MED24, a tail component of the Mediator complex, links FGF/ERK signaling to nuclear receptor engagement

Master Thesis in Biology (Cell Biology and Physiology)

Maria Cristina Pires Clérigo (qjl216)

Supervisor: Ala Trusina

       Joshua Brickman

Submitted on: 28 July 2021

Abstract

Embryonic stem cells (ESCs) are described as pluripotent, meaning they can differentiate into all cell lineages. In mouse, FGF/ERK signaling is known to regulate the formation of the primitive endoderm (PrE) lineage. MED24, a Mediator complex subunit, has an essential role in regulating ERK-mediated transcription in ESC, with the loss of MED24 compromising the formation of PrE. In this thesis, I investigated the role for MED24 in ESC transcriptional regulation. I found that a nuclear hormone receptor (NR) domain in the C-terminus of MED24 is required for nuclear localization, suggesting there are specific NRs that control MED24 function. I also found that shortly after MED24 depletion ground state pluripotency is compromised with *Nanog* expression being reduced. However, this effect can only be maintained in the presence of FGF/ERK pathway inhibition. As the *Nanog* enhancer and other FGF/ERK pluripotency enhancers are dependent on NR, this result offers a link between MED24 interactions and transcriptional regulation. Moreover, MED24 appears to be required for the prolonged assembly of new Mediator complex at the differentiation enhancers, which is required for PrE maturation.

In conclusion, MED24 regulation of transcription in ESCs requires engagement of a specific NR domain and its specificity is likely controlled by levels of ERK signaling. Further analysis will provide us with a better understating of how these mechanisms feed into a bigger picture of ESC transcription regulation.

Contents

Acknowledgments

This thesis was so far the biggest challenge to my resilience, and I would have not be able to go through everything without the amazing people from the Brickman group. I will be forever grateful for having the opportunity to learn with all of you.

First, I would like to thank to Professor Josh Brickman for giving me the opportunity to work on this project, for your support and tolerance. Thank you for always making me think about the big questions.

Nothing that I would write here will ever express enough my gratitude to a group of people in this lab. When I say that without them none of this would have been possible, I am not exaggerating. The biggest thank you to Rita, Fung, Molly and Tina. I will be forever grateful for having the opportunity to learn and to work with you. You made this journey worth it! You're my support in every aspect of this project.

Thank you to everyone in the Brickman lab, for all the nice moments, all the long conversations and for always keeping in the good mood. Thank you all for your patience, especially in the last months, I couldn't ask for a better group of people to work with.

Thank you to my office mates, Michaela, Martin and José, even in darker times you made come to work easier. Michaela, somehow you were the maternal support that I did not know I need it. Martin, Jas and José you are the best, thank you for all the fun moments and the non-sense conversations!

Thank you to all my amazing friends that were so essential in this journey, thank you for always taking my mind from the lab problems Gracias á La Mari Guapetona!

Mil obrigados à minha outra metade da laranja, Ana! Sem ti nada disto teria sido possível. És e serás sempre um dos meus role models em todos os aspectos da vida. Não posso estar mais grata por ter embarcado nesta aventura contigo.

Aos meus pais e à minha irmã, obrigado por aturarem sempre o meu mau humor e por terem sempre paciência para mim. As saudades marcaram este percurso que não foi fácil, mas tenho a certeza de que no fim valeu a pena. A resiliência que eu demonstrei durante este percurso aprendi com vocês. Espero que tenha conseguido deixar-vos minimamente orgulhosos

Abbreviations

AID auxin inducible degron
BMP Bone morphogenic protein
ChIP-seq Chromatin immunoprecipitation followed by whole genome sequencing
CHIR CHIR99021, CHIRON, inhibitor of GSK-3
CUT&RUN Cleavage Under Target & Release Using Nuclease
E Embryonic day
E14 E14JU, mouse ESC line
Epi Epiblast
EpiSCs Epiblast stem cells
ESCs Embryonic stem cells
FACS Fluorescence activated cell sorting
FBS Fetal bovine serum
FGF Fibroblast growth factor
FGFR FGF receptor
gDNA Genomic DNA
GAPDH Glyceraldehyde-3-Phosphate Dehydrogenase
GFP Green fluorescent protein
GSK-3 Glycogen synthase kinase-3
GTF General transcription factor
H3K27ac Histone H3 Lysine 27 acetylation
IAA Indole-3-acetic acid
ICM Inner cell mass
JAK Janus associated kinase
KO Knock-out
LIF Leukaemia Inhibitory Factor
MAPK Mitogen-activated protein kinase
MED Mediator
MED1 Mediator subunit 1
MED16 Mediator subunit 16
MED24 Mediator subunit 24
MED23 Mediator subunit 23
NR nuclear receptor
PBS Phosphate buffered saline
PD03 PD0325901, FGF-MAPK signaling inhibitor
PD17 PD173074, FGF-MAPK signaling inhibitor
PFA Paraformaldehyde
PIC Preinitiation complex
PrE Primitive endoderm
RNA Pol II RNA polymerase II
RT-qPCR Real-time quantitative Polymerase Chain Reaction
TR tyrosine hormone receptor
TE Trophectoderm
TF Transcription factor
TSS Transcription start site
UPL Universal Probe Library

List of Figures and Tables

1. Introduction

2. Materials and Methods

3. Results

4. Discussion

1. Introduction

1.1 Murine embryonic development

1.1.1 From Zygote to Blastocyst – Morphological and potency changes

Murine embryonic development of an organism starts at fertilization with a sperm and an egg fusing together to form a zygote, a single totipotent cell that has the capacity to produce all cell lineages and ultimately form a whole organism. Initially the mouse embryo relies on maternal information prevenient of the oocyte but by embryonic day 1.5 (E1.5), the zygotic genome is entirely activated; the maternal mRNA starts to be degraded and upon the first division the 2-cell embryo becomes a fully self-organizing structure (Fig. 1.1). The cell then undergoes a set number of mitotic cleavage divisions, also known as 'reductive divisions'; as the number of individual blastomeres (cells) increases, the size of the embryo remains approximately the same (Lehtonen, 1980). At the 8-cell stage (E2.5), after three rounds of divisions, the cells undergo compaction, a crucial event for mammalian cleavage (Peyrieras et al., 1983). E-cadherin, a cell adhesion protein evenly expressed throughout the cell membrane prior to compaction, gets restricted to the basolateral cell membrane. Cells once loosely arranged form functional adhesion junctions, resulting in a dense ball of cells (Larue et al., 1994). By this stage, the first lineage specification begins; cells start to exhibit polarity and form a 16-cell embryo, also known as the morula (E3.0). Cell polarity is based on structural and functional asymmetric distribution of cellular components between the two different cell membrane domains, apical and basolateral. The majority of the apolar internal cells give rise to the inner cell mass (ICM) via symmetric divisions, while the polarized outer cells, formed by asymmetric divisions, become the extra-embryonic trophectoderm (TE) (Barlow et al., 1972). Lineage tracing experiments showed that these morphologically distinct regions have specific transcription factor (TF) expression, with TE marked by CDX2 and ICM by POU5F1 (hereafter referred to as OCT4) (Niwa et al., 2005; Dietrich & Hiiragi, 2007; Korotkevich et al., 2017).

As mentioned, the initial totipotent cell has the potential to give rise to all cell types in the organism. This potency decreases as the embryo developments with cell fates becoming restricted. While ICM cells lose their totipotency, they remain pluripotent and can contribute to all three germ layers (ectoderm, mesoderm and endoderm). Contrastingly, TE cells lose the capacity to generate embryonic structures. Instead, TE cells responsible for forming structures that ensure implantation as well as the support of the embryo: the chorion, the extraembryonic membrane and a portion of

the placenta. TE cells pump Na^+ ions into the intracellular spaces through the Na^+/K^+ ATPase ion channel present in the basal membrane, creating an ionic gradient and causing water to diffuse into the embryo generating a fluid-filled space denominate 'blastocoel'. This process, known as cavitation, occurs between the 32-cell and 64-cell stage. Upon cavitation, the ICM becomes positioned to one side of the circle of trophoblast cells, producing the hallmark structure of early mammalian development, the blastocyst (Fig. 1.1). This asymmetrical positioning of the ICM represents the first axis formation, with the embryonic pole containing the ICM and the opposite region representing the aembryonic pole. At this point in development (64-128 cells), the ICM acquires a 'salt and pepper' TF expression pattern. The NANOG-expressing epiblast (EPI) precursor cells will go on to generate the embryo proper and the GATA6-expressing presumptive primitive endoderm (PrE) cells will go on to form the extraembryonic tissues: parietal endoderm (PE) and visceral endoderm (VE). At E4.5, this pattern evolves stochastically into two distinct cell layers, with EPI cells surrounded by TE cells on the outside and PrE cells on the inside. During these morphological and potency changes, the embryo travels through the oviduct, reaching the uterus. Finally, the embryo exits the pre-implantation stage by escaping the zona pellucida and adhering to the uterine wall.
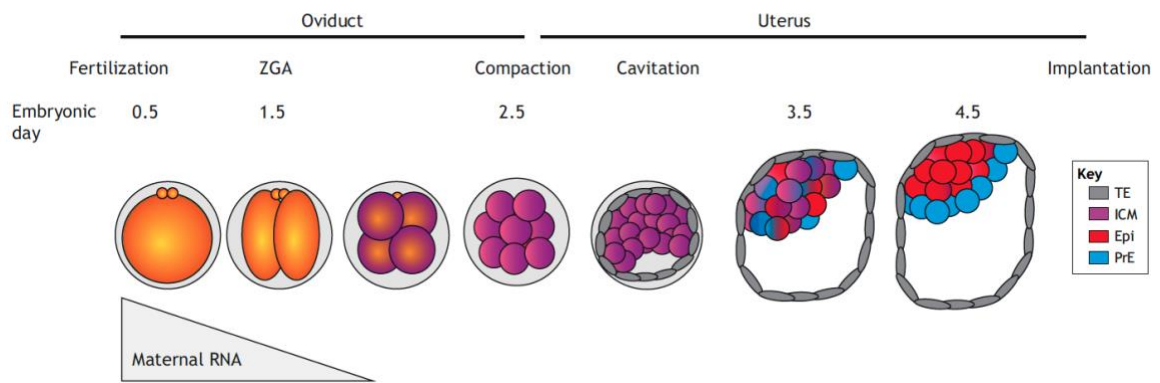
Figure 1.1 - Murine embryonic development from fertilisation to implantation – Schematic representation of morphological differences and lineages specification steps of mouse preimplantation embryo. The fertilised egg, zygote, divides into totipotent blastomeres (E0.0-E2.25). Soon after compaction stage (E2.5), two major segregation events occur: TE and ICM (E3.5); and EPI and PrE (E4.5). Figure adapted from Riveiro A.R and Brickman J., 2020

1.2 The role of TFs in fate specification of the pre-implantation embryo

1.2.1 First lineage choice – TE or ICM?

For many years, researchers have been trying to understand the molecular mechanisms involved in early lineage specification. A set of TFs have been identified as key molecules for this process, OCT4, NANOG and SOX2. This small but core transcriptional network works through autoregulation and mutual negative feedback loops, regulating expression of a larger network of TFs and epigenetic regulators that are implicated in pluripotency maintenance (Loh et al., 2006; Niwa et al., 2009).

OCT4 is considered a master regulator of pluripotency (Niwa et al., 2005) and is crucial for the first lineage segregation event involving the trophectoderm TF-specifier CDX2 (Stumpf et al., 2006). The mutual repression of these TFs leads to the segregation of the first two lineages in the embryo, cells with high CDX2 expression and loss of OCT4 become TE, whereas in presumptive ICM cells higher levels of OCT4 expression leads to inhibition of CDX2 (Niwa et al., 2000). TE specification also relies on the Hippo/YAP pathway, which responds to changes in cell polarisation. As mentioned before during compaction, the blastomeres divided asymmetrically, creating two distinct populations, inner and outer cells. Activation of the Hippo/YAP occurs in outside polar cells, resulting in the activation of differentiation genes and leading to the formation of TE cells (Fig. 1.2a). Knockout of *Cdx2* results in blastocysts with outer cells that maintain expression of OCT4 and NANOG (ICM profile) and are not able to activate TE key markers or upregulate the Hippo/YAP signalling pathway (Stumpf et al., 2006).

More specifically, YAP, a TF that can shuttle in and out of the nucleus, activates another TF TEAD4 when nuclear. These proteins can then work together to trigger the expression of TE-specific genes, like *Cdx2* (autoregulatory loop) and *Gata3* in the outer cells (Fig. 1.2a, b). In *Tead4* null embryos, a functional TE is not formed and upregulation of *Cdx2* does not occur (Nishioka et al., 2008). How cell polarisation is initiated remains unclear but studies have revealed that angiomotion (Amot) might have a role in tethering YAP at the cell membrane impeding its movement to the nucleus. Amot is inhibited by polarity, meaning that in TE cells YAP is free to travel to the nucleus and start the differentiation transcriptional program. In ICM cells, YAP remains in the cytoplasm and upregulation of Cdx2 does not happen (Nishioka et al., 2008) (Fig. 1.2a). The Hippo/YAP signalling pathway is transitory and terminates when a stable gene expression pattern is established for the ICM and TE cells (Lorthongpanich et al., 2013).

1.2.2 Second lineage choice – EPI or PrE?

The next lineage decision in the blastocyst defines the fate of the ICM. It involves a shift from mutual expression to antagonization of lineage-specific TFs and is not dependent on cell position. The key factors in this specification are the TFs mentioned earlier, NANOG and GATA6. NANOG and GATA6 are expressed simultaneously from the morula to early blastocyst stage (Guo et al., 2010; Plusa et al., 2008) (Fig. 1.2c). EPI precursors are characterized by higher expression of NANOG while PrE precursors lose NANOG expression and are marked by the sequential expression of GATA6, SOX17, GATA4 and SOX7 (Artus et al., 2011; Chazaud & Yamanaka, 2016; Niakan et al., 2010; Plusa et al., 2008). *Nanog* null mutants do not form the epiblast layer(Mitsui et al., 2003) and the ICM cells express *Gata6* (Frankenberg et al., 2011). Furthermore, these studies indicate that NANOG and GATA6 directly repress each other through mRNA decay, although there is evidence that this repression might also happen at the transcriptional level(Singh et al., 2007).

Recent studies revealed that the FGF/ERK pathway might play an essential role in this lineage-segregation step, more specifically in PrE formation (Fig. 1.2d). An early difference between EPI and PrE precursors is marked by gene expression of the ligand Fgf4 (EPI) and the membrane receptor, Fgfr2 (PrE) (Guo et al., 2010). At E3.25, Fgfr2 is expressed widely through all ICM cells but becomes restricted to PrE cells by E3.5 (Ohnishi et al., 2014). Blocking FGF signalling leads to the differentiation of ICM cells into EPI, but supplementation with FGF4 pushes ICM cells toward the PrE fate (Yamanaka et al., 2010). Although, it seems evident that FGF signalling has a direct role in controlling PrE fate about 70% of *Nanog* null embryos that were treated with an FGF inhibitor, retain the expression of GATA6. Furthermore, Fgf4 expression is lost in Nanog mutants(Frankenberg et al., 2011). Together these data indicates that FGF promotes PrE fate by repressing Nanog expression (Frankenberg et al., 2011) (Fig. 1.2d).

OCT4 also plays a role in PrE differentiation by regulating *Fgf4* levels (Ambrosetti et al., 1997) (Fig. 1.2d). In *Oct4* null embryos, Fgf4 levels are reduced, resulting in most cells expressing NANOG and the loss of cells with gene expression profiles characteristic of PrE cells (Frum et al., 2013). GATA6 levels cannot be maintained, and the cascade is not triggered, probably due to the decrease of Fgf4 levels. When FGF4 ligand was added to *Oct4* null embryos, PrE fate was not

rescued, indicating that Oct4 has an essential role in PrE segregation downstream of the FGF/ERK pathway.

During PrE differentiation, TFs such as GATA4, SOX17 and PDGFα (cell surface marker) are upregulated. This upregulation might be a direct effect of FGF4 or triggered indirectly by GATA6, SOX2 or OCT4 transcriptional activity. Altogether, these studies point to a network of transcription factors, NANOG, OCT4 and SOX2 that regulate the levels of FGF4 to promote the maturation of PrE and EPI cells (Fig. 1.2d).
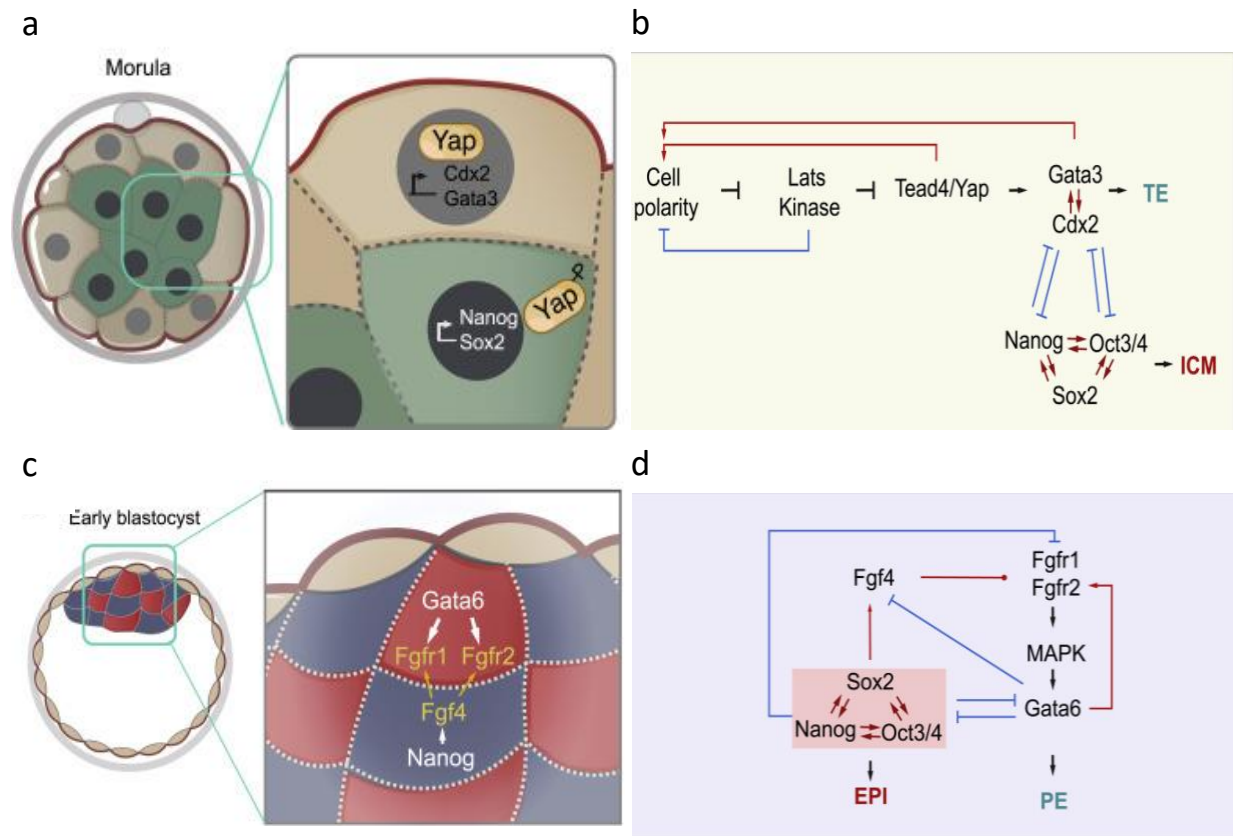
Figure 1.2 – Fate decision processes and TF networks - a) In the morula stage, differences in Hippo/YAP signalling due to cellular positioning (polarity) regulates ICM and TE segregation. In apolar cells, YAP remains in the cytoplasm, allowing the expression of pluripotency factors like NANOG and SOX2. In apical cells, YAP goes to the nucleus where activates the transcription of TE-factors CDX2 and GATA3. b) Cell polarity and its downstream effects are the key factors for the feedback network resulting in segregation of ICM and TE. c) Salt and pepper expression of the TFs involved in the second cell fate decision, NANOG and GATA6, which go on to form the EPI and PrE respectively. NANOG induces the expression of Fgf4 ligand whereas GATA6 leads to the expression of receptor Fgfr1/2. d) EPI and PrE segregation is regulated by an intricate network of positive and negative loops. Fgf4 production is supported by NANOG, OCT4 and SOX2 expression, while GATA6 expression is induced by FGFR activation. High levels of GATA6 lead to the repression of NANOG. Figure adapted from Zhu M. and Zernicka-Goetz M., 2020.

1.3 Embryonic stem cells

More than 30 years ago, murine embryonic stem cells (ESCs) from the inner cell mass of a developing blastocyst were isolated and grown for the first time *in vitro*(Evans & Kaufman, 1981; Martin, 1981). These cells maintain the capability to incorporate back into the developing embryo and can contribute to all the three germ lines, meaning they can maintain their potency outside the embryo (Bradley et al., 1984). The isolation of ESCs represents a major breakthrough for developmental biology, providing an *in vitro* model system to study processes of early embryonic development and cellular differentiation. While the pluripotent state of ESCs represents a short period of time *in vivo*, the power of the system comes from the ability to propagate them indefinitely *in vitro*.

The capturing of the undifferentiated state was accomplished by culturing ESCs on a feeder layer of mitotically inactivated murine embryonic fibroblasts (MEFs) in the presence of calf serum (Koopman & Cotton, 1984; A. G. Smith & Hooper, 1987; T. A. Smith & Hooper, 1983). Later, purification of one secreted protein showed to be responsible for the inhibition of differentiation, the cytokine leukaemia inhibitory factor (LIF)(A. G. Smith & Hooper, 1987; Williams et al., 1988.). LIF signals through a membrane receptor to activate Janus-associated kinases (JAKs), phosphorylating the TF STAT3 (Niwa et al., 1998) (Fig. 1.3a). LIF together with serum (Serum/LIF) allowed ESCs to be maintained and propagated in a feeder-free condition (A. G. Smith & Hooper, 1987). However, under these culture conditions ESCs appear heterogeneous, with cells having the tendency to adopt an EPI or a PrE-like state (Torres-Padilla & Chambers, 2014). Withdraw of serum from the culture medium promotes cell differentiation, implying that serum provides additional signal(s) required to a fully maintain the pluripotent state. Bone morphogenetic protein 4 (BMP4) can replace serum and sustain self-renewal along with LIF by inducing inhibitor-differentiation (Id) proteins (Ying et al., 2003).

To understand further the properties of the pluripotent state, studies turned back to investigating what triggers differentiation. Based on the fate decisions discussed earlier, a promising candidate was the fibroblast growth factor (FGF) pathway (Fig. 1.3b). As mentioned before, FGF4 plays a major role in the segregation of PrE and EPI and inhibition of FGF/ERK pathway supports ESC self-renewal. ESCs express FGFR2 and produce considerable amounts of FGF4, meaning the signal can act by stimulating the cells that produce the ligand (autocrine signaling) and/or stimulating neighboring cells (paracrine signaling), which contributes to the heterogeneity observed in the culture system(Stauber et al., 2000). Inactivation of downstream targets of FGF4,

like protein kinase ERK2, showed cells become resistant to differentiation and can be maintained in LIF alone (Kunath et al., 2007).

Another pathway implicated in ESC self-renewal is the Wnt/β-catenin pathway (Fig.1.3d) The downstream effector glycogen synthase kinase-3 (GSK3) acts as an antagonist in ESC self-renewal (Sato et al., 2004). Furthermore, GSK-3alpha/beta double knockout leads to hyperactivation of this pathway making cells resistant to differentiation (Sato et al., 2004). Similar to ERK2 ablation, inhibition of GSK3 along with LIF sustains ESC self-renewal. One the most striking observations was the clonal expansion of undifferentiated ESCs by combining inhibitors of ERK and GSK3, PD0325901 and CH99021 respectively, in a defined medium (N2B27) (Ying et al., 2008). Later, this medium condition would be termed 2i to signify the two inhibitors. In these conditions, the inhibitors shield ESCs from differentiation signals promoting a homogeneous early epiblast-like ground state (Nichols et al., 2009; Wray et al., 2012). 2i medium with LIF (2i/LIF), supports self-renewal but also promotes the expansion of trophoblast-like cells, thus resembling the pre-implantation embryo with totipotent cells present in the culture system (Poehlmann et al., 2005) (Fig.1.4).

