Activin/Nodal signalling before implantation: setting the stage for embryo patterning

Costis Papanayotou and Jérôme Collignon

Université Paris-Diderot, Sorbonne Paris Cité, Institut Jacques Monod, CNRS UMR 7592, 75205 Paris, France

Activins and Nodal are members of the transforming growth factor beta (TGF-β) family of growth factors. Their Smad2/3-dependent signalling pathway is well known for its implication in the patterning of the embryo after implantation. Although this pathway is active early on at preimplantation stages, embryonic phenotypes for loss-of-function mutations of prominent components of the pathway are not detected before implantation. It is only fairly recently that an understanding of the role of the Activin/Nodal signalling pathway at these stages has started to emerge, notably from studies detailing how it controls the expression of target genes in embryonic stem cells. We review here what is currently known of the TGF-β-related ligands that determine the activity of Activin/Nodal signalling at preimplantation stages, and recent advances in the elucidation of the Smad2/3-dependent mechanisms underlying developmental progression.

1. Introduction

The Smad2/3-dependent Activin/Nodal signalling pathway is known to play critical roles in the specification of cell identities in embryonic and extra-embryonic lineages of the postimplantation embryo, notably during processes such as the establishment of the anterior–posterior and left–right axes [1–5]. It has been known for a while that Activin/Nodal signalling is active well before implantation, but the absence of preimplantation defects when components of the pathway are mutated has delayed our understanding of the actual functions of Activin/Nodal signalling at these early stages. This dearth of notable phenotypes may reflect (i) the pathway’s robustness, derived from partial functional redundancies between some of its components, (ii) the possible rescue of zygotic deficiencies by molecules of maternal origin, and (iii) the possibility that molecular changes brought about by the inactivation of the pathway before implantation only become detectable at later stages. A review of recent and not so recent studies, conducted both in the embryo and in cultured pluripotent stem cells, allows the relative merits of these alternatives to be assessed, and provides valuable insights into how the Activin/Nodal signalling pathway is operating at preimplantation stages.

2. The Activin/Nodal signalling pathway

Activins and Nodal are secreted as dimerized precursors, which are then cleaved to generate an active ligand [6]. Activins are homo or heterodimers of βA or βB subunits and therefore come in three versions known as Activin A, Activin B and Activin AB, collectively designated as Activin thereafter. Activin and Nodal have in common that they signal via receptor complexes containing the same type I (ALK4 or 7) and type II (ActRIIA or B) serine/threonine kinase receptors, the activation of which leads to the phosphorylation of the cytoplasmic transducers Smad2 or Smad3 (figure 1) [7]. Upon phosphorylation, dimers of Smad2/3 form a ternary complex with Smad4 that translocates to the nucleus where it associates with tissue-specific transcription factors [8,9] to activate the expression of target genes. How this is achieved has been the focus of intense research activity in recent years. These studies all emphasize the role in this process of
chromatin modifiers recruited by Smad complexes and their cofactors. These advances have been the subject of several reviews [10–12].

Nodal differs from Activin in that it requires the presence of an EGF–CFC family co-receptor (Cripto or Cryptic) to be able to activate the receptor complex. This critical difference results in Activin and Nodal signals being subjected to distinct regulatory interactions. Among the targets of Smad2/3 signalling are Nodal itself, which can therefore amplify its own expression, and the Lefty1 and Lefty2 genes, which encode Nodal antagonists thus placing Nodal expression under the control of a powerful negative feedback mechanism that will limit how long and how far it can signal [13]. Lefty1,2 are also transforming growth factor beta (TGF-β) family members. They inhibit Nodal signalling by interacting with Nodal or with Cripto. Because Activin signals in a Cripto-independent fashion, it is not sensitive to inhibition by Lefty [14]. Activin is however antagonized by another secreted molecule, called Follistatin. There is also evidence that Cripto can inhibit signalling from Activin, TGF-β and Myostatin, which have in common to signal via Smad2/3 without requiring an EGF–CFC co-receptor to interact with their receptor complex [15–18].

