Impact of hypoxia on embryonic and extraembryonic stem cells and during differentiation via gastruloid formation

Inaugural-Dissertation
to obtain the academic degree
Doctor of Philosophy in Natural Science - PhD

submitted to the Department of Biology, Chemistry, Pharmacy
of Freie Universität Berlin

by

Natalia López Anguita

Berlin, 2022

The work was performed at the Max Planck Institute for Molecular Genetics in Berlin between October 2018 and September 2022 under the supervision of Dr. Aydan Bulut-Karslioglu.

Acknowledgments

First of all, I would like to thank Aydan who trusted me and allowed me to work in her lab. For supporting my work in the good and in the bad. For being always there to talk about any concern. Last, for growing (together) since the beginning of the ABK Lab existence. It was truly motivating and rewarding creating the ABK lab, especially together with Dhanur and Marina, it really felt like a family. Having fun while doing great science. I will never forget those first years.

Secondly, the whole ABK Lab, a group of very talented scientists who are also incredible people. I am truly grateful to have been working with you.

I would also like to thank my former colleague and supervisor Isabelle Bergiers who taught me all I needed to know before starting a PhD. From a colleague and mentor, she also became a friend. I can't thank enough for how impactful she has been and still is in my scientific and personal growth.

This long, challenging and life-changing journey wouldn't have been possible without the support of friends and family who always supported me no matter what. Coming from very different backgrounds it is hard sometimes to understand the challenges I could go through and yet, I haven't felt alone.

Last and extremely important, to my partner, who blindly supported me in whichever decision would make me happier, listened to me when I needed support, gave me advice, made me laugh and taught me what love is. Thanks for growing with me and choosing me every morning. Ti amo.

Declaration of Independence:

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me. Furthermore, I also declare than I have not submitted the dissertation in this or any other form to any other institution as a dissertation.

Summary

Environmental oxygen is crucial for mammalian life. Low oxygen levels –namely hypoxia—occur naturally in the developing embryo and cells adapt to it. Nonetheless, the role of oxygen as a factor influencing stem cell behavior and developmental trajectories is not well understood. In this study, I dissected the effects of acute and prolonged hypoxia in embryonic and extraembryonic stem cells as well as the functional impact on lineage choices and differentiation potential. I show that, while stem cells maintain their cellular identity in hypoxia, low oxygen levels promote a cell type-specific and temporal transcriptional response.

Specifically, in embryonic stem (ES) cells, hypoxia selectively induces a transcriptional early primitive streak signature with induction of mesendoderm marker genes, such as *Wnt3*, *T* and *Eomes,* without inducing spontaneous differentiation. Mechanistically, I show that HIF1α activation in normoxia recapitulates the induction of developmental genes as those observed in hypoxia. Additionally, low oxygen levels also alter the epigenetic landscape of ES cells, leading to global DNA demethylation and bivalent chromatin modification rewiring.

Last, using a 3D gastruloid differentiation model and in combination with scRNA-seq, I show that hypoxia-induced WNT pathway enables symmetry breaking, polarization and axial elongation in the absence of exogenous WNT activation. When combined with exogenous WNT activation, hypoxia enhances lineage representation by enriching the cell types that are otherwise absent or underrepresented in normoxic gastruloids, such as notochord and gut endoderm, respectively. Moreover, hypoxia provides morphological cues to gut endodermal cells which self-organize in tubular structures reminiscent of the embryonic gut tube.

Taken together, these findings reveal the impact of hypoxia on stem cell behavior and during gastrulation where it modulates morphogenesis and cellular composition in 3D gastrulation models. Hence, my investigation provides a direct link between microenvironmental factors and stem cell functions and strongly supports the use of physiologically relevant oxygen levels in models of embryo development.

Zusammenfassung

Sauerstoff ist für das Leben von Säugetieren von entscheidender Bedeutung. Niedrige Sauerstoffwerte - Hypoxie - kommen im sich entwickelnden Embryo natürlicherweise vor, und die Zellen des Embryos sind daran adaptiert. Jedoch ist die Rolle des Sauerstoffs als Entwicklungsfaktor, der das Verhalten und die Entwicklung von Stammzellen beeinflusst, noch weitgehend unerforscht. In dieser Studie habe ich die Auswirkungen von akuter und lang anhaltender Hypoxie in embryonalen und extraembryonalen Stammzellen sowie die daraus resultierenden funktionellen Auswirkungen auf die Zelllinien-Entscheidung und das Differenzierungspotenzial untersucht. Ich konnte zeigen, dass Stammzellen unter Hypoxie zwar ihre zelluläre Identität beibehalten, niedrige Sauerstoffwerte jedoch eine zelltypspezifische und zeitlich begrenzte Transkriptionsreaktion auslösen.

Insbesondere in embryonalen Stammzellen induziert Hypoxie selektiv eine transkriptionelle frühe Primitivstreifen-Signatur mit einhergehender Induktion von Mesendoderm-Markergenen wie *Wnt3*, *T* und *Eomes*, ohne eine spontane Differenzierung auszulösen. Auf der mechanistischen Ebene zeige ich, dass die HIF1α-Aktivierung in Normoxie die Induktion von Entwicklungsgenen rekapituliert, wie sie in Hypoxie beobachtet werden kann. Darüber hinaus verändern niedrige Sauerstoffwerte auch die epigenetische Landschaft von ES-Zellen, sodass es zu einer globalen DNA-Demethylierung kommt, was wiederum zu einer Neuverdrahtung bivalenter Chromatinmodifikationen führt.

Schließlich zeige ich anhand eines 3D-Gastruloid-Differenzierungsmodells und in Kombination mit scRNA-seq, dass der Hypoxie-induzierte WNT-Signalweg Symmetriebrechung, Polarisierung und axiale Dehnung in Abwesenheit exogener WNT-Aktivierung ermöglicht. In Kombination mit exogener WNT-Aktivierung verbessert Hypoxie die Repräsentation der Zelllinien durch Anreicherung von Zelltypen, die sonst in normoxischen Gastruloiden fehlen oder unterrepräsentiert sind, wie z. B. Notochord und Darmendoderm. Darüber hinaus liefert die Hypoxie morphologische Hinweise für Darmendodermzellen, die sich selbst in röhrenförmigen Strukturen organisieren, welche an das embryonale Darmrohr erinnern.

Zusammengefasst zeigen diese Ergebnisse den Einfluss von Hypoxie auf (i) das Verhalten von Stammzellen und (ii) die Gastrulation, wo die Morphogenese und die zelluläre Zusammensetzung in 3D-Gastrulationsmodellen moduliert werden. Meine Untersuchung stellt somit eine direkte Verbindung zwischen Umgebungsfaktoren und Stammzellfunktionen her und bekräftigt nachdrücklich die Verwendung von physiologisch angemessenen Sauerstoffwerten in Modellen der Embryonalentwicklung.

Contents

1. Introduction

1.1. Developmental biology

Mammalian embryonic development is a highly dynamic process that includes a large range of cellular and molecular mechanisms including cell differentiation, proliferation, movement, organization, migration and communication. These processes are molecularly guided by gene and protein expression, metabolism and epigenetics (Li and Belmonte, 2018; Perrimon et al., 2012; Shparberg et al., 2019).

1.1.1. Pre- and peri-implantation

Embryogenesis starts with the fertilization of the egg and the formation of the zygote (1 cell). This totipotent cell –that will generate all cells in the organism—undergoes rapid mitotic divisions with no significant embryo growth and forms the morula, consisting of 16 cells within the zona pellucida (a protective membrane) (Abbott, 1936). Then, compaction occurs followed by cell differentiation, resulting in the formation of a cavity (blastocoel) in the center. This event forms the so-called blastocyst, which is the end-point of the pre-implantation development (Fong et al., 1998) (**Figure 1A**).

Pre-implantation development in mouse, takes around 3 and a half days from fertilization and it is characterized by the first cell fate decision (**Figure 1B**): (1) a few cells in the pre-implantation embryo will differentiate towards the trophectoderm (TE) lineage. The TE will form the placenta and it is characterized by the expression of CDX2 (Schrode et al., 2013; Zernicka-Goetz et al., 2009). (2) Cells not committing to TE constitute the inner cell mass (ICM). Cells of the ICM are pluripotent and express pluripotency factors such as NANOG, OCT4 and SOX2 (Avilion et al., 2003; Mitsui et al., 2003; Nichols et al., 1998). ICM cells will ultimately form the embryonic germ layers (Chazaud et al., 2006).

The early embryo moves through the fallopian tubes and the zona pellucida disappears to initiate the attachment to the uterine walls (Forgács and Newman 2005) (**Figure 1A**). Right before implantation, the second cell fate decision takes place: cells from the ICM will differentiate into the epiblast (EPI) and primitive endoderm (PrE) (**Figure 1B**). The PrE will contribute mainly to extraembryonic endodermal tissues –the yolk sac and parietal endoderm—and partially to the embryonic endoderm (Nowotschin et al., 2019). It is characterized by the expression of GATA4 and GATA6 (Chazaud et al., 2006; Koutsourakis et al., 1999). The EPI, on the other hand, will give rise to the embryo proper (Schrode et al., 2013). The late-blastocyst (~E4.5-E5.0) implants into the uterine wall and the differentiation

1

events that will generate the future body plan follow (post-implantation development) (**Figure 1**).
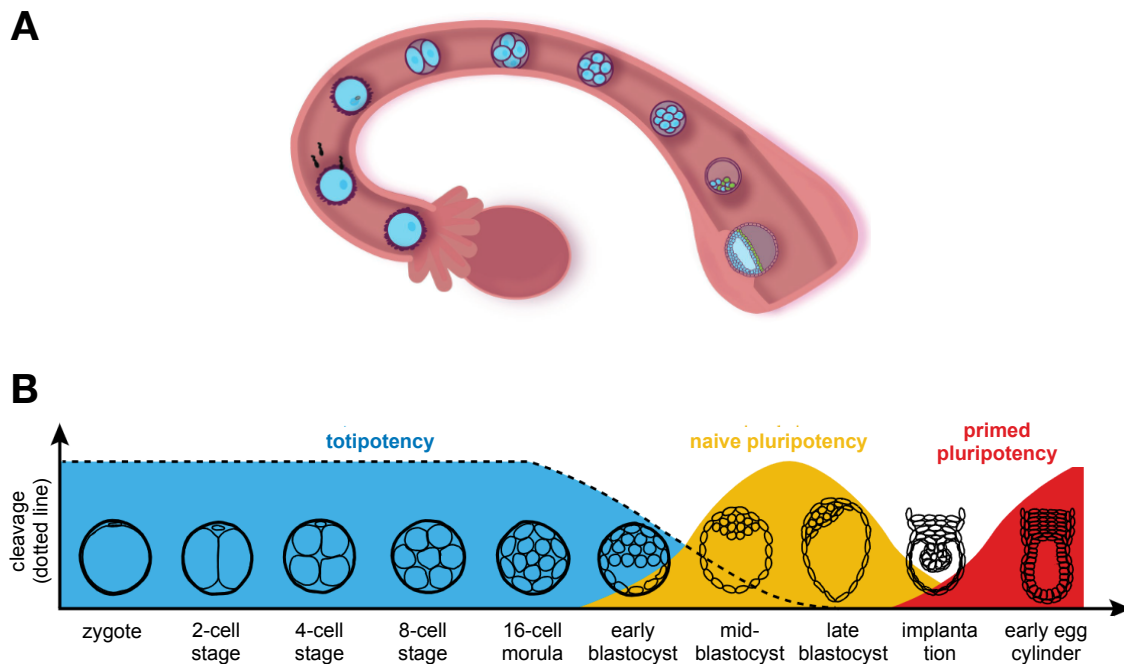
Figure 1. Early pre- and post-implantation mouse embryonic development.

(A) Developmental journey of the early embryo through the reproductive truck until implantation into the uterine walls. (B) Differentiation potential correspondent to the developmental stage from the formation of the zygote until the early egg cylinder. Modified figure from (Boroviak and Nichols, 2014; Shparberg et al., 2019)

1.1.2. Post-implantation development

Within the time window between implantation and gastrulation (E6.5), the blastocyst changes the morphology and transforms into a radially symmetric cylinder with a lumen. At this stage, the embryo is comprised of three epithelial cell populations: the pluripotent epiblast, the enveloping visceral endoderm (VE) and the extraembryonic ectoderm (ExE ectoderm), the latter being the position where the embryo is attached to the uterus (Bardot and Hadjantonakis, 2020) (**Figure 2**). This organization leads to a proximal-distal (P-D) axis, where the proximal is defined by the position of the ExE ectoderm (Rossant and Tam, 2009; Takaoka et al., 2007) (**Figure 2**).

2
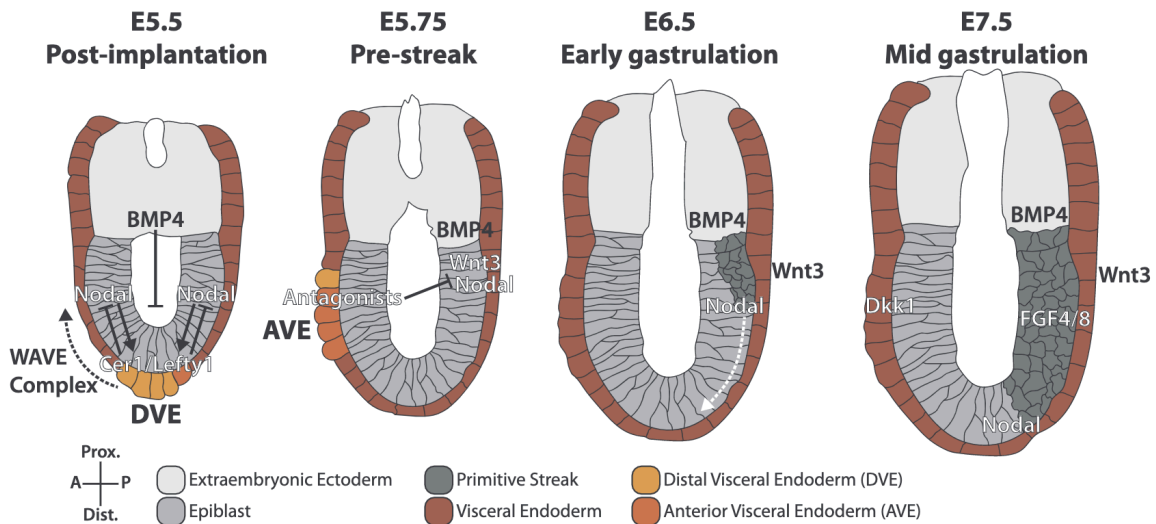
Figure 2. Mouse post-implantation and gastrulation.

Colors shows the distinct cell types present in the post-implantation and gastrulating embryo and their localization with respect the P-D and A-P axis. Arrows depict the signaling gradients and subsequent cellular responses co-occurring at these developmental stages. Image modified from (Bardot and Hadjantonakis, 2020)

At this stage (E5.5) the distal VE (DVE) migrates to the anterior region of the epiblast, forming the so-called anterior VE (AVE), and an important signaling center. The signaling center is generated by the DVE and AVE extraembryonic lineages and induces asymmetry in the epiblast. Thus, in the pre-gastrula (also known as pre-streak) stage (~E5.5-E6.0), the AVE secretes antagonists of BMP, Nodal and WNT signaling into the epiblast. At the most posterior side of the epiblast, from the ExE ectoderm, BMP4 initiates the expression of Nodal and WNT into the epiblast (Arias et al., 2022; Bardot and Hadjantonakis, 2020). As the AVE secretes antagonists of BMP, the signaling is restricted to the posterior axis of the embryo (Rivera-Pérez and Hadjantonakis, 2015). Thus, the anterior-posterior (A-P) axis is established and gastrulation initiates.

1.1.2.1.   Gastrulation

Gastrulation is the process by which the pluripotent epiblast differentiates into the three germ layers (ectoderm, mesoderm and endoderm) that will ultimately form all cell types found in the adult organism (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Tam and Behringer, 1997) (**Figure 3**).
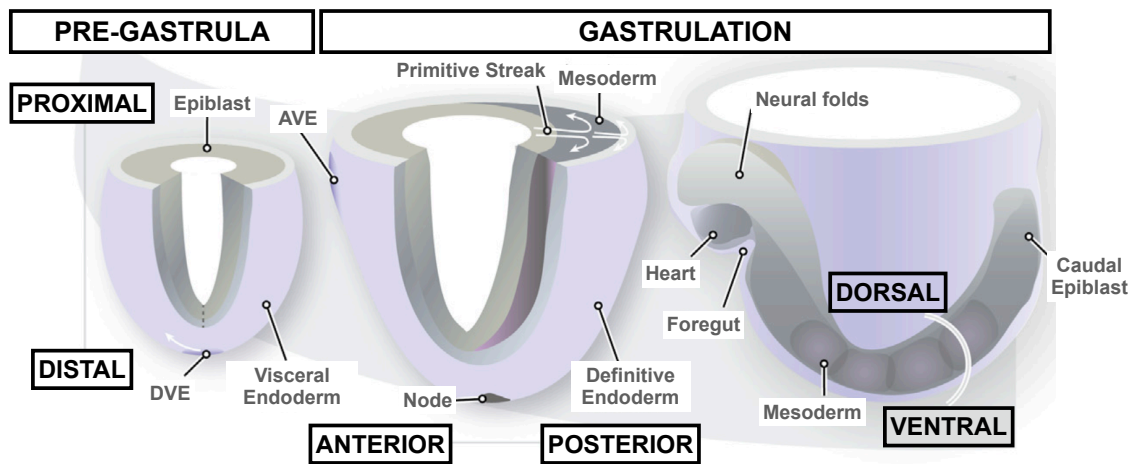
Figure 3. Mouse gastrulation.

Representation of mouse gastrulation, highlighting the differentiation of the pluripotent epiblast into the derivates of the three germ layers. Images shows the recapitulation of the cellular movement along the multiaxial system creating the body plan of the developing embryo. Modified figure from (Arias et al., 2022)

The high levels of WNT and Nodal in the posterior axis of the embryo induce local T expression. The main role of T –together with WNT signaling— is to drive gastrulation and to form the primitive streak (PS). Cells from the PS undergo epithelial-to-mesenchymal transition (EMT) forming the first mesenchymal cell population that will organize into the three germ layers (Williams et al., 2012).

At this stage (E6.5-E7.5) the embryo grows very fast, the ExE cells constrain and push cells from the epiblast through the PS (Williams et al., 2012). The coordination of this movement is highly regulated and controlled by WNT, BMP and FGF factors, shaping the outline of the body plan (Mohamed et al., 2004; Sun et al., 1999; Winnier et al., 1995).

The end of the EMT in the PS occurs and leads to the formation of the nascent mesoderm. The endodermal cells are also partially arising from the PS (E6.5) (Lawson and Pedersen, 1987). Yet, the exact origin of the endoderm germ layer is still unknown. On one hand, it has been hypothesized to arise from a bipotent population of cells able to give rise to both mesodermal and endodermal lineages, the so-called 'mesendoderm'. However, there is also evidence showing that the endoderm might originate from cells that are distinct from mesodermal ones (Probst et al., 2020; Scheibner et al., 2021).

Around E7.5, the PS meets the most distal region of the embryo, and the node appears (Arias et al., 2022) (**Figure 3**). The endoderm and mesoderm come together into an anterior region (also from the node), mediating the induction of neuroectoderm on the opposite side of the PS

(Camus et al., 2000). This is the anterior neuroectoderm that will generate the future head and brain (Tam and Behringer, 1997).

Between E7.5-E8.5, driven by cellular movement, the cylinder structure disappears and the embryo organizes into an elongated body plan (**Figure 3**). In parallel, a bipotent population of cells located proximal to the node, the so-called neuromesodermal progenitors (NMPs) will give rise to paraxial mesoderm and the future spinal cord cells (Wymeersch et al., 2021)

The process of gastrulation can be summarized in the following sequential events: (1) establishment of the multiaxial organizational system, (2) emergence of tissues organized according to the axial system and (3) subsequent elongation and expansion of the embryo, creating the body plan (Arias et al., 2022).

1.2. Stem cells and in vitro models of embryogenesis (Stembryos)

1.2.1.   Stem cells derived from the mouse pre-implantation blastocyst

At the pre-implantation stage, the late blastocyst (E4.5) comprises three distinct cell types: the epiblast (EPI), the primitive endoderm (PrE) and the trophoectoderm (TE) (see '*Developmental biology*' section for more information). The pre-implantation embryo is suited to study cell fate decisions in vivo, nonetheless, the limited cell number makes it challenging to find out molecular mechanisms. This limitation can be overcome by using stem cells that are derived from the early embryo (Garg et al., 2016)(**Figure 4**). Embryonic stem (ES) cells and epiblast stem cells (EpiSCs), derived from the pre- and post-implantation epiblast, respectively, provide invaluable sources to dissect mechanisms of pluripotency (Evans and Kaufman, 1981); (Martin, 1981) (Brons et al., 2007; Tesar et al., 2007) (**Figure 4**). Likewise, stem cells representing the extraembryonic lineages can also be derived from the preimplantation embryo. These are the extraembryonic endoderm (XEN) stem cells, resembling the PrE (Kunath et al., 2005) and the trophoblast stem (TS) cells resembling the TE (Tanaka et al., 1998)
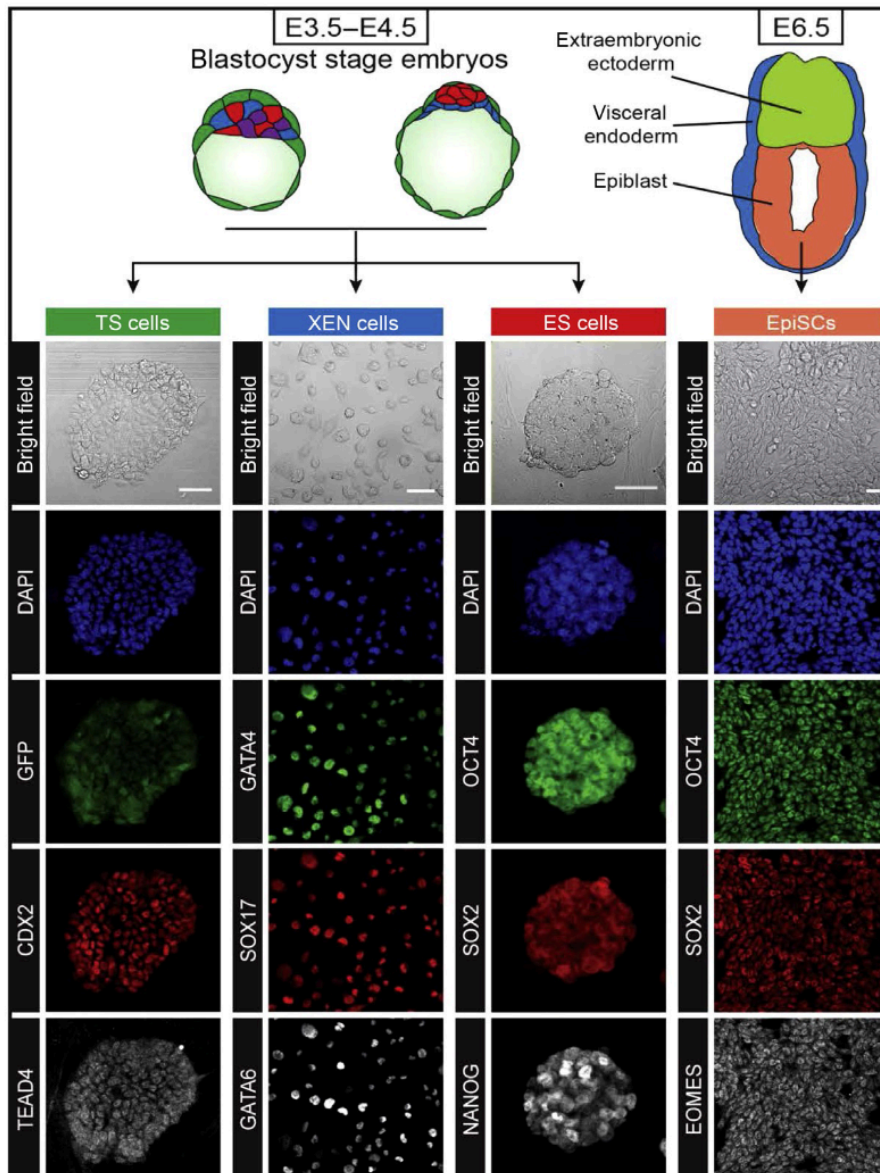
Figure 4. In vitro stem cell types derived from the pre-implantation and post-implantation blastocyst.

TS, XEN and ES cells represent the TE, PrE and EPI, respectively. EpiSCs represent the post-implantation epiblast. All stem cell types have self-renewal capacity, distinct morphology and a specific gene profile. Bright-field and immunofluorescence images show morphology and expression of marker genes. TS cells used in the analysis ubiquitously express GFP. Modified figure from (Garg et al., 2016)

1.2.1.1.   TS cells

TS cells are cells derived from the blastocyst or extraembryonic ectoderm of the post-implantation embryo (E6.5-E8.5) (Tanaka et al., 1998). They are similar to the TE at the transcriptomic level –expressing marker genes such as *Esrrb*, *Cdx2* and *Eomes*—and at the morphological and potential level (Beck et al., 1995) (**Figure 4**).

1.2.1.2.   XEN cells

XEN cells are the in vitro counterparts of the PrE (Kunath et al., 2005). They express markers of both the PE and the VE, and when introduced in the blastocyst, they preferentially contribute to the PrE (Fujikura et al., 2002). As such, they are more similar to the PrE morphologically (**Figure 4**).

1.2.1.3.   ES cells

ES cells are pluripotent and self-renewing cells derived from the pre-implantation EPI which can be cultured in vitro indefinitely (Beddington and Robertson, 1989; Boroviak et al., 2014).

Pluripotency is the capacity of a single cell to give rise to all specialized cell types present in the adult organism. ES cells retain pluripotency potential while undergoing self-renewal without losing their cell identity (Hackett and Surani, 2014) (**Figure 4**).
When subjected to cell differentiation cues, ES cells exit pluripotency, lose self-renewal capacity and differentiate (Bradley et al., 1984). Despite being generally categorized as pluripotent, ES cells with distinct transcriptional, epigenetic, and metabolic profiles vary along the pluripotency spectrum. Below, I focus on naïve pluripotency due to its relevance to our investigation.

1.2.1.3.1.   Naïve state of pluripotency

ES cells at the naïve state of pluripotency can be maintained under different cultured conditions. ES culture media is supplemented with leukemia inhibitor factor (LIF) and under Serum/LIF conditions are known as the 'conventional' or 'Serum/LIF' ES cells. As markers of pluripotency, they express SOX2, OCT4 and to some extent NANOG (Dunn et al., 2014; Young, 2011). On the other hand, ES cells can also be cultured with 2i/LIF instead of serum. 2i components correspond to a specific inhibitor of FGF/ERK and GSK3, which together act against differentiation (Ying et al., 2008;Jaenisch and Young, 2008). 2i/LIF ES cells present a distinct transcriptional and epigenetic state compared to serum/LIF ones, such as uniform expression of NANOG and PRD14, and global DNA hypomethylation (Hackett and Surani, 2014).

1.2.1.3.2.   Definition of the ground state of pluripotency

Both serum/LIF and 2i/LIF ES cells are able to maintain naïve pluripotency. Yet, serum/LIF ES cells present an altered transcription and epigenetic profile when compared to the in vivo pre-implantation epiblast cells. 2i/LIF ES cells, on the other hand, present a transcriptome that closely resembles the developmental ground state of pluripotency in vivo. Therefore, serum/LIF ES cells are considered to have a naïve but not ground state of pluripotency (Hackett and Surani, 2014; Marks et al., 2012). (**Figure 5**)
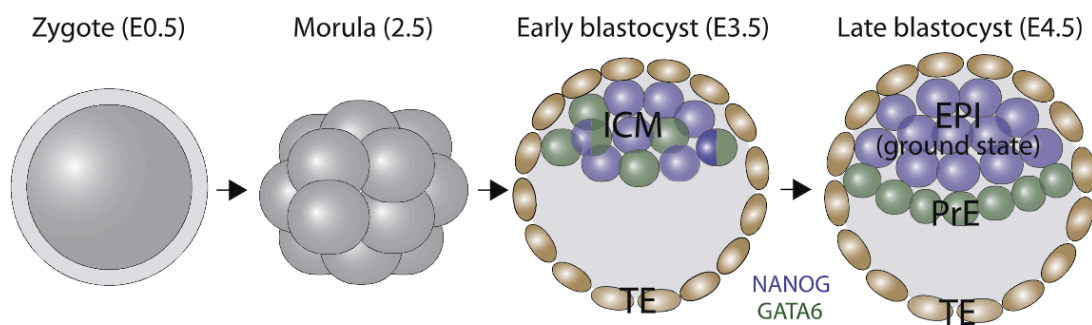
Figure 5. The ground state of pluripotency.

Development of the zygote gives rise to the early blastocyst (E3.5), which contains the TE and the ICM. The ICM cells express both NANOG (blue) and GATA6 (green) at E3.5. Then, the ICM of the late blastocyst (E4.5) segregates into the naïve pluripotent EPI and the extraembryonic PrE which express NANOG (blue) and GATA6 (green), respectively and in a mutually exclusive fashion. The naïve pluripotent EPI at the late blastocyst stage is the source of all embryonic lineages and concomitantly represents the "ground" state of pluripotency. Figure from (Hackett and Surani, 2014)

1.2.2.   In vitro models of embryogenesis (Stembryos)

Studying embryogenesis in vivo is sometimes challenging due to its small size, inaccessibility and experimental manipulation providing a limited readout. The appeal of in vitro models is based on the ability of scalability, trackability and reproducibility, providing the possibility of manipulation at different levels such as controlling signaling gradients, mechanical forces or environmental factors that would be more challenging in vivo. Therefore, developmental biologists have been working on the idea of developing in vitro models of differentiation. The initial idea was influenced by the observation that teratocarcinoma cell lines could aggregate and differentiate in vitro. This event was defined as the generation of embryoid bodies (EBs) (Martin and Evans, 1975; Pierce and Verney, 1961).

EBs have been used to study differentiation using mouse ES cells. They are three-dimensional (3D) aggregates in a free-floating culture fashion under a combination of controlled culture

media conditions and allowing the cells to interact within them. They are useful to generate specific cell lineages under distinct controlled culture conditions (Brickman and Serup, 2016). Yet, these structures lack patterning, leaving open questions on how the embryo achieves the body plan formation observed during the early stages of embryogenesis.

Only within the last 15 years, advances in synthetic biology have allowed the development of many in vitro models of mouse embryogenesis. Indeed, the number of systems that are emerging recently led to the generation of a unique discipline: 'stembryogenesis' –defined as the process by which 'stembryos' are formed, being 'stembryo' an umbrella term for all in vitro models of embryo development that starts with the use of stem cells (Veenvliet et al.).

1.2.2.1.   Gastruloids

Complex multi-tissue organoids can be generated by aggregating ES cells and providing signaling cues. Modification of the EB protocol by introducing a WNT activator resulted in the generation of the so-called 'gastruloids' (Brink et al., 2014). Gastruloids are 3D structures that under specific conditions recapitulate aspects of gastrulation such as polarization, symmetry breaking and axial elongation. Gastruloids are characterized by the emergence of the three germ layers and the establishment of the three embryonic body axis which shows spatial domains with specific gene expression patterns (Beccari et al., 2018; Brink et al., 2014; Turner et al., 2017) (**Figure 6**).
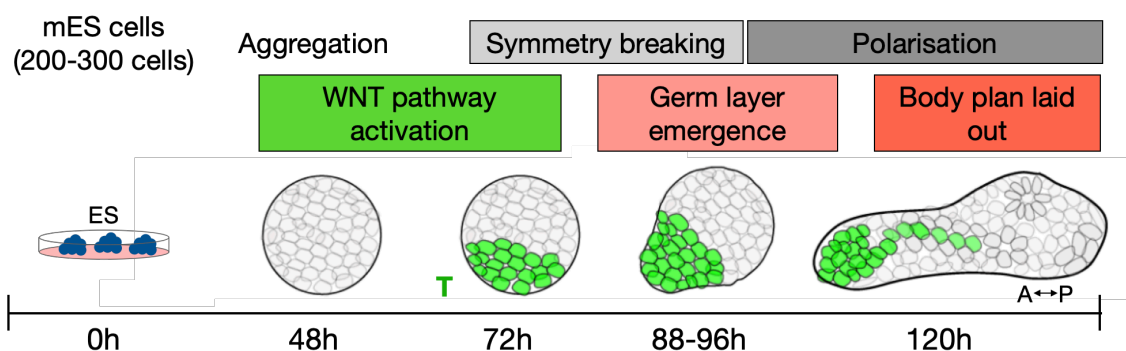
Figure 6. Conventional gastruloid development.

Schematic representation of the cellular and morphological events occurring during gastruloid formation. Mouse ES cells (200-300) are cultured in individual wells and allowed to aggregate for 48h. After aggregation, CHIRON is used for 24h to induce polarization of the aggregates with the localized expression of T in the future posterior axis. Gastruloids break symmetry and form the posterior domain and elongate resulting in the formation of an elongated structure resembling the body plan of the in vivo gastrulating embryo. A: anterior axis; P: posterior axis. Modified figure from https://www.embl.org/groups/trivedi/.

The original protocol begins with the aggregation of ~200-300 ES cells. After 48h, aggregates are exposed for 24h to chemical WNT activation using CHIR99021 –hereafter Chi—to mimic the downstream effects of the activation WNT pathway: the induction of the PS domain and polarization of T expression in the posterior axis (Beccari et al., 2018; Izpisúa-Belmonte et al., 1991). After the withdrawal of WNT signaling stimulation, structures break symmetry, self-organize and elongate and by 120h post-aggregation, they acquire the cellular and molecular complexity to resemble the E8.5 developing embryo (Brink et al., 2014) (**Figure 6**).

Single-cell RNA-sequencing (scRNA-seq) approaches identified the cell diversity of gastruloids where cell types belonging to the three germ layers were found. Yet, there is an absence of anterior neural (brain) cell types, likely linked to the absence of extraembryonic tissues (Rivera-Pérez and Hadjantonakis, 2015; Takaoka and Hamada, 2011). AVE secretes antagonists of WNT and Nodal signaling and promotes anterior neural development. The absence of VE and therefore AVE in gastruloids leaves the anterior axis unprotected from WNT-driven posteriorizing events, therefore inhibiting the development of anterior neural structures. This is supported by two studies where WNT inhibition in early aggregates led to the appearance of anterior neural structures (Girgin et al., 2021a) and co-culture approach, where two aggregates were used: a small one exposed to WNT activation and a big one with no signaling center. This approach generated gastruloids with both posterior and anterior domains with brain-like structures, and the appearance of notochordial-like cells which are absent in conventional gastruloids (Girgin et al., 2021b).

1.2.2.1.1.   Bringing gastruloids into shape

Even though gastruloids present the spatial domains and gene expression patterning of the post-implantation embryo, they lack the typical aspects of embryo architecture (Veenvliet et al.). For instance, the cardiac mesoderm does not form the heart tube (Rossi et al., 2021),the pre-somitic mesoderm (PSM) does not organize into somites (Beccari et al., 2018) and the neural cells do not form the neural tube (Beccari et al., 2018). Two recent approaches have shown that the complex architecture of gastruloids can be achieved by adding Matrigel, an extraembryonic extracellular matrix (ECM) of natural origin (Aisenbrey and Murphy, 2020). It provides biochemical and mechanical cues, resulting in complex morphogenesis in gastruloids (Bolondi et al., 2021a; Brink et al., 2020; Veenvliet et al., 2020a). In one of them, 10% of Matrigel added to the culture media at 96h post-aggregation led to a single string of somites-like structures organized along the A-P axis (Brink et al., 2020). In the other one, 5% of Matrigel added at the same time point, showed a better recapitulation of the embryo-like architecture

with somite-like structures adjacent to a neural tube-like structure, thus making so-called Trunk-like structures (TLS) (Veenvliet et al., 2020a) (**Figure 7**).
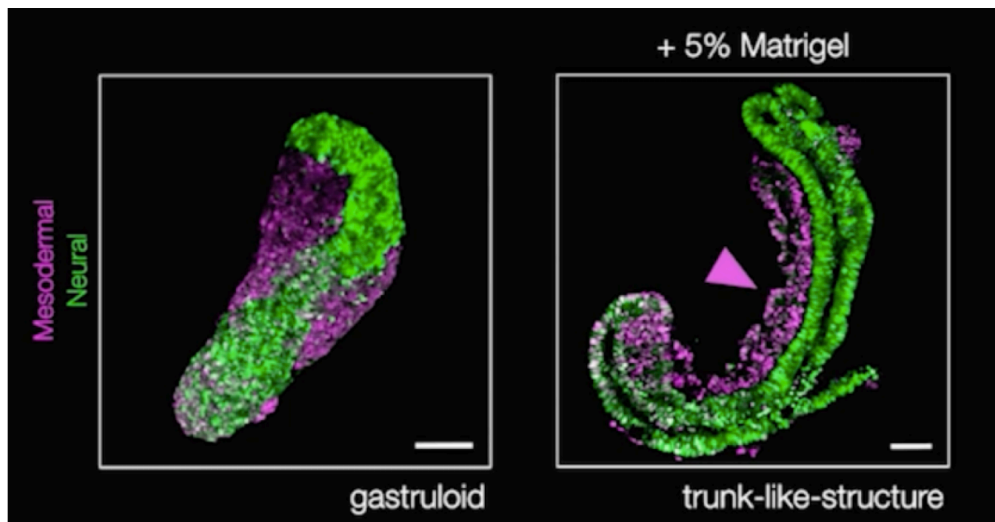
Figure 7. Comparative morphology of gastruloids and TLS.

A T::H2B-mCherry; Sox2::H2B-Venus double reporter mouse ES cell line was used to identify the mesodermal (magenta) and neural (green) progeny appearance, patterning and self-organization in gastruloids and TLS. Scale bars 100um. Modified figure from (Veenvliet et al., 2020a)

Nonetheless, these advances also come with some limitations. For instance, TLSs show high variability, with ~40% of the structures featuring multiple axes. Similarly, ~40% of the structures are biased towards the formation of the neural tube only, and among the structures presenting somites, only ~60% were bilateral somites (Veenvliet et al., 2020a). This high variability highlights the need to find a 'sweet spot' for a more robust TLS protocol, likely including other missing factors to better recapitulate architectural events during gastrulation.

