technische universität
dortmund

MAX PLANCK INSTITUTE
OF MOLECULAR PHYSIOLOGY

Differentiation of distinct mesoderm subtypes from regulatory interactions between FGF and BMP signaling

Marina Gattiglio

Zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften (Dr. rer. nat.)

bei der Fakultät für Chemie und Chemische Biologie
an der Technischen Universität Dortmund

August 2022

Declaration

Tag der Abgabe der Dissertation:
03. August 2022
von Marina Gattiglio

Gutachter:
Prof. Dr. Philippe I. H. Bastiaens
Prof. Dr. Jan Hengstler

The work presented in this thesis was performed in the group of Dr. Christian Schröter in the Department of Systemic Cell Biology lead by Prof. Dr. Philippe I. H. Bastiaens at the Max Planck Institute of Molecular Physiology, Dortmund, Germany.

Marina Gattiglio was affiliated to the International Max Planck Research School in Chemical and Molecular Biology, Dortmund, Germany

Table of Contents

Abstract

Distinct mesodermal precursors arise during gastrulation in a precise temporal and spatial manner. In the mouse, the cells differentiating earlier in a more proximal position form the precursors for the heart and vasculature, followed in a more distal position by precursors of the urogenital system and then the musculoskeletal system. Mesoderm patterning is controlled by signals such as BMP4, produced by the extraembryonic ectoderm, and FGFs and WNTs produced by the epiblast itself. How do these signals interact in this process and whether cell-cell crosstalk can influence mesoderm patterning is still largely unknown.

Here I address this question with a 2D cell culture model that employs primed pluripotent stem cells resembling mouse epiblast cells and differentiates them with precise concentrations of different signaling molecules in a chemically defined medium. By titrating BMP4, I observe that high concentrations of BMP4 promote the expression of proximal mesoderm markers, while intermediate to low concentrations favour distal and pan-mesoderm markers. The opposite effect is seen with a titration of FGF. A scRNAseq analysis revealed that endogenous FGF signaling lead to the differentiation of heterogeneous mixtures of cell types, and that their proportions were affected by the dose of exogenous FGF. Furthermore, higher doses of exogenous FGF reduced the expression of BMP ligands in the cells, and promoted the expression of selected WNT ligands and endogenous FGFs, indicating the existence of a positive FGF feedback loop. Cells associated with high FGF signaling levels were also found to be spatially clustered in the cultures, suggesting that this loop could generate coherent clusters of cells with discrete identities in a cell population. Based on these findings, I propose a molecular mechanism for the FGF feedback loop. Together, these results indicate that the patterning of the mesoderm is dependent on a combination of external signaling and cell-cell crosstalk.

Zusammenfassung

Während der Gastrulation entstehen distinkte mesodermale Vorläuferzellen in einer präzisen zeitlichen und räumlichen Abfolge. In der Maus bilden Zellen, die früh aus proximalen Positionen differenzieren, die Vorläufer von Herz und Gefäßsystem, während später aus distalen Positionen Vorläufer des urogenitalen Systems und anschließend des Bewegungsapparats gebildet werden. Die Identität mesodermaler Zellen wird durch Signale wie BMP4, die vom extraembryonalen Ektoderm produziert werden, sowie von FGFs und WNTs, die vom Epiblast selbst produziert werden, kontrolliert. Wie diese Signale interagieren und ob Zell-Zell-Kommunikation die Mesodermbildung beeinflusst, ist größtenteils unbekannt.

Um diese Frage zu beantworten verwende ich ein 2D-Zellkulturmodell aus epiblastartigen Mausstammzellen, die mit verschiedenen Konzentrationen bestimmter Signalmoleküle in einem chemisch definierten Medium differenziert werden. BMP4-Titrierung zeigte, dass hohe Konzentrationen von BMP4 die Expression proximaler Mesodermmarker unterstützen, während mittlere bis geringe Konzentrationen distale und panmesodermale Marker begünstigen. Das Gegenteil wird durch FGF-Titrierung ersichtlich. Die Analyse von scRNAseq-Daten zeigte, dass endogene FGF-Signale zur Differenzierung heterogener Zelltypen führen, deren Verhältnis von der Dosis von exogenem FGF abhängt. Zudem reduzierten höhere Dosen exogenen FGFs die Expression von BMP-Liganden in den Zellen und unterstützten die Expression einiger bestimmter WNT-Liganden und endogenen FGFs, was die Existenz eines positiven FGF-Feedback-Loops impliziert. Zellen, die mit einem hohen FGF-Signal-Level assoziiert sind, waren außerdem innerhalb der Kultur räumlich geclustert, was die Vermutung nahelegt, dass der FGF-Feedback-Loop kohärente Zellcluster mit diskreten Identitäten innerhalb einer Zellpopulation generiert. Basierend auf diesen Erkenntnissen schlage ich einen molekularen Mechanismus für den FGF-Feedback-Loop vor. Zusammengenommen zeigen diese Ergebnisse, dass die mesodermale Identität in Abhängigkeit einer Kombination von externen Signalen und Zell-Zell-Kommunikation bestimmt wird.

INTRODUCTION

The organogenesis of multicellular organisms requires the differentiation of groups of cells with discrete identities. In order to coordinate the differentiation of individual cell identities, embryonic cells need to integrate the mechanical and signaling cues that they receive from their environment. One critical example of a tightly regulated differentiation process is the differentiation of the mesoderm. Indeed, mesoderm differentiation marks the first time in the embryonic development that a group of cells with uniform developmental potential differentiates into one of several cell types, in a precise temporal and spatial process. Mesoderm, together with ectoderm and endoderm, is one of the three germ layers that arise during a process called gastrulation, which will determine the differentiation of all of the tissues and organs of the embryo (Tam and Loebel 2007). In particular, different mesoderm precursors will form tissues like muscles, connective tissues and most of the skeleton, as well as the kidneys, heart, blood vessels and some extraembryonic membranes (Gilbert and Barresi 2016).

Mesoderm is a relatively recent evolutionary invention, as species that arose very early during animal evolution are only formed by ectoderm and endoderm (Technau and Scholz 2003). These animals, the diploblasts, are characterized by more simple structures: examples of them are sponges and jellyfish. The appearance of mesoderm promoted the development of organs that facilitate movement and support and protect the body (the musculoskeletal system), exchange gas and nutrients in every cell (the circulatory system) and support the adaptation to different environments (the secretory and reproductive system). Metazoans with the three germ layers, the triploblasts, comprise the majority of the animal species: they encompass protostomes, such as insects and flatworms, and deuterostomes, such as starfish, reptiles and mammals (Technau and Scholz 2003; Ferretti and Hadjantonakis 2019).

Pattern formation can depend on external signaling or self-organized systems

Symmetry breaking events, axis formation and the emergence of patterns in precise spatial and temporal coordinated events are core processes in the field of developmental biology. At the base of these processes are the ability of cells to respond and interact with their environment through physical and chemical signals, such as tension forces or signaling molecules. The role of signaling molecules in the embryonic patterning has been extensively studied and can be conceptualized by two theories: the reaction-diffusion and the positional information theories (Turing 1952; Wolpert 1969). These two theories, although seemingly opposite, can be considered complementary and can cooperate in distinct embryonic stages and structures (Fig. 1, Green & Sharpe, 2015). One of the core elements of both theories is the concept of morphogen, a signaling molecule that can instruct different cell fates depending on the local concentration.

The positional information theory describes an extrinsically organized system, where the tissue asymmetries are generated outside of the tissue itself, in different tissues or signaling centers. An externally produced morphogen would determine this asymmetry through the gradual differences in its concentration across the tissue. According to this theory, cells are able to interpret the local concentration to form coordinated patterns over a whole tissue, for example stripes, dots or a sequence of cell fates (Wolpert 1969). The reaction-diffusion theory describes a locally self-organizing system. It is based on a two-equation system to explain how uniform field of cells can develop different patterns, like stripes or dots, thanks to a local enhancement of a signal and its lateral inhibition. Random fluctuations of the signal production determine a higher concentration of this ligand in certain cells. The signal would promote its own production, stabilizing itself and creating a peak of expression, but also determined the expression of its own fast diffusing inhibitor. The fast inhibitor is then able to travel to the neighboring areas, repressing the signal and inhibiting the formation of a close peak of expression (Turing 1952).

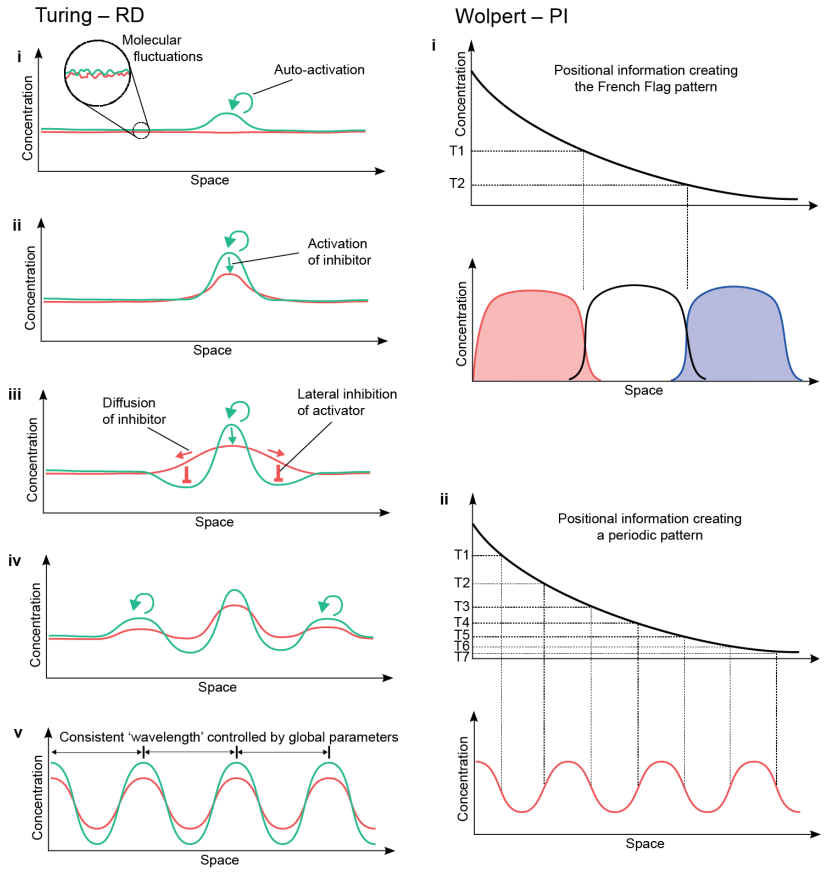


Fig. 1: Reaction-diffusion and positional information theory.

Green & Sharpe, 2015

The positional information theory is conceptually simpler and in the decades since the original paper, it has been found to apply in the differentiation of many organs and structures. The formation of patterns through cell-cell crosstalk highlighted in the reaction diffusion mechanism has been revived only in more recent times, as more cell-cell interactions are found to contribute to the formation of distinct cell lineages. The two mechanisms have even been suggested to cooperate in the formation of different structures: an example is the patterning of a 2D model of human gastrulation, where a reaction diffusion mechanism defines regions of signaling gradient and is then followed by a positional information fate (Tewary et al. 2017). A different cell-cell interaction mechanism, the community effect, has been reported to determine the differentiation of the muscle progenitors in frog. In this system, a large number of pluripotent cells in close proximity is required to activate the muscle differentiation program: only in this case the concentration of a specific signaling molecule, eFGF,

can reach a certain concentration and determine the differentiation of muscle cells (Gurdon 1988; Gurdon et al. 1993; Standley, Zorn, and Gurdon 2001).

Gastrulation patterns are highly conserved among vertebrates

The coordinate formation of different cell lineages is highly conserved even between distant vertebrates. The epiblast differentiation and fate map of avian species, for example, can be compared to mammal development (Kinder et al. 1999; Lawson, Meneses, and Pedersen 1991; Martinez Arias and Steventon 2018). Similar topological fate relationships have also been seen in a revised fate map of the development of another model animal, the frog Xenopus laevis (Lane and Sheets 2006). These similarities also include the signaling centers and the signaling networks during gastrulation. The organizer of the frog Xenopus laevis is a source of inhibitors of WNT, NODAL and BMP signaling and is found to be similar to the node in chick embryos and the AVE in the mouse. Indeed, signaling centers of the mouse embryo can induce anterior structures in chicken and frog embryos (Martinez Arias and Steventon 2018; Kintner and Dodd 1991; Knoetgen et al. 2000). These similarities allow to compare the signaling involved in the development of different species.

Given the ethical and practical problems in studying human gastrulation, a considerable amount of mammalian development studies has been carried out in the house mouse (Mus musculus): this animal model has many advantages, due to its small size, easiness to breed, fast development and large litters.

Mouse embryonic development

Early mouse development

In the mouse, after fertilization, the one-cell zygote goes into three rounds of cell division to form progressively smaller identical cells, the blastomeres. These identical blastomeres are only loosely arranged, but when they enter the 8-cell stage, cells start to produce cell adhesion molecules, such as E-cadherin, to compact the zygote (de Vries et al. 2004). This compact 8-cell embryo divide again to form the 16-cell morula, formed by a small group of internal cells surrounded by a larger group of cells (Fig. 2, Bedzhov, Graham, Leung, & Zernicka-Goetz, 2014).

At this point, around embryonic day 2.5-3.0 (abbreviated as E2.5-3.0), the cells that compose the embryo experience the first binary cell fate decision, with the descendants of the external cells giving rise to the trophectoderm (TE) and the internal ones becoming the central inner cells mass (ICM) (Bedzhov and Zernicka-Goetz 2014). The TE cells mediate the implant into the uterus and will later form embryonic parts of the placenta, whereas the ICM cells will give rise to the embryonic tissues, as well as some extraembryonic structures such as the amnion (Ralston and Rossant 2005).

With a process called cavitation, a fluid-filled cavity is formed in the morula and the ICM cells position themselves on one side of the TE cells. The mouse embryo is then in the blastocyst stage, where a second binary cell fate decision around E3.5 divides the ICM cells in primitive endoderm (PrE) and epiblast (Epi) (Bedzhov et al. 2014; Simon, Hadjantonakis, and Schröter 2018; Chazaud and Yamanaka 2016). The latter will form all tissues of the embryo proper, while the former is instrumental in surrounding and influencing the differentiation of the epiblast, and will later form the yolk sac. Around implantation into the uterus, circa E4.5 to E5.0, the epiblast and part of the TE cells form an elongated, cup-shaped structure surrounded by a layer of the PrE-derived visceral endoderm (VE, Bedzhov et al., 2014).

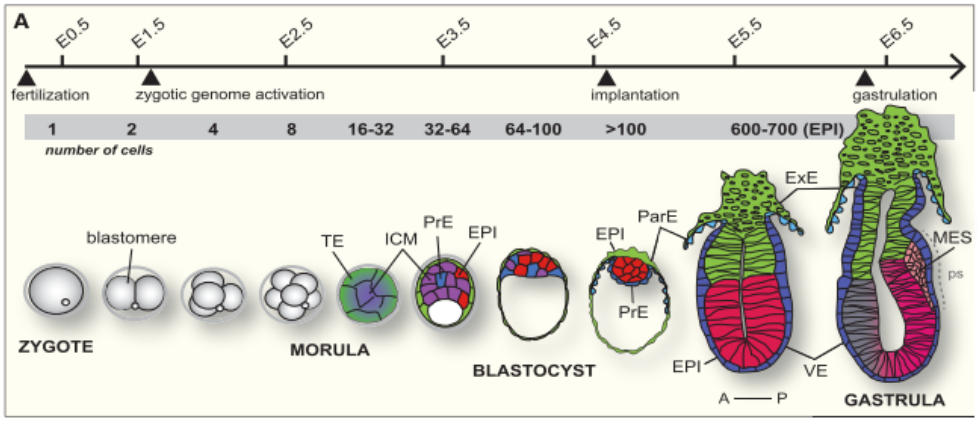


Fig 2: Early mouse embryonic development.

Modified from Hadjantonakis & Martinez Arias, 2016

At E5.5 the uniform epiblast cells undergo dramatic changes, where they differentiate and migrate to form a tridimensional embryo. This process is called gastrulation and it establishes the three germ layers that will form all of the tissues and organs of the

embryo proper: ectoderm, mesoderm and endoderm. Ectoderm cells will contribute to the central nervous system and the epidermis, the endoderm to the digestive and respiratory system and the mesoderm to the musculoskeletal, circulatory and secretory system (Gilbert and Barresi 2016).

Gastrulation and mesoderm differentiation

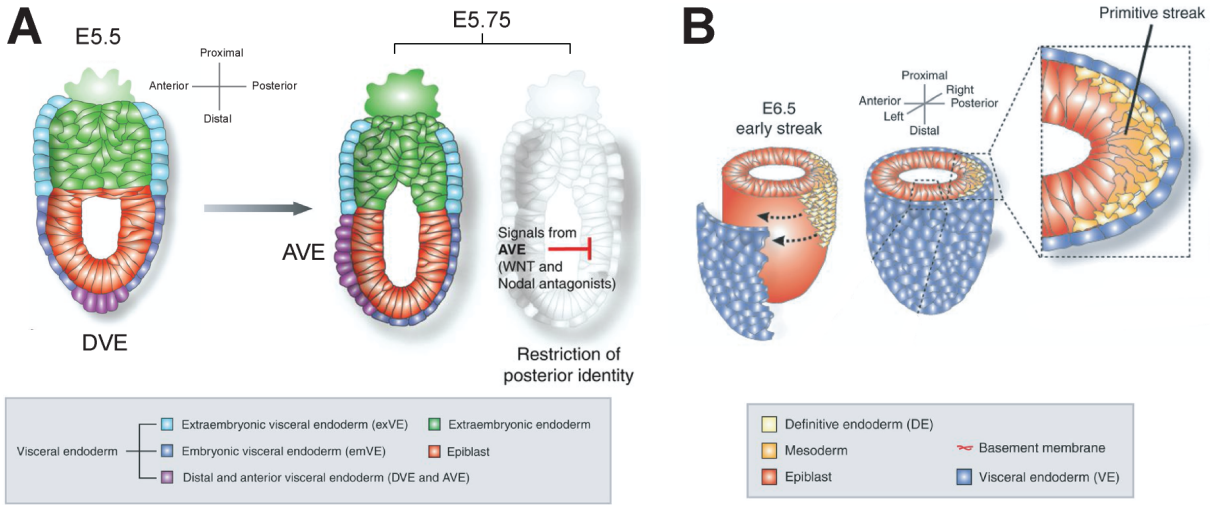


Fig. 3: Mouse gastrulation

Modified from Rivera-Perez & Hadjantonakis, 2015. A)Positioning and migration of DVE and AVE. B)Epiblast cells migration during gastrulation

At E5.5 in mouse development, the epiblast cells form a single layered cup-shaped structure, adjacent to the extraembryonic ectoderm and surrounded by the visceral endoderm (Fig. 3a). Complex patterns of molecular activity and precise crosstalk between extraembryonic and embryonic cells are necessary to establish signaling centers and define the proximal-distal and anterior-posterior axes of the embryo that will be essential for the differentiation and patterning of the mesoderm (Arnold and Robertson 2009; Tam and Loebel 2007). One key signaling center for the formation of the three germ layers is the distal/anterior visceral endoderm (DVE and AVE). It is initially located in the distal tip of the embryo, at which stage it is termed distal visceral endoderm (DVE), and its formation is inhibited in the proximal portion of the visceral endoderm by the extraembryonic ectoderm (ExE) (Rodriguez et al. 2005). At E6.0, the DVE migrates to the prospective anterior side to form the AVE (Fig. 3a). In this position, the AVE breaks the radial symmetry of the embryo, as it is a source of NODAL, WNT and BMP inhibitors (Tam and Loebel 2007). The anterior-posterior

gradient of these inhibitors generates tissues asymmetries in the epiblast and confer neuroectodermal identity to the neighbouring cells (Thomas and Beddington 1996; Lu, Brennan, and Robertson 2001). The epiblast cells on the opposite side of the embryo are free from the influence of the AVE-produced inhibitors and can differentiate into mesodermal and endodermal lineages. As they differentiate, they ingress and invade the space between the epiblast layer and the VE, marking the point of ingression as the primitive streak (Fig. 3b, Arnold & Robertson, 2009; Lawson et al., 1991; Tam & Loebel, 2007). To ingress in the primitive streak, cells need to undergo the epithelial-to-mesenchymal transition (EMT). During EMT, the expression of the epithelial adherent junctions (E-cadherin) is downregulated, while the transcription factor EOMES and the transmembrane protein N-cadherin are upregulated and cells experience a rearrangement of their cytoskeleton (Cano et al. 2000; Yang and Weinberg 2008; Arnold et al. 2008).

While the primitive streak is initially located in the posterior proximal side of the epiblast, the front of cell differentiation and ingression progressively moves towards the distal tip of the embryo. The precursors of the mesodermal subtypes and the endoderm originate from different areas of the posterior epiblast, specified along the proximal-distal axis (Tam and Behringer 1997). These precursors differentiate then at different timepoints and are subjected to a different signaling environment (Fig. 4, (Lawson, Meneses, and Pedersen 1991). Epiblast cells located in the more proximal side will give rise to extraembryonic mesoderm, which contributes to extraembryonic structures such as the amnion. Moving progressively towards the distal tip of the embryo, the epiblast cells differentiate into cardiac and lateral mesoderm that will form heart and limb buds, and paraxial and axial mesoderm that give rise to the trunk and notochord, essential for the formation of the vertebral column. The last tissue to migrate in the distal tip of the embryo is the endoderm. The cells that do not ingress the primitive streak differentiate as ectoderm (Arnold and Robertson 2009; Kinder et al. 1999; Lawson, Meneses, and Pedersen 1991).

Signaling molecules in mouse mesoderm differentiation

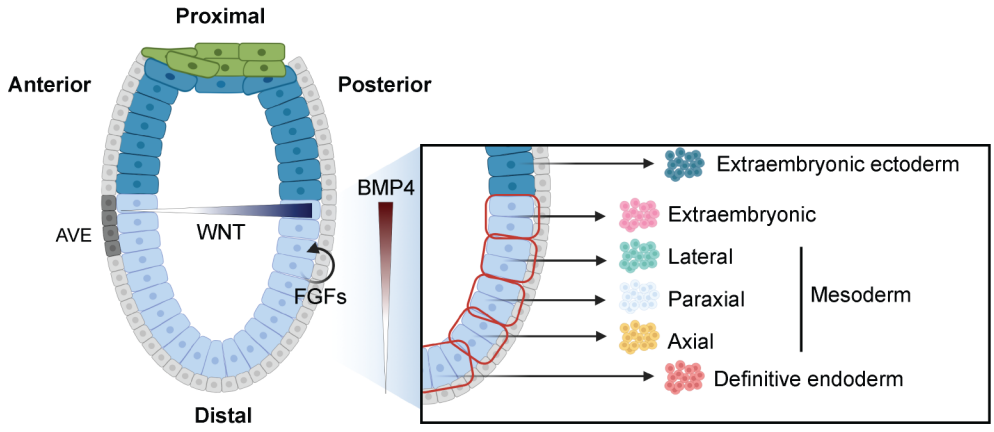


Fig. 4: Embryo fate map for mesoderm differentiation

Scheme of the pre-gastrulating E5.0 embryo with the fate map of the mesoderm subtypes. AVE: Anterior Visceral Endoderm. Created with BioRender.com.

During gastrulation, the embryonic cells are exposed and influenced by signaling molecules and inhibitors produced by the signaling centers. In the mouse, the AVE is only one of the tissues producing signaling molecules that act on the epiblast during gastrulation. The extraembryonic ectoderm is also a source of signaling molecules, in particular BMP4, which is thought to function as a positional information gradient in the posterior part of the epiblast (Fig. 4, Morgani & Hadjantonakis, 2019; Tam & Loebel, 2007). The visceral endoderm is instead a source of low levels of FGF and in the posterior part of the embryo it produces high levels WNTs before and during gastrulation (Crossley and Martin 1995; Haub and Goldfarb 1991; Rivera-Pérez and Magnuson 2005; Morgani and Hadjantonakis 2019).

While the epiblast itself is a source of signaling molecules, such as the uncleaved NODAL, WNTs and FGFs, it is not classically considered an active tissue. On the contrary, it is thought to be patterned from the surrounding tissues and produce its own regionalized signaling after receiving external cues (Tam and Loebel 2007; Morgani and Hadjantonakis 2019). There are some studies *in vivo* and *in vitro*, however, that suggest that cell-cell crosstalk and self-patterning might arise in uniform groups of epiblast cells (Standley, Zorn, and Gurdon 2001; Moris et al. 2020; Warmflash et al. 2014).

FGFs, BMP4 and WNTs are key signaling proteins involved in the differentiation and patterning of the mesoderm. They are expressed by specific extraembryonic tissues and the epiblast cells themselves at precise developmental stages to form the cell lineages precursors of the mesoderm-derived tissues and organs.

WNT signaling

The *Wingless iNTegrated* family (WNT) is composed by 11 conserved members among vertebrates, with most mammalian genomes possessing 19 WNT genes in 12 conserved subfamilies (Clevers and Nusse 2012). WNT are ~40kDa proteins produced in the endoplasmic reticulum, where they are modified by the addition of lipids, essential for their secretion and efficient signaling. After secretion, WNT proteins adhere to proteoglycans of the extracellular matrix, which reduce their diffusion: WNT rarely functions as a long-range morphogen and it appears to mediate signaling between cells in close proximity (Franch-Marro et al. 2008; Clevers and Nusse 2012).

In the canonical WNT signaling, the interaction between WNT proteins and the receptor complex formed by Frizzled and an LRP5/6 protein enables the receptors to bind and inhibit part of the β-catenin degradation complex, such as Axin, GSK3 and Disheveled. β-catenin is then able to translocate into the nucleus, bind to the LEF/TCF transcription factor and activate WNT targets. In the absence of WNT, β-catenin is degraded thanks to a destruction complex formed by Dvl, APC, Axin and Gsk3, which phosphorylate and ubiquitinate β-catenin (Fig. 5, Nusse, 2012). In two different non-canonical WNT signaling, WNT proteins can cause rearrangements in the actin and microtubule cytoskeleton or the release of calcium from intracellular stores. These alternative signaling are mediated by different receptors and intracellular proteins and are important for cell shape, division plane, movement and differentiation. WNTs are implicated in the activation of different cellular responses, which could intersect within the same cell to elicit a response to the WNT signal (van Amerongen and Nusse 2009; Nusse 2012).

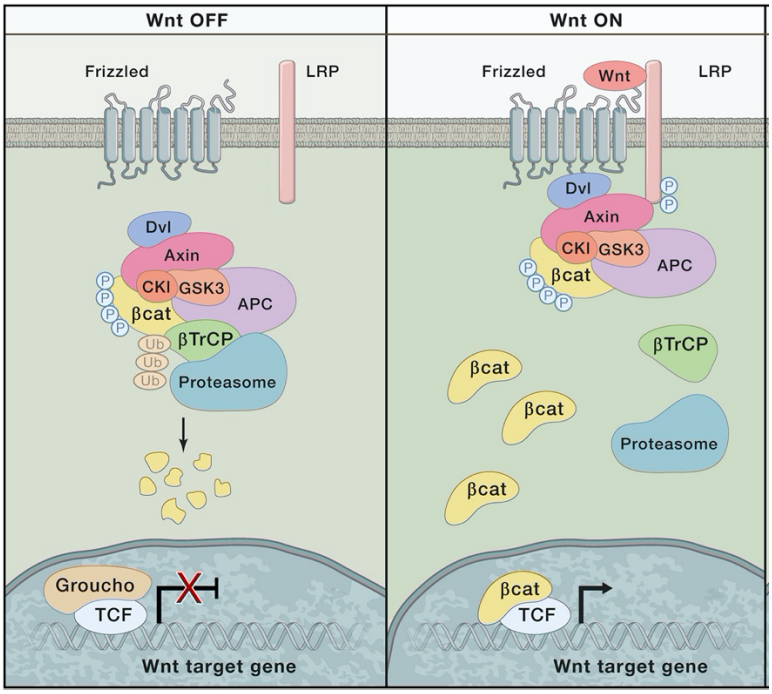


Fig. 5: Canonical WNT signaling.

Modified from Clevers & Nusse, 2012

WNT ligands presents a spatial pattern at the beginning of gastrulation, with a posterior to anterior gradient (Fig. 6). This gradient is sustained in the posterior side by a BMP4 signaling and inhibited in the anterior side by the AVE, which secretes the WNT inhibitors Cerberus and Dickkopf (Liu et al. 1999; Perea-Gomez, Rhinn, and Ang 2001; Tam and Loebel 2007; Ben-Haim et al. 2006; Yoon et al. 2015). The WNT ligands expressed in the posterior epiblast are WNT3a, WNT5a and WNT5b: WNT5a and WNT5b are restricted more proximally in the early primitive streak, in the cells fated to differentiate into extraembryonic mesoderm or proximal mesoderm, while WNT3a is expressed more distally in the streak, in the cells that will give rise to other mesoderm subtypes (Takada et al. 1994). The posterior visceral endoderm bordering the future primitive streak expresses as well one WNT ligand, WNT3, and this area progressively increases to include the posterior epiblast portion of the embryo (Fig. 6, Rivera-Pérez & Magnuson, 2005). This posterior WNT signaling sustains the expression of the pan-mesoderm market T/Bra (Yamaguchi, Takada, et al. 1999; Arnold et al. 2000).

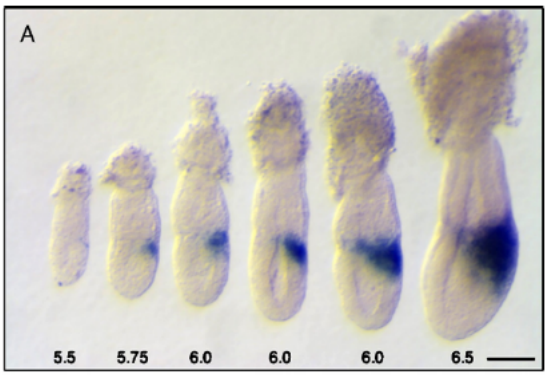


Fig. 6: WNT3 expression in the mouse embryo around gastrulation.
From Rivera-Pérez & Magnuson, 2005

WNT signaling is crucial for mesoderm differentiation, as seen when WNT receptors or specific ligands are inhibited. *In vivo* mutation of the WNT receptors LPR5 and LPR6 show the complete absence of mesoderm and definitive endoderm, while retaining at least one WT copy of LPR5 ($LPR5^{+/-}$ $LPR6^{-/-}$) still determines mesoderm migration failures and deficiency of paraxial mesoderm (Kelly, Pinson, and Skarnes 2004). Embryos with the mutation of the single ligand WNT3 also show a severe phenotype: they develop normally until the pre-gastrulation phase and have a normal AVE, but do not initiate gastrulation and the germ layer differentiation is blocked and the epiblast cells differentiate as ectoderm (Liu et al. 1999). In case WNT3 is present in the epiblast, but not in the VE, the gastrulation is delayed, but not impaired (Yoon et al. 2015). WNT3a mutant embryos do undergo gastrulation, but they show a severe phenotype with a shortened trunk, missing somites and a disorganized neural tube (Takada et al. 1994). In *in vitro* models, the inhibition of WNT hinders the expression of T/Bra and the differentiation of mesoderm, while its ectopic expression can stimulate EMT and the expansion of the primitive streak (Morgani, Metzger, et al. 2018).

BMP signaling

The *Bone Morphogenetic Proteins* family (BMP) is a subgroup of the Transforming Growth Factor β family, which also include NODAL and Activin. The BMPs are secreted growth factors and can act as morphogens in the developing embryos, but they are also involved to regulate cell division, migration and apoptosis (Hogan 1996). The BMP proteins bind to a type II TGF-β receptor, triggering its hetero-oligomerization with a type I TGF-β receptor and causing its phosphorylation (Fig. 7). The activated

type I receptor then phosphorylates a subgroup of the receptor-activated SMAD proteins (R-smad): for the BMP signaling, these are SMAD1, SMAD5 and SMAD8. The SMAD proteins can then translocate into the nucleus after forming a transcription factor complex with the common mediator SMAD4 (Co-smad, Miyazawa & Miyazono, 2017).