Overall, mouse ESCs are highly plastic and dynamic, making it possible to use defined culture conditions to mimic different stages of development, from ground state pluripotency, which resembles the pre-implantation blastocyst, to later stages with primed pluripotency, such as gastrulation. These dynamic states are mainly controlled in ESCs by activators/inhibitors of signaling pathways, culminating in TF profiles that produce cell type specific transcriptional programs (Fig. 1.4). ESCs have been a key factor in unveiling how signal transduction pathways modulate the activity of many types of TFs, leading to specific cell fate choices.
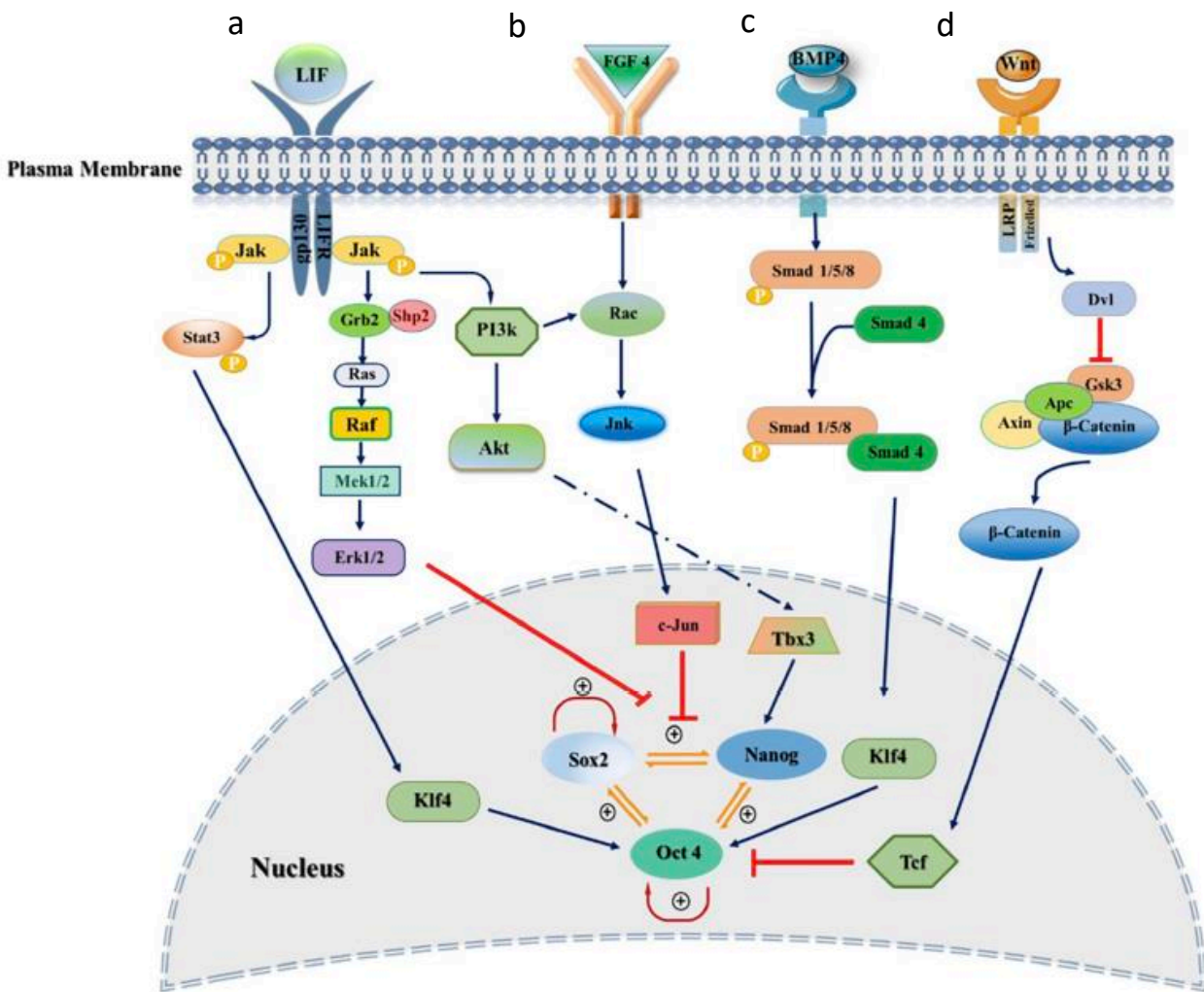
Figure 1.3 - Overview of important extrinsic pathways involved in pluripotency in mESCs - a) LIF binds to heterodimeric recepto complex, that when brought into proximity allows autophosphorylation and autoactivation of JAK molecules. Activated JAKs the phosphorylate downstream STATs, that then translocate to the nucleus and activate target genes associated with self-renewa Alternatively, JAK can activate PI3K pathway, which feeds into the pluripotency network to help maintain the self-renewal state o mESCs. b) FGF4 binds to FGFR2 receptor, leading to an autophosphorylation event that culminates in the recruiting of a scaffol protein, GRB2. Consequently, GRB2 recruits SOS, which then activates RAS. RAS activation leads to a several phosphorylatio events, culminating with phosphorylation and activation of ERK. ERK translocates to the nucleus to initiate the transcription programmes associated with differentiation. c) BMP4 binds to a heterotetrameric receptor complex leading to the activation o SMADs. SMADs then move to the nucleus where they activate target genes associated with self-renewal. d) WNT binds to it heterodimeric receptor leading to the DVL activation. DVL then mediates the breakdown of the destruction complex that i responsible for the degradation of β-CATENIN. Therefore, in the presence of WNT, β-CATENIN is not degradated and ca translocate to the nucleus to activate target genes. Figure adapted from Mossahebi-Mohammadi M. *et al.,* 2020
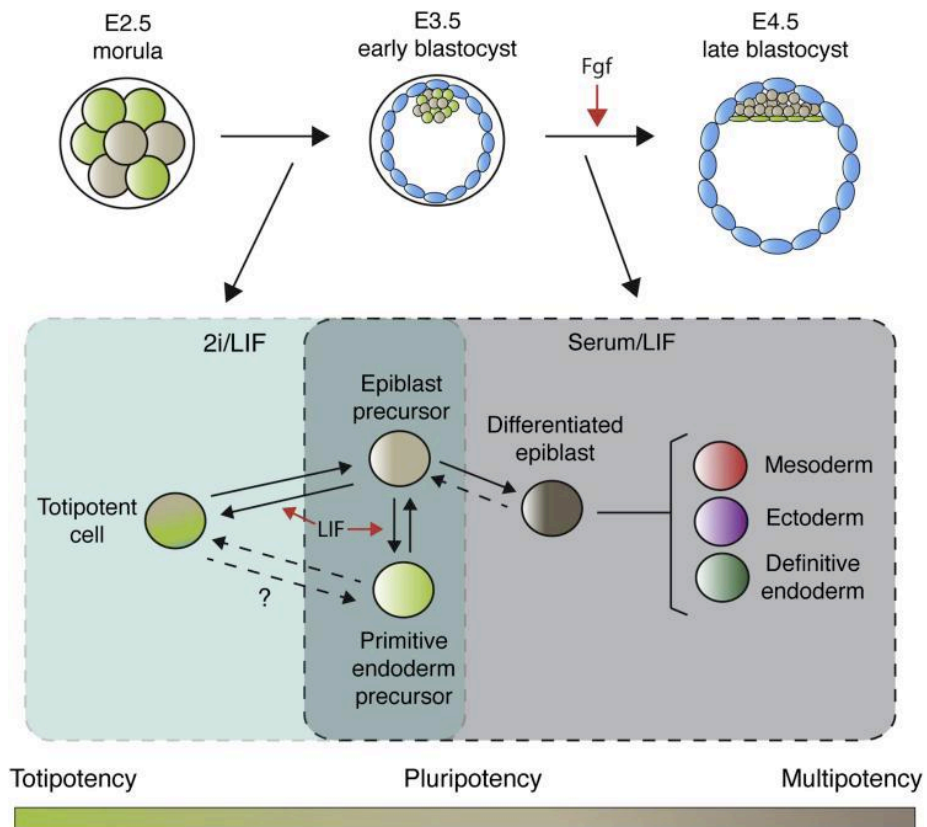
Figure 1.4 - Comparison between Serum/LIF and 2i/LIF culture – While 2i/LIF culture contains totipotent cells that resemble early pre-implantation stages, cells culture in Serum/LIF are reminiscent of late pre-implantation stages, representing a heterogeneous population of cells that are already restricted to certain cell fates. Morgani S. *et al.*, 2013

1.4 Regulation of Gene Expression

A fundamental element behind translating environmental cues to the induction of differentiation pathways i and cell identity maturation is gene expression regulation. Gene expression occurs through transcription, a process that is widely conserved through every branch of life making it essential for every cell. Transcription controls the production of protein-coding transcripts and other RNA products. Precise expression of these products is controlled across time and space through the integration of signals superimposed or encoded in DNA elements. Additionally, transcription is a highly complex and heavily regulated process that involves a wide range of proteins, ensuring that the expressed genetic information is precise and suitable to cell necessities. In eukaryotes, three enzymes are responsible for transcription: RNA polymerase I, transcribes rRNA precursors; RNA polymerase II (Pol II), transcribes all protein-coding genes and some non-coding genes, like small nuclear RNAs (snRNAs); and RNA polymerase III (Poll III), transcribes small non-coding RNAs, like tRNAs. Transcription can be divided into three main phases: initiation, elongation, and termination; with each phase being subject to regulation. Failures in these regulatory processes can lead to serious cell abnormalities and diseases.

1.4.1 RNA Polymerase II mediated transcription

The enzyme responsible for transcribing all protein-coding genes in eukaryotes is RNA polymerase II (Pol II). It is a multiprotein complex composed of 12 subunits, one of which is the key structural component Rpb1, containing a long carboxyl-terminal domain (CTD) with tandem heptamer repeats of the consensus sequence, YSPTSPS (Corden et al., 1985). The CTD is a target of several modifications that affect its conformation and ability to bind to other factors involved in different transcription processes and regulation.

RNA Pol II is not able to transcribe a gene by itself; it requires a concerted activity of transcriptional activators and repressors. The core, gene promoter region is the docking site for general TFs that together with the polymerase form the preinitiation complex (PIC). General TFs make contacts between themselves, the polymerase and DNA, producing a strong and stable bond that also presents an impediment to the moving of RNA polymerase. The assembly of the PIC is a multistep process that starts with Pol II binding to five general TFs, TFIIB, TFIID, TFIIE, TFIIF

and TFIIH at the promoter site (Matsui et al., 1980) (Fig. 1.5a). TFIID is a complex constituted by a TATA-binding protein (TBP, Sawadogo & Roeder, 1985), required for transcription in all promoters, and other 14 TBP-associated factors (TAFs), that have promoter-specific functions. After the assembly and ATP consumption, the PIC promotes the unravelling of the double-stranded DNA, creating a transcription bubble where the promoter template strand is positioned in proximity of the Pol II active site, forming the Open Complex (W. Wang et al., 1992). Transcription initiates when RNA Pol II CTD is phosphorylated by TFIIH on serine 5 (Komarnitsky et al., 2000), and synthesis of small RNAs (3-10 bp), also known as abortive products, happens (Luse & Jacob, 1987) (Fig. 1.5a). Then the polymerase breaks its connections with the core promoter and the rest of the transcriptional machinery and starts transcribing full-length RNAs, entering the elongation phase (Holstege et al., 1997). After the synthesis of 20-40 bp the RNA Pol II pauses, this phenomenon is called promoter-proximal pausing. It was first described in Drosophila melanogaster (Rougvie & Lis, 1988) and later shown to be conserved in mammalian cells (Strobl & Eick, 1992). RNA pol II needs an extra phosphorylation step on serine 2 of the CTD to resume transcription (Komarnitsky et al., 2000). After elongation, RNA pol II reaches the transcription termination site (TTS), where the transcript is released, and the polymerase dissociates from the DNA strand.

Normally, genes transcribed by Pol II have different *cis*-acting elements, such as the proximal promoter (localized <1kb from the transcription start site (TSS)) and distal enhancers, silencers, insulators, or locus control regions (>kb from TSS). The exact composition and location of these elements plays an important role in the establishment of promoter-enhancer interactions and in the initiation of transcription.

1.4.2 Transcription regulation by TFs and enhancer elements

Transcription in eukaryotic cells is a highly dynamic and heterogeneous process with respect to specificity and rate, resulting in tight regulation of gene expression. This is achieved through the regulation of transcription at several key steps, one being the initiation of transcription (Darnell, 1982). This mechanism has different layers of regulation, while the general TFs regulate the position of RNA pol II on the core promoter region, sequence-specific TFs also have the capacity to regulate TSS activity. TF activity can be modulated directly by signal transduction pathways (D.S. Latchman, 1993) acting as nuclear messengers. By binding to DNA regulatory sequences (promoters and enhancers), TFs can control the rate of transcription. TFs bind gene regulatory

elements by recognizing sequence motifs, meaning that a specific enhancer or promoter can be bound by several TFs, depending on the DNA sequence. Besides the sequence-specific DNA binding domains, TFs also possess regulatory domains that interact with other transcription regulators.

Enhancers are a crucial regulatory element of gene expression that were first characterised more than 40 years ago. The first enhancer ever described was SV40 from viral DNA, a 72 bp fragment that when localized upstream of the β-globin gene was capable of stimulating its transcription by more than two orders of magnitude independently of its direction and over distances over than 3000bp (Banerji et al., 1981). In the following years, diverse viral enhancers were described, and several ones were described in animal genomes (Serfling et al., 1985). Extensively studies have contributed to the current general definition of enhancers, a DNA sequence that is able to strongly stimulate gene transcription independently of its genomic orientation, location (upstream or downstream) and distance to the gene body. As referred to above, enhancers contain short DNA motifs that act as binding sites for specific TFs, that can recruit co-activators and co-repressors, chromatin remodelling enzymes and ultimately factors that can bind the transcription initiation apparatus (Soutourina, 2018). Hence, regulatory cues from all the bound factors are integrated to determine the enhancer activity (Shlyueva et al., 2014).

1.4.3 Chromatin structure

Normally DNA in eukaryotic cells is found in a complex known as chromatin, meaning that the DNA is bound to proteins and packed in compacted, dense structures. Dynamic restriction of chromatin structure of is necessary for processes such as DNA replication and gene expression. For example, the chromatin at enhancers needs to be open and accessible for the assembly of diverse proteins, like TFs. Consequently, active enhancers are usually devoid of nucleosomes making the chromatin very sensitive to DNase I, enabling the use of DNase assays to assess the chromatin state. Another hallmark of active enhancers is their close proximity to histones containing specific post-translational modifications, like histone H3 lysine 4 monomethylation (H3K4me1) and H3K27 acetylation (H3K27ac) at the N-terminus (Heintzman et al., 2007). Lastly, mapping of enhancers RNA (eRNAs) is also used to identify active enhancers, these molecules are a result of uni- or bi-directional transcription at the enhancer site by RNA Pol II (Core et al., 2014).

How enhancer regions with bound TFs and cofactors communicate over long distances to regulate RNA pol II activity on target genes is still poorly understood. More than 20 years ago, in bacteria, there was evidence suggesting that enhancers activate gene expression by bringing transcription machinery to the promoter region (Ptashne, M., Gann, 1997). Later, evidence in eukaryotic cells suggested that distal regulatory elements acted on transcription regulation by looping into close physical proximity to target promoters, forming a three-dimensional (3D) chromatin structure to facilitate enhancer-promoter (E-P) communication (Bulger & Groudine, 1999) (Fig. 1.5a)(Bulger & Groudine, 1999). These long-range chromosomal interactions are further stabilized by binding of enhancer-bound transcription factors to coactivators, like cohesin and Mediator complex (Kagey et al., 2010) (Fig. 1.5a). This describes the 'looping model', one of the diverse models that attemps to explain E-P communication. Another model, 'facilitated-tracking model', suggests that the transcriptional machinery first binds to the enhancer and then tracks through the DNA until it finds the promoter (Vernimmen & Bickmore, 2015).

DNA loops occur at sites delimited by insulator molecules (CTCF) and by the cohesin extruding factor. While cohesin deletion in mESCs leads to loss of loop domains, similarly to CTCF deletion, it only has a mild effect upon levels of transcription (Nora et al., 2017; Rao et al., 2017). CTCF and cohesin may serve as important architectural proteins to establish DNA loop extrusion between enhancers and promoters, but additional mechanisms must be involved in the maintenance of E-P communication. While CTCF and cohesion have been shown to be involved in acting on the chromatin structure the Mediator complex is thought to play a role in the direct communication of enhancer and promoter regions (El Khattabi et al., 2019).

1.4.4 Mediator complex

Transcription has been a subject of study for many years, with some components of transcription machinery conserved from single to multi-cellular organisms. They diverge on how RNA polymerase II is targeted by co-regulators. In bacteria the polymerase is directly regulated by the binding of activators or repressors, while in eukaryotes intermediary factors like chromatin modifiers and Mediator play a major role in transcription activation (Kornberg, 2005). While most studies and information about transcription regulation and its activators, like Mediator, originally

came from yeast, however, it seems that a high degree of similarity remains throughout eukaryotes evolution (Asturias et al., 1999; Malik & Roeder, 2000).

Mediator is a multi-subunit complex that was identified 30 years ago as a co-regulator of transcription in yeast (Kelleher et al., 1990; Lue et al., 1991). Mediator subunits were first identified as thyroid hormone receptor-associated proteins (TRAP); when associated with thyroid hormone receptor (TR) these proteins were shown to activate transcription from a promoter containing a T3-response elements (TREs) (Fondell et al., 1996). Later, it was shown that the complex also interacted with other activators, like sterol-responsive element-binding protein (SREBP) (Näär et al., 1998) and the vitamin D receptor (Rachez et al., 1999). Since then, several studies in eukaryotes suggest Mediator acts as a functional bridge between TFs and the basal machinery, promoting the assembly of the PIC complex (Kagey et al., 2010; Kornberg, 2005; Malik & Roeder, 2005).

Mediator is composed of 25 subunits in budding yeast cells, while it has up to 30 different subunits in mammalian cells. In both organisms, this multi-subunit complex is formed by four different modules: head, middle, tail and CDK8 (Dotson et al., 2000, Fig. 5b, c). The Mediator structure is not static, its subunits composition can change affecting the biological function of the complex. Biochemical studies suggest that Mediator composition changes during cell development, simplifying throughout differentiation (D'Alessio et al., 2009; Deato et al., 2008). Mediator has two main isoforms, binding to CDK8 module or to RNA Pol II (El Khattabi et al., 2019; G. Wang et al., 2001). Knockout of different mammalian subunits (*Med1*, Med21, *Med23*, *Med24* and *Cdk8*) was shown to be lethal in embryos, revealing that they are required for viability and likely play essential roles in transcription initiation (Ito et al., 1999, 2002; Stevens et al., 2002a; Tudor et al., 1999; Westerling et al., 2007). In yeast, MED17 was shown to be required for the expression of mostly protein-coding genes (Thompson & Young, 1995), indicating that Mediator might be necessary for Pol II activity and consequently controls gene expression patterns by being an integrator of signalling cascades.

Mediator conformation can also be affected by the binding of activators and nuclear receptors, suggesting these conformational changes could be used to regulate transcriptional activity (Taatjes et al., 2002, 2004). Several studies have reported that Mediator activity is regulated by post-translational modifications (PTM) of its subunits (Belakavadi et al., 2008; Pandey et al., 2005). In the last years, it has been shown that binding of TF activation domains was enough to induce conformational changes in the mammalian Mediator (Ebmeier & Taatjes, 2010; Taatjes et al., 2002, 2004).

Mediator interacts with RNA Pol II, forming a stable complex denominated RNA Pol II holoenzyme (Asturias et al., 1999). This stable complex is based on extensive contacts, most of them occurring between the head and the middle module of the Mediator and the back surface of the polymerase (Chadick & Asturias, 2005). The Mediator head module can adopt different conformations due to a movable jaw (Cai et al., 2010), a malleable structure revealed to be extremely important to accommodate the polymerase (Davis et al., 2002). In human cells, binding of Pol II CTD to Mediator triggers extensive structural changes in the head module (Bernecky & Taatjes, 2012; Näär et al., 2002). Structural analysis of the protein and profiling chromatin binding (ChIP-seq) data, supports the previous statement that the head module interacts with the RNA Pol II while the tail module is in contact with enhancer elements (Petrenko et al., 2016; Plaschka et al., 2015; Soutourina, 2018).

The specific mechanisms for how Mediator complex regulates E-P communication remains unclear. Recent studies suggest the complex acts as a functional bridge rather than an architectural bridge between the elements. Depletion of Med14, an essential Mediator subunit, did not affect the loops between E-P. While depletion of cohesin abolished DNA loops, Mediator and RNA Pol II could still be recruited to enhancer and promoter sites. These findings suggest that cohesin is essential for the physical formation of loops while Mediator ensures the functionality of E-P communications (El Khattabi et al., 2019).

Lastly, as mentioned before, Mediator complex can be a target of different PTMs, acting as an integrator of different signalling pathways that culminate in gene transcription responses. Phosphorylation led by kinase cascades, like FGF-ERK pathway, has been shown to act as a regulatory mechanism for the activation of Mediator subunits, like MED1 (Belakavadi et al., 2008; Pandey et al., 2005). Also, diverse Mediator subunits are involved in differentiation processes. For example, MED20 subunit is required for adipogenesis while *Med23* knockout (KO) enhances neural differentiation (Tang et al., 2021; Zhu et al., 2015). However how these signalling pathways control gene transcription and reach specificity remains elusive.
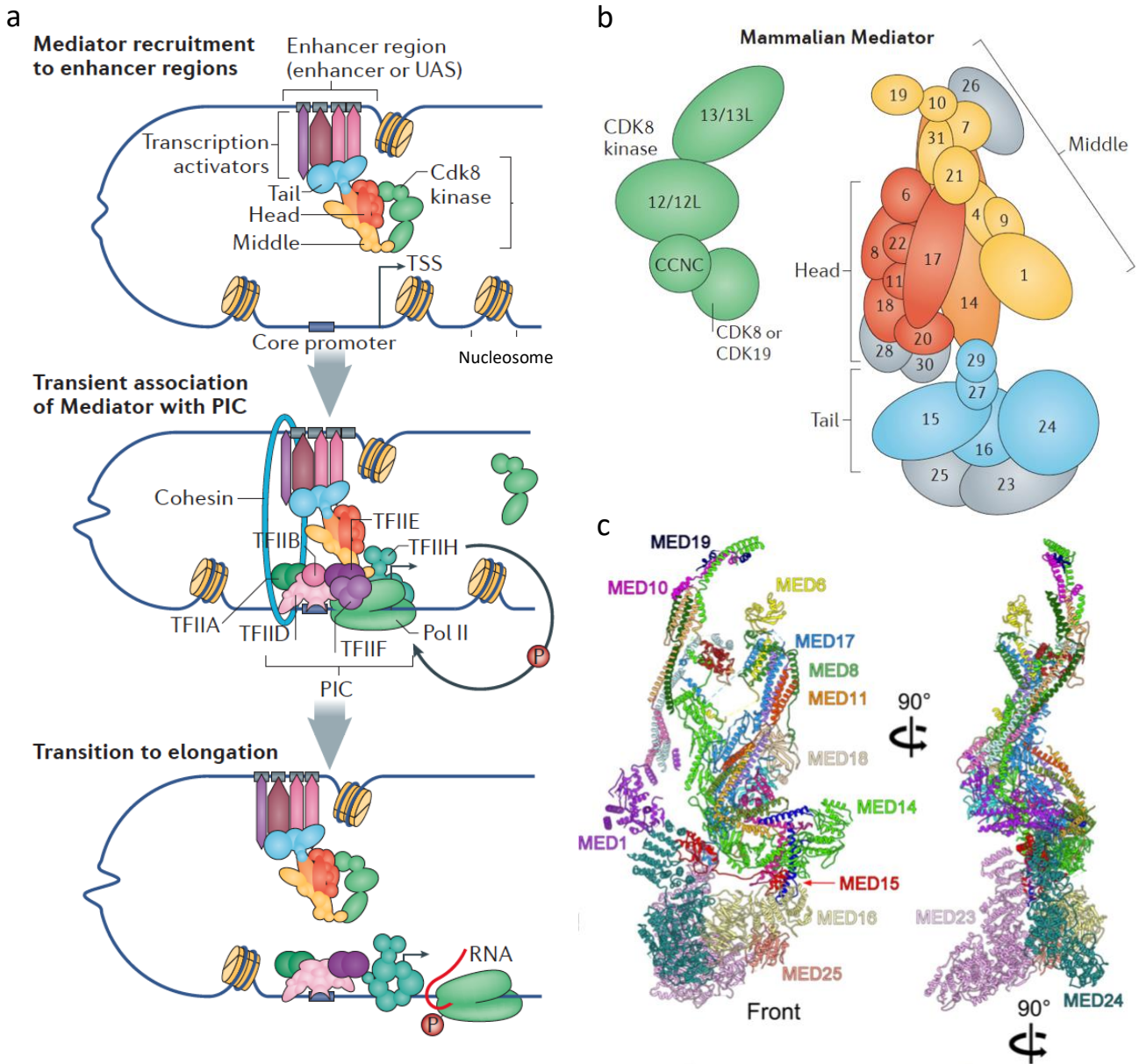
Figure 1.5 - Mediator complex acts as a bridge between enhancers and promoters - a) Mediator complex interacts with transcription factors bound to enhancers and with PIC components at promoter sites. These interactions are responsible for the looping of chromatin, bringing enhancers close to promoters. Mediator interacts with PIC components, as well as other structural components, like cohesin. Mediator exists mainly in two different conformations, associated with Cdk8 module, or associated with PIC components and without the CDK8 module. b) Mediator complex is a multisubunit complex that can be divided into 4 modules: Head (red), middle (yellow), tail(blue) CDK kinase (green). Med14 is coloured in orange representing an essential core subunit for Mediator functionality. c) Cryo-EM structure of Mediator complex, with each subunit color coded. Figure adapted from Soutourina J., 2018 and Zhao H. *et al.*, 2020.

1.4.5 FGF/ERK signalling in transcription regulation

Growth factor signalling pathways, modulated by extracellular stimuli or cell intrinsic regulators, are critical for regulating gene expression, cellular identity and tissue patterning during development .The fibroblast growth factors (FGFs) are a family of secreted proteins that were first described in the 1980s. FGFs are involved in many developmental processes such as patterning, cell proliferation, and differentiation (Fürthauer et al., 2004; Nutt et al., 2001; Slack et al., 2015). More specifically, it is involved in proliferation and differentiation of mESCs (Roux & Blenis, 2004; Thisse & Thisse, 2005). Upon binding to their receptors (FGFR, a tyrosine kinase receptor), FGFs trigger the activation of different effectors including Ras, mitogen-activated protein kinases (MAPK) and phospholipase-C. Most of these activations occur through a phosphorylation cascade, the sequential transfer of a phosphoryl group (PO_3^-) to downstream factors (Fig. 1.6). Enzymes, like kinases, usually carry out these chemical reactions.

In vertebrates, five families of MAPK cascades are extensively described: extracellular signal-regulated kinases (ERKs) 1 and 2 (ERK1/2), c-Jun amino-terminal kinases (JNKs) 1, 2, and 3, p38 isoforms α, β, γ and δ, ERKs 3 and 4, and ERK5 (Uhlik et al., 2004). Although each family has their own characteristics, some features are common throughout all MAPK pathways. Every pathway is comprised of three kinases that are responsible for establishing a sequential activation cascade, MAPK kinase kinase (MKKK), MAPK kinase (MKK) and MAPK (Fig. 1.6). The RAF-MEK-ERK cascade belongs to the MAPK family.

The ERK 1/2 family responds mainly to growth and mitogenic factors, these signals typically activate tyrosine kinases (RSK) or G protein-coupled receptors present in the cell membrane that trigger a cascade of phosphorylation. In the RAF-MEK-ERK pathway, proteins of the Raf family function as MKKK and it has been shown that they phosphorylate and activate MKKs called MEKs, which in turn activate MAPKs, namely ERKs (Fig. 1.6). Furthermore, when ERK (ERK1 and ERK2) is active it changes location, specifically it moves into the nucleus, where it activates a set of different targets (TFs), leading to changes in the gene expression.

FGF4 is produced in an autocrine fashion by the ICM cells (Niswander & Martin, 1992). *Fgf4* homozygous mutants are present at normal Medelian ratios until E3.5, but after this stage the survival rate decreases drastically, with embryos showing disorganized structures by E5.5. FGF4 is necessary for peri-implantation development, specifically for survival of ICM cells (Feldman et al., 1995). However, mESCs cultured in the presence of recombinant FGF4 do not increase in general cell number but do increase in the number of endoderm-like cells (Rappolee et al., 1994). ESCs *Fgf4* KO cells are viable and do not show any proliferative defect when cultured in self-

renewal conditions, but when LIF is withdrawn and retinoic acid is added, cells fail to differentiate and cell survival drops drastically (Wilder et al., 1997). As mentioned before, blocking FGF signalling leads to the differentiation of ICM cells into EPI, while supplementation with FGF4 pushes ICM cells toward the PrE fate (Yamanaka et al., 2010).These studies indicate that FGF4 is essential to promoting PrE differentiation.

Therefore, by manipulating FGF/ERK signalling it is possible to support the existence of multiple, interconvertible cell states. Nichols and colleagues showed that the segregation of the ICM into PrE and epiblast is dependent on MEK-ERK activity by using small molecules inhibitors of both FGFR and MEK1/2 (Nichols et al., 2009) (Fig. 1.7). Later, it was shown that segregation of the PrE layer could be driven in a reversible manner, by treating blastocyst with high doses of recombinant FGF4 (Yamanaka et al., 2010). On the other hand, it is possible to block PrE fate by culturing mESCs in the presence of MEK inhibitor (PD0325901) alone or alongside the FGFR inhibitor (PD173074) (Mohammadi et al., 1998).