1.2.2.2. Other stembryo models

From the first gastruloid generation, the stembryogenesis field has grown enormously to study in vitro distinct developmental stages during embryogenesis (**Figure 8**). Blastoids were generated to recapitulate the main developmental transitions occurring to form the blastocyst by using mouse ES and TS cells. (Rivron et al., 2018).
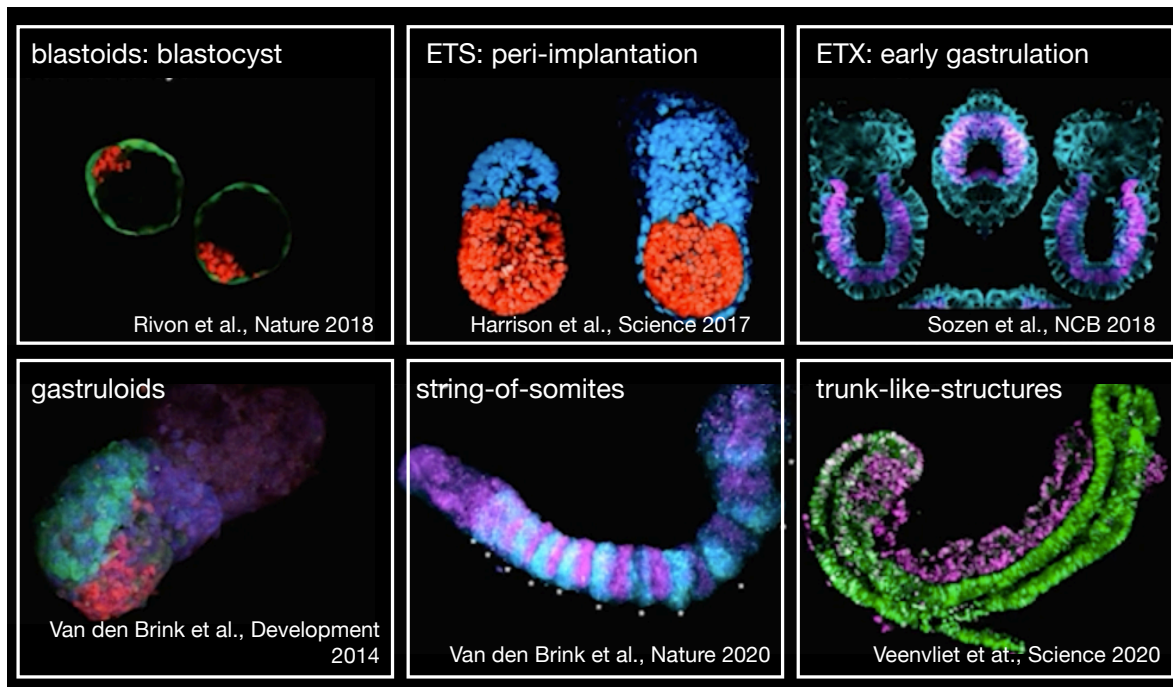
Figure 8. Mouse in vitro models of embryogenesis.

Schematic representation of the distinct in vitro models of embryogenesis recapitulating different stages of early mouse embryogenesis. Figure generated based on studies from (Brink et al., 2014; Brink et al., 2020; Harrison et al., 2017; Rivron et al., 2018; Sozen et al., 2018; Veenvliet et al., 2020a)

To study peri-implantation, a system using also mouse ES and TS cells, and in the presence of Matrigel, established the so-called ETS embryos (Harrison et al., 2017). A similar approach where Matrigel was substituted by the derivation and maintenance of XEN cells –mimicking the presence of PrE and VE during stembryogenesis—recapitulated peri-implantation too, these are known as ETX embryos (Sozen et al., 2018). This system is enhanced by the fact the ETX embryos undergo symmetry breaking and EMT without the need for exogenous WNT activation, indicating the relevance of the extraembryonic endoderm to instruct the epiblast to start gastrulation (Rivera-Pérez and Hadjantonakis, 2015).

The most recent advances have shown that synthetic embryos recapitulating pre-, peri- and post-implantation can be achieved starting solely from naïve ES cells and reaching the post-gastrulation stage (~E8.5) (Lau et al., 2022; Tarazi et al., 2022) (**Figure 9**). These synthetic embryos present all embryonic and extraembryonic compartments, form organ-specific progenitors and develop a beating heart and the primordium of the brain (Lau et al., 2022; Tarazi et al., 2022).
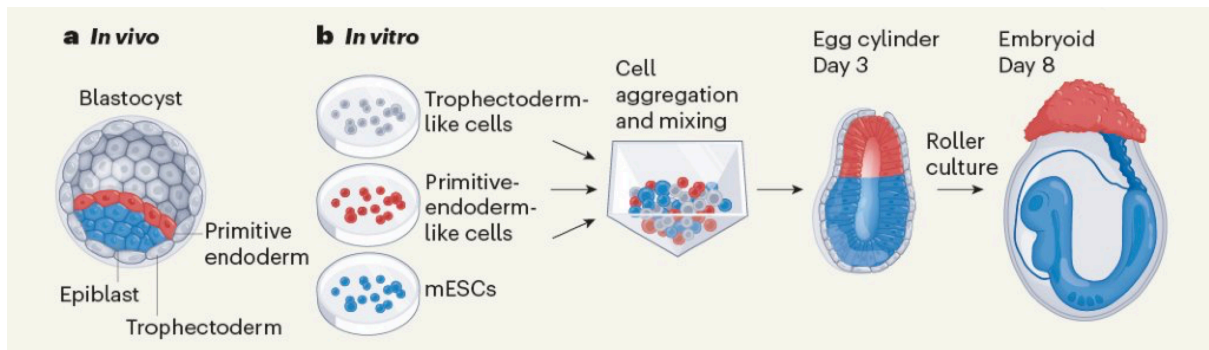
Figure 9. Recapitulating mouse early embryogenesis purely from stem cells in vitro.

Co-culture of ES, trophoectoderm-like and primitive endoderm-like cells under an optimized culture medium and the use of a rotating culture chamber, are able to recapitulate principles of self-organization giving rise to mid-gestation-like mouse embryos. Figure from (Amin and Pașca, 2022)based on (Amadei et al., 2022; Tarazi et al., 2022) studies.

They also present limitations in terms of survival. For example, using the same system, ex-uterus natural embryos were able to survive up to E11. On the other hand, synthetic embryos did not survive for longer than E8.5. Additionally, only 0.1%-0.5% of the initial aggregates would generate synthetic embryos until E8.5 (Lau et al., 2022; Tarazi et al., 2022).

1.3. The role of oxygen tension in embryogenesis

Environmental oxygen (O_2) is essential for mammalian life. Oxygen is an important regulator of cellular physiology and it is crucial to sustaining oxidative metabolism. Low oxygen levels occur naturally in the developing embryo and cells can adapt to it. When O_2 levels decrease, there is a cellular response to adapt to it. This response is driven by a family of transcription factors known as 'Hypoxia Inducible Factors' (HIFs) which are oxygen-sensing genes (Simon and Keith, 2008). This response to hypoxia leads to a metabolic switch to glycolysis and a wide range of other cellular responses at the transcriptomic, proteomic and epigenomic levels (Chakraborty et al., 2019; Houghton, 2021; Lee et al., 2020).

Hypoxia is defined as a state in which levels of oxygen are decreased when compared to atmospheric levels of oxygen (~20%). Atmospheric oxygen tensions (20% O_2) is therefore defined as normoxia. The term 'physiological oxygen levels' has been broadly used as an equivalent to atmospheric oxygen levels, even though during early embryogenesis, low levels of O_2 occur in a physiological manner (Houghton, 2021).

1.3.1. HIF transcription factors

In mammals, HIF1α is a transcription factor and master regulator of any hypoxic response (Semenza, 1999), activating the hypoxia response signaling pathway and regulating the expression of several genes. Some examples of hypoxia-responsive activated genes are VEGF, LOXL2, LDHA, PDK1 or KDM6A, coordinating therefore many biological processes such as angiogenesis, EMT, pyruvate metabolism, glycolysis and cell fate transitions and differentiation through chromatin remodeling (Batie et al., 2019; Chakraborty et al., 2019; Iyer et al., 1998; Yang et al., 2008).

HIFs are heterodimeric transcription factors consisting of two subunits: HIF-α (HIF1α, HIF2α and HIF3α) and HIF-ß (ARNT [HIF1ß] and ARNT2 [HIF2ß]) (Simon and Keith, 2008) (**Figure 10**).
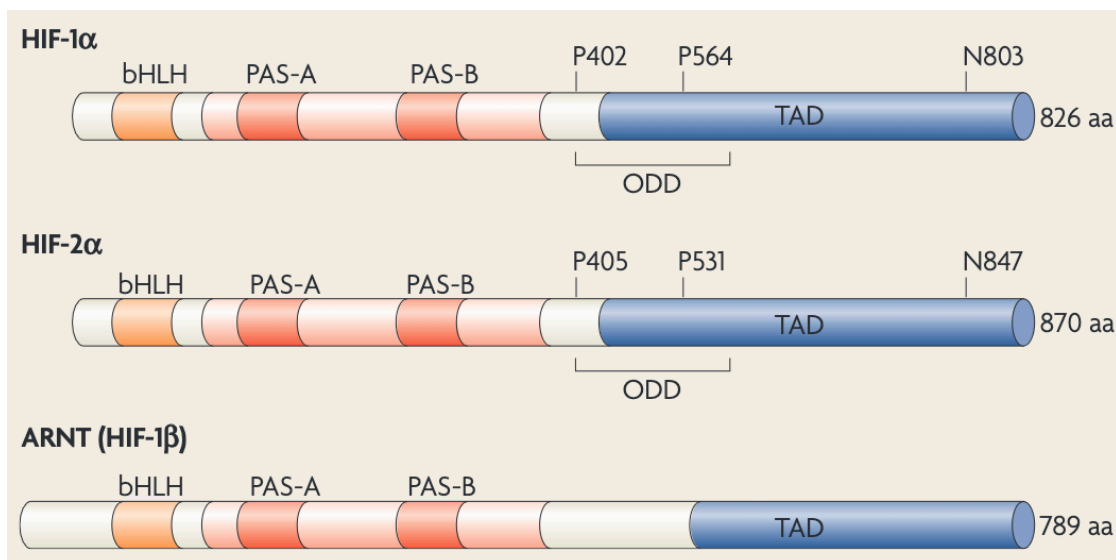
Figure 10. Hypoxia-inducible factors.

The figure shows the similarities and differences at the protein subunit level of two HIF-α (HIF1α and HIF2α) and one HIF-ß (ARNT [HIF1ß]). bHLH: basic helix-loop-helix; PAS: Per-Arnt-Sim; ODD: oxygen-dependent degradation domain: TAD: transcriptional transactivation domain. Modified figure from (Simon and Keith, 2008)

The subunits interact via two Per-Arnt-Sim (PAS) domains, bind DNA via N-terminal basic helix-loop-helix (bHLH) domains, and activate transcription with C-terminal transcriptional transactivation domains (TADs) (**Figure 10**). In mammals, HIF1α, is constitutively expressed, while the expression of HIF1ß is highly dependent on the cellular oxygen concentration (Lee et al., 2020).

1.3.2. Hypoxia signaling pathway

Under normoxia, HIF1α is hydroxylated at two conserved proline residues by specific Prolil Hydroxylases (PHDs) at the Oxygen-dependent Degradation Domain (ODD) **(Figure 10)** in a reaction that requires oxygen. HIF1α hydroxylation enables the binding of von Hippel-Lindau protein (pVHL) (Jaakkola et al., 2001). pVHL forms the recognition domain, targeting HIF1α poly-ubiquitination and proteasomal degradation (Ohh et al., 2000) **(Figure 11)**.
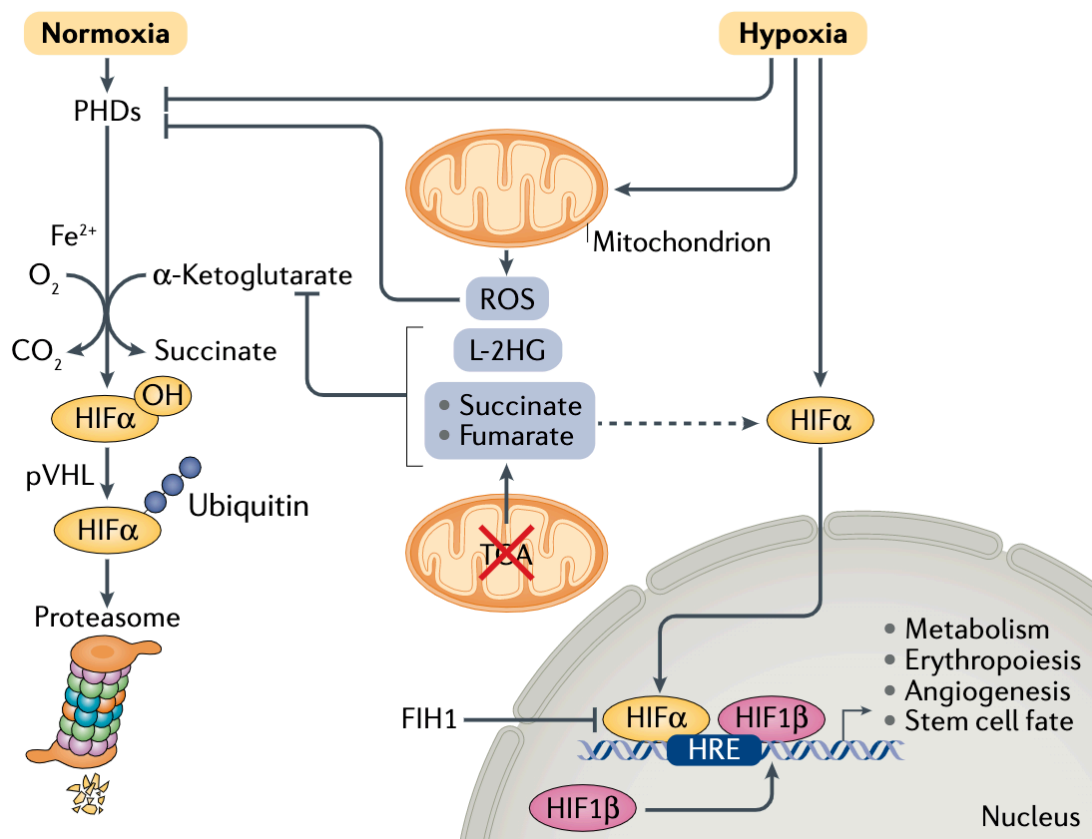
Figure 11. Hypoxia signaling pathway.

Under normoxia, there is an oxygen-dependent interaction between HIF1α and the PHDs. This interaction hydroxylases HIF1α, which will afterward be polyubiquitylated by pVHL and marked for proteasomal degradation. On the other hand, under hypoxia, PHDs are inhibited, leading to HIF1α stabilization and translocation to the nucleus where it dimerizes with HIF1ß and binds to HRE to modulate the expression of HRGs. Modified figure from (Lee et al., 2020).

In hypoxic conditions, PHDs are repressed, and pVHL binding is reduced. HIF1α is stabilized and translocated to the nucleus where it dimerizes with HIF1ß and binds to 'Hypoxia Responsive Elements' (HRE), to modulate the expression of HRGs (Lee et al., 2020; Maxwell et al., 2001) **(Figure 11)**.

1.3.3. Oxygen availability during embryo development

In vivo, the pre-implantation blastocyst is exposed to a hypoxic environment, with O_2 levels going from ~8% in the oviduct and gradually decreasing through the reproductive tract up to ~1.5% in the uterus, where the early embryo implants (Houghton, 2021). Additionally, early embryogenesis takes place before the establishment of blood circulation, indicating a microenvironment with low oxygen levels (Fischer and Bavister, 1993; Ottosen et al., 2006; Simon and Keith, 2008).

Importantly, it has been demonstrated that not only O_2 levels but also the timing at which embryos are exposed to distinct O_2 tensions, are important (Nguyen et al., 2020; Pabon et al., 1989). Recently, this was also observed when culturing embryos from E6.5 onwards ex utero – where a sequential increase of O_2 from 5% to 21% at E10.5 significantly increased the output of ex utero mouse development (Aguilera-Castrejon et al., 2021). These data together highlight the importance of O_2 tensions in a time-dependent manner during early embryogenesis and suggest that mimicking O_2 levels found in the oviduct and uterus might be beneficial in in vitro systems.

1.3.4. Biological processes occurring in a hypoxic context

1.3.4.1. Epithelial to mesenchymal transition (EMT)

During gastrulation, the PS undergoes gastrulation (see '*Gastrulation section*' for more information). A subset of epiblast cells moves to the midline to form it and internalization of cells from the PS forms the primary mesenchyme through EMT.

EMT is triggered by low oxygen tensions in various cancer contexts (Choi et al., 2017; Misra et al., 2012; Yeo et al., 2017). The EMT switch to differentiation and its behavior is mediated by transcription factors such as SNAI1 and TWIST, modulating the expression of epithelial and mesenchymal markers (Gheldof and Berx, 2013; Yang et al., 2008). HIFs promote the EMT process by regulating the expression of the above-mentioned genes together with other EMT regulators (Peinado et al., 2007). To date, there is no direct evidence showing that EMT is triggered by hypoxia in the context of development. Nonetheless, it is likely occurring in a similar manner as in cancer contexts.

1.3.4.2. Vasculature development and angiogenesis

Vasculogenesis is defined as the differentiation of precursor cells (angioblasts) into endothelial cells and the de novo formation of the primitive vasculature network. Angiogenesis is the process by which new blood vessels are formed from the existing blood vessels (Risau, 1997). Vasculature morphogenesis is regulated by factors including VEGF.

The importance of oxygen levels in vasculogenesis and angiogenesis was discovered when HIF1ß (-/-) mice were shown to be lethal at E10.5 due to inhibition of vasculogenesis and angiogenesis (Iyer et al., 1998; Ryan et al., 2005). Furthermore, HIF1ß (+/-) deficient mice showed decreased VEGF expression, an increased number of apoptotic cells and decreased number of hematopoietic stem cells (HSCs). Similarly, HIF1α (-/-) mice showed defective blood vessel formation (Iyer et al., 1998; Kozak et al., 1997). VEGF and therefore vasculogenesis and angiogenesis are modulated by low oxygen levels in the early embryo (Mahon et al., 2001; Rey and Semenza, 2010; Risau, 1997).

1.3.5. Contributors to the hypoxia response

1.3.5.1. Metabolic switch toward glycolysis

The choice of the metabolic pathway as a source of energy production depends on oxygen and nutrient availability. Glucose is indispensable during early stages of development, specifically at the morula to blastocyst transition. The blastocyst is characterized by a bivalent metabolic profile where both glycolysis and TCA cycle coupled to OxPhos are functional to generate energy in form of ATP (Folmes et al., 2012). Recent single-cell transcriptomic analysis has shown that there is a metabolic switch toward glycolysis during early embryonic development (Malkowska et al., 2022). Concomitantly, upon implantation, the embryo is exposed to a hypoxic environment that is likely to promote the switch from aerobic to anaerobic metabolism, due to the low oxygen availability (Folmes et al., 2012).

At the implantation stage and during the first days after implantation, glucose is converted into lactate (Johnson et al., 2003). This conversion is driven by HIF1 which increases glycolysis and promotes the conversion of pyruvate to lactate by increasing the expression of LDHA (Mathupala et al., 1997; Semenza et al., 1994). Thus, there is a clear link between environment and metabolism, where metabolic rewiring is driven by oxygen availability and therefore, it is essential in shaping embryo development and the adaptation to the maternal uterus.

1.3.5.2. Histone modifiers as oxygen sensors

Activation of HIFs upon hypoxia is mediated via inhibition of dioxygenases, being prolil-hydroxylases (PHDs) the most prominent. Yet, the dioxygenase family includes Jumanji C (JmjC) domain histone lysine demethylases (KDMs), the tet-eleven translocation (TET) DNA hydroxylases, among ~50 other enzymes (McDonough et al., 2010).

It is known that oxygen regulates histone methylation, as a form of epigenetic chromatin modification (Shmakova et al., 2014). Yet, it was not clear until recently, whether histone methylation was a downstream effect regulated directly by KDMs sensing oxygen or it was an indirect effect. In recent years, two specific lysine demethylases –KDM5A and KDM6A—have been identified as oxygen sensors (Batie et al., 2019; Chakraborty et al., 2019). For instance, KDM6A directly senses oxygen and its hypoxia-mediated inactivation leads to hypermethylation of H3K27 preventing cellular differentiation in several cell line model systems (Chakraborty et al., 2019). These complementary studies presented for the first time a direct link between hypoxia and cellular events by means of direct epigenetic sensing of oxygen availability.

1.3.5.3. Translation suppression under hypoxia

Protein synthesis is an energy-expensive process and it has been shown to be suppressed under stress conditions such as oxygen deprivation (Lee et al., 2020). Thus, preventing both the generation of "unnecessary" proteins and the accumulation of stressed-induced misfolded proteins.

Despite a global reduction of translation, overcoming translational repression also occurs and is necessary for genes that are required for an adaptive response (Harding et al., 2000; Harding et al., 2003). As such, de novo synthesis of proteins essential for the adaptive response, such as ATF4 (activating transcription factor 4) induces the expression of genes involved in protein synthesis or metabolism, enabling the survival response.

2. Aim of this study

Despite rich evidence pointing to hypoxia and the HIF response as essential components of mammalian embryonic development, how hypoxia impacts the cell types present in the pre- and post-implantation blastocyst is not fully resolved.

In my PhD, I set out to understand the transcriptional response to hypoxia in embryonic and extraembryonic stem cells and how its effect can be harnessed to enhance embryo models.

To this end, I first systematically studied the molecular effects of acute and prolonged hypoxia in ES, TS and XEN cells.

Then, I focused on ES cells to dissect components of the hypoxia response, including the role of HIF1α and epigenetic mechanisms including histone and DNA methylation.

Last, I applied this knowledge to gastruloids and to investigate hypoxia-driven changes in morphology and cellular diversity.

3. Results

FIRST PART: Impact of hypoxia on mouse embryonic and extraembryonic stem cells

To investigate the impact of hypoxia on these three stem cell types, we used ES cells ((Evans and Kaufman, 1981)), XEN cells (Kunath et al., 2005) and TS cells (Tanaka et al., 1998) to model them in vitro, respectively (**Figure 12**).

The acute response to hypoxia is very fast and occurs within minutes to hours (Pagé et al., 2002). As the embryo might be exposed to low oxygen tension for about 9-10 days, when the placentation begins (Woods et al., 2018), we aimed to study the temporal dynamics of oxygen levels for a prolonged period of time. Therefore, we cultured ES, TS and XEN cells in normoxia (20% oxygen) and hypoxia (2% oxygen) for 2 days (acute response) and for 7 days (prolonged response) (**Figure 12**).

Initial characterization performed by my former colleague Marina Typou showed that hypoxic stem cells are adaptable to hypoxia, without compromising their cell state and remaining undifferentiated.
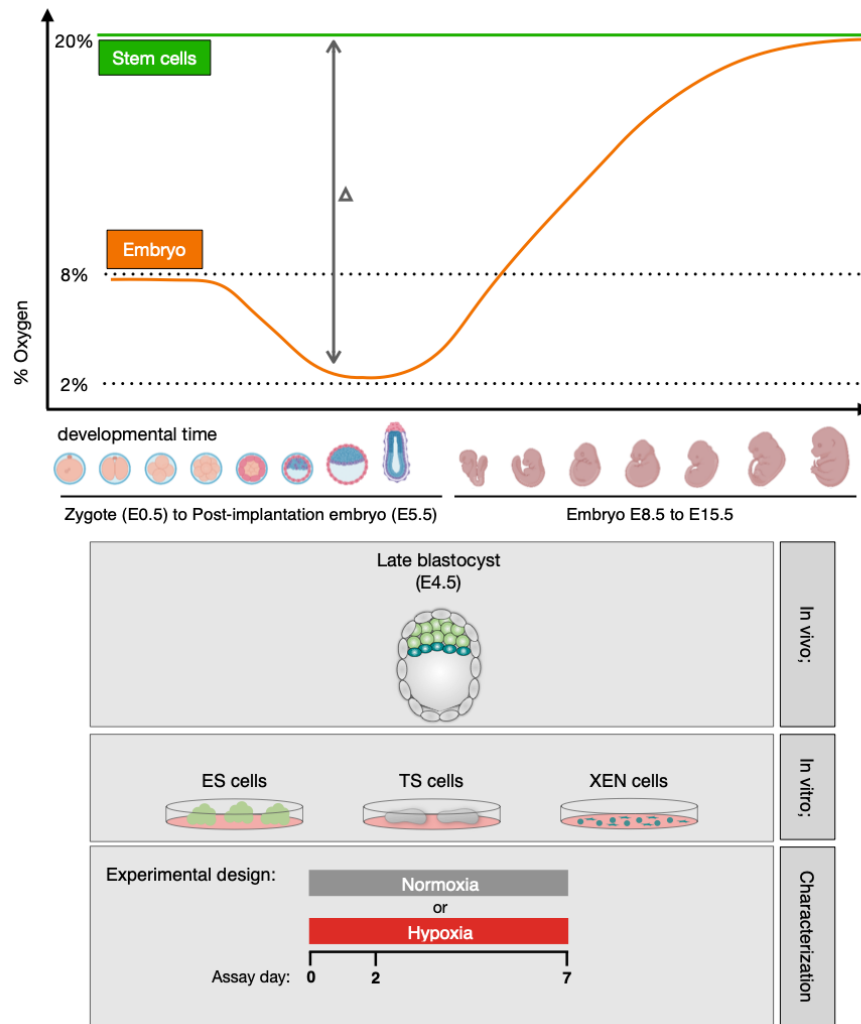
Figure 12. Studying hypoxia in embryonic and extraembryonic stem cells in vitro.

Comparison of the levels of oxygen used to culture stem cells in vitro and the oxygen level gradients observed during early embryogenesis (upper). Early stages of development from fertilization to term E18.5 (middle). Embryonic and extraembryonic stem cell lines –ES, TS and XEN cells—were used in this study to investigate the impact of acute (2 days) and prolonged (7 days) hypoxia (bottom).

3.1. Characterization of ES, TS and XEN cells in hypoxia at the transcriptomic level (bulk RNA-seq)

To better understand the impact of hypoxia on the transcriptional programs, we profiled gene expression in each stem cell type on day 2 and day 7 of hypoxia exposure by bulk RNA sequencing (RNA-seq). Principal component analysis (PCA) and hierarchical clustering based on the whole transcriptome profiles showed that the samples are clustered by cell type, indicating that hypoxia does not change cellular identities (**Figure 13A, B**). As shown in the heatmap (**Figure 13C**), differences are due to cell type rather than oxygen levels when it comes to global observations.
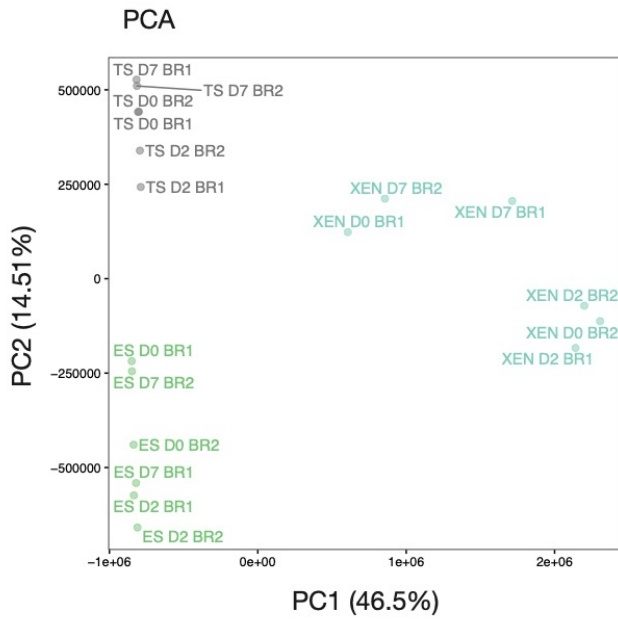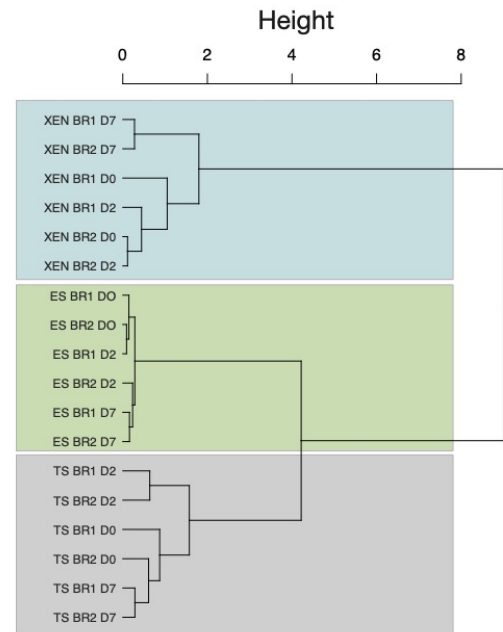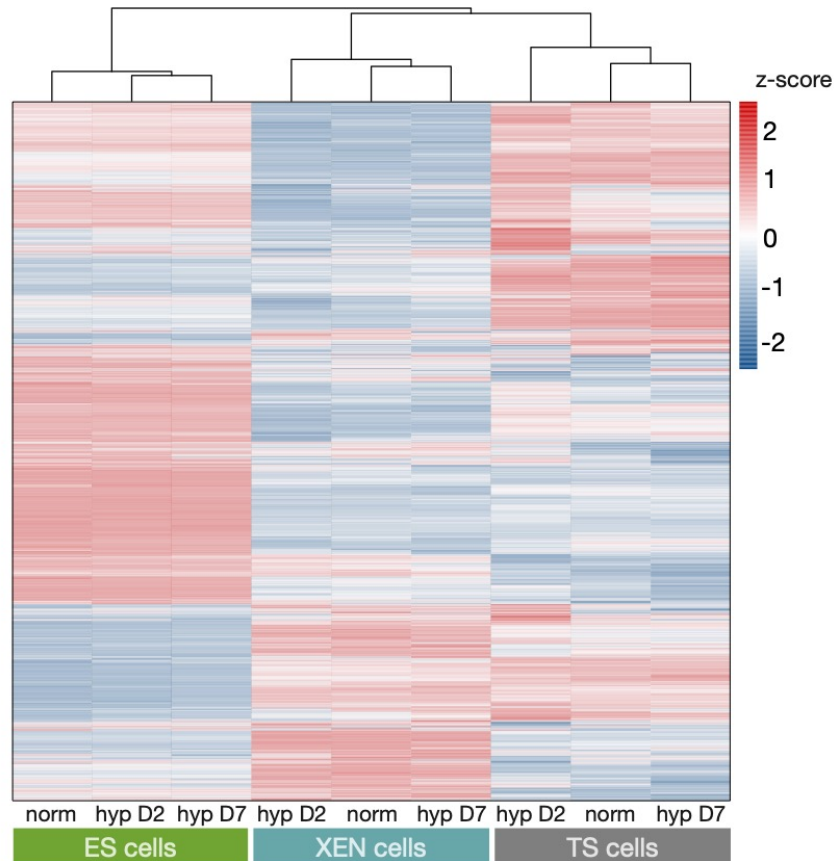
Figure 13. Global transcriptome comparison of ES, TS and XEN cells in hypoxia.

(A) Principal component analysis (PCA), (B) hierarchical clustering based on the global transcriptome and (C) heatmap showing expression levels of all identified transcripts of normoxic, hypoxic day 2 and day 7 ES cells.

A moderate number of genes was significantly differentially expressed (DE) in each cell type and time point when compared to normoxia (ES cells: 268 and 522 genes; TS cells: 677 and 606 genes; XEN cells: 366 and 659 genes on day 2 and day 7, respectively) (**Figure 14A**). DE genes were largely cell type-specific, with little overlap (**Figure 14B**) and largely time point independent with a small subset of genes being DE at day 2 and also being DE after 7 days of exposure to hypoxia (**Figure 14C**).
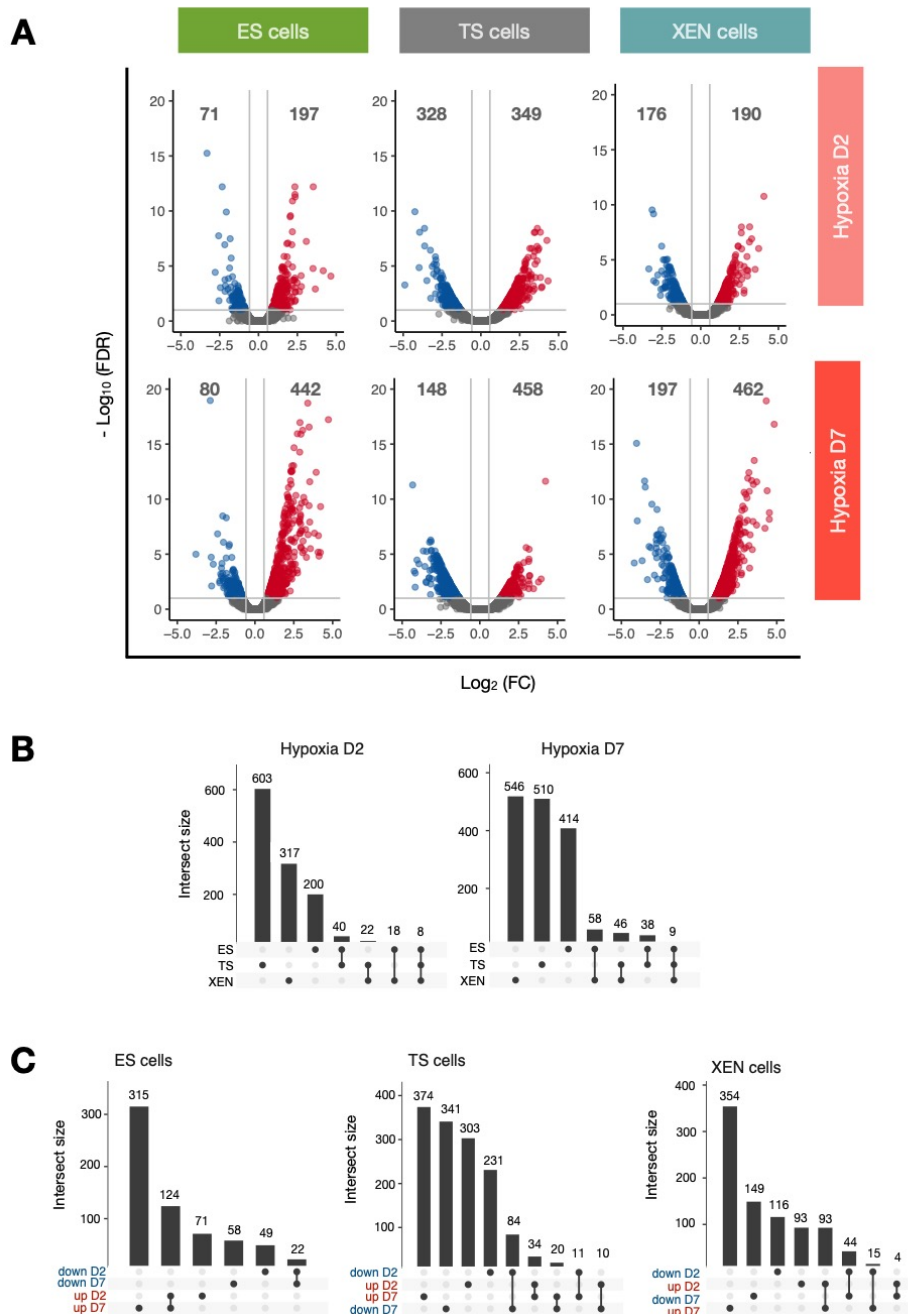
Figure 14. Differentially expressed genes of hypoxic vs. normoxic ES cells.

(A) Volcano plots showing the DE genes in hypoxia day 2 and day 7 relative to normoxia. The number of DE genes is indicated above each sub-group. (B) Overlap of DE genes across cell types on hypoxia day 2 and day 7 relative to normoxia. (C) Overlap of DE genes per cell type.

Gene ontology (GO) analysis in ES cells showed that the early response was mainly associated with upregulation of glycolysis and angiogenesis while the late response was associated with cell development and differentiation (**Figure 15A**).
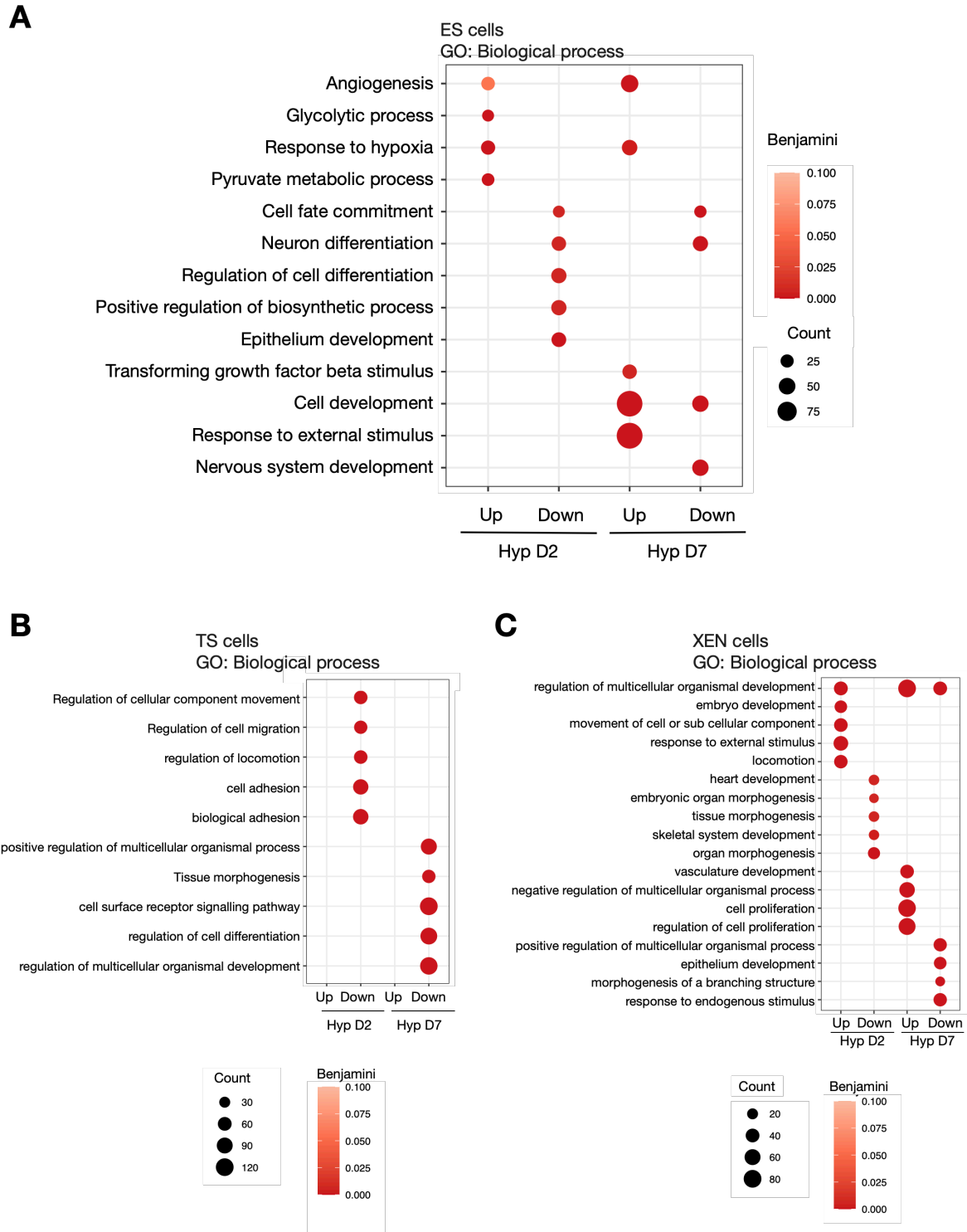
Figure 15. Biological process associated with DE genes upon hypoxic exposure.