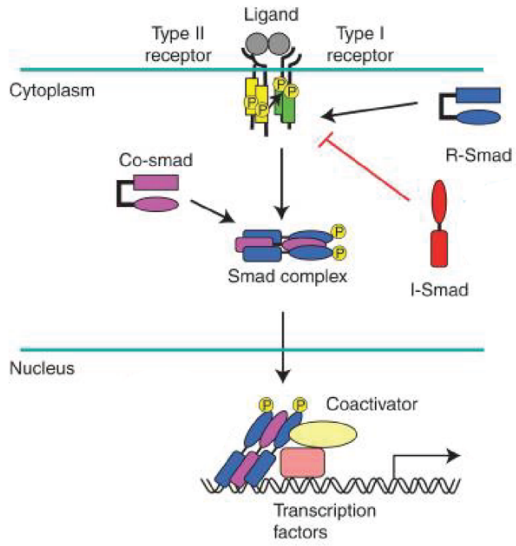


Fig. 7: BMP signaling.

From Miyazawa & Miyazono, 2017

Similarly to the WNT signaling, BMP is confined in the posterior part of the epiblast. However, BMP signaling also presents a proximal to distal gradient (Morgani and Hadjantonakis 2019). The formation of this spatial pattern is dependent on several factors, particularly the BMP4 source, the BMP inhibitors and the localization of BMP receptors. At the onset of gastrulation, BMP4 is only produced outside of the epiblast, in the extraembryonic ectoderm, and it is sustained by the NODAL precursor expressed by the epiblast cells (Ben-Haim et al. 2006; Tam and Loebel 2007). However, as gastrulation proceeds, BMP4 starts to be expressed by the epiblast cells themselves, in the posterior primitive streak (Winnier et al. 1995). BMP localization is restricted in the posterior side of the embryo by the BMP inhibitors produced by the AVE and it presents a proximal-distal gradient along the prospective primitive streak (Winnier et al. 1995; Tam and Loebel 2007; Z. Zhang et al. 2019).

A mechanism for the formation of BMP4 gradient relies on the BMPR localization and the tight junctions between epiblast cells. BMP4 is released by the ExE in the pre-

amnionic cavity, from which it can ingress in a channel between ExE and epiblast cells. Once the ligand has entered the channel, it diffuses in the interstitial space between epiblast and the VE, where the BMP receptors then capture the signaling proteins and determine the formation of the gradient (Fig. 8, Zhang et al., 2019).

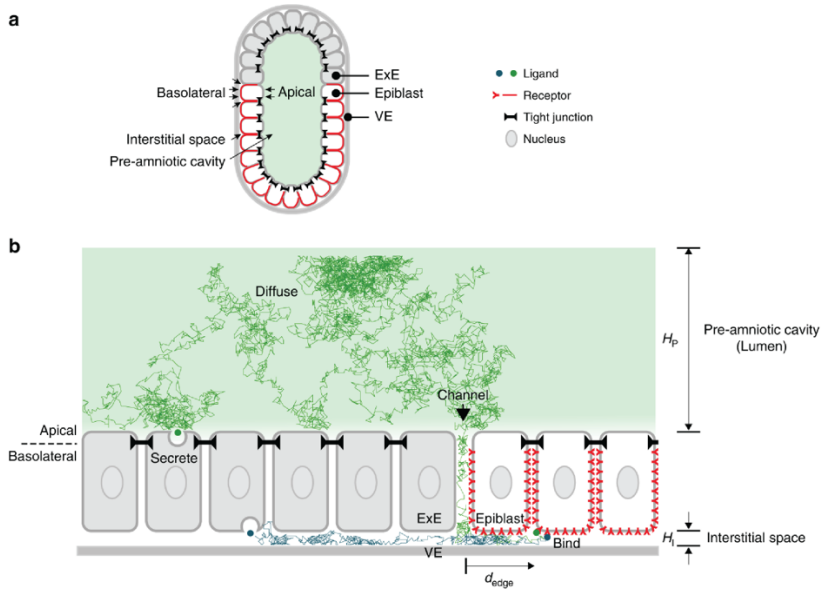


Fig. 8: BMP4 gradient formation in the mouse embryo.

From Zhang et al., 2019

BMP is essential for gastrulation: embryos with a mutation in several BMP ligands and receptors can develop normally only until the pre-gastrulation stage, when their development is arrested; most embryos do not form a primitive streak and have a reduced or no mesoderm (Winnier et al. 1995; Gu et al. 1999; Beppu et al. 2000; Mishina et al. 1995; H. Zhang and Bradley 1996). *In vitro*, BMP4 is required to specify proximal mesoderm, however, the distal mesoderm and endoderm can be differentiated by culturing cells with WNT and Activin A or NODAL (Morgani, Metzger, et al. 2018). BMP4 is then proposed to act as a morphogen in the differentiation of cells specified in the primitive streak, controlling trajectories of cell differentiation in a concentration-dependent manner (Manfrin et al. 2019; Tam and Loebel 2007).

FGFs signaling

The *Fibroblast Growth Factor* family (FGF) include 22 members, divided into seven subfamilies. The different proteins have a size of ~ 150-300 aa, and have a conserved

core of ~120 amino acids with 30-60% identity. Most of the FGF proteins act as secreted factors, generally acting as autocrine or paracrine signaling, with the exception of one subfamily possessing endocrine functions and one comprised of intracellular non-signaling proteins (Itoh and Ornitz 2010; Ornitz and Itoh 2015). The secreted FGF proteins can elicit a response in the target cell via different signaling cascades: two critical cascades are the Raf-MEK-ERK and the JAK-STAT cascade (Fig. 9). In the Raf-MEK-ERK cascade, the tyrosine kinase receptor recruits a GTP exchange factor to activate Ras, which in turns activates the protein Raf, which starts a phosphorylation cascade with MEK1/2 and ERK1/2. Activated ERK1/2 will then translocate into the nucleus and phosphorylate a variety of transcription factors (Lake, Corrêa, and Müller 2016). In the JAK-STAT cascade, the receptor is bound to a JAK kinase, which phosphorylates and activates one of the STAT transcription factors, which can translocate into the nucleus (Ornitz and Itoh 2015).

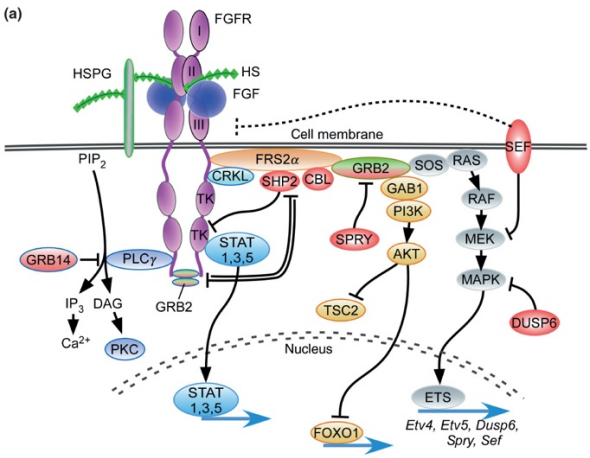


Fig. 9: FGF signaling.
From Ornitz & Itoh, 2015

FGF signaling is essential in early mouse development: cell-cell crosstalk mediated by FGF4 ensures that the right proportion of pluripotent cells differentiate into epiblast or primitive endoderm and its mutation determines the early death of the embryo (Schröter et al. 2015; Raina et al. 2021; Tam and Loebel 2007).

During gastrulation, FGF ligands, especially FGF8, FGF4, FGF3 and FGF5, are produced by epiblast cells themselves and are elevated in the primitive streak and nascent meso-endoderm in the posterior side of the embryo (Tam and Loebel 2007;

Morgani and Hadjantonakis 2019; Maruoka et al. 1998; Morgani, Saiz, et al. 2018). FGF8 in particular is strongly expressed in the prospective streak prior to gastrulation and in the epithelial component of the primitive streak during gastrulation. The expression of this ligand is then dramatically reduced as the cells experience EMT and migrate from the point of ingression in the primitive streak (Crossley and Martin 1995).

The chemical inhibition of FGF signaling during gastrulation suggests the importance of FGF in mesoderm patterning, as it causes the downregulation and displacement of several mesoderm markers, as well as the impairment of cell migration (Oki, Kitajima, & Meno, 2010). The EMT defects have been reported also for FGFR1 mutations *in vivo*, also causing the inhibition of WNT signaling (Ciruna and Rossant 2001; Yamaguchi et al. 1994). Additionally, FGFRi mutant embryos displayed strong defects in pre-somitic mesoderm formation and the elongation of the embryo along the anterior-posterior axis (Yamaguchi et al. 1994).

Out of the single FGFs involved in gastrulation, FGF8 exhibit a particularly strong phenotype: in $FGF8^{-/-}$ mutants, epiblast cells move into the primitive streak and undergo EMT, but most cells then fail to move away from the streak, creating an abnormal thickening in the posterior side (Sun et al. 1999). Targeted disruption of FGF8 has also been shown to impair FGF4 expression in the streak, but not FGF3 and FGF5 (Sun et al. 1999). Conditional inactivation of FGF8 and FGF4 around E8.5, a later stage corresponding to embryonic elongation, causes posterior truncation and strong defects in the paraxial mesoderm, while lateral mesoderm was not as affected; they also reported a reduction in the expression of WNT3a, WNT5b and T/BRA (Boulet and Capecchi 2012). FGF3 and FGF5 exhibit a milder phenotype in mutant mice: FGF3-deficient mice often do not survive into adulthood and exhibit defects in the mesoderm derived from the posterior primitive streak, which also causes a deform tail, whereas FGF5-deficient mice reveal abnormal hair growth (Mizuno et al. 2011; Mansour, Goddard, and Capecchi 1993; Hébert et al. 1994).

In the frog Xenopus Laevis, a homolog of FGF4 has been reported to mediate the differentiation of a paraxial mesoderm through a community effect: only when a large group of precursors of muscle cells are in close contact, their expression of the homolog of FGF4 can supply the cells a stable concentration of this signaling protein

over a period of several hours and differentiate into muscle cells (Standley, Zorn, and Gurdon 2001). Without this close contact, the cells do not express muscle genes, while being equally viable (Gurdon et al. 1993).

The role of FGFs in the differentiation of distinct mesoderm subtypes is still unclear, whether it acts to favour a cell type over another or if it simply drives the epithelial to mesenchymal transition and the differentiation of all mesoderm subtypes (Morgani and Hadjantonakis 2019).

Signaling interactions during mesoderm differentiation

FGFs, BMP4 and WNTs signaling are critical for gastrulation, as they have been implicated in the differentiation and patterning of the mesoderm, as well as the EMT and cell migration. Some interactions between these signaling have been established in the gastrulating embryo or in different cell types and species, but it's unclear what is their combined contribution to the differentiation of distinct mesoderm lineages.

BMP4 has been found to sustain the differentiation of mesoderm in the posterior side of the embryo by inducing the expression of WNT ligands and even activate β-catenin independently of the WNT ligand (Massey et al. 2019; Ben-Haim et al. 2006). Moreover, BMP4 expression in the extraembryonic ectoderm is maintained by NODAL precursors produced by the epiblast, creating a positive regulatory interaction composed by NODAL precursor -> BMP4 -> WNT3 (Ben-Haim et al. 2006; Massey et al. 2019; Morgani and Hadjantonakis 2019).

In vivo mutations of WNT and FGF receptors or ligands during gastrulation suggest a connection also between these signaling. Indeed, WNT may be downstream of FGF or independently regulated in related processes: mutations of the WNT signaling does not affect FGF signaling, but it creates mesoderm defects resembling those found in FGF mutants (Kelly, Pinson, and Skarnes 2004; Morgani and Hadjantonakis 2019). Furthermore, a mutation of FGFR, while not impairing the expression of WNT3a, has been reported to sequester β-catenin, downregulating the downstream WNT signaling and the expression of target genes (Ciruna and Rossant 2001).

During mesoderm differentiation, FGF and BMP signaling have as well been reported to interact. In *in vitro* differentiation protocols, BMP signaling is a key inducer of both mesoderm and extraembryonic lineages. However, culturing the cells with FGF2 can change the differentiation induced by BMP4 from extraembryonic to embryonic lineages through a prolongation of NANOG expression (Bernardo et al. 2011; Yu et al. 2011; Vallier et al. 2009). Moreover, BMP has been shown to interfere with the FGF signaling in the neuromesodermal progenitors in the zebrafish tailbud. Known BMP4 target genes, the ID proteins, can sequester and inhibit the activity of FGFs-dependent bHLH transcription factors. This inhibition of bHLH then causes a reduction of a distal mesoderm subtype, the paraxial mesoderm, and the increase of the more proximal lateral mesoderm (Row et al. 2018).

BMP4, FGFs and WNTs have also been shown to interact directly at the level of signal transduction: in presence of WNT signaling, the FGFs effector ERK1/2 determine the phosphorylation of the the linker region of BMP downstream effectors SMAD1 and SMAD4 and promote the peak transcriptional activity of SMAD4 and the duration of the BMP/Smad1 signal. However, if the WNT signaling is not active, one of the proteins of the β-catenin destruction complex, GSK3, recognizes and phosphorylates the pre-phosphorylated SMADs and primes them for polyubiquitination and degradation (Fig. 10, Demagny, Araki, & De Robertis, 2014; Fuentealba et al., 2007; Kretzschmar, Doody, & Massagué, 1997; Pera, Ikeda, Eivers, & De Robertis, 2003).

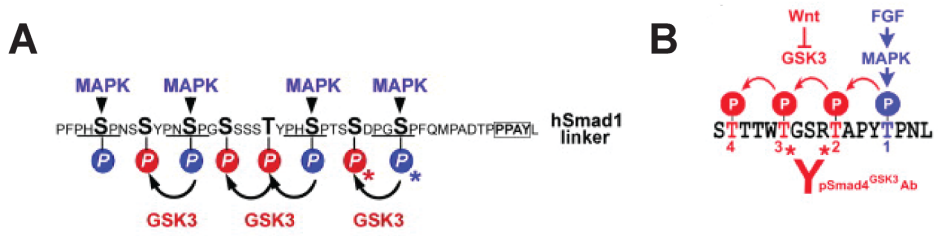


Fig. 10: Smad1 and Smad 4 phosphorylation sites in their linker region.
A) From Fuentealba et al., 2007. B) Demagny, Araki, & De Robertis, 2014

In vitro models for gastrulation

The role of cell signaling during gastrulation has been intensively studied *in vivo* through the use of mutagenesis, signaling reporters and the use of small molecules. However, the extreme complexity of *in vivo* embryo and the technical difficulties to

manipulate and study live embryos gave rise to diversified *in vitro* gastrulation models. *In vitro* systems offer more controlled and accessible setup: reporter or knockout cell lines are more available or more easily produced and cell density and spatial arrangement can be a priori defined, as well as the signals the cells are going to be exposed to, in terms of timings and concentrations. These settings also allow to more easily analyse the mRNA or protein content of the cells, but also acquire images, analyse markers expressed by the cells and possible cell clustering.

In recent years, several *in vitro* systems with a precise morphology have been established to recapitulate different aspects of gastrulation: examples are micropatterned colonies, embryoid bodies and stem cell aggregates (Warmflash et al. 2014; Moris et al. 2020; Harrison et al. 2017). Micropatterned colonies are established as pluripotent cells cultured in circular surfaces of defined size: these colonies display radially dependent responses dependent on cell-cell crosstalk, paracrine signaling and receptor localization (Manfrin et al. 2019; Warmflash et al. 2014; Nemashkalo et al. 2017). Embryoid bodies and stem cell aggregates are tridimensional cell aggregates respectively composed by only pluripotent stem cells and a mix of pluripotent stem cells and cells corresponding to one or more extraembryonic structures (Moris et al. 2020; Harrison et al. 2017; Sozen et al. 2018). Embryoid bodies are useful to study how pluripotent cells can self-organize to cavitate, change shape or differentiate derivatives of the three germ layers. Cell aggregates with more than one cell lineage, in contrast, are uniquely apt to study the interaction of embryonic and extraembryonic tissues and how this crosstalk can generate specific structures. The ability of these models to self-organize suggests that cell-cell crosstalk, together with external signals, strongly contribute to cell patterning and opens the possibility to study the mechanisms of these critical cell-cell communications.

While these *in vitro* systems can reproduce specific aspects of the embryo, many other models rely instead on a 2D cell culture to understand how endogenous and exogenous signals impact cell differentiation during gastrulation. They are, in fact, simple and versatile systems for studying developmental events that are less accessible to experimentation. Their signaling and differentiation outcomes can more easily be studied both at a single cell level and in cell populations (Mendjan et al. 2014; Kattman et al. 2010; Rao et al. 2016; Loh et al. 2016; Vallier et al. 2009). These

protocols can employ mouse embryonic stem cells (mESCs) or the primed pluripotent mouse epiblast stem cells (EpiSCs). mESCs are isolated from the inner cell mass of the blastocyst and can be differentiated into embryonic lineages, as well as extraembryonic lineages such as the primitive endoderm (Weinberger et al. 2016; Hermitte and Chazaud 2014). mEpiSCs are isolated from the post implantation epiblast and they have different characteristic from ESCs. The EpiSCs inactivate one X chromosome in female cells, upregulate the DNA methylation levels and have different set of signals that induce differentiation or the maintenance of pluripotency (Weinberger et al. 2016). These cells are able to differentiate into all three germ layers as well as the primordial germ cells and correspond to the epiblast cells just prior to gastrulation (Weinberger et al. 2016; Sumi et al. 2013).

The limitation of the *in vitro* systems is that the pluripotent cells and the phenomena observed cannot be completely equivalent to the embryonic cells and their full range of cell lineage interactions of the developing embryo (Rossant and Tam 2021). Whenever possible, it is important to validate and compare these findings to the *in vivo* embryos.

Objectives

The main objective of this thesis was to explore how do cells process and integrate distinct signals to pattern the developing mesoderm. I focused here the role of BMP4, produced *in vivo* from external signaling centers, and of FGF, produced *in vivo* by the cells themselves.

In this work I first set out to develop a simple 2D cell culture protocol based on WNT, BMP and FGF signaling to obtain differentiated mesoderm *in vitro*. Compared to *in vivo* or 3D models, this system would allow to more easily control the signals applied to the pluripotent cells, as well as analyse their impact on phenotype or signal transduction. This system can be used to explore if different concentrations of BMP4 would affect the differentiation of specific mesoderm subtypes, consistently with the proposed morphogenic effect of this ligand. From this frame of reference, I could then determine the possible role of exogenous and endogenous FGF signaling by titrating the ligand and employing a FGF signaling inhibitor. The cells differentiated with these culture conditions could be employed to create transcriptomic libraries with a single cell resolution. By analysing the expression of distinct lineage markers, ligands, and target genes, the transcriptomic libraries could uncover possible effects of the FGF concentrations on mesoderm patterning or cell-cell communication. The integration of the libraries with published reference datasets from gastrulating mouse embryos could aid the characterization of the identities and developmental stages of the cells obtained *in vitro*. This work intends to test the morphogenic effect of BMP4 on mesoderm differentiation *in vitro*, uncover a possible role of FGF signaling in the differentiation of mesoderm or on distinct mesoderm identities and propose a mechanism to explain the observed phenotypes.

MATERIALS AND METHODS

Cell culture

Mouse pluripotent cell lines

The wild type mouse epiblast cell line (EpiSCs) employed here was obtained from the laboratory of Dr. Jennifer Nichols in the Wellcome-MRC Cambridge Stem Cell Institute. The cells were directly derived from the epiblasts of E6.5 embryos with a 129 background, cultured on fibronectin-coated dishes with standard EpiSC chemically defined medium: N2B27 with FGF2 and ActivinA.

The mouse Sox1-T/Bra reporter cell line (SBR line) was obtained from the laboratory of Dr. David Suter from École polytechnique fédérale de Lausanne. This line contains a P2A-eGFP cassette downstream from the Sox1 sequence and a P2A-mCherry cassette downstream from the Brachyury sequence (Deluz et al. 2016). The cell line was acquired as embryonic stem cells (mESCs), cultured in serum/LIF + 2i medium.

The mouse Spry4 H2B-Venus reporter cell line (Morgani, Saiz, et al. 2018) was acquired as mESCs, cultured in serum/LIF medium.

Cell culture maintenance

Mouse embryonic stem cells (mESCs) were cultured on T25 tissue flasks coated with 0.1% gelatine (Sigma Aldrich). They were maintained with serum/LIF medium or serum/LIF + 2i medium, depending on their previous culture medium. Epiblast stem cells (EpiSCs) were cultured on T25 tissue flasks coated with 20 µg Fibronectin (20 µg/ml, Merck). EpiSC were cultured in a serum-free medium, the FAX medium, based on home-made N2B27. Cells were passaged every two days: 1 ml of Accutase (Merck) was used to detach the cells, which were then centrifuged and replated at a density of 8000 to 10000 cells per cm^2. For the cryo-preservation of EpiSCs, cell pellet was resuspended in freezing medium, containing 65% of maintenance medium, 25% KSR and 10% DMSO as cryo-protectant to prevent ice crystal formation (Wagner and Welch 2010). Cells were stored in CryoPure screw cap tubes and transferred in a CoolCell freezing container, placed in a -80°C freezer overnight and then transferred for long term storage in a -150°C freezer.

To differentiate mESCs to EpiSCs and create stable cell cultures, mESCs cells were initially thawed in gelatin-coated T25 flasks with the same ES medium they were cultured in before acquiring them. The cells were then passaged using 0.05% Trypsin (PAN Biotech) in gelatin-coated T25 flasks and the ES medium was replaced with FAX medium a day after splitting. To ensure a homogenous cell differentiation to EpiSCs, cells were kept in culture with FAX medium on Fibronectin-coated T25 flasks for 8-10 passages before starting experiments.

Cell culture media

Serum/LIF medium	500 ml GMEM (Thermo Fisher Scientific) supplemented with 10% FBS (Sigma Aldrich), 5 ml non-essential amino acids (Thermo Fisher Scientific), 2 mM GlutaMAX (Thermo Fisher Scientific), 1 mM sodium pyruvate (Thermo Fisher Scientific), 0.1 mM β-mercaptoethanol (Gibco), 10 ng/ml LIF (MPI protein expression facility)
Serum/LIF + 2i medium	Serum/LIF medium supplemented with 1 μM PD0325901 (Selleck Chem) and 3 μM Chiron (CHIR00201, Merck)
N2B27 medium	500 ml DMEM/F12 (PAN Biotech), 500 ml Neurobasal medium (PAN Biotech), supplemented with 5 ml N2 and 10ml B27 (Thermo Fisher Scientific), 335 μl BSA (7.5%, Gibco), 100 μM β-mercaptoethanol (Gibco) and 5 ml Glutamax (Thermo Fisher Scientific)
FAX medium	N2B27 with the addition of 12 ng/ml FGF2 (Cell Guidance System), 25 ng/ml ActivinA (Cell Guidance System) and 20 μM XAV939 (Cell Guidance System)

Mesoderm differentiation protocol

The mesoderm differentiation protocol employed EpiSCs cells. The cells were split the afternoon before the start of the experiment with a concentration of 1300-2000 cells/cm^2. The following morning the medium was changed to N2B27 with the appropriate supplements, according to the experiment layout: this media change marked the beginning of day 1 of the experiment. The experiments lasted 3 days, with a media change every morning. Commonly used supplements were: 1 μM Chiron

(CHIR99201, Merck), 0-128 ng/ml BMP4 (PeproTech), 0-96 ng/ml FGF2 (Cell Guidance System), 0-96 ng/ml FGF4 (PeproTech), 30 nM AZD4547 (Selleckchem).

Immunostaining

Cells prepared for immunostainings were cultured on µ-Slide 8 well Chamber (Ibidi).
At the end of the experiment, cells were washed with room temperature PBS, then fixed with 4% PFA for 15 minutes at room temperature. Cells were then washed 3 times with PBT-BSA for 15 minutes: this solution is composed by PBS with 0.1% Triton (Serva) and 1% BSA (Gibco) to permeabilize and block the cells. The primary antibodies were diluted in this solution at the appropriate concentration and the incubation was carried out overnight at 4°C on a rocking platform. The following day the cells were washed with PBT-BSA and incubated in the dark for 1-2 h with PBT-BSA supplemented with the appropriate secondary antibodies (Alexa Fluor Dyes, Thermo Fisher Scientific) and 1 µg/ml Hoechst 33342 to counter-stain the nuclei. Cells were then washed with PBS and covered with homemade mounting medium, composed of 80% glycerol and 4% N-propyl gallate.
The primary antibodies I used were: a-Nanog 1:200 (Thermo Fisher Scientific, eBioMLC-51), a-T/Bra 1:200 (R&D, AF2085), a-Hand1 1:200 (R&D, AF3168), a-Tbx6 1:200 (R&D, AF4744), a-Cdx2 1:250 (BioGenex, MU392A-5UC).

Microscopy

Live brightfield imaging was performed on a Leica MC 170 HD microscope.
The fluorescent imaging was performed with a confocal laser-scanning microscope (Leica Sp8 confocal microscope). The fluorescently labelled proteins or endogenous fluorescent proteins were excited with appropriate wavelength of the white light laser of the microscope and the fluorescence signal was collected through an oil immersion objective 63x, collecting the emission light of appropriate wavelength width. Bigger fields of view were imaged employing the multi-tile settings of the Leica software, with 10% overlap between single tiles. The images were later reconstructed using the BigStitcher plug-in in Fiji (https://github.com/PreibischLab/BigStitcher/ Schindelin, Arganda-carreras, et al., 2012).

Custom trained nuclear segmentation model

The analysis of the immunofluorescent images required to identify and segment the nuclei of the cells. Images often contained a high number of nuclei in close contact, so I automated the segmentation process using the Fiji plug-in StarDist (Schmidt et al. 2018). StarDist is a nuclei detection algorithm for microscopy images that assumes star-convex shapes objects, that can be used with pre-trained or custom-trained models. I trained my own model, starting with the generation of a paired training dataset made of images of Hoechst-stained nuclei and their corresponding masks. These masks were created by using the pre-trained StarDist model '*Versatile (fluorescent nuclei)*' and manually correcting the regions of interest found by the algorithm to obtain the training target (Fig. 11). I chose the pre-trained model that best fitted my data and later used the same model as a starting point to train the custom model.

The training was performed on the StarDist 2D ZeroCostDL4Mic platform (v 1.12), that allows to leverage the resources of a Google Colab notebook to run a deep learning network (https://github.com/HenriquesLab/ZeroCostDL4Mic/wiki, Von Chamier et al., 2020). The StarDist model was trained for 40 epochs on 7 paired image patches, starting from the pretrained model '*Versatile (fluorescent nuclei)*'.

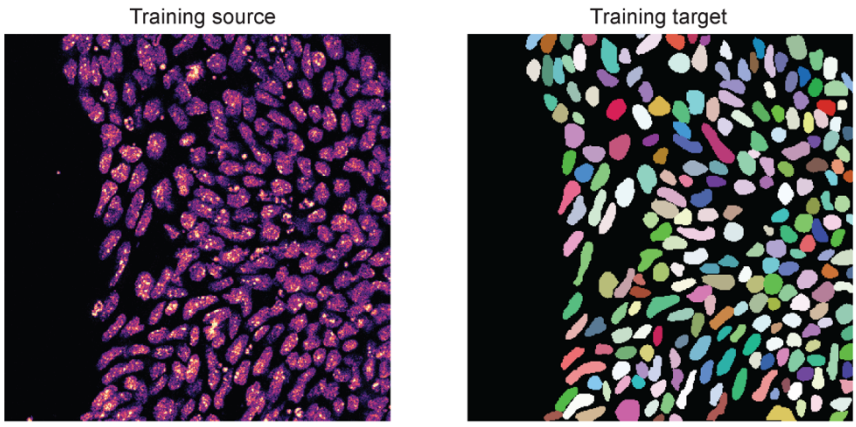


Fig. 11: Training pair example

Nuclear segmentation and analysis

The nuclear segmentation and the quantification of the fluorescence in each channel was automated with a custom script in Fiji/ImageJ. Briefly, the script made a copy of the channel with the nuclear staining and run the StarDist plugin with the custom-trained model to retrieve the regions of interest (ROI). The script then projected the

ROIs on the image with the nuclear channel and it paused to allow a manual check of the automatically-identified nuclei. Finally, it projected the mask onto all channels to retrieve the fluorescent measurements. Circa 8 images were retrieved per condition, capturing the variability of the expression of the chosen differentiation markers.

The recovered data was then analysed using custom scripts in Python, available at https://github.com/gattigli/mesoderm_diff. Briefly, for each labelled protein I retrieved the single cell fluorescence of every image and sample of the dataset. This joined dataset contained the entire spectrum of fluorescence in every sample and was ideal to define the threshold between positive and negative cells. To define the threshold, I used a Gaussian Mixture Model with two components: cells were defined as positive if they belonged to the positive group of cells with a probability > 70% (Fig. 12). The lowest fluorescence connected to this group was considered as threshold between positive and negative cells. That value was then employed to calculate, in each condition, the percentage of marker positive cells.

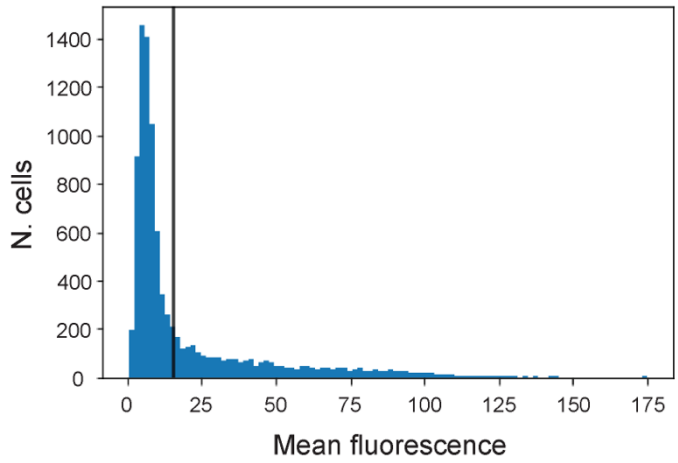


Fig. 12: Representative analysis and thresholding of Tbx6

Black line represents the lower threshold of the cells clustered in the 'positive' group.

RT-qPCR

RNA isolation

RNA was isolated with the TRIZOL reagent (Thermo Fisher Scientific). Live cells were detached using Accutase, then centrifuged to obtain a pellet. The pellet could be immediately processed or snap-frozen in liquid nitrogen and stored for long term at

-80°C. The second passage required the pellet to be resuspended in 1ml TRIZOL reagent and incubated for 5 minutes to allow the complete dissociation of nucleoprotein complexes. 0.2 ml Chloroform was added and then the samples were vigorously resuspended and centrifuged at 12.000 × g for 15 minutes at 8°C. The colorless upper aqueous phase containing RNA was retrieved and the RNA was precipitated using 0.5 ml of isopropyl alcohol. Solution was incubated for 10 minutes and centrifuged at 12.000 × g for 15 minutes at 8°C. the RNA pellet was the washed twice with 75% ethanol and then resuspended in DNase/RNase-free water.

Quality control and quantification was performed with a Spectrophotometer (Nanodrop, Thermo Fisher Scientific). The ratio A260/A280 assessed the eventual protein contamination and was deemed acceptable between 1.8 and 2.1, while the ratio A260/A230 showed the presence of salt and it was considered acceptable for values above 2. The RNA was dissolved to a concentration of 100 ng/µl and stored long term at -80°C.

Primer design, RT-qPCR and quantification

To determine the mRNA levels of specific genes of interest, I performed a reverse transcription quantitative polymerase chain reaction (RT-qPCR). When no suitable primer couple were found in Primerbank (Spandidos et al. 2009), I designed the primers with the online tools Primerblast or Primer3 (Ye et al. 2012; Untergasser et al. 2012), selecting primer pairs with a product length ~ 200bp, the lowest self-complementarity and self 3'complementarity (Table 1). RT-qPCR was performed with the Luna Universal One-Step RT-qPCR Kit (E3005, New England Biolabs) according to manufacturer's instruction and loaded on the iQ5 Real-Time PCR System (Bio-Rad). Every quantification was performed in three replicates per condition for the housekeeping genes and two replicates per condition for the marker genes. If the differences between Ct values for a given transcript and a matched negative control (NRT, same primer pair without reverse transcription) were not statistically significant using a two-tailed Student's t-test, then the transcript was classified as not detected. Transcript levels calculated as $efficiency^{-(condition-control)}$, then normalized to the geometric mean of the expression of the two house-keeping genes *Actin* and *TATA box–binding protein* (Tbp).

