Dynamic lineage priming is driven via direct enhancer regulation by ERK

DOI:
10.1038/s41586-019-1732-z

Citation for published version (APA):

Published in:
Nature

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [http://man.ac.uk/04Y6B0] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
Dynamic Lineage Priming is Driven via Direct Enhancer Regulation by ERK

William B. Hamilton¹, Yaron Mosesson²,⁵, Rita S. Monteiro¹,⁵, Kristina B. Emdal³, Teresa E. Knudsen¹, Chiara Francavilla³,⁴, Naama Barkai², Jesper V. Olsen³, and Joshua M. Brickman¹,⁶.

¹The Novo Nordisk Foundation Center for Stem Cell Biology, 3B Blegdamsvej, DK-2200 Copenhagen N, Denmark. ²Weizmann Institute of Science, Rehovot 76100 Israel. ³The Novo Nordisk Foundation Center for Protein Research, 3B Blegdamsvej, DK-2200 Copenhagen N, Denmark. ⁴Division of Molecular & Cellular Function, School of Biological Sciences, FBMH, The University of Manchester, Oxford Rd., Manchester, UK. ⁵These authors contributed equally to this manuscript. ⁶Correspondence to be addressed to joshua.brickman@sund.ku.dk
Summary

Central to understanding cellular behaviour in multi-cellular organisms is the question of how a cell exits one transcriptional state to adopt, and eventually become committed to another. FGF-ERK signalling drives differentiation of both embryonic stem cells (ESC) and pre-implantation embryos towards primitive endoderm (PrE) and inhibiting ERK supports ESC self-renewal\(^1\). Paracrine FGF-ERK signalling induces heterogeneity, where cells reversibly progress from pluripotency towards PrE, while retaining their capacity to re-enter self-renewal\(^2\). Here we find that ERK reversibly regulates ESC transcription by directly affecting enhancer activity without requiring a change in transcription factor (TF) binding. ERK triggers the reversible association and disassociation of RNAPII and associated co-factors from genes and enhancers with the Mediator component, MED24, playing an essential role in ERK-dependent transcriptional regulation. Though the binding of Mediator components responds directly to signalling, the persistent binding of pluripotency factors to both activated and repressed genes bookmarks them for activation, and/or reactivation in response to fluctuations in ERK activity. Amongst the repressed genes are several core components of the pluripotency network that all act to drive their own expression and maintain the ESC state, and if their binding is lost, the ability to reactivate transcription is compromised. Thus, as long as TF occupancy is maintained so is plasticity, enabling cells to distinguish between transient and sustained signals. If ERK signalling persists, pluripotency TF levels are reduced by protein-turnover and irreversible gene silencing and commitment can occur.
TFs bind to promoter and enhancer sequences to regulate transcriptional programs by recruiting the core RNA Polymerase II (RNAPII) complex alongside co-activators to assemble both a pre-initiation complex (PIC) and promote the active engagement of RNAPII. Co-activators such as Mediator and EP300 bridge the interaction between the polymerase and TFs resulting in enhanced transcription\(^3\). In ESCs, enhancer decommissioning has been shown to be an early event in differentiation and suppression of the pluripotency network\(^4\). This network forms a positive feedforward loop in which TFs positively regulate each other’s expression through reciprocal binding and gene activation\(^5\). Several studies have implicated ERK as a differentiation trigger, however nuclear targets of ERK, as well as how ERK regulates transcription outside of the mitogenic response, are poorly understood\(^6\).

We previously generated a rapid and reversible system for ERK activation in ESCs (Fig. 1a, Extended Data Fig. 1a-c and ref\(^7\)) that induces phospho-ERK to a similar magnitude as FGF4 stimulation on a single cell level, however, unlike FGF4 treatment, our system maintains ERK activity homogenously within the culture (Extended Data Fig. 1d), making it ideal for population level analysis. Activation-inhibition experiments showed that ERK reversibly induced canonical mitogen responsive genes, while it reversibly repressed key components of the ESC pluripotency network, in a non-selective manner (Fig. 1b, Extended Data Fig. 1e, f, Supplemental Video 1 and supplemental table 1). Between 24- and 36hrs of ERK signalling is sufficient to render a Nanog::EGFP protein reporter\(^8\) refractory to reactivation (Fig. 1c and Extended Data Fig. 1g), correlating with appearance of low levels of H3K27me3 (Extended Data Fig. 1h).

