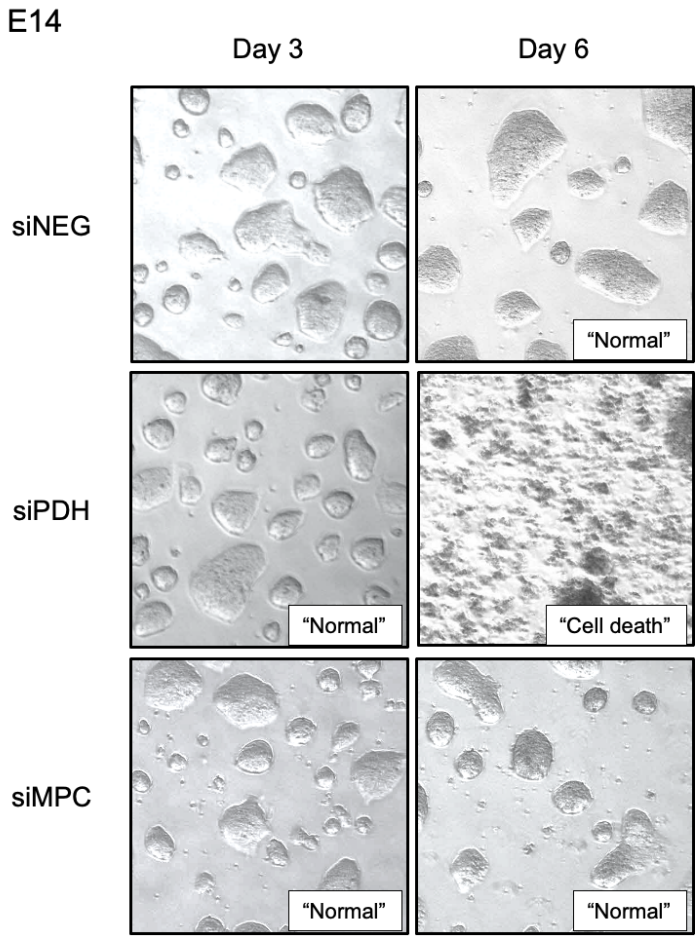


***Figure 41**. Brightfield microscopy analysis in E14 mES cells of the effects of siRNA depletion of PDHA1 and DLAT (siPDH) and MPC1 + MPC2 depletion (siMPC) vs control si Scramble (siNEG) at day 3 and day 6 of the depletions as indicated.*

To confirm that siMPC was successfully impairing mitochondrial PDH activity, we grew the cells in galactose media. Mitochondrial PDH role is crutial in the convertion of pyruvate to acetyl-CoA, which then enters the TCA cycle. The activity of pyruvate dehydrogenase is essential for cells utilizing oxidative metabolism. In cells grown in galactose media, the reliance on mitochondrial metabolism is higher, and therefore, the activity of pyruvate dehydrogenase converting pyruvate to acetyl-CoA becomes crucial.

In contrast, cells grown in glucose media primarily rely on glycolysis for energy production, and the entry of pyruvate into the mitochondria is not as critical for their

survival (Wallace et al., 1992). In these cells, pyruvate can be converted to lactate or undergo other metabolic fates without entering the TCA cycle.

When we cultured our ES cells in galactose media, we again observed that cells that lacked PDH were severely affected, but in this case, siMPC-knockdown also resulted in cell death, indicating that our MPC-depletion worked as expected. Moreover, when we blocked mitochondrial PDH activity in cells that rely in oxidative phosphorylation to obtain energy (in galactose media), they did not survive this intervention (Figure 42).

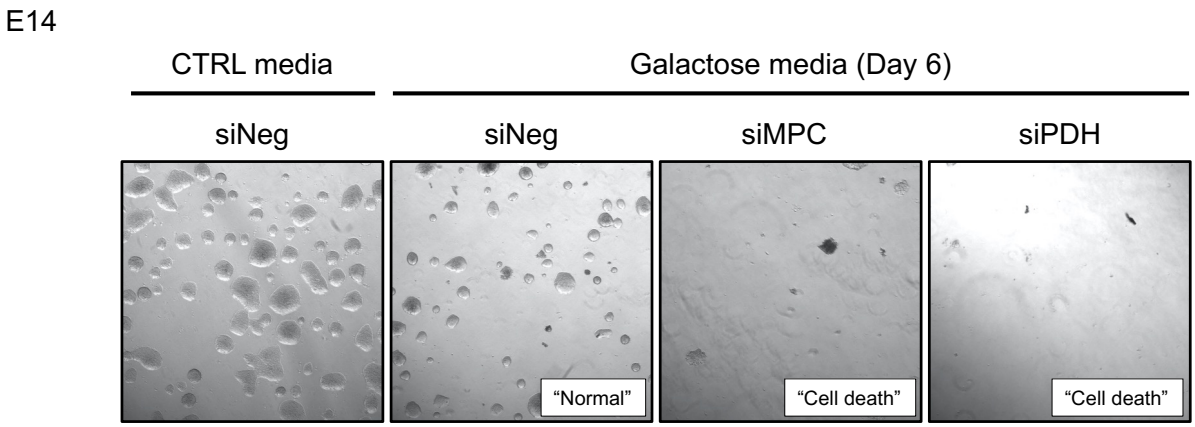


Figure 42. *Brightfield microscopy analysis in E14 mES cells of the effects of siRNA depletion of PDHA1 and DLAT (siPDH) and MPC1+ MPC2 (siMPC) vs control si Scramble (siNEG). Cells were grown in either control (day 6) or Galactose media (day 3) as indicated.*

We observed the same effect when we reproduced the experiment in different cell types, such as SKMEL cells (Figure 43).

Previous studies have shown that global histone hypoacetylation linked to a partially reduced PDH activity, could be rescued by an external supplementation of acetate (Y. Li et al., 2020).
We were able to rescue the lethality of total PDH silencing by supplementation of the culture media with 5mM acetate, providing the cells with an external source for nuclear acetylation. This allowed us to relate the absence of PDH with the absence of a main source of acetyl-CoA needed to maintain cell function and chromatin acetylation (Figure 43).

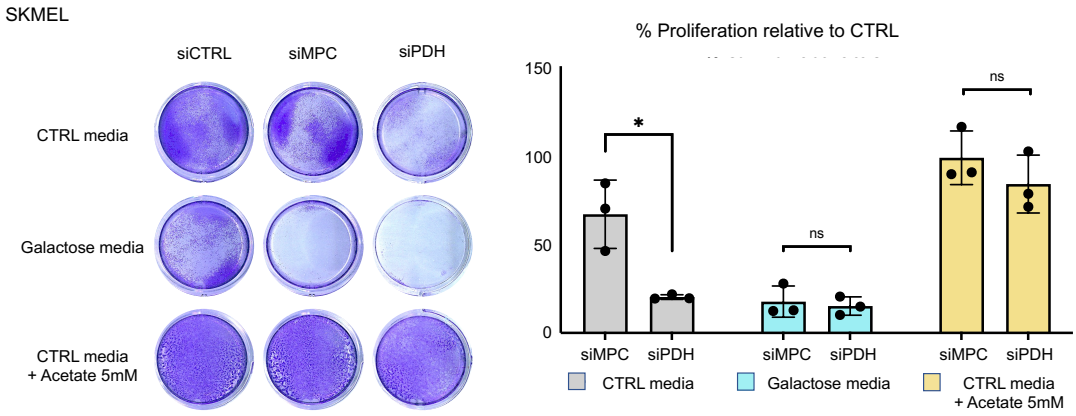


Figure 43. *Left: Viability assay at day 6 of control (siNEG/SCR), silenced MPC (siMPC1+siMPC2) or silenced PDH (siDLAT+siPDHA)1 cells in either control media, galactose media or control media supplemented with 5mM acetate as indicated. Right: Scale bar representation of the % of proliferation of the cells with the different indicated depletions relative to the levels of control (siNEG/CTRL) cells in each corresponding media.*

Discussion

PROXIMITY LIGATION ASSAY (PLA) REVEALS NEW INSIGHTS INTO THE DYNAMICS OF MEDIATOR-POL II INTERACTIONS

We used standard mammalian cell culture and molecular methods such as Western blotting, nuclear- cytoplasmic fractionation, and immunofluorescence, and also established a new technique in this laboratory (PLA) to study interactions among the transcriptional machinery. To our knowledge, the PLA assay has not been applied for this purpose previously. Moreover, we have demonstrated that quantification of the Mediator-Pol II interaction is possible using the PLA assay by standard confocal immunofluorescence, while previously super-resolution microscopy and advanced imaging analysis were required (Cho et al., 2018; Guo et al., 2019). We have found that PLA is a suitable method to observe and quantify endogenous POLR2A-MED1 interactions in a rapid and single-cell manner.

After establishing the PLA assay, we employed it to observe and quantify the interaction between the Mediator complex and RNAPol II in the context of several different cell identities. We have focused on their role in the establishment of the naïve and primed states ES cell pluripotency, a highly characterized prototypical in vitro model of cell identity transitions that faithfully reflects developmental events in the embryo (Weinberg et al., 2016; Hackett and Surani, 2014). Altogether, our data suggests that 2i-naïve ES cells in the naïve state show an increased amount of Mediator- RNA Pol II interactions compared to primed ES cells.

This increase can also be detected in CDK8i treated ES cells. These findings agree with our previous data, which suggested that both 2i and CDK8i treatments may hyper-activate Mediator activity (Lynch et al., 2019), resulting in increased recruitment of RNA Pol II to enhancer-target genes.

We previously observed the link between 2i and CDK8i effects by different means: we performed phospho-proteome analyses d in mESCs cultured in 2i or to CDK8i, observing a high overlap. Of note, we also observed that 2i down-regulates CDK8/19 activity, whereas CDK8i does not affect MEK activity (Lynch et al., 2020). This suggests a simple model by which the effects of MEK and GSK3 inhibition converge on the transcriptional machiner y, and that CDK8/19 may lie downstream of these pathways (Figure 44).

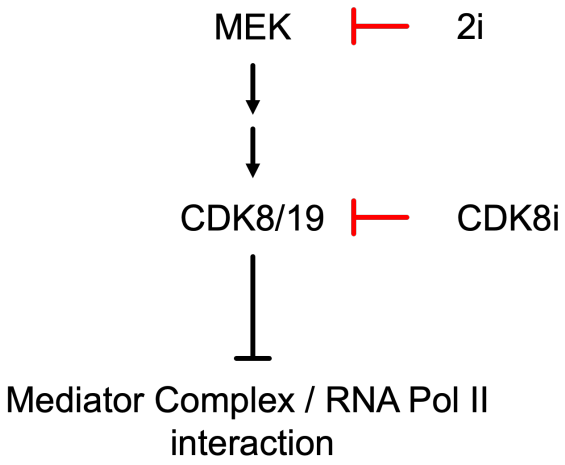


Figure 44. *Model of convergence of 2i and CDK8i in promoting Mediator complex/RNA Pol II interaction.*

We interpret the increased number of Mediator-Pol II interaction foci in 2i or CDK8i to be evidence that Mediator is indeed more active, more frequently recruiting Pol II. This would explain our previous observations that 2i or CDK8i can boost the ability of existing enhancers to drive transcription of their target genes, reinforcing cell identity circuitry, and in the case of ES cells, favor the stabilization of the naïve state (Figure 6). Thus, our quantification of the Mediator-Pol II interaction provides mechanistic insight into a transition in cell identity. Moreover, the naïve-primed developmental transition of ES cells is regarded as a standard model, and thus in the future, we predict that a similar approach would reveal insights in other transitions in cell identity.

However, some outstanding questions remain. Currently, we cannot exclude the possibility that the lower number of Mediator-Pol II interactions in primed ES cells could be due to structural rearrangements of the Mediator complex (Tsai et al., 2017), which could in theory move the MED1 subunit outside the proximity-detection limit of the PLA assay (~40 nm) (Bagchi et al., 2015). Nevertheless, combining the changes observed by PLA, with our previous data which shows that 2i or CDK8i can boost enhancer-driven gene expression, we would suggest that such putative structural rearrangements of Mediator may reflect changes in Mediator activity. Also, looking forwards, it would be informative to probe these putative changes in interaction frequency using antibodies for other subunits of Mediator and Pol II, an approach which could assess the possibility described above, where structural rearrangements of Mediator could exceed the 40 nm limit for the subunits chosen in the PLA. In another direction, we hypothesize that

modification of the length of the antibody-bound oligos might permit the PLA to detect protein interactions across specific distances.

In summary, our data suggests that PLA can be used to observe and quantify the Mediator- Pol II interaction, and that the frequency of this interaction may differ between cell types. We suggest that the changes we observe in Mediator-Pol II interactions may in fact be causal to the changes which occur in transcriptional programs as cells transition between alternate identities. It will be important to pursue this question further, given the central role of Mediator and Pol II in defining cell identity. In this context, changes in Mediator activity may not only stabilize the transcriptional program of naïve ES cells. In the future, this concept could also be applied to stabilize other unstable cellular identities by reinforcing the number of Mediator-Pol II interactions and thus, the transcriptional program of one of those identities. In essence, this would constitute a new approach to toggle plasticity and decisions in cellular identity. Recently, inhibition of CDK8 was reported to modulate T cell identity by inducing the conversion of antigen-specific effector/memory T cells into Foxp3-expressing T_{reg} cells, thus forming an therapy against auto-immunity (Akamatsu et al., 2019; Z. Guo et al., 2019). Moreover, there is an increasing number of studies that use CDK8i as a therapy to treat cancer, either by directly attacking cancer cells which are known to have a high plasticity potency or by modulating the activity of immune cells (Ding et al., 2022; Hofmann et al., 2020; Philip et al., 2018). The mechanisms involved in the modulation of cancer cell progression and aggressiveness could be related to the effects of hyperactivation of the interaction between Mediator and RNA Pol II. This hypothesis should be further explored.

IP-MS MED1 ANALYSIS EXPANDS THE MEDIATOR INTERACTION NETWORK

Following our interest in obtaining a better understanding of Mediator complex, we performed an interactome by using MED-immunoprecipitation, its largest subunit (Verger et al., 2019) by Mass Spectrometry (IP-MS) to identify accessory proteins associated with Mediator.

The high peptide counts for MED-related proteins in addition to other Mediator subunits and known interactors of Mediator (Quevedo et al., 2019; Uthe et al., 2017b)in the IP-MS analysis conducted on ES cells validates the robustness of the chosen approach

validating the obustness of the chosen approach in effectively capturing Mediator-associated proteins. We generated an updated interaction network of Mediator, incorporating novel proteins, with potential significance in Mediator complex functionality. While we found RNA Polymerase II complex elements within the Mediator interactome, the detection levels where moderate, suggesting that while the Mediator-RNA Pol II interaction is crucial, it might be transient. Indeed, this would be consistent with models of transcriptional bursting (Dar et al., 2012). This dynamics of transient but strong interactions fit well in the landscape of phase separated transcriptional condensates (Cho et al., 2018; Y. E. Guo et al., 2019a)

While we could not elaborate into all the interesting putative novel interactors, we provide the scientific community with a list of interesting candidates to study.

THE PDH COMPLEX STANDS OUT AS A NEW MEDIATOR INTERACTOR.

We found different novel and unexpected interactors through MED1 IP-MS, with a particularly strong interaction observed with several subunits of the Pyruvate Dehydrogenase complex. While the PDH complex had been found in the nucleus, and its critical role in histone acetylation has been reported (Chen et al., 2018; Srivastava et al., 2023; Sutendra et al., 2014), the precise subnuclear localization, its interacting neighbor proteins and its potential link to specific chromatin regions remained unknown.

We performed different techniques including fractionation of the different cellular compartments of the cell, immunofluorescence imaging in different cell types and proximity ligation assay to not only localize the enzymes of the pyruvate dehydrogenase complex in the nucleus, but also confirm the proximity of these enzymes and RNA Pol II, therefore situating the PDH complex at the core transcriptional machinery. This raised the provocative possibility that the nuclear PDH complex could generate acetyl groups proximal to sites of consumption for histone acetylation in transcription units.

Once we confirmed the interaction, we further explored wether this nuclear PDH was active. We confirmed the presence of the active isoform of DLAT which is lipoylated form of DLAT (Mathias et al., 2014; Rowland et al., 2018; Tsvetkov et al., 2022), indicating a local source for acetyl-CoA production.

PDH IS LOCATED AT ENHANCER REGIONS AND HIGHLY ACETYLATED AND ACTIVELY TRANSCRIBED CHROMATIN

Uncovering the specific localization of the Pyruvate Dehydrogenase (PDH) complex within chromatin and its associations with distinct genomic elements was the next challenge. Understanding the specific genomic regions where PDH binds chromatin represented by the DLAT subunit, localizes adds a layer of specificity to PDH functional implications in chromatin regulation. Ourdentification of PDH within proximity to enhancers, promoters, and transcription units underscores its regulatory role. In particular, alignment within Super-Enhancers (SEs) particularly stands out, given their crucial role in driving cell identity and transcriptional programs. The detection of active PDH in the nucleus suggests a local source of acetyl groups, potentially contributing to the acetylation landscape within chromatin. Notably, our Chromatin Immunoprecipitation (ChIP-Seq) analyses revealed a strong genomic coincidence of DLAT-bound chromatin regions with various histone acetylation marks.

PDH's potential to supply acetyl groups locally suggest that it could work as a partner with Mediator in orchestrating precise and localized chromatin modifications and allowing transcriptional control of cellular identity and plasticity. PDH's local acetyl group generation points to a role of PDH in facilitating the activity of CBP/E300 (Rahman et al., 2004) histone acetyltransferases (HATs) by locally providing acetyl CoA that will be used at high demand (Weinert et al., 2018).

CBP/E300 HATs could then use the acetyl groups to modify histones, leading to chromatin relaxation and creating a permissive environment for transcription (Hebbes et al., 1988; Marmorstein & Zhou, 2014; Yamada et al., 2004) (see suggested model Figure 45).

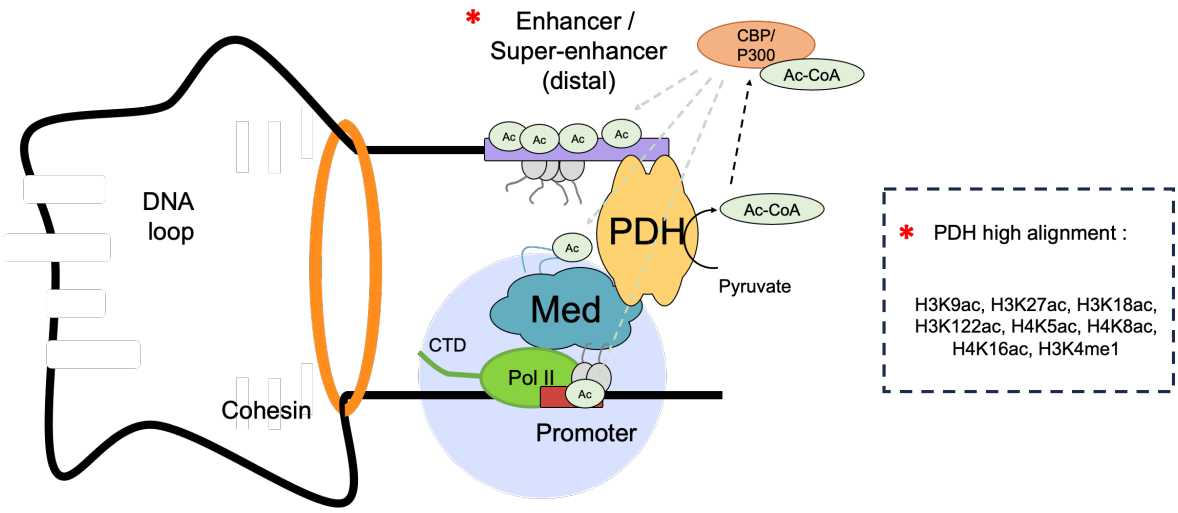


Figure 45. Suggested model: The Mediator Complex - PDH Complex Axis as a Central Hub for Acetyl Moiety Supply. *Our proposed working model places the Mediator complex-PDH complex axis as a central hub orchestrating the efficient supply of acetyl moieties to the transcriptional machinery. This regulatory network could play a pivotal role in chromatin opening, histone acetylation, and transcriptional regulation. Regions of the chromatin enriched in DLAT appear indicated.*

The interaction and the convergence of PDH and the Mediator complex within specific chromatin regions points towards a synergistic relationship in the fine-tuned control of transcription. Together, PDH and Mediator may coordinate their activities to influence the local chromatin landscape, ensuring precise control over genes requiring transcriptional activation. A possible hypotheses is that when Mediator-RNA Pol II are more actively transcribing genes, they also recruit more PDH acetyl enzymes close to the rest of the transcriptional machinery given that an insufficient amount of acetyl moiety supply to the active transcription sites could have a dramatic effect in transcription (Mallm et al., 2019; Sterner & Berger, 2000; J. P. Taylor et al., 2003). This recruitment could happen via phase separation, as acetylation isone of the few studied mechanisms of regulation in the formation and dissolution of LLPS transcriptional condensates. In this regard, it would be interesting to manipulate the presence of these acetyl-producing enzymes in a controlled manner inside the nucleus.

The manipulation of the Pyruvate Dehydrogenase (PDH) complex, specifically the knockdown of its core catalytic subunit DLAT, unraveled critical insights into the functional implications of PDH in cellular processes, particularly during cellular

differentiation. This suggests a specific requirement for PDH during dynamic cellular states, particularly when chromatin undergoes remodeling and new regions demand acetylation for transcriptional changes (Gregory et al., 2001; Rahman et al., 2004; Yamada et al., 2004).

We found that external acetate supplementation rescue of the lethality of total PDH knockdown in SKMEL cells. This underscores the crucial role of PDH as a major source of acetyl-CoA essential for cell function and chromatin acetylation.

Future investigations should delve into the specific molecular mechanisms through which PDH-generated acetyl groups influence chromatin dynamics.
The identification of PDH as a potential local acetyl group generator in collaboration with Mediator opens avenues for future research and potential therapeutic interventions. Modulating the local generation of acetyl groups could have implications for diseases involving dysregulated transcriptional control and cellular transitions, as cancer (Bradner et al., 2017), neurodegenerative diseases (Labadorf et al., 2018), autoimmune diseases (Wu et al., 2018), diabetes (Moin & Butler., 2019)and more .
In summary, the novel association between PDH and Mediator, coupled with the significant overlap of DLAT with enhancers and super-enhancers, emphasizes the importance of PDH in the regulation of gene expression and chromatin structure. This finding expands our understanding of the intricate network governing transcriptional control and highlights a previously unrecognized player in the complex landscape of cellular identity regulation. Our study not only elucidates the functional significance of PDH in cellular processes but also provides a foundation for future investigations into the intricate interplay between metabolic pathways and transcriptional regulation. The findings underscore the complexity of cellular identity transitions and highlight potential avenues for therapeutic interventions targeting PDH-mediated processes.

Conclusions

1. **Proximity ligation assay represents a novel visualization technique that allows us to comprehensively capture the dynamics of Mediator-Pol II interactions.**
   1.1. The use of PLA allows the observation and quantification of 10-50 foci per nucleus, suggesting the capture of the largest Mediator-RNA Pol II interactions.
   1.2. The number of Mediator-RNA Pol II interaction foci per nucleus significantly increase in number in 2i- and CDK8i- induced naïve ES cells compared to primed ES cells.
   1.3. There is a common mechanism influenced by both 2i and CDK8i, promoting the association of Mediator and RNA Pol II.

2. **Immunoprecipitation followed by Mass Spectrometry (IP-MS) analysis to encrich the interactors of the Mediator complex has provided new insights into the protein interaction network of the Mediator complex.**
   2.1. Known core and conserved Mediator interactors are detectable through IP-MS of MED1, confirming the representation of the previously reported Mediator interactome in mouse ES cells.
   2.2. Novel and potentially crucial proteins linked with the Mediator complex are found in our updated interaction network of Mediator.
   2.3. Mediator closely interacts with all subunits of the Pyruvate Dehydrogenase (PDH) complex.

3. **Mediator Interacts with a functional nuclear PDH Complex and its targeting affects cellular identity and viability.**
   3.1. There is a lipoylated/active form of nuclear DLAT, distinct from the mitochondrial form.
   3.2. The DLAT subunit of the PDH complex exhibits a high abundance in proximity to the POLR2A subunit of RNA Pol II, indicating a close association between the PDH complex and the core transcriptional machinery.
   3.3. ChIP-Seq analysis of the genomic distribution of DLAT as a surrogate of the rest of the PDH complex identifies binding loci within euchromatic regions of actively transcribed chromatin and distal intergenic regions associated with enhancer marks (H3K27ac and H3K4me1) and active transcription, aligning with the localization of Mediator.

3.4. DLAT/PDH complex-bound chromatin regions are linked to the control of pluripotency and cell proliferation, crucial features of ES cell identity.

3.5. Experimental manipulation of the PDH complex unveils the functional significance of this complex in the transcriptional regulation of cell identity and viability. These effects are independent of the mitochondrial PDH activity.

Bibliography

Akamatsu, M., Mikami, N., Ohkura, N., Kawakami, R., Kitagawa, Y., Sugimoto, A., Hirota, K., Nakamura, N., Ujihara, S., Kurosaki, T., Hamaguchi, H., Harada, H., Xia, G., Morita, Y., Aramori, I., Narumiya, S., & Sakaguchi, S. (2019). Conversion of antigen-specific effector/memory T cells into Foxp3-expressing Tregcells by inhibition of CDK8/19. *Science Immunology*, *4*(40), 1–17. https://doi.org/10.1126/SCIIMMUNOL.AAW2707

Akgol Oksuz, B., Yang, L., Abraham, S., Venev, S. V., Krietenstein, N., Parsi, K. M., Ozadam, H., Oomen, M. E., Nand, A., Mao, H., Genga, R. M. J., Maehr, R., Rando, O. J., Mirny, L. A., Gibcus, J. H., & Dekker, J. (2021). Systematic evaluation of chromosome conformation capture assays. *Nature Methods*, *18*(9), 1046–1055. https://doi.org/10.1038/s41592-021-01248-7

Amemiya, H. M., Kundaje, A., & Boyle, A. P. (2019). The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Scientific Reports*, *9*(1), 1–5. https://doi.org/10.1038/s41598-019-45839-z

Andrew J, B., Robert, S., & Tony, K. (2002). Histone methylation: dynamic or static?. *Cell*, *109*, 801–806.

Bailey, L. B., & Gregory, J. F. (2018). Folate Metabolism and Requirements. *Recent Advances in Nutritional*, *March*, 779–782.

Baltimore, D. (1970). Viral RNA-dependent DNA polymerase: RNA-dependent DNA polymerase in virions of RNA tumour viruses. In *Nature* (Vol. 226, Issue 5252, pp. 1209–1211). https://doi.org/10.1038/2261209a0

Becker, M., Baumann, C., John, S., Walker, D. A., Vigneron, M., McNally, J. G., & Hager, G. L. (2002). Dynamic behavior of transcription factors on a natural promoter in living cells. *EMBO Reports*, *3*(12), 1188–1194. https://doi.org/10.1093/embo-reports/kvf244

Berman, B. P., Nibu, Y., Pfeiffer, B. D., Tomancak, P., Celniker, S. E., Levine, M., Rubin, G. M., & Eisen, M. B. (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the Drosophila genome. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(2), 757–762. https://doi.org/10.1073/pnas.231608898

Bogu, G. K., Vizán, P., Stanton, L. W., Beato, M., Di Croce, L., & Marti-Renom, M. A. (2016). Chromatin and RNA Maps Reveal Regulatory Long Noncoding RNAs in Mouse. *Molecular and Cellular Biology*, *36*(5), 809–819. https://doi.org/10.1128/mcb.00955-15

Boroviak, T., Loos, R., Bertone, P., Smith, A., & Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nature Cell Biology*, *16*(6), 513–525. https://doi.org/10.1038/ncb2965

Boyle, S., Gilchrist, S., Bridger, J. M., Mahy, N. L., Ellis, J. A., & Bickmore, W. A. (2001). The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human Molecular Genetics*, *10*(3), 211–219. https://doi.org/10.1093/hmg/10.3.211

Bradner, J. E., Hnisz, D., & Young, R. A. (2017). Transcriptional Addiction in Cancer. *Cell*, *168*(4), 629–643. https://doi.org/10.1016/j.cell.2016.12.013

Brangwynne, C. P., Eckmann, C. R., Courson, D. S., Rybarska, A., Hoege, C., Gharakhani, J., Jülicher, F., & Hyman, A. A. (2009). Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, *324*(5935), 1729–1732. https://doi.org/10.1126/science.1172046

Butler, J. E. F., & Kadonaga, J. T. (2002). The RNA polymerase II core promoter: A key component in the regulation of gene expression. *Genes and Development*, *16*(20), 2583–2592. https://doi.org/10.1101/gad.1026202

Canham, M. A., Sharov, A. A., Ko, M. S. H., & Brickman, J. M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biology*, *8*(5). https://doi.org/10.1371/journal.pbio.1000379

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., & Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature*, *450*(7173), 1230–1234. https://doi.org/10.1038/nature06403

Chandel, N. S. (2015). Evolution of Mitochondria as Signaling Organelles. *Cell Metabolism*, *22*(2), 204–206. https://doi.org/10.1016/j.cmet.2015.05.013

Chen, J., Guccini, I., Mitri, D. Di, Brina, D., Revandkar, A., Sarti, M., Pasquini, E., Alajati, A., Pinton, S., Losa, M., Civenni, G., Catapano, C. V., Sgrignani, J., Cavalli, A., D'Antuono, R., Asara, J. M., Morandi, A., Chiarugi, P., Crotti, S., … Alimonti, A. (2018). Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer. *Nature Genetics*, *50*(2), 219–228. https://doi.org/10.1038/s41588-017-0026-3

Cho, W. K., Spille, J. H., Hecht, M., Lee, C., Li, C., Grube, V., & Cisse, I. I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science*, *361*(6400), 412–415. https://doi.org/10.1126/science.aar4199

Clark, A. D., Oldenbroek, M., & Boyer, T. G. (2015). Mediator kinase module and human tumorigenesis. *Critical Reviews in Biochemistry and Molecular Biology*, *50*(5), 393–426. https://doi.org/10.3109/10409238.2015.1064854

Clayton, A. L., Hazzalin, C. A., & Mahadevan, L. C. (2006). Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Molecular Cell*, *23*(3), 289–296. https://doi.org/10.1016/j.molcel.2006.06.017

Cramer, P. (2019). Organization and regulation of gene transcription. *Nature*, *573*(7772), 45–54. https://doi.org/10.1038/s41586-019-1517-4

Crick, F. (1970). Central Dogma of Molecular Biology. *Nature*, *228*, 726–734. http://www.mendeley.com/research/discreteness-conductance-chnge-n-bimolecular-lipid-membrane-presence-certin-antibiotics/

Dar, R. D., Razooky, B. S., Singh, A., Trimeloni, T. V., McCollum, J. M., Cox, C. D., Simpson, M. L., & Weinberger, L. S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(43), 17454–17459. https://doi.org/10.1073/pnas.1213530109

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H. G., Lensch, M. W., Lujan, E., Pei, D., Rossant, J., Wernig, M., Park, P. J., & Daley, G. Q. (2015). Hallmarks of pluripotency. *Nature*, *525*(7570), 469–478. https://doi.org/10.1038/nature15515

Dekker, J., & Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. *FEBS Letters*, *589*(20), 2877–2884. https://doi.org/10.1016/j.febslet.2015.08.044

Dignon, G. L., Best, R. B., & Mittal, J. (2020). Biomolecular phase separation: From molecular driving forces to macroscopic properties. *Annual Review of Physical Chemistry*, *71*, 53–75. https://doi.org/10.1146/annurev-physchem-071819-113553

Ding, X., Sharko, A. C., McDermott, M. S. J., Schools, G. P., Chumanevich, A., Ji, H., Li, J., Zhang, L., Mack, Z. T., Sikirzhytski, V., Shtutman, M., Ivers, L., O'Donovan, N., Crown, J., Gyorffy, B., Chen, M., Roninson, I. B., & Broude, E. V. (2022). Inhibition of CDK8/19 Mediator kinase potentiates HER2-targeting drugs and bypasses resistance to these agents in vitro and in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(32), 1–11. https://doi.org/10.1073/pnas.2201073119

Ducker, G. S., & Rabinowitz, J. D. (2017). One-Carbon Metabolism in Health and Disease. *Cell Metabolism*, *25*(1), 27–42. https://doi.org/10.1016/j.cmet.2016.08.009

Fant, C. B., & Taatjes, D. J. (2019). Regulatory functions of the Mediator kinases CDK8 and CDK19. *Transcription*, *10*(2), 76–90. https://doi.org/10.1080/21541264.2018.1556915

Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., Kriwacki, R. W., Pappu, R. V., & Brangwynne, C. P. (2016). Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell*, *165*(7), 1686–1697. https://doi.org/10.1016/j.cell.2016.04.047

Ficz, G., Hore, T. A., Santos, F., Lee, H. J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y. L., Walter, J., Cook, S. J., Andrews, S., Branco, M. R., & Reik, W. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell*, *13*(3), 351–359. https://doi.org/10.1016/j.stem.2013.06.004

Filtz, T. M., Vogel, W. K., & Leid, M. (2014). Regulation of transcription factor activity by interconnected post-translational modifications. *Trends in Pharmacological Sciences*, *35*(2), 76–85. https://doi.org/10.1016/j.tips.2013.11.005

Galbraith, M. D., Donner, A. J., & Espinosa, J. M. (2010). CDK8: A positive regulator of transcription. *Transcription*, *1*(1), 4–12. https://doi.org/10.4161/trns.1.1.12373

Galdieri, L., Zhang, T., Rogerson, D., Lleshi, R., & Vancura, A. (2014). Protein

acetylation and acetyl coenzyme a metabolism in Budding Yeast. *Eukaryotic Cell*, *13*(12), 1472–1483. https://doi.org/10.1128/EC.00189-14

Gibney, E. R., & Nolan, C. M. (2010). Epigenetics and gene expression. *Heredity*, *105*(1), 4–13. https://doi.org/10.1038/hdy.2010.54

Gibson, B. A., Doolittle, L. K., Schneider, M. W. G., Jensen, L. E., Gamarra, N., Henry, L., Gerlich, D. W., Redding, S., & Rosen, M. K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell*, *179*(2), 470-484.e21. https://doi.org/10.1016/j.cell.2019.08.037

Gregory, P. D., Wagner, K., & Hörz, W. (2001). Histone acetylation and chromatin remodeling. *Experimental Cell Research*, *265*(2), 195–202. https://doi.org/10.1006/excr.2001.5187

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature*, *389*(6649), 349–352. https://doi.org/10.1038/38664

Guo, Y. E., Manteiga, J. C., Henninger, J. E., Sabari, B. R., Dall'Agnese, A., Hannett, N. M., Spille, J. H., Afeyan, L. K., Zamudio, A. V., Shrinivas, K., Abraham, B. J., Boija, A., Decker, T. M., Rimel, J. K., Fant, C. B., Lee, T. I., Cisse, I. I., Sharp, P. A., Taatjes, D. J., & Young, R. A. (2019a). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*, *572*(7770), 543–548. https://doi.org/10.1038/s41586-019-1464-0

Guo, Y. E., Manteiga, J. C., Henninger, J. E., Sabari, B. R., Dall'Agnese, A., Hannett, N. M., Spille, J. H., Afeyan, L. K., Zamudio, A. V., Shrinivas, K., Abraham, B. J., Boija, A., Decker, T. M., Rimel, J. K., Fant, C. B., Lee, T. I., Cisse, I. I., Sharp, P. A., Taatjes, D. J., & Young, R. A. (2019b). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature*, *572*(7770), 543–548. https://doi.org/10.1038/s41586-019-1464-0

Guo, Z., Wang, G., Lv, Y., Wan, Y. Y., & Zheng, J. (2019). Inhibition of Cdk8/Cdk19 activity promotes treg cell differentiation and suppresses autoimmune diseases. *Frontiers in Immunology*, *10*(AUG), 1–10. https://doi.org/10.3389/fimmu.2019.01988

Hackett, J. A., & Azim Surani, M. (2014). Regulatory principles of pluripotency: From the ground state up. *Cell Stem Cell*, *15*(4), 416–430. https://doi.org/10.1016/j.stem.2014.09.015

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. *Nature Structural and Molecular Biology*, *11*(5), 394–403. https://doi.org/10.1038/nsmb763

Handy, D. E., Castro, R., & Loscalzo, J. (2011). Epigenetic modifications: Basic mechanisms and role in cardiovascular disease. *Circulation*, *123*(19), 2145–2156. https://doi.org/10.1161/CIRCULATIONAHA.110.956839

Hebbes, T. R., Thorne, A. W., & Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *The EMBO Journal*, *7*(5), 1395–1402. https://doi.org/10.1002/j.1460-2075.1988.tb02956.x

Heinz, S., Romanoski, C. E., Benner, C., & Glass, C. K. (2015). The selection and function of cell type-specific enhancers. *Nature Reviews Molecular Cell Biology*, *16*(3), 144–154. https://doi.org/10.1038/nrm3949

Hengartner, C. J., Myer, V. E., Liao, S. M., Wilson, C. J., Koh, S. S., & Young, R. A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Molecular Cell*, *2*(1), 43–53. https://doi.org/10.1016/S1097-2765(00)80112-4

Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., Hoke, H. A., & Young, R. A. (2013). XSuper-enhancers in the control of cell identity and disease. *Cell*, *155*(4), 934. https://doi.org/10.1016/j.cell.2013.09.053

Hofmann, M. H., Mani, R., Engelhardt, H., Impagnatiello, M. A., Carotta, S., Kerenyi, M., Lorenzo-Herrero, S., Böttcher, J., Scharn, D., Arnhof, H., Zoephel, A., Schnitzer, R., Gerstberger, T., Sanderson, M. P., Rajgolikar, G., Goswami, S., Vasu, S., Ettmayer, P., Gonzalez, S., … Moll, J. (2020). Selective and potent CDK8/19 inhibitors enhance NK-cell activity and promote tumor surveillance. *Molecular Cancer Therapeutics*, *19*(4), 1018–1030. https://doi.org/10.1158/1535-7163.MCT-19-0789

Hsin, J. P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes and Development*, *26*(19), 2119–2137. https://doi.org/10.1101/gad.200303.112

Hur, W., Kemp, J. P., Tarzia, M., Deneke, V. E., Marzluff, W. F., Duronio, R. J., & Di Talia, S. (2020). CDK-Regulated Phase Separation Seeded by Histone Genes Ensures Precise Growth and Function of Histone Locus Bodies. *Journal of Cleaner*

Production, 1–16. https://doi.org/10.1016/j.devcel.2020.06.003

Inouye, M., & Delihast, N. (1988). Small RNAs in the prokaryotes: A growing list of diverse roles. *Cell*, *53*(1), 5–7. https://doi.org/10.1016/0092-8674(88)90480-1

Jarvius, M., Paulsson, J., Weibrecht, I., Leuchowius, K. J., Anderson, A. C., Wählby, C., Gullberg, M., Botling, J., Sjöblom, T., Markova, B., Östman, A., Landegren, U., & Söderberg, O. (2007). In situ detection of phosphorylated platelet-derived growth factor receptor β using a generalized proximity ligation method. *Molecular and Cellular Proteomics*, *6*(9), 1500–1509. https://doi.org/10.1074/mcp.M700166-MCP200

Jeronimo, C., & Robert, F. (2017). The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. *Trends in Cell Biology*, *27*(10), 765–783. https://doi.org/10.1016/j.tcb.2017.07.001

Kaufmann, R., Schellenberger, P., Seiradake, E., Dobbie, I. M., Jones, E. Y., Davis, I., Hagen, C., & Grünewald, K. (2014). Super-resolution microscopy using standard fluorescent proteins in intact cells under cryo-conditions. *Nano Letters*, *14*(7), 4171–4175. https://doi.org/10.1021/nl501870p

Kornberg, R. (1974). Chromatin Structure : A Repeating Unit of Histones and DNA Chromatin structure is based on a repeating unit of eight. *Science*, *184*, 868–871.

