WNT signaling memory is required for ACTIVIN to function as a morphogen in human gastruloids

Anna Yoney[1,2], Fred Etoc[1,2], Albert Ruzo[1], Jakob J. Metzger[1,2], Iain Martyn[1,2], Shu Li[1], Christoph Kirst[2], Thomas Carroll[3], Eric D. Siggia[2,*], Ali H. Brivanlou[1,*]

[1]Laboratory of Stem Cell biology and Molecular Embryology, The Rockefeller University, New York, NY 10065, USA

[2]Center for Studies in Physics and Biology, The Rockefeller University, New York, NY 10065, USA

[3]Bioinformatics Resource Center, The Rockefeller University, New York, NY 10065, USA

[*]Correspondence: siggiae@rockefeller.edu (E.D.S), brvnlou@rockefeller.edu (A.H.B.)

ABSTRACT

Self-organization of discrete fates in human gastruloids is mediated by a hierarchy of signaling pathways. How these pathways are integrated in time, and whether cells maintain a memory of their signaling history remains obscure. Here, we dissect the temporal integration of two key pathways, WNT and ACTIVIN, which along with BMP control gastrulation. CRISPR/Cas9 live reporters of SMAD1, 2 and 4 demonstrate that in contrast to the stable signaling by SMAD1, signaling and transcriptional response by SMAD2 is transient, and while necessary for pluripotency, it is insufficient for differentiation. Pre-exposure to WNT, however, endows cells with the competence to respond to graded levels of ACTIVIN, which induces differentiation without changing SMAD2 dynamics. This cellular memory of WNT signaling is necessary for ACTIVIN morphogen activity. A re-evaluation of the evidence gathered over decades in model systems, re-enforces our conclusions and points to an evolutionarily conserved mechanism.

INTRODUCTION

In the early embryo, secreted morphogens regulating a handful of signaling pathways carry the task of instructing dynamic and coordinated cell differentiation across the developing tissue. In the day 6.5 mouse embryo, a hierarchy of signaling mediated by BMP, WNT, and NODAL carries the positional information in the epiblast leading to axis formation and germ layer specification (Robertson, 2014). Each of these morphogens induces the expression of the next ligand (BMP signaling induces WNT expression, and WNT signaling induces NODAL expression), and all induce the expression of their own inhibitors. To what extent this signaling cascade is involved in human gastrulation can now be investigated using *in vitro* models of early human embryos derived from human embryonic stem cells (hESCs).

We have previously shown that the first step of this signaling hierarchy is conserved in humans and that in response to BMP4, hESCs grown in geometrically confined colonies, self-organize to induce and pattern embryonic and extra-embryonic germ layers. Ectoderm was specified at the center of circular colonies, extra-embryonic tissue at the edge and mesendoderm in between. Molecular signatures of gastrulation, such as the induction of SNAIL and activation of pERK could also be observed (Warmflash, Sorre, Etoc, Siggia, & Brivanlou, 2014). Consistent with the conservation of the second step of the signaling hierarchy, BMP4, which signals through SMAD1/5/8, was shown to induce the expression of WNT ligands, which can also induce the emergence of a primitive streak and the self-organization of embryonic germ layers in this system (Etoc et al., 2016; Martyn, Kanno, Ruzo, Siggia, & Brivanlou, 2017). We, therefore, termed this system a human gastruloid.

The third pathway, SMAD2/3, which signals on behalf of ACTIVIN and NODAL ligands, has an intriguing role. On the one hand, SMAD2/3 signaling is required for hESC pluripotency maintenance (James, Levine, Besser, & Hemmati-Brivanlou, 2005; Vallier, Alexander, & Pedersen, 2005). On the other hand, the ACTIVIN/NODAL pathway is necessary for fate specification, acting as a morphogen to pattern the blastula of

vertebrate embryos from the amphibian to the mouse (McDowell & Gurdon, 1999; Robertson, 2014). The ACTIVIN/NODAL pathway is also necessary for the self-organization of human gastruloids as inhibition of SMAD2/3 signaling blocks primitive streak formation and mesendoderm induction by BMP4 and eliminates anterior mesendoderm fates induced by WNT3A (Martyn et al., 2017; Warmflash et al., 2014). Furthermore, co-presentation of ACTIVIN with WNT3A to micropatterned hESC colonies leads to the induction of the organizer specific marker GOOSECOID, and a functional human primitive streak, which when grafted into chick embryos induces the formation of a secondary axis (Martyn et al., 2017).

Despite these findings, the mechanism by which SMAD2/3 signaling can specify both hESC pluripotency and differentiation remains perplexing, and a number of key questions remain unanswered. Among them are: how can a single signaling pathway carry these two opposite functions before and after the onset of gastrulation? To what extent do the dynamics of SMAD signaling affect these readouts? Finally, to what extent do cells have a memory of past signaling?

In this study we aim to provide highly quantitative answers to these questions by following the dynamics of TGFβ signaling during the self-organization of human gastruloids. CRISPR/Cas9-mediated tagging of SMAD1, SMAD2, and SMAD4 demonstrated that each branch of the pathway has distinct signaling dynamics. In response to ACTIVIN, SMAD2 displayed a dramatic transient nuclear translocation, which stood in sharp contrast to the stable BMP4-induced SMAD1 response. ACTIVIN stimulation did induce transient mesendodermal gene transcription, which correlated with SMAD2 dynamics. This induction, however, was not sustained and cells reverted back to pluripotency at later times. Interestingly, pre-presentation of WNT3A to the cells, while not changing SMAD2 dynamics or expression of the pluripotency markers, stabilized the transcriptional response to ACTIVIN to maintain mesendodermal fates. This implies an unexpected ability of human embryonic stem cells to record their signaling history without overt changes in fate.

RESULTS

Gastruloids respond to ACTIVIN stimulation at the colony border

A functional SMAD2/3 pathway is necessary for both maintenance of pluripotency, as well as differentiation and self-organization of human gastruloids downstream of BMP4 and WNT3A (Martyn et al., 2017; Warmflash et al., 2014). In the context of BMP4 induced differentiation, treatment with a small molecule inhibitor of SMAD2/3 signaling, SB431542 (SB), eliminated all mesendodermal fates, as indicated by the loss of BRA and SOX17 positive cells, leaving only the putative extra-embryonic and ectoderm fates, marked by CDX2 and SOX2 expression, respectively (Figure 1A) (Warmflash et al., 2014). Presentation of the downstream morphogen WNT3A induced primitive streak formation at the colony border, and addition of SB in this context removed the SOX17 positive mesendodermal population (Figure 1A) (Martyn et al., 2017). In order to ask whether gastruloid self-organization can be initiated at the ACTIVIN/NODAL point in the signaling hierarchy, we stimulated micropatterned colonies with high concentrations of ACTIVIN A (referred to as ACTIVIN throughout this work). In contrast to stimulation with BMP4 and WNT3A, no differentiation was observed after 48 hours with ACTIVIN alone (Figure 1B). Surprisingly, the lack of differentiation in response to ACTIVIN was not due to a lack of signal sensing, as an increase in nuclear SMAD2/3 was detected by immunofluorescence at the colony border after one hour of stimulation (Figures 1C and 1D). We conclude that while SMAD2/3 signaling is necessary for mesendoderm induction downstream of BMP4 and WNT3A in human gastruloids, it is not sufficient to induce it.

Two branches of the TGFβ pathway display different signaling dynamics

The inability of ACTIVIN/SMAD2 to induce differentiation stood in stark contrast with the ability of BMP/SMAD1 to induce gastruloid self-organization, including patterning of the mesoderm and endoderm layers. This contrast in activity of the two branches of the TGFβ pathway prompted us to assess possible differences in their signaling dynamics. We used CRISPR/Cas9 genome engineering on RUES2 to fluorescently tag the N-

terminus of the endogenous receptor-associated, R-SMAD1, with tagRFP (RUES2-RFP-SMAD1) and R-SMAD2 with mCitrine (RUES2-mCit-SMAD2) (Figures 2A and S1). As activation of the pathway leads to the binding of SMAD1 and 2 to the co-SMAD, SMAD4, before nuclear translocation and regulation of gene expression, a GFP-tagged SMAD4 line (RUES2-GFP-SMAD4) was also included (Nemashkalo, Ruzo, Heemskerk, & Warmflash, 2017). Each line was also transfected with ePiggyBac transposable elements carrying a nuclear marker (H2B-mCitrine or H2B-mCherry) in order to analyze the response of individual cells (Figure S2A and STAR Methods). N-terminal SMAD fusion proteins were shown to function similarly to endogenous proteins in biochemical and cell-based assays (Schmierer & Hill, 2005). Additionally, the SMAD response dynamics measured with our reporter lines, matched the behavior by of the endogenous proteins measured by immunofluorescence and western blotting as discussed below.

We began our dynamic studies of the two branches of TGFβ signaling in RUES2 cells grown on micropatterned colonies in chemically defined medium (TeSR-E7), which is derived from serum-free E8 medium but lacks any TGFβ ligands (G. Chen et al., 2011). Therefore, in E7 the exogenous TGFβ levels could be precisely controlled. In response to BMP4, we detected an increase in the SMAD1 nuclear signal that was stable over 14 hours (Figures 2B-2D, S2B, and Movies S1 and S2). The SMAD1 response was observed at the colony edge, consistent with our previous immunofluorescence results and our discovery that the TGFβ receptors are localized to the apical surface only at the border of the colony (Etoc et al., 2016; Warmflash et al., 2014). In response to ACTIVIN, SMAD2 on the other hand, responded only transiently: a pulse of nuclear translocation over the first 1-2 hours was followed by a gradual decrease over the next 6 hours (Figure 2E-2G, S2C, and Movie S3). Interestingly, the long-term SMAD2 nuclear level did not return completely to the pre-stimulus level. As in the case of SMAD1, the SMAD2 response was highest at the colony edge, again consistent with the apical localization of the receptors (Etoc et al., 2016).

Single-cell response dynamics reflect the behavior at the edge of gastruloids

In order to study signaling dynamics at the single-cell level and to eliminate modifier influences on both SMAD branches coming from neighboring cells within the micropatterned colony, we performed the same experiment on dissociated cells grown under regular culture conditions (Figure 3A). The SMAD1 response to BMP was stable and the level of nuclear RFP-SMAD1 was dependent on the BMP4 ligand concentration in agreement with previous single-cell immunofluorescence measurements (Figures 3B and 3C) (Etoc et al., 2016). The SMAD2 response to ACTIVIN was again transient with a baseline nuclear signal that remained elevated at longer times as demonstrated by our reporter line and by immunofluorescence and western blot analysis of the parental RUES2 line (Figures 3D, 3E, S3A, and S3B). The SMAD2 peak response displayed a strong sigmoidal dependence on ACTIVIN concentration (Figure S3C). However, the post-stimulation baseline, as well as the time scale of the transient response, was not dependent on the ACTIVIN dose above 0.5 ng/mL (Figures S3C and S3D). This transient SMAD2 signaling behavior did not result from depletion of ACTIVIN from the medium, as culture medium recovered from cells that were incubated with ACTIVIN for 12 hours, still induced a SMAD2 response when presented to unstimulated cells (Figure S3E). SMAD4 followed the dynamics of both R-SMADs following presentation of BMP4 or ACTIVIN (Figures 3F and 3G). Thus, in response to their activating ligands, the two branches of the TGFβ pathway display distinct dynamics of signal transduction for both the R-SMAD and the co-SMAD. This also demonstrates that modifying signals do not influence the dynamics of the response at the edge of the micropatterned colonies.

ACTIVIN elicits a transient and stable transcriptional response

We had previously shown that BMP4 signaling induces a sustained transcriptional response leading to gastruloid differentiation (Etoc et al., 2016; Warmflash et al., 2014). This is consistent with the stable nature of SMAD1 signaling presented above. The adaptive behavior of SMAD2 signaling prompted us to ask whether the short SMAD2 signaling pulse was sufficient to elicit transcriptional response and fate changes in RUES2 cells exposed to ACTIVIN. RNA-seq analysis was performed on dissociated cells cultured in E7 and E7 + ACTIVIN at 1, 2.5, 4, 8 and 12 hours following stimulation.

3,529 genes showed a change in expression level of at least two-fold during the experimental time course. They fell into three distinct groups. The first, which consisted of the majority of transcripts (2,956), peaked at 2.5 hours and declined at later time points (Figure 4A, magenta box). This group matched the timing of the transient SMAD2 response and it included key regulators of mesendodermal differentiation, such as EOMES, HHEX, GATA2, and GATA3 (Table S1) (Loh et al., 2014; Teo et al., 2011). The second group, which consisted of 452 transcripts, showed stable induction (Figure 4A, orange box). This group included genes expressed during pluripotency, such as NANOG, NODAL, LEFTY1, LEFTY2 and SMAD7 (Table S2) (Sato et al., 2003). Finally, the third group, which consisted of 121 transcripts, represented genes that were stably or transiently down regulated upon ACTIVIN presentation and included genes that are involved in signaling pathways not previously associated with pluripotency or differentiation, such as insulin signaling and cAMP response (Figure 4A, gray box and Table S3). These results suggest that cells transiently activate differentiation in response to ACTIVIN.