Over the past 30 years an intense characterization has improved our understand around what happens downstream of ERK phosphorylation but the molecular role for ERK in transcription regulation is far from complete. ERK can directly phosphorylate a number of key pluripotency TFs: OCT4, NANOG, KLF4 and KLF2, inducing the degradation of the latter two (Fig.1.7) (Brumbaugh et al., 2014; Dhaliwal et al., 2018; Kim et al., 2012). Furthermore, ERK has been linked to the recruitment of Polycomb repressive complex 2 (PRC2), a chromatin-modifying enzyme, and phosphorylation of RNA polymerase II CTD, specifically on developmental genes. This specific phosphorylation was shown to be associated with poised RNA Pol II, a mechanism that was suggested to facilitate the interconvertible transcriptional states of ESCs (Tee et al., 2014). Finally, ERK has a permissive role in lineage specification acting as a licensing factor, by regulating the transcription response of mESCs.
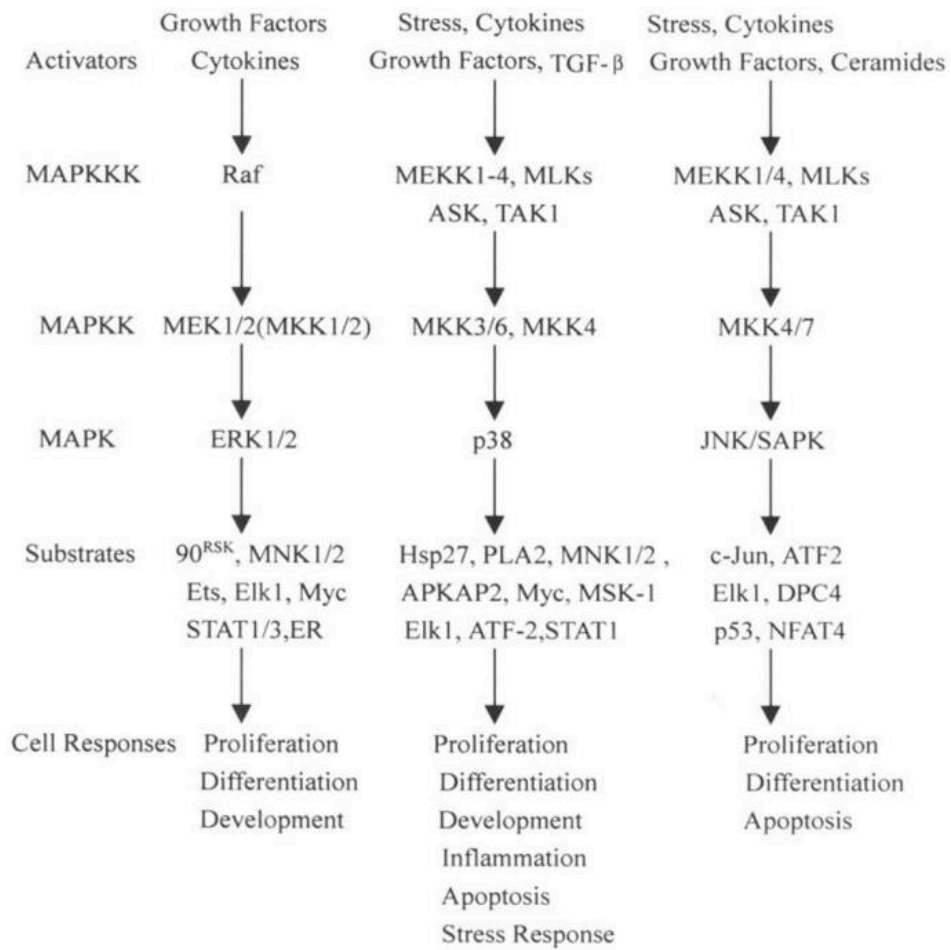
Figure 1.6 - **Principal MAP kinase cascades in mammalian cells** - Three major pathways constituted by a three-tiered kinase cascade: MAP kinase kinase kinase (MAPKKK), MAP kinase kinase (MAPKK) and MAPK, which mediates responses to specific stimuli. MEKK: mitogen-activated kinase kinase kinase; MLK: mixed lineage kinase; TAK: Tat-associated kinase; MTK: mitogen-activated protein kinase kinase kinase 4; RAF: RAF proto-oncogene serine/threonine kinase; MEK: mitogen-activated protein kinase kinase; MKK: mitogen-activated protein kinase kinase; ERK: extracellular regulated kinase; JNK: c-Jun N-terminal kinase. Zhang W. *et al.*, 2002.
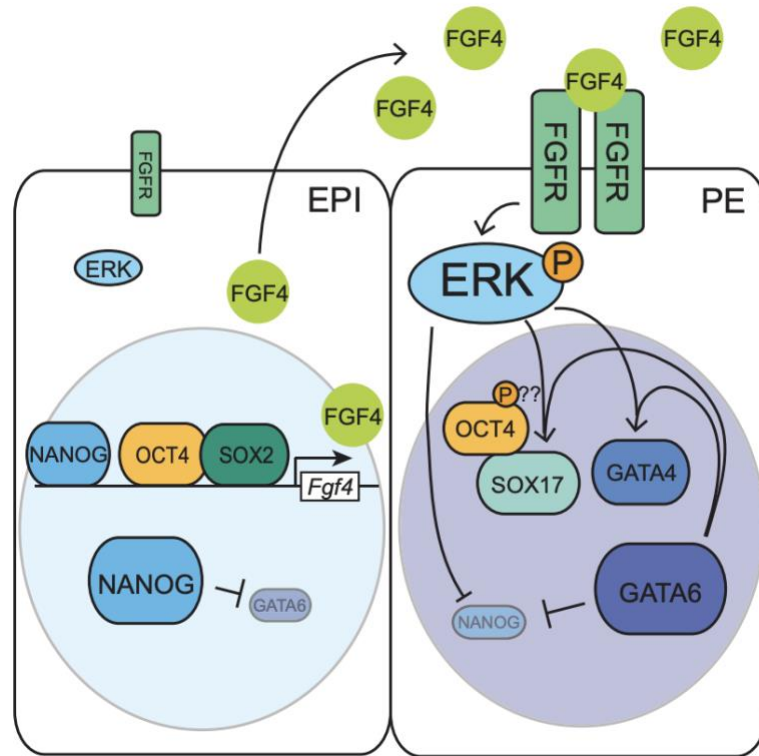
Figure 1.7 - **FGF/ERK pathway controls the segregation of EPI and PrE cells** - In EPI cells (left), NANOG expression prevails over GATA6, upregulating Fgf4 expression and therefore inhibiting FGF signalling and causing low levels of phosphorylated ERK. However, in PrE cells (right), the FGF receptor binds the Fgf4 secreted by EPI cells, activating ERK pathway. This activation leads to higher levels of GATA6 and repression of NANOG expression. FGF signalling together with GATA6 upregulation prime cells to a PrE cell fate by upregulating the expression of Gata4 and Sox17 genes. Chazaud and Yamanaka, 2016

1.5 Research at the Brickman Laboratory

In the Brickman Laboratory, an inducible gain-of-function system was generated to study the molecular mechanisms of the ERK pathway transcriptional regulation in ESCs. In this system a RAF construct missing the ERK repressive domain was engineered to create a constitutively active form of the protein ($cRAF^{\Delta 26-303}$). Additionally, a variant of the ligand binding domain of the estrogen receptor (ER^{T2}) that binds selectively to 4-hydroxy tamoxifen (4OHT), was fused to the C-terminus (Fig. 1.8a). The resulting fusion protein allowed for a homogeneous, rapid and reversible induction of ERK phosphorylation in response to 4OHT treatment (Fig 1.8b).

By combining this system with downstream inhibition of the ERK pathway via a MEK inhibitor (PD03), it was shown that this pathway in ESCs is involved in primitive endoderm specification, while neural differentiation proceeds normally (Hamilton & Brickman, 2014, Fig. 1.8.c).

Furthermore, ERK activation with this system showed an immediate and rapid effect on the transcriptional response. Within 2 hours of ERK induction, pluripotency genes (*Nanog*, *Tbx3*, *Klf4/2*) were downregulated while endoderm formation and canonical MAPK responses genes (*Myc*, *Egr1*, *Spry4*, *Dusp6*) were upregulated (Fig. 1.8d), as shown by nascent RNA-sequencing ((Hamilton et al., 2019). Hamilton et al. 2019 also showed that ERK defines lineage priming by directly regulating enhancer activity independently of transcription factor binding. ERK triggers the reversible association and dissociation of different factors of the transcription machinery from genes and enhancers. ERK-repressed enhancers are enriched for nuclear hormone receptors, particularly that of oestrogen-related receptor β (ESRRB), whereas activated enhancers were enriched for ETS and AP-1 motifs (Fig. 1.8e). RNA Pol II binding positively correlates with the expression state of ERK-regulated genes, indicating that paused RNA Pol II is not the mechanism by which ERK regulates transcription (Fig.1.9f) (Hamilton et al., 2019). These observations raised questions about the molecular mechanism of this transcriptional response. If transcription factors remain bound during repressed and activated states, how does ERK stimulation achieve specificity? Are there other factors involved in this regulatory response? Extensive analysis of the ERK phosphoproteome identified several potential factors. One of the factors identified was MED24, a subunit of the Mediator complex (Fig. 1.8f).

1.5.1 MED24 subunit

Several pathways responsible for cell growth, differentiation or tissue development have been shown to converge on one or more of Mediator subunits, suggesting that the Mediator might act

as a 'hub' or 'integrator' for transcription regulation (Malik & Roeder, 2010). Since Mediator is proposed to act as a 'bridge' between enhancer-bound TFs and the basal machinery at promoters, the regulation of its subunits will help us understand better how gene expression specificity is accomplished.

MED24, also known as TRAP100, is a tail subunit that has extensive contacts with MED23, MED16, MED15 and MED1 (Fig. 1.9a,b) (Zhao et al., 2020). MED24 plays an important role in connecting the tail module with the essential core module of Mediator (El Khattabi et al., 2019). Previously, MED24 knockout was shown to be lethal in mice during an early developmental stage, specifically mice died with severe cardiac hypoplasia and the development of the central nervous system was severely impaired (Ito et al., 2002). This lethal phenotype could be partly explained by the loss of a tail sub-complex MED16-MED24-MED23, which functions as a co-activator for nuclear hormone receptors and has been implicated in MAPK activity (Ito et al., 2002; Stevens et al., 2002b).

MED24 is specifically phosphorylated by ERK (Hamilton et al., 2019) (Hamilton et al 2019) and is upregulated in early development during EPI and PrE segregation (Boroviak et al., 2015). To study to role of MED24 in ERK-mediated transcriptional responses a conditional knockout cell line was generated (Med24 cKO). A FLAG-MED24 sequence under the control of a tetracycline responsive element (TRE) was inserted randomly into ESCs, followed by CRISPR-mutagenesis KO of the endogenous MED24 alleles (Fig. 1.9c). Using this system, FLAG-MED24 expression can be controlled by the addition of doxycycline, with cells expressing FLAG-MED24 behaving as wild-type cells (Fig. 1.9d)

Depletion of the MED24 subunit led to a significant attenuation in the transcriptional response to ERK, with respect to both activation and repression (Fig. 1.9e) (Hamilton et al., 2019). MED24 responded to ERK signalling, by abandoning repressed enhancers and loading onto induced enhancers together with other components of the PIC machinery (Fig. 1.9f). Knockout of MED24 showed that it was also essential for PrE differentiation, suggesting that MED24 plays a crucial role downstream of ERK in PrE maturation (Fig. 1.9.g) (Hamilton et al., 2019) . In contrast, MED24 KO did not affect neural differentiation, suggesting specificity of this subunit for PrE differentiation. Finally, MED24 phospho-mutants showed a slightly attenuated response to ERK and were unable to fully rescue the WT phenotype for the PrE differentiation. This suggests that ERK-induced phosphorylation is not the only mechanism by which MED24 is regulated, and other protein domains might be involved in the subunit functionality.

1.5.2 MED24 degron system – a rapid and efficient protein depletion system

Murine ESCs are a powerful tool to study the mammalian genome and protein function. The unlimited proliferative capacity and differentiation potential make them an ideal model for gain- and loss-of function (LOF) studies. Despite the versatility of ESCs for protein function studies, for a long-time protein depletion relied on homologous recombination to generate knockouts at the gene level.

The use of CRISPR-mutagenesis to generate loss of function (LOF) alleles or knock-in specific tags in the endogenous loci, is a powerful approach that allows for the study of numerous genes, and consequently proteins, in mammalian cells. Gene knockouts rely on mutation or removal of a gene or an essential part of a gene that ultimately will result in irreversible loss of protein (Housden et al., 2016). This issue can be overcome creating conditional KO, by inducing the expression of the protein of interest at specific time points. However, time required for protein production/depletion can span a few days and is dependent on mRNA stability and protein half-life. Long protein depletion times can lead cell lines to adaptation, preventing or hindering the study of a real loss-of-function phenotype in dynamic cell states like differentiation (Housden et al., 2016).

Ideally, systems to study the necessity of a target protein for cellular responses requires a rapid and efficient depletion at the levels of the protein. This allows for immediate and direct assessment of the depletion effects, without the phenotype being overshadowed by the accumulation of secondary effects. One of the mostly wide used protein degradation systems is the auxin-inducible degron technology. This technology co-opts an auxin-dependent degradation pathway from plants, called the auxin-inducible degron (AID) system to induce rapid and reversible depletion of a protein of interest in the presence of indole-3-acetic acid (IAA; natural auxin). IAA binds to the degron tag (AID) and the F-box transport inhibitor response 1 (TIR1) protein to promote the recruitment of different ubiquitin enzymes. The polyubiquitin IAA/AID complex then induces rapid degradation of the target protein by the proteosome complex (Nishimura et al., 2009).

To study MED24 function downstream of ERK signalling, the limitation associated with a conditional KO at the genomic level, taking almost 72 hours to have a total MED24 depletion, makes it difficult to study the acute effects of LOF. Thus, a reversible MED24-degron system offers the opportunity to study the acute effects of MED24 depletion.
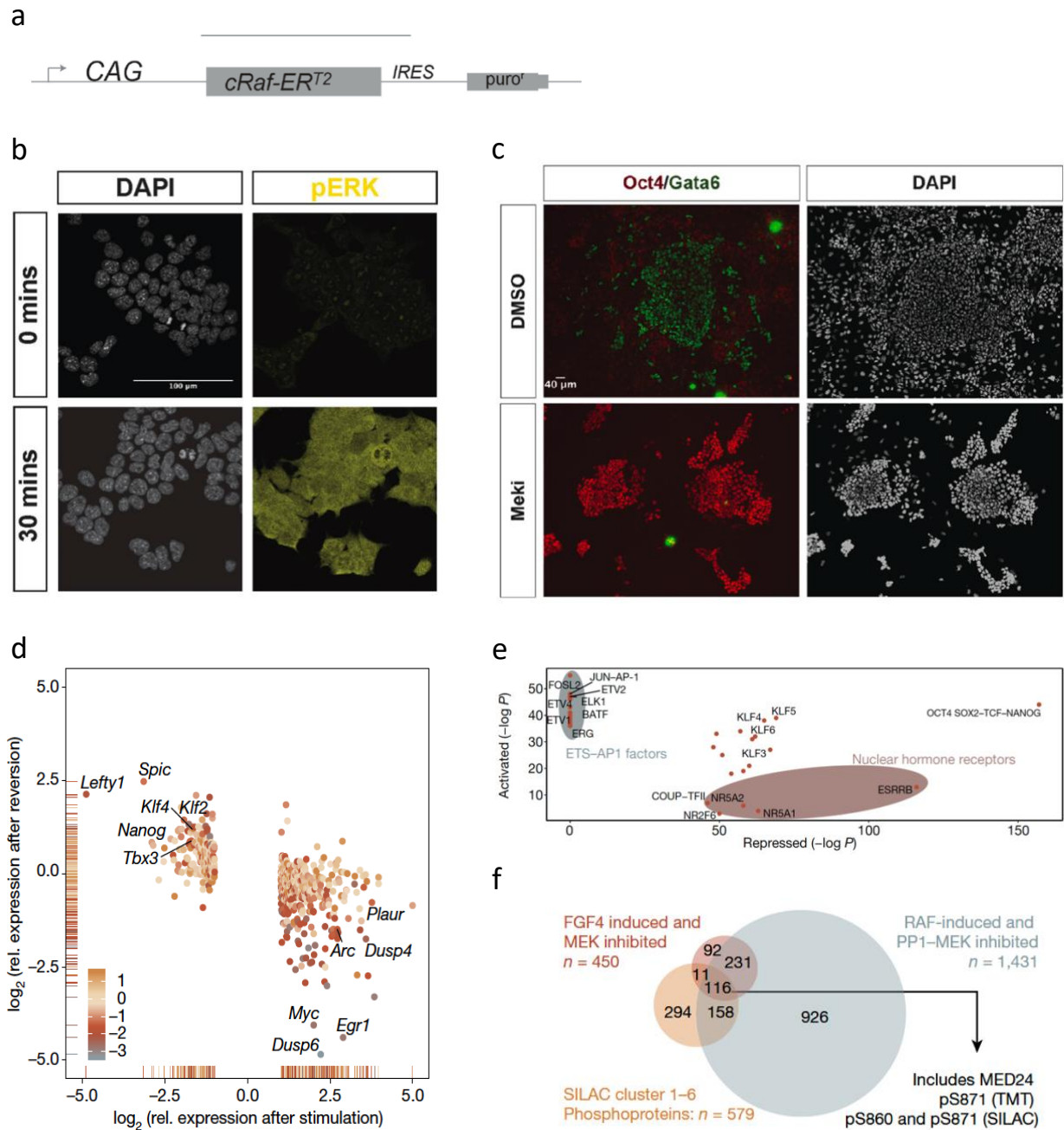
Figure 1.8 - ERK Induction system reveals Med24 as potential regulator – a) A schematic representation of the inducible cRAF-ERT2 construct. b) Immunostaining of ERK phosphorylation after 30 minutes of 4OHT treatment in cRAF-ERT2-expressing ES cells showing homogenous induction of pERK. c) Immunostaining of Oct4 and Gata6 TFs following 5 days of PrE differentiation. Inhibition of ERK pathway (MEKi) shows the presence of Oct4 positive and Gata6 negative cell population contrasting to the majority of Gata6 positive cells in DMSO treatment. d) Scatter plot comparing repressed and activated nascent transcripts after 8 hours stimulation or reversion with MEKi. ERK stimulation leads to the upregulation of PrE differentiation genes (e.g Egr1 and Dusp6) and downregulation of pluripotency genes (e.g Nanog and Klf4). e) Motif analysis on ERK-regulated enhancers. Activated enhancers are enrich in motifs of ETA and AP-1 factor while nuclear hormone receptors are enriched repressed enhancers. f) Venn Diagram showing MED24 as a candidate in three different phosphoproteomes. Figure adapted from Hamilton W. *et al.,* 2014; 2019
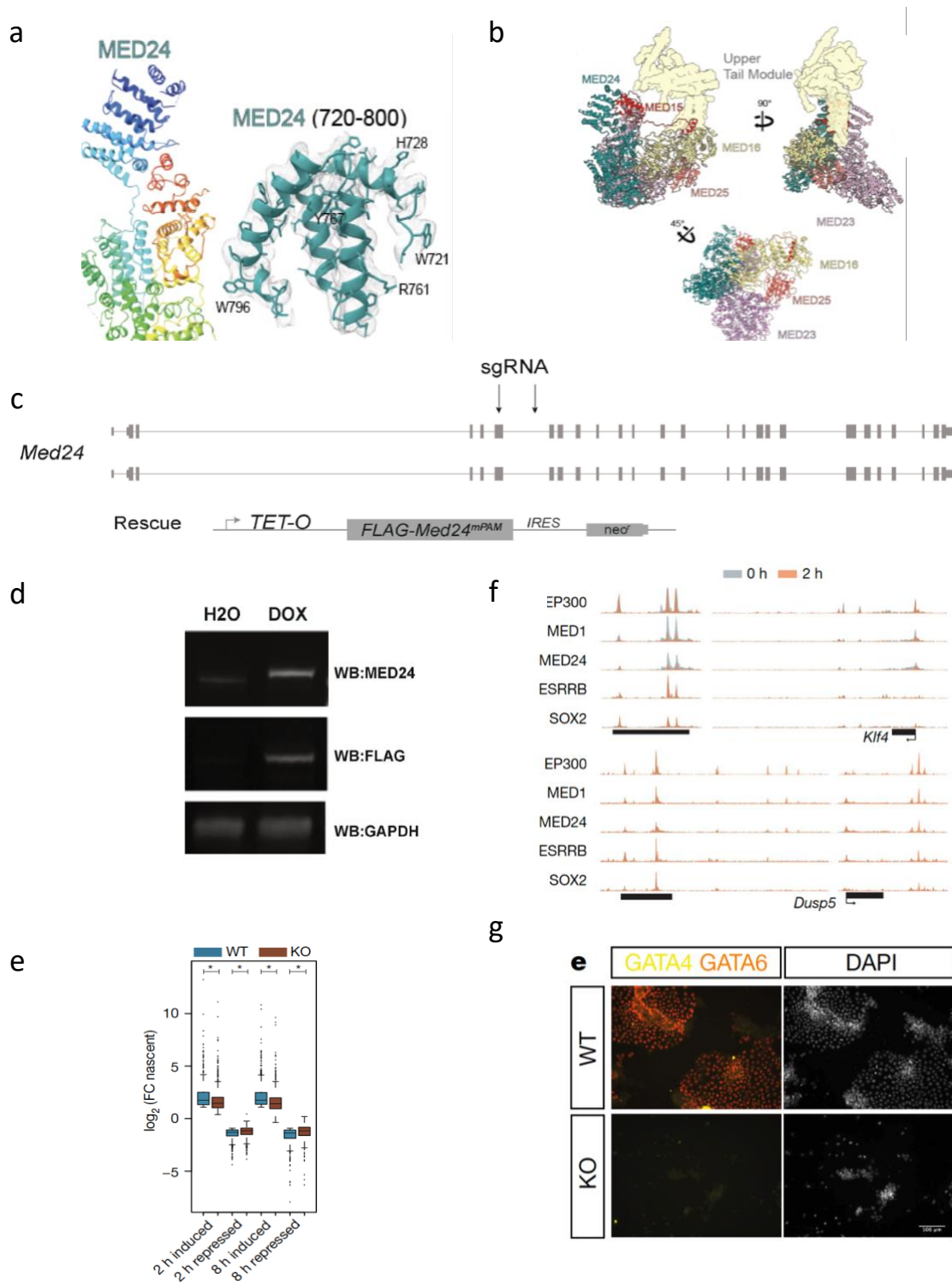
Figure 1.9 - MED24 subunit is necessary for a robust ERK response and PrE differentiation – a) Model of Med24 subunit (left). N-terminal depicted in blue while C-terminus is represented in red. Residues from the model and Cryo-EM structure are marked (right). b) Structure and organization of the different subunits that composed the lower tail module. Upper tail module is showed in yellow. c) A schematic overview of the strategy used to make the MED24 conditional KO ES line (MED24 cKO). The arrows represent where single guide RNAs are targeted to induce knockout of endogenous alleles. d) The expression of Flag-MED24 is maintained by the addition of doxycycline to the culture medium. Withdrawal of doxycycline creates Knockout MED24 ESCs. e) MED24 subunit is necessary for a robust transcriptional response to ERK in ESCs. f) ChIP-seq tracks showing the binding of the respective factors to the *Klf4* or *Dusp5* super-enhancers. Upon ERK activation, the binding of MED24 and MED1 shifts from repressed (*Klf4*) to activated (*Dusp5*) corresponding enhancers. g) Immunostaing of MED24 cKO cells for GATA6 and GATA4 following five days of PrE differentiation with DOX (WT) or without DOX (KO). Figure adapted from Hamilton W. *et al.*, 2019 and Zhao *et al.*, 2020.

1.6 Aims of the project

Our laboratory recently showed that MED24 has an essential role in PrE differentiation, and it is necessary for a robust ERK-transcriptional response. Thus, the principal aim of this thesis is to investigate the role of MED24 subunit in the regulation of ESCs transcriptional response. Additionally, I characterize a protein degradation system (mAID system) that is then used to deplete MED24 subunit. I hypothesize that regulation of MED24 plays an important role in ESC transcriptional response. I start by addressing this hypothesis by investigating which domains of MED24 are required for its functionality. Then, I aim to address this hypothesis by acutely deplete MED24 in ESCs and investigate how it affects pluripotency maintenance and differential potential. Further, to address this hypothesis, I specifically ask the following questions:

- Can truncated forms of MED24 protein be stably produced?
- What are the possible domains that are required for MED24 functionality and regulation?
- Determine if truncated versions of MED24 can rescue transcriptional response.
- Is auxin degron system suitable to study acute protein depletion?
- Does MED24 acute depletion affects self-renewal of ESCs?
- How MEKi affect cells response to MED24 depletion?
- Is PrE differentiation affected by MED24 depletion?

2. Materials and Methods

2.1 ES cell culture

2.1.1 Cell lines

In this study all cell lines used were derived from E14JU (from hereafter referred as 'E14'). The E14 cell line was derived from 129/Ola mice at the University of Edinburgh.

ERB cell line

Generated in the laboratory by William Hamilton. A copy of the truncated cRAF (cRAF$^{\Delta 26\text{-}303}$-ER^{T2}) (Fig. 1.8a) together with a sequence of the rtTA (reverse tetracycline-controlled transactivator) were randomly integrated in E14 cells.

TBR cell line

Generated in the laboratory by Christina Schuh using CRISPR/Cas9 mediated homologous recombination. A copy of the truncated cRAF (cRAF$^{\Delta 26\text{-}303}$- ER^{T2}) (Fig. 1.8a) was engineered into the *Tigre* locus.

MED24-mAID cell line

Generated in the laboratory by Rita Monteiro using CRISPR/Cas9 mediated homologous recombination. The mAID-GFP sequence was targeted to the *Med24* alleles to produce a homozygous fusion protein. Furthermore, the OSTIR sequence together with the ARF16 sequence was engineered together with the sequence of the truncated cRAF (cRAF$^{\Delta 26\text{-}303}$- ER^{T2})(Fig. 1.8a) and inserted into the *Tigre* locus. Three clones were generated: C2B, C7C and F4E.

F4ER cell line

Generated in the laboratory by Maria Clérigo. A sequence of the rtTA (reverse tetracycline-controlled transactivator) were randomly integrated in MED24-mAID cell line (F4E clone).

2.1.2 Cell culture

Table 2.1 – Cell culture medium composition

Cell medium	Components (Catalog #)
Serum/LIF	GMEM (Sigma-Aldrich; G5154) 10 % Fetal Bovine Serum (FBS, Gibco; F7524) 1x MEM non-essential amino acids (Gibco; 11140035) 10ng/ml leukaemia inhibitory factor (LIF) (made in house) 2 mM L-glutamine (Gibco, 25030024) 1 mM sodium pyruvate (Gibco; 11360070) 100 µM 2-mercaptoethanol (Sigma-Aldrich, M6250)
N2B27	1:1 Neurobasal and DMEM/F12 without glutamine (Gibco, 21103049 and 21331020 respectively) 1x B27 (Gibco, 17504044) 1x N2 (made in house) 2 mM L-glutamine (Gibco, 25030081) 100 µM 2-mercaptoethanol (Sigma-Aldrich, M6250)
2i/LIF	N2B27 supplemented with: 3 µM CHIR99021 (Axon, 1386) 10 ng/ml LIF (made in house) 1 µM PD0325901 (Sigma-Aldrich, PZ0162)
PrE basal medium	RPMI 1640 with GlutaMAX (Gibco, 61870044) 1x B27 minus insulin (Gibco, A1895601)
RACL	PrE base media supplemented with: 10 ng/ml LIF (made in house) 3 µM CHIR99021 (Axon, 1386) 100 ng/ml Activin A (Peprotech, 120-14E)

Table 2.2 Diverse reagents using during cell culture

Name	Working concentration (Catalog #)
Accutase	1x (Sigma-Aldrich; A6964)
Doxycycline	1 µg/ml (Sigma-Aldrich; D9891)
Gelatin	0.1% gelatin in sterile water (Sigma-Aldrich;
Geneticin (G418)	100 µg/ml (Sigma-Aldrich; 10131035)
Hygromycin B	100 µg/ml (Roche; 10843555001)
IAA	100 µM (Sigma-Aldrich; I5148)
PBS	1x (Sigma-Aldrich; D8537)
Penicillin-Streptomycin	5000 U/mL (Gibco, 15070063)
Puromycin	1 µg/ml (Sigma-Aldrich; P962)
Trypsin	1x (Made in house)

2.1.3 Routine ESC culture

All cell lines were cultured on gelatin (Table 2.2) coated culture flasks or plates (Corning) containing Serum/LIF medium (see Table 2.1). The cells were grown at 37 ºC, 5 % CO_2 with 90 % humidity and

passaged every two days or upon reaching approximately 80 % confluency. Cells were washed with phosphate buffered saline (PBS, Table 2.2) to remove cell debris and old medium, then accutase (Table 2.2) was added for 2-5 minutes at 37 ºC. Cells were collected by adding media. Cells were centrifuged at 4000 rpm for 2 minutes, the accutase was removed and the pellet was resuspended in fresh medium. A portion of the cells (1/8 or 1/10) were re-plated into fresh medium. At certain time points, ESCs were cultured in 2i/LIF medium (see Table 2.1). Cells were kept in this medium for at least two passages but no more than three passages before being used for experiments.

2.1.4 Primitive Endoderm differentiation

To differentiate ESCs into primitive endoderm, 3.5×10^5 cells previously cultured in 2i/LIF conditions were seeded in 6-well plates in PrE basal Medium (see Table 2.1). The next day, the medium was changed to RACL (see Table 2.1). For MED24 was degradation, 100 μM IAA (Table 2.2) was added to the Basal Medium and RACL. Cells were allowed to differentiate for seven days, and the medium was changed every two days. At day 7 cells were collected for flow cytometry analysis. For immunostaining, 3.5×10^4 cells were seeded in Ibidi 8-well slides with IbiTreat (Ramcon) and the same culture conditions were applied.

2.1.5 Neural differentiation

Similar to PrE differentiation, cells were previously cultured in 2i/LIF and 1.0×10^4 cells were seeded in Ibidi 8-well plates with IbiTreat (Ramcon) in N2B27 (see Table 2.1) supplemented with LIF. The next day, the medium was changed to N2B27 (see Table 2.1), and it was replaced every two days or whenever the medium became acidic. At day 7, cells were fixed and prepared to immunostaining analysis.

2.1.6 Freezing

For long-term storage cells were kept at -80 ºC, ES cells were dissociated using accutase as described in 2.1.3 and placed into cryovials (Thermo Scientific) with GMEM supplemented with 10 % DMSO (Sigma-Aldrich) and 40 % FBS. Each cryovials contained approximal 500 k cells. Cells were thawed by placing the cryovials in the incubator and then transferring the cells into 5 ml pre-warmed medium to wash away DMSO. Then, cells were centrifuged and plated into fresh medium.