Other TGF-β-related ligands are known to modulate Activin/Nodal signalling. Bone morphogenetic protein (BMP) signals are transduced via the Smad1/5/8 branch of TGF-β signalling. There are now two well-characterized instances where BMP/Smad1,5,8 signalling ensures the proper developmental outcome by countering Activin/Nodal/Smad2/3 signalling via competition for a limited pool of Smad4 or for a shared receptor, ActRIIB [19,20]. Gdf1 and Gdf3, two other TGF-β family members, were found to bind the same receptor complex as Nodal, with the same requirement for an EGF–CFC co-receptor [21]. They are, however, unable to activate the Smad2/3 pathway on their own at physiological concentrations [22–25]. Instead, they seem to act as heterodimers, either to increase the range or strength of Nodal signalling when combined with Nodal, or to inhibit BMP signalling when combined with Bmp4 [22,24–26]. Further interactions between the two signalling pathways involve the inhibitory Smads, Smad6 and Smad7, which are themselves targets of Smad signalling and act either as general inhibitors of Smad-mediated signalling (Smad7), or inhibit more specifically the BMP pathway (Smad6) [27]. Smads can also integrate the input of FGF/RTK signalling, via MAPK-mediated phosphorylation, which can, for example, promote Smad1 degradation, and reduce BMP signal transduction [12,28]. Further information on how Smads integrate input from other signalling pathways can be found in other reviews [7,10,29].

3. Phenotypes of Activin/Nodal signalling pathway mutants

At first glance, genetics seems to offer little support to the notion that Activin/Nodal signalling plays an important role at preimplantation stages. Loss-of-function mutations for components of the Activin/Nodal signalling pathway (reviewed in [4,13] and summarized in table 1) result in a
Table 1. Mutant phenotypes of genes encoding components of the Activin/Nodal signalling pathway. AVE, anterior visceral endoderm; DVE, distal visceral endoderm; Epi, epiblast; ExE, extraembryonic ectoderm; LR, left–right axis; ME, mesendoderm; PS, primitive streak; VE, visceral endoderm.

<table>
<thead>
<tr>
<th>gene</th>
<th>mutant phenotype</th>
<th>references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodal</td>
<td>small size; premature Epi differentiation (neuralization); no DVE, defective ExE; no PS; no ME</td>
<td>Conlon et al. [30], with Iannaccone et al. [31], Camus et al. [32], Mesnard et al. [33], Brennan et al. [34] and Lowe et al. [35]</td>
</tr>
<tr>
<td>Activin A</td>
<td>get born, die within 24 h; craniofacial defects</td>
<td>Matzuk et al. [36]</td>
</tr>
<tr>
<td>Activin B</td>
<td>get born; maternal Activin B not essential; defective eyelid and female reproduction</td>
<td>Vassalli et al. [37]</td>
</tr>
<tr>
<td>Activin A, B</td>
<td>additive phenotype; no new defect; no functional redundancy</td>
<td>Matzuk et al. [36]</td>
</tr>
<tr>
<td>Alk4</td>
<td>resembles Nodal $^{+/+}$; small; defective from E5.5; reduced ME</td>
<td>Gu et al. [38]</td>
</tr>
<tr>
<td>Alk7</td>
<td>no phenotype even when combined with Alk4 or Nodal mutation</td>
<td>Jornvall et al. [39]</td>
</tr>
<tr>
<td>ActRIIA</td>
<td>get born; craniofacial and reproductive defects; PS defects</td>
<td>Matzuk et al. [36] and Song et al. [40]</td>
</tr>
<tr>
<td>ActRIIB</td>
<td>get born; LR defects (right isomerism)</td>
<td>Oh et al. [41]</td>
</tr>
<tr>
<td>ActRIIA, B</td>
<td>resembles Nodal $^{+/+}$; small; no PS and ME; demonstrates partial redundancy</td>
<td>Song et al. [40]</td>
</tr>
<tr>
<td>Cripto</td>
<td>form DVE but no AVE; reduced ME</td>
<td>Ding et al. [42] and Xu et al. [43]</td>
</tr>
<tr>
<td>Cryptic</td>
<td>LR defects (right isomerism and heterotaxia)</td>
<td>Gao et al. [44] and Yan et al. [45]</td>
</tr>
<tr>
<td>Cripto, Cryptic</td>
<td>resembles Nodal $^{+/+}$; defects in Epi, ExE and AVE</td>
<td>Chu et al. [46]</td>
</tr>
<tr>
<td>Gdf1</td>
<td>LR defects (resembles Cryptic $^{+/+}$)</td>
<td>Rankin et al. [47] and Andersson et al. [48]</td>
</tr>
<tr>
<td>Gdf3</td>
<td>AVE defects; ME defects</td>
<td>Chen et al. [49]</td>
</tr>
<tr>
<td>FoxH1</td>
<td>no anterior PS and its derivatives</td>
<td>Hoodless et al. [50]</td>
</tr>
<tr>
<td>Smad2</td>
<td>Epi converted into ExE; no AVE; early patterning defects</td>
<td>Waldrip et al. [51], Heyer et al. [52] and Brennan et al. [34]</td>
</tr>
<tr>
<td>Smad3</td>
<td>get born; subtle developmental anomalies</td>
<td>Zhu et al. [53], Datto et al. [54] and Yang et al. [55]</td>
</tr>
<tr>
<td>Smad2/3</td>
<td>resembles Smad4 $^{+/+}$; very small; fail to gastrulate and lack ME</td>
<td>Dunn et al. [56]</td>
</tr>
<tr>
<td>Smad4</td>
<td>very small; fail to gastrulate</td>
<td>Sirard et al. [57]</td>
</tr>
</tbody>
</table>

