UNIVERSITY OF NAPLES FEDERICO II

Doctorate School in Biology

Ph.D. COURSE IN BIOLOGY

XXXII CYCLE

2017-2020

Characterization of an early-primed state of pluripotency competent to gastruloid formation and primordial germ cells induction

Tutor
Gabriella Minchiotti

Candidate
Federica Cermola

Coordinator
Prof. Salvatore Cozzolino

INDEX

SUMMARY

Mouse embryonic stem cells (mESCs) exposed to a high L-Proline regimen undergo a phenotypic transition named embryonic stem-to-mesenchymal transition (esMT). The resulting cells, named L-Proline-induced (PiCs), maintain some features of naïve pluripotent cells (LIF reliance, teratoma and chimera formation), and acquire other (morphology, energetic metabolism, and epigenetic profile) specific of primed pluripotent cells. These data suggest that PiCs represent an early-primed pluripotency state [1-3]. To further characterize the pluripotency state of PiCs, we set up a high-performing assay to measure the gastruloids formation efficiency (GFE) of mESCs, *i.e.* the ability of spherical aggregates of mESCs to develop into elongated/polarized embryonic organoids [4]. By exploiting this innovative experimental approach we provide morphological and molecular evidences that gastruloid development relies on *Cripto* gene, a key regulator of stem cells pluripotency [5, 6]. We also found that GFE assay discriminate different phenotypic/functional states of pluripotency. Indeed, while *naïve* cells efficiently give rise to aggregates, most of which elongate (GFE ≥ 95%), *primed* Epiblast stem cells (EpiSCs) fail to aggregate, and consequently to generate gastruloids (GFE = 0%). Moreover, although early-primed Epiblast-like cells (EpiLCs) efficiently aggregate, the resulting aggregates remain as undeveloped organoids (GFE = 0%). PiCs generate cell aggregates that elongate earlier and develop into smaller gastruloids, exhibiting highly differentiated areas (GFE ≥ 50%). Moreover, like EpiLCs, PiCs are competent to differentiate into primordial germ cell-like cells (PGCLCs). These results characterize PiCs as a unique pluripotency state, with competence for both gastruloid formation and PGCLCs differentiation.

How L-Pro impact ESC identity is not yet fully understood. It is known that L-Pro supplementation increases the biosynthesis of L-Pro-rich protein such as collagen [2, 7], and that Collagen prolyl hydroxylation (CPH), catalysed by prolyl 4-hydroxylase (P4H), requires Vitamin C (VitC). Here, we identify CPH as an epigenetic modulator. The induction of collagen synthesis, and thus of VitC-dependent CPH, increases global DNA/histone methylation level and promote cell state transition. Interfering with CPH, by either genetic ablation of P4H alpha subunits or pharmacologic treatment, avoid both epigenetic changes and cell state transition. Mechanistically, these results suggest that a

sudden increment in CPH can modify the epigenetic landscape by reducing VitC availability for DNA and histone demethylases enzymes. Our study provides mechanistic insights into how metabolic cues and epigenetic factors integrate to control cell state transition.

SOMMARIO

L'amminoacido non essenziale L-Prolina (L-Pro) è in grado di influenzare la metastabilità delle cellule staminali embrionali murine (ESCs) che, se esposte a un'alta disponibilità di L-Pro, vanno incontro a una transizione denominata *embryonic stem-to-mesenchymal transition* (esMT), dove le cellule derivanti, definite *L-Proline-induced cells (PiCs)*, acquisiscono caratteristiche (metaboliche, epigenetiche , ecc.) di uno stadio *primed* di pluripotenza [1-3]. Inoltre, i test standard di pluripotenza *in vivo* (formazione di teratomi e di chimere) suggeriscono che le PiCs presentano caratteristiche di cellule pluripotenti *naïve* [1]. Tali dati avvalorano l'ipotesi che l'aumento della disponibilità di L-Pro possa catturare le ESCs in uno stadio intermedio di pluripotenza. In questo contesto, lo sviluppo di organoidi embrionali polarizzati/elongati derivanti da ESCs, chiamati gastruloidi [4], può essere sfruttato come un potente strumento per valutare e discriminare diversi stadi di pluripotenza e, in particolare, per definire in maniera più approfondita la pluripotenza delle PiCs. In linea con questa idea abbiamo modificato il protocollo pubblicato [8] per misurare l'efficienza di formazione di gastruloidi (GFE), cioè la frazione di aggregomi che sviluppano un gastruloide. Sfruttando questa procedura, i nostri risultati forniscono prove morfologiche e molecolari che lo sviluppo del gastruloide dipende dall'espressione di *Cripto*, regolatore chiave della pluripotenza delle cellule staminali [5, 6]. Abbiamo dimostrato inoltre che la GFE diminuisce man mano che la pluripotenza delle cellule passa dallo stato *naïve* a quello *primed*. Infatti, aggregati derivati da ESC cresciute in condizioni *naïve* elongano in maniera efficiente (GFE ≥ 95%), mentre le *primed Epiblast stem cells* (EpiSCs) non riescono ad aggregarsi correttamente e di conseguenza a generare gastruloidi (GFE = 0%). Al contrario, le *early-primed Epiblast-like cells* (EpiLC) si aggregano correttamente, ma gli aggregomi derivati da esse sono per lo più abortivi (GFE = 0%). A differenza delle EpiLC, le PiCs generano gastruloidi (GFE ≥ 50%), che tuttavia iniziano ad allungarsi anticipatamente rispetto alle ESCs *naïve* e sono di dimensioni più piccole. Inoltre, come le EpiLCs, le PiCs sono competenti a differenziare in cellule germinali primordiali (PGCs), suggerendo che le PiCs catturano uno stadio di pluripotenza simile alle EpiLCs ma con la peculiare competenza di formare sia gastruloidi che differenziare in PGCs. Pertanto, i risultati di tale lavoro di tesi propongono GFE come un metodo *in vitro*

semplice e robusto capace di discriminare diversi stati fenotipici/funzionali del *continuum* della pluripotenza, ricapitolando l'andamento della formazione delle chimere.

Il meccanismo molecolare mediante il quale L-Pro cambia l'identità delle ESCs non è ancora completamente noto. È stato recentemente pubblicato che L-Pro aumenta la biosintesi di proteine, in particolare quelle ricche in residui di L-Pro come ad esempio il collagene [2, 7]. L'idrossilazione dei residui di prolina del collagene (CPH), che è catalizzata dalla prolil 4-idrossilasi (P4H), è la modifica post-traduzionale più diffusa nell'uomo e richiede vitamina C (VitC). Nel presente lavoro di tesi è stato identificato che CPH agisce come un modulatore epigenetico della plasticità cellulare. L'aumento della CPH induce la metilazione globale di DNA/istone nelle cellule staminali pluripotenti e favorisce la transizione indotta da prolina. Bloccando la CPH o mediante ablazione genetica delle subunità alfa di P4H o tramite trattamento farmacologico si ripristinano i cambiamenti epigenetici e si antagonizza la transizione. Meccanicisticamente, i risultati ottenuti suggeriscono che la CPH modifica il paesaggio epigenetico riducendo la VitC per il DNA e le demetilasi degli istoni. Il nostro studio fornisce approfondimenti meccanicistici su come i segnali metabolici e i fattori epigenetici si integrano per controllare la metastabilità delle cellule staminali embrionali.

1. INTRODUCTION

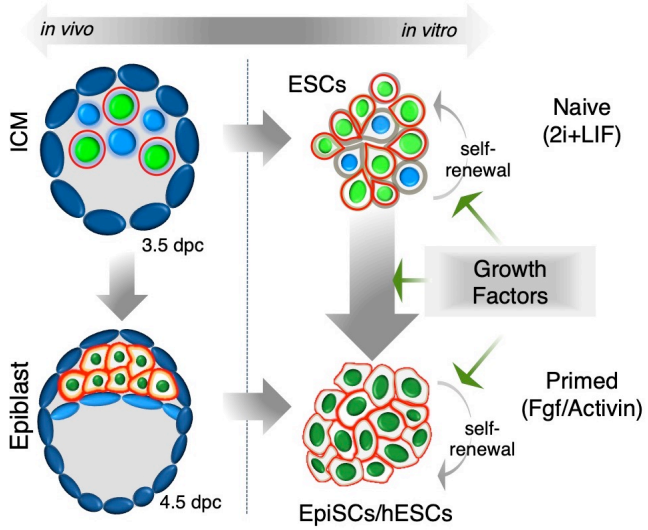
1.1. Stem cells pluripotency

Pluripotency is the capability of stem cells to differentiate and give rise to all somatic cell lineages as well as germ cells. Pluripotent stem cells are present from day (E) 3.5 to 8.0 of mouse development. Before implantation the inner cell mass (ICM) of a blastocyst generates the epiblast and the primitive endoderm. The epiblast is the precursor of all somatic cells and germ cells, while the primitive endoderm is the precursor of visceral and parietal yolk sacs. Epiblast stem cells progress from a ground/naïve state of pluripotency (pre-implantation epiblast) to a primed state of pluripotency (post-implantation epiblast) [9].

1.1.1. Naïve and primed pluripotency states

The *in vitro* counterpart of naïve and primed pluripotency states are the embryonic stem cells (ESCs), and the epiblast stem cells (EpiSCs), which derive from the inner cell mass (ICM) of the pre-implantation embryo [10, 11] and the epiblast of the post-implantation embryo [12, 13], respectively (**Fig. 1.1**). Naïve ESCs can be cultured *in vitro* in the so called 2i medium supplemented with two chemical inhibitors (PD0325901/MEK and CHIR99021/GSK3) and in the presence of the leukemia inhibitor factor (LIF) [14]. Under naïve-inducing conditions (2i+LIF): a) pluripotency genes like *Nanog* are expressed homogeneously in the ESC population, while their expression is heterogeneous and metastable in serum/LIF ESCs. Conversely, the expression of genes involved in stem cells differentiation is consistently reduced in 2i compared to serum/LIF ESCs; b) the cells grow as round-shaped domed colonies and show transcriptional and epigenetic profiles that resemble those of the naïve cells of the pre-implantation epiblast, including a low level of DNA methylation [15-18]; c) ESCs are characterized by a bivalent energetic metabolism, *i.e.* they obtain energy/ATP from both glycolysis and oxidative phosphorylation; d) ESCs have the potential to generate cells from the three germ layers of the embryo giving rise to teratomas *in vivo*, and can colonize the ICM of the blastocyst to generate chimeric animals [19]. Primed EpiSCs can be isolated from the embryo at E4.5-E5.5 [12] (**Fig. 1.1**) and differ from ESCs with respect to colony morphology, gene expression profile and growth factors requirements.

For instance, EpiSCs grow generating irregular flat-shaped cell colonies and requires *Fibroblast growth factor 2* or *basic* FGF (Fgf2 or bFGF), and Activin A to be isolated and maintained [12, 13].



Adapted from: Fiorenzano et al., Nature Communications vol. 7, 2016

Figure 1.1: Schematic representation of naïve and primed cells. During mouse development pluripotent cells progress from a naïve state in the Inner cell mass (ICM) of the pre-implantation blastocyst 3.5 *day post coitum* (3.5 dpc) to a primed state Epiblast of the embryo at 4.5-5.5 dpc. The *in vitro* counterparts of these two stages are the ESCs grown in 2i+LIF and EpiSCs in Fgf/Activin, respectively.

EpiSCs are a heterogeneous population of cells showing a transcriptome similar to the gastrula-stage epiblast [20]. Specifically, they are marked by the expression of post-implantation genes, and show female X chromosome inactivation, glycolytic metabolism and elevated DNA methylation [12, 13, 21]. EpiSCs express the pluripotent factors *Oct4* and *Sox2* but also some early-gastrulation markers like *Fgf5, Brachyury* (*T*)*, Sox17* and *Gata6* [22]. Contrarily to ESCs, EpiSCs are unable to colonize the ICM, and thus fail to form chimeric embryo if injected into a blastocyst [9].

Naïve (ESCs) and primed (EpiSCs) states are considered the two extremes of a *continuum* of pluripotency states, *i.e.* a gradual sequence of pluripotent states with distinct functional and transcriptional signatures. Of note, some intermediate states have been characterized in vitro, such as the Epiblast-like cells (EpiLCs), a transient

pluripotent state that is able to generate primordial germ cell-like cells (PGCLCs), a feature lost in both ESCs and EpiSCs [23].

1.1.2. Pluripotency *continuum* and "formative" state

In the last years it has been hypothesized the existence of a *continuum* of different states of pluripotency between naïve and primed. This hypothesis comes from the observation that during the transition from naïve to primed, the cells lose the naïve state's features and gradually acquire characters of primed state. Different intermediate states have been described [9]. The *Epiblast-like cells* (EpiLCs) represent a transient early primed state of pluripotency between naïve and primed that cannot be captured *in vitro*. EpiLCs are generated from naïve ESCs treated with bFgf and Activin (F/A) for 48h [23]. This resembles the time window in which the pre-gastrula epiblast cells acquire the competence for germ line induction [23]. In EpiLCs some naïve genes including *Stella*, *Rex1* and *Klf4* are downregulated, while some primed markers including *Fgf5* and *Oct6*, are induced at the same or higher levels than in primed EpiSCs. Combining naïve and primed culture conditions other pluripotent states, showing several features of both ESCs and EpiSCs have been captured *in vitro*. For instance, intermediate pluripotent stem cells (INTPSCs) [24] can be captured by culturing ESCs in a mix of F/A and CHIR, show a domed shape colony morphology and express both naïve markers such as *Klf4*, *Rex1* and *Esrrb* and some primed markers, including *Fgf5*, *Wnt3* and *Otx2*. Moreover, INTPSCs show X chromosome activation and when injected into a blastocyst they contribute to embryonic development [24]. Furthermore, FAB-SCs [bFGF, Activin, and BIO (Wnt/β-catenin signaling agonist)-derived stem cells] can be obtained in similar culture conditions as INTPSCs but they are derived from pre-implantation embryos [25]. FAB-SCs express pluripotency markers such as *Oct4*, *Nanog* and *Sox2* but are unable to differentiate *in vitro*, and *in vivo* (teratomas) as well as to contribute to embryonic development after blastocyst injection. Interestingly, when FAB-SCs are stimulated with LIF and BMP4, they acquire the ability to generate teratomas and give rise to chimeric mice [25]. Early primitive ectoderm-like (EPL) cells are characterized by primitive ectoderm-like gene expression (*e.g. Fgf5*), express pluripotency markers such as *Oct4,* SSEA1 and alkaline phosphatase, but are unable to contribute to chimeric mice when injected into a blastocyst [26]. Intermediate ESCs

(IESCs) express intermediate levels of naïve markers (*Rex1*, *Stella*, *Kfl4*, *Nanog*) and also primed markers of (*Fgf5*, *Nodal*, *Lefty*, *FoxA2*, *Otx2*, *Gata6*, *Brachyury*) [27]. IESCs grow in serum + Activin medium generating domed-shaped cell colonies. Similarly to EPL cells, IESCs are unable to contribute to embryonic development when injected into a blastocyst [27].

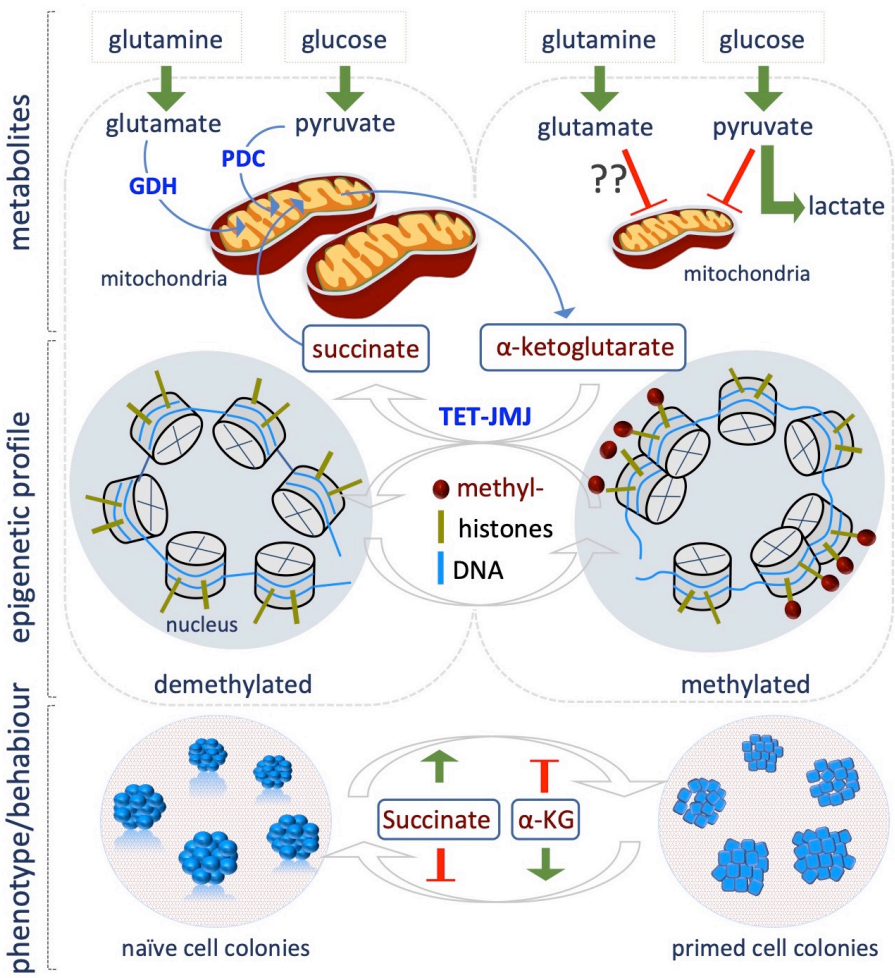
All these intermediate states are maintained in culture using growth factors (*e.g.* bFgf; Activin etc.) and/or in the presence of chemical inhibitors/small molecules (*e.g.* CHIR99021, PD0325901; etc). More recently, a "formative" state of pluripotency, obligatory for the genetic and epigenetic remodeling that the naïve cells must experience in order to become prone for lineage specification, have been hypothesized and described [19, 28]. In term of genes expression, the cells in the formative state should display an intermediate profile between naïve and primed, and should also be able to either acquire primed features or revert spontaneously to the naïve state and competent to PGC differentiation [29]. So far, pluripotent cells showing the features of this formative state of pluripotency have not been captured *in vitro*.

1.1.3. Plasticity/metastability of pluripotent stem cells

In vitro cultured pluripotent stem cells show a high phenotypic heterogeneity, *i.e.* are continuously fluctuating between different pluripotency states [30]. The plasticity/metastability of stem cells depends on the combination of extrinsic and intrinsic factors that influence the expression of key regulatory genes [30]. Indeed, different pluripotent genes, including transcription factors coding genes, display a fluctuating expression [31, 32]. One of these genes is *Stella* (*Dppa3*), which is expressed in the pre-implantation embryo, repressed in the epiblast and reactivated in primordial germ cells (PGCs). When cultured in standard conditions, *i.e.* DMEM, 15% fetal bovine serum (FBS) and LIF, some ESCs are Stella-positive (Stella+), while others are Stella-negative (Stella-), and this influences the differentiation potential of the cells. Pluripotency-related genes such as *Rex1*, *Klf4, SSEA1*, and *Nanog* also show variable levels of expression in cultured stem cells. Of note, ESCs that express low level of *Nanog* are more prone to differentiate and do not contribute to chimeric embryos [33].

1.2. Metabolites as regulators of stem cells plasticity

In the last years an increasing interest is emerging towards the role of metabolites in the control of pluripotency, metastability and differentiation of stem cells. Among metabolites a key role is played by those acting as epigenetic modifiers [34], such as S-adenosylmethionine (SAM), acetyl coenzyme A (acetyl-CoA) and α-ketoglutarate (αKG) [35]. These metabolites act as a substrate of enzymes that modify the chromatin structure and its methylation state, influences the pluripotency of stem cells and the reprogramming/de-differentiation efficiency of somatic cells (**Fig 1.2)** [36].



From D'Aniello et al., Epigenomes 2019

Figure 1.2 Alpha-ketoglutarate (αKG)/ succinate ratio modulates the naïve to primed transition. αKG abundance in the naïve state depends on glutamate, through the glutamate dehydrogenase (GDH), and on glucose-derived pyruvate. Pyruvate is converted into acetyl-coA by the pyruvate dehydrogenase complex (PDC), and enters the Krebs cycle in the mitochondria, generating citrate, isocitrate and then αKG. High levels of αKG, a key cofactor of the DNA and histone demethylases (TET and JMJ) guarantee DNA and histone demethylation, promoting a naïve state in mouse ESCs. Succinate, which is produced from αKG oxidation, pushes cells towards differentiation [35].

Another key metabolite is the Ascorbic acid (Vitamin C, VitC), which is required for the activity of α-KG/Fe2+-dependent dioxygenases, including the DNA and histone demethylases, and is emerging as a key regulator of pluripotency and differentiation [3, 7, 37, 38]. Other metabolites, including fatty acids and amino acids (AA) can influence cell differentiation and reprogramming [39, 40]. For instance, both ESCs and induced pluripotent stem cells (iPSC) are strictly dependent on L-Threonine and L-Methionine [41, 42]. L-Methionine, through its conversion into S-AdenosylMethionine (SAM), is essential for maintaining undifferentiated ESCs. In particular, the methylation level of lysine 4 of histone H3 [43] and the expression of Nanog rely on the availability L-Methionine [42]. Most recently, also the non-essential amino acid L-Proline (L-Pro) has emerged as a potent regulator of stem cells plasticity [35].

1.2.1. The non-essential amino acid L-Proline

(S)-Pyrrolidine-2-carboxylic acid, known as L-Proline (L-Pro) (CAS Number: 147-85-3, EC Number: 205-702-2, CHEBI:17203, HMDB0000162), is a cyclic, nonpolar, non-toxic, odorless, and sweet amino acid (**Fig. 1.3**) [44]. L-Pro is a non-essential amino acid (NEAA) and it is synthetized from Glutamate. In a first step, the Aldehyde Dehydrogenase 18 family member A1 (Aldh18a1) enzyme catalyzes the conversion of Glutamate into Pyrroline-5-carboxylate (P5C), subsequently the Pyrroline-5-carboxylate Reductase 1 (Pycr1) enzyme catalyzes the conversion of P5C into L-Pro. Of note, L-Pro is catabolized in the mitochondria through the activity of Prodh enzyme. L-Pro has numerous physicochemical proprieties. For instance, it enables a wide range of asymmetric chemical reactions by acting as an enantioselective organocatalyst. L-Pro acts as a chemical chaperone increasing proteins solubility and preventing aggregation/fibrillation, and indeed it is used to stabilize the monoclonal antibodies. Acting as an antioxidant, L-Pro, inhibits the corrosion/oxidation of the iron in steel treated with aggressive acidic solutions [45]. L-Pro is also a key building block for extracellular peptides/proteins, from antimicrobial peptides to salivary proteins and collagens. L-Pro is an efficient source of energy/ATP and carbon skeletons, able to sustain the growth/infectivity of pathogen microorganisms (bacteria, fungi, parasites) and metastatic cancer cells.

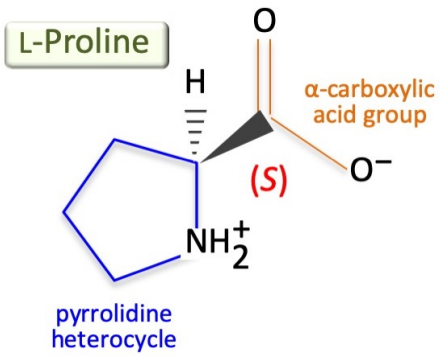


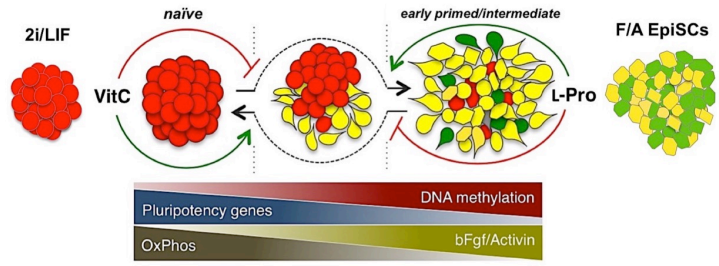
Figure 1.3: Molecular structure of the non-essential amino acid L-Proline. L-Proline is a nonpolar cyclic imino acid. Its structure contains an α-amino group, an α-carboxylic acid group, and a side chain pyrrolidine.

1.2.2. L-Proline is a modulator of stem cells pluripotency/identity

The availability of L-Pro influences ESC proliferation, morphology, motility, gene expression profile and epigenetic landscape [1-3]. L-Pro supplementation increases ESC proliferation and induces a morphological/molecular transition defined as embryonic-stem-cell to mesenchymal-like transition (esMT). Indeed, while ESCs normally grow generating domed/round-shaped colonies, L-Pro-induced cells (PiCs) grow as irregular flat-shaped colonies. This phenotypic transition is a complex process that occurs in 5 days, and is induced in ESCs growing in a complete medium (DMEM, fetal bovine serum_FBS and LIF) after addition of L-Pro at a physiological concentration (150-250 μM) [2]. Of note, esMT is fully reversible by removing L-Pro or adding VitC [2]. Several evidences indicate that exogenous provided L-Pro induces an early-primed state of pluripotency. Indeed, other than a different morphology, ESCs acquire epigenetic and metabolic features of an early-primed state, similar to that of Epiblast-like cells (EpiLCs) [3]. Of note, PiCs acquire mesenchymal features (motile and invasive), but maintain the expression of pluripotent genes such as *Oct4* and *Sox2*. Similarly to what observed in EpiSCs, the expression levels of ICM markers such as *Nanog*, *Rex1* and *Gbx2* is reduced whereas the expression of specific epiblast markers such as *Fgf5* and *Fut4*, is induced in PiCs [2]. PiCs are pluripotent stem cells, since they retain a high differentiation potential both *in vitro* and *in vivo* (**Fig. 1.4**). Notably, PiCs are able to differentiate into cardiomyocytes and neurons i*n vitro*, and to generate teratomas upon injection into immunocompromised mice; furthermore, they contribute to development of chimeric embryos after injected into a blastocyst [1]. A high L-Pro regimen also

induces a metabolic reprogramming in ESCs from bivalent (oxidative phosphorylation/ glycolysis) to glycolysis [3]. This resembles the metabolic reprogramming observed during the naïve to primed transition [46]. The transition from naïve to primed is accompanied by changes in the expression of *mir-290* and *mir-302* clusters. Specifically, while *mir-290* marks the naïve state (pre-implantation epiblast), *mir-302* marks the primed state (late post-implantation epiblast of the pre-gastrulating embryo); co-expression of the two clusters identifies an intermediate (early post-implantation epiblast) state [47].

A



B

Pluripotency-associated properties	ESCs (naïve)	PiCs (early primed)	EpiSCs (primed)
Colony morphology	domed	flat	flat
Global DNA hypomethylation	Yes	No	No
Pluripotency markers	↑	↓	↓↓
Priming markers	↓	↑	↑↑
Adhesion molecules	E-Cadherin	N-Cadherin	N-Cadherin
Sensibility to Trypsin	No	Yes	Yes
Reversibility to ESCs	--	High	Low
MEK_ERK dependence	No	? (yes generation)	Yes
Tgfβ-ActivinA dependence	No	? (yes generation)	Yes
Metabolism	OxPhos/Glycolytic	Glycolytic	Glycolytic
Contribution to chimeric embryos	High	High	Low
LIF dependent	Yes	Yes	No

Adapted from: D'Aniello C. et al., SCR, 2017

Figure 1.4: L-Proline availability influences the pluripotency state of ESCs.
(A) Schematic representation of L-Proline activity on ESCs. L-Proline supplementation in serum/LIF culture conditions induces a phenotypic switch and modifies the epigenetic, transcriptional and metabolic features of ESCs.
(B) Comparison of the molecular and functional features of ESCs, PiCs and EpiSCs.

PiCs show a *mir-290*$^{\pm}$/*mir-302*$^{+}$ signature similar to that of early post-implantation epiblast [3]. These data suggest that PiCs are an intermediate state of pluripotency with some features of the theorized "formative" pluripotency state (**Fig. 1.4**).