We next tested whether this transcriptional response was a function of ERK-regulated changes in mRNA and protein stability\(^9\)-\(^11\), by inducing ERK in the presence or absence of either actinomycin d, or MG132. We failed to observe any ERK-dependent change in mRNA degradation rates of target genes (Extended Data Fig. 1i, j, and supplemental table 2), or an effect of proteasomal inhibition on ERK transcriptional response (Extended Data Fig. k, l and supplemental table 3). Moreover though KLF2 and KLF4 have been reported to be phosphorylated and destabilized by ERK\(^10,11\), phospho-mimetic mutants failed to block ERK-mediated suppression of Nanog::EGFP (Extended Data Fig. 1m).

RNA-seq analysis showed that robust changes in mRNA levels are not evident until 4hrs post stimulation (Extended Data Fig. 1n and supplemental table 4), however, analysis of intronic reads, nascent 4SU-seq and cycloheximide experiments showed that ERK directly
regulates pluripotency transcription with rapid kinetics (Fig. 1d, Extended Data Fig. 2a, b and supplemental tables 5-6). As PRC2 has been implicated in ERK-mediated gene regulation\textsuperscript{12} we analysed H3K27me3 deposition across ERK-responsive genes, but found it was unaffected by ERK (Extended Data Fig. 2c, and supplemental table 7), nor was the catalytic activity of EZH2 necessary for transcriptional repression (Extended Data Fig. 2d, e). Taken together, we find that ERK regulates ESC transcription in a direct and rapidly reversible manner, and that the regulation of mRNA, protein turnover, or PRC2 activity, likely function downstream to stabilise this new transcriptional state.

Promoter-proximal RNAPII pausing has been proposed as a mechanism to maintain poised transcriptional states in development\textsuperscript{13,14}, and could explain the dynamic transcriptional responses to ERK. RNAPII occupancy was assessed at the transcriptional start site (TSS) and polyadenylation cleavage site (PAS) (Extended Data Fig. 2f) following 2- and 8hrs of ERK stimulation, and at 8 hours of stimulation followed by 2hrs of inhibition. We found that RNAPII binding positively correlates with the expression state of regulated genes (Fig. 2a, supplemental table 8, R2 and R3 for TSS and PAS, respectively), and that the changes in binding were similar at the PAS and TSS (R1), indicating that repression does not utilize TSS pausing, or activation TSS release. Changes in RNAPII binding were also rapidly reversed upon the inhibition of ERK signalling with MEKi (Fig. 2b, Extended Data Fig. 2g and supplemental table 8). We also observed ERK dependent changes in the binding of the general factor, Tata Binding Protein (TBP) (Extended Data Fig. 2h), implying that ERK regulates RNAPII and PIC binding at important ESC and differentiation genes.

Regions of high Mediator 1 (MED1) clustering, termed super enhancers (SEs)\textsuperscript{15} have been shown to drive robust expression of many pluripotency and mitogenic genes, that are also regulated by ERK (Extended Data Fig. 2i and supplemental table 9). Moreover, expression of non-coding enhancer RNA (eRNA) at these SEs is also regulated by both ERK and endogenous FGF4 signalling (Fig. 2c, Extended Data Fig. 3a and supplemental table 10) and is correlated with RNAPII binding (Extended Data Fig. 3b). To further explore enhancer regulation, we assessed EP300 binding and H3K27ac deposition to generate a list of not only SEs, but also putative traditional enhancers (TEs) active in our datasets. We found a discrete subset of these TEs were repressed (8.7%) and induced (3.2%) by ERK, and that changes in H3K27ac were tightly correlated with EP300 binding and adjacent gene expression (Extended Data Fig. 3c-e, and supplemental table 11). Though EP300 binding and H3K27ac deposition
were affected at regulated SEs, the magnitude was considerably less than at TEs, indicating SEs exhibit a high level of cooperativity and stability (Fig. 2d), also reflected in their chromatin accessibility (Extended Data Fig. 3f-i). Motif analysis showed that activated enhancers were enriched in binding sites for ETS/AP1 TFs, whereas repressed enhancers were dominated by nuclear hormone receptor motifs, particularly that of Esrrb. Surprisingly, sites for core pluripotency factors were enriched in both classes of enhancer (Fig. 2e), supporting their proposed role in promoting differentiation.