Gene	Forward primer	Reverse primer	efficiency
Actin*	GCAGGAGTACGATGAGTCCG	ACGCAGCTCAGTAACAGTCC	2.00
Tbp*	GTGCCAGATACATTCCGCCT	CAAGCTGCGTTTTTGTGCA	2.00
T/Bra	CTGGGAGCTCAGTTCTTTCG	GTCCACGAGGCTATGAGGAG	1.94
Hand1	GAGGAGAGGAAAGGACGCAG	CTCGGCGGGAAGTGAACATA	2.00
Gata6	TGGGAGCCATTTGGTCTATC	GACCTCAGATCAGCCACGTT	2.00
Tbx6	GGCAGCTCCATCTGTACCAT	ACCGAGGCTCAGTACATTGG	1.92
Msgn1	CCAGAAAGGCAGCAAGTCA	GAGGAGGTCTGTGAGTTCCC	2.00
Foxa2	CATTACGCCTTCAACCACCC	GGTAGTGCATGACCTGTTCG	1.92
Shh	AAAGCTGACCCCTTTAGCCTA	TTCGGAGTTTCTTGTGATCTTCC	2.00

Table 1: RT-qPCR primer couples

** Housekeeping genes*

Western blot

Cells in culture were washed twice with ice-cold PBS supplemented with 1mM orthovanadate and incubated with lysis buffer. The buffer was prepared fresh using the complete EDTA-free protease inhibitor cocktail (Roche), phosphatase inhibitor cocktail 2 and 3 (Sigma) and benzonase (Thermo Fisher Scientific), diluted in the commercially available lysis buffer (Cell Signaling). Cells were mechanically detached, collected and snap-frozen in liquid nitrogen. Samples were than analysed for quality control and quantification with a micro-BCA essay (Thermo Fisher Scientific). Lysates were denatured using a standard Laemelli buffer and boiled for 5 minutes at 95°C. Samples were then incubated into ice for 5 minutes and 10 to 20µg protein was loaded on a Bis-Tris SDS gels gel. Gels were run with 1x MOPS buffer (Thermo Fisher Scientific) with fresh sodium bisulphite. After the run, the gels were transferred on methanol-activated PVDF membranes (Millipore) at 40 V for 90 minutes with the NuPage transfer system (Thermo Fisher Scientific) and incubated with primary and secondary antibodies for 1 h each. The membranes were imaged with the Odyssey Infrared Imaging System (LI-COR Biosciences).

Primary antibodies: a-Tubulin 1:5000 (Sigma, T6074), a-pERK1/2 1:1000 (Cell Signaling, 4370S), a-total ERK1/2 1:1000 (Abcam, ab36991)

Secondary antibodies: donkey a-mouse 800 1:500 and donkey a-rabbit 680 1:500, both IRDyes (LI-COR Biosciences).

In-Cell Western and quantification

Cells were seeded at a concentration of 7000 cells/cm^2 to obtain a high cell density at the end of differentiation, needed for the quantification. Cells were grown on 96-well black plates with transparent polystyrene bottom (3340, Corning), coated with Fibronectin. Similarly to the immunostaining, at the end of the experiment the cells were fixed with 4% PFA, then permeabilized and blocked with a solution of PBS with 0.1% Triton (Serva) and 1% BSA (Gibco). Incubation with primary antibodies was performed overnight at 4°C on a rocking platform and the secondary antibodies were incubated in the dark for 1 h. The secondary antibodies were conjugated with infrared-fluorescent molecules and the intensity measurements were performed with the Odyssey Infrared Imaging System (LI-COR Biosciences). The selection of the region of interests in the images was performed with a Fiji plugin, the MicroArray Profile, and the expression of the protein of interest was normalized against the expression of the housekeeping gene Tubulin.
Primary antibodies: a-Hand1 1:500 (R&D, AF3168), a-Tbx6 1:500 (R&D, AF4744), a-Gata6 1:1000(R&D, AF1700), a-Tubulin 1:5000 (Sigma, T6074)
Secondary antibodies: donkey a-goat 800 1:500 and donkey a-mouse 680 1:500, both IRDyes (LI-COR Biosciences).

Flow cytometry

Cells prepared for flow cytometry were grown on 6 wells plates to obtain at least 20000 cells to analyse in each condition.
The SBR cell line expressed cytosolic fluorescent proteins and the flow cytometry was performed on live cells, as the fixation would have lowered the mean fluorescence of the cells. At the end of the experiment, cells were detached from the plate using Accutase, centrifuged to obtain cell pellet and resuspended in PBS with 1% BSA. Before the analysis, cells were passed through a strainer to obtain single cells and avoid clumps. Since the protein of interest was connected to the expression of the

fluorescent protein mCherry, the analysis was performed with the BD FACSAria Fusion flow cytometer (BD Biosciences).
The Spry4 cell line expressed the nuclear-bind fluorescent protein Venus and it was analysed on the flow cytometer as fixed cells. After detaching, cells were fixed with 4% PFA at room temperature for 15 minutes, then resuspended in PBS with 1% BSA and strained to obtain single cells. Since the protein of interest was connected to the expression of the fluorescent protein Venus, the analysis was performed with the BD LSRII flow cytometer (BD Biosciences).
Flow cytometry data was analysed in FlowJo (BD Biosciences).

Single Cell RNA Sequencing

Libraries production and sequencing depth

Cells were differentiated with the standard protocol, treating the cells with 1 μM Chiron, 8 ng/ml BMP4, and 5 different FGF conditions: FGF2 concentration of 0 ng/ml, 0.75 ng/ml, 3 ng/ml and 12 ng/ml or treatment with 30 nM of the FGF receptor inhibitor AZD4547. Sc-RNA sequencing was performed as per manufacturer's instruction (10x Genomics). Briefly, cells were detached with Accutase and counted on a Neubauer chamber. We aimed at 2000 cells recovered per sample, with a concentration of 900 cells/μl. As per manufacturer's instruction, an appropriate dilution of cells was incorporated to the reaction master mix containing enzymes, buffer and oligos, and loaded on the Chromium Next GEM Chip, together with the beads and the partitioning oil. The Chip was then inserted in the Chromium controller machine to encase the single cells in hydrophilic droplets with the barcoded gel beads.
The recovered samples were incubated in a thermocycler to lyse the cells and generate barcoded, full-length cDNA from poly-adenylated mRNA in each droplet. To recover the DNA oligos, the droplets were broken and the samples cleaned to get rid of the droplet reaction mixture, partitioning oil, primers and biochemical reagents. cDNAs were amplified with in a thermal cycler with 12 total number of cycles to generate enough material for sequencing and minimizing PCR amplification artifacts. The samples were then purified again and analyzed for quality control and quantification.
A fixed amount of the sample (255 μl, as per manufacturer's instruction) was then fragmented, end-repaired and A-tailed. The adaptor ligation was performed to optimize

the cDNA amplicon size and add the P5 and P7 oligos used in Illumina bridge amplification. The sample were then cleaned to retrieve the cDNAs. Libraries were analyzed to control the quality and quantify the cDNA before sending them to a company to perform the sequencing.

Sample were analyzed by Novogene with a first shallow sequence of $20*10^6$ reads per sample to assess the dataset and determine the number of cells retrieved per sample. This shallow sequencing was followed by a deeper one between 125 and $180*10^6$ reads per sample ($800*10^6$ total). The different sequencing depth was useful to correct for the different number of single cells retrieved during the libraries production and obtain a similar sequencing depth per cell in each sample. Before quality control analysis, I obtained ~50000 reads per cell in every sample, with a sequencing saturation of ~30% according to the sequencing company.

scRNA sequencing analysis

Using CellRanger (10x Genomics, v4.0.0), the sequenced libraries were demultiplexed, aligned to the mouse genome mm10 (GENCODE vM23/Ensembl 98, from 10x Genomics) and the reads were quantified. The libraries were saved as annotated dataset objects (*anndata*) with extension '.h5ad': *anndata* is a Python package for handling annotated data matrices. Quality control and further analyses were carried out with the Python package *Scanpy* 1.7.0rc1 (Wolf, Angerer, and Theis 2018) according to the template in https://scanpy-tutorials.readthedocs.io/en/latest/ /pbmc3k.html.

All scripts are available at https://github.com/gattigli/mesoderm_diff. Briefly, the 5 anndata files retrieved from CellRanger were first passed through an individual quality control preprocessing, then joined and normalized together. In the preprocessing, the datasets were first analyzed to obtain the number of genes per cell, the total amount of reads per cell and the percentage of mitochondrial DNA detected per cell (Fig. 13). Based on the violin plot distribution of those parameters, I filtered out cells that had less then 2500 genes per cell and more than 10% mitochondrial DNA.

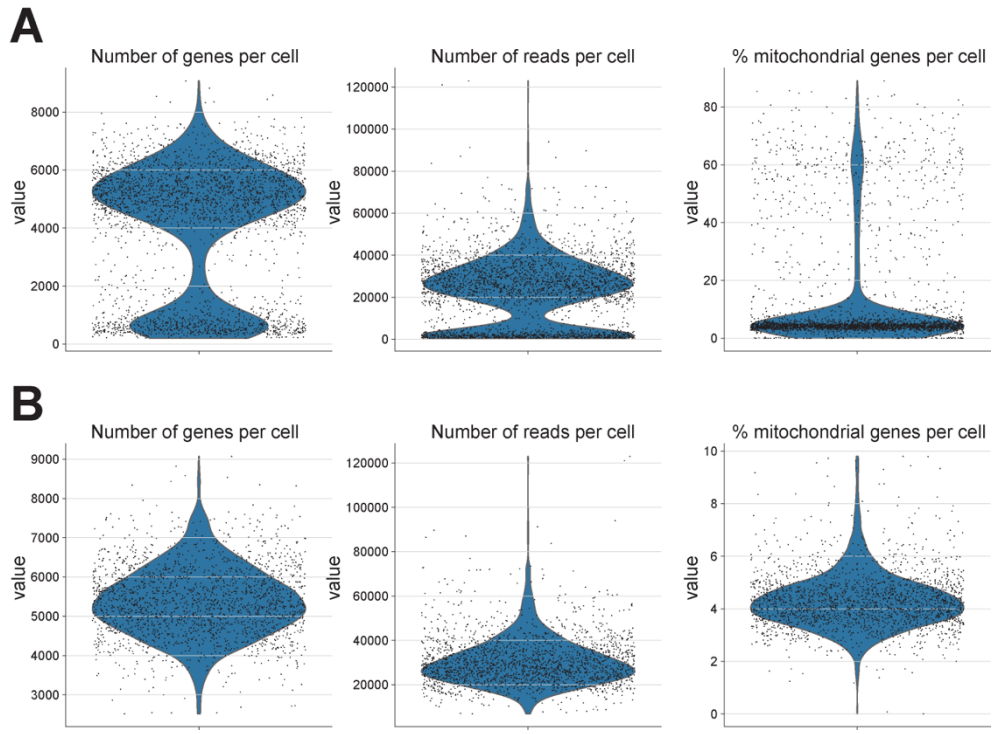


Fig. 13: Representative violin plots of the deep sequenced libraries, before and after applying the first filters.

a)First analysis of the datasets. b)After applying filters.

Dataset were then concatenated, and the resulting dataset normalized to a value of 50'000 and logarithmized to log1p. The highly variable genes were detected and the genes which were not considered highly variable were filtered out. The effects of total counts per cell were regressed out and the percentage of mitochondrial genes expressed. The data was then scaled to unit variance and values exceeding standard deviation 10 were clipped.

The dimensionality of the data was reduced by running a Principal Component Analysis (PCA, Fig. 14a), then the neighborhood analysis was performed and the graph was visualized in two dimensions using the Uniform Manifold Approximation and Projection technique (UMAP, Fig. 14b, McInnes, Healy, & Melville, 2018). In order to divide the cells into a discreet number of clusters and proceed with the dataset analysis, I performed a Leiden clustering with a resolution of 0.33 (Traag, Waltman, and van Eck 2019). The dataset I obtained was used to calculate correlation scores, density plots, visualize the expression of specific genes and create violin plots of the

gene expressions over the different conditions or clusters. All of these analyses were carried out with specific functions of *Scanpy*. Heatmaps were created with a *Prism 7* software from the primary data recovered from tables visualized in *Scanpy*.

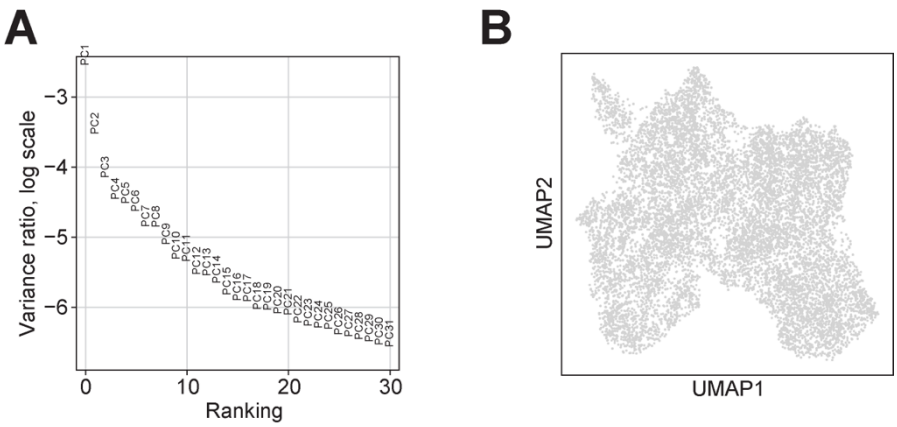


Fig. 14: Dimensionality reduction of the scRNAseq dataset

a)Contribution of single PCs to the total variance in the data. b)UMAP plot of the concatenated dataset.

<u>scRNA sequencing integration with published results</u>

In order to better understand the single cell dataset, I integrated it with two fully annotated datasets created by deep sequencing mouse embryos during gastrulation, from E6.5 to E8.5 (Pijuan-sala et al. 2019; Grosswendt et al. 2020). The Pijuan-Sala 2019 dataset was retrieved with R package *BiocManager* (https://bioconductor.org/ /packages/release/data/experiment/html/MouseGastrulationData.htm) and saved as a Python annotated dataset (*anndata*), to be able to open it with the Python package *Scanpy*. The Grosswendt 2020 dataset was obtained from one of the authors of the paper as a R object (extension '.Robj') and translated into an *anndata* object.

To compare the datasets, they were normalized the same way I had normalized my own dataset and before integration I filtered the genes to keep only those in common between the annotated embryo dataset and my *in vitro* dataset.

I performed the integration with the *Ingest* function of *Scanpy* considering the annotation for both the embryonic stage and the cell identities. The integration

projected the *in vitro* dataset onto the embryonic annotated datasets and automatically transferred the labels, that can be visualized on a UMAP plot or inspected with a table.

RESULTS

The first stages of mammalian embryogenesis aim to divide the precursors of extraembryonic structures from the cells that will eventually form the embryonic tissues. Forming the extraembryonic tissues first is important to create a complex signaling network between embryonic and extraembryonic cells, essential to differentiate the single layer of pluripotent epiblast cells into an organized tri-dimensional structure. The epiblast cells differentiate into the three germ layers: ectoderm, mesoderm, endoderm, that will form all of the organs of the embryo proper.

This complex signaling network is dependent upon signaling proteins produced by two extraembryonic structures, the visceral endoderm and the extraembryonic ectoderm, but also by paracrine or autocrine signaling from the epiblast cells themselves. Regionalized signaling environments are created within the epiblast and coordinate mesoderm patterning. Key signaling proteins involved in this process are BMP4, WNTs and FGFs (Fig. 15). BMP4 and WNT are inhibited in the anterior side of the embryo and present their highest concentration at the posterior side. At this stage, BMP is only produced by the ExE and it is believed to establish a proximal to distal gradient (Z. Zhang et al. 2019; Winnier et al. 1995; Rivera-Pérez and Magnuson 2005; Morgani and Hadjantonakis 2019). FGFs are produced by epiblast cells themselves and enriched in the primitive streak, however, how this signaling interacts with the other elements of the signaling network and its role in mesoderm patterning is still unclear (Morgani and Hadjantonakis 2019).

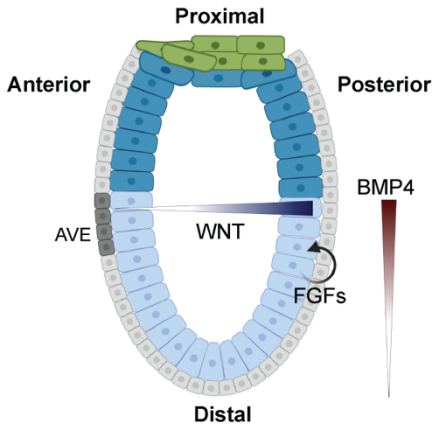


Fig. 15: Signaling network in the pre-gastrulating embryo

AVE: Anterior Visceral Endoderm; created with BioRender.com.

In vitro mesoderm protocol

The complex signaling network between the epiblast cells and extraembryonic tissues is difficult to tease apart *in vivo* in a mouse embryo, especially for the cross-regulation of the individual signals involved in the mesoderm patterning. In order to unravel how exogenous and cell-cell signaling organizes mesoderm differentiation, I developed an *in vitro* mesoderm differentiation model. This model would allow to more easily access and control the signals applied to the pluripotent cells, in terms of strength and timing, and design conditions to elucidate the role of cell-cell crosstalk in mesoderm patterning.

The mesoderm differentiation protocol employed mouse EpiSCs. These are primed pluripotent cells functionally equivalent to epiblast cells just prior to gastrulation, allowing a more direct differentiation than using embryonic stem cells. EpiSCs present heterogenous subpopulations of distinct pluripotential state when are maintained in FGF2/Activin A conditions, comparable to cells of early- and late-stage post implantation mouse embryos (Han et al. 2010; Weinberger et al. 2016). The expression of endogenous WNT signaling mediate the acquisition of primitive streak characteristics and its inhibition with the small molecule XAV939 maintains the EpiSCs in a homogenous undifferentiated state (Tsakiridis et al. 2014; Sumi et al. 2013). The EpiSCs cells were therefore cultured in presence of FGF2, Activin A and XAV939 (FAX medium) and represented the base for the mesoderm differentiation protocol. Cells cultured under these conditions are homogeneously $NANOG^{+}$ and do not express mesoderm markers (Fig. 16).

Building on existing protocols for the differentiation of mesoderm or specific mesoderm subtypes, I cultured the EpiSCs with 1 µM Chiron, a WNT signaling agonist, and 12 ng/ml FGF2 in the chemically defined N2B27 medium for three days, with daily media changes (Chal et al. 2015; Loh et al. 2016; Sudheer et al. 2016; Vallier et al. 2009). The simple differentiation protocol allows to titrate or pulse each component and explore the impact on phenotype or signal transduction.

Cells cultured under these conditions downregulated the pluripotency marker NANOG and promoted the expression of the pan-mesoderm marker T/BRACHYURY (Fig. 16).

I also analysed the expression of HAND1 and TBX6, mesoderm markers expressed respectively in a more proximal and distal region of the primitive streak (Peng et al. 2016): I found groups of cells positive for TBX6, while no cells expressed HAND1. When the cells were also treated with 10 ng/ml BMP4 during the three days of differentiation, they maintained the expression of T/BRA and TBX6, while also expressing HAND1. The addition of BMP4 in culture seemed to give a more proximal identity to the mesodermal cells. Nevertheless, the expression of TBX6 remained unchanged, and this might indicate that a higher dose could inhibit TBX6 or even the other selected markers.

This assay showed that the protocol effectively differentiates cells towards a mesoderm identity and that different culture conditions could control the differentiation of distinct mesoderm subtypes.

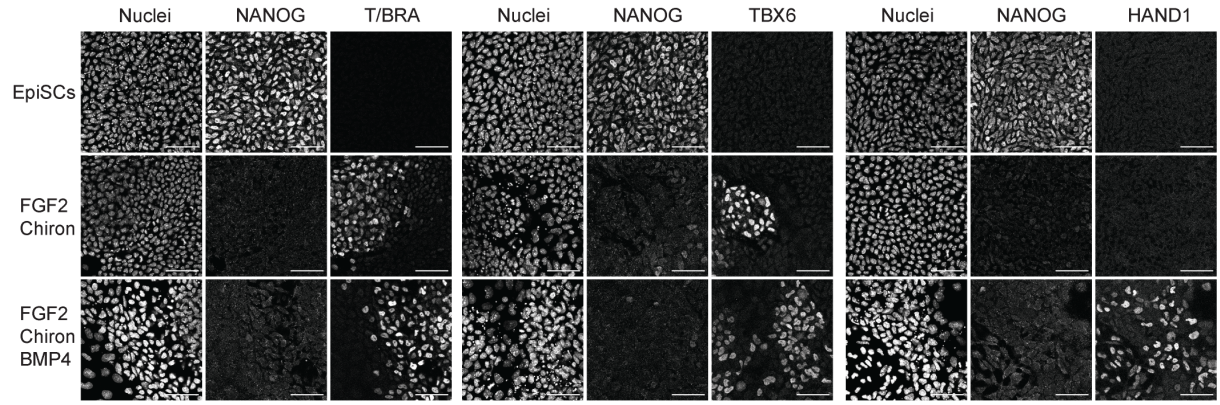


Fig. 16: FGF2 and Chiron promote mesoderm differentiation. BMP4 promotes the expression of the lateral mesoderm marker Hand1

Immunofluorescence of EpiSCs differentiated as described, stained with Hoechst, NANOG, T/BRA, TBX6, HAND1. Scale bar = 50 μm.

BMP4 titration shows a morphogenic effect in *in vitro* mesoderm differentiation

Mesoderm patterning is dependent upon signaling proteins produced by extraembryonic tissues surrounding the epiblast cells. At the beginning of gastrulation, BMP4 is inhibited in the anterior side of the embryo and present its highest concentration at the posterior side. At this stage, BMP4 is only produced by the ExE

and it is believed to establish a proximal to distal gradient (Z. Zhang et al. 2019; Winnier et al. 1995; Morgani and Hadjantonakis 2019), thus creating an externally generated signaling asymmetry in the posterior epiblast. To assess if the concentration of the BMP4 ligand would affect the differentiation of specific mesoderm subtypes, I titrated BMP4 and analysed the mRNA expression of marker genes that showed regionalized expression in the gastrulating embryo (Peng et al. 2016). FGF2 and Chiron were kept constant at the concentrations of 12 ng/ml and 1 µM.

The expression of Hand1 and Gata6, which define the more proximal region of the primitive streak, was highest at the higher BMP4 concentrations, 16 and 32 ng/ml. Markers of a more distal part of the primitive streak, Tbx6 and Msgn1, reached the highest expression at intermediate BMP4 concentrations, 2 to 8 ng/ml, while the most distal markers Foxa2 and Shh peaked at the lowest BMP4 concentrations, 1 to 4 ng/ml (Fig. 17a). The expression curve of the pan-mesoderm marker T/Bra resembled Tbx6 and reached its highest expression at 4 ng/ml.

To confirm these results on a protein level, I tested the expression of Hand1, Gata6 and Tbx6 on a population-based essay, an In Cell Western. The results largely supported the previous analysis, even though the signal-to-noise ratio was in this case higher. The expression of the proximal HAND1 and GATA6 increased with higher BMP4 concentrations, respectively until 64 and 16 ng/ml. The more distal TBX6 peaked its expression at 16 ng/ml. All of these markers were inhibited at the highest concentration of BMP4 (Fig. 17b,c).

These results together indicated that BMP4 can determine the expression of proximal mesoderm markers when it is used at high concentrations. They also suggested the existence of an ideal BMP4 concentration for the emergence of specific mesoderm markers, with intermediate levels being ideal for Tbx6 and Msgn1 and low levels for Foxa2 and Shh. This observation supported the hypothesis that the external gradient of BMP4 along the primitive streak drives the emergence of distinct mesoderm subtypes.

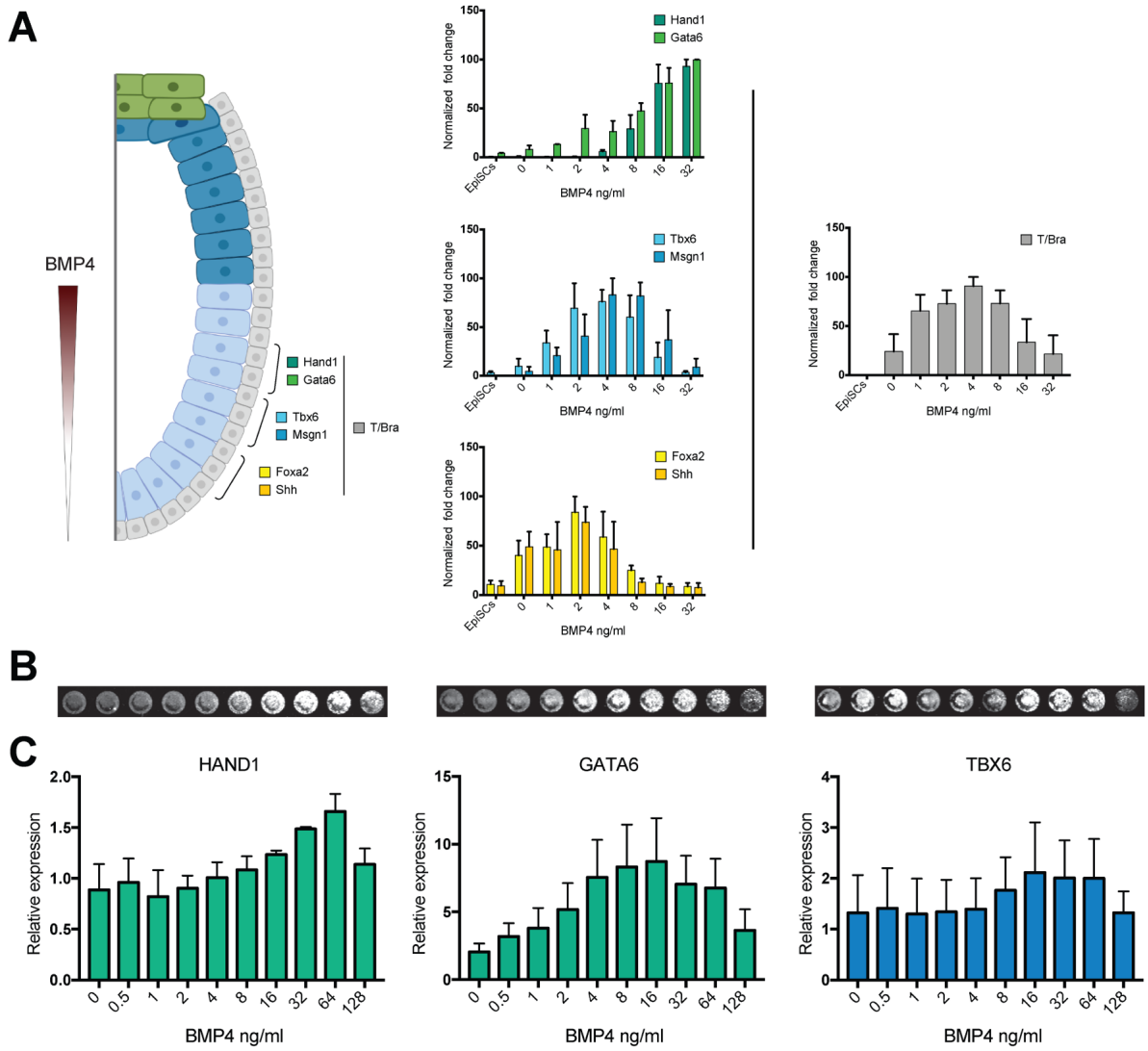


Fig 17: Differentiation of mesoderm subtype is affected by BMP4 concentration.

A)qPCR analysis of selected mesoderm markers in cells cultured with constant FGF2 and BMP4 titration; error bars indicate SEM (n. experiments = 3). B)Representative In Cell Western detection of HAND1, GATA6 and TBX6. C)In Cell Western analysis of HAND1, GATA6 and TBX6; data was normalized on the negative control cultured with no FGF or BMP; error bars indicate SEM (n. experiments = 3).

FGF affects the expression of mesoderm markers during *in vitro* differentiation

High concentrations of FGF2, but not FGF4, inhibit mesoderm differentiation

FGFs have been observed to be necessary for mesoderm differentiation *in vivo* and *in vitro* (Yu et al. 2011; Oki, Kitajima, and Meno 2010), but it is unknown if and how different concentrations of FGF could impact the differentiation of distinct mesoderm lineages, whether they determine a permissive condition for one or more specific cell lineages beyond specific thresholds or they determine a slow shift between them.

I assessed the effect of FGF titration on the pan-mesoderm marker T/Bra by employing a fluorescent Sox1-T/Bra reporter, called the SBR line (Deluz et al. 2016). This line expresses cytosol-localized eGFP and mCherry when the production of Sox1 and T/Brachyury are triggered. The mCherry expression of the differentiated reporter cells was analyzed with a flow cytometer to obtain a single cell level resolution of the expression of T/Bra. I tested two FGFs: FGF2 and FGF4. FGF2, although not present in the gastrulating embryo, is commonly used in mesoderm protocols and was used here in the previous experiments; in contrast, FGF4 is present in the embryo as a key signaling molecule expressed in the posterior epiblast during gastrulation (Boulet and Capecchi 2012). I conducted the titration experiment with constant 1μM Chiron, as previously done, and 8ng/ml BMP. This intermediate BMP4 concentration was chosen as it supported the differentiation of more than one mesoderm marker subtype (Fig. 17).

Data showed that at low concentrations of both FGF2 and FGF4, until 12 ng/ml, there was a positive correlation between the FGFs and the proportion of T/BRA^{+} cells. With higher FGF2 this proportion slowly declined (Fig. 18a,b), while with FGF4 it increased until a plateau at 48-96 ng/ml FGF4 (Fig. 18c,d). The decreased expression of T/BRA was observed at high doses of FGF2, but not FGF4, and recapitulated published results, with FGF2 displaying an autoinhibitory effect at high concentrations (Gharibi et al. 2020). A concentration of FGF2 lower than 12-24 ng/ml was used in all further experiments, since FGF2 mimicked the FGF4 effect at low concentrations. Moreover,

the common use of FGF2 in published protocols and studies could help to contextualize the results with the literature.

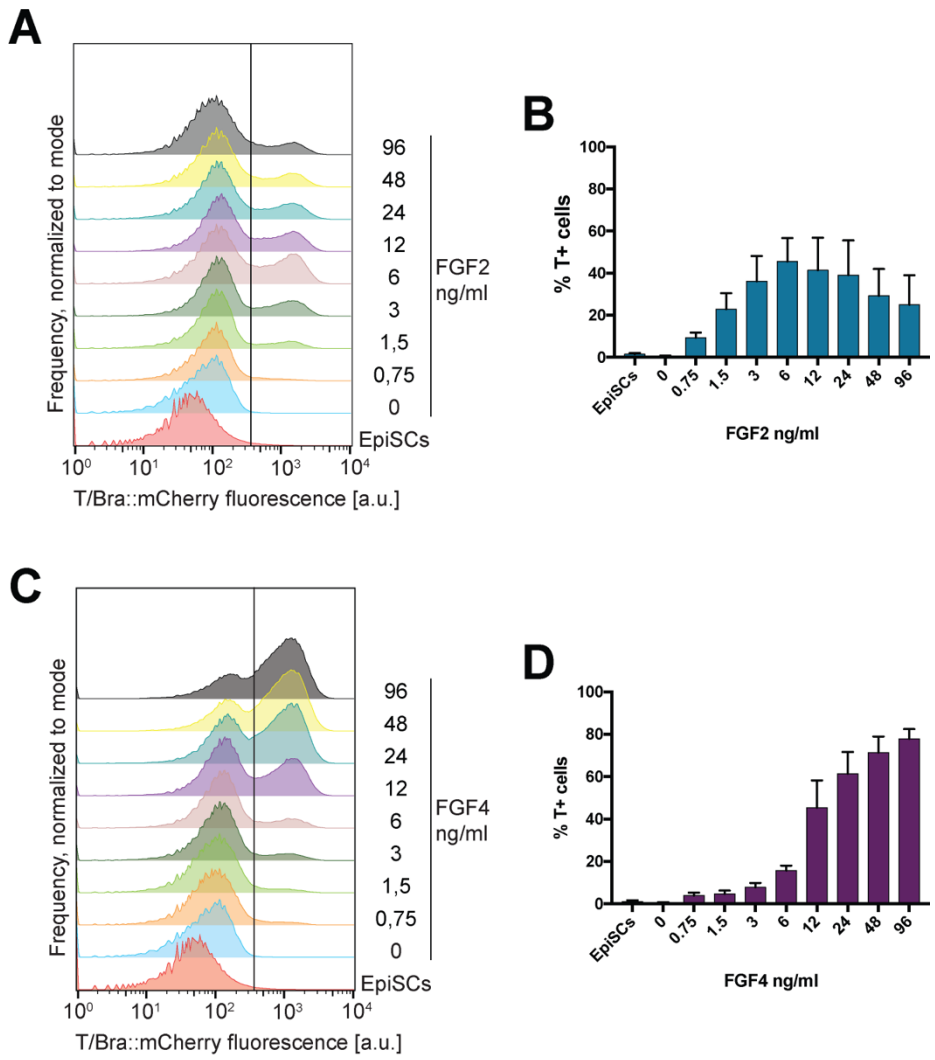


Fig. 18: FGF2, but not FGF4, shows a decreased T activation at high concentrations.
A,C)Representative flow cytometry of T/Bra:: mCherry reporter line (SBR line) differentiated with a range of FGF2 or FGF4 concentrations; threshold between positive and negative cells is displayed. B,D)Quantification of T/Bra:: mCherry fluorescence positive cells from flow cytometry; error bars indicate SEM (n. experiments = 3).