Examination of the signaling hierarchy involved in gastruloid self-organization revealed the presence of feedback loops at all three levels. ACTIVIN induced the expression of its own ligands and inhibitors, as well as those of the BMP and WNT pathway (Tables S1 and S2). However, despite the induction of the expression of the ligands and inhibitors, the overall threshold of signaling was not sufficient to induce and maintain mesendodermal fates from either the BMP or the WNT pathway.

As an additional approach to evaluate the transient and stable transcriptional responses following ACTIVIN treatment we performed motif enrichment analysis on our dataset. We selected 10 transcription factors that regulate primitive streak and mesendodermal differentiation: MIXL1, LEF1, BRACHYURY (BRA), GATA6, FOXH1, FOXA1, FOXA2, GOOSECOID (GSC), SOX17, and EOMES, and asked if their binding sites are enriched in the promoter region of the genes belonging to each of the dynamic groups (see STAR METHODS). In support of our hypothesis, the motifs were significantly enriched only in the transiently expressed genes of group 1 and not within group 2 or 3

(Figure 4B). This suggests that the gene regulatory network activated during the peak of SMAD2 signaling is associated with mesendodermal differentiation. Overall our data demonstrate that during the peak of SMAD2 nuclear accumulation, hESCs are *en route* for differentiation. However, the mesendodermal differentiation program is not maintained and cells return to pluripotency.

SMAD3 is dispensable for the response of the adaptive and stable gene classes to ACTIVIN

In the mouse embryo, SMAD3 is dispensable for early development as demonstrated by the fact that SMAD3 knockout mice make it to adulthood(Datto et al., 1999; Yang et al., 1999; Zhu, Richardson, Parada, & Graff, 1998). In the absence of SMAD2 in the epiblast, SMAD3 can mediate some mesoderm induction during gastrulation. However, more anterior mesendodermal lineages are completely eliminated and the embryos fail at gastrulation suggesting a critical role for SMAD2 in this process (Dunn, Vincent, Oxburgh, Robertson, & Bikoff, 2004; S. D. Vincent, Dunn, Hayashi, Norris, & Robertson, 2003). In order to decipher the relative influence of SMAD2 and SMAD3 in driving the transcriptional program downstream of ACTIVIN presentation in hESCs, we generated two independent RUES2 SMAD3 knockout lines (RUES2-SMAD3$^{-/-}$) using CRISPR-Cas9 mediated genome editing (Figures S4A and S4B). These lines maintained expression of pluripotency markers NANOG, OCT4, and SOX2, which is consistent with the previous finding that SMAD2, but not SMAD3, regulates NANOG expression to promote pluripotency in hESCs and mouse epiblast stem cells (Figure S4C) (Sakaki-Yumoto, Liu, Ramalho-Santos, Yoshida, & Derynck, 2013). In response to ACTIVIN, RUES2-SMAD3$^{-/-}$ cells displayed a transcriptional response identical to the parental RUES2 line for both pluripotency- and mesendoderm-associated ACTIVIN target genes (Figures 4C and 4D). From this we conclude that SMAD3 is not necessary for maintenance of pluripotency and that the transcriptional response dynamics following ACTIVIN presentation can be attributed to SMAD2.

Long-term, elevated SMAD2 baseline maintains pluripotency

We have shown that SMAD2 nuclear levels are transient following a step input of ACTIVIN, with a long-term baseline that remains elevated relative to the pre-stimulus level. Here we wanted to ask if the increase in the SMAD2 baseline after the peak response is required for the maintenance of pluripotency. In order to address this question, RUES2-mCit-SMAD2 cells were treated with SB in order to inhibit ACTIVIN signaling, 8 hours after stimulation, and analyzed for their ability to maintain pluripotency. SB treatment led to a decrease in SMAD2 nuclear levels back to the unstimulated baseline (Figure 5A). As observed previously, presentation of SB led to a loss of pluripotency, as indicated by the loss of NANOG expression in both RUES2-mCit-SMAD2 and the parental RUES2 line (Figures 5B, S5A, and S5B) (James et al., 2005; Vallier et al., 2005; Xu et al., 2008). We conclude that the elevated baseline at the tail of the SMAD2 response is ligand dependent and responsible for maintaining the pluripotency program long-term.

The fact that the SMAD2 post-stimulation baseline is the same regardless of ACTIVIN concentration above 0.5 ng/mL, suggests that pluripotency is insensitive to graded ligand levels above this threshold (Figures 3E and S3C). To test this hypothesis, we treated single cells with different levels of ACTIVIN and compared the expression of NANOG, OCT4, and SOX2 after 2 days of stimulation. Expression of all three markers was similar at three different concentrations of added ACTIVIN (1, 10, and 100 ng/mL) and was elevated relative to the –ACTIVIN condition (Figures 5C, 5D, and S5C). To test for zero ACTIVIN input we presented SB for the same amount of time. Pluripotency was not maintained under these conditions as demonstrated by down-regulation of NANOG and OCT4 (Figures 5C, 5D, and S5C). This confirms the insensitivity of the pluripotent state to graded ACTIVIN. Surprisingly, even at the highest ligand concentrations pluripotency is maintained and no differentiation is observed.

WNT priming unveils ACTIVIN-dependent mesendoderm differentiation

We have demonstrated that ACTIVIN alone cannot drive stable differentiation of human gastruloids. However, WNT can lead to differentiation and self-organization of

gastruloids in a SMAD2/3 dependent manner (Martyn et al., 2017). Since WNT is operating up-stream of ACTIVIN/NODAL in the proposed signaling hierarchy, we asked whether cells with a history of WNT signaling might respond differently to ACTIVIN treatment. To address this, cells were treated with WNT for 24 hours, washed to remove WNT, and then cultured with or without ACTIVIN for an additional 24 hours (Figure S6A). Surprisingly, when cells were treated with WNT alone in the absence of ACTIVIN no mesendodermal differentiation was observed and cells remained pluripotent (Figure 6A and 6B). However, if the cells were exposed to WNT before ACTIVIN stimulation mesendodermal fates were robustly induced (Figure 6C and 6D). The RUES2-SMAD3$^{-/-}$ line treated under the same conditions behaved as the parental RUES2 line, suggesting that SMAD3 is not necessary for ACTIVIN-dependent mesendoderm induction (Figure S6B). Differentiation was ACTIVIN concentration dependent, as demonstrated by the induction of BRA at low ACTIVIN and the induction of BRA, EOMES and GSC at high ACTIVIN concentrations (Figures S6C-S6E). Therefore, following WNT priming ACTIVIN functions as a morphogen to pattern mesendodermal fates and SMAD3 is again dispensable for this process.

WNT memory stabilizes the mesendodermal transcriptional response to ACTIVIN

In order to address the mechanism of WNT priming, we first asked whether the transcriptional effector of canonical WNT signaling, β-catenin, is involved in this process. We therefore treated cells with the small molecule inhibitor endo-IWR-1, which blocks β-catenin function through stabilization of its destruction complex. endo-IWR-1 was added with WNT and maintained throughout the 2-day protocol (Figure S6A). In the presence of endo-IWR-1 mesendoderm maker expression was eliminated and pluripotency marker expression was maintained, indicating a requirement for β-catenin in the differentiation process, which is consistent with previous findings (Figures 7A and S7A) (Funa et al., 2015). We then asked if endogenous WNT, induced by a possible positive feedback loop, is required for stable induction of mesendoderm, particularly during the ACTIVIN treatment phase when exogenous WNT has been removed (day 2, Figure S6A). To address this we blocked endogenous WNT secretion using the small molecule inhibitor IWP-2, which was again added with WNT and maintained throughout

the 2-day protocol. We find that addition IWP-2 does not affect mesendoderm differentiation (Figures 7A and S7A). Together, these results demonstrate the presence of an unexpected WNT signaling memory in cells that is mediated via β-catenin and is established prior to and is required for the morphogen activity of ACTIVIN.

Since we have shown that the SMAD response dynamics can be stable or transient depending on the branch being activated, we next asked if WNT memory affects the dynamics of SMAD2 signal transduction. The transient response of SMAD2 and SMAD4 and the elevation in the baseline post-stimulation was the same whether or not cells were previously exposed to WNT (Figures 7B and 7C). However, transcription of mesendodermal genes was stabilized in response to ACTIVIN following WNT priming (Figure 7D). As in the case of pluripotency, stable mesendodermal fate acquisition required on-going SMAD2 signaling in the elevated baseline as treatment with SB, 8 hours after ACTIVIN stimulation, eliminated mesendodermal differentiation at 24 h after ACTIVIN (Figure S7B). In contrast, and further arguing for a mechanism of WNT priming that is temporally up-stream of ACTIVIN, addition of SB during WNT priming did not block mesendoderm differentiation (Figure S7C). Overall, we conclude that cells maintain a memory of WNT exposure, which makes them competent to differentiate in response to ACTIVIN (Figure 7E). The molecular mechanism of this memory occurs through β-catenin-induced transcriptional changes and not as a modifier of SMAD2 signaling dynamics.

DISCUSSION

Following the tradition of experimental embryology, during the late 1960s, Nieuwkoop began a series of explant and transplant experiments in the amphibian blastula that ultimately led to the discovery of ACTIVIN as a mesoderm inducer in the 1980s (McDowell & Gurdon, 1999; Nieuwkoop, 1969). When presented to isolated animal cap explants that normally give rise only to ectoderm, ACTIVIN was sufficient to induce different types of mesendodermal cells based on its concentration, and was thus

qualified in principle as a morphogen (Green & Smith, 1990; Wilson & Melton, 1994). When the ACTIVIN receptors and the SMAD pathways were characterized in 1990s, it was shown that micro-injection of different amounts of synthetic mRNAs encoding SMAD2 into *Xenopus* animal caps, also recapitulates the mesendodermal-inducing effects of A ACTIVIN morphogen presentation, independently confirming that the activation of different thresholds of the pathway was sufficient for mesoderm induction (Shimizu & Gurdon, 1999). Inhibition of the pathway by microinjection of a dominant negative type II ACTIVIN receptor into *Xenopus* blastula led to complete loss of mesendodermal derivatives, demonstrating that ACTIVIN /SMAD2 signaling was necessary for mesoderm induction in the amphibian embryo (Hemmati-Brivanlou & Melton, 1992).

Loss-of-function analysis in the mouse later confirmed the frog conclusions about the pathway as SMAD2/3 double knockout mice failed to properly induce mesendoderm and gastrulate (Dunn et al., 2004). However another TGFβ ligand, NODAL, rather than ACTIVIN was shown to be the inducer in the mammalian embryo (Conlon et al., 1994). Elegant genetic experiments also performed in the mouse placed NODAL signaling in a positive feedback loop that drives gastrulation. The loop is initiated when NODAL signaling induces BMP4 expression in the extra-embryonic ectoderm. BMP signaling subsequently induces WNT3 expression, which in turn induces high levels of NODAL in the proximal-posterior part of the embryo, which marks the site of primitive streak initiation (Arnold & Robertson, 2009).

Using live reporters of SMAD1, SMAD2 and SMAD4, we find that BMP/SMAD1 signaling is stable, ACTIVIN/SMAD2 signaling is transient, and that co-SMAD4 follows the dynamics of the receptor-associated SMADs. In response to ACTIVIN, hESCs transiently induce mesendodermal gene expression, without generating mesendodermal fates, and the cells return to the state of pluripotency. Pre-presentation of WNT, however, stabilizes transcription and fate acquisition without modifying SMAD2 dynamics. Both pluripotency maintenance and mesendodermal fate acquisition are not affected by the loss of SMAD3. Our data provide evidence for a previously undetected

level of signal integration that implies the presence of a cellular memory. This memory operates not at the level of the SMAD2 signaling dynamics but by some other means to change the response of the cells to ACTIVIN. It is tempting to speculate that it might be mediated through epigenetic mechanisms, in a manner similar to what has been described in *Xenopus* experiments where β-catenin recruits PrMT2 to induce dorsal fates (Blythe, Cha, Tadjuidje, Heasman, & Klein, 2010).