Kunath, T., Arnaud, D., Uy, G. D., Okamoto, I., Chureau, C., Yamanaka, Y., Heard, E., Gardner, R. L., Avner, P., & Rossant, J. (2005). Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development*, *132*(7), 1649–1661. https://doi.org/10.1242/dev.01715

Kunath, T., Saba-El-Leil, M. K., Almousailleakh, M., Wray, J., Meloche, S., & Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, *134*(16), 2895–2902. https://doi.org/10.1242/dev.02880

Kutach, A. K., & Kadonaga, J. T. (2000). The Downstream Promoter Element DPE Appears To Be as Widely Used as the TATA Box in Drosophila Core Promoters . *Molecular and Cellular Biology*, *20*(13), 4754–4764. https://doi.org/10.1128/mcb.20.13.4754-4764.2000

Labadorf, A., Choi, S. H., & Myers, R. H. (2018). Evidence for a pan-

neurodegenerative disease response in Huntington's and Parkinson's disease expression profiles. *Frontiers in Molecular Neuroscience*, *10*(January), 1–12. https://doi.org/10.3389/fnmol.2017.00430

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S., & Narlikar, G. J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, *547*(7662), 236–240. https://doi.org/10.1038/nature22822

Leitch, H. G., Mcewen, K. R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J. G., Smith, A., Surani, M. A., & Hajkova, P. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nature Structural and Molecular Biology*, *20*(3), 311–316. https://doi.org/10.1038/nsmb.2510

Li, X., Yu, W., Qian, X., Xia, Y., Zheng, Y., Lee, J. H., Li, W., Lyu, J., Rao, G., Zhang, X., Qian, C. N., Rozen, S. G., Jiang, T., & Lu, Z. (2017). Nucleus-Translocated ACSS2 Promotes Gene Transcription for Lysosomal Biogenesis and Autophagy. *Molecular Cell*, *66*(5), 684-697.e9. https://doi.org/10.1016/j.molcel.2017.04.026

Li, Y., Gruber, J. J., Litzenburger, U. M., Zhou, Y., Miao, Y. R., LaGory, E. L., Li, A. M., Hu, Z., Yip, M., Hart, L. S., Maris, J. M., Chang, H. Y., Giaccia, A. J., & Ye, J. (2020). Acetate supplementation restores chromatin accessibility and promotes tumor cell differentiation under hypoxia. *Cell Death and Disease*, *11*(2). https://doi.org/10.1038/s41419-020-2303-9

Linden, C. Vander, & Corbet, C. (2018). Killing two birds with one stone: Blocking the mitochondrial pyruvate carrier to inhibit lactate uptake by cancer cells and radiosensitize tumors. *Molecular and Cellular Oncology*, *5*(4). https://doi.org/10.1080/23723556.2018.1465016

Linn, T. C., Pettit, F. H., & Reed, L. J. (1969). Alpha-keto acid dehydrogenase complexes. X. Regulation of the activity of the pyruvate dehydrogenase complex from beef kidney mitochondria by phosphorylation and dephosphorylation. *Proceedings of the National Academy of Sciences of the United States of America*, *62*(1), 234–241. https://doi.org/10.1073/pnas.62.1.234

Liu, Y., Ai, C., Gan, T., Wu, J., Jiang, Y., Liu, X., Lu, R., Gao, N., Li, Q., Ji, X., & Hu, J. (2021). Transcription shapes DNA replication initiation to preserve genome integrity. *Genome Biology*, *22*(1), 1–27. https://doi.org/10.1186/s13059-021-02390-3

Lu, H., Yu, D., Hansen, A. S., Ganguly, S., Liu, R., Heckert, A., Darzacq, X., & Zhou, Q. (2018). Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature*, *558*(7709), 318–323. https://doi.org/10.1038/s41586-018-0174-3

Lynch, C. J., Bernad, R., Martínez-Val, A., Shahbazi, M. N., Nóbrega-Pereira, S., Calvo, I., Blanco-Aparicio, C., Tarantino, C., Garreta, E., Richart-Ginés, L., Alcazar, N., Graña-Castro, O., Gómez-Lopez, G., Aksoy, I., Muñoz-Martín, M., Martinez, S., Ortega, S., Prieto, S., Simboeck, E., … Serrano, M. (2020). Global hyperactivation of enhancers stabilizes human and mouse naive pluripotency through inhibition of CDK8/19 Mediator kinases. *Nature Cell Biology*, *22*(10), 1223–1238. https://doi.org/10.1038/s41556-020-0573-1

Malik, S., & Roeder, R. G. (2016). Mediator: A Drawbridge across the Enhancer-Promoter Divide. *Molecular Cell*, *64*(3), 433–434. https://doi.org/10.1016/j.molcel.2016.10.024

Mallm, J., Iskar, M., Ishaque, N., Klett, L. C., Kugler, S. J., Muino, J. M., Teif, V. B., Poos, A. M., Großmann, S., Erdel, F., Tavernari, D., Koser, S. D., Schumacher, S., Brors, B., König, R., Remondini, D., Vingron, M., Stilgenbauer, S., Lichter, P., … Rippe, K. (2019). Linking aberrant chromatin features in chronic lymphocytic leukemia to transcription factor networks. *Molecular Systems Biology*, *15*(5), 1–20. https://doi.org/10.15252/msb.20188339

Manor, Y. S., Massarwa, R., & Hanna, J. H. (2015). Establishing the human naïve pluripotent state. *Current Opinion in Genetics and Development*, *34*, 35–45. https://doi.org/10.1016/j.gde.2015.07.005

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Francis Stewart, A., Smith, A., & Stunnenberg, H. G. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell*, *149*(3), 590–604. https://doi.org/10.1016/j.cell.2012.03.026

Marmorstein, R., & Zhou, M. M. (2014). Writers and readers of histone acetylation:

Structure, mechanism, and inhibition. *Cold Spring Harbor Perspectives in Biology*, *6*(7). https://doi.org/10.1101/cshperspect.a018762

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 II), 7634–7638. https://doi.org/10.1073/pnas.78.12.7634

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journa*, *17*, 10–12.

Martinez-Val, A., Lynch, C. J., Calvo, I., Ximénez-Embún, P., Garcia, F., Zarzuela, E., Serrano, M., & Munoz, J. (2021). Dissection of two routes to naïve pluripotency using different kinase inhibitors. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-22181-5

Mathias, R. A., Greco, T. M., Oberstein, A., Budayeva, H. G., Chakrabarti, R., Rowland, E. A., Kang, Y., Shenk, T., & Cristea, I. M. (2014). Sirtuin 4 is a lipoamidase regulating pyruvate dehydrogenase complex activity. *Cell*, *159*(7), 1615–1625. https://doi.org/10.1016/j.cell.2014.11.046

Mathieu, J., & Ruohola-Baker, H. (2017). Metabolic remodeling during the loss and acquisition of pluripotency. *Development (Cambridge)*, *144*(4), 541–551. https://doi.org/10.1242/dev.128389

Millán-Zambrano, G., Burton, A., Bannister, A. J., & Schneider, R. (2022). Histone post-translational modifications — cause and consequence of genome function. *Nature Reviews Genetics*, *23*(9), 563–580. https://doi.org/10.1038/s41576-022-00468-7

Mitchell, P. (1961). Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic type of Mechanism. *Nature*, *191*, 144–148.

Mizuno, T., Chou, M. Y., & Inouye, M. (1984). A unique mechanism regulating gene expression: Translational inhibition by a complementary RNA transcript (micRNA). *Proceedings of the National Academy of Sciences of the United States of America*, *81*(7 I), 1966–1970. https://doi.org/10.1073/pnas.81.7.1966

Mizutani, S., & Temin, H. M. (1970). An RNA-Dependent DNA Polymerase in Virions of Rous Sarcoma Virus. In *Cold Spring Harbor Symposia on Quantitative Biology*

(Vol. 35, Issue 0, pp. 847–849). https://doi.org/10.1101/sqb.1970.035.01.100

Moin, A. S. M., & Butler, A. E. (2019). Alterations in Beta Cell Identity in Type 1 and Type 2 Diabetes. *Current Diabetes Reports*, *19*(9), 1–12. https://doi.org/10.1007/s11892-019-1194-6

Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., & Taylor, J. P. (2015). Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell*, *163*(1), 123–133. https://doi.org/10.1016/j.cell.2015.09.015

Nagaraj, R., Sharpley, M. S., Chi, F., Braas, D., Zhou, Y., Kim, R., Clark, A. T., & Banerjee, U. (2017). Nuclear Localization of Mitochondrial TCA Cycle Enzymes as a Critical Step in Mammalian Zygotic Genome Activation. *Cell*, *168*(1–2), 210-223.e11. https://doi.org/10.1016/j.cell.2016.12.026

Nichols, J., & Smith, A. (2009). Naive and Primed Pluripotent States. *Cell Stem Cell*, *4*(6), 487–492. https://doi.org/10.1016/j.stem.2009.05.015

Nunnari, J., & Suomalainen, A. (2012). Mitochondria: In sickness and in health. *Cell*, *148*(6), 1145–1159. https://doi.org/10.1016/j.cell.2012.02.035

Orphanides, G., Lagrange, T., & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes and Development*, *10*(21), 2657–2683. https://doi.org/10.1101/gad.10.21.2657

Pelish, H. E., Liau, B. B., Nitulescu, I. I., Tangpeerachaikul, A., Poss, Z. C., Da Silva, D. H., Caruso, B. T., Arefolov, A., Fadeyi, O., Christie, A. L., Du, K., Banka, D., Schneider, E. V., Jestel, A., Zou, G., Si, C., Ebmeier, C. C., Bronson, R. T., Krivtsov, A. V., … Shair, M. D. (2015). Mediator kinase inhibition further activates super-enhancer-associated genes in AML. *Nature*, *526*(7572), 273–276. https://doi.org/10.1038/nature14904

Philip, S., Kumarasiri, M., Teo, T., Yu, M., & Wang, S. (2018). Cyclin-Dependent Kinase 8: A New Hope in Targeted Cancer Therapy? [Review-article]. *Journal of Medicinal Chemistry*, *61*(12), 5073–5092. https://doi.org/10.1021/acs.jmedchem.7b00901

Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B. A., Solcan, N., Cerase, A., Castello, A., Mohammed, S., Moindrot, B., Nesterova, T. B., & Brockdorff, N. (2017).

hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish Polycomb-Mediated Chromosomal Silencing. *Molecular Cell*, *68*(5), 955-969.e10. https://doi.org/10.1016/j.molcel.2017.11.013

Poss, Z. C., Ebmeier, C. C., Odell, A. T., Tangpeerachaikul, A., Lee, T., Pelish, H. E., Shair, M. D., Dowell, R. D., Old, W. M., & Taatjes, D. J. (2016). Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Reports*, *15*(2), 436–450. https://doi.org/10.1016/j.celrep.2016.03.030

Puchalska, P., & Crawford, P. A. (2017). Multi-dimensional Roles of Ketone Bodies in Fuel Metabolism, Signaling, and Therapeutics. *Cell Metabolism*, *25*(2), 262–284. https://doi.org/10.1016/j.cmet.2016.12.022

Quevedo, M., Meert, L., Dekker, M. R., Dekkers, D. H. W., Brandsma, J. H., van den Berg, D. L. C., Ozgür, Z., IJcken, W. F. J. va., Demmers, J., Fornerod, M., & Poot, R. A. (2019). Mediator complex interaction partners organize the transcriptional network that defines neural stem cells. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-10502-8

Radzisheuskaya, A., Shliaha, P. V., Grinev, V. V., Shlyueva, D., Damhofer, H., Koche, R., Gorshkov, V., Kovalchuk, S., Zhan, Y., Rodriguez, K. L., Johnstone, A. L., Keogh, M. C., Hendrickson, R. C., Jensen, O. N., & Helin, K. (2021). Complex-dependent histone acetyltransferase activity of KAT8 determines its role in transcription and cellular homeostasis. *Molecular Cell*, *81*(8), 1749-1765.e8. https://doi.org/10.1016/j.molcel.2021.02.012

Rahman, I., Marwick, J., & Kirkham, P. (2004). Redox modulation of chromatin remodeling: Impact on histone acetylation and deacetylation, NF-κB and pro-inflammatory gene expression. *Biochemical Pharmacology*, *68*(6), 1255–1267. https://doi.org/10.1016/j.bcp.2004.05.042

Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, *42*(W1), 187–191. https://doi.org/10.1093/nar/gku365

Roeder, R. G., & Rutter, W. J. (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature*, *224*(5216), 234–237. https://doi.org/10.1038/224234a0

Rowland, E. A., Snowden, C. K., & Cristea, I. M. (2018). Protein lipoylation: an evolutionarily conserved metabolic regulator of health and disease. *Current Opinion in Chemical Biology*, *42*, 76–85. https://doi.org/10.1016/j.cbpa.2017.11.003

Rust, M. J., Bates, M., & Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, *3*(10), 793–795. https://doi.org/10.1038/nmeth929

Sabari, B. R., Dall'Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., Abraham, B. J., Hannett, N. M., Zamudio, A. V., Manteiga, J. C., Li, C. H., Guo, Y. E., Day, D. S., Schuijers, J., Vasile, E., Malik, S., Hnisz, D., Lee, T. I., Cisse, I. I., … Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, *361*(6400). https://doi.org/10.1126/science.aar3958

Sabari, B. R., Dall'Agnese, A., & Young, R. A. (2020). Biomolecular Condensates in the Nucleus. *Trends in Biochemical Sciences*, *45*(11), 961–977. https://doi.org/10.1016/j.tibs.2020.06.007

Sanderson, S. M., Gao, X., Dai, Z., & Locasale, J. W. (2019). Methionine metabolism in health and cancer: a nexus of diet and precision medicine. *Nature Reviews Cancer*, *19*(11), 625–637. https://doi.org/10.1038/s41568-019-0187-8

Schneider, R., & Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes and Development*, *21*(23), 3027–3043. https://doi.org/10.1101/gad.1604607

Schuster, S., Fell, D. A., & Dandekar, T. (2000). *A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks*. *18*(March), 1–7. papers2://publication/uuid/00A659FC-63C2-438F-880F-F7D7070A1E4B

Sdelci, S., Rendeiro, A. F., Rathert, P., You, W., Lin, J. M. G., Ringler, A., Hofstätter, G., Moll, H. P., Gürtl, B., Farlik, M., Schick, S., Klepsch, F., Oldach, M., Buphamalai, P., Schischlik, F., Májek, P., Parapatics, K., Schmidl, C., Schuster, M., … Kubicek, S. (2019). MTHFD1 interaction with BRD4 links folate metabolism to transcriptional regulation. *Nature Genetics*, *51*(6), 990–998. https://doi.org/10.1038/s41588-019-0413-z

Sekar, R. B., & Periasamy, A. (2003). Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations. *Journal of Cell Biology*, *160*(5), 629–633. https://doi.org/10.1083/jcb.200210140

Selhub, J. (1999). Homocysteine metabolism. *Annual Review of Nutrition*, *19*, 217–246. https://doi.org/10.1248/yakushi.127.1579

Serfling, E., Jasin, M., & Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. *Trends in Genetics*, *1*(C), 224–230. https://doi.org/10.1016/0168-9525(85)90088-5

Shahbazi, M. N., & Zernicka-Goetz, M. (2018). Deconstructing and reconstructing the mouse and human early embryo. *Nature Cell Biology*, *20*(8), 878–887. https://doi.org/10.1038/s41556-018-0144-x

Shi, L., & Tu, B. P. (2015). Acetyl-CoA and the regulation of metabolism: Mechanisms and consequences. *Current Opinion in Cell Biology*, *33*, 125–131. https://doi.org/10.1016/j.ceb.2015.02.003

Sivanand, S., Rhoades, S., Jiang, Q., Lee, J. V., Benci, J., Zhang, J., Yuan, S., Viney, I., Zhao, S., Carrer, A., Bennett, M. J., Minn, A. J., Weljie, A. M., Greenberg, R. A., & Wellen, K. E. (2017). Nuclear Acetyl-CoA Production by ACLY Promotes Homologous Recombination. *Molecular Cell*, *67*(2), 252-265.e6. https://doi.org/10.1016/j.molcel.2017.06.008

Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews Molecular Cell Biology*, *19*(4), 262–274. https://doi.org/10.1038/nrm.2017.115

Srivastava, S., Gajwani, P., Jousma, J., Miyamoto, H., Kwon, Y., Jana, A., Toth, P. T., Yan, G., Ong, S. G., & Rehman, J. (2023). Nuclear translocation of mitochondrial dehydrogenases as an adaptive cardioprotective mechanism. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-40084-5

Sterner, D. E., & Berger, S. L. (2000). Acetylation of Histones and Transcription-Related Factors. *Microbiology and Molecular Biology Reviews*, *64*(2), 435–459. https://doi.org/10.1128/mmbr.64.2.435-459.2000

Suganuma, T., & Workman, J. L. (2008). Crosstalk among Histone Modifications. *Cell*, *135*(4), 604–607. https://doi.org/10.1016/j.cell.2008.10.036

Sun, F., Sun, T., Kronenberg, M., Tan, X., Huang, C., & Carey, M. F. (2021). The Pol II preinitiation complex (PIC) influences Mediator binding but not promoter-enhancer looping. *Genes and Development*, *35*(15–16), 1175–1189. https://doi.org/10.1101/gad.348471.121

Sun, L., Fu, X., Ma, G., & Hutchins, A. P. (2021). Chromatin and Epigenetic Rearrangements in Embryonic Stem Cell Fate Transitions. *Frontiers in Cell and Developmental Biology*, *9*(February), 1–20. https://doi.org/10.3389/fcell.2021.637309

Sutendra, G., Kinnaird, A., Dromparis, P., Paulin, R., Stenson, T. H., Haromy, A., Hashimoto, K., Zhang, N., Flaim, E., & Michelakis, E. D. (2014). A nuclear pyruvate dehydrogenase complex is important for the generation of Acetyl-CoA and histone acetylation. *Cell*, *158*(1), 84–97. https://doi.org/10.1016/j.cell.2014.04.046

Takahashi, H., McCaffery, J. M., Irizarry, R. A., & Boeke, J. D. (2006). Nucleocytosolic Acetyl-Coenzyme A Synthetase Is Required for Histone Acetylation and Global Transcription. *Molecular Cell*, *23*(2), 207–217. https://doi.org/10.1016/j.molcel.2006.05.040

Tanaka, S., Kunath, T., & Hadjantonakis, A. (1998). Promotion of Trophoblast Stem Cell Proliferation by FGF4. *Science (New York, N.Y.)*, *282*(December), 2072–2075.

Taylor, G. C. A., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M. M., & Bickmore, W. A. (2013). H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Research*, *23*(12), 2053–2065. https://doi.org/10.1101/gr.155028.113

Taylor, J. P., Taye, A. A., Campbell, C., Kazemi-Esfarjani, P., Fischbeck, K. H., & Min, K. T. (2003). Aberrant histone acetylation, altered transcription, and retinal degeneration in a Drosophila model of polyglutamine disease are rescued by CREB-binding protein. *Genes and Development*, *17*(12), 1463–1468. https://doi.org/10.1101/gad.1087503

Torres-Padilla, M. E., & Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: A stochastic advantage. *Development (Cambridge)*, *141*(11), 2173–2181. https://doi.org/10.1242/dev.102624

Tsai, K. L., Yu, X., Gopalan, S., Chao, T. C., Zhang, Y., Florens, L., Washburn, M. P., Murakami, K., Conaway, R. C., Conaway, J. W., & Asturias, F. J. (2017). Mediator structure and rearrangements required for holoenzyme formation. *Nature*, *544*(7649), 196–201. https://doi.org/10.1038/nature21393

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., Rossen, J., Joesch-Cohen, L., Humeidi, R., Spangler, R. D., Eaton, J. K., Frenkel, E., Kocak, M., Corsello, S. M., Lutsenko, S., Kanarek, N., Santagata, S., & Golub, T. R. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science*, *375*(6586), 1254–1261. https://doi.org/10.1126/science.abf0529

Uthe, H., Vanselow, J. T., & Schlosser, A. (2017a). Proteomic Analysis of the Mediator Complex Interactome in Saccharomyces cerevisiae. *Scientific Reports*, *7*(January), 1–11. https://doi.org/10.1038/srep43584

Uthe, H., Vanselow, J. T., & Schlosser, A. (2017b). Proteomic Analysis of the Mediator Complex Interactome in Saccharomyces cerevisiae. *Scientific Reports*, *7*(January), 1–11. https://doi.org/10.1038/srep43584

Verger, A., Monté, D., & Villeret, V. (2019). Twenty years of Mediator complex structural studies. *Biochemical Society Transactions*, *47*(1), 399–410. https://doi.org/10.1042/BST20180608

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J., & Grunstein, M. (2002). Histone Acetylation Regulates the Time of Replication Origin Firing. *Molecular Cell*, *10*(5), 1223–1233. http://www.ncbi.nlm.nih.gov/pubmed/12453428

Weinert, B. T., Narita, T., Satpathy, S., Srinivasan, B., Hansen, B. K., Schölz, C., Hamilton, W. B., Zucconi, B. E., Wang, W. W., Liu, W. R., Brickman, J. M., Kesicki, E. A., Lai, A., Bromberg, K. D., Cole, P. A., & Choudhary, C. (2018). Time-Resolved Analysis Reveals Rapid Dynamics and Broad Scope of the CBP/p300 Acetylome. *Cell*, *174*(1), 231-244.e12. https://doi.org/10.1016/j.cell.2018.04.033

Wellen, K. E., Hatzivassiliou, G., Sachdeva, U. M., Bui, T. V., Cross, J. R., & Thompson, C. B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. *Science*, *324*(5930), 1076–1080. https://doi.org/10.1126/science.1164097

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H.,

Rahl, P. B., Lee, T. I., & Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, *153*(2), 307–319. https://doi.org/10.1016/j.cell.2013.03.035

Wright, R. H. G., Lioutas, A., Dily, F. Le, Soronellas, D., Pohl, A., Bonet, J., Nacht, A. S., Samino, S., Font-Mateu, J., Vicent, G. P., Wierer, M., Trabado, M. A., Schelhorn, C., Carolis, C., Macias, M. J., Yanes, O., Oliva, B., & Beato, M. (2016). ADP-ribose-derived nuclear ATP synthesis by NUDIX5 is required for chromatin remodeling. *Science*, *352*(6290), 1221–1225. https://doi.org/10.1126/science.aad9335

Wu, H., Deng, Y., Feng, Y., Long, D., Ma, K., Wang, X., Zhao, M., Lu, L., & Lu, Q. (2018). Epigenetic regulation in B-cell maturation and its dysregulation in autoimmunity. *Cellular and Molecular Immunology*, *15*(7), 676–684. https://doi.org/10.1038/cmi.2017.133

Yamada, T., Mizuno, K. I., Hirota, K., Kon, N., Wahls, W. P., Hartsuiker, E., Murofushi, H., Shibata, T., & Ohta, K. (2004). Roles of histone acetylation and chromatin remodeling factor in a meiotic recombination hotspot. *EMBO Journal*, *23*(8), 1792–1803. https://doi.org/10.1038/sj.emboj.7600138

Yankulov, K., Todorov, I., Romanowski, P., Licatalosi, D., Cilli, K., McCracken, S., Laskey, R., & Bentley, D. L. (1999). MCM Proteins Are Associated with RNA Polymerase II Holoenzyme. *Molecular and Cellular Biology*, *19*(9), 6154–6163. https://doi.org/10.1128/mcb.19.9.6154

Yeo, J. C., Jiang, J., Tan, Z. Y., Yim, G. R., Ng, J. H., Göke, J., Kraus, P., Liang, H., Gonzales, K. A. U., Chong, H. C., Tan, C. P., Lim, Y. S., Tan, N. S., Lufkin, T., & Ng, H. H. (2014). Klf2 is an essential factor that sustains ground state pluripotency. *Cell Stem Cell*, *14*(6), 864–872. https://doi.org/10.1016/j.stem.2014.04.015

Ying, Q. L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., & Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature*, *453*(7194), 519–523. https://doi.org/10.1038/nature06968

Yu, M., & Ren, B. (2017). The three-dimensional organization of mammalian genomes. *Annual Review of Cell and Developmental Biology*, *33*, 265–289. https://doi.org/10.1146/annurev-cellbio-100616-060531

Zhang, T., Zhang, Z., Dong, Q., Xiong, J., & Zhu, B. (2020). Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biology*, *21*(1), 1–7. https://doi.org/10.1186/s13059-020-01957-w

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nussbaum, C., Myers, R. M., Brown, M., Li, W., & Shirley, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biology*, *9*(9). https://doi.org/10.1186/gb-2008-9-9-r137

Zhao, W., Xu, Y., Wang, Y., Gao, D., King, J., Xu, Y., & Liang, F. Sen. (2021). Investigating crosstalk between H3K27 acetylation and H3K4 trimethylation in CRISPR/dCas-based epigenome editing and gene activation. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-95398-5

Zhong, Y., Li, X., Yu, D., Li, X., Li, Y., Long, Y., Yuan, Y., Ji, Z., Zhang, M., Wen, J. G., Nesland, J. M., & Suo, Z. (2015). Application of mitochondrial pyruvate carrier blocker UK5099 creates metabolic reprogram and greater stem-like properties in LnCap prostate cancer cells in vitro. *Oncotarget*, *6*(35), 37758–37769. https://doi.org/10.18632/oncotarget.5386

Annex

Published and in preparation articles directly related to this PhD thesis:

- Lynch CJ, Bernad R, **Calvo I**, Serrano M. **Manipulating the Mediator complex to induce naïve pluripotency**. *Exp Cell Res*. 2020 Oct 15;395(2):112215. doi: 10.1016/j.yexcr.2020.112215. Epub 2020 Aug 6. PMID: 32771524; PMCID: PMC7584500.

- Lynch CJ, Bernad R, Martínez-Val A, Shahbazi MN, Nóbrega-Pereira S, **Calvo I**, Blanco-Aparicio C, Tarantino C, Garreta E, Richart-Ginés L, Alcazar N, Graña-Castro O, Gómez-Lopez G, Aksoy I, Muñoz-Martín M, Martinez S, Ortega S, Prieto S, Simboeck E, Camasses A, Stephan-Otto Attolini C, Fernandez AF, Sierra MI, Fraga MF, Pastor J, Fisher D, Montserrat N, Savatier P, Muñoz J, Zernicka-Goetz M, Serrano M. **Global hyperactivation of enhancers stabilizes human and mouse naive pluripotency through inhibition of CDK8/19 Mediator kinases**. *Nat Cell Biol*. 2020 Oct;22(10):1223-1238. doi: 10.1038/s41556-020-0573-1. Epub 2020 Sep 28. PMID: 32989249.

- Martinez-Val A, Lynch CJ, **Calvo I**, Ximénez-Embún P, Garcia F, Zarzuela E, Serrano M, Munoz J. **Dissection of two routes to naïve pluripotency using different kinase inhibitors.** *Nat Commun*. 2021 Mar 25;12(1):1863. doi: 10.1038/s41467-021-22181-5. PMID: 33767186; PMCID: PMC7994667.

- The article that contains the data presented in this thesis is currently in preparation.

Other published articles:

- Lynch CJ, Bernad R, **Calvo I**, Nóbrega-Pereira S, Ruiz S, Ibarz N, Martinez-Val A, Graña-Castro O, Gómez-López G, Andrés-León E, Espinosa Angarica V, Del Sol A, Ortega S, Fernandez-Capetillo O, Rojo E, Munoz J, Serrano M. **The RNA Polymerase II Factor RPAP1 Is Critical for Mediator-Driven Transcription and Cell Identity**. *Cell Rep*. 2018 Jan 9;22(2):396-410. doi:

10.1016/j.celrep.2017.12.062. PMID: 29320736; PMCID: PMC5775503.

- Batlle C, **Calvo I**, Iglesias V, J Lynch C, Gil-Garcia M, Serrano M, Ventura S. **MED15 prion-like domain forms a coiled-coil responsible for its amyloid conversion and propagation**. *Commun Biol*. 2021 Mar 26;4(1):414. doi: 10.1038/s42003-021-01930-8. PMID: 33772081; PMCID: PMC7997880.

Experimental Cell Research 395 (2020) 112215

Contents lists available at ScienceDirect

Experimental Cell Research

journal homepage: www.elsevier.com/locate/yexcr

Manipulating the Mediator complex to induce naïve pluripotency

Cian J. Lynch [a,**], Raquel Bernad [a], Isabel Calvo [a], Manuel Serrano [a,b,*]

[a] Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
[b] Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, 08010, Spain

ARTICLE INFO

Keywords:
Pluripotency
Naïve
Primed
Mediator
Super-enhancers
Phase separation
CDK8
MEK
2i
Demethylation
Embryonic stem
Reprogramming
RNA polymerase

ABSTRACT

Human naïve pluripotent stem cells (PSCs) represent an optimal homogenous starting point for molecular interventions and differentiation strategies. This is in contrast to the standard primed PSCs which fluctuate in identity and are transcriptionally heterogeneous. However, despite many efforts, the maintenance and expansion of human naïve PSCs remains a challenge. Here, we discuss our recent strategy for the stabilization of human PSC in the naïve state based on the use of a single chemical inhibitor of the related kinases CDK8 and CDK19. These kinases phosphorylate and negatively regulate the multiprotein Mediator complex, which is critical for enhancer-driven recruitment of RNA Pol II. The net effect of CDK8/19 inhibition is a global stimulation of enhancers, which in turn reinforces transcriptional programs including those related to cellular identity. In the case of pluripotent cells, the presence of CDK8/19i efficiently stabilizes the naïve state. Importantly, in contrast to previous chemical methods to induced the naïve state based on the inhibition of the FGF-MEK-ERK pathway, CDK8/19i-naïve human PSCs are chromosomally stable and retain developmental potential after long-term expansion. We suggest this could be related to the fact that CDK8/19 inhibition does not induce DNA demethylation. These principles may apply to other fate decisions.

1. A new approach to human pluripotency

Human embryonic pluripotent stem cells (PSCs) provide a challenging parallel to our understanding of pluripotency based in rodents. The extent to which human PSCs can be captured in a stable naïve state, its similarity to rodent naïve pluripotency, and even its clinical usefulness, remain in question [1–3]. Inhibition of the FGF-MEK-ERK signaling pathway is central to naïve pluripotency across multiple mammalian species, however the application of this in culture media appears to require species-specific adaptations to avoid genomic instability [4–7]. Understanding the mechanisms by which MEK-inhibition leads to the transition from primed to naïve pluripotency remains in progress.

The Mediator complex operates at enhancers and is a pivotal orchestrator of the transcriptional program of cell identity [8,9] (Fig. 1). Moreover, several data suggest that Mediator function may lie downstream of the MEK-ERK pathway (see below). Indeed, we recently identified an important role for the Mediator complex in stabilizing naïve pluripotency both in mouse and human PSCs, which recapitulates many aspects of MEK-inhibition [10]. In this review, we introduce the role of Mediator in cell identity, in particular during pluripotency, and we summarize our recent report of a pharmacological strategy for the manipulation of Mediator that stabilizes the naïve state in mouse, human, and non-human primate PSCs [10]. We also discuss the implications for other states of cellular plasticity beyond pluripotency, and its application in directing the resolution of cellular fate decisions.

2. Enhancers, super-enhancers, and cell identity

Each cell-type contains a unique repertoire of active enhancer complexes at specific DNA regions, which arise by high concentration of lineage-specific transcription factors [11–14]. Collectively, the transcription factors, associated chromatin regulators, and the epigentic marks that they generate, serve as the platform to recruit a single large multi-protein complex known as Mediator [8,9,15] (Fig. 1). In a sense, the diverse combinatorial information input present at enhancers is reduced into a single output, namely, the Mediator complex, which thereby can be considered the universal transducer of enhancers. The main function of Mediator is, in turn, the recruitment of RNA polymerase II (RNA Pol II) to nearby and distant promoters, thus having a major contribution to the transcriptional program characteristic of each

* Corresponding author. Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Barcelona, Spain.
** Corresponding author.
E-mail addresses: cian.lynch@irbbarcelona.org (C.J. Lynch), manuel.serrano@irbbarcelona.org (M. Serrano).

https://doi.org/10.1016/j.yexcr.2020.112215
Received 28 May 2020; Received in revised form 8 July 2020; Accepted 1 August 2020
Available online 6 August 2020

cell type [8,9,13,16] (Fig. 1).

Super-enhancers (SEs) constitute a relatively novel concept that refers to a small fraction of unusually large and powerful enhancers [12,13]. SE's are multipartite, as they arise by coallescing clusters of smaller typical enhancer constituents in close proximity, and these large aggregates display emergent properties that distinguish them functionally from the smaller typical enhancers (TEs) [17] (Fig. 1). In particular, SEs are extraordinarily potent in driving transcription and they rely on phase separation to confer their stability and flexibility to change [17]. Interestingly, the most influential genes determining cell identity are often found under the control of SEs [11–13]. SEs drive high expression of master transcription factors, which in turn, enrich within the enhancer loci, completing a positive-feedback loop and establishing the regulatory network that maintains cell identity [11–13]. In agreement with its critical role in recruiting RNA Pol II, Mediator is highly abundant at SEs [13] and therefore, we reasoned that it could be an actionable point to manipulate SEs and cell identity (Fig. 1). Below, we summarize the current understanding of Mediator function, and regulation by CDK8 and CDK19 kinases (abbreviated as CDK8/19).

3. Mediator: a bridge between enhancers and promoters

There have been great advances in recent years regarding the structure and mechanistic functioning of Mediator [8,9,15,18–20]. The 30 subunits of Mediator are organized in four general domains: the "head", "middle", and "tail" domains that constitute "core-Mediator", plus a fourth accessory domain known as the "CDK8/19-module" (Fig. 1). The Mediator tail-domain interacts with the enhancer chromatin, including transcription factors and cofactors, while the middle- and head-domains form contacts with RNA Pol II and the pre-initiation complex at target promoters [8,18,19]. While the many subunits of Mediator can undergo extensive structural re-arrangements, the CDK8/19-module contains the only enzymatic activity of Mediator, namely the kinase CDK8 or its highly similar, but poorly studied, paralog CDK19 [9,21,22]. Completing the kinase module are: cyclin C (CCNC), and subunits MED12 and MED13. CDK8/19 activity appears restricted by its requirement for proper quaternary structure [23]. For full activity, CDK8/19 must associate simultaneously with cyclin C and with MED12, the latter embraces CDK8 allowing it to attain proper opening of its T-loop [23]. Evidence suggests that only CDK8, or CDK19, can occupy Mediator at one time, and the same is true for the paralog subunits MED12L and MED13L (reviewed elsewhere) [8,9,22]. Collectively, these may represent alternate combinations of the CDK8/19-module, providing opportunity for subtle modulation of Mediator function.

The Mediator complex must associate with the many different transcription factors that define an enhancer locus. This constitutes one arm of the Mediator bridge at the enhancer. Recent evidence hints at how this may be achieved. Mediator subunits in its tail sub-module have been shown to engage in a high number of low specificity and weak interactions via phase separation, thereby achieving a "fuzzy" interaction between Mediator and the activation domains of many different transcription factors [24–27]. The second arm of the Mediator bridge extends to interact with the pre-initiation complex, inducing the recruitment of RNA Pol II at target promoters [20] (Fig. 1). Comprehensive cryo-EM studies and *in vitro* biochemical assays have collectively revealed how Mediator undergoes extensive structural re-arrangements to achieve this [8,9,18,19]. Notably, the CDK8/19-kinase module plays a key role, associating with core-Mediator and repressing RNA Pol II recruitment [8,9,18,19] (see below). In addition to its structural aspects, the Mediator bridge is dynamic and programmable, integrating many upstream signals [8,28]. Precisely how upstream signals modulate Mediator function is only

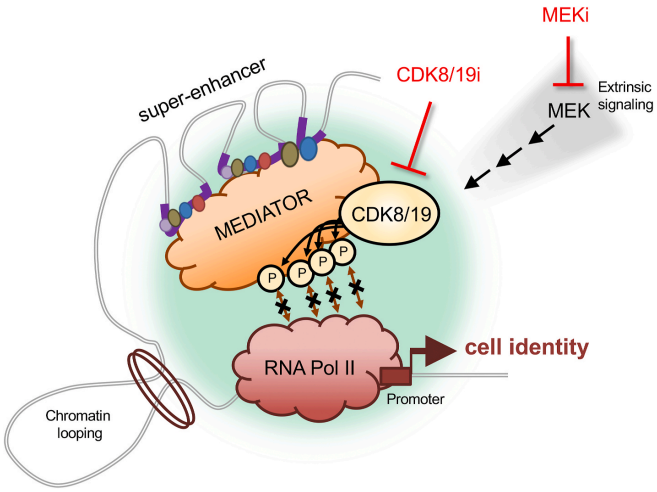


Fig. 1. The Mediator complex integrates multiple signals and bridges enahcers to promoters to regulate pluripotent cell identity. Mediator acts as a central hub which links enhancers and target promoters, integrates combinatorial signals, and recruits RNA Pol II, to guide the transcriptional program of cell identity. The core transcriptional machinery is shown simplified, within a phase-separated transcriptional condensate (light green, membrane-less region). The largest enhancers, known as super-enhancers emerge by coalescing multiple smaller constituent enhancer regions (purple thick lines), each enriched in lineage-specifying transcription factors, histone marks, and chromatin regulators (indicated by small coloured circles). Mediator associates with these many factors and receives upstream signaling inputs from pathways, including MEK-ERK. Mediator recruits RNA Pol II to the pre-initiation complex, a function which can be hindered by the CDK8/19 kinase-module, via structural or kinase-dependent activities of CDK8/19. In PSCs, the 2i inhibitor cocktail is known to rapidly alter the transcriptional program to promote naïve pluripotency, of which, MEK-inhibition is the key feature. However, how MEK-inhibition resets the transcriptional machinery to the naïve program remains to be clarified. In PSCs, we find that MEK-inhibition leads to global up-regulation of enhancer activity via increased RNA Pol II recruitment. We find that the reorganization of the transcriptional machinery by MEK inhibition operates largely via controling the activity of CDK8/19 downstream. In this way, both treatments induce a highly overlapping set of phospho-changes focused on the transcriptional machinery, hyper-activating enhancers and Mediator, triggering increased RNA Pol II recruitment, and promoting the transcriptional program of naïve pluripotency.

recently emerging. Acetylation [29] and phosphorylation [10,30–33] of multiple sites and subunits of Mediator have been detected, implicating p300 acetyl-transferase activity, and the MEK-ERK pathway, as major inputs.

4. CDK8/19: the sub-module of Mediator that represses RNA Pol II recruitment

The CDK8/19-kinase module (CDK8/19-CCNC-MED12/L-MED13/L) associates with core-Mediator and plays a central role in the regulation of how Mediator recruits RNA Pol II during each transcription cycle [8,9,22]. The detailed mechanisms of how CDK8/19 operates during the transcription cycle remain to be fully understood [22,34], and the relative contribution of CDK8/19 kinase activity versus its kinase-independent structural effects are unclear. However, we and others have observed that the ultimate outcome of CDK8/19 kinase-inhibition is increased recruitment of RNA Pol II by Mediator, and stimulation of transcription [10,35]. In the context of leukemia, this global hyperactivation of Mediator function results in cancer cell death [35]. Intriguingly, in pluripotent stem cells, we found that CDK8/19-inhibition and global hyperactivation of Mediator did not result in transcriptional dysfunction, instead, a coherent shift occurred in the transcriptional program, transitioning from one cell identity to another [10]. This raises the possibility of pharmacologically-modulating cell identity of non-cancer cells in a non-deleterious manner.

As mentioned above, currently it is not possible to fully understand the mechanism of CDK8/19 during the transcription cycle. Nevertheless, it is useful to consider the temporal order of events when interpreting the literature, and we summarize the role of CDK8/19 below in this order. CDK8/19 can sterically hinder the initial approach of Pol II to the Mediator complex based on several data and crystal structures [18,19,36], and indeed, this was how CDK8 was first identified in yeast [37]. Thus, CDK8/19 must transiently vacate its initial position to permit Mediator-Pol II interaction [8,9,18,19,38]. In yeast, this process involves CDK11-catalyzed phosphorylation of Mediator subunits MED4 and MED27 [39]. Also, we and others have found that CDK8/19 can phosphorylate itself and several other Mediator subunits [10,30,[36] 37], raising the possibility that the kinase activity of CDK8/19 may somehow influence its structural role hindering the approach of Pol II to the Mediator complex. In reciprocal, extensive Mediator structural rearrangements re-position CDK8/19 with respect to RNA Pol II [18,19,23].

Later in the transcription cycle, following recruitment of RNA Pol II to Mediator, the pre-initiation complex is completed, and CDK8/19-Mediator undergoes further structural re-arrangements to participate in the release of RNA Pol II into productive elongation [8,9,18,19]. CDK8/19 has been reported to phosphorylate RNA Pol II directly *in vitro* [36,37], and to affect Pol II-CTD phosphorylation *in vivo* [40–45]. Also, CDK8/19 is reported to phosphorylate and orchestrate the function of positive and negative elongations factors including NELF, DSIF, TFIIH and the super-elongation complex [8,9]. However, we suggest that these reported roles of CDK8/19-mediated phosphorylation in elongation may be to some extent redundant, gene-specific, and/or compensated by other kinases such as CDK7 and CDK9. In this regard, we and others have used potent CDK8/19 small molecule inhibitors where transcription is modulated but can still proceed and cells are viable [10,30,35,43–46]. Here, it is notable that while the other transcriptional CDKs, namely CDK7 and CDK9, play general roles positively promoting transcriptional elongation, the role of CDK8/19 appears more nuanced. Thus, while chemical inhibition of CDK7 and CDK9 are lethal, abolishing transcription [47–49], CDK8/19 chemical inhibition appears to modulate the transcriptional program, triggering global hyperactivation of Mediator, and increased Mediator-RNA Pol II interaction [10].

In summary, CDK8/19 has an overall effect of repressing RNA Pol II recruitment and, thereby, CDK8/19 kinase inhibition increases recruitment of RNA Pol II by Mediator and stimulates global hyperactivation of transcription of enhancer target genes [10,35].

5. Mediator, intrinsically disordered regions, and transcriptional condensates

RNA Pol II is known to participate in discrete membrane-less protein aggregates previously described as transcription factories [50–52]. More recently, studies have elucidated how these factors efficiently co-segregate from the nuclear milieu based on a shared propensity for phase separation, leading to the term transcriptional condensates [24,25,50–53]. Mediator and several other protein elements of the transcriptional machinery contain intrinsically-disordered regions (IDRs) that confer their phase separation [24,25,50,52–54]. Thus, IDRs act as biophysical addresses to localize each protein, and they contribute an essential structural component to stabilize the transcriptional condensate. In particular, approximately half of Mediator's 30-subunits contain conserved IDRs, including CDK8 and CDK19 [54]. This shift in our understanding has been essential to explain the efficient assembly and structural stability of transcriptional condensates, but also their dynamic flexibility to changes in signaling. Specifically, transcriptional condensates possess vulnerability to perturbation via sharp transitions in their phase separation [17], a regulatory feature conferring dual capacity to stably drive high gene expression, or dissolve rapidly, as required. The ability of transcriptional condensates to switch between formation or dissolution may be essential to rapid transcriptional changes [17,22,55]. Pluripotent cells contain a discrete number of particularly large condensates containing Mediator and RNA Pol II that are thought to correspond to SEs [45,47,48]. Post-translational modifications within protein IDRs have been shown to control transitions in phase separation [53]. Thus, given the central role of the CDK8/19-module for Mediator, we speculate that its kinase activity could affect the entry or exit of proteins into the condensates, or even the formation and dissolution of the condensates themselves. In support of this, we recently identified CDK8/19-dependent phosphosites in the Mediator subunit Med1-IDR domain in mouse PSCs [10], a particularly large IDR previously shown to play a crucial role in transcriptional condensates in mouse PSCs [50].

6. The naïve and primed PSC states reflect developmental stages

Pluripotent stem cells (PSCs) transition between cell states *in vitro* which faithfully reflect developmental stages in the early embryo [1,2,6,7] (Fig. 2A–C). Culture conditions to shield mouse PSCs from extra-cellular differentiation stimuli involve chemical inhibition of MEK and GSK3 kinases with a two inhibitor cocktail known as "2i" [6,56]. Mouse PSCs cultured in 2i (referred to as "2i-naïve" cells) phenocopy the stable and homogenous state of undifferentiated naïve pluripotency which exists transiently *in vivo* around the E4.5 mouse pre-implantation embryo epiblast (Fig. 2A–C). In contrast, culture of PSCs with standard serum/LIF media, in the absence of 2i, triggers a shift in cell identity towards post-implantation epiblast, also known as the primed state, analogous to the mouse ~ E5.5–6.5 stage. It is important to clarify that in serum/LIF cells are more properly described as heterogenous, fluctuating between the naïve and primed state [6]. Full primed identity is achieved in FGF/Activin media and this recapitulates the cellular identity of the E6.5 mouse embryo epiblast [1,6,7]. Importantly, PSCs can be maintained during long-term passage in each of these developmental stages. Also, by simply changing the media to alter the signaling inputs to the transcriptional machinery, cells can be directed forward or reverse in developmental identity. In this way, PSCs provide a prototypical model of cellular plasticity, whose transcriptional program can be stabilized, extinguished or re-captured. Notably however, much remains unclear regarding how changes in extrinsic stimuli reset the transcriptional machinery to a new program.