Enhancer activity is thought to be determined by TF binding, the regulation of which is implicated in determining cell states in development\(^3,15\). As ESC TFs drive their own and each other’s expression, we explored how the expression and binding of these TFs responded to ERK. We found protein expression of these factors persisted in early stages of differentiation, with the exception of NANOG (Extended Data Fig. 3j, k). However we have previously shown that misexpression of Nanog does not block the transcriptional response to ERK\(^7\). We next assessed the binding of two key factors that occupy most ESC enhancers; Esrrb as it is enriched at ERK-repressed enhancers and is itself regulated by ERK, and Sox2 because it is unaffected by ERK signalling (Extended Data Fig. 3j, l). Almost all regulated enhancers showed little difference in TF binding following 2hrs ERK activity (Fig. 3a), indicating that changes in TF binding are secondary to changes in EP300 binding. Moreover, even after 8hrs signalling, the majority of TFs remained associated with repressed TEs (>70%, Fig. 3b, Extended Data Fig. 3m and supplemental table 12), and almost all SEs (Fig. 3c and Extended Data Fig. 4a-c). Analysis of regions with reduced binding showed that they lack Sox2 and Esrrb motifs indicating these TFs only vacate low affinity binding sites (Extended Data Fig. 4d). Mobilised TFs were recruited to ETS/AP1-regulated enhancers, also lacking specific recognition sequences (Fig. 3d and Extended Data Fig. 4e, f), therefore likely through protein-protein interactions at regions of robust transcription (Extended Data Fig. 4g). While Esrrb is not required for induction of these genes (Extended Data Fig. 4g, and supplemental table 13), its persistent binding was required for transcriptional reactivation following ERK-mediated repression (Fig. 3e). Taken together this suggests that plasticity involves regulation on at least two timescales; early co-factor and RNAPII disassociation correlating with changes in expression, and subsequent to this, the loss of TF binding that correlates with the reduced reactivation rates associated with commitment.
We employed quantitative phosphoproteomics (Extended Data Fig. 5a) identifying many potential ERK substrates possessing either a canonical MAPK (PXSP, cluster 2)\textsuperscript{17}, or AKT/AGC kinase motif (RXRXXS/T, cluster 4) (Fig. 4a and supplemental table 14), with cluster 2 being enriched in transcriptional regulators (Extended Data Fig. 5b). Depletion of PDK1, essential for most AKT/AGC kinase activity in ESCS\textsuperscript{18} failed to impact in transcriptional regulation by ERK (Extended Data Fig. 5c-f and supplemental table 15), therefore we focused only on peptides possessing proline-directed serine phosphorylations. We next employed two complimentary phosphoproteomics approaches to refine the list of targets; firstly, by stimulating FGF4-/MEKi cells with rFGF4 in the presence or absence of MEKi; and secondly, by generating ERK\textsuperscript{2AS} analogue sensitive cells (ERK\textsuperscript{2AS}) that can be inhibited by PP1 analogues\textsuperscript{19,20} (Extended Data Fig. 5g, h and 6a-c). With these tools we detected 784 FGF4-induced, MEKi inhibited, and 3478 4OHT induced, MEKi and PP1 inhibited phosphorylation sites, again, enriched in proline-directed serine phosphorylation (Fig. 4b, Extended Data Fig. 6d, e and supplemental table 16). Subclusters containing FGF4-regulated and MEK-independent phosphorylated peptides (n=919) display enrichment for substrates of Akt, p38, PKC and p70S6K (supplemental table 16). Thus, FGF4 stimulation induces a broader range and diversity of phosphorylation than that generated by focused RAF induction. Moreover, analysis of nascent transcription in ERK\textsuperscript{2AS} cells showed that the major transcriptional response was dependent on the kinase activity of MEK and ERK (Fig. 4c, Extended Data Fig. 6f and supplemental table 17). Additionally, we observed hyperactivated MEK in the ERK\textsuperscript{2AS} cells (Extended Data Fig. 6g), however robust repression of pluripotency genes and induction of mitogenic targets required ERK2 rescue (Extended Data Fig. 6h). Taken together, both phosphoproteomic and transcriptional profiling indicate that ERK is the primary kinase mediating FGF4/RAF signalling and that MEK activation alone is insufficient to regulate transcription.