2.2 Vector design and implementation

Table 2.3 - Buffers and reagents used in Molecular Biology

Buffers and reagents for molecular biolo	Components (Catalog #)
5x First strand buffer	Life Technologies
Cutsmart Buffer	NEB (B7204)
dNTPs	ThermoFisher (18427013)
DTT	Life Technologies (D1532)
Formaldehyde	Fisher Scientific (PI-28906)
Generuler 1kb plus DNA ladder	ThermoFisher (SM1332)
KaryoMAX	ThermoFisher (15212012)
Lipofectamine 2000 transfection reagen	ThermoFisher (11668030)
Opti-MEM™ Reduced Serum Medium	ThermoFisher (31985062)
Phusion MM High Fidelity	NEB (M0531)
Random hexamer primers	Invitrogen (N8080127)
RNase OUT	ThermoFisher (10777019)
SuperscriptIII reverse transcriptase	ThermoFisher (18080085)
SYBR Safe DNA gel stain	Invitrogen (S33102)
TAE	40 mM Tris, 1 mM EDTA, pH 8.0 with glacial acetic acid (Sigma-Aldrich)
Ultrapure Agarose powder	LifeTecnhologies (16500500)

2.2.1 Polymerase Chain Reaction (PCR)

Primers for each truncation were design (Table 2.5) and the sequence of interest was amplified from pTET_MED24_PAM_IN vector using Polymerase Chain Reaction (PCR). 500 ηg of DNA was mixed with 2x Phusion MM High Fidelity (Table 2.3), 10 μM of forward primer, 10 μM of reverse primer and nuclease free water was added until the volume of 25 μl was reached. The PCR program (T100™ Thermal Cycler) used was as follows:

1. 98 ºC – 3 minutes
2. 98 ºC – 25 seconds
3. 58 ºC/63 ºC – 25 seconds
4. 72 ºC – 45 seconds
5. Go to step 2 24 times (or 34 times)
6. 72 ºC – 5 minutes

2.2.2 Agarose gel electrophoresis

For DNA analysis, 0.8 % or 1 % agarose gels were used throughout this study. They were made by dissolving Ultrapure Agarose powder into TAE buffer, SYBR Safe DNA gel stain was added to the solution (see Table 2.3). DNA was mixed with 5x Orange loading buffer and loaded onto the gel. To determine the fragment size, 1 kb plus DNA ladder (Table 2.3) was loaded onto every gel. DNA was visualized with UV light using a Gel Doc XR+ (BioRad).

2.2.3 DNA extraction

The DNA fragments amplified for each truncation were isolated from the agarose gel and extracted using ZymoClean DNA recovery kit (Zymo Research) according to the manufacturer's instructions. The DNA was extracted from the agarose gel by adding ADB buffer, it was cleaned on-column with washing buffer, and finally eluted in Nuclease Free Water. The concentration and purity were measured using the Nanadrop2000 spectrometer (ThermoFisher).

2.2.4 Restriction cloning

The fragments amplified and pTET_MED24_PAM_IN vector were double digested for 4 hours at 37 ºC. The volume of fragments obtained from 2.2.2 were mixed with 10xCutsmart Buffer (NEB), restriction enzymes 10 U SpeI (NEB) and 10 U AgeI (NEB) and a volume of nuclease free water to total 20 µl. For the vector digestion, 5 µg of DNA was added to 10x CutSmart (NEB), 25 U SpeI (NEB), 25 U AgeI (NEB) and nuclease free water. The ligation reaction mix was set up using NEBioCalculator (https://nebiocalculator.neb.com/) and the sequence of interest was ligated to the backbone vector using T4 DNA ligase enzyme (NEB). The mix was incubated at RT for at least

30 minutes. Competent bacteria were transformed with the ligation mix and colonies formation were checked the next day. Colonies from successful ligations were picked and grown overnight, followed by DNA extracting using QIAprep spin miniprep kit (QIAGEN; #27104).

In order to screen, F4E cells that had insert the rTta sequence a plasmid carrying the mCherry sequence under the TRE system was creating from the pTET_MED24_PAM_IN vector. The result can see seen on Figure 2.1.

After obtaining the different MED24 truncations plasmids, unexpectedly, between the ampicillin resistance and the TetO promoter a small fragment that aligned with the enhancer of *Tfcp2l1* gene was found. In order to remove this fragment and to change the selection cassette from Neomycin to Hygromycin (Table 2.3), three different digestion were set: to remove the *Tfcp2l1* enhancer fragment, all the plasmids were digested with 10 U SphI (NEB) and 10 U SalI (NEB); the Hygromycin sequence was obtaining by digesting a plasmid made in the house (WBH99) with 10 U KpnI (NEB) and 10 U SphI (NEB); the MED24 truncations sequence were obtained by digested with 10 U XhoI (NEB) and 10 U KpnI (NEB). The DNA was extracted from the agarose gel and a ligation reaction was performed as described above. The constructs obtained were sequenced.

Table 2.4 - Different MED24 truncated forms

MED24 truncations	Description	Fragment Size (bp)	Protein Weigh (kDa)
#1	N-terminus removed	2585	95.69
#2	C-terminus removed	2580	96.25
#3	N-terminus and C-terminus removed	2197	83.29
#4	C-terminus partially removed but with the two ERK-phosphosites	2616	97.44
#5	C-terminus partially removed but with only one ERK-phosphosite	2586	96.42
#6	C-terminus, the two ERK-phosphosites and motif 6 removed	2470	91.99

Table 2.5 - Primer pairs used to amplify the different MED24 truncated forms

| MED24 truncations	Forward Primer	Reverse Primer
#1	ACTAGTCTCCACTGGCTGCTT	ACCGGTGAGTGCAGCAATGGC
#2	GGAACTAGTTATGAAGGTGGTGAACCTG	CCGCACCGGTTACAGGAGTCGCAT
#3	ACTAGTCTCCACTGGCTGCTT	AGCACCGGTCAGGAGTCGGATCAG
#4	GGAACTAGTTATGAAGGTGGTGAACCTG	CGCACCGGTTAACTTGACAGGATGT
#5	GGAACTAGTTATGAAGGTGGTGAACCTG	CGCACCGGTTAGGAGCTCAGGAGT
#6	ACTAGTTATGAAGGTGGTGAACCTG	ACCGGTCTTGTGGGAAGAGTAAG
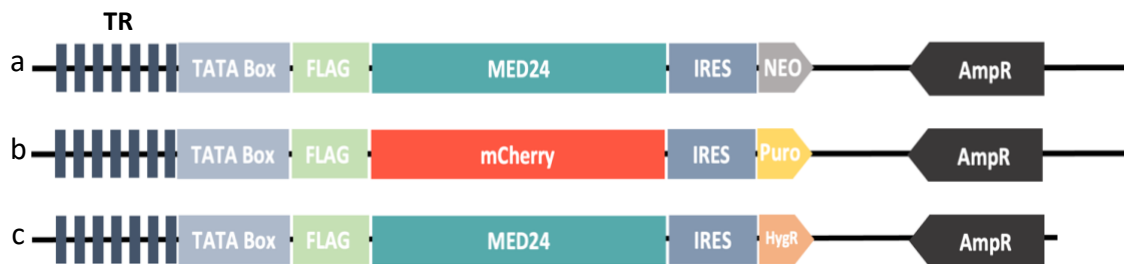
Figure 2.1 - Simplified schematic representation of used plasmids – a) Plasmid containing MED24 wild-type (WT) sequence or the different MED24 truncation sequences under the control of TRE system. All sequences were tagged with a FLAG tag. The plasmid had a neomycin (NEO) selection cassette. b) mCherry was cloned into the same vector as the previous one, but a puromycin selection cassette was used. c) MED24 WT and truncations were cloned into a vector without the small fragment belonged to the enhancer of Tfcp2l1 gene and with a Hygromycin (HygR) resistance. IRES - Internal ribosome entry site; AmpR – Ampicillin resistance

2.3 Generation of cell lines and validation

2.3.1 Transfection and Electroporation

ES cells were plated in a 6-well plate with a confluency of 500 k cells per well and transfected with the different vectors. Two different mixes were made, one with 3 μg of DNA and 250 μL of the Opti-MEM medium, and the other with 3 μL Lipofectamine 2000 transfection reagent with 250 μL Opti-MEM medium (Table 2.3). Before the transfection, the mixes were combined and incubated for 30 minutes at room temperature. The mix was added to each well and the cells were incubated at 37ºC and 5 % of CO_2. After 24 hours, the medium was supplemented with the appropriated selection drug. When performed in 24-well plate, the reactions and cell number were scale down, 0,8 μg of DNA was mixed with 50 μL of the Opti-MEM medium, and 2 μL Lipofectamine 2000 transfection reagent with 50 μL Opti-MEM medium. This solution was added to 125 k cells.

For electroporation, 5×10^6 cells were transfected with the 25 ug of linear DNA. Cells free of medium were combined with the DNA into a Gene Pulser electroporation cuvette (0.4cm gap, BioRad). The cuvette containing the mix was inserted in the electroporator (250 V, 500 µF). The mix was carefully transferred to complete medium and 1×10^6 cells were plated per 10cm dish. After 24 hours, fresh medium was added, and selection started with the appropriated selection drug. The medium was changed every other day. After 10 days of selection, clones were picked using trypsin into a 48-well. The next day, the medium was changed to fresh medium and the surviving clones were expanded for further analysis.

2.3.2 Western Blotting

For protein detection, cells were collected and resuspended in laemelli lysis buffer (4 % SDS, 20 % glycerol and 120 mM Tris pH 7.4) and sonicated (40 % amplitude; 10 seconds; Brandelin Ultrasonic homogenizer mini20). 70 ug – 150 ug of protein sample together with bromophenol blue solution and DTT were loaded into a 4 %-10 % SDS-PAGE gel (Life Technologies) and the proteins transferred to nitrocellulose membrane. The membrane was blocked for 1-2 hours using a 5 % skim milk solution. The membrane was washed several times with TBST (see table 2.3) before incubating with the respective primary antibodies diluted in a 5 % BSA solution. Primary antibodies were detected using secondary antibodies that are conjugated with fluorophores (see Table 2.6). The secondary antibodies diluted in a 5 % skim milk solution were incubated for 2 hours at RT. The membrane was visualised using Chemidoc MP (BioRad).

2.3.3 Immunofluorescence Staining

Cells were fixed with 4 % formaldehyde for 10 minutes at 37 ºC, blocked in PBS, 5 % donkey serum and 0.3 % Triton, for 2 hours at RT.

Primary antibodies were incubated in PBS with 1 % BSA and 0.3 % Triton overnight at 4 ºC. On the next day, plates were washed three times with PBS. Secondary antibodies were incubated for 1-2 hours at RT in the same solution as the primary antibody. The primary antibodies used can be seen in Table 2.6.

Plates were washed three times with PBS and nuclei were stained with DAPI (1:10000 in PBS) for 5 minutes. All images were acquired using confocal Zeiss LSM 780.

2.3.4 Flow cytometry

For flow cytometry analysis, cells were detached with accutase, incubated with the respective antibody (see table 2.6) in PBS with 10 % FBS for 30 minutes at 4 ºC, washed three times. Then cells were resuspended in PBS with 10 % FBS and 1 μg/ml DAPI. Cells were analyzed using LSRFortessa (BD Bioscience) and dead cells were excluded based on DAPI inclusion.

Table 2.6 Antibodies used for Western Blot (WB), Immunofluorescence staining (IF) and Flow Cytometry (FACS). Respective dilutions and suppliers.

Antibody	Dilution (WB)	Dilution (IF)	Dilution (FACS)	Supplier (Catalog #)
ERK total	1:1000	-	-	Cell Signalling (4696)
FLAG	1:500-1:1000	1:200	-	Sigma-Aldrich (F3165)
GFP	1:1000	-	-	Aves labs (GFP-1020)
H3	1:1000-1:5000	-	-	Abcam (10799)
MED24	1:500-1:1000	-	-	Bethyl Laboratories (A301-472)
NANOG	1:1000	1:200	-	eBioscience (14-5761)
OCT4	-	1:200	-	SantaCruz Biotechnologies (sc5279)
pERK	1:1000	-	-	Cell Signalling (4376)
pERK XP	-	1:200	-	Cell Signalling (4370)
PECAM	-	-	1:200	eBioscience (25-0311-82)
PDGFα	-	-	1:200	eBioscience (17-1401-81)
Donkey anti-mouse Alexa Fluor 647	1:2000	1:2000	-	ThermoFisher (A-31571)
Donkey anti-rabbit Alexa Fluor 647	1:2000	1:2000	-	ThermoFisher (A-31573)
Goat anti-chicken Alexa Fluor 647	1:2000		-	ThermoFisher (A-21449)
Donkey anti-rat Alexa Fluor 488	1:2000	1:2000	-	ThermoFisher (A-21208)

2.3.5 RNA extraction

Cells were collected in RTL buffer supplemented with β-mercaptethanol and RNA was extracted using RNeasy Mini Kit (Qiagen, 74104). To remove genomic DNA, on-column DNase digestion was performed using RNase-Free DNase Set (Qiagen, 79254), following manufacture's instruction. RNA was eluted in nuclease free water and concentration determined using a NanoDrop2000 spectofrometer (Thermo Scientific).

2.3.6 cDNA synthesis

cDNA was synthesized using SuperscriptIII reverse transcriptase according to manufacturer's guidelines. 1 µg RNA was incubated at 65 °C for 5 minutes with 1 µg random hexamer primers and 5 mM dNTPs in a volume of 13 µl. Samples were then incubated at 4 ºC for 5 minutes before adding 5x First strand buffer, 0.1 M DTT, 40,000 U/ml RNaseOUT and SuperscriptIII enzyme (see Table 2.3). After the synthesis, cDNA was diluted 1:10 in nuclease free water.

2.3.7 Real- Time quantitative PCR

Real-time quantitative PCR (RT-qPCR) was performed by using Universal Probe Library (UPL) system (Roche), using LightCycler 480 Instrument II (Roche Life Science). Primers were designed using the UPL assay design center (https://www.lifescience.roche.com). Probes were chosen using the Roche Probe Library. The level of probe fluorescence corresponds to the level of product generated in the PCR reaction. cDNA corresponding to 10 ng of RNA was combined with 500 nM of probes, 450 nM of primers and 2 ul of nuclease-free water. First, a mix containing the primers, probe and water was pipetted into a 384-well plate, centrifuged for 1 minute, lastly the cDNA was added in the opposite side of the well to prevent cross-contamination. Standard curves for each gene assay were generated by making serial dilutions of cDNA containing the gene of interest. The samples were normalized for the amount of cDNA in each sample by calculating the values of each gene relative to the housekeeping genes. In each run (?), two housekeeping genes were used to normalise the data. Details of the primers and probes used in this study can be found in Table 2.7.

Table 2.7 UPL primer pairs and respective probe

Gene	Forward Primer	Reverse Primer	Probe
Dusp6	CTGGTGGAGAGTCGGTCCT	CGGCCTGGAACTTACTGAAG	66
Egr1	CCCTATGAGCACCTGACCAC	TCGTTTGGCTGGGATAACTC	22
Gapdh	GGGTTCCTATAAATACGGACT	CCATTTTGTCTACGGGACGA	52
Klf4	GAGTTCCTCACGCCAACG	CGGGAAGGGAGAAGACACT	62
Nanog	CCTCCAGCAGATGCAAGAA	GCTTGCACTTCATCCTTTGG	25
Pbgd	AAAGTTCCCCAACCTGGA	CCAGGACAATGGCACTGAAT	42
Sdha	TGTTCAGTTCCACCCCACA	TCTCCACGACACCCTTCTG	71

Spry4	GTGGAGCGATGCTTGTGAC	CACCAAGGGACAGGCTTCTA	17
Tbx3	TTGCAAAGGGTTTTCGAGAC	TGCAGTGTGAGCTGCTTTCT	51

2.3.8 Karyotyping

A quick method to check if cell lines have the right karyotype is chromosome counting. To do this, ESCs were expanded until they reached 70-80 % of confluency in a 25 cm^2 plate. Cells were washed with PBS and to arrest cells in metaphase they were incubated in cell medium containing 0.1 ug/ml KaryoMAX for 2 hours at 37 ºC. After, cells were washed with PBS and trypsin was added for 30 seconds. The cells were collected by adding medium and centrifuged at 1300 rpm for 5 minutes. The supernatant was removed, 2.5 ml of 0.56 % (w/v) KCl was added and incubated for exactly 6 minutes at RT. After slowly adding 1ml fixative solution (75 % (v/v) methanol and 25 % (v/v) acetic acid), the cells were centrifuged at 1300 rpm for 3 min. The supernatant was replaced with 1 ml fresh fixative and incubated at 4 °C for ON. After washing the cells another time with fresh fixative without disturbing the pellet, the supernatant was carefully removed, and the cells were resuspended in 200 μl fixative. Single drops were tipped down onto polylysine coated glass slides. After air drying, chromosomes were stained with 10 % Giemsa solution for 30 min. The final chromosome counting was done with a regular light microscope. The chromosomes of at least 10 randomly chosen cells were counted for each cell line.

2.3.9 Alkaline Phosphatase Staining

The undifferentiated state of ES cells can be characterized by a high level of expression of Alkaline Phosphatase (AP). Cells were cultured in the different media for at least one passage and then 600 cells were plate into a gelatin-coated 6-well plate. Cells were allowed to expand for 5 days and then fixed and stained using an Alkaline Phosphatase Detection Kit (Sigma-Aldrich). Cells were washed with PBS and fixed for approximately 1 minute. After being washed with MiliQ water, cells were stained for 1 hour and the plates were allowed to dry for one day before colony counting.

2.3.10 Alphafold

Alphafold was used to model a 3D structure of MED24 WT. Appendix 1

3. Results

3.1 C-terminal NR region of MED24 is necessary for nuclear localization of the protein

Before 2020, little was known about MED24 structure and function or its specific interactions with different Mediator subunits. Therefore, one of the aims for this project was to investigate which specific part of the protein is essential for ESC transcriptional regulation and PrE differentiation. To characterize the protein regions and to understand how the subunit responds to ERK activation, different MED24 truncations were designed. The N-terminus and C-terminus represent large regulatory regions carrying important signals and residues. These regions can be targets of post-translational modifications and used for binding to other proteins, ultimately affecting protein structure and function (Fig. 3.1.1a,b). These important domains were removed, with ΔMED24_1 missing the N-terminus, ΔMED24_2 missing C-terminus and ΔMED24_3 missing both N-terminus and C-terminus (Fig. 3.1.1c). It is also known that MED24 protein has six different LxxLL motifs along with the two ERK-phosphosites, S860 and S870 (Fig.3.1.1b). Of note, Hamilton et al. showed that while the two phosphosites of MED24 were extensively phosphorylated by ERK, corresponding phosphomutants had only a minimal effect on ERK-regulated transcription. Therefore, in order to eliminate the possibility of alternative proximal phosphosites, differing portions of the C-terminus were deleted: a section containing the two phosphosites together with the remaining C-terminus (ΔMED24_4), a section missing only one phosphosite and the C-terminus (ΔMED24_5) and a section missing the two phosphosites and the C-terminal LxxLL motif (ΔMED24_6) (Fig. 3.1.1c).

All truncations were cloned using a vector containing a FLAG-tagged cDNA sequence of *Med24* under a TRE transactivator system (rtTA-dependent) and a Neomycin selection cassette (Fig. 3.1.2a). Primers for each truncation were designed and a restriction cloning strategy was used to obtain the truncation vectors (see section 2.2). All truncations vectors, along with MED24 WT vector, were inserted into an ERB cell line that constitutively expresses rtTA (Fig. 3.1.2a). Western Blot analysis confirmed that the truncations were being expressed with the expected size, except for ΔMED24_2 which was smaller than expected and ΔMED24_3 which showed no expression (Fig. 3.1.2b). Since MED24 subunit is a nuclear protein; localization of the truncated subunits was assessed. Stable cell lines were generated for MED24 WT, ΔMED24_1, and ΔMED24_4-6. Immunostaining for the FLAG tag was used to detect MED24 and the truncations. Most of the

truncated proteins retained a nuclear localization, except for ΔMED24_6 which appears to have lost its nuclear localization signal (NLS) and is exclusively cytoplasmatic (Fig. 3.1.3).
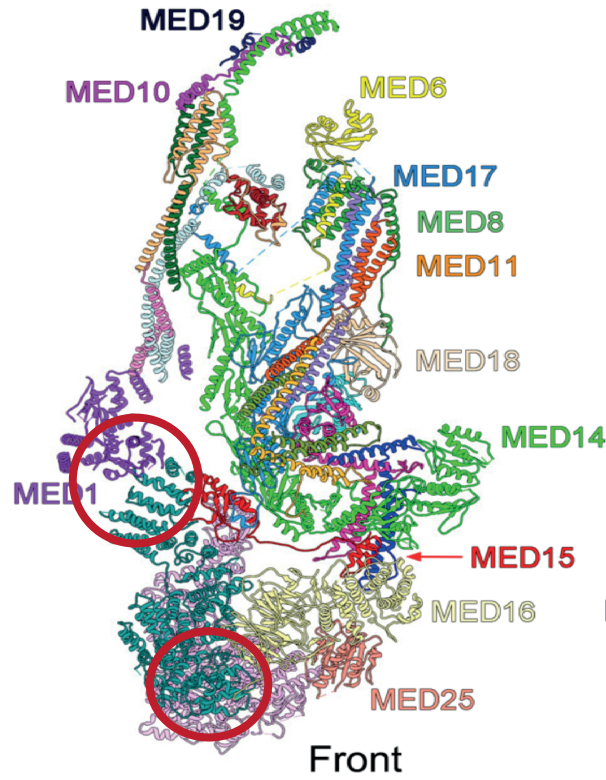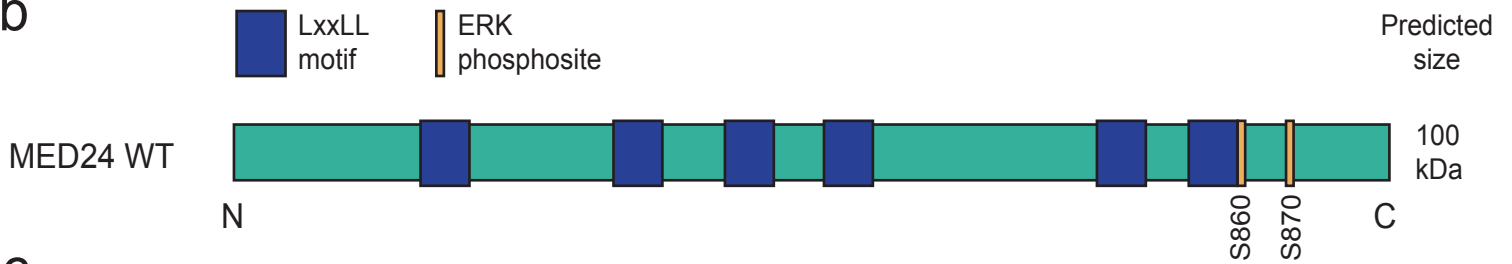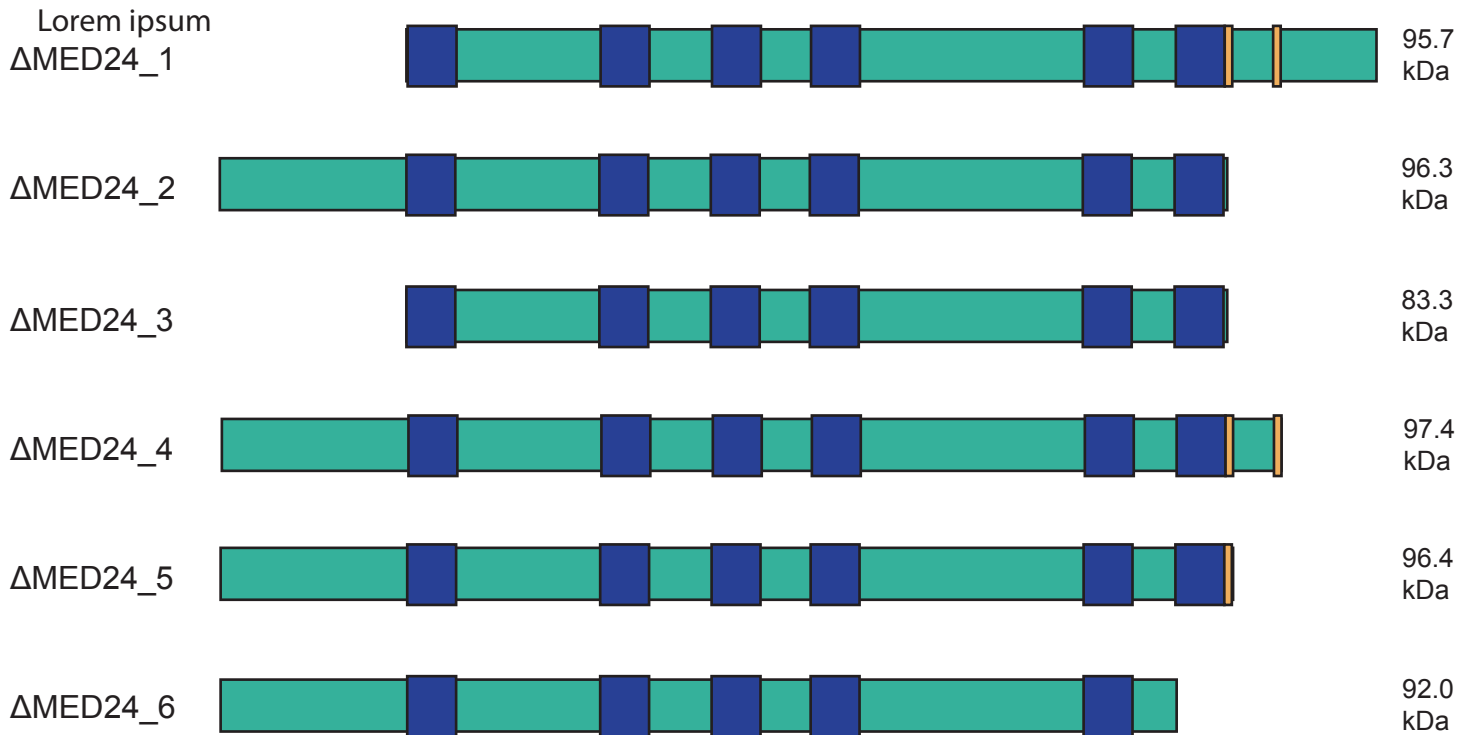
Figure 3.1.1 – Schematic representation of MED24 WT and truncations - a) Cryo-EM structure from Zhao H. et al., 2020, N-terminus and C-terminus are marked with red circles. b) MED24 WT sequence. The six LxxLL motifs are indicated by blue boxes.ERK phosphorylation sites are indicated by yellow vertical bars. c) The six diffrent truncations are represented along with the predicted size.
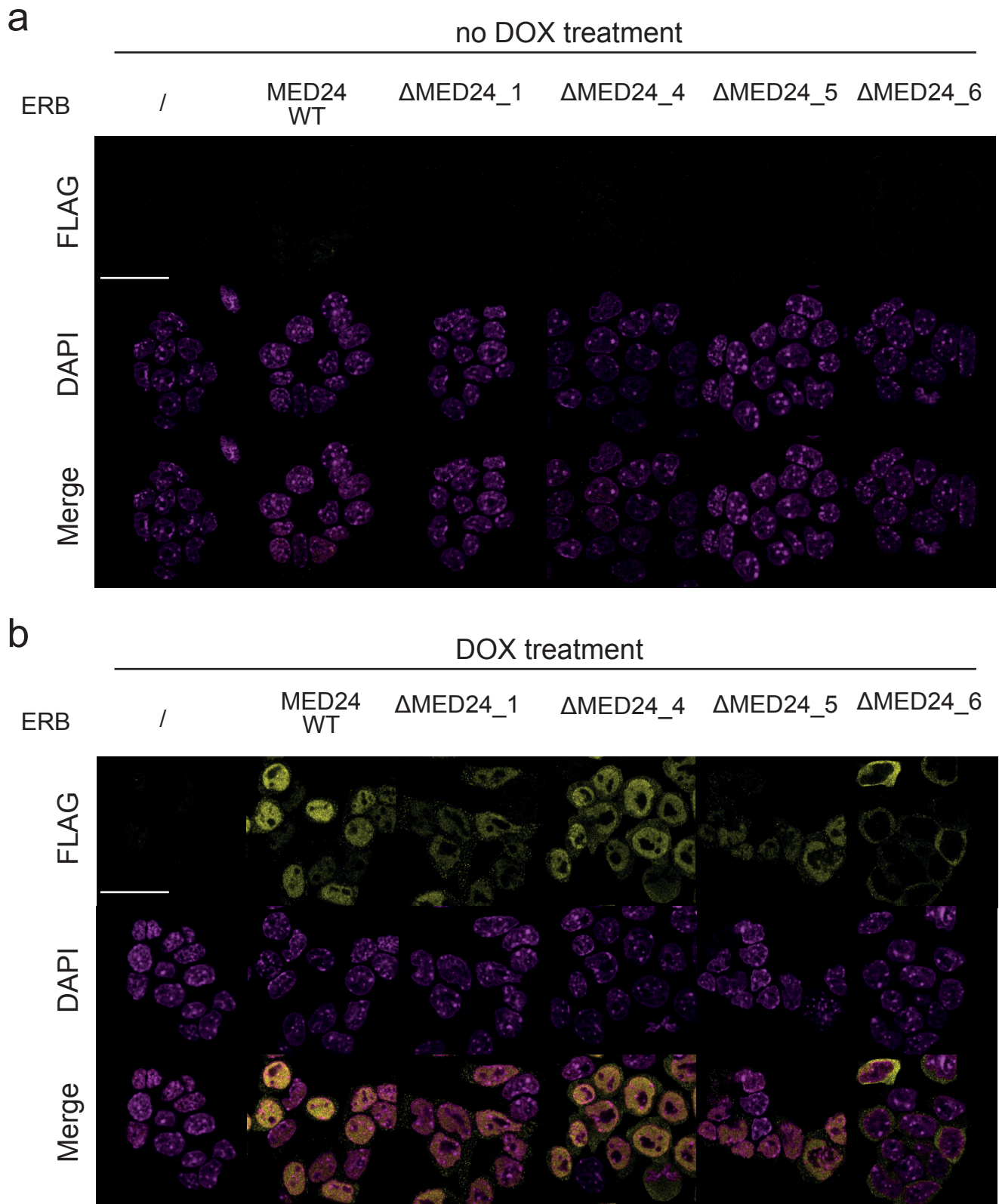
Figure 3.1.3 - C-terminal NR region of MED24 is necessary for nuclear localization of the protein - Immunostaing against FLAG tag in ERB cell lines containing the mentioned vector without dox a) and with 3 days of dox treatment. Scale bars equal to 100 μm.