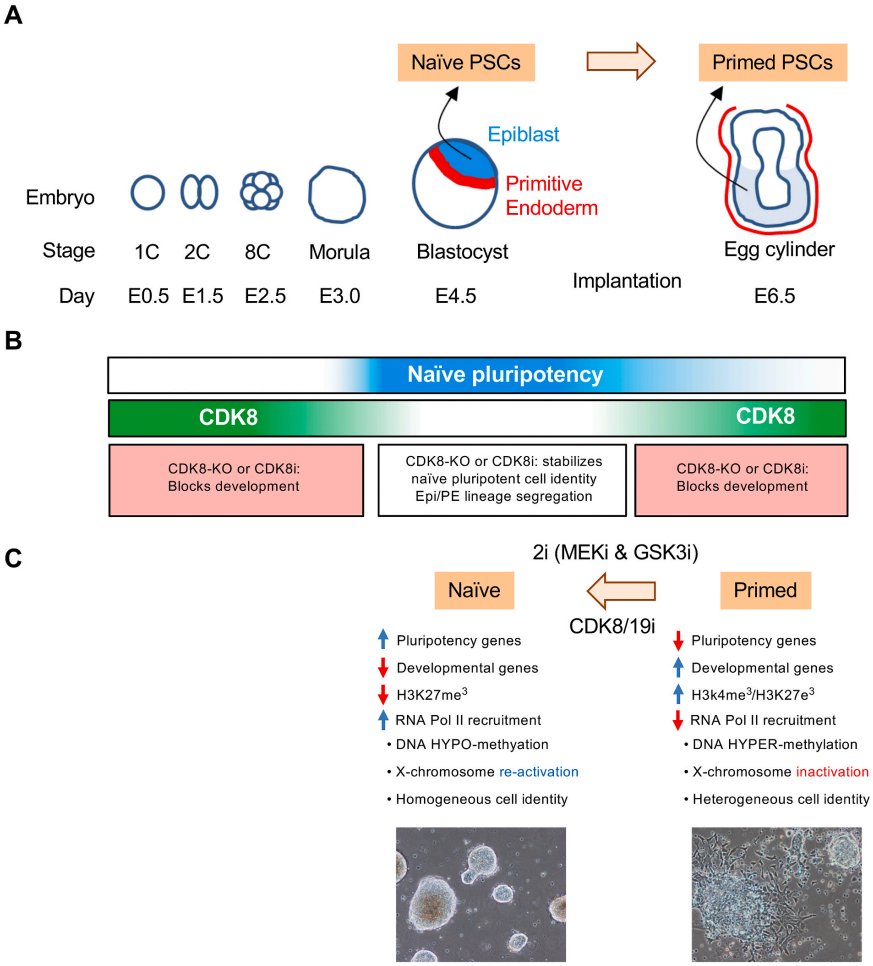


Fig. 2. A summary of CDK8/19 function in early embryo cell identity transitions. **(A)** Schematic overview of mouse early embryo development. Naïve PSCs can be isolated from the E4.5 epiblast, and they can retain this cell identity homogenously *in vitro*. Primed PSCs can be isolated from the E6.5 epiblast, and they can be maintained in the fluctuating and heterogenous primed state *in vitro*. The 2i inhibitor cocktail can stabilize naïve identity, and promote primed PSCs to transition into the naïve state. **(B)** Similar to small molecule 2i-treatment of PSCs *in vitro*, in the embryo, MEK-signaling is naturally repressed by down-regulation of the FGF receptor. We find that CDK8/19 kinase inhibition, by genetic or pharmacological methods, can phenocopy these effects of 2i: stabilizing naïve identity, and promoting primed cells to transition into the naïve state. Interestingly, we observe that in the embryo, CDK8 activity appears to be naturally repressed, with down-regulation of CDK8 levels and decreased nuclear availability of its essential binding partner Cyclin C. The table below summarizes the developmental phenotypes observed by genetic- or pharmacological-inhibition of CDK8 in the mouse. **(C)** A summary of the molecular and morphological differences reported for naïve and primed pluripotent states.

7. The transcriptional brain during cellular fate decisions in pluripotency

Here we consider the transcriptional decision-making process. Cell identity is the aggregate outcome of the entire set of active enhancers in a cell at any one moment. During identity transitions, enhancers are activated and decommissioned individually. Thus, the events at each enhancer can be rapid (consistent with the sharp transitions conferred by phase separation), while the aggregate shift in cell identity (all active enhancers) appears relatively slower and incremental. We use as a model the developmental window of naïve-to-primed pluripotency, outlined above. Extensive analyses over the past ~15 years have shown how this path in embryos is followed in close parallel by PSCs *in vitro* [1,2,6,7]. Pluripotency first arises at ~ E3.5–4.5 in mice, and at this initial

boundary, cellular identity is very well defined in the naïve state, where cells are effectively deaf to developmental signals and lack any trace of differentiation programs. This is analogous to 2i-naïve PSCs *in vitro* [1,2,6,7,56,57]. Subsequently, cells progress through a gradient of intermediate pluripotency stages, where they gradually become receptive to differentiation stimuli, collectively referred to as being "primed" for development [1,2,6,7]. This period reflects mouse embryo epiblast development from E4.5–5.5. A multitude of studies have captured and defined specific intermediate states with PSCs *in vitro*, referring sequentially, for example, to formative, primed-like, epiblast-like, or rosette states [58] of pluripotency (thoroughly reviewed elsewhere [1,2,6,59]. However, now, and with the benefit of hindsight, we favour the idea that a continuum of pluripotency states exists, marked by incremental changes in gene expression and other detailed molecular characteristics.

Ultimately, around E5.5–6.5, mouse embryo epiblast cells exit from pluripotency, and choose between three competing transcriptional programs for lineage-specification, into one of the embryonic germ layers: ectoderm, mesoderm, and endoderm [1,6,7,60] (Fig. 3). This decision-making process remains to be fully understood, however evidence suggests that while naïve cells possess a single homogenous undifferentiated transcriptional program, cells in the primed stages begin to receive simultaneous and competing signals drawing them to activate lineage-specific transcriptional programs of one of the 3 germ layers [61–66] (Fig. 3).

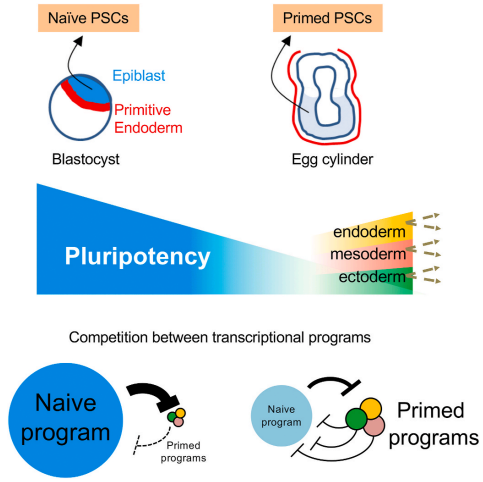


Fig. 3. Emergence and dissolution of transcriptional programs in the naïve to prime transition. A putative model of dominance and antagonism in transcriptional programs. Binary cell fate choices are common though development and regeneration. In the case of the naïve-primed developmental window of pluripotency, global hyperactivation of enhancers and Mediator-recruitment of RNA Pol II produces a coherent outcome: up-regulation and stabilization of naïve epiblast identity. Here, we suggest a putative mechanism. In primed identity, cells are plastic and heterogenous due to co-existence of lineage-specifying transcriptional programs for forward development into the 3 embryoinc germ layers, however, they are moderately expressed, antagonistic, and weakly established. Powerful naïve-specifiying transcription factors and super-enhancers have begun their decline, but are not yet decommissioned. Upon global hyperactivation of enhancers and Mediator-recruitment of RNA Pol II (by treatment with 2i or CDK8/19i), the naïve-specifiying transcription factors and super-enhancers become dominant, and suppress the nascent germ layer programs.

Crucially, lineage-specifying transcriptional programs attempt to reinforce themselves, and repress other competing programs [67–70]. The antagonistic and heterogenous behaviour of transcriptional programs within individual cells at this time is thought to represent a prototypical example of plasticity. Here, cells exist in a super-position, simultaneously expressing the early lineage-specifiers of opposing future fates [61–66]. Thus the cell population is also heterogenous; with individual cells expressing genes that suggest a moderate bias toward one or other fate. Such stochastic decision-making may have evolved to allow cells to sense and respond to diverse environmetal stimuli, ensuring that from one population of cells, three developmental fates emerge. *In vivo*, this is transient, and resolves in a few hours, with each cell choosing a single germ-layer program. *In vitro*, it is possible to capture this decision-making process, as a population of primed PSCs, with heterogenous and fluctuating gene expression programs, literally "priming" themselves to fully-activate ectoderm, mesoderm, or endoderm [1,2,6,7]. It is interesting to consider the events that occur at super-enhancers, and Mediator function, during this competition between transcriptional programs, as these processes may represent common features of many, if not all, transitions in cellular identity. For example, similar co-expression of competing lineage-specifiers has been observed in other fate decisions throughout development and regeneration [71,72]. Thus, the naïve-primed pluripotency window provides a model to elucidate the transcriptional basis of cell plasticity. The molecular lessons learned here are likely to apply to other fate decisions.

8. The role of Mediator in pluripotent cell identity

We recently hypothesized that manipulation of the transcriptional machinery during states of plasticity, might toggle cellular identity toward one path or another. We chose PSCs in the naïve-primed pluripotency window as our model of plasticity. As an actionable strategy, we began by targeting the transcriptional CDKs, CDK7, CDK8/19, and CDK9, since these kinases represent direct regulators of the transcriptional machinery, they are amenable to pharmacological inhibition, and indeed, several highly selective small molecule inhibitors exist [10]. We found that selective inhibition of CDK7 and CDK9 produced a general inhibition of transcription, and was ultimately deleterious, consistent with their known general roles in RNA Pol II transcription. In contrast, small molecule kinase inhibition of CDK8/19 (CDK8/19i) produced a striking and characteristic shift in mouse PSC identity, from primed to naïve.

We explored a possible mechanism, following the role of CDK8/19 as a transcriptional CDK, located within the Mediator complex [10]. Also, we compared the induction and stabilization of naïve pluripotency by CDK8/19i versus the well-characterized 2i inhibitor cocktail. 2i is highly-effective in mouse, but not in human PSCs [1,4,73,74], and the effects of 2i on the transcriptional machinery have not been explored. Importantly, we found that 2i-naïve and CDK8/19i-naïve mouse PSCs were highly similar according to their transcriptome and proteome [10]. Collectively, our data suggest that 2i and CDK8/19i rapidly induce a highly overlapping set of phospho-changes focused on the transcriptional machinery (Fig. 1). We also found that Mediator activity was increased in 2i and CDK8/19i, consistent with a repressive role of CDK8/19 in the ability of Mediator to recruit RNA Pol II. Thus 2i and CDK8/19i triggered enhancer hyperactivation, global increase in RNA Pol II recruitment to promoters, and resetting of gene expression. This included the upregulation of enhancer-derived RNAs (eRNAs), and the resetting of endogenous retroviral and repeat element expression. In sum, the ability of 2i and CDK8/19i to induce naïve features appears to originate from their common effect on Mediator and RNA Pol II transcriptional activity. This model is consistent with the concept that transitions in cell identity are driven by early reconfiguration of the active enhancer network, which resets the transcriptional machinery to the new program [14,67,70,75–77]. In summary, targeting Mediator through its kinase module selectively stabilizes an early pluripotent cell

identity, repressing differentiation, favoring self-renewal, and up-regulating pre-implantation naïve epiblast gene expression patterns.

As mentioned before, the induction of naïve identity using 2i or CDK8/19i treatments can stimulate Mediator function, which we detect by a global increase in RNA Pol II recruitment, global hyper-activation of existing mouse PSC enhancer loci, and upregulation of enhancer-driven transcription. We propose that this reinforces the pluripotency network underlying naïve PSC identity. In agreement with a recent report [78], we observe that in 2i, mouse naïve-specific enhancer activity is partially resistant to enhancer/Mediator destabilization by BRD4-inhibition. Importantly, this property can also be conferred by expression of CDK8-kinase dead mutant protein. This suggests a simple mechanism where removal of the inhibitory influence of CDK8/19, hyperactivates Mediator function at enhancers, and that this occurs similarly in 2i or via CDK8/19 inhibition. In support of global activation of super-enhancers in the naïve state, a recent study of chromatin looping has revealed that super-enhancers interact with more target promoters, and engage in more long-range interactions, during mouse naïve pluripotency compared to primed pluripotent cells [79], while furthermore, a state of global hyper-transcription has been suggested in PSCs [80,81]. Lastly, we note that the induction of naïve features in human PSCs was recently associated with a significant global increase in the enhancer mark H3K27ac, including at SEs [82]. These data may be consistent with our mechanistic insights, where CDK8/19 inhibition up-regulates enhancer activity, in particular at super-enhancers, thereby driving the stabilization of naïve pluripotency.

9. DNA methylation separates CDK8/19i from 2i

As outlined above, CDK8/19i recapitulates a significant proportion of 2i-associated effects on mouse naïve cell identity, however, global DNA hypomethylation was an exception [10] (Fig. 4). The ability of 2i to trigger global DNA hypomethylation is thought to be due to the inhibition of MEK signaling [1,83–87]. It is important to mention that transient DNA hypomethylation is a characteristic of embryonic naïve pluripotency at E4.5 [6]. However, both mouse and human 2i-naïve PSCs differ significantly in the genomic methylation pattern compared to pre-implantation epiblast cells [6,73]. Moreover, it has been reported that 2i-naïve mouse PSCs lose developmental potential after long-term passage [88,89]. This loss of developmental potential has been attributed to the continued inhibition of DNA methylation that resulted in chromosomal instability and loss of imprinting. Moreover, female cells turned out to be more sensitive to the long-term effects of MEK inhibition, consistent with the expression of MEK inhibitory factors from the two active X chromosomes. Transcriptional mechanisms have been proposed to connect MEK-inhibition with DNA demethylation, specifically through upregulation of PRDM14, a transcriptional repressor of the DNA methyl-transferase DNMT3 gene family, and activity of the Tet-family of dioxygenases [83,84,86,87,90].

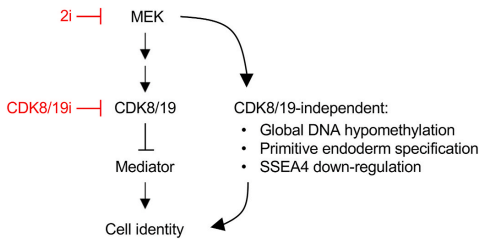


Fig. 4. Unified model of 2i and CDK8/19i naïve pluripotency. A summary of the signaling hierarchy between MEK, CDK8/19-kinase, and Mediator. A large proportion of the effect of MEK signaling channels through CDK8/19-Mediator to affect the transcriptional program of cell identity. CDK8/19-independent effects of MEK inhibition that we have identified include global DNA hypo-methylation in mouse and human PSCs, specification of the primitive endoderm in the mouse pre-implnatation blastocyst, and, in human PSCs, SSEA4 down-regulation.

In contrast to 2i, CDK8/19i recapitulates the same transcriptional changes as 2i, but does not trigger global DNA hypomethylation [10] (Fig. 4). In CDK8/19i-naïve mouse PSCs, DNA methylation remains unchanged. We speculate that MEK inhibition must contribute to DNA demethylation through direct phosphorylation events, such as direct phosphorylation of DNMTs [91]. The DNA hypo-methylation following inhibition of MEK signaling was recently attributed to the down-regulation of UHRF1 protein levels, resulting in a failure to recruit DNMT1 to the replication fork for methylation maintenence [85]. We have confirmed that 2i (which includes MEK-inhibitor) leads to UHRF1 down-regulation, and that this is associated with global DNA hypo-methylation as expected [10]. Interestingly, CDK8/19i treatment did not down-regulate UHRF1 protein levels, thus providing a likely mechanism to explain how CDK81/9i can preserve global DNA methylation [10]. The ability of CDK8/19i to implement naïve pluripotency, without global DNA hypomethylation, may avoid the detrimental side effects of imprint erasure recently reported during extended passage of mouse PSCs under conditions of MEK-inhibition [73,88,89]. Indeed, we have observed that mouse ES cells long-term adapted to CDK8/19i retain full developmental potential upon withdrawal of CDK8/19i, as demonstrated by mouse chimera generation assays, where highly chimeric off-spring developed to adulthood and achieved germline transmission [10]. Since DNA hypo-methylation is part of the naïve stage of pluripotency and is required for normal development [92], the possibility remains that upon withdrawal of CDK8/19i and aggregation of the pre-treated PSCs within a developing blastocyst inner cell mass, a transient wave of DNA hypo-methylation may occur, permitting successful development.

10. Application of CDK8/19i for human pluripotency

Species-specific differences exist between human PSCs and our prior understanding of pluripotency based on rodent models [1,5]. In particular, while 2i treatment has a dramatic efficiency in stabilizing naïve features in mouse PSCs, 2i-based media cocktails generally produce extensive cell death during adaption of human and primate PSCs, and are associated with impaired developmental capacity [4,73]. Importantly, CDK8/19i upregulates and stabilizes multiple features of the naïve state in human PSCs, with minimal cell death, shifting the identity of all cells gradually. Cultivation of human PSCs in the presence of a chemical inhibitor of CDK8/19 is sufficient to recapitulate the majority of molecular characteristics associated with a transition from the primed to the naïve state. Other molecular features associated with the more naïve end of this spectrum in human naïve PSCs include SSEA4 down-regulation [73]. However, SSEA4 down-regulation may not be a strict requirement, since there are chemical cocktails that induce a naïve state in human PSCs withouth downregulating SSEA4 [4,93]. Similarly, CDK8/19i also installs naïve features in human PSCs while maintaining SSEA4 (Fig. 4).

11. CDK8/19i-naïve human pluripotent cells retain long-term developmental potential

Stabilization of the human naïve pluripotent state *in vitro* has proven to be challenging and remains to be optimized [73,74,94]. In particular the genomic instability and loss of imprinting observed with 2i-induction of induce naïve features are associated with striking loss of developmental potential in human PSCs [73,94,95]. This includes inability, or heavy bias, in forming all 3 embryonic germ layers upon differentiation, for example in embryoid body or teratoma assays, which are normally straight-forward methods for primed human PSCs to express their

developmental potency. All this is probably caused, as in mouse, by MEKi-driven DNA demethylation. Indeed, 2i-naïve human PSCs present reduced global methylation, however, the methylation patterns differ significantly from the methylation pattern in pre-implantation epiblast cells [73].

Similar to our observations in mouse PSCs, CDK8/19i does not induce MEK-inhibition or global DNA hypo-methylation in human PSCs [10]. Accordingly, multiple CDK8/19i-naïve human PSCs preserved a normal karyotype after >16 passages [10]. Importantly, upon removal of CDK81/9i chemical inhibition, full developmental potency is maintained [10]. Specifically, following prolonged adaption to CDK8/19i, human PSCs could develop into all three embryonic germ layers in embryoid body and teratomas assays. Moreover, CDK8/19i-pre-treated human PSCs displayed clonal survival and low level chimerism in human-rabbit blastocyst interspecies assays [10]. Altogether, these data suggest that the role of CDK8/19 in pluripotency is conserved in mouse and human, and therefore perhaps across mammalian species.

12. The CDK8/19-kinase during early development

Since CDK8/19 kinase inhibition favours up-regulation of naïve pluripotency *in vitro*, we have explored the possible role of CDK8/19 in the embryo, where the naïve state arises naturally in the absence of chemical inhibitors [10] (Fig. 2A). Firstly, we found that CDK8 mRNA and protein expression is ~5×–20× fold higher than CDK19 in mouse and human early embryo development up to day ~ E6.5, suggesting that CDK8, rather than CDK19, is the major player at this stage. Thus we focused on CDK8 function across early mouse embryonic development (Fig. 2B), and its role may be summarized in three periods [10]:

(i) CDK8 is required during 1C to morula development, where its expression is high. In support of this, CDK8-knockout is embryonic lethal before the 4C stage [96], and CDK8/19i blocked development at the 2C stage [10].
(ii) During morula to blastocyst pre-implantation development, CDK8 and cyclin C expression declines. This coincides with the emergence of the E4.5 pre-implantation naïve epiblast and, accordingly, small molecule CDK8/19i does not interfere with naïve epiblast specification [10]. Also, in contrast to MEK inhibition, CDK8/19i does not affect the epiblast/primitive endoderm (EPI/PE) lineage segregation [10] (Fig. 4). In agreement, CDK8-knockout starting in ~E3.5 embryos permitted naïve epiblast specification, and EPI/PE segregation [97]. Specification of the PE is highly sensitive to MEK inhibition [98–100]. The phosphorylation of the transcription factor GATA6 by MEK has been recently shown to be a key event in the determination of the PE [101]. Since CDK8/19i does not affect the kinase activity of MEK [10], it is possible that the presence of an active MEK/-GATA6 circuit is sufficient to determine PE formation in the face of CDK8/19 inhibition.
(iii) During the subsequent developmental transition of pre-implantation naïve epiblast to the post-implantation primed state, CDK8 expression becomes increased and its activity is required for the morphogenic events during this transition [10]. Moreover, genetic evidence has very recently emerged of an essential role for CDK8 in post-implantation development around ~ E5.5-E10.5 [97], and we speculate that here upstream signals such as FGF-MEK-ERK may guide CDK8 function (as discussed below in the section on MEK signaling). Therefore, overall, CDK8 function in early embryonic development mirrors its expression pattern (Fig. 2B), with elevated expression and essential roles at the 1–2C stage and during post-implantation development, In contrast, between these periods, a physiological minima in CDK8 expression and function exists around E4.5 that equates to naïve PSCs *in vitro* [10]. In sum, we conclude that the physiological minimum in CDK8 function coincides with the emergence of naïve pluripotent epiblast identity *in vivo*, a feature which can be exploited to stabilize naïve PSC culture by CDK8/19i *in vitro*.

In further support of these data, mouse PSCs are known to include an additional form of transcriptional heterogeneity involving infrequent, but dramatic and transient, re-activation of the gene expression program of the 2C embryo stage [102–104]. It has been shown that 2i treatment represses this 2C fluctuation [102,104], and we have confirmed that CDK8/19i also represses this additional example of transcriptional heterogeneity [10]. We suggest an analogy between the high expression and developmental requirement for CDK8 in the 2C embryo stage, and repression of the 2C transcriptional fluctuation in mouse PSCs by CDK8/19i.

13. A signaling hierarchy: CDK8 inhibition downstream of MEK inhibition

Better understanding of the downstream signaling from MEK-ERK, and how it links to the transcriptional machinery, may improve our control over cellular plasticity and aide stabilization of cell identity *ex vivo*. The high degree of overlap in phospho-changes and RNA Pol II regulation induced by 2i and CDK8/19i kinase inhibitors immediately suggests that these chemical inducers of the naïve state might operate within the same pathway [10]. Importantly, analysis of kinase activity has implied a signaling hierarchy (Figs. 1 and 4), where 2i treatment decreases CDK8/19 kinase activity, yet CDK8/19i has no effect on MEK-mediated phosphorylation of ERK. Thus, downstream of MEK-ERK signaling, CDK8/19 activity is down-regulated, explaining how 2i and CDK81/9i treatments functionally overlap in the control of Mediator and the transcriptional machinery [10].

Significant evidence supports the direct input of MEK-ERK signaling into Mediator and the control of the core transcriptional machinery in PSCs. Consistent with this model, it has been recently reported that MEK signalling during exit from the naïve state, directly or indirectly, results in phosphorylation of Mediator subunits and alter eRNA transcription within PSC super-enhancers [32]. All current PSC media cocktails that stabilize the naïve state contain small molecule inhibitors targeting one or more factors in the MEK signalling pathway (FGFRi, RAFi, SRCi, PKCi, p38i, JNKi, MEKi) [1]. It is notable that many components of the MEK pathway have also been shown to regulate CDK8 activity, including KRAS, RAF, SRC, PKC, p38, JNK, MEK, and ERK [105–108]. Thus, CDK8/19-inhibition may be a common feature of media cocktails for stabilizing PSCs in the naïve state. Further studies are required to reveal the precise mechanism by which MEK-ERK signaling regulates CDK8/19 activity in PSCs.

14. Model: dominant programs resolve transcriptional decisions in cell identity

A better understanding of the intracellular competition between transcriptional programs may apply widely to human disease. In this regard, global enhancer hyperactivation was recently identified as a common feature across all human cancers tested [109], while addiction to globally up-regulated transcription also appears to be a unifying aspect of cancer [110]. Using the model of PSCs to study these phenomena, we observe that following exposure to 2i or CDK8/19i, the *trans*-activating potential of existing enhancers and Mediator complexes become globally hyper-activated [10]. This raises an important question: why is the naïve pluripotent state favoured by global enhancer/-Mediator hyperactivation? A general feature of lineage-specifying transcription factors is that they act to promote their own lineage, but repress alternative fates. This is known to also apply for the transcription factors that specify for the 3 embryonic germ layers as cells exit pluripotency [61–66,111–113]. For example, the pluripotency factors OCT4 and SOX2 are known to repress ectodermal- and mesendodermal-lineage enhancers, respectively [65,99,111–113].

Within individual primed cells, these lineage-specifying gene expression programs are in competition with each-other, and they are also moderately expressed and not fully established. Indeed, there is evidence that nascent germ layer programs rely on relatively weak enhancers which are not yet fully established [70,82,114,115], while the naïve enhancers retain a capacity to rapidly return to full strength, retaining their plasticity [10]. Thus, upon global hyperactivation of enhancers by 2i or CDK8/19i, the naïve program becomes dominant, quickly suppressing forward differentiation of the nascent germ layer programs (Fig. 3). In this way, the transcriptional landscape of naïve pluripotency can be up-regulated, establishes dominance, and remains stabilized, by Mediator stimulation, and this can be directly achieved by chemical inhibition of CDK8/19 [10].

15. Other applications of Mediator stimulation via CDK8/19i

We note that a similar mechanism of Mediator hyperactivation via CDK8/19 inhibition has been reported in cancer cells [35]. However, intriguingly, this resulted in cell death in acute myeloid leukemia (AML) cells [35], while we find that a similar approach in PSC reinforces naive cell identity [10]. Cancer cells commonly develop novel oncogenic SEs [110,116,117] that can result in addiction to a defined range of enhancer-driven transcription. Thus cancer cell oncogenic SEs may be sensitive to perturbation, either when hyperactivated, as in the case of CDK8 inhibition [35], or when inhibited, as in the case of BRD4 inhibition [78]. This provides an interesting parallel with MEK inhibition, which is also detrimental to many cancer cells, but is beneficial to the naïve state.

Binary fate decisions are common throughout development and regeneration. We speculate that other cell fate decisions may operate via processes in the transcriptional machinery similar to the naïve-primed pluripotency equilibrium. Thus, we suggest that other examples of CDK8/19i influencing cell identity may exist. Recently, CDK8/19-kinase inhibition was found to produce an anti-inflammatory effect in mouse tissues. This was attributed to the ability of genetic and small molecule CDK8/19-inhibition to promote the production of immuno-suppressive regulatory T cells (T-regs) from their progenitor population of naïve T cells [118–120]. We hypothesize that this process may operate, at least in part, through hyper-activation of enhancers, biasing the outcome of naïve T-cell differentiation toward the production of T-reg cells. It will be interesting to apply CDK8/19i and global hyperactivation of enhancers and Mediator in other examples of cellular plasticity to toggle cell fate, where each system may display a naturally dominant fate.

16. Conclusions

Mediator is a central hub (Fig. 1) which we can target pharmacologically using small molecule inhibition of CDK8/19. CDK8/19i removes a repressive influence from Mediator function, effectively triggering hyperactivation of enhancers via the ability of Mediator to more efficiently recruit RNA Pol II [10]. *In vitro*, we identify that CDK8/19i shifts the equilibrium between naïve and primed pluripotent states, favouring naïve features in mouse and human pluripotent cells (Fig. 2). *In vivo*, we observe CDK8 down-regulation in the E4.5 epiblast, coinciding with the natural emergence of the naïve pluripotent state (Fig. 2). Thus CDK8 down-regulation *in vivo* appears to parallel the ability of CDK8/19i to stabilze the naïve state in PSCs *in vitro*.

The effects of CDK8/19i are not dependent on direct MEK-inhibition and this has the advantage of avoiding some deleterious effects of MEK-inhibition such as DNA hypo-methylation and genomic instability in PSCs [10] (Fig. 4). Nevertheless the majority of the other effects of MEK-inhibition can be phenocopied by CDK8/19i[10]. Thus, current evidence suggests that the control of CDK8/19 activity lies downstream of MEK-ERK signaling (Fig. 4). Further studies are required to reveal precisely how MEK-ERK signaling regulates CDK8/19 activity in PSCs. However, we suggest this model to explain how CDK8/19-inhibition can recapitulate many, but not all, molecular events typically observed during the induction of the naïve state downstream of MEK inhibition. Thus chemical inhibition of CDK8/19 offers a new approach that may help to solve remaining challenges in human naïve PSC culture associated with direct MEK-inhibition [10].

Molecular analyses reveal how the RNA Pol II transcriptional machinery is reorganized by CDK8/19i to coordinate cell identity conversion [10]. In the primed state, multiple lineage-specifiying transcription programs are in competition, and we suggest that CDK8/19i resolves this by up-regulating enhancers and forcing one dominant program to suppress the others (Fig. 3). In this way, heterogenous gene expression constituting a plastic cell state, is resolved into a single homogenous expression program and a stable cell identity. This may reveal insights into how the transcriptional machinery resolves other cell fate decisions, and surprisingly, how it can be directly manipulated to produce a coherent outcome, favouring one local cell identity over others. We hypothesize that CDK8/19i may similarly toggle between cell fate outcomes in other systems of cellular plasticity. Indeed, there is some evidence to support this in AML cancer cells, and in T cell differentiation pathways *in vivo*.

The extent to which CDK8/19i mimics 2i suggests a central role of Mediator during the induction of naïve pluripotency, and it provides a mechanism by which naïve pluripotency may arise *in vivo* [10]. Lastly, chemical inhibition of CDK8/19 may help to stabilize other intrinsically unstable cell states, and it will be of interest to transfer these principles to other contexts of cellular plasticity.

CRediT authorship contribution statement

Cian J. Lynch: Methodology, Conceptualization, Writing - original draft, participated in all discussions. **Raquel Bernad:** Methodology, Conceptualization, helped to improve the text and figures. **Isabel Calvo:** Methodology, Conceptualization, helped to improve the text and figures. **Manuel Serrano:** Methodology, Conceptualization, helped to improve the text and figures, Funding acquisition.

Acknowledgements

We thank our multiple collaborators, and the authors of our references, who have significantly contributed to knowledge of pluripotency and the transcriptional engine. I.C was funded by Secretaria d'Universitats i Recerca de la Generalitat de Catalunya and European Social Fund. Work in the laboratory of M.S. was funded by the IRB, and by the Spanish Ministry of Science co-funded by European Regional Development Fund (ERDF) (SAF2013-48256-R), the European Research Council (ERC-2014-AdG/669622), "laCaixa" Foundation, and Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of Catalonia (Grup de Recerca consolidat 2017 SGR 282).

References

[1] L. Weinberger, M. Ayyash, N. Novershtern, J.H. Hanna, Dynamic stem cell states: naive to primed pluripotency in rodents and humans, Nat. Rev. Mol. Cell Biol. 17 (2016) 155–169.
[2] J. Wu, J.C. Izpisua Belmonte, Dynamic pluripotent stem cell states and their applications, Cell Stem Cell 17 (2015) 509–525.
[3] J. Wu, J.C. Izpisua Belmonte, Stem cells: a renaissance in human biology Research, Cell 165 (2016) 1572–1585.
[4] X. Liu, et al., Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming, Nat. Methods 11 (2017) 1055–1062.
[5] P. Savatier, P. Osteil, P.P.L. Tam, Pluripotency of embryo-derived stem cells from rodents, lagomorphs, and primates: slippery slope, terrace and cliff, Stem Cell Res. 19 (2017) 104–112.
[6] J.A. Hackett, M.A. Surani, Regulatory principles of pluripotency: from the ground state up, Cell Stem Cell 15 (2014) 416–430.
[7] J. Nichols, A. Smith, Naive and primed pluripotent states, Cell Stem Cell 4 (2009) 487–492.
[8] J. Soutourina, Transcription regulation by the Mediator complex, Nat. Rev. Mol. Cell Biol. 4 (2017) 262–274.

[9] C. Jeronimo, F. Robert, The Mediator complex: at the nexus of RNA Polymerase II transcription, Trends Cell Biol. 27 (2017) 765–783.
[10] C.J. Lynch, et al., Global hyperactivation of enhancers stabilizes human and mouse naïve pluripotency through inhibition of Mediator kinases CDK8/19, Nat. Cell Biol. (2020). In press.
[11] S. Heinz, C.E. Romanoski, C. Benner, C.K. Glass, The selection and function of cell type-specific enhancers, Nat. Rev. Mol. Cell Biol. 16 (2015) 144–154.
[12] D. Hnisz, et al., Super-enhancers in the control of cell identity and disease, Cell 155 (2013) 934–947.
[13] W.A. Whyte, et al., Master transcription factors and mediator establish super-enhancers at key cell identity genes, Cell 153 (2013) 307–319.
[14] H.K. Long, S.L. Prescott, J. Wysocka, Ever-changing landscapes: transcriptional enhancers in development and evolution, Cell 167 (2016) 1170–1187.
[15] B.L. Allen, D.J. Taatjes, The Mediator complex: a central integrator of transcription, Nat. Rev. Mol. Cell Biol. 16 (2015) 155–166.
[16] S. Malik, R.G. Roeder, The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation, Nat. Rev. Genet. 11 (2010) 761–772.
[17] D. Hnisz, K. Shrinivas, R.A. Young, A.K. Chakraborty, P.A. Sharp, A phase separation model for transcriptional control, Cell 169 (2017) 13–23.
[18] K.-L. Tsai, et al., Mediator structure and rearrangements required for holoenzyme formation, Nature 544 (2017) 196–201.
[19] P.J. Robinson, et al., Structure of a complete mediator-RNA polymerase II pre-initiation complex, e16, Cell 166 (2016) 1411–1422.
[20] S. Malik, R.G. Roeder, Mediator: a drawbridge across the enhancer-promoter divide, Mol. Cell 64 (2016) 433–434.
[21] A.D. Clark, M. Oldenbroek, T.G. Boyer, Mediator kinase module and human tumorigenesis, Crit. Rev. Biochem. Mol. Biol. 50 (2015) 393–426.
[22] C.B. Fant, D.J. Taatjes, Regulatory functions of the Mediator kinases CDK8 and CDK19, Transcription 10 (2018) 76–90.
[23] F. Klatt, et al., A precisely positioned MED12 activation helix stimulates CDK8 kinase activity, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 2894–2905.
[24] A. Boija, et al., Transcription factors activate genes through the phase-separation capacity of their activation domains, e16, Cell 175 (2018) 1842–1855.
[25] A.V. Zamudio, et al., Mediator condensates localize signaling factors to key cell identity genes, Mol. Cell 76 (2019) 753–766.
[26] A. Erijman, et al., A high-throughput screen for transcription activation domains reveals their sequence features and permits prediction by deep learning, Mol. Cell 78 (2020) 890–902.
[27] L.M. Tuttle, et al., Gcn4-Mediator specificity is mediated by a large and dynamic fuzzy protein-protein complex, Cell Rep. 22 (2018) 3251–3264.
[28] L. El Khattabi, et al., Article A pliable mediator acts as a functional rather than an architectural bridge between promoters and article A pliable mediator acts as a functional rather than an architectural bridge between promoters and enhancers, Cell 178 (2019) 1145–1158.
[29] B.T. Weinert, et al., Time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome resource time-resolved analysis reveals rapid dynamics and broad scope of the CBP/p300 acetylome, Cell 174 (2018) 231–244.
[30] Z.C. Poss, et al., Identification of mediator kinase substrates in human cells using cortistatin A and quantitative phosphoproteomics, Cell Rep. 15 (2016) 436–450.
[31] J. van de Peppel, et al., Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets, Mol. Cell 19 (2005) 511–522.
[32] W.B. Hamilton, et al., Dynamic lineage priming is driven via direct enhancer regulation by ERK, Cell 174 (2018) 231–244.
[33] D. Gonzalez, et al., Suppression of mediator is regulated by cdk8-dependent Grr1 turnover of the Med3 coactivator, Proc. Natl. Acad. Sci. U.S.A. 111 (2014) 2500–2505.
[34] J. Nemet, B. Jelicic, I. Rubelj, M. Sopta, The two faces of Cdk8, a positive/negative regulator of transcription, Biochimie 97 (2014) 22–27.
[35] H.E. Pelish, et al., Mediator kinase inhibition further activates super-enhancer-associated genes in AML, Nature 526 (2015) 273–276.
[36] M.T. Knuesel, K.D. Meyer, C. Bernecky, D.J. Taatjes, The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function, Genes Dev. 23 (2009) 439–451.
[37] C.J. Hengartner, et al., Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases, Mol. Cell 2 (1998) 43–53.
[38] H. Uthe, J.T. Vanselow, A. Schlosser, Proteomic analysis of the mediator complex interactome in Saccharomyces cerevisiae, Sci. Rep. 7 (2017) 43584.
[39] J. Drogat, et al., Cdk11-CyclinL controls the assembly of the RNA polymerase II mediator complex, Cell Rep. 2 (2012) 1068–1076.
[40] M.D. Galbraith, et al., CDK8 kinase activity promotes glycolysis, Cell Rep. 21 (2017) 1495–1506.
[41] M.D. Galbraith, et al., HIF1A employs CDK8-Mediator to stimulate RNAPII elongation, Cell 153 (2014) 1327–1339.
[42] M.D. Galbraith, A.J. Donner, J.M.C.D.K.8 Espinosa, A positive regulator of transcription, Transcription 1 (2010) 4–12.
[43] I. Steinparzer, et al., Transcriptional responses to IFN-γ require mediator kinase-dependent pause release and mechanistically distinct CDK8 and CDK19 functions, e8, Mol. Cell 76 (2019) 485–499.
[44] M.S.J. McDermott, et al., Inhibition of CDK8 mediator kinase suppresses estrogen dependent transcription and the growth of estrogen receptor positive breast cancer, Oncotarget 8 (2017) 12558–12575.
[45] M. Chen, et al., CDK8/19 Mediator kinases potentiate induction of transcription by NFκB, Proc. Natl. Acad. Sci. Unit. States Am. (2017) 201710467.
[46] T. Dale, et al., A selective chemical probe for exploring the role of CDK8 and CDK19 in human disease, Nat. Chem. Biol. 11 (2015) 973–980.
[47] N. Kwiatkowski, et al., Targeting transcription regulation in cancer with a covalent CDK7 inhibitor, Nature 511 (2014) 616–620.
[48] K.A. Nilson, et al., THZ1 reveals roles for Cdk7 in Co-transcriptional capping and pausing, Mol. Cell 59 (2015) 576–587.
[49] P.B. Rahl, et al., C-Myc regulates transcriptional pause release, Cell 141 (2010) 432–445.
[50] B.R. Sabari, et al., Coactivator condensation at super-enhancers links phase separation and gene control, Science 361 (6400) (2018) 3958.
[51] J.A. Mitchell, P. Fraser, Transcription factories are nuclear subcompartments that remain in the absence of transcription, Genes Dev. 22 (2008) 20–25.
[52] W. Cho, et al., Mediator and RNA polymerase II clusters associate in transcription-dependent condensates, 4199, Science 361 (6400) (2018) 412–415.
[53] Y.E. Guo, et al., Pol II phosphorylation regulates a switch between transcriptional and splicing condensates, Nature 572 (2019) 543–548.
[54] M. Nagulapalli, S. Maji, N. Dwivedi, P. Dahiya, J.K. Thakur, Evolution of disorder in Mediator complex and its functional relevance 44 (2016) 56–58.
[55] M.V. Dannappel, D. Sooraj, J.J. Loh, R. Firestein, G. Chan, Molecular and in vivo functions of the CDK8 and CDK19 kinase modules, Front Cell Dev Biol 6 (2019) 171.
[56] Q.-L. Ying, et al., The ground state of embryonic stem cell self-renewal, Nature 453 (2008) 519–523.
[57] H. Marks, et al., The transcriptional and epigenomic foundations of ground state pluripotency, Cell 149 (2012) 590–604.
[58] A. Neagu, et al., In vitro capture and characterization of embryonic rosette-stage pluripotency between naive and primed states, Nat. Cell Biol. 22 (2020) 534–545.
[59] S. Morgani, J. Nichols, A.-K. Hadjantonakis, The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states, BMC Dev. Biol. 17 (2017) 7.
[60] M.N. Shahbazi, E.D. Siggia, M. Zernicka-Goetz, Self-organization of stem cells into embryos: a window on early mammalian development, Science 364 (2019) 948–951.
[61] G. Guo, et al., Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst, Dev. Cell 18 (2010) 675–685.
[62] T. Boroviak, J. Nichols, The birth of embryonic pluripotency, Philos. Trans. R. Soc. 369 (2014), 20130541.
[63] T. Nakamura, et al., A developmental coordinate of pluripotency among mice, monkeys and humans, Nature 537 (2016) 57–62.
[64] S. Petropoulos, et al., Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos, Cell 165 (2016) 1012–1026.
[65] K.M. Loh, B. Lim, A precarious balance: pluripotency factors as lineage specifiers, Cell Stem Cell 8 (2011) 363–369.
[66] K.M. Loh, et al., Mapping the pairwise choices leading from pluripotency to human bone, heart, and other mesoderm cell types, Cell 166 (2016) 451–468.
[67] C. Buecker, et al., Reorganization of enhancer patterns in transition from naive to primed pluripotency, Cell Stem Cell 14 (2014) 838–853.
[68] D. Acampora, L.G. Di Giovannantonio, A. Simeone, Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition, Development 140 (2013) 43–55.
[69] S.H. Yang, et al., Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency, Cell Rep. 7 (2014) 1968–1981.
[70] D.C. Factor, et al., Epigenomic comparison reveals activation of 'seed' enhancers during transition from naive to primed pluripotency, Cell Stem Cell 14 (2014) 854–863.
[71] C. Blanpain, E. Fuchs, Plasticity of epithelial stem cells in tissue regeneration, Science 344 (6189) (2014) 1242281.
[72] A. Wuidart, et al., Early lineage segregation of multipotent embryonic mammary gland progenitors, Nat. Cell Biol. 20 (2018) 666–676.
[73] W.A. Pastor, et al., Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory, Cell Stem Cell 18 (2016) 323–329.
[74] B. Di Stefano, et al., Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells, Nat. Methods 15 (2018) 732–740.
[75] E. Arner, et al., Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells, Science 347 (2015) 1010–1014.
[76] R. Andersson, et al., An atlas of active enhancers across human cell types and tissues, Nature 507 (2014) 455–461.
[77] P. Respuela, et al., Foxd3 promotes exit from naive pluripotency through enhancer decommissioning and inhibits germline specification, Cell Stem Cell 18 (2016) 118–133.
[78] L.W.S. Finley, et al., Pluripotency transcription factors and Tet1/2 maintain Brd4-independent stem cell identity, Nat. Cell Biol. 20 (2018) 565–574.
[79] C. Lopes Novo, et al., Long-range enhancer interactions are prevalent in mouse embryonic stem cells and are reorganized upon pluripotent state transition, Cell Rep. 22 (2018) 2615–2627.
[80] M. Percharde, A. Bulut-Karslioglu, M. Ramalho-Santos, Hypertranscription in development, stem cells, and regeneration, Dev. Cell 40 (2017) 9–21.
[81] S. Efroni, et al., Global transcription in pluripotent embryonic stem cells, Cell Stem Cell 2 (2008) 437–447.
[82] B. Di Stefano, et al., The RNA helicase DDX6 controls cellular plasticity by modulating P-body homeostasis, Cell Stem Cell 25 (2019) 622–638.
[83] J.A. Hackett, M.A. Surani, DNA methylation dynamics during the mammalian life cycle, Philos. Trans. R. Soc. B Biol. Sci. 368 (2012), 20110328.
[84] E. Habibi, et al., Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells, Cell Stem Cell 13 (2013) 360–369.