As we had identified Esrrb binding as enriched, stably associated with, and required for reactivation of repressed enhancers, we looked specifically for factors that interact with this TF (Extended Data Fig. 6i, data from ref\textsuperscript{21}). One of these factors, MED24, is part of a subcomplex that had been implicated in MAPK activity\textsuperscript{22-24}. MED24, identified in all three screens, (Fig. 4d) is robustly phosphorylated by ERK and its expression is upregulated in peri-implantation development (Extended Data Fig. 7a, b). We therefore deleted endogenous Med24 and rescued function with a FLAG-tagged cDNA under the control of a TET inducible
promoter to generate a conditional allele (Extended Data Fig. 7c-g). ChIP-seq for FLAG-MED24 (hereafter referred to as wild type (WT)) showed MED24 was bound to regions associated with early post-implantation embryonic development, and in the vicinity of Esrrb and ETS/AP1-factor motifs at ERK-regulated enhancers (Extended Data Fig. 7h, i). Depletion of MED24 resulted in a reduction in Nanog::EGFP expression and RNAPII binding to the Nanog enhancer (Extended Data Fig. 7j, k), suggesting it is a transcriptional activator.

Transcriptional profiling of nascent transcription showed that both robust activation and repression of ERK targets was attenuated upon MED24 depletion, with the dependency on MED24 being particularly evident for a subset of SE-driven genes (Fig. 5a, b and Extended Data Fig. 7l and 8a). ChIP-seq analysis of MED24 and MED1 showed they colocalise with Esrrb, that MED1 is rarely found with Esrrb in the absence of MED24 and that their binding is regulated only at subset of enhancers (Extended Data Fig. 8b, c). Furthermore, both Mediator components were readily lost from repressed, and loaded onto activated SEs at magnitudes greater than that of EP300 (Fig. 5c and Extended Data Fig 8d). Additionally TEs with changes in both Mediator and EP300 binding were more significantly enriched for ESRRB and ETS/AP1 motifs than those with changes in only EP300 (Extended Data Fig. 8e-h, and supplemental tables 19-21). Taken together this indicates that Mediator binding is more closely correlated with the activity of ERK-regulated SEs than either EP300 or TFs.

We have previously shown that ERK is required specifically for PrE, but not neural differentiation\(^7\). We therefore tested the requirement for MED24 in ESC differentiation to both lineages. While neural specification proceeded normally in the absence of MED24 differentiation to PrE was compromised (Fig. 5d, e, and Extended Data Fig. 8i, j and 9a-e). This defect occurs during the transition from ESC to PrE as MED24 is dispensable for naïve extra-embryonic endoderm (nEnd)\(^25\) viability and expansion (Extended Data Fig. 9f, g). Mutation of both serine phosphorylation sites identified in MED24 also attenuated PrE differentiation (Extended Data Fig. 9h, i). Thus, ERK regulates the transition from pluripotency/epiblast identity into PrE by direct and reversible regulation of enhancers associated with both pluripotency and MAPK/differentiation via specific co-factors that include MED24.

A fascinating feature of early mammalian development is the ability of cells to revise developmental decisions in response to changing environmental cues. This phenomenon underpins regulative development\(^26-29\) and requires plasticity in the underlying gene regulatory network (GRN). The heterogeneities and plasticity observed in pre-implantation
development manifest transiently, however the *ex vivo* expansion of these cells as ESCs traps them in a perpetual cycle of reversible priming. We found that the plasticity of the pluripotency network required to maintain this state is dependent on the inherent stability of its components, with their expression determining the time frame preceding commitment. Through shifting the transcriptional machinery and selected co-factors from pluripotency genes to lineage and mitogenic ones, ERK both primes differentiation and ensures that reactivation of the ESC GRN can occur within a discrete time window (Extended Data Fig. 10). Thus, signalling triggers a molecular count-down clock, which can be rewound upon inhibition and runs out when the pluripotency TFs drop below a functional level that supports reactivation. That transcription can be repressed, but TF occupancy of enhancers left intact suggests a general mode for signalling response that enables cells to preserve developmental potency. This suggests that TFs provide the architecture for alternative lineage choices, outlining GRNs that can be activated when the appropriate signalling pathway induces co-factors to choose between sets of preordained enhancers to drive cells towards specific states in development and differentiation.
**Author contribution**