3.2 The mAID system can be used to induce rapid depletion of MED24 in ESCs

To investigate the role of MED24 in ESCs, the Brickman Laboratory generated a conditional MED24 KO cell line(Hamilton et al., 2019). Unfortunately, this cell line had limitations, like total protein degradation taking up to 72 hours, making it impossible to induce MED24 depletion in a time-sensitive manner in order to study acute effects.

The addition of an auxin-inducible degron (AID) system, which destabilizes expression of the target protein, has been shown to be applicable for mESC (mAID) (Nishimura et al., 2009). Additionally, Mediator subunits have been successfully targeted with mAID system or similar protein degradation systems (El Khattabi et al., 2019; Jaeger et al., 2020). Therefore, this system for acute protein depletion was used for the generation of a homozygous MED24-GFP-mAID cell line (Fig. 3.2.1a) and an extensive characterization was performed. This line contains the Oryza sativa TIR1 (OsTIR1) sequence, the auxin-response transcription factor (ARF16) sequence and the cRAF$^{\Delta 26\text{-}303}$- ER^{T2} sequence all in one allele of *Tigre locus* (Fig. 3.2.1b). Consequently, auxin-binding to the OsTIR1 receptor targets the mAID-tagged MED24-GFP for degradation by the proteasome complex (Fig. 3.2.1c). ARF16 represents a second layer of regulation for this system after it was found that some proteins tagged with the mAID were degraded in the absence of auxin (IAA) (Sathyan et al., 2019; Yesbolatova et al., 2020). ARF16 interacts with the mAID tag in the absence of IAA, repressing the constitutive degradation of tagged proteins.

Rita Monteiro generated this cell line and based on flow cytometry analysis of the GFP signal from the fusion protein (MED24-GFP), three isogenic clones were chosen: C2B, C7C and F4E (data not shown). These clones were then used for characterization of this system. Since this line underwent extensive genetic modifications, I assessed if the cells maintained a normal karyotype and an ESCs phenotype under normal culture conditions (2iL). Chromosome counting was performed, and the karyotyping revealed that two of the three clones were chromosomally normal (n=40) (Fig. 3.1.2). The ESC phenotype (pluripotency state) was assessed using an Alkaline Phosphate (AP) assay. All three clones showed a large number of AP$^+$ colonies (Fig. 3.1.3).

Then, to determine the optimal IAA concentration for inducing rapid degradation of MED24-GFP, a titration of IAA treatment was performed (1 ρM, 1 ηM, 100 ηM, 1 μM, 100 μM, 250 μM and 500 μM) (Fig. 3.2.4a). After 24 hours of 100 μM IAA treatment the GFP signal was no longer detected. TBR, a cell line that does not respond to IAA treatment was used as a control (see 2.1.1). Next, a time-course to assess the duration of MED24-GFP degradation in this system was performed. After 30 minutes of IAA treatment, the MED24-GFP signal was reduced by half, while

after 3 hours the signal was barely above background in all three clones (Fig. 3.2.4b). Western Blot analysis confirms that 24 hours after 100 μM IAA treatment, the MED24-GFP signal was undetectable in all three clones (Fig. 3.2.4c). Unfortunately, unlike TBR cells, no signal was detected in the MED24-GFP-mAID line when the MED24 antibody was used for detection. However, a GFP antibody showed a clean and strong signal.

To probe the cellular response to ERK activation with or without MED24, cells were cultured with PD17 (FGFRi), preventing extracellular signals, and treated with IAA, for MED24 degradation, 48 hours prior to ERK induction (Fig. 3.2.5a, b). Then, cells were treated with 4OHT for 8 hours to induce ERK and RNA was collected for analysis (Fig. 3.2.5b). Hamilton *et al.* observed that ERK response was impaired with MED24 depletion, specifically the repression of pluripotency genes, *Nanog* and *Tbx3* was attenuated. However, depletion of MED24 in ERK-induced MED24-GFP-mAID clones did not appear to affect the downregulation of pluripotency genes (Fig. 3.2.5c). This result was supported by Western Blot analysis of NANOG protein with also showed no difference (Fig. 3.2.6). Upon ERK induction for 8 hours, NANOG was still present in all MED24-GFP-mAID clones treated with IAA, while in the control cell line TBR, a visible degradation of the protein occured. The lower levels of pERK protein might explain the weak downregulation of pluripotency genes (Fig. 3.2.5c). Even though the total levels of pan-ERK were similar in all cell lines, the pERK levels observed upon 4OHT treatment appeared different between TBR and MED24-GFP-mAID cell lines (Fig. 3.2.6). This result was supported by the pERK immunostaining, with TBR cells showing a strong homogeneous signal while MED24-GFP-mAID clones showed a weak and heterogenous signal (Fig. 3.2.7). Even with heterogeneous induction, MED24 depleted cells showed a reduced upregulation of differentiation genes, *Egr1* and *Spry4*, as compared to controls, although this effect was less pronounced in MED24-GFP-mAID clones when compared to the MED24 KO cell line (Fig. 3.2.6d). Several attempts to insert an extra copy of cRAF$^{\Delta26\text{-}303}$- ER^{T2} -rtTa to have a more homogenous induction of pERK were made using different transfection techniques but to date none of them was successful.

Although, ERK induction in this cell line was not ideal, I showed that the degron lines could be used for functional studies, such as differentiation assays, since the cells maintained normal ESC qualities. I showed that F4E behaved most similarly to control cells and showed a better potential for ERK induction than the other two clones; for that reason further analysis proceed with just the F4E clone.
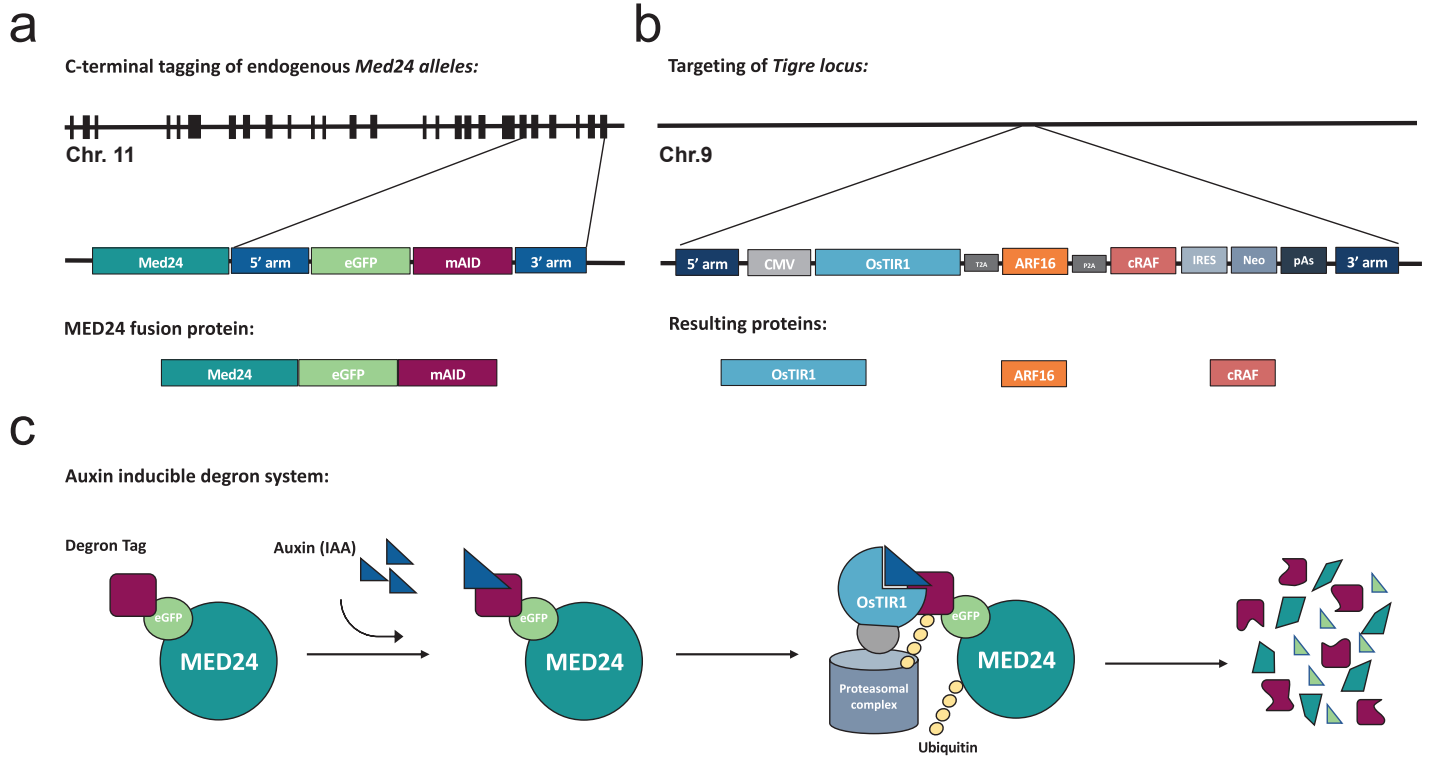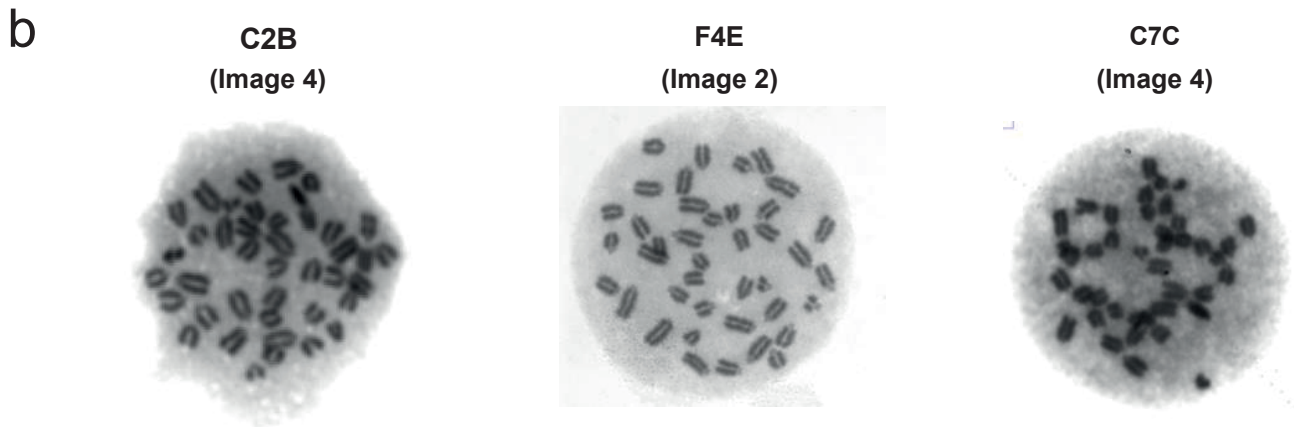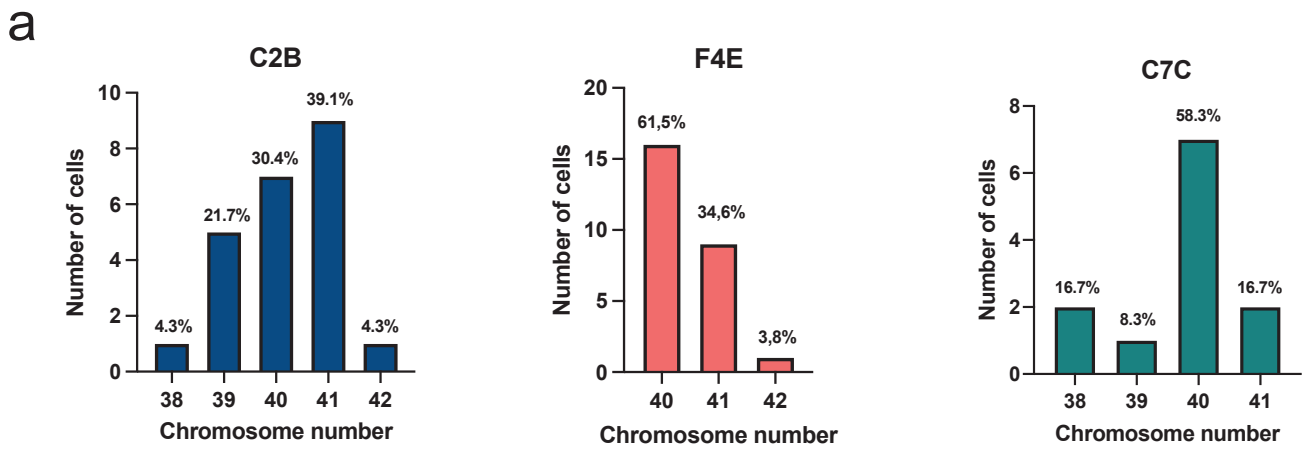
Figure 3.2.1 – Strategy of inducible MED24-GFP-mAID protein degradation – a) Schematic illustration of how MED24-GFP-mAID cell line was generated. Firstly, both alleles of Med24 gene were targeted using CRISPR/Cas9 mediated homologous recombination to generate a fusion protein. b) The same technique was used to target an allele of Tigre locus with the following sequences: OsTIR1 (auxin binding element), ARF16 (auxin-responsive transcription factor) and cRAF$_\Delta$26-303- ER^{T2}. c) Schematic representation of mAID system. The fusion protein exists in the cell as the WT in the absence of IAA. When IAA is added to the culture medium, it will bind the degron tag together with OsTIR1 protein, therefore this binding promotes the recruitment of proteasomal complex leading to the degradation of the target protein, in this case, degradation of MED24 protein.

a

C2B

Number of cells vs Chromosome number
- 38: 4.3%
- 39: 21.7%
- 40: 30.4%
- 41: 39.1%
- 42: 4.3%

F4E

Number of cells vs Chromosome number
- 40: 61,5%
- 41: 34,6%
- 42: 3,8%

C7C

Number of cells vs Chromosome number
- 38: 16.7%
- 39: 8.3%
- 40: 58.3%
- 41: 16.7%

b

C2B
(Image 4)

F4E
(Image 2)

C7C
(Image 4)

c

C2B	Chromosome number
Image 1	40
Image 2	39
Image 3	40
Image 4	41
Image 5	41
Image 6	42
Image 7	41
Image 8	41
Image 14	41
Image 16	40
Image 17	40
Image 18	39
Image 24	40
Image 27	41
Image 33	41
Image 37	39
Image 42	40
Image 44	40
Image 46	39
Image 50	38
Image 51	41
Image 53	39
Mode	**41**

F4E	Chromosomes number
Image 1	41
Image 2	40
Image 6	42
Image 7	41
Image 8	40
Image 9	41
Image 10	40
Image 11	41
Image 12	41
Image 13	40
Image 14	41
Image 15	40
Image 16	40
Image 21	40
Image 24	40
Image 25	40
Image 26	40
Image 27	41
Image 28	41
Image 29	40
Image 30	40
Image 33	40
Image 34	40
Image 36	41
Image 37	40
Mode	**40**

C7C	Chromosomes number
Image 2	38
Image 3	39
Image 4	40
Image 5	40
Image 6	40
Image 7	41
Image 8	40
Image 9	41
Image 10	40
Image 11	38
Image 13	40
Image 14	40
Mode	**40**

Figure 3.2.2 – Kayrotyping of three different MED24- GFP-mAID clones. a. Bar charts show the percentage of chromosomes present in the sample of analyzed cells from images. b. Representative images of cells that were used to count chromosomes for each clone. c. Tables show the number of chromosomes in cell from the images taken.
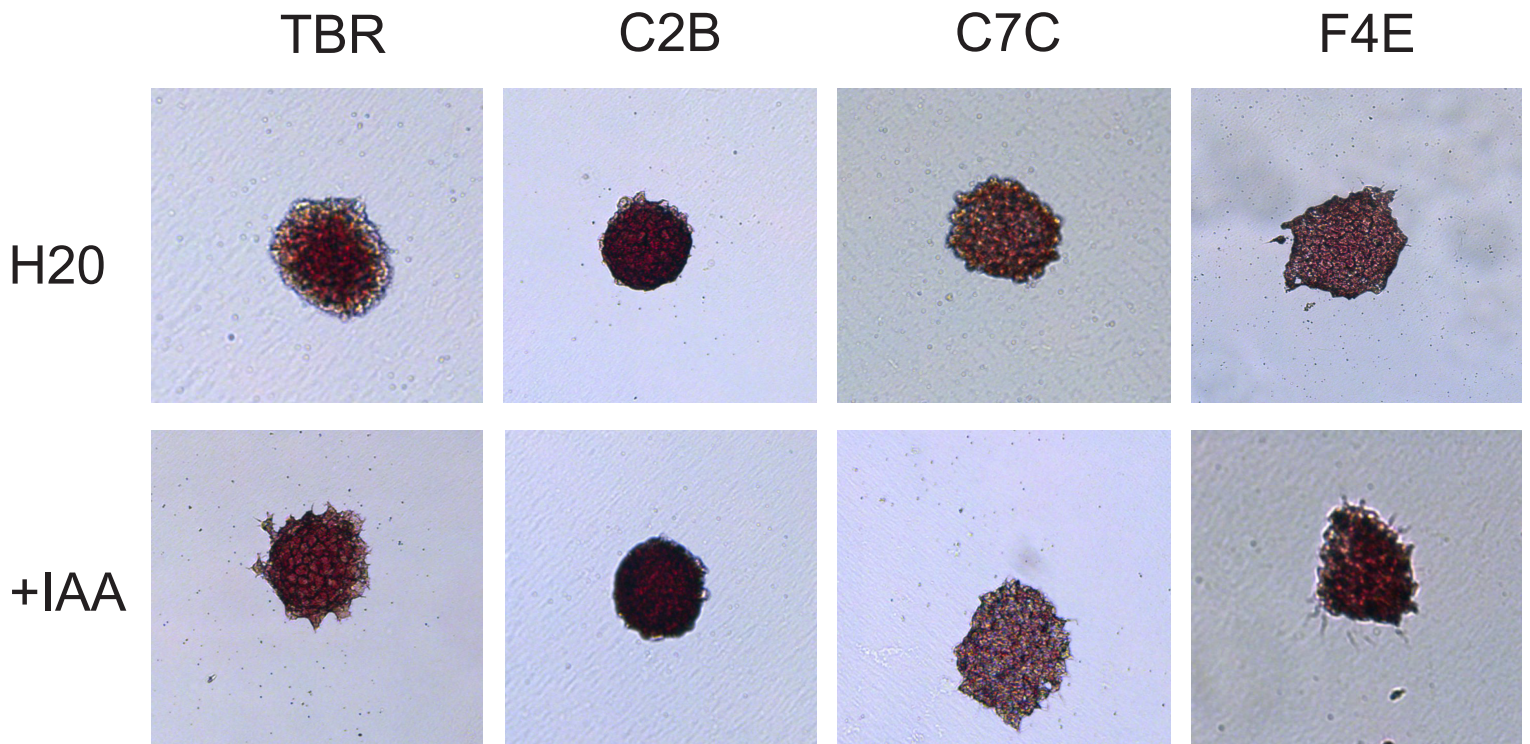
Figure 3.2.3 - AP staining for TBR cell line and MED24-GFP-mAID cell line - Cells were cultured in 2iL and expanded for 5 days before the staining. No significant difference was observed between control and IAA treatment. Images are representative of 3 biological replicated.
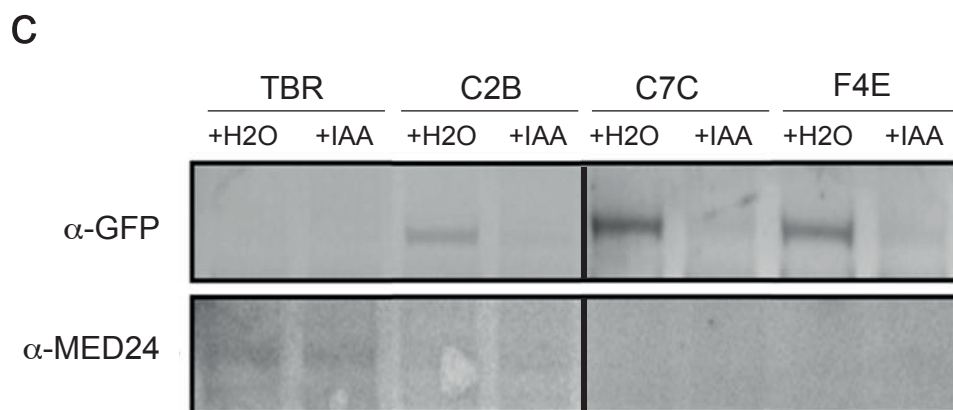
a) **IAA titration on MED24-GFP-mAID cell line (24 hours)**

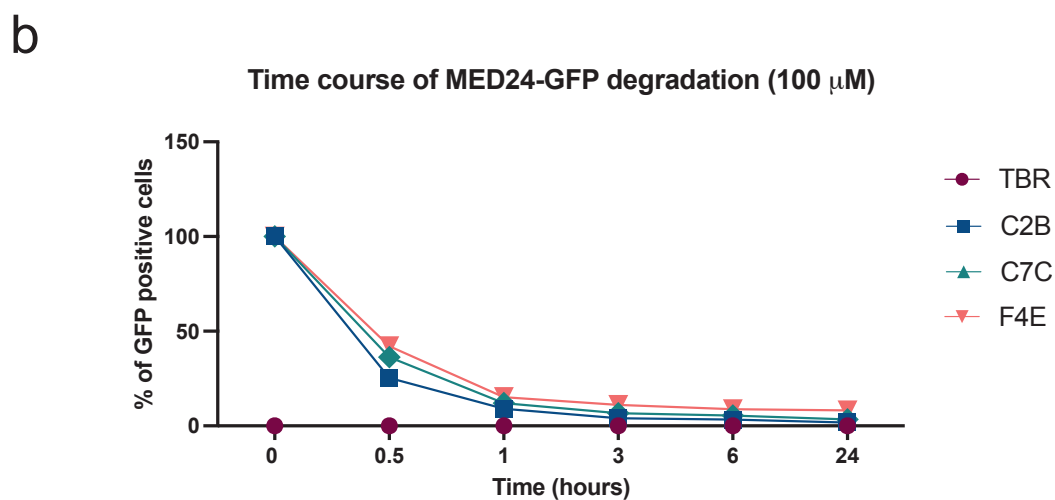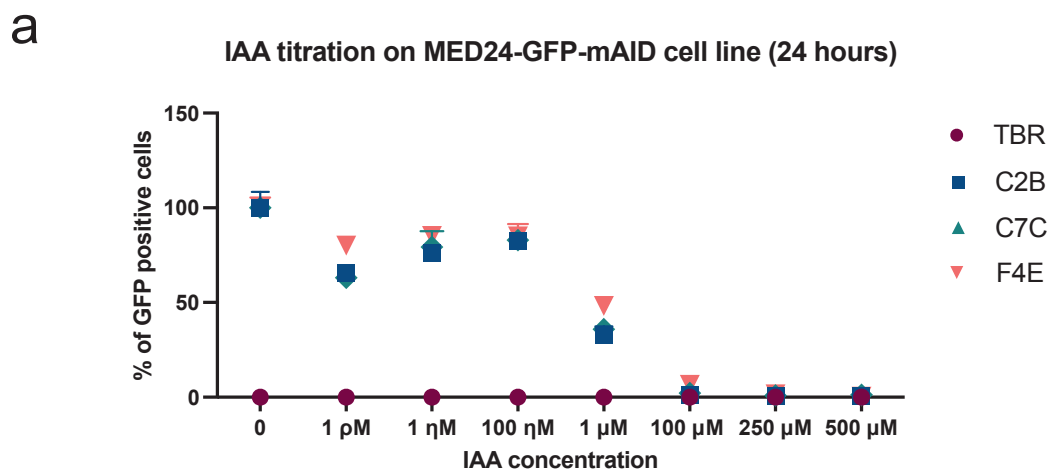b) **Time course of MED24-GFP degradation (100 μM)**

c)