[85] F. von Meyenn, et al., Impairment of DNA methylation maintenance is the main cause of global demethylation in naive embryonic stem cells, Mol. Cell 62 (2016) 848–861.
[86] H.G. Leitch, et al., Naive pluripotency is associated with global DNA hypomethylation, Nat. Struct. Mol. Biol. 20 (2013) 311–316.
[87] G. Ficz, et al., FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency, Cell Stem Cell 13 (2013) 351–359.
[88] J. Choi, et al., Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells, Nature 548 (2017) 219–223.
[89] M. Yagi, et al., Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation, Nature 548 (2017) 224–227.
[90] P.A. Ginno, et al., A genome-scale map of DNA methylation turnover identifies site-specific dependencies of DNMT and TET activity, Nat. Commun. 11 (2020) 2680.
[91] A. Jeltsch, R.Z. Jurkowska, Allosteric control of mammalian DNA methyltransferases - a new regulatory paradigm, Nucleic Acids Res. 44 (2016) 8556–8575.
[92] K. Kim, et al., Epigenetic memory in induced pluripotent stem cells, Nature 467 (2010) 285–290.
[93] O. Gafni, et al., Derivation of novel human ground state naive pluripotent stem cells, Nature 504 (2013) 282–286.
[94] X. Li, et al., Direct reprogramming of fibroblasts via a chemically induced XEN-like state, Cell Stem Cell 21 (2017) 264–273.
[95] J. Lee, et al., Lineage-specific differentiation is influenced by state of human pluripotency article lineage-specific differentiation is influenced by state of human pluripotency, Cell Rep. 19 (2017) 20–35.
[96] T. Westerling, E. Kuuluvainen, T.P. Makela, Cdk8 is essential for preimplantation mouse development, Mol. Cell Biol. 27 (2007) 6177–6182.
[97] A. Postlmayr, C.E. Dumeau, A. Wutz, Cdk8 is required for establishment of H3K27me3 and gene repression by Xist and mouse development, Development 11 (2020), 147 dev175141.
[98] C. Chazaud, Y. Yamanaka, T. Pawson, J. Rossant, Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the grb2-MAPK pathway, Dev. Cell 10 (2006) 615–624.
[99] S. Bessonnard, et al., Gata6, Nanog and Erk signaling control cell fate in the inner cell mass through a tristable regulatory network, Development 141 (2014) 3637–3648.
[100] Y. Yamanaka, F. Lanner, J. Rossant, FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst, Development 137 (2010) 715–724.
[101] Y. Meng, et al., GATA6 phosphorylation by Erk1/2 propels exit from pluripotency and commitment to primitive endoderm, Dev. Biol. 436 (2018) 55–65.
[102] T.S. Macfarlan, et al., Embryonic stem cell potency fluctuates with endogenous retrovirus activity, Nature 487 (2012) 57–63.
[103] M. Zalzman, et al., Zscan4 regulates telomere elongation and genomic stability in ES cells, Nature 464 (2010) 858–863.
[104] M.A. Eckersley-Maslin, et al., MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs, Cell Rep. 17 (2016) 179–192.
[105] J. Staab, C. Herrmann-Lingen, T. Meyer, CDK8 as the STAT1 serine 727 kinase? JAK-STAT 2 (2013), e24275.
[106] M.S.J. McDermott, et al., Inhibition of CDK8 mediator kinase suppresses estrogen dependent transcription and the growth of estrogen receptor positive breast cancer, Oncotarget 8 (2017) 12558–12575.
[107] W. Xu, et al., Mutated K-ras activates CDK8 to stimulate the epithelial-to-mesenchymal transition in pancreatic cancer in part via the Wnt/β-catenin signaling pathway, Canc. Lett. 356 (2015) 613–627.
[108] N. Li, J.E. Mclaren, D.R. Michael, C.A. Fielding, D.P. Ramji, Uptake of modified lipoproteins by human genes implicated in atherosclerosis, and the activation of STAT1, the expression of key ERK is integral to the IFN-g-Mediated macrophages 185 (2010) 3041–3048.
[109] H. Chen, et al., A pan-cancer analysis of enhancer expression in nearly 9000 patient samples, Cell 173 (2018) 386–399.
[110] J.E. Bradner, D. Hnisz, R.A. Young, Transcriptional addiction in cancer, Cell 168 (2017) 629–643.
[111] J. Shu, et al., Induction of pluripotency in mouse somatic cells with lineage specifiers, Cell 153 (2013) 963–975.
[112] N. Montserrat, et al., Reprogramming of human fibroblasts to pluripotency with lineage specifiers, Cell Stem Cell 13 (2013) 341–350.
[113] I. Sancho-Martinez, A. Ocampo, J.C. Izpisua Belmonte, Reprogramming by lineage specifiers: blurring the lines between pluripotency and differentiation, Curr. Opin. Genet. Dev. 28 (2014) 57–63.
[114] C. Lopes Novo, et al., Long-range enhancer interactions are prevalent in mouse embryonic stem cells and are reorganized upon pluripotent state transition, Cell Rep. 22 (2018) 2615–2627.
[115] S.L. Battle, et al., Enhancer chromatin and 3D genome architecture changes from naive to primed human embryonic stem cell states, Stem Cell Reports 12 (2019) 1129–1144.
[116] C.Y. Lin, et al., Transcriptional amplification in tumor cells with elevated c-Myc, Cell 151 (2012) 56–67.
[117] E. Chipumuro, et al., CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer, Cell 159 (2014) 1126–1139.
[118] N. Ohkura, S. Sakaguchi, Transcriptional and epigenetic basis of Treg cell development and function : its genetic anomalies or variations in autoimmune diseases, Cell Res. 30 (2020) 465–474.
[119] M. Akamatsu, et al., Conversion of antigen-specific effector/memory T cells into Foxp3-expressing Treg cells by inhibition of CDK8/19, Sci. Immunol 4 (2019) 1–17.
[120] Z. Guo, G. Wang, Y. Lv, Y.Y. Wan, Zheng, J. Inhibition of Cdk8/Cdk19 activity promotes treg cell differentiation and suppresses autoimmune diseases 10 (2019) 1–10.

nature cell biology

ARTICLES

https://doi.org/10.1038/s41556-020-0573-1

Check for updates

Global hyperactivation of enhancers stabilizes human and mouse naive pluripotency through inhibition of CDK8/19 Mediator kinases

Cian J. Lynch[1,2], Raquel Bernad[1,2], Ana Martínez-Val[3], Marta N. Shahbazi[4,5], Sandrina Nóbrega-Pereira[6], Isabel Calvo[1,2], Carmen Blanco-Aparicio[7], Carolina Tarantino[8], Elena Garreta[8], Laia Richart-Ginés[9], Noelia Alcazar[1,2], Osvaldo Graña-Castro[10], Gonzalo Gómez-Lopez[10], Irene Aksoy[11], Maribel Muñoz-Martín[1,2], Sonia Martinez[7], Sagrario Ortega[12], Susana Prieto[13], Elisabeth Simboeck[13], Alain Camasses[13], Camille Stephan-Otto Attolini[14], Agustin F. Fernandez[15], Marta I. Sierra[15], Mario F. Fraga[15], Joaquin Pastor[7], Daniel Fisher[13], Nuria Montserrat[8,16,17], Pierre Savatier[11], Javier Muñoz[3], Magdalena Zernicka-Goetz[4,18] and Manuel Serrano[1,2,17] ✉

Pluripotent stem cells (PSCs) transition between cell states in vitro, reflecting developmental changes in the early embryo. PSCs can be stabilized in the naive state by blocking extracellular differentiation stimuli, particularly FGF–MEK signalling. Here, we report that multiple features of the naive state in human and mouse PSCs can be recapitulated without affecting FGF–MEK signalling or global DNA methylation. Mechanistically, chemical inhibition of CDK8 and CDK19 (hereafter CDK8/19) kinases removes their ability to repress the Mediator complex at enhancers. CDK8/19 inhibition therefore increases Mediator-driven recruitment of RNA polymerase II (RNA Pol II) to promoters and enhancers. This efficiently stabilizes the naive transcriptional program and confers resistance to enhancer perturbation by BRD4 inhibition. Moreover, naive pluripotency during embryonic development coincides with a reduction in CDK8/19. We conclude that global hyperactivation of enhancers drives naive pluripotency, and this can be achieved in vitro by inhibiting CDK8/19 kinase activity. These principles may apply to other contexts of cellular plasticity.

Each cell type contains a unique repertoire of active enhancer complexes at specific DNA regions that arise due to a high concentration of lineage-specific transcription factors and signalling pathways[1–3]. The Mediator complex is enriched at enhancers, where it integrates multiple upstream signals and recruits RNA Pol II to nearby and distant promoters[4–6]. A large fraction of Mediator clusters within a small number of unusually long multipartite enhancers, which are known as super enhancers (SEs)[2,3,7]. SEs drive high expression of master transcription factors that maintain cell identity, yet SEs are also vulnerable to perturbation through sharp transitions in their phase separation[2,3,7]. Given the central role of Mediator in enhancer-driven transcription, modulation of its activity may influence cellular identity and plasticity[7,8]. Indeed, global enhancer activation was identified across multiple types of cancer in humans[9]. The kinase CDK8 and its similar, albeit poorly studied, paralogue CDK19 represent the only enzymatic activity of the thirty-subunit Mediator complex[1–6]. This CDK8/19 submodule negatively regulates the recruitment of RNA Pol II through its kinase activity, and it may also sterically hinder the association between Mediator and RNA Pol II[5,6,10–13]. Accordingly, chemical inhibition of CDK8/19 results in global hyperactivation of enhancer function within cancer cells[14]. There are additional layers of

[1]Tumour Suppression Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [2]Cellular Plasticity and Disease Group, Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [3]ProteoRed-ISCIII Proteomics Unit, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [4]Mammalian Embryo and Stem Cell Group, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK. [5]MRC Laboratory of Molecular Biology, Biomedical Campus, Cambridge, UK. [6]Department of Medical Sciences and Institute of Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal. [7]Experimental Therapeutics Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [8]Pluripotency for Organ Regeneration, Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute for Science and Technology (BIST), Barcelona, Spain. [9]Maintenance of Transcriptional Repression by Polycomb Proteins, Institut Curie, Paris, France. [10]Bioinformatics Unit, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [11]Stem Cell and Brain Research Institute, Univ Lyon, Université Lyon 1, INSERM U1208, Bron, France. [12]Transgenic Mice Unit, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [13]IGMM, University of Montpellier, CNRS, Inserm, Montpellier, France. [14]Bioinformatics-Biostatistics Unit, Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain. [15]Cancer Epigenetics and Nanomedicine Laboratory, Nanomaterials and Nanotechnology Research Center (CINN-CSIC), Institute of Oncology of Asturias (IUOPA), ISPA-Hospital Universitario Central de Asturias (HUCA), Universidad de Oviedo, Oviedo, Spain. [16]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Barcelona, Spain. [17]Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain. [18]Division of Biology and Biological Engineering, Caltech, Pasadena, CA, USA. ✉e-mail: manuel.serrano@irbbarcelona.org

complexity—CDK8 can phosphorylate multiple Mediator subunits, the RNA Pol II C-terminal regulatory domain, chromatin regulators and transcription factors[4–6,11–17].

PSCs provide a prototypical model of cellular plasticity, the transcriptional program of which can be stabilized, extinguished or recaptured[18–23]. Although human PSCs offer great therapeutic promise, successful clinical applications remain limited, as human pluripotency is less characterized and less stable in vitro compared with in mice[21,22,24]. Chemical inhibition of MEK and GSK3 kinases with a two-inhibitor cocktail known as 2i shields mouse PSCs from extracellular differentiation in a state that is known as naive pluripotency[25]. Mouse PSCs cultured in 2i (referred to as 2i-naive cells) phenocopy the stable and homogenous state of undifferentiated naive pluripotency that exists transiently in embryonic day 4.5 (E4.5) preimplantation embryo epiblast[18,19,25]. By contrast, culture of PSCs without 2i shifts cell identity towards the postimplantation epiblast at ~E6.5, also known as primed pluripotency[18,19,23]. Enhancer destabilization by chemical blockade of BRD4, which is a key component of enhancers and SEs, triggers the loss of Mediator-driven gene expression in many cell types and induces differentiation in primed PSCs[26–28]. Notably, 2i-naive PSCs are highly resistant to enhancer destabilization[28], indicating that there is an association between naive pluripotency and enhancer stability/resilience. MEK inhibition has been implicated upstream of potent and rapid reconfiguration of the transcriptome, proteome and DNA methylome, within embryonic or 2i-naive pluripotency[18–20,23]. However, the molecular mediators of 2i that are responsible for enhancer stabilization remain unclear.

Here we assessed the effect of inhibiting the activity of the Mediator CDK8/19 kinases, in order to elucidate the transcriptional basis of PSC identity and their plasticity. In summary, stimulating Mediator through its kinase module represses differentiation, favours self-renewal and upregulates preimplantation naive epiblast gene expression in mouse and in human.

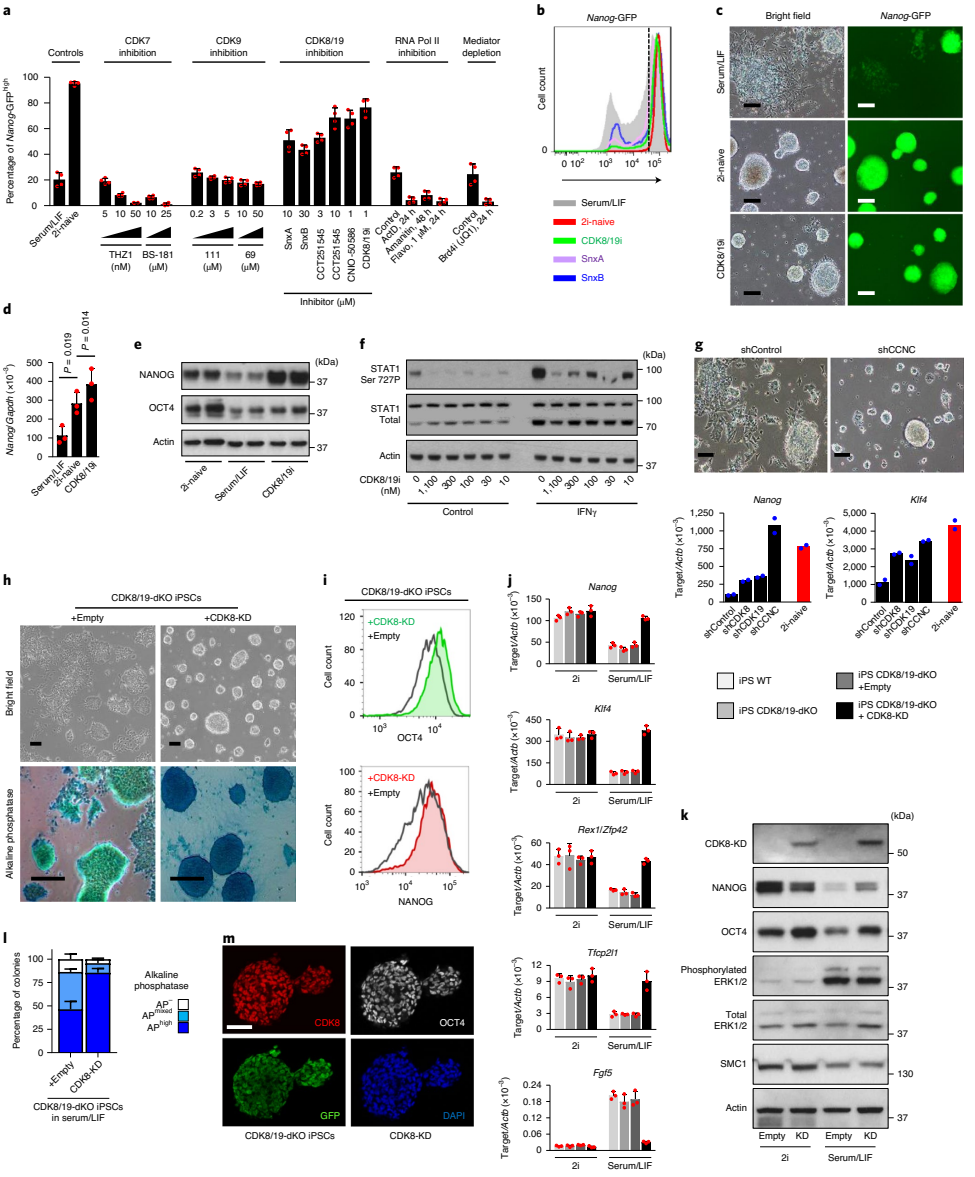
Results

Inhibition of Mediator kinase stabilizes mouse naive pluripotency. GFP knock-in reporters at key stem cell marker genes such as *Nanog* represent well-established and precise indicators of the naive (GFP^{high}) and primed states (GFP^{low})[18,22,29]. For example, in the 2i-naive state, *Nanog* promoter activity is enhanced, yielding a characteristically homogenous *Nanog*-GFP^{high} cell expression pattern and uniform dome-shaped colonies (Fig. 1a–c and Extended Data Fig. 1a). By contrast, the *Nanog* promoter is metastable in primed-state PSCs, reversibly oscillating between high and low activity, presenting a heterogeneous *Nanog*-GFP expression pattern and flattened diffuse colonies, indicative of a general underlying switch in the transcriptional program[18,20,23,29,30]. The BRD4 inhibitor JQ1 destabilizes enhancers, resulting in colony flattening and a GFP^{low} status (Fig. 1a), as reported previously[26–28]. In this experimental setting, we tested the effect of manipulating the transcriptional cyclin-dependent kinases (CDK7, CDK8/19 and CDK9) with a panel of small-molecule inhibitors. Several potent and structurally unrelated CDK8/19 inhibitors had a positive effect, inducing the formation of homogenous dome-shaped colonies, and upregulating both the *Nanog*-GFP reporter and endogenous *Nanog* expression, similar to PSCs in the 2i-naive state (Fig. 1a–e, Extended Data Fig. 1a and Supplementary Table 1), whereas inhibition of CDK7 or CDK9 did not. We assessed the potency and selectivity of CDK8/19 inhibitors, commercially available or developed in-house, using multiple methods: (1) selectivity was suggested using a KinomeScan panel of 456 kinases; (2) LanthaScreen assays demonstrated inhibitory activity at nanomolar concentrations against pure recombinant CDK8–CCNC and CDK19–CCNC; (3) luciferase reporter cell assays (TOP-FLASH); and (4) potent inhibition of phosphorylation of STAT1 at Ser 727 in human PSCs, which is a well-documented CDK8 target site[11,14,16,31] (Fig. 1f, Extended Data Fig. 1b, Supplementary Table 1 and Supplementary Information). On the basis of these data, we focused on the molecule that was generated at the Centro Nacional de Investigaciones Oncológicas (CNIO)—CDK8/19i-ETP-47799 (hereafter, CDK8/19i), which was the most effective at improving mouse PSCs (Fig.1a,b and Extended Data Fig. 1a; information about the structure and characterization of this inhibitor, as well as a comparison with other inhibitors used in this study, is provided in Supplementary Table 1 and Supplementary Information). In addition to improvements in the *Nanog*-GFP profile and colony morphology described above, the effect of CDK8/19i on mouse PSCs resembled 2i in three other ways: (1) it was observed in serum-containing and serum-free-based media (Fig. 1a and Extended Data Fig. 1a); (2) it was reversible after CDK8/19i withdrawal, with kinetics similar to that of 2i removal (Extended Data Fig. 1c); and (3) after removal of LIF or inhibition of LIF signalling using a JAK inhibitor, the presence of CDK8/19i delayed the downregulation of *Nanog*-GFP expression (Extended Data Fig. 1d,e). We conclude that inhibiting Mediator kinase CDK8/19 shifts mouse PSC morphology and *Nanog* expression towards their characteristic status in the naive state[18,23,29].

As a genetic validation, depletion of CDK8, CDK19 and, most successfully, their regulatory partner cyclin C (CCNC; which is

Fig. 1 | An inhibitor screen for factors that promote the PSC naive state identifies a distinct role for Mediator kinase activity. a, Fluorescence-activated cell sorting (FACS) analysis of the effect of the indicated treatments on *Nanog* expression (*Nanog*-GFP^{high}) per cell, using a mouse *Nanog*-GFP knockin reporter PSC line[29] in standard serum/LIF base medium. Data are mean ± s.d. of $n = 4$ independent experiments. ActD, Actinomycin D; Flavo, Flavopiridol; compounds 69 and 111 are specific CDK9 inhibitors. Details about inhibitor characterization are provided in Supplementary Table 1 and the Supplementary Information **b**, FACS analysis of *Nanog*-GFP expression with three different CDK8/19i inhibitor molecules. *Nanog*-GFP^{low} and *Nanog*-GFP^{high} cell populations in the serum/LIF population (grey). The dotted line indicates the threshold at which >95% cells are *Nanog*-GFP^{High} in 2i-naive culture conditions. Data are representative of three experiments. **c**, PSC colony morphology in the indicated treatments. Bright-field images and *Nanog*-GFP expression are shown. Data are representative of six experiments. **d**,**e**, Endogenous *Nanog* mRNA (**d**) or protein (**e**) expression levels in mouse PSCs adapted to the indicated conditions. Data are representative of three experiments. Data are mean ± s.d. Statistical analysis was performed using unpaired two-tailed *t*-tests; $^{*}P < 0.05$. **f**, The levels of the CDK8-target STAT1 phosphorylation at Ser 727 (Ser 727P). HERVH human iPSCs treated with CDK8/19i concentrations for 3 h with or without the simultaneous induction of STAT1-Ser 727P by interferon-γ for 3 h. Data are representative of two experiments. **g**, Cell morphology and qPCR with reverse transcription (RT–qPCR) analysis of mouse PSCs after 7 d of shRNA-mediated knockdown of *CDK8*, *CDK19* and *CCNC* (which encodes cyclin C). Data are the mean values from two experiments. **h**,**i**, Cell morphology and alkaline phosphatase staining (**h**) and FACS analysis of endogenous NANOG and OCT4 protein levels (**i**) in CDK8/19-dKO iPSCs stably expressing pMSCV-Empty (Empty) or pMSCV-CDK8-kinase dead (CDK8-KD). Data are representative of three independent clones. **j**,**k**, RT–qPCR analysis (**j**; data are mean ± s.d. from $n = 3$ independent clones) and western blot analysis of protein expression (**k**; data are representative of two experiments) in WT iPSCs or CDK8/19-dKO iPSCs stably expressing pMSCV-Empty or pMSCV-CDK8-KD, adapted to the indicated medium conditions. **l**, Alkaline phosphatase (AP) staining. Cells were fixed and stained 14 d after retroviral expression of pMSCV-Empty or pMSCV-CDK8-KD. Staining intensity was scored visually for each colony using ten fields of view. Data are mean ± s.d. from three experiments. **m**, Immunofluorescence in CDK8/19-dKO iPSCs expressing pMSCV-CDK8-KD-puro-IRES-GFP. Data are representative of four experiments. For **c**, **g**, **h** and **m**, scale bars, 100 μm.

essential for full kinase activity[8]) by short-hairpin RNA (shRNA) knockdown led to upregulation of *Nanog* expression and naive-like colony morphology (Fig. 1g and Extended Data Fig. 1f,g). In another genetic approach, we generated CDK8/19 double-knockout (dKO) mouse PSCs (Extended Data Fig. 1h–k). CDK8/19-dKO PSCs could self-renew indefinitely, but CDK8/19-dKO was insufficient to confer naive morphological features or *Nanog* upregulation. Importantly, CDK8/19-dKO PSCs no longer responded to CDK8/19 inhibitors (Extended Data Fig. 1l,m). Together, this suggested that the beneficial effects observed may require the physical



presence of the inactive kinase. In agreement, we found that reconstituting CDK8/19-dKO PSCs with exogenous CDK8 rescued the ability of these cells to respond to CDK8/19i, observed by naive morphological features and *Nanog, Klf4* and *Oct4* upregulation (Extended Data Fig. 1l,m). Moreover, CDK8/19-dKO PSCs that were reconstituted with a CDK8 kinase-dead mutant (CDK8-KD; D173A) displayed homogenous naive colony morphology, high expression of naive-state markers (Fig. 1h–m) and downregulation of *Fgf5*, which is a key marker of the primed state[18,19,23] (Fig. 1j), all without the need for any chemical inhibitor and despite maintaining active MEK–ERK signalling (Fig. 1k). Thus, CDK8/19-dKO cells expressing CDK8-KD phenocopy the effects of chemical inhibition of CDK8/19. Finally, postimplantation epiblast stem cells (EpiSCs; cultured with FGF2 and activin) are in a more developmentally advanced primed state than mouse PSCs in serum/LIF[18,19]. Interestingly, EpiSCs exogenously expressing CDK8-KD lost *Fgf5*, upregulated *Nanog, Rex1* and *Klf4*, and formed dome-shaped colonies with high alkaline phosphatase staining; together, these are characteristic of conversion to the naive state (Fig. 2a and Extended Data Fig. 2a). In summary, CDK8/19 kinase inhibition is sufficient to promote key characteristics of naive pluripotency, despite the continued presence of MEK–ERK signalling.

Long-term culture of mouse PSCs (>10 passages) in CDK8/19i maintained PSC naive features, including colony morphology, high alkaline phosphatase, *Nanog*-GFP^{high}, high endogenous *Nanog*, high ICAM1 cell surface expression and nuclear localization of TFE3 (refs. [21,24,32–34]; Fig. 2b–d and Extended Data Fig. 2b,c). Long-term CDK8/19i-adapted PSCs displayed typical developmental capacity after inhibitor withdrawal, specifically, retinoic-acid-induced differentiation, embryoid body cardiac centre formation, spheroid polarization and lumenogenesis[35], generation of teratomas containing three germ layers, and robust chimaera contribution after morula aggregation and blastocyst microinjection assays (traced by constitutive GFP or RFP) evaluated at E4.5, E7.5, E14.5 and in fully developed adults that subsequently completed germline transmission (Fig. 2d–i and Extended Data Fig. 2d–f). Notably, the continued presence of CDK8/19i impaired the early developmental events[35] of polarization and lumenogenesis in vitro (Fig. 2e), an observation that is discussed below. Thus, PSCs that are long-term adapted to CDK8/19i maintain upregulation of naive features, self-renewal and developmental capacity.

CDK8/19i induces and stabilizes the naive state in human PSCs. We tested the effect of CDK8/19i on human stem cell identity. STAT3 overexpression plus 2i induces the human naive state[36], and we observed that CDK8/19i could replace 2i in this system (Fig. 2j). Even in the absence of STAT3 overexpression, other transgenes or chemicals, CDK8/19i treatment progressively converted human induced PSC (iPSC) colonies from flat and primed-like, to dome-shaped naive-like birefringent morphology. This was observed for a total of 7 human PSC lines treated with 0.4 µM or 1.1 µM CDK8/19i/LIF for 2–3 weeks (Fig. 2k and Extended Data Fig. 2g), including human iPSCs carrying a specific *HERVH*-GFP reporter insertion that marks human naive cell identity[37] (Fig. 2k and Extended Data Fig. 2h). A 2i-based chemical cocktail (hereafter 2i p38iJNKi) induced naive colony morphology, as expected[35,37], and combined with selection by cell sorting yielded cultures with homogeneous *HERVH*-GFP^{high} (Extended Data Fig. 2h). Interestingly, treatment with CDK8/19 inhibitors (CDK8/19i or SnxA) also produced morphological conversion and increased GFP, similar to 2i p38iJNKi (Fig. 2l and Extended Data Fig. 2h). The changes induced by CDK8/19 inhibition were gradual, required no selection after passage (sorting or manual picking), required no additional supplements except for rhLIF and were stable in the continuous presence of the inhibitor. By contrast, CDK7 inhibition failed to change colony morphology or GFP fluorescence, and produced cell death (Extended Data Fig. 2h). Culturing human PSCs in CDK8/19i, with or without p38iJNKi, increased their clonogenicity, alkaline phosphatase intensity and pluripotency markers[32–34,38] NANOG, OCT4, SSEA4, TRA1–81, TFCP2L1 and KLF17 (Figs. 2m and 3a, and Extended Data Figs. 2i and 3a–c). MYC, which is known to be reduced in naive cells[25,32], was also reduced in cells maintained in CDK8/19i (Fig. 3a). Thus, similar to the observations in mouse PSCs described above, treatment of human PSCs with CDK8/19i establishes features that are characteristic of the naive state.

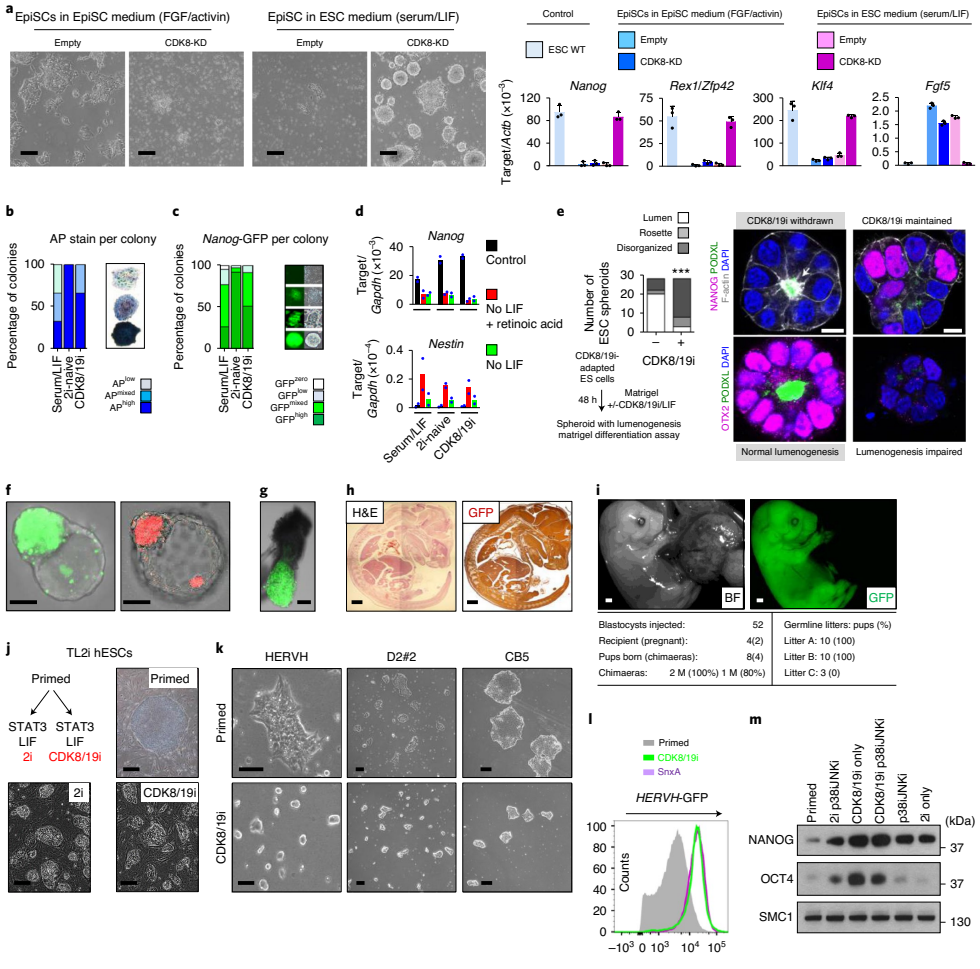
Developmental potential of CDK8/19i-adapted human PSCs. Chemical induction of the human naive state can trigger genomic instability, severely impairing developmental potential[24,39]. We found that CDK8/19i-adapted human PSCs (five lines) had a normal karyotype over >16 passages (Extended Data Fig. 3d) and, after inhibitor withdrawal, maintained the capacity to contribute towards all three embryonic germ layers by embryoid-body differentiation in vitro and by teratoma assay in vivo (Fig. 3b–d and Supplementary Table 1), comparable to control primed cells. Preimplantation interspecies chimerism tests for naive-specific properties, namely, capacity for clonal survival in a host embryo[40,41]. We tested CDK8/19i-adapted human iPSCs carrying a constitutive

Fig. 2 | Positive effect of long-term treatment of CDK8/19i on PSC self-renewal and pluripotency. a, Morphology and mRNA expression of mouse EpiSCs expressing pMSCV-Empty or pMSCV-CDK8-KD, then 7 d in EpiSC medium or standard embryonic stem cell (ESC) medium serum/LIF. Data are mean ± s.d. from $n = 3$ experiments. Scale bars, 100 µm. **b**,**c**, Clonogenicity of mouse PSCs. *Nanog*-GFP PSCs were FACS-sorted one cell per well, cultured for 7 d and then stained for alkaline phosphatase (**b**) or scored for *Nanog*-GFP intensity (**c**) to assess the pluripotent status of each colony, in standard medium serum/LIF, 2i-naive or CDK8/19i conditions. Data are representative of three experiments. **d**,**e**, The differentiation capacity of mouse PSCs that were previously adapted to serum/LIF, 2i or CDK8/19i. **d**, PSCs differentiated as indicated in two-dimensional culture. Analysis of PSC exit from pluripotency (*Nanog* downregulation) and differentiation (*Nestin* upregulation) using RT–qPCR. Data are the mean values of two experiments. **e**, Pluripotency exit assessed using immunofluorescence after PSC culture in three-dimensional Matrigel with or without CDK8/19i/LIF to observe early epiblast development (rosette formation and lumenogenesis) in PSC spheroids[35]. 'Disorganized' indicates differentiation failure. Data are representative of three experiments, $n = 30$ spheroids per condition. Statistical analysis was performed using unpaired two-tailed *t*-tests; $^{***}P = 0.0097$. Scale bars, 10 µm. **f**–**i**, In vivo assays of developmental capacity. Mouse CDK8/19i-treated PSCs, constitutively labelled with ROSA26–GFP or Tg.CAG–Katushka, were aggregated with, or microinjected into, host E2.5 morulae. Embryo chimerism was assessed visually. E4.5 blastocyst, $n = 10$ (**f**); E6.5 egg cylinder, $n = 10$ (**g**); E14.5, $n = 2$ (**h**); and perinatal E19.5, $n = 4$ (**i**). In **i**, three male (M) adult chimaeras (bottom left, the percentage of chimerism on the basis of coat colour is indicated) displayed germline transmission, generating three litters (bottom right, coat colour confirmed germline transmission per litter). Scale bars, 25 µm (**f**), 100 µm (**g**), 1 mm (**h**), 1 mm (**i**). H&E, haematoxylin and eosin; BF, bright field. **j**, Induction of naive colony morphology in human OSCAR ESCs. Tamoxifen-inducible constitutively active STAT3/LIF/2i (TL2i)[36], or substituting CDK8/19i for 2i (TLCDK8/19i). Data are representative of three experiments. Scale bars, 100 µm. **k**, Induction of naive colony morphology in three human PSC lines; primed or cultured for 14 d with CDK8/19i. Scale bars, 100 µm. **l**, Cytometry analysis of *HERVH*-GFP intensity per cell in human PSCs; primed or cultured for 14 d with CDK8/19i. For **k** and **l**, data are representative of >5 experiments. **m**, Western blots of pluripotency markers in human PSCs; primed or cultured for 14 d with 2i-based or CDK8/19i-based medium, with or without p38iJNKi. SMC1 was used as a loading control.

Tomato-red marker for human–rabbit interspecies chimerism by microinjecting them into E2.5 rabbit morulae. Interestingly, the presence of human cells (Tomato$^+$) was detected 72 h later in up to 50% of the injected rabbit blastocysts (Fig. 3e). By contrast, human PSCs in the primed state were unable to integrate or survive in rabbit embryos (0 out of 24 rabbit embryos), similar to previous reports for primed state human PSCs within the embryos of mice, pigs and cattle[40,41]. In summary, long-term adaptation of human PSCs to CDK8/19i stabilizes naive pluripotency while preserving their developmental potential. We conclude that the role of CDK8/19 in pluripotency is conserved in mice and humans and, therefore, presumably across mammals.

CDK8/19i resets the transcriptome and proteome similar to 2i. Using RNA-seq, we compared global gene expression in mouse PSCs that were long-term adapted to CDK8/19i versus 2i. Overall, CDK8/19i altered gene expression with a magnitude similar to the magnitude in 2i conditions, and with a highly significant overlap in the identity and biological functions of genes that were up- or downregulated in both serum-containing and serum-free media (Fig. 4a, Extended Data Fig. 3e,f and Supplementary Table 2). Compared with control primed conditions, naive pluripotency markers were enhanced in CDK8/19i and 2i (Fig. 4b and Extended Data Fig. 3g), whereas differentiation markers were globally downregulated in CDK8/19i and 2i conditions (Supplementary Table 2).

Endogenous retrovirus (ERV) expression is highly stage-specific during mammalian preimplantation and precisely defines naive and primed PSC identity[42–46]. The transcriptomic overlap between CDK8/19i or 2i treatments extended to ERVs; similar viral families were significantly up- or downregulated in mouse PSCs (Fig. 4c and Supplementary Table 2). In particular, LINE L1 families, each with thousands of copies across the genome, were regulated



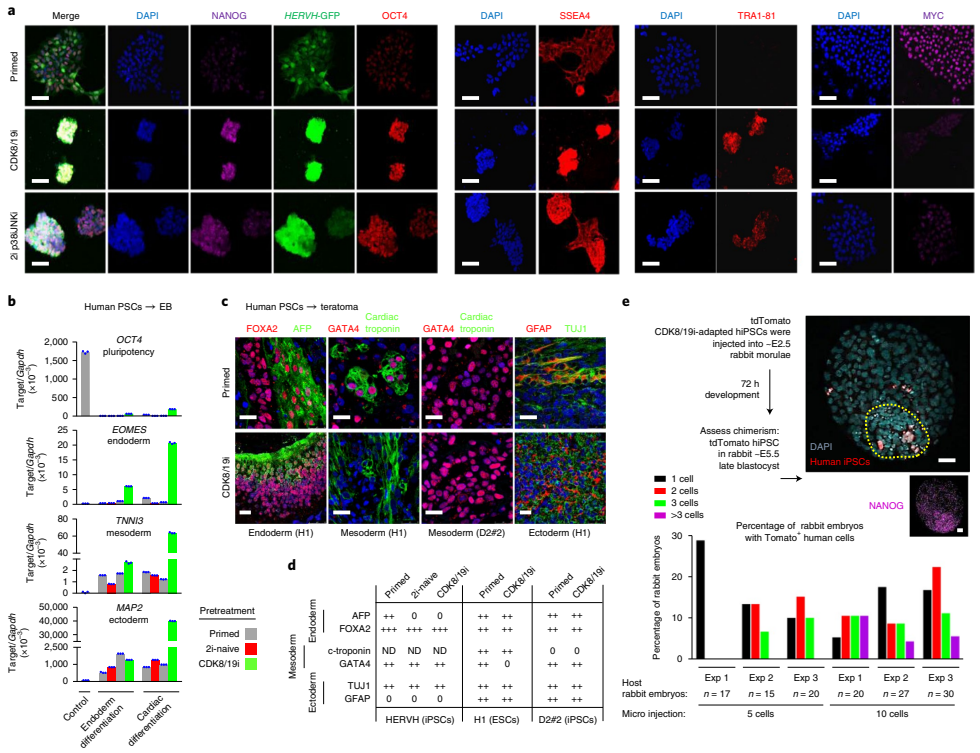


Fig. 3 | Positive effect of long-term treatment of CDK8/19i on the self-renewal and pluripotency of human PSCs. a, Immunofluorescence analysis of pluripotency markers. Human PSCs were cultured as indicated. Data are representative of two experiments. Scale bars, 100 µm. **b**, Embryoid body differentiation assay using endoderm-directed or cardiac-directed protocols with human PSCs. mRNA expression of pluripotency or embryonic germ layer markers on the basis of RT–qPCR analysis. Mean values of $n = 4$ technical replicates. Data for the H9 cell line are shown, representative of two experiments (human PSC lines H1 and H9: a summary of all tested lineage markers determined using RT–qPCR ($n = 17$) or immunofluorescence ($n = 6$) is provided in the Supplementary Information). EB, embryoid body. **c**,**d**, Human PSCs were adapted to primed or CDK8/19i conditions and then analysed using a teratoma differentiation assay. Data are representative of three experiments (human PSC lines: H1, D2#2 and HERVH). **c**, Immunofluorescence imaging shows markers for three embryonic germ layers in the H1 and D2#2 cell lines, as indicated. Scale bars, 50 µm. **d**, A summary of all of the tested lineage markers (6) determined on the basis of immunofluorescence in teratomas generated from the three human PSC lines described in **c**. The plus symbol (+) indicates detected; '0' indicates not detected. ND, not determined. A summary of all of the differentiation markers tested for all three cell lines in **c** and **d** is provided as source data. c-troponin, cardiac troponin. **e**, Interspecies chimaera assay in vivo to assess the developmental capacity of human PSCs that were adapted to primed or CDK8/19i conditions. Constitutively labelled human iPSCs (hiPSCs) (tdTomato (red); HERVH iPSC line) were introduced into host rabbit morulae of ~E2.5. Chimerism was assessed visually 72 h later in ~E5.5 rabbit blastocysts. The number of human cells introduced (5 or 10) and the number of embryos (n) in each of the three experiments (Exp 1–3) are indicated at the bottom. Quantification of the number of human cells observed in rabbit embryos is shown; data are from three independent experiments. A representative image (top) shows immunofluorescence of E5.5 rabbit blastocysts; the ICM is indicated by a dotted yellow line (determined by NANOG staining (inset)). Human PSCs that were adapted to CDK8/19i displayed a moderate contribution to human–rabbit chimaeras. Scale bars, 20 µm.

in close parallel, displaying highly similar pattern of expression in the CDK8/19i and 2i-naive states (Extended Data Fig. 3g,h). Another aspect of the plasticity of mouse PSCs is their ability to transition to a two-cell-like (2C) state, specifically marked by hyper-activation of the MERVL family of ERVs and by *Zscan4c* expression[46,47]. Stabilization of the naive state with 2i impairs the 2C-like fluctuation[46,47]. We also observed this in CDK8/19i-treated PSCs using multiple 2C markers, including *MERVL* and *Zscan4c*, and *MERVL*-Tomato and *Zscan4c*-eGFP 2C-reporter models (Fig. 4d, and Extended Data Figs. 3i–m and 4a–c). Finally, our CDK8/19i and 2i transcriptomic data correlated with published transcriptomes[48–51] from independent studies of 2i-naive mouse PSCs and the transcriptome of E4.5 epiblast single cells[52] (Fig. 4e and Extended Data Fig. 4d,e).