range of phenotypes that broadly fall in three groups: in the first, homozygotes get born, and display either craniofacial defects that lead them to an early death, or milder defects or an absence of defects that let them reach adulthood (Activin A, Activin B, Alk7, Smad3); in the second, homozygotes have primitive streak and/or left–right axis defects, which are sometimes mild enough to allow them to develop to term (Gdf1, ActRIIA, ActRIIB, Cryptic); in the third, homozygotes have early patterning defects and fail to gastrulate normally (Nodal, Gdf3, Alk4, Cripto, Smad2, Smad4). The Nodal$^{+/+}$ phenotype is the one that has been most extensively characterized and therefore provides a convenient standard for comparison with the other mutant phenotypes in that last group. In Nodal$^{+/+}$ embryos, the earliest defects appear shortly after implantation. They are smaller than littermates [30,31], and their epiblast, the pluripotent tissue that gives rise to the embryo proper, differentiates prematurely towards an anterior neural identity [32,33]. In addition, their visceral endoderm (VE), the extraembryonic layer that surrounds both the epiblast and the extraembryonic ectoderm (ExE), is improperly regionalized and fails notably to differentiate distal visceral endoderm (DVE) cells [34], which are essential for the establishment of anterior–posterior polarity. These embryos do not gastrulate [1,31]. Cripto, Cryptic double mutant embryos have a phenotype similar to that of Nodal$^{+/+}$ [46]. Consequently, although there have been suggestions that Cripto and Cryptic may have Nodal-independent functions, their compound mutant phenotype is consistent with the notion that Nodal absolutely requires them to signal. By contrast, the fact that Smad2,3 double mutant embryos [56], as well as Smad4$^{+/+}$ embryos [57], are even smaller than Nodal$^{+/+}$ embryos suggests that zygotic Nodal may not be the only ligand capable of activating the Smad2/3 pathway early on. Consistent with this view, the expression of a reporter transgene for the autoregulatory Smad2/3-dependent Nodal enhancer ASE, called ASE–YFP, was found to be maintained up to embryonic day (E)4.5 in Nodal$^{+/+}$ embryos [23]. In other animal models, there is broad evidence of another TGF-β family member acting upstream of early Nodal expression [58–62]. Vg1 in Xenopus is the prototype of a maternally deposited TGF-β-related ligand that is required to form the organizer and the mesoderm [58], and Vg1-related molecules of maternal origin identified in zebrafish and in sea-urchin appear to have similar properties [59–61]. Gdf1 and Gdf3 in the mouse are the two factors identified as Vg-1-related, however, as we saw, indications are that it is as Nodal partner and BMP antagonist that they are playing a role in the regulation of Activin/Nodal signalling and Nodal expression, not as inducers. Better candidate ligands for the early activation of the Activin/Nodal signalling pathway may thus be activins, which are present in the oviduct, uterine epithelia and blastocysts prior to implantation [63–65]. In any case, what the analysis of the earliest mutant phenotypes indicates is that the Activin/Nodal signalling pathway is not required for the formation of the epiblast.
and the VE, but that it is important to ensure their proper growth and patterning.