W.B.H. and J.M.B. conceived the study. W.B.H., J.M.B., C.F., K.B.E., J.V.O., R.S.M. and T.E.K. designed and interpreted experiments. W.B.H. performed all experiments except ATAC-seq, which was performed by R.S.M., and phosphopeptide screens and analysis, which were performed by C.F. and K.B.E. R.S.M. prepared samples for sequencing. The initial analysis of enhancer and RNAPII binding was performed by Y.M. under the guidance of N.B. with all subsequent bioinformatics analysis carried out by W.B.H., R.S.M. and T.E.K. W.B.H and J.M.B. wrote the manuscript with input from all other authors.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints)

Correspondence and requests for materials should be addressed to joshua.brickman@sund.ku.dk

**Acknowledgements**

We thank M. Thomson for the NANOG–eGFP ES cells, T. Kunath for Erk2-KO cells, N. Festuccia for Esrrb-KO cells, H.H. Ng for the KLF2 and TFCP2L1 antibodies, the Brickman laboratory members for critical discussions, Y. Spector for sequencing assistance, H. Neil, M. Michaut and the DanStem Genomics Platform for technical expertise, support, and the use of instruments, S. Pozzi, N. Festuccia and P. Navarro Gil for advice on ChIP protocols, K. Stewart-Morgan for help with ATAC-seq, A. Azad and J. A. R. Herrera for bioinformatics advice and P. van Dieken for technical support and proof reading. This work was funded by grants from the Novo Nordisk Foundation, Danish Council for Independent Research (8020-00100B), Danish National Research Foundation (DNRF116) and Human Frontiers in Science (RGP00008/2012). The Novo Nordisk Foundation Centers for Stem Cell Biology and Protein Research are supported by a Novo Nordisk Foundation grant numbers NNF17CC0027852 and
NNF14CC0001. R.S.M. is supported by a fellowship from the Lundbeck Foundation (R303-2018-2939).
Fig. 1 | Transcriptional regulation by ERK is rapid and direct. 

a, A cartoon depicting the experimental outline for ERK activation by cRAF-ErT2. 

b, Scatter plot depicting genes either repressed or induced by ERK following 8h stimulation and the extent to which they revert back to t=0 expression ((abs(log2fc) >1, ANOVA FDR<= 0.01, n = 4; 0h, n = 5; 8h, n =5; 8h2h biologically independent samples). Colours are scaled by half-life, (log2(hrs)), from ref31. 

c, Flow cytometry histograms of NANOG::EGFP expression following either a 24- or 36h stimulation of ERK and a subsequent 4hrs treatment with MEKi, data is representative of 6 biologically independent samples. 

d, Volcano plot of the changes in nascent RNA levels following 2hrs ERK induction. 916 genes show significant induction, whereas 1412 genes are being repressed ((abs(log2fc) >1, padj <= 0.01, n = 3 biologically independent samples). Colour scale represents the -log10 adjusted p-value.
**Fig. 2 |** ERK controls both RNAPII binding and enhancer activity at regulated genes. 

**a,** A scatter plot contrasting the changes in RNAPII binding between TSS and PAS of regulated genes. The loss and gain of RNAPII at TSSs and PASs are positively correlated, R1 (Spearman's correlation coefficient), with changes in nascent transcription most closely correlating with loss at the TSS (R3) than at PAS (R2). Colours are scaled by log2fc in nascent transcription from Figure 1d. n = 3 biologically independent samples. 

**b,** ChIP-seq profiles of Nanog and Spry4, data is representative of 3 biologically independent samples. 

**c,** Volcano plot showing the ERK-dependent changes in eRNA expression at SEs. See materials and methods for details. n = 3 biologically independent samples. 