BP-GO terms associated with DE genes in (A) ES, (B) TS and (C) XEN cells exposed to acute (day 2) and prolonged (day 7) hypoxia.

Interestingly, TS and XEN cells did not show changes at genes related to metabolism but did show altered differentiation-related transcriptional profiles where TS cells showed downregulated genes after 7 days and XEN cells already after 2 days and maintained at least for 7 days (**Figure 15B, C**). These results thus suggest that hypoxia leads to cell type-specific and temporal dynamics transcriptomic changes in terms of differentiation capacities in embryonic and extraembryonic stem cells.

3.2. Investigation of hypoxia-associated effects and their functional implications in ES cells

In general, ES cells seem to be the stem cell type that present most profound changes when exposed to hypoxia. In addition to expected changes, such as an increase of glycolysis (Muz et al., 2015), Strikingly, data shows a time-dependent upregulation of genes related to development and differentiation. Taking this into consideration, and the fact that ES cells resemble the EPI (able to give rise to any cell type in the adult organism), we decided to focus our study to investigate the hypoxia-associated effect on lineage-specification and its functional implications in ES cells.

3.2.1.   Transcriptional expression levels of lineage-specific marker genes in hypoxic ES cells

To better understand the transcriptional profile of hypoxic ES cells in terms of pluripotency and differentiation capacity, we checked the expression levels of lineage-specific marker genes in the early differentiating embryo (E6.5-E7.5) (Grosswendt et al., 2020a). Early mesoderm and endoderm marker genes, such as *T*, *Fgf8* and *Tmprss2*, were upregulated in hypoxia (2- to 4-fold compared to normoxia), whereas ectoderm and pluripotency genes showed either no change or a less dynamic response with little progression over time (**Figure 16**).
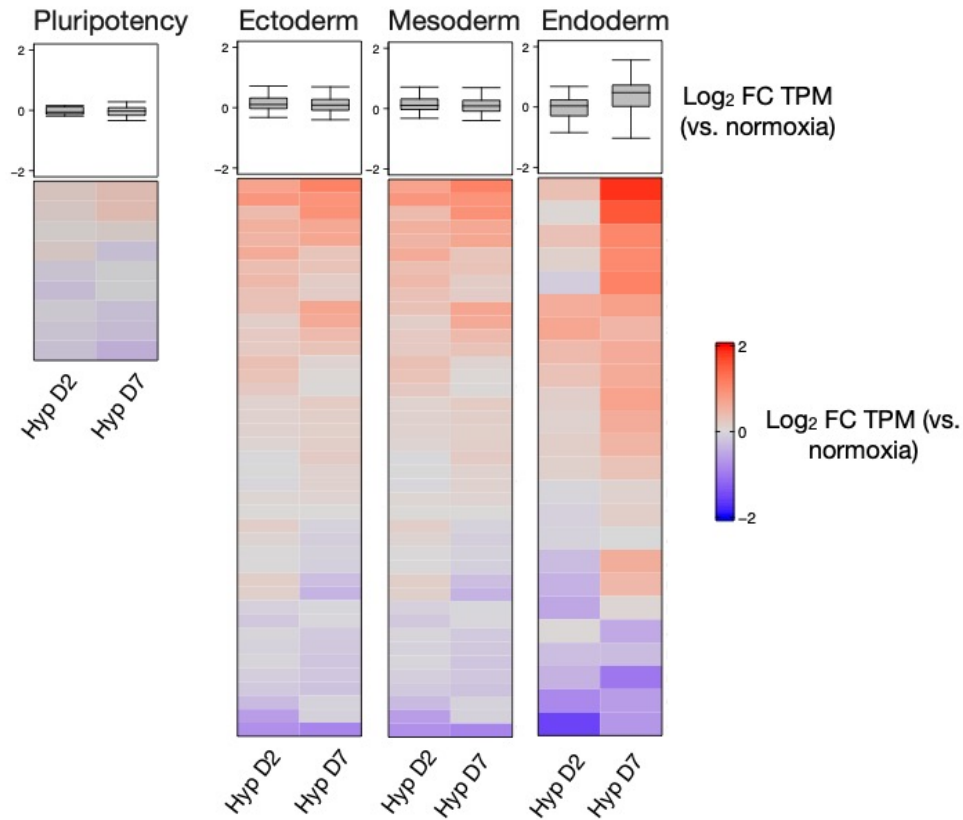
Figure 16. Hypoxia induces the expression of early 'mesendoderm' marker genes.

Heatmaps showing the expression levels of the indicated genes related to pluripotency, ectoderm, mesoderm and endoderm lineages in hypoxic relative to normoxic ES cells.

Genes such as *Nog*, *Krt19* and *Krt7* are associated with more advanced developmental stages (E7.5-E8.5) and are marker genes of the node, notochord and primitive gut. They were also selectively upregulated in hypoxic ES cells, especially after 7 days of hypoxia (**Figure 17**).
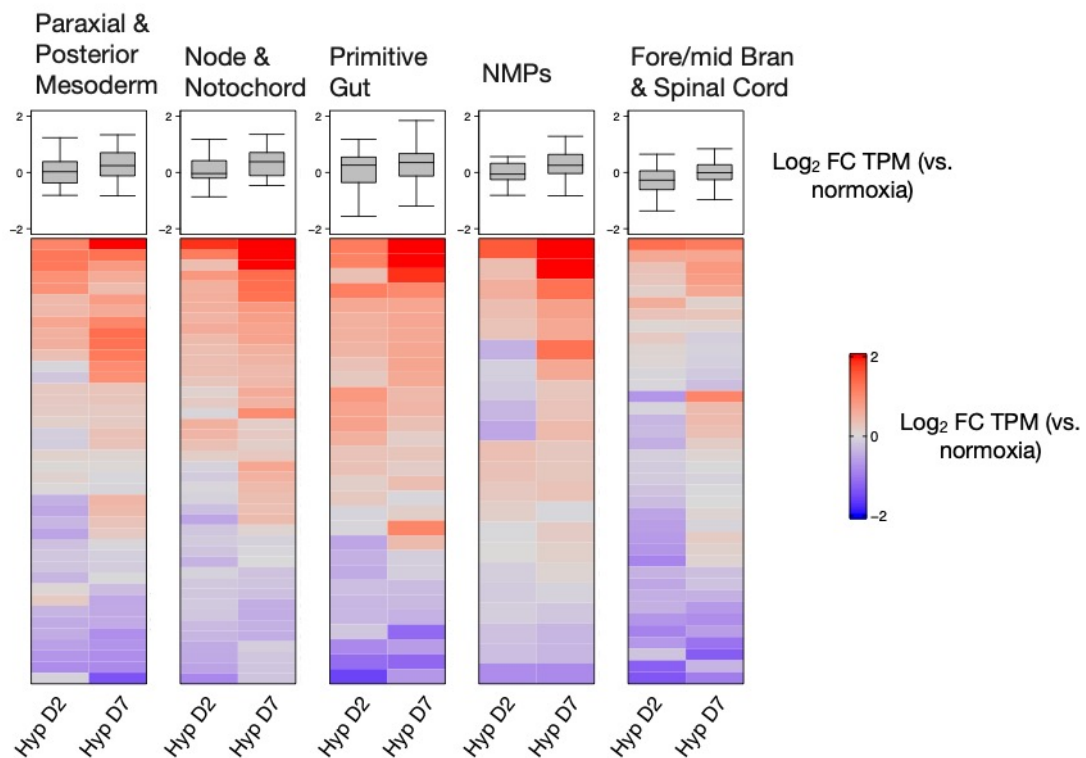
Figure 17. Hypoxia induces the expression of several cell-type specific marker genes.

Heatmaps showing the expression levels of the indicated genes related to several "advanced" (E7.0-E8.5) cell types in hypoxic relative to normoxic ES cells.

Thus, hypoxic ES cells seem to selectively induce key mesoderm and endoderm lineage regulators. Importantly, even if these genes are induced at the transcript level, ES cells are not spontaneously differentiating, yet the observed selective gene upregulation may affect developmental trajectories in the presence of the right differentiation cues.

3.2.2. WNT pathway-related genes are gradually upregulated during prolonged hypoxia

The induction of mesendodermal genes, such as *T*, *Eomes*, *Tbx6* and *Lef1*, are known to be activated by WNT pathway activity (Lindsley et al., 2006). Specifically, WNT3 mediates EMT in the posterior primitive streak during gastrulation via the upregulation of such genes (Arnold and Robertson, 2009; Bardot and Hadjantonakis, 2020; Tam and Behringer, 1997). We found that not only *Wnt3*, *T*, *Eomes*, *Tbx6* or *Lef1* are significantly upregulated under hypoxia conditions (**Figure 18**), canonical WNT pathway and EMT-related genes were strongly deregulated too (**Figure 19**).

ES cells (RNA-seq)
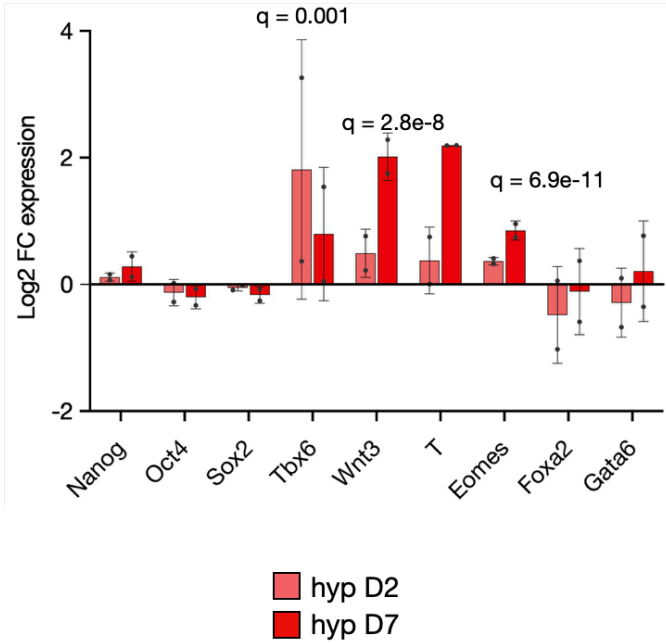
Figure 18. Hypoxia induces the expression of genes related to PS signature.

Bar plot showing the expression levels of the displayed genes in acute (hyp D2) and prolonged (hyp D7) hypoxia related to normoxia. Data represent log_2FC over normoxia of TMP expression values and standard deviation. Q values correspond to adjusted p-values.
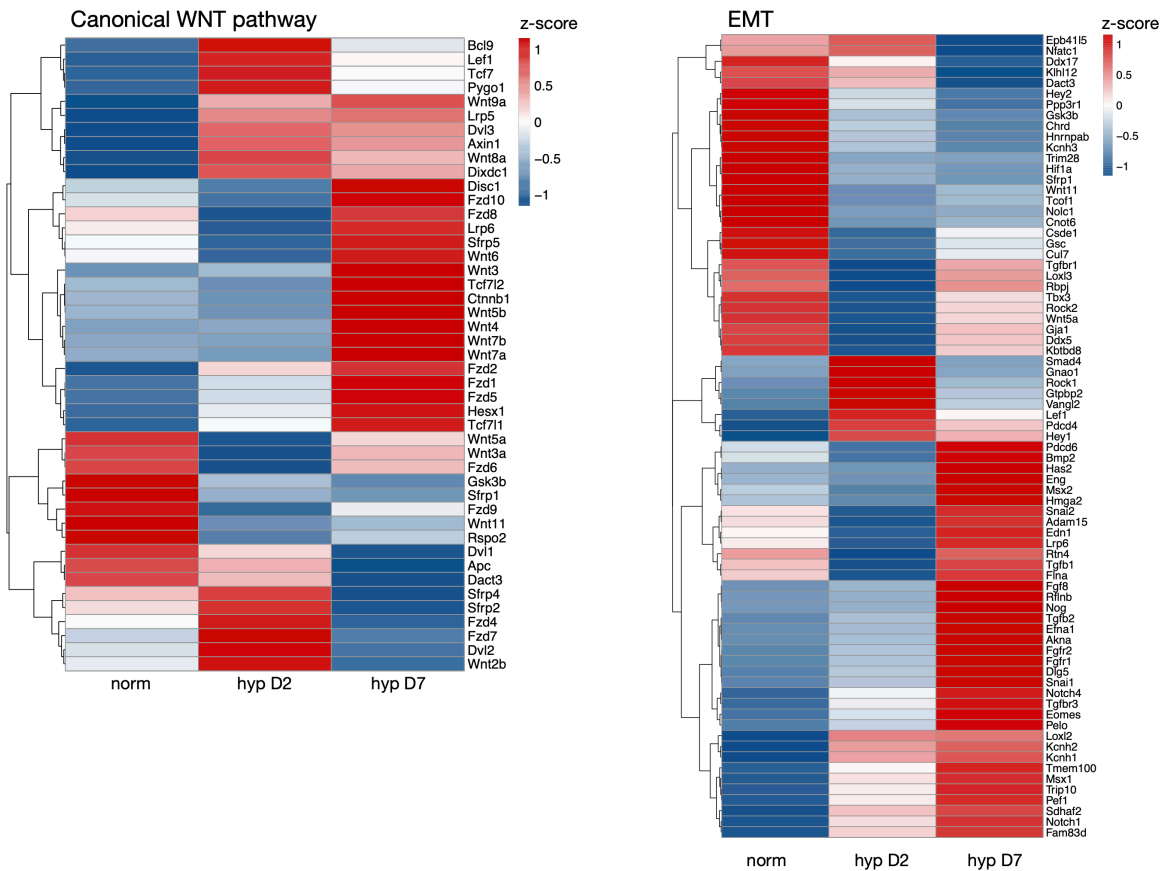
Figure 19. Hypoxia deregulates genes related to canonical WNT pathway and EMT.

Heatmaps showing expression levels of genes related to canonical WNT pathway and EMT in acute (hyp D2) and prolonged (hyp D7) hypoxia compared to normoxia.

Of note, data shows the upregulation of many WNT genes (*Wnt3*, *Wnt4*, *Wnt7a*, and *Wnt7b* significantly upregulated [fold change >1.5, FDR ≤ 0.1]), as well as other important WNT pathway-related genes such as *Ctnnb1* (ß-catenin), *Axin1* and *Tcf7*. Last, data also shows the induction of EMT-related genes, such as *Eomes*, *Snai1* and *Loxl2.*

3.2.3. Temporal dynamics of WNT pathway and its relation to oxygen levels

To better understand the relationship between WNT-related gene activation upon exposure to hypoxia, we cultured ES cells in normoxia and hypoxia collecting samples every two days (on days 1, 3, 5 and 7) as well as with variations of the oxygen levels (5%, 2% and 0.5%).

The data shows that 2% oxygen levels gradually increase the expression of *Wnt3* and *T* up to ~4-fold compared to normoxia by day 7. *Eomes* and *Tbx6* were on the other hand modestly upregulated up to ~1.5-fold compared to normoxia. In contrast, pluripotency and endoderm markers did not change (**Figure 20A**).

Strikingly, *Wnt3* and *Eomes* expression levels were inversely correlated with oxygen availability, associating their activation with an oxygen-mediated process (**Figure 20B**). Glycolytic and oxidative phosphorylation (OxPhos) genes were also checked: OxPhos genes showed no correlation with oxygen levels whereas glycolytic gene expression gradually increased with lower levels of oxygen (**Figure 20B**).
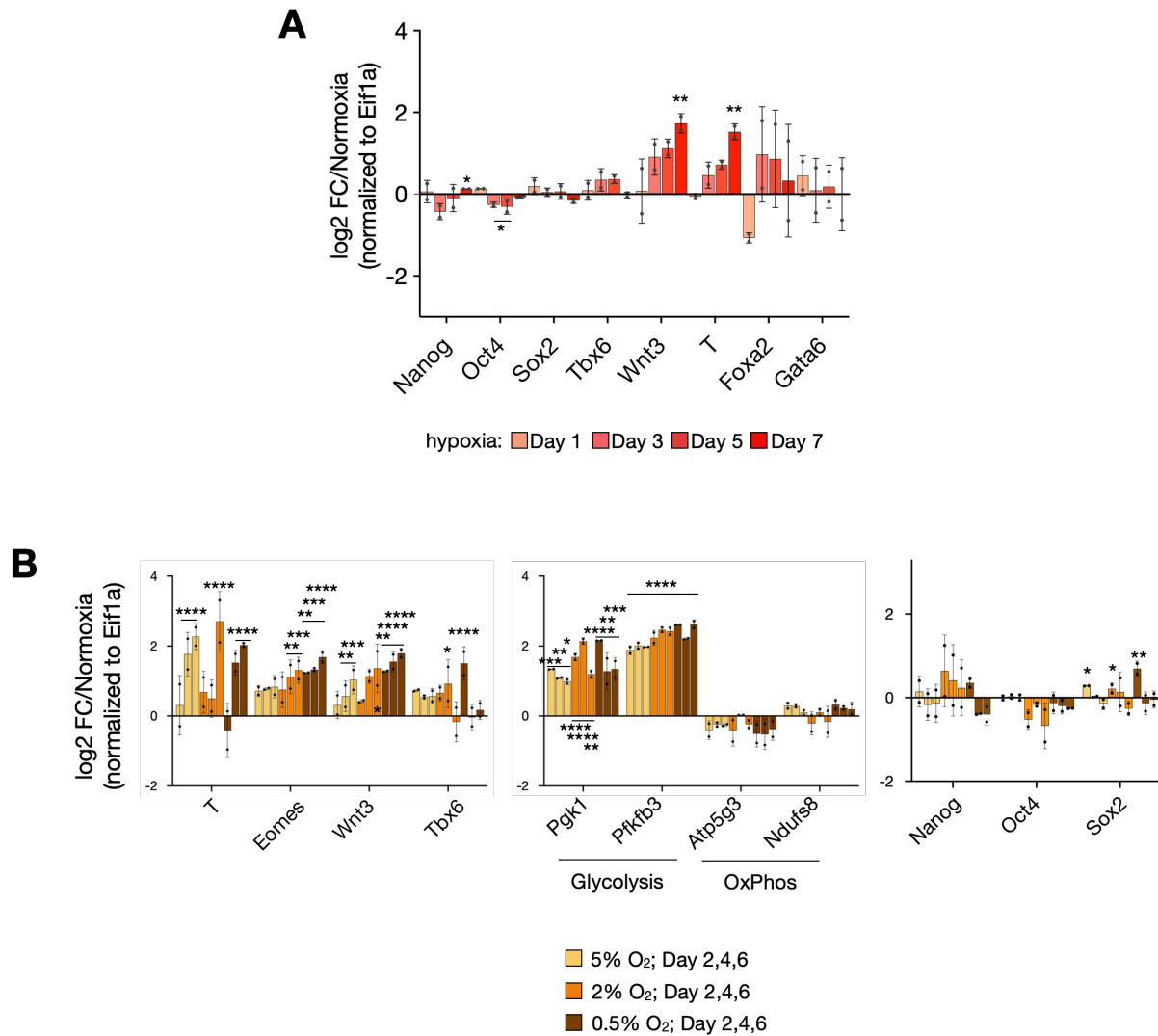
Figure 20. Transcriptional priming of ES cells is inversely correlated with oxygen levels.

(A) RT-qPCR analysis of the indicated genes in hypoxic ES cells (A) 2% oxygen over time from day 1 to day 7 and (B) measuring different levels of oxygen: 5,2 and 0.5%. The statistical test was two-way ANOVA. Error bars represent s.d. Two biological replicates are shown.

3.2.4. Prolonged hypoxia culture

The expression of *Wnt3* and *T* presented a gradual increase in normoxia compared to hypoxia. When culturing ES cells for a longer hypoxia exposure (up to 16 days), we observed that not only *Wnt3* and *T* but also *Eomes* further increased their expression levels at least until 16 days. Yet, pluripotency markers were not modified. Interestingly, glycolytic genes were expressed at similar levels during the whole period from day 2. When returning hypoxic cells (for 7 days) to normoxia (for another 7 days), the expression pattern of target genes was restored to those observed in normoxia (**Figure 21**). Hence, showing the reversibility of hypoxia, and WNT pathway-related activity.
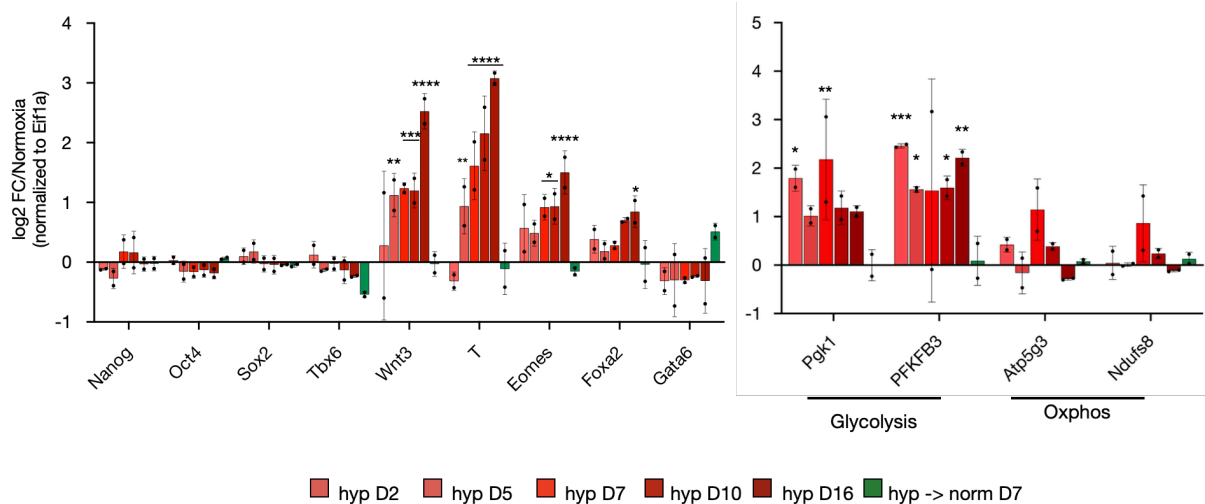
Figure 21. Hypoxia gradually upregulates WNT pathway related and developmental genes.

RT-qPCR analysis of the indicated genes in prolonged hypoxia (up to 16 days) and hypoxic cells returned to normoxia (hyp > norm D7) related to normoxia. Statistical test is two-way ANOVA. Error bars represent s.d. Two biological replicates are shown.

3.2.5. Proteome profile of hypoxic ES cells

To investigate if the transcriptomic changes were also reflected at the protein level, we performed label-free mass spectrometry in normoxic and hypoxic ES cells on day 2 and day 7. A total number of 4,260 proteins were identified. Among all protein-coding genes identified at the transcript level, only ~22% were also identified at the protein level, which is a common limitation of mass spectrometry (**Figure 22A**). In general, acute and prolonged hypoxic ES cells retain a similar proteome compared to normoxic ES cells with few DE proteins (92 and 106 DE proteins [FC>1.5, q-value < 0.05] in hypoxia day 2 and day 7, respectively) (**Figure 22B**). DE proteins on day 7 were mainly enriched for metabolic functions. In contrast, cell signaling, development and differentiation functions were only enriched at the transcript level (**Figure 22C**). This may be associated with different reasons: (1) Limited sensitivity of mass spectrometry to detect lowly expressed proteins, such as transcription factors, (2) or alternatively, due to selective post-transcriptional or translation degradation, that in fact, has been associated to oxygen depletion in order to reduce energy-costing processes (Koumenis et al., 2002; Lee et al., 2020; Pettersen et al., 1986).
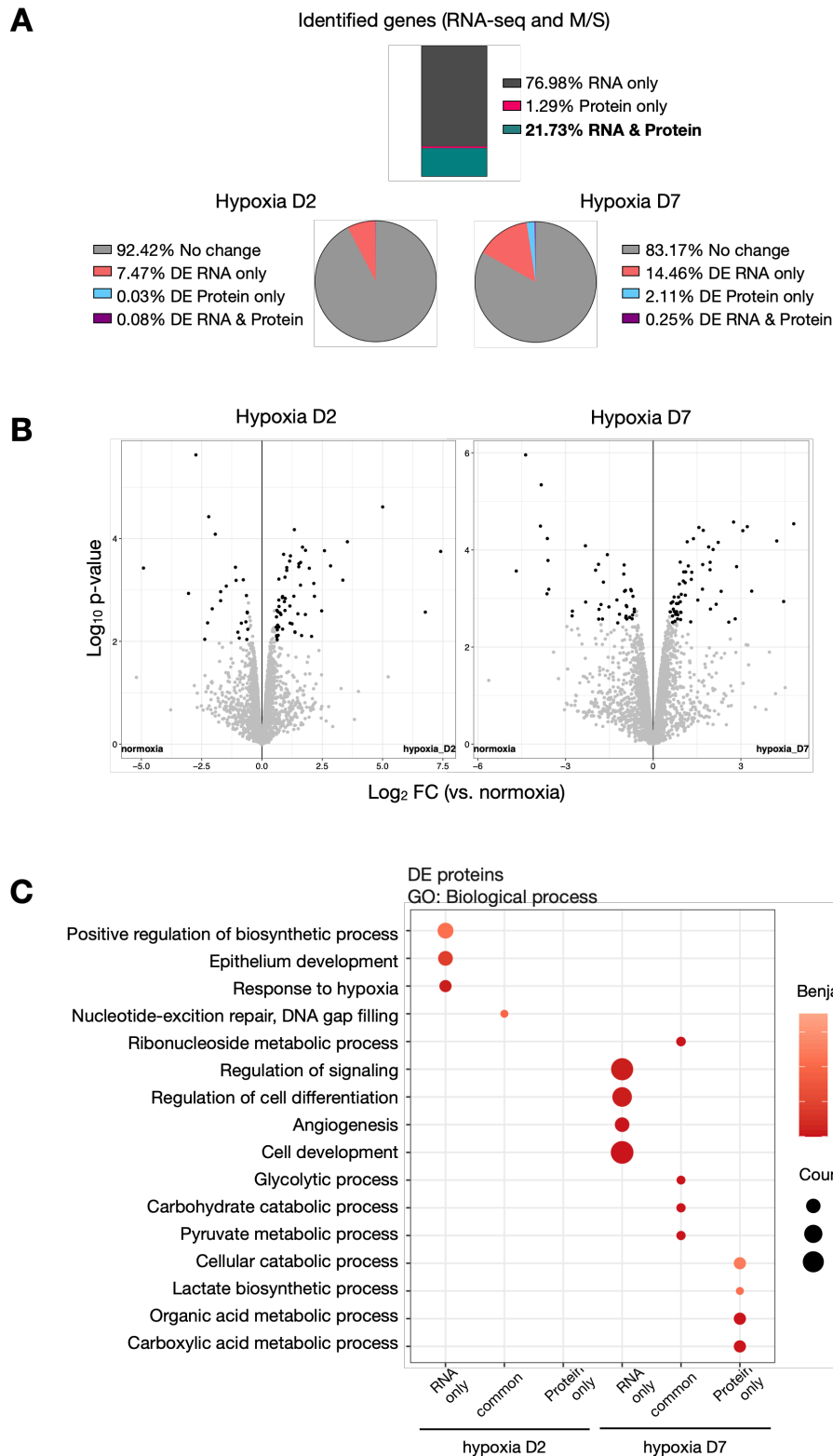
Figure 22. Hypoxia-related effects at the proteomic level reflect the enrichment of metabolic functions.

(A) Bar plot showing the percentage of genes identified at the transcript and protein level (*top*) and pie charts showing the DE genes at hypoxia day 2 and day 7 compared to normoxia at the transcript and protein level (*bottom*). (B) Volcano plots showing the DE proteins in hypoxia day 2 and hypoxia day 7 vs normoxia. (C) BP-GO terms associated with the DE genes at RNA-, protein-only and RNA & protein levels.

To address it, we developed Western blotting on whole cell extracts and detected a low-level but gradually increasing expression of T and EOMES in hypoxic ES cells. For WNT3 though, we were unable to detect its expression (**Figure 23A**).

3.2.6. T target genes are not differentially expressed

We reasoned that, if T is functional at the protein level in hypoxic ES cells, then the expression of T target genes would change accordingly. To study the expression of T target genes, we used publicly available data corresponding to in vitro primitive streak differentiation (Lolas et al., 2014a). This publication contains (1) T ChIP-seq data that we used to identify T target genes and (2) bulk RNA-seq data that we used to find the expression of direct T target genes and to classify them into activated or repressed genes. Then, we checked the expression of T target genes in hypoxic ES cells (**Figure 23B**). A very small subset (~4%) of T target genes were found to be DE genes in hypoxic ES cells. Additionally, some of them were already upregulated on day 2 (where we found little upregulation of *T*, ~1.2-fold).
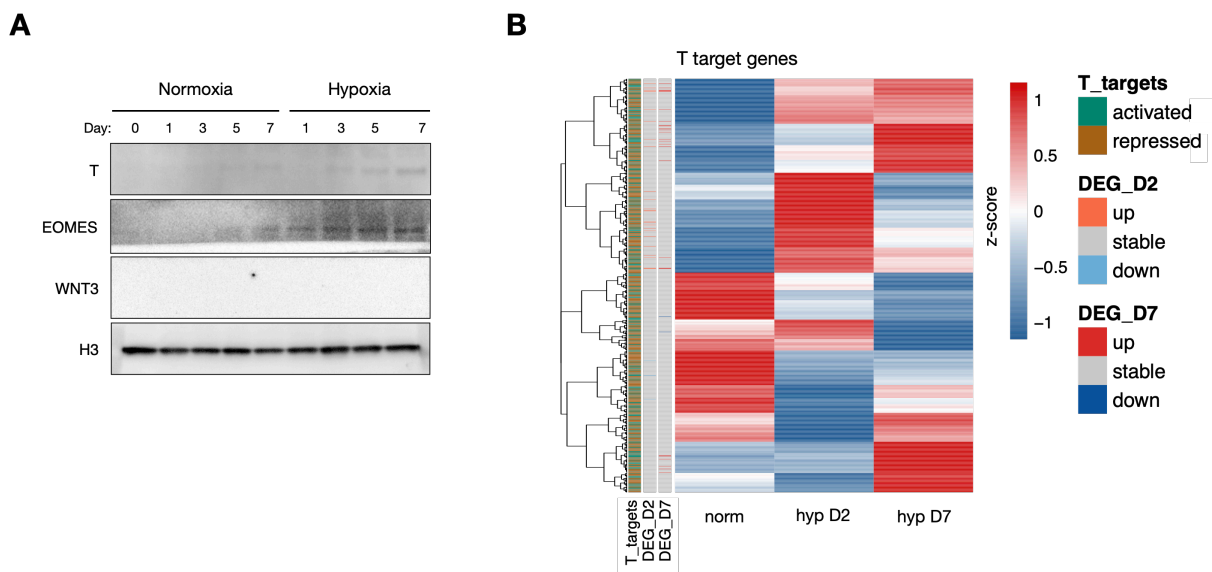
Figure 23. Hypoxia-mediated T expression is not functional at the protein level.

(A) Western blot showing expression levels of the indicated genes in normoxic and hypoxic ES cells. H3 was used as the loading control. (B) Heatmap showing expression levels of T target genes in normoxic and hypoxic ES cells. Bar annotations show: (1) T_targets: T-activated (green) and T-repressed (brown) target genes in in vitro primitive streak differentiated cells, (2) DE genes in hypoxia day 2 (D2) and day 7 (D7): up- (red), down- (blue) regulated and stable (gray) genes on hypoxia day 2 and/or day 7.

Yet, it is worth mentioning that DE proteins in hypoxia contained epigenetic regulators (**Figure 24**) such as UHRF1, involved in the maintenance of DNA methylation (Bostick et al., 2007) and histone demethylases which are known to sense cellular oxygen levels, such as KDM4C.

To a lesser extent, some proteins belonging to the polycomb repressive complex 2 (PRC2) such as EZH2 also presented an altered expression pattern at the protein level. PRC2 has been associated as a strong regulator during cell fate decisions and differentiation of ES cells via chromatin rewiring. Therefore, they are very likely potential regulators involved in the hypoxic response and were studied further. Further investigation and results can be found in the second part of the result section.
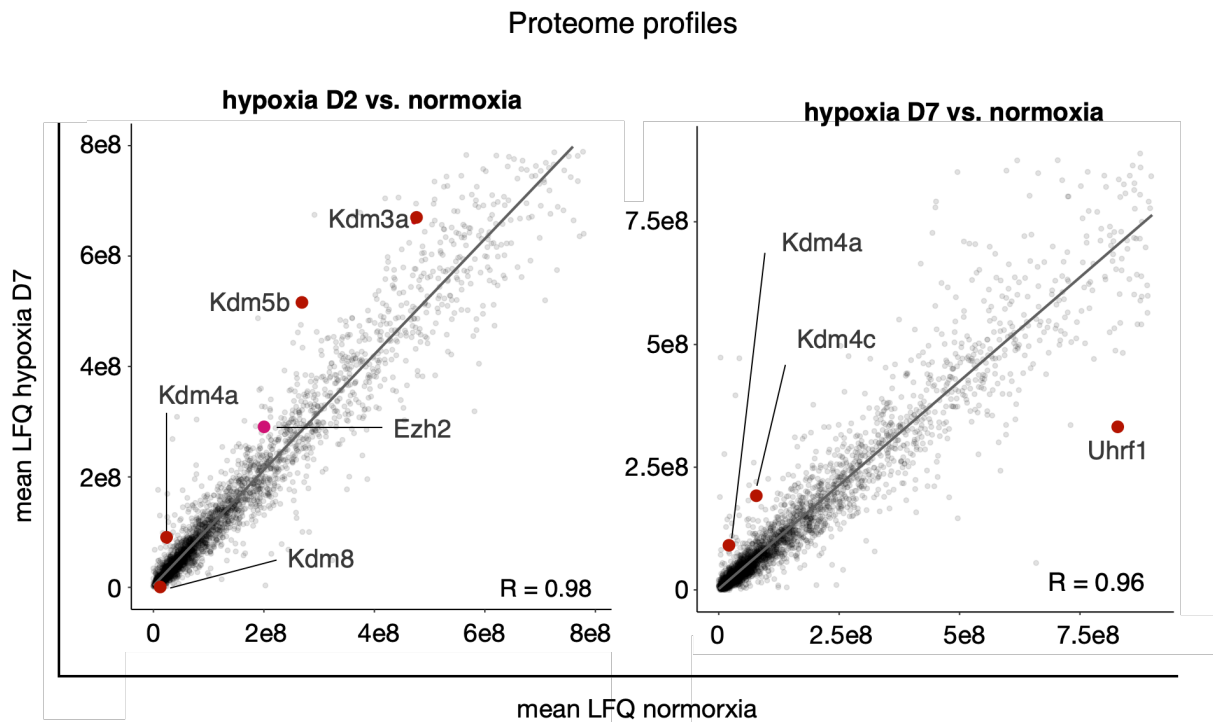
Figure 24. Hypoxia induces the deregulation of epigenetic regulators at the protein level.

Scatter plot highlighting the deregulation of the indicated genes. In dark red genes are DE and in pink, genes that are not DE at the protein level.

Taking this data together, hypoxia mediates a selective induction of WNT pathway and EMT-related genes, and their targets which have a role in determining an early transcriptional primitive streak signature in ES cells, that is not directly reflected at the protein level. Additionally, ES cells do not resemble epiblast stem cells derived from the post-implantation embryo, which have a T-, SOX2- and FOXA2- triple-positive cell profile and anterior primitive streak characteristics (Kojima et al., 2014; Tsakiridis et al., 2014). From a mechanistic point of view, epigenetic regulators involved in DNA methylation and chromatin rewiring might be involved to shape the phenotypical observation found in ES cells in response to hypoxia.