4. Expression dynamics of transforming growth factor beta-related ligands and associated developmental events

After fertilization, the first three rounds of divisions lead the mouse embryo to the 8-cell stage. Symmetric and asymmetric divisions then generate inner and outer cells within the 16- and 32-cell stage embryos [66]. At E3.5, during blastocyst formation, outer cells differentiate into trophoblasts to form the trophectoderm (TE), an extraembryonic tissue that encloses inner cell mass (ICM) cells and the blastocoel cavity. At E4.0, shortly before implantation, the ICM gives rise to the epiblast and to the primitive endoderm (PrE), another extraembryonic layer from which the VE will later derive. Around E4.5, the TE mediates the implantation of the blastocyst in the uterine wall.

Phosphorylated forms of Smad2/3 are detected in cell nuclei of mouse embryos from the 4-cell stage onward [67]. Studies of Activin expression in the mouse embryo showed that there is a stock of maternal Activin present in the egg that gets depleted during early cleavage stages. Zygotic expression is detected in the compacted morula (figure 2). By E3.5, this expression is confined to the ICM of the blastocyst, but at E4.5 it is not detected in the ICM-derived epiblast and PrE and is instead present in the TE [63,68]. Activin is also produced by the oviduct and the uterus at these stages [65], which may explain the lack of an early phenotype in embryos mutant for the βA and βB subunits [36,37]. An indication that this Activin of maternal origin does contribute to embryo development comes from studies showing that cleavage stage embryos developed better in vitro when cultured in the presence of Activin [69,70].

In contrast to this presence of Activin in the embryo from a very early stage, other TGF-β family members of interest all see their embryonic expression begin in the ICM of the E3.5 blastocyst, but from there they follow different dynamics. *Nodal* expression persists in the epiblast and the PrE when these tissues segregate [23,71]. It is also detected in the endometrium of pregnant females at E3.5 [72], a potential source of Nodal of maternal origin to the blastocyst. *Lefty1* expression is detected in a cluster of cells of the PrE when this layer forms [71,73]. The expression of *Bmp4* persists in the epiblast but not in the PrE when they appear, and is induced in the polar trophectoderm [74]. *Gdf3* expression strengthens in the epiblast, but is not detected in the PrE [23,24,49,71].

The onset of the expression of these genes and their subsequent diverging dynamics mark key transitions in the development of the epiblast and the PrE, events in which Activin/Nodal signalling appears to be implicated. Analysis of *Nodal*−/− embryos had revealed their failure to properly differentiate the part of the VE overlying the epiblast just after implantation and their subsequent failure to form DVE cells [33,34], thus establishing a requirement for *Nodal* in the patterning of this extraembryonic endodermal layer. Use of a conditional gene inactivation strategy showed that it is *Nodal* expressed in the epiblast that drives the regionalization of this layer [75]. However, the realization that two Nodal-dependent DVE markers, *Lefty1* and *Cerl*, had an earlier phase of expression in a small cluster of PrE cells [71,76], and that *Nodal* was expressed in the ICM and in the preimplantation epiblast [23,71], suggested that the regionalization of the VE had its origin in earlier events. Lineage-tracing studies indeed showed that descendants from *Lefty1* and *Cerl*-expressing PrE cells give rise to the DVE at E5.5, thus possibly placing the specification of this particular cell identity shortly before implantation [5,76,77]. For *Lefty1*, the enhancer driving this early PrE expression was shown to be dependent on FoxH1, a well-known effector of Activin/Nodal signalling. Likewise, ASE–YFP, the Smad2/3-dependent reporter transgene, was found to be transiently expressed in a subpopulation of PrE cells before implantation [23]. However, as previously mentioned, this expression is maintained in *Nodal*−/− embryos. So, although evidence shows that the Activin/Nodal signalling pathway is
implicated in the regionalization of the extraembryonic endodermal layer as soon as the PrE is established, these data indicate that, at this early stage, it involves a ligand other than zygotic Nodal, possibly Activin.