The ACTIVIN/SMAD2 pathway has been shown to be necessary for maintenance of pluripotency, but not sufficient to induce mesendoderm in hESCs, except when influenced by modifiers of other signaling pathways (D'Amour et al., 2005; Funa et al., 2015; McLean et al., 2007; Singh et al., 2012). Our study shows that mere pre-treatment of hESCs with WNT, endows them with a memory that enables a graded response to subsequently applied ACTIVIN; simultaneous exposure to WNT and ACTIVIN is not required. Furthermore, the WNT treatment alone is not sufficient to elicit a stable fate change since in minimal media, without ACTIVIN, the cells revert to pluripotency. The ability of embryonic cells to record WNT signals may be a broadly conserved and fundamental aspect of animal development, as shown by elegant experiments in *Drosophilia* where the authors demonstrate a morphogen effect for Dpp (a BMP homologue) after cells have lost contact with the source of WNT (Alexandre, Baena-Lopez, & Vincent, 2013).

In the context of vertebrate development, our findings force the reevaluation of the traditional literature regarding the sufficiency of the ACTIVIN/SMAD2 pathway for mesendodermal induction and patterning. A review of the evidence in model systems for the ACTIVIN /NODAL-mediated morphogen effect reveals that in all the experimental settings, at least some of the cells were either still under the influence, or had been previously exposed to WNT signaling before or during ACTIVIN/NODAL signaling. For example, cells of the *Xenopus* animal cap derived from the blastula stage embryo have been under maternal WNT influence as early as the two-cell stage, hours before ACTIVIN is presented to the explants. This influence is the consequence of cortical rotation that occurs after sperm entry, and activates the WNT pathway on the

dorsal side of the embryo, as evidenced by the dorsal-specific nuclear localization of β-catenin (Larabell et al., 1997; Rowning et al., 1997; Schneider, Steinbeisser, Warga, & Hausen, 1996). It is clear from the asymmetric elongation of the animal cap explants in response to ACTIVIN that the prior WNT exposure sets up a dorsal-ventral pre-pattern that affects the response. It is also clear that the dorsal and ventral regions of the animal cap when separated respond differently to ACTIVIN (Bolce, Hemmati-Brivanlou, Kushner, & Harland, 1992; Sokol & Melton, 1991). Although some mesoderm is induced in the ventral caps, only cells of the dorsal cap that have seen WNT signals undergo ACTIVIN-mediated induction of GSC+ mesendoderm. These experiments demonstrate that both pathways are required for complete mesendodermal patterning without explicitly distinguishing their temporal relationship.

Our study revises this point of view by bringing a temporal order. We suggest that only cells endowed with a WNT memory can respond to ACTIVIN to specify the full range of mesendodermal fates. Even in the extensively studied *Xenopus* embryo, it is still debated whether in the late blastula stage marginal zone, prior to the onset of gastrulation, an ACTIVIN/NODAL gradient already defines the medial to lateral mesendodermal fates or whether there is simply a bipartite division into dorsal organizer and ventral mesoderm (Harland & Gerhart, 1997; Smith, 2009). This debate also translates into whether ACTIVIN/NODAL is sufficient to pre-pattern the mesendoderm or whether complete patterning requires prior WNT exposure present only in the dorsal part of the embryo. Controversies persist, since the embryo is rapidly developing, there are no live reporters of signaling, and fates are often assayed in early gastrulation rather than in the late blastula. In the mouse the earliest manifestation of the streak is the proximal-posterior expression of WNT, which extends distally. By mid-streak stage, NODAL signaling is highest in the Node (Tam & Loebel, 2007). Thus, cells that leave the streak at various proximal-distal positions and NODAL levels, manifestly have been exposed to WNT first (Tam & Behringer, 1997). This interpretation, however, does not eliminate the co-requirement for ACTIVIN and WNT, but rather suggests that instructive WNT, required for mesendodermal differentiation, occurs temporally up-stream of ACTIVIN/SMAD2 signaling.

In conclusion, a cellular memory of WNT signaling, endows embryonic cells, including hESCs, with the competence to maintain ACTIVIN/SMAD2-mediated mesendodermal fate specification and patterning.

ACKNOWLEDGEMENTS

We are thankful to all the members of the Brivanlou and Siggia groups for helpful discussions and to C. Zierhut for critical reading of the manuscript. Our research was supported by NIH grants R01 HD080699 and R01 GM101653 to A.H.B. and E.D.S and NSF grant DGE 132526 to A.Y. Imaging was performed at The Rockefeller University Bio-Imaging Resource Center.

AUTHOR CONTRIBUTIONS

Conceptualization and writing, A.Y., F.E., A.H.B., and E.D.S; Investigation, A.Y., F.E., I.M., and S.L.; Formal analysis, A.Y., A.R., J.J.M, and T.C.; Software, C.K. and J.J.M.; Resources, A.R. All authors reviewed the manuscript.

MATERIAL AND METHODS

Human embryonic stem cell culture

Experiments were performed with the RUES2 hESC line (XX female; US National Institutes of Health, human ESC registry no. 0013) or CRISPR/Cas9 edited cell lines based on RUES2. hESCs were grown in HUESM medium that was conditioned by mouse embryonic fibroblasts and supplemented with 20 ng/mL bFGF (MEF-CM). Cells were grown on tissue culture dishes (BD Biosciences, San Jose, CA) for maintenance and expansion at 37 °C and 5% CO_2. Dishes were coated overnight at 4 °C with Geltrex (Thermo Fisher Scientific, Waltham, MA) diluted 1:40 in DMEM/F12 and then incubated at 37 °C for at least 20 minutes before passaging. Cells were passaged as aggregates

using Gentle Cell Dissociation Reagent (STEMCELL Technologies, Vancouver, Canada). Cells were tested for mycoplasma infection before beginning the experiments.

Micropatterned cell culture

Individual coverslips (CYTOO, Grenoble, France) were washed one time with water for 5 minutes at room temperature (RT) to activate the surface according to the manufacturers recommendation. Coverslips were then coated at 37 °C for 2 hours with 20 µg/mL Laminin-521 (BioLamina, Sundbyberg, Sweden) in 0.5 mL PBS +Mg/+Ca. The laminin was then removed by serial dilutions without allowing the coverslip to dry (1:4 dilution in PBS –Mg/–Ca, six times). Chips were seeded immediately or stored overnight at 4 °C in 2 mL PBS –/– and seeded on the following day.

Cell seeding was performed as follows. Cells growing in MEF-CM were washed once with PBS +/+ and dissociated to single cells with Accutase (STEMCELL Technologies). Cells were centrifuged and 600,000 cells were resuspended in 2 mL of MEF-CM supplemented with 10 µM Rock-inhibitor (Y-27632, Abcam, Cambridge, MA). The cell suspension was then placed over the coverslip in a 35-mm tissue culture dish. The sample was left unperturbed for 10 minutes at RT in order to achieve homogeneous seeding of the cells throughout the chip before being moved to the incubator. After 2 hours, the medium was replaced with MEF-CM without Rock-inhibitor and the cells were incubated overnight. On the following day, the medium was changed to MEF-CM with additional ligands, BMP4, WNT3A, or ACTIVIN A (R&D Systems, Minneapolis, MN) with or without 10 µM SB431542 (Stemgent, Lexington, MA). For experiments in TeSR-E7 (STEMCELL Technologies), cells were moved to E7 medium the day after seeding and incubated for an additional 24 hours before adding ligands. Live imaging was carried out in E7 imaging medium, which was prepared with FluoroBrite DMEM (Thermo Fisher Scientific) according the published protocol for E8 (Beers et al., 2012).

Single-cell culture

Optical-quality plastic tissue culture dishes (ibidi, Martinsried, Germany) were coated with 10 µg/mL Laminin-521 in PBS +/+ at 37 °C for 2 hours or overnight at 4 °C. Single

cells were collected as described for micropatterned cell culture and dishes were seeded with 50,000 single cells resuspended in 2 mL E7 supplemented with 10 μM Rock-inhibitor. The samples were incubated overnight. On the following day, the medium was changed to E7 supplemented with 10 μM Rock-inhibitor and ligands or small molecules. For the 2-day protocol in which cells were switched from E7 +/– 100 ng/mL WNT3A to E7 +/– 10 ng/mL ACTIVIN A, the samples were washed with PBS +/+ before adding fresh medium. Live imaging was carried out in E7 imaging medium. Small molecules were used at the following concentrations: 10 μM SB431542 (Stemgent), 1 μM IWP-2 (Stemgent), and 1 μM endo-IWR-1 (Tocris).

Immunofluorescence and western blotting

Cells on dishes or coverslips were rinsed once with PBS –/–, fixed with 4% paraformaldehyde (Alfa Aesar, Thermo Fisher Scientific, Tewksbury, MA) for 20 minutes at RT, and then rinsed twice and stored in PBS –/–. Cells were blocked and permeabilized with blocking buffer (2% bovine serum albumin and 0.1% Triton X-100 in PBS –/–) for 30 minutes at RT. Cells were incubated with primary antibodies in blocking buffer overnight at 4 °C and then washed three times with 0.1% Tween-20 in PBS –/– (PBST). Cells were incubated with secondary antibodies (diluted 1:1000): Alexa Fluor 488, 555, or 647-conjugated from (Invitrogen Molecular Probes, Thermo Fisher Scientific) and DAPI nuclear stain (1:10,000 dilution) in blocking buffer for 30 minutes at RT, and then washed twice with PBST and once with PBS –/–. Dishes were stored and imaged in PBS –/–. Coverslips were mounted on slides using Fluoromount-G mounting medium (SouthernBiotech, Birmingham, AL).

Western blotting was performed using standard techniques. The nuclear fractions presented in Figure S3 were obtained using the NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific) and following the manufacturers protocol. Following SDS-PAGE and Western transfer, membranes were stained with Smad2 (86F7) Rabbit mAb (Cell Signaling Technology, Danvers, MA) at 1:1000 dilution. Membranes were stripped for 15 minutes at RT using Restore PLUS Western Blot

Stripping Buffer (ThermoFisher Scientific) and re-stained with Lamin B1 Rabbit polyAb (Proteintech, Rosemont, IL) at 1:2000 dilution.

Imaging

Wide-field images of fixed samples were acquired on an Olympus IX-70 inverted microscope with a 10x/0.4 numerical aperture objective lens. Tiled image acquisition was used to acquire images of large areas of dishes or coverslips in four channels corresponding to DAPI and Alexa Fluor 488, 555, and 647. Live imaging was performed on a spinning disk confocal microscope equipped with 405-, 488-, and 561-nm lasers and an environmental chamber (CellVoyager CV1000, Yokogawa). Images were acquired every 10 minutes with a 20x/0.75 numerical aperture objective lens. The cells were maintained at 37 °C and 5% CO_2 during live imaging.

Generation of SMAD1 and SMAD2 reporter cell lines

For the tagRFP-SMAD1 reporter cells, CRISPR/Cas9 mediated genome engineering was used to fuse a cassette containing a blasticidin resistance gene (BsdR), a T2A self-cleaving peptide, and a tagRFP fluorescent protein onto the N-terminus of SMAD1, so that the locus produces both a tagRFP-SMAD1 fusion protein together with BsdR. Similarly, for the mCitrine-SMAD2 reporter line, CRISPR/Cas9 was used to fuse a cassette containing a puromycin resistance gene (PuroR), a T2A self-cleaving peptide, and an mCitrine fluorescent protein onto the N-terminus of SMAD2, so that the locus produces both an mCitrine-SMAD2 fusion protein together with PuroR.

RUES2 hESCs were nucleofected with a pX335 plasmid (Cong et al., 2013) that co-expresses the nickase version of Cas9 and the specific sgRNA of interest (protospacer sequences: 5'-GCAGCACTAGTTATACTCCT-3' for SMAD1 and 5'-GGACGACATGTTCTTACCAA-3' for SMAD2), as well as the appropriate homology donor plasmid. Nucleofection was carried out using the Cell Line Nucelofector Kit L (Lonza, Walkersville, MD) and the B-016 setting of a Nucleofector II instrument. Nucleofected cells were plated into MEF-CM supplemented with 10 μM Rock-inhibitor. After 4 days, blasticidin or puromycin was added for 7 days to select for cells that had

been targeted. Cells that survived selection were passaged as single cells using Accutase, plated in MEF-CM supplemented with 10 µM Rock-inhibitor, and allowed to grow into colonies. Colonies arising from a single cell were handpicked, expanded, and screened for correct targeting by PCR amplification of the genomic region and Sanger sequencing. Correctly targeted clones were subsequently transfected with ePiggyBac plasmids containing either H2B-mCitrine or H2B-mCherry cassettes to enable nuclear labeling for cell tracking (Lacoste, Berenshteyn, & Brivanlou, 2009). Individual clones were again isolated and controlled for normal karyotype (G-banding) and pluripotency maintenance.