Both FGF2 and FGF4 appeared to positively influence the differentiation of mesoderm in a concentration-dependent manner until 12-24 ng/ml. From these results, I decided to use FGF2 to explore the effect of FGF signaling on the differentiation of distinct mesoderm subtypes.

FGF2 affects the expression of distinct mesoderm markers in a concentration-dependent manner

To determine the effect of an increasing concentration of FGF2 on proximal or distal mesoderm markers, I performed an immunofluorescence on wild-type cells cultured with a titration of FGF2 and a constant 8ng/ml BMP and 1µM Chiron. The expression of the pluripotency marker NANOG was strongly decreased in every condition (Fig. 19a,b), as also seen in the immunofluorescence in Fig. 16. Quantification of the T/BRA protein demonstrated a positive correlation with the increase of FGF2 concentration, with a stronger increase in the number of positive cells between 0.75 and 1.5 ng/ml (Fig. 19a,c). While the increase in T/BRA$^+$ cells with the exogenous FGF2 was consistent with the previous experiment (Fig. 18b), this experiment did not show a decline in the proportion of T/BRA$^+$ cells, but a slower increase between 3 and 24 ng/ml FGF2. TBX6 expression also revealed a positive correlation with higher FGF2 concentration, with a strong increase between 0.75 and 1.5 ng/ml (Fig. 19d,e), while HAND1 was negatively correlated with the exogenous FGF2 (Fig. 19f,g). CDX2 was found to be expressed in every differentiated condition, but did not show a consistent trend with changing FGF2 concentrations (Fig. 19f,h).

These results suggested that FGF inhibited the expression of more proximal mesoderm markers and promoted markers of more distal fates in a concentration-dependent manner. This was the inverse effect observed for BMP in this *in vitro* differentiation model. Furthermore, it represented a novel signaling interaction involved in mesoderm patterning.

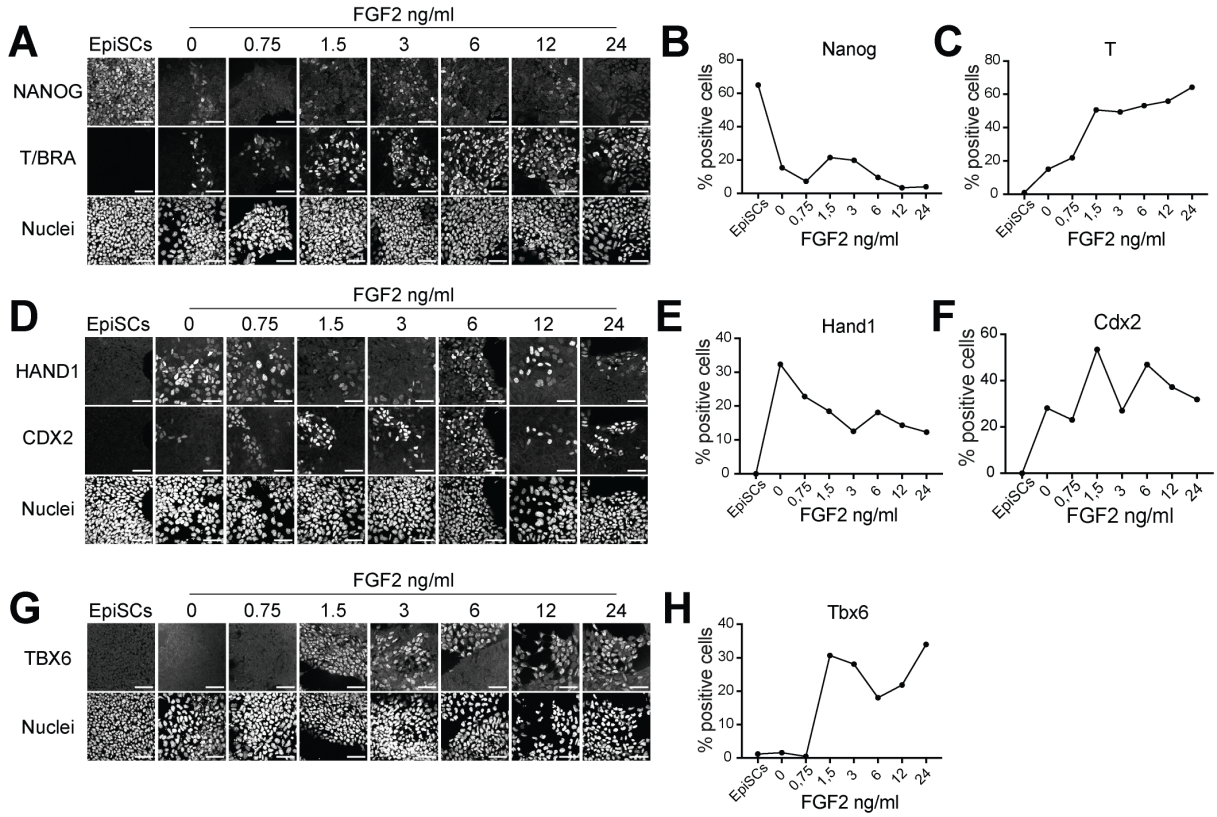


Fig. 19: Immunofluorescent imaging and quantification of selected mesoderm markers

A,D,G)Immunofluorescence of EpiSCs differentiated as described, stained with Hoechst, NANOG, T/BRA, HAND1, CDX2 and TBX6. Scale bar = 50 μm. B,C,E,F,H)Quantification of the immunofluorescence; graph show the pertantage of cells positive for distinct mesoderm markers according to culture condition (n. experiments = 1).

A FGFR inhibitor could inhibit autocrine and paracrine FGF signaling

Since the epiblast cells close to the primitive streak upregulated the expression of FGF ligands, it was possible that the cells differentiating *in vitro* also retained a basal level of expression of FGFs. I then verified whether it was possible to inhibit the FGF-mediated cell-cell crosstalk, as this would allow to evaluate the influence of autocrine and paracrine FGF signaling on mesoderm patterning.

I tested the toxicity and effectiveness of a selective FGFR inhibitor (AZD4547) by adding a titration of the molecule in the mesoderm protocol. Results indicate that micromolar concentrations previously used in the embryo, down to 62.5nM of the

molecule, lead to complete cell death (Fig. 20b, Morgani, Saiz, et al., 2018), whereas a concentration of 31.25 nM allowed cells to survive and form small colonies (Fig. 20a,b). This concentration of the FGFR inhibitor provoked a decrease in ppErk after one day of differentiation (Fig. 20c), but not a complete inhibition, thus suggesting that low levels of basal FGF signaling are essential for cell survival during mesoderm differentiation. A concentration of 30 nM was used in all subsequent experiments, since it consented cell survival and decreased ppErk.

The addition of a FGFR inhibitor in the cell culture media would allow to restrict the possible FGF-mediated cell-cell crosstalk and investigate its effect on mesoderm differentiation.

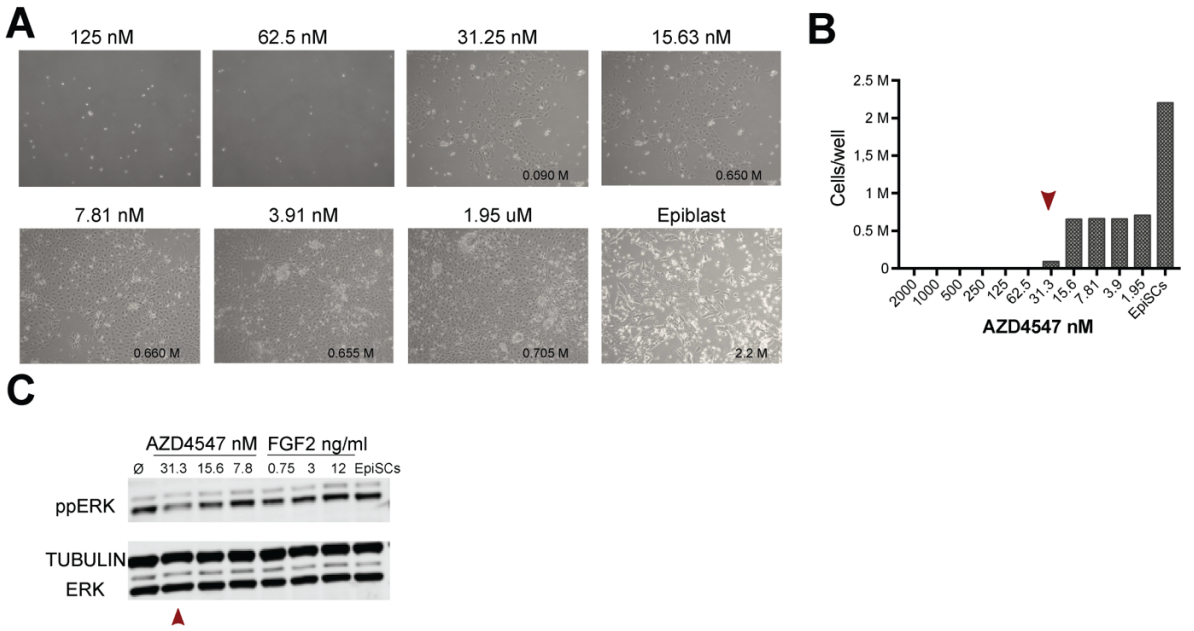


Fig. 20: Working concentration of FGFRi in mesoderm differentiation protocol

A,B)Brightfield images and cell count of cells treated for 3 days of treatment with 1 μM Chiron, 8 ng/ml BMP4 and titration of AZD4547 (n. experiments = 1). C)Western blot of ppERK and ERK of cells treated with 1 μM Chiron, 8 ng/ml BMP4 and titration of AZD4547 or FGF2 for 1 day.

Single cell transcriptomic analysis

Endogenous and exogenous FGF signaling affects the transcriptional identity of differentiated cells

The results from the qPCR and immunostaining indicated that the *in vitro* differentiation protocol here employed often determined the co-expression of markers of distinct mesoderm subtypes, especially at intermediate levels of FGF2 and BMP4

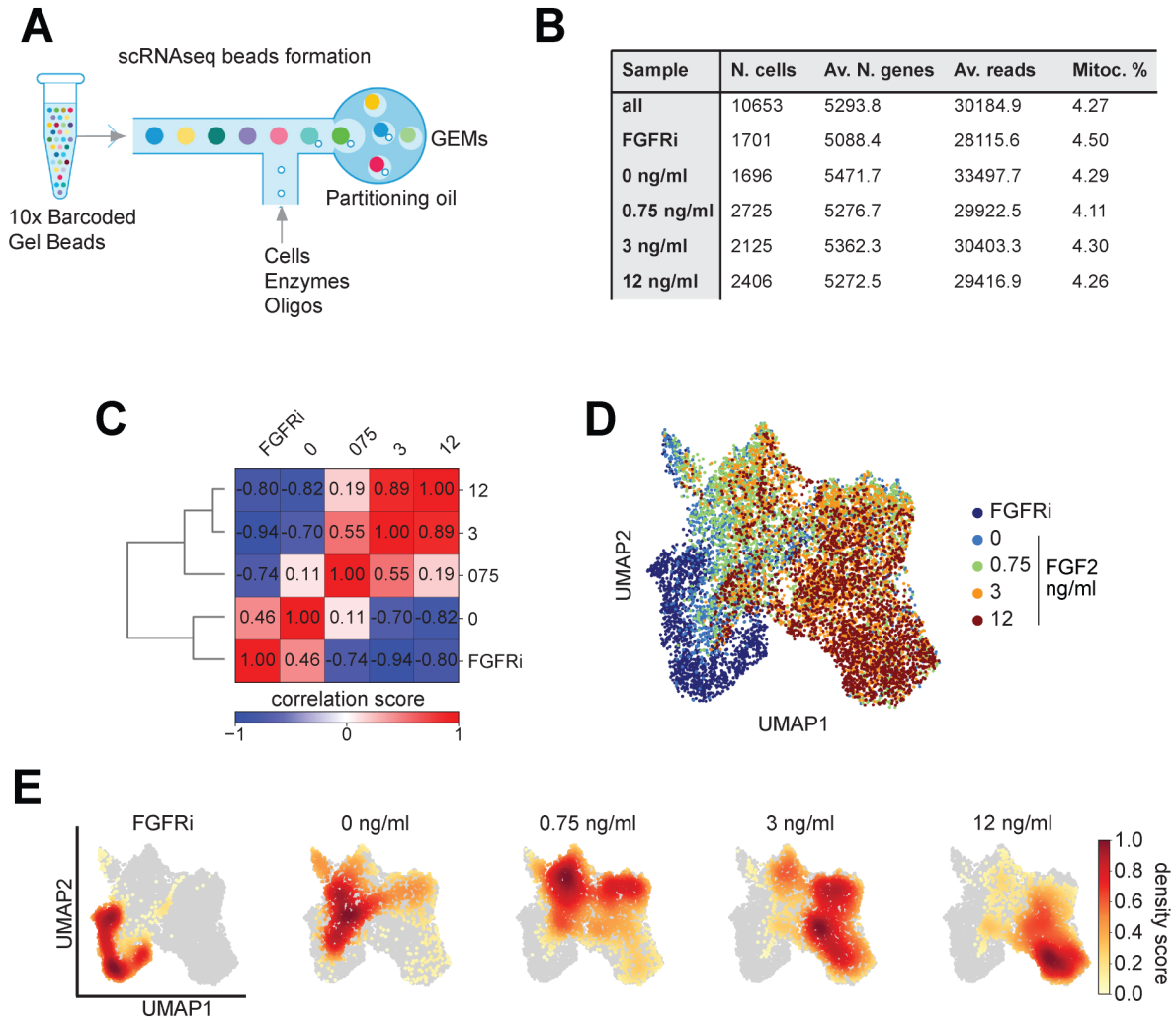
(Fig. 16, 17, 19). This observation suggested the existence of a heterogenous mixture of mesoderm subtypes in these conditions. To identify the cell types represented in this heterogenous mixture, their proportion and how would they change depending on FGF signaling, I performed a single cell RNA sequencing (scRNAseq) experiment on cells differentiated with a critical range of the FGF concentrations established above: FGFR inhibition, no added FGFs or increasing exogenous FGF2: 0.75, 3.0 or 12.0 ng/ml (Fig. 21a). The differentiation conditions tested with scRNAseq had constant BMP and WNT signaling, specifically 8ng/ml BMP4 and 1μM Chiron, as the previous experiments.

At the end of differentiation, the cells were detached and resuspended to obtain single cells in solution. I aimed to analyze approximately 2000 cells per sample, so I mixed an appropriate amount of the cells in solution with a mix of enzymes, oligos and buffer as per manufacture's instruction. This solution was partitioned in nanoliter-scale Gel Beads-in-emulsion (GEMs) using the Chromium controller machine (Fig. 21a). Each GEM serve as individual reaction chamber for a single cell to generate cDNAs and tag them with a unique barcode that identifies the original mRNA, the cell and the sample. GEMs are then dissolved, the cDNA cleaned and amplified to obtain the libraries. Each library was sequenced with a high sequencing depth of ~30000 reads per cell and an average of ~5300 detected genes per cell (Fig. 21b). The analysis was conducted with a dedicated Python package, *Scanpy* (Wolf, Angerer, and Theis 2018).

The correlation matrix and the dendrogram calculated from *in silico* generated bulk transcriptomes revealed the strong correlations between samples treated with similar FGF signaling strength (Fig. 21c). The two most similar samples are the ones cultured with the highest FGF2 concentrations, 3 and 12 ng/ml, followed by the 0.75 and 3 ng/ml conditions and the 0 and FGFRi. The biggest step was found between cells with no exogenous FGF2 and 0.5 ng/ml. This difference, together with the similarity between all samples cultured with exogenous FGF2, suggests the existence of two clusters of samples: the 0.75, 3 and 12, with exogenous FGF2, and the FGFRi and 0, without exogenous FGF2 or with an inhibition of FGFR.

These results indicated that the different FGF signaling strength of the cell culture media lead to distinct transcriptional cell states and that the addition of any

concentration of exogenous FGF2 determined a substantial difference in the transcriptional cell state.



Sample	N. cells	Av. N. genes	Av. reads	Mitoc. %
all	10653	5293.8	30184.9	4.27
FGFRi	1701	5088.4	28115.6	4.50
0 ng/ml	1696	5471.7	33497.7	4.29
0.75 ng/ml	2725	5276.7	29922.5	4.11
3 ng/ml	2125	5362.3	30403.3	4.30
12 ng/ml	2406	5272.5	29416.9	4.26

Fig. 21: Quality control of scRNAseq analysis and single cell visualization with UMAP dimensionality reduction technique.

A)GEM beads formation: GEMs generated the uniquely tagged cDNA that later formed the libraries. B)Summary statistics for the analysed samples, detailing the number of cells, average number of genes per cells, average number of reads per cell and average percentage of mitochondrial genes. C)Correlation matrix of the sequenced samples, considered as pseudo-bulk samples; on the left, the dendrogram of the hierarchical clustering of the samples. D) UMAP plot of the 5 samples: AZD4547, no FGFs and 3 FGF2 concentrations, 0.75, 3 and 12 ng/ml. E)Density plots of the single samples.

The entire dataset was then visualized at a single cell resolution with a dimensionality reduction technique, the Uniform Manifold Approximation and Projection plot (UMAP McInnes et al., 2018) to display the transcriptional states of the singles cells in each sample (Fig. 21d). The combined UMAP plot showed that all conditions, except the FGFRi treated cells, occupied overlapping territories in gene expression space (Fig. 21d,e). The cells cultured with FGFRi, albeit in contact to the no FGF condition, were cleanly separated, indicating that basal endogenous FGF signalling affected gene transcription and cell identities in the cells cultured with no FGF. Increasing concentrations of FGF2 shifted the most densely populated areas in the UMAP plot towards the opposite side, populated by the cells exposed to 12 ng/ml FGF2.

This indicated that cells in all conditions that allowed endogenous FGFs were highly heterogeneous, whereas exogenous FGF2 strength affected cell response and identity.

FGF2 promotes distal mesoderm markers and supports EMT

To confirm the concentration-dependent effect of exogenous FGF2 on the mesoderm markers previously analyzed (Fig. 19), I visualized their expression over the UMAP plot and in violin plots, grouped by sample condition. Cells expressing distinct mesoderm markers formed cohesive groups in the UMAP plot rather than a salt and pepper distribution, suggesting that these cells showed similar transcriptional identity (Fig. 22a). The quantification of T/Bra, Tbx6 and Cdx2 confirmed the positive correlation with increasing exogenous FGF2, showing a particularly strong effect in Cdx2 (Fig. 22b). The number of cells expressing Hand1 also confirmed previous findings (Fig. 19), as they decreased with higher exogenous FGF2 treatments.

I also analyzed the expression of the transmembrane proteins Cdh1 and Cdh2 (also known as E-cadherin and N-cadherin), as the replacement of Cdh1 with Cdh2 indicates epithelial to mesenchymal transition during gastrulation. Both cadherins were found in each condition, implying that the cells obtained were a heterogeneous mixture of gastrulating and non-gastrulating cells (Fig. 22). However, higher exogenous FGF2 was correlated with a decrease of the number of $Cdh1^{+}$ cells and an increase of $Cdh2^{+}$ cells,. The effect of increasing FGF signaling on the two cadherins

is consistent with the proposed role of FGF in promoting mesoderm migration and the presence of mesoderm migrations defects in FGF-signaling deficient embryos (Oki, Kitajima, and Meno 2010).

The results confirmed the previous findings that suggested that FGF could inhibit the expression of more proximal mesoderm markers and promote markers of more distal fates in a concentration-dependent manner. They also indicated that FGF could support the epithelial-to-mesenchymal transition (EMT) in *in vitro* mesoderm differentiation. This result was compatible with the *in vivo* findings that FGF signaling mutations determine EMT defects (Ciruna and Rossant 2001; Yamaguchi et al. 1994).

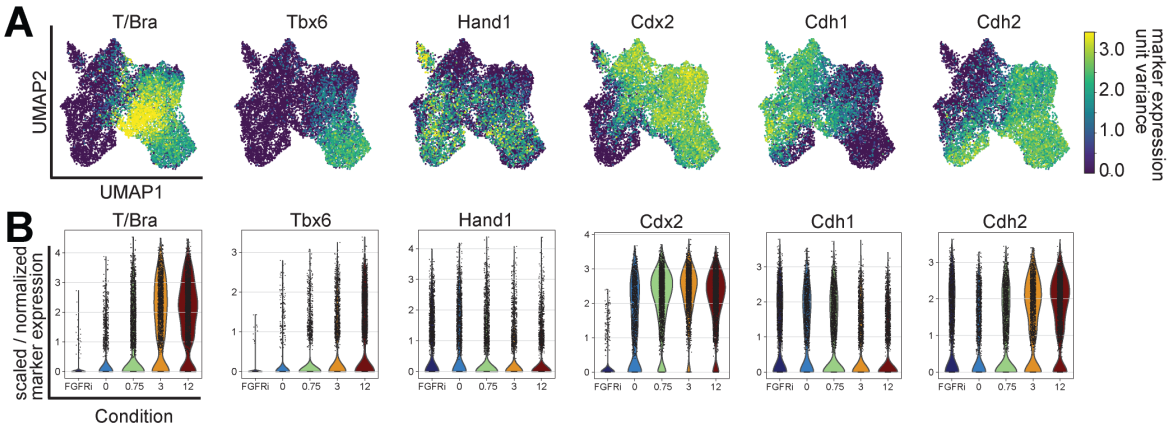


Fig 22: Marker expression and clustering in the UMAP plot expression space.
A)UMAP plots coloured by the expression level of T/Bra, Tbx6, Hand1, Cdx2, Cdh1 and Cdh2 in gene expression space; brighter coloured regions exhibit higher expression levels. B)Violin plots of the expression levels of T/Bra, Tbx6, Hand1, Cdx2, Cdh1 and Cdh2, grouped by sample; width of the violin plots is scaled per number of observations.

FGF2 shifts the allocation of cells in different clusters

In order to characterize the cell types differentiated *in vitro*, I clustered the dataset with a Leiden algorithm and chose a resolution consistent with the expression of the marker genes and with a number of clusters equal or greater than the number of conditions (Fig. 23a). Each condition preferentially populated different clusters: the FGFRi sample equally divided and almost exclusively found in cluster 5 and 6; cells exposed to no FGF2 or 0.75 ng/ml FGF2 mostly contributed to cluster 3 and 2 or cluster 2 and 0; cells exposed to the high FGF2, 3 and 12 ng/ml, were mostly found in clusters 0, 4 and 1, with the highest fraction of cells of the 12 ng/ml sample belonging to cluster 1

(Fig. 23b). Almost all samples had a small number of cells mapped outside of the main clusters of the condition, suggesting that FGF signaling shape the proportions of cells identifying as one out of a range of cell states during *in vitro* mesoderm differentiation.

Single clusters were shown to have an upregulation of specific gene markers previously quantified (Fig. 23c). T/Bra was upregulated in clusters connected to high exogenous FGF2: cluster 0, 1 and especially 4. Tbx6 expression was particularly high in cluster 1 and Cdx2 expression increased constantly from clusters mainly connected to cells not exposed to FGF2 (cluster 3) to clusters associated with high FGF2 (clusters 1, 0 and 4). Expression of Cdh1 and Cdh2 differed especially in clusters 0, 1 and 6.

The different clusters divided the dataset into groups of cells with similar transcriptional profiles and were preferentially populated by different samples. They also showed distinct expression of the selected proteins, suggesting that they could be used to identify the different cell types.

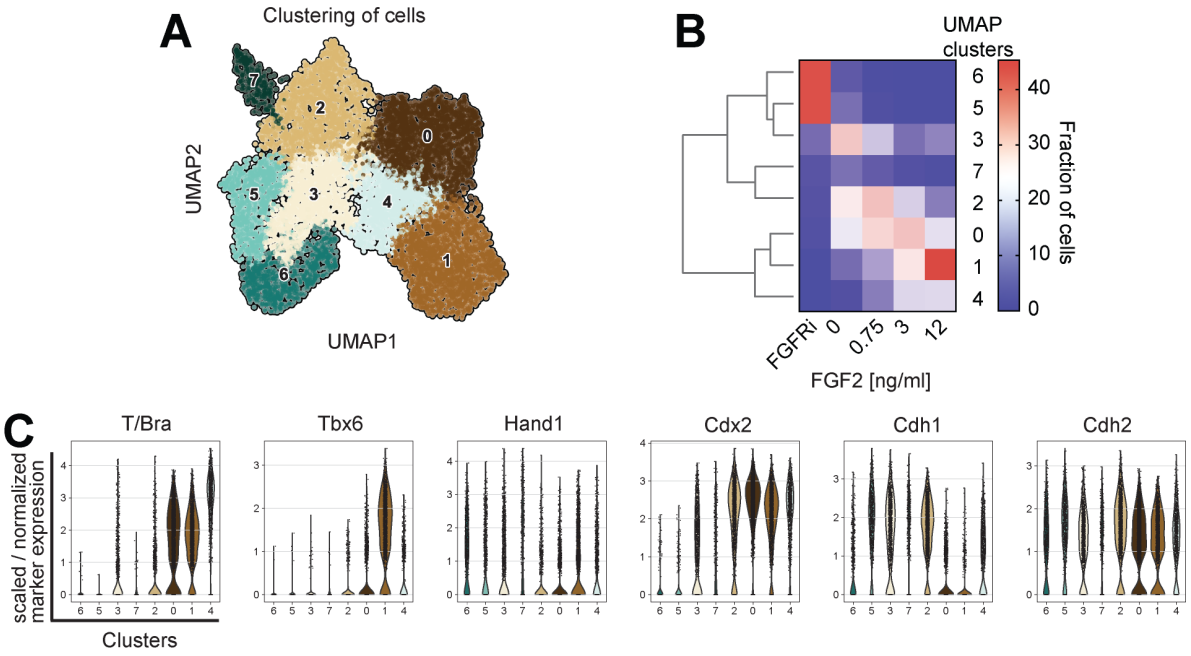


Fig. 23: Dataset clustering with Leiden algorithm

A)UMAP plot showing computed cell clustering. B) proportion of cells of each sample in the clusters; the clusters have been ordered based on computed hierarchical clustering; dendrogram on the left side of the heatmap indicates transcriptional similarity between clusters. *C)Violin plots of the expression levels of T/Bra, Tbx6, Hand1, Cdx2, Cdh1 and Cdh2, grouped by cluster; width of the violin plots is scaled per number of observations.*

Integration with fully annotated embryo reference datasets

The heterogeneity and the expression of specific mesoderm marker in scRNAseq dataset suggested the existence of a discrete number of cell identities. I integrated the *in vitro* dataset with two fully annotated reference datasets from gastrulating mouse embryos to determine which cell lineages were differentiated *in vitro*, as well as their developmental stage, and assessed whether FGF signaling affected these parameters (Pijuan-sala et al. 2019; Grosswendt et al. 2020). These two datasets have been created by different scientific groups, with differences in the biological samples and the bioinformatic analysis. However, they both sampled embryos just around gastrulation, between E6.5 and E8.5, and used the same 10x genomics technology used in this work to perform the scRNAseq. The commonalities and the differences of these datasets make them valuable refences that could confirm any shared results found with their integration with the *in vitro* dataset created in this work. An asymmetric dataset integration was performed with the *Scanpy* function *ingest* to compare the datasets and transfer the annotations of cell identity and developmental stage (from *Scanpy*, Wolf et al., 2018).

The Pijuan-Sala 2019 annotated embryo dataset employed the C57BL/6 mouse strain. They harvested embryos every 6 h between E6.5 and E8.5 and carefully staged them to exclude outliers before processing them, since development can proceed at different speeds even between the same litter. They analyzed 116312 cells in total, employed the CellRanger analysis pipeline to demultiplex their dataset and R and python packages for quality control, normalization and clustering (Pijuan-sala et al. 2019).

When projected onto the annotated Pijuan-Sala 2019 dataset, the *in vitro* differentiated cells occupied a vast area, suggesting that the cellular heterogeneity previously observed represented different cell identities within the embryonic development (Fig. 24a). Cells belonging to the FGFRi condition populated with high density two discrete areas in the UMAP space. The endogenous FGF signaling and the increasing exogenous FGF2 dose during differentiation determined a shift in the areas of the UMAP plot occupied by the *in vitro* cells, suggesting a progressive shift towards a specific FGF-induced cell identity (Fig. 24b,c).

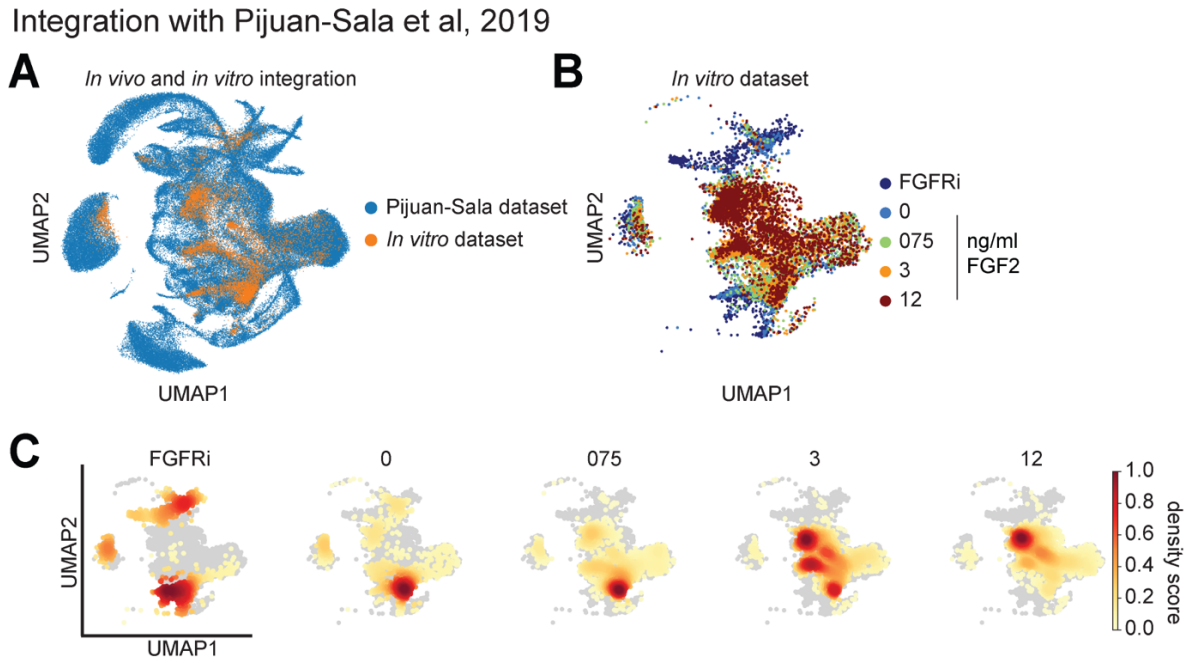


Fig. 24: Dataset integration with single cell transcriptome from the embryo

Dataset integration with Pijuan-Sala et al, 2019. A)Superimposition between the recalculated UMAP plots of the in vitro *dataset and the reference dataset. B)UMAP plot of the* in vitro *dataset, color-coded for the sample identity. C)Density plots of the single samples.*

The Grosswendt 2020 annotated embryo dataset employed the B6D2F1 mouse strain. To create this dataset, they harvested embryos every 12 h between E6.5 and E8.5 and processed them without any morphology-based pre-selection. They analyzed 88779 cells in total, employed the CellRanger analysis pipeline to demultiplex their dataset and the R package Seurat for quality control, normalization and clustering (Grosswendt et al. 2020).