There are indications that something similar is taking place in the epiblast. The epiblast that first emerges from the ICM, the preimplantation epiblast, is quite different from that of the egg cylinder, the postimplantation epiblast. Preimplantation epiblast is composed of apolar cells, post-implantation epiblast becomes organized in an epithelium when the proamniotic cavity forms at E5.0. This correlates with marked changes in the properties of the tissue. Single preimplantation epiblast cells injected in a host blastocyst can contribute to all fetal lineages of the resulting chimera [78,79], whereas epiblast cells isolated at postimplantation stages cannot [80]. These differences extend to the distinct pluripotent stem cell lines derived from the epiblast at these stages: blastocyst-derived embryonic stem cells (ESCs) can contribute to all fetal lineages [81], whereas E5.5 to E6.5 embryo-derived epiblast stem cells (EpiSCs) can differentiate in multiple lineages in vitro but are unable to participate in the formation of a chimera when injected in a blastocyst [82]. Unsurprisingly, the physiology of these cells is also different. Although ESCs express Nodal and have an active Activin/Nodal signalling pathway, this is not essential to their maintenance [67,83]. In contrast, EpiSCs' capacity to self-renew depends critically on Activin/Nodal signalling [82,84]. Inhibition of the pathway in EpiSCs triggers a drastic downregulation of the pluripotency factor Nanog, which results in the cells differentiating towards a neural identity [82,85]. ESCs can be converted into EpiSCs in vitro when cultured in the presence of Activin and FGF, suggesting that the maturation of the epiblast is dependent on the same signals in vivo [86]. This differentiation has been described as a transition from a ground state of pluripotency to a primed state of pluripotency [81]. The dynamics of certain molecular markers allow us to visualize the beginning of this process in the embryo. Nascent epiblast cells maintain Nanog expression as they emerge from the ICM, but they start to downregulate it, in a salt and paper fashion, as they approach implantation. At the same time, the expression of the ASE–YFP transgene gets activated in epiblast cells with low or no Nanog, so that the two markers briefly display a somewhat complementary pattern in the epiblast of the implanting blastocyst [23]. Nanog then disappears while ASE–YFP is found in all epiblast cells. The observation that the expression of the Smad2/3-dependent ASE–YFP transgene is still present in the epiblast of implanting Nodal$^{-/-}$ embryos suggests that the transition to a primed state of pluripotency is correctly initiated in these mutants, and that it involves a TGF-β-related ligand other than zygotic Nodal. This interpretation is supported by the fact that Nodal$^{-/-}$ epiblast cells prematurely differentiate along the neural pathway, just as EpiSCs do when deprived of Activin/Nodal signalling [32,82,85].

5. Embryonic stem cells as a model to study early function and regulation of Activin/Nodal signalling

Studying what happens when ESCs differentiate into EpiSCs has recently led to important advances in our understanding of epiblast maturation and of the part Activin/Nodal signalling is playing in it. Transcriptomic and epigenomic comparisons between the two cell types has revealed that a global rearrangement of enhancer chromatin landscape is taking place, that leads to a shift in enhancer usage not just for the few genes found to be differentially expressed, but also for those that do not see a change in their expression levels [87,88]. For the latter, enhancers specifically active in ESCs, which tend to be enriched for DNA binding motifs of Smad2/3 and Smad4, are decommissioned once their EpiSC-specific enhancers are activated in differentiating ESCs [88]. A genome-wide characterization of Smad3 binding in different cell types led to the surprising finding that a small set of cell-type-specific master transcription factors direct Smad3 to cell-type-specific binding sites and determines cell-type-specific responses to TGF-β signalling [8]. Thus, in ESCs, it is with the pluripotency factor Oct4 that Smad3 co-occupies the genome. Oct4 is known to act as a pioneer factor at enhancers, opening up the chromatin to allow other factors to access their binding sites. Recent studies have shown that the capacity of ESCs to differentiate is critically dependent on the level of Oct4 not being too low [89,90], indicating the function of Oct4 involves more than allowing the establishment of pluripotency, an assessment supported by its extensive relocalization on the genome during the ESC to EpiSC transition [87].