Generation of SMAD3 knockout cell lines

RUES2 hESCs were nucleofected with a pX330 plasmid (Cong et al., 2013) that co-expressed the wild-type version of Cas9 and one of two sgRNAs targeting SMAD3 (protospacer sequences: 5'-CCACCAGATGAACCACAGCA-3' for sgRNA #1 and 5'-TTATTATGTGCTGGGGACAT-3' for sgRNA #2). The sgRNAs were designed to target the first or second coding exon shared by all SMAD3 isoforms, a strategy that was successfully used to knockout SMAD3 in human primary cell lines (Voets et al., 2017). Two different sgRNAs were used to control for off-target effects. We modified the pX330 plasmid to also express a puromycin-2A-EGFP cassette to enrich for cells that had been successfully nucleofected. Nucleofection was carried out as described above and cells were plated in MEF-CM supplemented with 10 µM Rock-inhibitor. On the following day puromycin was added for 24 hours. Cells that survived selection were allowed to recover for several days. Cells were then passaged as single cells using Accutase, plated in MEF-CM supplemented with 10 µM Rock-inhibitor, and allowed to grow into colonies. Colonies arising from a single cell were handpicked, expanded, and screened for correct targeting by PCR amplification of the genomic region and Sanger sequencing. The resulting chromatograms were decomposed using the Tide web-based tool (Brinkman, Chen, Amendola, & van Steensel, 2014).

RT-PCR

RUES2 cells were seeded in 6-well plates (100,000 single cells per well) in E7 medium supplemented with 10 μM Rock-inhibitor and incubated overnight. On the following day the medium was changed to E7 supplemented with 10 μM Rock-inhibitor with or without 10 ng/mL ACTIVIN A. Samples (3 pooled-wells) were collected in 1 mL Trizol at 0, 1, 2.5, 4 and 8 hours. The 2-day protocol was carried out similarly: the medium was changed on the day after seeding to E7 supplemented with 10 μM Rock-inhibitor with or without 100 ng/mL WNT3A. On the second day following seeding the wells were washed once with PBS and the media was changed to E7 supplemented with 10 μM Rock-inhibitor with or without 10 ng/mL ACTIVIN A. Samples were collected in 1 mL Trizol before the addition of WNT3A, after WNT3A treatment (referred to as 0 h), and at 2.5, 4, 6, and 12 h after the addition of ACTIVIN A. Total cellular RNA for each sample was extracted using the RNeasy Mini Kit (QIAGEN, Germantown, MD) and cDNA was synthesized using the Transcriptor First Strand cDNA Synthesis Kit (Roche). RT-PCR for selected genes was performed using the LightCycler 480 SYBR Green I Master mix in a LightCycler 480 instrument (Roche, Basel, Switzerland). Primers were designed using Primer-BLAST (Ye et al., 2012) or obtained from qPrimerDepot (Cui, Taub, & Gardner, 2007) or from previously published sequences (Mendjan et al., 2014). Primer sequences and source are listed in Table S4.

Total RNA-sequencing

RUES2 cells were seeded in 6-well plates (200,000 single cells per well) in E7 medium supplemented with 10 μM Rock-inhibitor and incubated overnight. On the following day the media was changed to E7 supplemented with 10 μM Rock-inhibitor with or without 10 ng/mL ACTIVIN A. Samples (3 pooled-wells) were collected in 1 mL Trizol at 1, 2.5, 4, 8 and 12 hours for the ACTIVIN-treated conditions and after 0, 6 and 12 hours for the no-ACTIVIN conditions (to be used as negative controls). Total cellular RNA for each sample was extracted using RNeasy Mini Kit (QIAGEN) and 2ug of total RNA was used to prepare each individual RNA-seq library. RNA-seq library construction was conducted with the TruSeq RNA Library Preparation Kit (Illumina, San Diego, CA) as per the manufacturer's instructions and sequenced in an Illumina HiSeq 2500

apparatus. Raw reads were mapped to hg19 using STAR aligner, and the gene read counts were normalized using the DESeq2 Bioconductor package (Love, Huber, & Anders, 2014). Library preparation, sequencing, and mapping were performed by the New York Genome Center (New York, NY, USA). All raw data files are available from the GEO database (accession number GSE111717).

Image analysis

For images acquired from micropatterned cell culture experiments, stitching and colony detection were carried out as described previously using custom software written in MATLAB (Etoc et al., 2016). For analysis of the SMAD reporter lines on micropatterned colonies a single z-plane through the middle of the colony was analyzed at each time point. Background in each channel was removed by subtracting a minimum intensity image that was generated by taking the minimum value at each pixel over all images in that channel in the same z-plane. Vignetting was corrected by dividing by a flat-field image that was generated by normalizing the intensity of the minimum image to 1. This procedure corrects the intensity drop-off at the border of the images without further altering the average image intensity.

Nuclei segmentation and signal quantification were performed on the corrected images as follows. The H2B image was thresholded to generate a binary image separating the foreground (nuclei) from the background. The original, corrected H2B image was then filtered with a median and sphere filter with parameters matching the expected size of individual nuclei. Local maxima corresponding to individual nuclei were detected using the MATLAB extended-maxima transform function. Maxima were dilated to increase the likelihood of obtaining a single maximum per nuclei. Maxima falling within the foreground were used as seeds for watershed segmentation, which was also restricted to the foreground and was used to obtain a labeled object corresponding to each nucleus within the image. The segmented objects were further processed to eliminate objects much larger or smaller than the expected size of individual nuclei. The results of the segmentation were used as a mask to obtain the median per cell nuclear intensity in each channel.

For the mCitrine-SMAD2 reporter line an enrichment of the mCitrine signal in the cytoplasm prior to ligand presentation could be detected, which decreased concomitantly with an increase in the nuclear signal following ACTIVIN presentation. Therefore, the mCitrine-SMAD2 response was quantified as the nuclear-to-cytoplasmic ratio, which was used previously as a readout for TGFβ pathway activity (Warmflash et al., 2012). In order to estimate the cytoplasmic signal for each cell, a narrow donut surrounding the nuclear mask was formed by dilating the nuclear mask once by an inner radius and a second time by an outer radius and subtracting the first dilated object from the second. The donut, which formed the cytoplasmic mask, was not restricted to the foreground pixels (H2B signal), but it was prevented from overlapping with the masks of neighboring nuclei. The median mCitrine intensity within the nuclear mask was divided by the median intensity within the cytoplasmic mask on a per cell basis. The qualitative behavior of the SMAD2 nuclear-to-cytoplasmic response was not sensitive to the exact size of the inner and outer radius of the donut. Therefore, the values were chosen manually and kept fixed throughout all analyses. For the RFP-SMAD1 reporter line a faint nuclear signal could be detected prior to ligand presentation. However, no cytoplasmic signal could be detected above the background. Therefore, the RFP-SMAD1 response was quantified as the median nuclear signal normalized to the median H2B signal to normalize for cells moving in and out of the z-plane. In order to analyze the SMAD response as a function of radial position within the micropatterned colonies, the cells within a single colony were binned based on the radial position of their center and the average response per cell within each bin was calculated. The radial profile for individual colonies was then averaged over several colonies. Analysis of the SMAD reporter lines in single-cell culture was carried out in a similar manner. The maximum intensity projection image, rather than a single z-plane was analyzed. Since cells remained flat under these conditions it was not necessary to normalize the SMAD1 signal by H2B.

Images acquired from fixed samples that were stained by immunofluorescence were similarly segmented and analyzed. The DAPI channel was used to perform nuclei

segmentation, and the segmented objects were subsequently used to obtain the median per cell signal intensity in each channel. Only samples stained on the same day and imaged together were compared.

RNAseq analysis and clustering

For each gene, a baseline expression profile, which was calculated using a linear interpolation between the 0, 6 and 12 h control samples, was subtracted from the expression values of the ACTIVIN-treated samples. The gene list was filtered to contain only those genes that: 1) had at least one time point with an absolute fold-change larger than 2 (up- or down-regulated) compared to 0 h and 2) had at least one time point with a normalized read count higher than 100. That generated a list of 3,529 genes of interest, which was then hierarchically clustered by their z-scored expression values, using Cluster 3.0 (de Hoon, Imoto, Nolan, & Miyano, 2004) with the following options: centered correlation as the similarity metric and average linkage as clustering method. The resulting hierarchical tree was visualized using Java TreeView (Saldanha, 2004) to identify the minimal clusters of interest.

Motif analysis

To identify motifs enriched within gene clusters, the 2000 bp upstream sequences for all genes were extracted using the PWMEnrich Bioconductor package. Motif enrichment of each cluster's sequence set was performed using AME (Bailey, Johnson, Grant, & Noble, 2015) with the HOCOMOCOv10 database (Kulakovskiy et al., 2016) against the background (upstream sequences of all genes).

DATA AND SOFTWARE AVAILABILITY

The RNA-sequencing data reported in Figure 4 is available is available under the GEO accession number GSE111717.

REFERENCES

Alexandre, C., Baena-Lopez, A., & Vincent, J.-P. (2013). Patterning and growth control by membrane-tethered Wingless. *Nature*, *505*(7482), 180–185. http://doi.org/10.1038/nature12879

Arnold, S. J., & Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nature Reviews Molecular Cell Biology*, *10*(2), 91–103. http://doi.org/10.1038/nrm2618

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, *43*(W1), W39–W49. http://doi.org/10.1093/nar/gkv416

Beers, J., Gulbranson, D. R., George, N., Siniscalchi, L. I., Jones, J., Thomson, J. A., & Chen, G. (2012). Passaging and colony expansion of human pluripotent stem cells by enzyme-free dissociation in chemically defined culture conditions. *Nature Protocols*, *7*(11), 2029–2040. http://doi.org/10.1038/nprot.2012.130

Blythe, S. A., Cha, S.-W., Tadjuidje, E., Heasman, J., & Klein, P. S. (2010). β-Catenin Primes Organizer Gene Expression by Recruiting a Histone H3 Arginine 8 Methyltransferase, Prmt2. *Developmental Cell*, *19*(2), 220–231. http://doi.org/10.1016/j.devcel.2010.07.007

Bolce, M. E., Hemmati-Brivanlou, A., Kushner, P. D., & Harland, R. M. (1992). Ventral ectoderm of Xenopus forms neural tissue, including hindbrain, in response to activin. *Development*, *115*(3), 681–688.

Brinkman, E. K., Chen, T., Amendola, M., & van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Research*, *42*(22), e168–e168. http://doi.org/10.1093/nar/gku936

Chen, G., Gulbranson, D. R., Hou, Z., Bolin, J. M., Ruotti, V., Probasco, M. D., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nature Methods*, *8*(5), 424–429. http://doi.org/10.1038/nmeth.1593

Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science*, *339*(6121), 819–823. http://doi.org/10.1126/science.1231143

Conlon, F. L., Lyons, K. M., Takaesu, N., Barth, K. S., Kispert, A., Herrmann, B., &

Robertson, E. J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development*, *120*(7), 1919–1928.

