The SMAD2/3 interactome reveals that TGFβ controls m^6A mRNA methylation in pluripotency

Bertero A. et al.,

Supplementary Information

Supplementary Discussion

Optimization of the SMAD2/3 co-immunoprecipitation protocol

The first step towards defining the SMAD2/3 interactome was to identify a co-immunoprecipitation protocol allowing specific identification of the greatest number of interactors following mass spectrometry analysis. Indeed, the biochemical conditions used for protein extraction and immunoprecipitation have a profound effect on the stability of various protein-protein interactions[71]. Of note, SMAD2 and SMAD3 interact with several of their known binding partners through a set of contiguous hydrophobic patches, referred to as the "hydrophobic corridor", which is located on the surface of the MH2 domain[72]. Therefore, biochemical conditions that stabilize hydrophobic interactions might be preferable. To test this, we compared two SMAD2/3 co-immunoprecipitation (co-IP) methods that are expected to preferentially preserve different types of protein-protein interactions.

First, we tested the protocol that we had recently used to demonstrate the interaction of SMAD2/3 with COMPASS complexes[10]. This method, which we named co-IP1, relies on an isotonic buffer with low concentration of a mild detergent (0.1% Tween-20) both to solubilize nuclear proteins and to minimize background binding during immunoprecipitation. As such, these biochemical conditions are likely to preserve salt-sensitive hydrophilic bonds, while the presence of detergent might partially interfere with hydrophobic interactions. Secondly, we developed a different protocol that avoids the use of detergent and in which nuclear proteins are extracted using a high-salt buffer followed by dialysis of the lysate to re-adjust the salt content to physiological levels (co-IP2). In this case, hydrophobic interactions should be better preserved due to the lack of detergent. However, the high-salt conditions used for the nuclear extraction could disrupt certain hydrophilic bonds, which might only be partially re-established following dialysis.

Initial comparisons in hESCs demonstrated that both conditions allowed the detection of well-characterized SMAD2/3 binding factors by Western blot, with co-IP2 being slightly more efficient than co-IP1 (Extended Data Fig. 1a). We therefore performed small-scale pilot mass spectrometry analyses of SMAD2/3 co-IPs from hESCs using both methods (Extended Data Fig. 1b). Remarkably, co-IP2 allowed the identification of a larger number of SMAD2/3 interacting proteins following statistical analysis (23, compared to 12 for co-IP1; Extended Data Fig. 1b). Interestingly, roughly half of the proteins significantly enriched in co-IP1 experiments were also identified as specific binders in co-IP2 ones. On the other side, co-IP2-specific interactors included important transcription factors (such as SOX13, ETV6, and SMAD4), epigenetic regulators (like SETDB1 and ATF7IP), and RNA-binding proteins (for instance WTAP and CPSF6). Overall, these results showed that the co-IP2 protocol is more suitable for the large-scale analysis of SMAD2/3 interacting proteins. We therefore chose this method for subsequent experiments.

Functional roles of the SMAD2/3 interactome

Having identified the SMAD2/3 interactome in hPSCs, we decided to validate its functional role in pluripotency and endoderm differentiation. First, we focused on selected transcriptional and epigenetic cofactors (the transcription factor FOXH1, the histone acetyltransferases EP300 and CREBBP, and the histone methyltransferase SETDB1), as the function of these proteins in hPSCs is not fully understood. We confirmed their interactions with SMAD2/3 by co-IPs followed by Western Blot (Extended Data Fig. 2a). We then took advantage of our recently established OPTimized inducible gene KnockDown system (OPTiKD[27]; Extended Data Fig. 2b) to decrease the expression of these factors in hESCs and during differentiation (Extended Data Fig. 2c and 3a). Knockdown of SMAD2 was used as a positive control in these experiments as this factor is necessary for both pluripotency and endoderm specification[30,73].