1.2.3. L-Proline induces a stem cell-to-mesenchymal-like transition (esMT)

The phenotypic transition induced by L-Pro (esMT) show key features of the epithelial-mesenchymal transition (EMT). The morphological transition starts to be evident 3 days after L-Pro supplementation, with the cells that acquire motile and invasive features (**Fig. 1.5**). In the absence of exogenously supplied L-Pro, ESCs develop compacted colonies, and the cells maintain strong cell-cell interaction. Upon L-Pro supplementation, the expression of genes related to cell shape, adhesion and motility is induced and the cells tend to lose the cell-cell interactions. Specifically, a high L-Pro regimen favors cell-substrate adhesions and cell migration and destabilizes cell-cell adhesions. At molecular level, L-Pro induces the expression of *N-cadherin*, *Vimentin*, and *Brachyury* (*T*) [2]. N-cadherin transcripts are detected from day 2 onward during the esMT, while *T* transcripts are induced at day 3, in a time-window that correlates with the onset of the motile phenotype. Of note, while N-cadherin is induced at both RNA and protein levels, E-cadherin is expressed at comparable levels in ESCs and PiCs. Accordingly, the expression of the E-cadherin transcriptional repressors, Snail and Slug, are not significantly altered in PiCs. However, E-cadherin protein becomes delocalized from the cell membrane to the Golgi in PiCs. PiCs acquire invasive and metastatic properties *in vivo*. Specifically, unlike ESCs, PiCs generate lung metastasis when injected intravenously into the tail bud of immunodeficient mice [2] (**Fig. 1.5**). Remarkably, esMT is fully reversible, either after L-Pro withdrawal or by addition of ascorbic acid (Vitamin C, VitC) [1, 2], raising the hypothesis that these two metabolites play antagonistic roles in controlling ESC identity. Besides VitC, supplementation of other antioxidants, such as N-Acetyl Cysteine (NAC) and Glutathione, are unable to antagonize L-Pro effects (proliferation, esMT), suggesting an anti-oxidant independent activity of VitC in counteracting L-Pro effects

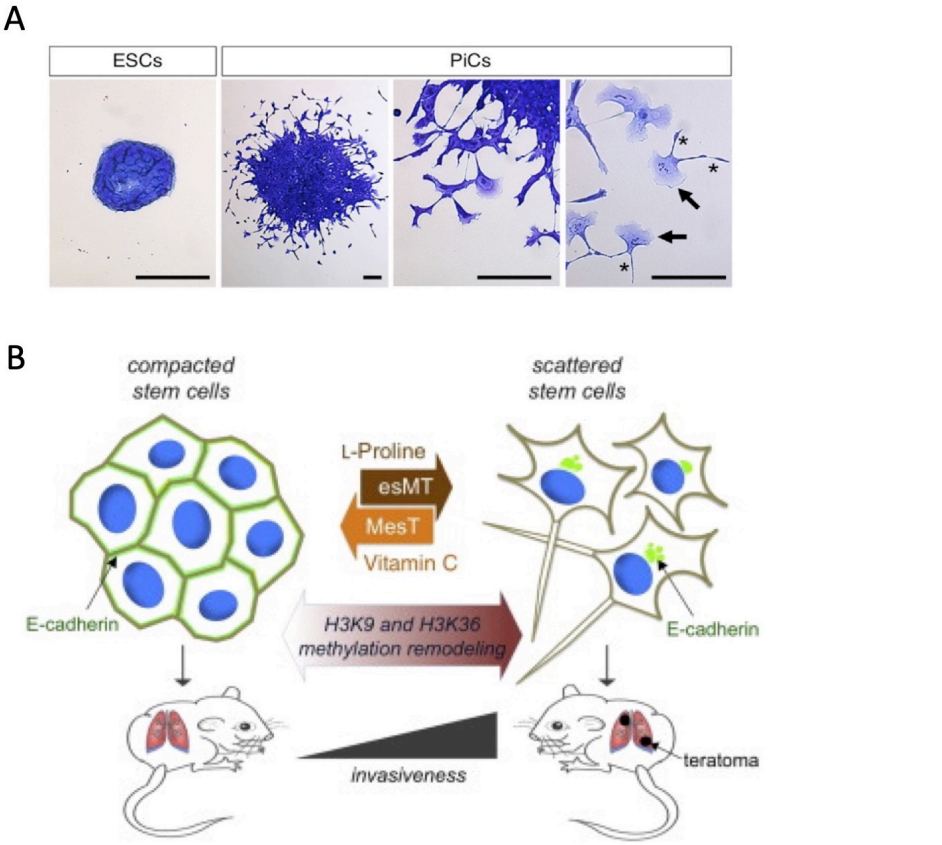


Figure 1.5: Embryonic stem cell to mesenchymal transition induced by L-Pro supplementation.
(A) Representative pictures of ESCs and PiCs colonies at day 5 of treatment. Cell colonies were fixed and stained with a mix of Glutaraldehyde and Crystal Violet. Black arrows show the presence in PiCs colonies of migrating cells.
(B) Schematic representation of ESCs ←→ PiCs transition. The L-Pro-induced transition is followed by global epigenetic (de-regulation DNA/histones levels) and transcriptome remodeling.

1.2.4. L-Proline modifies the transcription and epigenetic profiles

L-Pro supplementation modifies the transcriptional profile of ESCs, altering the expression of a significant number (>1500) of protein-coding genes [2, 48], including genes involved in cytoskeleton remodeling, cell adhesion and extracellular matrix proteins degradation. Moreover, L-Pro influences the expression of growth factors, cytokines or their receptors, among which members of the STAT, PI3K and MAPK pathways, as well as the TGF-beta pathway. It is known that epigenetic changes control stem cells differentiation, and that epigenetic changes may be linked to metabolic alterations. L-Pro supplementation induces a global and genome-wide remodeling of histone and DNA methylation in ESCs [2, 3, 7].

Histone methylation: L-Pro supplementation increases the global level of H3K9me3/me2 and H3K36me3 in a significant number (~ 16500) of sites, and with the highest methylation increment occurring at the noncoding intergenic regions of the genome [2]. Chromatin immunoprecipitation sequencing (ChIP-seq) analysis revealed that H3K9me3 frequently marks repetitive elements, suggesting an induced local heterochromatin spreading into nearby genes that influences their expression. The highest H3K9me3 enrichment at constitutive heterochromatin, *i.e.* pericentromeres and gene deserts, suggest that L-Pro induces heterochromatin reorganization toward more dense structures in the nucleus [2]. Histone demethylation reactions are catalyzed by members of the JMJ enzymes, which belong to a large family of dioxygenases enzymes that require VitC as cofactor. Indeed, at appropriate stoichiometric ratios, VitC fully counteracts L-Pro-induced histone methylation in ESCs. Of note, upon silencing of Jmjd1a (H3K9 demethylase) ESCs adopt PiCs-like phenotypic features, *i.e.* flat morphology, sensitivity to trypsin digestion and upregulation of *Fgf5* and *Brachyury/T* genes [49].

DNA methylation: L-Pro supplementation modifies DNA methylation at global (quantification by LC–MS) and genome-wide levels (RRB-Seq). Similar changes occur during the naïve to primed transition [50]. Specifically, L-Pro supplementation increases 5mC and reduces 5hmC levels and results in ~1000 differentially methylated regions (DMRs) distributed throughout all chromosomes. Of note, this effect is fully counteracted by VitC supplementation, which in turns reduces DNA methylation levels. The findings that the majority of the DMRs that are hypomethylated in VitC-treated ESCs are conversely hypermethylated in PiCs (96% at day 5) suggest a common mechanism of action of the two metabolites [3].

1.2.5. L-Proline releases the amino acid stress (AAR/Atf4) pathway

In ESCs L-Pro availability modulates the expression of Activating transcription factor 4 (*Atf4*), which in turn controls L-Pro biosynthesis [48]. Interestingly, even though ESCs are able to synthesize L-Pro, suffer of a finely regulated intrinsic starvation of L-Pro, even when cultured in complete rich medium (serum/LIF). Intrinsic L-Pro starvation activates the Gcn2-Eif2α-Atf4 amino acid stress response (AAR) pathway and limits,

without blocking, ESC proliferation [48]. Briefly, in the absence of exogenously supplied L-Pro, the serine/threonine kinase Gcn2 (general control nonderepressible 2) phosphorylates Eif2a (eukaryotic translation initiation factor 2), which results in, increased expression of the transcription factor Atf4 also known as cAMP-response element binding protein 2 (Creb-2) (**Fig. 1.6**) [48]. Atf4 induces the expression of genes involved in the transport and synthesis of amino acids, including L-Pro biosynthesis genes (*Aldh18a1* and *Pycr1*). The exogenous addition of L-Pro in a physiological range of concentrations (50-150µM) alleviates AAR stress, reduces the expression of *Aldh18a1* and *Pycr1*, and promotes ESCs proliferation (**Fig. 1.6**). This auto-regulatory feedback loop generates and sustains the intrinsic shortage of L-Pro in mESCs [48].

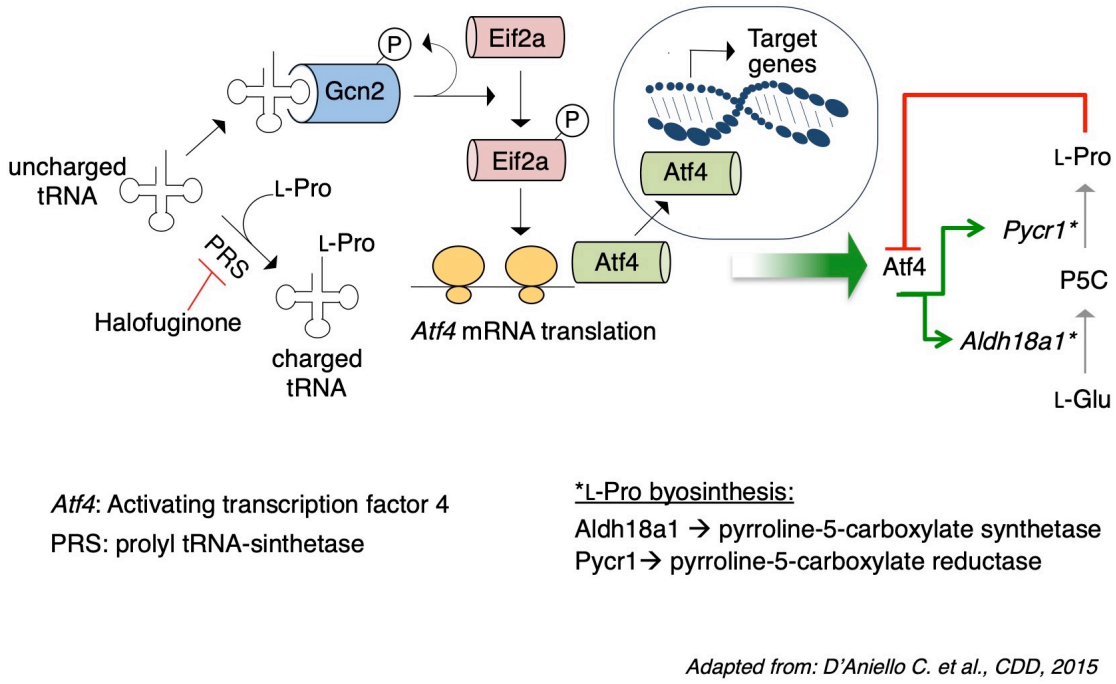


Figure 1.6: L-Proline addition modulates the AAR pathway in ESCs. Schematic representation of the AAR (Gcn2-Eif2α-ATF4) pathway (*left*). Schematic representation of the L-Pro–AAR/Atf4 regulatory feedback loop (*right*).

1.2.6. L-Proline induces FGF and TGF-β pathways

A high L-Pro regimen (> 150µM) forces an autocrine induction of Fibroblast Growth Factor (FGF) and Transforming Growth Factor beta (TGFβ) signaling pathways in ESCs [3]. Indeed, L-Pro supplementation promptly induces the phosphorylation of

signaling transducers, such as SMAD2 (TGFβ) and the extracellular signal-regulated kinase (ERK). Hence, the gene expression profile of L-Pro-treated cells becomes highly enriched in genes related to the ERK super-pathway [3]. Accordingly, the pro-proliferative effect of L-Pro abundance is counteracted using chemical inhibitors of FGF and TGFβ pathways, *i.e.* PD0325901 and SB431542, respectively. Moreover, L-Pro effects are highly specific since none of the non-essential amino acids (NEAAs) other than L-Pro, including L-Ala, L-Asn, L-Asp, L-Glu, L-Gly, and L-Ser, induces neither cell proliferation nor the induction of TGFβ-related genes (*Lefty1*, *Lefty2*) [48]. The mechanism linking L-Pro with activation of ERK pathway is not completely understood. However, under L-Pro abundance an increased synthesis of L-Pro-rich proteins such as collagens may contribute to the induction of *Fgf5*, *Fgf8* and *Fgf13* genes [2], which in turn may activate the ERK pathway. Under such conditions, E-cadherin, a key component of the cell-cell adherent junctions, is delocalized from the cytoplasmic membrane to the Golgi in a Protein kinase domain containing, cytoplasmic (Pkdcc)-dependent manner [2]. Indeed, WNT (Wingless and INT-1) signaling is known to influence cell migration [51], and a potent WNT agonist CHIR99021, which selectively inhibits the glycogen synthase kinase 3 (GSK3), counteracts L-Pro-induced E-cadherin delocalization and cell migration [2]. Thus L-Pro abundance may antagonize, directly or indirectly, the WNT pathway in ESCs.

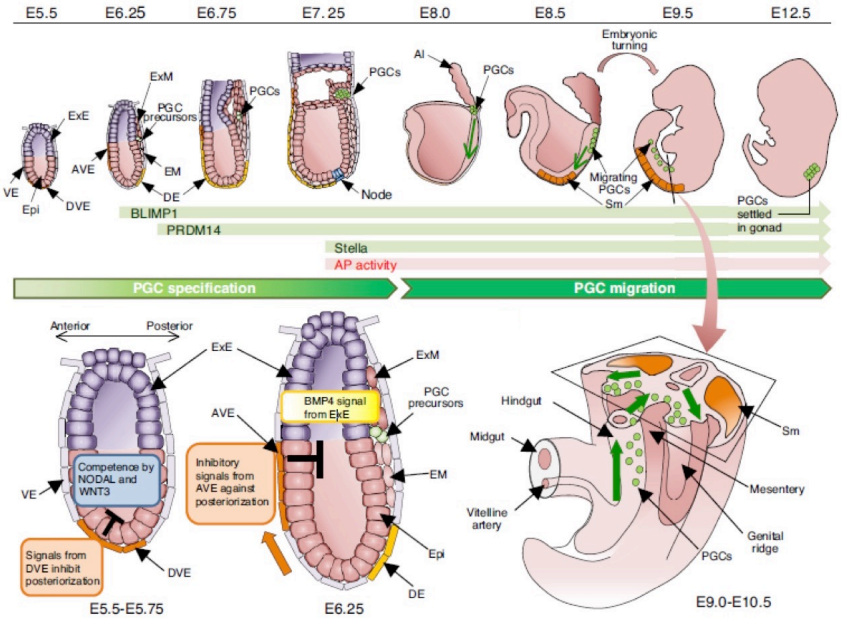
1.3. Gastruloids formation assay: *in vitro* gastrulation

Naïve pluripotent cells forced to grow under non-adherent culture conditions are inclined to establish cell-cell adhesive interactions and, eventually, to generate floating three-dimensional cells aggregates (hereafter called ‘aggregomes’). Tightly packed and globular aggregomes of ESCs can undergo symmetry breaking and grow in a polarized manner with regard to three orthogonal axes (anterior-posterior, dorsal-ventral and left-right) [4, 8, 52-56], resulting in elongated organoids called ‘gastruloids’ [54]. The aggregome to gastruloid transition has been shown to exhibit some features of mouse embryo gastrulation [4, 57, 58]. Gastruloids development depends on: i) the cell line examined; ii) the diameter (150-200 μm) of the initial cell aggregomes, which is determined by the number of cells aggregated (200-400 cells); and iii) a transient (24 h)

treatment with a WNT (Wingless/Integrated) signaling inducer [4, 8]. Nevertheless, a significant fraction (20-30%) of the initial aggregomes developed into atypical organoids, including large spherical cell masses (proliferation without symmetry breaking), spherical structures with protrusions (multiple elongation attempts), and elongated gastruloid-like organoids but with one or few short ectopic protrusions [4, 8]. Although the reason for these failures in gastruloid development remains largely unknown, gastruloid development provide a great opportunity to answer important questions regarding embryo development/gastrulation [57, 58].

1.4. Primordial Germ Cell (PGC) differentiation assay

Primordial Germ Cells (PGCs) are the precursor of the gametes, both oocytes and spermatozoa, and are detectable in the early mouse development around embryonic day E6.25 [59] (**Fig. 1.7**). Early germ cells are unique since express pluripotency specific markers after gastrulation [59]. PGCs express the pluripotency transcription factors *Oct4*, *Sox2* and *Nanog* as well as some lineage specific markers, including *Fragilis*, *Nanos3* and *Dazl.* Furthermore, *Blimp1* and *Prdm14* genes are key regulators of the PGCs fate. PGCs undergo also epigenetic reprogramming, particularly genome-wide DNA demethylation, that is fundamental for X inactivation, genome imprinting and gene expression [60]. So far, the capability to differentiate into primordial germ cell-like cells (PGCLCs) has only been found in early-primed (EpiLCs; 2d_Fgf/Activin A) pluripotent stem cells. Actually, neither naïve (2i_LIF or FBS_LIF) nor primed (EpiSCs, 5d_F/A) are able to undergo PGC differentiation [9, 23]. This perfectly correlates with the fact that PGCs derived from the E5.5-6.0 epiblast [61]. The mechanisms underlying PGCs specification are not easy to explore mainly due to their recalcitrant capability to proliferate *in vitro* [62]. However, the PGCLCs differentiation assay is a valid tool for the evaluation of intermediate functional states of pluripotency.



Adapted from Saitou M et al., Development 2012

Figure 1.7: Schematic representation of germ cells differentiation and migration during mouse embryo development.
(*Top*) Primordial germ cell (PGC) schematic specification and migration during embryo development. PGCs precursors (E6.5) and PGCs are shown as green dot. Key expression genes (*Blimp1*, *Prdm14* and *Stella*) and alkaline-phosphatase activity (AP) are shown in green and red arrows, respectively.
(*Bottom left*) Signalling induction at E5.5-5.75 and E6.25 of PGC inducing pathway.
(*Bottom right*) Schematic representation of PGCs migration form the hindgut to genital ridge at E9.0-10.5. E, embryonic day, l, allantois; AVE, anterior visceral endoderm; DE, distal endoderm; DVE, distal visceral endoderm; EM, embryonic mesoderm; Epi, epiblast; ExE, extra-embryonic ectoderm; ExM, extra-embryonic mesoderm; PGCs, primordial germ cells; Sm,somite; VE, visceral endoderm.

2. AIM OF THE PROJECT

It has been recently reported that mouse *naïve* embryonic stem cells (ESCs) exposed to a high L-Pro regimen undergo an embryonic stem-to-mesenchymal transition (esMT), and that the resulting L-Proline-induced cells (PiCs):

a. maintain some features of naïve pluripotent cells including LIF reliance, the ability to generate teratomas and to colonize a blastocyst [1, 2];

b. acquire some features (metabolic, epigenetic, ecc.) of early *primed* pluripotent cells [2, 3];

c. revert efficiently to naïve pluripotency state [1, 2].

Thus, the evidences collected so far suggest that PiCs have features of the recently hypothesized "formative" state of pluripotency [19]. Yet, conclusive evidence is still missing, also due to limitations of *in vivo* and *in vitro* experimental models. In this context, the development of well-structured embryonic organoids from naive ESCs, named gastruloids [4], can be exploited as a powerful tool to overcome, at least in part these limitations. Indeed, we reasoned that an optimized/adapted version of this *in vitro* method could be useful to evaluate and discriminate different pluripotency states and, in particular, to ultimately define PiCs pluripotency.

Mechanistically, emerging evidence lead to hypothesise that PiCs may be functionally linked to increased biosynthesis of L-Pro-rich proteins such as collagens [2, 7]; however, direct experimental evidence for this mechanism is still lacking.

Based on the above considerations, the aims of my PhD project were the followings:

1. set up an improved protocol to quantify the Gastruloid Formation Efficiency (GFE) of ESCs as *in vitro* model to functionally discriminate different states of pluripotency (naïve, EpiLC/early-primed, EpiSCs/primed), including PiCs.

2. investigate the ability of PiCs to differentiate into primordial germ-like cells (PGLCs), which so far is a unique prerogative of early-primed cells.

3. evaluate the role of L-Pro-induced collagen biosynthesis in the determination of ESCs identity.

3. MATERIALS AND METHODS

3.1. Culture of ESCs, reagents, and treatments

Wild type mouse TBV2 (129/SvP) and R1 ESCs and two independent *Cripto* KO ESC clones, DE7 and DE39 (named Cl.1 and Cl.2) [63, 64], were used throughout the study. ESC lines were maintained on a feeder layer of mitomycin C-treated primary MEFs according to standard procedures. Feeder-independent ESCs (E14) were cultured onto gelatin-coated plates. ESCs were cultured in high glucose Dulbecco's modified Eagle's medium (DMEM, Invitrogen, Life Technologies) supplemented with 15% ES-screened Fetal Bovine Serum (FBS, Euroclone), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), 1 mM sodium pyruvate, 2 mM glutamine, 100 U/ml penicillin/streptomycin (all from GIBCO), and 1000 U/ml recombinant LIF (ESGRO, Millipore). 2i Medium (N2B27) was supplemented with PD0325901 (1 μM), CHIR99021 (3 μM) and LIF (ESGRO, Millipore). Recombinant soluble Cripto protein (sCripto) was used at the concentration of 10μg/ml. All cell lines were routinely tested and confirmed to be free of mycoplasma.

3.2. *In vitro* capture of primed state

For ESC to EpiLC and EpiSC transition, 2i mESCs were plated at 1500 cells/cm^2 onto FBS-coated plates and grown in N2B27 medium supplemented with 20 ng/ml Activin A (Invitrogen) and 12 ng/ml bFgf (Provitro) and cultured for 2 days (2d_F/A, EpiLCs) and 5 days (5d_F/A, EpiSCs), respectively.

3.3. *In vitro* generation of L-Proline- induced cells (PiCs)

To induce ESC to PiC transition ESCs were plated at low density (50–250 cells/cm^2) in complete medium (DMEM/15%FBS/LIF) on gelatin-coated plates and grown for 5 days in the presence/absence of L-Pro (250–500 μM) (Sigma-Aldrich), with a medium change at day 3. PiCs were harvested using accutase (Sigma-Aldrich).

3.4. Colony Formation Assay

Colony-formation assay was performed as previously described [1]. Briefly, ESCs were trypsinized to obtain a single cell suspension and plated at low density (300 cells/cm^2). After 5 days, colonies were washed twice with phosphate-buffered saline (PBS) fixed/stained with a solution of PBS 1x/6% glutaraldehyde/0.15% crystal violet (CV). Images were collected on a DMI6000B12 microscope (Leica Microsystems). The morphological classification (domed/flat) was performed blinded by two investigators. Experiments were performed in triplicate.

3.5. Generation of *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} ESCs

For CRISPR/Cas9-mediated P4ha2 knockout, TBV2 ESCs were infected with lentiviral particles (1 MOI) carrying the guide RNA (gRNA), which targets P4ha2 exon 6 (ATCCGGACACGATTTCCAGA), followed by the Cas9 and the puromycin resistance [transEDIT Lentiviral gRNA plus Cas9 (pCLIP-All) Target Gene; transOMIC Technologies]. Following puromycin selection (1 μg/mL; 6 days), resistant cells were dissociated and single cells were FACS-sorted (FACS Aria; Becton Dickinson) on the basis of forward scatter and side scatter parameters and seeded into 96-well plates. Genomic DNA was analysed by PCR using primers designed on intron 5–6 (forward) and 6–7 (reverse) sequences that amplify a product of 735 bp. IN/DEL mutations were identified by digesting PCR products with the T7 endonuclease and confirmed by sequencing. The absence of the wild-type allele was confirmed by RT-PCR (see Primers Table). *P4ha2*^{KO}/*P4ha1*^{KD} ESCs were obtained by transfecting *P4ha2*^{KO} ESCs (Clone #1) with short hairpin RNAs (shRNA) plasmids targeting different nonoverlapping P4ha1 mRNA sequences followed by a GFP reporter (shERWOOD UltramiR Lentiviral shRNA, pZIP-mEF1a, TransOMIC). Forty-eight hours after electroporation, single GFP-positive cells were FACS sorted into 96-well plates and the cells were used for further analysis.

3.6. Gastruloid formation assay

For gastruloids formation assay we modified the previously published protocol [8, 54]. Briefly, cells grown at low density, the resulting colony were dissociated with Accutase (Sigma-Aldrich) and then an exact number of cells ($2.5x10^2$ or $3.0x10^2$) were FACS sorted/seeded in U-shaped ultra-low attachment 96-multiwell (Corning Costar) and allowed to aggregate for 48h in N2B27 medium without supplemental growth factors. After 48h in culture, CHIR was added at the concentration of 3µM and maintained for 24 hours. From 72h onwards, N2B27 medium (150 µl) was refreshed daily until 120h. When the addition of specific factors to Gastruloids was required 24h after aggregation (AA), 20 µl medium was carefully removed with a multichannel pipette, and 20 µl of N2B27 containing required factors, at the appropriate concentration, was added as previously described [8]. Gastruloids were collected randomly at the different time-points analysed and were imaged using the Evos Cell imaging systems.

3.7. Primordial Germ Cell-like cell (PGCLCs) differentiation

PGCLCs differentiation was performed under a floating condition as previously described [65]. Briefly, $2.5x10^4$ cells/well were plated in U-shaped ultra-low attachment 96-multiwell (Corning Costar) in a serum-free GK15 medium containing Glasgow minimal essential medium (GMEM) supplemented with 15% KSR, 0.1 mM NEAA, 1 mM sodium pyruvate, 0.1 mM 2-mercaptoethanol, 100 U/ml penicillin, 0.1 mg/ml streptomycin, and 2 mM L-glutamine. To induce PGCLC differentiation the medium was supplemented with growth factors/ cytokines including BMP4 (500 ng/ml; R&D Systems), LIF (1000 u/ml; Invitrogen), SCF (100 ng/ml; R&D Systems), and EGF (50 ng/ml; R&D Systems). After 4 days the cell aggregates were dissociated with Trypsin and the cells were used for further analysis.

3.8. Flow cytometry and cell sorting

TBV2 mESCs and E14 cells were dissociated with accutase mix (1x for 5 min at 37°C) to obtain a single cell suspension and sorted with a FACS ARIAIII (Becton Dickinson) on the basis of Forward (FSC-A) and side scatter (SSC-A) parameters, excluding cellular debris and cells death.

3.9. High-throughput screening

The Prestwick Chemical Library 1200 compounds (ww.prestwickchemical.com/libraries-screening-lib-pcl.html) was used for the screening. The screening was performed using the Cellmaker robotic platform [1]. TBV2 ESCs were plated at 250 cells/cm2 in gelatin-coated plates (96-multiwell) in DMEM/10% FBS/LIF. Five hours after plating, cells were supplemented with L-Pro (250 µM) or left untreated as control, and the compounds (Prestwick chemical library) were singularly added at a final concentration of 10 µM. At day 4 after treatment, the resulted colonies were fixed/stained in PBS1x/6% glutaraldehyde/0.15% crystal violet for 30 minutes at room temperature. The plates were scanned/imaged and the cell colonies (if any) were analysed using the ImageJ 1.46r software (https://imagej.nih.gov/ij/). Chemical structure information for budesonide are available at PubChem (https://pubchem.ncbi.nlm.nih.gov/compound/5281004), DrugBank(https://www.drugbank.ca/drugs/DB01222) and Therapeutic Target Database (https://db.idrblab.org/ttd/drug/d0y7iu).

3.10. Preparation of cytospin samples

Cells were dissociated with Accutase or Trypsin-EDTA for 5 min at 37°C, re-suspended in 15% FBS/1x PBS and centrifuged at 800 rpm for 8 min onto glass slides (2 spots, $5x10^5$ cells/spot) using a Thermo Shandon Cytocentrifuge (CytoSpinTM 4). Specimens were fixed with PFA 4% for further analysis.

3.11. Immunofluorescence analysis and quantification

Gastruloids were fixed (4% PFA) and the immunofluorescence was performed as previously described [8]. The following primary antibodies were used overnight at 4°C T (1:200; Santa Cruz), Sox17 (1:200; R&D), Cdx2 (1:100; Cell Signaling), Sox2 (1:100; Cell Signaling), Oct4 (1:100; Santa Cruz), Blimp1 (1:50; Santa Cruz), Ki67 (1:50; Invitrogen), Nestin (1:500; Santa Cruz), Oct4 (1:100; Santa Cruz); Nanog (1:1000; Cell Signaling); anti-AP2γ (1:100; Santa Cruz). After washing cells were incubated with the appropriate secondary antibodies (Alexa Fluor DAR-594, 1:400, Molecular Probes #A21207; Alexa Fluor GAM-488, 1:400, Molecular Probes #A11001,

Alexa Fluor DAG-594 1:400, Molecular Probes #). Cell nuclei were counterstained with DAPI (Invitrogen).

Cytospin cell samples were fixed in 4% PFA and permeabilized (0.1% Triton X-100) for 10 minutes at RT and incubated with primary antibodies [anti-Oct4 (1:100; Santa Cruz), anti-Blimp1 (1:50; Santa Cruz), anti-AP2γ (1:100; Santa Cruz), anti-Col1a1 (1:100; abcam), anti-PDI (1:100; Abcam) and anti-C-HyP (1:100; Abcam)] overnight at 4°C. After washing in 0,5% Tween-1x PBS, cells were incubated with the appropriate secondary antibodies. Cell nuclei were counterstained with DAPI (Invitrogen).

Images were obtained using the DMI6000B microscope and the DFC 350FX B/W digital camera (Leica Microsystems). Confocal images were obtained on a Nikon A1 microscope. The AF6000 (Leica Microsystems) and NIS Element C (Nikon, Tokyo) software were used for image acquisition/elaboration.