Cui, W., Taub, D. D., & Gardner, K. (2007). qPrimerDepot: a primer database for quantitative real time PCR. *Nucleic Acids Research*, *35*(Database), D805–D809. http://doi.org/10.1093/nar/gkl767

D'Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E., & Baetge, E. E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nature Biotechnology*, *23*(12), 1534–1541. http://doi.org/10.1038/nbt1163

Datto, M. B., Frederick, J. P., Pan, L., Borton, A. J., Zhuang, Y., & Wang, X.-F. (1999). Targeted Disruption of Smad3 Reveals an Essential Role in Transforming Growth Factor β-Mediated Signal Transduction. *Molecular and Cellular Biology*, *19*(4), 2495–2504. http://doi.org/10.1128/MCB.19.4.2495

de Hoon, M. J. L., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. *Bioinformatics*, *20*(9), 1453–1454. http://doi.org/10.1093/bioinformatics/bth078

Dunn, N. R., Vincent, S. D., Oxburgh, L., Robertson, E. J., & Bikoff, E. K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development*, *131*(8), 1717–1728. http://doi.org/10.1242/dev.01072

Etoc, F., Metzger, J., Ruzo, A., Kirst, C., Yoney, A., Ozair, M. Z., et al. (2016). A Balance between Secreted Inhibitors and Edge Sensing Controls Gastruloid Self-Organization. *Developmental Cell*, *39*(3), 302–315. http://doi.org/10.1016/j.devcel.2016.09.016

Funa, N. S., Schachter, K. A., Lerdrup, M., Ekberg, J., Hess, K., Dietrich, N., et al. (2015). β-Catenin Regulates Primitive Streak Induction through Collaborative Interactions with SMAD2/SMAD3 and OCT4. *Cell Stem Cell*, *16*(6), 639–652. http://doi.org/10.1016/j.stem.2015.03.008

Green, J. B. A., & Smith, J. C. (1990). Graded changes in dose of a Xenopus activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature*, *347*(6291), 391–394. http://doi.org/10.1038/347391a0

Harland, R., & Gerhart, J. (1997). FORMATION AND FUNCTION OF SPEMANN'S ORGANIZER. *Annual Review of Cell and Developmental Biology*, *13*(1), 611–667. http://doi.org/10.1146/annurev.cellbio.13.1.611

Hemmati-Brivanlou, A., & Melton, D. A. (1992). A truncated activin receptor inhibits mesoderm induction and formation of axial structures in Xenopus embryos. *Nature*, *359*(6396), 609–614. http://doi.org/10.1038/359609a0

James, D., Levine, A. J., Besser, D., & Hemmati-Brivanlou, A. (2005). TGFbeta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development*, *132*(6), 1273–1282. http://doi.org/10.1242/dev.01706

Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Soboleva, A. V., Kasianov, A. S., Ashoor, H., et al. (2016). HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models. *Nucleic Acids Research*, *44*(D1), D116–25. http://doi.org/10.1093/nar/gkv1249

Lacoste, A., Berenshteyn, F., & Brivanlou, A. H. (2009). An Efficient and Reversible Transposable System for Gene Delivery and Lineage-Specific Differentiation in Human Embryonic Stem Cells. *Cell Stem Cell*, *5*(3), 332–342. http://doi.org/10.1016/j.stem.2009.07.011

Larabell, C. A., Torres, M., Rowning, B. A., Yost, C., Miller, J. R., Wu, M., et al. (1997). Establishment of the Dorso-ventral Axis in Xenopus Embryos Is Presaged by Early Asymmetries in β-Catenin That Are Modulated by the Wnt Signaling Pathway. *The Journal of Cell Biology*, *136*(5), 1123–1136. http://doi.org/10.1083/jcb.136.5.1123

Loh, K. M., Ang, L. T., Zhang, J., Kumar, V., Ang, J., Auyeong, J. Q., et al. (2014). Efficient endoderm induction from human pluripotent stem cells by logically directing signals controlling lineage bifurcations. *Cell Stem Cell*, *14*(2), 237–252. http://doi.org/10.1016/j.stem.2013.12.007

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 31. http://doi.org/10.1186/s13059-014-0550-8

Martyn, I., Kanno, T., Ruzo, A., Siggia, E. D., & Brivanlou, A. (2017). Self-organization of a functional human organizer by combined WNT and NODAL signalling. *bioRxiv*,

234633. http://doi.org/10.1101/234633

McDowell, N., & Gurdon, J. B. (1999). Activin as a morphogen in Xenopus mesoderm induction. *Seminars in Cell & Developmental Biology*, *10*(3), 311–317. http://doi.org/10.1006/scdb.1999.0307

McLean, A. B., D'Amour, K. A., Jones, K. L., Krishnamoorthy, M., Kulik, M. J., Reynolds, D. M., et al. (2007). Activin a efficiently specifies definitive endoderm from human embryonic stem cells only when phosphatidylinositol 3-kinase signaling is suppressed. *Stem Cells*, *25*(1), 29–38. http://doi.org/10.1634/stemcells.2006-0219

Mendjan, S., Mascetti, V. L., Ortmann, D., Ortiz, M., Karjosukarso, D. W., Ng, Y., et al. (2014). NANOG and CDX2 pattern distinct subtypes of human mesoderm during exit from pluripotency. *Cell Stem Cell*, *15*(3), 310–325. http://doi.org/10.1016/j.stem.2014.06.006

Nemashkalo, A., Ruzo, A., Heemskerk, I., & Warmflash, A. (2017). Morphogen and community effects determine cell fates in response to BMP4 signaling in human embryonic stem cells. *Development*, *144*(17), 3042–3053. http://doi.org/10.1242/dev.153239

Nieuwkoop, P. D. (1969). The Formation of the Mesoderm in Urodelean Amphibians. *Wilhelm Roux Archiv Für Entwicklungsmechanik Der Organismen*, *163*(4), 298–315. http://doi.org/10.1007/BF00577017

Robertson, E. J. (2014). Dose-dependent Nodal/Smad signals pattern the early mouse embryo. *Seminars in Cell & Developmental Biology*, *32*, 73–79. http://doi.org/10.1016/j.semcdb.2014.03.028

Rowning, B. A., Wells, J., Wu, M., Gerhart, J. C., Moon, R. T., & Larabell, C. A. (1997). Microtubule-mediated transport of organelles and localization of -catenin to the future dorsal side of Xenopus eggs. *Proceedings of the National Academy of Sciences*, *94*(4), 1224–1229. http://doi.org/10.1073/pnas.94.4.1224

Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N., & Derynck, R. (2013). Smad2 is essential for maintenance of the human and mouse primed pluripotent stem cell state. *The Journal of Biological Chemistry*, *288*(25), 18546–18560. http://doi.org/10.1074/jbc.M112.446591

Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data.

Bioinformatics, *20*(17), 3246–3248. http://doi.org/10.1093/bioinformatics/bth349

Sato, N., Sanjuan, I. M., Heke, M., Uchida, M., Naef, F., & Brivanlou, A. H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. *Developmental Biology*, *260*(2), 404–413.

Schmierer, B., & Hill, C. S. (2005). Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor beta-dependent nuclear accumulation of Smads. *Molecular and Cellular Biology*, *25*(22), 9845–9858. http://doi.org/10.1128/MCB.25.22.9845-9858.2005

Schneider, S., Steinbeisser, H., Warga, R. M., & Hausen, P. (1996). β-catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mechanisms of Development*, *57*(2), 191–198. http://doi.org/10.1016/0925-4773(96)00546-1

Shimizu, K., & Gurdon, J. B. (1999). A quantitative analysis of signal transduction from activin receptor to nucleus and its relevance to morphogen gradient interpretation. *Proceedings of the National Academy of Sciences*, *96*(12), 6791–6796. http://doi.org/10.1073/pnas.96.12.6791

Singh, A. M., Reynolds, D., Cliff, T., Ohtsuka, S., Mattheyses, A. L., Sun, Y., et al. (2012). Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell*, *10*(3), 312–326. http://doi.org/10.1016/j.stem.2012.01.014

Smith, J. C. (2009). Forming and Interpreting Gradients in the Early Xenopus Embryo. *Cold Spring Harbor Perspectives in Biology*, *1*(1), a002477–a002477. http://doi.org/10.1101/cshperspect.a002477

Sokol, S., & Melton, D. A. (1991). Pre-existent pattern in Xenopus animal pole cells revealed by induction with activin. *Nature*, *351*(6325), 409–411. http://doi.org/10.1038/351409a0

Tam, P. P. L., & Behringer, R. R. (1997). Mouse gastrulation: the formation of a mammalian body plan. *Mechanisms of Development*, *68*(1-2), 3–25. http://doi.org/10.1016/S0925-4773(97)00123-8

Tam, P. P. L., & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nature Reviews Genetics*, *8*(5), 368–381.

http://doi.org/10.1038/nrg2084

Teo, A. K. K., Arnold, S. J., Trotter, M. W. B., Brown, S., Ang, L. T., Chng, Z., et al. (2011). Pluripotency factors regulate definitive endoderm specification through eomesodermin. *Genes \& Development*, *25*(3), 238–250. http://doi.org/10.1101/gad.607311

Vallier, L., Alexander, M., & Pedersen, R. A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *Journal of Cell Science*, *118*(19), 4495–4509. http://doi.org/10.1242/jcs.02553

Vincent, S. D., Dunn, N. R., Hayashi, S., Norris, D. P., & Robertson, E. J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes \& Development*, *17*(13), 1646–1662. http://doi.org/10.1101/gad.1100503

Voets, O., Tielen, F., Elstak, E., Benschop, J., Grimbergen, M., Stallen, J., et al. (2017). Highly efficient gene inactivation by adenoviral CRISPR/Cas9 in human primary cells. *PLoS One*, *12*(8), e0182974. http://doi.org/10.1371/journal.pone.0182974

Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D., & Brivanlou, A. H. (2014). A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nature Methods*, *11*(8), 847–854. http://doi.org/10.1038/nmeth.3016

Warmflash, A., Zhang, Q., Sorre, B., Vonica, A., Siggia, E. D., & Brivanlou, A. H. (2012). Dynamics of TGF- signaling reveal adaptive and pulsatile behaviors reflected in the nuclear localization of transcription factor Smad4. *Proceedings of the National Academy of Sciences*, *109*(28), E1947–E1956. http://doi.org/10.1073/pnas.1207607109

Wilson, P. A., & Melton, D. A. (1994). Mesodermal patterning by an inducer gradient depends on secondary cell–cell communication. *Current Biology*, *4*(8), 676–686. http://doi.org/10.1016/S0960-9822(00)00152-4

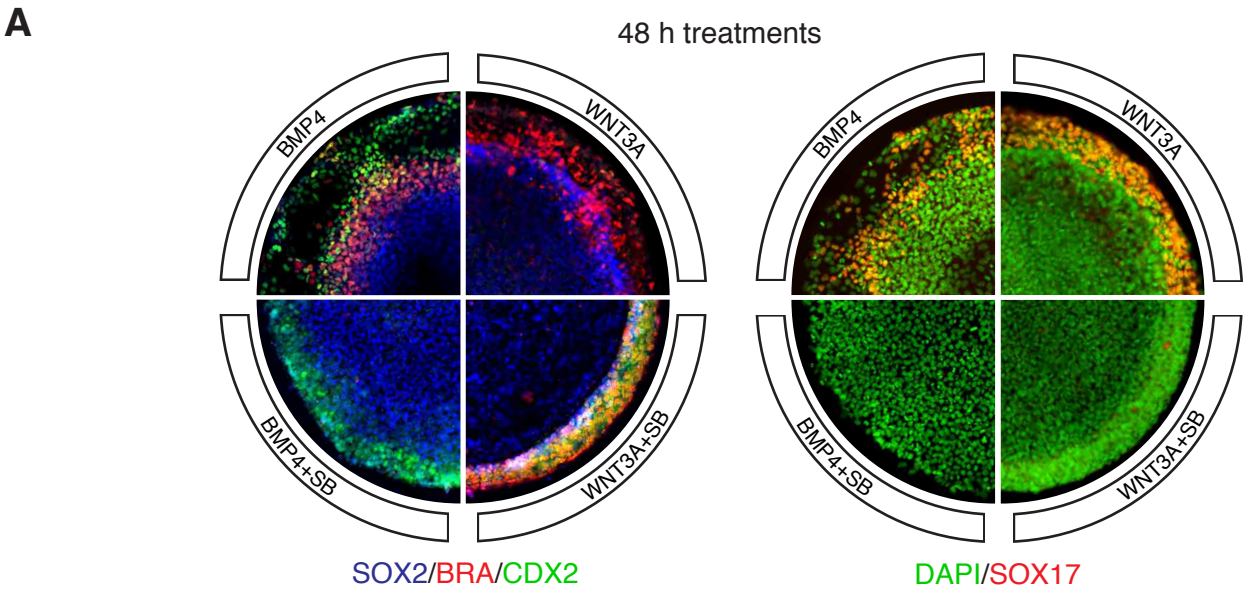
Xu, R.-H., Sampsell-Barron, T. L., Gu, F., Root, S., Peck, R. M., Pan, G., et al. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell*, *3*(2), 196–206. http://doi.org/10.1016/j.stem.2008.07.001

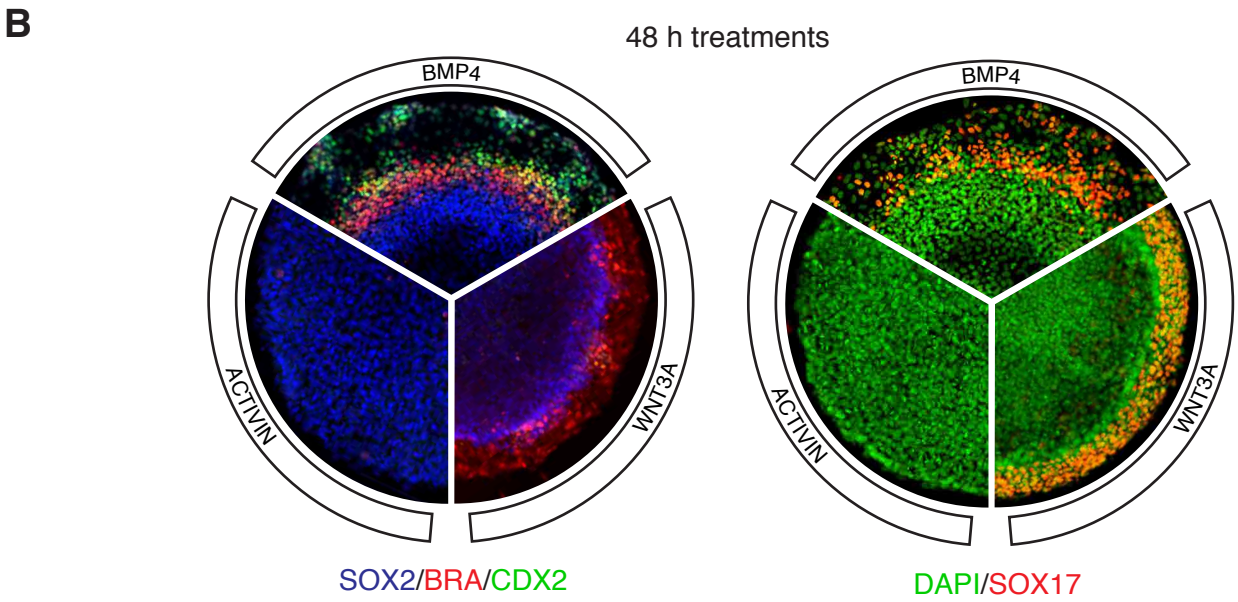
Yang, X., Letterio, J. J., Lechleider, R. J., Chen, L., Hayman, R., Gu, H., et al. (1999). Targeted disruption of SMAD3 results in impaired mucosal immunity and diminished T cell responsiveness to TGF-β. *The EMBO Journal*, *18*(5), 1280–1291.