Figure 3.2.2 – MED24- GFP-mAID cell line shows a rapid and inducible protein degradation – a) IAA concentration titration on MED24-GFP-mAID cell line. IAA was added and cells were collected 24 hours after treatment. GFP signal was measured by flow cytometry analysis. b) Time course of GFP degradation using 100 μM of IAA. After 3 hours most of GFP signal is gone in all three clones. c) Western Blot analysis of GFP and MED24 signal in cells treated and not treated with IAA. No signal was obtained for MED24 antibody.
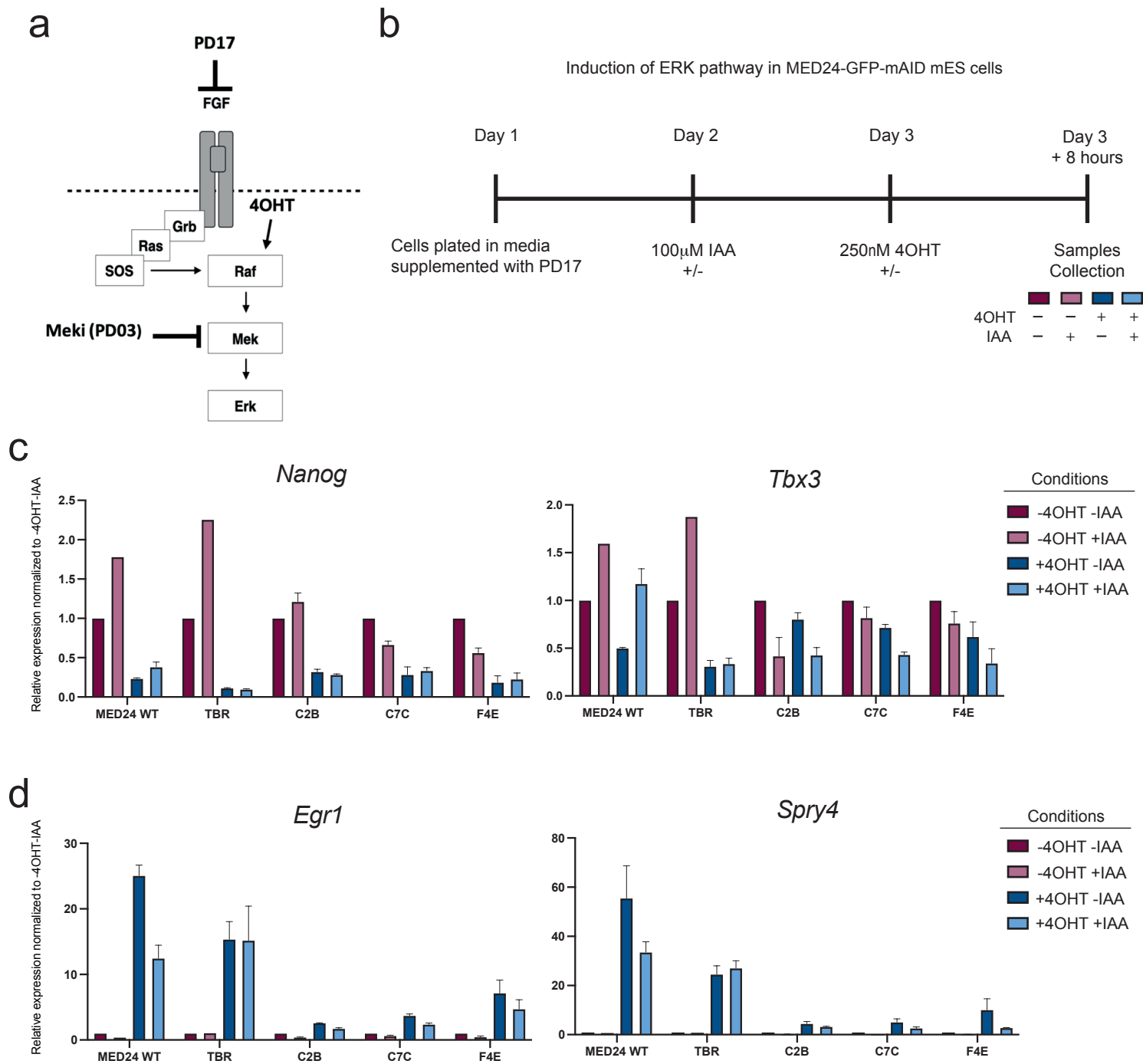
Figure 3.2.5 - Induction of ERK pathway in MED24-GFP-mAID cell line – a) Schematic illustration on where PD17 inhibitor acts and where 4OHT (tamoxifen) acts to activate the pathway. b) Timeline of ERK induction. Cells are plated in medium supplemented with PD17, the next day IAA was added to the cell. At day 2, 4OHT was added to the cells and 8 hours later cells were collected, and RNA or protein were extracted in all four conditions indicated. Expression levels of differentiation genes, *Egr1* and *Spry4* (c), and pluripotency genes, *Nanog* and *Tbx3* (d) in TBR and MED24-GFP-mAID clones were evaluated by qPCR. Bar plots show the relative expression normalized to -4OHT-IAA in each cell line.
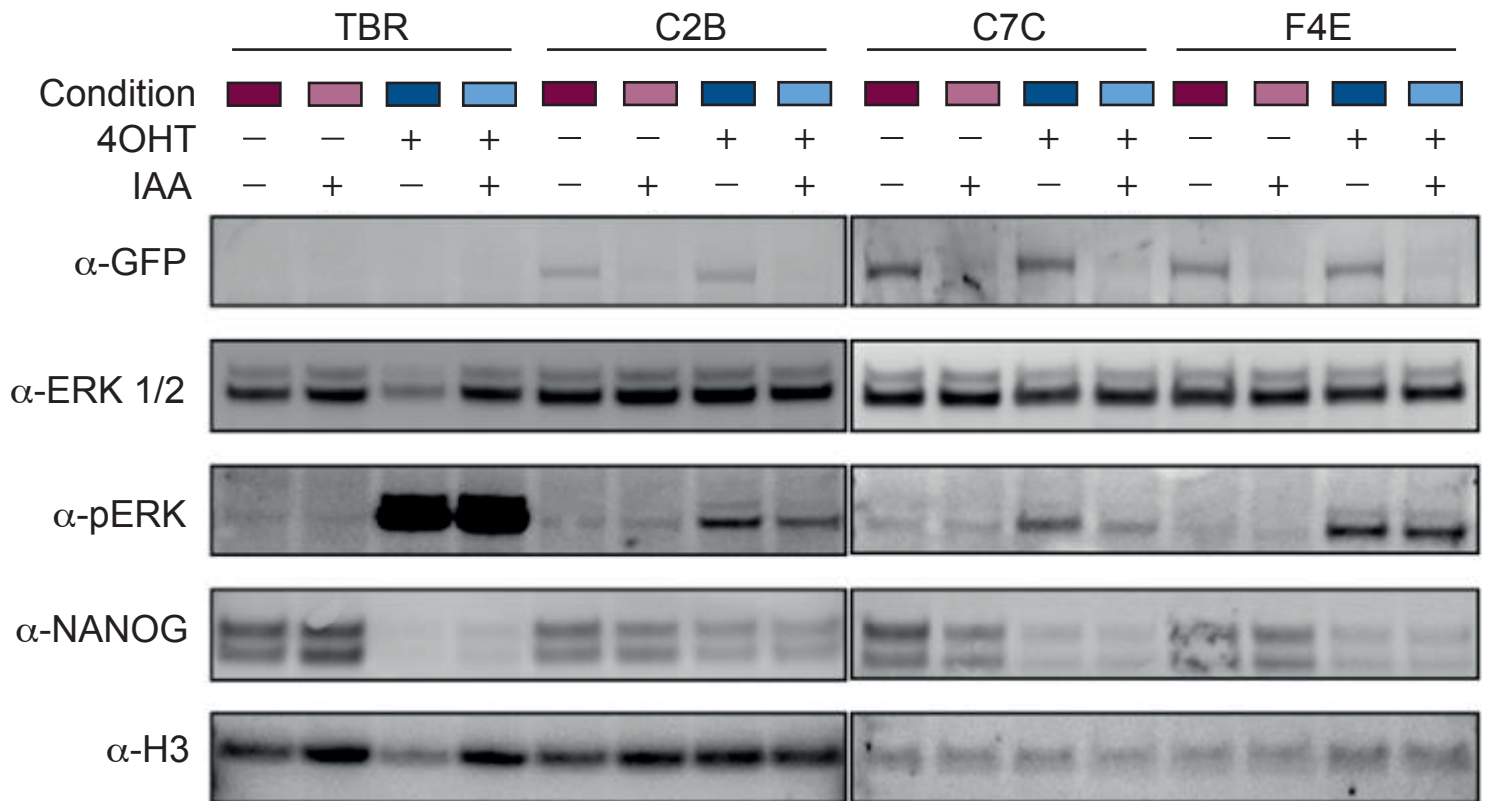
Figure 3.2.6- Western Blot analysis for the expression levels of GFP, ERK, pERK and NANOG in the TBR and MED24-GFP-mAID clones after MED24 depletion and ERK induction. Histone H3 was used as loading control.
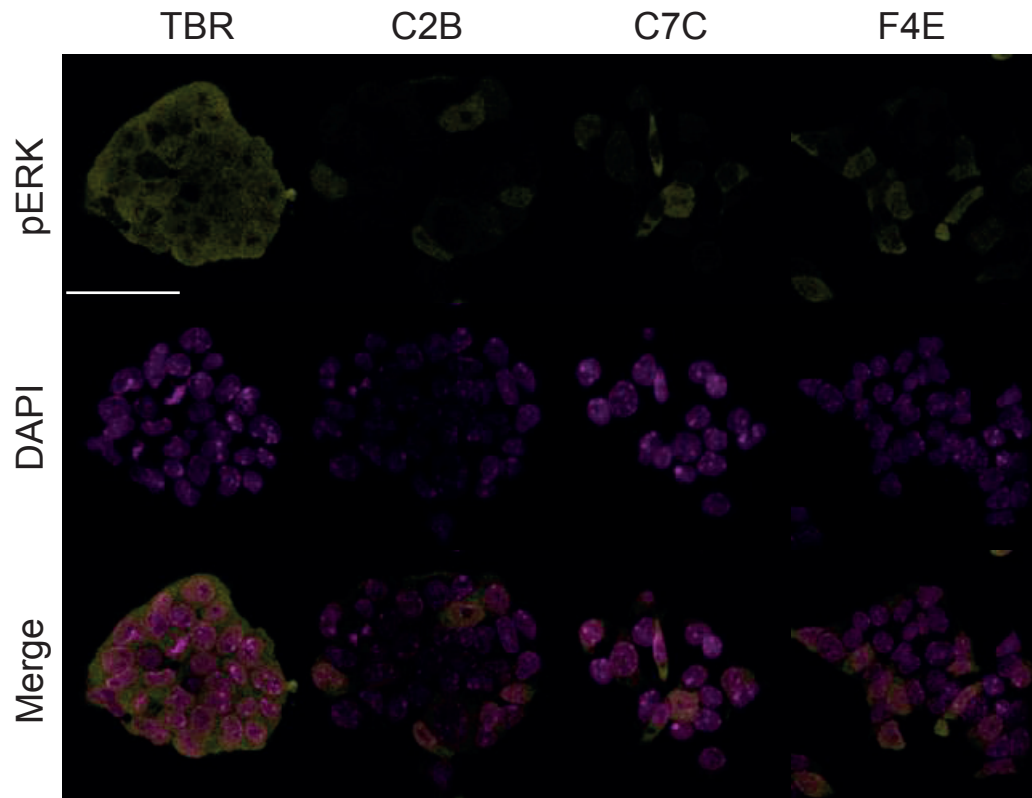
Figure 3.2.7 - Immunostaining of pERK (citrine) in the TBR and MED24-GFP-mAID clones following 2 hours of 4OHT treatment, including DAPI (magenta). Scale bars equal to 100μm.

3.3 *Nanog* expression in ground state ESCs is regulated by MED24 and MEKi

PD03 (MEKi) is known to be a potent inhibitor of the FGF/ERK pathway and is used to keep ESCs in a more pluripotent state. To investigate the interplay between MED24 depletion, MEKi and pluripotency maintenance I cultured cells in defined mediums: 2iL and ChirL (only CHIR99021 and LIF were added to the medium) and performed a self-renewal assay (AP staining). Depletion of MED24 in cells cultured in 2iL did not seem to affect the pluripotent state with most cells being AP$^+$ (Fig.3.3.1a, b). However, when cells were cultured in ChirL a difference in cell survival was noticeable and cells without MED24 appeared AP$^-$, indicating that they exited the ground state (Fig.3.3.1a, b).

Furthermore, to assess expression of pluripotency genes, MED24 was degraded at different time points and gene transcription was measured. After 3 and 6 hours of MED24 depletion in 2iL a downregulation of *Nanog* gene transcription was observed, with this downregulation disappearing after 24 hours of treatment (Fig. 3.3.1c)). This result explains why no difference was observed in the self-renewal assay.

Then, to assess crosstalk between MED24 and the FGF/ERK pathway cells, were first adapted to ChirL medium and then MED24 was depleted. Gene expression analysis revealed that without MEKi, cells seem to lose the ability to recover *Nanog* expression after 24 hours of IAA treatment (Fig. 3.3.1c). Again, the self-renewal assay supported this result, showing that cells could not maintain the pluripotent state when MED24 loss was coupled with FGF/ERK signaling (removal of MEKi).
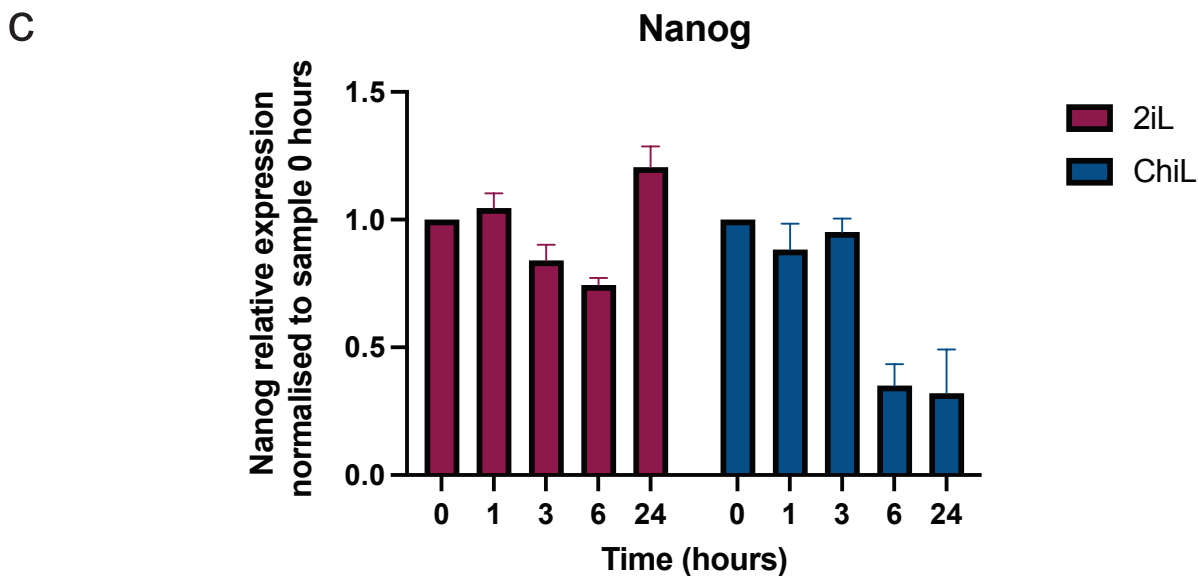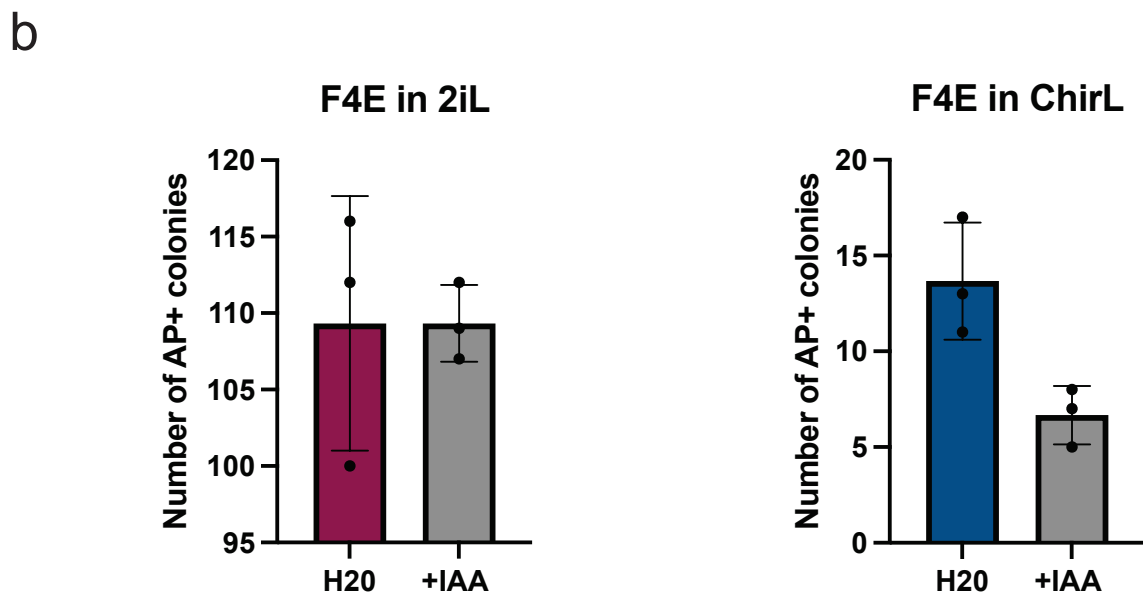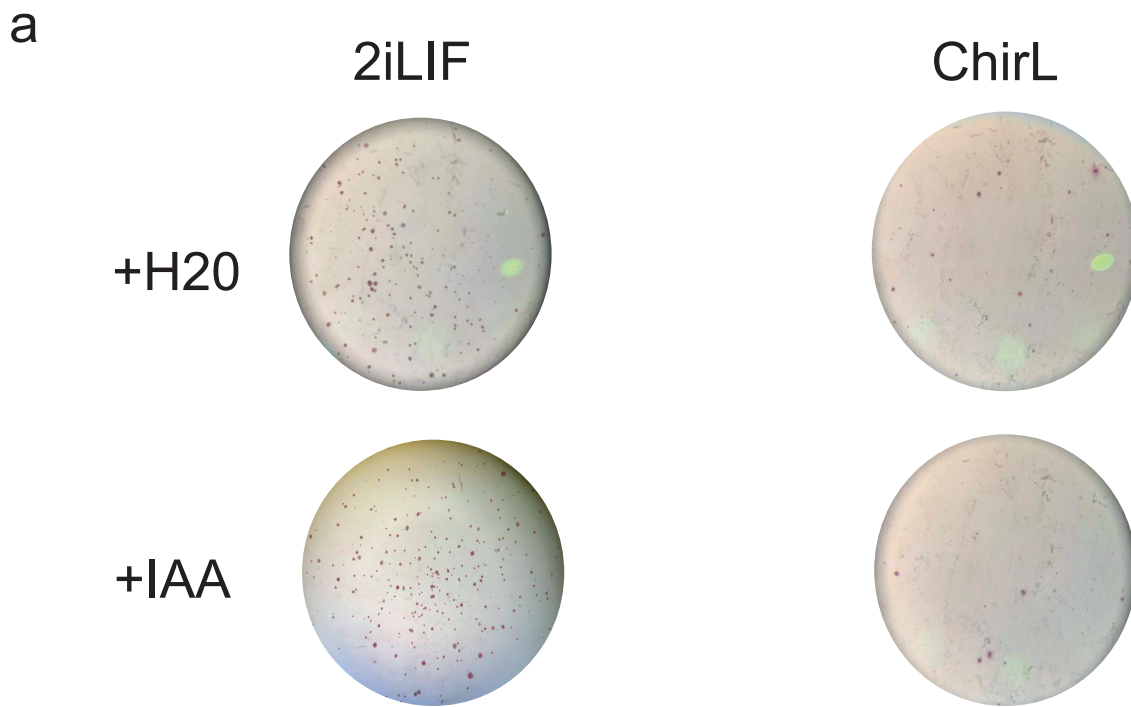
Figure 3.3.1 – MED24 and MEKi are necessary to keep Nanog expression in ground state cells - a) Alkaline Phosphatase (AP) staining shows that the MEF24-GFP-mAID cell line (F4E) keep the pluripotent stage when cultured in 2iL, no difference is observed when IAA is added to the culture medium. In the right panel, a difference in the number of AP+ colonies between control and cells treated with IAA is visible. b) Quantification of AP+ colonies in the medium indicated. Data shown is representative of 3 biological replicates. c) RT-qPCR analysis of *Nanog* expression in different time points for MED24 depletion in the presense and absence of MEKi.

3.4 MED24 is essential for PrE endoderm but not Neural differentiation

As mentioned earlier, Hamilton et al. showed that the MED24 subunit is essential for normal PrE differentiation while it is dispensable for neural differentiation. To confirm that the same phenotype could be observed using a cleaner degradation system, differentiation potential of the MED24-GFP-mAID cell line was tested. Neural differentiation (cells cultured in N2B27 for 7 days) did not appear to be affected by MED24 removal (100 μM IAA treatment), with both treated and non-treated cells upregulating TUJ1, a neural marker (Fig. 3.4.1). Alternatively, by day 7 of PrE differentiation (cells cultured in RACL) (see Table 2.1), cell morphology analysis showed that PrE differentiation was compromised in MED24 depleted cells, while TBR cells formed the characteristic PrE patches (experimental schematic, Fig. 3.4.2a; Fig. 3.4.2b). MED24 loss was confirmed by GFP expression, showing a diminished signal in the MED24-GFP-mAID line at day 5 (Fig 3.4.2c). A difference in the expression of a PrE cell surface marker (PDGFRα) between treated and non-treated cells was also observed (Fig. 3.4.2d). Immunostaining at day 5 further supported that MED24 depleted cells failed to differentiate to PrE since they were unable to upregulate GATA6 (Fig. 3.4.3). These results indicate that this system presents a similar phenotype as the previously described for the MED24 cKO line, reinforcing that MED24 is necessary for PrE but not neural differentiation.
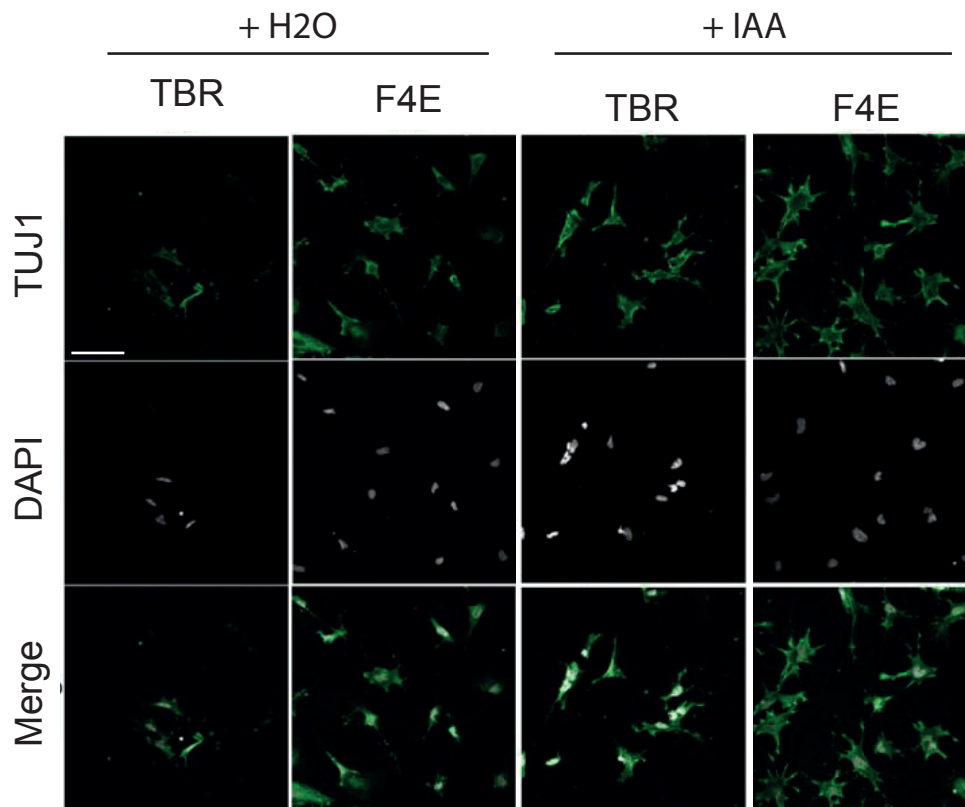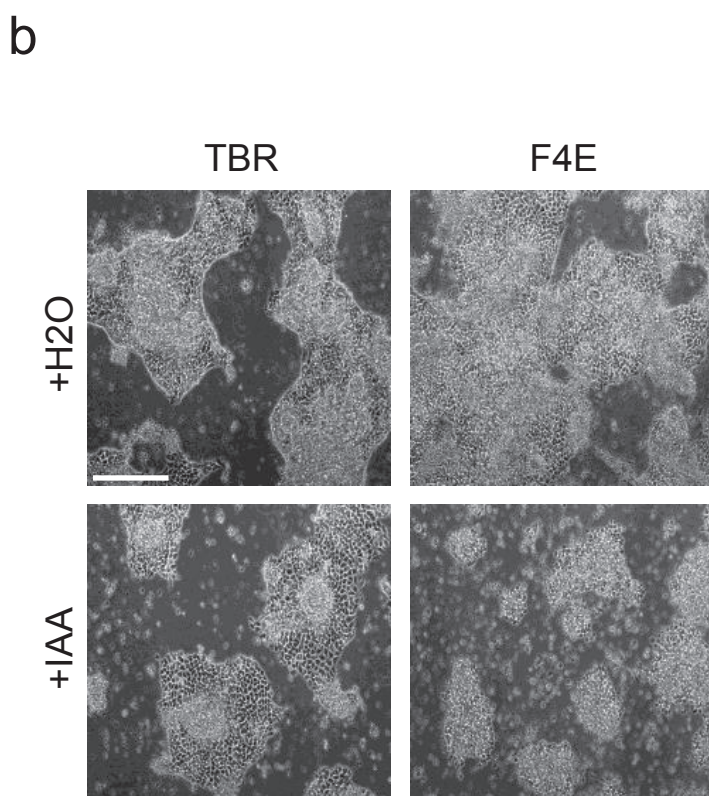
Figure 3.4.1 - Expression of TUJ1 is not affected in MED24-depleted cells undergoing neural differentiation. Immunostaining of TUJ1 (green) of TBR and MED24-GFP-mAID clones at day 5 of neural differentiation, in the absence or presence of IAA (100μM), including DAPI (grey). Scale bars equal to 100μm.

a Primitive Endoderm (PrE) differentiation of MED24-GFP-mAID mES cells

Day 1 — Day 3 — Day 5 — Day 7

Cells plated in RACL media supplemented with 100μM IAA

- Cell morphology
- PDGFRa expression
- MED24-GFP expression

b

TBR   F4E

+H2O

+IAA

c + H2O   + IAA (100μM)

% of GFP +ve cells

d % of PDGFRα +ve cells
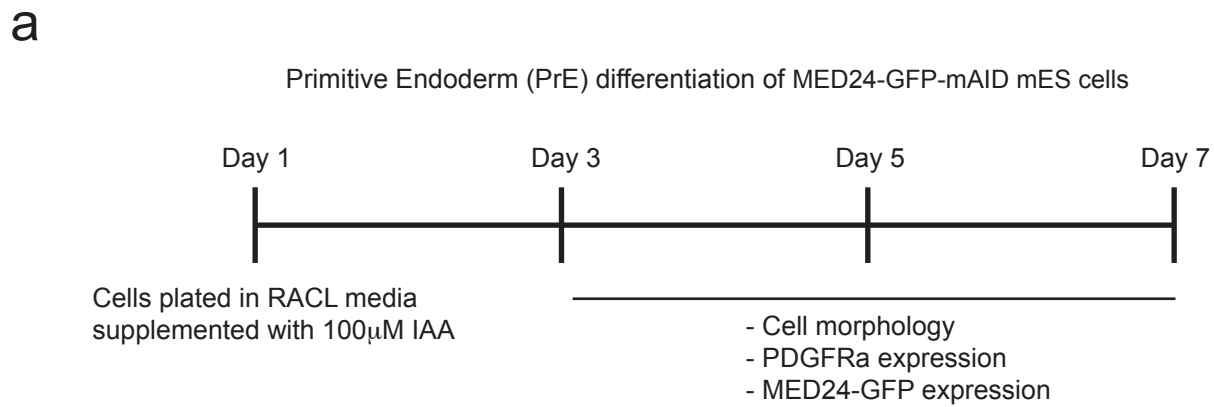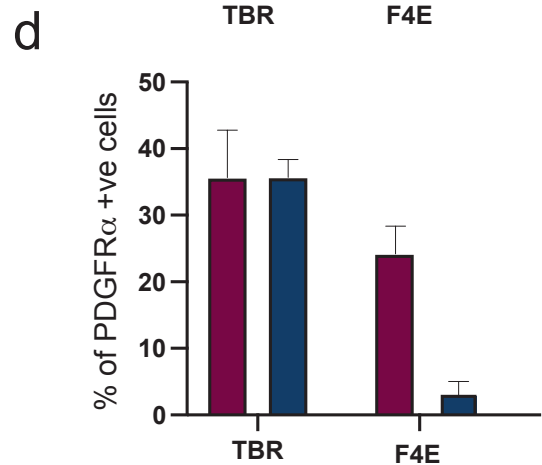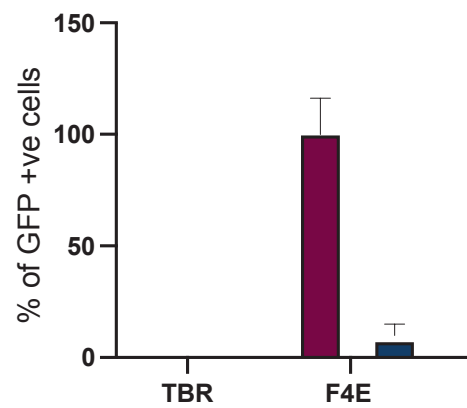
Figure 3.4.2 - MED24 is essential for PrE differentiation – a) A schematic diagram for Primitive Endoderm (PrE) differentiation of MED24-GFP-mAID mES cells. b) Brightfield images of TBR and MED24-GFP-mAID clones at day 7 of PrE differentiation in the absence or presence of IAA (100μM). Scale bar equals to 100μm.

Figure 3.4.3 - Expression of GATA6 is reduced in MED24-depleted PrE. Immunostaining of GATA6 (cyan) of TBR and MED24-GFP-mAID clones at day 5 of PrE differentiation, in the absence (a) or presence of IAA (100μM) (b), including DAPI (grey). Scale bars equal to 100μm.