C-HyP positive area (pixel/cm^2) was measured over the total nuclei area (100–300 cells/conditions) using ImageJ software (https://imagej.nih.gov/ij/docs/guide/146-30.html#toc-Subsection-30.2).

3.12. Quantification of 5hmC

Gelatin plated ESCs (15 × 103/cm^2) were pre-treated with VitC (50 μM; 24 hours) before addition of L-Pro ± budesonide for 24 hours. For ELISA-based quantification, genomic DNA was extracted (Wizard Genomic DNA Purification Kit; Promega) and 5hmC levels were measured using MethylFlash Global DNA Hydroxymethylation (5hmC) ELISA Easy Kit (Epigentek) and normalized to DNA content, following the manufacturer instructions.

For immunofluorescence, cells were fixed (4% PFA) and permeabilized (0.4% Triton X-100; 15 minutes at room temperature). After a denaturation step (2N HCl, 15 minutes at room temperature) and a neutralization step (100 mmol/L Tris-HCl pH 8.5, 10 minutes at room temperature), cells were incubated with 0.1% Triton X-100/5% BSA for 1 hour. Anti-5hmC mAb was incubated overnight at 4°C (1:200; HMC/4D9 A1018; Epigentek) followed by anti-mouse 594 secondary antibody (Alexa Fluor Molecular Probes). Images were obtained using the DM6000B microscope equipped with DFC 350FX B/W digital camera (Leica Microsystems). The AF6000 (Leica Microsystems) software was used for image acquisition/elaboration.

3.13. Cell proliferation assays

Cell proliferation was evaluated using FACS-based EdU incorporation (Click-iT EdU Flow Cytometry Assay Kit; Molecular Probes), and the CCK-8 (Dojindo Laboratories) assays following manufacturer's instructions.

3.14. RNA extraction and quantitative real-time PCR

Total RNAs were isolated using either the RNeasy kit or Trizol reagent (Invitrogen) and reverse transcribed using QuantiTect Reverse Transcription kit (Qiagen). qPCR was performed using SYBR Green PCR master mix (FluoCycle II TM SYBR, EuroClone). Primers are listed in below.

Gene symbol	Primer Forward	Primer Reverse	Application
Atf4	TCCACTCCAGAGCATTCCTT	AAGCAGCAGAGTCAGGCTTC	qPCR
Trib3	GCTGGCAGATACCCATTCCA	GGCCGTGATCCTCTCTCAAC	qPCR
Nupr1	GAGAAGCTGCTGCCAATACC	GGTGTGGTGTCTGTGGTCTG	qPCR
Col1a1	GAAGAACTGGACTGTCCCAAC	CCTCGACTCCTACATCTTCTG	qPCR
P4ha1	TGAGCTATGCGGTGTACCAG	CGCTGACTTATTGGCATCTTT	qPCR
P4ha2	TCCTTTTCTCTCCGCTCCTC	AAACACTGGCTCCTCCAATG	qPCR
P4ha3	GCTGAGACTGCAATTCAGAGG	GGGCTGGAATTGGTCTCATA	qPCR
Nanog	AAGTACCTCAGCCTCCAGCA	GTGCTGAGCCCTTCTGAATC	qPCR
Oct3/4	TCAGCTTGGGCTAGAGAAGG	TGACGGGAACAGAGGGAAAG	qPCR
Sox2	CACAACTCGGAGATCAGCAA	CTCCGGGAAGCGTGTACTTA	qPCR
Cdx2	AGGAAGCCAAGTGAAAACCA	CAGCCAGCTCACTTTTCCTC	qPCR
Sox17	AGCTAAGCAAGATGCTAGGCAAG	TCTCTGCCAAGGTCAACGC	qPCR
Stella (Dppa3)	TTCCGAGCTAGCTTTTGAGG	ACACCGGGGTTTAGGGTTAG	qPCR
Foxa2	TGTAGCTGCGTCGGTATGTC	ACCTGAGTCCGAGTCTGAGC	qPCR
T/Bra	GAACCTCGGATTCACATCGT	TTCTTTGGCATCAAGGAAGG	qPCR
Cer1	CAGGCCGTGACTCAGCCAGCAG	CCGGGAAAACGAATGGAACTGC	qPCR
Rex1	TTGCCTCGTCTTGCTTTAGG	AAAATGAATGAACAAATGAAGAAAA	qPCR
Gbx2	GCTGCTCGCTTTCTCTGC	GCTGTAATCCACATCGCTCTC	qPCR
Fgf5	CAAAGTCAATGGCTCCCACGAAG	CTACAATCCCCTGAGACACAGCAAATA	qPCR
N-Cad	AGCGCAGTCTTACCGAAGG	TCGCTGCTTTCATACTGAACTTT	qPCR
E-Cad (Cdh1)	GGAGGAGAACGGTGGTCAAA	TGTGCAGCTGGCTCAAATCAA	qPCR
Prdm1(Blimp1)	AACTGGATGCGCTACGTGAA	TTGCTTTGGGTTGCTTTCCG	qPCR
Prdm14	AAGGCACACAGGGACAACTC	CTGGTTCCGCTGGATGTCTC	qPCR
Nanos3	GGGTGCTGTGTCCCATTTTG	ACCTGCATAGACACCTGCTG	qPCR
Tet1	GCTGGATTGAAGGAACAGGA	CCCTTCTGGCCAAACCTAGT	qPCR
Tet2	CTCAAGCAACCAAAAGCACA	GGACCAGCTCCTAGATGGGTA	qPCR
Tet3	TCCGGATTGAGAAGGTCATC	CCAGGCCAGGATCAAGATAA	qPCR
Cyr61	AAGGTGCTGGAGCTCATTGT	GCACCTCGAGAGAAGGACAC	qPCR
Jmjd2c	TGGATGTGGTTGAAGAAGAGTG	CGCAAATGTTGTTTTCCACA	qPCR
Gapdh	TGCACCACCAACTGCTTAGC	TCTTCTGGGTGGCAGTGATG	qPCR
P4ha2	CGCTTCAGGGGTCAAATCTG	AGTTTCCTGAGCTACGACCT	PCR-Screening CRISPR clones
P4ha2	AGACAGGTGTTTCATCACTG	AGTTTCCTGAGCTACGACCT	Sequencing CRISPR clones
P4ha2	TTCCAGAGGGGAACTTCCAG	CAGCATAGCTCAGGTAGTCC	RT-PCR

3.15. Western blot analysis

Total proteins were extracted in 100 mmol/L Tris pH 8, 140 mmol/L NaCl, 20 mmol/L EDTA, 0,2% SDS, 1% Nonidet P-40 lysis buffer, resolved on SDS-PAGE gels and transferred onto PVDF membranes (iBlot dry Transfer System; Life Technologies). Histones proteins were prepared as follows: cell pellets were resuspended in triton extraction buffer [TEB: PBS containing 0.5% Triton X-100 (v/v), 2 mmol/L phenylmethylsulfonylfluoride (PMSF), 0.02% (w/v) NaN3] at a cell density of 10^7 cells/mL. Then, the pellets were lysed overnight in 0.2N HCl at a density of 4×10^7 cells/mL. Primary Antibodies [anti-Col1a1 (1:200; Abcam); anti-C-HyP (1:500; Abcam); anti-P4ha1 (1:1000; Abcam); anti-Gapdh (1:10000; Abcam); anti-H3K9m3 (1:1000; Abcam); anti-H3K36me3 (1:1000; Abcam); anti-H3 (1:1000; Abcam)] were used overnight at 4°C followed by the appropriate HRP-conjugated secondary antibodies. Detection was performed with ECL reagents (Pierce, Thermo Scientific). ImageJ software was used for the densitometric analysis.

3.16. Histological analysis

Gastruloid were fixed in 2% glutaraldehyde/4% paraformaldehyde, post-fixed in osmium tetraoxide, dehydrated, and embedded in Epon 812 (Polyscience, Niles, IL, USA). Thin (5 µm) sections were cut using a Leica ultracut UCT ultramicrotome (Leica Microsystems) and stained with toluidine blue (1% in water) for 10 min at RT. Images were obtained with ECLIPSE Ni-E microscope.

3.17. Statistical analysis

Experiments were all performed with a minimum of three biological samples. Replicate number is indicated in figure legends as n=. Results are presented as the mean ± SD or SEM (Standard Deviation or Standard Error of the Mean, respectively) as indicated in the figure legends or as box plot displaying minimum, first quartile, median, third quartile and maximum. To determine significance between two groups, comparisons were made by a two-tailed paired Student's. $P<0.01$ was considered statistically significant.

4. RESULTS

PART I. Exploring the pluripotency state of PiCs:

4.1. Gastruloids formation ability of ESCs

We first evaluated the gastruloid formation ability of the feeder-dependent TBV2 ESCs that are currently used in our laboratory, using the protocol recently described by Baillie-Johnson and co-workers [8]. Briefly, TBV2 ESCs were grown on gelatin-coated plates in naïve inducing-medium (FBS/LIF), detached with trypsin and plated at the density of 300 cells/40µl in ultra-low attachment plates to force aggregation. As expected, 48 hours (h) after seeding, spherically shaped cell aggregates (hereafter 'aggregomes'), with a diameter ranging from 125 to 195 µm (mean = 156 µm) were observed (**Fig. 4.1A**).

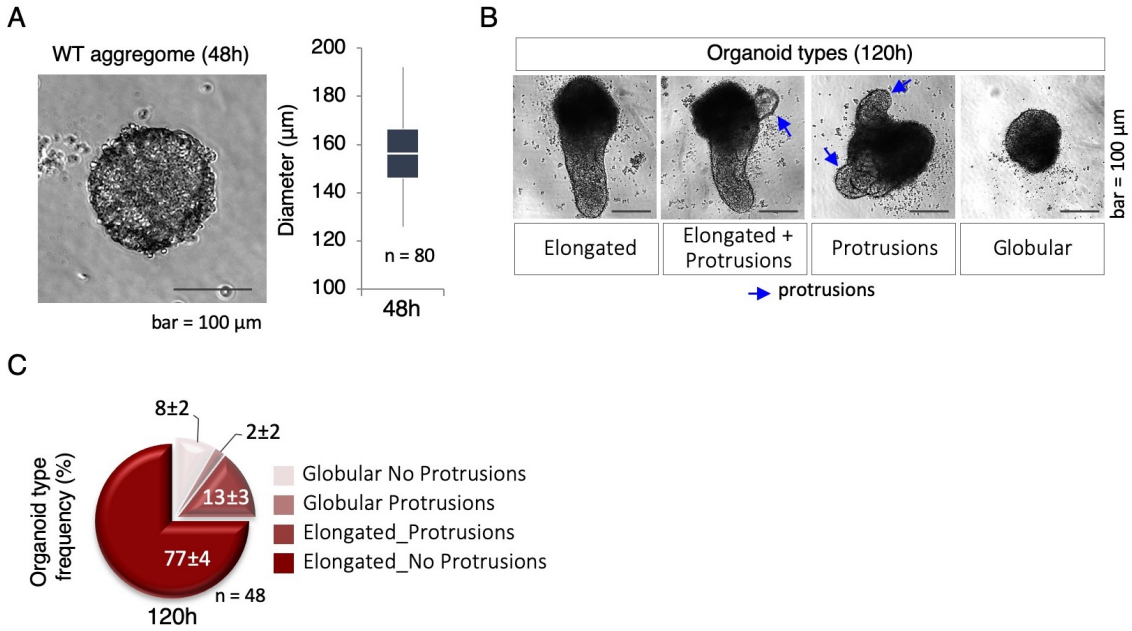


Figure 4.1. Handmade gastruloids formation assay
(A) Representative pictures of WT ESC aggregomes at 48h AA (*left*) and box plot diagram (*right*) of the distribution of the aggregomes' diameter (n=80)
(B) Representative pictures of different organoids' phenotypes at 120h AA. Blue arrows indicate protrusions. (bar=100 µm).
(C) Pie chart representing the percentage of the different phenotypic outcomes showed in panel B. (n=48).

Three days later, *i.e.* 120h after aggregation (AA), a major fraction (~75%) of the primary cell aggregomes displayed an elongated morphology (0,5 to 1 mm long),

whereas a minor fraction developed into abnormal/aberrant organoids (~15%) displaying ectopic elongation zones (ectopic protrusions) or remained as unstructured globular cell masses (~10%) (**Fig. 4.1B-C**). These results confirmed previous reports showing that a significant fraction (up to 30%) of the initial cell aggregomes undergo developmental failures [4, 8]. The fraction of efficient aggregomes, *i.e.* those able to develop into a fully elongated gastruloid, was defined as gastruloid formation efficiency (GFE).

4.1.1. Optimization of the gastruloid formation assay

To minimize the fraction of abnormal/abortive events, and thus to improve the GFE of mESCs, specific cell culture conditions and procedures to detach and plate the cells, were established (**Fig. 4.2A**).

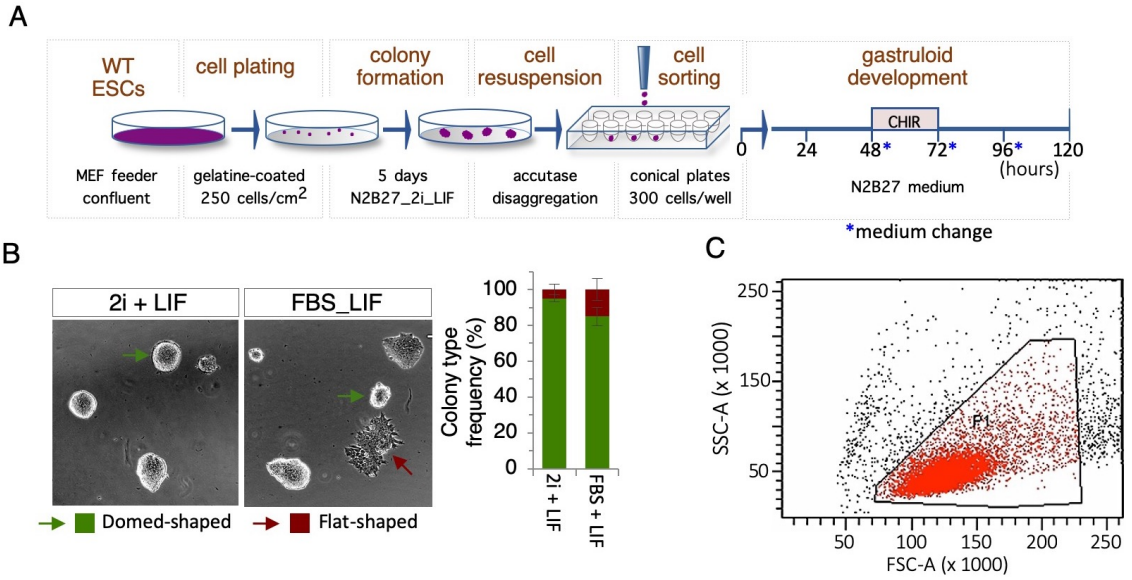


Figure 4.2. Set up of optimised conditions for the gastruloid formation assay.
(A) Schematic representation of the experimental design. Low-density colony formation, accutase dissociation and FACS seeding were introduced to standardize the protocol. 2i_LIF TBV2 mESCs (WT) were seeded in ultra-low attachment 96-well plate and cultured in N2B27 medium for 120h. CHIR was added at 48h.
(B) Representative pictures (*left*) and colony type quantification (*right*) of colonies generated from 2i_LIF and FBS_LIF cells. Arrows indicated domed- (green) and flat- (red) shaped colonies.
(C) Flow cytometry scatter plot (FSC vs SSC) of 2i_LIF ESCs. Dead cells and debris are excluded based on FSC/SSC parameters.

1. Cell culture: since the morphology of the cell colonies (round/domed vs irregular/flat) reflect their pluripotency state (**Fig. 4.2B**), mESCs were grown at low density (cell colonies) rather than as confluent monolayer (>60%) as reported by Baillie-Johnson and

co-workers [8]. FBS/LIF TBV2 mESCs grown on feeder layer were dissociated by trypsin digestion and then seeded at low density (250 cells/cm^2) on gelatin-coated plates in naïve state-inducing medium, *i.e.* N2B27 supplemented with LIF and the 2i inhibitors (PD0325901/CHIR99021). Under this culture conditions, mESCs gave rise to 90-95% of naïve/round-domed cell colonies (**Fig. 4.2B**).

2. Cell dissociation: given that trypsin-mediated cell dissociation reduces cell-cell adhesion capability of mESCs [2, 12], which may affect the cell aggregation step and eventually reduce the GFE, the cells were dissociated using a milder treatment with accutase (**Fig. 4.2A**).

3. Cell plating: Since proper development of gastruloid relies on the initial number of aggregated cells [8], FACS cell sorting was exploited to seed a precise number of living cells per well, usually 250-350 cells/well, as well as to exclude dead cells and cellular debris from the initial cell aggregomes (**Fig. 4.2A and 4.2C**).

The subsequent steps, including a pulse of CHIR between 48h and 72h AA (**Fig. 4.2A**), were carried out as previously described by Baillie-Johnson and co-workers [8].

Following this new experimental approach, only two different phenotypes emerged at 120h AA; *i.e.* a prevalent fraction (95-98%) of fully developed elongated-shaped organoids and a minor fraction (2-5%) of elongated organoids but with at least one ectopic protrusion (**Fig. 4.3A-B**).

Different parameters analysed, including the diameter of aggregomes (48h AA), and the length_volume of the developed organoids (120h AA), displayed a low dispersion around the corresponding mean value (**Fig. 4.3C**), thus suggesting a high accuracy of the method used. For instance, the diameter of FACS-plated aggregomes fluctuated in a significant lower range (mean=166 μm; min=153 μm; max=180 μm) compared to the handmade aggregomes (mean=156 μm; min=125 μm; max=195 μm) (**Fig. 4.3C; Fig. 4.1A**, respectively).

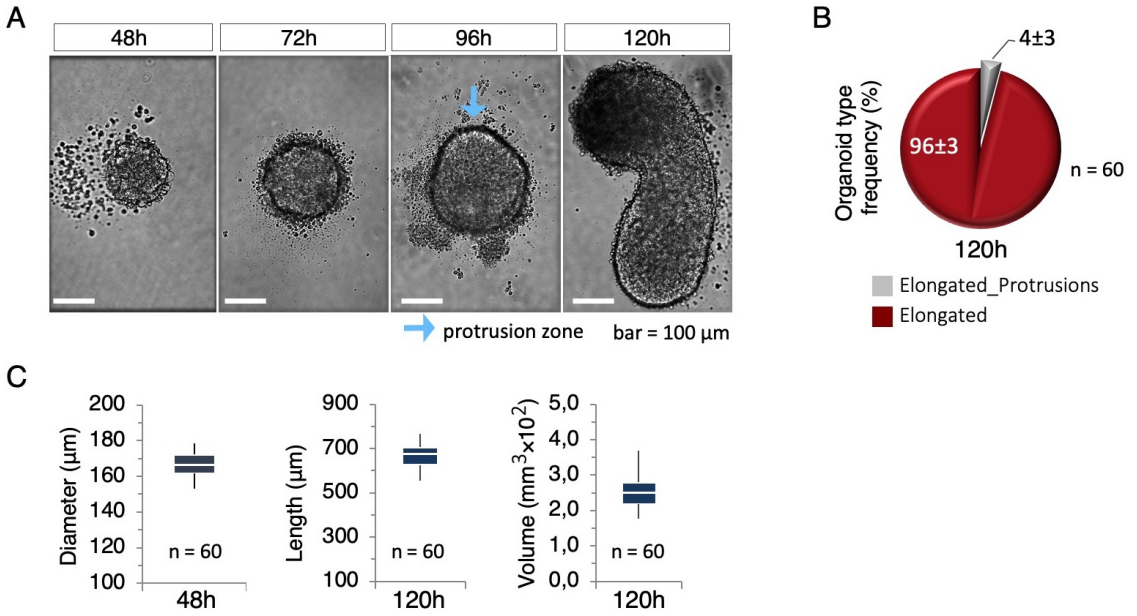


Figure 4.3. Efficiency of the optimised gastruloid formation assay.
(A) Time-course representative bright field images of aggregomes to gastruloids transition showing the morphology of aggregomes from 48h to 120h after aggregation (AA). Light blue arrow indicates the protrusion zone (bar=100 µm).
(B) Pie chart representing the percentage of the different phenotypic outcomes including organoids with a defined AP axis and without protrusions ('Elongated') or organoids with a defined AP axis but with ectopic protrusions ('Elongated_Protrusions') (n=60).
(C) Box plot diagram of the aggregomes diameter at 48h AA (*left*), and the length (*middle*) and volume (*right*) of the gastruloid at 120h AA (n=60/time point).

Furthermore, we evaluated by immunofluorescence analysis the expression profile of selected developmental markers that are induced during the aggregome to gastruloid transition in a time- and space- controlled manner (**Fig. 4.4A**), including Brachyury (T) (primitive streak, mesoderm), Sox2 (progenitor cells, stemness), Sox17 (primitive and definitive endoderm), Cdx2 (posterior mesoderm), and Nestin (neuronal precursor cells), and found that they were properly induced, thus providing molecular evidence to the establishment of cell lineages and antero-posterior (A-P) axis. Additionally, immunofluorescence analysis of the proliferation marker Ki67 revealed the presence of actively dividing cells (**Fig. 4.4B**).

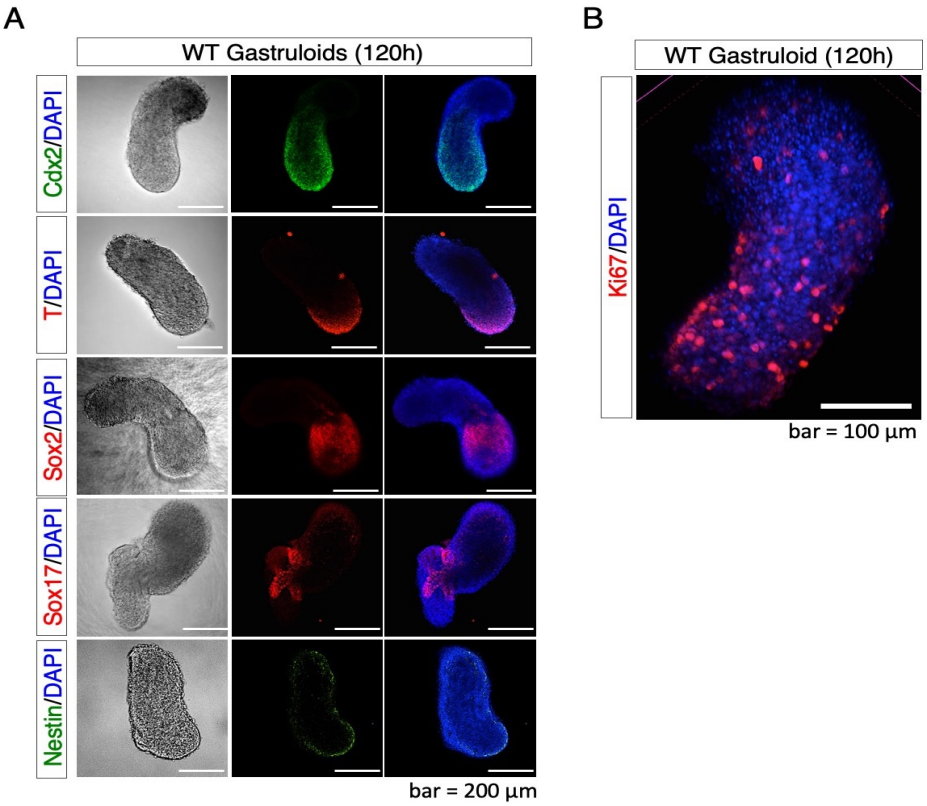


Figure 4.4. Expression of specific markers in gastruloids.
(A) Representative pictures of bright-field and confocal immunofluorescence analysis with Cdx2 (green), T/Bra, Sox2, Sox17 (red) and Nestin (green) of 2i_LIF mESC gastruloids at 120h AA. Nuclei were counterstained with DAPI (blue) (bar=200 µm).
(B)Representative pictures of confocal immunofluorescence with Ki67 of gastruloid at 120 AA (red). Nuclei were counterstained with DAPI (blue).

In order to better characterize the effect(s) of cell sorting on the GFE, the gastruloid formation assays was performed using different stem cell subpopulation. Surprisingly, we observed that excessive narrowing of the forward-scattered light (FSC) value of the cells sorted reduced GFE by increasing the fraction of undeveloped aggregomes up to 5-fold (**Fig. 4.5**).

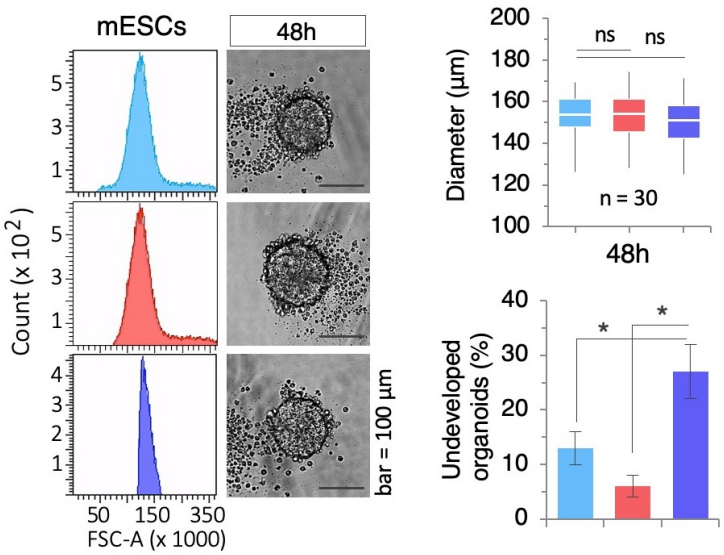


Figure 4.5. **Setting up the best condition in the optimised gastruloid formation assay.**
(*Left*) Flow cytometry histograms of the different sorted populations from untreated FBS/LIF ESCs.
(*Middle*) Bright field representative pictures of the aggregomes.
(*Right*) Aggregomes diameter distribution (48h AA) and percentage of undeveloped organoids (120h AA) of the different selected population (total population=light blue; population without dead cells and debris=red; narrow population = violet) (bar= 100µm).

These data suggested that a certain degree of phenotypic heterogeneity in the stem cell population is essential to generate functional aggregomes. Gastruloid formation assays were then performed using a different stem cell line, the feeder-free E14 mESCs. To this end, the cells were grown at low density (250 cells/cm^2) on gelatin-coated plates in naïve-inducing (N2B27/2i/LIF) medium, rising up to 95% of round-domed cell colonies (**Fig. 4.6A**). The resulting cells were sorted and seeded (300 cells/well) by FACS (**Fig. 4.6B**), and after 5 days incubation (120h AA) almost all aggregomes became elongated-shaped gastruloids (**Fig. 4.6C**).

All together these data lead to conclude that the modified protocol results in improved GFE of both feeder-free and feeder-dependent mESCs.

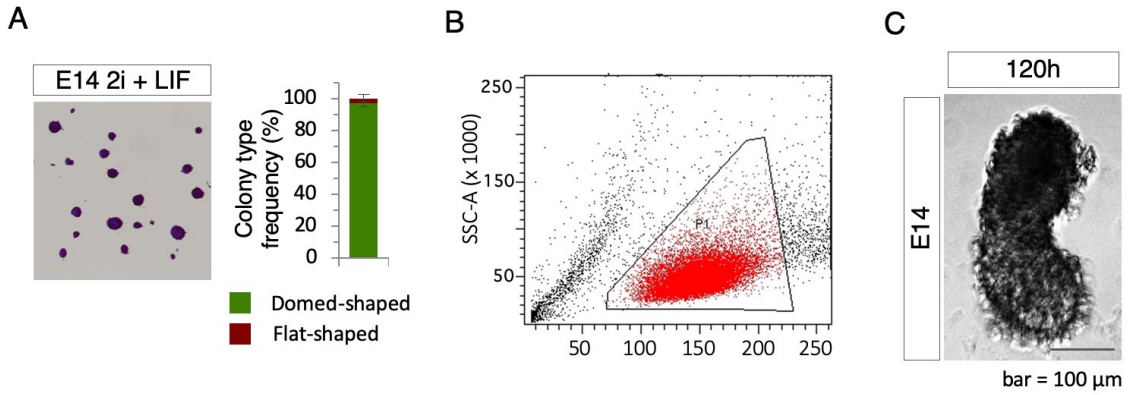


Figure 4.6. Gastruloids formation efficiency of a feeder-free ESC line.
(A)Representative pictures (*left*) and colony type quantification (*right*) of colonies generated from 2i_LIF E14 feeder-free cells line.
(B) Flow cytometry histograms of the E14 2i+LIF sorted population.
(C) Bright field representative picture of E14 gastruloid at 120h AA.