Interestingly, knockdown of FOXH1 did not result in overt hESC differentiation, while this gene was required for endoderm differentiation (Extended Data Fig. 2d-f and 3b-d). This indicates that this well-known SMAD2/3 co-factor regulates only a specific subset of the transcriptional responses to Activin/Nodal signalling and it is predominantly involved in the expression of endoderm genes. This is in agreement with previous ChIP-seq results, which showed that FOXH1 and SMAD2/3 only weakly colocalize in pluripotent cells, while their genomic binding largely overlaps during endoderm differentiation[74]. Finally, loss of Foxh1 in the mouse embryo does not cause overt defects in the post-implantation epiblast, while it specifically impairs

patterning of the anterior primitive streak and formation of the node, prechordal mesoderm, notochord, and definitive endoderm[75–77].

In contrast to this, decreased expression of SETDB1, EP300, or CREBBP induced hESC differentiation, while having only moderate (SETDB1), little (EP300), or no effect (CREBBP) on endoderm specification (Extended Data Fig. 2d-f and 3b-d). Of note, gastrulation is not affected in mice knockout for Ep300 or Crebbp, which only show later embryonic defects such as heart malformations, defective neurulation, and impaired haematopoiesis and vasculogenesis[78–81]. This suggests that EP300 and CREBBP might be redundant during Activin/Nodal-induced endoderm specification, either because they compensate for each other or because other epigenetic regulators play a more significant role (for instance the COMPASS complexes[10]). Further studies involving conditional single and double conditional knockout for EP300 and CREBBP during hPSC differentiation will be required to clarify their role in early cell-fate choices.

Moving beyond the functional validation presented here, our data show that SMAD2/3 interacts with more than a dozen of different transcription factors and a similar number of epigenetic modifiers. Aside from well-known SMAD2/3 cofactors (such as SMAD4, SKI, and SNON), most of these proteins have never been previously reported to interact with SMAD2/3. These include multiple transcription factors (for instance ETV6, NFAT5, and SOX13) whose role in hPSCs is unknown despite being crucial for other developmental processes[82–86]. We anticipate that future studies will take advantage of the dataset we present to further dissect the transcriptional and epigenetic regulations involving SMAD2/3 in hPSCs.

To our surprise, the interactome of SMAD2/3 proved remarkably similar in undifferentiated hPSCs and hPSCs differentiating into endoderm (Extended Data Fig. 1e). Nevertheless, a few factors appear to differentially bind to SMAD2/3 in each condition. The most remarkable example is FOXH1, which as discussed above functionally interacts with SMAD2/3 to regulate expression of endoderm. Overall, only limited differences in the SMAD2/3 interactome could be sufficient to substantially modify the outcome of Activin/Nodal signalling in hPSCs. On the other hand, it is possible that at later stages of hPSC differentiation the changes in the SMAD2/3 interactome might become more significant, with novel partners such as EOMES driving yet other SMAD2/3-dependent transcriptional responses[13].

Optimization of the m6A methylated RNA immunoprecipitation (MeRIP) experiments

Since our data showed that SMAD2/3 physically interacts with the m6A methyltransferase complex (Figs. 1 and 2a-d), we hypothesized that Activin/Nodal signalling might regulate m6A deposition. To test this notion, we performed m6A methylated RNA immunoprecipitation (MeRIP) followed by qPCR to monitor the level of m6A onto Activin/Nodal-regulated transcripts following short-term signalling inhibition. Interestingly, treatment of hESCs for 2h with the Activin/Nodal inhibitor SB-431542 (SB) decreased m6A levels of specific nuclear transcripts (such as *NANOG* and *LEFTY1*, but not *DPPA4*), while cytoplasmic mRNAs were unaffected (Extended Data Fig. 5a-b). This suggested that SMAD2/3 might promote m6A deposition onto certain transcripts at the nuclear level, in agreement with its known localization. Therefore, in our following experiments we decided to focus on nuclear transcript by performing nuclear-enriched MeRIP, a method which we named NeMeRIP (Fig. 2e and Extended Data Fig. 5c-j).

RNA-seq experiments in WTAP inducible knockdown cells.