The dataset integration was performed with ingest as before and the annotations of cell identity and developmental stage were transferred. As seen with the Pijuan-Sala 2019 dataset integration, the *in vitro* differentiated cells occupied a vast area, with cells from the FGFRi sample divided in two discrete portions of the UMAP space (Fig. 25a, c). Other samples occupied broad areas of the UMAP plot, experiencing a slow shift in the most densely populated areas of the graph (Fig. 25b,c).

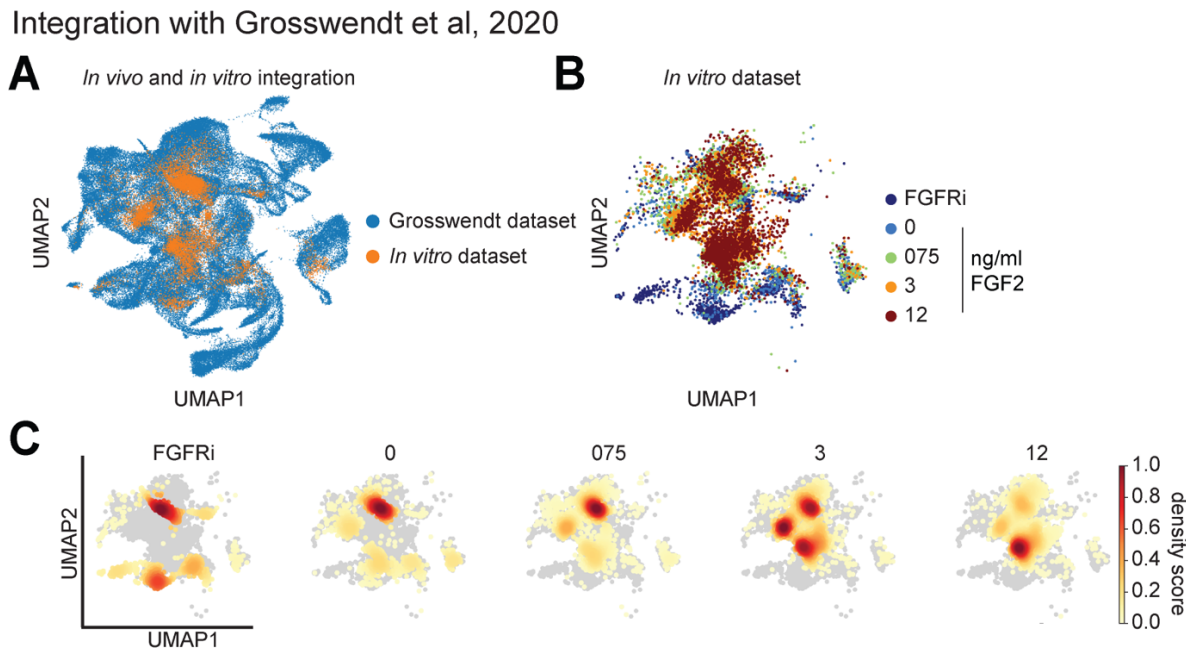


Fig. 25: Dataset integration with second independent single cell transcriptome from the embryo

Dataset integration with Grosswendt et al, 2020. A)Superimposition between the recalculated UMAP plots of the in vitro *dataset and the reference dataset. B)UMAP plot of the* in vitro *dataset, color-coded for the sample identity. C)Density plots of the single samples.*

<u>High exogenous FGF2 promotes a subgroup of younger cell types</u>

I first focused on the developmental stage labels transferred from the embryo reference datasets. The two datasets sequenced whole mouse embryos from E6.5 to E8.5, with a resolution of 6 h or 12 h. When visualized in UMAP plots, cells from younger embryos tended to populate one side of the graph, while cells taken from progressively older embryo slowly shifted toward the opposite side, consistently with the published analysis of these datasets (Fig. 26a,d, Grosswendt et al., 2020; Pijuan-sala et al., 2019). Notably, the labels of the *in vitro* differentiated cells appeared to originate from a wide range of developmental stages, covering almost all timepoints (Fig. 26b,e). However, when considering the specific labels transfered from each reference dataset, it became apparent that not all timepoints were equally represented by the *in vitro* differentiated cells.

The integration with the study with the higher time resolution (Pijuan-sala et al. 2019) revealed that, for all culture conditions, the most represented stage label corresponded

to the E8.0 embryo, which was assigned to a minimum of 33.2% of the cells of each sample (Fig. 26c). Except for the FGFRi condition, the large majority of cells from all conditions corresponded to E8.0 and the temporally close timepoints E7.75 and E8.25. The EpiSCs cells used as the starting point for mesoderm differentiation should correspond to the pre-gastrulation epiblast, around E5.0: considering the three days of the *in vitro* differentiation, this analysis indicated that the pace of differentiation for different FGF signaling is similar to the embryonic development. However, a large proportion of the FGFRi treated cells, 32.5%, corresponded to the older E8.5 label. On the contrary, increasing FGF signaling strength was correlated to a larger proportion of cells labelled as E7.0 and E7.25: 21.2% for the highest FGF2 concentration. This shift towards earlier developmental stages suggests that FGF signaling directly or indirectly regulated the differentiation pace and that higher FGF signaling could delay the differentiation of a portion of cells. The label transfer was as well analyzed by dividing the cells in the clusters defined in figure 23c and d: clusters corresponding to the FGFRi condition, cluster 5 and especially 6, had a higher proportion of cells identifying as E8.5, while clusters 3 and 4 had an overrepresentation of cells classified as less advanced (Fig. 26c').

The analysis of the integration with the second embryo dataset supported the hypothesis that the cells *in vitro* largely had the same differentiation pace of the cells in the embryo, since a large majority of the cells corresponded to E7.5 and E8.0, irrespective of the cell culture condition (Fig. 26f, Grosswendt et al., 2020). This dataset showed as well a shift of a proportion of cells from more advanced stages to earlier ones in relation to a higher FGF signaling strength.

These findings of the integration with the two datasets indicated that the pace of differentiation of the cells obtained through the *in vitro* mesoderm differentiation broadly resembled the pace of the cells in the embryo, while FGF signaling promoted a slower differentiation of a subset of cells.

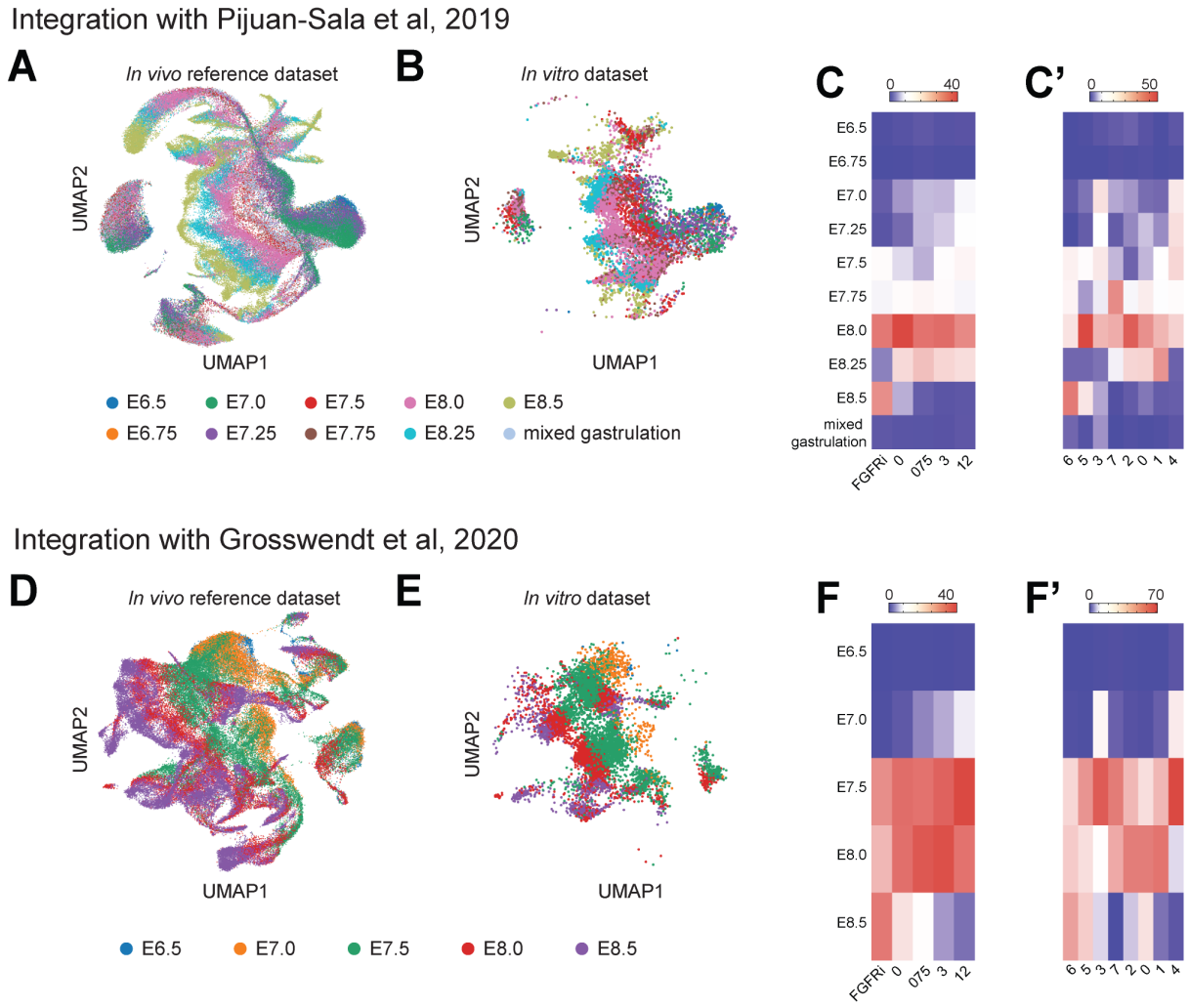


Fig. 26: Label transfer of developmental stages from two fully annotated embryo datasets. *A-C')Dataset integration with Pijuan-Sala et al, 2019. D-F')Dataset integration with Grosswendt et al, 2020. A,D)Recalculated UMAP plots of the reference dataset, color-coded for embryonic development stage. B,E)UMAP plot of the* in vitro *differentiated dataset, color-coded for embryonic development stage. C,F)Heatmap of the percentage of cells with a specific developmental stage label for different FGF signaling strength. C',F') Heatmap of the percentage of cells with a specific developmental stage label for different clusters.*

FGF signaling influences the cell types obtained *in vitro*

I then examined the transferred annotations of cell identity to determine which cell types were obtained *in vitro* and the effect that the different FGF signaling strengths had on the proportion and type of these identities. Each study employed a different strategy and slightly different nomenclature to classify their cells based on their

transcriptional identity. However, these labels can be compared and, when integrated to an *in vitro* dataset, they can reinforce the confidence in the results.

When the labels of the cell identities were visualize onto the UMAP plots of the published datasets, cells of the same type formed cohesive groups (Fig. 27a,d), as did the labels transferred on the *in vitro* differentiated cells (Fig. 27b,e). The cells that formed the *in vitro* dataset appeared to be transcriptionally very diverse when confronted with the embryo datasets and comprised many different cell identities. I plotted the integration results as heatmaps in order to determine the identity and proportion of cell types identified in each culture condition (Fig. 27c,f) and cluster (Fig. 27c',f'), defined as in figure 23c and d. To avoid overcrowding the results with underrepresented cell types, I only considered labels assigned to a minimum of 4% of the cells in at least one culture condition or cluster. I also ordered the cell types to facilitate their interpretation, starting at the top with pluripotent epiblast-like cell types, then primordial germ cells (PGCs), mesoderm subtypes, endoderm, and finally the ectoderm-related cell types.

The integration with the Pijuan-Sala et al. indicated that the pluripotent subgroups of epiblast-like cell types were predominantly found in conditions with higher FGF signaling, decreasing with no or low FGF2 and absent in the FGFRi condition (Fig. 27c). Consistently, these pluripotent cells were absent from clusters that corresponded to the FGFRi condition, cluster 5 and 6, but interestingly also from cluster 1, mostly populated by the highest FGF2 condition. Cells labeled as epiblast were mostly found in clusters 3 and 2, predominantly populated by cells cultured with 0 or 0,75 ng/ml FGF2, while cells identified as caudal epiblast or primitive streak were mostly found in clusters 0 and 4, populated by cells cultured with at least 0,75ng/ml FGF2 (Fig. 27c'). Two of these clusters, cluster 3 and 4, were also found to be associated with earlier developmental stages (Fig. 27c',f'). These results suggest that the maintenance of a pool of pluripotent cells require basal FGF signaling and some of its subtypes are assisted by the exogenous FGF2. Higher exogenous FGF2 also favored the differentiation of a small subset of PGCs (primordial germ cells), almost exclusively found in cluster 4.

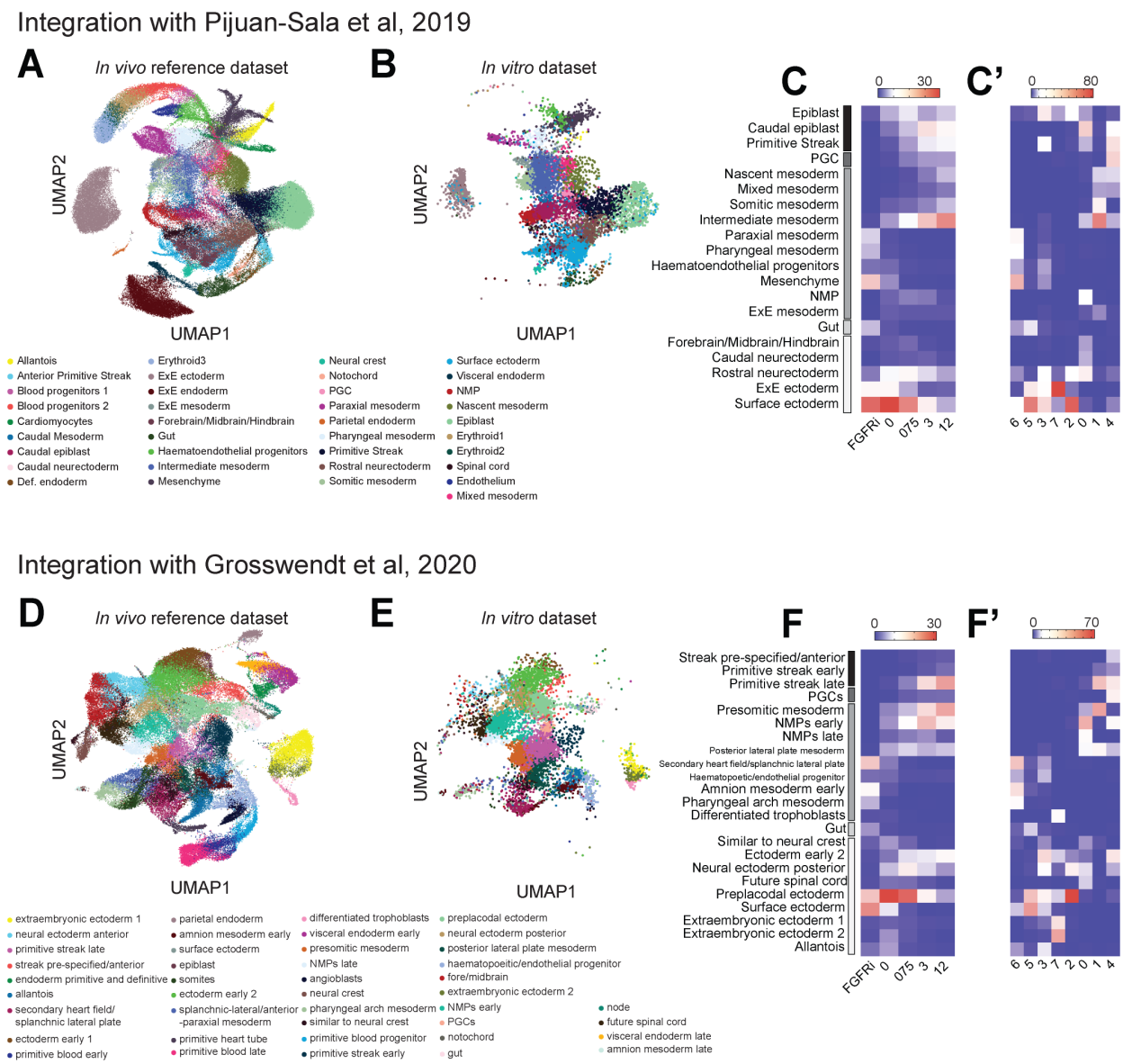


Fig. 27: Label transfer of cell identities from two fully annotated embryo datasets. *A-C')Dataset integration with Pijuan-Sala et al, 2019. D-F')Dataset integration with Grosswendt et al, 2020. A,D)Recalculated UMAP plots of the reference dataset, color-coded for cell identity. B,E)UMAP plot of the* in vitro *dataset, color-coded for cell identity. C,F)Heatmap of the percentage of cells with a specific cell identity label for different FGF signaling strength; cell types are clustered depending on their origin, from top to bottom: pluripotent epiblast-like cell types, primordial germ cells (PGCs), mesoderm subtypes, endoderm, and ectoderm-related cell types. C',F')Heatmap of the percentage of cells with a specific cell identity label for different clusters; cell types are clustered depending on their origin, from top to bottom: naïve epiblast-like cell types, primordial germ cells (PGCs), mesoderm subtypes, endoderm, and ectoderm-related cell types.*

Cells identifying as mesoderm belonged to two different groups of mesoderm labels, sitting at the opposite sides of the FGF signaling strength. The nascent, mixed, somitic and intermediate mesoderm, as well as the NMPs (neuro-mesodermal progenitors), were favored by increasing FGF2 dose and were mainly found in the clusters connected to higher exogenous FGF2: cluster 0, 1 and 4. In contrast, the paraxial and pharyngeal mesoderm, together with the hematoendothelial progenitors and mesenchyme, were mostly present in the FGFRi condition. Out of the 8 clusters, they were predominantly found in cluster 6, one of the two FGFRi clusters, and to a lower level in cluster 3. The mesoderm subtypes appeared then to be divided into a FGF dependent and independent group. A small group of cells in the FGFRi condition also identified as gut, and they all belonged to cluster 5.

The last group, the ectoderm-related cell identities, were mostly found in samples with low FGF to FGFR inhibition, and their corresponding clusters. Notable exception was the cluster 0, with a small percentage of caudal related neuroectoderm types (29,2%): given the caudal epiblast and neuro-mesodermal progenitor identity, these cells might be similar to the extending posterior section of the embryo.

The integration with the orthogonal embryo dataset of Grosswendt et al showed the transfer of broadly consistent labels on the samples and clusters (Fig. 27f,f'). Here, pluripotent cells were mostly identified as being related to the primitive streak and were more restricted to the samples and clusters characterized by high exogenous FGF2. PGCs and gut were, as before, primarily connected to one cluster, respectively cluster 4 and 5. The mesodermal subtypes identified by the integration were again found to be split into an FGF-dependent and independent group. Presomitic mesoderm, NMPs early and posterior lateral plate mesoderm were mostly increased with the dose of FGF2 and were found in the high FGF clusters: 0, 1 and 4. The secondary heart field/splanchnic lateral plate, amnion mesoderm early, pharyngeal arch mesoderm, and hematopoietic/endothelial progenitor were mostly found in cluster 6, one of the two clusters the FGFRi condition is mainly composed of. Also consistently with Pijuan-Sala 2019, the ectoderm-related cell identities were mostly found in samples with low FGF signaling strength.

The integration with the two annotated embryo datasets revealed that exogenous FGF2 supported the maintenance of a younger subgroup of cells and mesoderm identities that differentiate in more proximal parts of the primitive streak. Mesodermal cells that differentiate in a more proximal parts of the primitive streak were proven to differentiate only when FGF signaling was inhibited, proving to be FGF-independent. Low to no FGF signaling promoted ectoderm cell types.

FGF2 concentration affects endogenous cell signaling

Exogenous FGF2 promotes the endogenous expression of FGF ligands

Comparing the *in vitro* differentiated cells with annotated embryo datasets showed that FGF signaling affected the identity the cells obtained during differentiation and, to a lower level, the pace of differentiation. However, all experimental conditions presented a high heterogeneity of cell types, from the FGFRi, presenting two main clusters with a proximal mesodermal or ectodermal identity, to the cells with high exogenous FGF2, with the coexistence of distal mesoderm and younger pluripotent cells. Since the difference in cell type differentiation could be due to how the cells responded to their signaling environment, I first analyzed the expression of a panel of FGF target genes (Fig. 28a-c). The expression of these genes increased as expected with the exogenous FGF2 dose (Fig. 28a,b) and high expression levels were mainly restricted to the high FGF signaling clusters (0, 1 and 4, Fig 28c). The expression of FGF target expression was then visualized separately for each sample, subdividing the cells depending on their assigned cluster (Fig 28e): cells exposed to the same culture condition showed a variability of expression to the same exogenous FGF signaling. This result suggests that while the cell culture condition affected the number of cells in each cluster, cells exhibited a heterogenous response to the FGFRi or the FGF2 titration. This heterogenous response was not dependent on the differential expression of FGF receptors, as they were uniformly expressed across culture conditions and clusters (Fig. 28b-d). The main receptor Fgfr1 showed a high uniform expression across conditions, Fgfr3 and Fgfr4 had a uniformly low expression and Fgfr2 was higher in cluster 5.

I then considered the production of FGF ligands to ascertain whether their endogenous signaling could be correlated with the observed heterogeneous response. The

expression of Fgf4, Fgf3, Fgf17 and especially Fgf8 was found to be upregulated with the increase of FGF signaling strength and was particularly associated with cluster 0, 1 and 4 (Fig. 28b-d). The cluster association was maintained even when the expression of the FGFs was visualized separately for each culture condition (Fig. 28e). The clusters with higher levels of FGF target expression were the same clusters that upregulated the FGF ligands, suggesting an autocrine or paracrine mechanism.

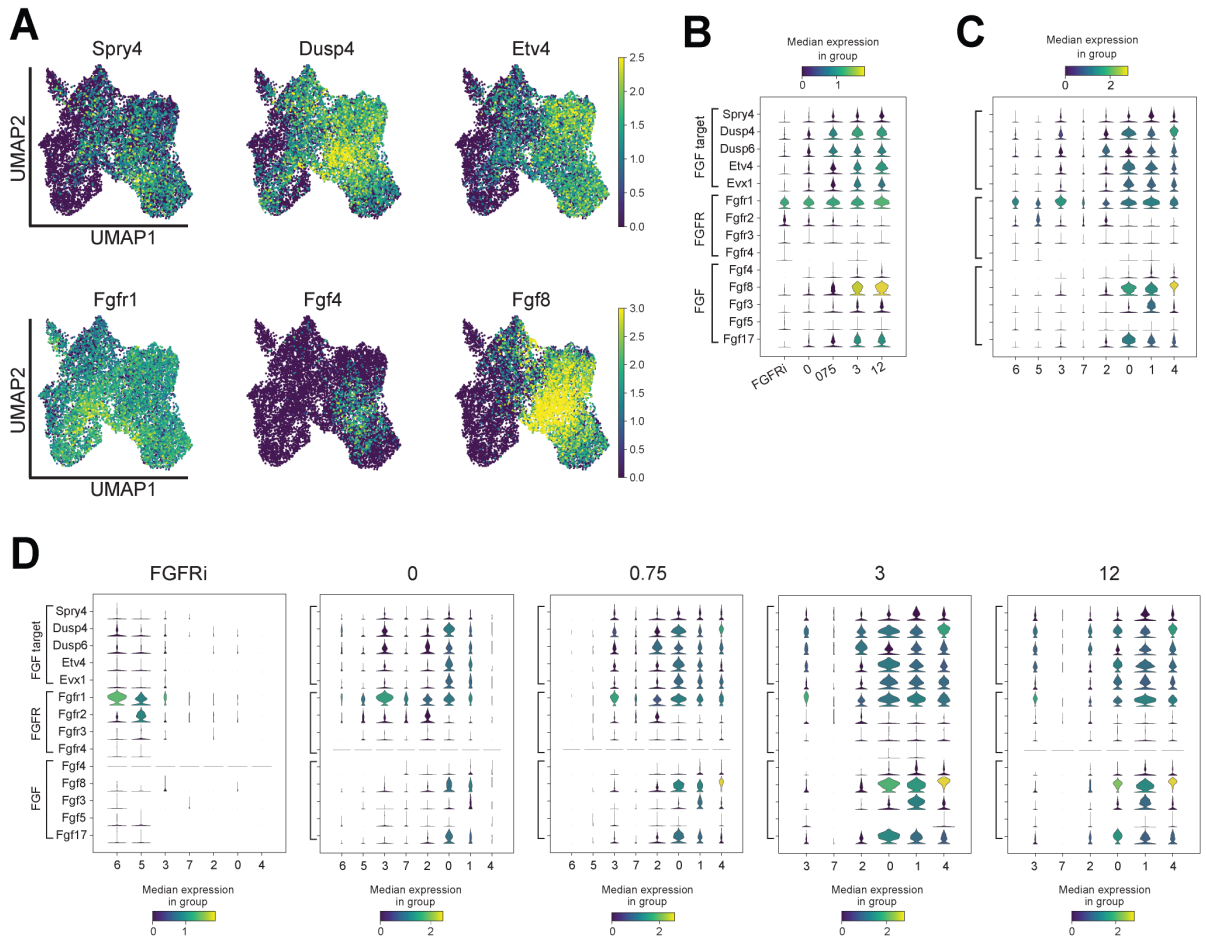


Fig. 28: Expression of FGF signaling-related proteins in* in vitro *differentiated cells.
A)UMAP plots coloured by the expression level of Spry4, Dusp4, Etv4, Fgfr1, Fgf4 and Fgf8 in gene expression space; brighter coloured regions exhibit higher expression levels. B,C,D)Stacked violin plots of the expression levels of FGF signaling proteins and receptors; width of the violin plots indicates the number of observations; the colour of the violin plot reflects the median expression of the selected gene for the cells contained in each observation. B)Stacked violin plots grouped by treatment condition. C)Stacked violin plots grouped by cluster. D)Stacked violin plots divided by treatment condition and grouped by cluster.

I concluded that exogenous FGF2 in the culture condition appeared to stimulate the endogenous expression of FGF ligands. These ligands could, in turn, amplify initial niches of FGF expression and determine different differentiation trajectories that generated the observed cell heterogeneity.

FGF signaling has diverging effects on WNT ligands and might affect Nodal signaling

The difference in FGF ligands expressed by the cells themselves suggested possible signaling niches in each condition, so I expanded the analysis to include other essential signaling in mouse gastrulation. In this experiment I tested the effect of increasing FGF signaling strength, from a small molecule inhibition of the FGF receptors to 12 ng/ml of exogenous FGF2. However, all other parameters were kept constant, including the concentration of the WNT agonist, Chiron, and the BMP4. Any difference between samples could be dependent on an interaction of the FGF signaling, or lack thereof, with the constant WNT and BMP.

While the expression of the WNT receptors Lrp5/6 did not vary across samples (Fig. 29a,b), the analysis of the WNT ligands revealed instead a opposite response of two groups of ligands (Fig. 29a). Wnt3, Wnt4 and Wnt6 were negatively correlated to the exogenous FGF2, exhibiting high expression for FGF2 $\leq$ 0,75 ng/ml: Wnt3 was connected to the clusters 6, 5 and 3, while Wnt4 and Wnt6 were more broadly expressed, including cluster 7 and 2 (Fig. 29a,b,c). All of these three ligands had their highest expression in cluster 5, connected to a more ectodermal identity (Fig. 29c',f'). The ligands Wnt3a, Wnt5a and Wnt5b were positively correlated to the exogenous FGF2, exhibiting high expression in clusters 0, 1 and 4, while Wnt5a expression extended to cluster 2, 3 and 6 (Fig. 29a,b,c). These WNT ligands are produced in the epiblast cells of the primitive streak during gastrulation and their presence in the gastrulating cells of the high FGF clusters supported their identification as mesoderm cells. The expression of inhibitors of WNT signaling seemed to be restricted to only specific clusters: Sfrp1 to cluster 1, Dickkopf1 (Dkk1) to clusters 6, 5 and 1 and Sfrp2 to clusters 0 and 6 (Fig. 29b,c).

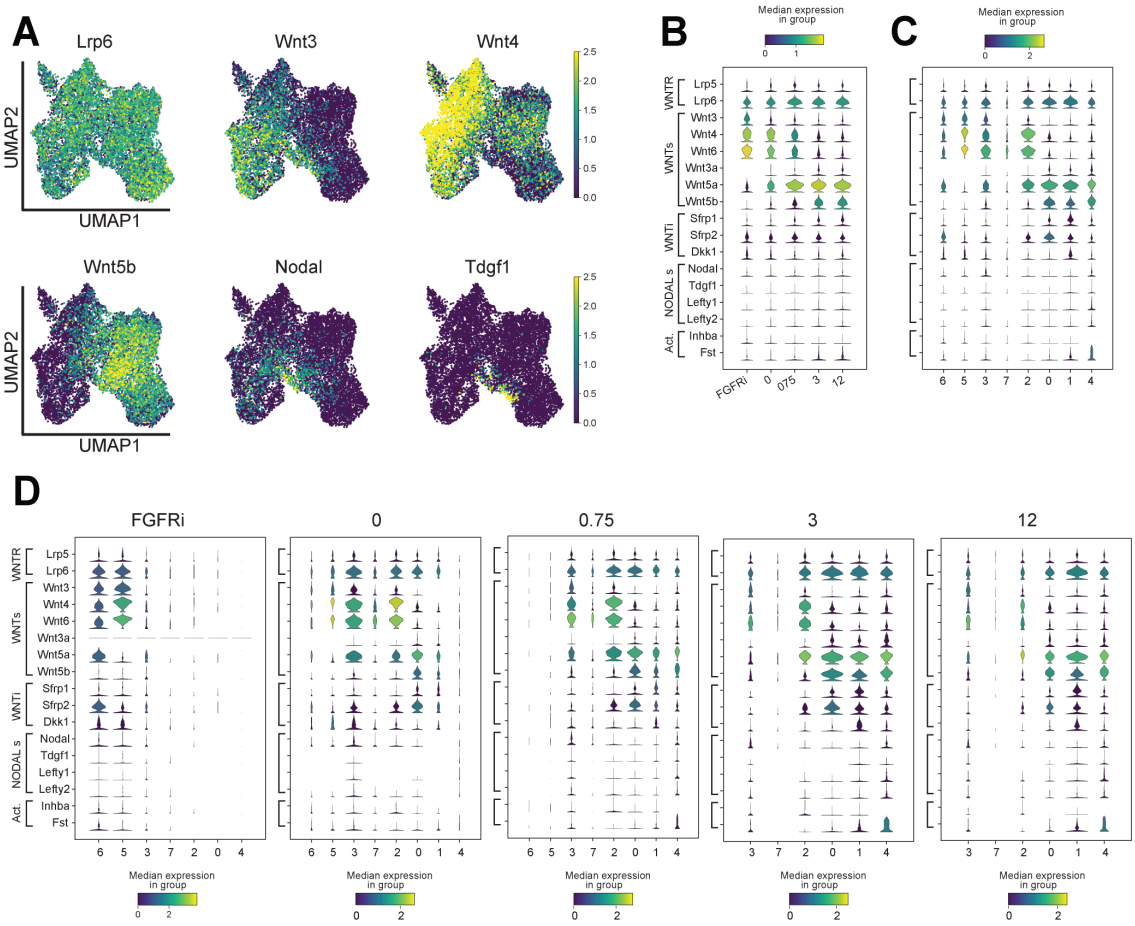


Fig. 29: Expression of WNT, NODAL and ACTIVIN signaling-related proteins in* in vitro *differentiated cells.