4.1.2. *Cripto* genetic ablation impairs gastruloid formation ability of ESCs

In order to validate the optimised protocol, we used the *Cripto* Knockout (KO) mESCs. *Cripto* (also known as *Teratocarcinoma-Derived Growth Factor 1*, *Tdgf1*) is a key regulator of stem cells pluripotency [5, 6], and is required for the anterior-posterior axis formation during mouse embryo development [66, 67]. To this end, wild-type (R1) and two independent *Cripto* KO cell lines (hereafter Cl.#1 and Cl.#2) [63], were grown at low density on gelatin-coated plates in naïve state-inducing medium (N2B27/2i/LIF). Following 5 days in culture, the resulting colonies were dissociated with accutase mix and the disaggregated cells were seeded in 96-well ultra-low attachment plates (300 cells/40 µl) (**Fig. 4.2A)**. Following two days of incubation (48h AA) both WT and *Cripto* KO cells were able to generate almost spherical cell aggregomes with a diameter of around 160 µm (**Fig. 4.7A**). Later on, at 120h AA, almost all the aggregomes generated by control WT cells became fully elongated organoids, while a small fraction (<5%) remained as undeveloped spheroids (**Fig. 4.7B**). Conversely, most of the *Cripto* KO aggregomes (≥95%) maintained an unstructured spheroidal morphology (**Fig. 4.7B-C**), providing evidence that in the absence of *Cripto* the aggregome to gastruloid transition was halted.

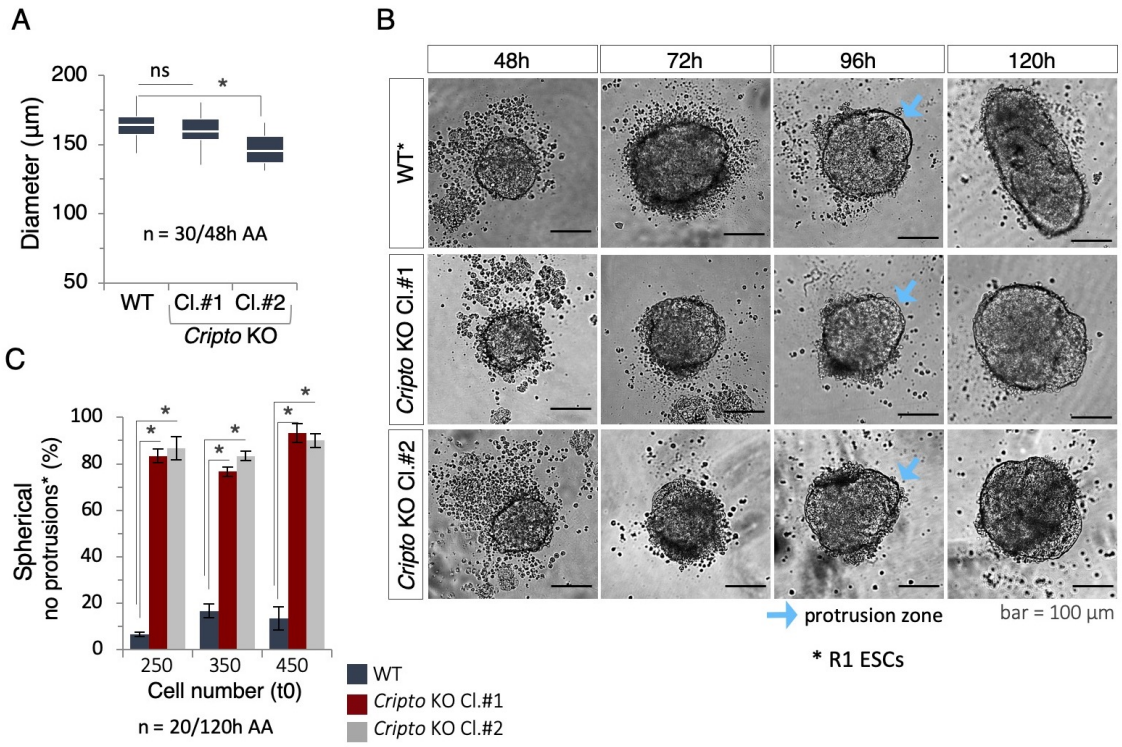


Figure 4.7. *Cripto* genetic ablation impairs gastruloids formation.
(A) Diameter distribution of R1 (WT) and *Cripto* KO mESCs (Cl.#1 and Cl.#2) derived aggregomes at 48h AA (n=30; bar=100 µm).
(B) Bright field representative pictures of time course of WT and *Cripto* KO ESC-derived gastruloids showing that *Cripto* KO ESCs failed to develop elongated structures.
(C) Percentage of undeveloped gastruloids (*spherical no protusions) of WT and *Cripto* KO clone #1 (Cl.#1, red) and clone #2 (Cl.#2, grey) derived organoids at 120h (n=20, *p<0.01, bar=100 µm).

Cripto is a glycosylphosphatidylinositol (GPI)-anchored extracellular protein [68], which is biologically active also as a soluble protein [63]. Thus, to further validate our protocol and investigate the role of *Cripto* in gastruloid development, rescue experiments were performed by adding a recombinant soluble active form of Cripto protein (sCripto) [63, 69]. *Cripto* KO aggregomes were thus treated with sCripto at 10 µg/ml (from 24 to 48h AA) or left untreated as Control and the GFE was analysed at 120h AA (**Fig. 4.8A**). Interestingly, a significant fraction (>40%) of sCripto-treated *Cripto* KO aggregomes was able to generate elongated gastruloids (**Fig. 4.8B**). To further investigate this phenotype, we also analysed the expression pattern of different developmental associated genes that are induced during the aggregome to gastruloid transition in a time- and space- controlled manner, including *Brachyury* (T), *Cdx2*, *Sox2* and *Sox17* [8, 13, 57].

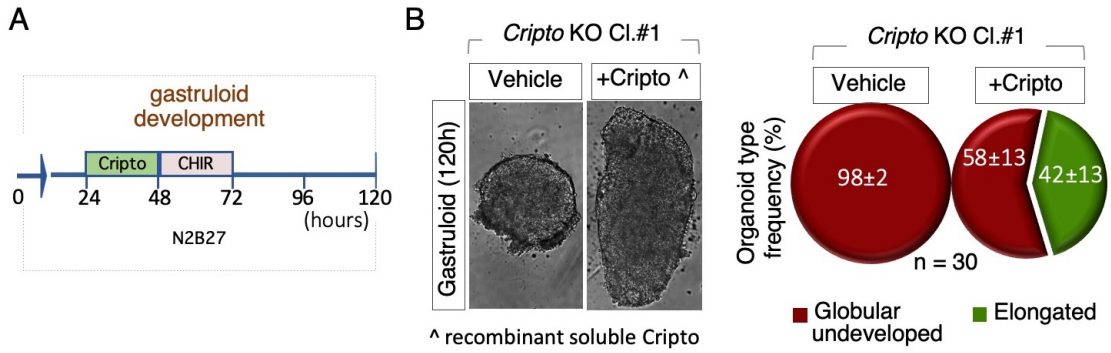


Figure 4.8. sCripto rescue elongation in *Cripto* KO gastruloids.
(A) Schematic representation of the rescue experiment. Soluble Cripto protein (sCripto, 10 µg/ml) was added at 24h AA.
(B) Bright field representative pictures (*left*) of *Cripto* KO rescued gastruloids at 120h AA. Pie chart of the percentage of organoid type frequency (*right*) in *Cripto* KO treated with sCripto or vehicle as Control.

Immunofluorescence analysis revealed a similar spatial pattern of cells expressing *Brachyury* (T) and *Cdx2* in fully elongated WT and *Cripto* KO rescued (*Cripto* KO + sCripto) gastruloids; which, conversely were almost absent in the *Cripto* KO undeveloped organoids (**Fig. 4.9A**). In the rescued mutant the expression domain of both genes was larger compared to that of WT gastruloids, suggesting that sCripto treatment expanded the posterior mesoderm fate of gastruloids, likely due to increased Nodal activity [70]. Interestingly, *Sox2* and *Sox17* were expressed at different extent in the undeveloped *Cripto* KO aggregomes (**Fig. 4.9A**). qPCR analysis performed on 120h AA organoids confirmed and extended these findings; for instance, expression of *Brachyury* and *Cdx2* was undetectable in *Cripto* KO organoids, whereas were expressed as usual in sCripto-treated *Cripto* KO gastruloids (**Fig. 4.9B**). Expression of the neural marker *Sox2* was significantly higher in *Cripto* KO than in Control (**Fig. 4.9B**), correlating with the presence of a large group of Sox2 positive cells in the central region of the undeveloped *Cripto* KO spheroids (**Fig. 4.9A**). All together these findings supported previous *in vivo* reports showing that *Cripto*-null mutant embryos mostly consist of anterior neuroectoderm and lack posterior structures (Ding et al., 1998), and provided evidence that *Cripto* is essential for the induction of symmetry breaking (A-P axis formation) in developing mESCs aggregomes.

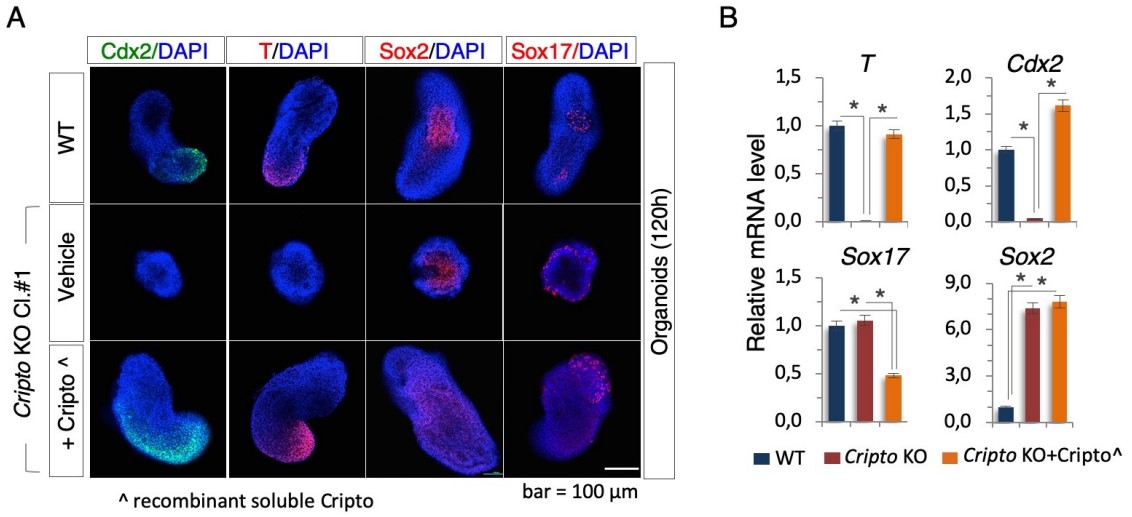


Figure 4.9. sCripto rescue *Cdx2* and *Brachyury* (*T*) expression in *Cripto* KO gastruloids.
(A) Representative confocal pictures of immunofluorescence analysis with Cdx2 (green), T (red), Sox2 (red) and Sox17 (red) on from WT and *Cripto* KO ± sCripto-derived organoids at 120h AA. Nuclei were counterstained with DAPI (blue) (bar=100 µm).
(B) qPCR analysis of *Cdx2*, *T*, *Sox2* and *Sox17* expression in WT and *Cripto* KO ± sCripto-derived organoids at 120h AA.

Cripto is an obligate co-receptor of Nodal [6], and to better define the role of Nodal signalling in the inability of *Cripto* KO cells to generate gastruloids, the effect of Activin A, a member of the transforming growth factor beta (TGF-β) family of proteins and a potent inducer of Nodal pathway [71], was investigated. Recent findings showed that Activin signalling can be induced in *Cripto* KO mESCs [5]. A transient pulse (24-48h AA) of Activin A (20 ng/ml) induced both the polarization/elongation of a significant fraction (~60%) of *Cripto* KO aggregomes (**Fig. 4.10A-B**) and the expression of *Brachyury* at the posterior end of the gastruloid (**Fig. 4.10C**). However, unlike WT gastruloids, the elongated gastruloid-like organoids developed by Activin A-treated *Cripto* KO aggregomes, were unable to induce the expression of *Cdx2* and *Sox17* markers (**Fig. 4.10C**) and displayed large zones of Nestin (neuronal precursor marker) expressing cells, thus suggesting a deregulation of neural differentiation (**Fig. 4.10C-D**).

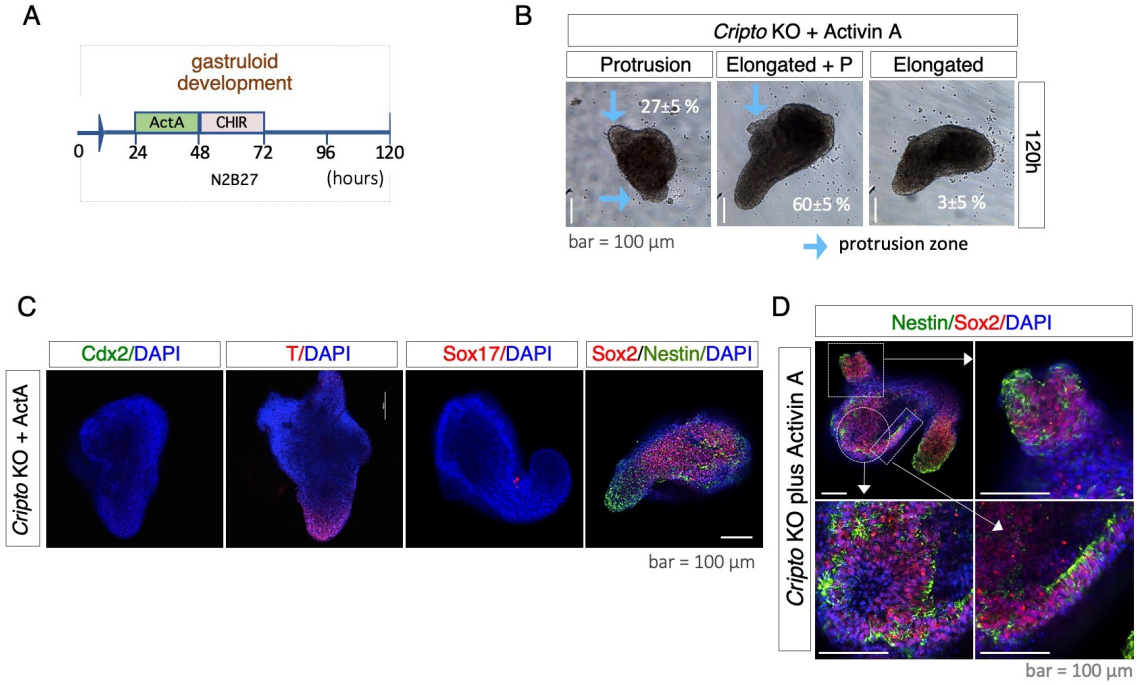


Figure 4.10. Activin A partially rescue *Cripto* KO gastruloids
(A) Schematic representation of the experimental design. Activin A (ActA) protein was added at 24h AA.
(B) Bright field representative pictures and percentage distribution of the different organoids generated by *Cripto* KO+ActA at 120h AA.
(C) Representative confocal pictures of immunofluorescence with Cdx2 (green), T (red), Sox17 (red) and Sox2/Nestin (red and green respectively) of *Cripto* KO organoids treated as indicated at 120h AA.
(D) Representative confocal pictures of immunofluorescence with Sox2 (red) and Nestin (green) on Cripto KO + ActA derived gastruloids at 120h AA. Nuclei were counterstained with DAPI (blue) (bar=100 µm).

Thus, Activin A supplementation, at least at the time and concentration used, rescued WT polarization/elongation but not proper gastruloid development of *Cripto* KO aggregomes. In conclusion, the abortive development of *Cripto* KO aggregomes validate the new experimental approach set up here and further support the idea that gastruloid development mimics early embryo gastrulation.

4.1.3. Primed pluripotent cells are unable to generate gastruloids

In order to investigate the impact of the pluripotency state on GFE, the naïve to primed transition was induced as previously described [72] by treating mESCs with a mixture of Fibroblast Growth Factor (12 ng/ml) and Activin A (20 ng/ml) (F/A mix), after LIF withdrawal. As schematized in **Figure 4.11A**, naïve (2i_LIF) TBV2 cells were seeded at low density (1500 cells/cm^2) on serum-coated plates and incubated in N2B27/1% Knockout Serum Replacement (KSR) medium supplemented with F/A. The cells were

cultured for two (2d_F/A) and five (5d_F/A) days and characterized at molecular and phenotypic level. First, the expression profile of different pluripotency and lineage specific markers was analysed by qPCR. As expected, 2d_F/A cells expressed lower levels of the pluripotency markers *Nanog*, *Sox2*, *Rex1* and *Dppa5a*, and higher levels of differentiation markers *Apelin* receptor (*Aplnr* or *ApJ*), *Cerberus* (*Cer1*), and *Brachyury* (*T*), compared to naïve ESCs (**Fig. 4.11B**). Conversely, 2d_F/A cells expressed higher levels of pluripotency and lower levels of differentiation markers compared to 5d_F/A cells (**Fig. 4.11B**).

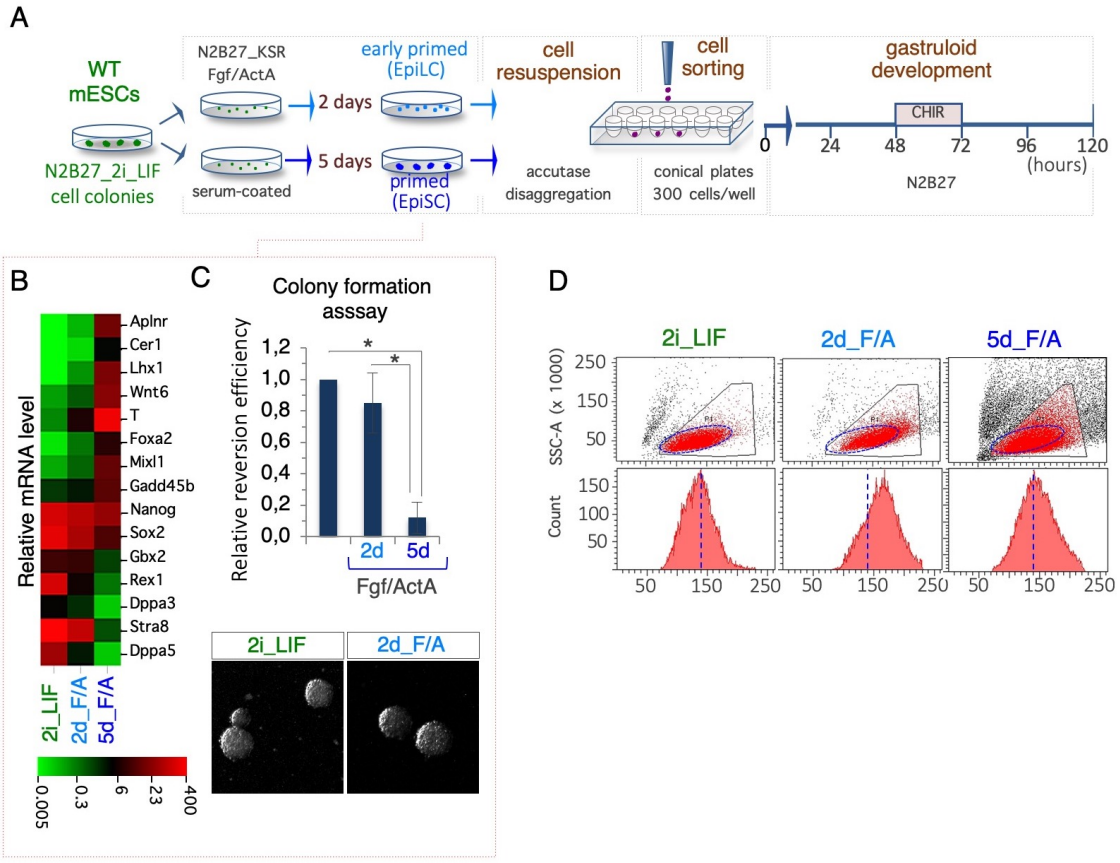


Figure 4.11. F/A primed pluripotent cells characterization.
(A) Schematic representation of the experimental design using EpiLCs (2d_F/A) and EpiSCs (5d_F/A) cells.
(B) Heat-map of mRNA levels of the indicated genes in naïve cells (2i_LIF), EpiLCs (2d_F/A) and EpiSCs (5d_F/A).
(C) Representative pictures of colony formation assay (*bottom*) and relative reversion efficiency level (*top*) of naïve cells, EpiLCs (2d_F/A) and EpiSCs (5d_F/A)
(D) Flow cytometry scatter plot (FSC vs SSC) and histogram of the 2i+LIF, EpiLCs (2d_F/A) and EpiSCs (5d_F/A) FACS-sorted population.

Second, the ability of primed cells to revert to the naïve state was assessed by performing colony formation assays (CFA). To this end, F/A-treated cells were

dissociated with accutase mix and plated at low density on gelatin-coated plates in naïve-inducing (N2B27/2i/LIF) medium. Under such conditions, 2d_F/A but not 5d_F/A cells, generated naïve-like domed round-shaped colonies (**Fig. 4.11C**). Thus, 2i_LIF (naïve, Control), 2d_F/A (early-primed EpiLCs) and 5d_F/A (primed EpiSCs) were used to perform the GFE assay using the optimized protocol. Naïve and primed cells showed remarkable differences in both the cytometric parameters (**Fig. 4.11D**), and the ability to generate aggregomes (**Fig. 4.12A**). Indeed, at 48h AA, both 2d_F/A (EpiLCs) and 2i_LIF (naïve) cells generated aggregomes similar in shape (spheroidal) (**Fig. 4.12A, *left***) and size (diameter mean ranging around 170-180 µm) (**Fig. 4.12A, *right***).

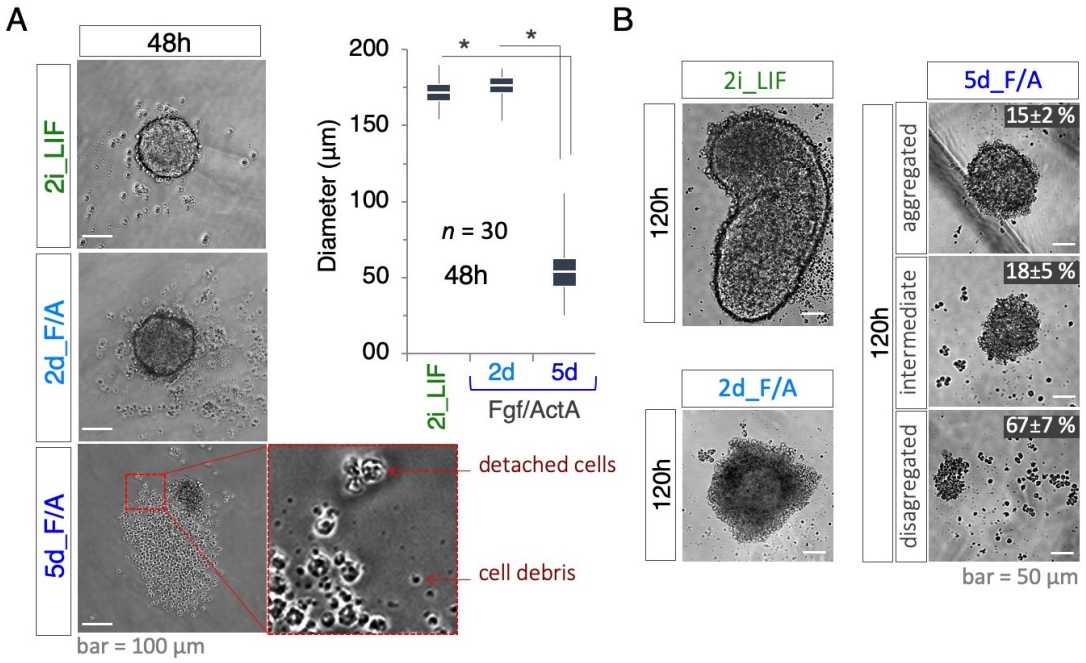


Figure 4.12. F/A primed pluripotent cells fails to generate gastruloids.
(A) Bright field representative pictures of different aggregomes at 48h AA (*left*) and box plot diagram (*right*) of diameter distribution. Red arrows indicate detached cells and debris (n=30; bar=100 µm).
(B) Representative bright field images of (*left*) 120h AA organoids and (*right*) different aggregomes phenotypes with percentage at the indicated conditions (*left*).

In contrast, the aggregomes generated by 5d_F/A (EpiSCs) were significantly smaller in size (~50 µm) (**Fig. 4.12A**). Later on, at 120h AA, while Control aggregomes developed fully elongated gastruloids (**Fig. 4.12B**), almost all the aggregomes generated by EpiLCs maintained a spheroidal morphology, with a central globular aggregate of cells surrounded by a disorganized and irregular tissue-like structure (**Fig. 4.12B**), suggesting that they failed to undergo proper aggregome to gastruloid transition. Finally, 5d_F/A

(EpiSCs) aggregomes, which were extremely small at 48h AA, became only abortive structures whose morphology ranged from spheroidal cell aggregates (~15%) to small irregular cell clusters (~70%) (**Fig. 4.12B**). Since none of the EpiLC- and EpiSC-derived aggregomes scored (n≥300) developed into an elongated gastruloid-like organoid, we concluded that 2 days incubation in primed state-inducing medium (N2B27_KSR plus F/A mix) and without LIF is sufficient to abolish the gastruloid formation ability of mESCs. Surprisingly, after transient LIF supplementation (from 0 to 48h AA) more than 70% of the EpiLCs aggregomes generated a gastruloid-like organoids at 120h AA (**Fig. 4.13A-C**). All together these data suggested that, under the experimental conditions used, primed stem cells are unable to generate elongated gastruloids (GFE value = 0%).

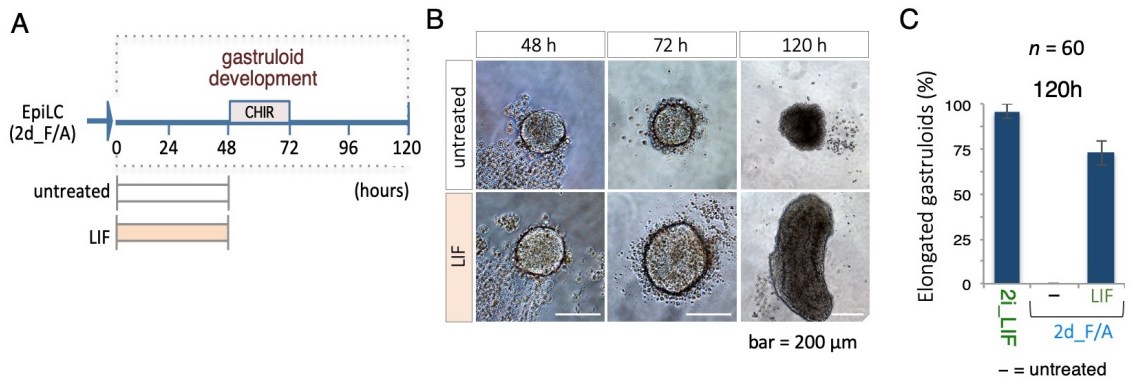


Figure 4.13. Rescue of EpiLCs-derived gastruloids.
(A) Schematic representation of the experimental design.
(B) Bright field representative pictures of time course of *Cripto* KO ESC-derived organoids showing that *Cripto* KO ESCs rescue the elongated phenotype after LIF treatment for 48h.
(C) Percentage (%) of elongated gastruloids of 2i+LIF (WT) and *Cripto* KO ± LIF derived organoids at 120h.

4.1.4. Primed pluripotent cells display a reduced propensity to aggregate

The precise dimension of the mESCs aggregomes (48h AA) is a critical parameter for gastruloid induction [8]; thus, the abortive development of the small aggregomes generated by 5d_F/A (EpiSCs) was almost expected. A reduced aggregome size can be a consequence of either a defective cell aggregation process or a disaggregation subsequent to a normal aggregation process. To discriminate these possibilities, aggregome generation was imaged at early time points. Recent findings reported that, under low-attachment conditions, mESCs (250-300 cells/well) spontaneously

aggregates in 8-10h [4]; in line with this observation, irregular-shaped cell aggregates were generated by 2i_LIF (naïve) TBV2 cells already at 7h after seeding (**Fig. 4.14A-B**).