SECOND PART: Mechanisms involved in the hypoxia response in ES cells

3.3. Chemical activation of HIF1α in normoxia

HIF1α is the main transcription factor involved in the hypoxia response (Semenza, 1999). To test whether it is directly involved in the activation of primitive streak related genes, IOX2 was used to chemically activate HIF1α in normoxia. IOX2 inhibits the HIF1α destabilizer prolyl hydroxylase 2, leading to HIF1α stabilization, translocation to the nucleus and activation of its target genes. Therefore, we treated ES cells in normoxia with IOX2 in two different concentrations (50 µM and 100µM). As control, we saw dose-dependent upregulation of glycolytic genes. EMT-related genes did get upregulated upon IOX2 treatment in a dose-dependent manner (**Figure 25**). While the levels of glycolysis genes was did not increase beyond day 2, *Wnt3*, *Eomes* and *T* levels were significantly higher on day 7 (**Figure 25**). Pluripotency genes remained largely unchanged. Hence, chemical activation of HIF1α mimics WNT pathway induction observed in hypoxia.
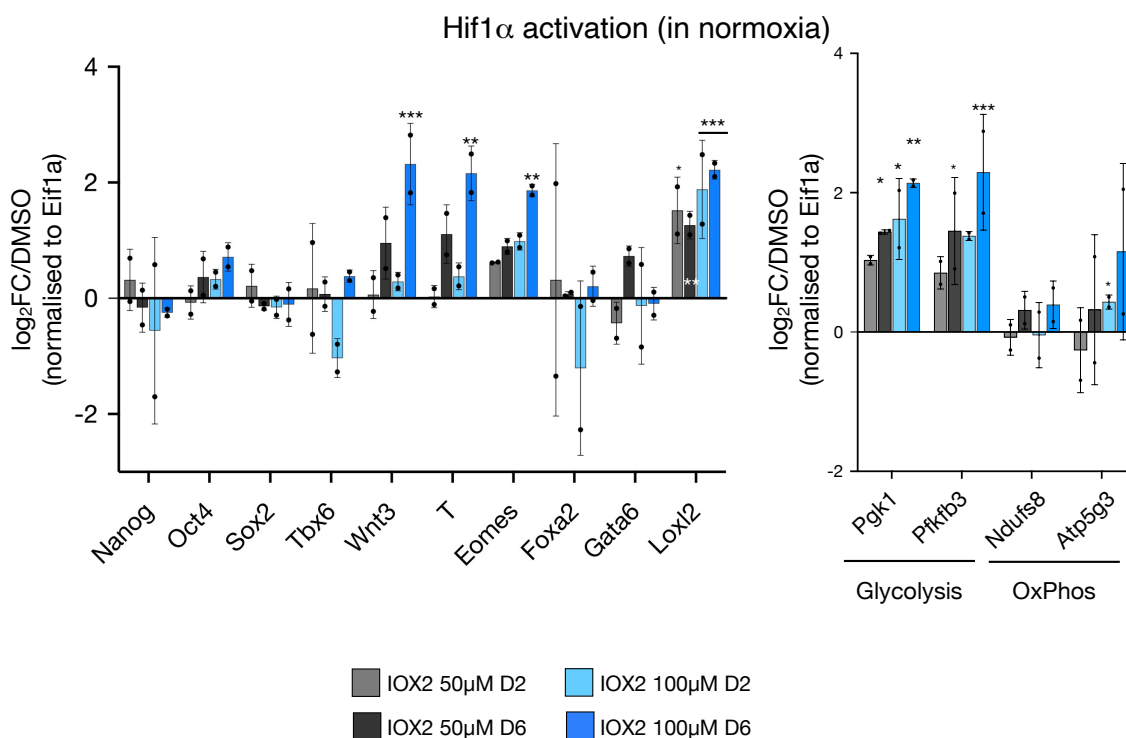
Figure 25. HIF1α chemical activation in normoxia mimics induction of developmental related genes as in hypoxia.

RT-qPCR analysis of the indicated genes in normoxic ES cells treated with IOX2 at the indicated concentrations for 2 and 6 days. Data represents Log2FC over DMSO-treated ES cells.

3.4. Genome-wide profiling of HIF1α genomic occupancy

The previous results pointed out to the direct role of HIF1α in activating the expression of the developmental genes induced under hypoxic conditions. To investigate whether developmental genes are directly induced by HIF1α binding at promoters, we profiled HIF1α genome-wide occupancy by ChIP-seq. To do so, we used not only ES cells cultured in hypoxia (day 2 and day 6) but also in normoxia. This helped us identify the potential and unknown binding profile of HIF1α also in its presumably inactive form.

As control regions, we checked promoters of *Pgk1* and *Ldha* glycolytic genes which are known to be targeted by HIF1α (Iyer et al., 1998). Indeed, promoters of these genes were bound by HIF1α in hypoxia, at both day 2 and day 6 (**Figure 26**).
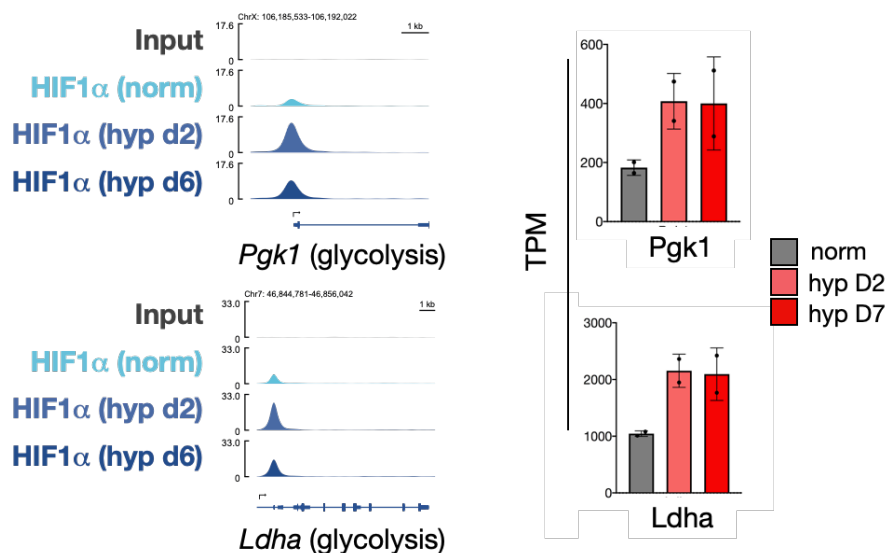
Figure 26. HIF1α binds to promoter regions of metabolism-related genes upon hypoxia.

Genome browser views of HIF1α-bound sites of the given genes (*left*). Expression values of the indicated genes in normoxic and hypoxic ES cells were measured by RNA-seq (*right*).

At promoter regions, we could identify many genes that were already bound by HIF1α in normoxia and the majority of those were also bound in hypoxia day 2 and 6 (137 genes) (**Figure 27A, B**). Yet, the majority of the target genes were found both on day 2 and day 6 or only on day 2 (297 and 214 genes, respectively) (**Figure 27B**). On the other hand, few genes were bound by HIF1α only at day 6 (**Figure 27B**), which signal seems very similar at day 2, very likely being a weak binding (**Figure 27A**). In total, 692 HIF1α target genes were identified. In general, the data shows that HIF1α regulates mainly metabolic genes by direct binding at promoters (**Figure 27C**).
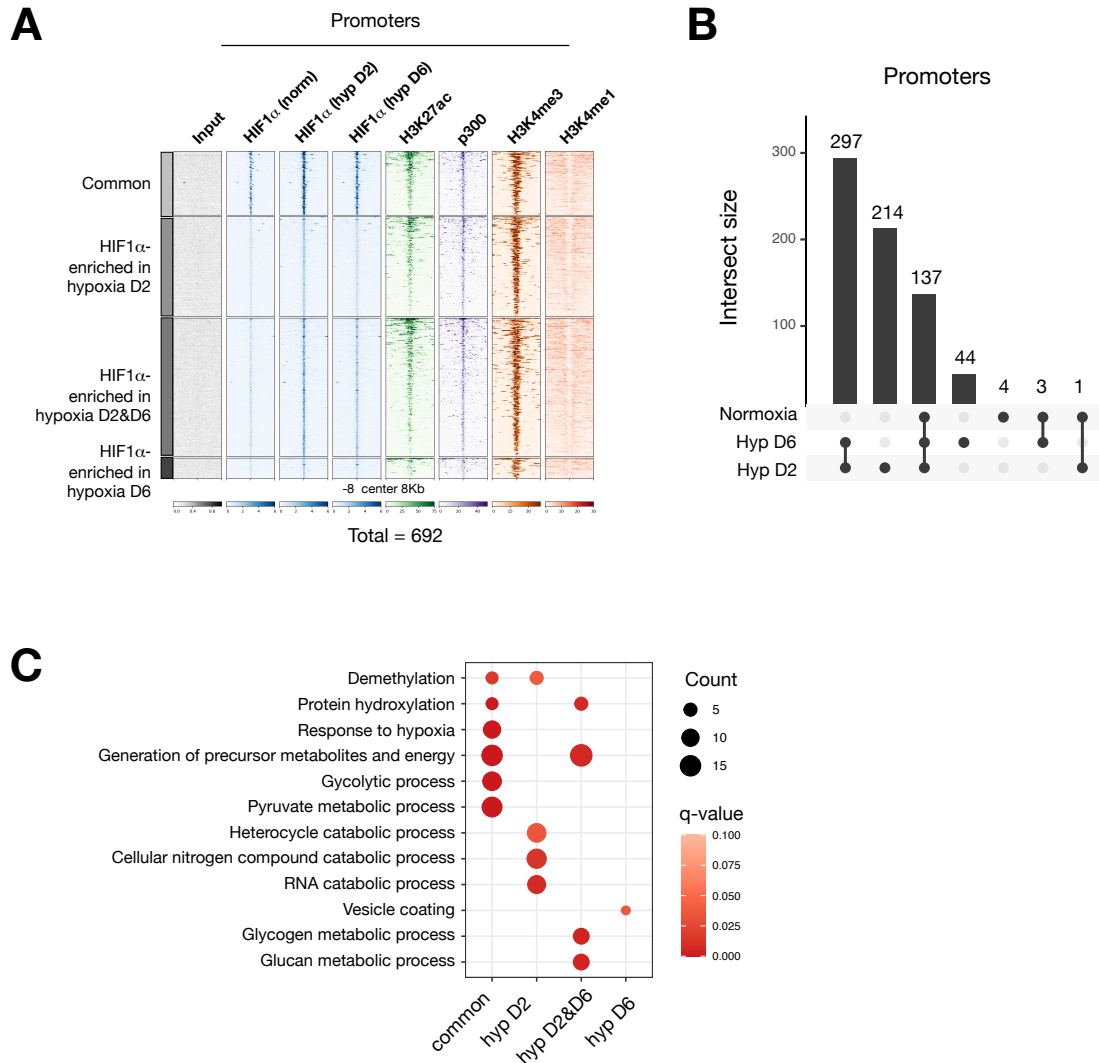
Figure 27. HIF1α binds to protomer regions of mainly metabolic genes.

(A) Density plots showing enrichment of HIF1α and the indicated histone modifications at HIF1α-bound promoters. ±8 kb surrounding the peak center is shown. Total depicts the total number of HIF1α-bound sites. (B) Overlap of HIF1α-bound promoters in normoxic and hypoxic ES cells. (C) BP-GO terms associated with HIF1α target genes with HIF1α binding at promoters.

Unexpectedly, HIF1α did not bind directly at promoters of developmental genes or WNT pathway effectors (**Figure 28**). Only at *Ctnnb1* (ß-catenin), we identified a peak at the promoter region (see H3K4me3 signal colocalizing with HIF1α peak). Nonetheless, *Ctnnb1* expression at the transcript level did not seem to increase in hypoxia compared to normoxia (**Figure 28**).
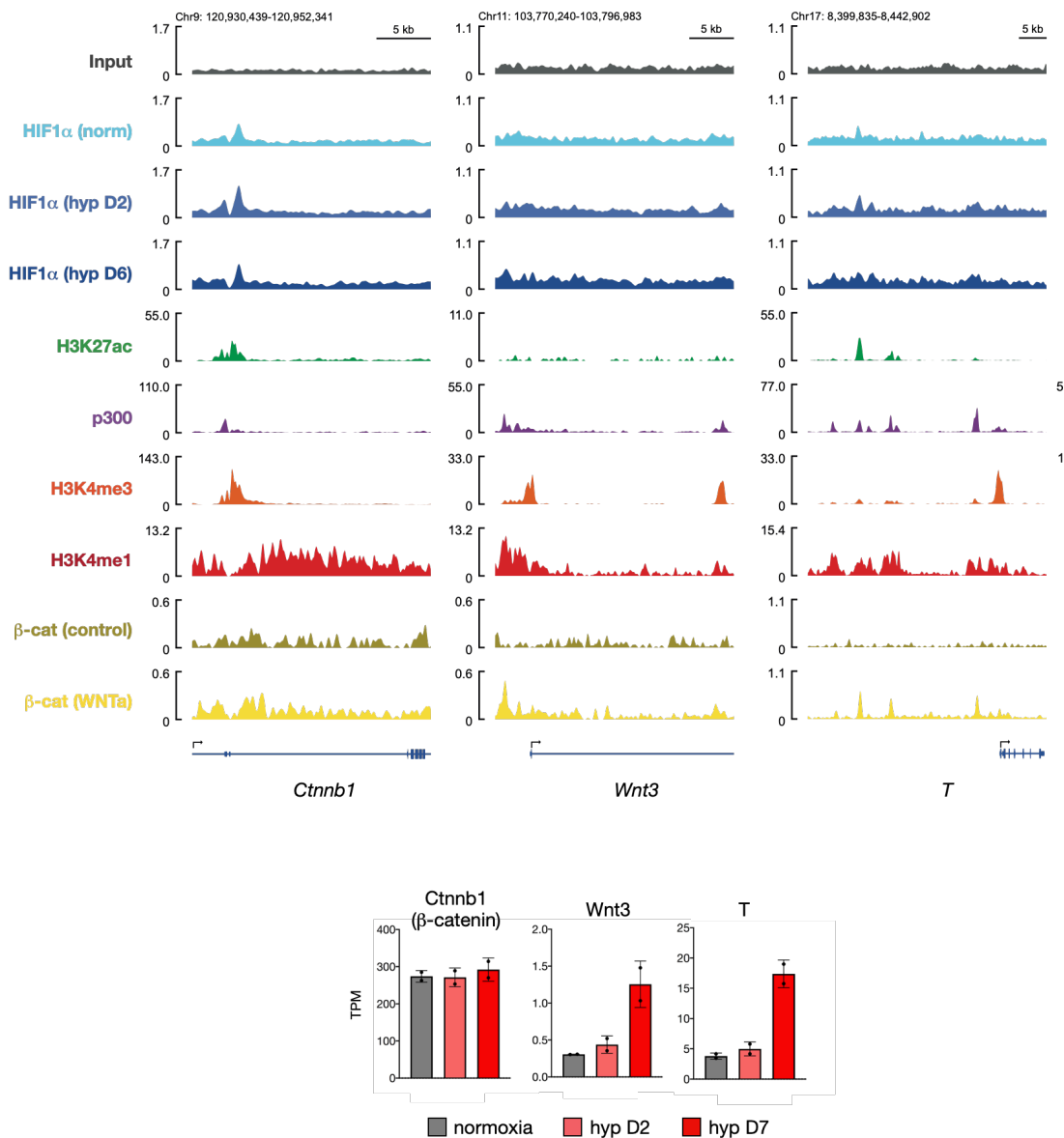
Figure 28. HIF1α does not bind directly to the promoter region of *Wnt3* or *T*.

Genome browser views of chromatin occupancy at canonical WNT pathway genes (*upper*) and bar plot showing transcript expression of the indicated genes in normoxic and hypoxic ES cells (*bottom*). Error bars represent s.d.

These results suggest that the transcriptional primitive streak profile in hypoxic ES cells is mediated by HIF1α but not by direct binding at promoters.

In addition to promoters, HIF1α also binds to distal regions. Similar to promoters, we found many regions where HIF1α bound already in normoxia (59 distal regions). The majority of the distal peaks were nevertheless found at hypoxia day 2 and day 6 (101 distal regions) (**Figure 29A, B**).
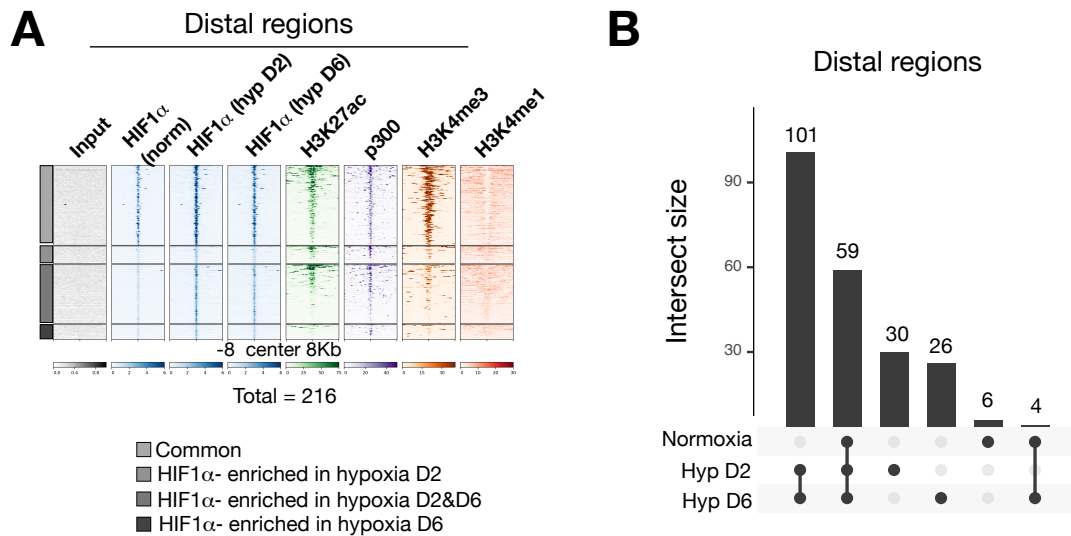
Figure 29. HIF1α binds to distal regions across the genome.

(A) Density plots showing enrichment of HIF1α and the indicated histone modifications at HIF1α-bound distal regions. ±8 kb surrounding the peak center is shown. Total depicts the total number of HIF1α-bound sites. (B) Overlap of HIF1α-bound distal regions in normoxic and hypoxic ES cells.

To identify active enhancer regions in ES cells –defined by enrichment of H3K27ac, H3K4me1 and p300 and lack of H3K4me3— a publicly available dataset was used (Cruz-Molina et al., 2017a). We were able to identify HIF1α binding to 70 regions that have an active enhancer signature in ES cells (**Figure 30A**). Interestingly, HIF1α bound to enhancer regions in close proximity to several developmental and WNT effector transcription factors such as *Lef1*, which is also upregulated in hypoxia day 2 (**Figure 30B**). Interestingly, this distal region is bound by ß-CATENIN in WNT-stimulated ES cells but not in normal ES cell culture conditions (**Figure 30B**; compare ß-CATENIN (WNTa) versus ß-CATENIN (control); data from (Zhang et al., 2013a)). Indeed, we identified HIF1α / ß-CATENIN colocalization at 27 active enhancer regions and 52 promoter regions (**Figure 30A**), suggesting a potential colocalization role to selectively induce developmental genes in hypoxia.
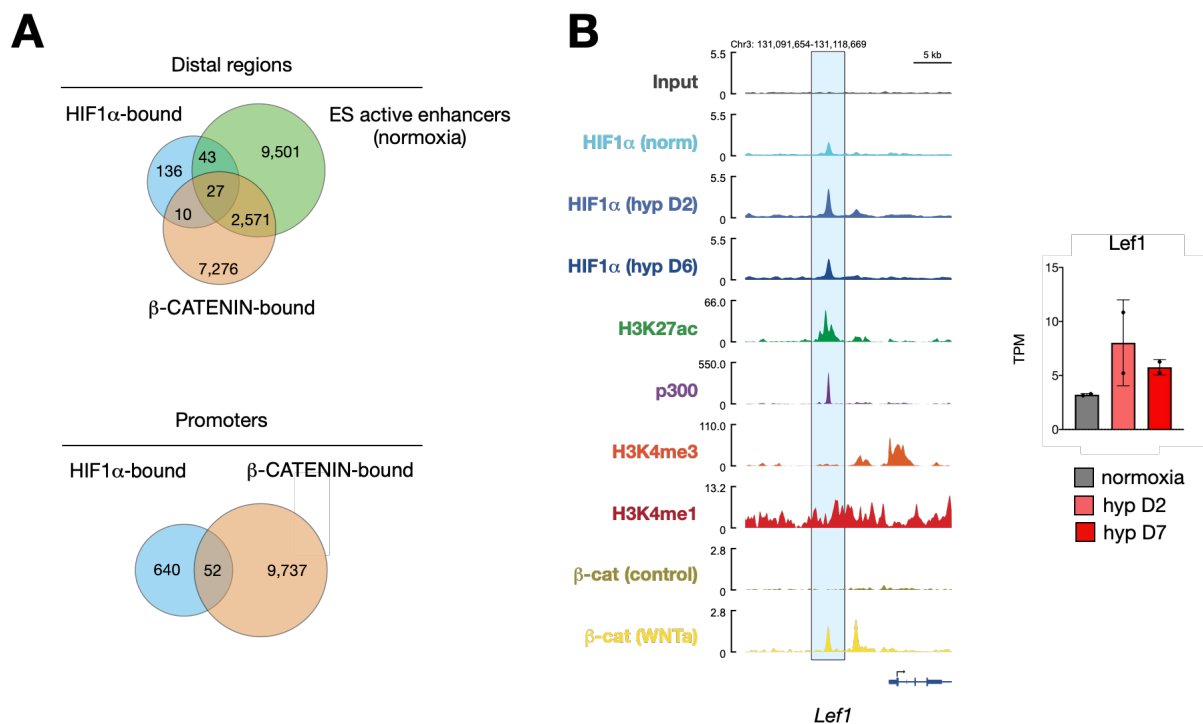
Figure 30. HIF1α and ß-CATENIN colocalization across the genome.

(A) Venn diagrams showing the number of overlapping HIF1α, ß-CATENIN and ES active enhancers at distal regions (upper) and HIF1α and ß-CATENIN at promoters (bottom). (B) Genome browser views of chromatin occupancy at *Lef1* (*left*) and bar plot showing its transcript expression levels in normoxic and hypoxic ES cells (*right*). Error bars represent s.d.

Interestingly, HIF1α also bound several chromatin regulations, presumably also in normoxia, but with a clearer signal in hypoxia. Examples of those, are PRC2 components such as *Eed* and *Ezh2*, histone deacetylase *Kdm4c* and DNA methylation-related gene *Tet2* (**Figure 31A**). It is intriguing to see that some of these epigenetic regulators, also presented an altered expression profile at the protein level (**Figure 24**), even though we did not see deregulation at the transcript level (**Figure 31B**). Nonetheless, they are likely involved in regulating the epigenetic landscape in hypoxic conditions, by direct binding of HIF1α either at promoter or/and enhancer regions.
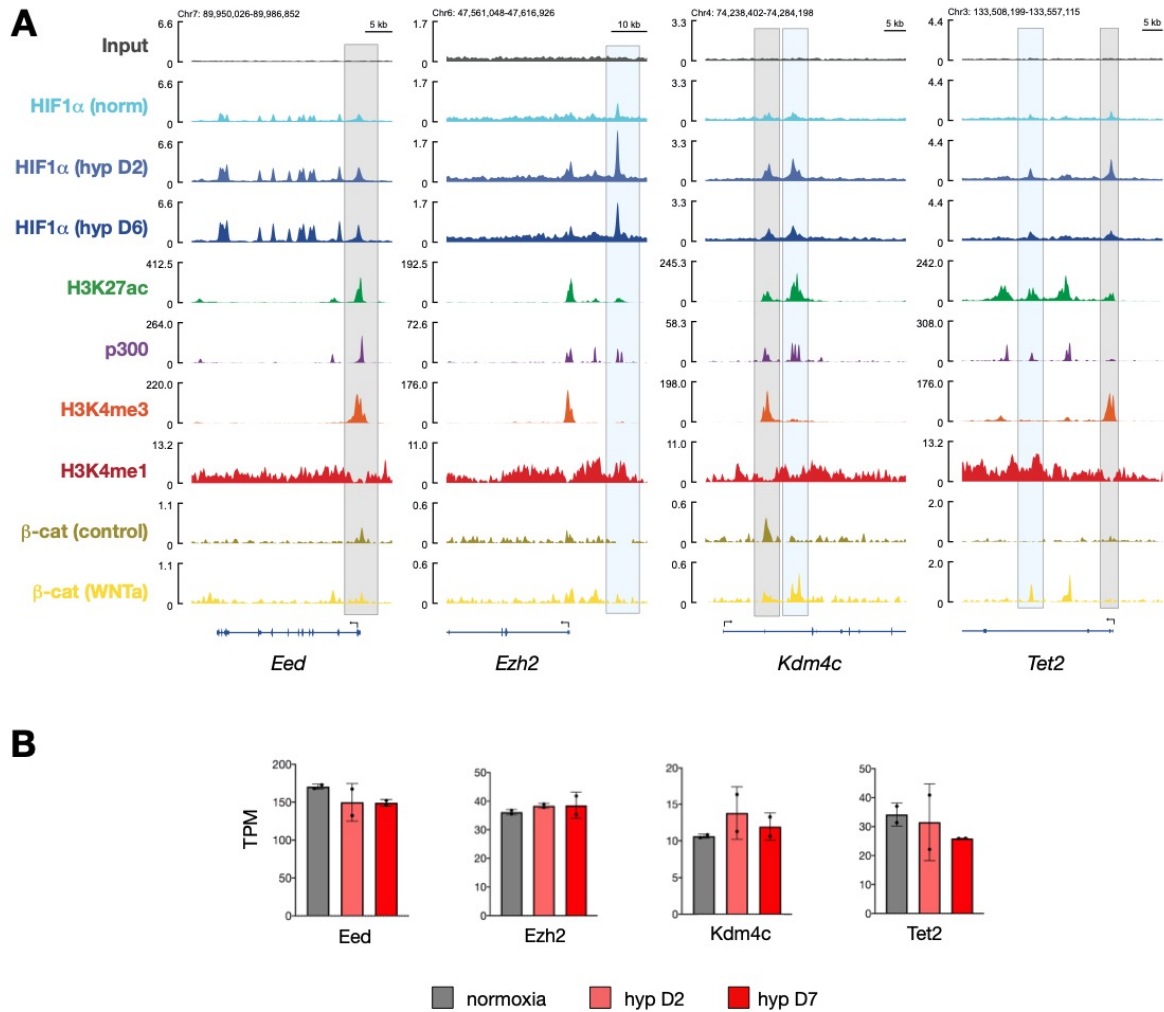
Figure 31. HIF1α binds to promoters of chromatin regulators and to distal regions in close proximity to those.

(A) Genome browser views of chromatin occupancy at indicated genes. (B) Bar plots showing transcript expression levels of the indicated genes in normoxic and hypoxic ES cells. Error bars represent s.d.

3.5. DNA methylation profile of hypoxic ES cells.

Using mass spectrometry, we identified UHRF1 as one of the most significantly downregulated proteins in hypoxic ES cells (**Figure 24**). UHRF1 partners with the DNA methyltransferase DNMT1 for the maintenance of DNA methylation during cell division. Even though CG-rich gene promoters, including many developmental genes, are already depleted of DNA methylation, gene expression may be controlled by the methylation of other regulatory elements such as enhancers (Feldmann et al., 2013; Stadler et al., 2011). Therefore, we reasoned that DNA methylation could be a major epigenetic regulator involved in the hypoxia response.

First, we were able to confirm the downregulation of UHRF1 over time in hypoxic ES cells. This downregulation was already notable from day 1 in hypoxia (**Figure 32A**). Additionally, chemical activation of HIF1α also led to downregulation of UHRF1, but to a lesser extend if compared to the hypoxia-mediated downregulation (**Figure 32B**). This data together shows that HIF1α might be implicated in the downregulation of UHRF1, but its downregulation does not entirely depend on HIF1α activation and other hypoxic cues might be implicated too.
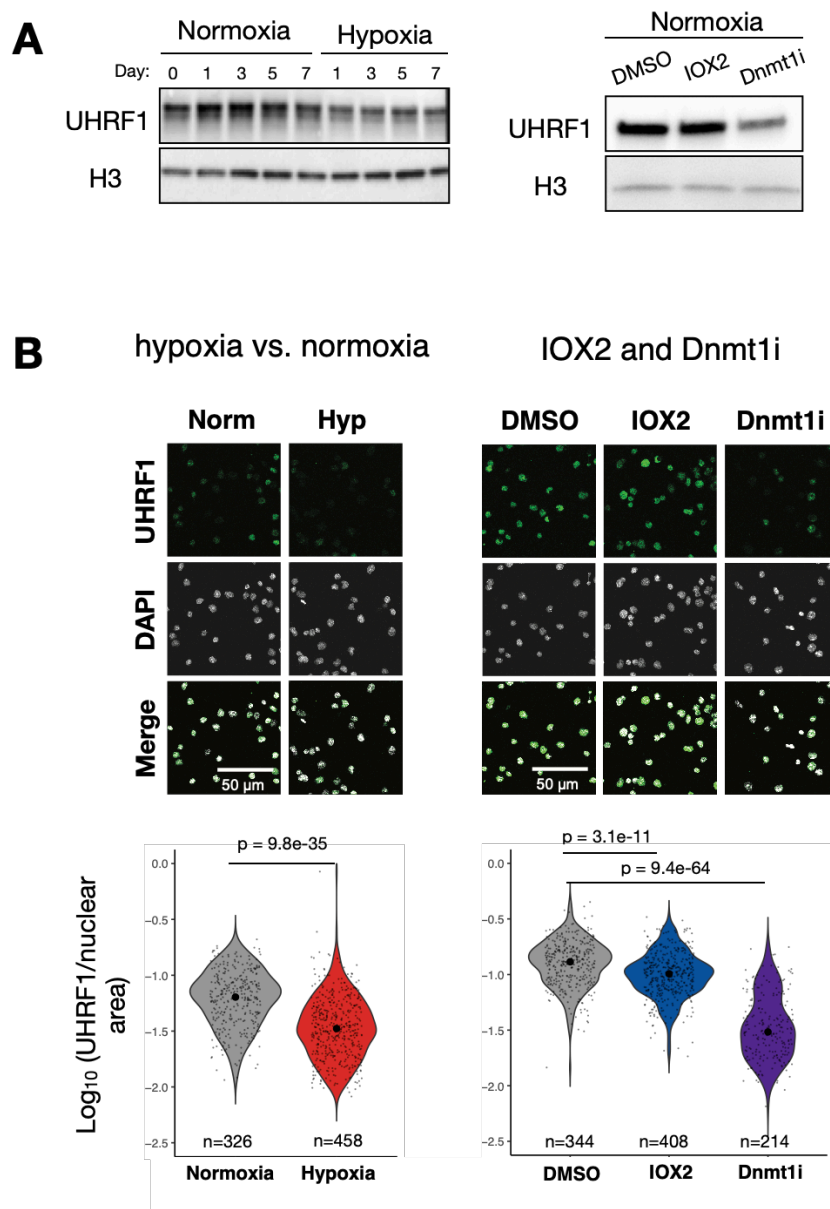
Figure 32. Hypoxia downregulates UHRF1 in a HIF1α–mediated manner.

(A) IF images and quantifications of Uhrf1 in hypoxic (d6) and normoxic ES cells (*left*) and normoxic ES cells treated with the HIF1α activator IOX2, Dnmt1 inhibitor, or DMSO as control (*right*). Fluorescence intensity of each nucleus was measured and normalized to the nuclear area. Each dot represents a cell. N represents the number of cells quantified per sample. Statistical tests performed are unpaired two-sample Wilcoxon tests. IF data generated by Maximilian Stötzel.

Due to the significant downregulation of UHRF1 in hypoxia and likely being regulated to some extent by activation of HIF1α, we investigated DNA methylation levels in ES cells exposed to acute (day 2) or prolonged hypoxia (day 7) by performing whole genome bisulfite sequencing (WGBS). As expected, hypoxic ES cells had globally reduced levels of DNA methylation compared to normoxic cells (**Figure 33A**). DNA methylation was reduced on day 2 of hypoxia (~25% demethylation), which then was largely regained on day 7 (~5-10% demethylation), potentially due to increased expression of the de novo DNA methyltransferase DNMT3B (**Figure 33B**). Expression of the TET DNA demethylases did not change, although their activity might be altered due to the inherent oxygen sensitivity (Burr et al., 2017) (**Figure 33B**). HIF1α-bound regions were already demethylated in normoxia (**Figure 33A**) and DNA demethylation was not specific to differentially expressed genes in hypoxia (**Figure 33C**).
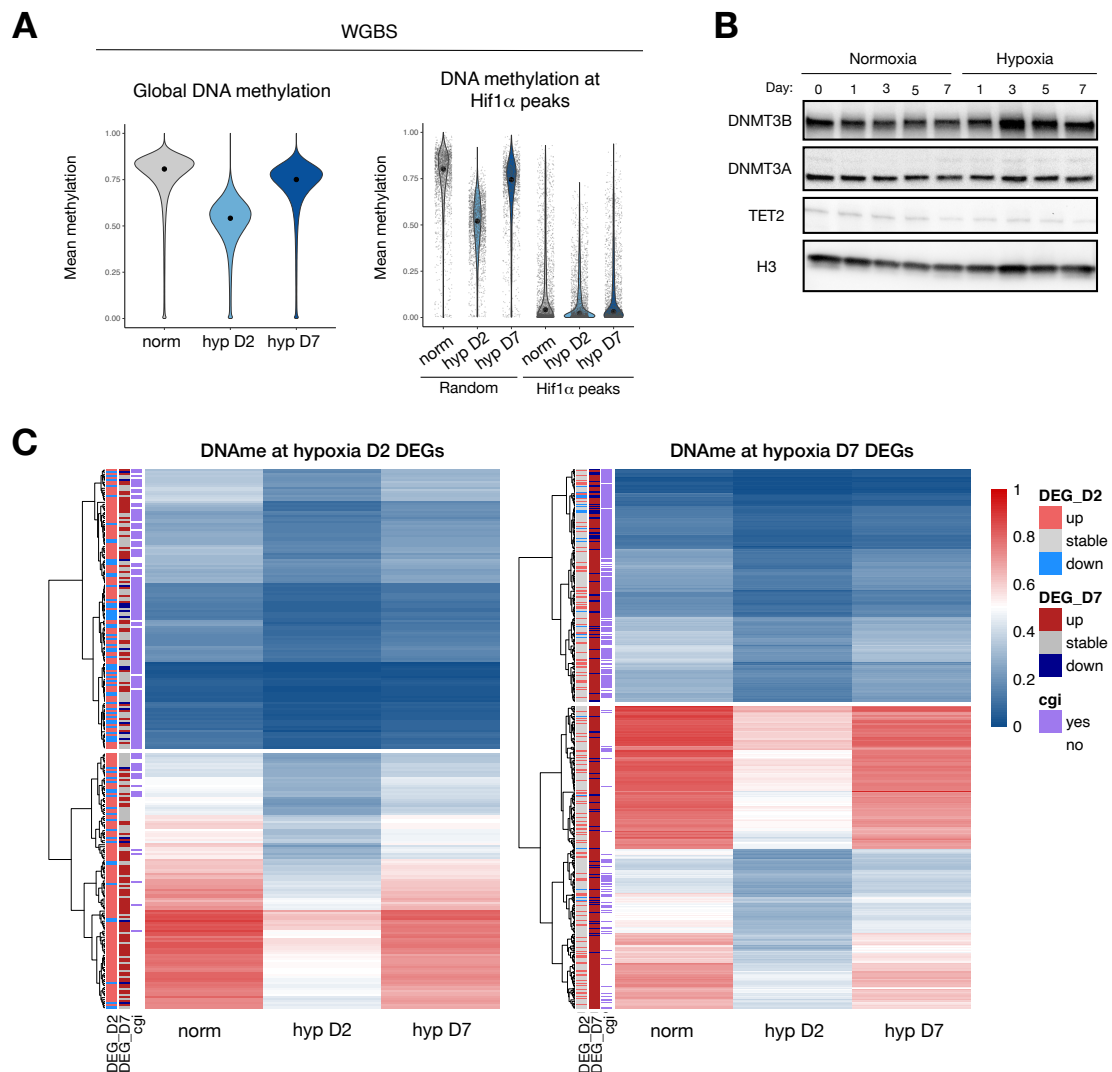
Figure 33. Hypoxia induces global DNA demethylation in ES cells.

(A) Global DNA methylation levels in normoxic and hypoxic ES cells determined by WGBS (*left*) and mean DNA methylation levels in normoxic and hypoxic ES cells at HIF1α bound peaks. As a

control, an equal number of random regions from the genome was used (*right*). (B) Western blot showing expression levels of DNMT3B, DNMT3A and TET2 in normoxic and hypoxic ES cells. H3 was used as loading control. (C) Heatmap showing mean DNA methylation levels of DE gene promoters identified on day 2 and day 7 in hypoxia. Bar annotations (reported as DEG_D2 and DEG_D7) show the up- (red), down- (blue) regulated and stable (gray) genes in hypoxia day 2 or day 7. Bar annotation (reported as cgi) depicts CG island. Data was generated in collaboration with Alexandra Mattei and Sara Hetzel.

Chemical HIF1α activation in normoxia led to DNA demethylation at levels compared to hypoxia but less than direct Dnmt1 inhibition (**Figure 34A**). Therefore, in ES cells, DNA demethylation is part of the HIF1α−dependent hypoxic response but is not a major determinant of gene regulation.

To dissect the contribution of DNA demethylation to the upregulation of key mesendoderm genes in hypoxia, we cultured normoxic ES cells with a DNMT1 inhibitor for 2 or 6 days (**Figure 34B**). DNA demethylation did not by itself lead to upregulation of the tested mesendoderm genes except for *Wnt3*. Its upregulation on day 2 may nevertheless be an initial trigger for downstream upregulation of mesendoderm genes in hypoxia.
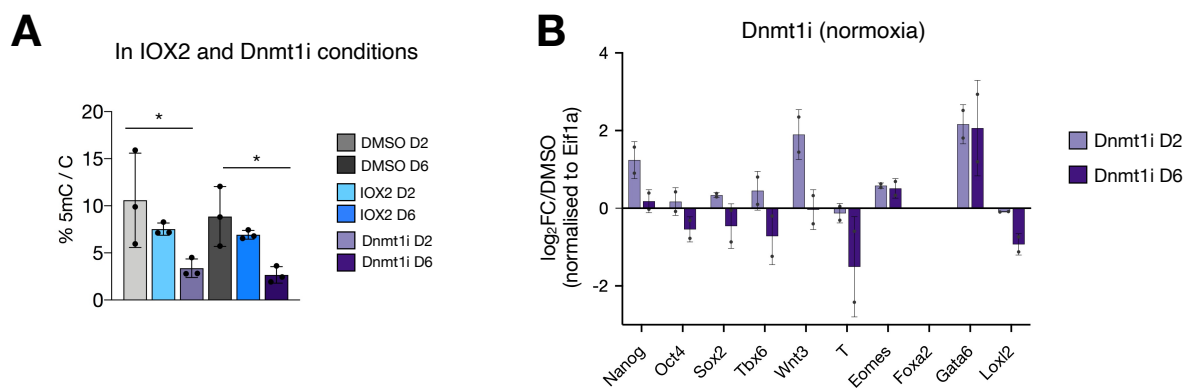
Figure 34. DNA demethylation in hypoxia is not the major contributor to the hypoxia-mediated transcriptional changes.