**d,** Box and whisker plots showing that the magnitude of change at ERK-regulated TEs (defined in Extended Data Fig. 3c) is considerably less at SEs, (two sided Wilcoxon test), with only activated SEs showing significant changes in EP300 binding (n = 19675; all TEs, n = 1726; repressed TEs, n = 636; activated TEs, n = 231; all SEs, n = 30; repressed SEs, n = 27; activated SEs, **p < 2.2e-16, *p = 0.01). Data is derived from two biologically independent samples. 

**e,** Motif analysis on ERK-regulated TEs showed the enrichment of motifs associated with canonical ETS/AP1 factors at activated, nuclear hormone receptors at repressed, and canonical pluripotency factors at both classes of enhancer.
**Fig. 3 | Pluripotency TF binding to enhancers is stable.**

**a,** Chow Ruskey diagrams showing the overlap between repressed or activated enhancers with changes in EP300, Esrrb and Sox2 binding. n = 2 biologically independent samples. Overlap is shown as relative to all ERK-regulated enhancers identified in Extended Data Figure 3c. **b,** The percentage of all ERK-regulated enhancers that exhibited significant changes across the 8hr time course. n = 2 biologically independent samples. **c,** ChIP-seq tracks for Esrrb and Sox2 at the Klf2 and Arc loci. Data is representative of 2 biologically independent samples. **d,** Motif analysis at enhancers with changing TF binding at 8hrs. Enhancers with increasing Esrrb or Sox2 binding show an enrichment only in ETS/AP1 factor motifs. n = 138; Sox2, n = 103; Esrrb. Data is derived from 2 biologically independent samples. **e,** Box and whisker plots showing the rate of reactivation following 8h ERK stimulation and subsequent 2h MEKi treatment in Esrrb WT or KO cells. The loss of Esrrb results in a significant reduction in reactivation of both SE and TE driven genes (two-sided Wilcoxon test, **p < 2.2e-16, *p = 0.01).** n = 6712; all genes, n = 61; SE-driven. Data is derived from 3 biologically independent samples.
Fig. 4 | ERK phosphorylates key transcriptional regulators. **a,** Clustering analysis of ERK-dependent phosphopeptides. Cluster 2 is enriched for the canonical MAPK motif. The location of the phosphorylation is centred at 0 (red), with enrichment non-serine residues highlighted (see clusters 3 and 4). n = 2 biologically independent samples. **b,** Clustering of FGF4- or 4OHT-induced phosphorylations that are dependent on the kinase activity of MEK/ERK from both TMT11 multiplexed experiments. n = 3; (FGF4)-DMSO, n = 3; (FGF4)-FGF4, n = 3; (FGF4)-MEKi, n = 2; (FGF4)-FGF4/MEKi, n = 3; (ERK2<sup>AS</sup>)-DMSO, n = 3; (ERK2<sup>AS</sup>)-4OHT, n = 3; (ERK2<sup>AS</sup>)-4OHT/PP1, n = 2; (ERK2<sup>AS</sup>)-4OHT/MEKi biologically independent samples. **c,** PCA of nascent transcription changes following 2hrs ERK induction in ERK2<sup>AS</sup> cells. n = 2; 0h DMSO, n = 3; 0h PP1, n = 3; 0h MEKi, n = 3; 2h DMSO, n = 3; 2h PP1, n = 3; 2h MEKi biologically independent samples. **d,** Venn diagrams showing the overlap of peptides from Figures 4a, and 4b. Phosphorylation of MED24 was detected in all experiments.
**Fig. 5 | Med24 is necessary for robust ERK-regulated transcription and primitive endoderm differentiation.** a, Transcriptional response to ERK is reduced in the absence of MED24. Box and whisker plots showing the fold change of all ERK regulated genes in WT or MED24 KO cells, (i) = induced, (r) = repressed. Both the magnitude of activation and repression is reduced in the absence of MED24 (two-sided Wilcoxon test, *p < 2.2e-16). n = 707; 2hr induced, n = 793; 2hr repressed, n = 705; 8hr induced, n = 608; 8hr repressed. Data is derived from 3 biologically independent samples. b, Ordered expression changes of SE-driven genes showing considerable MED24-dependent regulation for both repressed and activated genes. n = 3 biologically independent samples c, ChIP-seq tracks of the binding of the indicated factors to either the *Klf4*, or *Dusp5* SE. TF, and EP300 binding is largely unaffected by ERK, however the binding of both MED24 and MED1 reflect SE activity. Data is representative of two biologically independent samples and is scaled as mapped reads per million (Sox2;12, Esrrb;12, Med24;15, Med1;15, EP300;15). d, Immunofluorescence of MED24 WT or KO cells following 5 days of neural differentiation for the early neural marker SOX1. Data is representative of three biologically independent samples. e, Immunofluorescence of MED24 WT or KO cells for the PrE markers GATA6 and GATA4 following 5 days of PrE-differentiation. Data is representative of three biologically independent samples.