A)UMAP plots coloured by the expression level of Lpr6, Wnt3, Wnt4, Wnt3a, Wnt5a, Nodal in gene expression space; brighter coloured regions exhibit higher expression levels. B,C,D)Stacked violin plots of the expression levels of FGF signaling proteins and receptors; width of the violin plots indicates the number of observations; the colour of the violin plot reflects the median expression of the selected gene for the cells contained in each observation. B)Stacked violin plots grouped by treatment condition. C)Stacked violin plots grouped by cluster. D)Stacked violin plots divided by treatment condition and grouped by cluster.

The fraction of Nodal expressing cells did not increase or decrease in response to the FGF signaling strength and it comprised only a small subset of cells (Fig. 29a, b). However, except for the FGFRi condition, those Nodal^{+} cells were not uniformly distributed in every cluster: they were instead grouped in cluster 3 and 4 (Fig. 29c), characterized by having less differentiated cells (Fig. 26c',f'). This result was

consistent with the role of Nodal in maintaining pluripotency (Mesnard, Guzman-Ayala, and Constam 2006; Guzman-Ayala et al. 2004). The same cluster 3 and 4 also expressed the Nodal cofactor Tdgf1, also known as Cripto, and the inhibitors Lefty1 and Lefty2, whose gene expression was activated by Nodal itself. While Nodal was expressed in presence of FGFR inhibition, Tdgf1, Lefty1 and Lefty2 were not, suggesting that Nodal signaling might be active only in presence of endogenous or exogenous FGF signaling.

Activin A expression was low overall, as expected, since the epiblast cells do not produce Activin A during gastrulation, but its inhibitor Follistatin exhibited a high expression in cluster 4 and to a lower level in cluster 1 (Fig. 29b,c). Follistatin has been reported to be expressed at low levels in pre-implantation mouse embryo and ES cells and to increase greatly in the primitive streak and later in somites, identities similar to those of the clusters 4 and 1 (Albano and Smith 1994; Albano et al. 1994).

The results indicated that FGF signaling strength had opposite effect on two different groups of WNT ligands, increasing the expression of Wnt3a, Wnt5a and Wnt5b and decreasing the expression of Wnt3, Wnt4 and Wnt6. They also indicated that, while Nodal expression was constant through the samples, the expression of the cofactor Tdgf1 and Nodal inhibitors Lefty1 and Lefty2 were correlated to intermediate to high exogenous FGF2.

FGF signaling inhibits the expression of endogenous BMPs

I then analyzed the BMP signaling receptors and ligands. The receptor Bmpr1a had a constant high expression through culture conditions and clusters, while Bmpr2 showed a slightly negative correlation to FGF2 dose, with its expression slightly higher in clusters 5 and 6 (Fig 30a,b,c). However, expression of the BMP ligands was negatively affected by the exogenous FGF2: Bmp1, Bmp2 and Bmp4 expression was the highest in the cell sample treated with the inhibitor of FGFR and was then anticorrelated with the increase in FGF signaling strength. Even at the cluster level, clusters 6, 5 and 3 showed the highest expression, with Bmp4 extending to clusters 2 and 4 (Fig 30 a,b,c). The expression of BMP ligands appeared to be anticorrelated to the endogenous production of FGF ligands, with the exception of Bmp4 in cluster 4. Inhibitors of BMP,

Smad7 and Noggin (Nog), which are also target genes of this signaling, exhibited the same behavior as the BMP ligands and are downregulated with the increase in exogenous FGF2 (Fig 30 b,c).

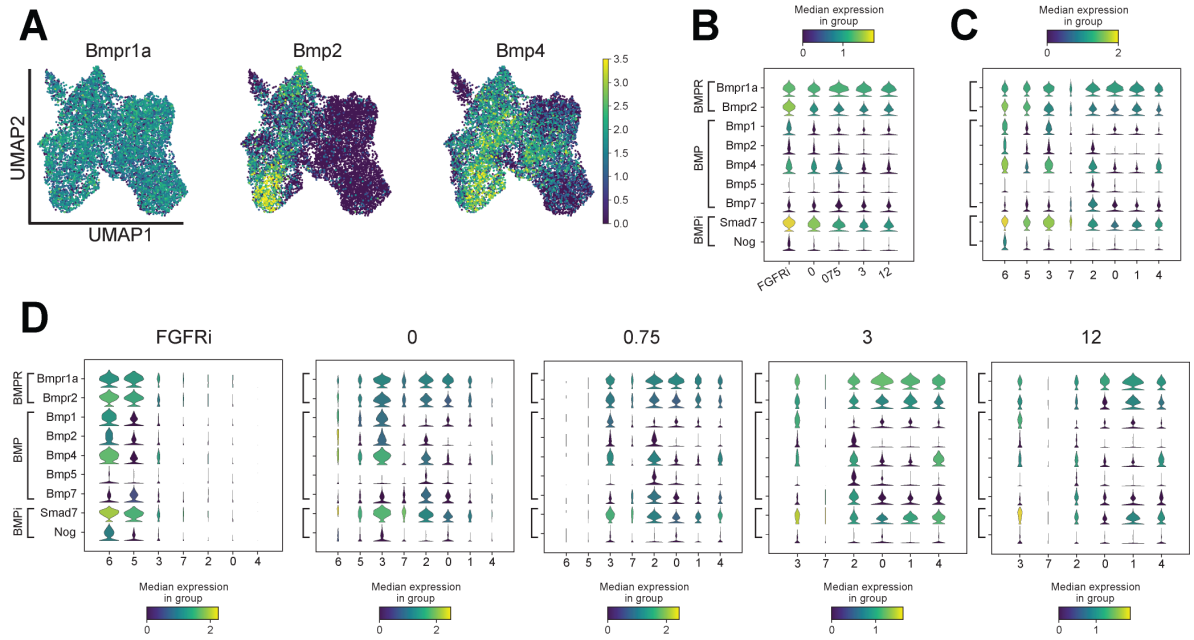


Fig. 30: Expression of BMP signaling-related proteins in* in vitro *differentiated cells.
A)UMAP plots coloured by the expression level of Bmpr1a, Bmp2 and Bmp4 in gene expression space; brighter coloured regions exhibit higher expression levels. B,C,D)Stacked violin plots of the expression levels of FGF signaling proteins and receptors; width of the violin plots indicates the number of observations; the colour of the violin plot reflects the median expression of the selected gene for the cells contained in each observation. B)Stacked violin plots grouped by treatment condition. C)Stacked violin plots grouped by cluster. D)Stacked violin plots divided by treatment condition and grouped by cluster.

The expression of endogenous BMP responded negatively to the FGF signaling and was anticorrelated with the expression of endogenous FGF. The mutually exclusive expression of BMP and FGF proteins by individual cells could act to locally divide the cells expressing these signaling proteins to determine more proximal mesoderm and ectodermal fates for the BMP expressing cell niches and more distal mesoderm and pluripotent cells fates for the FGF expressing cell niches. Under these experimental conditions, these distinct niches could be autonomously generated by populations of differentiating cells.

A timed FGF exposure can trigger a FGF positive feedback loop

A pulse of FGF on the second day is sufficient to maintain the expression of the T/Bra marker

The single cell RNA sequencing analysis suggested that exogenous FGF2 signalling could promote the endogenous production of FGF ligands in cells with a higher expression of FGF target genes, in a positive feedback loop. If the exogenous FGF2 in the culture medium could give rise to such positive feedback loop, then a limited exposure to the ligand would be sufficient to trigger a self-sustaining FGF loop and determine the expression of the same differentiation markers as the continuous exposures previously tested.

To verify this hypothesis and to identify the window of FGF2 sensitivity, I exposed the cells to 24 or 48 h pulses of 6 ng/ml FGF2 at different times during the three days of the differentiation protocol (Fig. 31a). The concentration of FGF2 was chosen based on the expression curve of T/Bra in figure 18, in order to maximise the expression range in the experiment. As in previous experiments, BMP4 and the WNT agonist Chiron were kept constant at 8 ng/ml and 1 μM respectively. The effect of the FGF pulsed treatment was analysed with a T/Bra reporter cell line (SBR line, Deluz et al., 2016), as this pan-mesoderm marker was shown to be highly correlated with the exogenous FGF2 and the expression of FGF-ligands (Fig. 18, 23c, 28c). Treatment with FGF2 on day 1 determine a slightly higher proportion of T/Bra$^+$ cells then cells exposed in day 3 (respective average: 13.1% and 4.8%), however, these conditions were both comparable to those of the cells not exposed to FGF2 (Fig. 31b,c). In contrast, FGF2 treatment during day 2 strongly increased the proportion of T/Bra$^+$ cells to levels comparable to the cells of the continuous treatment regimen (respective average: 43.3% and 59.8%).

Cells could then respond with a strong upregulation of T/Bra only when the 24h treatment was on day 2. Moreover, the T/Bra expression in the 48h treatment during day 2 and day 3 was only marginally higher than in the 24h treatment in day 2, suggesting that the FGF signaling could be self-maintaining after a pulse stimulation or that cells did not need FGF signaling on day 3 to differentiate into T/Bra^{+} cells.

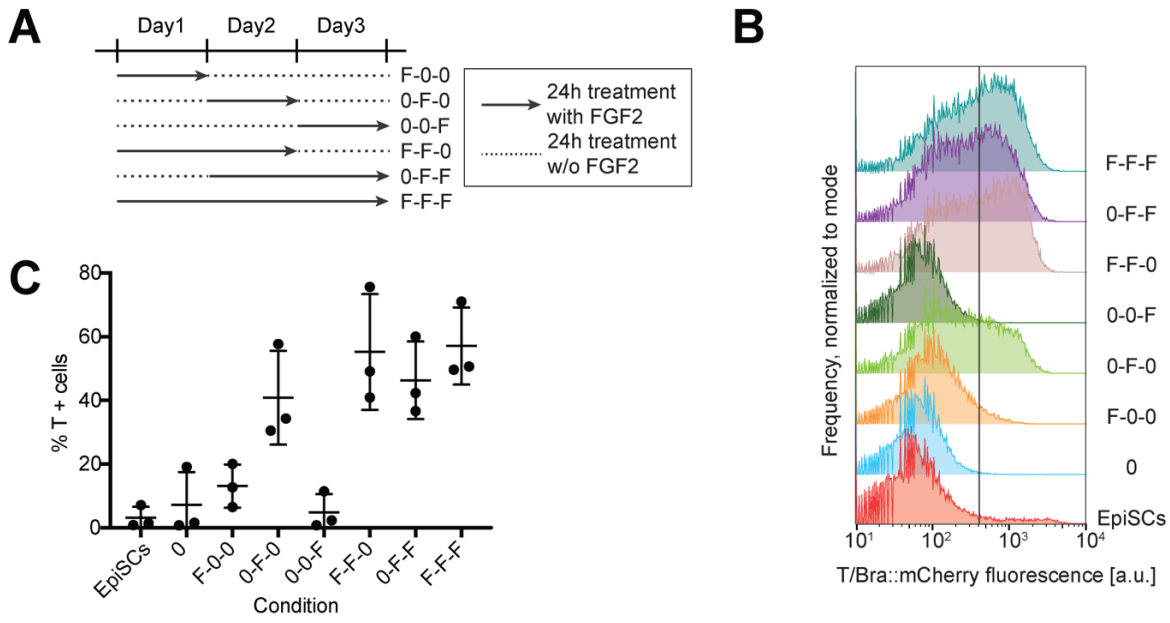


Fig. 31: T/Bra expression on a timed exposure of FGF2

A)Experimental scheme: uniform lines represent FGF2 treatment, dashed lines FGF2-free treatment; F = 24h treatment with FGF2, 0 = 24h without FGF2. B)Representative flow cytometry of T/Bra:: mCherry reporter line (SBR line) differentiated according to the experimental scheme; threshold between positive and negative cells is displayed. C)Quantification of T/Bra:: mCherry fluorescence positive cells from flow cytometry; error bars indicate SEM (n. experiments = 3).

<u>The FGF pulse on day 2 triggers a FGF positive signaling loop</u>

In order to directly test if FGF signaling could be self-maintaining after a pulse of exogenous FGF, I employed a Spry4::H2B-Venus line (Morgani, Saiz, et al. 2018), a reporter cell line for Spry4, a target gene of FGF signaling. Here I tested FGF2 and FGF4 to see whether distinct FGFs could differently affect the cell-cell communication. I employed here the same FGF2 concentration used in the timed exposure experiment and a FGF4 concentration that determined a similar T/BRA expression at the end of the differentiation protocol, based on the expression curve of T/Bra in figure 18. Pulses

of 6 ng/ml FGF2 or 12 ng/ml FGF4 on day 2 were compared to the continuous exposure to the ligands and with a FGFR inhibition following the FGFs pulse; as the previous experiment, BMP4 and WNT agonist Chiron were kept constant at 8 ng/ml and 1 µM.

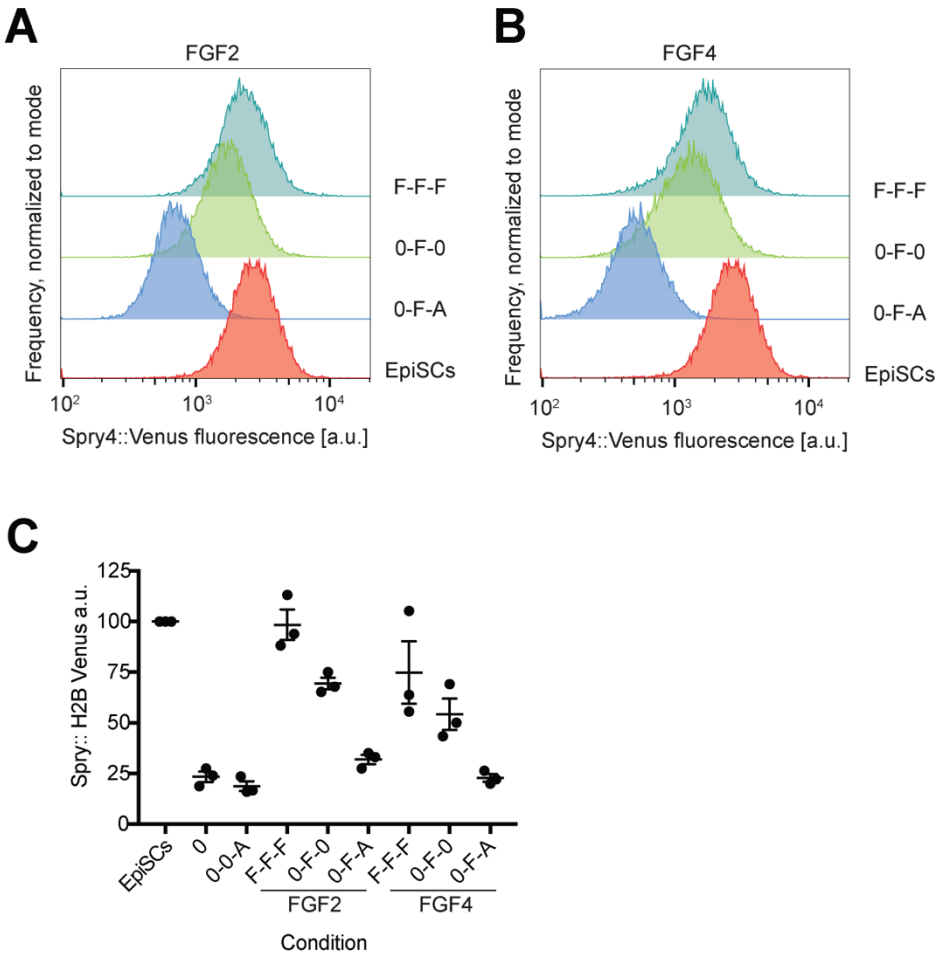


Fig. 32: Spry4 expression on a timed exposure of FGF2 or FGF4

A,B)Representative flow cytometry of Spry4::H2B-Venus reporter line differentiated with timed exposure to FGF2 or FGF4; F = 24h treatment with 6 ng/ml FGF2 or 12 ng/ml FGF4, 0 = 24h without FGFs, A = 24h treatment with 30 nM AZD4547 (FGFRi). C)Quantification of Spry4::H2B-Venus fluorescence positive cells from flow cytometry; y axis represents mean fluorescence of the cells, normalized to the EpiSCs value; error bars indicate SEM (n. experiments = 3).

The expression levels of the Spry4 reporter were found to be higher in cells differentiated with a continuous treatment with FGF2 or FGF4 (Fig. 32a-c). A high Spry4 expression was also found in the EpiSCs cells, likely determined by the 12 ng/ml FGF2 contained in the maintenance medium. Cells not exposed to FGFs exhibited a strongly reduced mean fluorescence and the FGFRi treatment on day 3 only

marginally reduced those expression levels. Pulse treatment on day 2 with FGF2 or FGF4 could maintain the reporter fluorescence to approximately 72% of the respective continuous treatment. The additional treatment on day 3 with the inhibitor of FGFR reduced this fluorescence to approximately 31% of the respective continuous treatment. This observation revealed that FGF signaling was self-maintained by the cells following a pulse of FGF2 or FGF4 treatment. It is possible that residual FGF exogenous ligand bound to the extracellular matrix would contribute to the signaling maintenance, but the strong difference between the FGF pulse conditions with or without the following inhibition of FGFR supported the existence of a FGF feedback loop.

These results indicated that endogenous FGF signaling was maintained after a pulse exposure to FGF2 or FGF4, confirming existence of a self-sustaining FGF feedback loop during in vitro mesoderm differentiation.

Endogenous FGF signaling is maintained in local cell clusters and suggests a community effect

The scRNAseq analysis showed that cells within the same culture could respond differently to the same signaling and indicated the existence of an intense cell-cell communication that affected the differentiation into a specific cell type (Fig. 27-30). These analyses further demonstrated that cells belonging to clusters with high expression of FGF targets, not only upregulated the endogenous production of FGF ligands but were also identified as specific mesoderm lineages (Fig. 27c',f', 28c). Since FGF has also been reported to act as a short-range signal (Raina et al. 2021), the FGF positive feedback loop could act locally, in signaling niches, and promote the differentiation of specific mesodermal cells.

I then performed an immunofluorescence imaging of the Spry4::H2B-Venus reporter cells differentiated with a pulse or a continuous exposure to 6 ng/ml FGF2; BMP4 and WNT agonist Chiron were kept constant at 8 ng/ml and 1 µM. I tested different culture conditions to investigate whether they affected the spatial localization of the cells expressing the FGF reporter gene. I also stained the cells for T/BRA, to determine whether the pan-mesoderm marker was expressed in the vicinity of the SPRY4^{+} cells.

In the culture conditions here tested, the Spry4::H2B-Venus^{high} cells were spatially clustered and contained nested groups of T/BRA^{+} cells (Fig. 33). Even though these clusters of positive cells may be formed from a common ancestor or could spatially cluster through cell adhesion, the observed clustering is consistent with a spatially restricted signaling environment created by a positive feedback based on short range endogenous FGF signaling.

The result revealed the existence of FGF expressing cell niches. The nested T/BRA^{+} cells also suggested a community effect, where a critical number of FGF expressing cells could be needed to induce the differentiation of specific mesoderm subtypes.

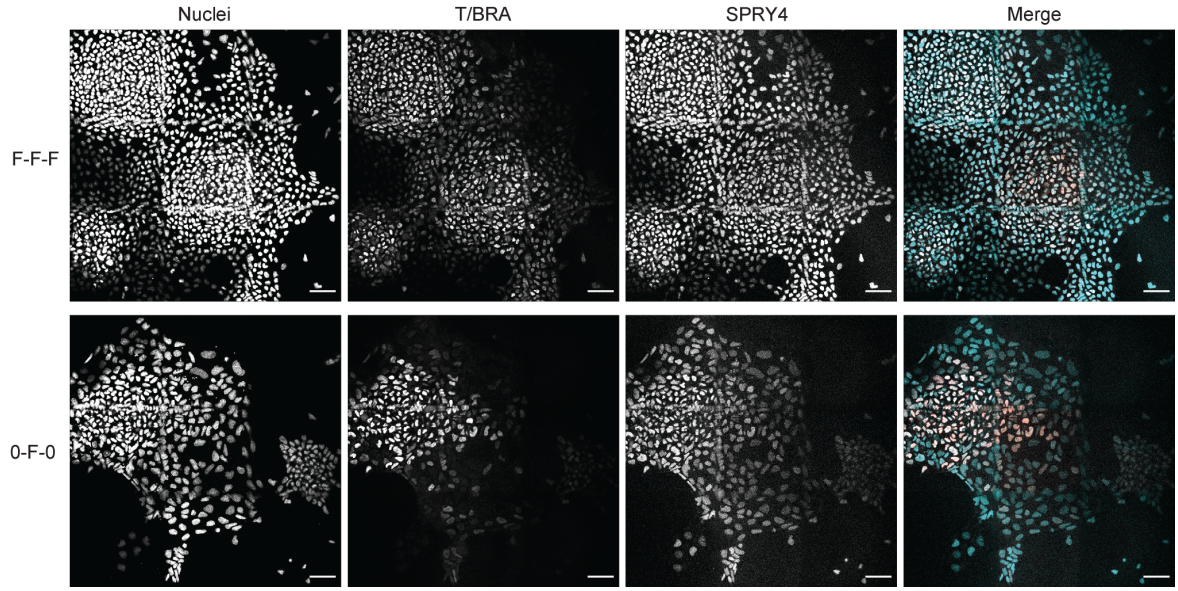


Fig. 33: Spry4^{+} cells form a spatial cluster and colocalize with a smaller cluster of T/Bra^{+} cells

Immunofluorescence of Spry4::H2B-Venus reporter line differentiated with timed exposure or continuous treatment with FGF2, stained with Hoechst and T/BRA. Images were reconstructed from a 3x3 multi-tile scan. T/BRA staining is shown in red and Spry4 in cyan. Scale bar = 500 µm.

DISCUSSION

A classical view in biology considers the epiblast cells as passively receiving instructions from the signaling centers and surrounding tissues to differentiate according to a specific fate map. The signaling landscape the cells are subjected to is then viewed as externally imposed. This is also reflected in many *in vitro* differentiation protocols in the literature, as they aim to push cells towards a specific differentiation by exposing cells to distinct signaling at precise timing (Loh et al. 2016; Rao et al. 2016; Mendjan et al. 2014).

In this work I have explored how do cells process externally given signaling proteins or chemicals to uncover the effect of BMP and FGF signaling in developing mesoderm. The population-based essays I have performed on titrations of BMP4 are compatible with the view of an externally imposed signal: cells seem to respond to a BMP4 gradient by obtaining a specific cell identity. However, in the single cell essays performed on the FGF titration I observed a high heterogeneity of cell identities and expression of signaling proteins. This suggested that cells differentiate in a specific lineage depending on the exogenously given signals, as well as the autocrine and paracrine signals from the neighboring cells. Those essays and following experiments also uncovered how the exogenous and endogenous FGF signaling might have influenced the endogenous production of BMP and FGF ligands in *in vitro* mesoderm differentiation.

In vitro mesoderm differentiation protocol

In vitro models for mammalian gastrulation are common in the field of developmental biology, as they allow to more easily control and analyze the differentiating cells. In this work, I employed a 2D differentiation protocol to identify how BMP and FGF signaling affect mesoderm differentiation. In this protocol I employed mouse EpiSCs, primed pluripotent stem cells that correspond to the epiblast cells prior to gastrulation.

Cells treated with FGF2, the WNT agonist Chiron and BMP4, downregulated the pluripotent marker Nanog and upregulated several mesoderm markers (Fig. 16, 17, 19, 22). The mesodermal identity was confirmed when the transcriptomic libraries

were integrated with fully annotated embryonic datasets (Fig. 27). Unlike previous studies (Vallier et al. 2009; Chal et al. 2015; Loh et al. 2016; Manfrin et al. 2019; Sudheer et al. 2016) here I carefully titrated the two signaling to see if any phenotypic effect or any change in cell-cell crosstalk were gradual or showed the existence of a threshold concentration. The protocol was also used to assess if and how a timed exposure could affect mesoderm differentiation or the cell-cell communication, showing consistent results when used across different cell lines. My work has also benefitted from a single cell and a transcriptomic approach in the analysis of the cell identities and cell-cell signaling.

BMP4 has a morphogenic effect in *in vitro* mesoderm differentiation

BMP4 is critical for the formation of the primitive streak and the development of mesoderm (Winnier et al. 1995; Lawson et al. 1999; Beppu et al. 2000; H. Zhang and Bradley 1996). It is a signaling protein produced at the beginning of gastrulation by the neighboring extraembryonic ectoderm and it is thought to establish a concentration gradient along the prospective primitive streak (Z. Zhang et al. 2019; Winnier et al. 1995; Ben-Haim et al. 2006; Morgani and Hadjantonakis 2019). BMP4 is considered as an externally imposed signaling that affect the mesoderm formation and patterning according to the local concentration experienced by the differentiating cells (Tam and Loebel 2007; Morgani and Hadjantonakis 2019).

Consistently with this hypothesis, I determined that at constant FGF and WNT signaling, cells could react to distinct concentrations of BMP4 by upregulating the marker genes of different mesoderm subtypes (Fig. 16, 17). Cell treated with high BMP4 concentration expressed the proximal mesoderm markers Hand1 and Gata6, while intermediate and low BMP4 concentrations upregulated the distal Tbx6 and Msgn1 and the more distal Foxa2 and Shh. The upregulation of Tbx6 and Msgn1 for intermediate concentration of BMP4 is also consistent with an *in vivo* study showing that a reduction in the BMP signaling causes the prospective proximal mesoderm cells to contribute to the formation of the somites derived from a more distal mesoderm subtype (Miura et al. 2006).

The BMP4 titration experiments also showed that the increase in BMP4 revealed a gradual change in the proportion of each mesoderm marker expression, rather than uncover concentration thresholds. The gradual change in the expression of different mesoderm markers suggests that the identity of the cells might not be steered completely in a specific cell lineage, possibly due to signaling interactions between the cells. Moreover, at each BMP4 concentration, cells expressed more than one mesoderm marker, indicating the coexistence of multiple cell types in the same cell culture condition. This heterogeneity could be the result of cells responding differently to the same culture conditions or an interaction between cells in culture (Fig. 16, 17).

These results confirm that BMP4 acts as a morphogen in the differentiation of mesoderm cells specified in the primitive streak and supports the hypothesis that BMP4 follows the positional information model theorised by Wolpert (Wolpert 1969).

Exogenous FGF promotes distal mesoderm and inhibits proximal mesoderm

Specific FGF ligands have been shown to be enriched in the posterior part of the embryo, and especially in the primitive streak (Maruoka et al. 1998; Morgani, Saiz, et al. 2018; Morgani and Hadjantonakis 2019). It has also been reported that in *in vivo* and *in vitro* studies, FGF signaling has appeared to be critical for the differentiation of mesoderm (Oki, Kitajima, and Meno 2010; Yu et al. 2011; Bernardo et al. 2011). However, it was unclear how different levels of exogenous FGF signaling would affect the differentiation of distinct mesoderm subtypes and how endogenous FGF ligands could also influence this patterning.

In order to identify the possible influence of FGF signaling, I titrated FGF2 on an intermediate BMP4 concentration that supported the differentiation of more than one mesoderm marker subtype (Fig. 17). Differentiating EpiSCs responded to this titration by gradually increasing the percentage of cells expressing the pan-mesoderm and distal mesoderm markers T/Bra and Tbx6, while decreasing the percentage of cells positive for the more proximal mesoderm marker Hand1 (Fig. 19, 22). The positive effect on the distal marker Tbx6 is consistent with a previous *in vivo* study, where a

mutation of the FGF receptor caused strong defects in the pre-somitic mesoderm, a lineage derived from distal mesoderm (Yamaguchi et al. 1994). This *in vitro* result indicated that the effect can be also observed in isolated EpiSCs and it is therefore not dependent on the interaction with extraembryonic tissues. As previously observed in the BMP4 titration, the FGF2 effect on mesoderm patterning was gradual and did not exhibit concentration thresholds. Intermediate to high FGF2 concentration exhibited as well the co-expression of more than one mesoderm marker, confirming a cell heterogeneity of the cultured cells. Taken together, these findings support the idea that that the diverging identity of the differentiating cells might be influenced by endogenously produced cell signaling proteins.

The integration of the single cell RNA sequencing dataset with fully annotated embryonic datasets (Pijuan-sala et al. 2019; Grosswendt et al. 2020) shed light on the identity of the cells differentiated *in vitro* and demonstrated that the FGF2 titration gradually shifted the proportion of cells classified with specific cell identities. Cells differentiated with intermediate to high concentrations of FGF2 (3 and 12 ng/ml) were mostly identified in the scRNAseq as pre-somitic, NMPs or intermediate mesoderm (clusters 0, 1 and 4, Fig. 27c,f). These were cell types that differentiated in a more distal position of the streak, confirming the earlier findings that FGF2 promoted distal mesoderm subtypes. It also represented an opposite effect observed for BMP in this *in vitro* differentiation model. The intermediate to high concentrations of FGF2 also saw the emergence of a subpopulation of cells with a more pluripotent identity, closer to the epiblast or the primitive streak (cluster 4, Fig. 27c, c', f, f').

Cells cultured with low FGF2 concentration or without exogenous FGF (0.75 and 0 ng/ml) revealed a gradual shift towards more ectodermal identities: while a low percentage of cells retained a distal mesodermal lineage, the majority identified as surface ectoderm or preplacodal ectoderm, which respectively form the epidermis or part of the eyes, and smaller proportions of neural ectoderm and extraembryonic ectoderm (clusters 2, 3 and 7, Fig. 27c',f'). In the mouse embryo, the cells that differentiate into surface ectoderm have been reported to derive from a proximal position between the anterior and the posterior sections of the embryo (Tam and Behringer 1997). There, cells could still be subjected to the BMP signaling from the ExE, without BMP inhibitors from the AVE and with low or no FGFs and WNTs from

the posterior primitive streak. *In vitro*, surface ectoderm can be obtained by culturing pluripotent cells with BMP4, which inhibits neural differentiation, yet without any WNT agonist or ligand (Qiao et al. 2012; Hemmati-Brivanlou and Meltont 1997; Muñoz-Sanjuán and Brivanlou 2002). In this work, BMP4 indeed inhibited the neural differentiation of most of the cells, pushing them into an epidermal identity, but ~ 10% of the cells still obtained a neural ectoderm fate (Fig. 27). Without high exogenous FGF, the WNT agonist that these cells were cultured in was not enough to efficiently promote a mesoderm differentiation or prevent an ectodermal identity, except for a cell subgroup of mesodermal cells. The emergence of the mesodermal cells is possibly caused by an endogenous FGF signaling, since a proportion of these cells belong to the mesodermal cluster 0, characterized by a higher endogenous expression of FGF ligands (Fig. 23b, 27, 28c). Together, these results refuse a previous study showing that the removal of FGF2 can change the differentiation induced by BMP4 from embryonic to extraembryonic lineages, not ectoderm (Yu et al. 2011).

When cells were cultured with an inhibitor of FGFR, cells acquired different transcriptional identities compared to the other samples (Fig. 21e). These cells differentiated into two very distinct cell types: ectoderm and mesoderm cells (Fig. 27c,f). The ectoderm cells in cluster 5 were mainly identified as surface ectoderm by the integration with the embryo datasets (Fig. 27c',f'). The mesodermal cells in cluster 6 were identified as a mix of mesenchyme, heart field and amnion mesoderm. These mesoderm subtypes, contrary to those obtained in high FGF2 cell cultures, differentiate in a more proximal part of the embryo and therefore might be more subjected to exogenous BMP4 *in vivo*. The emergence of a subpopulation of proximal mesoderm when cells are cultured with WNT, BMP4 and FGFRi was unexpected and it was in stark contrast with cells cultured without either FGFR inhibition or FGF2 (0 ng/ml condition). Those cells, however, had the highest endogenous expression of BMP4: it is possible that this had a role in the differentiation of more proximal mesoderms, that in previous experiment were shown to depend more strongly on BMP signaling (Fig. 16, 17).