(A) DNA 5mC methylation levels in normoxic ES cells treated with IOX2, a DNMT1 inhibitor, or DMSO as control measured by mass spectrometry. Data represent 5mC normalized to total cytosine. Three biological replicates were used. (B) Relative expression levels of indicated genes in ES cells treated with a DNMT1 inhibitor for 2 or 6 days. Data represent log2FC over DMSO-treated ES cells normalized to Eif1a. Two biological replicates were used. Error bars represent s.d. Statistical test performed is two-way ANOVA.

3.6. Chromatin rewiring in response to hypoxia

To a lower extent, we also found PRC2 components being affected by a hypoxic environment in ES cells. At the protein level, EZH2 was downregulated in hypoxia day 2 (**Figure 24**). Additionally, HIF1α bound directly to its promoter region and to an enhancer region in close

proximity (**Figure 31A**). As many developmental genes are bivalent– present both the H3K4me3 active histone mark and H3K27me3 repressive histone mark— we considered the possibility that hypoxia-induced upregulation of key mesendoderm genes could occur as a result of altered H3K27me3 and H3K4me3 levels.

To explore this possibility, we profiled genomic distributions of H3K4me3 and H3K27me3 in acute and prolonged hypoxia compared to normoxia to identify locus-specific changes and their potential impact on gene expression. Quantitative ChIP-seq using Drosophila spike-in chromatin was performed to be able to reflect the global changes in abundance of these modifications. Quantitative ChIP-seq showed a global increase in H3K4me3 and H3K27me3 levels on day 2 of hypoxia, which then decreased on day 7 (**Figure 35A**). H3K27me3 levels on day 7 were on average 1.3-fold higher than in normoxia, whereas H3K4me3 was depleted compared to normoxia (on average -1.3-fold). We then calculated H3K4me3 or H3K27me3 enrichment at different genomic regions as categorized by ChromHMM (Ernst and Kellis, 2012) (**Figure 35B**). This analysis revealed a ~1.5-fold increase in H3K27me3 across the genome. Interestingly, bivalent regions were initially resistant to H3K27me3 hypermethylation on day 2, despite the depletion of H3K4me3. H3K4me3 was depleted from promoters on day 2 and intergenic and heterochromatic regions on day 7 of hypoxia exposure, suggesting that global H3K27me3 hypermethylation may reduce spurious transcription across the genome.

A

H3K4me3, H3K27me3 & DNAme
density genome-wide

B

Chromatin states

		Enhancer	Weak enhancer	Strong enhancer	Active promoter	Transcription transition	Transcription elongation	Bivalent chromatin	Intergenic	Insulator	Repressed chromatin	Heterochromatin
H3K4me3	hyp D2	0.31	0.23	0.03	−0.36	−0.33	0.2	−0.27	0.2	0.24	0.17	0.19
	hyp D7	0.09	0.07	0.28	0.13	0.09	−0.23	−0.28	−0.33	−0.23	−0.32	−0.28
H3K27me3	hyp D2	0.52	0.55	0.44	0.28	0.42	0.64	0.18	0.57	0.51	0.46	0.62
	hyp D7	0.47	0.47	0.4	0.3	0.35	0.48	0.47	0.49	0.47	0.53	0.47

log2 Fold−change

Figure 35. Hypoxia leads to global chromatin rewiring without major influence on transcription levels.

(A) Genome-wide levels of H3K4me3 and H3K27me3 and mean DNA methylation levels in normoxic and hypoxic ES cells. The genome was split into 1-Kb tiles and mean methylation levels of two replicates in each tile was computed. Box plots show the distribution of these tiles. Line indicates mean. (B) Heatmap showing log_2FC of H3K4me3 and H3K27me3 densities across 11 chromatin states in hypoxic vs normoxic ES cells.

These results show that hypoxia impacts global and local chromatin landscapes, potentially in crosstalk with global DNA demethylation.

THIRD PART: Functional relevance of the hypoxia-induced early primitive streak signature via gastruloid formation

We have shown that hypoxic ES cells present a transcriptional early primitive streak while retaining their pluripotent cell state. It is likely that, under differentiation cues, the hypoxia-induced signature enables the emergence of the cell states and tissues that arise from the

primitive streak in vivo. To better understand the functional relevance of hypoxia during the early stages of differentiation in ES cells, we used the gastruloid model, to study gastrulation in vitro.

3.7. The gastruloid: an in vitro model of gastrulation

To generate gastruloids, we used a defined number of ES cells (400 cells) and allowed them to aggregate for 48h. Then, between 48h and 72h, aggregates were treated with an exogenous WNT activator (CHIR99021 (Chi)). Under these conditions, aggregates undergo symmetry breaking, elongation and self-organization of the body plan (Brink et al., 2014) (**Figure 36**).
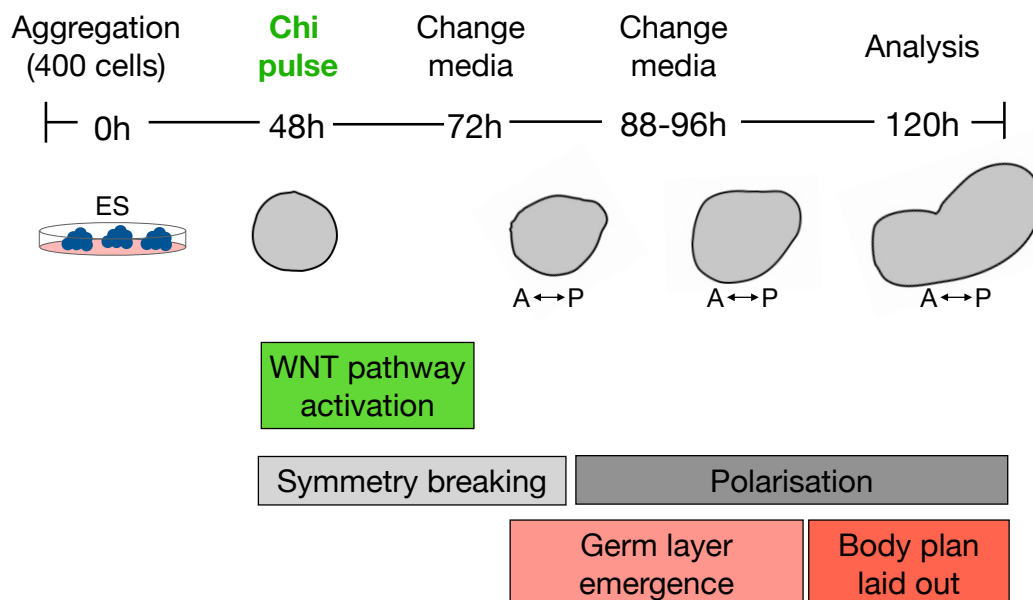
Figure 36. Schematic of the experimental setup for conventional gastruloid generation.

400 mouse ES cells were cultured in individual wells and allowed to aggregate for 48h. After aggregation, CHIRON was used for 24h to induce polarization that in the next steps will define the posterior axis by the expression of T. Gastruloids break symmetry and forms the posterior domain and elongate resulting in the formation of an elongated structure resembling the body plan of the in vivo gastrulating embryo. A: anterior axis; P: posterior axis.

In addition, they show polarized T expression at the posterior end. Simultaneously and mediated by WNT/T activation, the three germ layers emerge (Brink et al., 2014; Moris et al., 2020; Turner et al., 2017; Vianello and Lutolf, 2021).

Given that hypoxia induces expression of WNT pathway-related genes and T in ES cells, we reasoned that the hypoxia-mediated induction of a transcriptional early primitive streak signature might be enough to enable symmetry breaking and axial elongation in the absence

of exogenous WNT activation. Indeed, initial experiments performed by my colleague Ipek Gassaloglu, showed that gastruloids generated in hypoxia and in the absence of any exogenous WNT activation were able to break symmetry, showed T polarization and axial elongation (**Figure 37**).
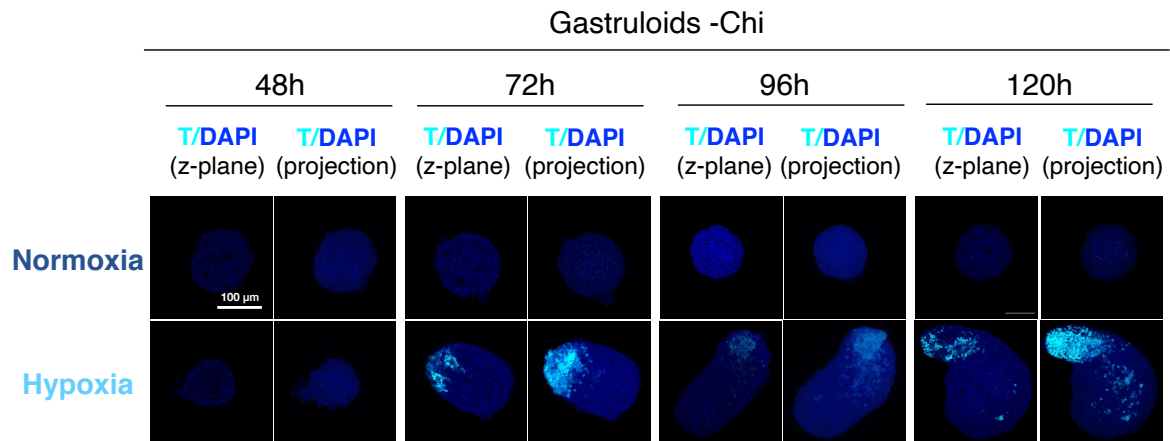
Figure 37. Hypoxia can induce spontaneous elongation of gastruloids in the absence of exogenous WNT activation.

Confocal fluorescence microscopy images of gastruloids from 48h to 120h post-aggregation and at the indicated conditions. 'Projection' depicts three-dimensional projection. Scale bar: 100um. Data generated by Ipek Gassaloglu.

3.7.1. Hypoxia-induced spontaneous elongation is WNT-dependent in hypoxic gastruloids

To test whether hypoxia induces elongation of gastruloids via WNT activation, we treated them (during the whole gastrulation formation) with two WNT pathway inhibitors (the PORCN inhibitor LFK-974 and the tankyrase inhibitor XAV939). We could see that inhibition of WNT pathway results in loss of spontaneous elongation (**Figure 38A**) and no induction of *T* and *Eomes* target genes (**Figure 38B**).
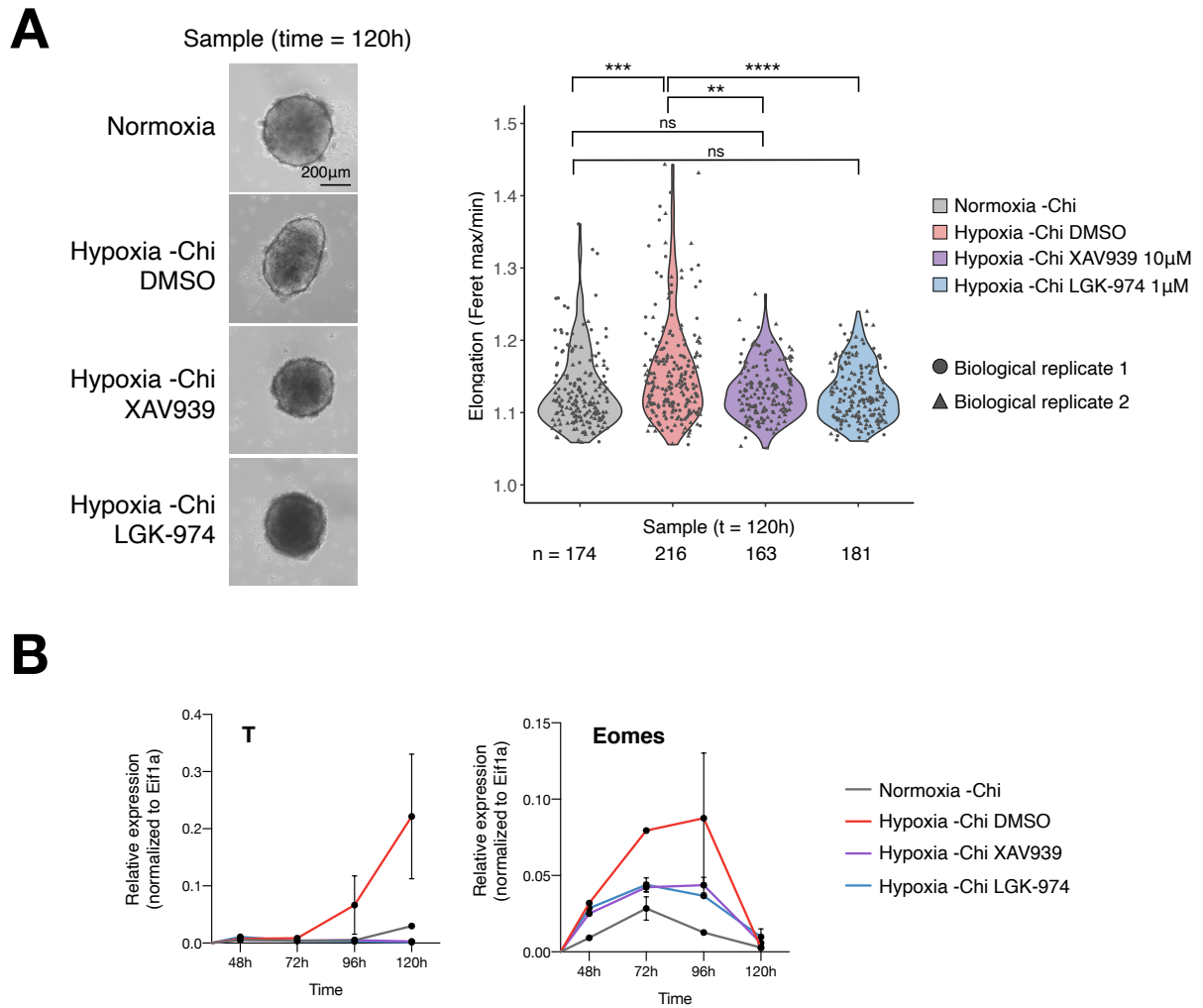
Figure 38. Hypoxia-induced spontaneous elongation in hypoxic gastruloids -Chi is WNT-dependent.

(A) Bright file images of gastruloids under the indicated conditions at 120h post-aggregation. Scale bar: 200um (*left*). Elongation index (defined as aspect ratio Feret max/Feret min) of gastruloids at the indicated conditions at 120h. Data from two biological replicates (*right*). (B) Relative expression levels of T and Eomes over time. Data represent two biological replicates. Error bars represent s.d.

All this data together shows that hypoxia equips ES cell aggregates with the ability to self-initiate the developmental events that characterize gastrulation, including symmetry breaking, polarization and axial elongation.

3.7.2. Role of hypoxia in conventional gastruloids (with exogenous WNT activation)

Although gastruloids in the absence of exogenous WNT activation could elongate in hypoxia (**Figure 37**), the efficiency was limited, where only ~30% of the structures did elongate. Therefore, we also studied the effect of hypoxia-induced T activity in gastruloids under the presence of exogenous WNT activation. Initial characterization carried out by my colleague

49

Ipek Gassaloglu showed that while hypoxic gastruloids presented a similar elongation and T expression posteriorization (**Figure 39**), the cellular composition was different. Specifically, hypoxic gastruloids showed a less mature neuroectodermal lineage and enhanced representation of the endodermal lineage.

Gastruloids +Chi

T/DAPI (z-plane)
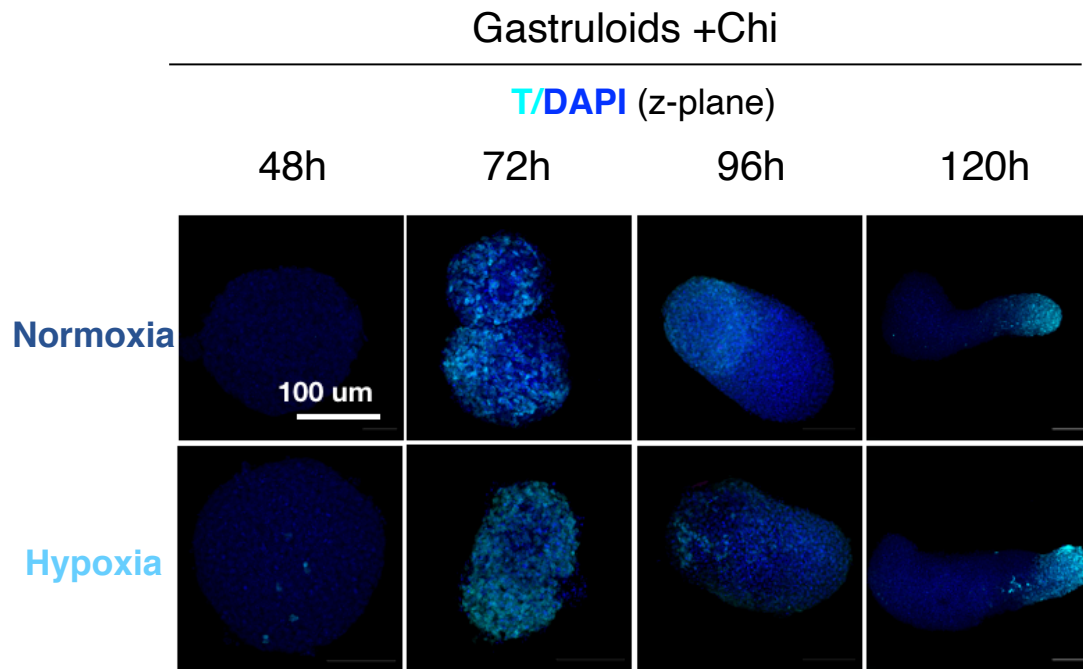
Figure 39. Hypoxia can induce the elongation of gastruloids.

Confocal fluorescence microscopy images of gastruloids from 48h to 120h post-aggregation and at the indicated conditions (*left*). Scale bar: 100um. Data generated by Ipek Gassaloglu.

Strikingly, endodermal cells self-organized in primitive lumen structures and in some special cases, we could also observe cells self-organizing around these lumen openings, forming tissue-like structures resembling the embryo gut tube (**Figure 40**).

Collectively, the data suggest that hypoxia is a potent modulator during gastruloid formation. Hypoxia not only induces symmetry breaking, polarization and axial elongation in a WNT-dependent manner, but it also shapes lineage trajectories and patterning. Additionally, it provides architecture to some cells to form tissue-like structures, reminiscent of the embryonic gut tube.
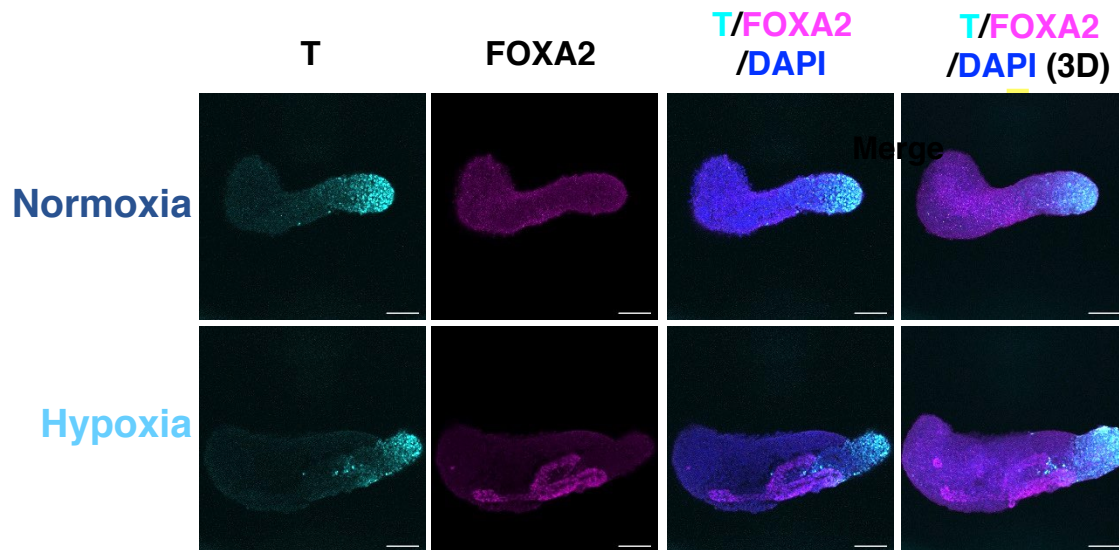
Figure 40. Endodermal cells self-organize in gut tube-like structures in hypoxic gastruloids.

Confocal fluorescence microscopy images of gastruloids at 120h post-aggregation and at the indicated conditions. Staining against T and FOXA2. Scale bar= 100um. Data generated by Ipek Gassaloglu.

3.7.3. Single-cell RNA-sequencing to explore the cell-type composition of hypoxic gastruloids

To better characterize the transcriptional changes that occur when implementing a hypoxic environment during gastruloid formation, we performed single-cell RNA-sequencing (scRNA-seq). scRNA-seq library was performed by my colleague Adriano Bolondi. As a control, we used normoxic gastruloid with Chi together with our hypoxic conditions in the presence and absence of Chi (Hypoxia +Chi or Hypoxia -Chi). This would give a direct comparison of the cellular composition of hypoxic gastruloids formed and elongated in the absence of WNT activation compared to conventional hypoxic gastruloids. We pooled 28 structures per condition and used lipid-indices bases multiplexing technology (MULTI-seq) (McGinnis et al., 2019a). Briefly, MULTI-seq allows the multiplexing of different samples by using lipid-based sample barcodes. Samples are pooled together before sequencing, allowing not only pooled sequencing but also a reduction of costs and batch effect related to sample handling during library preparation of different samples (McGinnis et al., 2019a).

To demultiplex sample barcodes, deMULTIplex R package was used (https://github.com/chris-mcginnis-ucsf/MULTI-seq). It allowed the generation of a sample barcode UMI count matrix where cells without MULTI-seq (sample) barcode and cells with more than a MULTI-seq barcode (doublets) were removed (**Figure 41**; **Table 1**)
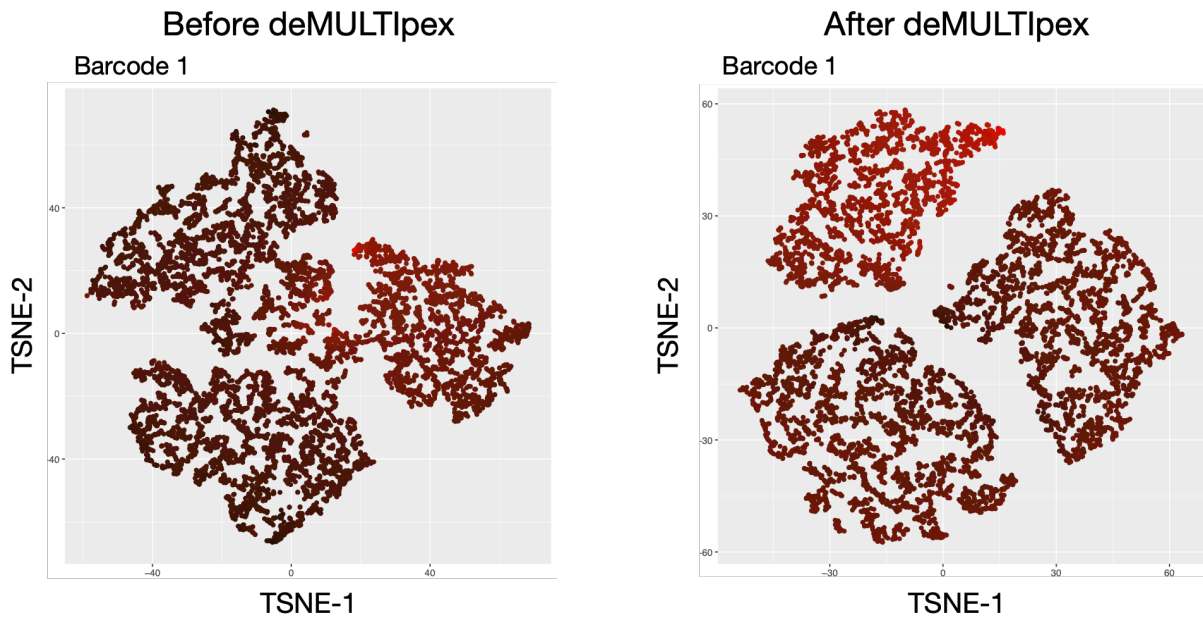
Figure 41. Demultiplexing MULTI-seq data.

T-SNE plots showing cell distribution according to the presence (red) or absence (black) of cell barcode 1 before and after demultiplexing.

	Normoxia +Chi	Hypoxia +Chi	Hypoxia -Chi	Negative	Doublets
After QC	1,979	2,477	2,692	650	1,210

Table 1. Number of cells and its indicated category after demultiplexing approach.

Following demultiplexing, downstream analyses were performed using Seurat R package, see '*Bioinformatic analyses*' in the '*Methods*' section for further details. After quality control, a total number of 7,014 cells were used for downstream analysis.

Then, a conventional scRNA-seq workflow was performed (see '*Bioinformatic analyses*'). The integrated data set was used to perform Seurat clustering and a total of ten clusters were identified. UMAP dimension reduction technique and coloring by sample showed distinct cell type contributions depending on the experimental condition (**Figure 42A, B**).
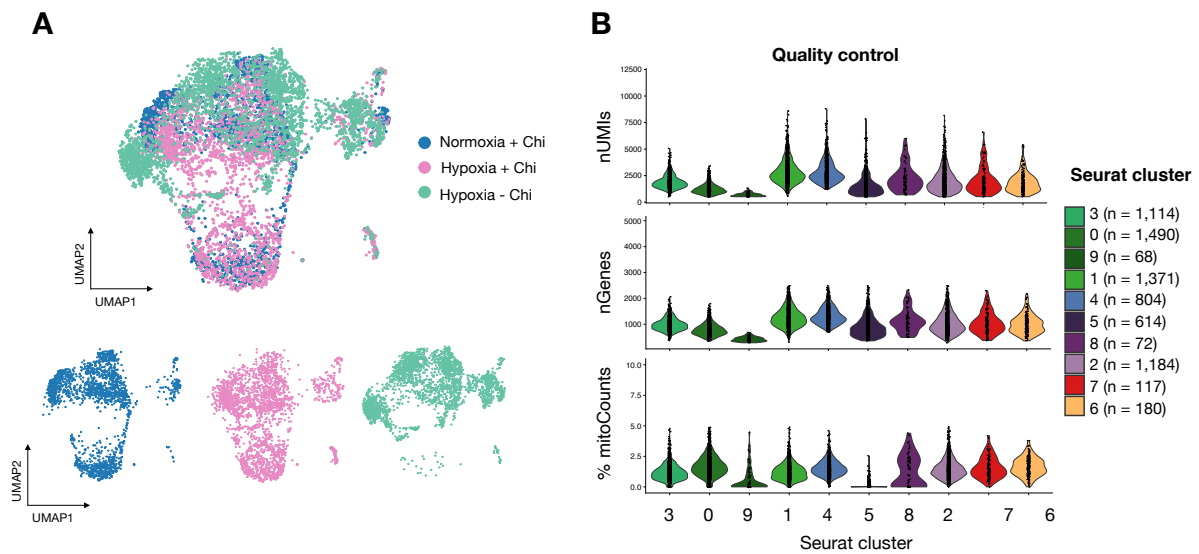
Figure 42. Integration and Seurat clustering of scRNA-seq data.

(A) Integrated gastruloid datasets UMAP colored by condition (*top*) and split by condition (*bottom*).
(B) Quality control violin plots by identified Seurat clustering.

Next, gastruloid cells were annotated using two distinct reference scRNA-seq atlases of embryonic stages E6.5-E8.5 (Grosswendt et al., 2020a; Pijuan-Sala et al., 2018). Prediction scores were used as a measurement for the cell-type annotation call (**Figure 43**; *Grosswendt et al., 2020 reference atlas as an example*).
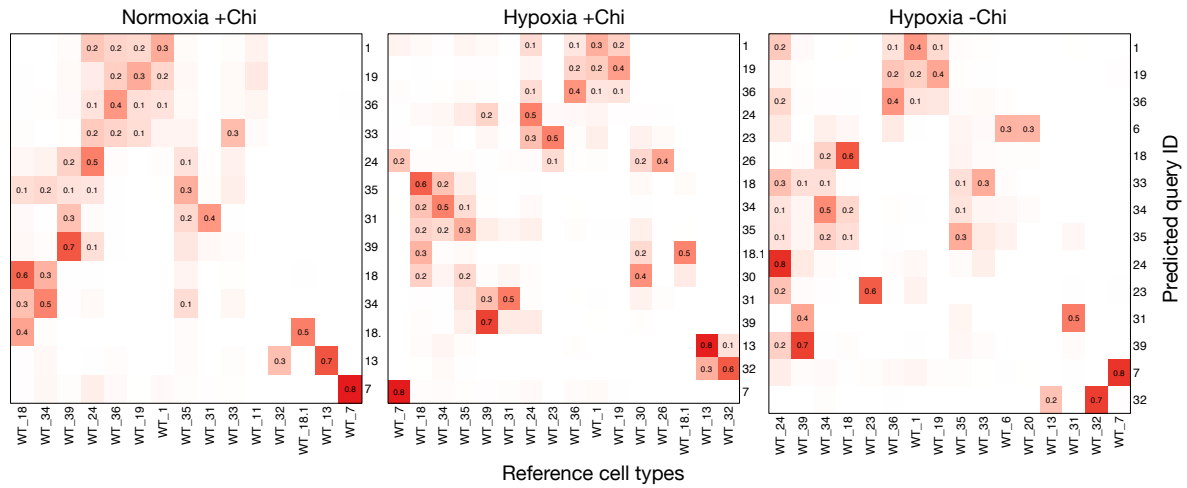
Reference cell types

1. Early ectoderm 1	18.1 Posterior presomitic mesoderm
19. Early ectoderm 2	18. Somite
33. Neural crest	34. Pharyingeal arch mesoderm
35. Similar neural crest	6. Secondary heart field/splanchnic lateral plate mesoderm
24. Fore / mid brain	13. Hematopoietic and endothelial progenitors
39. Future spinal cord	32. Angioblast
26. NMPs early	36. Primitive blood progenitors
31. NMPs late	7. Gut endoderm
30. Anterior presomic mesoderm	23. Notochord

Figure 43. Prediction score as a measure for cell-type annotations.

Heatmaps show the predicted query score by reference cell type.

Of note, it is worth mentioning the relevance at this stage in the selection of a reference atlas. By comparing the results of using two independent atlases covering the same developmental stages, we observed highly similar results, yet some differences were found in some cell type assignments (**Figure 44A, B, C, D**). These results highlight the importance of the generation of a 'universal' reference atlas with reduced human biased in terms of manual annotations and universal cell state names.
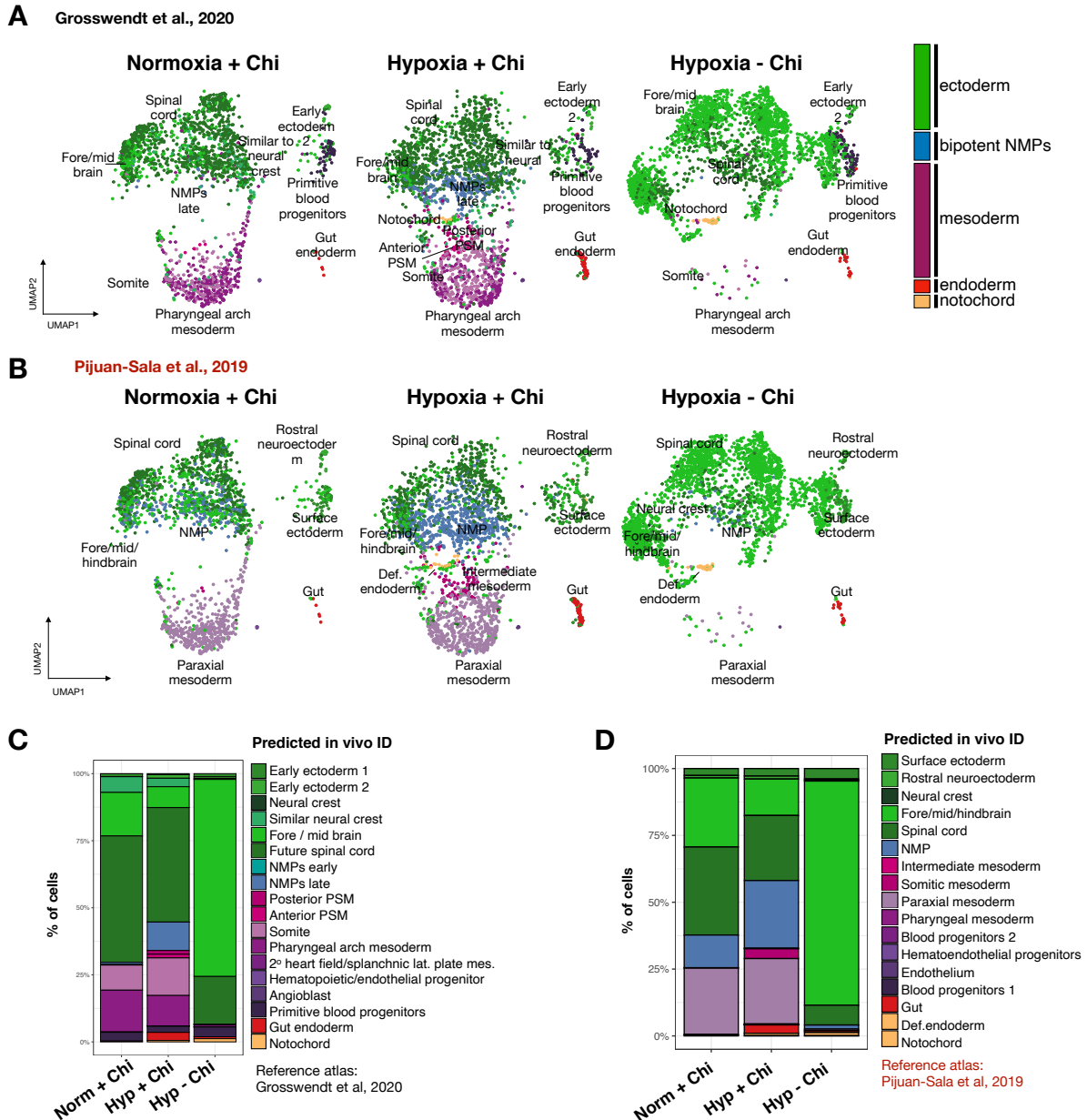
Figure 44. scRNA-seq reveals enhanced lineage representation of diverse cell types in hypoxic gastruloids.

(A) UMAPs by sample with Seurat clusters colored by germ layer and according to in vivo predicted cell type with reference atlas from Grosswendt et al., 2020. (B) UMAPs by sample with Seurat clusters colored by germ layer and according to in vivo predicted cell type with reference atlas from Pijuan-Sala et al., 2019. (C) Percentage of cells assigned to the in vivo predicted cell types with reference atlas from Grosswendt et al., 2020. (D) Percentage of cells assigned to the in vivo predicted cell types with reference atlas from Pijuan-Sala et al., 2019.

Coloring of the annotated cell states by germ layer of origin revealed substantial differences between conditions (**Figure 44A, B, C, D**). The data shows a clear increase of neural cells and a nearly complete depletion of mesodermal lineage in Hypoxia -Chi gastruloids. These results, therefore, suggest that spontaneous T induction under hypoxia is not enough to generate or

sustain mesodermal fate in the absence of exogenous WNT activation. On the other hand, Hypoxia +Chi gastruloids presented a substantial increase of the mesodermal lineage and were able to generate all cell types that are present under normoxic conditions. (**Figure 44A, B, C, D**). For the purpose of this investigation, we unbiasedly decided to use the Grosswendt et al., 2020 reference atlas for further analysis.

To better understand the differences in cellular composition, we calculated the differential abundance of each cell type over Normoxia +Chi (**Figure 45**).
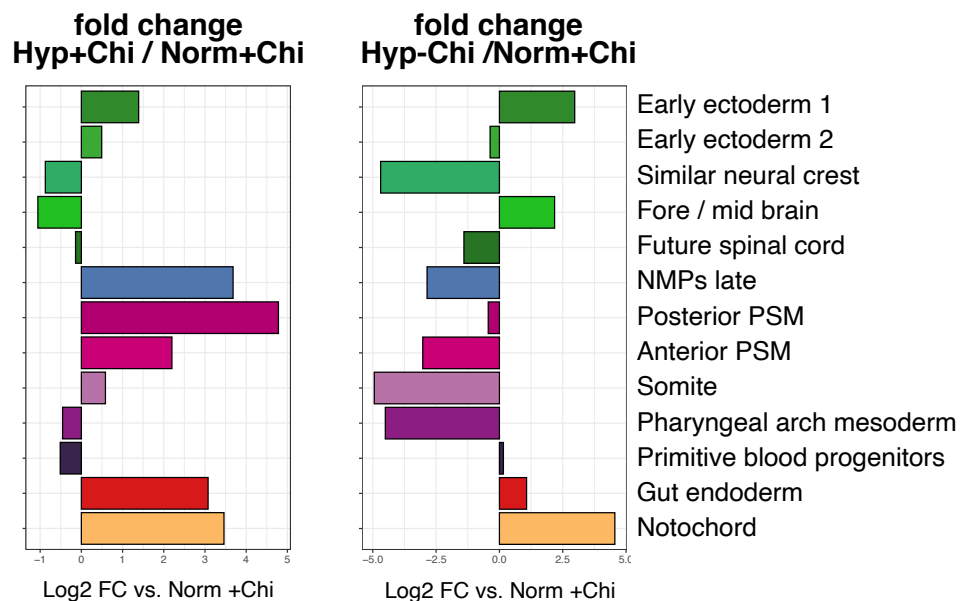
Figure 45. Hypoxia induces redistribution of cellular composition in several cell types.