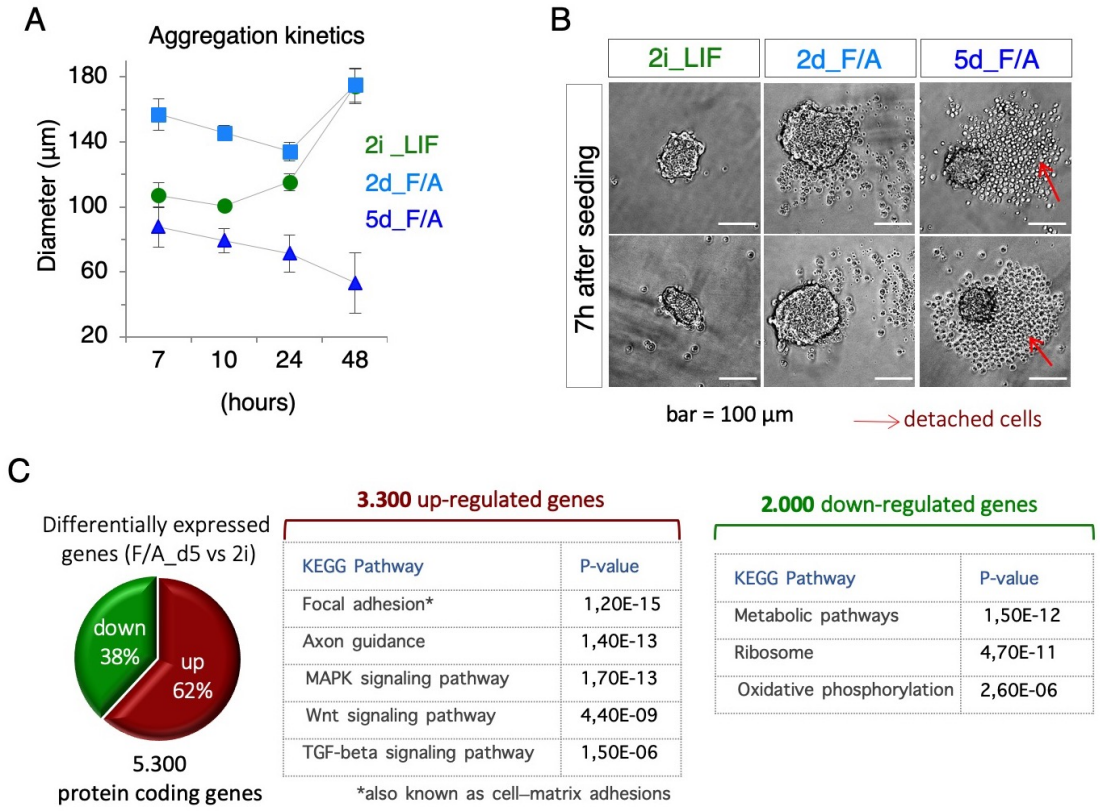


Figure 4.14. Aggregation kinetics analysis of 2i_LIF, d2_F/A and d5_F/A aggregates.
(A) Aggregation kinetics of naïve (2i_LIF), EpiLCs (2d_F/A) and EpiSCs (5d_F/A) aggregomes measured by diameter distribution (µm).
(B) Representative bright field images of aggregomes at 7h after seeding generated from the indicated cells. Red arrow shows detached cells.
(C) Pie chart of 5300 differentially expressed genes (d5_ F/A vs 2i_LIF) (*left*). KEGG pathway of 3300 up-regulated genes (*middle*) and 2000 down-regulated genes (*right*).

Later on (10h) the aggregates of naïve cells became highly compacted and displayed a spheroidal-shaped morphology (mean diameter ~100 µm) (**Fig. 4.14A**). Within this time window only few non-adherent free cells were observed, thus confirming the extraordinary propensity of naïve cells to spontaneously aggregate and generate stable cell-cell interactions. Similarly, d2_F/A (EpiLCs) early aggregomes (7h after seeding) were spherical, although less compacted (mean diameter ~150 µm) and surrounded by few non-adherent cells (**Fig. 4.14B**). By 10h onwards, the dimension of EpiLC (d2_F/A) aggregomes decreased, reaching its minimal (~135 µm) around 24h after seeding and without increasing the number of detached cells (**Fig. 4.14A**). Thus, we

speculated that the generation of gastruloid-developing aggregomes requires the stabilization of cell-cell contacts (compaction) in the initial cell aggregates. Conversely, at 7h after seeding, 5d_F/A (EpiSCs) aggregates were smaller in size (50-70 μm) (**Fig. 4.14A**) and, most relevantly, were surrounded by a large number of non-adherent cells and cellular debris (**Fig. 4.14B**). The analysis of a previously reported gene expression profile of 5d_F/A treated TBV2 cells (GEO accession: GSE84373) revealed that their reduced aggregation ability correlates with a significant up-regulation of focal adhesion related genes (**Fig. 4.14C**). Thus, during the ESC to EpiSC transition mESCs undergo a substantial loss of their capacity to generate functional aggregomes.

4.1.5. L-Proline-induced cells undergo premature gastruloid formation

We then evaluated the GFE of L-Proline induced cells (PiCs). PiCs were obtained by seeding naïve (2i_LIF) mESCs at low clonal density on gelatin-coated plates in DMEM/FBS medium supplemented with LIF and L-Proline (250-500 μM), as previously described [1, 2].

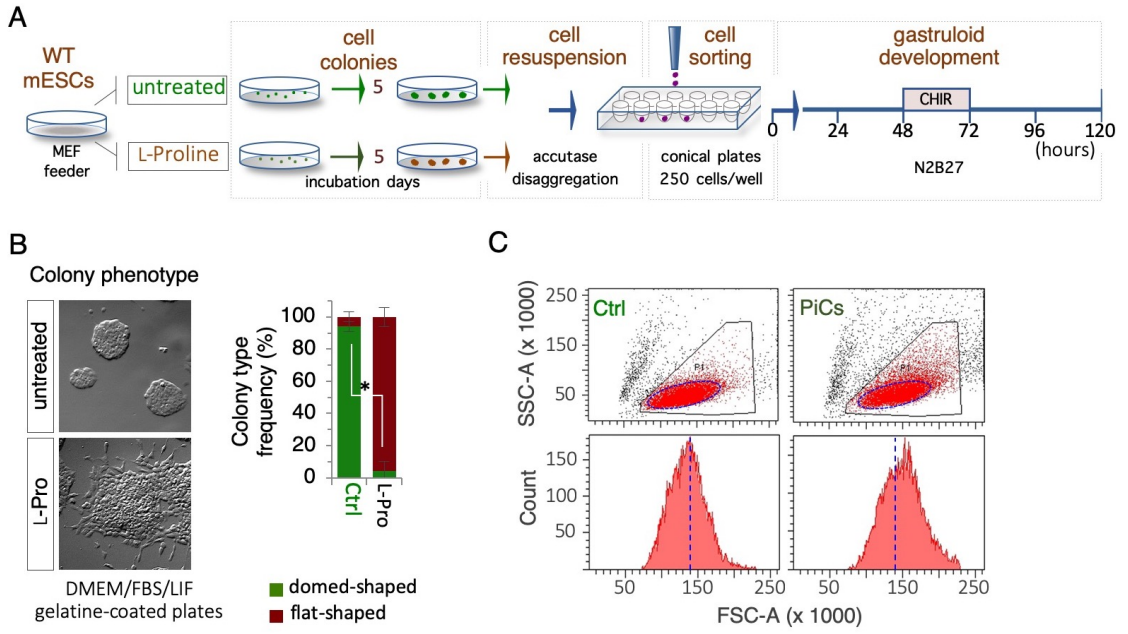


Figure 4.15. Set up of PiCs-derived gastruloids.
(A) Schematic representation of experimental scheme using L-Proline-treated cells (PiCs) and untreated control (Ctrl).
(B) Bright field representative pictures (*left*) and colony type quantification (*right*) of colonies generated from Control (Ctrl) ESCs and PiCs.
(C) Flow cytometry 2D scatter plot (FSC vs SSC) and histogram (FSC) of Ctrl and PiCs sorted population.

As expected, more than 90% of the cell colonies displayed a typical flat and irregular morphology after 4-5 days in culture. PiCs were then dissociated using the accutase mix and the resulting cells were seeded by FACS in 96-well plates and incubated to induced gastruloids formation (**Fig. 4.15A-C**). The diameter of two-day-old aggregomes (48h AA) generated by PiCs was smaller compared to Ctrl aggregomes (**Fig. 4.16A**).

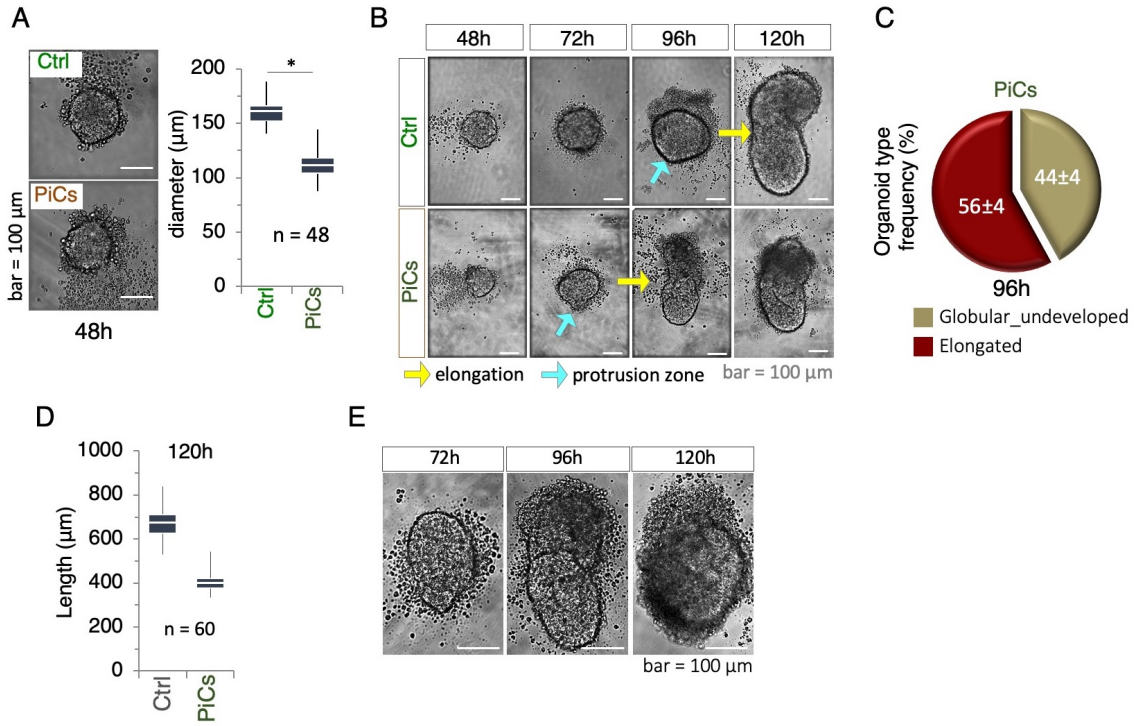


Figure 4.16. PiCs-derived organoids show a peculiar developmental process.
(A) Bright field representative pictures (*left*) of Ctrl and PiCs aggregomes at 48h AA and box plot diagram (*right*) of diameter distribution (n=48; bar=100 µm).
(B) Bright filed representative pictures of the aggregome to gastruloid transition of PiCs and Ctrl cells at the indicated time points (bar=100 µm). Light blue and yellow arrows indicate the protrusion zone, and the switch from ovoidal to elongated shape, respectively.
(C) Pie chart of PiCs derived organoids' types frequency.
(D) Length distribution of 120h Ctrl and PiCs-derived gastruloids.
(E) Bright field representative pictures of PiCs derived gastruloids at the indicated time points (bar=100 µm).

Remarkably, by 72h onwards, PiCs aggregomes started to undergo elongation; whereas, as expected, Control aggregomes (Ctrl) still exhibited a round-shaped morphology (**Fig. 4.16B**). At 96h AA, the majority of PiCs-derived organoids (~60%) displayed a gastruloid-like fully elongated shape, while the remaining (~40%) maintained the spheroidal morphology of the initial aggregates (**Fig. 4.16C**). Interestingly, the mean length of PiCs-derived gastruloids was smaller compared to the Control (**Fig. 4.16D**). At 120h AA a significant fraction (>50%) of PiCs-derived organoids showed a

disorganized structure surrounded by a large number of detached cells (**Fig. 4.16E**), suggesting that underwent early disaggregation compared to naïve-derived Control gastruloids. To evaluate if this peculiar developmental process (advanced lengthening, reduced dimension and early disaggregation) correlated with an altered expression of key developmental genes, we analyzed their expression pattern in PiCs elongated organoids (**Fig. 4.17A**).

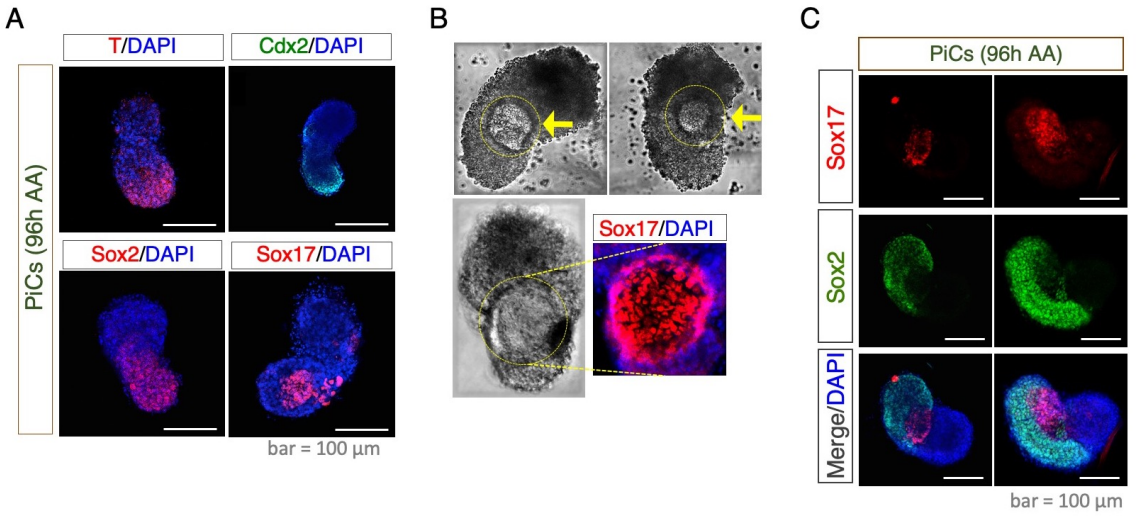


Figure 4.17. Characterization of L-Pro-derived gastruloids.
(A) Representative confocal pictures of immunofluorescence with T/Bra (red), Cdx2 (green), Sox2 and Sox17 (red) on PiCs derived gastruloids at 96h AA. Nuclei were counterstained with DAPI (blue) (bar=200 µm).
(B) Bright-field representative pictures of different PiCs derived gastruloids at 96h AA showing the Sox17 positive area (yellow circle).
(C) Representative confocal pictures of immunofluorescence with Sox17 (red) and Sox2 (green) on PiCs derived gastruloids at 96h AA. Nuclei were counterstained with DAPI (blue) (bar=100 µm).

Cdx2 and *Brachyury* genes were both expressed in specific territories of PiCs elongated organoids. Frequently, the PiCs-derived organoids exhibited a prominent round-shaped zone of histological heterogeneity, located in the middle/central region of the gastruloid, and stained positive for the mesoendodermal marker Sox17 (**Fig. 4.17B**). Interestingly, large areas of cells expressing the pan-neuronal marker Sox2 were observed adjacent to and surrounded the central region of Sox17 expressing cells (**Fig. 4.17C**). To better analyze their histological organization, resin-embedded naïve- and PiCs-derived gastruloids (120 and 96h AA respectively) were sectioned and stained with toluidine blue. Histological examination revealed that the gastruloids generated by naïve (2i_LIF) TBV2 cells (120h AA) displayed a tightly compacted structure mainly formed by

actively dividing cells. Interestingly, the gastruloids generated by PiCs showed the presence of actively dividing cells, as expected in a developing gastruloids, but also areas of highly differentiated cells/tissues (**Fig. 4.18A**). Immunofluorescence analysis of Ki67 confirmed the presence of actively dividing cells (**Fig. 4.18B**).

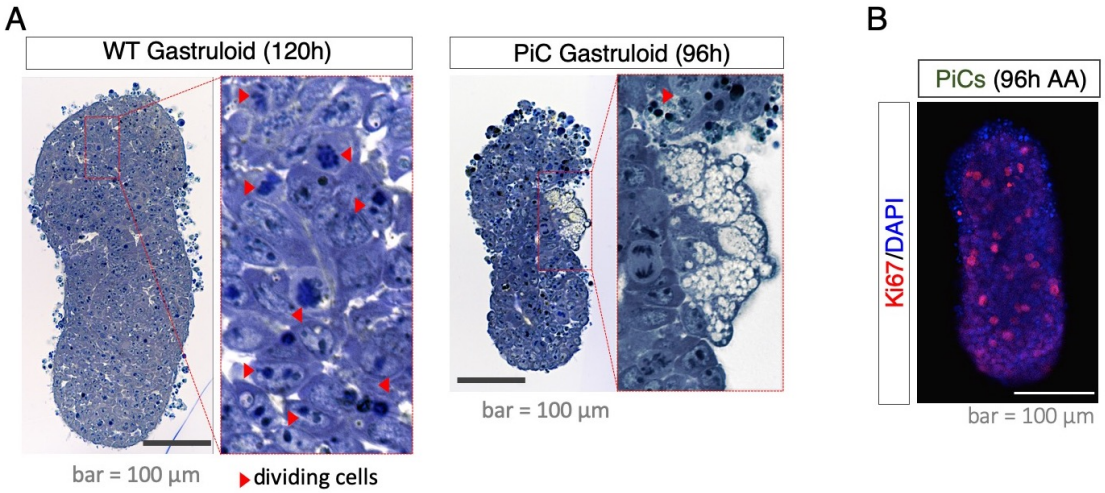


Figure 4.18. Histological analysis of gastruloids.
(A) Representative pictures of blue toluidine staining sections of WT and PiC derived gastruloids.
(B) Representative pictures of confocal immunofluorescence with Ki67 of PiCs derived gastruloid at 96h AA (red). Nuclei were counterstained with DAPI (blue) (bar=100 µm).

All together these findings support the idea that L-Pro supplementation in the presence of LIF capture an early primed and reversible state of pluripotency, which maintained the competence to generate gastruloid-like organoids.

4.2. PGCs differentiation

4.2.1. PiCs are competent to differentiate into precursors of primordial germ cells

The Primordial Germ Cells (PGCs) are the precursor of the gametes and are generated early during mouse development around E6.5. The capability to differentiate *in vitro* into primordial germ cell-like cells (PGCLCs) is a unique feature of EpiLCs (2d_F/A). Actually, neither naïve (2i_LIF or FBS_LIF) nor primed (EpiSCs, 5d_F/A) are able to undergo PGC differentiation [9, 23]. To better define the pluripotency state of PiCs we evaluated their ability/competence to generate PGCLC, using naïve and primed (EpiLCs and EpiSCs) pluripotent cells as control. Naïve (2i_LIF), early-primed (2d_F/A, EpiLC), primed (5d_F/A, EpiSC), and PiCs (L-Proline_LIF) were generated and their PGCLC formation capability was evaluated (**Fig. 4.19A**). Briefly, after seeding the

resulting cell aggregates (aggregomes) were incubated for 4 days of in the standard mix of PGC-inducing factors as previously described [65]. We first noted that the size of aggregomes vary among the different pluripotent states analysed (**Fig. 4.19B**).

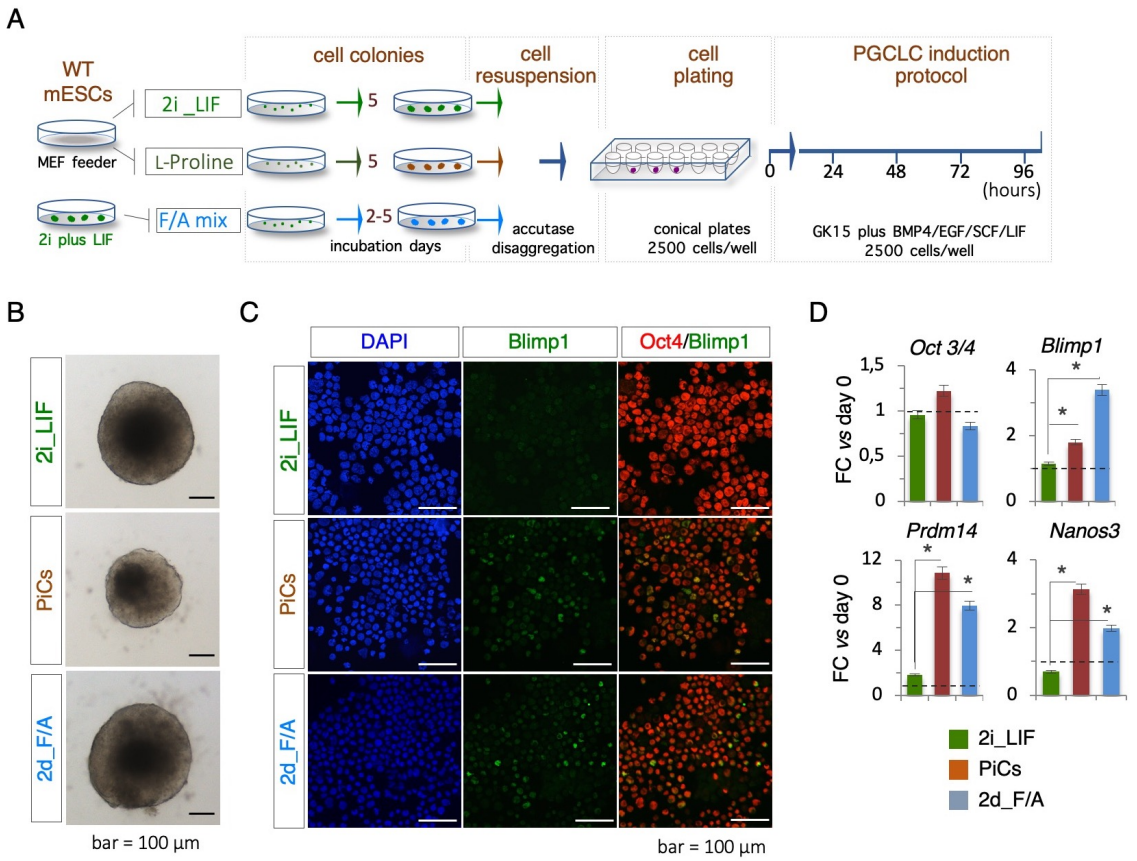


Figure 4.19. Primordial germ cell-like cells (PGCLCs) competence induction discriminate different pluripotent states.
(A) Schematic representation of the experimental strategy to induce PCGLCs differentiation using naïve (2i_LIF), EpiLCs (2d_F/A) and EpiSCs (5d_F/A).
(B) Bright field representative pictures of 96h cell aggregomes derived from naïve, PiCs, and EpiLCs cells, generated as indicated in panel (A) (bar=100 µm).
(C) Representative confocal pictures of immunofluorescence with Oct4 (red) and Blimp1 (green) on cytospined cells derived from aggregomes of naïve, PiCs and EpiLCs. Nuclei were counterstained with DAPI (blue) (bar=100 µm)
(D) RT-qPCR analysis of *Oct3/4*, *Blimp1*, *Prdm14* and *Nanos3* genes in the indicated conditions. Data are normalized to *Gapdh* and are mean ± SD and represent of fold change vs 0h (n=3; **p*<0.001).

In particular, the aggregomes generated by PiCs were smaller compared to those generated by naïve and early primed EpiLCs (**Fig. 4.19B**). Immunofluorescence analysis revealed the presence of *Blimp1*/*Oct4* double positive cells in aggregomes generated by EpiLCs and PiCs (**Fig. 4.19C**), thus, suggesting PGCLC differentiation [73]. To validate these results, we analysed the expression profile of different markers

including *Prdm14* and *Nanos3* by qPCR analysis. Remarkably, these genes, including *Blimp1* were significantly induced in EpiLCs- and PiCs-derived aggregomes (**Fig. 4.19D**). Moreover, we evaluated the expression of the transcription factor AP2γ (encoded by *Tfap2C* gene), a highly specific PGC marker [74]. Immunofluorescence analysis revealed that like EpiLCs-derived aggregomes, PiCs-derived aggregomes display a high percentage of AP2γ positive cells (**Fig. 4.20**).

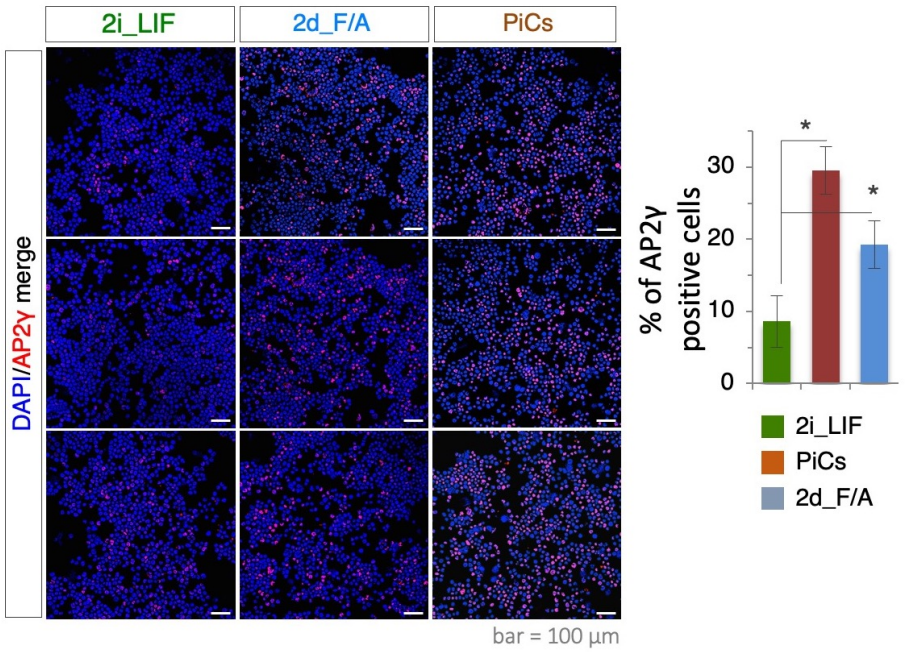


Figure 4.20. AP2γ different expression in naïve (2i), early primed (F/A d2) and PiCs cell aggregates. (*Left*) Representative confocal pictures of immunofluorescence with AP2γ (red). Nuclei were counterstained with DAPI (blue) (bar=100 µm); (*Right*) percentage (%) of AP2γ positive cells at indicated conditions (n=3; *p<0.01).

These findings provide further evidence that a high L-Proline regimen drives mESCs towards an early-primed state of pluripotency, which exhibits the unique competence for both gastruloid formation and differentiation into PGC-like cells.

4.2.2. PiCs gastruloids show a premature expression of PGCs markers

The ability of mESCs to differentiate into PGCLCs during gastruloid development was further assessed by using a Blimp1:GFP mESC line. To this end, naïve (2i_LIF) cells were seeded at 350 cells/well, and the resulting aggregomes were incubated and treated to induce gastruloid development. Co-expression of Blimp1/Oct4/Ap2γ was assessed by immunofluorescence analysis in 120h AA gastruloids (**Fig. 4.21A**), and revealed the

presence of Blimp1/Oct4/Ap2γ triple positive cells, which further indicate that gastruloids are competent to generate PGCLCs (**Fig. 4.21A**).

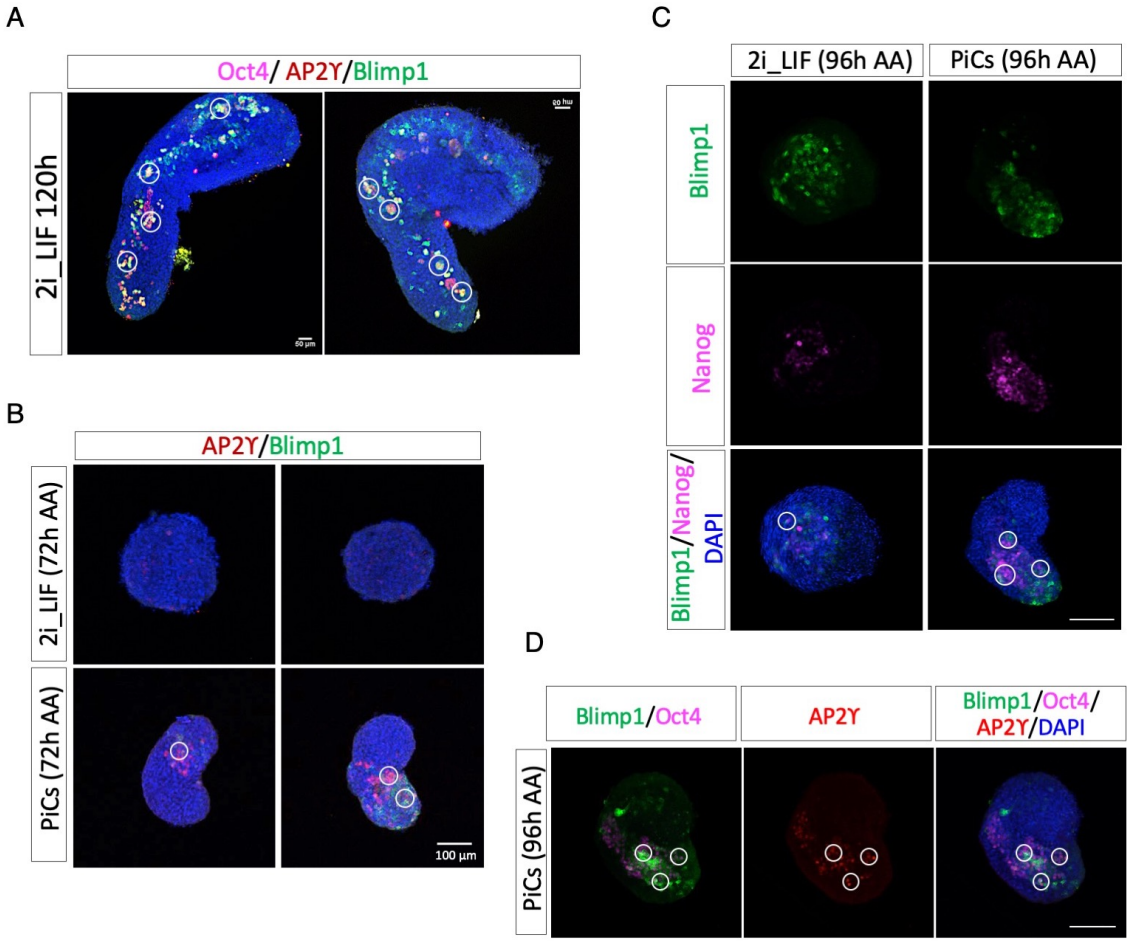


Figure 4.21. Formation of PGCLCs in gastruloids.
(A) Representative confocal pictures of immunofluorescence with AP2γ/Oct4 (magenta and red) in gastruloids at 120h AA using Blimp1:GFP cell line (bar = 50 µm).
(B) Representative confocal pictures of immunofluorescence with AP2γ (red) in 2i_LIF- and PiCs-derived gastruloids at 72h AA using a Blimp1:GFP (green) cell line. Nuclei were counterstained with DAPI (blue) (bar=100 µm).
(C) Representative confocal pictures of immunofluorescence with Nanog (magenta) in 2i_LIF and PiCs derived gastruloids at 96h AA using a Blimp1:GFP (green). Nuclei were counterstained with DAPI (blue) (bar=100 µm).
(D) Representative confocal pictures of immunofluorescence with Oct4 (magenta) and AP2γ (red) in PiCs-derived gastruloids at 96h AA using a Blimp1:GFP (green). Nuclei were counterstained with DAPI (blue) (bar=100 µm).