5. Discussion

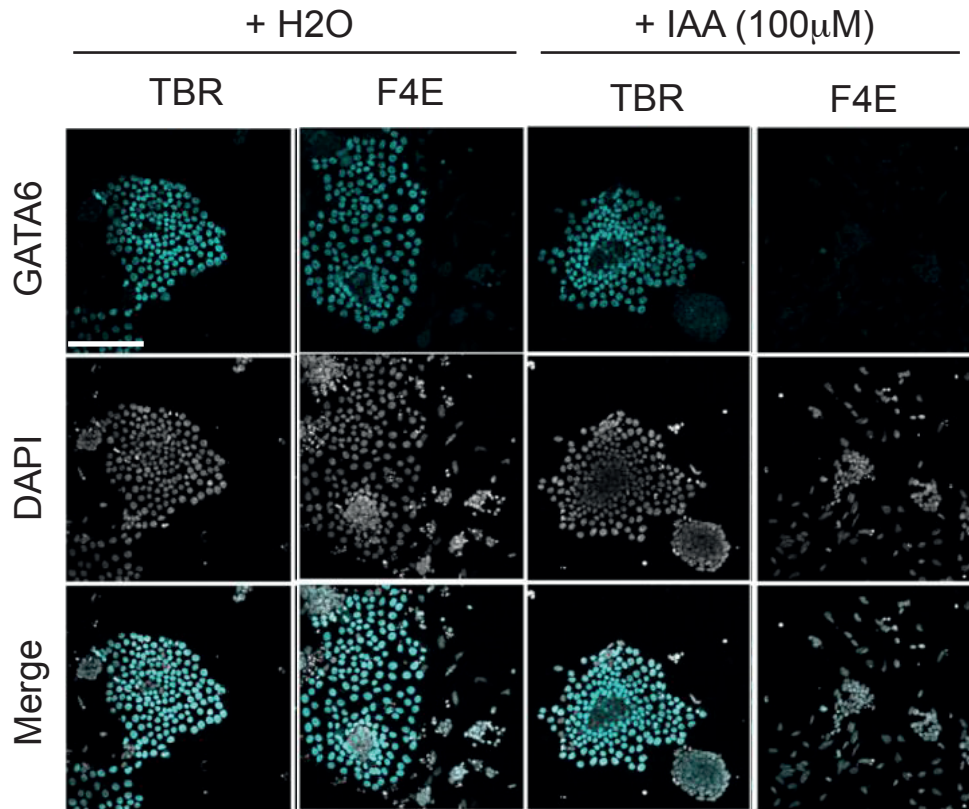In this thesis, I addressed the hypothesis that regulation of MED24 plays an important role in ESC transcriptional response.

First, I set out to investigate which domains of MED24 are required for its functionality, beginning with a general analysis of structural components and motifs. Different LxxLL motifs were identified, which have been described as necessary to mediate the binding of proteins to nuclear hormone receptors (NRs) (Heery et al., 1997). Furthermore, adjacent to one of the LxxLL motifs, MED24 has two serine residues that have been shown to be specifically phosphorylated by ERK (Hamilton et al., 2019). I designed six different truncations of the MED24 protein to decipher how crucial the regions around the phosphosites were for function as well as assess potential crosstalk with NR binding (Fig. 2.1 and Fig. 3.1.1c). I showed that the truncated forms were stably produced and could be used to investigate MED24 structure. I found that ΔMED24_6, deletion of the NR-binding motif adjacent to the ERK-phosphosites, lost its nuclear localization, resulting in a protein mainly expressed in the cytoplasm (Fig. 3.1.3). Of note, motif analysis on ERK-repressed enhancers, such as *Nanog*, showed an enrichment for proteins with nuclear hormone receptor motifs, which are part of the NR superfamily (Hamilton et al., 2019). Specifically, oestrogen-related receptor β (ESRRB) binding motifs were enriched in ERK-repressed enhancers, coinciding with MED24 binding sites. Therefore, the interaction of ESRRB with MED24 through the NR motif might be necessary for MED24 nuclear localization. I could test this hypothesis by assessing MED24 localization in an *Esrrb* KO cell line. Of course, other NRs, such as Nr5a2, could compensate for Esrrb loss, therefore it would also be beneficial to characterize co-binding of NRs (ESRRB and NR5A2) with the MED24 truncated proteins. The co-binding assays could also address the importance of the adjacent phosphosites, with the possibility of reduced binding with truncation missing these sites. This analysis could be improved by the addition of a MED24 mutant harboring a specific mutation in the NR binding motif and specific phospho-mutants.

The protein structure field is advancing quickly with the implementation of new technology such as CryoEM and AI modeling, resulting in novel information about protein structures coming out every day. As an example, at the start of this project very little was known about MED24 structure and about the mammalian Mediator complex. This reality changed during just the last year, with novel information about the Mediator structure and its subunits becoming known (El Khattabi et al., 2019; Zhao et al., 2020). These new analyses revealed critical residues in the mediator subunits,

offering even more possibilities for potential mutations to assess MED24 function. Furthermore, I was able to visualize MED24 using the brand new AI AlphaFold software database (Jumper, J *et al*. *Nature* (2021); https://alphafold.ebi.ac.uk/), which provides 3D structures of proteins based on sequence alone (Appendix 1). I believe that this tool offers a novel way to model my MED24 truncations, producing a more comprehensive analysis of how the loss of specific regions and residues affects the overall structure of MED24, providing more information about how it binds other TFs or how it fits into the mediator complex as a whole.

Secondly, I demonstrated that the MED24-GFP-mAID cell line can be used as a time-sensitive system to study the effects of MED24 loss (Fig. 3.1a, b). I showed that an efficient and rapid degradation of MED24 occurs after IAA addition, with 30 minutes reducing the GFP signal by half and loss of a detectable signal achieved after 3 hours (Fig 3.2.4). Depletion of MED24 protein was confirmed by Western Blot analysis of GFP levels (Fig. 3.2.6). A specific MED24 antibody was also used to detect the protein but unfortunately, no signal was obtained. Addition of a tag, like GFP-mAID, to proteins might interfere with stability of the protein leading to a possible decrease in the levels of protein expression, making it difficult to obtain a signal.

Moreover, the MED24-GFP-mAID cell line contained a $cRAF^{\Delta 26-303}$- ER^{T2} sequence, but unfortunately its activity was heterogeneous and not being ideal for studying ERK induction. A possible explanation for reduced ERK induction could be the difference in how the $cRAF^{\Delta 26-303}$-ER^{T2} sequence was inserted into the genome. As explained earlier, the *Tigre locus* was targeted with a piece of DNA containing the OsTIR1, ARF16 and $cRAF^{\Delta 26-303}$- ER^{T2}. Consequently, poor cleavage efficiency from the two different self-cleavage peptides sites (T2A and P2A; Fig.3.2.1b) could be leading to instability of the $cRAF^{\Delta 26-303}$- ER^{T2} protein, resulting in a reduced ERK cell induction. Solutions to address this issue include increasing the concentration of tamoxifen (4OHT) or re-targeting the *Tigre locus* with a different approach, where the OsTIR and ARF16 sequence are integrated in one allele and the $cRAF^{\Delta 26-303}$- ER^{T2} sequence into the other allele. Additionally, I inserted in the MED24-GFP-mAID cell line with an rtTA sequence (TRE transactivator) and showed that this cell line can express all the MED24 truncations vectors (Appendix 2). Together this suggests that it is possible to obtain a cell line where I can acutely deplete the endogenous MED24 protein, induce ERK phosphorylation homogeneously and conditionally express specific MED24 truncations. This tool will allow me to investigate the role of different MED24 domains in the context of FGF/ERK signaling. For example, do MED24

structural changes affect ERK-dependent enhancer activity? A new technique implemented in the Brickman Laboratory, Cut&Run, could be used to build a comprehensive map of TF binding and chromatin marks, like the active enhancer marker H3K27ac. This method would allow me to determine if specific domains of MED24 are necessary for enhancer activation. Specifically I could look at ERK activated and repressed enhancers and see how its activity (H3K27ac levels) is affected by the different MED24 truncations.

Next, I looked at the importance of MED24 for maintaining ground state pluripotency, in other words in the presence of MEK inhibition. I showed that in the absence of MED24, MEKi is required to maintain pluripotency. More specifically, MED24 depletion in 2iL showed that short term loss (3-6 hours) resulted in downregulation of *Nanog* expression while a longer depletion (24 hours) recovered levels (Fig. 3.3.1c). AP staining of ESC colonies grown in 2iL also showed no detectable difference between depleted and non-depleted cells, supporting the *Nanog* expression observation. What possible mechanism could cells possess to overcome the loss of MED24? MED24 has been described as part of a tail sub-complex with MED23 and MED16, that is implicated in MAPK activity (Ito et al., 2002; Stevens et al., 2002b). As mentioned, recent cryo-EM structures revealed extensive binding sites between these three subunits, supporting the possibility of an intricate relation between this sub-complex. Additional protein structures revealed a high degree of similarity between MED24 and MED23 subunits (Zhao et al., 2020). This information leads me to hypothesize that Mediator complex is destabilized upon short depletion of MED24 in 2iL conditions, but when this loss is maintained for longer periods, Mediator adapts, forming an alternative complex to resume expression of pluripotency genes. To determine if the MED23 subunit could replace MED24 in this alternative complex I could make double knockouts of these subunits and additionally of MED16 and assess the transcriptional response in the ground state. However, when cells were released from MEKi, the recovery of *Nanog* expression was not detected. Could this mean that in the presence of FGF signals, Mediator complex cannot recover from the MED24 loss? This could possibly be explained by the presence of basal levels of FGF signaling, producing enough phosphorylation to sufficiently destabilize the Mediator complex, blocking the formation of an alternative structure. Therefore, loss of MED24 leads to the amplification of the FGF repressing effect.

Finally, I challenged cells to differentiate into neurons or primitive endoderm in the presence or absence of MED24, using my time-sensitive system for MED24 depletion. I showed that while MED24 was not necessary for neural differentiation it was required for PrE differentiation,

validating what Hamilton et al showed previously. Therefore, acute depletion of MED24 had the same effect on differentiation potential as long-term loss that was necessary in the previous system. These observations show that while Mediator at pluripotency enhancers in the ground state can recover from MED24 loss, MED24 appears to be required for the prolonged assembly of new Mediator complex at the differentiation enhancers, which is required for PrE maturation.

These observations lead me to propose a model (Figure 4.1):

- In 2iL conditions: short-term loss of MED24 destabilizes the Mediator complex causing it to fall off chromatin. Longer-term loss of MED24 gives the cells time to adopt an alternative structure that is functional and able to recover *Nanog* levels (Fig. 4.1a).
- When cells are released from MEKi, MED24 is necessary to keep the holoenzyme at nuclear hormone enhancers. Basal phosphorylation levels, as a result of FGF signaling, are sufficient to fully destabilize the Mediator complex, blocking recovery of *Nanog* expression (Fig. 4.1a).
- In differentiation conditions, where high levels of FGF are involved, activated genes are not able to be transcribed in the absence of MED24. Mediator complex cannot assembly *de novo* in MED24 is absence (Fig.4.1b).

I could validate aspects of this model using RIME (Rapid Immunoprecipitation Mass spectrometry of Endogenous proteins), a method recently implemented in the laboratory. This would be done by pulling down a core protein (i.e. MED1) and characterizing Mediator complex composition in the presence or absence of MED24. Is stability of the complex affected or is the connection to the core module loss leading to a non-functional complex? (El Khattabi et al., 2019). How is the binding of PIC components affected, specifically RNA Pol II? Furthermore, by using a method based in mass spectrometry analysis I have the opportunity to assess phosphorylation levels of residues within the subunits, with or without FGF/ERK signaling. Finally, this technique could be applied to cells expressing MED24 truncations, in order to test if the integration of MED24 into the core module and/or the tail module is dependent on specific domains. Where do TFs bind MED24? Additionally, using this technique with the different MED24 truncations will give us a better understanding of how MED24 interacts with other subunits, TFs and cofactors of the PIC complex. In the context of ERK-mediated response, we known that MED24 is present at activated and repressed genes but the motifs differ, with ESRRB and NR motifs more specific to ERK-

repressed enhancers. So how is specificity accomplished? Do activators bind to a specific domain of MED24 while repressors prefer a different domain?

For a broader analysis of how depletion of MED24 affects E-P interactions, we could compare chromatin confirmation using Hi-C analysis or Micro-C for fine structures. Does acute depletion of MED24 have a strong effect on chromatin structure? Are the loops maintained with acute depletion and destroyed with long-term depletion? Do the loops change to accommodate a possible alternative Mediator complex? Overall, this could provide a better understanding of how Mediator complex composition affects chromatin structures at specific regions of the genome.
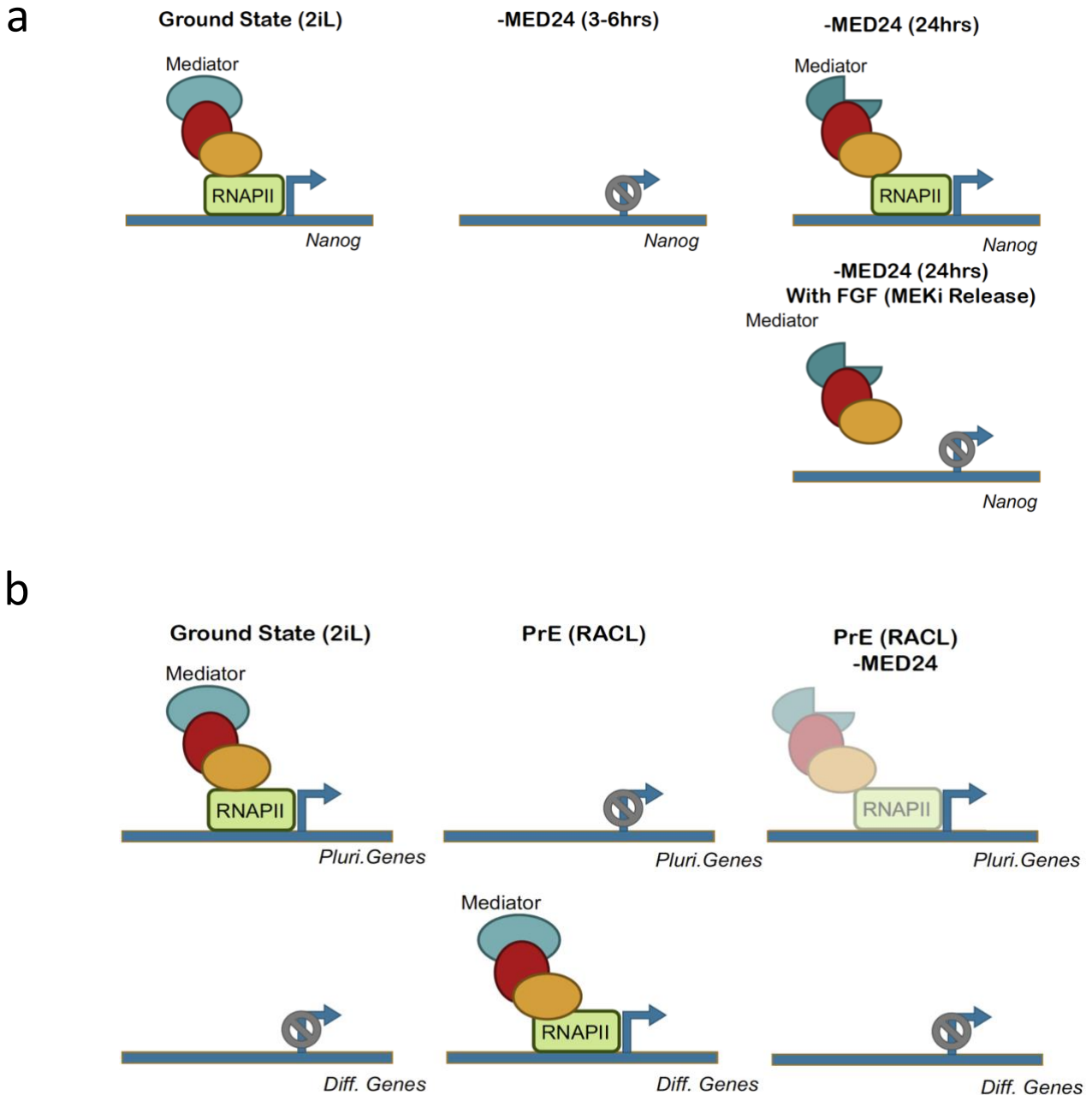
Figure 4.1 – Model proposed to explain MED24 requirement in the gene transcription - a) In 2iL (top) short-term loss of MED24 causes the complex to fall off of chromatin. Longer periods of MED24 depletion leads to the formation of an alternative complex that allows the recovery of Nanog. In the presence of FGF (lower), phosphorylation causes instability in the Mediator complex and consequently it leaves the pluripotency genes. b) In differentiation conditions, the PIC complex leaves the pluripotency genes. Differentiation genes are not able to be transcribed in the absence of MED24. Mediator complex cannot assembly de novo in MED24 is absence.

1. Conclusion

Gene transcription is a complex and intricate mechanism that is extensively investigated. Different factors involve in the regulation have been identified. Thus, a good system to study these factors is necessary. LOF systems represent a functional tool that allows us to understand better the role of specific genes and proteins. MED24 was identify as a co-factor involved in ERK-mediated transcriptional response. In this thesis, I showed that study MED24 structure has the potential to give us great comprehension about protein function and regulation. This knowledge will also allow us to understand better regulation of gene transcription. I also show that an auxin-inducible degron (AID) system can be used to study protein depletion. Specifically, I show that MED24-GFP-mAID cell line can be used to study acute protein depletion. This tool combine with MED24 truncations will allow us to get an insight about MED24 structure and how MED24 is regulated by ERK signaling. Here, I explored the effects of MED24 acute depletion in pluripotency maintenance and differential potential. I showed that in a ground state, *Nanog*, a pluripotency transcription factor is regulated by MED24 and MEKi. I show that while Mediator at pluripotency enhancers in the ground state can recover from MED24 loss, MED24 appears to be required for the prolonged assembly of new Mediator complex at the differentiation enhancers, which is required for PrE maturation.

Concluding, MED24 seems to be a major player in transcription regulation in ESCs and more analysis will provide us with better understating about this subunit acting mechanism.

Appendix 1
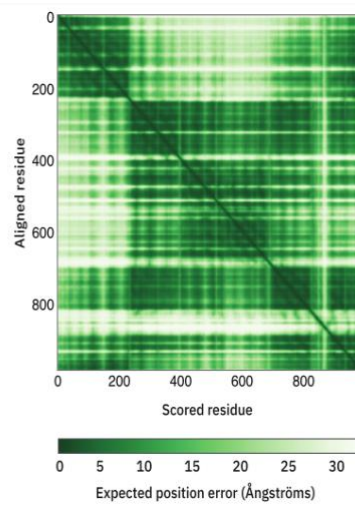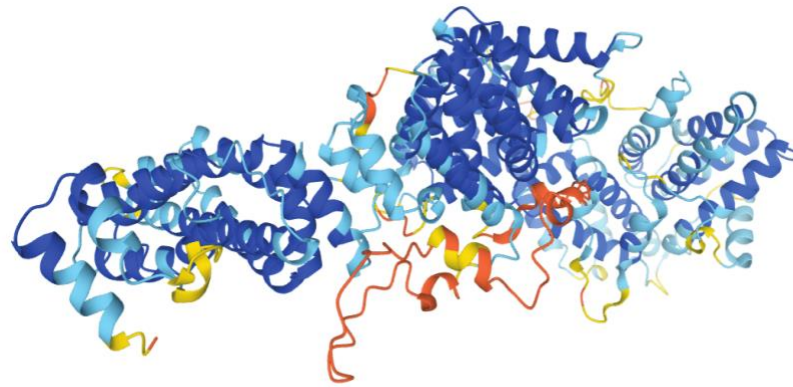
Appendix 1 - 3D structure predicted based on AAs sequences by AlphaFold. Predicted aligned error The color at position (x, y) indicates AlphaFold's expected error at residue x, when the predicted and true structures are aligned on residue y.

Appendix 2
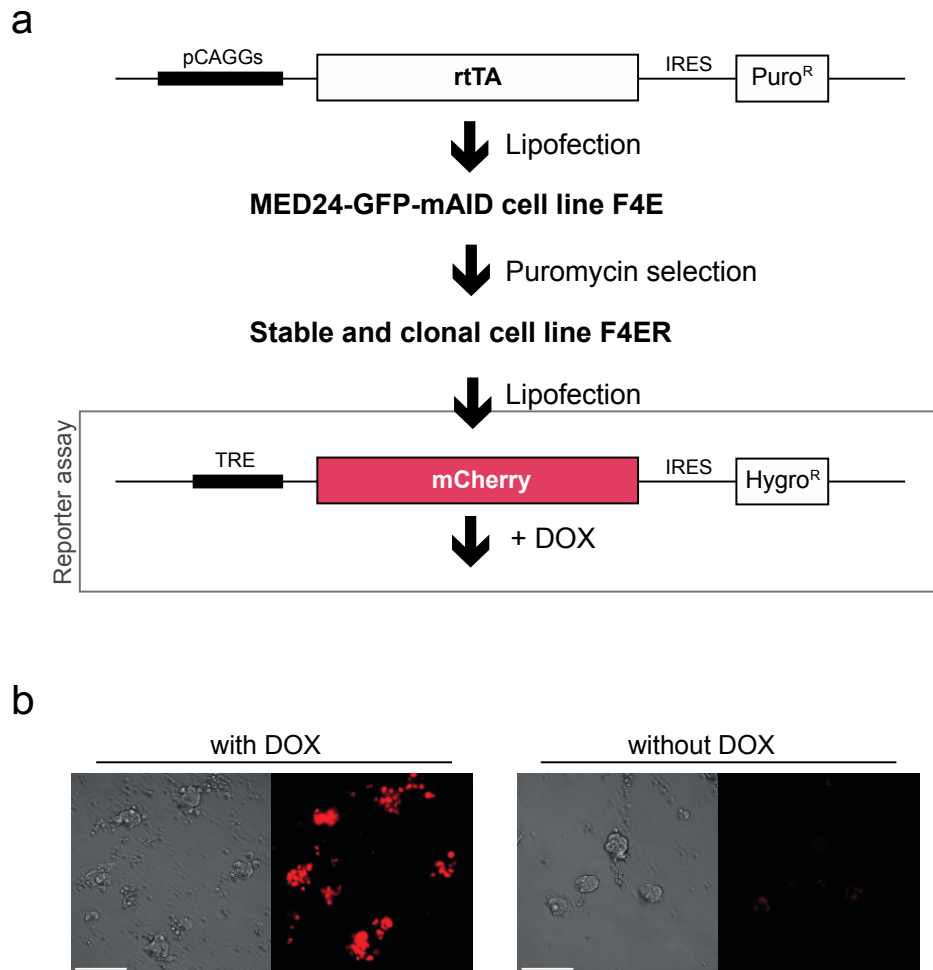
Appendix 2 – a) Strategy used to generate F4ER cell line - a) rtTA sequence was insert in F4E clone of MED24-GFP-mAID cell line. Three different clones were screen using the expression of a mCherry plasmid under a TRE system. b) mCherry expression was confirmed using live imaging. Scale bars equal to 100 ㎜
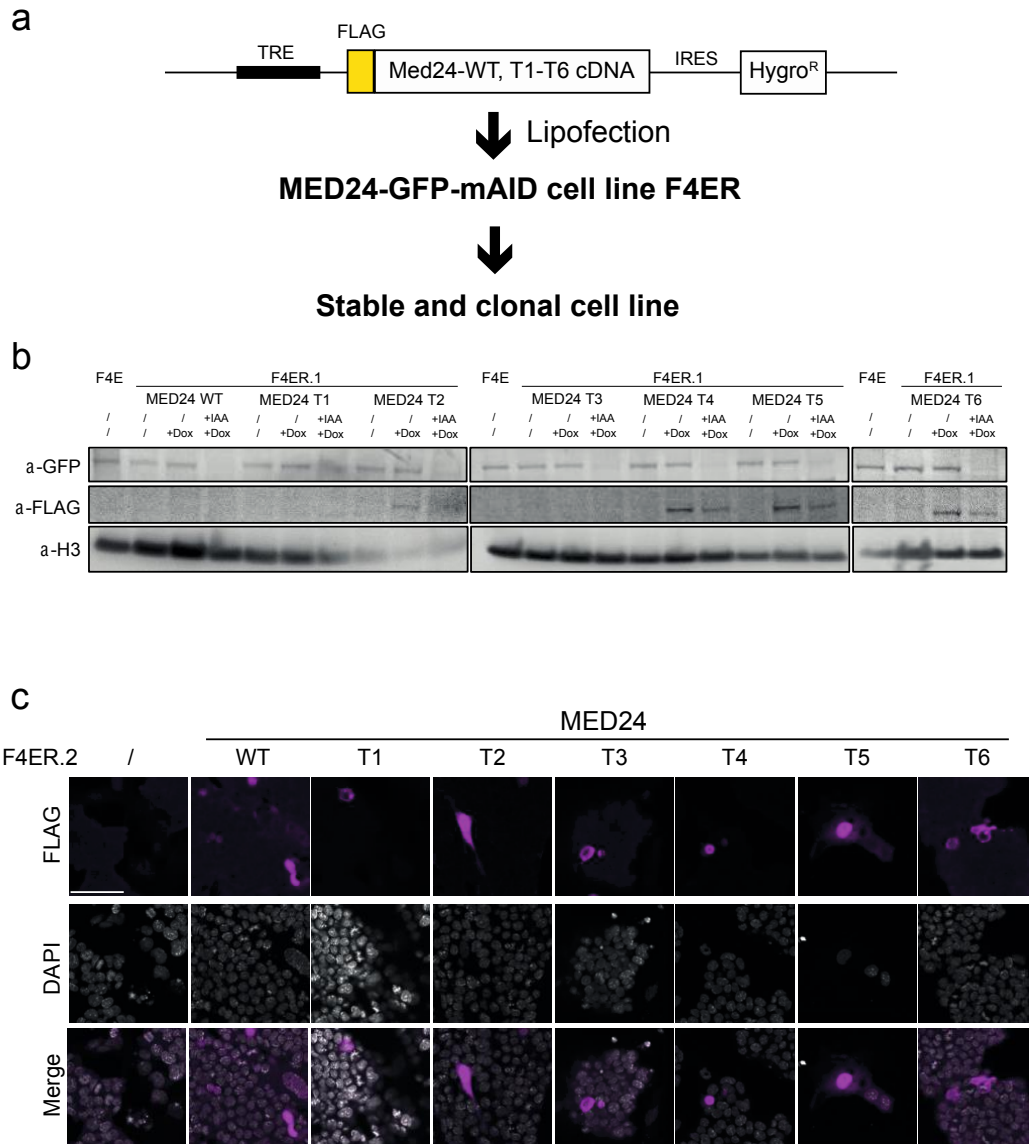
Appendix 3

Appendix 3 - a) Linearized vectors containing the MED24 WT sequence and the six different truncations were transfected to F4ER cell line. b) Western Blot analysis for GPF and FLAG-tag confirm the functionality of the mAID system and the expression of ΔMED24_2, ΔMED24_4, ΔMED24_5 and ΔMED24_6. c) Immunostaining against FLAG shows positive cells in all truncations, These results indicate that a stable cell line expressing all 6 truncations is possible

References

Ambrosetti, D. C., Basilico, C., & Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Molecular and Cellular Biology*, *17*(11), 6321–6329. https://doi.org/10.1128/mcb.17.11.6321

Artus, J., Piliszek, A., & Hadjantonakis, A. K. (2011). The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. *Developmental Biology*, *350*(2), 393–404. https://doi.org/10.1016/j.ydbio.2010.12.007

Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M., & Kornberg, R. D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. *Science*, *283*(5404), 985–987. https://doi.org/10.1126/science.283.5404.985

Banerji, J., Rusconi, S., & Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, *27*(2 PART 1), 299–308. https://doi.org/10.1016/0092-8674(81)90413-X

Barlow, P., Owen, D. A., & Graham, C. (1972). DNA synthesis in the preimplantation mouse embryo. *Journal of Embryology and Experimental Morphology*, *27*(2), 431–445.