RNA-seq analyses of human PSCs that were adapted to CDK8/19i or a 2i-based naive cocktail overlapped significantly (Fig. 4f and

Supplementary Table 3). Markers of human and primate preimplantation epiblasts and in vitro naive human PSCs[32–39] were upregulated by CDK8/19i, whereas differentiation markers were repressed[53–59] (Fig. 4g, Extended Data Fig. 4f–j and Supplementary Table 3). Moreover, the global human ERV transcriptomes of CDK8/19i- and 2i-adapted cells overlapped extensively, including upregulation of the SVA, LTR7 and HERV families (Fig. 4h–j and Supplementary Table 3), consistent with reports of ERV expression in human and primate naive PSCs and preimplantation epiblast[43–45]. Finally, we observed that there is a high correlation between RNA expression data of CDK8/19i-adapted human PSCs and seven independent studies in human and primate PSCs from the in vitro naive state and from embryo naive epiblast single-cell analyses[39,53–60] (Fig. 4k).

Although PSC plasticity has been examined in terms of RNA expression, its proteome remains relatively poorly defined. We analysed the proteome of mouse PSCs in serum/LIF versus 2i-naive or CDK8/19i-adapted conditions. Across five mouse PSC lines, CDK8/19i altered the expression levels of 465 proteins, of which 159 (34%) changed in the same direction in 2i conditions (Fig. 4l,m, Extended Data Fig. 5a,b and Supplementary Table 4). Importantly, among the overlapping changes in both 2i-naive and CDK8/19i conditions, we noted that key pluripotency regulators, such as KLF4, and metabolic pathways, such as oxidative phosphorylation, featured among the most upregulated; by contrast, LIN28A, MYC-target genes and differentiation markers were downregulated (Fig. 4m, Extended Data Fig. 5c and Supplementary Table 4). Furthermore, proteomic changes in 2i and CDK8/19i conditions were significantly correlated with the transcriptomic changes observed (Extended Data Fig. 5d,e).

In summary, CDK8/19i upregulates pluripotency markers, reshapes the endogenous retroviral transcriptome and represses differentiation markers in a manner that is similar to the transcriptomic and proteomic resetting that was observed in previous studies of naive pluripotency, in vitro and in vivo, in mice and humans.

CDK8/19i does not reset global DNA methylation levels. Many 2i-based chemical cocktails induce global DNA hypomethylation, both in mouse and human PSCs[21]. This has been attributed to MEK-dependent stabilization of UHRF1, which is a critical factor for recruiting DNMT1 to the DNA[61]. Importantly, the pattern of demethylation induced by 2i diverges significantly compared with the preimplantation naive epiblast state, and is associated with PSCs exhibiting genomic instability, chromosomal defects and loss of pluripotency[24,39,62,63]. Recent 2i-variant cocktails (with partial MEK inhibition) offer the advantage of largely preserving global DNA methylation[62–64]. Importantly, neither mouse nor human CDK8/19i-adapted PSCs showed evidence of global DNA hypomethylation (Fig. 5a,b). Moreover, 2i or MEK inhibition alone induced demethylation of LINE L1 repeat regions (Fig. 5c) and major satellite regions (Extended Data Fig. 5f) but had no effect on the methylation of IAP repeats (Extended Data Fig. 5g), all as previously reported[65]. By contrast, CDK8/19i did not reduce methylation at any of these mouse repeat elements (Fig. 5c and Extended Data Fig. 5f,g) or UHRF1 protein levels (Supplementary Table 4). Thus, CDK8/19i induces naive features in the absence of global DNA hypomethylation, and this is probably due to its lack of MEK inhibition (Fig. 1k and see below) or UHRF1 downregulation. By not recapitulating the partial demethylation of the naive epiblast, CDK8/19i has the advantage of preserving chromosomal stability and pluripotency after cell expansion (see above), which is particularly relevant for naive human PSCs. This is consistent with variant medium cocktails that are based on minimizing MEK inhibition both in mouse and human naive PSCs[62–64].

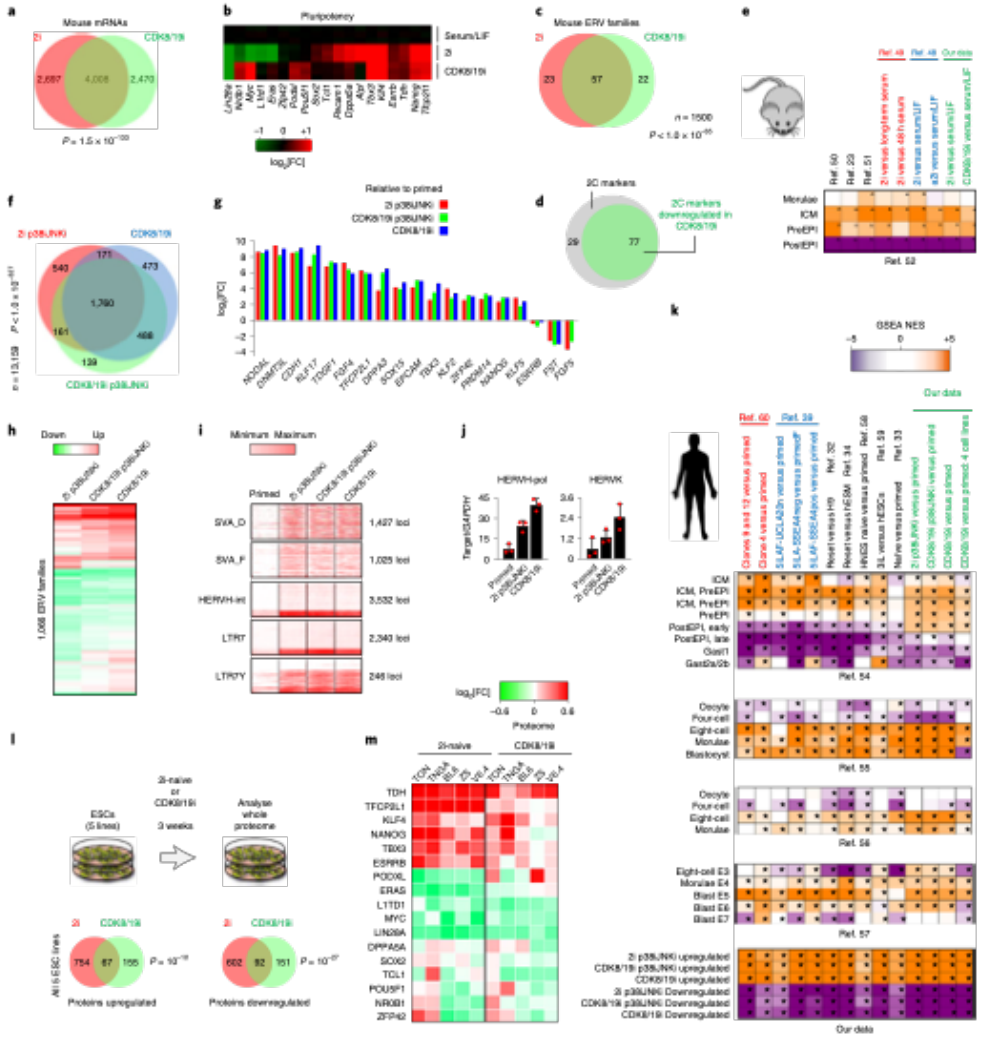
X-chromosome reactivation status is another molecular signature that has been reported in human naive pluripotency during MEK inhibition[21,66,67], which may be inferred by assessing XIST RNA expression in female cells. However, analysis using quantitative PCR (qPCR) revealed that there was very low XIST expression in our primed human PSCs (Extended Data Fig. 5h), suggesting that erosion of X-chromosome silencing may have already occurred in the parental cells under primed conditions, as observed previously[67]. Notably, some 2i-based cocktails reactivate XIST expression even in primed human PSCs displaying erosion of X-chromosome-silencing[66,67], but this was not the case for our CDK8/19i-adapted cells (Extended Data Fig. 5h). In summary, CDK8/19i treatment does not recapitulate the reactivation of XIST in X-chromosome-silencing-eroded primed cells, indicating another distinction between CDK8/19i and most human medium cocktails that are based on MEK inhibition.

CDK8/19i induces phosphorylation changes similar to 2i. We assessed the phosphoproteome of mouse PSCs after only 15 min exposure to CDK8/19i or 2i to explain their phenotypic similarity. Strikingly, out of 622 altered phosphorylation sites, 495 (79.6%) were similarly regulated by CDK8/19i and 2i (Fig. 5d,e). The co-regulated phosphorylation sites occurred on proteins that

Fig. 4 | Gene expression in mouse and human PSCs adapted to 2i or CDK8/19i. a,**b**, Overlap and hypergeometric significance (**a**) of differentially expressed mRNAs in mouse PSCs in 2i-naive or CDK8/19i conditions versus serum/LIF (RNA-seq; $n = 3$ biological replicates; false-discovery rate (FDR)-adjusted P (q) < 0.01). **b**, Heat map of changes in selected pluripotency regulators. **c**,**d**, Overlap and hypergeometric significance of differentially expressed ERV families (**c**) and overlap of 2C fluctuation markers (**d**) in mouse PSCs in 2i-naive or CDK8/19i versus serum/LIF. $n = 3$ biological replicates; $q < 0.05$. **e**, Heat map of normalized enrichment scores (NES) from a gene set enrichment analysis (GSEA) comparison of the mouse RNA-seq data shown in **a** ($n = 3$ biological replicates), or five other studies[23,48–51], versus stage-specific marker gene sets that were identified during mouse preimplantation development[52]. PreEPI/PostEPI, refer to pre- or post-implnatation epiblast cells. Significant comparisons are indicated by asterisks; $P < 0.05$ and $q < 0.05$. **f**,**g**, Overlap and hypergeometric significance (**f**) of differentially expressed mRNAs in human PSCs cultured as indicated versus control primed cells (RNA-seq; $n = 3$ biological replicates; $q < 0.05$, threshold >2× fold change). **g**, Selected human naive pluripotency markers (upregulated), or postimplantation primed epiblast markers (downregulated) in human PSCs cultured as indicated. **h**,**i**, Heat maps of the correlation in RNA expression changes for ERV families (**h**; $n = 1{,}066$ families, each row shows the mean value of the family) or individual loci of selected ERV families (**i**; genomic loci n values are indicated) in human PSCs cultured as indicated versus control primed cells. HERVH-int refers to the internal transcribed region of the HERVH genome. **j**, RT–qPCR analysis of RNA expression of the ERVs HERVH or HERVK, in human PSCs cultured as described in **f**. $n = 3$ biological replicates. Data are mean ± s.d. **k**, Heat map of NES scores from a GSEA comparison of our human PSC RNA-seq ($n = 3$ biological replicates), or seven other studies (indicated at the top)[32–34,39,58–60], versus stage-specific marker gene sets identified during human preimplantation development (below each heat map)[54–57]. Bottom, comparison between our data and the other studies. Significant comparisons are indicated by asterisks; $P < 0.05$ and $q < 0.05$. Colour scale as in **e** above. **l**,**m**, Proteomic overview. **l**, Overlap and hypergeometric significance of differentially expressed proteins averaged across five mouse PSC lines in 2i-naive or CDK8/19i versus standard serum/LIF culture. $n = 5$ biological replicates; $q < 0.05$. **m**, Heat map of protein changes in selected pluripotency regulators per cell line. Data for each cell line are provided in Extended Data Fig. 5a,b. Full gene lists, ERV lists, fold changes and statistical tests in **a**–**i** and **k** are provided in Supplementary Table 2. A list of differentially expressed proteins in **l** and **m** is provided in Supplementary Table 4.

were heavily enriched for functions in transcriptional regulation and key stem cell signalling pathways (Fig. 5e, Extended Data Fig. 5i and Supplementary Table 5). Note that CDK8/19i did not inhibit the kinase activity of purified recombinant GSK3 or MEK (Supplementary Table 1), and CDK8/19 inhibition did not reduce the relative levels of phosphorylated ERK (Figs. 1k and 5f,g and Extended Data Fig. 5j). However, 2i treatment reduced CDK8/19 kinase activity (Fig. 5f and Extended Data Fig. 5k) and moderately downregulated CDK8 protein levels (Extended Data Fig. 5l). These data suggest that CDK8/19 inhibition occurs downstream of 2i, such that both treatments result in highly overlapping changes to phosphorylation sites.

CDK8/19i resets global RNA Pol II loading similar to 2i. 2i and CDK8/19i induce similar phosphoproteomic changes (converging on transcriptional machinery) and similar transcriptomic changes. Thus, to understand how CDK8/19 inhibition phenocopies the transcriptome of 2i-induced naive pluripotency, we investigated the global regulation of RNA Pol II abundance on chromatin using chromatin immunoprecipitation with sequencing (ChIP–seq) in mouse PSCs in 2i or CDK8/19i, versus serum/LIF as control. Overall, total and Ser5-phosphorylated RNA Pol II genomic distribution was consistent with published resources[23,48,49] (Encode, https://www.encodeproject.org/). We observed that 2i globally increases RNA Pol II binding to promoters (Fig. 5h,i and Extended



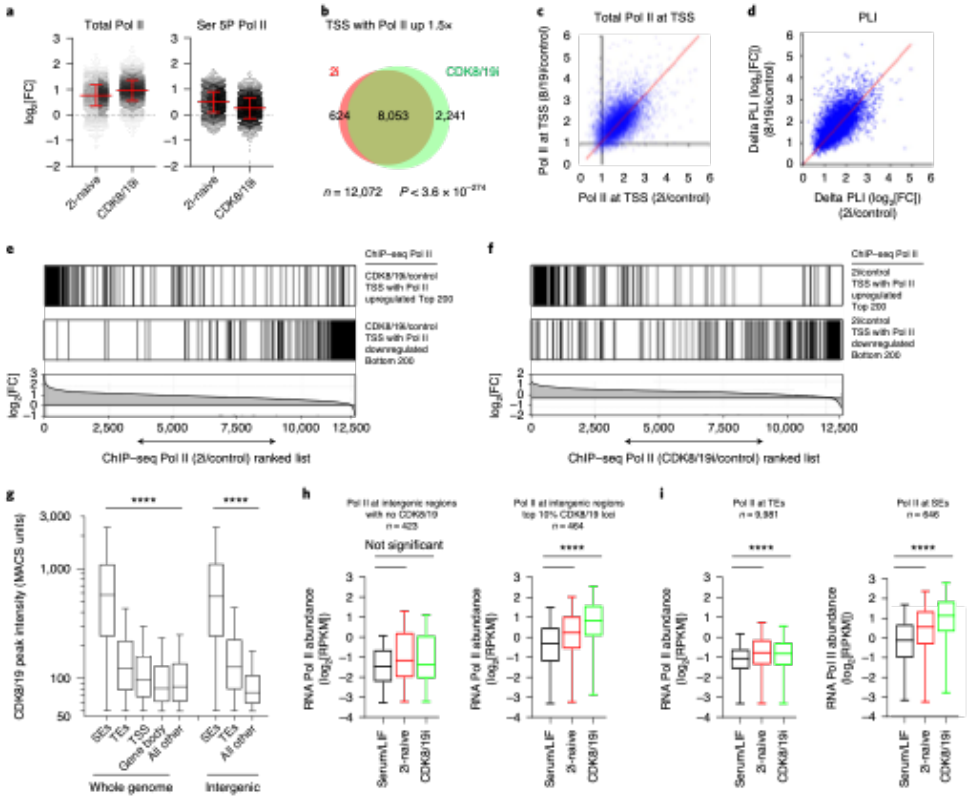


Fig. 6 | 2i and CDK8/19i exert similar effects on Mediator, RNA Pol II loading and enhancer activity. a, The change in total or phosphorylated Ser 5 (Ser 5P) RNA Pol II abundance at TSS, determined on the basis of ChIP-seq analysis in mouse PSCs in 2i or CDK8/19i conditions versus serum/LIF control. Data are mean ± s.d. Total Pol II, $n = 12,693$ TSS; Ser 5P, $n = 4,470$ TSS. **b**, Overlap and hypergeometric significance (the P value is indicated) of genes where RNA Pol II abundance at TSS increased by >1.5-fold in 2i and CDK8/19i (as described in **a**). TSS, $n = 12,072$. **c**,**d**, The fold change in RNA Pol II abundance at TSS (**c**) and the fold change in RNA Pol II PLI (**d**) at genes ($n = 12,693$) where RNA Pol II was detected in mouse PSCs in 2i or CDK8/19i versus control serum/LIF conditions. **e**,**f**, Genes with the largest changes in ChIP–seq RNA Pol II abundance are correlated between CDK8/19i and 2i-naive conditions. **e**, The individual promoter TSS with RNA Pol II loading altered in CDK8/19i by the greatest amount (top panel) or least amount (middle panel) versus serum/LIF control (top-200 TSS in each case) were compared against a ranked list (bottom panel) of differential Pol II loading on all TSS for 2i-naive versus serum/LIF control. **f**, Reverse comparison of the top-200 altered TSS in 2i-naive versus a ranked list of Pol II changes in CDK8/19i. **g**, CDK8/19 abundance in mouse PSCs, defined on the basis of ChIP-seq[2,3], peak-calling and grouped by genomic localization. The promoter TSS includes the TSS ±1 kb. Gene body includes exons, introns and the TTS ±1 kb. Statistical analysis was performed using unpaired two-tailed t-tests; ****$P < 0.0001$. **h**, RNA Pol II abundance in mouse PSC genomic regions without CDK8/19 binding (left, $n = 423$), or with the top-10% strongest CDK8/19-binding signals (right, $n = 464$) as defined in Extended Data Fig. 8d. Statistical analysis was performed using unpaired two-tailed t-tests with Welch's correction; ****$P < 0.0001$. RPKM, reads per kb of transcript per million mapped reads. **i**, RNA Pol II abundance in mouse PSCs at TEs (left, $n = 9981$) or SEs (right, $n = 646$), as defined previously[2,3]. Statistical analysis was performed using unpaired two-tailed t-tests with Welch's correction; ****$P < 0.0001$. For **a**–**f**, **h** and **i**, RNA Pol II abundance was measured using ChIP-seq (three pooled replicates). For the Tukey box plots in **g**–**i**, the centre lines show the median values, the box limits represent the upper and lower quartiles, and the whiskers show 1.5x the interquartile range. The number of regions per group, P value calculations, a list of defined loci and HOMER functional annotations are provided in Supplementary Tables 3 and 7.

Data Fig. 6a–c), which was confirmed by reanalysing independent data[25,68]. Notably, this global effect of 2i was phenocopied by CDK8/19i, regarding both total and Ser 5-phosphorylated RNA Pol II (Fig. 5h,i and Extended Data Fig. 6a–c). We measured the abundance of RNA Pol II at the promoter, gene body and transcription termination site (TTS) for each gene (Fig. 5j, Extended Data Fig. 6d and Supplementary Table 6). Consistent with previous analyses in mouse PSCs[21,68], most genes (90%) possessed a ratio of promoter to gene body loading of >2.0 (Fig. 5j and Supplementary Table 6). A comparison of RNA Pol II ratios between the promoter, gene body

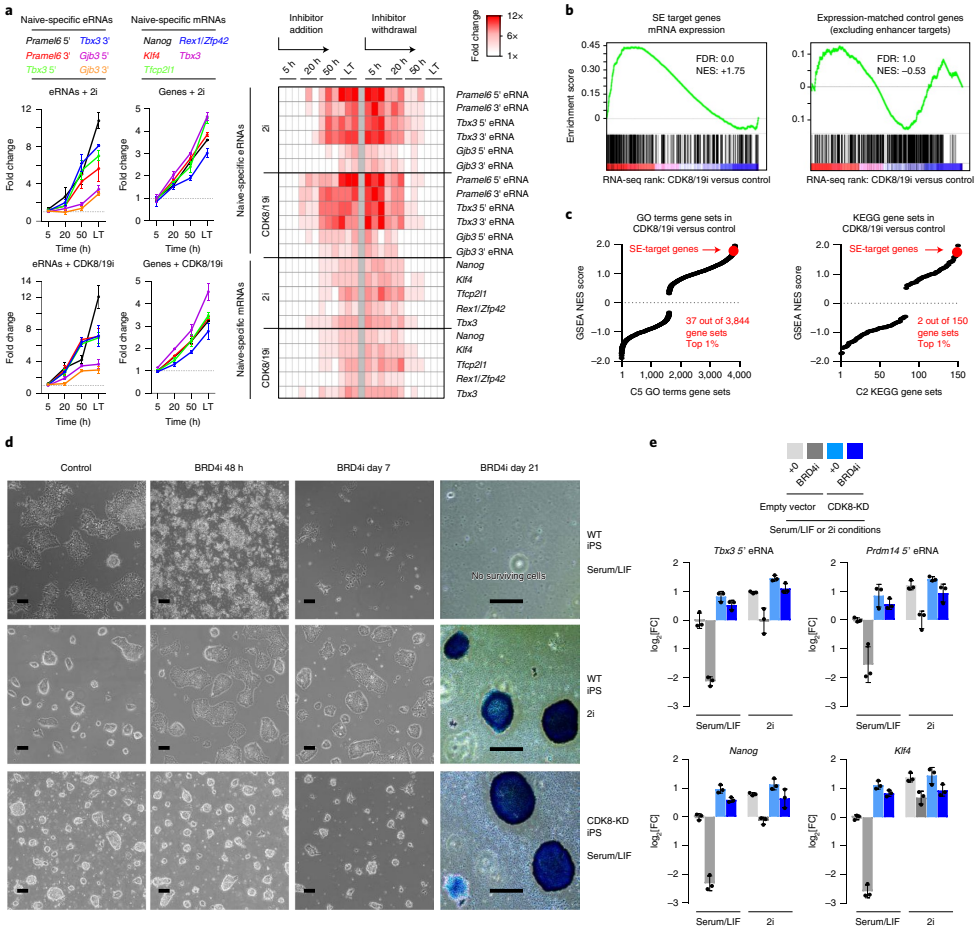


Fig. 7 | 2i and CDK8/19i hyperactivate naive-state enhancer activity, conferring resistance to enhancer destabilization. a, RT–qPCR analysis of pluripotency genes and naive-specific eRNA[71] abundance in mouse PSCs at time intervals after exposure to 2i or CDK8/19i, relative to standard serum/LIF control (left). LT, Long-term treatment. Data are mean ± s.e.m.; $n = 3$ independent experiments. Right, heat map summarizing the mean fold change in expression of the replicate experiments determined on the basis of RT–qPCR data shown on the left and from Extended Data Fig. 8g during inhibitor withdrawal. **b**,**c**, Selective upregulation of SE-target genes. **b**, GSEA of SE-target gene mRNAs in mouse PSCs adapted to CDK8/19i (left). $n = 3$ biological replicates. SE-target genes were defined as the single nearest gene on the basis of GREAT analysis (Methods) that were significantly upregulated (408 genes, $q < 0.001$). Right, expression-matched control genes (which were not predicted to be enhancer targets, but which are highly expressed similar to SE-target genes) show no significant change in levels (464 genes; $q = 1.0$). **c**, The relative specificity of SE-target gene upregulation by CDK8/19i was determined by comparison with databases of many other gene sets using GSEA. C5 GO terms ($n = 3,844$ gene sets) and C2 KEGG ($n = 150$ gene sets) as defined by the Broad Institute (Methods). SE-target genes lie within the top-1% most-significantly upregulated gene sets relative to these GO term or KEGG databases. For **b** and **c**, SE-target and expression-matched SE-non-target genes were defined as the single nearest gene on the basis of GREAT analysis (Methods). The SE-target gene list (408 genes) and the expression-matched control gene set (464 genes) are provided in Supplementary Table 7. **d**, Bright-field images of typical cell morphology after treatment with 500 nM BRD4i/JQ1 for 48 h (left) or 7 d (right) in WT iPSCs or in CDK8/19-dKO iPSCs stably expressing pMSCV-CDK8-kinase dead (CDK8-KD). Right, bright-field images of colonies that were fixed and stained for alkaline phosphatase at day 21/passage 5 of treatment with 500 nM BRD4i/JQ1. Data are representative of three independent cell experiments. **e**, RT–qPCR analysis of the expression of naive-specific eRNA[71] and marker genes after treatment with 500 nM BRD4i/JQ1 for 48 h. CDK8/19-dKO iPSCs with or without CDK8-KD were cultured with 2i or standard serum/LIF, as indicated. Data are mean ± s.d.; $n = 3$ clones.

or termination subregions of each gene indicated that 2i induces an increase in RNA Pol II binding selectively to the promoter region (Fig. 5j and Extended Data Fig. 6d). Importantly, this was recapitulated by CDK8/19i, increasing RNA Pol II binding to promoters at a similar magnitude to that observed in 2i-induced naive pluripotency, following a gene-specific pattern (Figs. 5j and 6a–f and Supplementary Table 6). Thus, 2i- and CDK8/19i-induced naive pluripotency is accompanied by widespread accumulation of RNA Pol II abundance at promoters. We also observed a correlation between changes in the abundance of RNA Pol II at promoters in 2i or in CDK8/19i conditions, as well as changes in mRNA expression for each gene (Extended Data Fig. 6e–i). In summary, gene-specific changes in RNA Pol II promoter loading may explain a significant proportion of the mRNA expression profile characteristic of 2i- or CDK8/19i-induced naive pluripotency.

CDK8/19i and 2i trigger activation of SEs. The primary role of Mediator is at enhancers, regulating the recruitment of RNA Pol II to promoters[4–6]. Using published ChIP–seq datasets[2,3] (Supplementary Table 7), we confirmed that CDK8/19 was enriched at promoter, typical enhancer (TE) and SE regions as previously defined in mouse PSCs[2] (Fig. 6g and Extended Data Fig. 7a–c). There is a strong correlation between the abundance of CDK8/19, Mediator subunits and other factors that are critical for enhancer activity[1,6] (such as, p300, CBP, Pol II or BRD4; Extended Data Fig. 7d); the highest levels of CDK8/19 occurred within SE regions (Fig. 6g and Supplementary Table 7); and, finally, putative target genes proximal to genomic CDK8/19-binding loci were highly enriched in preimplantation functions characteristic of pluripotent cell identity (Extended Data Fig. 8a–c). We therefore hypothesized that, in mouse PSCs, CDK8/19 inhibition might act through Mediator to trigger changes in enhancer activity, explaining the observed increase in RNA Pol II loading at promoters and regulation of pluripotent states. As CDK8/19 protein was particularly enriched at SE regions, we examined the impact of CDK8/19i on SE function. Enhancers contain RNA Pol II, which transcribes enhancer-derived RNAs (eRNAs)—a process that faithfully reflects enhancer activity[4,8,70]. We therefore measured the effect of CDK8/19i or 2i on the levels of RNA Pol II and eRNAs at SEs. Importantly, the abundance of RNA Pol II was selectively increased at CDK8/19-binding sites and, accordingly, RNA Pol II recruitment was also preferentially increased at SEs compared with TEs (Fig. 6h,i and Extended Data Fig. 8d). Consistent with this, mouse PSCs treated with 2i or CDK8/19i showed an increase in enhancer-derived eRNA levels, as well as RNA Pol II abundance, within enhancers specific for the naive state[7,1] (Fig. 7a and Extended Data Fig. 8e,f). Induction of naive-specific eRNAs and naive marker genes was an early event, occurring within 48 h after adding 2i or CDK8/19i, and it was rapidly reversible (Fig. 7a and Extended Data Fig. 8g). Finally, consistent with naive-specific enhancer activation, the expression levels of SE target genes were preferentially upregulated in both 2i and CDK8/19i conditions (Fig. 7b,c and Extended Data Fig. 8h). We conclude that, in PSCs, CDK8/19i and 2i hyperactivate existing SEs and upregulate SE-target genes in a manner that reinforces naive pluripotency.

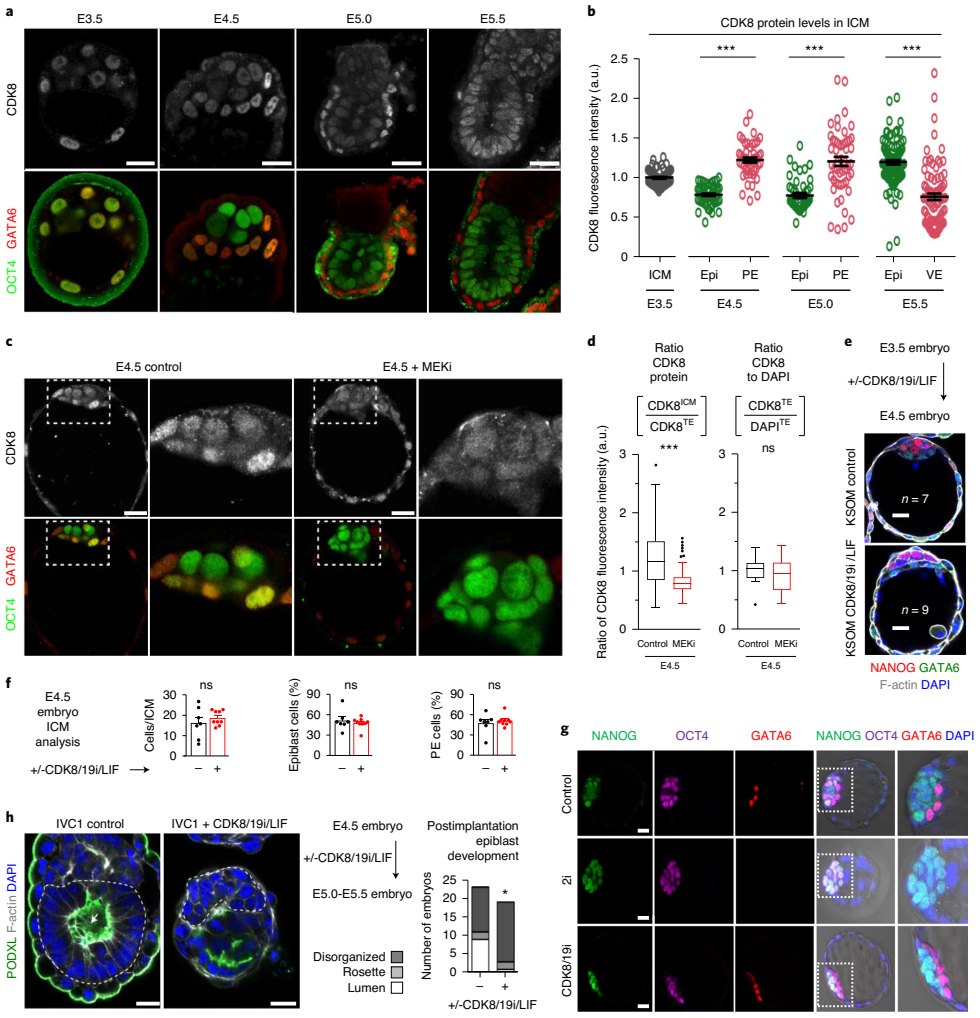
CDK8/19 inhibition compensates for BRD4 inhibition. Loss of Mediator function preferentially decreases the expression of enhancer target genes across multiple cell types[4–6,26–28]. In particular, BRD4 inhibition in primed state PSCs decreases the ability of Mediator to recruit RNA Pol II, and this results in the loss of Mediator-driven transcription, a collapse in pluripotency gene expression and differentiation[27,28] (Fig. 1a). Compared with primed PSCs, naive PSCs are highly resistant to the decreased Mediator activity and enhancer destabilization induced by BRD4 inhibition[28]. Interestingly, mouse PSCs lacking endogenous CDK8/19 and reconstituted with kinase-dead CDK8 were resistant to enhancer destabilization by BRD4 inhibition for ten passages (>3 weeks), maintaining naive morphology, and showed high expression of alkaline phosphatase, naive-specific pluripotency markers and naive-specific eRNAs, similar to 2i-naive PSCs (Fig. 7d,e and Extended Data Fig. 8i,j). Thus, PSCs expressing kinase-dead CDK8 phenocopy the robust resistance to enhancer destabilization that is characteristic of 2i-naive PSCs.

The roles of CDK8/19 during early embryonic development. Given our observations that CDK8/19 inhibition stabilizes naive pluripotency, we investigated CDK8/19 function during early embryonic development. We focused on CDK8, which we found is highly expressed compared with CDK19 in both mouse and human PSCs (Extended Data Figs. 1k and 9a). Using a CDK8-specific antibody (Extended Data Fig. 1j), we detected CDK8 protein from the mouse zygote to morula (Extended Data Fig. 9b). Consistent

Fig. 8 | CDK8 expression in vivo and the role of Mediator during mouse preimplantation development. a,**b**, Immunofluorescence imaging and quantification of CDK8 expression at the indicated time points during early mouse embryo development. **a**, A single *z* section. **b**, CDK8 protein levels per nucleus, quantified per time point, relative to the internal controls. OCT4/GATA6 coexpression marks all ICM cells at E3.5. OCT4/GATA6 segregation from E4.5 to E5.5 marks epiblast (Epi/OCT4$^+$) and PE (PE/GATA6$^+$), which later forms visceral endoderm (VE) at E5.5. Embryo staging, CDK8 quantification and normalization (Methods). Data represent two independent experiments (representative images); ICM, $n=6$ embryos, $n=64$ nuclei; E4.5, $n=5$ embryos, nuclei: $n=51$ (Epi), $n=48$ (PE); E5.0, $n=5$ embryos, nuclei: $n=48$ (Epi), $n=52$ (PE); E5.5, $n=6$ embryos, nuclei: $n=100$ (Epi), $n=84$ (PE/VE). Data are mean ± s.d. Statistical significance was assessed using Kruskal–Wallis tests with Dunn's multiple-comparison test. **c**,**d**, CDK8 expression is repressed by MEK inhibition in vivo. Embryos were incubated for 48 h with or without MEKi (8-cell stage to E4.5 blastocyst). **c**, Immunofluorescence analysis of CDK8 protein expression in E4.5 blastocysts. Data are representative of three experiments. **d**, CDK8 protein levels per cell quantified in ICM or trophectoderm (TE), relative to internal controls, per *z* slice/image. Left, control, $n=7$ embryos, $n=55$ images; MEKi, $n=12$ embryos, $n=44$ images. Right, control, $n=22$ embryos, $n=22$ images; MEKi, $n=27$ embryos, $n=27$ images. For the Tukey box plot, the centre lines show the median values, the box limits represent the upper and lower quartiles, and the whiskers show 1.5× the interquartile range. Statistical significance was assessed using unpaired two-tailed *t*-tests; ***$P<0.001$; ns, not significant. **e**,**f**, CDK8/19 inhibition does not prevent Epi/PE segregation. E3.5 embryos were incubated for 24 h with or without CDK8/19i/LIF during Epi/PE segregation, and then assessed using immunofluorescence (representative images; **e**). **f**, Quantification of ICM cell number, and lineage allocation in ICM, defined as: Epi/NANOG$^+$; PE/GATA6$^+$; ICM: NANOG$^+$ or GATA6$^+$. Data are mean ± s.e.m. from two experiments. Statistical significance was determined using unpaired two-tailed *t*-tests. Number of embryos: control, $n=7$; CDK8/19i, $n=9$. **g**, 2i prevents PE formation in the ICM, but CDK8/19i does not. E3.5 embryos were incubated for 24 h with or without 2i or CDK8/19i during Epi/PE segregation, and then assessed using immunofluorescence with brightfield overlap (representative images; $n=8$ embryos per condition, from two independent experiments). Lineage allocation defined as: ICM/OCT4+, Epi/NANOG+, PE/GATA6+. **h**, CDK8/19 inhibition interrupts pre- to postimplantation morphogenic events. Preimplantation E4.5 embryos were cultured until E5.0 in vitro[35] with or without CDK8/19i/LIF. PODXL and F-actin staining determines the emergence of the epiblast lumen (future pro-amniotic cavity). Epiblast indicated by the white dashed lines in the left and right image, respectively). Right, Lumenogenesis was quantified. Statistical significance was determined using the χ^2 test; *P < 0.05. Data are from two experiments. Number of embryos: control, $n=15$; CDK8/19i, $n=16$. For **a**, **c**, **e**, **g** and **h**, scale bars, 20 µm.

with this, CDK8-KO zygotes cannot progress beyond the 4–8-cell stage[72], and we observed that CDK8/19i impaired the progression of zygotes to the two-cell stage (Extended Data Fig. 9c). CDK8 activity is therefore essential for the zygote to morula transition.

We next investigated the role of CDK8 post-morula. CDK8 mRNA expression declines until the blastocyst stage, both in mouse and human preimplantation embryos (Extended Data Fig. 9d–f). CDK8 protein expression per cell was homogenous in the mouse inner cell mass (ICM) at E3.5 (Fig. 8a,b). Interestingly—at E4.5, when the ICM segregates into the naive epiblast and the primitive endoderm (PE)—CDK8 protein levels diverged, with lower levels in the epiblast compared to PE (Fig. 8a,b and Extended Data Fig. 10a). This pattern was transient, and it reversed in postimplantation epiblast at E5.5 (Fig. 8a,b and Extended Data Fig. 10a). To further document that CDK8 levels are reduced in the naive epiblast, embryos were cultured from E3.5–E4.5 with MEK inhibitor (MEKi), which blocks PE formation and permits only the development of naive epiblast[73,74]. As expected, in MEKi, the ICM contained only naive epiblast cells and not PE, this simplified the quantification of CDK8 in the ICM and trophectoderm and we confirmed reduced CDK8 expression in the ICM (Fig. 8c,d). The CDK8-binding partner and essential activating subunit cyclin C also altered its nuclear–cytoplasmic ratio during this developmental window. Specifically, E4.5 epiblast contained significantly less nuclear cyclin C compared



with E5.5 epiblast in vivo (Extended Data Fig. 10b,c), and a similar pattern was observed when comparing between 2i-naive and primed state PSCs in vitro (Extended Data Fig. 10d). In summary, the emergence of naive pluripotency during embryo development at E4.5 coincides with decreased CDK8 expression and decreased availability of its essential subunit cyclin C. This parallels the effect of MEKi on CDK8 expression and stabilization of naive epiblast identity in PSCs in vitro (Extended Data Fig. 5).

We wondered whether inhibition of CDK8/19 affects the emergence of naive pluripotency. Similar to MEKi, CDK8/19i treatment during E3.5–E4.5 did not interfere with embryo naive epiblast development (Fig. 8e–g) and enabled the derivation of PSC lines. In contrast to MEKi, CDK8/19i permitted PE formation (Fig. 8e–g). This suggests that the critical roles of MEK for PE segregation are independent of CDK8/19, and agrees with our observation that MEK activity is unaffected by CDK8/19i (Fig. 5f,g).

Finally, we examined the importance of CDK8/19 activity during preimplantation-to-postimplantation epiblast developmental progression. We focused on lumen formation within the epiblast, which marks the initiation of morphogenesis downstream of naive pluripotency exit[35]. We found that CDK8/19i treatment during E4.5–E5.5 impaired embryo epiblast lumenogenesis (Fig. 8h; for spheroids, Fig. 2e). This indicates that CDK8/19 activity is required to support epiblast development, from the naive preimplantation to primed postimplantation embryonic stages, consistent with the significant increase in CDK8 expression that we observed at this time (Fig. 8a,b and Extended Data Fig. 9e).

These data suggest that CDK8/19 expression in early embryonic development mirrors its function—the transition from zygote to morula, and the formation of the postimplantation epiblast require CDK8/19 activity; the intervening naive ICM has low CDK8 expression and reduced nuclear cyclin C (a summary is provided in Extended Data Fig. 10e).

Discussion

Here we uncovered a role for the Mediator kinases CDK8/19 in defining the equilibrium between naive and primed pluripotent states, in both mouse and human pluripotent cells. Collectively, our data indicate the following model: 2i and CDK8/19i rapidly induce a highly overlapping set of phosphorylation changes focused on the transcriptional machinery, triggering enhancer hyperactivation, a global increase in RNA Pol II recruitment to promoters and resetting gene expression. This includes upregulation of eRNAs, as well as resetting the expression of endogenous retroviral and repeat elements, as part of this cell identity conversion. Further evidence supporting transcriptional stabilization of naive pluripotency includes repression of 2C fluctuation in PSC identity, similar to in 2i conditions. Thus, the ability of 2i and CDK8/19i to induce naive features seems to originate from their common effect on Mediator and RNA Pol II transcriptional activity. In support of this, SEs interact with more target promoters[75], engage in more long-range interactions[75] and display increased H3K27ac[76] in the naive state versus primed. Our model agrees with the concept that transitions in cell identity are driven by early reconfiguration of the active enhancer network, which resets the transcriptional machinery to the new program[70,71,77].

The evidence presented here suggests a signalling hierarchy; in particular, MEK inhibition results in CDK8/19 inhibition, whereas inhibition of CDK8/19 does not affect MEK activity. Accordingly, we observed that (1) the ability of MEK and GSK inhibition (2i) to induce naive features in PSCs in vitro is recapitulated by CDK8/19i; (2) 2i and CDK8/19i exhibit a 79.6% overlap in downstream phosphorylation changes; (3) CDK8/19 downregulation coincides with the emergence of naive state in vivo, when MEK–ERK signalling is decreased[18,19]; and (4) both MEK–ERK activation[18,19,30,73,74] and CDK8/19 (refs. [78,79]) drive postimplantation epiblast differentiation, a process that we found is impaired by CDK8/19i, and a period during which CDK8 is upregulated. Thus, we propose that CDK8/19 inhibition is a common downstream feature of naive-inducing medium cocktails. Further studies will elucidate how MEK–ERK signalling regulates CDK8/19 Mediator activity in PSCs. Interestingly, Mediator hyperactivation through CDK8/19 inhibition triggers cancer cell death[14], while we find a similar approach reinforces naive pluripotent identity. Cancer cells commonly develop novel oncogenic SEs, becoming addicted to a defined range of enhancer-driven transcription that seems to be sensitive to perturbation[9,80]. This provides an interesting parallel with MEK inhibition, which is also detrimental to many cancer cells, but beneficial to pluripotency.

Stabilization of the human naive pluripotent state in vitro is challenging and remains to be optimized[21,24]. Our understanding of stem cell identity indicates a continuum of molecular changes along a spectrum from naive to primed states, which also reflects the developmental path in early embryos[18,19,21,22]. Where does CDK8/19i position PSCs along this gradient? We found that CDK8/19 inhibition recapitulates the majority of the molecular characteristics that are associated with the primed-to-naive transition. However, other molecular features that are associated with the more-naive end of this spectrum are not recapitulated by CDK8/19 inhibition, particularly, global DNA hypomethylation, X-chromosome reactivation[66,67] and SSEA4 downregulation[24,39]. Achieving these last features of naive pluripotency seems to come at a price. Naive-inducing medium cocktails that are dependent on MEK inhibition can generate harmful side effects, specifically acute chromosomal instability and imprinting erasure[24,39,62,63]. Interestingly, those other cocktails which do not downregulate SSEA4 and produce modest DNA demethylation, are not associated with genomic instability[24,33,64]. Similarly, CDK8/19i installs many naive features in human cells while maintaining SSEA4, DNA global methylation and genomic stability. CDK8/19i-treated cells retain a normal karyotype after prolonged culture. We suggest that these important differences are due to CDK8/19i not impinging directly on MEK signalling.