The differentiation of mesoderm and endoderm from the primitive streak is also characterized by the epithelial-to-mesenchymal transition (EMT), where cells lose the cell-cell adhesion and gain migratory and invasive properties. EMT has been

previously correlated with FGF signaling (Ciruna and Rossant 2001; Sun et al. 1999). Consistently with previous studies that report EMT defects for FGF signaling mutation, here I observed a correlation between the exogenous FGF2 and an increase of Cdh2 expression at the expense of Cdh1 (Fig. 22). This change in adhesion properties has been observed in all clusters that in the integration with the embryo dataset had the higher percentage of mesoderm cells, specifically clusters 6, 0, 1 and 4. These findings indicate that the mesoderm differentiation protocol here employed also triggers the epithelial-to-mesenchymal transition.

Exogenous FGF might delay the differentiation of a subgroup of cells through Nodal signaling

The integration with the fully annotated embryonic datasets (Pijuan-sala et al. 2019; Grosswendt et al. 2020) also revealed the developmental stages of the *in vitro* differentiated cells, demonstrating that the majority of the cells identified with a E7.75-E8.25 timepoint. Since the cells employed for the *in vitro* differentiation protocol are equivalent to epiblast cells just prior to gastrulation, around E5.0, cells are observed to differentiate with a similar pace *in vivo* and *in vitro*. Nevertheless, the FGF signaling had an effect on a cell subset, as exogenous FGF2 was correlated to the appearance of a younger subgroup of cells (Fig. 26c,f). These younger cells were mostly identified within cluster 3, ~15 to 30% of the total, and cluster 4, ~18 to 36% of the total (Fig. 26c',f'). These clusters are identified by very different types of cells. Cluster 3 was mostly formed by cells exposed to low or no exogenous FGF2 and was populated by ectodermal cells and subpopulations of cells with an early ectoderm or primitive streak identity (Fig. 27c',f'). In contrast, cluster 4 was mostly populated by cells exposed to intermediate to high exogenous FGF2 and contained distal mesoderm cells and a smaller group of primitive streak or epiblast cells (Fig. 27c',f').

The scRNAseq analysis also revealed that, excluding the FGFRi condition, the low number of Nodal^{+} cells converged in the same clusters 3 and 4 (Fig. 29). Nodal^{+} cells are also found in the FGFRi-rich clusters (cluster 5 and 6), but appeared to be evenly distributed. The expression of the Nodal cofactor Tdgf1 (also known as Cripto), as well as the Nodal targets and inhibitors Lefty1 and Lefty2, was also restricted to clusters 3

and 4, suggesting that Nodal signaling might be active when FGF signaling is not inhibited (Fig. 29). As Nodal is known in other embryonic stages and cell types to promote stemness, I speculate that Nodal activity might be involved in the maintenance of the pluripotent pool of cells observed *in vitro* (Mesnard, Guzman-Ayala, and Constam 2006; Guzman-Ayala et al. 2004). No exogenous Nodal or Activin has been given to the cells and the FGF signaling strength did not affect the fraction of Nodal^{+} cells, but it appears to affect directly or indirectly the activity of Nodal signaling. FGF4 and FGF8 have also been suggested to have a role, *in vivo*, in maintaining a pluripotent pool of cells used for the elongation of the animal trunk (Yamaguchi et al. 1994; Boulet and Capecchi 2012). Nodal and FGF signalling could then influence a subset of cells to remain less differentiated or differentiate with a slower pace.

On the contrary, the FGFRi condition comprised an older subgroup of cells identifying as E8.5. These cells were found to have higher concentration in the mesodermal cluster 6, making up ~45% of the cells, while the ectodermal cluster 5 contained ~ 20-25% E8.5 cells (Fig. 26c',f'). Since the proximal mesoderm subtypes differentiated in cluster 6 are differentiated earlier *in vivo*, it raises the possibility that the lateral mesoderm cells *in vitro* could also display a faster differentiation. Moreover, a previous study has demonstrated that a reduction of BMP signaling in mouse embryos can cause a delay in the recruitment of the prospective mesoderm cells to the primitive streak and a delay in the primitive streak extension (Miura et al. 2006). Since the FGFRi cells endogenously express BMP ligands, it is possible that higher BMP signaling could speed the differentiation pace of both cluster 5 and cluster 6 cells.

Effect of exogenous FGF on cell-cell signaling

<u>High FGF2 and FGFR inhibition promote different WNTs</u>

In the scRNAseq analysis, FGF2 treatment was correlated with the expression of specific WNT ligands: Wnt3a, Wnt5a and Wnt5b (Fig. 29). The expression of these ligands is elevated in clusters with a high percentage of distal mesoderm cells, identified as cluster 0, 1 and especially cluster 4, which also contained a subgroup of pluripotent cells. These results were consistent with the *in vivo* analysis of the expression patterns of these WNT ligands, as they have been found to be connected

to the distal mesoderm formation (Takada et al. 1994; Yamaguchi, Bradley, et al. 1999). Moreover, the higher expression in cluster 4 could be connected with the younger subgroup of cells it contains, as well as the older NMP cells. Indeed, Wnt3a, Wnt5a and Wnt5b were found to be maintained in the tailbud region of the embryo, a region that continues to generate new mesoderm and elongates the trunk, producing neuronal and mesodermal progenitors (Takada et al. 1994).

This result also confirms previous studies that suggest that FGF could be upstream of the WNT signaling and upregulate it (Ciruna and Rossant 2001; Kelly, Pinson, and Skarnes 2004; Tam and Loebel 2007; Morgani and Hadjantonakis 2019). This result is not dependent of the expression of the WNT receptors, as Lrp6 is constantly expressed in every condition. It is also not connected to a change in the exogenous WNT signaling, as a constant concentration of the WNT agonist Chiron has been added to the cell culture medium in every cell culture condition here tested.

Exogenous FGF2 demonstrated a opposite effect on other WNT ligands, specifically Wnt3, Wnt4 and Wnt6 (Fig. 29). This suggests a negative correlation between their expression and the endogenous or exogenous FGF signaling. The Wnt3 expression is in contrast with previous studies, as its mutation has been reported to impair the initiation of gastrulation and the differentiation of mesoderm and endoderm (Liu et al. 1999). The expression of Wnt4 and Wnt6 was particularly high in clusters with a high percentage in ectoderm subtypes: cluster 5, followed by 7, 3 and 2. It is possible that these WNT ligands could have a role in the differentiation of the ectoderm: the Pijuan-Sala 2019 dataset partly confirms this possibility, as it shows an overexpression of Wnt6 in surface ectoderm and extraembryonic ectoderm, while Wnt4 had a low overall expression (Pijuan-sala et al. 2019).

FGF inhibits the expression of BMP ligands

In the scRNAseq analysis, the exogenous FGF2 also appeared to inhibit the endogenous expression of several BMP ligands, especially Bmp1, 2 and 4 (Fig. 29). Their expression was found to be higher in conditions and clusters with no exogenous FGF or FGFR inhibition, identified as proximal mesoderm or ectodermal subtypes. An outlier is cluster 4, where both signals are expressed: the balance of different signals

is consistent with the similarities between this cluster and the tailbud region of the embryo (Row et al. 2018). The inhibition of endogenous BMP ligands is highly dependent on the cell cluster, indicating a strong effect of the endogenously produced FGF ligands, as well as the exogenously provided FGF2 (Fig. 27b,c, 29b,c). The FGF-mediated inhibition of BMPs has been reported in the elongation of the mouse tail, where the FGF3 ligand downregulates the endogenous production of BMP in the neuroepithelium (Anderson, Schimmang, and Lewandoski 2016). My results indicate instead that this mechanism could be more general, as it can apply to more FGF ligands and to another differentiation process.

The opposite mechanism has also been observed in the zebrafish embryo, where BMPs can inhibit the activity of transcriptional regulators of the FGF signaling at the level of transcriptional targets, promoting the differentiation of tissues derived from proximal mesoderm (Row et al. 2018). When the two directions of inhibitions between FGF and BMP signaling are combined at the cell population level, this would lead to a toggle switch mechanism between different cells (Fig. 34).

The toggle switch mechanism between FGF and BMP might lead to a generation of signaling niches with discrete cell identities, as it was observed in figure 33.

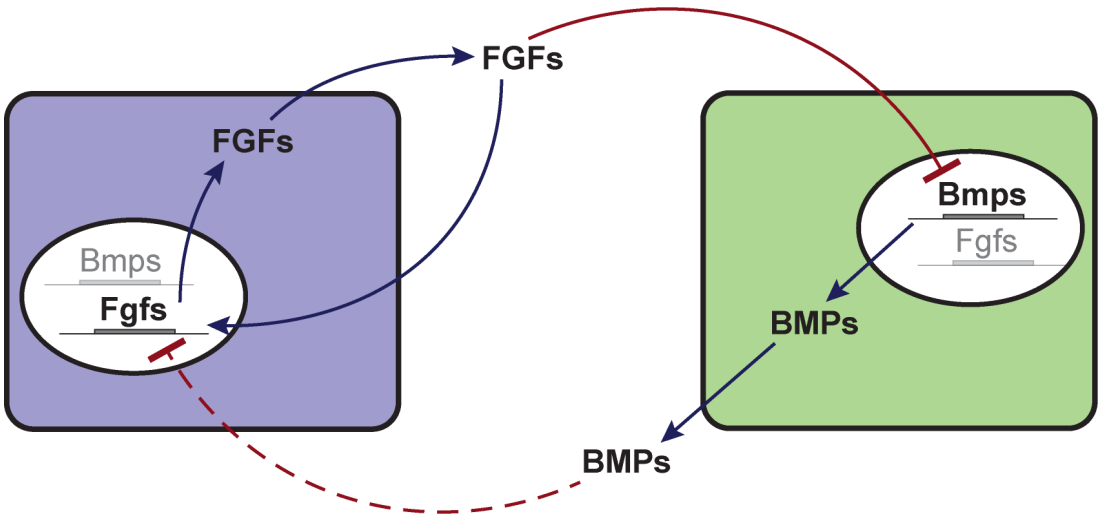


Fig. 34: Model of the effect of FGF and BMP signaling on neighboring cells

Dotted line is based on published results (Row et al, 2018)

Exogenous FGF2 triggers a FGF positive feedback loop upon a 24 h exposure on day 2

The scRNAseq experiment revealed differences in cell identities and stages between cells cultured without exogenous FGFs or in presence of a FGFR inhibitor (Fig. 26-30). These phenotypic differences suggested the existence of a basal endogenous FGF-mediated crosstalk, confirmed by the expression analysis of the FGF ligands. Indeed, cells cultured without exogenous FGF expressed low levels of selected FGFs, in particular Fgf3, Fgf4, Fgf8 and Fgf17. These FGF ligands are found *in vivo* to be upregulated in the primitive streak, indicating that the EpiSCs responded to the *in vitro* differentiation with a cell-cell crosstalk similar to the *in vivo* primitive streak (Maruoka et al. 1998; Crossley and Martin 1995; Morgani and Hadjantonakis 2019).

The scRNAseq experiment also showed a strong correlation between a higher expression of these FGF ligands, the increasing concentration of exogenous FGF2 in the culture medium and a higher expression of FGF targets (Fig. 28a,b). This result implies that cells that were exposed to and responded to FGF signaling would also produce FGF ligands, suggesting a positive FGF signaling feedback loop that could work in an autocrine or paracrine fashion (Fig. 34). The effect was not correlated with a difference in the FGFRs, as the main receptor Fgfr1 showed a constant high expression throughout the conditions (Fig. 28). This was a novel and unexpected result in the cell-cell communication during mesoderm differentiation.

The experiments run on a Sox1-T/Bra reporter cell line tested the existence of the FGF positive feedback loop. They demonstrated that a limited exposure to the exogenous FGF2 would be sufficient to trigger a similar phenotypic response, with a high percentage of T/BRA$^+$ cells. The window of sensitivity was identified to be equal to 24h on day 2 of the differentiation protocol. Indeed, the treatment almost recapitulated the fraction of T/BRA$^+$ cells found when cells were continuously exposed to FGF2 (Fig. 31). This high sensitivity was exclusive to day 2: I speculate that the differentiation induced by BMP4 and the WNT agonist Chiron on day 1 could determine the right conditions to induce FGF sensitivity in the differentiating cells. Cells on the day 3 of differentiation, instead, could be already committed to a different identity.

The same 24h FGF exposure on day 2 was also sufficient to maintain high expression of the FGF target Spry4, confirming the existence of a self-maintaining FGF loop (Fig. 32, 34). The effect was the same when two different FGFs were used, the FGF2 used in this work or the FGF4 also expressed in the gastrulating mouse embryo.

In culture, the Spry4$^+$ cells were organized in tight spatial clusters, containing T/Bra$^+$ cells at their center (Fig. 33). This tight spatial arrangement is consistent with the Fgf signaling acting as a short-range signal (Raina et al. 2021) and suggests the existence of a community effect dependent on the positive FGF feedback loop. The community effect describes how a critical number of cells in close contact could produce and respond to a paracrine signaling by differentiating into the same lineage (Gurdon et al. 1993). Previous research have demonstrated that a community effect based on eFGF crosstalk is critical, in Xenopus, for the differentiation of muscle progenitors, a line close to the somitic mesoderm and the NMP differentiated in this work with high exogenous FGF2 (Gurdon et al. 1993; Standley, Zorn, and Gurdon 2001).
In this work, the community effect might be dependent on FGF8. This ligand was shown to be critical for mesoderm differentiation (Sun et al. 1999; Boulet and Capecchi 2012) and was found here to have the highest overexpression when cells were treated with high FGF2 (3 and 12 ng/ml) and in particular in the clusters 0, 1 and 4, exhibiting high proportion of distal mesoderm identity (Fig. 28b,c).

The shortest treatment pulse tested was equal to 24h and I did not investigate shorter FGF pulses. However, previous research confirms that the cells would need a long exposure to FGF, at least 3 to 5h, to trigger the community effect, speculating that it was critical to consolidate its differentiation pathway and to enable all the cells in the community to differentiate co-ordinately (Standley, Zorn, and Gurdon 2001). This is unlike the rapid induction of Activin A or BMP, as they have been shown to need an incubation of only 10 to 30 minutes to induce a specific cell differentiation (Gunne-Braden et al. 2020; Dyson and Gurdon 1998).

Proposed molecular basis of the FGF positive feedback loop

Experiments run in this work indicate the existence of a novel positive FGF loop acting in the differentiation of distal mesoderm (Fig. 28, 31, 32, 33). They do not however

suggest whether this might be a direct effect of FGF ligands or if other signaling molecules or transcription factors might be involved in a paracrine/autocrine fashion or within the cell itself.

It has been observed that in different tissues of the mouse embryo, Nodal signaling could promote the expression of Fgf4 and Fgf8 (Guzman-Ayala et al. 2004; Mesnard, Guzman-Ayala, and Constam 2006), but this mechanism is not involved in the observed FGF positive feedback loop, as the expression of Nodal and Tdgf1 is not correlated with the FGF ligands expression (Fig. 28b, 29b).

However, the expression of FGF ligands is observed in the scRNAseq to be correlated with other signaling molecules or transcription factors, in particular with selected WNT ligands and the pan-mesoderm marker and transcription factor T/Bra (Fig. 22, 23c, 28, 29). It has indeed been reported that FGF signaling could be upstream of WNT signaling (Ciruna and Rossant 2001; Kelly, Pinson, and Skarnes 2004) and that WNT signaling can upregulate T/Bra (Arnold et al. 2000; Yamaguchi, Takada, et al. 1999). The transcription factor T/Bra, in turn, has been shown to have Fgf8 between its targets (Evans et al. 2012). Considering these previous studies and the correlation I have observed, I then speculate that the positive FGF feedback loop could act through selected WNT ligands and T/Bra (Fig. 35). A similar correlation has also been noted *in vivo*, where a double mutation of Fgf4 and Fgf8 caused the inhibition of Wnt3a, Wnt5b and T/Bra (Boulet and Capecchi 2012). Moreover, FGF and T/Bra have been reported to activate the expression of each other in the Xenopus embryo before gastrulation (Isaacs, Pownall, and Slack 1994; Schulte-Merker and Smith 1995).

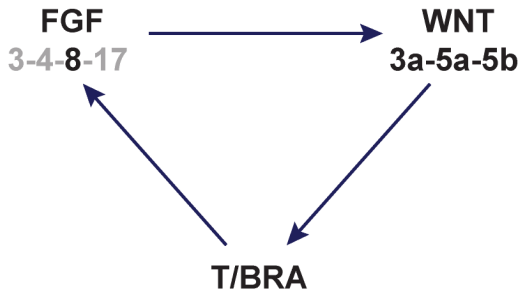


Fig. 35: Proposed molecular basis for the FGF positive feedback loop

The ligands and transcription factors here highlighted were found to have a correlated expression in the scRNAseq analysis and each separate interaction is reported in literature (Ciruna et al, 2001, Kelly et al, 2004; Arnold et al, 2000, Yamaguchi et al, 1999; Evans et al, 2012; Boulet and Capecchi, 2012; Isaacs et al, 1994; Schulte-Merker and Smith, 1995). The transcription factor T/BRA is established to have FGF8 as target, while I infer the interaction with the other FGF ligands.

The restriction of the FGF sensibility on day 2 of the differentiation protocol might be due to the changes operated by Chiron and BMP4 during the first 24 h of differentiation, while after 48 h, cells might be locked into a specific differentiation trajectory, analyzed in this work at endpoint in the 0 ng/ml FGF2 in the scRNAseq. There, cells display a high expression of Bmp4 ligands, suggested to inhibit FGF signaling (Row et al. 2018), and different WNTs ligands. The concentration of the Wnt agonist Chiron is constant in all conditions, but it appears to be insufficient to start this positive feedback loop, as it might need a reinforcement from FGF exogenous or endogenous signaling.

Conclusions and future directions

With this thesis I investigated the FGF and BMP signaling function in the differentiation and patterning of mesoderm. I first developed a cell model for mesoderm differentiation, which I used to investigate the effect of increasing concentrations of different ligands and their timed exposure on mesoderm patterning. With this model, I found that increasing concentrations of the BMP4 ligand promoted different mesoderm identities that *in vivo* are found along the proximal-distal axis of the embryo, the same axis along which BMP4 is reported to establish a gradient. This result confirmed that

BMP4 can control the trajectories of cell differentiation in a concentration-dependent manner, compatible with the effect of an externally imposed morphogen.
The experiments conducted on the FGF signaling showed an opposite effect compared to BMP4, as FGF promoted distal mesoderm subtypes and inhibited proximal mesoderm and surface ectoderm. This represented a novel role for FGF signaling.
While BMP4 and FGF influenced the type of mesoderm produced, I found that the epiblast cells were behaving as an active tissue, contributing to the formation and maintenance of a signaling network *in vitro*. This cell-cell crosstalk had a phenotypic effect, connecting specific signaling to cell identities or differentiation pace. Indeed, FGF endogenous crosstalk inhibited the differentiation of proximal mesoderm; FGF signaling niches promoted the expression of the mesoderm marker T/Bra and likely even distal mesoderm identities. FGF has been shown to interact with directly or indirectly with different endogenous signaling, such as WNT, BMP, NODAL and even determined a FGF positive feedback loop. The interactions between FGF and BMP suggested a toggle switch mechanism, while WNT ligands and the transcription factor T/Bra could be key elements of the FGF feedback loop.

Building on this work, the 2D mesoderm differentiation protocol could be employed to characterize the cell-cell signaling I have observed with the transcriptomic analysis.
It could be used to investigate the spatial arrangement of Spry4^{+} cells: the size of the cell clusters, the relative distance between them and the effect of the exogenous FGF2. The presence of markers for proximal or distal mesoderm lineages could be explored to assess if they form clusters for mesoderm marker^{+} cells and if they are connected with the Spry4^{+} clusters. This could also be a starting point to explore the FGF-BMP toggle switch mechanism.
The FGF feedback loop could be tested with specific knockout cell lines. FGF8 could be a key ligand for the loop I have observed: it is essential for gastrulation (Sun et al. 1999), it is a target of T/Bra (Evans et al. 2012) and I found it to have a particularly high expression in response to exogenous FGF2 (Fig. 28b,c). FGF8 mutation could impair or eliminate the spatial patterning I have observed and require additional exogenous FGF to rescue the differentiation of distal mesoderm. The proposed molecular mechanism for the FGF feedback loop could instead be tested with mutation

in the T/Bra gene or the WNT ligands WNT5a and WNT5b, given their high expression and role in mesoderm differentiation (Takada et al. 1994).

REFERENCES

Albano, R. M., R. Arkell, Rosa S.P. Beddington, and J. C. Smith. 1994. "Expression of Inhibin Subunits and Follistatin during Postimplantation Mouse Development: Decidual Expression of Activin and Expression of Follistatin in Primitive Streak, Somites and Hindbrain." *Development* 120 (4): 803–13. https://doi.org/10.1242/dev.120.4.803.

Albano, R. M., and J. C. Smith. 1994. "Follistatin Expression in ES and F9 Cells and in Preimplantation Mouse Embryos." *International Journal of Developmental Biology* 38 (3): 543–47.

Amerongen, Renée van, and Roel Nusse. 2009. "Towards an Integrated View of Wnt Signaling in Development." *Development* 136 (19): 3205–14. https://doi.org/10.1242/dev.033910.

Anderson, Matthew J., Thomas Schimmang, and Mark Lewandoski. 2016. "An FGF3-BMP Signaling Axis Regulates Caudal Neural Tube Closure, Neural Crest Specification and Anterior-Posterior Axis Extension." *PLoS Genetics* 12 (5): 1–30. https://doi.org/10.1371/journal.pgen.1006018.

Arnold, Sebastian J., Ulf K. Hofmann, Elizabeth K. Bikoff, and Elizabeth J. Robertson. 2008. "Pivotal Roles for Eomesodermin during Axis Formation, Epithelium-to-Mesenchyme Transition and Endoderm Specification in the Mouse." *Development* 135 (3): 501–11. https://doi.org/10.1242/dev.014357.

Arnold, Sebastian J., and Elizabeth J Robertson. 2009. "Making a Commitment: Cell Lineage Allocation and Axis Patterning in the Early Mouse Embryo." *Nature Reviews. Molecular Cell Biology* 10 (2): 91–103. https://doi.org/10.1038/nrm2618.

Arnold, Sebastian J., Joèrg Stappert, Andreas Bauer, Andreas Kispert, Bernhard G Herrmann, and Rolf Kemler. 2000. "Brachyury Is a Target Gene of the Wnt/b -Catenin Signaling Pathway." *Mechanisms of Development* 91 (1–2): 249–58. https://doi.org/doi: 10.1016/s0925-4773(99)00309-3.

Bedzhov, Ivan, Sarah J L Graham, Chuen Yan Leung, and Magdalena Zernicka-Goetz. 2014. "Developmental Plasticity, Cell Fate Specification and Morphogenesis in the Early Mouse Embryo." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 369 (1657): 20130538-. https://doi.org/10.1098/rstb.2013.0538.

Bedzhov, Ivan, and Magdalena Zernicka-Goetz. 2014. "Self-Organizing Properties of Mouse Pluripotent Cells Initiate Morphogenesis upon Implantation." *Cell* 156 (5): 1032–44. https://doi.org/10.1016/j.cell.2014.01.023.

Ben-Haim, Nadav, Cindy Lu, Marcela Guzman-Ayala, Luca Pescatore, Daniel Mesnard, Mirko Bischofberger, Felix Naef, Elizabeth J Robertson, and Daniel B. Constam. 2006. "The Nodal Precursor Acting via Activin Receptors Induces Mesoderm by Maintaining a Source of Its Convertases and BMP4." *Developmental Cell* 11 (3): 313–23. https://doi.org/10.1016/j.devcel.2006.07.005.

Beppu, Hideyuki, Masahiro Kawabata, Toshiaki Hamamoto, Anna Chytil, Osamu Minowa, Tetsuo Noda, and Kohei Miyazono. 2000. "BMP Type II Receptor Is Required for Gastrulation and Early Development of Mouse Embryos." *Developmental Biology* 221 (1): 249–58. https://doi.org/10.1006/dbio.2000.9670.

Bernardo, Andreia S., Tiago Faial, Lucy Gardner, Kathy K. Niakan, Daniel Ortmann, Claire E. Senner, Elizabeth M. Callery, et al. 2011. "BRACHYURY and CDX2 Mediate BMP-Induced Differentiation of Human and Mouse Pluripotent Stem Cells into Embryonic and Extraembryonic Lineages." *Cell Stem Cell* 9 (2): 144–55. https://doi.org/10.1016/j.stem.2011.06.015.

Boulet, Anne M., and Mario R. Capecchi. 2012. "Signaling by FGF4 and FGF8 Is Required for Axial Elongation of the Mouse Embryo." *Developmental Biology* 371 (2): 235–45. https://doi.org/10.1016/j.ydbio.2012.08.017.

Cano, Amparo, Mirna A. Pérez-Moreno, Isabel Rodrigo, Annamaria Locascio, María J. Blanco, Marta G. Del Barrio, Francisco Portillo, and M. Angela Nieto. 2000. "The Transcription Factor Snail Controls Epithelial-Mesenchymal Transitions by Repressing E-Cadherin Expression." *Nature Cell Biology* 2 (2): 76–83. https://doi.org/10.1038/35000025.

Chal, Jérome, Masayuki Oginuma, Ziad Al Tanoury, Bénédicte Gobert, Olga Sumara, Aurore Hick, Fanny Bousson, et al. 2015. "Differentiation of Pluripotent Stem Cells to Muscle Fiber to Model Duchenne Muscular Dystrophy." *Nature Biotechnology* 33 (August): 962–69. https://doi.org/10.1038/nbt.3297.

Chamier, Lucas Von, Johanna Jukkala, Christoph Spahn, Martina Lerche, Sara Hernández-pérez, Pieta K Mattila, Eleni Karinou, et al. 2020. "ZeroCostDL4Mic : An Open Platform to Simplify Access and Use of Deep-Learning in Microscopy." *BioRxiv*.

Chazaud, Claire, and Yojiro Yamanaka. 2016. "Lineage Specification in the Mouse Preimplantation Embryo." *Development (Cambridge, England)* 143 (7): 1063–74. https://doi.org/10.1242/dev.128314.

Ciruna, Brian, and Janet Rossant. 2001. "FGF Signaling Regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak." *Developmental Cell*

1 (1): 37–49. https://doi.org/10.1016/S1534-5807(01)00017-X.

Clevers, Hans, and Roel Nusse. 2012. “Wnt/β-Catenin Signaling and Disease.” *Cell* 149 (6): 1192–1205. https://doi.org/10.1016/j.cell.2012.05.012.

Crossley, P. H., and G. R. Martin. 1995. “The Mouse Fgf8 Gene Encodes a Family of Polypeptides and Is Expressed in Regions That Direct Outgrowth and Patterning in the Developing Embryo.” *Development* 121 (2): 439–51.

Deluz, Cédric, Elias T. Friman, Daniel Strebinger, Alexander Benke, Mahé Raccaud, Andrea Callegari, Marion Leleu, Suliana Manley, and David M. Suter. 2016. “A Role for Mitotic Bookmarking of SOX2 in Pluripotency and Differentiation.” *Genes and Development* 30 (22): 2538–50. https://doi.org/10.1101/gad.289256.116.

Demagny, Hadrien, Tatsuya Araki, and Edward M. De Robertis. 2014. “The Tumor Suppressor Smad4/DPC4 Is Regulated by Phosphorylations That Integrate FGF, Wnt, and TGF-β Signaling.” *Cell Reports* 9 (2): 688–700. https://doi.org/10.1016/j.celrep.2014.09.020.

Dyson, Steven, and J. B. Gurdon. 1998. “The Interpretation of Position in a Morphogen Gradient as Revealed by Occupancy of Activin Receptors.” *Cell* 93 (4): 557–68. https://doi.org/10.1016/S0092-8674(00)81185-X.

Evans, Amanda L., Tiago Faial, Michael J. Gilchrist, Thomas Down, Ludovic Vallier, Roger A. Pedersen, Fiona C. Wardle, and James C. Smith. 2012. “Genomic Targets of Brachyury (T) in Differentiating Mouse Embryonic Stem Cells.” *PLoS ONE* 7 (3). https://doi.org/10.1371/journal.pone.0033346.

Ferretti, Elisabetta, and Anna Katerina Hadjantonakis. 2019. “Mesoderm Specification and Diversification: From Single Cells to Emergent Tissues.” *Current Opinion in Cell Biology* 61: 110–16. https://doi.org/10.1016/j.ceb.2019.07.012.

Franch-Marro, Xavier, Franz Wendler, Janice Griffith, Madelon M. Maurice, and Jean Paul Vincent. 2008. “In Vivo Role of Lipid Adducts on Wingless.” *Journal of Cell Science* 121 (10): 1587–92. https://doi.org/10.1242/jcs.015958.

Fuentealba, Luis C., Edward Eivers, Atsushi Ikeda, Cecilia Hurtado, Hiroki Kuroda, Edgar M. Pera, and Edward M. De Robertis. 2007. “Integrating Patterning Signals: Wnt/GSK3 Regulates the Duration of the BMP/Smad1 Signal.” *Cell* 131 (5): 980–93. https://doi.org/10.1016/j.cell.2007.09.027.

Gharibi, Borzo, Emanuel Gonçalves, Buhe Nashun, Alex Montoya, Rahuman S M Sheriff, Alessandro Cicarrelli, Joana Carvalho, et al. 2020. “A FGF2-Mediated Incoherent Feedforward Loop Induces Long-Term Erk Inhibition and Sustains Naïve Pluripotency.”

Gilbert, Scott F., and Michael J. F. Barresi. 2016. “Developmental Biology. 11th Edition.”

Green, Jeremy B A, and James Sharpe. 2015. “Positional Information and Reaction-Diffusion: Two Big Ideas in Developmental Biology Combine.” *Development (Cambridge, England)* 142 (7): 1203–11. https://doi.org/10.1242/dev.114991.

Grosswendt, Stefanie, Helene Kretzmer, Zachary D. Smith, Abhishek Sampath Kumar, Sara Hetzel, Lars Wittler, Sven Klages, Bernd Timmermann, Shankar Mukherji, and Alexander Meissner. 2020. *Epigenetic Regulator Function through Mouse Gastrulation*. *Nature*. Vol. 584. Springer US. https://doi.org/10.1038/s41586-020-2552-x.

Gu, Zhenyu, Ellen M. Reynolds, Jihwan Song, Hong Lei, Alie Feijen, Lian Yu, Weiwu He, et al. 1999. “The Type I Serine/Threonine Kinase Receptor ActRIA (ALK2) Is Required for Gastrulation of the Mouse Embryo.” *Development* 126 (11): 2551–61. https://doi.org/10.1242/dev.126.11.2551.

Gunne-Braden, Alexandra, Adrienne Sullivan, Borzo Gharibi, Roy Wollman, Philip East, Silvia D M Santos, Alexandra Gunne-braden, et al. 2020. “GATA3 Mediates a Fast , Irreversible Commitment to BMP4-Driven Differentiation in Human Embryonic Stem Cells Article GATA3 Mediates a Fast , Irreversible Commitment to BMP4-Driven Differentiation in Human Embryonic Stem Cells.” *Stem Cell*, 1–14. https://doi.org/10.1016/j.stem.2020.03.005.

Gurdon, J. B. 1988. “A Community Effect in Animal Development.” *Nature* 336 (6201): 772–74. https://doi.org/10.1038/336772a0.

Gurdon, J. B., E. Tiller, J. Roberts, and K. Kato. 1993. “A Community Effect in Muscle Development.” *Current Biology* 3 (1): 1–11. https://doi.org/10.1016/0960-9822(93)90139-F.

Guzman-Ayala, Marcela, Nadav Ben-Haim, Séverine Beck, and Daniel B. Constam. 2004. “Nodal Protein Processing and Fibroblast Growth Factor 4 Synergize to Maintain a Trophoblast Stem Cell Microenvironment.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (44): 15656–60. https://doi.org/10.1073/pnas.0405429101.

Han, Dong Wook, Natalia Tapia, Jin Young Joo, Boris Greber, Marcos J. Araúzo-Bravo, Christof Bernemann, Kinarm Ko, et al. 2010. “Epiblast Stem Cell Subpopulations Represent Mouse Embryos of Distinct Pregastrulation Stages.” *Cell* 143 (4): 617–27. https://doi.org/10.1016/j.cell.2010.10.015.