Other evidence for the implication of Activin/Nodal signalling during this transition has emerged from particular examples of the two sets of genes described above, the ones that are differentially expressed and the ones that are similarly expressed. Gsc, which encodes the transcription factor and mesendoderm regulator Goosecoid, is a poised gene whose expression is induced once differentiation is underway [91]. Work in the Massagué laboratory showed that the stimulation of ESCs with Activin leads to the formation of companion Smad4–Smad2/3 and TRIM33–Smad2/3 complexes, which are both required to activate Gsc expression. The repressive histone mark H3K9me3 at the poised promoter of the gene is bound by TRIM33–Smad2/3, which in turn displaces the chromatin-compacting factor HP1γ to make neighbouring Smad binding elements (SBEs) accessible to the Smad4–Smad2/3 complex, presumably associated with FoxH1, and allows PolII recruitment [91]. Smad4 then promotes further binding of the TRIM33 complex via chromatin modification, ensuring robust gene expression.

Nodal is expressed in both ESCs and EpiSCs. As both a target and an inducer of Activin/Nodal signalling, it has been extensively studied and much is known about its regulation. Recent work led to the identification of a novel Nodal enhancer called HBE that is a hotspot for the binding of pluripotency factors, and to the characterization of its implication in the regulatory shift taking place at the Nodal locus. HBE is the only active Nodal enhancer in ESCs, while it is ASE that is the most active one in EpiSCs [23,92]. Both enhancers are dependent on Activin/Nodal signalling, but their activation relies on pSmad2/3 interacting with distinct transcription factors: the pluripotency factor Oct4 in the case of HBE (figure 3) [8,93], and FoxH1 in that of ASE [94]. Deletion of HBE in ESCs eliminates Nodal expression, confirming that HBE is essential to the expression of the gene in these cells. Deletion of HBE in EpiSCs does not affect Nodal expression, which is consistent with Nodal expression being dependent on ASE in these cells. However, ESCs carrying an HBE-deleted allele of Nodal fail to express this allele when induced to
differentiate into EpiSCs, revealing that the activation of ASE is dependent on HBE being present when the cells differentiate. Furthermore, HBE deletion results in the repressive histone mark H3K27me3 accumulating in the vicinity of the ASE [92]. This suggests that multi-transcription factors binding loci (MTLs, as the hotspots for the binding of pluripotency factors are called) may mediate the influence of the pluripotency gene regulatory network by determining the status of adjacent regulatory elements and the timing of their activation. The transition from ESC to EpiSC, from an HBE-driven phase to an ASE-driven phase, correlates with a decrease in the expression of master pluripotency factors known to bind HBE (such as Nanog and Klf4), and an upregulation of Nodal downstream targets [23,82,84]. The exposure of ESCs to FGF and Activin, and the resulting surge in Activin/Nodal signalling, triggers a cascade of events leading them to reach a new equilibrium, which defines their new identity. The molecular mechanisms underlying this transition, notably the role played by physical interactions between components of the Activin/Nodal signalling pathway, Smad2/3-associated transcription factors and chromatin modifiers and specific regulatory sequences, remain to be elucidated.

6. Conclusions

Current data therefore suggest that Activin/Nodal signalling, although not required for the formation of the epiblast and the PrE, plays an essential role in their development and regionalization as soon as these lineages emerge. Modulations in the activity of Activin/Nodal signalling trigger changes in gene expression via the recruitment of cell-type specific Smad2/3 complexes at cognate regulatory sequences, and the remodelling of adjacent chromatin. Studying the dynamics and the molecular bases of these events in pluripotent stem cells is bound to further our understanding of the role of Activin/Nodal signalling in the early mouse embryo.

Funding statement. Work in J.C.’s laboratory is supported by the CNRS, the ANR and the Ile-de-France Committees of Ligue Nationale contre le Cancer. C.P. was supported by a fellowship from the FRM and by the ANR.

Competing interests. The authors declare no competing interests.

References


47. Matzuk MM, Kumar TR, Bradley A. 1995 Different phenotypes for mice deficient in either activin or activin receptor type II. *Nature* 374, 356 – 360. (doi:10.1038/374355a0)