Relative abundance of each predicted cell type in Hypoxia+Chi and Hypoxia-Chi gastruloids compared to Normoxia+Chi. Cell types with less than ten cells across conditions are not shown.

The data shows a clear increase in endodermal lineage (**Figure 45**), both with and without exogenous WNT activation. This was corroborated by the presence of *Foxa2*- and *Sox17*-positive cells (**Figure 45; Figure 46A**, annotated in red). Strikingly, also in both with and without exogenous WNT activation, we detected the emergence of cells with a notochordial transcriptional signature. These cells were co-expressing *Shh* (well-known marker of the notochord (Echelard et al., 1993), *T* and *Sox9* (**Figure 45; Figure 46B**, annotated in orange). Unfortunately, we were not able to detect the expression of *Noto* notochord marker gene. The increase (of gut endoderm-like cells) and emergence (of notochord-like cells) in both hypoxic conditions point to a WNT-independent route.
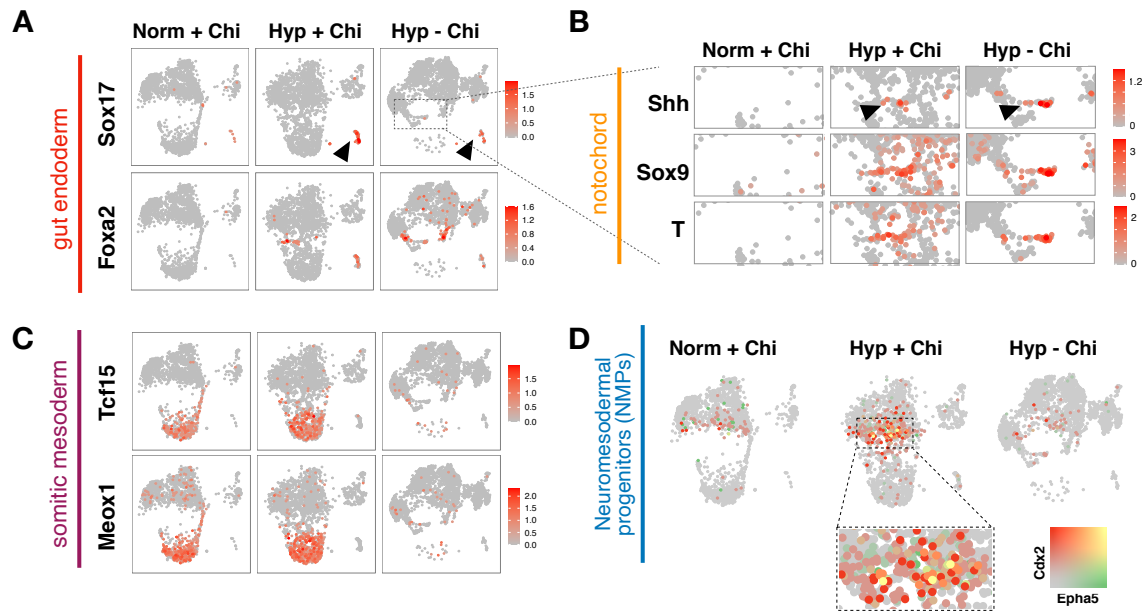
Figure 46. Enhanced and depleted representation of different cell types in hypoxic gastruloids.

(A-C) UMAP feature plots colored by expression of marker genes for gut endoderm, notochord and somitic mesoderm. (D) UMAP feature plot colored by co-expression of NMPs marker genes.

While Hypoxia -Chi gastruloids are biased towards neural identity, Hypoxia +Chi gastruloids presented an enhanced mesodermal signature (**Figure 46C**), especially with increase and appearance on anterior and posterior pre-somitic mesoderm (PSM), respectively (**Figure 45A, B; Figure 46**). We believe that the balance between neural and mesodermal lineages found in Hypoxia +Chi gastruloids is likely supported by an increased population of NMPs (**Figure 46D**), generating a more sustainable balance in the specification of both neural and mesodermal lineages.

To better investigate the neural cells, we used the Hox code (Neijts et al., 2014) as a proxy for the study of the A-P identity. To do so, all cells annotated as spinal cord-like and mid/hind brain-like cells were used (**Figure 47A**).

B

Future spinal cord

Hox+ Hox-

spinal cord-like

spinal cord-like
<5/30 brain markers expressed

brain-like
≥5/30 brain markers expressed & no expression of Fgfbp3 spinal cord marker

C

| Norm + Chi | Hyp + Chi | Hyp - Chi |

Hox+
Hox- spinal cord-like
Hox- brain-like

Hoxb3 Hoxa9 Hoxb6 Hoxd4 Fgfbp3 Crabp2 Tmsb4x Mest En1 Otx2

Spinal cord — Brain

Expression
maximum
minimum

% detected cells
· 0
○ 10
○ 20
○ 30
○ 40
○ 50

Figure 47. Characterization of the neural cell identity in .... **loids**.

(A) Schematic overview of the classification of neural cells. .... of cells assigned to the indicated categories. (I) Dot plot showing the expression of .... spinal cord- and brain-like cell types.

Then we performed a classification according to the expression of Hox genes and several brain- and spinal cord-marker genes. More details can be found in '*Bioinformatic analysis*' in the '*methods section*'. In general, we found decreased Hox-positive (spinal cord) cells in both hypoxic conditions, especially in Hypoxia -Chi gastruloids (**Figure 47B**). Concomitantly, a strong increase of Hox-negative and especially brain-like cells was observed in Hypoxia -Chi gastruloids (**Figure 47B**). This analysis suggests that in hypoxic conditions and substantially in the absence of exogenous WNT activation, the neural cells have a more brain-like signature, which is however immature due to the lack of brain markers such as *Otx2* and *En1* (**Figure 47C**).

It is known that notochord-secreted Shh confers ventral identity to spinal cord cells (Sagner and Briscoe, 2019a). With the emergence of notochordial-like cells expressing *Shh* (**Figure 48B**) in hypoxic gastruloids, we hypothesized that it might be possible that hypoxia-induced *Shh* expression could lead to the appearance of ventral identity. To study dorsal-ventral (D-V) identity, spinal cord cells were categorized into dorsal, midline or ventral according to the expression markers or known D-V genes.
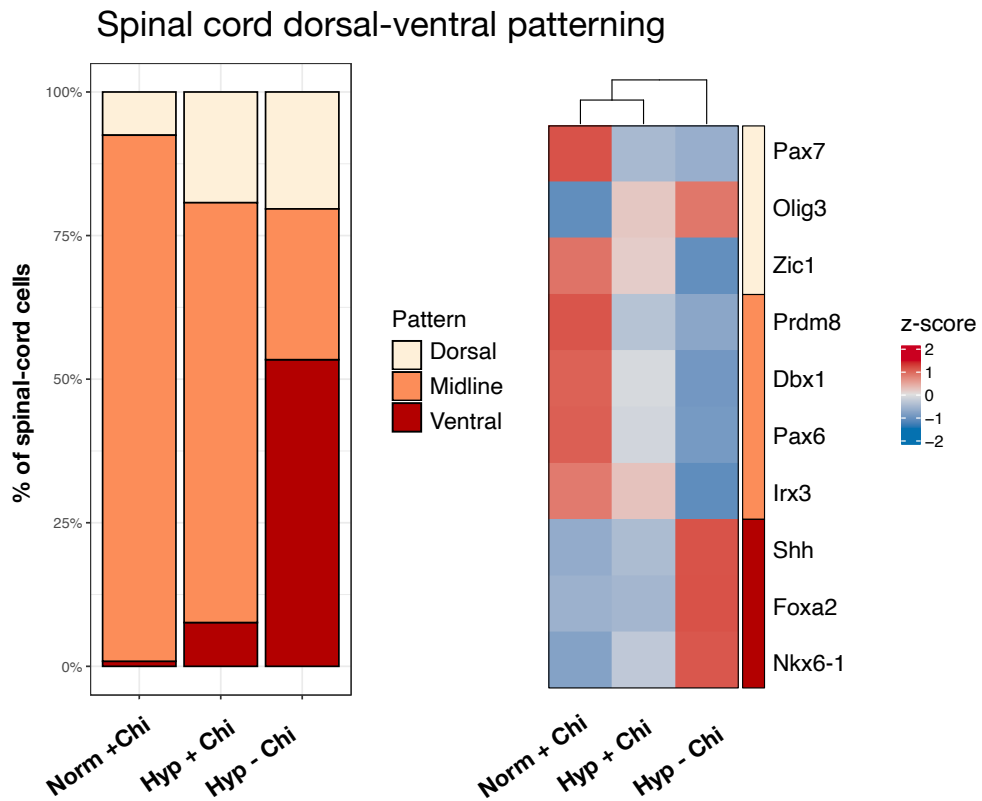
Figure 48. Characterization of spinal cord dorsal-ventral patterning in hypoxic gastruloids.

Stacked bar plot showing the categorization of the spinal cord into dorsal-ventral patterning (*left*), and heatmap showing pseudo-bulk z-score expression values of the indicated genes used for the categorization (*right*).

D-V categorization showed that in both hypoxic gastruloids, spinal cord cells with dorsal identity increased. Importantly, both hypoxic gastruloids and especially, those without exogenous WNT activation, dramatically present an increase in ventral spinal cord identity. (**Figure 48**). This identity, as hypothesized, is conferred by the notochord-secreted *Shh* under hypoxic conditions.

Taking scRNA-seq data together, we show that (1) endoderm is highly enriched in hypoxia and may form independently of WNT activity; (2) similarly notochordial-like cells emerge in hypoxia and are otherwise not present under normoxic conditions; (3) hypoxia enhances lineage representation in conventional gastruloids, likely due to the presence of NMPs and especially enhancing mesodermal cell types; and (4) spontaneously elongating gastruloids largely lack mesodermal lineages but contains more neural cells with a cephalic identity.

4. Discussion

4.1. Impact of hypoxia on embryonic and extraembryonic stem cells

In hypoxia, ES, TS and XEN cells are able to maintain their cell state profile. They present an unaltered morphology, and stem cell marker levels either do not change (in ES cells) or significantly increase in extraembryonic cell types. A deeper characterization at the transcriptomic level shows that a moderate number of genes are differentially expressed in hypoxia when compared to normoxia. Interestingly, the hypoxic transcriptional response is cell type- and time-specific.

In ES cells, data show a temporal and constant upregulation of genes involved in response to hypoxia and angiogenesis, the latter known to be activated by low oxygen tensions (Iyer et al., 1998; Ryan et al., 2005). On the other hand, the time-dependent response is composed of an early upregulation of the glycolytic pathway in line with the metabolic switch from oxidative phosphorylation to glycolysis in the absence or reduction of oxygen concentrations (Muz et al., 2015). Afterward, temporal changes in genes related to cell fate, differentiation and development come into play (**Figure 15A**). Intriguingly, in extraembryonic cell types, metabolism seems to be unaltered while differentiation genes are (**Figure 15B, C**). Thus, hypoxia promotes a transcriptional response that is cell type- and time-specific towards the differentiation of stem cells.

Indeed, it has been shown that hypoxia promotes the differentiation of certain types of stem or progenitor cells while inhibiting the differentiation of others, depending on the context. For example, HSCs reside in the most hypoxic niche in the adult bone marrow. It has been shown that hypoxia increases the proliferation of hemangioblast (the bipotent progenitor population of endothelial and hematopoietic cells) (Parmar et al., 2007). In contrast, hypoxia has been described to promote stemness and pluripotency of human ES cells in prolonged hypoxic cultures and reduce the appearance of '*differentiating*' colonies (Ramírez-Bergeron et al., 2004). In my thesis, I present hypoxia as a factor that indeed modulates cell fate and differentiation, in a context-dependent manner. In agreement with what has been observed in human ES cells, mouse ES cells in hypoxia do not present an altered pluripotent profile. In contrast, the morphology of hypoxic and normoxic ES cell colonies was rather similar.

Yet, it is remarkable that many genes related to cell development and differentiation are altered. A closer look at the expression pattern of pluripotency and early lineage-specification genes confirms the unaltered pluripotency capacity but a tendency towards upregulation of

mesendoderm markers (**Figure 16**). Mesendodermal genes are induced by WNT pathway activity (Tortelote et al., 2013). Specifically, WNT pathway induces local T expression in the posterior axis, driving and forming the PS (**Figure 2; Figure 3**). Cells from the PS undergo EMT to later form and organize into the three germ layers (Williams et al., 2012). In accordance with it, hypoxic ES cells at day 7 activate WNT pathway and EMT-related genes. Particularly, the upregulation of *Wnt3* and *T* is gradual and inversely correlates with oxygen levels (**Figure 20A, B**), revealing a direct effect of oxygen tensions regulating key factors involved in early embryogenesis events.

Yet, the transcriptomic observations are not entirely recapitulated at the protein level. Even though this observation can be related to a known limited sensitivity of mass spectrometry to detect lowly abundant proteins (such as transcription factors), it is likely that under hypoxia, there is an increase of oxidative stress that leads to global reduction as a mechanism to save energy while selectively inducing the expression of proteins that are necessary for the adaptive response, such as metabolic related ones (Lee et al., 2020; Harding et al., 2000; Harding et al., 2003).

In the first part of my thesis, I aimed to determine the hypoxia-related response in embryonic and extraembryonic cells. The initial characterization showed that both embryonic and extraembryonic cell types maintain their cell state characteristics. Yet, at the transcriptional profile, they show a cell-type and time-dependent response. In particular, hypoxic ES cells induce WNT pathway-related genes, leading to the activation of a transcriptional PS signature that does not lead to spontaneous differentiation. Nonetheless, the functionality of this signature might come into play under the proper developmental and differentiation cues.

4.2. Mechanisms involved in the hypoxic response in ES cells

HIF1α is the transcription factor master regulator of any hypoxic response (Semenza, 1999). In normoxia, it is marked to be degraded. In contrast, in hypoxia, it is stabilized and translocated to the nucleus where together with HIF1ß binds to HER and modulates the expression of HRGs (Lee et al., 2020; Maxwell et al., 2001). Recent studies have shown the existing link between hypoxia, HIFs and molecules that are crucial for the regulation of developmental and differentiation events of stem and progenitor cells.

First, Kaidi et al., 2007 showed that hypoxia modulates WNT pathway. In short, hypoxia leads to the downregulation of ß-CATENIN and induces the expression of TCF4 which competes with HIF1α to bind to *Ctnnb1* (ß-Catenin) and promotes a hypoxia-mediated transcriptional

response. In agreement with the literature, I show that hypoxia induces the expression of WNT-related genes such as *Wnt3* and its downstream effectors, *T* and *Eomes* and that HIF1α in a normoxic context is enough to recapitulate their expression pattern observed in hypoxia (compared log$_2$FC over DMSO to normoxia, respectively in **Figures 20A** and **Figure 25**). Profiling of HIF1α occupancy genome-wide by ChIP-sequencing shows that HIF1α does not bind directly to those genes. Indeed, the main HIF1α target genes bound at promoter regions are metabolism-related genes (**Figure 27C**), showing the direct role of HIF1α in shaping the metabolic profile of ES cells under hypoxia. Yet, even though I observed HIF1α binding at the *Ctnnb1* promoter, its transcript levels remained unchanged (**Figure 28**).

Interestingly, HIF1α and ß-CATENIN colocalize at both promoter and ES active enhancer regions upon WNT activation (**Figure 30A**). Indeed, as an example, HIF1α in hypoxia and ß-CATENIN colocalize at an ES enhancer region in close proximity to *Lef1*, an important WNT pathway effector which is upregulated at the transcript level in hypoxia day 2 (**Figure 30B**). Together, these data show that HIF1α likely works together with ß-CATENIN to selectively induce the activation of developmental genes in hypoxia and promote a hypoxia-mediated transcriptional response in cooperation with WNT pathway molecules to regulate developmental and differentiation events of stem cells (Simon and Keith, 2008).

Oxygen has been shown to modulate the epigenetic status of the embryo. For instance, in bovine embryos, hypoxia leads to a decreased DNA methylation profile at the blastocyst stage (Li et al., 2014). In mouse ES cells, I show that the DNA methylation levels were decreased in hypoxia compared to normoxia, being global rather than local (**Figure 33A, C**). By mass spectrometry, I identified many deregulated proteins including some epigenetic regulators (**Figure 24**). Among them, UHRF1 –a counterpart of DNMT1—is significantly downregulated in a HIF1α-modulated manner (**Figure 32A, B**) and very likely being the responsible effector in driving the DNA demethylation profile on hypoxic ES cells.

EZH2 (PRC2 component) is downregulated to a lesser extent in hypoxia (**Figure 24**). Interestingly, HIF1α binds at its promoter region as well as at an ES enhancer region in close proximity (**Figure 31A**). Furthermore, another PRC2 component, *Eed*, is also bound by HIF1α at its promoter region and *Eed* transcript levels are reduced in hypoxia (**Figure 31A, B**). Developmental genes are known to be regulated by bivalent chromatin marks (Bernstein et al., 2006). These are H3K4me3 and H3K27me3, the latter, deposited by EZH2, and highly regulated by PRC2 (Loh and Veenstra, 2022). In hypoxia, H3K4me3 and H3K27me3 present global alterations rather than local ones (**Figure 35A**). Yet, ChromHMM analysis shows that bivalent regions upon 2 days of the hypoxic environment are resistant to gain of H3K27me3,

even though H3K4me3 is already decreased (**Figure 35B**). Instead, H3K27me3 is highly enriched at heterochromatin regions (**Figure 35B**), likely to overcome DNA demethylation showing possible crosstalk between DNA demethylation and H3K27me3 redistribution.

These data collectively reveal the role of environmental oxygen in altering the chromatin landscape of the embryo. Importantly, hypoxia also plays a role in controlling disease phenotypes. An example of that is endometriosis, where hypoxia has been shown to promote the disease by altering the epigenetic landscape in endometrial cells (Li et al., 2021). In agreement with my data, hypoxia leads to the degradation of DNMT1 leading to global DNA demethylation and regulating a downstream global gene activation (Hsiao et al., 2015). Additionally, PRC2 proteins such as EED, SUZ12 and EZH2 were found downregulated. Thus, leading to a global increase of the repressive histone modification H3K27me3 and inducing genes involved in proliferation and angiogenesis (Lin et al., 2019).

In the second part of my thesis, I aimed to decipher the mechanisms by which hypoxia leads to the observed transcriptional response. In particular to better understand the mechanisms regulating the induction of WNT pathway-related- and mesendoderm marker- genes. Here, I identified that activation of HIF1α in normoxia is able to recapitulate the induction of developmental genes at the transcript levels of those found in hypoxia. Yet, this activation is not by direct binding of HIF1α to their promoter regions. In contrast, it is likely that HIF1α in cooperation with ß-CATENIN binds to promoter regions and distal regions, such as ES active enhancers to modulate downstream events. In terms of the epigenetic landscape, I identified DNA methylation and chromatin rewiring as important contributors to the hypoxia-mediated response, in agreement and in a conserved manner as in disease-related phenotypes such as endometriosis.

4.3. Functional role of the hypoxia-mediated transcriptional PS signature in ES cells via gastruloid formation

In the absence of the right differentiation culture conditions, hypoxia induces a transcriptional early PS signature in ES cells. 3D in vitro differentiation, via gastruloid formation, shows the relevance of a hypoxic environment during early embryogenesis. In the conventional gastruloid formation protocol, a short pulse of exogenous WNT activation is required for gastruloids to form (Brink et al., 2014). It is of great relevance that hypoxic gastruloids in the absence of exogenous WNT activation are able to break symmetry, present polarization and elongate – the developmental events that characterize gastrulation (**Figure 37**). My data reveal the previous knowledge on the crucial role of WNT pathway activation that leads to PS formation

in vivo and to undergo symmetry breaking and axial elongation in in vitro models such as gastruloids (Williams et al., 2012; Brink et al., 2014). I further show that this spontaneous elongation of hypoxic gastruloids is WNT-dependent (**Figure 38**), revealing the hypoxic mode of action: activation of WNT pathway to modulate the developmental events happening during gastrulation. This result further agrees with the direct link between hypoxia and WNT activation in modulating stem cell fate or differentiation depending on the context ((Simon and Keith, 2008); Kaidi et al., 2007).

Yet, the formation of gastruloids in the absence of exogenous WNT activation is limited in hypoxia (~30% of the structures are able to elongate). In the future, it will be of great interest to address how this efficiency could be enhanced. Modulation of the levels of oxygen or/and the timing of the hypoxic exposure would help address it.

Further, investigation on the role of hypoxia in conventional gastruloids (with exogenous WNT activation) shows that hypoxic gastruloids are able to elongate similarly to their normoxic counterparts.

Analysis of the neuroectodermal and endodermal lineages in conventional hypoxic gastruloids shows that hypoxia generates a less mature population of neuroectodermal cells, suggesting a bigger pool of progenitor cells. On the other hand, hypoxic gastruloids present a significant increase of endodermal cells (FOXA2+ cells), that self-organize in primitive lumen structures, forming in some cases tubular structures, reminiscent of the embryonic gut tube (**Figure 40**). These data together present hypoxia as a potent modulator of lineage trajectories, which also provides morphological cues for some groups of cells to form tissue-like structures. Indeed, when it comes to embryo architecture, it is clear that there are crucial inputs that are missing in the conventional gastruloid model. As such, I propose hypoxia as an environmental factor able to alter cell-intrinsic determinants to promote tissue-like architecture, in a similar way as previously shown by treating gastruloids with FGF2 which inducted robust FOXA2+ tubular structures (Gharibi et al., 2020).

ScRNA-seq in hypoxic (+/-Chi) and normoxic (+Chi) gastruloids allows for a better characterization of how hypoxia modulates lineage choices, at the transcriptomic level (**Figure 44**). Annotation of cells to an in vivo reference atlas of the gastrulating mouse (E6.5-E8.5) confirms the potent modulation role of hypoxia. In summary, both hypoxic +/-Chi gastruloids present an enriched population of endodermal and notochordial cells (**Figure 45**). To my knowledge, there is only one more piece of evidence where notochordial cells were found in a gastruloid model. To achieve it, scientists used two different aggregates where only one of

them was exposed to chemical WNT activation, showing the relevance of the WNT mode of activation (Xu et al., 2021).

In conventional gastruloids, hypoxia presents an enhanced lineage representation, which might be due to an increase in the bipotent NMP population of cells (**Figure 45**) known to contribute to both the neuroectodermal and mesodermal lineages in vivo (Koch et al., 2017; Tzouanacou et al., 2009). On the other hand, spontaneously elongating gastruloids lack mesodermal lineages and are biased towards the formation of more neural cells with a large proportion of cells with a cephalic identity (**Figure 47**). In agreement with Girgin et al., 2021a, altering the WNT activity during gastruloid formation, in this case through WNT inhibition, reshape cellular identity by achieving more anterior neural cell types.

In the third part of my thesis, I aimed to investigate the functional impact of the hypoxia-mediated PS signature in ES cells. By employing gastruloid formation, I identified hypoxia as a potent microenvironmental factor able to promote the developmental events characteristic of gastrulation: symmetry breaking, polarization and axial elongation. Additionally, I showed that hypoxia does not only provide patterning but acts to guide cells to self-organize into tissue-like structures, providing architecture too. By using scRNA-seq, I provide a transcriptomic picture of the role of hypoxia in shaping cell fate trajectories. Hypoxia enhances lineage representation in conventional gastruloids and modulates the enrichment and appearance of cell types poorly found or completely missing in normoxia, such as endodermal and notochordial cells, respectively. Last, categorization analyses of neural cells provide a clearer role of the importance of hypoxia in generating notochordial-like cells that express and secrete *Shh*. Likely, neural cells in the hypoxic gastruloids show an enriched ventral identity due to the notochord-secreted *Shh* (Sagner and Briscoe, 2019a).

4.4. Open questions and future perspectives

My thesis presents hypoxia as a potent microenvironmental factor shaping cell fate decisions in stem cells and the post-implantation embryo. Indeed, the implanted embryo in the uterus is exposed to as low as ~1.5% oxygen levels in vivo (Houghton, 2021). Future approaches should investigate deeper the role of hypoxia in lineage choices, whether hypoxia has a general role, for instance in promoting a metabolic switch towards glycolysis as a mode of cellular adaptation, or it, also, has a tissue and/or context-dependent role. Our data suggest that some tissues, such as endoderm and notochord are more sensitive to a hypoxic environment (**Figure 49A**) and intriguingly, their hypoxia-mediated transcriptional response is likely cell type-specific (**Figure 49B**).
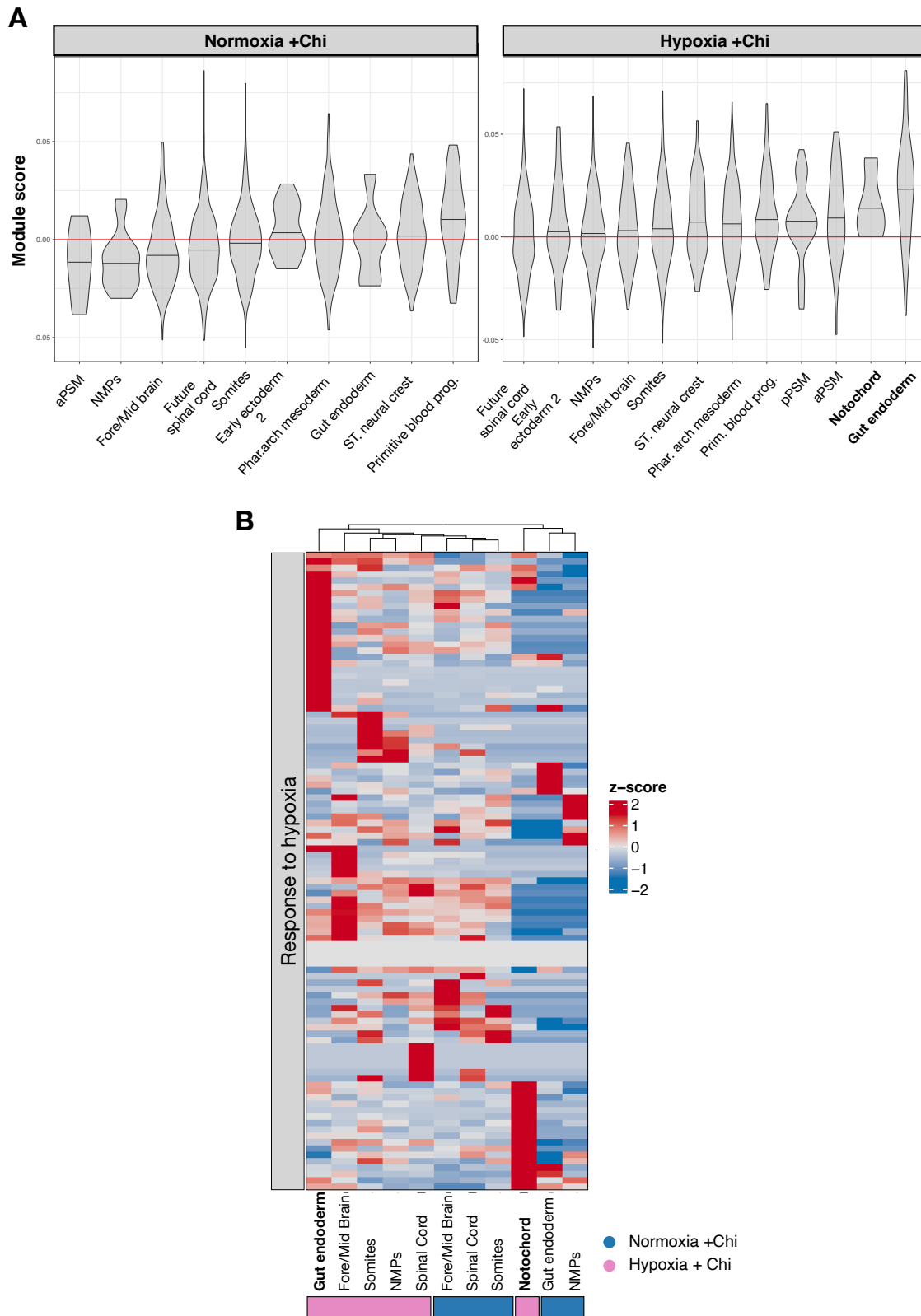
... responsive to hypoxia.

... s found in Normoxia+Chi and Hypoxia+Chi
... e. (B) Heatmap showing pseudo-bulk z-score

... ere, by using a scRNA-seq reference atlas
... ssect and classify the hypoxic profile of the distinct

cell types over time. Preliminary results show that in vivo, endoderm and notochord tissues at E8.5 are more responsive to hypoxia than other tissues, such as the epiblast (**Figure 50**). On the other hand, fore/midbrain and NMP cell types are not responsive to hypoxia (**Figure 50**) as well as in the gastruloids model, where hypoxia seems to barely affect the hypoxic profile of fore/midbrain and NMP cells (**Figure 49A**).
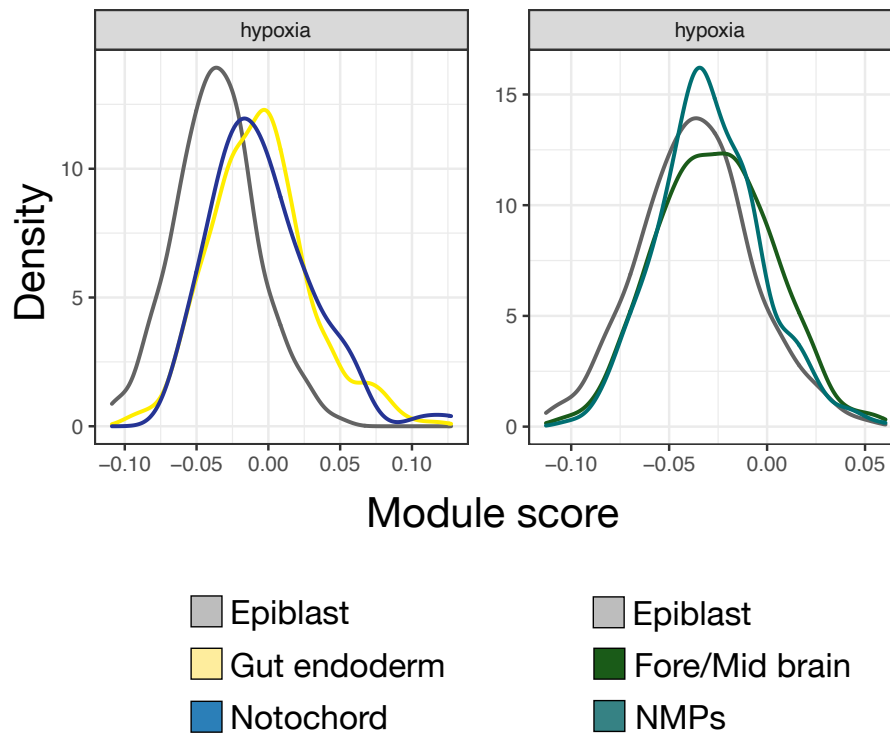
Figure 50. In vivo ... **hypoxia**.

Density plot showi... ... ...s o... embryo (E8.5).

A complete classification using a curated hypoxia set of genes will provide a near complete picture of the lineage roots that are more affected or modulated by the hypoxic environment or if on the other hand, hypoxia acts in a general manner, affecting all cell types similarly.

One of the drawbacks of gastruloids is the lack of architecture in the structures (Veenvliet et al.) (see '*Bringing gastruloids into shape*' section in the introduction for more information). Recent advances have shown that matrigel provides biochemical and mechanical cues providing embryo-like architecture –the so-called TLS— which are gastruloids with somites-like structures juxtaposed to a neural tube-like structure (Veenvliet et al., 2020a). Nonetheless, TLS formation presents high levels of variability (for instance, 40% of a generation presents neural bias, with TLS resulting in only a neural tube-like structure and absence of somites)

(Veenvliet et al., 2020a). My thesis results show that hypoxia can act as a factor providing architecture as well as enhancing the formation of somitic cells. I propose that hypoxia should be harnessed in the TLS model to possibly increase the robustness of the generation by reducing levels of variability and enhancing the formation and organization of somitic cells. Preliminary results show that hypoxia drastically modifies TLS morphology when their formation occurs in hypoxia (**Figure 51**).
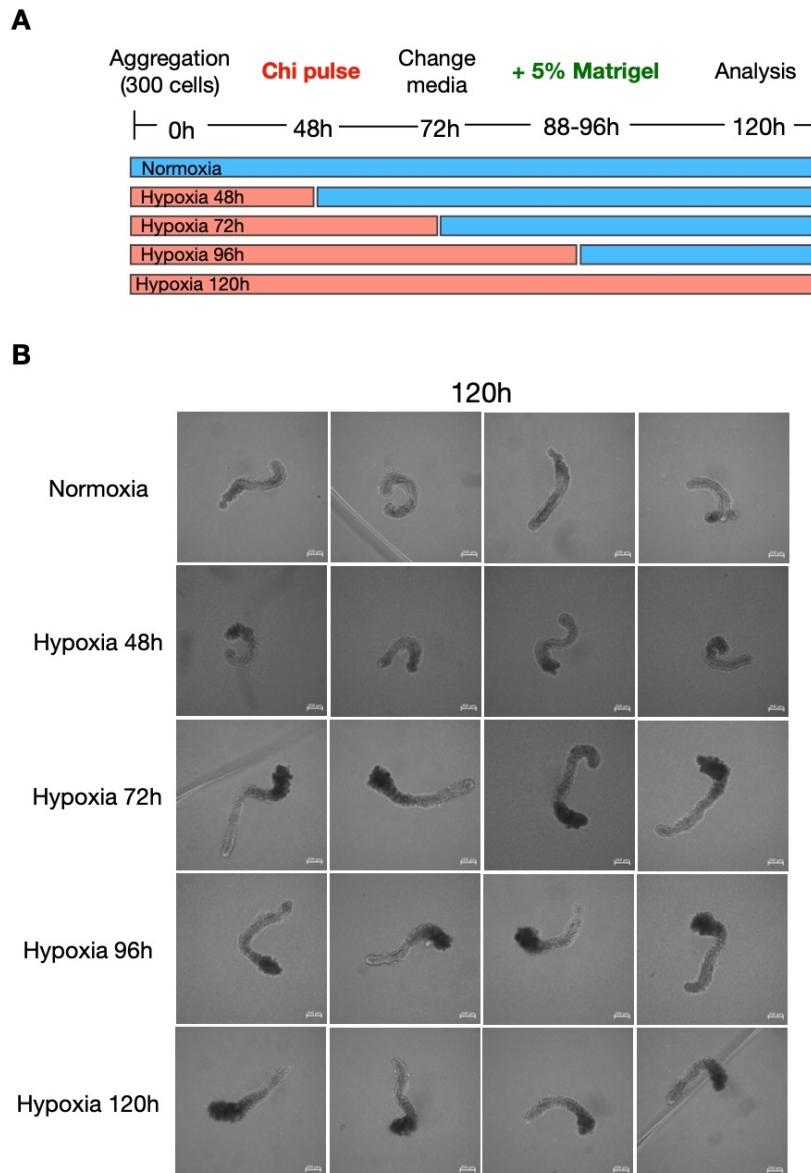
Figure 51. Hypoxia drastically affects the morphology of TLS.

(A) Schematic representation of the experimental setup. 300 ES cells were used for aggregation. After 48h, a 24h Chi pulse allows polarization and symmetry breaking. Around 88-96h, 5% Matrigel is added to the media to allow the formation of TLS. (B) Bright-field images of TLS at 120h post-aggregation at the indicated experimental conditions.

In general, TLS robustness is increased: all structures (1) elongate (compared to a ~9% non-elongated structures in normoxia) (**Figure 51; Figure 52A, B**) and (2) present a neural tube-like organization surrounded by self-organized somitic cells (**Figure 51**).
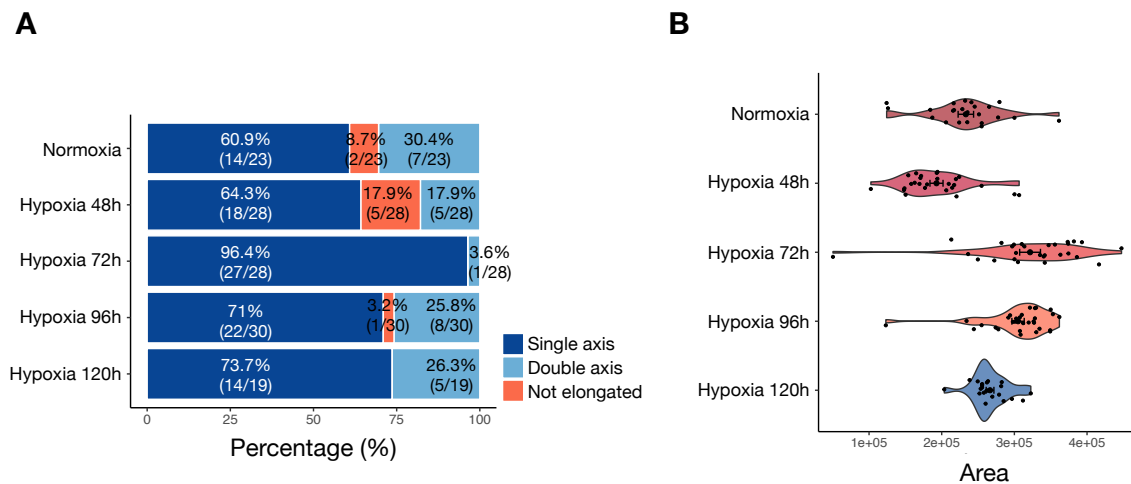
Figure 52. Modulation of oxygen levels during TLS formation enhances the efficiency of their output.

(A) Stacked bar plot showing the number and percentage of distinct TLS morphologies at the indicated experimental conditions. (B) Area of TLSs at the indicated experimental conditions.

In my thesis, hypoxia remains constant at 2% oxygen levels during gastruloid formation. Nonetheless, in vivo embryos are exposed to variable levels of oxygen. Additionally, ex-utero culturing of mouse embryo has shown that a gradual increase of oxygen from E6.5 to E10.5 increased the survival of embryos reaching E10.5 (Aguilera-Castrejon et al., 2021). Therefore, it will be important to investigate how the gradual increase of oxygen during gastruloid/TLS formation impacts its generations in terms of variability, morphology, cellular composition and self-organization into tissue-like structures. Initial analysis shows that modulation of hypoxia timing during TLS formation also affects the robustness and morphology of the generation. For instance, hypoxia during the first 48h decreased the number of structures generated and those were biased toward neural tube with nearly complete depletion of somitic cells. On the other hand, hypoxia for the first 72h significantly increased the robustness of the generation and reduced the heterogeneity of the structures within a generation (**Figure 51B; Figure 52A**).