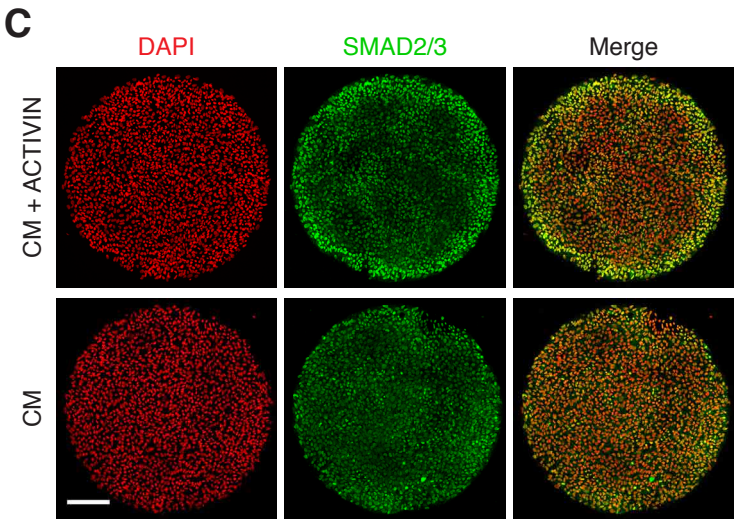
http://doi.org/10.1093/emboj/18.5.1280

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*(1), 134. http://doi.org/10.1186/1471-2105-13-134

Zhu, Y., Richardson, J. A., Parada, L. F., & Graff, J. M. (1998). Smad3 Mutant Mice Develop Metastatic Colorectal Cancer. *Cell*, *94*(6), 703–714. http://doi.org/10.1016/S0092-8674(00)81730-4







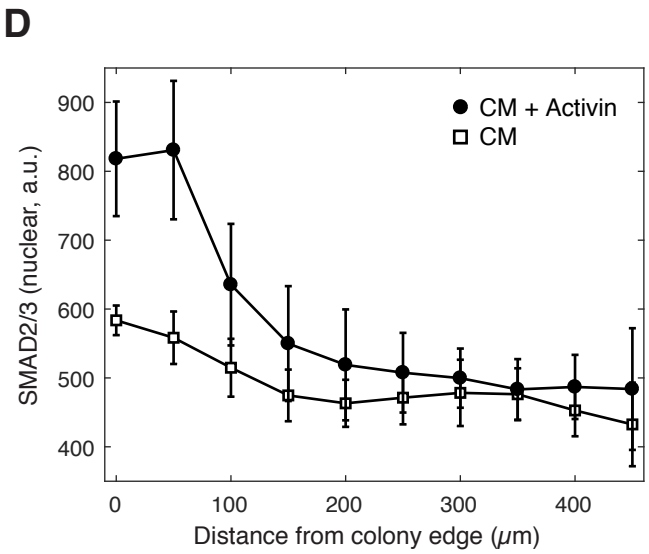


Figure 1

Figure 1. ACTIVIN modifies gastruloid differentiation but it cannot induce it

(A) Micropatterned colonies in conditioned media (CM) treated for with BMP4 (50 ng/mL), BMP4 (50 ng/mL) + SB (10 μM), WNT3A (100 ng/mL), or WNT3A (100 ng/mL) + SB (10 μM) for 48 hours. The colonies were fixed and analyzed by immunofluorescence. Left: CDX2 (green), BRA (red), SOX2 (blue). Right: SOX17 (red), DAPI (green).

(B) Micropatterned colonies in conditioned media (CM) treated with BMP4 (50 ng/mL), WNT3A (100 ng/mL), or ACTIVIN (100 ng/mL) for 48 hours. The colonies were fixed and analyzed by immunofluorescence. Left: CDX2 (green), BRA (red), SOX2 (blue). Right: SOX17 (red), DAPI (green).

(C) Micropatterned colonies in conditioned media (CM) with or without treatment with ACTIVIN (100 ng/mL) for 1 hour. The colonies were fixed and analyzed by immunofluorescence: DAPI (red), SMAD2/3 (green). Scale bar represents 200 μm.

(D) Quantification of the average SMAD2/3 nuclear fluorescence as a function of radial position from the colony edge for the same experiment shown in (C): CM + ACTIVIN (close circles), CM (open squares). Error bars represent the standard deviation across n = 5 (CM) and n = 6 (CM + ACTIVIN) colonies.

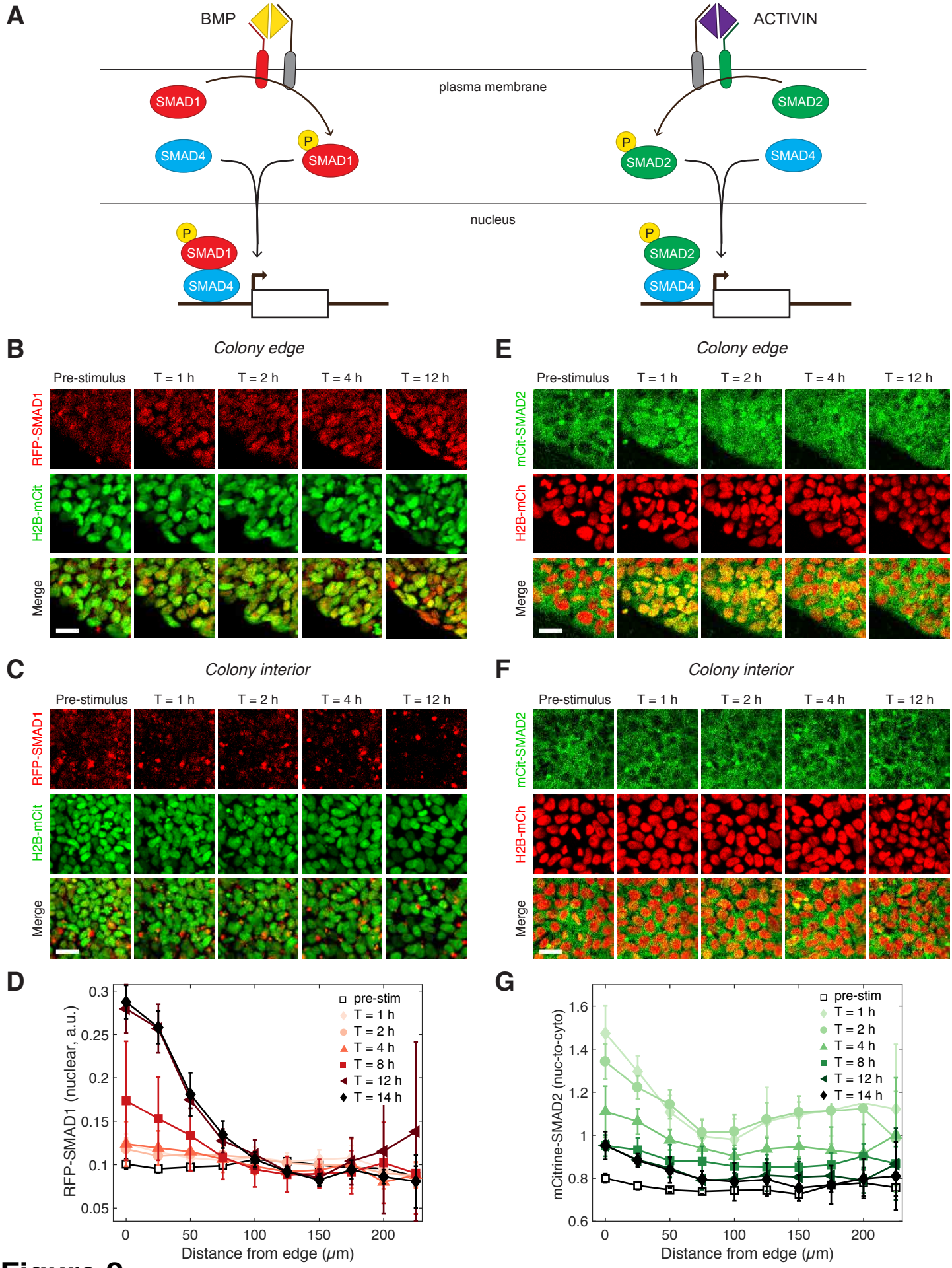


Figure 2

Figure 2. Two branches of the TGFβ pathway display different signaling dynamics

(A) BMP and ACTIVIN signaling represent the two branches of the TGFβ pathway. BMP signals through R-SMAD1 and ACTIVIN signals through R-SMAD2. R-SMADs form a complex with co-SMAD4, which regulates target gene expression.

(B-C) RUES2-RFP-SMAD1 grown on micropatterned colonies in E7 and stimulated with BMP4 (50 ng/mL). The response of cells at the colony edge (B) and near the colony center (C) as a function of time following BMP4 presentation. The intensity range was adjusted to the same minimum and maximum values in all images in both (B) and (C). Scale bars represent 25 μm. See also Movies S1 and S2.

(D) Average RFP-SMAD1 nuclear signal as a function of radial position within the colony at different time points following BMP4 treatment. The single-cell nuclear RFP-SMAD1 intensity was quantified and normalized to the single-cell H2B-mCitrine signal. Error bars represent the standard deviation over n = 5 colonies.

(E-F) RUES2-mCit-SMAD2 grown on micropatterned colonies in E7 and stimulated with ACTIVIN (10 ng/mL). The response of cells at the colony edge (E) and near the colony center (F) as a function of time following ACTIVIN presentation. The intensity range was adjusted to the same minimum and maximum values in all images in both (E) and (F). Scale bars represent 25 μm. See also Movie S3.

(G) Average mCitrine-SMAD2 nuclear-to-cytoplasmic signal as a function of radial position within the colony at different time points following ACTIVIN treatment. Since a cytoplasmic mCitrine signal could be detected in the pre-stimulus condition, the single-cell nuclear mCitrine intensity was normalized to the single-cell cytoplasmic mCitrine signal. Error bars represent the standard deviation over n = 5 colonies.

See also Figures S1 and S2.