Extended Data Fig. 1 | ERK induction by cRAF-ErT2 mimics endogenous FGF activity and directly regulates ESC transcription. a, A schematic representation of the inducible c-RAF-ErT2 construct. Versions with and without the FKBP<sup>L106P</sup> were used in this study and were found to be identical with respect to leakiness and induction. Stable cell lines were generated by random integration, all clones were screened for homogenous induction and maintained in selection for routine culture to prevent transgene variegation. b, Immunofluorescence analysis of pERK following 30 minutes 4OHT of treatment in cRAF-ErT2 cells showing homogenous induction of ERK phosphorylation. Data is representative of more than three biologically independent samples. c, Western blot showing loss of ERK and p90RSK phosphorylation following a 2-hour induction and a subsequent 30-minute treatment with 1 μM MEKi. All cell treatments were performed without media changes to avoid confounding effects of serum stimulation. Data is representative of greater than three biologically independent samples. d, Immunofluorescence analysis of pERK in FGF4<sup>-/-</sup> ESCs expressing the c-RAF-ErT2 construct stimulated with either rFGF4 (40ng/ml) or 4OHT (250 nM) for the indicated times. Arrowheads show cells with persistent pERK staining in FGF4 stimulated cells. Data is representative of three biologically independent samples. e, Schematic of a single ERK pulse. f, GO analysis of ERK-regulated genes at 8hrs. Data is derived from 2 biologically independent samples. g, Comparison of the ERK-dependent half-life of WT Nanog and Nanog::EGFP showing no significant differences resulting from the EGFP tag. WT or Nanog::EGFP cells were subjected to ERK activity and Nanog expression was monitored at 3 or 6 hrs. The half-life was calculated by \( t_x \cdot \log(0.5)/\log(N(t_x)/N(t_0)) \), where \( t_x \) is the duration of treatment, \( N(t_x) \) is protein concentration at time \( t_x \), and \( N(t_0) \) is initial protein concentration, and significance was tested using two-tailed Student t-test, data depicts the mean and standard error. n = 6 derived from three biologically independent samples measured at 3- and 6hrs. h, ChIP-QPCR analysis of changes in H3K27me3 deposition across the Nanog locus between 24- and 36hrs ERK induction, significant increases between 24- and 36hrs are denoted by \( *p = 0.05 \), \( **p = 0.02 \) (one tailed student t-test, data depicts the mean and standard error). n = 3 biologically independent samples. i, A heatmap showing effective suppression of a panel of ERK-induced genes upon treatment in the presence of actinomycin d. Values are log2 transformed and displayed between the range of -5 and 5. n = 3 biologically independent samples. j, A box and whisker plot comparing the mRNA degradation rates of
ERK target genes following a 4h treatment with actinomycin d, with (4OHT) or without (EtOH) ERK activation showing no significant contribution from ERK (two-sided Wilcoxon test). n = 415, data is derived from n = 3; 4OHT, and n = 2; EtOH biologically independent samples. k, PCA of gene expression in cells stimulated with ERK (4OHT) for 6hrs with or without the proteasome inhibitor MG132. PC1 includes the variance caused by proteasomal inhibition, whereas PC2 captures the effect of ERK stimulation, n = 3 biologically independent samples. l, Scaled gene expression values for either repressed genes (top) or induced genes (bottom) that contribute to the variance in PC2 overlaid with the respective eigenvectors from 1k. The correlation threshold was set at 0.7. PC analysis and correlation scores were derived using ExAtlas. Data is derived from 3 biologically independent samples. m, Expression changes in Nanog::EGFP level s in response to ERK stimulation in cells expressing WT KLF2 or KLF4, or their respective putative ERK-regulated phosphorylation-site mutants (serine to alanine). Stable cell lines were derived for each gene and an empty vector served as a negative control. Expression of each protein (WT and SA), was confirmed by western blot analysis (data not shown). A modest, but significant reduction in the magnitude of NANOG::EGFP repression was observed only in cells expressing KLF4S123A at 4hrs (one-tailed Students t-test p=0.02). Data is presented as the mean and standard error. n = 3 biologically independent samples. n, mRNA-seq time-course analysis of the first 4hrs of ERK induction with robust mRNA differences only being observed at 4 hours (top, y-axis represent the number of significantly changing genes), and heatmap of row-scaled values (bottom). n = 3 biologically independent samples.
Extended Data Fig. 2 | ERK regulates RNAPII and TBP binding to control nascent transcription independently of PRC2 activity. 