Overall, our data show the relevance of using physiologically relevant oxygen levels in models of embryo development. I firmly believe hypoxia is a potent modulator during early embryogenesis and it should be harnessed in stembryogenesis to better understand (1) its individual role and (2) the crosstalk with other layers of regulation during early post-implantation and gastrulation of the mouse embryo.

5. Methods

5.1. Cell line and culture conditions

Wild type E14 mouse ES cells (Sarah Kinkley Lab) were cultured in a feeder-independent context in plates coated with gelatin (0.1% gelatin, Sigma-Aldrich G1393). The ES cell culture media was changed every day and cells were passaged to new plates every 2 days in a 1:6 ratio. At each passage, TrypLE (Thermo Fisher 12604-021) was used for dissociation and cells were re-plated.

TS cells (Magdalena Zernicka-Goetz Lab) were culture in a feeder dependent context with TS cell culture media. This media was changed every day and cells were passaged every 4 to 6 days in a 1:10 to 1:20 ratio.

XEN cells (Magdalena Zernicka-Goetz Lab) were cultured on feeders and 0.1% gelatin-coated plates with XEN cell culture media. Media was changed every day and cells were passaged in 1:20 to 1:40 ration every 2 to 3 days.

Cells in normoxia were cultured with 20% oxygen while cells in hypoxia were cultured with 2% oxygen unless otherwise specified for experimental purposes.

5.2. Media and supplements

ES cell culture media: DMEM/High glucose with Glutamax media (Thermo, 31966047) supplemented with 15% FBS (Thermo, 2206648RP), 1x non-essential amino acids, 1x Penicillin/streptomycin, 1x ß-mercaptoethanol, 1000U/mL LIF (homemade).

TS cell culture media: RPMI 1640 (Gibco, 61870010) supplemented with 1x Glutamax, 20% FBS, 1 mM sodium pyruvate (Thermo, 11360070), 1.3x ß-mercaptoethanol, 1x Penicillin/streptomycin, 25 ng/ml FGF4 (R&D systems, 235-F4-025) and 1 µg/ml Heparin (Sigma, H3149).

XEN cell culture media: DMEM + GlutaMAX (Gibco, 10566016) supplemented with 15% FBS, 1x non-essential amino acids, 1x Penicillin/streptomycin, 1x ß-mercaptoethanol, 1 mM Sodium pyruvate, 20 mM HEPES (Sigma, H0887). Media was conditioned on mouse embryonic fibroblasts for 3 days and filtered before use. Then, feeder-conditioned media was mixed in a 7:3 ratio with unconditioned XEN media.

5.3. Inhibitor treatments

For chemical activation of Hif1α, cells were cultured in normoxia and treated with 100µM IOX2 (Sigma-Aldrich, SML0652) dissolved in DMSO (Sigma-Aldrich, D2650).

For DNA demethylation experiments, cells were cultured in normoxia and treated with 1µM Dnmt1 inhibitor (MedChem Tronica, HY-135146) dissolved in DMSO.

For WNT pathway inhibition, cells were treated with either 10 µM XAV939 (Sigma X3004) or 1 µM LGK-974 (Biotrend 331-21058-1) dissolved in DMSO (Sigma-Aldrich, D2650). The inhibition started at aggregation day (0 hours) and ended at timepoint 120h during gastruloid formation process.

5.4. RNA extraction and RT-qPCR

In order to prevent the phenotype from being influenced by cell density, cells were cultured at comparable confluency and matching densities. For time-course hypoxia analysis, corresponding normoxia samples were gathered. Total RNA was extracted using the Qiagen RNeasy kit (Qiagen, 74004), following the manufacturer's instructions. 5 µg of RNA was used as input to generate complementary DNA (cDNA) with the High-Capacity cDNA Reverse Transcription Kit (Thermo, 4368814) using random primers. As a control, a reaction without reverse transcriptase enzyme (-RT control) was performed. cDNA was diluted 1:5 and was used to perform real-time quantitative PCR (qPCR) using primers designed to amplify specific target genes (Table 3) using KAPA SYBR FAST qPCR Master Mix (2X) ABI Prism (Thermo, KK4617) on the QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems) thermal cycler. As a control, a reaction without cDNA was performed. In RT-qPCR experiments, data represent log$_2$FC over normoxia normalized to Eif1a and standard deviation for two biological replicates, and a two-tailed paired Student's t-test is applied, unless otherwise indicated.

5.5. Western blotting
5.5.1. Whole-cell extracts

Cell pellets were resuspended in RIPA buffer (Thermo, 89900) containing 1x protease inhibitor cocktail (Thermo, 78425), incubated at 4°C for 30 minutes and afterwards centrifugated at maximum speed at 4°C for 20 minutes. The supernatant was collected and protein concentration was determined using Pierce™ BCA™ Protein-Assay (Thermo, 23225). 10-20 µg of protein was used for subsequent steps.

5.5.2. Subcellular fragmentation of cytoplasm and nucleus

Cell were washed with buffer A (10mM HEPES pH7.9 (Gibco, 15630-080), 5mM $MgCl_2$ (Sigma, M8266), 0.25M Sucrose (Sigma, S7903), 0.1% Igepal 630 (Merck, 56741), 1x protease inhibitor cocktail, 1mM PMSF, 1mM $NaVO_3$ and incubated on ice for 10 minutes. Then, samples were passed through 18G needles and centrifuged again for 10 minutes. Pellet corresponding to the nuclear fraction was resuspended in cold buffer B (10mM HEPES pH7.9, 1mM $MgCl_2$, 0.1mM EDTA (Jena Bioscience, BU-105), 25% glycerol (Sigma, G5516), 0.5M NaCl (Invitrogen, AM9760G) and 0.5% Triton X-100) and incubated on ice for 30 minutes. Samples were passed through 18G needles and sonication was performed using Bioruptor with the settings "30 seconds on 30 seconds off" for 5 minutes with samples on ice. Subcellular fractions were quantified using Pierce™ BCA™ Protein-Assay. 10-20 µg protein was used for subsequent steps.

5.5.3. SDS-PAGE

Samples were mixed with 4x ROTI loading buffer (Carlroth, K929.2), heated at 98°C for 5 minutes, and loaded on 4–15% Mini-PROTEAN®TGX™ precast protein gels (Biorad, 4561083). Proteins were separated by electrophoresis at 70V for 15 minutes followed by 100V for approximately 1 hour using 10x Tris/Glycine/SDS running buffer (BioRad, 1610772).

5.5.4. Blotting and detection

Proteins were transferred to a PVDF membrane (Thermo, IB24001) using the iBlot 2 dry blotting system (Thermo, IB21001) and run at 20V for 7 minutes. After blotting, membranes were blocked with 5% milk in TBS-T buffer (Thermo, 28360) at RT for 1 hour. To detect protein of interest, membranes were incubated at 4°C overnight with primary antibody (Table 4) in 5% milk in TBS-T buffer, followed by secondary antibody at RT for 1 hour. For detection, membranes were incubated with ECL Western Blotting Substrate (Thermo, 32106) for 1 minute prior to imaging with a ChemiDoc system (BioRad).

5.6. Bulk RNA-sequencing
5.6.1. Sample collection

Cells were cultured at comparable confluency and matching densities to avoid contribution of different densities to the phenotype. Total RNA was extracted using the Qiagen RNeasy kit following the manufacturer's instructions and quantified using Qubit 3.0.

5.6.2. Library preparation and sequencing

Library preparation was performed using KAPA RNA HyperPrep Kit (Kapa Biosystems, KR1350) (input amount 500ng RNA, adapter concentration 1.5 µM and PCR cycle number = 10)), and samples were sequenced using Illumina HiSeq4000, in 75-bp, paired-end format.

5.6.3. Analysis in *"bioinformatic analysis section"*

5.7. Global proteomics
5.7.1. Sample preparation

Samples for global proteomics were prepared according to a published protocol with minor modifications (Kulak et al., 2014). Briefly, 2 biological replicates were used and for each sample 5 million cells of 2 days and 7 days in hypoxia and normoxia controls were lysed under denaturing conditions in 500 µl of a buffer containing 3 M guanidinium chloride (GdmCl), 10 mM tris(2-carboxyethyl)phosphine, 40 mM chloroacetamide, and 100 mM Tris-HCl pH 8.5. Lysates were denatured at 95°C for 10 min shaking at 1000 rpm in a thermal shaker and sonicated in a water bath for 10 min. 100 µl lysate was diluted with a dilution buffer containing 10% acetonitrile and 25 mM Tris-HCl, pH 8.0, to reach a 1 M GdmCl concentration. Then, proteins were digested with LysC (Roche, Basel, Switzerland; enzyme to protein ratio 1:50, MS-grade) shaking at 700 rpm at 37°C for 2 hours. The digestion mixture was diluted again with the same dilution buffer to reach 0.5 M GdmCl, followed by tryptic digestion (Roche, enzyme to protein ratio 1:50, MS-grade) and incubation at 37°C overnight in a thermal shaker at 700 rpm. Peptide desalting was performed according to the manufacturer's instructions (Pierce C18 Tips, Thermo Scientific, Waltham, MA). Desalted peptides were reconstituted in 0.1% formic acid in water and further separated into four fractions by strong cation exchange chromatography (SCX, 3M Purification, Meriden, CT). Eluates were first dried in a SpeedVac, then dissolved in 5% acetonitrile and 2% formic acid in water, briefly vortexed, and sonicated in a water bath for 30 seconds prior injection to nano-LC-MS/MS.

5.7.2. Run parameters

LC-MS/MS was carried out by nanoflow reverse-phase liquid chromatography (Dionex Ultimate 3000, Thermo Scientific) coupled online to a Q-Exactive HF Orbitrap mass spectrometer (Thermo Scientific), as reported previously (Ni et al., 2019). Briefly, the LC separation was performed using a PicoFrit analytical column (75 µm ID × 50 cm long, 15 µm Tip ID; New Objectives, Woburn, MA) in-house packed with 3-µm C18 resin (Reprosil-AQ Pur, Dr. Maisch, Ammerbuch, Germany).

5.7.3. Peptide analysis

Raw MS data were processed with MaxQuant software (v1.6.10.43) and searched against the mouse proteome database UniProtKB with 55,153 entries, released in August 2019.

5.7.4. Analysis in *"bioinformatic analysis section"*

5.8. ChIP-sequencing
5.8.1. Sample preparation

ES cells cultured in normoxia, hypoxia for 2 days (2d) and hypoxia for 6 days (6d) were used. Chromatin extracts were prepared as described in (Brookes et al., 2012). Briefly, cells were fixed with 1% formaldehyde (Thermo, 28906) in PBS at RT for 10 minutes. Then, 0.125M glycine (Sigma, 50046) was added to quench formaldehyde at RT for 5 minutes. Cells were washed twice with ice-cold PBS. To lyse, fixed cells were treated with swelling buffer (25mM HEPES pH 7.9, 1.5mM $MgCl_2$, 10mM KCl (Invitrogen, AM9640G), 0.1% Igepal 630, 1x protease inhibitor cocktail, 1mM PMSF, 2mM $NaVO_3$, 5mM NaF) at 4°C for 10 minutes. Afterwards, cells were scraped on ice and passed through 18G needles before centrifugation at 3000g, 4°C for 5 minutes. Cell pellet, corresponding to nuclei was carefully resuspended in sonication buffer (50mM HEPES pH 7.9, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% Sodium-deoxycholade (Thermo, 89904), 0.1% SDS (Invitrogen, AM9822), 1mM PMSF, 2mM $NaVO_3$, 5mM NaF) and incubated on ice for 10 minutes before sonication. Chromatin was sheared to an average size of 200-300bp with E220 Evolution Covaris sonicator for 6 cycles, 1 minute each. Shearing efficiency was checked by agarose gel.

5.8.1.1.   HIF1α ChIP

10µg chromatin was incubated with 5µL of HIF1α antibody (concentration not provided) (CST, 36169, lot:2), 20 µL of Protein A/G dynabeads (Thermo, 10002D/10004D) in sonication buffer overnight. Beads were washed twice with sonication buffer, followed by a wash with high salt buffer (sonication buffer with 500mM NaCl instead) and TE buffer (10mM Tris-HCl pH 8.5, (Teknova, T5085), 1mM EDTA, 1% SDS). Then, the sample was resuspended in elution buffer (50mM Tris-HCl pH 7.5 (Sigma, T2319), 1mM EDTA, 1% SDS). Samples were treated with RNase A (Thermo, EN0531) at 37°C for 30 minutes, followed by Proteinase K (NEB, P8107S) at 65°C overnight. Genomic DNA was purified using MinElute PCR purification kit (Qiagen, 28004) and quantified by Qubit 3.0. Two biological replicates were collected per condition.

5.8.1.2. H3K4me3 and H3K27me3 quantitative ChIP

Chromatin was extracted as explained above. Additionally, 2% vol/vol *Droshophila* spike-in chromatin (Active Motif, 53083) was added. Based on preliminary experiments to determine the saturation point, 4 µL of H3K4me3 (concentration not provided, Millipore, 05-745, lot: 3158071) and 10 µg of H3K27me3 (Millipore, 07-449, lot: 3517402) antibodies were used for 5ug of chromatin, respectively.

5.8.2. Library preparation and sequencing

Libraries were prepared using the KAPA Hyper Prep Kit (Kapa Biosystems, KR0961) (input amount 10ng DNA, adapter concentration 1.5 µM, and size selection of 200-700-bp after PCR with cycle number = 15). Samples were sequenced using NovaSeq 6000, in 100-bp, paired-end format.

5.8.3. Analysis in *"bioinformatic analysis section"*

5.9. Simultaneous determination of cytidine modifications by targeted LC-MS/MS
5.9.1. Sample preparation

Genomic DNA was extracted using the PureLink Genomic DNA Mini Kit following manufacturer's instructions. 300 ng genomic DNA were used for profiling the following cytidine modifications in E14 mouse ES cells samples: 2′-deoxycytidine (dC), 5-methyl-2′-deoxycytidine (5-mdC), 5-hydroxymethyl-2′-deoxycytidine (5-hmdC), 5-formyl-2′-deoxycytidine (5-fodC), and 5-carboxyl-2′-deoxycytidine (5-cadC), as well as the other three bases 2′-deoxyguanosine (dG), 2′-deoxyadenosine (dA), thymidine (T). A multiple reaction monitoring (MRM) method with three transitions was established by using pure compounds. Furthermore, a dilution series of these standards was used for absolute quantification. 300 ng DNA was dissolved in a total of 21.5 µL water in Protein LoBind Tubes, 2.5 µL DNA Degradase buffer and 1 µL DNA Degradase Plus enzyme (Zymo Research, Freiburg, Germany) were added. The DNA digestion efficiency was monitored by using 100 ng 5-methylcytosine and 5-hydroxymethylcytosine DNA standard sets (Zymo Research) in parallel under identical conditions. The digest was carried out at 37°C on a rocking platform (600 rpm) for 2 hours. 1 µL of 100 µM chloramphenicol was added as internal standard as well as 1 µL of 5% formic acid to inactivate the digest.

5.9.2. Run parameters

180 ng of the DNA digests were used for LC-MS/MS analysis. Cytidine separation was performed on an LC instrument (1290 series UHPLC; Agilent, Santa Clara, CA), online coupled to a triple quadrupole hybrid ion trap mass spectrometer QTrap 6500 (Sciex, Foster City, CA). Cytidines were eluted from a Reprosil-PUR C18-AQ (1.9 µm, 120 Å, 150 × 2 mm ID; Dr. Maisch; Ammerbuch, Germany) column at a controlled temperature of 30 °C, using a gradient from 2 to 98% solvent B in solvent A over 10 min at 250 µL per minute flow rate. Solvent A was 10 mM ammonium acetate, pH 3.5 (adjusted with acetic acid) and solvent B was 0.1% formic acid in acetonitrile. Transitions were monitored in a 240 second window of the expected elution time and acquired at unit resolution (peak width at 50% was 0.7 ± 0.1 Da tolerance) in quadrupole Q1 and Q3. Data acquisition was performed with an ion spray voltage of 5.5 kV in positive mode of the ESI source, N2 as the collision gas was set to high, curtain gas was set to 30 psi, ion source gas 1 and 2 were set to 50 and 70 psi, respectively, and an interface heater temperature of 350 °C was used.

5.9.3. Analysis

Relative quantification of the peaks was performed using MultiQuant software v.2.1.1 (Sciex, Foster City, CA). The integration settings were a peak-splitting factor of 5 and a Gaussian smoothing width of 2 All peaks were reviewed manually. Only the average peak area of the first transition was used for calculations. Normalization was done according to IS. Results were plotted using GraphPad Prism (version 8).

5.10. Whole genome bisulfite sequencing (WGBS)
5.10.1. Sample preparation

Genomic DNA was extracted from 1 million cells using the PureLink Genomic DNA Mini Kit (Thermo, K182002) following manufacturer's instructions. gDNA was sheared in Covaris micro TUBE AFA Fiber Pre-Slit Snap-Cap tubes (SKU, 520045) and purified with the Zymo DNA Clean & Concentrator-5 kit (D4013) according to manufacturer's instructions. The purified DNA was then bisulfite converted using the EZ DNA Methylation-Gold Kit (Zymo, D5005).

5.10.2. Library preparation and sequencing

WGBS libraries were processed using the Accel-NGS Methyl-seq DNA library kit (Swift Biosciences, DL-ILMMS-12) following manufacturer's recommendations. Libraries were

prepared and cleaned using Agencourt AMPure XP beads (Beckman Coulter, A63881). The absence of adapters from the final libraries was verified using the Agilent TapeStation. The WGBS libraries were sequenced on the NovaSeq6000, in 150-bp, paired-end format.

5.10.3. Analysis in *"bioinformatic analysis section"*

5.11. Gastruloid formation

Male F1G4 mouse ES cells (George et al., 2007) were cultured on 6 cm plates (Corning 430166) with 0.1% gelatin and coated in a feeder dependent context with ES cell culture media. ES cells were cultured as described in *"Cell line and culture conditions"* section.
Gastruloids were then generated as described previously (Veenvliet et al., 2020b) with some minor modifications. Briefly, ES cells were first depleted of feeders, then washed once in 5ml pre-heated (37°C) PBS containing MgCl2 and CaCl2 (Sigma D8662) and once in 5ml NDiff227 medium (Takara Y40002) pre-conditioned in normoxia or hypoxia. ESCs were then pelleted by centrifugation for 5 minutes at 1000rpm and resuspended in 250µl of NDiff227. 10µl of the cell suspension was mixed with 10µl of Trypanblue (Bio-Rad 1450021) for automated cell counting with Luna Automated Cell Counter. 400 live cells were plated in a volume of 35µl NDiff227 into each well of a 96-well round bottom, low attachment plate (Costar 7007 ultra-low attachment 96 well plate (7007)). Cells were then allowed to aggregate for 48 hours under normoxic or hypoxic conditions. After 48h, the aggregates were treated with 3µM CHIR99021 (CHIR, Merck Millipore) in 150µl NDiff227 for 24h to induce robust gastruloid formation. For "-Chi" aggregates, 150µl NDiff227 without CHIR was added. Between 72 and 120h, medium was chaged every 24h by removing 150µl of the old media and adding the same volume of new, pre-conditioned NDiff227. Hypoxic gastruloids were formed at 2% O_2.

5.12. Trunk-like structure formation

Male F1G4 mouse ES cells (George et al., 2007) were cultured on 6 cm plates (Corning 430166) with 0.1% gelatin and coated in a feeder dependent context with ES cell culture media. ES cells were cultured as described in *"Cell line and culture conditions"* section. TLS were then generated as described previously (Bolondi et al., 2021a; Veenvliet et al., 2020b) with some minor modifications and as described for gastruloid formation. The only differences were that 300 live cells were plated for aggregation and Matrigel (Corning 356231, 5% final concentration) was added at 88-96h post-aggregation when the majority of the structures showed the 'teardrop' shape.

5.13. Imaging analysis

Brightfield images from individual gastruloids were used for morphometric analysis. For images acquired using the Olympus IXPlore SpinSR, Z-stacks were first patched into a single focused image using the Gaussian-based stack focuser plugin. Images were then auto-thresholded using Otsu algorithm and binary processed to fill holes. Manual thresholding was carried out in samples where auto-thresholding failed. Gastruloids were segmented, and masks with registered regions-of-interests (ROIs) were created. For generation of the movie, the projected mCherry images' intensity range of display was set to a minimum of 0 and a maximum of 400, before being merged with the corresponding focused brightfield image. All bioimage analyses were executed using Fiji v1.53s.

5.14. Multiplexed single-cell RNA-sequencing

Single-cell transcriptome profiling of gastruloids using MULTI-seq was performed by combining the protocols described in (Bolondi et al., 2021b; McGinnis et al., 2019b; Veenvliet et al., 2020b) with minor modifications.

5.14.1. Sample preparation

Gastruloids in NN+CHI, NH+CHI and NH-CHI conditions were generated as described in *"Gastruloid formation"* section. From each condition, 28 elongated structures were pooled together. Gastruloids were picked with a p200 pipette with the pipette tip cut-off at the 50µl mark, and pooled in a 1.5 ml tube filled with ice-cold PBS. The three pools were then washed twice with ice-cold PBS. Next, structures were dissociated in 50 µl TrypLE Express for 20 minutes at 37ºC, with pipetting after 10 minutes.

5.14.2. Library preparation and sequencing

While gastruloid structures dissociation (described aboved), a 10x lipid-modified oligonucleotides (LMOs) – Barcode (BC) oligo solution was prepared (1:1; 2 µM each) in PBS for each of the three samples on ice. From here on, every step has been performed on ice. After cell dissociation, LMOs–BC solutions were added to each sample separately (final concentration 200 nM), and samples were incubated for 5 min on ice. In the meantime, a 10x co-anchor solution was prepared (2µM) in PBS. Following the 5 min incubation, the co-anchor solution was added to each sample (final concentration 200 nM), and samples were again incubated for 5 min on ice. The reaction was quenched by adding 1 ml PBS + 1% BSA to each

tube. Cells were washed twice with 1 ml PBS + 1% BSA with centrifugation steps performed for 5 minutes at 300 g and 4°C in low DNAbind Eppendorf tubes. Cell pellets were resuspended in 100 µl PBS + 0.4% BSA and pooled together in one new tube. The cell suspension was filtered using Scienceware Flowmi Cell Strainers, 40 µm. Cells where then centrifuged for 5 minutes at 300 g at 4°C, and resuspended in 45 µl PBS+0,4% BSA and the cell concentration was determined using a hemocytometer. Cells were subjected to single-cell RNA sequencing (10x Genomics, Chromium™ Single Cell 3' v3; one reaction). Single-cell libraries were generated according to manufacturer's instructions, with one modification: Multi-seq additive primer (5'-CTTGGCACCCGAGAATTCC-3') was added at the cDNA amplification step (see (McGinnis et al., 2019b) for detailed protocol). During amplified cDNA cleanup, the Multi-seq BCs fraction was isolated and a separate library was performed as described. cDNA library was sequenced with a minimum of 400 million fragments and Multi-seq BCs library was sequenced with a minimum of 50 million fragments.

5.14.3. Analysis in *"bioinformatic analysis section"*

5.15.   Bioinformatic analyses
5.15.1. Bulk RNA-sequencing
5.15.1.1.   Mapping and primary analysis

Raw reads were subjected to adapter and quality trimming with cutadapt (version 2.4; parameters: --nextseq-trim 20 --overlap 5 --minimum-length 25 --adapter AGATCGGAAGAGC -A AGATCGGAAGAGC), followed by poly-A trimming with cutadapt (parameters: --overlap 20 --minimum-length 25 --adapter "A[100]" --adapter "T[100]"). Reads were aligned to the mouse reference genome (mm10) using STAR (version 2.7.5a; parameters: --runMode alignReads --chimSegmentMin 20 --outSAMstrandField intronMotif --quantMode GeneCounts) and transcripts were assembled using StringTie (version 2.0.6; parameters: -e) with GENCODE annotation (VM19). Table of read counts was generated using featureCounts function from the Rsubread package (version 1.32.4). To filter lowly expressed genes, count-per-million (CPM) was calculated using EdgeR package (version 3.26.8). Genes having > 0.4 CPM in at least 2 out of 6 samples were kept. Additionally, only coding protein genes were used for further analysis. Transcripts Per Kilobase Million (TPM) was used to generate normalized expression values. Differentially expressed genes (DEG) were determined using EdgeR package, applying multiple-testing adjusted p-value (FDR) <= 0.1 significance threshold and absolute fold-change > 1.5.

Principal component analysis (PCA) was performed with the built-in R functions "prcomp" and hierarchical clustering was made with dendextend library (version 1.14.0) using "dist" function

to compute distance between sample and hclust to perform hierarchical clustering. PCA and volcano plots were generated using ggplot2 (version 3.3.0). Heatmaps were generated using ComplexHeatmap package (version 2.11.1).

5.15.1.2. Selection of germ layers and lineage markers

Marker genes from different cell states were selected following the single-cell transcriptional reference profile of early post-implantation development from (Grosswendt et al., 2020b). Early ecto-, endo- and mesoderm germ layers represents marker genes that were reliably assigned to cell states within these three lineage cell fates from E6.5 to E7.5.

For later cell states, lineage specific marker genes from node & notochord (E7.0-8.0, mesoderm lineage), neuromesodermal progenitors (E7.5-8.0, mesoderm lineage), paraxial & posterior mesoderm (E7.5-8.5, mesoderm lineage), primitive gut (E7.5-8.5, endoderm lineage) and for mid brain & spinal cord (E7.5-8.5, ectoderm lineage) were selected.

5.15.1.3. Gene ontology-biological process (GO-BP) analysis

GO-BP analyses were performed using DAVID web-tools (Dennis et al., 2003). Selected terms were plotted using ggplot2.

5.15.1.4. Analysis of T target genes

To identify T target genes which are either activated or repressed in the presence of T binding, publicly available data from T ChIP-seq and RNA-seq datasets were retrieved from (Lolas et al., 2014b). Briefly, T-activated (FC>2) or T-repressed genes (FC<-2) in *in vitro* primitive streak differentiated cells on day 4 vs. ES cells were selected. Expression levels of these T target genes in hypoxia day 2, day 7 and normoxia was plotted in heatmap using pheatmap package (version 1.0.12) showing row z-score normalized expression values. Annotation bars were added to show which of these genes are differentially expressed in hypoxia day 2 and/or day 7 compared to normoxia.

5.15.2. Global proteomics

To identify the differentially expressed proteins in hypoxia compared to normoxia, the DEP package (version 1.6.1) was used. First, duplicate proteins, contaminants, and proteins that were not found in at least 2 out of the total number of samples (n=6) were filtered out. A total number of 4,260 proteins were identified. The LFQ values were normalized using the

background correction variance stabilizing transformation (VSN). Missing values were imputed using a left-censored imputation method as the proteins with missing values were biased to low expression values. Differential expression analysis was run with test_diff function from DEP package which uses limma (Ritchie et al., 2015). Differentially expressed proteins were defined applying multiple-testing adjusted P (Benjamini-Hochberg) < 0.05 significant threshold and fold-change with an absolute value of > 1.5. Scatter plots were generated to show the mean LFQ values in hypoxia versus normoxia and were generated using ggplot2. GO-BP analysis was performed using DAVID web tools (Dennis et al., 2003).

5.15.3. ChIP-sequencing
5.15.3.1.   Mapping and analysis: HIF1α ChIP-sequencing

Raw reads of treatment and input samples were subjected to adapter and quality trimming with cutadapt (Martin, 2011) (version 2.4; parameters: --nextseq-trim 20 --overlap 5 --minimum-length 25 --adapter AGATCGGAAGAGC -A AGATCGGAAGAGC). Reads were aligned to the mouse genome (mm10) using BWA with the 'mem' command (version 0.7.17, default parameters). A sorted BAM file was obtained and indexed using samtools with the 'sort' and 'index' commands (version 1.10). Duplicate reads were identified and removed using GATK (version 4.1.4.1) with the 'MarkDuplicates' command and default parameters. Peaks were called with reads aligning to the mouse genome only using MACS2 'callpeak' (version 2.1.2; parameters --bdg --SPMR) using the input samples as control samples. For downstream analyses, after validation of reproducibility, replicates were pooled using samtools 'merge'. Genome-wide coverage tracks (signal files) for merged replicates normalized by library size were computed using samtools bamCoverage (version 3.4.3) (parameters: --normalizeUsing CPM --extendReads) When required, bedgraph files were also generated using bigWigToBedGraph from Kent utils tools (Kent et al., 2010). For identification of consistent HIF1α peaks, only those that were identified in both replicates were used for downstream analyses. Peaks were annotated using ChIPseeker package (version 1.20.0) using default parameters (TSS region ±3-Kb) and subdivided into peaks at promoter or distal regions. Density plots were generated using computeMatrix (reference-point --referencePoint center) and plotHeatmap functions from Deeptools. Enrichment of different features was at ± 8 kb of centered HIF1α peaks at promoters or distal regions. H3K27ac, p300, H3K4me3 and H3K4me1 ChIP-seq data generated on E14 ESCs were retrieved from (Cruz-Molina et al., 2017b).

Genome view of selected loci were generated using SparK (version 2.6.2). ß-CATENIN ChIP-seq data generated on V6.5 ESCs was retrieved from (Zhang et al., 2013b) and analyzed following our ChIP-seq workflow.

Overlap of HIF1α-, ß-CATENIN- and ES enhancer peaks was generated using intersect function (-wa -wb) from Bedtools (version 2.29.2). Venn diagrams were generated manually.

5.15.3.2.   Mapping and analysis: H3K4me3 and H3K27me3 ChIP-seq

H3K4me3 and H3K27me3 quantitative ChIP-seq analyses were performed as explained above. Additionally, reads were aligned to the Drosophila genome (dm6) using BWA with the 'mem' command (version 0.7.17, default parameters). For spike-in normalization, the percentage of dm6 mapping reads over mm10. Secondly, the ratio of the experimental conditions (hypoxia day 2 and day 7) were calculated over control condition (normoxia). These values were used as scaling factors to quantitatively compare among the three different conditions. See Table 5 for more details. Quantitative genome-wide coverage tracks (signal files) for merged replicates normalized by library size were computed using deeptools bamCoverage (--scaleFactor X parameters: --normalizeUsing CPM --extendReads).

For global quantification of H3K4me3 and H3K27me3, 1-Kb tiles signal files were used to quantify the levels of H3K4me3, H3K27me3 genome-wide. Aditionally, DNA mean methylation 1-Kb tiles signal files (see "WGBS, Genomic feature annotations section) where in a boxplot to show $\log_{10}$ of the density of the different features was generated in R.

For ChromHMM (Ernst and Kellis, 2012) analysis, the pre-defined mouse ESC ChromHMM (https://github.com/guifengwei/ChromHMM_mESC_mm10) was used to map H3K4me3 and H3K27me3 spike-in normalized signal to the different chromatin states. Heatmap was generated using ggplot2 to show the $\log_2$FC of the average spike-in normalized density over normoxia across 11 chromatin states.

Density plots were generated using computeMatrix (reference-point --referencePoint) for TSS and (scale-regions) for gene body resolution and plotHeatmap functions from Deeptools. Enrichment of different features was defined at 1) ±2.5-Kb of DE genes at promoters identified by RNA-seq.

5.15.4. WGBS

Raw reads were subjected to adapter and quality trimming with cutadapt (version 2.4; parameters: --quality-cutoff 20 --overlap 5 --minimum-length 25; Illumina TruSeq adapter clipped from both reads) followed by trimming of 10 nucleotides from the 5' and 3' end of both reads. The trimmed reads were aligned to the mouse genome (mm10) using BSMAP (version 2.90; parameters: -v 0.1 -s 16 -q 20 -w 100 -S 1 -u -R). Duplicates were removed using the 'MarkDuplicates' command from GATK (version 4.1.4.1; --VALIDATION_STRINGENCY=LENIENT --REMOVE_DUPLICATES=true). Methylation rates

were called using mcall from the MOABS package (version 1.3.2; default parameters). Only CpGs covered by at least 10 and at most 150 reads located on autosomes or chromosome X were considered for all downstream analysis. Replicates were merged by calculating the average for each CpG covered by at least one of the two samples.

Genomic feature annotations: Genomic tiles were defined by segmenting the genome into regions of one kb size. The annotation of CGIs was downloaded from UCSC for the mouse reference genome mm10. CGI shores were defined as the two kb regions flanking each CGI on either side. CGI shelves were defined as the two kb flanking the outer sides of CGI shores. Annotation of repeat elements was downloaded from RepeatMasker. The gene annotation was downloaded from GENCODE (VM19) and promoters were defined as the regions 1.5 kb upstream and 500 bp downstream of the TSS.

Feature-wise DNA methylation: Average methylation levels per feature were calculated using the arithmetic mean based on the CpGs commonly covered by all samples (merged replicates). Only features with at least three covered CpGs were considered. For methylation differences per feature, the arithmetic mean of one sample was subtracted by the arithmetic mean of another sample.

Heatmaps of average promoter methylation were prepared using the 'pheatmap' R package. Promoters were classified as overlapping with a CGI if at least 20% of the promoter or CGI were overlapping with each other.

DNA methylation levels at HIF1α peaks: HIF1α peaks at both promoter and distal regions were merged. 1-Kb tiles signal files were used to quantify the mean methylation levels at HIF1α peaks using map function (-c 4 - mean) from Deeptools. Mean methylation levels were used to generate violin plot using ggplot2.

5.15.5. Single-cell RNA-sequencing

All analyses and plots were generated using R version 4.1.0 "Camp Pontanezen" and Seurat (version 4.0.5) (Hao et al., 2021).

5.15.5.1.   Preprocessing of the raw data

Cell Ranger pipeline version 3 (10x Genomics Inc.) was used for scRNA-seq dataset to de-multiplex the raw base call files, generate fastq files, perform the mapping to the mouse (reference genome mm10), filter the alignment and count barcodes and UMIs.

To de-multiplex samples within our single MULTI-seq scRNA-seq dataset, the deMULTIplex R package (version 1.0.2) (McGinnis et al., 2019b) was used. Briefly, sample IDs, "GGAGAAGA" for NN + Chi, "CCACAATG" for NH + Chi and "TGAGACCT" for NH – Chi, were

assigned to cells. Then, cells with no associated sample barcode and cells with more than one barcode (doublets) were discarded for downstream analysis.

5.15.5.2.   Quality control

The initial quality control was performed with Seurat. Single cell data generated was loaded with a minimum requirement of 3 cells and 200 features (default parameters). Cells with less than 200 or more than 2,500 unique feature counts and a mitochondrial fraction above 5% were removed from the analysis. Quality control features were checked for each individual condition.

5.15.5.3.   Data integration and cluster determination

After quality control, the Seurat object was divided by condition (NN + Chi, NH + Chi and NH – Chi). Subsequently, the expression data were independently normalized, variable features within the three datasets were detected, log-normalized and scaled to 10,000 (default settings). For integration of the three conditions, integration anchors within our three data sets were calculated and a PCA (reduction "rpca", dims = 1:30) was run. Finally, these anchors were used to integrate the data. A list of cell cycle markers was used to score for cell cycle stage and to subsequent scale the data with regression out (var.to.regress) of S and G2M phase-related genes (Macosko et al., 2015). For downstream analysis and visualization of the integrated dataset, a PCA followed by a UMAP (dims=1:30, n.neighbors=10) were run. A total number of 10 clusters were identified applying standard parameters (FindNeighbors, dims = 1:20 and FindClusters, resolution = 0.5). These Seurat clusters were further annotated to germ layer origin taking into account cell type annotation to mouse embryo.

5.15.5.4.   Cell type annotation to mouse embryo

To confidently predict the in vivo cell ID of our integrated gastruloid dataset, two independent publicly available mouse reference atlases (from E6.5 to E8.5) were used (Grosswendt et al., 2020b; Pijuan-Sala et al., 2019). These were also used for proportion comparisons. The reference atlases were filtered to include only the relevant time points for our study (gastruloids at 120h resemble the in vivo embryo at E8.5, so we used the closest in time developmental stages E7.5, E8.0 and E8.5). We also excluded all extra-embryonic cell states. The mouse reference atlases and our integrated gastruloid dataset were normalized (SCTransform, with default parameters). Cell type classification was afterwards performed by finding anchors to transfer (dims = 1:30) and adding these predictions to our integrated gastruloid dataset.

Prediction scores and percentage of cells assigned to each cluster were used as measurements for the cell type annotation call (data not shown). Finally, gastruloid cells were matched to their in vivo counterparts. All analyses described below were done with gastruloid cells assigned to their in vivo counterparts according to (Grosswendt et al., 2020b).

5.15.5.5. Neural cell categorization

Neural cells assigned to future spinal cord and fore/mid brain were extracted from the integrated dataset. To classify them, first neural cells were assigned to Hox positive module if expressing at least one Hox gene out of a subset of 28 (Duboule and Dollé, 1989). Among Hox-negative cells, putative brain-like module was defined as cells (1) not expressing any Hox gene, (2) expressing at least five fore/mid brain marker genes out of the top 30 in vivo fore/mid brain marker genes, and (3) not expressing spinal cord marker gene *Fgf3bp* (Grosswendt et al., 2020b). Lastly, spinal cord-like module corresponded to the remain fraction of cells which are Hox-negative and are expressing less than five fore/mid brain marker genes.

5.15.5.6. Dorsal-ventral patterning categorization

Spinal cord assigned cells were extracted from the integrated dataset. A list of marker genes belonging to dorsal, midline and ventral identity (Sagner and Briscoe, 2019b). Ventral module was defined as cells (1) expressing at least one out of three ventral genes and (2) not expressing midline or dorsal markers. Same rule applied for midline and dorsal module assignment. Around 71.5% of spinal cord assigned cells passed the dorsal-ventral patterning categorization.

5.15.5.7. Pseudo-bulk expression analysis

Average gene expression values per cell state and sample were calculated using the "AverageExpression" function from Seurat was used.