We then evaluated the presence of PGCLCs in PiCs-derived gastruloids. Cells that co-express Blimp1/Ap2γ were detected in PiCs-derived but not in Control (2i_LIF) gastruloids at 72h AA (**Fig. 4.21B**); which, however became evident also in Control gastruloids later on at 96-120h AA. Moreover, PiCs-derived gastruloids co-express Blimp1/Nanog and Blimp1/Oct4/Ap2γ positive cells at 96h AA (**Fig. 4.21C-D**). All

together these data suggest that ESCs are able not only to self-organize into embryo-like structures but also to differentiate into PGCLCs in the absence of exogenous added cytokines. Moreover, we found that PiCs gastruloids generate PGCLCs earlier compared to 2i_LIF.

PART II. Molecular mechanism underlying L-Proline effect on mESCs

4.3. Identification of inhibitors of L-Proline activity on mESCs

To characterize the molecular mechanisms regulated by a high L-Pro regimen, we first set up a fully automated unbiased high-throughput screening (HTS) in a 96-well format based on a robust cell colony phenotype assay (**Fig. 22A-B**).

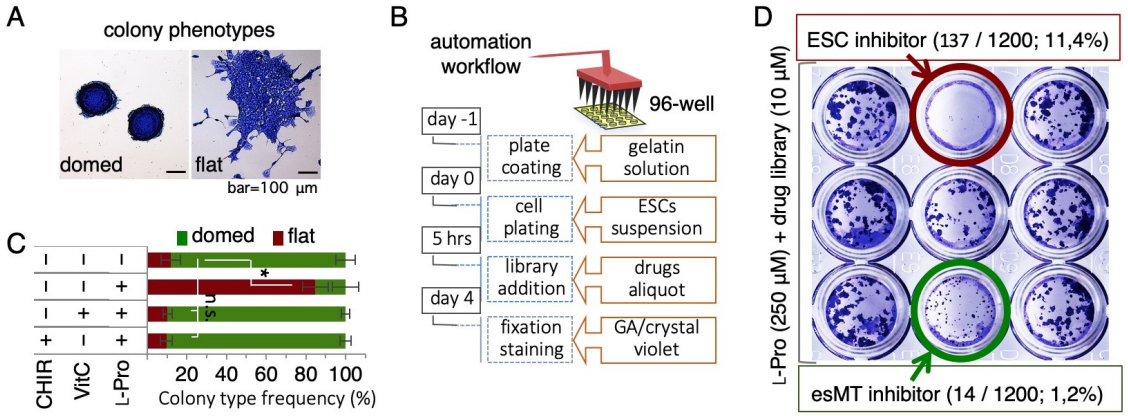


Figure 4.22. High-throughput screening-based identification of drugs that block esMT.
(A) Photomicrographs of domed (*left*) and flat (*right*) colony types.
(B) Schematic representation of the colony phenotype-based drug screening protocol
(C) Colony type quantification (domed vs. flat) of ESCs treated with L-Pro (250 µM) ± VitC (150 µM) or CHIR99021 (CHIR, 3 µM), or left untreated (day 5). Data are mean ± SD., *$p \leq 0.001$; ~100 colonies scored/condition n = 5; n.s. = nonsignificant.
(D) Representative pictures of the drug screening result. Red and green circles indicate ESC inhibitors (no colonies) and esMT inhibitors (domed-shaped colonies), respectively.

Both Ascorbic acid (Vitamin C, VitC) and CHIR99021 (GSK3 inhibitor, WNT agonist) were used as positive controls (esMT inhibitors). As expected, VitC and CHIR reduced the fraction of flat-shaped colonies to the control level (**Fig. 4.22C**). Briefly, mESCs were plated in the presence of L-Pro and exposed to 1,200 FDA-approved small molecules (Prestwick Chemical Library) in complete rich medium (FBS/LIF). After 4 days of incubation the resulting colonies were simultaneously fixed and stained with a glutaraldehyde/crystal violet mix and the colony shape (domed round *vs* flat irregular) was analyzed. The screening identified 137 drugs (**Table 1**), uniformly distributed throughout the plates that completely inhibited cell growth (ESC inhibitors, ESCi) and 14 drugs, which were able to counteract L-Pro effect and significantly reduce the

fraction of flat colonies (esMT inhibitors, esMTi). Among the ESCi the large majority (80%) have anticancer activities, while nearly 50% of the esMTi, including Budesonide, are inhibitors of collagen accumulation/fibrosis *in vitro* and/or *in vivo* (**Table 2**) [75].

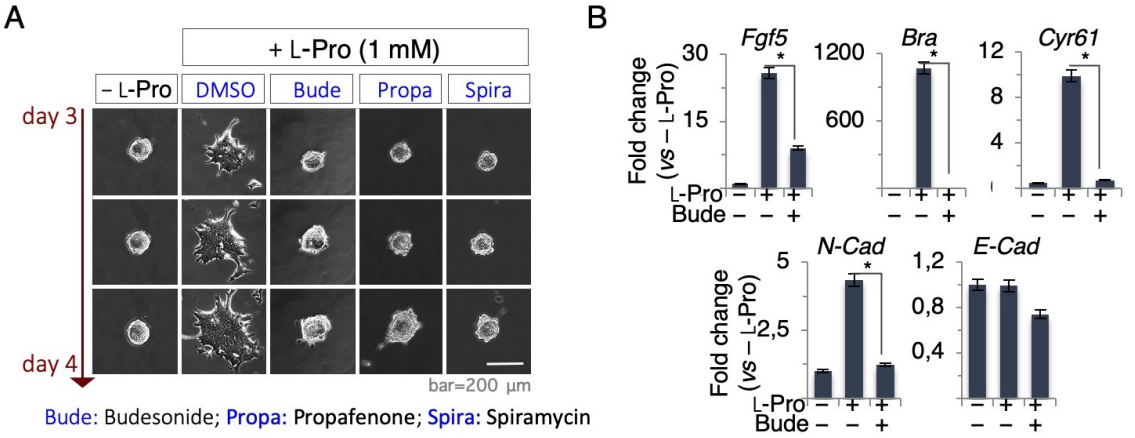


Figure 4.23. **Effect of Budesonide, Propafenone, Spiramycin on L-Pro–induced esMT.**
(A) Representative frames of time lapse series from untreated (-L-Pro) and + L-Pro + Budesonide (Bude) or Propafenone (Propa) or Spiramycin (Spira). Images were captured from day 3 to 4 post plating. DMSO was used as control.
(B) qPCR analysis of selected esMT markers [*Fgf5, Cyr61, Brachyury (Bra), N-Cadherin (N-Cad) e E-Cadherin (E-Cad)*] in ESCs treated with L-Pro (1 mM) ± budesonide (10 µM) or left untreated as Control. mRNA expression levels were normalized to *Gapdh*. Data are shown as fold change *vs* untreated and are mean ± SEM (**p*≤0.01; n=3).

Three drugs, namely Budesonide (Bude), Propafenone (Propa), and Spiramycin (Spira) were selected and further validated in cell motility and colony formation assays (**Fig. 4.23A-B**).

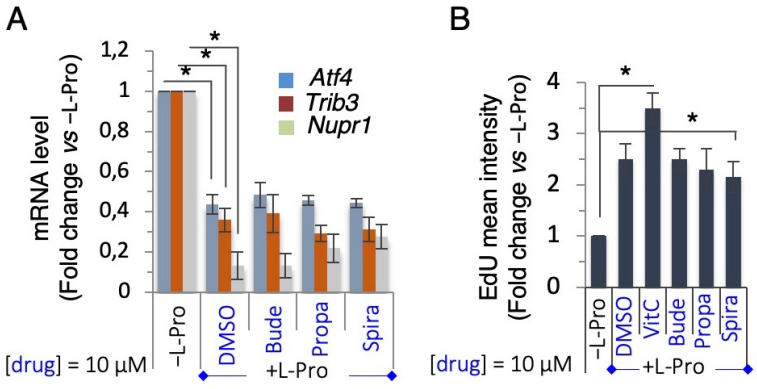


Figure 4.24. L-Proline effect on AAR and proliferation was not prevented by Budesonide, Propafenone, Spiramycin.
(A) qPCR analysis of *Atf4*, *Trib3*, and *Nupr1* expression. Data are fold change versus untreated after normalization to *Gapdh* and are mean ± SEM (*P ≤ 0.01; n = 3)
(B) FACS analysis of EdU incorporation. Data are shown as fold change versus untreated and are mean ± SEM (*P ≤ 0.01; n = 3)

Remarkably, Budesonide, Spiramycin, and Propafenone did not interfere either with L-Pro-induced downregulation of amino acid stress response (AAR) pathway markers (**Fig. 4.24A**) nor with cell proliferation (**Fig. 4.24B**). These results suggested that esMTi block L-Pro-induced esMT acting downstream of L-Pro-tRNA loading.

4.4. esMT is induced and sustained by collagen accumulation

Given that L-Pro and 4-*trans*-hydroxyproline (HyP) are the major components of collagen, and that around 50% of the esMTi are inhibitors of collagen accumulation, we hypothesized that L-Pro supplementation may increase collagen synthesis/hydroxylation in ESCs, and that collagen accumulation and esMT are functionally linked. In line with our idea, Collagen1a1 (Col1a1) and collagen hydroxyproline (C-Hyp) levels both increased in L-Pro-treated ESCs, accumulating primarily in the endoplasmic reticulum (ER) (**Fig. 4.25A-B**).

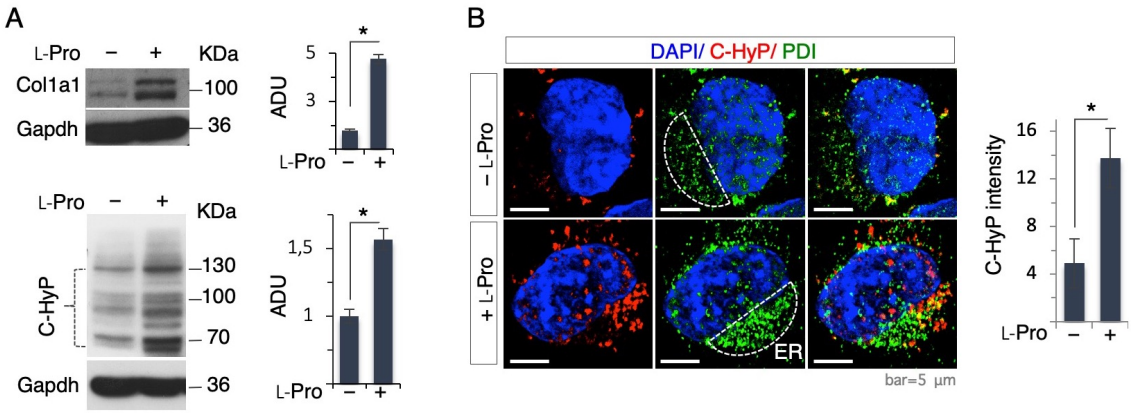


Figure 4.25. L-Proline supplementation induce collagen accumulation in ESCs.
(A) Western blot analysis of Col1a1 (*top*) and collagen hydroxyproline (C-HyP; *bottom*) in ESCs ± L-Pro (1 mM). Densitometric analysis (ADU) is shown as fold change versus untreated ESCs. Data are mean ± SEM (* $p \leq 0.01$; n = 3) after normalization to Gapdh.
(B) Confocal images of C-HyP (red) and Protein Disulfide Isomerase (PDI; green) immunofluorescence on ESCs ± L-Pro (*left*). Half circles, endoplasmic reticulum (ER). Quantification of C-HyP intensity (*right*). Data are mean ± SD (* $p \leq 0.01$; n = 3).

Interestingly, the esMTi Budesonide prevented L-Pro induced Col1a1 and C-Hyp RNA and protein (**Fig. 4.26A-D**). Together all these results supported the hypothesis that L-Pro-induced esMT depends on L-Pro-mediated induction of collagen synthesis/hydroxylation and that Budesonide counteract this effect (**Fig. 4.26E**).

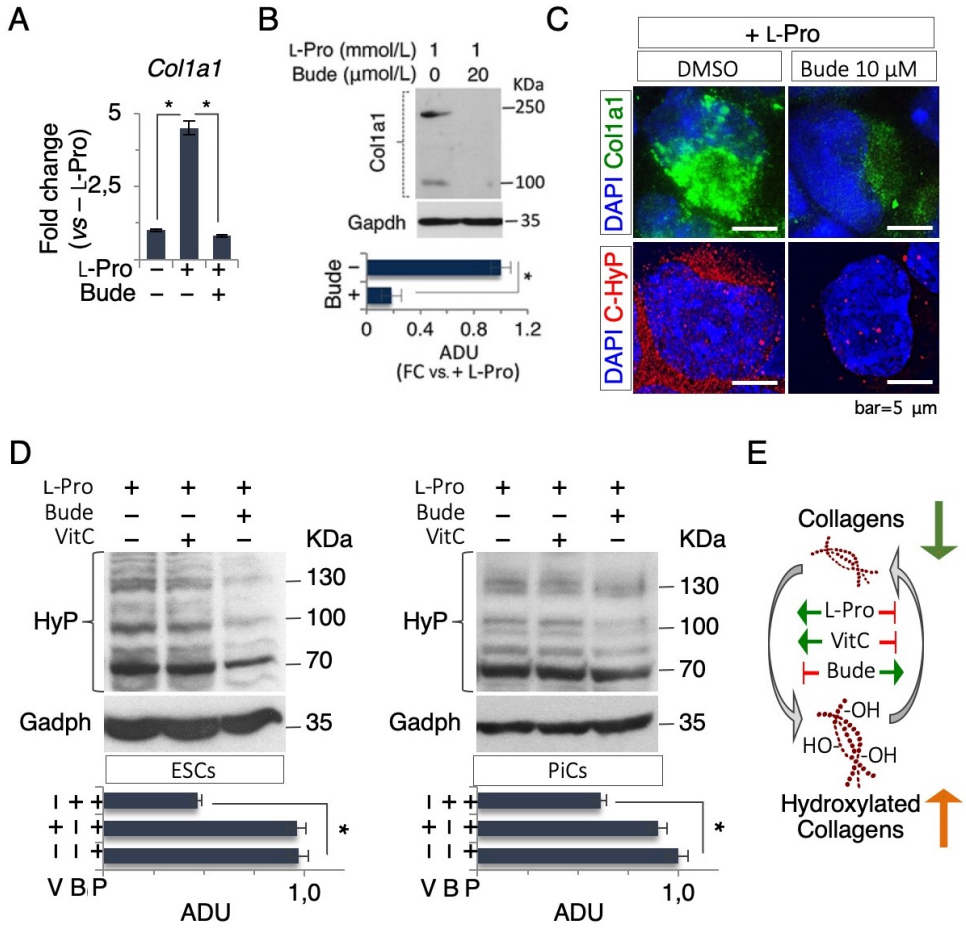


Figure 4.26. Molecular characterization of budesonide effect on L-Pro-induced esMT.
(A) qPCR analysis of *Col1a1* in ESCs treated with L-Pro (1 mM) ± Budesonide (10 µM) or untreated (–) as Control. mRNA expression levels were normalized to *Gapdh*. Data are shown as fold change *vs* untreated cells and are mean ± SEM (*$p \leq 0.01$; n=3).
(B) Western blot analysis of Col1a1 in ESCs + L-Pro (1 mM) ± Budesonide (10 µM). Densitometric analysis (ADU) is shown as fold change versus ESCs + L-Pro. Data are mean ± SEM (*$p \leq 0.01$; n=3).
(C) Fluorescence pictures of Col1a1 (green) and confocal images of C-HyP (red) immunofluorescence on L-Pro–treated cells ± Budesonide. Nuclei were counterstained with DAPI.
(C) Western blot analysis of C-HyP in ESCs (*left*) and PiCs (*right*) treated with L-Pro (1 mM) ± Budesonide (10 µM) or VitC (500 µM). ADU is shown as C-HyP/Gapdh. Data are mean ± SEM. (*$p \leq 0.01$; n=3).
(D) Schematic representation of L-Pro, VitC, and Budesonide effects on C-HyP.

4.5. Genetic inhibition of collagen hydroxylation prevents esMT

To test our hypothesis and given that it was not feasible to knock-out all the members of the large family of collagen coding genes, we interfered with collagen hydroxylation by targeting the prolyl-4hydroxylases (P4h). P4h enzymes are members of VitC/Fe^{+2}/KG-dependent dioxygenases, and form a tetrameric complex able to catalyze the hydroxylation of L-Pro residues in nascent collagens in the ER. The alpha subunit of

Collagen prolyl hydroxylase is encoded by three genes: *P4ha3* is not expressed in ESCs and PiCs, whereas *P4ha1* and *P4ha2* are expressed at comparable levels (**Fig. 4.27A**).

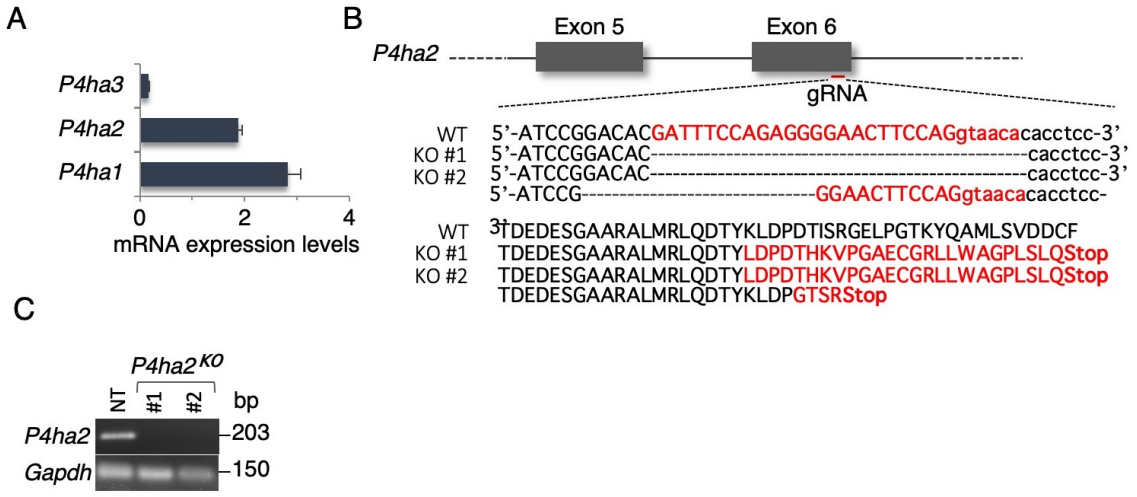


Figure 4.27. Generation of *P4ha2*^{KO} ESCs.
(A) qPCR analysis of *P4ha1, 2* and *3* expression in ESCs. Data are normalized to *Gapdh* and are mean±SEM (n=3).
(B) CRISPR/Cas9 mediated *P4ha2* knockout. Two independent *P4ha2*^{KO} ESC clones were selected (KO #1 and #2) carrying different mutations that lead to a frameshift and the formation of a premature stop codon. *P4ha2* KO #1 carries a deletion of 29 bp spanning nucleotide 13328 to 13357; *P4ha2* KO #2 carries a deletion of 29 bp spanning nucleotide 13328 to 13357 on allele 1, and a deletion of 21 bp, spanning nucleotide 13323 to 13344, on allele 2. Numbering is based on the entire *P4ha2* gene sequence using the reference gene RefSeq NM_001136076. DNA (*top*) and protein (*bottom*) sequences of wild-type and mutated alleles (*P4ha2*^{KO} #1 and #2) are shown. Upper and Lower letters represent exons and introns, respectively.
(C) RT-PCR of *P4ha2* from CRISPR-Cas9 nontargeting (NT) and *P4ha2*^{KO} ESCs.

Interestingly, P4HA2 has been recently found to be involved in human breast cancer progression [76]. We thus first inactivated *P4ha2* in ESCs by CRISPR/Cas9 and selected two independent clones (*P4ha2*^{KO} #1 and #2) with two different genomic deletions between the exon 6 and the intron 7 (**Fig. 4.27B**). These deletions generated a frameshift and the formation of a premature stop codon (**Fig. 4.27C**).

To avoid the possibility that *P4ha1* may at least partially compensate for the lack of *P4ha2*, we knocked down *P4ha1* in *P4ha2*^{KO} ESCs using two shRNAs, which target nonoverlapping *P4ha1* mRNA sequences (**Fig. 4.28A and B**). The *P4ha2*^{KO} ESC colonies (*P4ha2*^{KO} #1 and #2) kept a domed shaped phenotype, were AP positive and preserved the expression of pluripotency markers (**Fig. 4.29A-C**). This was similarly observed in *P4ha2*^{KO}/*P4ha1*^{KD} ESCs (**Fig. 4.29A-C**).

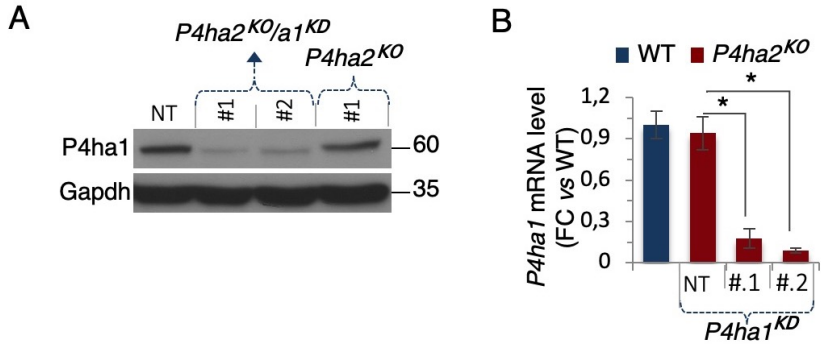


Figure 4.28. Generation of *P4ha2*$^{KO/KD}$ ESCs.
(A) Western blot analysis of P4ha1 levels in control (NT), *P4ha2*^{KO}/a1^{KD} and P4ha2^{KO} ESCs. Anti-Gapdh antibody was used as loading control.
(B) qPCR analysis of *P4ha1* in Wild-Type (WT), Non-Targeting (NT), *P4ha2*^{KO}/*a1*^{KD} and in *P4ha2*^{KO} ESCs. Data are normalized to *Gapdh* and shown as fold change *vs* WT cells. Data are mean±SEM (*p<0.01; n=3).

We then examined the effect of *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} on L-Pro–induced esMT. Control ESCs (NT) retained high frequency of atypical colonies in response to L-Pro, whereas *P4ha2*^{KO} ESCs mostly developed typical domed colonies lacking the crown of mesenchymal-like cells scattered around the colony core (**Fig. 4.30A**).

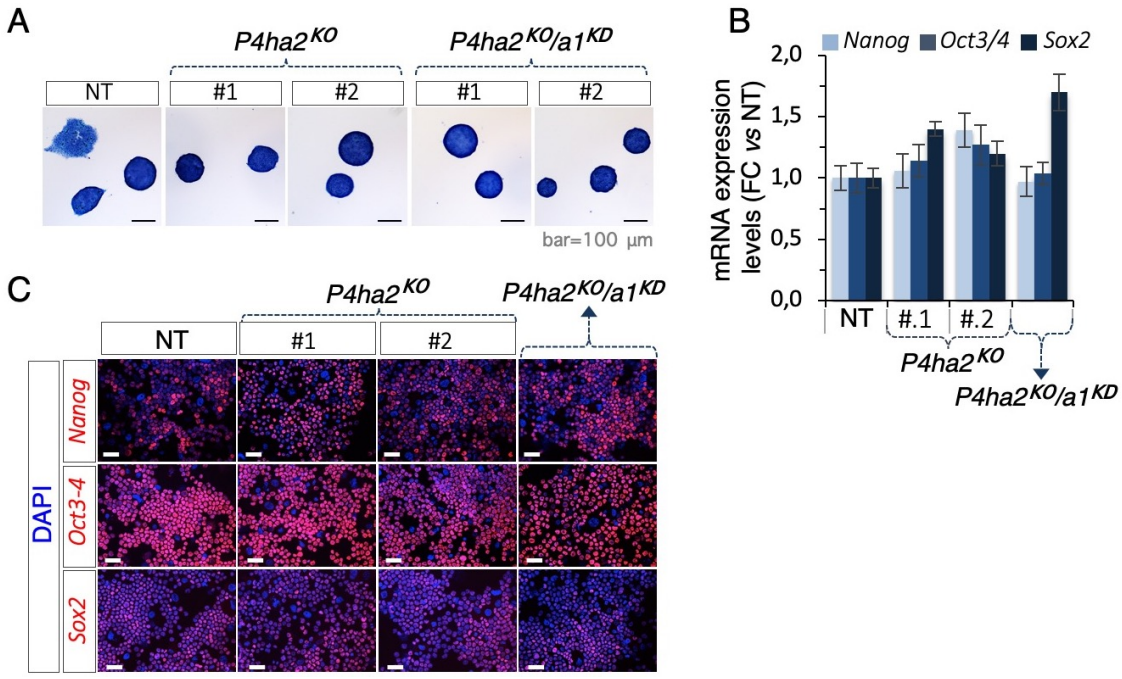


Figure 4.29. Pluripotency characterization of *P4ha2*^{KO} and *P4ha2*$^{KO/KD}$ ESCs
(A) Representative pictures of AP staining of NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESC colonies.
(B) qPCR analysis of *Nanog*, *Oct3/4* and *Sox2* pluripotency markers in NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs. Data are normalized to *Gapdh* and are shown as fold change *vs* NT cells. Data are mean±SEM (n=3).
(C) Representative pictures of Nanog, Oct4 and Sox2 immunofluorescence on NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs. Nuclei are counterstained with DAPI.

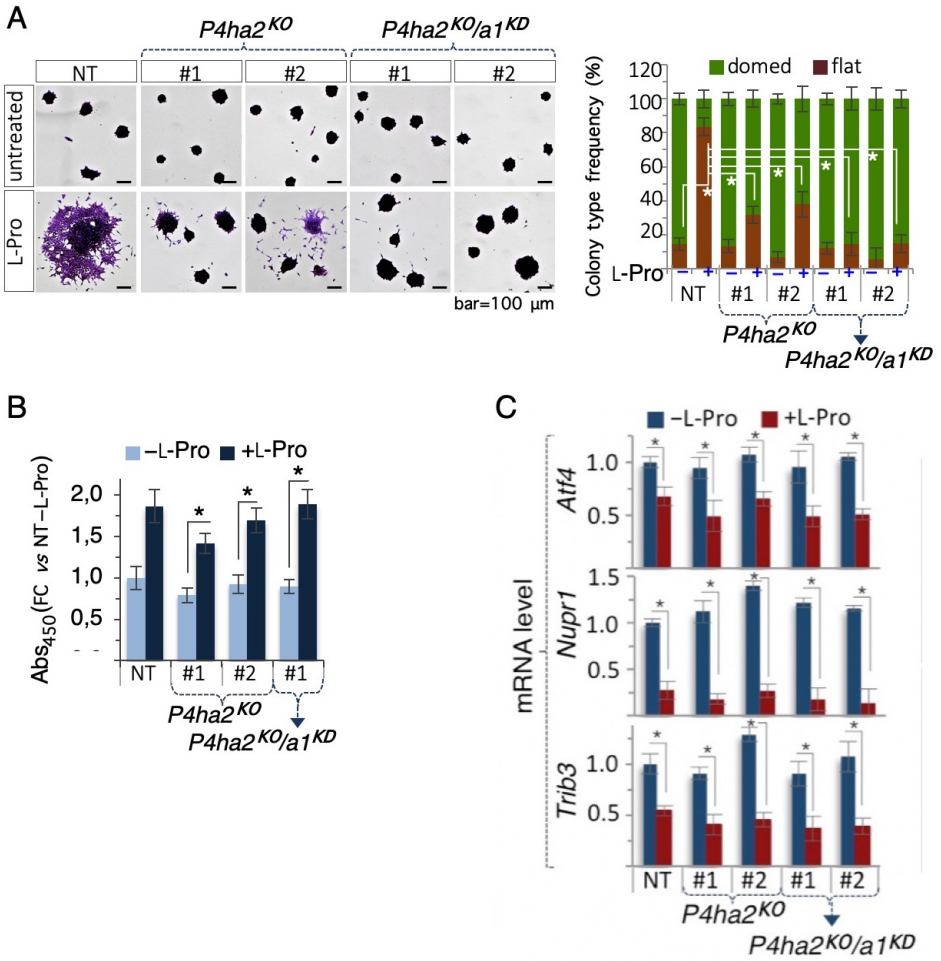


Figure 4.30. Fuctional characterization of *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs.
(A) Representative pictures (*left*) and colony type quantification (*right*) of colonies generated from NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM) at day 5. Data are mean ± SD. (*$p \leq 0.01$, ~100 colonies scored/condition; n = 5).
(B) L-Pro–induced proliferation of NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM) analyzed by CCK-8 assay. Data are shown as fold change versus NT minus L-Pro. Data are mean ± SEM. (*$p \leq 0.01$; n = 3).
(Cs) qPCR analysis of *Atf4* and selected Atf4-target genes (*Nupr1* and *Trib3*) in NT, *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro for 24 hours. mRNA expression levels were normalized to *Gapdh*. Data are shown as fold change *vs* untreated NT cells and are mean ± SEM (*$p \leq 0.01$; n=3).