In summary, CDK8/19i stimulates the recruitment of RNA Pol II by Mediator. This hyperactivates enhancers and stabilizes the transcriptional program of naive pluripotent cell identity. Thus, chemical inhibition of CDK8/19 may help to solve remaining challenges in unstable human naive PSC culture. Similarly, these principles of stabilizing cellular identity may apply to other contexts of cellular plasticity.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41556-020-0573-1.

Received: 2 May 2019; Accepted: 7 August 2020;
Published online: 28 September 2020

References

1. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
2. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
3. Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
4. Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
5. Jeronimo, C. & Robert, F. The Mediator complex: at the nexus of RNA Polymerase II transcription. *Trends Cell Biol.* **27**, 765–783 (2017).
6. Soutourina, J. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **4**, 262–274 (2017).

7. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
8. Fant, C. B. & Taatjes, D. J. Regulatory functions of the Mediator kinases CDK8 and CDK19. *Transcription* **2**, 76–90 (2019).
9. Chen, H. et al. A pan-cancer analysis of enhancer expression in nearly 9000 patient samples. *Cell* **173**, 386–399 (2018).
10. Clark, A. D., Oldenbroek, M. & Boyer, T. G. Mediator kinase module and human tumorigenesis. *Crit. Rev. Biochem. Mol. Biol.* **50**, 393–426 (2015).
11. Poss, Z. C. et al. Identification of mediator kinase substrates in human cells using cortistatin a and quantitative phosphoproteomics. *Cell Rep.* **15**, 436–450 (2016).
12. Knuesel, M. T., Meyer, K. D., Bernecky, C. & Taatjes, D. J. The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* **23**, 439–451 (2009).
13. van de Peppel, J. et al. Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol. Cell* **19**, 511–522 (2005).
14. Pelish, H. E. et al. Mediator kinase inhibition further activates super-enhancer-associated genes in AML. *Nature* **526**, 273–276 (2015).
15. Gonzalez, D. et al. Suppression of Mediator is regulated by Cdk8-dependent Grr1 turnover of the Med3 coactivator. *Proc. Natl Acad. Sci. USA* **111**, 2500–2505 (2014).
16. Bancerek, J. et al. CDK8 kinase phosphorylates transcription factor STAT1 to selectively regulate the interferon response. *Immunity* **38**, 250–262 (2013).
17. Galbraith, M. D., Donner, A. J. & Espinosa, J. M. CDK8: a positive regulator of transcription. *Transcription* **1**, 4–12 (2010).
18. Hackett, J. A. & Surani, M. A. Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **15**, 416–430 (2014).
19. Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
20. Galonska, C., Ziller, M. J., Karnik, R. & Meissner, A. Ground state conditions induce rapid reorganization of core pluripotency factor binding before global epigenetic reprogramming. *Cell Stem Cell* **17**, 462–470 (2015).
21. Weinberger, L., Ayyash, M., Novershtern, N. & Hanna, J. H. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. *Nat. Rev. Mol. Cell Biol.* **17**, 155–169 (2016).
22. Wu, J. & Izpisua Belmonte, J. C. Dynamic pluripotent stem cell states and their applications. *Cell Stem Cell* **17**, 509–525 (2015).
23. Marks, H. et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).
24. Liu, X. et al. Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. *Nat. Methods* **11**, 1055–1062 (2017).
25. Ying, Q.-L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
26. Bhagwat, A. S. et al. BET bromodomain inhibition releases the Mediator complex from select *cis*-regulatory elements. *Cell Rep.* **15**, 519–530 (2016).
27. Di Micco, R. et al. Control of embryonic stem cell identity by BRD4-dependent transcriptional elongation of super-enhancer-associated pluripotency genes. *Cell Rep.* **9**, 234–247 (2014).
28. Finley, L. W. S. et al. Pluripotency transcription factors and Tet1/2 maintain Brd4-independent stem cell identity. *Nat. Cell Biol.* **5**, 565–574 (2018).
29. Chambers, I. et al. Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
30. Ficz, G. et al. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
31. Dale, T. et al. A selective chemical probe for exploring the role of CDK8 and CDK19 in human disease. *Nat. Chem. Biol.* **11**, 973–980 (2015).
32. Takashima, Y. et al. Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
33. Gafni, O. et al. Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).
34. Theunissen, T. W. et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
35. Shahbazi, M. N. et al. Pluripotent state transitions coordinate morphogenesis in mouse and human embryos. *Nature* **552**, 239–243 (2017).
36. Chen, H. et al. Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat. Commun.* **6**, 7095 (2015).
37. Wang, J. et al. Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405–409 (2014).
38. Stirparo, G. G. et al. Integrated analysis of single-cell embryo data yields a unified transcriptome signature for the human pre-implantation epiblast. *Development* **145**, dev158501 (2018).
39. Pastor, W. A. et al. Naive human pluripotent cells feature a methylation landscape devoid of blastocyst or germline memory. *Cell Stem Cell* **18**, 323–329 (2016).
40. Wu, J. et al. Interspecies chimerism with mammalian pluripotent stem cells. *Cell* **168**, 473–486 (2017).
41. Wu, J. et al. Stem cells and interspecies chimaeras. *Nature* **540**, 51–59 (2016).
42. Gifford, W. D., Pfaff, S. L. & MacFarlan, T. S. Transposable elements as genetic regulatory substrates in early development. *Trends Cell Biol.* **23**, 218–226 (2013).
43. Göke, J. et al. Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell* **16**, 135–141 (2015).
44. Grow, E. J. et al. Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature* **522**, 221–225 (2015).
45. Theunissen, T. W. et al. Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell* **19**, 502–515 (2016).
46. Macfarlan, T. S. et al. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
47. Eckersley-Maslin, M. A. et al. MERVL/Zscan4 network activation results in transient genome-wide DNA demethylation of mESCs. *Cell Rep.* **17**, 179–192 (2016).
48. Kolodziejczyk, A. A. et al. Single cell RNA-sequencing of pluripotent states unlocks modular transcriptional variation. *Cell Stem Cell* **17**, 471–485 (2015).
49. Fidalgo, M. et al. Zfp281 coordinates opposing functions of Tet1 and Tet2 in pluripotent states. *Cell Stem Cell* **19**, 355–369 (2016).
50. Buecker, C. et al. Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 838–853 (2014).
51. Bulut-Karslioglu, A. et al. Inhibition of mTOR induces a paused pluripotent state. *Nature* **540**, 119–123 (2016).
52. Boroviak, T. et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. *Dev. Cell* **35**, 366–382 (2015).
53. Blakeley, P. et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development* **142**, 3613–3613 (2015).
54. Nakamura, T. et al. A developmental coordinate of pluripotency among mice, monkeys and humans. *Nature* **537**, 57–62 (2016).
55. Yan, L. et al. Single-cell RNA-seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131–1139 (2013).
56. Xue, Z. et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593–597 (2013).
57. Petropoulos, S. et al. Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell* **165**, 1012–1026 (2016).
58. Guo, G. et al. Naive pluripotent stem cells derived directly from isolated cells of the human inner cell mass. *Stem Cell Rep.* **6**, 437–446 (2016).
59. Chan, Y. S. et al. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
60. Sahakyan, A. et al. Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell* **20**, 87–101 (2017).
61. von Meyenn, F. et al. Impairment of DNA Methylation maintenance is the main cause of global demethylation in naive embryonic stem cells. *Mol. Cell* **62**, 848–861 (2016).
62. Yagi, M. et al. Derivation of ground-state female ES cells maintaining gamete-derived DNA methylation. *Nature* **548**, 224–227 (2017).
63. Choi, J. et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature* **548**, 219–223 (2017).
64. Di Stefano, B. et al. Reduced MEK inhibition preserves genomic stability in naive human embryonic stem cells. *Nat. Methods* **15**, 732–740 (2018).
65. Leitch, H. G. et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311–316 (2013).
66. Sahakyan, A. et al. Human naive pluripotent stem cells model X chromosome dampening and X inactivation human naive pluripotent stem cells model. *Stem Cell* 87–101 (2017).
67. Khan, S. A., Audergon, P. N. C. B. & Payer, B. X-chromosome activity in naive human pluripotent stem cells—are we there yet? *Stem Cell Investig.* **4**, 54–54 (2017).
68. Rahl, P. B. et al. c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
69. Williams, L. H. et al. Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. *Mol. Cell* **58**, 311–322 (2015).
70. Arner, E. et al. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* **347**, 1010–1014 (2015).
71. Respuela, P. et al. Foxd3 Promotes exit from naive pluripotency through enhancer decommissioning and inhibits germline specification. *Cell Stem Cell* **18**, 118–133 (2016).
72. Westerling, T., Kuuluvainen, E. & Makela, T. P. Cdk8 is essential for preimplantation mouse development. *Mol. Cell. Biol.* **27**, 6177–6182 (2007).
73. Yamanaka, Y., Lanner, F. & Rossant, J. FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**, 715–724 (2010).

74. Nichols, J., Silva, J., Roode, M. & Smith, A. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**, 3215–3222 (2009).
75. Lopes Novo, C. et al. Long-range enhancer interactions are prevalent in mouse embryonic stem cells and are reorganized upon pluripotent state transition. *Cell Rep.* **22**, 2615–2627 (2018).
76. Di Stefano, B. et al. The RNA helicase DDX6 controls cellular plasticity by modulating P-body homeostasis. *Cell Stem Cell* **25**, 622–638 (2019).
77. Factor, D. C. et al. Epigenomic comparison reveals activation of 'seed' enhancers during transition from naive to primed pluripotency. *Cell Stem Cell* **14**, 854–863 (2014).
78. Postlmayr, A., Dumeau, C. E. & Wutz, A. Cdk8 is required for establishment of H3K27me3 and gene repression by Xist and mouse development. Development 147, dev175141 (2020).
79. Feldmann, A., Dimitrova, E., Kenney, A., Lastuvkova, A. & Klose, R. J. CDK-Mediator and FBXL19 prime developmental genes for activation by promoting atypical regulatory interactions. *Nucleic Acids Res.* **48**, 2942–2955 (2020).
80. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional addiction in cancer. *Cell* **168**, 629–643 (2017).

Methods

A list of primers, antibodies, shRNAs and CRISPR–Cas9 gRNAs is provided in Supplementary Table 8.

Pluripotency and differentiation assays. *Mouse and human research.* Animal experimentation at the CNIO was performed according to protocols that were approved by the CNIO-ISCIII Ethics Committee for Research and Animal Welfare (CEIyBA). Animal experimentation at the University of Cambridge was approved by the Home Office, performed according to Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 and reviewed by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). Cdk8-flox/flox RERT-Cre mice were generated by D. Fisher (IGMM, Montpellier).

Human pluripotent stem cell studies were ethically approved at the CNIO, Madrid, by the Comisión de Garantías para la Donación y Utilización de Células y Tejidos Humanos and signed by Director of Instituto de Salud Carlos III (Nuevas fronteras en la Reprogramación Celular: Explotando la plasticidad cellular; 303). Studies at the IRB Barcelona, were approved by the Ethics Committee of the CMRB, and by Comisión de Seguimiento y Control de la Donación de Células y Tejidos Humanos del Instituto de Salud Carlos III and the Ministry of Health from Government of Catalonia (project numbers: 0336S/11730/2015; 0336S/11220/2016; 0336S/2473/2017; and 0336/747/2018).

Mouse cells and culture conditions. The following mouse PSCs were used: E14Tg2a.4 (wild-type parental, 129/Ola background) from BayGenomics/MMRRC resource, University of California; wild-type PSCs were derived at the Transgenic Mouse Unit of the CNIO from E3.5 C57BL6 blastocysts, or mixed background C57BL6/129 blastocysts; Rosa26–GFP and Tg.CAG–Katushka-red PSC lines were derived from 129-*Gt(ROSA)26Sor*$^{tm1(CAG-EGFP)Luo}$/J mice (Jackson, 006053) and from Tg.CAG-Katushka mice[81], respectively. *Nanog*-GFP knock-in mouse PSCs (TNGA, TON) were derived as described previously[29] and shared by the laboratory of A. Smith. MERVL-td:Tomato mouse 2C-reporter PSCs were shared by the laboratory of T. Macfarlan[46]. ZS mouse 2C-reporter PSCs were shared by the laboratory of M. Ko[82]. Mouse PSCs (ESCs and iPSCs) were routinely cultured on 0.1% gelatin-coated plates in a base medium of either serum/LIF (15% FBS) or knockout serum replacement (KSR, Invitrogen) 'KSR/LIF' (15% KSR) in DMEM (high glucose) basal medium with LIF (1,000 U ml^{-1}), non-essential amino acids, Glutamax and β-mercaptoethanol plus antibiotics. In cases in which the 2i two-inhibitor cocktail was used with mouse PSCs, 2i comprised 1 μM MEK inhibitor (PD0325901, Axon Medchem, 1408) plus 3 μM GSK3β inhibitor (CHIR 99021, Axon Medchem, 1386) as described previously[25]. Cultures were routinely tested for mycoplasma. Primary mouse embryo fibroblasts (MEFs; wild type, passage 2) were obtained at E13.5 from pure inbred C57BL6 background mice or from CDK8-flox/flox RERT-Cre mice. Human HEK293T cells were obtained from ATCC. All of the above-mentioned cells were maintained in DMEM and 10% FBS (Gibco), supplemented with antibiotics (penicillin–streptomycin, 100 U ml^{-1}). Reprogrammed iPSCs were initially derived and expanded on mitomycin-C-inactivated feeder cells on gelatin-coated plates, before transfer to gelatin only.

Human PSC resources. HERVH iPSCs were shared by the laboratory of Z. Izsvak (Max Delbruck Centre for Molecular Medicine)[37]. WIBR3 ESCs were shared by the laboratory of J. Hanna (Weizmann Institute of Science). OSCAR ESCs carrying inducible STAT3 were shared by the laboratory of P. Savatier (SBRI, Stem Cell and Brian Research Institute)[36]. H1 and H9 human ESCs, and CB5, D2#2 and D2#4 human iPSCs were shared by the laboratory of N. Montserrat (IBEC, Institute for Bioengineering).

Human PSC cell culture in the primed state. Human PSCs (H1, H9, WIBR3, HERVH, CB5, D2#2, D2#4 and OSCAR) were maintained in conventional primed conditions as described previously[33,36,37], specifically, by culture on growth-factor-reduced phenol-red-free Matrigel (BD Biosciences, 356231) with mTeSR1 medium (Stem Cell Technologies). Cultures were passaged every 5–7 d manually using either 2 mg ml^{-1} dispase (Gibco), 0.5 μM EDTA/1× PBS or Accutase (Gibco).

Resetting human PSCs from primed to naive state using 2i-based medium cocktail. The naive human pluripotent state was obtained using two methods. OSCAR PSCs were reset to the naive state with 2i (TL2i) or CDK8/19i (1.1 μM or 0.4 μM) plus rhLIF and STAT3 transgene induction, as described previously[36]. In a transgene-free approach, human PSCs were cultured in a 2i-based chemical cocktail[33] referred to here as 2i p38iJNKi. Cells were maintained on Matrigel (BD Biosciences, 356231) using mTeSR1 (Stem Cell Technologies). Medium was supplemented with 20 ng ml^{-1} of recombinant human LIF (Peprotech, as described previously[33]), 1 μM PD0325901 (MEKi, Axon Medchem), 1.5 μM CHIR 99021 (GSK3i, Axon Medchem), 10 μM SP600125 (JNKi, TOCRIS) plus 2 μM BIRB796 (p38i, Axon Medchem). To obtain and maintain the naive state using the 2i p38iJNKi medium cocktail, cells were selected at each passage by cytometry sorting for the top-10% *HERVH*-GFP levels or by repeated manual picking to select colonies with dome-shaped morphology. Conversion of human PSCs from primed to naive required three passages/rounds of selection over 14–18 d.

Resetting human PSCs from primed to naive pluripotent state using CDK8/19i. To adapt and maintain human PSCs to CDK8/19i culture (CDK8/19i-adapted), cells were maintained on Matrigel (BD Biosciences, 356231) using mTeSR1 (Stem Cell Technologies). Medium was supplemented with 20 ng ml^{-1} of recombinant human LIF (Peprotech), as described previously[33], plus 0.4 μM or 1.1 μM of CNIO-CDK8/19 inhibitor, or plus 10 μM of SenexinA-CDK8/19 inhibitor[83]. This adaption process was also successful with 10 μM SP600125 (JNKi, TOCRIS) plus 2 μM BIRB796 (p38i, Axon Medchem), but these additives were not required with CDK8/19i. Following background cell death in the first passage, colonies gradually became dome shaped within 10–14 d without additional selection, and could be expanded by passage, using 3–5 min treatment with Accutase (Gibco) or 0.5 μM EDTA/1× PBS to avoid confluency, usually every 5–7 d due to a slowdown in proliferation. The optimal CDK8/19i concentration was 1.1 μM for *HERVH*-GFP human iPSCs, whereas the optimal CDK8/19i concentration was 0.4 μM for all other human PSC lines.

Derivation of mouse ESCs. ESC line derivation was performed using standard methods. Eight-cell stage mouse embryos obtained from the oviducts of pregnant female mice were cultured in serum/LIF on mitomycin-C-inactivated MEF feeders plus 2i or CDK8/19i (added fresh every 2 d) until the emergence of colonies from hatched blastocysts. The feeders were not compatible with several days of CDK8/19i treatment; cells were therefore passed to fresh feeders every 2 d, then transferred to 0.1% gelatin only.

Derivation of mouse EpiSCs. PSCs in 2i/LIF cultured on gelatin were first induced to differentiate into epiblast-like cells over 48 h by seeding on fibronectin-coated plates (10 ng ml^{-1}) and switching to medium containing 1% KSR, N2B27, FGF2 (12 ng ml^{-1}) and activin A (20 ng ml^{-1})[30]. After 48 h, the cells were in a flat epiblast-like cell state, and the medium was switched to include 20% KSR, and expanded for five passages to stabilize the EpiSC primed state, which was confirmed by typical flat colony morphology and Fgf5 expression. EpiSC colonies were passaged as clumps.

Analysis of PSC self-renewal. Mouse or human PSC self-renewal and pluripotency was scored by colony morphology, cytometry (in mouse cells, *Nanog*-GFP heterogeneity and overall intensity, and costaining for ICAM1; in human cells, *HERVH*-GFP intensity, and assessing the expression of NANOG, OCT4, SSEA4 and TRA1-81), alkaline phosphatase staining (fixed cells; Promega, S3771), and using immunofluorescence and RT–qPCR (for pluripotency markers, indicated in each figure). The intensity of alkaline phosphatase staining was quantified by scoring colonies observed using bright-field microscopy in ten random fields of view per well.

Mouse PSC differentiation with LIF removal and retinoic acid. LIF was first removed for 24 h by culturing in LIF-free differentiation medium (as described for serum/LIF medium, except LIF was omitted). Next, retinoic acid was added (10 μM) from 24 h to 72 h, followed by LIF-free differentiation medium alone from 72 h to 96 h. Differentiation was also assessed using the same protocol of LIF-withdrawal except without adding retinoic acid.

Mouse PSC differentiation by hanging-drop culture and as embryoid bodies. PSCs were transferred to LIF-free differentiation medium (as described above) and suspended in hanging-drop culture at 1,000–5,000 cells per 20 μl for 48 h to form embryoid bodies (EBs), followed by transfer to suspension culture in low-adherence Petri dishes. Fresh medium was added every 3 d, and development of beating cells in cardiac centres was scored daily.

Morula aggregation and blastocyst microinjection in mouse chimaera assays. After ten passages in serum/LIF, 2i or CDK8/19i, mouse PSCs labelled constitutively with Rosa26–GFP or Tg.CAG–Katushka[81] underwent morula aggregation at E2.5 or blastocyst microinjection at E3.5 as described previously[25]. The extent of GFP$^+$ or Katushka-red$^+$ cell chimeric contribution was assessed on the basis of confocal fluorescence at E4.5 or embryos were introduced into CD1 pseudopregnant females for implantation, and collected at the following postimplantation time points: E6.5, E14.5 or E19.5. Chimaeras that developed to adulthood were assessed by coat colour contribution and capacity for germline transmission.

Cardiac-tissue- and endoderm-directed differentiation of EBs derived from hPSCs. Human PSC colonies were dissociated and cultured in suspension for 3 d to form EBs in DMEM/F12, 15% FBS, 2 mM L-glutamine, non-essential amino acids and penicillin–streptomycin. To generate endoderm, EBs were transferred to 0.1% gelatin-coated plates for 2 weeks in differentiation medium (DMEM, 20% FBS, 2 mM L-glutamine, 0.1 mM 2-mercaptoethanol, non-essential amino acids and penicillin–streptomycin). To generate cardiac tissue, differentiation medium was supplemented with 100 μM ascorbic acid (Sigma-Aldrich). In all conditions, EBs spontaneously gave rise to neural cell clusters.

Teratoma assays. For mouse PSCs, 10^6 cells in 100 μl were injected subcutaneously in nude mice. For human PSCs, 2×10^6 cells in 30 μl were injected into the testis of male SCID beige mice.

Mouse embryo manipulation and analysis. Mouse embryo collection, culture for preimplantation embryo development in vitro and fixation for immunofluorescence analysis was performed as described previously[74,35]. Pre- to postimplantation embryo development in vitro, immunofluorescence analysis of CDK8 and cyclin C levels in preimplantation mouse embryos, and lumenogenesis by mouse PSC embryoid formation in Matrigel were performed as described previously[35]; further details are available from the corresponding author on request.

Viral production and iPSC reprogramming. Viral production and iPSC reprogramming were performed as described previously[84]. In brief, retroviral and lentiviral supernatants were produced in HEK293T cells. Filtered supernatants were collected after 48 h, and added to recipient cells in four infections. Retroviral supernatants delivered exogenous CDK8 expression constructs and iPSC reprogramming vectors. Lentivirus supernatants delivered shRNA knockdown vectors and CRISPR–Cas9 vectors. A list of plasmids is provided in Supplementary Table 8.

Interspecies chimaera developmental potency. Primed human iPSCs were precultured with ROCK inhibitor for 24 h, prepared as a unicellular suspension and electroporated (Neon Transfection System; Invitrogen; 1 pulse at 1,400 V for 20 ms) with 10 µg of DNA constructs for constitutive tdTomato expression (PB-Hygro-PGK-CAG-tdTomato and PBase pCMV-Transposase). Cells were subsequently plated on Matrigel in mTeSR1 medium supplemented with ROCK inhibitor for 24 h, then antibiotic selection with 20 µg ml^{-1} hygromycin was applied for 12 d, before cytometric sorting for tdTomato constitutively labelled cells.

Sexually mature NZW female rabbits (HyPharm) were superovulated. Then, 60 h after artificial insemination, fertilized embryos (eight-cell stage; E1.5) were flushed from explanted oviducts using Euroflush (IMV Technologies) and cultured in RDH medium (1/3 volume of DMEM–GlutaMAX, 1/3 volume of RPMI–GlutaMAX and 1/3 volume of Ham's F10–GlutaMAX; Life Technologies) at 38 °C and 5% CO_2.

Human PSCs were dissociated into single-cell suspension with trypsin, and 5–10 cells were microinjected under the mucus coat and zona pellucida of morula eight-cell stage rabbit embryos, the day after collection. After microinjection, embryos were sequentially cultured in CDK8/19i medium for 4 h, followed by 20 h incubation with a 1:1 mixture of RDH:CDK8i medium and finally in RDH medium for extended in vitro culturing. After 24 h of in vitro culture, early blastocyst stage embryos (E3.5) were rinsed three times in embryo-holding medium (IMV Technologies) and treated with 5 mg ml^{-1} protease E (Sigma-Aldrich) for 3 min at 37 °C to digest the mucus coat and weaken the zona pellucida. Embryos were then rinsed three times in 199 HEPES medium (Sigma-Aldrich) and cultured in RDH medium for 3 d until the late-blastocyst stage (E5.5). Rabbit embryos were fixed in 2% paraformaldehyde for 20 min at room temperature, washed in PBS + 0.1% Tween-20 and permeabilized in PBS + 1% Tween-20 overnight at 4 °C. After 1 h blocking with 5% donkey serum, embryo immunofluorescence was performed as described previously[36]. A list of the antibodies used is provided in Supplementary Table 8.

Molecular methods. *Transcriptional CDK inhibitors.* The structure and characterization of the CNIO CDK8/19 inhibitor (CDK8/19i-47799) as well as notes on all of the other transcriptional CDK inhibitors used in this study are provided in Supplementary Table 1.

Small-molecule inhibitor characterization assays. In vitro quantitative assays of enzyme inhibitors with recombinant proteins were performed using the LanthaScreen Eu-Kinase Binding Assay (Invitrogen) for CDK8–cyclin C, CDK9–cyclin T; CDKs, DYRK1A, GSK3β, mTOR, PI3K, PIM1/2, FLT3, KDR, KIT, PDGR-α and SRC. A summary of data from the small molecule inhibitor characterization assays is provided in Supplementary Table 1.

Generation of CDK8/19-dKO iPSCs. To target mouse CDK19, we designed sgRNA against CDK19 exon 1, targeting 76 bp downstream of the ATG translation start site to generate indels (a schematic of which is provided in Extended Data Fig. 1i). sgRNA sequences and plasmid details are provided in Supplementary Table 8. Primary CDK8-flox/flox RERT-Cre MEFs of passage 1–4 were infected with lenti-CRISPR-Cas9 containing the CDK19 sgRNA (pLenti-CRISPRV2; Addgene, 52961) followed by selection with puromycin (1 µg ml^{-1}). CDK19-KO was assessed using western blot. The MEFs were reprogrammed to iPSCs, single clones were picked and expanded, and CRISPR-induced indels were characterized by sequencing the CDK19 target region for frameshift mutations. CDK19-KO iPSC clones were compared versus iPSC clones that retained wild-type CDK19 expression, and no effect of CDK19-KO was observed in MEFs or in iPSCs. CDK8-KO was induced by 6 d of culture with 0.5 µM 4-hydroxy-tamoxifen to induce Cre-mediated deletion of CDK8 exon 2 (a schematic of which is provided in Extended Data Fig. 1j). CDK8-KO was confirmed using allele-specific PCR (to demonstrate deletion of exon 2; Extended Data Fig. 1h) and using western blot (to demonstrate complete loss of CDK8 protein; Extended Data Fig. 1j,k).

Stable exogenous expression of CDK8. Wild-type CDK8 (CDK8-WT) and catalytically inactivated kinase-dead CDK8 (CDK8-KD; D173A) were cloned into pMSCV-puro-IRES-GFP (Addgene, 21654) using the BglII and HpaI restriction enzymes, and confirmed by sequencing. Retroviral supernatants were generated in HEK293T cells with the packaging plasmid pCl-Eco (Addgene, 12371), followed by retroviral expression into CDK8/19-dKO iPSCs. Two rounds of FACS-selection by GFP expression were performed to enrich for expressing cells, and CDK8-WT or CDK8-KD protein expression was confirmed by western blot (a schematic of which in addition to western blot data are provided in Fig. 1k and Extended Data Fig. 1l).

FACS cytometry. FACS analysis of SSEA1 or ICAM1 was performed using FlowJo v.9.6.2 as described previously[84]. Live-cell analysis of the *Nanog*-GFP used 2i-adapted mouse PSCs to define the threshold (95% of cells) for the homogenous *Nanog*-GFP^{high} population, against which other treatment groups were compared (Fig. 1a,b). Live-cell sorting for human PSCs carrying *HERVH*-GFP selected the top 10% GFP-expressing cells, as previously described[37]. The FACS gating strategy for live/dead cell discrimination is provided in Extended Data Fig. 9g.

Cell lysis, fractionation and western blot. Cell lysis, fractionation and western blots were performed as described previously[85]. A list of the antibodies used is provided in Supplementary Table 8. Nuclear/cytoplasmic fractionation was performed using the NE-PER kit (Thermo Fisher Scientific, 78833).

G-banding karyotype methodology. Subconfluent mouse and human PSC lines were arrested in metaphase by adding 0.02 µg ml^{-1} KaryoMax Colcemid (Gibco). Twenty metaphase spreads were analysed per condition.

Immunohistochemistry and immunofluorescence analysis of mouse embryos. Mouse tissues were fixed in formalin at 4 °C, embedded in paraffin block and sectioned at a thickness of 5 µm. Staining was performed using standard methods. A list of the antibodies is provided in Supplementary Table 8.

Mouse teratoma and embryoid body immunohistochemistry. Mouse teratoma and embryoid body immunohistochemistry analysis was performed as described previously[84]. A list of the antibodies used is provided in Supplementary Table 8.

Cell immunofluorescence. PSCs were grown on chamber slides using culture conditions indicated in each experiment. Confocal immunofluorescence staining and microscopy was performed as described previously[35,84] using a Leica SP5 microscope. A list of the antibodies is provided in Supplementary Table 8.

DNA methylation. Global DNA methylation status was quantified by mass spectrometry (MS). CpG methylation status at individual CpG sites of repeat DNA regions was assessed by DNA bisulphite conversion and pyrosequencing. A list of the primers used for PCR amplification and sequencing is provided in Supplementary Table 8.

Image analysis. All image analyses were performed using Fiji (http://fiji.sc).

Proteomics. *Full proteome quantitative analysis of five mouse ESC lines.* Five mouse ESC lines (ZS, TNGA, TON, BL6 and V6.4) were cultured in serum/LIF (as a control) or, additionally, with either 2i or CDK8/19i for >2 weeks. Cell pellets were collected by trypsinization, washed with cold 1× PBS and preserved immediately at −80 °C for further analysis. Protein sample preparation for MS, protein digestion, our scheme for isobaric labelling with iTRAQ8plex, detailed settings for high pH reverse-phase fractionation, detailed settings for the whole-proteome analysis using liquid chromatography coupled with tandem MS (LC–MS/MS) and bioinformatics analyses with the whole-proteome data were reported previously[84].

Phosphoproteome analysis of mouse PSC lines after 15 min of inhibitor treatment. Two mouse ES PSC lines (TON and ZS) were cultured in serum/LIF (as a control) or, additionally, with either 2i or CDK8/19i. Cells were treated with inhibitor for precisely 15 min, after which the cells were collected rapidly by scraping in ice-cold PBS, washed with ice-cold PBS, snap-frozen on dry ice and preserved at −80 °C for further analysis. Sample preparation for MS, protein digestion, isobaric labelling, phosphopeptide enrichment, micro high pH reverse-phase fractionation, settings used for phosphoproteome LC–MS/MS and bioinformatics analyses with phosphoproteomic data were performed as described previously[11,84].

Transcriptomics. *RNA isolation and RT–qPCR.* Total RNA was extracted (on-column; RNeasy kit with DNA digestion; Qiagen, 74104, 79254) and retrotranscribed into cDNA (Superscript Reverse Transcriptase; Biorad, 170–889). RT–qPCR was performed using the GoTaq qPCR Master Mix (Promega A6002) in an ABI PRISM 7700 thermocycler (Applied Biosystem). Input normalization of all RT–qPCR data was performed using the $2^{-\Delta\Delta C_t}$ method, using housekeeping genes *Actb* or *Gapdh* as indicated in each figure. A list of the primers used is provided in Supplementary Table 8.

RNA-seq transcriptomic analyses. The complete set of reads has been deposited in GEO (GSE112208 and GSE127186). A complete list of meta-analyses expression comparisons between this study and multiple mouse and human published datasets, in vitro and in vivo, is provided in Supplementary Table 3.

For RNA-seq analysis in mice, samples of 1 µg of total RNA (RIN numbers: 9.8–10; Agilent 2100 Bioanalyzer) were used. PolyA+ fractions were processed using the TruSeq Stranded mRNA Sample Preparation Kit (Agilent). Adapter-ligated library was completed by PCR with Illumina PE primers (8 cycles) and sequenced for 40 bases in a single-read format (Genome Analyzer IIx, Illumina).

For RNA-seq analysis in human cells, samples of total RNA (RIN numbers: 9.0–10; Agilent 2100 Bioanalyzer) were used. For library construction, 10 ng of total RNA samples were processed using the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech) according to the manufacturer's instructions. The resulting cDNA was processed using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, E7645). Adapter-ligated libraries were completed by PCR (8 cycles), and sequenced for 50 bases in a single-read format, (Illumina HiSeq2500).

Reads were aligned to the reference mouse genome (GRCm38/mm10) or the human genome (GRCh37/hg19) using TopHat-2.0.4 (using Bowtie v.0.12.7 and Samtools v.0.1.16, allowing for two mismatches and five multihits). Transcript assembly, estimation of abundance and differential expression were calculated using Cufflinks v.1.3.0. When comparing samples, total read numbers were normalized and visualized using SeqMiner v.1.3.3e or Integrated Genome Viewer from the Broad Institute (http://software.broadinstitute.org/software/igv/)

Functional analyses of differential gene expression. Lists of differentially expressed genes are provided in Supplementary Table 2 for mouse PSCs adapted to control serum/LIF, +2i or +CDK8/19i; and Supplementary Table 3, for human PSCs adapted to control/primed, +2i or +CDK8/19i. Genes were ranked using the FDR q-value statistic to identify significant genes ($q < 0.05$), and then by fold change in expression. Venn diagrams and hypergeometric testing were performed to assess any significant overlaps. GSEA (GSEA_Pre-ranked) was performed with MsigDB Hallmarks, C5 GO terms, C2 Curated, KEGG, Reactome and NCI databases, using the standard settings, and with 1,000 permutations for Kolmogorov–Smirnoff correction for multiple testing. GSEA enrichment data were obtained and ranked according to FDR q value (significance threshold, $q < 0.25$). Heat maps of expression data were generated using GenePattern. Rank–rank hypergeometric overlap (RRHO) analysis was performed using the ranked list of $\log_2$-transformed fold changes in gene expression or RNA Pol II abundance using the standard settings[85]. The colour intensity of the RRHO heat map indicates the $-\log_{10}$-transformed P value after Benjamini–Yekutieli correction of the hypergeometric overlap (http://systems.crump.ucla.edu/rankrank/rankranksimple.php)[85].

Analysis of repeat sequences and ERV expression was performed using Repbase datasets for rodent or human repeat elements and featureCounts. In Extended Data Fig. 3h, the total fragments per kb of transcript per million mapped reads for RNA expression of LINE L1 subtypes was calculated by grouping and summing by family, and was then arranged by evolutionary age[86]. A full list of three biological replicates for each viral subtype and the calculation for the summary of each viral LINE L1 family are provided in Supplementary Table 2.

Differential gene expression comparison of published mouse and human studies. Gene expression changes have been comprehensively characterized in mouse, primate and human PSCs in response to overexpression of transcription factors after culture in various medium cocktails or in vivo during the development of the mouse or human embryos[87,88]; a full list of datasets and references used here is provided in Supplementary Tables 2 and 3. We used the marker gene sets for each developmental stage to perform GSEA on the ranked list of genes that were up- or downregulated in the cellular studies of mice and humans. We also performed the analysis in reverse, comparing the gene sets of significantly differentially expressed mRNAs that were up- or downregulated in our cells versus the complete ranked list of differential gene expression in other studies. GSEA results are shown in Fig. 4e (mouse) and Fig. 4k (human). The readout is the NES. Data with $P < 0.05$ and $q < 0.05$ were considered to be significant and are indicated by asterisks in heat maps of GSEA NES scores.

ChIP–seq and genomic analyses. ChIP–qPCR was performed as described previously[84]; a list of the primers and antibodies used is provided in Supplementary Table 8. ChIP–seq was performed as described previously[2,3,68,84]. We performed six biological replicates for each condition (three conditions: serum/LIF, 2i and CDK8/19i) and for each antibody (three antibodies: anti-total RNA Pol II, anti-Ser 5P-RNA Pol II and control IgG). Three replicates were used for ChIP–qPCR validations, and the other three replicates were pooled for sequencing. Note that our RNA Pol II ChIP–seq data in this study for serum/LIF and 2i-naive cells very closely match previous ChIP–seq data involving the same comparison, that is, mouse PSCs in primed versus 2i-naive states[23,69] (compare Fig. 5h,i with Extended Data Fig. 6c).

Promoter and gene body regions were defined and RNA Pol II total and Ser 5P abundance along genes was calculated as described previously by Young and colleagues[68,84] (a schematic of which is provided in Fig. 5; Pol II abundance data are provided in Supplementary Table 6). RNA Pol II abundance was assessed by normalizing the total number of reads between treatments, and using featureCounts to calculate the background-subtracted $\log_2$-transformed RPKM of RNA Pol II abundance in the indicated regions. TSS and the transcription termination zone were identified using the Database of Transcriptional Start Sites (http://dbtss.hgc.jp). Metagenes were aligned to ±5 kb or ±2 kb around the TSS, and visualized using SeqMiner.

The promoter, gene body and transcription termination zone, as well as the ratios between these three regions for each gene (Fig. 5j, Extended Data Fig. 6d and Supplementary Table 6), were defined similar to that described previously[68,84]. Total and Ser 5P RNA Pol II abundance were quantified at the promoter, gene body and transcription termination zone for 31,167 RefSeq gene loci in which the transcription start and stop sites are known (Supplementary Table 6), in four steps, similar to previous reports[68]: (1) the number of reads per nucleotide was computed using BEDTools genomecov; (2) to extend this number to the number of reads per gene promoter or gene body, BEDTools map was used; (3) to correct for region size, the RNA Pol II abundance was calculated as follows: ((number of reads in region/region size) × scaling factor) × 10^5, where the scaling factor = (total number of reads in sample/genome length); (4) for the analysis of Pol II abundance according to inhibitor treatment, genes were first filtered for high-confidence Pol II detected at a threshold of >3,000 units at the promoter, and detected in all three conditions (serum/LIF, 2i or CDK8/19i), yielding 12,072 genes (Supplementary Table 6, for filtering and calculations). In Fig. 5h and Extended Data Fig. 6a,c, genes were arranged in rank by the abundance of RNA Pol II at the promoter region in the control serum/LIF condition.

CDK8/19 enrichment across the genome of wild-type mixed background V6.5 (C57BL/6–129) mouse PSCs was determined using a published dataset (GSE44286, GSM1082346) as previously described[2,3]; peak calling was performed with MACS v.1.4.1 using the standard settings and compared to the input negative control. Note that the ChIP antibody for this ChIP–seq analysis (Santa Cruz, sc-1521) is reported to bind to both CDK8 and CDK19 (ref. [89]). Peak annotation within local genomic features was performed using HOMER and the enhancer regions previously defined as constituent regions of TEs ($n = 9{,}981$) or SEs ($n = 646$)[2,3], and of SE extended regions ($n = 231$) as defined previously[2,3], where enhancers were defined by coenrichment for OCT4, SOX2, NANOG and MED1. Details about peak calls, CDK8/19 abundance at called peaks and loci annotations are provided in Supplementary Table 7. Naive-specific or primed-specific enhancer regions were defined by filtering the PREStige database of enhancers[77], which identifies enhancers by enrichment of H3K4me1 monomethylation in multiple tissues and lineages. Using the PREStige dataset, we identified enhancer regions with H3K4me1 enrichment of >20 units, and that were specific to only preimplantation naive PSCs or postimplantation EpiSCs versus all other tissue-specific enhancer regions listed in the database (~120,000), by subtracting overlapping enhancers (1 bp overhang threshold) as outlined in Extended Data Fig. 8e. Source data are available online, including lists of naive ESC-specific enhancers ($n = 1{,}424$) and EpiSC-specific enhancers ($n = 1{,}005$). To identify the single nearest target gene to each PSC SE and analyse their biological functions, we performed an analysis using GREAT v.3.0.0 (ref. [90]) with the standard settings, using the list of CDK8/19 peaks identified above (Supplementary Table 7). We used GREAT v.3.0.0 for GO analysis of target-gene functions, reporting the $-\log_{10}$-transformed binomial P value with conservative Bonferroni correction for multiple-hypothesis testing[90]. Correlation matrix of ChIP–seq data in Extended Data Fig. 7d was produced using Morpheus, which is available from the Broad Institute (https://software.broadinstitute.org/morpheus/).

For Fig. 7b,c, GSEA was run with a gene set of the single nearest genes to SEs (as identified in the GREAT analysis described above using the standard settings (GREAT v.3.0.0)[90], using the SEs that were previously described in mouse PSCs[2,3] versus the ranked list of differential gene expression determined by RNA-seq for serum/LIF control compared with CDK8/19i-adapted mouse ES PSCs. Source data are available online, including lists of SE-target and expression-matched control genes.

For Extended Data Fig. 8h, GREAT analysis using the standard settings (GREAT v.3.0.0)[90] was used to identify the set of single nearest genes ($n = 3{,}553$ genes) to enhancer regions that were previously identified in mouse PSCs ($n = 10{,}627$)[2,3]. The $\log_2$-transformed fold change in RNA expression of these genes from this study was then ranked from high to low (serum/LIF versus 2i; serum/LIF versus CDK8/19i), and the extent of the overlap, calculated using a hypergeometric test of significance of these two ranked lists, is shown as a heat map in Extended Data Fig. 8h, performed using RRHO[85] with the standard settings (http://systems.crump.ucla.edu/rankrank/rankranksimple.php). The colour intensity of the RRHO heat map indicates the $-\log_{10}$-transformed P value after Benjamini–Yekutieli correction of the hypergeometric overlap.

Statistics and reproducibility. Unless otherwise specified, quantitative data are presented as mean ± s.d. and significance was assessed using two-tailed Student's t-tests; $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. RRHO was performed as described previously[85] using the standard settings and after Benjamini–Yekutieli correction of the hypergeometric overlap. Except when annotated otherwise, each experiment shown was performed three times with similar outcomes. Statistical analyses are described in detail for each panel. No statistical methods were used to predetermine the sample size. In brief, for differential gene expression using RNA-seq analysis, a threshold of $q < 0.05$ or $q < 0.01$ was applied, as indicated in each case. For GSEA, the standard threshold for significance was applied ($P < 0.05$ and $q < 0.25$). Genes that were differentially expressed in the RNA-seq

analysis were called using DESeq2 or Cufflinks v.1.3.0 (as described above). Immunofluorescence image analysis is described in detail above in the section about embryo analysis. Statistics were performed using MACS for peak calling of the ChIP–seq experiments. Statistical analyses of ChIP–qPCR, quantitative RT–qPCR and cell culture experiments was performed using Prism (v.7.03; GraphPad) or Microsoft Excel.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

RNA-seq and ChIP–seq data are available from the GEO database under accession numbers GSE112208 and GSE127186. The MS proteomics data are available from the ProteomeXchange Consortium/PRIDE repository under the dataset identifier PXD009200. Details on the published datasets used in Fig. 4e,k are provided in Supplementary Table 3. All other data supporting the findings of this study are available from the corresponding author on reasonable request. Source data are provided with this paper.