5.15.5.8. Module score

Hypoxia related gene list was compiled based on Gene Ontology term 'Response to Hypoxia: GO:0001666' and filtered for Mus musculus. Then, to generate the hypoxia module score, the 'AddModuleScore' function from Seurat package was used.

5.16. Statistical tests

For all distinct statistical tests performed in this study, P-value follows the next style: $p > 0.05$ (ns), $p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***), $p < 0.0001$ (****).

Bibliography

Abbott, M. E. (1936). Men and Books. *Can Med Assoc J.*

Aguilera-Castrejon, A., Oldak, B., Shani, T., Ghanem, N., Itzkovich, C., Slomovich, S., Tarazi, S., Bayerl, J., Chugaeva, V., Ayyash, M., et al. (2021). Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature* 593, 119–124.

Aisenbrey, E. A. and Murphy, W. L. (2020). Synthetic alternatives to Matrigel. *Nat Rev Mater* 5, 539–551.

Amadei, G., Handford, C. E., Qiu, C., Jonghe, J. D., Greenfeld, H., Tran, M., Martin, B. K., Chen, D.-Y., Aguilera-Castrejon, A., Hanna, J. H., et al. (2022). Embryo model completes gastrulation to neurulation and organogenesis. *Nature* 610, 143–153.

Amin, N. D. and Pașca, S. P. (2022). Mouse embryo models built from stem cells take shape in a dish. *Nature* 610, 39–40.

Arias, A. M., Marikawa, Y. and Moris, N. (2022). Gastruloids: Pluripotent stem cell models of mammalian gastrulation and embryo engineering. *Dev Biol*.

Arnold, S. J. and Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Bio* 10, 91–103.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Gene Dev* 17, 126–140.

Bardot, E. S. and Hadjantonakis, A.-K. (2020). Mouse gastrulation: Coordination of tissue patterning, specification and diversification of cell fate. *Mech Develop* 163, 103617.

Batie, M., Frost, J., Frost, M., Wilson, J. W., Schofield, P. and Rocha, S. (2019). Hypoxia induces rapid changes to histone methylation and reprograms chromatin. *Science* 363, 1222–1226.

Beccari, L., Moris, N., Girgin, M., Turner, D. A., Baillie-Johnson, P., Cossy, A.-C., Lutolf, M. P., Duboule, D. and Arias, A. M. (2018). Multi-axial self-organization properties of mouse embryonic stem cells into gastruloids. *Nature* 562, 272–276.

Beck, F., Erler, T., Russell, A. and James, R. (1995). Expression of Cdx-2 in the mouse embryo and placenta: Possible role in patterning of the extra-embryonic membranes. *Dev Dynam* 204, 219–227.

Beddington, R. S. and Robertson, E. J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Dev Camb Engl* 105, 733–7.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* 125, 315–326.

Bolondi, A., Haut, L., Gassaloglu, S., Burton, P., Kretzmer, H., Buschow, R., Meissner, A., Herrmann, B. and Veenvliet, J. (2021a). Generation of Mouse Pluripotent Stem Cell-derived Trunk-like Structures: An in vitro Model of Post-implantation Embryogenesis. *Bio-protocol* 11, e4042.

Bolondi, A., Haut, L., Gassaloglu, S., Burton, P., Kretzmer, H., Buschow, R., Meissner, A., Herrmann, B. and Veenvliet, J. (2021b). Generation of Mouse Pluripotent Stem Cell-derived Trunk-like Structures: An in vitro Model of Post-implantation Embryogenesis. *Bio-protocol* 11, e4042.

Boroviak, T. and Nichols, J. (2014). The birth of embryonic pluripotency. *Philosophical Transactions Royal Soc B Biological Sci* 369, 20130541.

Boroviak, T., Loos, R., Bertone, P., Smith, A. and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. *Nat Cell Biol* 16, 513–525.

Bostick, M., Kim, J. K., Estève, P.-O., Clark, A., Pradhan, S. and Jacobsen, S. E. (2007). UHRF1 Plays a Role in Maintaining DNA Methylation in Mammalian Cells. *Science* 317, 1760–1764.

Bradley, A., Evans, M., Kaufman, M. H. and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* 309, 255–256.

Brickman, J. M. and Serup, P. (2016). Properties of embryoid bodies. *Wires Dev Biology* 6,.

Brink, S. C. van den, Baillie-Johnson, P., Balayo, T., Hadjantonakis, A.-K., Nowotschin, S., Turner, D. A. and Arias, A. M. (2014). Symmetry breaking, germ layer specification and axial organisation in aggregates of mouse embryonic stem cells. *Dev Camb Engl* 141, 4231–4242.

Brink, S. C. van den, Alemany, A., Batenburg, V. van, Moris, N., Blotenburg, M., Vivié, J., Baillie-Johnson, P., Nichols, J., Sonnen, K. F., Arias, A. M., et al. (2020). Single-cell and spatial transcriptomics reveal somitogenesis in gastruloids. *Nature* 582, 405–409.

Brons, I. G. M., Smithers, L. E., Trotter, M. W. B., Rugg-Gunn, P., Sun, B., Lopes, S. M. C. de S., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195.

Brookes, E., Santiago, I. de, Hebenstreit, D., Morris, K. J., Carroll, T., Xie, S. Q., Stock, J. K., Heidemann, M., Eick, D., Nozaki, N., et al. (2012). Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. *Cell Stem Cell* 10, 157–170.

Burr, S., Caldwell, A., Chong, M., Beretta, M., Metcalf, S., Hancock, M., Arno, M., Balu, S., Kropf, V. L., Mistry, R. K., et al. (2017). Oxygen gradients can determine epigenetic asymmetry and cellular differentiation via differential regulation of Tet activity in embryonic stem cells. *Nucleic Acids Res* 46, gkx1197-.

Camus, A., Davidson, B. P., Billiards, S., Khoo, P., Rivera-Pérez, J. A., Wakamiya, M., Behringer, R. R. and Tam, P. P. (2000). The morphogenetic role of midline mesendoderm and ectoderm in the development of the forebrain and the midbrain of the mouse embryo. *Dev Camb Engl* 127, 1799–813.

Chakraborty, A. A., Laukka, T., Myllykoski, M., Ringel, A. E., Booker, M. A., Tolstorukov, M. Y., Meng, Y. J., Meier, S. R., Jennings, R. B., Creech, A. L., et al. (2019). Histone demethylase KDM6A directly senses oxygen to control chromatin and cell fate. *Science* 363, 1217–1222.

Chazaud, C., Yamanaka, Y., Pawson, T. and Rossant, J. (2006). Early Lineage Segregation between Epiblast and Primitive Endoderm in Mouse Blastocysts through the Grb2-MAPK Pathway. *Dev Cell* 10, 615–624.

Choi, B.-J., Park, S.-A., Lee, S.-Y., Cha, Y. N. and Surh, Y.-J. (2017). Hypoxia induces epithelial-mesenchymal transition in colorectal cancer cells through ubiquitin-specific protease 47-mediated stabilization of Snail: A potential role of Sox9. *Sci Rep-uk* 7, 15918.

Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., Ijcken, W. F. J. van, Grosveld, F., Frommolt, P., Bazzi, H., et al. (2017a). PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* 20, 689-705.e9.

Cruz-Molina, S., Respuela, P., Tebartz, C., Kolovos, P., Nikolic, M., Fueyo, R., Ijcken, W. F. J. van, Grosveld, F., Frommolt, P., Bazzi, H., et al. (2017b). PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* 20, 689-705.e9.

Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C. and Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4, R60.

Duboule, D. and Dollé, P. (1989). The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *Embo J* 8, 1497–1505.

Dunn, S.-J., Martello, G., Yordanov, B., Emmott, S. and Smith, A. G. (2014). Defining an essential transcription factor program for naïve pluripotency. *Science* 344, 1156–1160.

Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* 75, 1417–1430.

Ernst, J. and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* 9, 215–216.

Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.

Feldmann, A., Ivanek, R., Murr, R., Gaidatzis, D., Burger, L. and Schübeler, D. (2013). Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions. *Plos Genet* 9, e1003994.

Fischer, B. and Bavister, B. D. (1993). Oxygen tension in the oviduct and uterus of rhesus monkeys, hamsters and rabbits. *Reproduction* 99, 673–679.

Folmes, C. D. L., Dzeja, P. P., Nelson, T. J. and Terzic, A. (2012). Metabolic Plasticity in Stem Cell Homeostasis and Differentiation. *Cell Stem Cell* 11, 596–606.

Fong, C. Y., Bongso, A., Ng, S. C., Kumar, J., Trounson, A. and Ratnam, S. (1998). Blastocyst transfer after enzymatic treatment of the zona pellucida: improving in-vitro fertilization and understanding implantation. *Hum Reprod* 13, 2926–2932.

Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki, J. and Niwa, H. (2002). Differentiation of embryonic stem cells is induced by GATA factors. *Gene Dev* 16, 784–789.

Garg, V., Morgani, S. and Hadjantonakis, A.-K. (2016). Chapter Ten Capturing Identity and Fate Ex Vivo Stem Cells from the Mouse Blastocyst. *Curr Top Dev Biol* 120, 361–400.

George, S. H. L., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens, M. E., Haigh, J. J. and Nagy, A. (2007). Developmental and adult phenotyping directly from mutant embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4455–4460.

Gharibi, B., Gonçalves, E., Nashun, B., Montoya, A., Mankalow, K., Strohbuecker, S., Sheriff, R. S. M., Cicarrelli, A., Carvalho, J., Nye, E., et al. (2020). A FGF2-mediated incoherent feedforward loop induces Erk inhibition and promotes naïve pluripotency. *Biorxiv* 2020.11.11.378869.

Gheldof, A. and Berx, G. (2013). Cadherins and epithelial-to-mesenchymal transition. *Prog Mol Biol Transl* 116, 317–36.

Girgin, M. U., Broguiere, N., Mattolini, L. and Lutolf, M. P. (2021a). Gastruloids generated without exogenous Wnt activation develop anterior neural tissues. *Stem Cell Rep* 16, 1143–1155.

Girgin, M. U., Broguiere, N., Hoehnel, S., Brandenberg, N., Mercier, B., Arias, A. M. and Lutolf, M. P. (2021b). Bioengineered embryoids mimic post-implantation development in vitro. *Nat Commun* 12, 5140.

Grosswendt, S., Kretzmer, H., Smith, Z. D., Kumar, A. S., Hetzel, S., Wittler, L., Klages, S., Timmermann, B., Mukherji, S. and Meissner, A. (2020a). Epigenetic regulator function through mouse gastrulation. *Nature* 584, 102–108.

Grosswendt, S., Kretzmer, H., Smith, Z. D., Kumar, A. S., Hetzel, S., Wittler, L., Klages, S., Timmermann, B., Mukherji, S. and Meissner, A. (2020b). Epigenetic regulator function through mouse gastrulation. *Nature* 584, 102–108.

Hackett, J. A. and Surani, M. A. (2014). Regulatory Principles of Pluripotency: From the Ground State Up. *Cell Stem Cell* 15, 416–430.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W. M., Zheng, S., Butler, A., Lee, M. J., Wilk, A. J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573-3587.e29.

Harding, H. P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M. and Ron, D. (2000). Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. *Mol Cell* 6, 1099–1108.

Harding, H. P., Zhang, Y., Zeng, H., Novoa, I., Lu, P. D., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., et al. (2003). An Integrated Stress Response Regulates Amino Acid Metabolism and Resistance to Oxidative Stress. *Mol Cell* 11, 619–633.

Harrison, S. E., Sozen, B., Christodoulou, N., Kyprianou, C. and Zernicka-Goetz, M. (2017). Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science* 356,.

Houghton, F. D. (2021). HYPOXIA AND REPRODUCTIVE HEALTH: Hypoxic regulation of preimplantation embryos: lessons from human embryonic stem cells. *Reproduction* 161, F41–F51.

Hsiao, K.-Y., Wu, M.-H., Chang, N., Yang, S.-H., Wu, C.-W., Sun, H. S. and Tsai, S.-J. (2015). Coordination of AUF1 and miR-148a destabilizes DNA methyltransferase 1 mRNA under hypoxia in endometriosis. *Mhr Basic Sci Reproductive Medicine* 21, 894–904.

Iyer, N. V., Kotch, L. E., Agani, F., Leung, S. W., Laughner, E., Wenger, R. H., Gassmann, M., Gearhart, J. D., Lawler, A. M., Yu, A. Y., et al. (1998). Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha. *Gene Dev* 12, 149–62.

Izpisúa-Belmonte, J. C., Falkenstein, H., Dollé, P., Renucci, A. and Duboule, D. (1991). Murine genes related to the Drosophila AbdB homeotic genes are sequentially expressed during development of the posterior part of the body. *Embo J* 10, 2279–2289.

Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim, A. von, Hebestreit, H. F., Mukherji, M., Schofield, C. J., et al. (2001). Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Sci New York N Y* 292, 468–72.

Jaenisch, R. and Young, R. (2008). Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell* 132, 567–582.

Johnson, M. T., Mahmood, S. and Patel, M. S. (2003). Intermediary Metabolism and Energetics during Murine Early Embryogenesis*. *J Biol Chem* 278, 31457–31460.

Kaidi, A., Williams, A. C. and Paraskeva, C. (2007). Interaction between β-catenin and HIF-1 promotes cellular adaptation to hypoxia. *Nat Cell Biol* 9, 210–217.

Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. and Karolchik, D. (2010). BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26, 2204–2207.

Koch, F., Scholze, M., Wittler, L., Schifferl, D., Sudheer, S., Grote, P., Timmermann, B., Macura, K. and Herrmann, B. G. (2017). Antagonistic Activities of Sox2 and Brachyury Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev Cell* 42, 514-526.e7.

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

Koumenis, C., Naczki, C., Koritzinsky, M., Rastani, S., Diehl, A., Sonenberg, N., Koromilas, A. and Wouters, B. G. (2002). Regulation of Protein Synthesis by Hypoxia via Activation of the Endoplasmic Reticulum Kinase PERK and Phosphorylation of the Translation Initiation Factor eIF2α. *Mol Cell Biol* 22, 7405–7416.

Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R. and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. *Development* 126, 723–732.

Kozak, K. R., Abbott, B. and Hankinson, O. (1997). ARNT-Deficient Mice and Placental Differentiation. *Dev Biol* 191, 297–305.

Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* 11, 319–324.

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

Lau, K. Y. C., Rubinstein, H., Gantner, C. W., Hadas, R., Amadei, G., Stelzer, Y. and Zernicka-Goetz, M. (2022). Mouse embryo model derived exclusively from embryonic stem cells undergoes neurulation and heart development. *Cell Stem Cell* 29, 1445-1458.e8.

Lawson, K. A. and Pedersen, R. A. (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Dev Camb Engl* 101, 627–52.

Lee, P., Chandel, N. S. and Simon, M. C. (2020). Cellular adaptation to hypoxia through hypoxia inducible factors and beyond. *Nat Rev Mol Cell Bio* 21, 268–283.

Li, M. and Belmonte, J. C. I. (2018). Deconstructing the pluripotency gene regulatory network. *Nat Cell Biol* 20, 382–392.

Li, W., Goossens, K., Poucke, M. V., Forier, K., Braeckmans, K., Soom, A. V. and Peelman, L. J. (2014). High oxygen tension increases global methylation in bovine 4-cell embryos and blastocysts but does not affect general retrotransposon expression. *Reproduction Fertility Dev* 28, 948–959.

Li, W.-N., Wu, M.-H. and Tsai, S.-J. (2021). HYPOXIA AND REPRODUCTIVE HEALTH: The role of hypoxia in the development and progression of endometriosis. *Reproduction* 161, F19–F31.

Lin, S.-C., Lee, H.-C., Hsu, C.-T., Huang, Y.-H., Li, W.-N., Hsu, P.-L., Wu, M.-H. and Tsai, S.-J. (2019). Targeting Anthrax Toxin Receptor 2 Ameliorates Endometriosis Progression. *Theranostics* 9, 620–632.

Lindsley, R. C., Gill, J. G., Kyba, M., Murphy, T. L. and Murphy, K. M. (2006). Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development* 133, 3787–3796.

Loh, C. H. and Veenstra, G. J. C. (2022). The Role of Polycomb Proteins in Cell Lineage Commitment and Embryonic Development. *Epigenomes* 6, 23.

Lolas, M., Valenzuela, P. D. T., Tjian, R. and Liu, Z. (2014a). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proc National Acad Sci* 111, 4478–4483.

Lolas, M., Valenzuela, P. D. T., Tjian, R. and Liu, Z. (2014b). Charting Brachyury-mediated developmental pathways during early mouse embryogenesis. *Proc National Acad Sci* 111, 4478–4483.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161, 1202–1214.

Mahon, P. C., Hirota, K. and Semenza, G. L. (2001). FIH-1: a novel protein that interacts with HIF-1α and VHL to mediate repression of HIF-1 transcriptional activity. *Gene Dev* 15, 2675–2686.

Malkowska, A., Penfold, C., Bergmann, S. and Boroviak, T. E. (2022). A hexa-species transcriptome atlas of mammalian embryogenesis delineates metabolic regulation across three different implantation modes. *Nat Commun* 13, 3407.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A. F., Smith, A., et al. (2012). The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. *Cell* 149, 590–604.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc National Acad Sci* 78, 7634–7638.

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

Martin, G. R. and Evans, M. J. (1975). Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc National Acad Sci* 72, 1441–1445.

Mathupala, S. P., Heese, C. and Pedersen, P. L. (1997). Glucose Catabolism in Cancer Cells. *J Biol Chem* 272, 22776–22780.

Maxwell, P. H., Pugh, C. W. and Ratcliffe, P. J. (2001). Activation of the HIF pathway in cancer. *Curr Opin Genet Dev* 11, 293–299.

McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. and Schofield, C. J. (2010). Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr Opin Struc Biol* 20, 659–672.

McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., Hu, J. L., Murrow, L. M., Weissman, J. S., Werb, Z., et al. (2019a). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 16, 619–626.

McGinnis, C. S., Patterson, D. M., Winkler, J., Conrad, D. N., Hein, M. Y., Srivastava, V., Hu, J. L., Murrow, L. M., Weissman, J. S., Werb, Z., et al. (2019b). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 16, 619–626.

Misra, A., Pandey, C., Sze, S. K. and Thanabalu, T. (2012). Hypoxia Activated EGFR Signaling Induces Epithelial to Mesenchymal Transition (EMT). *Plos One* 7, e49766.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S. (2003). The Homeoprotein Nanog Is Required for Maintenance of Pluripotency in Mouse Epiblast and ES Cells. *Cell* 113, 631–642.

Mohamed, O. A., Clarke, H. J. and Dufort, D. (2004). β-catenin signaling marks the prospective site of primitive streak formation in the mouse embryo. *Dev Dynam* 231, 416–424.

Moris, N., Anlas, K., Schroeder, J., Ghimire, S., Balayo, T., Brink, S. C. van den, Alemany, A., Oudenaarden, A. van and Arias, A. M. (2020). Generating Human Gastruloids from Human Embryonic Stem Cells.

Muz, B., Puente, P. de la, Azab, F. and Azab, A. K. (2015). The role of hypoxia in cancer progression, angiogenesis, metastasis, and resistance to therapy. *Adv Exp Med Biol* 3, 83–92.

Neijts, R., Simmini, S., Giuliani, F., Rooijen, C. and Deschamps, J. (2014). Region-specific regulation of posterior axial elongation during vertebrate embryogenesis. *Dev Dynam* 243, 88–98.

Nguyen, A. Q., Bardua, I., Greene, B., Wrenzycki, C., Wagner, U. and Ziller, V. (2020). Mouse embryos exposed to oxygen concentrations that mimic changes in the oviduct and uterus show improvement in blastocyst rate, blastocyst size, and accelerated cell division. *Reprod Biology* 20, 147–153.

Ni, Y., Hagras, M. A., Konstantopoulou, V., Mayr, J. A., Stuchebrukhov, A. A. and Meierhofer, D. (2019). Mutations in NDUFS1 Cause Metabolic Reprogramming and Disruption of the Electron Transfer. *Cells* 8, 1149.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H. and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell* 95, 379–391.

Nowotschin, S., Setty, M., Kuo, Y.-Y., Liu, V., Garg, V., Sharma, R., Simon, C. S., Saiz, N., Gardner, R., Boutet, S. C., et al. (2019). The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature* 569, 361–367.

Ohh, M., Park, C. W., Ivan, M., Hoffman, M. A., Kim, T.-Y., Huang, L. E., Pavletich, N., Chau, V. and Kaelin, W. G. (2000). Ubiquitination of hypoxia-inducible factor requires direct binding to the β-domain of the von Hippel–Lindau protein. *Nat Cell Biol* 2, 423–427.

Ottosen, L. D., Hindkjær, J., Husth, M., Petersen, D. E., Kirk, J. and Ingerslev, H. J. (2006). Observations on intrauterine oxygen tension measured by fibre-optic microsensors. *Reprod Biomed Online* 13, 380–385.

Pabon, J. E., Findley, W. E. and Gibbons, W. E. (1989). The toxic effect of short exposures to the atmospheric oxygen concentration on early mouse embryonic development. *Fertil Steril* 51, 896–900.

Pagé, E. L., Robitaille, G. A., Pouysségur, J. and Richard, D. E. (2002). Induction of Hypoxia-inducible Factor-1α by Transcriptional and Translational Mechanisms*. *J Biol Chem* 277, 48403–48409.

Parmar, K., Mauch, P., Vergilio, J.-A., Sackstein, R. and Down, J. D. (2007). Distribution of hematopoietic stem cells in the bone marrow according to regional hypoxia. *Proc National Acad Sci* 104, 5431–5436.

Peinado, H., Olmeda, D. and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* 7, 415–428.

Perrimon, N., Pitsouli, C. and Shilo, B.-Z. (2012). Signaling Mechanisms Controlling Cell Fate and Embryonic Patterning. *Csh Perspect Biol* 4, a005975.

Pettersen, E. O., Juul, N. O. and Rønning, O. W. (1986). Regulation of protein metabolism of human cells during and after acute hypoxia. *Cancer Res* 46, 4346–51.

Pierce, G. B. and Verney, E. L. (1961). An in vitro and in vivo study of differentiation in teratocarcinomas. *Cancer* 14, 1017–1029.

Pijuan-Sala, B., Guibentif, C. and Göttgens, B. (2018). Single-cell transcriptional profiling: a window into embryonic cell-type specification. *Nat Rev Mol Cell Bio* 19, 399–412.

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., Mulas, C., Ibarra-Soria, X., Tyser, R. C. V., Ho, D. L. L., et al. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* 566, 490–495.

Probst, S., Sagar, S., Tosic, J., Schwan, C., Grün, D. and Arnold, S. J. (2020). Spatiotemporal sequence of mesoderm and endoderm lineage segregation during mouse gastrulation. *Development* 148, dev193789.

Ramírez-Bergeron, D. L., Runge, A., Dahl, K. D. C., Fehling, H. J., Keller, G. and Simon, M. C. (2004). Hypoxia affects mesoderm and enhances hemangioblast specification during early development. *Development* 131, 4623–4634.

Rey, S. and Semenza, G. L. (2010). Hypoxia-inducible factor-1-dependent mechanisms of vascularization and vascular remodelling. *Cardiovasc Res* 86, 236–242.

Risau, W. (1997). Mechanisms of angiogenesis. *Nature* 386, 671–674.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47–e47.

Rivera-Pérez, J. A. and Hadjantonakis, A.-K. (2015). The Dynamics of Morphogenesis in the Early Mouse Embryo. *Csh Perspect Biol* 7, a015867.

Rivron, N. C., Frias-Aldeguer, J., Vrij, E. J., Boisset, J.-C., Korving, J., Vivié, J., Truckenmüller, R. K., Oudenaarden, A. van, Blitterswijk, C. A. van and Geijsen, N. (2018). Blastocyst-like structures generated solely from stem cells. *Nature* 557, 106–111.

Rossant, J. and Tam, P. P. L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136, 701–713.

Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R. G., Kwon, C. and Lutolf, M. P. (2021). Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell* 28, 230-240.e6.

Ryan, J. M., Barry, F. P., Murphy, J. M. and Mahon, B. P. (2005). Mesenchymal stem cells avoid allogeneic rejection. *J Inflamm* 2, 8.

Sagner, A. and Briscoe, J. (2019a). Establishing neuronal diversity in the spinal cord: a time and a place. *Development* 146, dev182154.

Sagner, A. and Briscoe, J. (2019b). Establishing neuronal diversity in the spinal cord: a time and a place. *Development* 146, dev182154.

Scheibner, K., Schirge, S., Burtscher, I., Büttner, M., Sterr, M., Yang, D., Böttcher, A., Ansarullah, Irmler, M., Beckers, J., et al. (2021). Epithelial cell plasticity drives endoderm formation during gastrulation. *Nat Cell Biol* 23, 692–703.

Schrode, N., Xenopoulos, P., Piliszek, A., Frankenberg, S., Plusa, B. and Hadjantonakis, A. (2013). Anatomy of a blastocyst: Cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. *Genesis* 51, 219–233.

Semenza, G. L. (1999). REGULATION OF MAMMALIAN O2 HOMEOSTASIS BY HYPOXIA-INDUCIBLE FACTOR 1. *Annu Rev Cell Dev Bi* 15, 551–578.

Semenza, G. L., Roth, P. H., Fang, H. M. and Wang, G. L. (1994). Transcriptional regulation of genes encoding glycolytic enzymes by hypoxia-inducible factor 1. *J Biol Chem* 269, 23757–23763.

Shmakova, A., Batie, M., Druker, J. and Rocha, S. (2014). Chromatin and oxygen sensing in the context of JmjC histone demethylases. *Biochem J* 462, 385–395.

Shparberg, R. A., Glover, H. J. and Morris, M. B. (2019). Modeling Mammalian Commitment to the Neural Lineage Using Embryos and Embryonic Stem Cells. *Front Physiol* 10, 705.

Simon, M. C. and Keith, B. (2008). The role of oxygen availability in embryonic development and stem cell function. *Nat Rev Mol Cell Bio* 9, 285–296.

Sozen, B., Amadei, G., Cox, A., Wang, R., Na, E., Czukiewska, S., Chappell, L., Voet, T., Michel, G., Jing, N., et al. (2018). Self-assembly of embryonic and two extra-embryonic stem cell types into gastrulating embryo-like structures. *Nat Cell Biol* 20, 979–989.

Stadler, M. B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., Nimwegen, E. van, Wirbelauer, C., Oakeley, E. J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 480, 490–495.

Sun, X., Meyers, E. N., Lewandoski, M. and Martin, G. R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Gene Dev* 13, 1834–1846.

Takaoka, K. and Hamada, H. (2011). Cell fate decisions and axis determination in the early mouse embryo. *Development* 139, 3–14.

Takaoka, K., Yamamoto, M. and Hamada, H. (2007). Origin of body axes in the mouse embryo. *Curr Opin Genet Dev* 17, 344–350.

Tam, P. P. L. and Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mech Develop* 68, 3–25.

Tanaka, S., Kunath, T., Hadjantonakis, A.-K., Nagy, A. and Rossant, J. (1998). Promotion of Trophoblast Stem Cell Proliferation by FGF4. *Science* 282, 2072–2075.

Tarazi, S., Aguilera-Castrejon, A., Joubran, C., Ghanem, N., Ashouokhi, S., Roncato, F., Wildschutz, E., Haddad, M., Oldak, B., Gomez-Cesar, E., et al. (2022). Post-gastrulation synthetic embryos generated ex utero from mouse naive ESCs. *Cell* 185, 3290-3306.e25.

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

Tortelote, G. G., Hernández-Hernández, J. M., Quaresma, A. J. C., Nickerson, J. A., Imbalzano, A. N. and Rivera-Pérez, J. A. (2013). Wnt3 function in the epiblast is required for the maintenance but not the initiation of gastrulation in mice. *Dev Biol* 374, 164–173.

Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* 141, 1209–1221.

Turner, D. A., Girgin, M., Alonso-Crisostomo, L., Trivedi, V., Baillie-Johnson, P., Glodowski, C. R., Hayward, P. C., Collignon, J., Gustavsen, C., Serup, P., et al. (2017). Anteroposterior polarity and elongation in the absence of extra-embryonic tissues and of spatially localised signalling in gastruloids: mammalian embryonic organoids. *Development* 144, 3894–3906.

Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. and Nicolas, J.-F. (2009). Redefining the Progression of Lineage Segregations during Mammalian Embryogenesis by Clonal Analysis. *Dev Cell* 17, 365–376.

Veenvliet, J. V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch, F., Guignard, L., Kumar, A. S., Pustet, M., et al. (2020a). Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* 370,.

Veenvliet, J. V., Bolondi, A., Kretzmer, H., Haut, L., Scholze-Wittler, M., Schifferl, D., Koch, F., Guignard, L., Kumar, A. S., Pustet, M., et al. (2020b). Mouse embryonic stem cells self-organize into trunk-like structures with neural tube and somites. *Science* 370, eaba4937.

Veenvliet, J. V., Lenne, P.-F., Turner, D. A., Trivedi, I. N. and V. and Trivedi, V. Sculpting with stem cells: how models of embryo development take shape. *Development*.

Vianello, S. and Lutolf, M. P. (2021). In vitro endoderm emergence and self-organisation in the absence of extraembryonic tissues and embryonic architecture. *Biorxiv* 2020.06.07.138883.

Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M. and Sutherland, A. (2012). Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev Dynam* 241, 270–283.

Winnier, G., Blessing, M., Labosky, P. A. and Hogan, B. L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Gene Dev* 9, 2105–2116.

Woods, L., Perez-Garcia, V. and Hemberger, M. (2018). Regulation of Placental Development and Its Impact on Fetal Growth—New Insights From Mouse Models. *Front Endocrinol* 9, 570.

Wymeersch, F. J., Wilson, V. and Tsakiridis, A. (2021). Understanding axial progenitor biology in vivo and in vitro. *Development* 148, dev180612.

Xu, P.-F., Borges, R. M., Fillatre, J., Oliveira-Melo, M. de, Cheng, T., Thisse, B. and Thisse, C. (2021). Construction of a mammalian embryo model from stem cells organized by a morphogen signalling centre. *Nat Commun* 12, 3277.

Yang, M.-H., Wu, M.-Z., Chiou, S.-H., Chen, P.-M., Chang, S.-Y., Liu, C.-J., Teng, S.-C. and Wu, K.-J. (2008). Direct regulation of TWIST by HIF-1α promotes metastasis. *Nat Cell Biol* 10, 295–305.

Yeo, C. D., Kang, N., Choi, S. Y., Kim, B. N., Park, C. K., Kim, J. W., Kim, Y. K. and Kim, S. J. (2017). The role of hypoxia on the acquisition of epithelial-mesenchymal transition and cancer stemness: a possible link to epigenetic regulation. *Korean J Intern Medicine* 32, 589–599.

Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519–523.

Young, R. A. (2011). Control of the Embryonic Stem Cell State. *Cell* 144, 940–954.

Zernicka-Goetz, M., Morris, S. A. and Bruce, A. W. (2009). Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat Rev Genet* 10, 467–477.

Zhang, X., Peterson, K. A., Liu, X. S., McMahon, A. P. and Ohba, S. (2013a). Gene Regulatory Networks Mediating Canonical Wnt Signal-Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. *Stem Cells* 31, 2667–2679.

Zhang, X., Peterson, K. A., Liu, X. S., McMahon, A. P. and Ohba, S. (2013b). Gene Regulatory Networks Mediating Canonical Wnt Signal-Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. *Stem Cells* 31, 2667–2679.

List of publications

López-Anguita, N., Gassaloglu, S. I.*, Stötzel, M.*, Bolondi, A., Conkar, D., Typou, M., Buschow, R., Veenvliet, J. V. and Bulut-Karslioglu, A. "Hypoxia induces an early primitive streak signature, enhancing spontaneous elongation and lineage representation in gastruloids". (2022). *Development* DOI:10.1242/dev.200679

* Gassaloglu and Stötzel contributed equally to this study.

Appendix

Supplementary tables

Target gene	Forward sequence	Reverse sequence
Oct4	GGGCTAGAGAAGGATGTGGTT	AAGGTGTCCCTGTAGCCTCA
Nanog	GAAGTACCTCAGCCTCCAGC	TTATGGAGCGGAGCAGCATT
Sox2	GATCAGCATGTACCTCCCCG	TCCTCTTTTTGCACCCCTCC
T	TGCCTGTGAGTCATAACGCC	GTCTGTGACTGTAGCAGCC
Eomes	TGTGACGGCCTACCAAAACA	ATCTAGGGGAATCCGTGGGA
Gata6	CGCTGTTTGTTTAGGGCTCG	AAGACGAGATGGGGGAAAGC
Foxa2	TAAGCGAGCTAAAGGGAGCA	GTGGTTGAAGGCGTAATGGT
Tbx6	CCGATTTCCTGAGACCACATTC	CCCGCTCCCTCTTACAGTTTC
Wnt3	CAAGCACAACAATGAAGCAGGC	TCGGGACTCACGGTGTTTCTC
Pgk1	CCCTTCCTGGCTATCTTGGG	ATCTGCTTAGCTCGACCCAC
Pfkfb3	CGAGATCGATGCTGGTGTGT	CAGCCTGGTGACAGATGACC
Atp5g3	GGAAGAAAGATGTTCGCCTGC	TGCCAAAGACTGTTCCAATACCA
Ndufs8	GGCATGACCCTAAGTTACCTCT	CTTGGCTCAGCCTCAATGGTG
Loxl2	CACAGGCACTACCACAGCAT	ATGCGGTAGCCATCATAGCG
Eif1a	TTTGTAGCTCTCAGAAGCCAGG	GCATACTCCTGCCCATCCTC

Table 3. List of primers.

Target protein	Technique	Company	Order number	Concentration
Foxa2	IF	Santacruz	sc-6554	1:500
T	IF	Cell Signaling Technology	D2Z3J	1:500
T	WB	R&Dsystems	AF2085	1:1000
H3	WB	Abcam	ab1791	1:2000
ß -catenin	WB	Cell Signalling	CST 8480T	1:1000
Gapdh	WB	Cell Signaling Technology	2118S	1:1000
Hif1α	ChIPseq	Cell Signaling Technology	36169S	Unknown concentration
Uhrf1	WB/IF	Merck	MABE308	1:1000/1:500
5-mC	IF	Diagenode	C15200003	1:200
Dntm3b	WB	Cell Signaling Technology	CST48488	1:1000
Dnmt3a	WB	Abcam	ab13888	1:1000
Tet2	WB	CST	45010S	1:1000
HRP Donkey anti-mouse	WB	Active Motif	15014	1:5000
HRP Donkey anti-rabbit	WB	Jackson	711-035-152	1:5000

Alexa Fluor 647	IF	Invitrogen	A31573	1:500
Alexa Fluor 546	IF	Invitrogen	A11056	1:500
Eomes	WB	Abcam	ab23345	1:1000
Wnt3	WB	CST	74537	1:1000
donkey anti-rabbit	IF	Invitrogen	A31573	1:250
donkey anti-goat	IF	Invitrogen	A11056	1:250
donkey anti-mouse	IF	Life technologies - Thermofisher	A21202	1:250

Table 4. List of antibodies

Sample	Norm. Factor
H3K27me3_hypD7	1.28
H3K27me3_hypD2	1.41
H3K27me3_norm	1.00
H3K4me3_hypD7	0.89
H3K4me3_hypD2	1.13
H3K4me3_norm	1.00

Table 5. Scaling factors for quantitative ChIP-seq

List of Figures

List of Tables

List of abbreviations

2i	2 inhibitors
3D	3 Dimensional
A	Anterior
A-P	Anterior-Posterior
ATF4	Activating transcription factor 4
ATP	Adenosine 5'-triphosphate
AVE	Anterior Ventral Endoderm
bHLH	basic helix-loop-helix
BP	Biological Process
CGI	Cytosine Guanine Island
Chi	CHIRON/ CHIR99021
ChIP-seq	Chromatin Immunoprecipitation Sequencing
D	Day
D-V	Dorsal-Ventral
DE	Differentially Expressed
Dist.	Distal
DNA	Deoxyribonucleic Acid
DVE	Dorsal Ventral Endoderm
E	Embryonic day
EBs	Embryoid Bodies
ECM	Extracellular Matrix
EMT	Epithelial to Mesenchymal Transition
EPI	Epiblast
EpiSC	Epliblast Stem Cell
ES	Embryonic Stem
ExE	Extraembryonic
FC	Fold Change
GFP	Green Fluorescent Protein
GO	Gene Ontology
H	Hour
HIFs	Hypoxia Inducible Factors
HRE	Hypoxia Responsive Element
HRGs	Hypoxia Responsive Genes
HSCs	Hematopoietic Stem Cells
Hyp	Hypoxia

ICM	Inner Cell Mass
IF	Immunofluorescence
JmjC	Jumanji C
Kb	Kilo base pair
KDM	Histone lysine demethylase
LFQ	Label Free Quantification
LIF	Leukemia Inhibitor Factor
Log	Logarithm
M/S	Mass Spectrometry
Max	Maximum
Min	Minimum
Neg.	Negative
NH	Normoxia-Hypoxia
NMPs	Neuromesodermal Progenitors
Norm	Normoxia
O_2	Oxygen
ODD	Oxygen-dependent Degradation Domain
OxPhos	Oxidative Phospholylation
P	p-value
P	Posterior
P-D	Proximal-Distal
PAS	Per-Arnt-Sim
PCA	Principal Component Analysis
PRC2	Polycomb Repressive Complex 2
PE	Parietal Endoderm
PHDs	Prolil Hydroxylases
PORCN	Porcupine O-Acyltransferase
PrE	Primitive Endoderm
Prox.	Proximal
PS	Primitive Streak
PSM	Pre-somitic mesoderm
pVHL	von Hippel-Lindau
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RT	Room Temperature
RT-qPCR	Real time-quantitative Polymerase Chain Reaction
s.d.	Standard deviation

scRNA-seq   Single-cell RNA Sequencing

T-SNE   T-distributed stochastic neighbor embedding

TADs   Transcriptional Transactivation Domains

TCA   Tricarboxyl acid

TE   Trophoectoderm

TET   Tet-eleven translocation

TLS   Trunk-like Structure

TPM   Transcript Per Million

TS   Trophoblast Stem

UMAP   Uniform Manifold Approximation and Projection

VE   Ventral Endoderm

Vs.   Versus

WGBS   Whole Genome Bisulfite Sequencing

WNTa   WNT activated

XEN   Extraembryonic Endoderm