Having shown that WTAP regulates expression of several Activin/Nodal target genes (Fig. 3 and Extended Data Figs. 9 and 10a-b), we validated these findings at a genome-wide level. Therefore, we performed RNA-seq in WTAP inducible knockdown cells cultured in presence of Activin or following Activin/Nodal signalling inhibition (Extended Data Fig. 10c-e) These results confirmed that knockdown of WTAP globally alters the response to Activin/Nodal signalling by: (1) upregulating a large cohort of developmental regulators whose expression is maintained by Activin/Nodal in the pluripotent state, and by delaying the downregulation of such genes upon Activin/Nodal inhibition (cluster 2); (2) impairing the upregulation of neuroectoderm genes induced following inhibition of Activin/Nodal (cluster 3). Additionally, WTAP knockdown resulted in up- and downregulation of additional factors whose expression is largely independent from Activin/Nodal signalling, and which are not associated to developmental regulations (cluster 1 and 4, respectively). This showed that WTAP has additional functions other than modulating the response to Activin/Nodal signalling, in agreement with its role as a general regulator of the epitranscriptome. Overall, these findings demonstrated that WTAP has an important role in modulating the gene expression network controlled by Activin/Nodal signalling.

Supplementary Information specific references

71. Sambrook, J. & Russell, D. W. Identification of associated proteins by coimmunoprecipitation. *CSH Protoc.* **2006,** (2006).
72. Massagué, J., Seoane, J. & Wotton, D. Smad transcription factors. *Genes Dev.* **19,** 2783–810 (2005).
73. Sakaki-Yumoto, M., Liu, J., Ramalho-Santos, M., Yoshida, N. & Derynck, R. Smad2 Is essential for maintenance of the human and mouse primed pluripotent stem cell state. *J. Biol. Chem.* **288,** 18546–18560 (2013).
74. Kim, S. W. *et al.* Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. *Dev. Biol.* **357,** 492–504 (2011).
75. von Both, I. *et al.* Foxh1 is essential for development of the anterior heart field. *Dev. Cell* **7,** 331–45 (2004).
76. Yamamoto, M. *et al.* The transcription factor FoxH1 (FAST) mediates Nodal signaling during anterior-posterior patterning and node formation in the mouse. *Genes Dev.* **15,** 1242–56 (2001).
77. Hoodless, P. a *et al.* FoxH1 (Fast) functions to specify the anterior primitive streak in the mouse. *Genes Dev.* **15,** 1257–71 (2001).
78. Oike, Y. *et al.* Mice homozygous for a truncated form of CREB-binding protein exhibit defects in hematopoiesis and vasculo-angiogenesis. *Blood* **93,** 2771–9 (1999).
79. Kung, A. L. *et al.* Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP. *Genes Dev.* **14,** 272–7 (2000).
80. Yao, T. P. *et al.* Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300. *Cell* **93,** 361–72 (1998).
81. Shikama, N. *et al.* Essential function of p300 acetyltransferase activity in heart, lung and small intestine formation. *EMBO J.* **22,** 5175–85 (2003).
82. Tsuzuki, S. & Seto, M. TEL (ETV6)-AML1 (RUNX1) initiates self-renewing fetal pro-B cells in association with a transcriptional program shared with embryonic stem cells in mice. *Stem Cells* **31,** 236–47 (2013).
83. Ford, A. M. *et al.* The TEL-AML1 leukemia fusion gene dysregulates the TGF-beta pathway in early B lineage progenitor cells. *J. Clin. Invest.* **119,** 826–36 (2009).
84. Melichar, H. J. *et al.* Regulation of gammadelta versus alphabeta T lymphocyte differentiation by the transcription factor SOX13. *Science* **315,** 230–3 (2007).
85. Wang, Y., Bagheri-Fam, S. & Harley, V. R. SOX13 is up-regulated in the developing mouse neuroepithelium and identifies a sub-population of differentiating neurons. *Dev. brain Res.* **157,** 201–8 (2005).
86. Berga-Bolaños, R., Alberdi, M., Buxadé, M., Aramburu, J. & López-Rodríguez, C. NFAT5 induction by the pre-T-cell receptor serves as a selective survival signal in T-lymphocyte development. *PNAS* **110,** 16091–6 (2013).