References

81. Diéguez-Hurtado, R. et al. A Cre-reporter transgenic mouse expressing the far-red fluorescent protein Katushka. *Genesis* **49**, 36–45 (2011).
82. Zalzman, M. et al. *Zscan4* regulates telomere elongation and genomic stability in ES cells. *Nature* **464**, 858–863 (2010).
83. Porter, D. C. et al. Cyclin-dependent kinase 8 mediates chemotherapy-induced tumor-promoting paracrine activities. *Proc. Natl Acad. Sci. USA* **109**, 13799–13804 (2012).
84. Lynch, C. J. et al. The RNA polymerase II factor RPAP1 is critical for mediator-driven transcription and cell identity. *Cell Rep.* **22**, 396–410 (2018).
85. Plaisier, S. B., Taschereau, R., Wong, J. A. & Graeber, T. G. Rank–rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res.* **38**, e169 (2010).
86. Castro-Diaz, N. et al. Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev.* **28**, 1397–1409 (2014).
87. Papadopoulou, T., Kaymak, A., Sayols, S. & Richly, H. Dual role of Med12 in PRC1-dependent gene repression and ncRNA-mediated transcriptional activation. *Cell Cycle* **15**, 1479–1493 (2016).
88. Xiang, L. et al. A developmental landscape of 3D-cultured human pre-gastrulation embryos. *Nature* **577**, 537–542 (2020).
89. Broude, E. V. et al. Expression of CDK8 and CDK8-interacting genes as potential biomarkers in breast cancer. *Curr. Cancer Drug Targets* **15**, 739–749 (2015).
90. McLean, C. Y. et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

Acknowledgements

We thank A. Smith, T. MacFarlan, Z. Izsvák, M. Ko, J. Hanna, D. Grégoire and U. Hibner for gifts of reagents; N. Prats and L. César Fernández for assistance, as well as staff at the CNIO and IRB core facilities. M.N.S. was funded by a Leverhulme Trust early career fellowship. Work in the laboratory of M.Z.-G. was funded by the Wellcome Trust (098287/Z/12/Z) and the European Research Council (ERC) (669198). I.C. was funded by the Secretaria d'Universitats i Recerca de la Generalitat de Catalunya and European Social Fund. I.A. and P.S. were supported by the Fondation pour la Recherche Medicale (DEQ20170336757), Infrastructure Nationale en Biologie et Santé INGESTEM (ANR-11-INBS-0009), IHU-B CESAME (ANR-10-IBHU-003), LabEx REVIVE (ANR-10-LABX-73), LabEx DEVweCAN (ANR-10-LABX-0061) and LabEx CORTEX (ANR-11-LABX-0042) of University of Lyon within the programme 'Investissements d'Avenir' (ANR-11-IDEX-0007). Research by J.P., S.M. and C.B.-A. was supported in part by a grant from the Spanish Ministry of Economy and Competitiveness (SAF2013-44267-R) and by the CNIO. Work in the laboratory of D.F. was funded by the Institut National du Cancer (PLBIO10-068 and PLBIO15-005) and the Ligue National Contre le Cancer (EL2018.LNCC/DF). Work in the laboratory of N.M. was funded by the ERC, under the European Union Horizon 2020 research and innovation programme (StG-2014–640525_REGMAMKID), the Spanish Association Against Cancer (AECC/LABAE16006), Carlos III Health Institute (Red TerCel, CardioCel, RD16/0011/0027), Ministry of Economy and Competitiveness (MINECO) projects SAF2017–89782-R, SAF2015–72617-EXP and RYC-2014–16242, and the CERCA/Government of Catalonia (2017 SGR 1306). Work in the laboratory of S.O. was funded by SAF2013–44866-R from MINECO Spain. Work in the laboratory of M.F.F. was funded by Plan Nacional de I+D+I 2013–2016/FEDER (PI15/00892, to M.F.F. and A.F.F.); the ISCIII-Subdireccion General de Evaluación y Fomento de la Investigación and Plan Nacional de I+D+I 2008–2011/FEDER (CP11/00131, to A.F.F.); IUOPA (to M.I.S.); and the Asturias Regional Government (GRUPIN14–052, to M.F.F.). The IUOPA is supported by the Obra Social Liberbank-Cajastur, Spain. Work in the laboratory of M.S. was funded by the CNIO, the IRB and by grants from Spanish Ministry of Economy co-funded by the European Regional Development Fund (SAF2017-82613-R), ERC (ERC-2014-AdG/669622), Botin Foundation, Banco Santander (Santander Universities Global Division), laCaixa Foundation and Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of Catalonia (Grup de Recerca consolidat 2017 SGR 282). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

C.J.L. designed and performed most of the experiments with mouse cells and embryos, contributed to bioinformatics data analysis and cowrote the manuscript. R.B. designed and performed most of the experiments with human cells, and provided general experimental support. A.M.-V. performed proteomic and bioinformatics analysis. M.N.S. performed embryo experiments, immunofluorescence and data analysis. S.N.-P., I.C., L.R.-G., N.A. and M.M.-M. contributed to experimental work and data analysis. C.T. and E.G. contributed to research with human PSCs and performed differentiation, immunofluorescence and confocal analysis of these experiments, supervised by N.M.; O.G.-C., G.G.-L. and C.S.-O.A., contributed to bioinformatics analyses. C.B.-A., S.M. and J.P. selected, synthesized and characterized small-molecule inhibitors. S.O. provided reagents, contributed to experimental design and supervised mouse embryo research. I.A. and P.S. performed human–rabbit interspecies chimaera and STAT3 assays. S.P., E.S., A.C. and D.F. generated the CDK8-KO mouse, provided reagents and performed additional inhibitor analyses. A.F.F., M.I.S. and M.F.F. performed DNA methylation analysis. P.S., D.F., J.M. and M.Z.-G. provided reagents, discussion and revisions. M.S. designed and supervised the study, secured funding, analysed the data and cowrote the manuscript. All of the authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

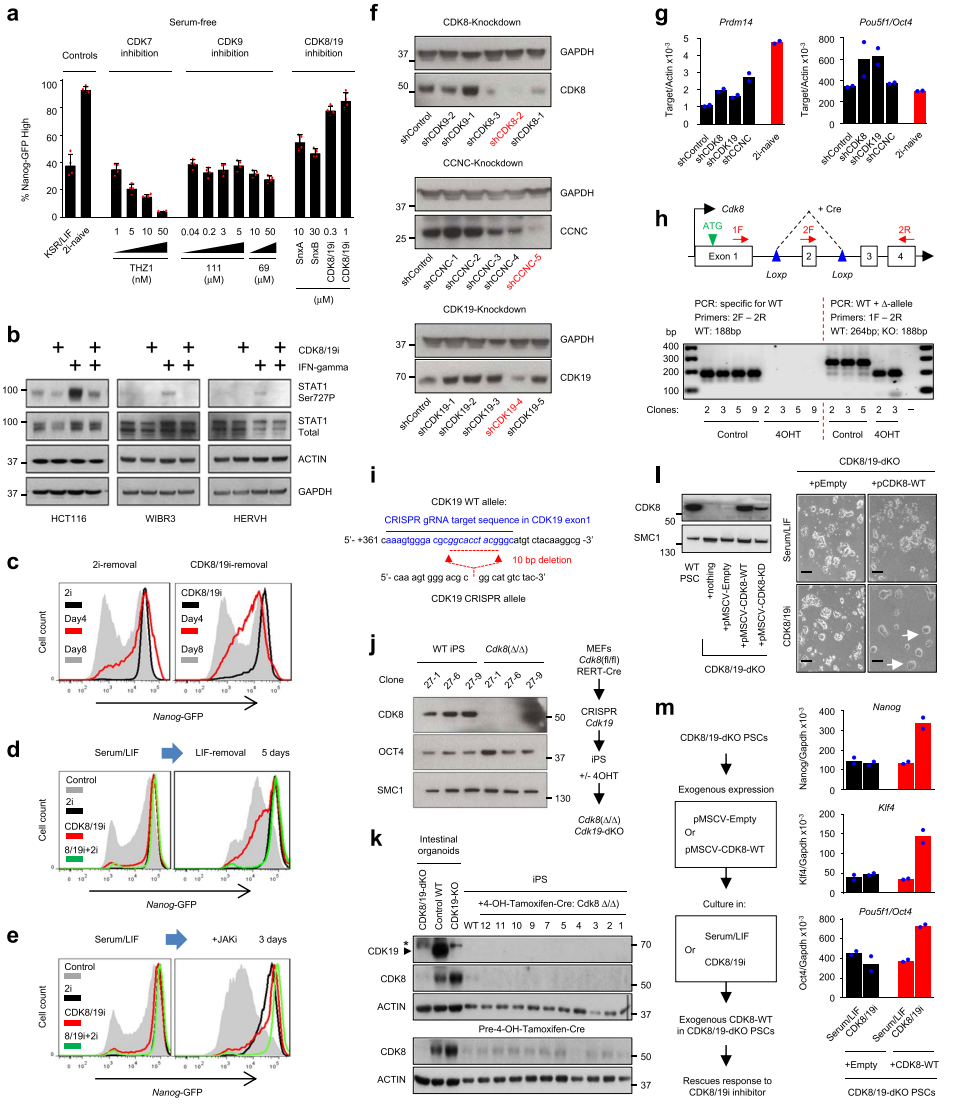
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41556-020-0573-1.

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-020-0573-1.

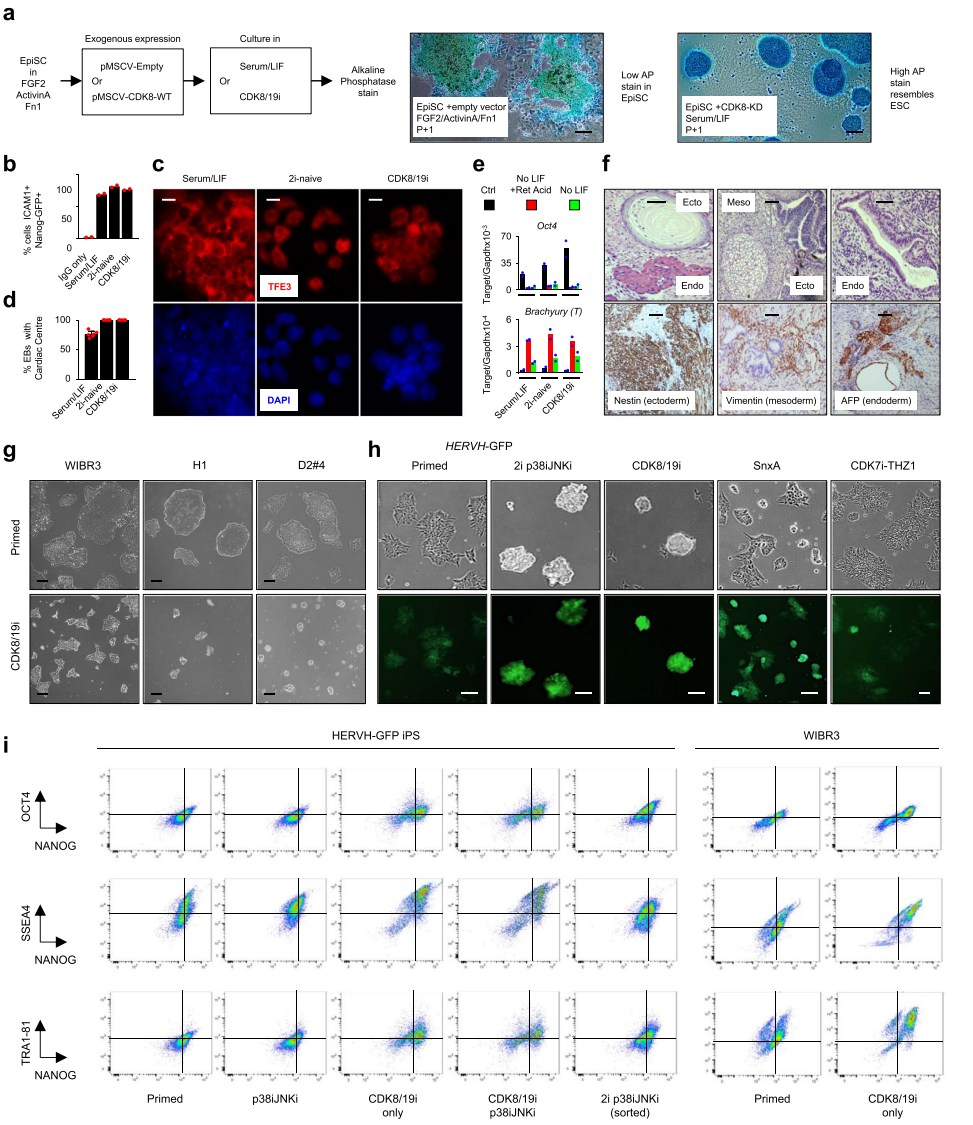
Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.



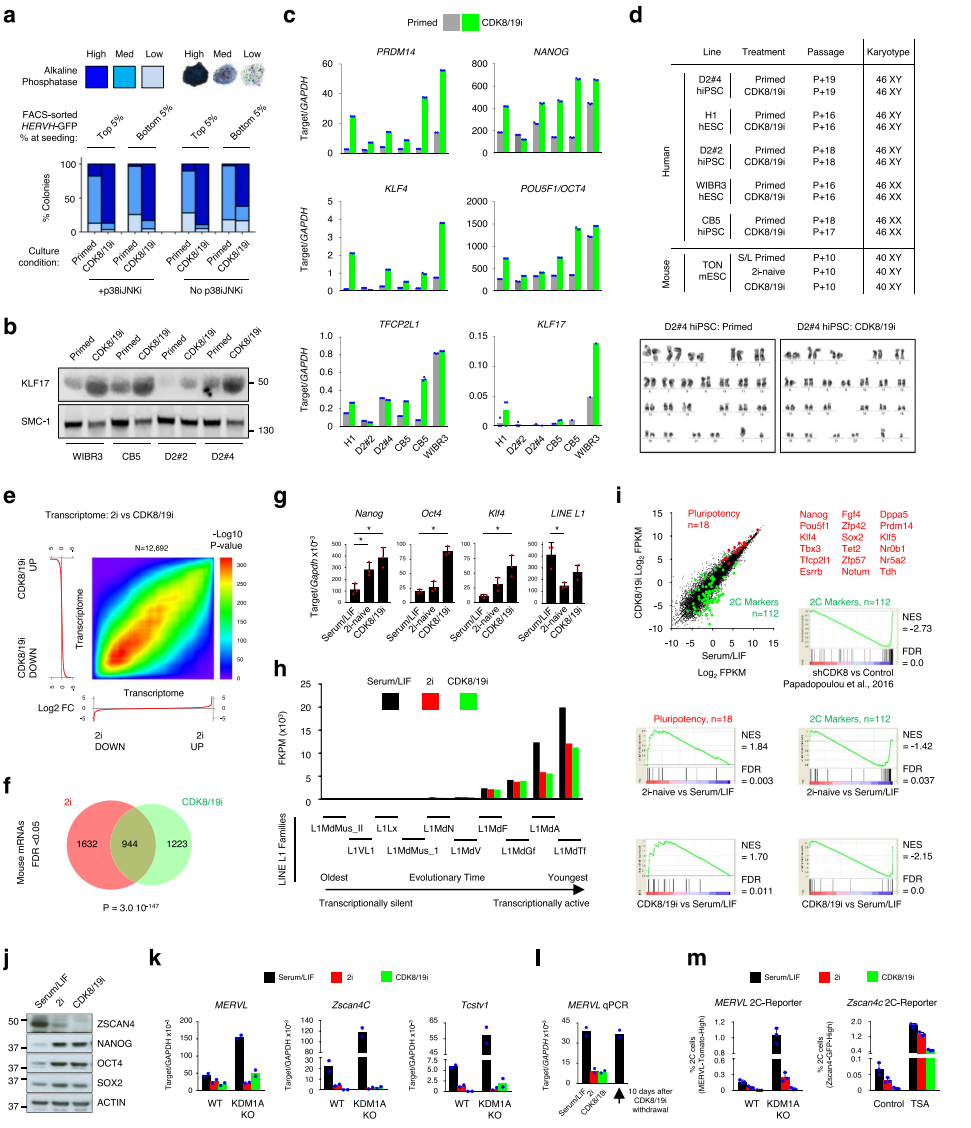
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | An Inhibitor Screen for Promotion of ES Naïve State identifies a distinct role for Mediator kinase activity. a, FACS: Percent cells Nanog-GFP^{high} in indicated treatments with serum-free KSR/LIF culture. Mean±SD, n=4 independent experiments. **b**, Western blot: CDK8-target STAT1-Ser727-Phosphorylation; n=3 experiments with indicated cell lines. STAT1-Ser727P induction by 3h Interferon-γ, ± simultaneous 1μM CDK8/19i. **c**–**e**, FACS histograms: mouse *Nanog*-GFP knockin reporter PSC previously adapted to 2i or CDK8/19i, tested at intervals. Decreased proportion of *Nanog*-GFP^{High} cells indicates loss of naïve state. **c**, Changes similar by 2i-removal or CDK8/19i-removal. **d**, 2i protects Nanog-GFP^{High} cells longer than CDK8/19i following LIF removal. **e**, Only [CDK8/19i+2i] protects *Nanog*-GFP^{High} naïve-cells completely following JAK-STAT inhibition. Representative of n=2 experiments. **f**, Western blots: lentiviral shRNA-mediated knockdown of CDK8, CDK19 or Cyclin C (CCNC) in mouse PSC. Efficient shRNAs (red). Representative of n=2 experiments. **g**, Pluripotency marker mRNA expression (qRT-PCR) in mouse PSCs following 7d lentiviral shRNA-mediated knockdown of CDK8, CDK19, or CCNC. Data are the mean of two experiments. **h**, Upper schematic: inducible-CDK8-knockout. 4-hydroxy-tamoxifen (4OHT)- inducible Cre drives excision of Exon2. Lower: PCR confirmation of CDK8 Exon2 deletion using indicated primers. Mouse *Cdk8*(fl/fl)-RERT-Cre iPS (n=3–4 independent clones) treated 6d with 0.5μM 4OHT. **i**, Indel mutation in one mouse CDK19-KO iPS clone, induced by indicated CRISPR guide-RNA against CDK19 Exon1, using lentiviral CRISPR-Cas9. Indel is 10 bp deletion at predicted CRIPSR target site, generating a frameshift immediately downstream of ATG start codon. **j**, Western blots: 4-OHT-inducible 8*Cdk8*-knockout mouse iPS clones as in (**h**). Schematic summarizes generation of these cells. Data representative of 4 experiments with n=3 iPS clones. **k**, Western blots: 4-OHT-inducible CDK8/19-double-knockout mouse iPS clones, generated as in (**J**). CDK8-knockout confirmed at protein level after 4-OHT-inducible Cre treatment. CDK19 protein was undetectable in PSCs, but readily detectable in intestinal organoid controls. Arrow indicates CDK19, confirmed by CRISPR-knockout as shown; *non-specific band. Data: CDK8/19-knockout with n=10 independent iPS clones. **l**, Left, Western blot: CDK8/19-double knockout (dKO) iPSCs ± empty-vector, catalytically-active CDK8 wild-type (WT), or CDK8 Kinase-Dead (KD). Right, bright field images: mouse iPS lines as indicated. Arrows: naïve-like colony morphology in cells expressing CDK8-WT plus treatment with CDK8/19i. Importantly, CDK8/19-dKO iPS with empty vector do not respond to CDK8/19-inhibitor. Images representative n=3 iPS clones. Scale bars 100 μm. **m**, Re-expression of CDK8-WT in null background rescues response to CDK8/19i. Left, overview. Right, mRNA expression of naïve pluripotency markers (qRT-PCR; Mean, 2 experiments) in CDK8/19-dKO mouse iPS with empty-vector, or catalytically-active CDK8-WT, in serum/LIF or CDK8/19i.



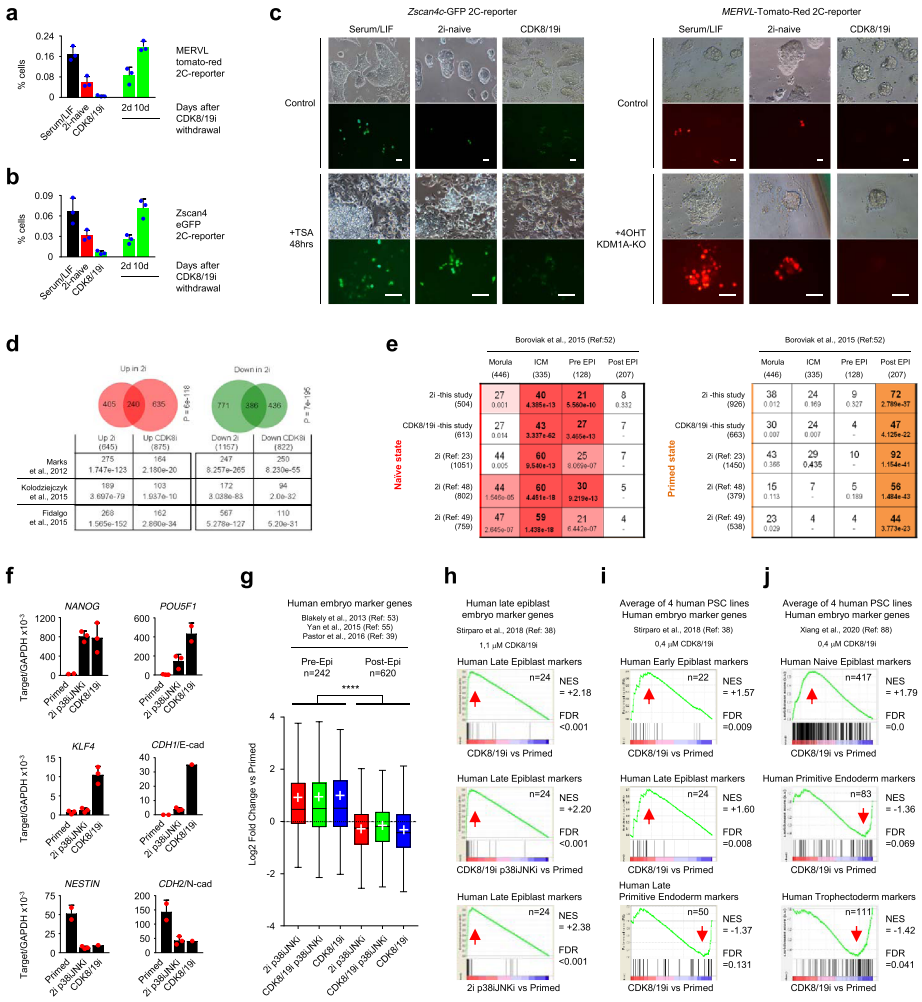
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Positive effect of long- term CDK8/19i on mammalian ES self- renewal and pluripotency. a, Cell morphology and alkaline phosphatase staining of WT mouse EpiSC infected with pMSCV-Empty or pMSCV-CDK8-KD, plus 7d/1 passage in either EpiSC media (Fgf2/ActivinA/ fibronectin; Methods) or serum/LIF ES media (see schematic). Images representative of day 7 after pMSCV infection and selection, n=2 cell experiments. Scale bars 100 μm. **b**, FACS: maintenance of pluripotency markers in mouse PSCs. Percentage of double- positive Nanog-GFP+/ICAM1+ PSCs following 3 weeks culture in control serum/LIF, 2i- naïve, or CDK8/19i. Data: Mean, n=2 cell experiments. **c**, Immunofluorescence: TFE3 expression and localization in mouse PSCs adapted to control serum/LIF, 2i-naïve, or CDK8/19i, as in (**b**). Scale bar 10 μm. **d**, Embryoid body (EB) differentiation with beating cardiac centre demonstrates developmental capacity of mouse PSCs previously adapted to indicated conditions. Data: Mean±SD, n=3 cell experiments, two technical replicates each. **e**, Differentiation in vitro demonstrates developmental capacity of mouse PSCs previously adapted to control, 2i or CDK8/19i conditions. PSC differentiation was by LIF-removal or LIF-removal plus retinoic acid, and assessed by qRT-PCR to show loss of pluripotency (Oct4) and induction of differentiation (Brachyury, T). Data: Mean, 2 experiments. **f**, Teratoma assay in vivo demonstrates developmental capacity of mouse PSCs previously adapted to CDK8/19i conditions. Three embryonic germ layers confirmed in teratomas using histology (H+E stain; upper panels), and staining for germ layer markers: NESTIN (ectoderm), VIMENTIN (mesoderm), and Alpha-feto-protein (AFP, endoderm). Data representative of n=6 teratomas. Scale bar 200 μm. **g**, Brightfield images showing colony morphology in 3 human PSC lines in primed state (upper panels), or 14d treatment with CDK8/19i. Images representative n=5 independent experiments. Scale bar 100μm. **h**, Brightfield and live-cell GFP-fluorescence images of human iPS cells (HERVH-GFP reporter) in primed conditions, or following 14d treatment with indicated media cocktails. To derive and maintain the 2i p38iJNKi condition. Images representative of n=5 independent experiments. Scale bar 100 μm. **i**, FACS analysis of pluripotency markers in human PSCs (HERVH iPS or WIBR3 ES), following 3 weeks adaption to indicated culture conditions, as in (**g**, **h**). Data represent one experiment with n=2 independent PSC lines. Primed or CDK8/19i PSCs were routinely passaged in bulk using collagenase. In contrast, for the 2i p38iJNKi condition, cytometric sorting was required to select the top 10% HERVH-GFP cells at each passage, for 3 passages, before fixing the cells 4d after third passage/selection- round.



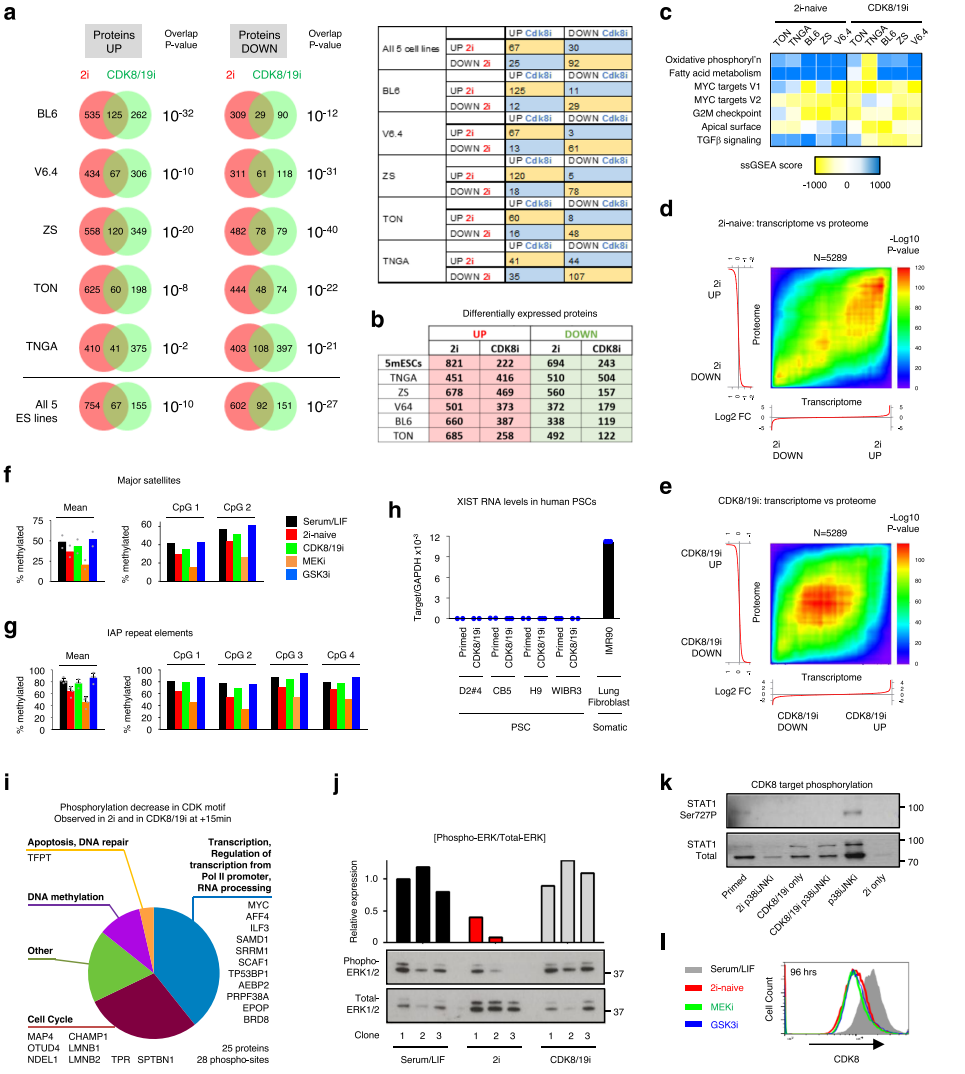
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Self-renewal, genomic stability, and gene expression analysis in mouse and human PSCs in CDK8/19i. a, Single cell clonogenicity: human PSCs in primed or CDK8/19i conditions. Cells were FACS-selected for top or bottom 5% HERVH-GFP intensity, seeded at clonal density in primed or CDK8/19i culture (±p38iJNKi) for 7d (individual colonies arise separately), then Alkaline phosphatase stained (inset, example colonies) to visually score maintenance of pluripotent status. Data: n=10 fields of view, multiple colonies per view, one experiment. **b**, Western blots: naïve pluripotency marker KLF17 in primed or CDK8/19i conditions. SMC1: nuclear internal loading control; n=4 human PSC lines. **c**, Mean pluripotency marker mRNA expression levels (qRT-PCR) in each of n=5 human PSC lines, in primed or CDK8/19i conditions >14d. **d**, Karyotyping indicates genomic stability: Human PSC lines (n=5), 16–19 passages in primed or CDK8/19i conditions. Inset, representative example: karyotype maintenance. **e**, Rank-Rank Hypergeometric Overlap (RRHO):[85] RNA-seq differential expression in mouse PSCs adapted to 2i-naïve or CDK8/19i conditions, versus serum/LIF (n=3 biological replicates; n=12,629 genes). Genes arranged by magnitude-change, then assessed for overlap by RRHO sliding window of 100 genes. Colour intensity: -log10 p- value after Benjamini-Yekutieli correction of hypergeometric overlap. **f**, Overlap and hypergeometric significance of differentially-expressed mRNAs in mouse PSCs in 2i or CDK8/19i, versus control KSR/LIF (serum-free conditions) RNA- seq; n=3 biological replicates; FDR<0.05. **g**, Pluripotency marker and LINE L1 repeat RNA expression (qRT-PCR) in mouse PSCs cultured as in (**e**). Data: n=3 experiments, Mean±SD, t-test, unpaired, two-tailed, *P<0.05. **h**, Effect of 2i or CDK8/19i on LINE L1 super-family expression (RNA-seq) in mouse PSCs cultured as in (**e**). Data: Mean, n=3 biological replicates, for each LINE L1 family, arranged by evolutionary age, which also reflects transcriptional activity and regulatory mechanisms[86]. 2i and CDK8/19i similarly regulate the youngest and most transcriptionally-active families (calculations, notes: Source Data). **i**, Dot plot: RNA-seq expression, in mouse PSCs, cultured as in (**e**). Pluripotency markers (red, n=18); 2C-fluctuation markers (green, n=112);[46,47,82] lists: Source Data). Below: effect of 2i or CDK8/19i (current study), or CDK8-knockdown[87], on these genesets. Significance: GSEA FDR q-values<0.25, indicated. **j**, Western blots, mouse PSCs. Markers of pluripotency, or 2C-fluctuation (ZSCAN4). Representative: n=2 experiments. **k**, **l**, RNA expression: 2C-fluctuation markers (qRT-PCR), mouse PSCs cultured as in (**e**), or after 10d CDK8/19i-withdrawal (**l**). Data: Mean, n=2 experiments. **m**, FACS quantification, percent fluorescence^{high} cells in 2C-fluctuation in two independent mouse PSC 2C-reporter lines, cultured as in (**e**). Induction of 2C- fluctuation: inducible-Kdm1a-knockout[46], or 48h TSA[46]. Data: Mean±SD, n=3 experiments.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Comparison of CDK8/19i and 2i in the current work versus published studies. a, **b**, FACS quantification, percent fluorescence^{high} cells in 2C-fluctuation: n=2 independent mouse PSC 2C-reporter lines[46,82], cultured as indicated. 2i and CDK8/19i repress the 2C-fluctuation, reversible by 2–10d inhibitor-removal. Data: Mean, n=3 experiments. **c**, Bright field or fluorescence^{high} cells in 2C-fluctuation: n=2 independent mouse PSC 2C-reporter lines[46,82], cultured as indicated. Induction of 2C-fluctuation: 7d inducible- Kdm1a-knockout[46], or 48h TSA[46]. **d**, **e**, Overlaps between RNAseq published datasets versus current study in mouse PSCs adapted to 2i or CDK8/19i versus control serum/LIF. Differentially expressed mRNAs in current study (FDR<0.01 and 2-fold change), versus three published studies[23,48,49] (**d**), or, versus developmental stage-specific marker genesets from preimplantation mouse embryos[52] (**e**). (D,E) Overlap hypergeometric significance, and number of genes changing in same direction, are reported below each 2-way comparison. **f**, Pluripotency (NANOG, POU5F1/OCT4, KLF4, CDH1/E-cadherin) or differentiation (CDH2, NESTIN) marker expression (qRT-PCR) in human PSCs adapted to indicated conditions. Mean±SD, n=1–3 biological experiments. **g**, Human embryo stage-specific developmental genesets (scRNA-seq), defining pre- implantation naïve epiblast (n=242 genes) and post-implantation primed epiblast (n=620 genes), can distinguish human PSCs between naïve and primed pluripotent states in vitro by up/down-regulation[39,53,55] (listed in: Source Data). RNA-seq expression of these genesets is shown in the current study in human PSCs adapted to indicated conditions. Tukey box plots; box reflects 25th -75th percentile; horizontal line is median; white-cross indicates mean. Data: n=3 biological replicates, Mean± SD, t-test, unpaired, two-tailed, ****P < 0.0001. **h**, **i**, Human embryo lineage-specific genesets[38] defining early/late pre-implantation naïve epiblast (n=22/24 genes), and late primitive endoderm (n=50 genes), were assessed by GSEA in our human PSC cultured in 2i-naïve or CDK8/19i. **h**, 1.1 μM CDK8/19i, one PSC line. **i**, 0.4 μM CDK8/19i, 4 PSC lines. Significance indicated by GSEA FDR q-values<0.25, and up/down-regulation (red arrows), in each panel. **j**, Human embryo lineage-specific scRNA-seq genesets[88], defining pre-implantation naïve epiblast (n=417 genes), primitive endoderm (n=83 genes), or trophectoderm (n=111 genes), were assessed by GSEA in our human PSC cultured in 0.4 μM CDK8/19i versus primed culture (n=4 PSC lines). Significance indicated by GSEA FDR q- values<0.25, and up/down-regulation (red arrows), in each panel.

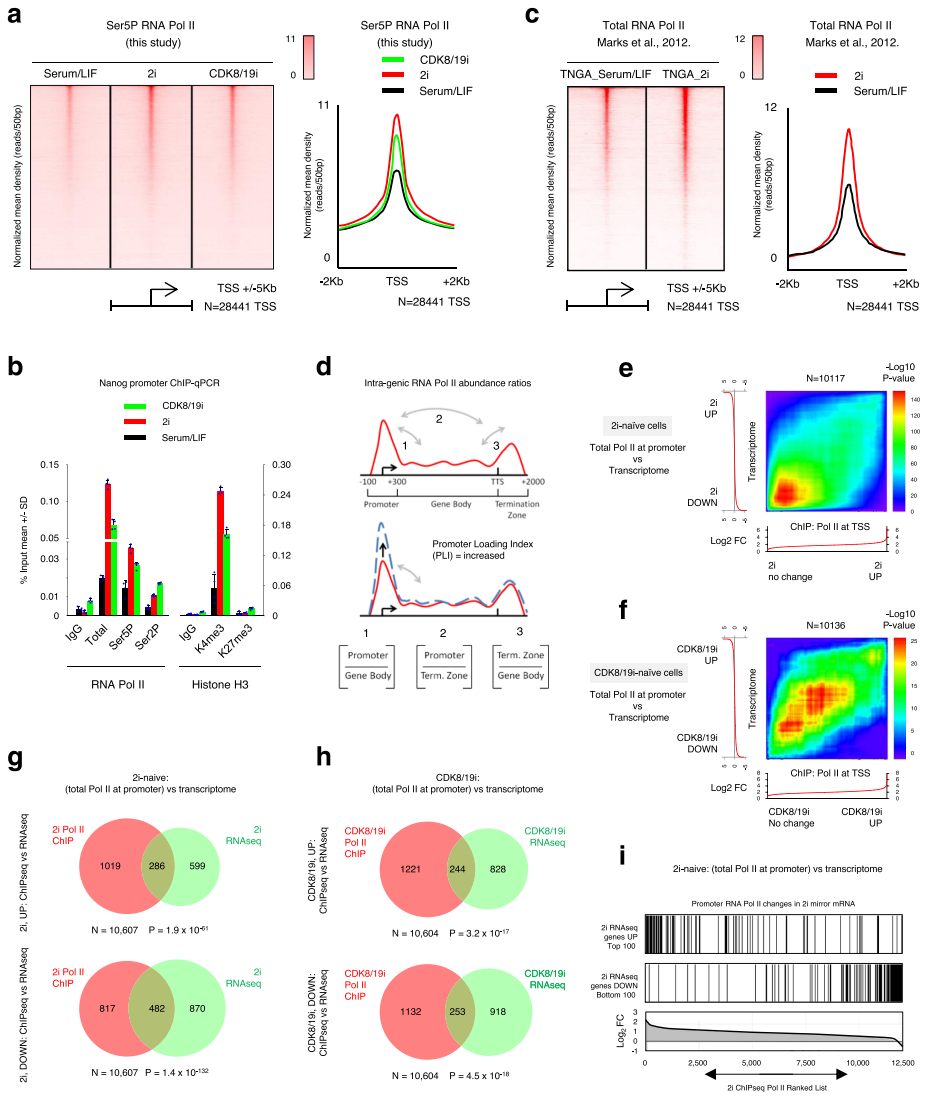


		UP Cdk8i	DOWN Cdk8i
All 5 cell lines	UP 2i	67	30
	DOWN 2i	25	92
BL6	UP 2i	125	11
	DOWN 2i	12	29
V6.4	UP 2i	67	3
	DOWN 2i	13	61
ZS	UP 2i	120	5
	DOWN 2i	18	78
TON	UP 2i	60	8
	DOWN 2i	16	48
TNGA	UP 2i	41	44
	DOWN 2i	35	107

	UP		DOWN	
	2i	CDK8i	2i	CDK8i
5mESCs	821	222	694	243
TNGA	451	416	510	504
ZS	678	469	560	157
V64	501	373	372	179
BL6	660	387	338	119
TON	685	258	492	122

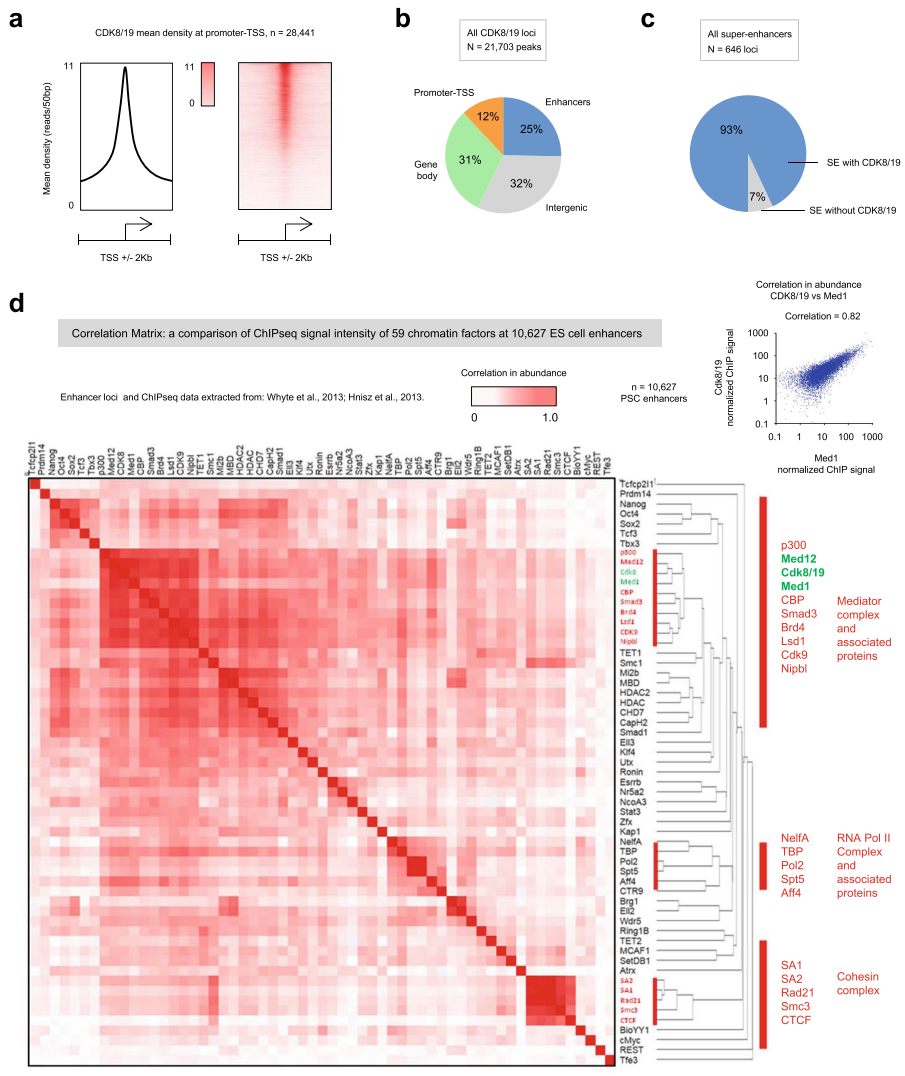
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | CDK8/19i regulates the phospho-proteome and proteome similar to 2i- naïve pluripotency, but not DNA methylation. a, Left: Overlap and hypergeometric significance (P-value) of differentially expressed proteins (FDR<0.05), in mouse PSC lines adapted to 2i-naïve or CDK8/19i, versus standard serum/LIF (n=5 lines, displayed individually). Right: table compares overlap in proteins up/down-regulated, per cell line and per condition, to highlight that the positive correlation (proteins changing in same direction) is greater than the negative correlation in all PSC lines. Supplementary Table 4: list of differentially expressed proteins. **b**, Summary, all proteomic changes, shown per mouse PSC line, adapted to 2i-naïve or CDK8/19i conditions, versus control serum/LIF, as in (**a**). **c**, Heatmap: normalized enrichment of biological pathways identified as significantly up/down-regulated (blue/yellow), by GSEA of proteomic changes, shown per cell line, in mouse PSCs cultured as in (**a**). Significance was confirmed in all pathways shown (GSEA FDR q-values<0.25). **d**, **e**, Rank-Rank Hypergeometric Overlap (RRHO):[85] differential mRNA expression (X- axis) in mouse PSCs adapted to 2i-naïve conditions (**d**) or CDK8/19i (**e**), versus, differential protein expression (Y-axis), for the same genes (n = 5289). Genes arranged by magnitude-change, then assessed for overlap by RRHO sliding window of 100 genes. Colour intensity: -log10 p-value after Benjamini-Yekutieli correction of hypergeometric overlap. **f**, **g**, DNA methylation changes (5-methyl-cytosine; pyrosequencing) in n=4 mouse PSC lines adapted to 2i or CDK8/19i. **f**, Methylation levels at two CpG sites in Major Satellite repeats, shown independently (right), or Mean±SD across the CpG loci (left). **g**, Methylation levels at four CpG sites in IAP repeats, shown independently (right), or Mean±SD across the CpG loci (left). **h**, XIST RNA levels in human PSC lines in this study (qRT-PCR). Female (n=3: WIBR3, CB5 and H9) and Male (n=1: D2#2) PSCs display low/undetectable XIST expression compared to control adult female human somatic cells (lung fibroblasts), suggesting X-silencing erosion may have already occurred in parental cells, as previously observed[67]. Data: n=3 technical replicates. **i**, Functional analysis of proteins containing a CDK phospho-target motif that displays phosphorylation decrease (FDR<0.05) within 15 min treatment of mouse PSCs with 2i or CDK8/19i. Data: n=2 PSC lines. **j**, Western blots: ERK1/2 phosphorylation after long-term adaption (3 weeks) of mouse PSCs to serum/LIF, 2i, or CDK8/19i. Above: relative ERK1/2 phospho-levels, normalized by total ERK1/2 levels. **k**, Western blots: CDK8 kinase-target STAT1-phospho-serine727, in human PSCs, with indicated culture media. **l**, CDK8 protein levels per cell measured by cytometry in mouse PSCs treated with indicated inhibitors. **j–l** Representative, n=2 independent experiments.