Interestingly, this phenotype was robustly enhanced in the two independent *P4ha2*^{KO}/*P4ha1*^{KD} clones (**Fig. 4.30A**), indicating that *P4ha1* expression partially compensates the lack of *P4ha2*. In P4h-deficient ESCs also the cell motility was impaired. Conversely, both L-Pro–induced proliferation (**Fig. 4.30B**) and alleviation of AAR pathway (**Fig. 4.30C**) was preserved, indicating that P4h activity is specifically required for L-Pro–induced esMT rather than for a general activity of L-Pro. Finally, as expected, in P4h-deficient ESCs, the C-Hyp accumulation was strongly reduced (**Fig. 4.31**).

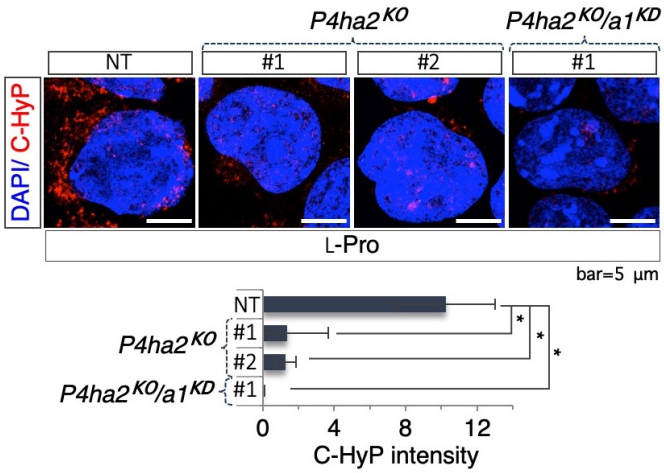


Figure 4.31. Effect of *P4ha2*^{KO} and *P4ha2*^{KO}/*a1*^{KD} on Collagen Hydroxyproline (C-HyP) accumulation in ESCs.
(Top) confocal images of C-HyP immunofluorescence on NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs treated with L-Pro (500 µM). Nuclei were counterstained with DAPI. (Bottom) C-Hyp intensity quantification. Data are mean ± SD. (*$p \leq 0.01$; n = 3).

Together, these findings show that the genetic ablation of *P4h* prevents collagen hydroxylation and antagonizes esMT mimicking the effects of budesonide, and supporting the idea that the induction of collagen synthesis/hydroxylation is essential for the induction of esMT.

4.6. Collagen hydroxylation influences the epigenetic landscape of mESCs

P4h is a dioxygenase enzyme that requires VitC to be active, *i.e.* to catalyze the hydroxylation of L-Pro residues in collagens. VitC is essential also for the activity of other VitC/Fe^{+2}/KG-dependent dioxygenases, which localize in different cellular compartments and have different substrates. For instance, Tet enzymes catalyzes the DNA hydroxylation/demethylation in the nucleus, while in the same cellular compartment the Jmj enzymes catalyzes histones (Jmj) hydroxylation/demethylation. We thus hypothesized that a sudden increase of collagen synthesis might channel VitC to the ER for the activity of collagen hydroxylases (P4hs). This would concomitantly provoke its stoichiometric reduction in other cellular compartments, which translates into a reduced activity of other VitC-dependent dioxygenases. Of note, such a substantial perturbation of VitC homeostasis might occur only when collagen synthesis is induced in a VitC-limiting environment. To test this hypothesis, we assessed the impact of *P4h* KO on the global levels of Tet-dependent DNA hydroxymethylation (5hmC). Indeed, in cells subjected to a high-L-Pro/low-VitC regimen, the level of 5hmC rapidly decreases in Control but not in P4h-deficient ESCs (**Fig. 4.32A-B**).

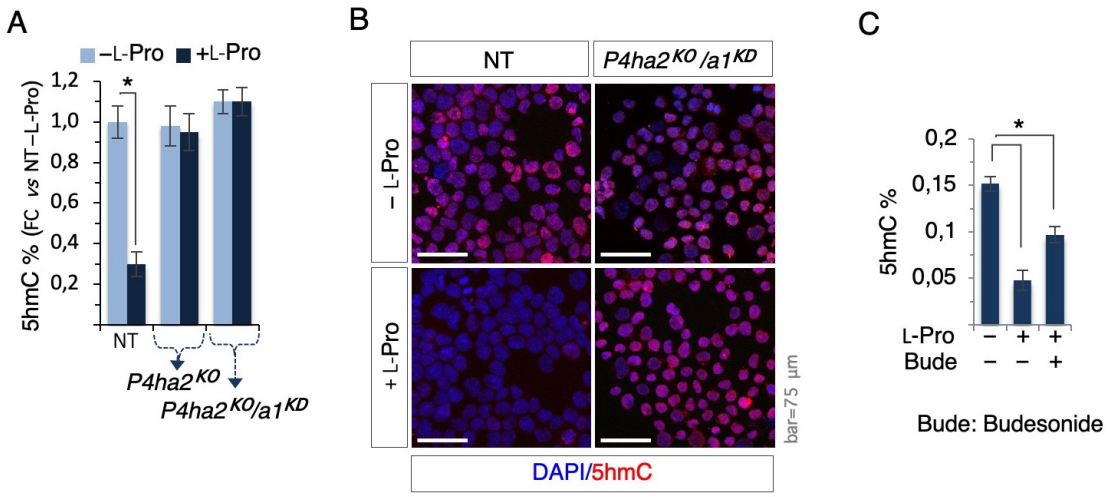


Figure 4.32. Effect of *P4ha2*^{KO} and budesonide on 5hmC in ESCs ± L-Pro.
(A) Elisa-based quantification of 5hmC in control (NT), *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM; 24 hours). 5hmC levels are shown as fold change versus NT minus L-Pro. Data are mean ± SEM. (**p* ≤ 0.01; n = 3).
(B) Representative pictures of 5hmC immunofluorescence on NT and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM). Nuclei are counterstained with DAPI.
(C) ELISA-based quantification of 5hmC levels in wild-type ESCs ± L-Pro (1mM), ± Budesonide (Bude, 10 µM) for 24 hours. Data are mean ± SEM (**p*≤0.01; n=3).

Budesonide addition counteracted the reduction of 5hmC in WT mESCs (**Fig. 4.32C**).

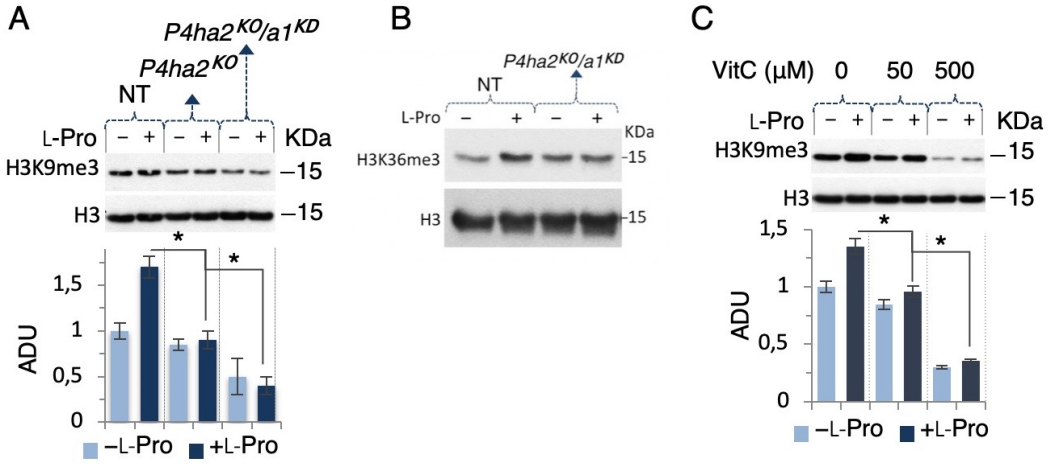


Figure 4.33. Effect of *P4ha2*^{KO} on histones methylation.
(A) Western blot analysis of H3K9me3 in NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM). ADU is shown as fold change versus untreated (minus L-Pro) controls after normalization to total histone H3. Data are mean ± SEM (**p*≤0.01; n=3).
(B) Western blot analysis of H3K36me3 in NT, *P4ha2*^{KO}, and *P4ha2*^{KO}/*a1*^{KD} ESCs ± L-Pro (500 µM). ADU is shown as fold change versus untreated (minus L-Pro) controls after normalization to total histone H3. Data are mean ± SEM (**p*≤0.01; n=3)
(C) Western blot analysis of H3K9me3 in wild-type ESCs ± L-Pro (500 µM) ± VitC. ADU is shown as fold change versus untreated (minus L-Pro) controls after normalization to total histone H3. Data are mean ± SEM (**p*≤0.01; n=3).

We then evaluated the impact of *P4h* KO on Jmj dioxygenases, which catalyze the removal of H3K9/36 methylation, and also assessed whether VitC supplementation may blunt the effect of L-Pro supplementation. In a high-L-Pro/low-VitC regimen, the global levels of both H3K9me3 and H3K36me3 increased in Control but not in *P4ha2*^{KO} and *P4ha2*^{KO}/*P4ha1*^{KD} mESCs (**Fig. 4.33A-B**). Consistent with the idea that VitC is a limiting metabolite in ESCs [3], VitC supplementation not only prevented L-Pro–dependent increase of H3K9me3 level but also further reduced H3K9me3 global levels in a low L-Pro environment (**Fig. 4.33C**). Interestingly, Budesonide mimicked the antagonistic effect of VitC, increasing of H3K9me3 levels (**Fig. 4.34A**).

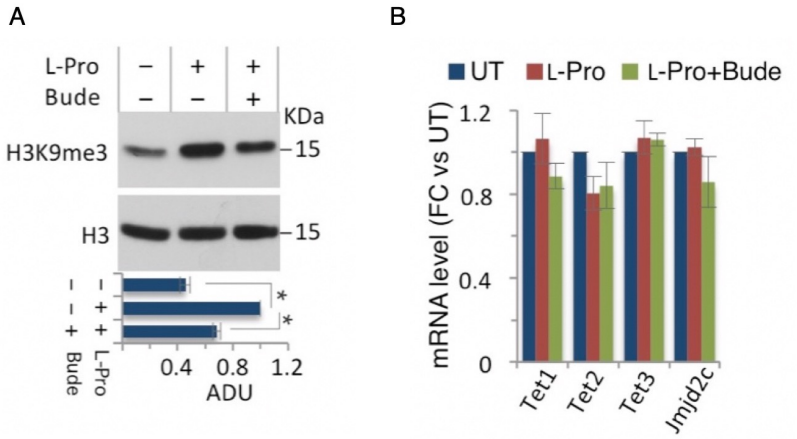


Figure 4.34. Effect of Budesonide on histones methylation.
(A) Western blot analysis of H3K9me3 in ESCs ± L-Pro (500µM) ± Budesonide (10µM). Anti-H3 antibody was used as a control.
(B) qPCR analysis of *Tet1, 2, 3 and Jmdj2c* expression in ESCs ± L-Pro (500µM) ± Budesonide (10µM). Data are normalized to *Gapdh* and are mean ± SEM (n=3).

The expression of specific DNA/histone demethylases genes, including *Tet1*, *Tet2*, *Tet3* and *Jmjd2c* (H3K9me2/me3 and H3K36me2/me3 demethylation; **Fig. 4.34B**) was not modified by L-Pro or Budesonide treatment. These results indicate that the epigenetic changes observed are not dependent on the expression of the enzymes. These findings suggest that increased P4h-dependent collagen hydroxylation increase DNA and histones methylation levels, most likely by altering VitC homeostasis and reducing the activity of nuclear hydroxylases/demethylases (**Fig. 4.35A**).

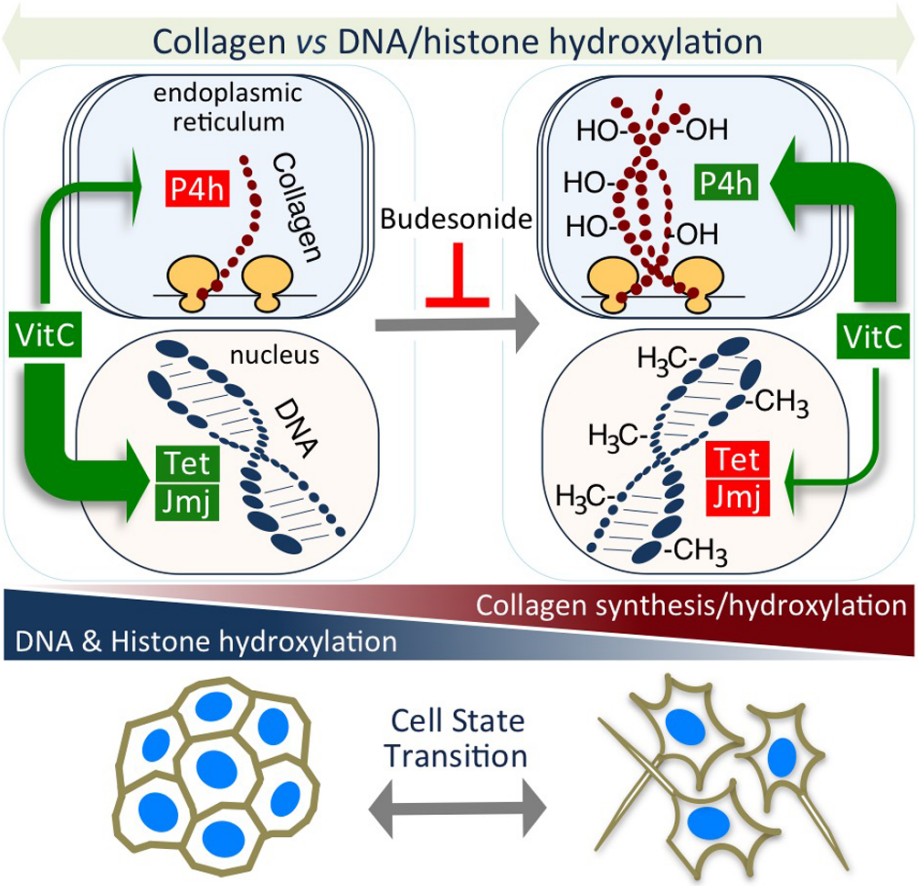


Figure 4.35 P4H-dependent vitamin C-consuming collagen hydroxylation increases global DNA/histone methylation by reducing the activity of TET/JMJ demethylase, resulting in pluripotent stem cells transition.

5. DISCUSSION

Stem cell pluripotency can be examined at a molecular and/or phenotypic level. However, while the development of global molecular profilings' applications (*i.e.* transcriptomics, epigenomics, metabolomics, proteomics) have provided comprehensive sets of pluripotency states associated molecular signatures [16, 77], the functional assays used to define the different pluripotency state-associated phenotypes remained unchanged and currently require the use of *in vivo* models (blastocyst colonization and teratomas formation competence). For instance, the contribution to chimeras is considered the gold-standard method to discriminate naïve (ground state) from primed cells [78], and results can be biased by the heterogeneity of the population analysed and the number of cells injected into blastocysts. Similarly, in the teratoma formation assay the identification of cells derived from the three primary germ layers into histological complex tumours is imprecise. In this context, an alternative emerges from the gastruloid formation efficiency (GFE) assay (**Fig. 4.2-4.6**). Indeed, here we propose GFE as a cost-effective, quantitative, animal-free and robust experimental approach to distinguish unambiguously different functional states of pluripotency from naïve (GFE>95%) to primed (GFE=0).

For a genetic validation of the system we have used *Cripto* KO mESCs, revealing that *Cripto* ablation impairs the induction of A-P axis development, and thus the aggregome to gastruloid transition (**Fig. 4.7-4.9**). Our results are in line with previous reports showing that *Cripto*-null mutant embryos die at E7.5 due to their inability to gastrulate [67]. Of particular relevance, the persistent expression of the endodermal marker Sox17 in undeveloped *Cripto* KO organoids (**Fig. 4.9**), supports the idea that *Cripto* is dispensable for endodermal fate induction [79]. Moreover, our observation that recombinant Cripto and Activin A are both able to re-establish the proper elongation of *Cripto* KO aggregomes (**Fig. 4.10**), well correlates with the observation that both proteins restore the WT phenotype of zebrafish oep (one-eyed pinhead) mutants [80]. Therefore, based on our findings and data in the literature that: i) *Cripto* induction makes ESCs competent to respond to Nodal [63], ii) Nodal is essential for gastruloid development [4] and, iii) Activin A can activate the same receptors and effectors of Nodal [71], it is reasonable to speculate that a Cripto-Activin/Nodal pathway underlies

the induction of A-P axis formation during gastruloid development. Moreover, our results highlight a potential discrepancy regarding the role of Cripto/Activin in gastruloid development. Specifically, unlike WT and Cripto-treated *Cripto* KO gastruloids, the elongated organoids generated by Activin A-treated *Cripto* KO aggregomes are mostly composed of cells expressing neural markers (*Sox2* and *Nestin*) (**Fig. 4.10**). This correlates with the observation that genetic ablation of *Cripto* results in increased neuralization and dopaminergic differentiation of mESCs [63, 81]. Thus, our findings suggest that Activin A overcome only in part the lack of *Cripto* in gastruloids development, and that *Cripto* activity is essential to restrain the formation of neural precursors during gastruloid development.

We have exploited the potential of the improved gastruloid formation assay to evaluate the impact of the pluripotency state on GFE, and show that unlike 2i_LIF (naïve) cells, dissociated primed 5d_F/A cells (EpiSCs) fail to re-aggregate and thus to generate aggregomes of correct size and shape (**Fig. 4.12**). In line with these findings, Hayashi and co-workers recently showed that F/A treatment strongly weakens the intrinsic propensity of naïve mESCs to aggregate, which eventually leads to failure of differentiation in Primordial Germ Cell-like Cells (PGCLCs) [23]. Indeed, already after 3 days of F/A treatment only small and poorly compacted cell aggregates are generated [23]. Notably, re-aggregation of dissociated/singularized cells, *i.e.* the formation of functional aggregomes, is a critical initial event for development of gastruloid [4, 8] and many different kinds of organoids [82, 83]. The molecular mechanism(s) underlying aggregomes formation, and the reason for a reduced re-aggregation competence of EpiSCs remains unknown. It is known that cell adhesive interactions rely on the expression and affinities of adhesion molecules, and we showed that a significant up-regulation of focal adhesion-related genes occurs in ESCs after 5 days of F/A treatment (EpiSCs) (**Fig. 4.14**). Accordingly, primed ESCs (5d_F/A) preferentially generate flat- rather that domed- shaped cell colonies, thus suggesting that these cells exhibit a high propensity to generate cell-substrate rather than cell-cell adhesive interactions. E-Cadherin (encoded by *Cdh1*) plays an essential role in cell-cell adhesion, and it has been reported that forced expression of *Cdh1* improved the chimeric formation ability of EpiSCs (integration into the ICM after blastocyst injection) [84]. Thus, it is reasonable to speculate that the loss of re-aggregation ability of EpiSCs (5d_F/A) is caused by an

F/A-induced down-regulation, activity inhibition and/or delocalization of cell-cell adhesion molecules such as E-Cadherin. After 2 days of F/A treatment in the absence of LIF (early primed EpiLCs), and despite the fact that cells maintain the capacity to aggregate, the generated aggregomes fail to undergo proper morphological elongation and gastruloid development (**Fig. 4.12; 4.14**). The reason of the developmental failure of EpiLCs aggregomes is unknown. Compaction/stiffening is required for the intrinsic induction of TGFβ signalling and chondrogenic differentiation in mesenchymal stem cell aggregates [85], and we have observed that aggregomes of EpiLCs display a smaller and less compacted structure compared to naïve-derived aggregates (**Fig. 4.14**). Thus, it is tempting to reason that a reduced compaction efficiency might generate abortive aggregomes, most likely by preventing the intrinsic activation of a crucial signalling pathway(s). Turner and co-workers showed that gastruloid development strictly relies on the activation of WNT and Nodal signalling pathways, whereas BMP pathway appears dispensable [4]. Since a WNT signalling agonist is exogenously provided during GFE assay, it is possible that alteration of Nodal signalling could underlie, at least in part, the EpiLC phenotype. In line with this idea, EpiLCs (**Fig. 4.12**), *Nodal* KO cells [4] and *Cripto* KO cells (**Fig. 4.7**) aggregomes display a similar undeveloped phenotype.

L-Proline-induced cells (PiCs) exhibit some features of the early-primed EpiLCs. Specifically, PiCs generate teratomas, colonize mouse blastocyst and spontaneously and efficiently revert to naïve pluripotency state [1-3], and here we showed that similar to EpiLCs, are able to differentiate into PGCLCs (**Fig. 4.19-4.20**). However, while EpiLCs represent a transient pluripotency state that cannot be captured *in vitro*, our findings support the idea that L-Pro supplementation captures ESCs in a LIF dependent early-primed state of pluripotency. Interestingly, we found that while EpiLCs generate only abortive aggregomes (**Fig. 4.12**), PiCs generate effective although smaller aggregomes (**Fig. 4.16**). We hypothesise that the developmental failure of EpiLCs is due to specific culture conditions. Indeed, while early-primed EpiLCs are generated from ESCs in the presence of Fgf/Activin and in the absence of LIF, early-primed PiCs are generated and maintained in LIF medium and rely on the activation of autocrine bFGF and TGF-β signaling pathways. A key role of LIF induced genes in GFE is supported by the observation that LIF supplementation during the first 2 days of incubation is sufficient

to restore the elongation of abortive EpiLC aggregomes (**Fig. 4.13**). PiCs aggregomes elongate earlier than control, generating elongate-shaped gastruloid-like organoids, which express markers of A-P axis formation (**Fig. 4.17**). Why PiCs aggregomes undergo symmetry breaking before control is still unclear. However, cultured PiCs are heterogeneous with a large majority of cells displaying features of an intermediate state of pluripotency, and two minor fractions of cells displaying features of the extreme naïve or primed states respectively [3]. Thus, it is reasonable to speculate that the aggregation of cells at different state of pluripotency may contribute to anticipate the aggregome to gastruloid transition, although this requires further investigations.

Gastruloids are a promising model to explore a number of novel research avenues about how cells form tissues and organs that cannot be explored with embryos [57]. Here we show that gastruloids are also competent to generate cells that express PGCs markers, and are *bona fide* PGCLCs. PGCs are specified during embryogenesis at around E7.5. The mechanism for PGC specification/development has been difficult to explore, mainly because PGCs are small in number and refractory to proliferation *in vitro* [62]. Indeed, here we observed the formation of a cluster of cells positive to PGCs markers during gastruloid development (**Fig. 4.21**). These promising results suggest that gastruloids development may represent a powerful experimental model to shed light on the particular nature of these cells that maintain pluripotency in a high differentiating environment. Of relevance, the development of PGCLCs in gastruloids occurs without the presence of extra-embryonic tissue [4] and without the mix of cytokines that is added to the standard protocol [23]. Interestingly, we show that expression of PGCs specific markers is induced at earlier time point in PiCs-derived gastruloids compared to naïve-derived gastruloids, suggesting a specific pluripotency potential of PiCs. A question remains open: does L-Pro regulate pluripotency in embryo development? Although no direct evidence is available so far, it is plausible to speculate that L-Pro availability may influence stem cells behaviour and plasticity *in vivo*. For instance, it has been reported that *Aldh18a1* (L-Pro biosynthesis), *Col4a1* (Collagen synthesis) and *P4ha2* (Collagen hydroxylation) genes are specifically induced in the Primitive Endoderm (PrE), exactly when the pluripotent cells are specified in the inner cell mass (ICM) of the blastocyst [86]. Moreover, embryo implantation induces tissue

remodelling and collagen degradation, likely raising the level of free L-Pro that might eventually influence stem cells plasticity in the peri-implantation embryo.

However, further characterisation is needed to shed light on a role of L-Pro *in vivo*. In conclusion, here we reported the optimization of an *in vitro* gastruloid formation efficiency (GFE) assay, useful to functionally characterize mouse stem cells pluripotency and elucidate some events occurring during embryo development. Besides being a valid (low-cost, animal-free) alternative to the *in vivo* chimera assay, the particular configuration (FACS combined with, 96 well-plates) and robustness (output >95%) of our GFE assay makes it suitable for large scale (high-throughput) phenotype-based genetic screenings (shRNA, miRNA) for the identification of genes involved in gastruloids/embryo development. Moreover, the screening of metabolite and drug libraries could allow the identification of beneficial metabolites and potential toxic teratogens drugs.

Another important AIM of my PhD project focused on exploring the molecular mechanism(s) underlying L-Pro-induced esMT. By using a combination of drug screening and genetic editing approaches, we have shown that L-Pro-induced esMT relies on a previously unexplored collagen-epigenetic functional axis. In particular, we suggest that exogenous addition of L-Pro induces a rapid increase of collagen synthesis/hydroxylation in ESCs (**Fig. 4.25**), which affects DNA/Histone methylation. Mechanistically, we identify a previously unexplored antagonist interplay between collagen and DNA/Histone hydroxylation, which relies on the availability VitC that is an essential cofactor for the activity of both collagen hydroxylases (P4Hs) and DNA/Histone demethylases (Tet/Jmj). Particularly, increased collagen synthesis/hydroxylation under VitC limiting conditions rapidly reduces 5hmC and increases H3K9/K36 histone methylation. In line with the idea that the increase in collagen synthesis/hydroxylation causes a reduced availability of VitC, which in turn reduces the activity of Tet/Jmj demethylases, a high VitC regimen antagonizes the phenotypic transition induced by L-Pro and prevents increased methylation levels. By blocking the synthesis/hydroxylation of collagen through the knock out/knock down of the *P4ha2* and *P4ha1* genes, and with a pharmacological approach, using Budesonide, ESCs do not undergo esMT. Interestingly, Budesonide has been recently identified in a combined connectivity mapping and pharmaco-epidemiological approach as a drug with

potential breast cancer-preventing properties [87]. Moreover, recently asthma was associated with reduced risk of developing pancreatic ductal adenocarcinoma [88]. Certainly, Budesonide is universally employed to treat asthma, underlying a possible connection with a reduced risk of pancreatic cancer. Thus, these unexpected findings that Budesonide blocks mesenchymal transition interfering with collagen accumulation provide novel insights into its therapeutic effects. Further investigations will require to completely elucidate Budesonide effects and mechanism. However, it is reasonable to speculate that this mechanism involved the Glucocorticoid Receptors (GRs), that is expressed in all the cell of the body, and it controls the expression of genes involved in development, metabolism and immune response. Moreover, it is not possible to exclude Budesonide's off-target effects at the concentration used, including for instance inhibition of P4H activity. Interestingly, epigenetic mechanisms may govern the accumulation of collagen by reducing promoter methylation and increasing expression of collagen genes in pathological conditions such as fibrosis [89, 90]. Of note, various collagens (collagen I, II, III, V and IX) show an increase in deposition during tumorigenesis and aging [91]. It is reasonable to assert that all these pathological conditions could be correlated with an alteration of the epigenetic→collagen axis described here. Furthermore, we identified budesonide as a potent inhibitor of this mechanism and a novel drug candidate for the treatment of a broad spectrum of diseases in which cell plasticity and fibrosis are involved, including metastatic cancer [75].

6. APPENDIX

Table 1. List of ESC inhibitors; plate position and therapeutic use are indicated.