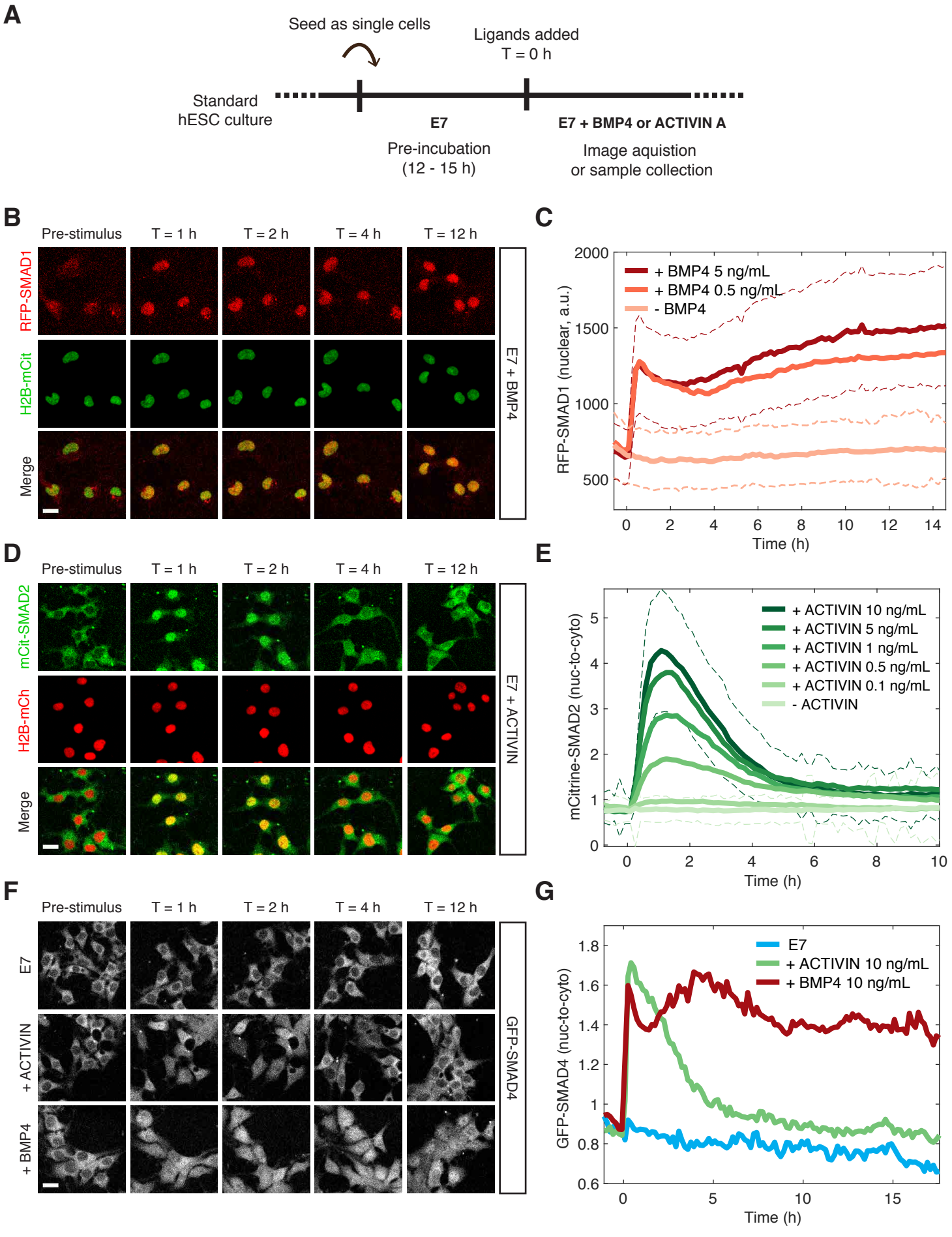


Figure 3

Figure 3. BMP and ACTIVIN elicit different single-cell SMAD response dynamics

(A) Schematic outlining the single-cell experimental protocol.

(B) RFP-SMAD1 response in single cells to BMP4 (5 ng/mL) added at T = 0 h.

(C) Quantification of the RFP-SMAD1 nuclear signal as a function of time and BMP4 concentration. Images were acquired every 10 minutes. Solid lines represent the average response at each time point (n > 200 cells per time point). Dashed lines represent the population standard deviation for – BMP4 (light red) and + BMP4 (5 ng/mL, dark red) conditions.

(D) mCitrine-SMAD2 response in single cells to ACTIVIN (10 ng/mL) added at T = 0 h.

(E) Quantification of the mCitrine-SMAD2 nuclear-to-cytoplasmic ratio as a function of time and ACTIVIN concentration. Images were acquired every 10 minutes. Solid lines represent the average response at each time point (n > 200 cells per time point). Dashed lines represent the population standard deviation for – ACTIVIN (lightest green) and + ACTIVIN (10 ng/mL, darkest green) conditions.

(F) GFP-SMAD4 response in single cells to ACTIVIN (10 ng/mL) or BMP4 (10 ng/mL) or cells that were left untreated (E7). Ligands were added at T = 0 h.

(G) Quantification of the GFP-SMAD4 nuclear-to-cytoplasmic ratio as a function of time in E7 (blue), E7 + BMP4 (10 ng/mL, red), or E7 + ACTIVIN (10 ng/mL, green). Images were acquired every 10 minutes. Solid lines represent the average response at each time point (n > 200 cells per time point).

Scale bars represent 25 μm. See also Figure S3.

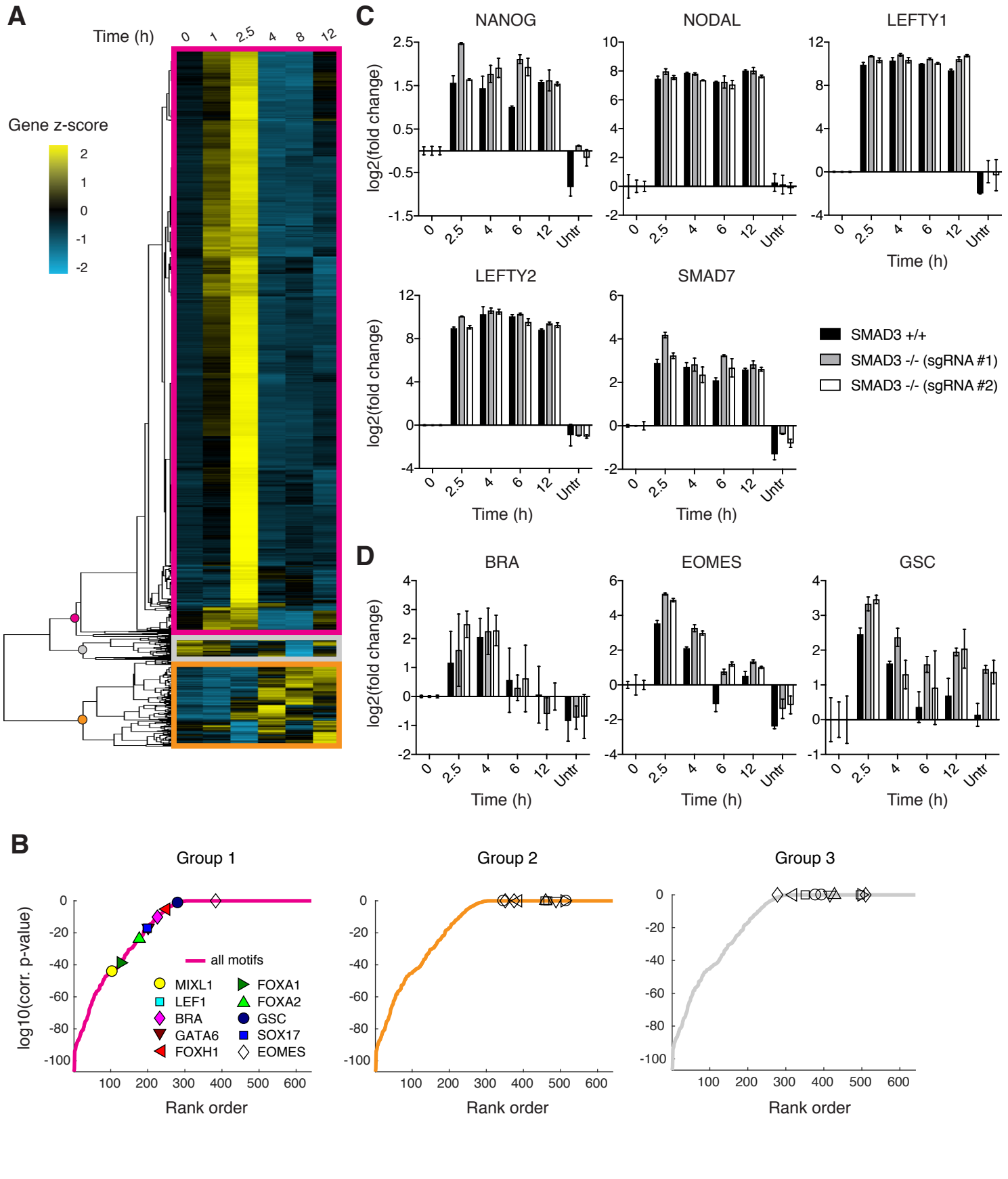


Figure 4

Figure 4. ACTIVIN/SMAD2 signaling results in stable and transient transcriptional responses

(A) Hierarchical clustering of RNA-seq time course data for genes showing a fold-change > 2 in response to ACTIVIN (10 ng/mL) relative to E7 (T = 0 h). The gene z-score at each time point was calculated by subtracting the average and dividing by the standard deviation of the normalized read counts across all time points. See also Tables S1, S2, and S3.

(B) Motif analysis of the gene groups obtained from hierarchical clustering shown in (A). The solid lines represent all motifs analyzed and ordered by the p-value of their enrichment corrected for multiple tests. Symbols indicate select motifs that are associated with primitive streak and mesendodermal differentiation: corrected p-value < 10^{-1} (filled symbols) and corrected p-value > 10^{-1} (empty symbols).

(C) RT-PCR analysis of pluripotency-associated SMAD2 target genes following presentation of ACTIVIN (10 ng/mL) in the parental RUES2 line (black bars) and the RUES2-SMAD3$^{-/-}$ lines (gray and white bars). An additional sample was collected that was left untreated for the 12 h time course (untr). Expression in each sample was normalized to GAPDH and then to the pre-stimulus level (T = 0 h).

(D) RT-PCR analysis of mesendoderm-associated SMAD2 target genes following presentation of ACTIVIN (10 ng/mL) in the parental RUES2 line (black bars) and the RUES2-SMAD3$^{-/-}$ lines (gray and white bars). An additional sample was collected that was left untreated for the 12 h time course (untr). Expression in each sample was normalized to GAPDH and then to the pre-stimulus level (T = 0 h).

See also Figure S4.

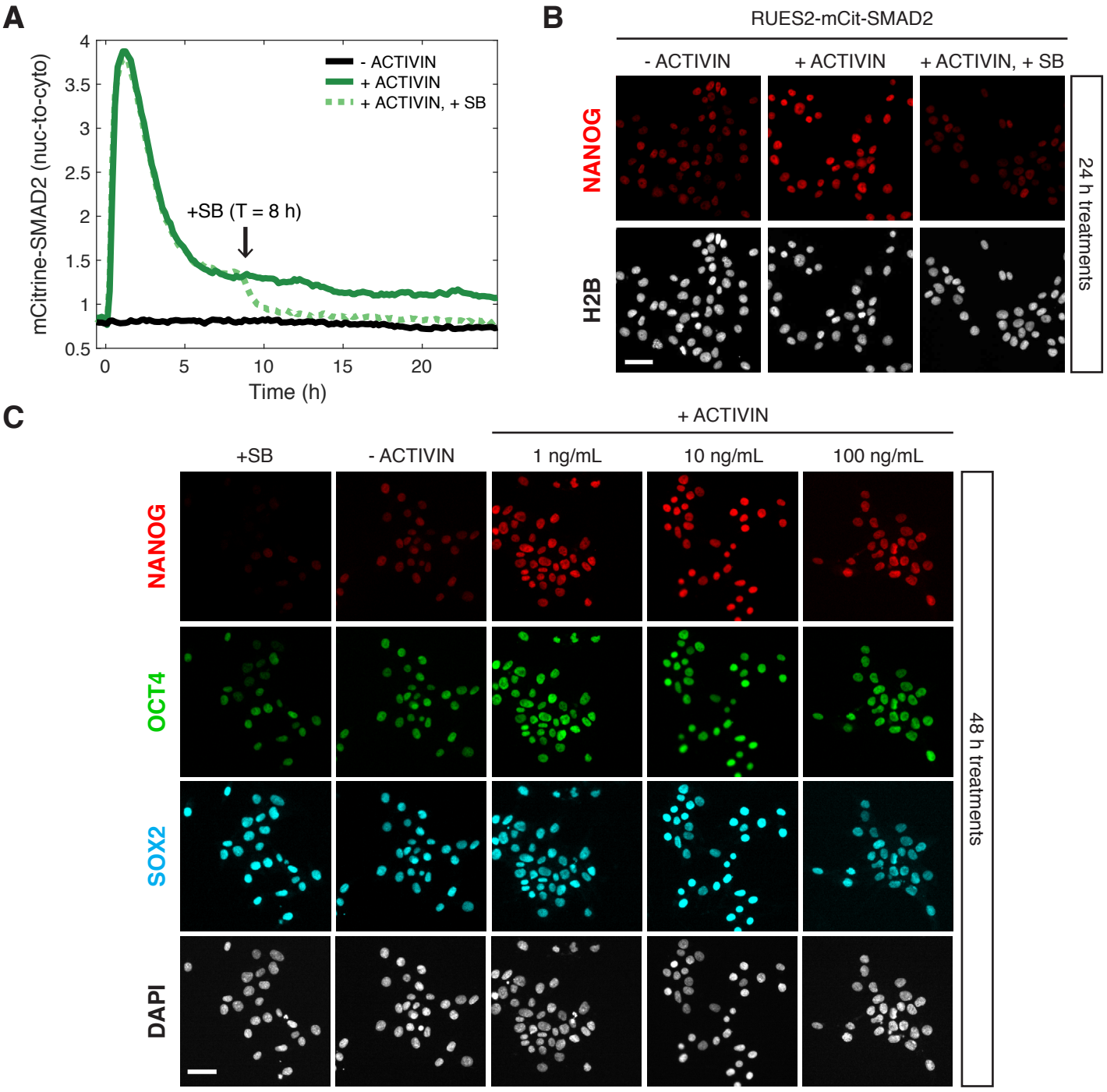


Figure 5

Figure 5. Long-term, elevated SMAD2 baseline maintains pluripotency

(A) Quantification of the mCitrine-SMAD2 nuclear-to-cytoplasmic ratio following treatment with ACTIVIN (10 ng/mL added at T = 0 h, solid green line). Following the transient SMAD2 response, SB (10 μM) was added to one of the samples (dotted green line, added at T = 8 h). A third sample was left untreated in E7 (-ACTIVIN) for the duration of the experiment (solid black line). Images were acquired every 10 minutes. Lines represent the average response at each time point (n > 200 cells per time point).