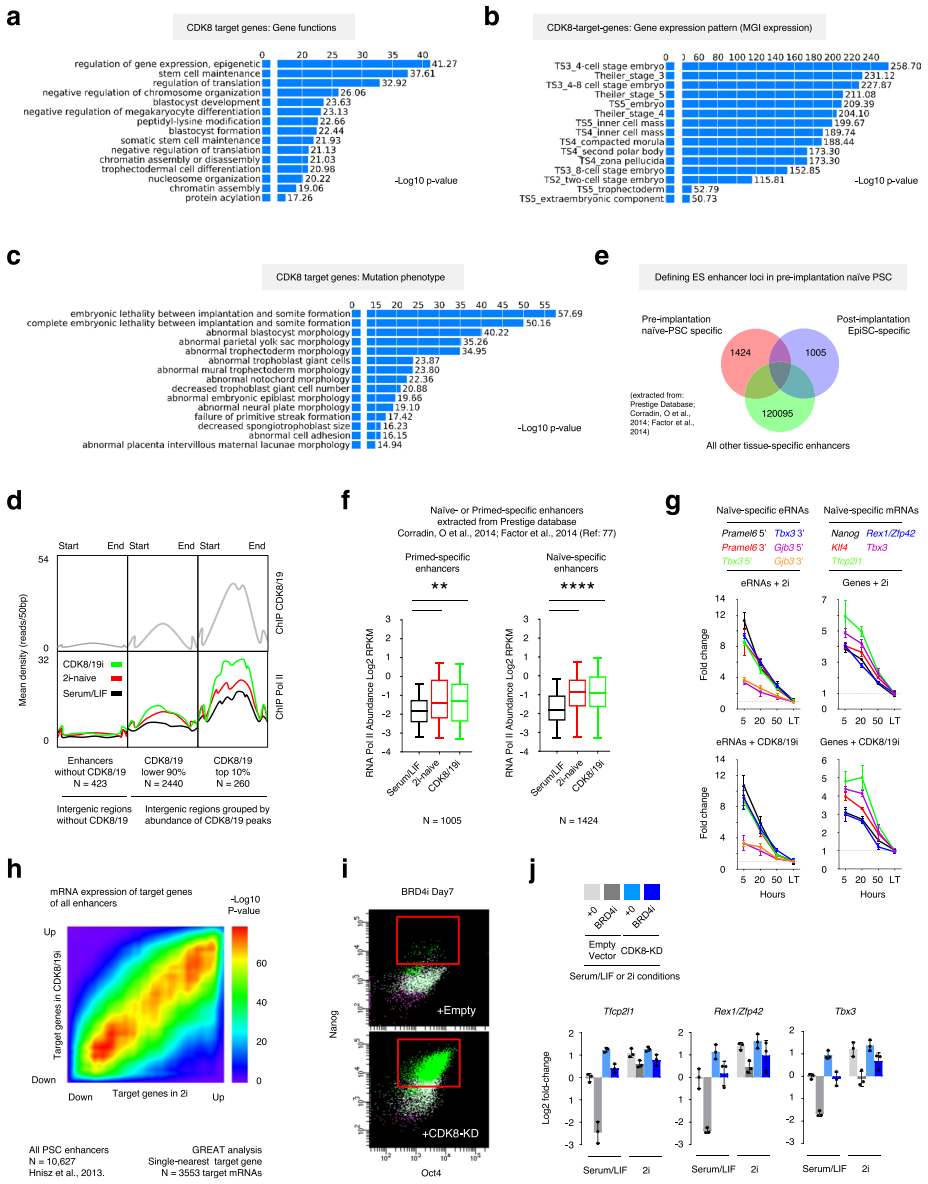
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Analysis of RNA Pol II genomic distribution and corellation with RNA gene expression. a, ChIP-seq: RNA Pol II Serine 5 phosphorylation (Ser5P) abundance at all Refseq Transcription Start Sites (TSS; n=28,441), in mouse PSCs, treated as indicated. Left: Heatmap, TSS +/−5 Kb. Right: Metagene average +/−2 Kb. ChIP-seq n=3 pooled replicates. **b**, ChIP-qPCR: RNA Pol II and histone marks at the Nanog TSS. RNA Pol II and H3K4me^3 (active euchromatin) are increased. Data: Mean±SD, n=4 ChIP replicates. **c**, Re-analysis of published[23,69] ChIP-seq: RNA Pol II abundance at all Refseq Transcription Start Sites (TSS; n=28,441), in mouse PSCs, treated as indicated, (similar conditions to current study, compare with Fig. 5h, i). Left: Heatmap, TSS +/−5Kb. Right: Metagene average +/−2Kb. ChIP-seq n=3 pooled replicates. **d**, Schematic: defining gene regions and Pol II loading ratios used in this study, similar to previous reports[68]. Lower panel: schematic summarizing results in Fig. 5h, i, where Promoter Loading Index is increased (Promoter/Body). **e**, **f**, Rank-Rank Hypergeometric Overlap (RRHO):[65] differential mRNA expression (RNA-seq data; Y-axis) in mouse PSCs adapted to 2i-naïve conditions (**e**; n=10,117) or CDK8/19i (**f**; n=10,136), versus, differential RNA Pol II abundance at promoter-TSS (ChIP-seq data; X-axis), for the same genes. Genes arranged by magnitude-change, then assessed for overlap by RRHO sliding window of 100 genes. Colour intensity: - log10 p-value after Benjamini-Yekutieli correction of hypergeometric overlap. **g**, **h**, Venn diagrams: genes with differential mRNA expression up/down-regulated (green circles; FDR<0.01) in 2i (**g**) or CDK8/19i (**h**) overlap significantly with genes where the promoter has the greatest/least change in RNA Pol II abundance (red circles; promoters with fold change > one standard deviation from mean). Overlap significance: hypergeometric test; P-values, and number of genes "n", indicated in each panel. Genes up (top Venn diagrams), and genes down (lower Venn diagrams), refer to inhibitor-treated cells versus control serum/LIF conditions. **i**, Genes with the greatest change in RNA Pol II abundance (lower panel; ranked list of promoters by magnitude of RNA Pol II abundance fold-change; n=12,693; ChIP-seq) correlate with the top 100 most differentially expressed mRNAs up/down-regulated in 2i-naïve conditions (upper two panels; RNA-seq). All changes refer to 2i-treated cells versus control serum/LIF conditions.



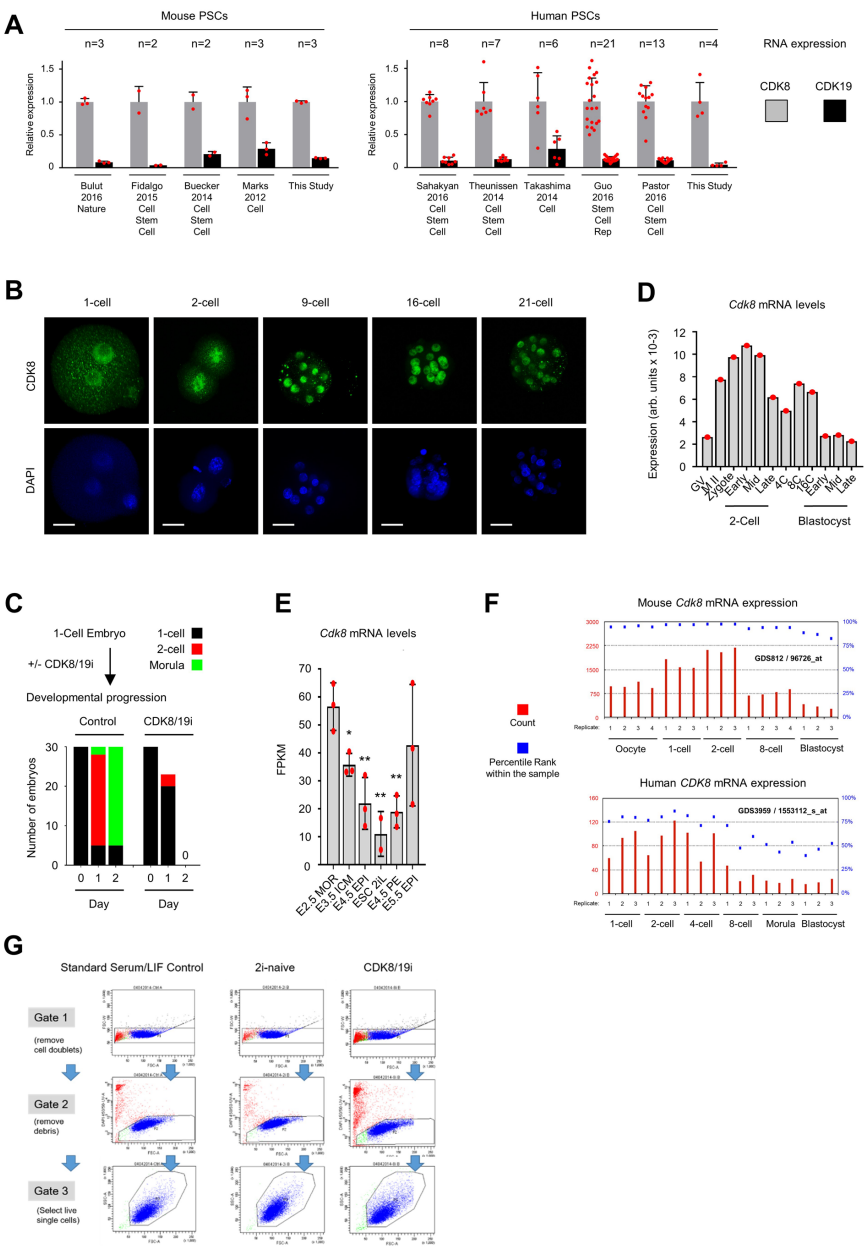
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | ChIP-seq for CDK8 and analysis of its genomic distribution. a, CDK8/19 average ChIP-seq enrichment[2,3] density in mouse PSCs at Promoter-TSS regions +/− 2 Kb, n=28,441 TSS (Refseq). **b**, CDK8/19 binding loci defined in mouse PSCs by ChIP-seq[2,3], MACS peak calling, and categorized by functional annotation of the region by HOMER (n=21,703, see Supplementary Table 7). Note: ChIP antibody binds both CDK8 and CDK19, see Methods. Promoter-TSS: TSS+/− 1Kb. Gene Body: Exons, Introns, and transcription termination site TTS +/− 1Kb. Enhancer constituent regions, as defined[2,3]. **c**, Percentage of SE-constituent regions[2,3] enriched for CDK8/19 binding (see also Supplementary Table 7). **d**, CDK8 is an integral part of Enhancer-Mediator in mouse PSCs. Pearson correlation Matrix summarizes correlation/co-occupancy between 59 factors in 10627 Enhancers in mouse PSCs, based on comparison of ChIP-seq signal intensity in published datasets. Enhancer loci and ChIPseq data extracted from[2,3]. The 59 factors indicated are a range of chromatin modifiers and transcription factors. Each square of the matrix represents a comparison between the corresponding pair of factors for their similarity in ChIP signal ranking across the 10,627 enhancer regions, to calculate a r^2 correlation of their similarity, where 1.0 = exactly similar. An example of a single correlation between two factors is shown for the Mediator subunit Med1 and CDK8/19 abundance within mouse PSC enhancers, in the upper-right of the panel. Hierarchical clustering groups those factors by similarity in ChIP signal pattern across all 10,627 enhancers. Thus, high correlation between two factors (red), indicates co- enrichment to similar levels and at the same set of enhancers, which is suggestive of functional co-operation. Co-enrichment patterns for subunits and co-factors of the Mediator, RNA Pol II and Cohesin complexes can be observed (indicated), consistent with their reported combinatorial roles at enhancers. CDK8/19 clusters most closely with the Mediator complex and other critical regulators of enhancer function. See Methods, Supplementary Table 7, and Source Data for analysis of the published ChIP datasets and enhancer loci defined by[2,3].



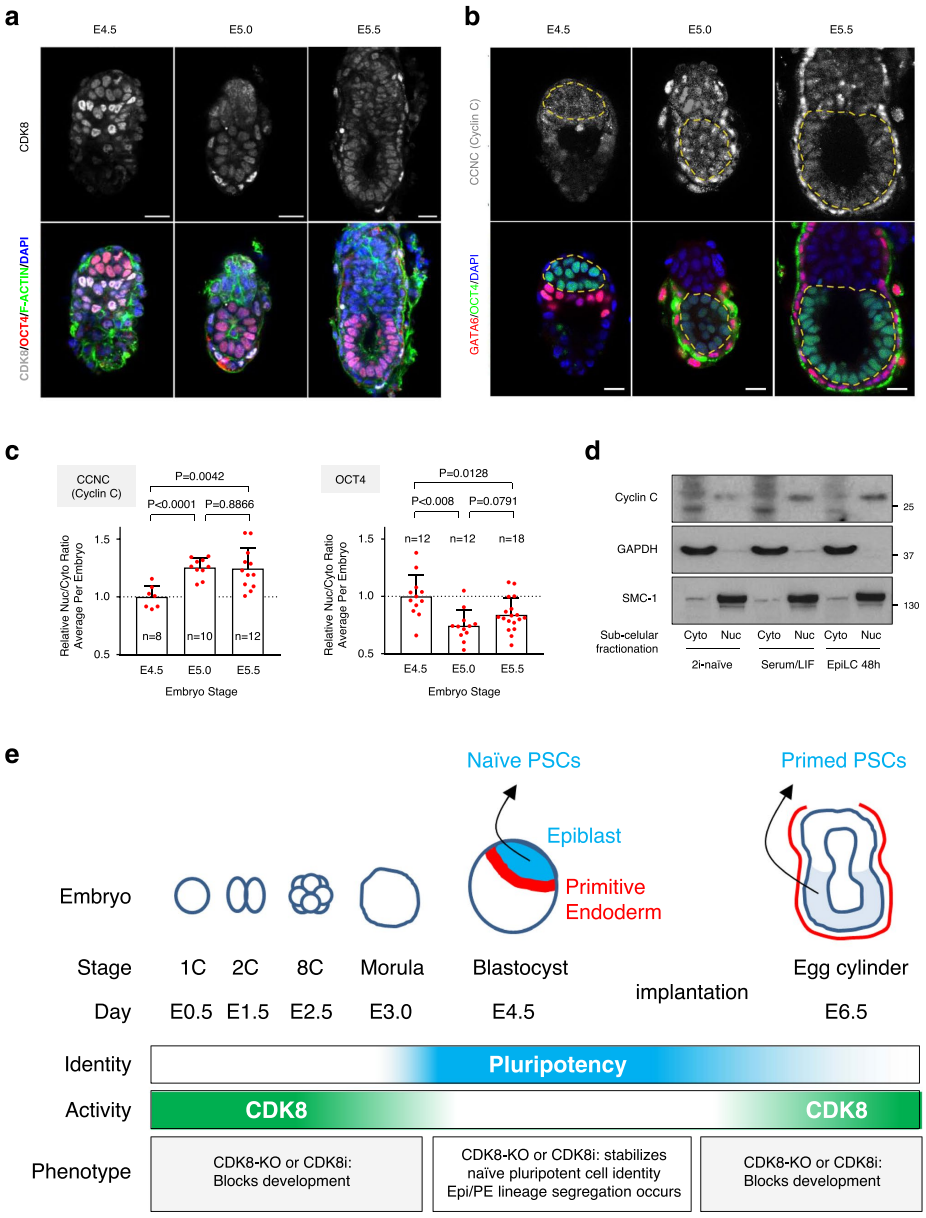
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | RNA Pol II and CDK8 genomic distribution. 2i and CDK8/19i hyper- actívate naïve-state enhancer activity. a–c, Gene ontology enrichment and functional annotation of CDK8/19-target genes identified by the single-nearest gene to each CDK8/19 binding site in mouse PSCs (ChIP-seq, n=21,703 peaks: Supplementary Table 7) using GREAT analysis[80]. Data: -log 10 binomial P-value with Bonferroni correction for multiple hypothesis testing indicating the significance of each gene ontology category. **d**, Metagene enrichment in the indicated genomic regions for CDK8/19 or RNA Pol II abundance, as determined by ChIP-seq in mouse PSCs above. Genomic regions were defined in groups by the CDK8/19 peak intensity as defined in (**a**). **e**, Identification of mouse PSC super-enhancer (SE) loci specific to pre-implantation naïve epiblast, or post-implantation primed epiblast. Enhancer loci were extracted from the Prestige Database[77]. The SEs in naïve or primed epiblast were first identified, and then any SEs common to a panel of 16 somatic tissues (threshold: 1 bp overlap) were subtracted (Methods). Enhancer loci lists: Source Data. **f**, RNA Pol II abundance in mouse PSC primed-specific super-enhancers (on left, n=1005), or naïve-specific super-enhancers (on right, n=1424), as defined in (**e**). RNA Pol II levels are significantly higher in 2i or CDK8/19i conditions versus serum/LIF control: t-test, unpaired, two-tailed, Welchs correction, **P<0.01; ****P<0.0001). Tukey box plot centre lines show median values, box limits represent upper and lower quartiles, and whiskers show 1.5× interquartile range. **g**, qRT-PCR: pluripotency marker genes and naïve-specific eRNA[71] abundance in mouse PSC at short time intervals after withdrawal of 2i or CDK8/19i from culture. Data: Mean±SEM, n=3 experiments. **h**, Rank-Rank Hypergeometric Overlap (RRHO)[85]: heatmap shows differential mRNA expression (RNA-seq) of the single-nearest target genes (n=3,553; GREAT analysis) identified for all PSC enhancers[2,3] (n=10,627) in 2i-naïve conditions (X-axis) or CDK8/19i (Y-axis), compared to control serum/LIF conditions. The enhancer-target mRNA expression changes are arranged by magnitude, then assessed for overlap by RRHO sliding window of 100 genes. Colour intensity: -$\log_{10}$ *P* value after Benjamini- Yekutieli correction of hypergeometric overlap. Highly significant overlap along the diagonal indicates similar regulation of enhancer-target gene mRNA expression in 2i and CDK8/19i. **i**, FACS: NANOG and OCT4 protein expression following 7d treatment with 500 nM BRD4i(JQ1) in CDK8/19-dKO iPS clones expressing pMSCV-Empty or pMSCV-CDK8- Kinase-Dead (CDK8-KD). Representative of n=3 cell experiments. **j**, qRT-PCR: expression of naïve marker genes following 48h treatment with 500 nM BRD4i(JQ1). CDK8/19-dKO iPS ± CDK8-KD were cultured in 2i or standard serum/LIF, as indicated. Mean±SD, n=3 independent clones.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | CDK8 expression in vivo and the role of Mediator during mouse preimplantation development. a, CDK8 and CDK19 mRNA relative expression levels in PSCs, as detected by RNAseq in 5 mouse datasets[23,49–51] and 6 human datasets[32,34,39,58,66] (including the current study). Mean±SD of independent RNAseq replicates in each published study (see: n replicates indicated in panel; data calculation, see: Source Data). **b**, Immunofluorescence for CDK8 protein levels during mouse preimplantation development from 1-Cell to early blastocyst stage (E3.5), representing one set of embryos. Scale bars 25μm. **c**, CDK8/19-inhibition blocks embryo development at 1–2 Cell stage. Day E0.5 zygotes were harvested from females and immediately cultured in vitro in KSOM ± CDK8/19i for 2 days, with assessment of their developmental progression by visual inspection of cell number and morphology at intervals. Data represents n=30 embryos per condition, across two independent experiments. **d**, **e**, CDK8 mRNA expression levels in specific embryo stages and lineages during mouse preimplantation development. CDK8 mRNA expression declines until blastocyst stage, both in mouse and human pre-implantation embryos. In (**d**), mean data values from published microarray studies (Methods). In (**e**), CDK8 mRNA expression levels detected by RNA-seq in specific embryo stages and lineages during mouse preimplantation development; data from[52], Mean±SD, n=2–3 replicates per time point, significance assessed by one-way ANOVA unpaired T-test. *P<0.05; **P<0.01. (see Source Data). **f**, CDK8 mRNA expression levels during mouse or human embryo pre-implantation development, as detected by microarray in published datasets. Mouse: https://www.ncbi.nlm.nih.gov/geo/tools/profileGraph.cgi?ID=GDS812:96726_at Human: https://www.ncbi.nlm.nih.gov/geo/tools/profileGraph.cgi?ID=GDS3959:1553112_s_at (**g**) Schematic showing example of FACS gating strategy in this study. DAPI was added to live cell suspension 2mins before analysis, as a live/dead discriminator. Gates 1, 2, and 3, sequentially act to exclude cell doublets, and debris, thus selecting live single cells for analysis of Nanog-GFP profile in mouse PSCs.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Cyclin C expression localization during mouse preimplantation development. a, Immunofluorescence: CDK8, OCT4, and F-ACTIN in mouse early embryos from E4.5 to E5.5. Scale bars 20 µm. Images representative of n=3 experiments. **b**, **c**, Immunofluorescence (**b**): cyclin C protein levels during mouse development from preimplantation blastocyst stage (E4.5) to post-implantation cylinder stage (E5.5). Co-staining with OCT4 to mark epiblast, and GATA6, to mark primitive endoderm at E4.5 and its maturation into post-implantation visceral endoderm. **c**,, Cyclin C nuclear-cytoplasmic ratio was quantified and plotted, where each data point represents the mean Nuc-Cyto ratio for epiblast cells of one embryo (Source Data). As internal control, the Nuc-Cyto ratio for OCT4 was also quantified. Nuclear abundance of cyclin C increases in epiblast cells during development from E4.5 to E5.5. In contrast, OCT4 nuclear-cytoplasmic ratio does not follow this pattern. This implies that cyclin C pattern is not related to staining or imaging artefacts. **b**,, representative examples shown; n=2 experiments. **c**,, data: Mean±SD, T-test, unpaired, two-tailed, P-values and number of embryos "n" is indicated. **d**, Western blot: cyclin C localization by sub-cellular localization. Nuclear and cytoplasmic fractions were prepared from mouse cells across developmental spectrum: naïve (adapted to 2i), primed (adapted to serum/LIF), or epiblast-like stem cells (abbreviated EpiLC; derived by treating PSCs for 48 h with EpiSC media[59]) (Methods). Relative abundance of nuclear cyclin C is greater in primed state EpiLC and in serum/LIF conditions, compared to 2i-naïve. Data represent n=2 experiments. **e**, Summary: The developmental requirement for CDK8 activity mirrors its embryonic expression pattern. Maxima in CDK8 expression coincide with a requirement for development around the zygote-morula and post-implantation stages. Between these two periods, a transient minima in CDK8 expression occurs during emergence of naïve epiblast, where CDK8 function appears dispensable. We suggest that CDK8/19 chemical inhibition in vitro mimics CDK8 downregulation during pre-implantation development in vivo, providing mechanistic insight on how naïve pluripotency may arise in the embryo: (i) CDK8/19 is required during zygote-to-morula development, where its expression is high. (ii) During morula-to-blastocyst pre-implantation development, CDK8 expression declines, and nuclear cyclin C decreases. This coincides with the emergence of E4.5 pre-implantation naïve epiblast and, accordingly, CDK8/19 inhibition does not interfere with naïve epiblast specification. In contrast to MEK inhibition, CDK8/19 inhibition does not affect the epiblast/PE lineage segregation. (iii) During the subsequent developmental transition of pre-implantation naïve epiblast to post-implantation primed state, CDK8 expression becomes increased and CDK8/19 activity is required for morphogenic events.

nature research

Corresponding author(s): Manuel Serrano
manuel.serrano@irbbarcelona.org

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	- qRT-PCR data were obtained using an ABI PRISM 7700 thermocycler (Applied Biosystem) and data were analyzed using Graph Pad Prism (version 9). - Cytometric data were collected using either of these two cytometers: FACSscalibur (BD Biosciences, Franklin Lakes, NJ) and FACS CANTO II (BD, San Jose, CA); and analyzed using FlowJo 9.6.2 software. - Immunofluorescence images were obtained using a Leica SP5 microscope and analyzed with Fiji (version 1.0). -RNAseq data were obtained in an Illumina HiSeq2500 sequencer; alignments were performed using with STAR software with default parameters. - Mass spectrometry data were collected using a Q Exactive HF mass spectrometer (ThermoFisher Scientific) and analysis was done using MaxQuant (v 1.6.0.16) software.
Data analysis	- Cytometric data were analyzed with FlowJo 9.6.2 software. - Immunofluorescence images: analyzed with Fiji (version 1.0). - Statistical analyses (excluding the sequencing data): analyzed with Graph Pad Prism (version 9). -RNAseq alignments also with STAR software with default parameters. RNAseq: number of reads per gene were calculated using the featureCounts function of the Rsubread package of the R statistical software. Gene annotations were performed using biomaRt with version "may2015.archive.ensembl.org". Perseus v1.5.5.2, Limma, Prostar package, Differential analysis was done with the R package DESeq2. RRHO available at: http://systems.crump.ucla.edu/rankrank/rankranksimple.php. Correlation matrix of ChIP-seq data was produced using Morpheus software, available from the Broad Institute: https://software.broadinstitute.org/morpheus/. -Analysis of deep-sequencing data: analyzed with R and with standard programs and packages as detailed step-by-step in the Methods

section. Briefly, ChIP-seq analysis was performed with the RUbioSeq pipeline (v3.8), using the following tools: FastQC v0.11.5, BWA v0.7.10, SAMtools v0.1.19, Picard tools vl.107, Bedtools v2.16.2 and MACS2 v2.1.1.20160309, SeqMiner 1.3.3e.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

DATA AVAILABILITY
RNA-seq and ChIP-seq data are available from the GEO database under accession numbers GSE112208 and GSE127186. The mass spectrometry proteomics data are available from the ProteomeXchange Consortium/PRIDE repository with the dataset identifier PXD009200. Published datasets included in this study: Mouse: GSE56138, GSE23943, GSE81285, GSE81045, GSE81045, E-MTAB-2600; Human/Primate: GSE76970, GSE76970, GSE76970, GSE87239, GSE87239, GSE60945, GSE59435, E-MTAB-4461, E-MTAB-2031, GSE44183, GSE36552, E-MTAB-3929; see publication References in Source Data Figure 4E and 4K. All other data supporting the findings of this study are available from the corresponding author on reasonable request.

List of Figures with associated Raw data:
Fig1
Fig2
Fig3
Fig4
Fig5
Fig6
Fig7
Fig8

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was determined based on our previous experience and the work of other groups using embryos and stem cells as experimental model systems.
Data exclusions	FACS analysis: live/dead staining was used to exclude dead cells and debris from the analysis for the quantifiaction of Nanog-GFP, HERVH-GFP, Zscan4-GFP, MERVL-tomator-Red reproter PSC cell lines. FACS gating strategy is in Extended Data Fig.9G.
Replication	The experimental findings were reliably reproduced. Key inhibitor treatment experiments with CDK8/19-inhibitors with mouse and human PSCs were reproduced in independent laboratories. All experimental data was replicated at least in two independent experiments.
Randomization	Embryos and cell cultures were randomly allocated to control and experimental groups before experimental treatments (Methods section).
Blinding	The investigators were not blinded to group allocation. This report relies on studying the differences between specific cell types and culture conditions. Therefore blinding was not applied in these cases

Reporting for specific materials, systems and methods

Materials & experimental systems	
n/a	Involved in the study
☐	☒ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants

Methods	
n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Unique biological materials

Policy information about availability of materials

Obtaining unique materials	Two of the five CDK8/19-inhibitors used in this study (see Table S1) were developed in-house at the CNIO Experimental Therapeutics program. Samples of these compounds (ETP-47799; ETP50586) are available upon reasonable request to the Corresponding Author.

Antibodies

Antibodies used	The following antibodies were used for Western blotting, Cytometry, cell culture immunofluorescence and ChIPseq: (Target; Provider; Clone; Supplier-product-code; dilutions used. Note: where possible, antibody clone and dilutions are listed): Nanog (Western, IF) Chemicon/Millipore #AB5731 (1/5000) Nanog (FACS, IF) eBiosciences #51-5761 (1/1000) Pou5fl/Oct4 BD Biosciences/Pharmingen #611203 (1/5000) Total RNA Pol II (RPBl) Santa Cruz #sc-899x (N-20) (1/5000) RNA Pol II Ser-SP Abeam #ab5131 (1/5000) RNA Pol II Ser-2P Abeam #ab5095 (1/5000) SMCl Bethyl Laboratories #A300-055A (1/5000) Gapdh Sigma #G8795 (1/10000) Beta-Actin Sigma #A5441 (1/50000) gamma-Tubulin Sigma #T6557, CLONE GTU-88 ascites fluid (1/50000) Lamin A/C Santa Cruz #sc-6215 (N-18) (1/1000) CDK8/19 Cell Signaling #4106 (P455) Atlas Antibodies #HPA007053 CDK8/19 (both) Santa Cruz #SC-1521 specificity for both. (1/1000) Cyclin C Santa Cruz #sc-1061 (1/500) STATl total Cell Signaling #9172 (1/1000) STATl Ser727-phospho Cell Signaling #9177 (1/1000) ERKl/2 total (p44/42) Cell Signaling #9102 (1/2000) ERKl/2 phospho (Thr202/Tyr204) Cell Signaling #9101 (1/2000) SSEA4 Stem Cell Technologies #60062 Clone MC-813-70 (1/500) Tral-81 Millipore #MAB4381 cMyc Santa Cruz #sc-40 {9E10) (1/200) Sox2 Chemicon/Millipore #AB5603 (1/5000) Zscan4c Millipore #AB4340 GATA6 R+D Systems #AF1700 (1/1000) Tfe3 Atlas Antibodies #HPA023881 (1/500) ICAM1/CD54 eBiosciences #13-0541 (1/500) H3K9me3 Upstate/Millipore #07-442 (1/5000) Alpha-fetoprotein (AFP) Abeam #ab46799 (1/500) Vimentin Santa Cruz #sc-6260 (1/5000) Nestin ThermoFisher Scientific #MAl-110 (10C2) (1/500) Antibodies used for embryo/embryoid immunofluorescence in this study: Target, Code, Company, Dilution Rabbit pAb anti-Nanog #ab80892 Abeam 1:200 Goat pAb anti-OTX2 #AF1979 R&D Systems 1:200 Goat pAb anti-GATA6 #AF1700 R&D Systems 1:200 Mouse mAb anti-Oct-3/4 #sc-5279 Santa Cruz Biotechnology 1:200 Rat mAb anti-Podocalyxin Clone 192703 #MAB1556 R&D Systems 1:500 Alexa Fluor 488 Phalloidin (F-actin) #A12379 ThermoFisher Scientific 1:500 Alexa Fluor 568 Donkey anti-Rabbit #A10042 ThermoFisher Scientific 1:500 Alexa Fluor 488 Donkey anti-Mouse #A21202 ThermoFisher Scientific 1:500 Alexa Fluor 647 Goat anti-Rat #A21247 ThermoFisher Scientific 1:500 Alexa Fluor 647 Donkey anti-Goat #A21447 Thermo Fisher Scientific 1:500
Validation	The expression patterns and/or subcellular localization of all the proteins analyzed in this study has been previously reported. This was used to validate the specificity of the antibody. Also, CDK8, CDK19, and CyclinC knockout and/or shRNA-knockdown cell lines were used to validate the respective CDK8, CDK19 and CyclinC antibodies, in this study. See Extended Data Fig.1.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Mouse Cells and culture conditions Mouse ES cells: E14Tg2a.4 (wild-type parental, 129/Ola background) were from BayGenomics/MMRRC resource, University of California; Wild-type ES cells were derived at the Transgenic Mouse Unit of CNIO from E3.5 C57BL6 blastocysts, or mixed background C57BL6/129 blastocysts; Rosa26-GFP and Tg.CAG-Katushka-red ES cell lines were derived at the Transgenic Mouse Unit of CNIO from 129- Gt(ROSA)26Sortm1(CAG-EGFP)Luo/J mice (Jackson 006053) and from Tg.CAG-Katushka mice (Diéguez-Hurtado et al., 2011), respectively. Nanog-GFP knock-in mouse ES cells (TNGA, TON) were previously described (Chambers et al., 2007) and were shared by the laboratory of Austin Smith; The MERVL-td:Tomato mouse ES line was a 2C-reporter were shared by the laboratory of Todd Macfarlan (Macfarlan et al., 2012); The ZS mouse ES line was a 2C-reporter shared by the laboratory of Minoru Ko (Zalzman et al., 2010). Primary mouse embryo fibroblasts (wild-type, MEFs, passage 2) were obtained at E13.5 from pure inbred C57BL6 background mice as described previously (Palmero and Serrano, 2001), or from CDK8 flox/flox RERT-Cre mice. Human 293T cells were from ATCC. Human PSC (HERVH-GFP reporter human iPS previously described: (Wang et al., 2014); and WIBR3 human ES cells: (Gafni et al., 2013). For References, please see Main Text of this manuscript.
Authentication	The self-renewal properties and pluripotency markers of mouse and human PSCs were confirmed by RT-PCR, FACS, immunofluorescence, RNAseq, proteomics, and chimera developmental assays (see Figs 1-4 and S1-S4). CDK8/19-double KO mouse PSCs were authenticated by PCR, sequencing, and/or Western blotting (primers provided in the Methods section).
Mycoplasma contamination	Cell lines were routinely tested for mycoplasma contamination by PCR and confirmed negative
Commonly misidentified lines (See ICLAC register)	No commonly mis-identified cell lines were used in this study

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mice (Mus musculus) were used to obtain mouse embryos for this study. The following strains and genetically-modified models were used: F1 (C57Bl6xCBA), CD1, CAG-GFP (both males and females). Nude mice were used for teratoma developmental assays of PSC pluripotency. Superovulated females for chimera experiments were used at 6 +- 1 week of age. Females for natural matings were used at 3 +- 1 month of age.
Wild animals	No wild animals were used.
Field-collected samples	There are no field-collected samples.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	GEO database: Series: GSE112208.
Files in database submission	Raw data files (fastq) and Processed data files (BED format) are included in the GEOaccession series: GSE112208. The list of pooled and processed data files is: GSM3061003 Input GSM3061004 Ser2P-2i GSM3061005 Ser2P-8i GSM3061006 Ser2P-Control GSM3061007 Ser5P-2i GSM3061008 Ser5P-8i GSM3061009 Ser5P-Control GSM3061010 Total-Pol-II-2i GSM3061011 Total-Pol-II-8i GSM3061012 Total-Pol-II-Control
Genome browser session (e.g. UCSC)	GEO database: Series: GSE112208.

Methodology

Replicates	Mouse ES cells were cultured in three conditions: "Serum/LIF" (control), "2i", or "CDK8/19i" (see text for details of these three conditions.

In each culture condition, ChIP-seq was performed separately with three antibodies: (i) anti-Total RNA Pol II (the central subunit, RPB1); (ii) anti-RNA Pol II phosphorylated at Serine5 of the C-Terminal Domain; (iii) anti-RNA Pol II phosphorylated at Serine2 of the C-Terminal Domain.

Three ChIP replicates were performed for each target and for each culture condition (27 replicates: 9 ChIP-seq targets with 3 technical replicates each). A pooled Input reference was included as a background control. The three ChIP replicates of 20-25 million reads each were pooled to make >60 million reads for each target and for each culture condition. From this, after quality control and read normalization between samples, >42 million reads per condition remained.

Sequencing depth

Reads were single-end and 50bp in length. Q20 sequencing quality was 95-99% for all samples. Three ChIP replicates of 20-25 million reads each were pooled to make >60 million total reads for each target and for each culture condition. From this, after quality control and read normalization between samples, >42 million reads per condition remained.

Table below summarizes the number of reads per 1 replicate representative of each condition:

Sample // Number sequenced reads // N. aligned reads // N. aligned reads without duplicates // N. reads after random sampling to balance read number
Input-2i 25737344 24575560 21876817 15281123
Input-8i 22303891 20402885 18128523 15284967
Input-Control 22755044 20889372 18198596 15286208
Ser2P-2i 23819011 22830756 16282619 15286714
Ser2P-8i 23950726 23154997 17037805 15287258
Ser2P-Control 23726068 22637170 16654134 15286369
Ser5P-2i 23766079 22130771 15285221 15285221
Ser5P-8i 23879973 22714294 16240914 15284796
Ser5P-Control 23943661 22708214 16396186 15285861
Total-Pol-II-2i 23662069 22923122 17729866 15284573
Total-Pol-II-8i 24008598 23260015 17853240 15287660
Total-Pol-II-Control 23901412 23196118 17298552 15285794

Antibodies

ChIP-seq was performed separately with three antibodies: (i) anti-Total RNA Pol II (the central subunit, RPB1) Santa Cruz, #sc-899x (N-20); (ii) anti-RNA Pol II phosphorylated at Serine5 of the C-Terminal Domain, Abcam #ab5131; (iii) anti-RNA Pol II phosphorylated at Serine2 of the C-Terminal Domain, Abcam #ab5095. All the antibodies were used at 1/1000 dilution

Peak calling parameters

Analysis of deep-sequecing data: analyzed with R and with standard programs and packages as detailed step-by-step in the Methods section, and briefly here:
BIOINFORMATIC METHODS: ChIP-seq analysis was performed with the RUbioSeq pipeline (v3.8), using the following tools: FastQC v0.11.5, BWA v0.7.10, SAMtools v0.1.19, Picard tools v1.107, Bedtools v2.16.2 and MACS2 v2.1.1.20160309. Sequencing quality for ChIP-seq samples was analyzed with FastQC (Andrews, 2011). Reads were aligned with Bwa 0.7.5a (Li and Durbin, 2009) to the mouse reference genome (NCBIm37/mm9) using the default seed length (32) and allowing 1 mismatch in the seed. SAMtools 0.1.16 (Li et al., 2009b) was used to convert the output alignment SAM files to the BAM file format, sort the alignments and eliminate duplicated reads. BEDTools 2.23.0 (Quinlan, 2014) was used to convert the resulting files to the BED format. All ChIP and input samples were randomly normalized to the same number of reads. Peak calling was performed with MACS 2.0.10.20130712 (Feng et al., 2012) using the input sample as control for each one of the ChIP samples. BED files containing aligned reads for each sample, with duplicates removed and with a balanced number of reads, used as input for MACS2. BigWig files were obtained with bedGraphToBigWig (Kent et al., 2010) from the BedGraph files generated with MACS. Resulting peaks were annotated with PeakAnalyzer 1.4 (Salmon-Divon et al., 2010) and the distribution of peaks was plotted with SeqMiner 1.3.3e (Ye et al., 2014). Genome build NCBIm37/mm9.

Data quality

Sequencing quality was assessed by Fastqc quality control checks.
MACS settings for peak calling:
tag size = 25
band width = 300
model fold = 32
pvalue cutoff = 1.00e-05
Ranges for calculating regional lambda are : peak_region,1000,5000,10000

Software

ChIP-seq analysis was performed with the RUbioSeq pipeline (v3.8), using the following tools: FastQC v0.11.5, BWA v0.7.10, SAMtools v0.1.19, Picard tools v1.107, Bedtools v2.16.2 and MACS2 v2.1.1.20160309. Sequencing quality for ChIP-seq samples was analyzed with FastQC (Andrews, 2011). Reads were aligned with Bwa 0.7.5a (Li and Durbin, 2009) to the mouse reference genome (NCBIm37/mm9) using the default seed length (32) and allowing 1 mismatch in the seed. SAMtools 0.1.16 (Li et al., 2009b) was used to convert the output alignment SAM files to the BAM file format, sort the alignments and eliminate duplicated reads. BEDTools 2.23.0 (Quinlan, 2014) was used to convert the resulting files to the BED format. All ChIP and input samples were randomly normalized to the same number of reads. Peak calling was performed with MACS 2.0.10.20130712 (Feng et al., 2012) using the input sample as control for each one of the ChIP samples. BED files containing aligned reads for each sample, with duplicates removed and with a balanced number of reads, used as input for MACS2. BigWig files were obtained with bedGraphToBigWig (Kent et al., 2010) from the BedGraph files generated with MACS. Resulting peaks were annotated with PeakAnalyzer 1.4 (Salmon-Divon et al., 2010) and the distribution of peaks was plotted with SeqMiner 1.3.3e (Ye et al., 2014). Genome build NCBIm37/mm9.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cell cultures were harvested by trypsinization to a single cell suspension (ensured by filtration through 0,7 micro-metre pore filters) and then either analyzed by FACS directly (for for fluorescent reporter lines), alternatively labelled with antibodies for pluripotency markers (as indicated in each Figure panel).
Instrument	Becton-Dickinson FACScalibur -Cytometric analyses. FACS-Aria -for live-cell sorting.
Software	Data were analyzed with FlowJo 9.6.2 software.
Cell population abundance	The abundance of the selected cell populations is reported in Figure panels where relevant
Gating strategy	A live/dead stain was used to exclude dead cells and debris from the analysis of live-cell fluorescent reporter cell lines. See Extended Data Figure 9G for FACS cytometry images of gating strategy employed here.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.