Harrison, Sarah Ellys, Berna Sozen, Neophytos Christodoulou, Christos Kyprianou, and Magdalena Zernicka-goetz. 2017. “Assembly of Embryonic and Extra-Embryonic Stem Cells to Mimic Embryogenesis in Vitro.” https://doi.org/10.1126/science.aal1810.

Haub, O., and M. Goldfarb. 1991. “Expression of the Fibroblast Growth Factor-5 Gene in the Mouse Embryo.” *Development* 112 (2): 397–406.

Hébert, Jean M., Thomas Rosenquist, Jürgen Götz, and Gail R. Martin. 1994. “FGF5 as a Regulator of the Hair Growth Cycle: Evidence from Targeted and Spontaneous Mutations.” *Cell* 78 (6): 1017–25. https://doi.org/10.1016/0092-8674(94)90276-3.

Hemmati-Brivanlou, Ali, and Douglas Meltont. 1997. “Vertebrate Embryonic Cells Will Become Nerve Cells Unless Told Otherwise.” *Cell* 88 (1): 13–17. https://doi.org/10.1016/S0092-8674(00)81853-X.

Hermitte, Stéphanie, and Claire Chazaud. 2014. “Primitive Endoderm Differentiation: From Specification to Epithelium Formation.” *Philosophical Transactions of the Royal Society of London. Series B.* 369 (1657): 20130537. https://doi.org/10.1098/rstb.2013.0537.

Hogan, Brigid L.M. 1996. “Bone Morphogenetic Proteins in Development.” *Current Opinion in Genetics and Development* 6 (4): 432–38. https://doi.org/10.1016/S0959-437X(96)80064-5.

Isaacs, H. V., M. E. Pownall, and J. M.W. Slack. 1994. “EFGF Regulates Xbra Expression during Xenopus Gastrulation.” *EMBO Journal* 13 (19): 4469–81. https://doi.org/10.1002/j.1460-2075.1994.tb06769.x.

Itoh, Nobuyuki, and David M. Ornitz. 2010. “Fibroblast Growth Factors: From Molecular Evolution to Roles in Development, Metabolism and Disease.” *Journal of Biochemistry* 149 (2): 121–30. https://doi.org/10.1093/jb/mvq121.

Kattman, Steven J, Alec D Witty, Mark Gagliardi, Nicole C Dubois, Maryam Niapour, Akitsu Hotta, and James Ellis. 2010. “Resource Stage-Specific Optimization of Activin / Nodal and BMP Signaling Promotes Cardiac Differentiation of Mouse and Human Pluripotent Stem Cell Lines.” *Stem Cell* 8 (2): 228–40. https://doi.org/10.1016/j.stem.2010.12.008.

Kelly, Olivia G., Kathy I. Pinson, and William C. Skarnes. 2004. “The Wnt Co-Receptors Lrp5 and Lrp6 Are Essential for Gastrulation in Mice.” *Development* 131 (12): 2803–15. https://doi.org/10.1242/dev.01137.

Kinder, Simon J., Tania E. Tsang, Gabriel A. Quinlan, Anna Katerina Hadjantonakis, Andras Nagy, and Patrick P L Tam. 1999. “The Orderly Allocation of Mesodermal Cells to the Extraembryonic Structures and the Anteroposterior Axis during Gastrulation of the Mouse Embryo.” *Development* 126 (21): 4691–4701.

Kintner, C. R., and J. Dodd. 1991. “Hensen’s Node Induces Neural Tissue in Xenopus Ectoderm: Implications for the Action of the Organizer in Neural Induction.” *Development* 113 (4): 1495–1505. https://doi.org/10.1242/dev.113.4.1495.

Knoetgen, Hendrik, Ulrike Teichmann, Lars Wittler, Christoph Viebahn, and Michael Kessel. 2000. “Anterior Neural Induction by Nodes from Rabbits and Mice.” *Developmental Biology* 225 (2): 370–80. https://doi.org/10.1006/dbio.2000.9834.

Kretzschmar, Marcus, Jacqueline Doody, and Joan Massagué. 1997. “Opposing BMP and EGF Signalling Pathways Converge on the TGF-β Family Mediator Smad1.” *Nature* 389 (6651): 618–22. https://doi.org/10.1038/39348.

Lake, David, Sonia A.L. Corrêa, and Jürgen Müller. 2016. “Negative Feedback Regulation of the ERK1/2 MAPK Pathway.” *Cellular and Molecular Life Sciences* 73 (23): 4397–4413. https://doi.org/10.1007/s00018-016-2297-8.

Lane, Mary Constance, and Michael D. Sheets. 2006. “Heading in a New Direction: Implications of the Revised Fate Map for Understanding Xenopus Laevis Development.” *Developmental Biology* 296 (1): 12–28. https://doi.org/10.1016/j.ydbio.2006.04.447.

Lawson, Kirstie A., N. Ray Dunn, Bernard A.J. Roelen, Laura M. Zeinstra, Angela M. Davis, Christopher V.E. Wright, Jeroen P.W.F.M. Korving, and Brigid L.M. Hogan. 1999. “Bmp4 Is Required for the Generation of Primordial Germ Cells in the Mouse Embryo.” *Genes and Development* 13 (4): 424–36. https://doi.org/10.1101/gad.13.4.424.

Lawson, Kirstie A., Juanito J Meneses, and Roger A Pedersen. 1991. “Clonal Analysis of Epiblast Fate during Germ Layer Formation in the Mouse Embryo.” *Development* 113: 891–911. https://doi.org/VL - 113.

Liu, Pentao, Maki Wakamiya, Martin J. Shea, Urs Albrecht, Richard R. Behringer, and Allan Bradley. 1999. “Requirement for Wnt3 in Vertebrate Axis Formation.” *Nature Genetics* 22 (4): 361–65. https://doi.org/10.1038/11932.

Loh, Kyle M M., Angela Chen, Pang Wei W. Koh, Tianda Z Z. Deng, Rahul Sinha, Jonathan M M. Tsai, Amira A A. Barkal, et al. 2016. “Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types.” *Cell* 166 (2): 451–68. https://doi.org/10.1016/j.cell.2016.06.011.

Lu, C. C., J. Brennan, and E. J. Robertson. 2001. “From Fertilization to Gastrulation: Axis Formation in the Mouse Embryo.” *Current Opinion in Genetics and Development* 11 (4): 384–92. https://doi.org/10.1016/S0959-437X(00)00208-2.

Manfrin, Andrea, Yoji Tabata, Eric R. Paquet, Ambroise R. Vuaridel, François R. Rivest, Felix Naef, and Matthias P. Lutolf. 2019. “Engineered Signaling Centers for the Spatially Controlled Patterning of Human Pluripotent Stem Cells.” *Nature Methods* 16 (7): 640–48. https://doi.org/10.1038/s41592-019-0455-2.

Mansour, Suzanne L, Judy M Goddard, and Mario R Capecchi. 1993. “Mice Homozygous for

a Targeted Disruption of the Proto-Oncogene Int-2 Have Developmental Defects in the Tail and Inner Ear." *Development* 117: 13–28.

Martinez Arias, Alfonso, and Ben Steventon. 2018. "On the Nature and Function of Organizers." *Development (Cambridge)* 145 (5). https://doi.org/10.1242/dev.159525.

Maruoka, Yutaka, Norihiko Ohbayashi, Masamitsu Hoshikawa, Nobuyuki Itoh, Brigid L.m. Hogan, and Yasuhide Furuta. 1998. "Comparison of the Expression of Three Highly Related Genes, Fgf8, Fgf17 and Fgf18, in the Mouse Embryo." *Mechanisms of Development* 74 (1–2): 175–77. https://doi.org/10.1016/S0925-4773(98)00061-6.

Massey, Joseph, Yida Liu, Omar Alvarenga, Teresa Saez, Matthew Schmerer, and Aryeh Warmflash. 2019. "Synergy with TGFβ Ligands Switches WNT Pathway Dynamics from Transient to Sustained during Human Pluripotent Cell Differentiation." *Proceedings of the National Academy of Sciences of the United States of America* 116 (11): 4989–98. https://doi.org/10.1073/pnas.1815363116.

McInnes, Leland, John Healy, and James Melville. 2018. "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction." http://arxiv.org/abs/1802.03426.

Mendjan, Sasha, Victoria L. Mascetti, Daniel Ortmann, Mariaestela Ortiz, Dyah W. Karjosukarso, Yifan Ng, Thomas Moreau, and Roger A Pedersen. 2014. "NANOG and CDX2 Pattern Distinct Subtypes of Human Mesoderm during Exit from Pluripotency. Supplementary." *Cell Stem Cell* 15 (3): 310–25. https://doi.org/10.1016/j.stem.2014.06.006.

Mesnard, Daniel, Marcela Guzman-Ayala, and Daniel B. Constam. 2006. "Nodal Specifies Embryonic Visceral Endoderm and Sustains Pluripotent Cells in the Epiblast before Overt Axial Patterning." *Development* 133 (13): 2497–2505. https://doi.org/10.1242/dev.02413.

Mishina, Yuji, Atsushi Suzuki, Naoto Ueno, and Richard R. Behringer. 1995. "Bmpr Encodes a Type I Bone Morphogenetic Protein Receptor That Is Essential for Gastrulation during Mouse Embryogenesis." *Genes and Development* 9 (24): 3027–37. https://doi.org/10.1101/gad.9.24.3027.

Miura, Shigeto, Shannon Davis, John Klingensmith, and Yuji Mishina. 2006. "BMP Signaling in the Epiblast Is Required for Proper Recruitment of the Prospective Paraxial Mesoderm and Development of the Somites." *Development* 133 (19): 3767–75. https://doi.org/10.1242/dev.02552.

Miyazawa, Keiji, and Kohei Miyazono. 2017. "Regulation of TGF-β Family Signaling by Inhibitory Smads." *Cold Spring Harbor Perspectives in Biology* 9 (3).

https://doi.org/10.1101/cshperspect.a022095.

Mizuno, Seiya, Saori Iijima, Tomoko Okano, Noriko Kajiwara, Satoshi Kunita, Fumihiro Sugiyama, and Ken ichi Yagami. 2011. "Retrotransposon-Mediated Fgf5go-Utr Mutant Mice with Long Pelage Hair." *Experimental Animals* 60 (2): 161–67. https://doi.org/10.1538/expanim.60.161.

Morgani, Sophie M., and Anna Katerina Hadjantonakis. 2019. "Signaling Regulation during Gastrulation: Insights from Mouse Embryos and in Vitro Systems." In *Current Topics in Developmental Biology*. Academic Press Inc. https://doi.org/10.1016/bs.ctdb.2019.11.011.

Morgani, Sophie M., Jakob J. Metzger, Jennifer Nichols, Eric D. Siggia, and Anna Katerina Hadjantonakis. 2018. "Micropattern Differentiation of Mouse Pluripotent Stem Cells Recapitulates Embryo Regionalized Cell Fate Patterning." *ELife* 7: 1–35. https://doi.org/10.7554/eLife.32839.

Morgani, Sophie M., Nestor Saiz, Vidur Garg, Dhruv Raina, Claire S. Simon, Minjung Kang, Alfonso Martinez Arias, Jennifer N. Nichols, Christian Schröter, and Anna Katerina Hadjantonakis. 2018. "A Sprouty4 Reporter to Monitor FGF/ERK Signaling Activity in ESCs and Mice." *Developmental Biology* 441 (1): 104–26. https://doi.org/10.1097/CCM.0b013e31823da96d.Hydrogen.

Moris, Naomi, Kerim Anlas, Susanne C. van den Brink, Anna Alemany, Julia Schröder, Sabitri Ghimire, Tina Balayo, Alexander van Oudenaarden, and Alfonso Martinez Arias. 2020. "An in Vitro Model of Early Anteroposterior Organization during Human Development." *Nature* 582 (7812): 410–15. https://doi.org/10.1038/s41586-020-2383-9.

Muñoz-Sanjuán, Ignacio, and Ali H. Brivanlou. 2002. "Neural Induction, the Default Model and Embryonic Stem Cells." *Nature Reviews Neuroscience* 3 (4): 271–80. https://doi.org/10.1038/nrn786.

Nemashkalo, Anastasiia, Albert Ruzo, Idse Heemskerk, and Aryeh Warmflash. 2017. "Morphogens and Community Effects Determine Cell Fates in Response to BMP4 Signaling in Human Embryonic Stem Cells. SUPPLEMENTARY." *Development* 2 (C): 1–9.

Nusse, Roel. 2012. "Wnt Signaling." *Cold Spring Harb Perspect Bio* 4 (5): 21–53. https://doi.org/10.1101/cshperspect.a011163.

Oki, Shinya, Keiko Kitajima, and Chikara Meno. 2010. "Dissecting the Role of Fgf Signaling during Gastrulation and Left-Right Axis Formation in Mouse Embryos Using Chemical Inhibitors." *Developmental Dynamics* 239 (6): 1768–78.

https://doi.org/10.1002/dvdy.22282.

Ornitz, David M., and Nobuyuki Itoh. 2015. “The Fibroblast Growth Factor Signaling Pathway.” *Wiley Interdisciplinary Reviews: Developmental Biology* 4 (3): 215–66. https://doi.org/10.1002/wdev.176.

Peng, Guangdun, Shengbao Suo, Jun Chen, Weiyang Chen, Chang Liu, Fang Yu, Ran Wang, et al. 2016. “Spatial Transcriptome for the Molecular Annotation of Lineage Fates and Cell Identity in Mid-Gastrula Mouse Embryo.” *Developmental Cell* 36 (6): 681–97. https://doi.org/10.1016/j.devcel.2016.02.020.

Pera, Edgar M., Atsushi Ikeda, Edward Eivers, and Edward M. De Robertis. 2003. “Integration of IGF, FGF, and Anti-BMP Signals via Smad1 Phosphorylation in Neural Induction.” *Genes and Development* 17 (24): 3023–28. https://doi.org/10.1101/gad.1153603.

Perea-Gomez, A., M. Rhinn, and S. L. Ang. 2001. “Role of the Anterior Visceral Endoderm in Restricting Posterior Signals in the Mouse Embryo.” *International Journal of Developmental Biology* 45 (1): 311–20.

Pijuan-sala, Blanca, Jonathan A Griffiths, Carolina Guibentif, Tom W Hiscock, Wajid Jawaid, Fernando J Calero-nieto, Carla Mulas, et al. 2019. “A Single-Cell Molecular Map of Mouse Gastrulation and Early Organogenesis.” *Nature* 566 (7745): 490–95. https://doi.org/10.1038/s41586-019-0933-9.A.

Qiao, Yunbo, Yue Zhu, Nengyin Sheng, Jun Chen, Ran Tao, Qingqing Zhu, Ting Zhang, Cheng Qian, and Naihe Jing. 2012. “AP2γ Regulates Neural and Epidermal Development Downstream of the BMP Pathway at Early Stages of Ectodermal Patterning.” *Cell Research* 22 (11): 1546–61. https://doi.org/10.1038/cr.2012.122.

Raina, Dhruv, Azra Bahadori, Angel Stanoev, Michelle Protzek, Aneta Koseska, and Christian Schröter. 2021. “Cell-Cell Communication through FGF4 Generates and Maintains Robust Proportions of Differentiated Cell Types in Embryonic Stem Cells.” *Development (Cambridge)* 148 (21). https://doi.org/10.1242/dev.199926.

Ralston, A., and Janet Rossant. 2005. “Genetic Regulation of Stem Cell Origins in the Mouse Embryo.” *Clinical Genetics* 68 (2): 106–12. https://doi.org/10.1111/j.1399-0004.2005.00478.x.

Rao, Jyoti, Martin J. Pfeiffer, Stefan Frank, Kenjiro Adachi, Ilaria Piccini, Roberto Quaranta, Marcos Arauzo-Bravo, et al. 2016. “Stepwise Clearance of Repressive Roadblocks Drives Cardiac Induction in Human ESCs.” *Cell Stem Cell* 18 (3): 341–53. https://doi.org/10.1016/j.stem.2015.11.019.

Rivera-Pérez, Jaime A., and Terry Magnuson. 2005. “Primitive Streak Formation in Mice Is

Preceded by Localized Activation of Brachyury and Wnt3." *Developmental Biology* 288 (2): 363–71. https://doi.org/10.1016/j.ydbio.2005.09.012.

Rodriguez, Tristan A., Shankar Srinivas, Melanie P. Clements, James C. Smith, and Rosa S.P. Beddington. 2005. "Induction and Migration of the Anterior Visceral Endoderm Is Regulated by the Extra-Embryonic Ectoderm." *Development* 132 (11): 2513–20. https://doi.org/10.1242/dev.01847.

Rossant, Janet, and Patrick P L Tam. 2021. "Opportunities and Challenges with Stem Cell-Based Embryo Models." *Stem Cell Reports* 16: 1–8. https://doi.org/10.1016/j.stemcr.2021.02.002.

Row, Richard H., Amy Pegg, Brian A. Kinney, Gist H. Farr, Lisa Maves, Sally Lowell, Valerie Wilson, and Benjamin L. Martin. 2018. "BMP and FGF Signaling Interact to Pattern Mesoderm by Controlling Basic Helix-Loop-Helix Transcription Factor Activity." *ELife* 7: 1–27. https://doi.org/10.7554/eLife.31018.

Schindelin, Johannes, Ignacio Arganda-carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, et al. 2012. "Fiji : An Open-Source Platform for Biological-Image Analysis" 9 (7). https://doi.org/10.1038/nmeth.2019.

Schmidt, Uwe, Martin Weigert, Coleman Broaddus, and Gene Myers. 2018. "Cell Detection with Star-Convex Polygons." *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11071 LNCS: 265–73. https://doi.org/10.1007/978-3-030-00934-2_30.

Schröter, Christian, Pau Rué, Jonathan Peter Mackenzie, and Alfonso Martinez Arias. 2015. "FGF/MAPK Signaling Sets the Switching Threshold of a Bistable Circuit Controlling Cell Fate Decisions in Embryonic Stem Cells." *Development (Cambridge, England)* 142 (24): 4205–16. https://doi.org/10.1242/dev.127530.

Schulte-Merker, S., and J. C. Smith. 1995. "Mesoderm Formation in Response to Brachyury Requires FGF Signalling." *Current Biology* 5 (1): 62–67. https://doi.org/10.1016/S0960-9822(95)00017-0.

Simon, Claire S., Anna Katerina Hadjantonakis, and Christian Schröter. 2018. "Making Lineage Decisions with Biological Noise: Lessons from the Early Mouse Embryo." *Wiley Interdisciplinary Reviews: Developmental Biology* 7 (4): 1–16. https://doi.org/10.1002/wdev.319.

Sozen, Berna, Gianluca Amadei, Andy Cox, Ran Wang, Ellen Na, Sylwia Czukiewska, Lia Chappell, et al. 2018. "Self-Assembly of Embryonic and Two Extra-Embryonic Stem Cell Types into Gastrulating Embryo-like Structures." *Nature Cell Biology* 20 (8): 979–89.

https://doi.org/10.1038/s41556-018-0147-7.

Spandidos, Athanasia, Xiaowei Wang, Huajun Wang, and Brian Seed. 2009. "PrimerBank: A Resource of Human and Mouse PCR Primer Pairs for Gene Expression Detection and Quantification." *Nucleic Acids Research* 38 (SUPPL.1): 792–99. https://doi.org/10.1093/nar/gkp1005.

Standley, H. J., A. M. Zorn, and J. B. Gurdon. 2001. "EFGF and Its Mode of Action in the Community Effect during Xenopus Myogenesis." *Development* 128 (8): 1347–57. https://doi.org/10.1242/dev.128.8.1347.

Sudheer, Smita, Jinhua Liu, Matthias Marks, Frederic Koch, Anna Anurin, Manuela Scholze, Anna Dorothea Senft, et al. 2016. "Different Concentrations of FGF Ligands, FGF2 or FGF8 Determine Distinct States of WNT-Induced Presomitic Mesoderm." *Stem Cells* 34 (7): 1790–1800. https://doi.org/10.1002/stem.2371.

Sumi, Tomoyuki, Shinya Oki, Keiko Kitajima, and Chikara Meno. 2013. "Epiblast Ground State Is Controlled by Canonical Wnt/β-Catenin Signaling in the Postimplantation Mouse Embryo and Epiblast Stem Cells." *PLoS ONE* 8 (5). https://doi.org/10.1371/journal.pone.0063378.

Sun, Xin, Erik N. Meyers, Mark Lewandoski, and Gail R. Martin. 1999. "Targeted Disruption of Fgf8 Causes Failure of Cell Migration in the Gastrulating Mouse Embryo." *Genes and Development* 13 (14): 1834–46. https://doi.org/10.1101/gad.13.14.1834.

Takada, Shinji, Kevin L. Stark, Martin J. Shea, Galya Vassileva, Jill A. McMahon, and Andrew P. McMahon. 1994. "Wnt-3a Regulates Somite and Tailbud Formation in the Mouse Embryo." *Genes and Development* 8 (2): 174–89. https://doi.org/10.1101/gad.8.2.174.

Tam, Patrick P L, and Richard R. Behringer. 1997. "Mouse Gastrulation: The Formation of a Mammalian Body Plan." *Mechanisms of Development* 68 (1–2): 3–25. https://doi.org/10.1016/S0925-4773(97)00123-8.

Tam, Patrick P L, and David a F Loebel. 2007. "Gene Function in Mouse Embryogenesis: Get Set for Gastrulation." *Nat Rev Genet* 8 (5): 368–81. https://doi.org/10.1038/nrg2084.

Technau, Ulrich, and Corinna B. Scholz. 2003. "Origin and Evolution of Endoderm and Mesoderm." *International Journal of Developmental Biology* 47 (7–8): 531–39.

Tewary, Mukul, Joel Ostblom, Laura Prochazka, Teresa Zulueta-Coarasa, Nika Shakiba, Rodrigo Fernandez-Gonzalez, and Peter W. Zandstra. 2017. "A Stepwise Model of Reaction-Diffusion and Positional Information Governs Self-Organized Human Peri-Gastrulation-like Patterning." *Development (Cambridge)* 144 (23): 4298–4312. https://doi.org/10.1242/dev.149658.

Thomas, Paul, and Rosa Beddington. 1996. “Anterior Primitive Endoderm May Be Responsible for Patterning the Anterior Neural Plate in the Mouse Embryo.” *Current Biology* 6 (11): 1487–96. https://doi.org/10.1016/S0960-9822(96)00753-1.

Traag, V. A., L. Waltman, and N. J. van Eck. 2019. “From Louvain to Leiden: Guaranteeing Well-Connected Communities.” *Scientific Reports* 9 (1). https://doi.org/10.1038/s41598-019-41695-z.

Tsakiridis, Anestis, Yali Huang, Guillaume Blin, Stavroula Skylaki, Filip Wymeersch, Rodrigo Rogrigo Osorno, Costas Economou, et al. 2014. “Distinct Wnt-Driven Primitive Streak-like Populations Reflect in Vivo Lineage Precursors.” *Development* 141 (6): 1209–21. https://doi.org/10.1242/dev.101014.

Turing, A. 1952. “The Chemical Basis of Morphogenesis.” *Philosophical Transactions of the Royal Society of London B.* 237 (641): 37–72.

Untergasser, Andreas, Ioana Cutcutache, Triinu Koressaar, Jian Ye, Brant C. Faircloth, Maido Remm, and Steven G. Rozen. 2012. “Primer3-New Capabilities and Interfaces.” *Nucleic Acids Research* 40 (15): 1–12. https://doi.org/10.1093/nar/gks596.

Vallier, Ludovic, T Touboul, Z Chng, M Brimpari, N Hannan, E Millan, L E Smithers, et al. 2009. “Early Cell Fate Decisions of Human Embryonic Stem Cells and Mouse Epiblast Stem Cells Are Controlled by the Same Signalling Pathways.” *PloS One* 4 (6): e6082. https://doi.org/10.1371/journal.pone.0006082.

Vries, Wilhelmine N. de, Alexei V. Evsikov, Bryce E. Haac, Karen S. Fancher, Andrea E. Holbrook, Rolf Kemler, Davor Solter, and Barbara B. Knowles. 2004. “Maternal β-Catenin and E-Cadherin in Mouse Development.” *Development* 131 (18): 4435–45. https://doi.org/10.1242/dev.01316.

Wagner, Kate, and David Welch. 2010. “Cryopreserving and Recovering of Human IPS Cells Using Complete Knockout Serum Replacement Feeder-Free Medium.” *Journal of Visualized Experiments : JoVE* 76 (41): 10–13. https://doi.org/10.3791/2237.

Warmflash, Aryeh, Benoit Sorre, Fred Etoc, Eric D Siggia, and Ali H Brivanlou. 2014. “A Method to Recapitulate Early Embryonic Spatial Patterning in Human Embryonic Stem Cells.” *Nature Methods* 11 (8): 847–54. https://doi.org/10.1038/nmeth.3016.

Weinberger, Leehee, Muneef Ayyash, Noa Novershtern, and Jacob H. Hanna. 2016. “Dynamic Stem Cell States: Naive to Primed Pluripotency in Rodents and Humans.” *Nature Reviews Molecular Cell Biology* 17 (3): 155–69. https://doi.org/10.1038/nrm.2015.28.

Winnier, G., M. Blessing, P. A. Labosky, and B. L.M. Hogan. 1995. “Bone Morphogenetic Protein-4 Is Required for Mesoderm Formation and Patterning in the Mouse.” *Genes and*

Development 9 (17): 2105–16. https://doi.org/10.1101/gad.9.17.2105.

Wolf, F. Alexander, Philipp Angerer, and Fabian J. Theis. 2018. "SCANPY: Large-Scale Single-Cell Gene Expression Data Analysis." *Environmental Microbiology* 19 (8): 2926–34. https://doi.org/10.1111/1462-2920.13787.

Wolpert, L. 1969. "Positional Information and the Spatial Pattern of Cellular Differentiation." *Journal of Theoretical Biology* 25 (1): 1–47.

Yamaguchi, Terry P., Allan Bradley, Andrew P. McMahon, and Steven Jones. 1999. "A Wnt5a Pathway Underlies Outgrowth of Multiple Structures in the Vertebrate Embryo." *Development* 126 (6): 1211–23. https://doi.org/10.1242/dev.126.6.1211.

Yamaguchi, Terry P., Kendraprasad Harpal, Mark Henkemeyer, and Janet Rossant. 1994. "Fgfr-1 Is Required for Embryonic Growth and Mesodermal Patterning During Mouse Gastrulation." *Genes and Development* 8 (24): 3032–44. https://doi.org/10.1101/gad.8.24.3032.

Yamaguchi, Terry P., Shinji Takada, Yoshiaki Yoshikawa, Nongying Wu, and Andrew P. McMahon. 1999. "T (Brachyury) Is a Direct Target of Wnt3a during Paraxial Mesoderm Specification." *Genes and Development* 13 (24): 3185–90. https://doi.org/10.1101/gad.13.24.3185.

Yang, Jing, and Robert A. Weinberg. 2008. "Epithelial-Mesenchymal Transition: At the Crossroads of Development and Tumor Metastasis." *Developmental Cell* 14 (6): 818–29. https://doi.org/10.1016/j.devcel.2008.05.009.

Ye, Jian, George Coulouris, Irena Zaretskaya, Ioana Cutcutache, Steve Rozen, and Thomas L. Madden. 2012. "Primer-BLAST: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction." *BMC Bioinformatics* 13: 134. https://doi.org/10.1186/1471-2105-13-134.

Yoon, Yeonsoo, Tingting Huang, Giovane G. Tortelote, Maki Wakamiya, Anna Katerina Hadjantonakis, Richard R. Behringer, and Jaime A. Rivera-Pérez. 2015. "Extra-Embryonic Wnt3 Regulates the Establishment of the Primitive Streak in Mice." *Developmental Biology* 403 (1): 80–88. https://doi.org/10.1016/j.ydbio.2015.04.008.Extra-embryonic.

Yu, Pengzhi, Guangjin Pan, Junying Yu, and James A. Thomson. 2011. "FGF2 Sustains NANOG and Switches the Outcome of BMP4-Induced Human Embryonic Stem Cell Differentiation." *Cell Stem Cell* 8 (3): 326–34. https://doi.org/10.1016/j.stem.2011.01.001.

Zhang, Hongbing, and Allan Bradley. 1996. "Mice Deficient for BMP2 Are Nonviable and

Have Defects in Amnion/Chorion and Cardiac Development." *Development* 122 (10): 2977–86. https://doi.org/10.1242/dev.122.10.2977.

Zhang, Zhechun, Steven Zwick, Ethan Loew, Joshua S. Grimley, and Sharad Ramanathan. 2019. "Mouse Embryo Geometry Drives Formation of Robust Signaling Gradients through Receptor Localization." *Nature Communications* 10 (1). https://doi.org/10.1038/s41467-019-12533-7.

Acknowledgements

I really want to thank Dr. Christian Schröter for your mentoring and support during these years. Our deep and far-ranging discussions helped me get out of my comfort zone and shape this project into a work I am proud of. I want to thank you for your encouragements that helped me when I felt stuck and I am grateful of the opportunity of working in your group.

Prof. Dr. Philippe Bastiaens, thank you for our challenging scientific discussions, for your support and for always believing in me and this project.
Dr. Peter Bieling and Dr. Aneta Koseska, thank you for the inspiring scientific and non-scientific talks and for contributing to the creation of this incredible and vibrant department.

To Christa and Lucia, you are the heart of the IMPRS program. I want to thank you for always having an open door whenever there was a question, a problem or just need for a coffee.
Thanks to Michelle Krause, for helping me with the administrative work and bureaucratic papers over these years, I would have been lost without you!

Michelle, my copilot and lab hero, thank you for your invaluable help and especially for being my right hand in the last year, but also for dealing with my lab bench and for providing me with tips and tricks in the lab.
To Dhruv, Fiorella, Azra, Max, Sina and Julia, thank you for creating a stimulating, weird and fun environment where we could discuss science, theories and techniques, but also grow lab pets, draw cows and drive me mad with octordle. I really want to thank Max, my UMAP plot buddy, for helping me with the R programming and Julia for the german translations.

To Sarah, Jana, Björn, Wayne, Marjia, Holger, Christian K., Bruno, Mai, Amit, Konstantin, Johanna and Akhilesh: thank you for welcoming me in the department, introducing me to the PhD life, help me overcome lab challenges, but also the random conversations and occasional drinks. Hans, Mike, Lena, Yannick, Ankit, Kitso and Manuel, thank you for our scientific talks, Friday evenings and our crazy lunches, you kept me sane in the writing of this thesis! Farid, Angel and Ilaria, you have been a real family here in Dortmund, I really have no words to thank you for your friendship.

To Svenja, Lisaweta, Manuela, Frau Schütz, Jutta, Kirsten, Michael R., Sabrina, Malte, Astrid and Tanja, thank you for keeping the lab running, for your help with everything science or bureaucracy related and for being pillars this department can build on.
Michael S. and Sven, thank you for invaluable help with flow cytometry machines and microscopes and for always being there for every minor or major issues of very complicated instruments.

To all the members of department II: thank you for these years and the mountains of cake and cookies!

Bhagy, Vicky, Valentina, Marion, Davide, Neha, Caitlin and many other MPI people: thank you for sharing this PhD adventure and exploring Dortmund. Ennio, thank you for getting my keys, the cooked meals and the much-needed sarcasm that lightened the occasional bad mood.

To Sofia, Silvia, Giuditta and Elisa, my maremman group, thank you for giving me a lifeline to home. You have the unique ability to ground me and at the same time make me laugh until I cry.

To Michael R., Chris M, Tanja, Felix, Sarah and Eva: thank you for welcoming me into the group, make me get in touch with the area, sometimes forcing me to speak german and being always so amazing.

To Dr. A.M, Dr. U.K., Dr. I.S. and Dr. G.K. in the city of Dortmund, Dr. S.K., Dr. H.H., Dr. V.B., Dr. V.V., Dr. M.R. in the Evangelic Hospital of Essen Mitte and to all of the doctors and nurses that helped me in a very challenging period of my life. Your thoughtfulness and professionality have been incredible and I cannot thank you enough for taking care of me.

I want to thank Moritz, for being my quiet place and my confidant, and also the source of the worst puns I have ever heard in my life.

Ai miei genitori, mio fratello e Federica: vi dedico questa tesi. Grazie per il vostro affetto e il vostro supporto in questi anni. Senza di voi non questa tesi non avrebbe visto la luce e io non ce l'avrei fatta ad arrivare dove sono ora.