(B) The samples in (A) were fixed 24 h after ACTIVIN addition and analyzed for NANOG expression by immunofluorescence. See Figure S5A for quantification.

(C) RUES2 cells were cultured in E7 with different levels of Activin (0, 1, 10, 100 ng/mL) or SB (10 μM) for 2 days. Cells were fixed and analyzed by immunofluorescence (IF). Images: DAPI (gray), NANOG (red), OCT4 (green), SOX2 (cyan).

(D) Histograms showing the nuclear IF signal quantified in single cells corresponding to the experiment shown in D (n > 5,000 cells per condition).

Scale bars represent 50 μM. See also Figure S5.

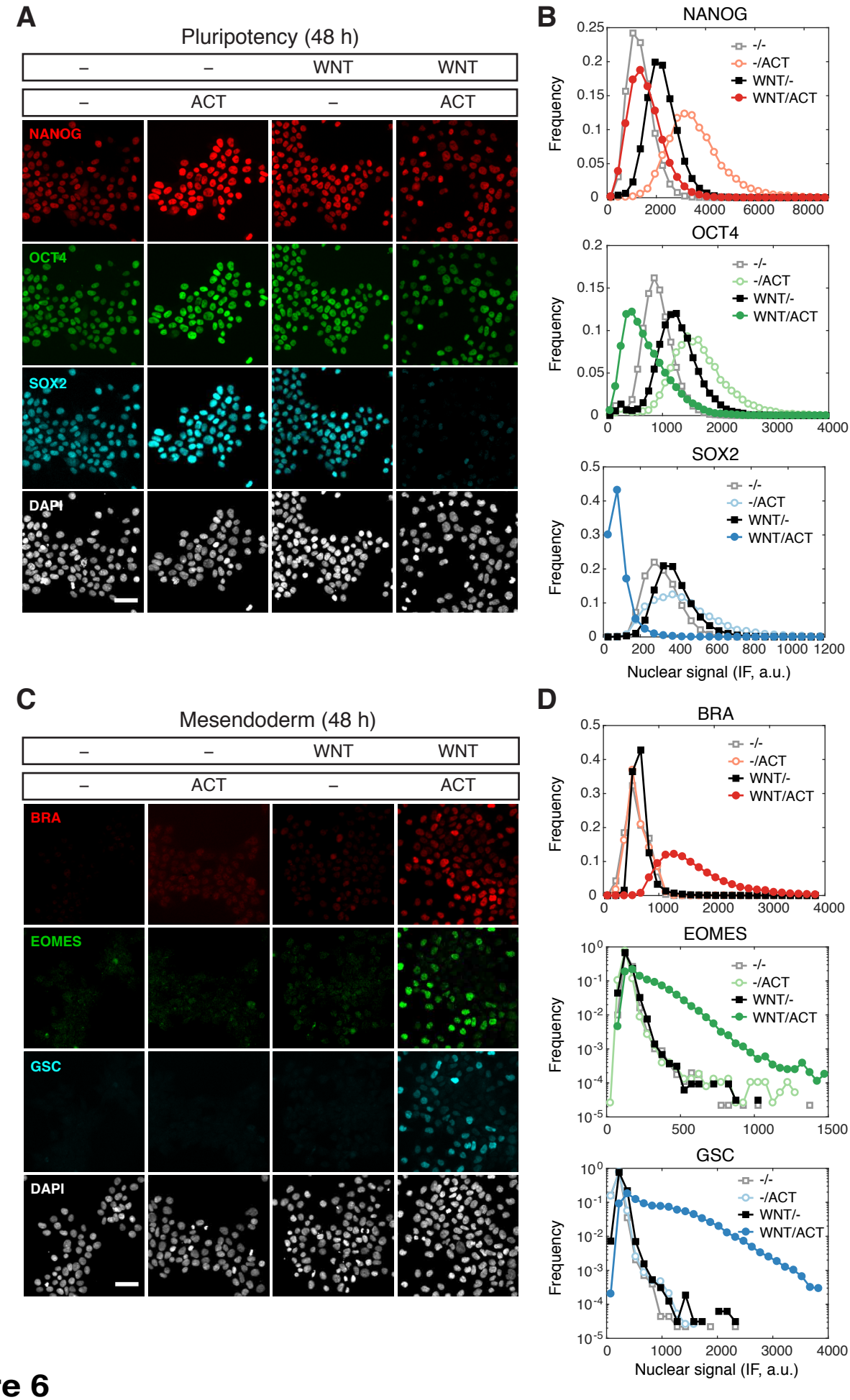


Figure 6

Figure 6. WNT priming unveils ACTIVIN-dependent mesendoderm differentiation

(A) Cells were cultured for one day with or without WNT3A (100 ng/mL, top bar). On the second day they were washed to remove WNT and treated with or without ACTIVIN (10 ng/mL, bottom bar). After the second day cells were fixed and analyzed by immunofluorescence: NANOG (red), OCT4 (green), SOX2 (cyan), DAPI (gray).

(B) Histograms showing the nuclear IF signal quantified in single cells corresponding to the experiment shown in (A). Quantification for n > 5,000 cells per condition.

(C) Cells were cultured for one day with or without WNT3A (100 ng/mL, top bar). On the second day they were washed to remove WNT and treated with or without ACTIVIN (10 ng/mL, bottom bar). After the second day cells were fixed and analyzed by immunofluorescence: BRA (red), EOMES (green), GSC (cyan), DAPI (gray).

(D) Histograms showing the nuclear IF signal quantified in single cells corresponding to the experiment shown in C. Quantification for n > 5,000 cells per condition.

Scale bars represent 50 μM. See also Figure S6.

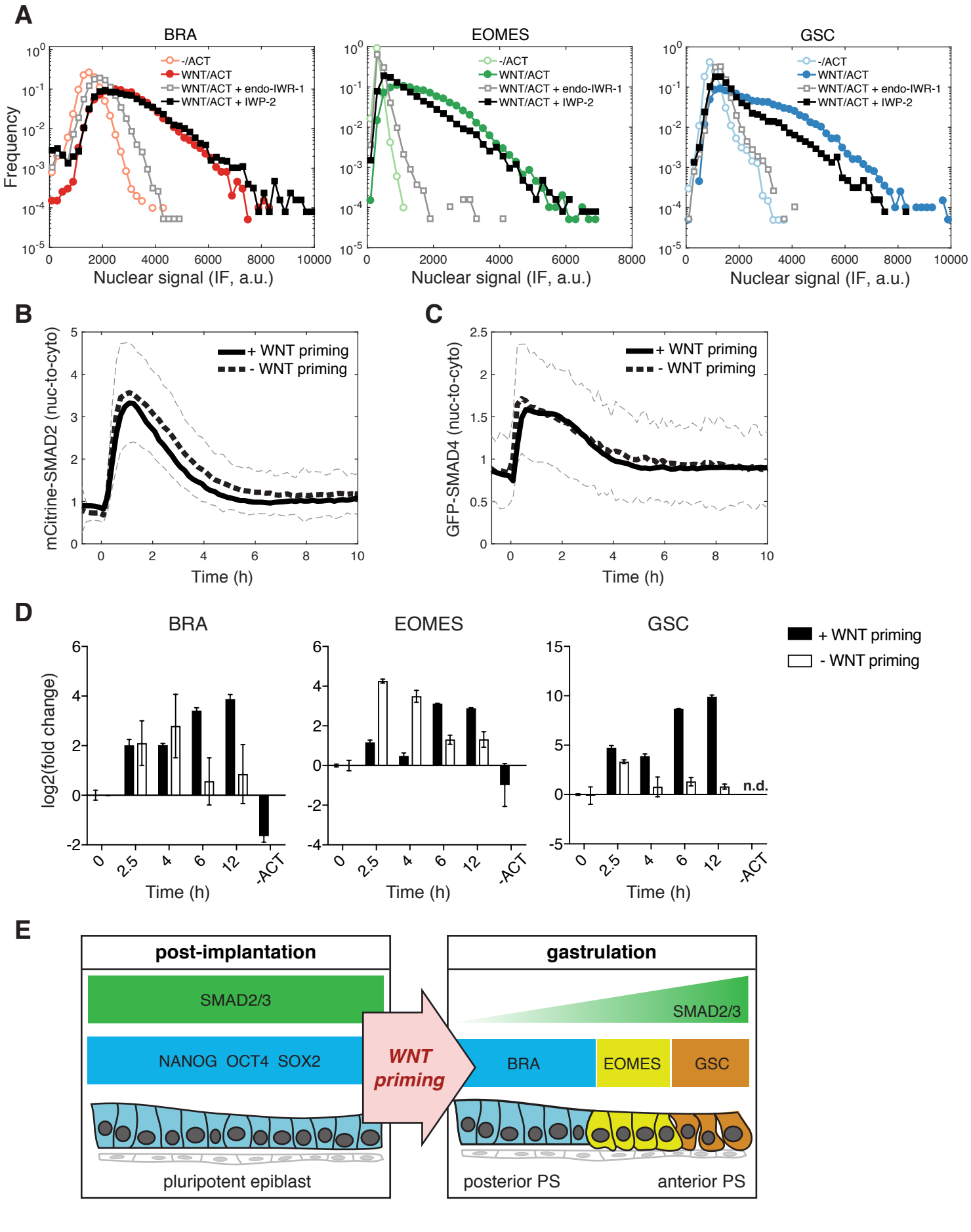


Figure 7

Figure 7. WNT signaling memory does not change SMAD2 response dynamics

(A) Cells were cultured for one day with WNT3A (100 ng/mL). On the second day they were washed to remove WNT and treated with ACTIVIN (10 ng/mL). In order to block WNT signaling through β-catenin cells were treated with endo-IWR-1 (1 μM), which was added with WNT on the first day and again with the ACTIVIN on the second day. In order to block WNT secretion, cells were similarly treated with IWP-2 (1 μM).
After the second day cells were fixed and analyzed by immunofluorescence. Histograms show the nuclear IF signal quantified in single cells (n > 5,000 cells per condition). Cells cultured without WNT (-/ACT) represent background signal levels.

(B) mCitrine-SMAD2 response in single cells to ACTIVIN (10 ng/mL) with or without WNT priming (black solid and black dashed lines, respectively). ACTIVIN was added at T = 0 h. Images were acquired every 10 minutes. The black solid and dashed lines represent the average response at each time point and the gray dashed lines represent the population standard deviation of the SMAD2 response without WNT priming (10 ng/mL ACTIVIN response from Figure 3E).

(C) GFP-SMAD4 response in single cells to ACTIVIN (10 ng/mL) with or without WNT priming (black solid and black dashed lines, respectively). ACTIVIN was added at T = 0 h. Images were acquired every 10 minutes. The black solid and dashed lines represent the average response at each time point and the gray dashed lines represent the population standard deviation of the SMAD4 response without WNT priming (ACTIVIN response from Figure 3G).

(D) Transcriptional response of mesendoderm genes to ACTIVIN with or without WNT priming. An additional sample was collected that was only treated with WNT and left untreated with ACTIVIN for the 12 h time course (-ACT). Expression in each sample was normalized to GAPDH and then to the level prior to ACTIVIN addition (T = 0 h).
n.d.: not detected.

(E) Model for WNT priming as a requirement for ACTIVIN to function as a morphogen during gastrulation. Following implantation ACTIVIN functions in a dose-independent manner to maintain pluripotency of the epiblast. At the onset of gastrulation, WNT signaling primes cells respond to graded ACTIVIN/SMAD2 signaling, which leads to the anterior-posterior patterning of cells emerging from the primitive streak (PS).

See also Figure S7.