Preswick N°	Chemical name	Plate Position	MW *	CAS ** number	Therapeutic effect
Pw1233	Auranofin	11B02	678,49	34031-32-8	Analgesic
Pw1082	(-)-Eseroline	14E03	334,38	104015-29-4	
Pw1269	Haloprogin	06A03	361,40	777-11-7	Anesthetic
Pw97	Disulfiram	02B08	296,54	97-77-8	Antabuse effect
Pw550	Parthenolide	07G11	248,32	20554-84-1	Anti-inflammatory
Pw740	Ebselen	10B11	274,18	60940-34-3	
Pw886	Clioquinol	12A07	305,50	130-26-7	Antiamebic
Pw63	Nifedipine	01G04	346,34	21829-25-4	Antianginal
Pw270	Fendiline	04C11	351,92	13636-18-5	
Pw286	Perhexiline	04E07	393,57	6724-53-4	
Pw368	Bepridil	05E09	403,01	74764-40-2	
Pw409	Amiodarone	06A10	681,78	19774-82-4	
Pw560	Prenylamine	07H11	419,57	69-43-2	
Pw666	Nisoldipine	09C07	388,42	63675-72-9	
Pw1043	Penbutolol	14A04	680,95	38363-32-5	
Pw319	Clofilium tosylate	04H10	510,18	92953-10-1	Antiarrhythmic
Pw143	Chlorhexidine	02G04	505,46	55-56-1	Antibacterial
Pw1474	Chloroxine	02H03	214,05	773-76-2	
Pw154	Josamycin	02H05	828,02	16846-24-5	
Pw388	Dequalinium	05G09	527,59	522-51-0	
Pw438	Doxorubicin	06D09	579,99	25316-40-9	
Pw487	Daunorubicin	07A08	563,99	23541-50-6	
Pw522	Thiostrepton	07E03	1664,92	1393-48-2	
Pw541	Ciclopirox	07G02	268,36	41621-49-2	
Pw705	Methyl benzethonium	09G06	462,12	25155-18-4	
Pw708	Benzethonium	09G09	448,09	121-54-0	
Pw748	Monensin	10C09	692,87	22373-78-0	
Pw777	Alexidine	10F08	581,73	22573-93-9	
Pw790	Cycloheximide	10G11	281,35	66-81-9	
Pw792	Propidium iodide	10H03	668,41	25535-16-4	
Pw1045	Sertaconazole	14A06	500,79	99592-39-9	
Pw1234	Azithromycin	15F04	749,00	83905-01-5	
Pw312	Flunarizine	04H03	477,43	30484-77-6	Anticonvulsant
Pw99	Mianserine	02B10	300,83	21535-47-7	Antidepressant
Pw102	Amoxapine	02C03	313,79	14028-44-5	
Pw254	Nortriptyline	04B05	299,85	894-71-3	
Pw346	Maprotiline	05C07	313,87	10347-81-6	
Pw386	GBR 12909	05G07	523,50	67469-78-7	
Pw851	Paroxetine	11F02	365,84	110429-35-1	

Pw1602	Indatraline	13A11	328,67	86939-10-8	
Pw144	Loperamide	02G05	513,51	34552-83-5	Antidiarrheal
Pw125	Perphenazine	02E06	403,98	58-39-9	Antiemetic
Pw149	Thioproperazine	02G10	638,85	2347-80-0	
Pw306	Clemastine	04G07	459,97	14976-57-9	
Pw313	Trifluoperazine	04H04	480,43	440-17-5	
Pw399	Prochlorperazine	05H10	606,10	84-02-6	
Pw706	Chlorcyclizine	09G07	337,30	1620-21-9	
Pw1600	Aprepitant	10C08	534,44	170729-80-3	
Pw1068	Thiethylperazine	14C09	667,80	52239-63-1	
Pw495	Terconazole	07B06	532,47	67915-31-5	Antifungal
Pw551	Hexetidine	07H02	339,61	141-94-6	
Pw553	Pentamidine	07H04	592,69	140-64-7	
Pw1139	Itraconazole	08A08	705,65	84625-61-6	
Pw363	Colchicine	05E04	399,45	64-86-8	Antigout
Pw40	Niclosamide	01D11	327,13	50-65-7	Antihelmintic
Pw156	Ivermectin	02H07	875,12	70288-86-7	
Pw210	Fenbendazole	03E11	299,35	43210-67-9	
Pw217	Mebendazole	03F08	295,30	31431-39-7	
Pw247	Albendazole	04A08	265,34	54965-21-8	
Pw318	Quinacrine	04H09	508,92	6151-30-0	
Pw355	Hycanthone	05D06	356,49	3105-97-3	
Pw1460	Oxibendazol	08E09	249,27	20559-55-1	
Pw660	Avermectin B1a	09B11	873,10	71751-41-2	
Pw1459	Oxfendazol	13C05	315,35	53716-50-0	
Pw1040	Pyrvinium	13H11	1151,00	3546-41-6	
Pw1080	Methiazole	14D11	265,34	108579-67-5	
Pw59	Flubendazol	12B03	313,29	31430-15-6	
Pw103	Parbendazole	14G11	247,30	14255-87-9	
Pw239	Astemizole	02F07	458,58	68844-77-9	Antihistaminic
Pw258	Terfenadine	02F09	471,69	50679-08-8	
Pw294	Homochlorcyclizine	04E04	387,78	1982-36-1	
Pw1102	Deptropine	13F04	525,60	2169-75-7	
Pw657	Lacidipine	08B10	455,56	103890-78-4	Antihypertensive
Pw994	Prazosin	12G08	419,87	19237-84-4	
Pw1258	Carvedilol	15B08	406,49	72956-09-3	
Pw498	Mefloquine	02E07	414,78	51773-92-3	Antimalarial
Pw534	Primaquine	06H07	455,34	63-45-6	Antimalarial
Pw1	Azaguanine-8	01A02	152,12	134-58-7	Antineoplastic
Pw94	Azathioprine	02B05	277,27	446-86-6	
Pw100	Nocodazole	02B11	301,33	31430-18-9	
Pw146	Tamoxifen citrate	02G07	563,65	54965-24-1	
Pw155	Paclitaxel	02H06	853,93	33069-62-4	
Pw200	Camptothecine (S,+)	03D11	348,36	7689-03-4	
Pw347	Thioguanosine	05C08	299,31	85-31-4	
Pw385	Mitoxantrone	05G06	517,41	70476-82-3	

Pw396	Etoposide	05H07	588,57	33419-42-0	
Pw1134	Cytarabine	08G06	243,22	147-94-4	
Pw1270	Gefitinib	09A03	446,91	184475-35-2	
Pw866	Azacytidine-5	11G07	244,21	320-67-2	
Pw1197	Toremifene	12H03	405,97	89778-26-7	
Pw1415	Floxuridine	13A06	246,20	50-91-9	
Pw1362	Vorinostat	14C11	264,33	149647-78-9	
Pw1079	Chlorambucil	14D10	304,22	305-03-3	
Pw1483	Fludarabine	15C08	285,24	21679-14-1	
Pw1484	Cladribine	15C09	285,69	4291-63-8	
Pw1203	5-fluorouracil	15D02	130,08	51-21-8	
Pw1196	Topotecan	15F02	421,46	123948-87-8	
Pw1266	Gemcitabine	15G04	263,20	95058-81-4	
Pw1180	Docetaxel	15G08	807,90	114977-28-5	
Pw1290	Imatinib	15H02	493,62	152459-95-5	
Pw1285	Ibandronate	09D10	341,22	114084-78-5	Antiosteoporetic
Pw101	R(-) Apomorphine	02C02	625,60	41372-20-7	Antiparkinsonian
Pw491	Metixene	07B02	345,94	1553-34-0	
Pw1236	Benztropine	15F07	403,54	132-17-2	
Pw421	Suloctidil	06C02	337,57	54063-56-8	Antiplatelet
Pw364	Metergoline	05E05	403,53	17692-51-2	Antiprolactin
Pw78	Thioridazine	01H09	407,04	130-61-0	Antipsychotic
Pw308	Pimozide	04G09	461,56	2062-78-4	
Pw320	Fluphenazine	04H11	510,45	146-56-5	
Pw375	Methiothepin	05F06	472,63	19728-88-2	
Pw1368	Zotepine	06C03	331,87	26615-21-4	
Pw1341	Sertindole	08H11	440,95	106516-24-9	
Pw693	Promazine	09F04	320,89	53-60-1	
Pw906	Fluspirilen	12C07	475,59	1841-19-6	
Pw932	S(-)Eticlopride	12F03	377,31	97612-24-3	
Pw998	Zuclopenthixol	13D09	473,90	633-59-0	
Pw1048	Piperacetazine	14A09	410,58	3819-00-9	
Pw1194	Thimerosal	09E07	404,81	54-64-8	Antiseptic
Pw925	Thonzonium	12E06	591,73	553-08-2	
Pw1601	Pinaverium	09A11	591,43	53251-94-8	Antispastic
Pw18	Idoxuridine	01B09	354,10	54-42-2	Antiviral
Pw782	Podophyllotoxin	10G03	414,42	518-28-5	
Pw1056	Trifluridine	14B07	296,20	70-00-8	
Pw862	Raloxifene	11G03	510,06	82640-04-8	Bone mineralization
Pw725	Deferoxamine	10A06	656,80	138-14-7	Chelating
Pw1473	Estradiol	15A11	356,51	979-32-8	Contraceptive
Pw912	Monobenzone	12D03	200,24	103-16-2	Depigmentation
Pw128	Spironolactone	02E09	416,58	52-01-7	Diuretic
Pw397	Clomiphene	05H08	598,10	50-41-9	Estrogen Antagonist
Pw516	Lovastatin	07D07	404,55	75330-75-5	Hypocholesterolemic
Pw1383	Simvastatin	11G06	418,58	79902-63-9	

Pw1441	Mevastatin	08A03	390,52	73573-88-3	
Pw1469	Mercaptopurine	08G05	152,18	50-44-2	Immunosuppressant
Pw199	Prilocaine	03D10	256,78	1786-81-8	Local anesthetic
Pw1105	Verteporfin	14G06	718,81	129497-78-5	Photosensitizing
Pw311	Darifenacin	13A10	507,48	133099-07-7	Urological Agent
Pw1244	Eburnamonine (-)	08E08	294,40	4880-88-0	Vasodilator

* Molecular Weight
** A division of the American Chemical Society

Table 2. List of esMT inhibitors with described anti-collagen/fibrotic effects.

*in PubChem budesonide is classified as glucocorticoid receptor (GR) antagonist

Drug	Chemical Class (Therapeutic Use)	Organism, Tissue, Disease	Reference
Budesonide	steroid, glucocorticoid GR agonist* (antinflammatory)	Human colon mucosa, colitis airways, asthma	[92-96]
Nicardipine **Cilnidipine**	1,4-dihydropidyridines blocks calcium channels (antianginal, antihypertension)	Rat kidney, renal fibrosis smooth muscle cells	[97, 98]
Chlorpromazine	tricyclic phenothiazine blocks many receptors (antipsychotic)	Human lung fibroblasts	[99]
Promethazine	tricyclic phenothiazine blocks many receptors (antihistaminic)	Rat skin, wound healing	[100]
Desipramine	tricyclic dibenzazepine blocks many receptors (antidepresant)	Mouse breast cancer cells	[101]

NOTES

The study presented in this Thesis is subject of the following publications:

- **Cermola F**, D'Aniello C, Taté R, De Cesare D, Martinez-Arias A, Minchiotti G and Patriarca EJ. *Gastruloid development competence discriminates different states of pluripotency between naïve and primed.* July 2019 bioRxiv 664920; doi: https://doi.org/10.1101/664920

- D'Aniello C, **Cermola F**, Palamidessi A, Wanderlingh L, Gagliardi M, Migliaccio A, Varrone F, Casalino L, Matarazzo M, De Cesare D, Scita G, Patriarca EJ and Minchiotti G. *Role of collagen prolyl hydroxylation-dependent metabolic perturbation in governing epigenetic remodeling and mesenchymal transition in pluripotent and cancer cells*. Cancer Research. 2019 Jul 1;79(13):3235-3250. doi: 10.1158/0008-5472.CAN-18-2070. IF: 9.130

7. REFERENCES

1. Casalino, L., et al., *Control of embryonic stem cell metastability by L-proline catabolism.* J Mol Cell Biol, 2011. **3**(2): p. 108-22.
2. Comes, S., et al., *L-Proline induces a mesenchymal-like invasive program in embryonic stem cells by remodeling H3K9 and H3K36 methylation.* Stem Cell Reports, 2013. **1**(4): p. 307-21.
3. D'Aniello, C., et al., *Vitamin C and L-Proline Antagonistic Effects Capture Alternative States in the Pluripotency Continuum.* Stem Cell Reports, 2017. **8**(1): p. 1-10.
4. Turner, D.A., et al., *Anteroposterior polarity and elongation in the absence of extra-embryonic tissues and of spatially localised signalling in gastruloids: mammalian embryonic organoids.* Development, 2017. **144**(21): p. 3894-3906.
5. Fiorenzano, A., et al., *Cripto is essential to capture mouse epiblast stem cell and human embryonic stem cell pluripotency.* Nat Commun, 2016. **7**: p. 12589.
6. Minchiotti, G., *Nodal-dependant Cripto signaling in ES cells: from stem cells to tumor biology.* Oncogene, 2005. **24**(37): p. 5668-75.
7. D'Aniello, C., et al., *Vitamin C in Stem Cell Biology: Impact on Extracellular Matrix Homeostasis and Epigenetics.* Stem Cells Int, 2017. **2017**: p. 8936156.
8. Baillie-Johnson, P., et al., *Generation of Aggregates of Mouse Embryonic Stem Cells that Show Symmetry Breaking, Polarization and Emergent Collective Behaviour In Vitro.* J Vis Exp, 2015(105).
9. Morgani, S., J. Nichols, and A.K. Hadjantonakis, *The many faces of Pluripotency: in vitro adaptations of a continuum of in vivo states.* BMC Dev Biol, 2017. **17**(1): p. 7.
10. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7634-8.
11. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.
12. Brons, I.G., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-5.
13. Tesar, P.J., et al., *New cell lines from mouse epiblast share defining features with human embryonic stem cells.* Nature, 2007. **448**(7150): p. 196-9.
14. Hirai, H., P. Karian, and N. Kikyo, *Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor.* Biochem J, 2011. **438**(1): p. 11-23.
15. Ficz, G., et al., *FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency.* Cell Stem Cell, 2013. **13**(3): p. 351-9.
16. Habibi, E., et al., *Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells.* Cell Stem Cell, 2013. **13**(3): p. 360-9.
17. Nichols, J. and A. Smith, *Pluripotency in the embryo and in culture.* Cold Spring Harb Perspect Biol, 2012. **4**(8): p. a008128.
18. Leitch, H.G., et al., *Rebuilding pluripotency from primordial germ cells.* Stem Cell Reports, 2013. **1**(1): p. 66-78.

19. Smith, A., *Formative pluripotency: the executive phase in a developmental continuum.* Development, 2017. **144**(3): p. 365-373.
20. Kojima, Y., et al., *The transcriptional and functional properties of mouse epiblast stem cells resemble the anterior primitive streak.* Cell Stem Cell, 2014. **14**(1): p. 107-20.
21. Hackett, J.A. and M.A. Surani, *Regulatory principles of pluripotency: from the ground state up.* Cell Stem Cell, 2014. **15**(4): p. 416-430.
22. Martello, G. and A. Smith, *The nature of embryonic stem cells.* Annu Rev Cell Dev Biol, 2014. **30**: p. 647-75.
23. Hayashi, K., et al., *Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells.* Cell, 2011. **146**(4): p. 519-32.
24. Tsukiyama, T. and Y. Ohinata, *A modified EpiSC culture condition containing a GSK3 inhibitor can support germline-competent pluripotency in mice.* PLoS One, 2014. **9**(4): p. e95329.
25. Chou, Y.F., et al., *The growth factor environment defines distinct pluripotent ground states in novel blastocyst-derived stem cells.* Cell, 2008. **135**(3): p. 449-61.
26. Rathjen, J., et al., *Formation of a primitive ectoderm like cell population, EPL cells, from ES cells in response to biologically derived factors.* J Cell Sci, 1999. **112 (Pt 5)**: p. 601-12.
27. Chang, K.H. and M. Li, *Clonal isolation of an intermediate pluripotent stem cell state.* Stem Cells, 2013. **31**(5): p. 918-27.
28. Kalkan, T. and A. Smith, *Mapping the route from naive pluripotency to lineage specification.* Philos Trans R Soc Lond B Biol Sci, 2014. **369**(1657).
29. Cheng, S., et al., *Single-Cell RNA-Seq Reveals Cellular Heterogeneity of Pluripotency Transition and X Chromosome Dynamics during Early Mouse Development.* Cell Rep, 2019. **26**(10): p. 2593-2607 e3.
30. Pera, M.F. and P.P. Tam, *Extrinsic regulation of pluripotent stem cells.* Nature, 2010. **465**(7299): p. 713-20.
31. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.* Nature, 2007. **450**(7173): p. 1230-4.
32. Toyooka, Y., et al., *Identification and characterization of subpopulations in undifferentiated ES cell culture.* Development, 2008. **135**(5): p. 909-18.
33. Torres-Padilla, M.E. and I. Chambers, *Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage.* Development, 2014. **141**(11): p. 2173-81.
34. Harvey, A.J., J. Rathjen, and D.K. Gardner, *Metaboloepigenetic Regulation of Pluripotent Stem Cells.* Stem Cells Int, 2016. **2016**: p. 1816525.
35. D'Aniello, C.C., F.; Patriarca, E. J.; Minchiotti, G., *Metabolic - epigenetic axis in pluripotent state transitions.* Epigenomes, 2019.
36. Tran, K.A., C.M. Dillingham, and R. Sridharan, *The role of alpha-ketoglutarate-dependent proteins in pluripotency acquisition and maintenance.* J Biol Chem, 2019. **294**(14): p. 5408-5419.
37. Cao, N., et al., *Ascorbic acid enhances the cardiac differentiation of induced pluripotent stem cells through promoting the proliferation of cardiac progenitor cells.* Cell Res, 2012. **22**(1): p. 219-36.

38. Hore, T.A., et al., *Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naive pluripotency by complementary mechanisms.* Proc Natl Acad Sci U S A, 2016. **113**(43): p. 12202-12207.
39. Mali, P., et al., *Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes.* Stem Cells, 2010. **28**(4): p. 713-20.
40. Yanes, O., et al., *Metabolic oxidation regulates embryonic stem cell differentiation.* Nat Chem Biol, 2010. **6**(6): p. 411-7.
41. Wang, J., et al., *Dependence of mouse embryonic stem cells on threonine catabolism.* Science, 2009. **325**(5939): p. 435-9.
42. Shiraki, N., et al., *Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells.* Cell Metab, 2014. **19**(5): p. 780-94.
43. Shyh-Chang, N., et al., *Influence of threonine metabolism on S-adenosylmethionine and histone methylation.* Science, 2013. **339**(6116): p. 222-6.
44. Wishart, D.S., et al., *HMDB 4.0: the human metabolome database for 2018.* Nucleic Acids Res, 2018. **46**(D1): p. D608-D617.
45. Parveen, M., et al., *L-proline mixed with sodium benzoate as sustainable inhibitor for mild steel corrosion in 1M HCl: An experimental and theoretical approach.* Sci Rep, 2018. **8**(1): p. 7489.
46. Zhou, W., et al., *HIF1alpha induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition.* EMBO J, 2012. **31**(9): p. 2103-16.
47. Parchem, R.J., et al., *Two miRNA clusters reveal alternative paths in late-stage reprogramming.* Cell Stem Cell, 2014. **14**(5): p. 617-31.
48. D'Aniello, C., et al., *A novel autoregulatory loop between the Gcn2-Atf4 pathway and (L)-Proline [corrected] metabolism controls stem cell identity.* Cell Death Differ, 2015. **22**(7): p. 1094-105.
49. Loh, Y.H., et al., *Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells.* Genes Dev, 2007. **21**(20): p. 2545-57.
50. Maruotti, J., et al., *Nuclear transfer-derived epiblast stem cells are transcriptionally and epigenetically distinguishable from their fertilized-derived counterparts.* Stem Cells, 2010. **28**(4): p. 743-52.
51. Clevers, H., *Wnt/beta-catenin signaling in development and disease.* Cell, 2006. **127**(3): p. 469-80.
52. Marikawa, Y., et al., *Aggregated P19 mouse embryonal carcinoma cells as a simple in vitro model to study the molecular regulations of mesoderm formation and axial elongation morphogenesis.* Genesis, 2009. **47**(2): p. 93-106.
53. Turner, D.A., P. Baillie-Johnson, and A. Martinez Arias, *Organoids and the genetically encoded self-assembly of embryonic stem cells.* Bioessays, 2016. **38**(2): p. 181-91.
54. van den Brink, S.C., et al., *Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells.* Development, 2014. **141**(22): p. 4231-42.
55. Vianello, S. and M.P. Lutolf, *Understanding the Mechanobiology of Early Mammalian Development through Bioengineered Models.* Dev Cell, 2019. **48**(6): p. 751-763.

56. Simunovic, M. and A.H. Brivanlou, *Embryoids, organoids and gastruloids: new approaches to understanding embryogenesis.* Development, 2017. **144**(6): p. 976-985.
57. Beccari, L., et al., *Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids.* Nature, 2018. **562**(7726): p. 272-276.
58. Rossi, G.B., A.; Guiet, R.; Girgin, M.; Kelly, R. G.; Lutolf, M. P., *Embryonic organoids recapitulate early heart organogenesis.*
59. Surani, M.A., K. Hayashi, and P. Hajkova, *Genetic and epigenetic regulators of pluripotency.* Cell, 2007. **128**(4): p. 747-62.
60. Saitou, M., S. Kagiwada, and K. Kurimoto, *Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells.* Development, 2012. **139**(1): p. 15-31.
61. Ohinata, Y., et al., *A signaling principle for the specification of the germ cell lineage in mice.* Cell, 2009. **137**(3): p. 571-84.
62. Saitou, M. and M. Yamaji, *Germ cell specification in mice: signaling, transcription regulation, and epigenetic consequences.* Reproduction, 2010. **139**(6): p. 931-42.
63. Parisi, S., et al., *Nodal-dependent Cripto signaling promotes cardiomyogenesis and redirects the neural fate of embryonic stem cells.* J Cell Biol, 2003. **163**(2): p. 303-14.
64. Persico, M.G., et al., *Cripto in tumors and embryo development.* Biochim Biophys Acta, 2001. **1552**(2): p. 87-93.
65. Hayashi, K. and M. Saitou, *Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells.* Nat Protoc, 2013. **8**(8): p. 1513-24.
66. D'Andrea, D., et al., *Cripto promotes A-P axis specification independently of its stimulatory effect on Nodal autoinduction.* J Cell Biol, 2008. **180**(3): p. 597-605.
67. Ding, J., et al., *Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo.* Nature, 1998. **395**(6703): p. 702-7.
68. Minchiotti, G., et al., *Membrane-anchorage of Cripto protein by glycosylphosphatidylinositol and its distribution during early mouse development.* Mech Dev, 2000. **90**(2): p. 133-42.
69. Guardiola, O., et al., *Cripto regulates skeletal muscle regeneration and modulates satellite cell determination by antagonizing myostatin.* Proc Natl Acad Sci U S A, 2012. **109**(47): p. E3231-40.
70. Minchiotti, G., et al., *Role of the EGF-CFC gene cripto in cell differentiation and embryo development.* Gene, 2002. **287**(1-2): p. 33-7.
71. Pauklin, S. and L. Vallier, *Activin/Nodal signalling in stem cells.* Development, 2015. **142**(4): p. 607-19.
72. Ying, Q.L., et al., *The ground state of embryonic stem cell self-renewal.* Nature, 2008. **453**(7194): p. 519-23.
73. Hayashi, M., et al., *Generation of germ cells from pluripotent stem cells in mammals.* Reprod Med Biol, 2018. **17**(2): p. 107-114.
74. Weber, S., et al., *Critical function of AP-2 gamma/TCFAP2C in mouse embryonic germ cell maintenance.* Biol Reprod, 2010. **82**(1): p. 214-23.
75. D'Aniello, C., et al., *Collagen Prolyl Hydroxylation-Dependent Metabolic Perturbation Governs Epigenetic Remodeling and Mesenchymal Transition in Pluripotent and Cancer Cells.* Cancer Res, 2019. **79**(13): p. 3235-3250.

76. Xiong, G., et al., *Prolyl-4-hydroxylase alpha subunit 2 promotes breast cancer progression and metastasis by regulating collagen deposition.* BMC Cancer, 2014. **14**: p. 1.
77. Pijuan-Sala, B., et al., *A single-cell molecular map of mouse gastrulation and early organogenesis.* Nature, 2019. **566**(7745): p. 490-495.
78. Mascetti, V.L. and R.A. Pedersen, *Contributions of Mammalian Chimeras to Pluripotent Stem Cell Research.* Cell Stem Cell, 2016. **19**(2): p. 163-175.
79. Jin, J.Z. and J. Ding, *Cripto is required for mesoderm and endoderm cell allocation during mouse gastrulation.* Dev Biol, 2013. **381**(1): p. 170-8.
80. Gritsman, K., et al., *The EGF-CFC protein one-eyed pinhead is essential for nodal signaling.* Cell, 1999. **97**(1): p. 121-32.
81. Lonardo, E., et al., *A small synthetic cripto blocking Peptide improves neural induction, dopaminergic differentiation, and functional integration of mouse embryonic stem cells in a rat model of Parkinson's disease.* Stem Cells, 2010. **28**(8): p. 1326-37.
82. Takebe, T., et al., *Self-organization of human hepatic organoid by recapitulating organogenesis in vitro.* Transplant Proc, 2012. **44**(4): p. 1018-20.
83. Takebe, T., et al., *Vascularized and Complex Organ Buds from Diverse Tissues via Mesenchymal Cell-Driven Condensation.* Cell Stem Cell, 2015. **16**(5): p. 556-65.
84. Ohtsuka, S., S. Nishikawa-Torikai, and H. Niwa, *E-cadherin promotes incorporation of mouse epiblast stem cells into normal development.* PLoS One, 2012. **7**(9): p. e45220.
85. Sarem, M., et al., *Cell number in mesenchymal stem cell aggregates dictates cell stiffness and chondrogenesis.* Stem Cell Res Ther, 2019. **10**(1): p. 10.
86. Ohnishi, Y., et al., *Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages.* Nat Cell Biol, 2014. **16**(1): p. 27-37.
87. Busby, J., et al., *A combined connectivity mapping and pharmacoepidemiology approach to identify existing medications with breast cancer causing or preventing properties.* Pharmacoepidemiol Drug Saf, 2018. **27**(1): p. 78-86.
88. Gomez-Rubio, P., et al., *Reduced risk of pancreatic cancer associated with asthma and nasal allergies.* Gut, 2017. **66**(2): p. 314-322.
89. O'Reilly, S., *Epigenetics in fibrosis.* Mol Aspects Med, 2017. **54**: p. 89-102.
90. Zhang, X., et al., *DNA methylation regulated gene expression in organ fibrosis.* Biochim Biophys Acta Mol Basis Dis, 2017. **1863**(9): p. 2389-2397.
91. Lu, P., V.M. Weaver, and Z. Werb, *The extracellular matrix: a dynamic niche in cancer progression.* J Cell Biol, 2012. **196**(4): p. 395-406.
92. Yu, F., et al., *ORMDL3 is associated with airway remodeling in asthma via the ERK/MMP-9 pathway.* Mol Med Rep, 2017. **15**(5): p. 2969-2976.
93. Munch, A., et al., *Low-dose budesonide for maintenance of clinical remission in collagenous colitis: a randomised, placebo-controlled, 12-month trial.* Gut, 2016. **65**(1): p. 47-56.
94. Bohr, J., et al., *Diagnosis and management of microscopic colitis: current perspectives.* Clin Exp Gastroenterol, 2014. **7**: p. 273-84.
95. Miehlke, S., et al., *Budesonide is more effective than mesalamine or placebo in short-term treatment of collagenous colitis.* Gastroenterology, 2014. **146**(5): p. 1222-30 e1-2.

96. Todorova, L., et al., *TGFbeta-induced matrix production by bronchial fibroblasts in asthma: budesonide and formoterol effects.* Respir Med, 2011. **105**(9): p. 1296-307.
97. Mishima, K., et al., *Involvement of N-type Ca(2+) channels in the fibrotic process of the kidney in rats.* Am J Physiol Renal Physiol, 2013. **304**(6): p. F665-73.
98. Zhao, J., et al., *Urotensin II-induced collagen synthesis in cultured smooth muscle cells from rat aortic media and a possible involvement of transforming growth factor-beta1/Smad2/3 signaling pathway.* Regul Pept, 2013. **182**: p. 53-8.
99. Liu, H.L., et al., *Effects of tetrandrine and chlorpromazine on synthesis of collagen and hyaluronic acid in cultured human lung fibroblasts.* Zhongguo Yao Li Xue Bao, 1995. **16**(5): p. 412-4.
100. Bairy, K.L., et al., *Effects of antihistamines on wound healing.* Indian J Exp Biol, 1991. **29**(4): p. 398-9.
101. Szpunar, M.J., et al., *The antidepressant desipramine and alpha2-adrenergic receptor activation promote breast tumor progression in association with altered collagen structure.* Cancer Prev Res (Phila), 2013. **6**(12): p. 1262-72.