Belakavadi, M., Pandey, P. K., Vijayvargia, R., & Fondell, J. D. (2008). MED1 Phosphorylation Promotes Its Association with Mediator: Implications for Nuclear Receptor Signaling. *Molecular and Cellular Biology*, *28*(12), 3932–3942. https://doi.org/10.1128/mcb.02191-07

Bernecky, C., & Taatjes, D. J. (2012). Activator-mediator binding stabilizes RNA polymerase II orientation within the human mediator-RNA polymerase II-TFIIF assembly. *Journal of Molecular Biology*, *417*(5), 387–394. https://doi.org/10.1016/j.jmb.2012.02.014

Boroviak, T., Loos, R., Lombard, P., Okahara, J., Behr, R., Sasaki, E., Nichols, J., Smith, A., & Bertone, P. (2015). Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Developmental Cell*, *35*(3), 366–382. https://doi.org/10.1016/j.devcel.2015.10.011

Bradley, A., Evans, M., Kaufman, M. H., & Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, *309*(5965), 255–256. https://doi.org/10.1038/309255a0

Brumbaugh, J., Russell, J. D., Yu, P., Westphall, M. S., Coon, J. J., & Thomson, J. A. (2014). NANOG is multiply phosphorylated and directly modified by ERK2 and CDK1 in vitro.

Stem Cell Reports, *2*(1), 18–25. https://doi.org/10.1016/j.stemcr.2013.12.005

Bulger, M., & Groudine, M. (1999). Looping versus linking: Toward a model for long-distance gene activation. *Genes and Development*, *13*(19), 2465–2477. https://doi.org/10.1101/gad.13.19.2465

Cai, G., Imasaki, T., Yamada, K., Cardelli, F., Takagi, Y., & Asturias, F. J. (2010). Mediator Head module structure and functional interactions. *Nature Structural and Molecular Biology*, *17*(3), 273–279. https://doi.org/10.1038/nsmb.1757

Chadick, J. Z., & Asturias, F. J. (2005). Structure of eukaryotic Mediator complexes. *Trends in Biochemical Sciences*, *30*(5), 264–271. https://doi.org/10.1016/j.tibs.2005.03.001

Chazaud, C., & Yamanaka, Y. (2016). Lineage specification in the mouse preimplantation embryo. *Development (Cambridge)*, *143*(7), 1063–1074. https://doi.org/10.1242/dev.128314

Corden, J. L., Cadenat, D. L., Ahearn, J. M., & Dahmust, M. E. (1985). A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II (transcription/gene expression/protein phosphorylation). *Biochemistry*, *82*(December), 7934–7938.

Core, L. J., Martins, A. L., Danko, C. G., Waters, C. T., Siepel, A., & Lis, J. T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nature Genetics*, *46*(12), 1311–1320. https://doi.org/10.1038/ng.3142

D.S. Latchman. (1993). Transcription factors: an overview Function of transcription factors. *Int. J. Exp. Path*, *74*, 417–422. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2002184/pdf/ijexpath00017-0003.pdf

D'Alessio, J. A., Wright, K. J., & Tjian, R. (2009). Shifting Players and Paradigms in Cell-Specific Transcription. *Molecular Cell*, *36*(6), 924–931. https://doi.org/10.1016/j.molcel.2009.12.011

Darnell, J. E. (1982). Variety in the level of gene control in eukaryotic cells. *Nature*, *297*(5865), 365–371. https://doi.org/10.1038/297365a0

Davis, J. A., Takagi, Y., Kornberg, R. D., & Asturias, F. J. (2002). Structure of the Yeast RNA Polymerase II Holoenzyme. *Molecular Cell*, *10*(2), 409–415. https://doi.org/10.1016/s1097-2765(02)00598-1

Deato, M. D. E., Marr, M. T., Sottero, T., Inouye, C., Hu, P., & Tjian, R. (2008). MyoD Targets TAF3/TRF3 to Activate Myogenin Transcription. *Molecular Cell*, *32*(1), 96–105. https://doi.org/10.1016/j.molcel.2008.09.009

Dhaliwal, N. K., Miri, K., Davidson, S., Tamim El Jarkass, H., & Mitchell, J. A. (2018). KLF4

Nuclear Export Requires ERK Activation and Initiates Exit from Naive Pluripotency. *Stem Cell Reports*, *10*(4), 1308–1323. https://doi.org/10.1016/j.stemcr.2018.02.007

Dietrich, J. E., & Hiiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development*, *134*(23), 4219–4231. https://doi.org/10.1242/dev.003798

Dotson, M. R., Yuan, C. X., Roeder, R. G., Myers, L. C., Gustafsson, C. M., Jiang, Y. W., Li, Y., Kornberg, R. D., & Asturias, F. J. (2000). Structural organization of yeast and mammalian mediator complexes. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14307–14310. https://doi.org/10.1073/pnas.260489497

Ebmeier, C. C., & Taatjes, D. J. (2010). Activator-mediator binding regulates mediator-cofactor interactions. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(25), 11283–11288. https://doi.org/10.1073/pnas.0914215107

El Khattabi, L., Zhao, H., Kalchschmidt, J., Young, N., Jung, S., Van Blerkom, P., Kieffer-Kwon, P., Kieffer-Kwon, K. R., Park, S., Wang, X., Krebs, J., Tripathi, S., Sakabe, N., Sobreira, D. R., Huang, S. C., Rao, S. S. P., Pruett, N., Chauss, D., Sadler, E., … Casellas, R. (2019). A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell*, *178*(5), 1145-1158.e20. https://doi.org/10.1016/j.cell.2019.07.011

Evans, M. J., & Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. In *Nature* (Vol. 292, Issue 5819, pp. 154–156). https://doi.org/10.1038/292154a0

Feldman, B., Poueymirou, W., Papaioannou, V. E., DeChiara, T. M., & Goldfarb, M. (1995). Requirement of FGF-4 for postimplantatior mouse development. *Science*, *267*(5195), 246–249. https://doi.org/10.1126/science.7809630

Fondell, J. D., Ge, H., & Roeder, R. G. (1996). Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(16), 8329–8333. https://doi.org/10.1073/pnas.93.16.8329

Frankenberg, S., Gerbe, F., Bessonnard, S., Belville, C., Pouchin, P., Bardot, O., & Chazaud, C. (2011). Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. *Developmental Cell*, *21*(6), 1005–1013. https://doi.org/10.1016/j.devcel.2011.10.019

Frum, T., Halbisen, M. A., Wang, C., Amiri, H., Robson, P., & Ralston, A. (2013). Oct4 Cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Developmental Cell*, *25*(6), 610–622. https://doi.org/10.1016/j.devcel.2013.05.004

Fürthauer, M., Van Celst, J., Thisse, C., & Thisse, B. (2004). Fgf signalling controls the dorsoventral pattering of the zebrafish embryo. *Development, 131*(12), 2853–2864. https://doi.org/10.1242/dev.01156

Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D., & Robson, P. (2010). Resolution of Cell Fate Decisions Revealed by Single-Cell Gene Expression Analysis from Zygote to Blastocyst. *Developmental Cell, 18*(4), 675–685. https://doi.org/10.1016/j.devcel.2010.02.012

Hamazaki, T., Oka, M., Yamanaka, S., & Terada, N. (2004). Aggregation of embryonic stem cells induces Nanog repression and primitive endoderm differentiation. *Journal of Cell Science, 117*(23), 5681–5686. https://doi.org/10.1242/jcs.01489

Hamilton, W. B., & Brickman, J. M. (2014). Erk Signaling Suppresses Embryonic Stem Cell Self-Renewal to Specify Endoderm. *Cell Reports, 9*(6), 2056–2070. https://doi.org/10.1016/j.celrep.2014.11.032

Hamilton, W. B., Mosesson, Y., Monteiro, R. S., Emdal, K. B., Knudsen, T. E., Francavilla, C., Barkai, N., Olsen, J. V., & Brickman, J. M. (2019). Dynamic lineage priming is driven via direct enhancer regulation by ERK. *Nature, 575*(7782), 355–360. https://doi.org/10.1038/s41586-019-1732-z

Heery, D. M., Kalkhoven, E., Hoare, S., & Parker, M. G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature, 387*(6634), 733–736. https://doi.org/10.1038/42750

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., Wang, W., Weng, Z., Green, R. D., Crawford, G. E., & Ren, B. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics, 39*(3), 311–318. https://doi.org/10.1038/ng1966

Holstege, F. C. P., Fiedler, U., & Timmers, H. T. M. (1997). Three transitions in the RNA polymerase II transcription complex during initiation. *EMBO Journal, 16*(24), 7468–7480. https://doi.org/10.1093/emboj/16.24.7468

Housden, B. E., Muhar, M., Gemberling, M., Gersbach, C. A., Stainier, D. Y. R., Seydoux, G., Mohr, S. E., Zuber, J., & Perrimon, N. (2016). Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nature Reviews Genetics, 18*(1), 24–40. https://doi.org/10.1038/nrg.2016.118

Ito, M., Okano, H. J., Darnell, R. B., & Roeder, R. G. (2002). The TRAP100 component of the TRAP/Mediator complex is essential in broad transcriptional events and development.

EMBO Journal, *21*(13), 3464–3475. https://doi.org/10.1093/emboj/cdf348

Ito, M., Yuan, C. X., Malik, S., Gu, W., Fondell, J. D., Yamamura, S., Fu, Z. Y., Zhang, X., Qin, J., & Roeder, R. G. (1999). Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Molecular Cell*, *3*(3), 361–370. https://doi.org/10.1016/S1097-2765(00)80463-3

Jaeger, M. G., Schwalb, B., Mackowiak, S. D., Velychko, T., Hanzl, A., Imrichova, H., Brand, M., Agerer, B., Chorn, S., Nabet, B., Ferguson, F. M., Müller, A. C., Bergthaler, A., Gray, N. S., Bradner, J. E., Bock, C., Hnisz, D., Cramer, P., & Winter, G. E. (2020). Selective Mediator dependence of cell-type-specifying transcription. *Nature Genetics*, *52*(7), 719–727. https://doi.org/10.1038/s41588-020-0635-0

Kagey, M. H., Newman, J. J., Bilodeau, S., Zhan, Y., Orlando, D. A., Van Berkum, N. L., Ebmeier, C. C., Goossens, J., Rahl, P. B., Levine, S. S., Taatjes, D. J., Dekker, J., & Young, R. A. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, *467*(7314), 430–435. https://doi.org/10.1038/nature09380

Kelleher, R. J., Flanagan, P. M., & Kornberg, R. D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell*, *61*(7), 1209–1215. https://doi.org/10.1016/0092-8674(90)90685-8

Kim, M. O., Kim, S. H., Cho, Y. Y., Nadas, J., Jeong, C. H., Yao, K., Kim, D. J., Yu, D. H., Keum, Y. S., Lee, K. Y., Huang, Z., Bode, A. M., & Dong, Z. (2012). ERK1 and ERK2 regulate embryonic stem cell self-renewal through phosphorylation of Klf4. *Nature Structural and Molecular Biology*, *19*(3), 283–290. https://doi.org/10.1038/nsmb.2217

Komarnitsky, P., Cho, E. J., & Buratowski, S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes and Development*, *14*(19), 2452–2460. https://doi.org/10.1101/gad.824700

Koopman, P., & Cotton, R. G. H. (1984). A factor produced by feeder cells which inhibits embryonal carcinoma cell differentiation. Characterization and partial purification. *Experimental Cell Research*, *154*(1), 233–242. https://doi.org/10.1016/0014-4827(84)90683-9

Kornberg, R. D. (2005). Mediator and the mechanism of transcriptional activation. *Trends in Biochemical Sciences*, *30*(5), 235–239. https://doi.org/10.1016/j.tibs.2005.03.011

Korotkevich, E., Niwayama, R., Courtois, A., Friese, S., Berger, N., Buchholz, F., & Hiiragi, T. (2017). The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse Embryo. *Developmental Cell*, *40*(3), 235-247.e7. https://doi.org/10.1016/j.devcel.2017.01.006

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

Larue, L., Ohsugi, M., Hirchenhain, J., & Kemler, R. (1994). E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(17), 8263–8267. https://doi.org/10.1073/pnas.91.17.8263

Lehtonen, E. (1980). Changes in cell dimensions and intercellular contacts during cleavage-stage cell cycles in mouse embryonic cells. *Journal of Embryology and Experimental Morphology*, *Vol. 58*, 231–249.

Loh, Y. H., Wu, Q., Chew, J. L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K. Y., Sung, K. W., Lee, C. W. H., Zhao, X. D., Chiu, K. P., Lipovich, L., Kuznetsov, V. A., Robson, P., Stanton, L. W., … Ng, H. H. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics*, *38*(4), 431–440. https://doi.org/10.1038/ng1760

Lorthongpanich, C., Messerschmidt, D. M., Chan, S. W., Hong, W., Knowles, B. B., & Solter, D. (2013). Temporal reduction of LATS kinases in the early preimplantation embryo prevents ICM lineage differentiation. *Genes and Development*, *27*(13), 1441–1446. https://doi.org/10.1101/gad.219618.113

Lue, N. F., Flanagan, P. M., Kelleher, R. J., Edwards, A. M., & Kornberg, R. D. (1991). RNA polymerase II transcription in vitro. *Methods in Enzymology*, *194*(C), 545–550. https://doi.org/10.1016/0076-6879(91)94041-A

Luse, D. S., & Jacob, G. A. (1987). Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *The Journal of Biological Chemistry*, *262*(31), 14990–14997. https://doi.org/10.1016/s0021-9258(18)48127-6

Malik, S., & Roeder, R. G. (2000). Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends in Biochemical Sciences*, *25*(6), 277–283. https://doi.org/10.1016/S0968-0004(00)01596-6

Malik, S., & Roeder, R. G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends in Biochemical Sciences*, *30*(5), 256–263. https://doi.org/10.1016/j.tibs.2005.03.009

Malik, S., & Roeder, R. G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Reviews Genetics*, *11*(11), 761–772. https://doi.org/10.1038/nrg2901

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

Matsui, T., Segall, J., Weil, P. A., & Roeder, R. G. (1980). Multiple factors required for accurate initiation of transcription by purified RNA polymerase II. *Journal of Biological Chemistry*, *255*(24), 11992–11996. https://doi.org/10.1016/s0021-9258(19)70232-4

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., & Yamanaka, S. (2003). The homeoprotein nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, *113*(5), 631–642. https://doi.org/10.1016/S0092-8674(03)00393-3

Mohammadi, M., Froum, S., Hamby, J. M., Schroeder, M. C., Panek, R. L., Lu, G. H., Eiiseenkova, A. V., Green, D., Schlessinger, J., & Hubbard, S. R. (1998). Crystal structure of an angiogenesis inhibitor bound to the FGF receptor tyrosine kinase domain. *EMBO Journal*, *17*(20), 5896–5904. https://doi.org/10.1093/emboj/17.20.5896

Näär, A. M., Beaurang, P. A., Robinson, K. M., Oliner, J. D., Avizonis, D., Scheek, S., Zwicker, J., Kadonaga, J. T., & Tjian, R. (1998). Chromatin, TAFs, and a novel multiprotein coactivator are required for synergistic activation by Sp1 and SREBP-1a in vitro. *Genes and Development*, *12*(19), 3020–3031. https://doi.org/10.1101/gad.12.19.3020

Näär, A. M., Taatjes, D. J., Zhai, W., Nogales, E., & Tjian, R. (2002). Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes and Development*, *16*(11), 1339–1344. https://doi.org/10.1101/gad.987602

Niakan, K. K., Ji, H., Maehr, R., Vokes, S. A., Rodolfa, K. T., Sherwood, R. I., Yamaki, M., Dimos, J. T., Chen, A. E., Melton, D. A., McMahon, A. P., & Eggan, K. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes and Development*, *24*(3), 312–326. https://doi.org/10.1101/gad.1833510

Nichols, J., Silva, J., Roode, M., & Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development*, *136*(19), 3215–3222. https://doi.org/10.1242/dev.038893

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., & Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nature Methods*, *6*(12), 917–922. https://doi.org/10.1038/nmeth.1401

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., & Sasaki,

H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. *Mechanisms of Development*, *125*(3–4), 270–283. https://doi.org/10.1016/j.mod.2007.11.002

Niswander, L. E. E., & Martin, G. R. (1992). *<1992 Fgf-4 expression during.pdf>*. *768*, 755–768.

Niwa, H., Burdon, T., Chambers, I., & Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes and Development*, *12*(13), 2048–2060. https://doi.org/10.1101/gad.12.13.2048

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

Niwa, H., Ogawa, K., Shimosato, D., & Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature*, *460*(7251), 118–122. https://doi.org/10.1038/nature08113

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., & Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell*, *123*(5), 917–929. https://doi.org/10.1016/j.cell.2005.08.040

Nora, E. P., Goloborodko, A., Valton, A. L., Gibcus, J. H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L. A., & Bruneau, B. G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell*, *169*(5), 930-944.e22. https://doi.org/10.1016/j.cell.2017.05.004

Nutt, S. L., Dingwell, K. S., Holt, C. E., & Amaya, E. (2001). Xenopus sprouty2 inhibits FGF-mediated gastrulation movements but does not affect mesoderm induction and patterning. *Genes and Development*, *15*(9), 1152–1166. https://doi.org/10.1101/gad.191301

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oleå, A. K., Araúzo-Bravo, M. J., Saitou, M., Hadjantonakis, A. K., & Hiiragi, T. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nature Cell Biology*, *16*(1), 27–37. https://doi.org/10.1038/ncb2881

Pandey, P. K., Udayakumar, T. S., Lin, X., Sharma, D., Shapiro, P. S., & Fondell, J. D. (2005). *Activation of TRAP / Mediator Subunit TRAP220 / Med1 Is Regulated by Mitogen-Activated Protein Kinase-Dependent Phosphorylation*. *25*(24), 10695–10710. https://doi.org/10.1128/MCB.25.24.10695

Petrenko, N., Jin, Y., Wong, K. H., & Struhl, K. (2016). Mediator Undergoes a Compositional Change during Transcriptional Activation. *Molecular Cell*, *64*(3), 443–454. https://doi.org/10.1016/j.molcel.2016.09.015

Peyrieras, N., Hyafil, F., Louvard, D., Ploegh, H. L., & Jacob, F. (1983). Uvomorulin: A nonintegral membrane protein of early mouse embryo. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(20 I), 6274–6277. https://doi.org/10.1073/pnas.80.20.6274

Plaschka, C., Larivière, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E. V., Borchers, C. H., Baumeister, W., Herzog, F., Villa, E., & Cramer, P. (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. *Nature*, *518*(7539), 376–380. https://doi.org/10.1038/nature14229

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., & Hadjantonakis, A. K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development*, *135*(18), 3081–3091. https://doi.org/10.1242/dev.021519

Poehlmann, T. G., Fitzgerald, J. S., Meissner, A., Wengenmayer, T., Schleussner, E., Friedrich, K., & Markert, U. R. (2005). Trophoblast invasion: Tuning through LIF, signalling via Stat3. *Placenta*, *26*(SUPPL.). https://doi.org/10.1016/j.placenta.2005.01.007

Ptashne, M., Gann, A. (1997). *Transcriptional activation by recruitment. 386*, 569–577.

Rachez, C., Lemon, B. D., Suldan, Z., Bromleigh, V., Gamble, M., Näär, A. M., Erdjument-Bromage, H., Tempst, P., & Freedman, L. P. (1999). Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex. *Nature*, *398*(6730), 824–828. https://doi.org/10.1038/19783

Rao, S. S. P., Huang, S. C., Glenn St Hilaire, B., Engreitz, J. M., Perez, E. M., Kieffer-Kwon, K. R., Sanborn, A. L., Johnstone, S. E., Bascom, G. D., Bochkov, I. D., Huang, X., Shamim, M. S., Shin, J., Turner, D., Ye, Z., Omer, A. D., Robinson, J. T., Schlick, T., Bernstein, B. E., … Aiden, E. L. (2017). Cohesin Loss Eliminates All Loop Domains. *Cell*, *171*(2), 305-320.e24. https://doi.org/10.1016/j.cell.2017.09.026

Rappolee, D. A., Basilico, C., Patel, Y., & Werb, Z. (1994). Expression and function of FGF-4 in peri-implantation development in mouse embryos. *Development*, *120*(8), 2259–2269. https://doi.org/10.1242/dev.120.8.2259

Rougvie, A. E., & Lis, J. T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced hsp70 gene of D. melanogaster is transcriptionally engaged. *Cell*, *54*(6), 795–804. https://doi.org/10.1016/S0092-8674(88)91087-2

Roux, P. P., & Blenis, J. (2004). ERK and p38 MAPK-Activated Protein Kinases: a Family of Protein Kinases with Diverse Biological Functions. *Microbiology and Molecular Biology Reviews*, *68*(2), 320–344. https://doi.org/10.1128/mmbr.68.2.320-344.2004

Sathyan, K. M., McKenna, B. D., Anderson, W. D., Duarte, F. M., Core, L., & Guertin, M. J.

(2019). An improved auxin-inducible degron system preserves native protein levels and enables rapid and specific protein depletion. *Genes and Development*, *33*(19–20), 1441–1455. https://doi.org/10.1101/gad.328237.119

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., & Brivanlou, A. H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nature Medicine*, *10*(1), 55–63. https://doi.org/10.1038/nm979

Sawadogo, M., & Roeder, R. G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. *Cell*, *43*(1), 165–175. https://doi.org/10.1016/0092-8674(85)90021-2

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

Shlyueva, D., Stampfel, G., & Stark, A. (2014). Transcriptional enhancers: From properties to genome-wide predictions. *Nature Reviews Genetics*, *15*(4), 272–286. https://doi.org/10.1038/nrg3682

Singh, A. M., Hamazaki, T., Hankowski, K. E., & Terada, N. (2007). A Heterogeneous Expression Pattern for Nanog in Embryonic Stem Cells. *Stem Cells*, *25*(10), 2534–2542. https://doi.org/10.1634/stemcells.2007-0126

Slack, C., Alic, N., Foley, A., Cabecinha, M., Hoddinott, M. P., & Partridge, L. (2015). The Ras-Erk-ETS-Signaling Pathway Is a Drug Target for Longevity. *Cell*, *162*(1), 72–83. https://doi.org/10.1016/j.cell.2015.06.023

Smith, A. G., & Hooper, M. L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. *Developmental Biology*, *121*(1), 1–9. https://doi.org/10.1016/0012-1606(87)90132-1

Smith, T. A., & Hooper, M. L. (1983). Medium conditioned by feeder cells inhibits the differentiation of embryonal carcinoma cultures. *Experimental Cell Research*, *145*(2), 458–462. https://doi.org/10.1016/0014-4827(83)90025-3

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

Stauber, D. J., DiGabriele, A. D., & Hendrickson, W. A. (2000). Structural interactions of fibroblast growth factor receptor with its ligands. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(1), 49–54. https://doi.org/10.1073/pnas.97.1.49

Stevens, J. L., Cantin, G. T., Wang, G., Shevchenko, A., Shevchenko, A., & Berk, A. J. (2002a).

Transcription control by E1A and MAP kinase pathway via Sur2 Mediator subunit. *Science*, *296*(5568), 755–758. https://doi.org/10.1126/science.1068943

Stevens, J. L., Cantin, G. T., Wang, G., Shevchenko, A., Shevchenko, A., & Berk, A. J. (2002b). Transcription Control by E1A and MAP Kinase Pathway via Sur2 Mediator Subunit. *Science*, *296*(5568), 755 LP – 758. https://doi.org/10.1126/science.1068943

Strobl, L. J., & Eick, D. (1992). Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-myc in vivo. *EMBO Journal*, *11*(9), 3307–3314. https://doi.org/10.1002/j.1460-2075.1992.tb05409.x

Stumpf, M., Waskow, C., Krötschel, M., Van Essen, D., Rodriguez, P., Zhang, X., Guyot, B., Roeder, R. G., & Borggrefe, T. (2006). The mediator complex functions as a coactivator for GATA-1 in erythropoiesis via subunit Med1/TRAP220. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18504–18509. https://doi.org/10.1073/pnas.0604494103

Taatjes, D. J., Näär, A. M., Andel, F., Nogales, E., & Tjian, R. (2002). Structure, function, and activator-induced conformations of the CRSP coactivator. *Science*, *295*(5557), 1058–1062. https://doi.org/10.1126/science.1065249

Taatjes, D. J., Schneider-Poetsch, T., & Tjian, R. (2004). Distinct conformational states of nuclear receptor-bound CRSP-Med complexes. *Nature Structural and Molecular Biology*, *11*(7), 664–671. https://doi.org/10.1038/nsmb789

Tang, W.-S., Weng, L., Wang, X., Liu, C.-Q., Hu, G.-S., Yin, S.-T., Tao, Y., Hong, N.-N., Guo, H., Liu, W., Wang, H.-R., & Zhao, T.-J. (2021). The Mediator subunit MED20 organizes the early adipogenic complex to promote development of adipose tissues and diet-induced obesity. *Cell Reports*, *36*(1), 109314. https://doi.org/10.1016/j.celrep.2021.109314

Tee, W. W., Shen, S. S., Oksuz, O., Narendra, V., & Reinberg, D. (2014). Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. *Cell*, *156*(4), 678–690. https://doi.org/10.1016/j.cell.2014.01.009

Thisse, B., & Thisse, C. (2005). Functions and regulations of fibroblast growth factor signaling during embryonic development. *Developmental Biology*, *287*(2), 390–402. https://doi.org/10.1016/j.ydbio.2005.09.011

Thompson, C. M., & Young, R. A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(10), 4587–4590. https://doi.org/10.1073/pnas.92.10.4587

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

Tudor, M., Murray, P. J., Onufryk, C., Jaenisch, R., & Young, R. A. (1999). Ubiquitous expression and embryonic requirement for RNA polymerase II coactivator subunit Srb7 in mice. *Genes and Development*, *13*(18), 2365–2368. https://doi.org/10.1101/gad.13.18.2365

Uhlik, M. T., Abell, A. N., Cuevas, B. D., Nakamura, K., & Johnson, G. L. (2004). Wiring diagrams of MAPK regulation by MEKK1, 2, and 3. *Biochemistry and Cell Biology*, *82*(6), 658–663. https://doi.org/10.1139/o04-114

Vernimmen, D., & Bickmore, W. A. (2015). The Hierarchy of Transcriptional Activation: From Enhancer to Promoter. *Trends in Genetics*, *31*(12), 696–708. https://doi.org/10.1016/j.tig.2015.10.004

Wang, G., Cantin, G. T., Stevens, J. L., & Berk, A. J. (2001). Characterization of Mediator Complexes from HeLa Cell Nuclear Extract. *Molecular and Cellular Biology*, *21*(14), 4604–4613. https://doi.org/10.1128/mcb.21.14.4604-4613.2001

Wang, W., Carey, M., & Gralla, J. a Y. D. (1992). Polymerase II Promoter Activation : Closed Complex. *Science*, *255*(5043), 450–453.

Westerling, T., Kuuluvainen, E., & Mäkelä, T. P. (2007). Cdk8 Is Essential for Preimplantation Mouse Development. *Molecular and Cellular Biology*, *27*(17), 6177–6182. https://doi.org/10.1128/mcb.01302-06

Wilder, P. J., Kelly, D., Brigman, K., Peterson, C. L., Nowling, T., Gao, Q. S., McComb, R. D., Capecchi, M. R., & Rizzino, A. (1997). Inactivation of the FGF-4 gene in embryonic stem cells alters the growth and/or the survival of their early differentiated progeny. *Developmental Biology*, *192*(2), 614–629. https://doi.org/10.1006/dbio.1997.8777

Williams, R. L., Hiltont, D. J., Pease, S., Willsont, T. A., Stewart, C. L., Gearingt, D. P., Wagner, E. F., Metcalft, D., Nicolat, N. A., & Gought, N. M. (1988). *336684a0*. 6–9.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., & Smith, A. (2012). Erratum: Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation (Nature Cell Biology (2011) 13 (838-845)). *Nature Cell Biology*, *14*(5), 555. https://doi.org/10.1038/ncb0512-555a

Yamanaka, Y., Lanner, F., & Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development*, *137*(5), 715–724. https://doi.org/10.1242/dev.043471

Yesbolatova, A., Saito, Y., Kitamoto, N., Makino-Itou, H., Ajima, R., Nakano, R., Nakaoka, H., Fukui, K., Gamo, K., Tominari, Y., Takeuchi, H., Saga, Y., Hayashi, K. ichiro, &

Kanemaki, M. T. (2020). The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-19532-z

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

Zhao, H., Young, N., Kalchschmidt, J., Lieberman, J., Khattabi, L. El, Casellas, R., & Asturias, F. J. (2020). Structure of the Mammalian Mediator. *BioRxiv*, 2020.10.05.326918. https://doi.org/10.1101/2020.10.05.326918

Zhu, W., Yao, X., Liang, Y., Liang, D., Song, L., Jing, N., Li, J., & Wang, G. (2015). Mediator med23 deficiency enhances neural differentiation of murine embryonic stem cells through modulating bmp signaling. *Development (Cambridge)*, *142*(3), 465–476. https://doi.org/10.1242/dev.112946