**a,** Quantification of the changes in intronic reads of the RNA-seq time course shown in Extended Data Fig. 1n showing rapid repression of nascent transcription of a panel of TFs. The curves represent a cubic spline interpolation of the discrete points (using the MATLAB function, `cscvn`). The maximal fold changes are for the 4hr period tested and are taken from the fitted spline curves. n = 3 biologically independent samples. **b,** QRTPCR analysis of nascent (intronic) RNA for the indicated genes in the presence or absence of the translation inhibitor cycloheximide. Cells were pre-treated with cycloheximide (20µg/ml) for 30-minutes prior to 4OHT addition. Data is presented as the mean and standard error. n = 3 biologically independent samples. **c,** Box and whisker plots showing H3K27me3 deposition across gene-bodies of ERK induced or repressed genes following 2hrs ERK stimulation. No significant changes in response to ERK were observed (significance was tested using two-permutation tests with 2000 simulations). n = 62426; global, n = 882; induced, n = 1362; repressed. Data is derived from 2 biologically independent samples. **d,** Immunofluorescence analysis showing H3K27me3 staining in cRAF-ErT2 cells treated for 24hrs with DMSO, or the EZH2 inhibitor EPZ-6438, and subsequently stimulated for 2hrs with 4OHT. Data is representative of 3 biologically independent samples. **e,** QRTPCR analysis of a panel of induced and repressed genes showing little dependency on PRC2 enzymatic activity for either induced or repressed genes other than Myc and cFos both of which showed slightly increased induction in cells depleted of H3K27me3 (two-tailed Students t-test *p = 0.02, **p = 0.001, data is presented as mean and standard error). n = 3 biologically independent samples. **f,** A schematic depicting the strategy to define changes in RNAPII binding. **g,** A scatter plot showing the correlation of RNAPII binding in 227 differentially bound genes. The differentially bound genes were identified by counting the number of RNAPII ChIP reads that uniquely mapped to gene bodies in the 0h, 8h and 8h-2hMeki conditions. Pairwise comparisons, using DESeq2, identified 180 genes with significant changes in RNAPII binding between 8h and 8h-2hMeki and 126 between 8h and 0h. Concatenation of the two gene lists resulted in 227 genes bodies that show a positive correlation in RNAPII binding (Spearman correlation R 0.79, p <2.2e-16) induced by ERK and reversed by inhibition of the pathway. n = 3; 0h, n = 3; 8h, n = 2; 8h:2hMEKi biologically independent samples. **h,** Binding profiles of TBP to the TSSs of genes with changing RNAPII (left), and ChIP-seq tracks for the TSS of both Nanog and Spry4 (right). Data is derived from 2 biologically independent
samples. i, A heatmap showing the changes in nascent RNA expression of ERK-regulated genes associated with super enhancers (abs(log2fc) >1, padj <=0.01). n = 3 biologically independent samples. Super enhancer gene association was performed using GREAT (http://great.stanford.edu/).
**Extended Data Fig. 3 | ERK regulates ESC enhancer activity.** *a,* qRT-PCR analysing the expression of eRNA at pluripotency SEs. FGF4−/− ESCs were stimulated for either 30- or 90m with 40ng/ml rFGF4 prior to analysis. (*p<0.05,**p<0.01, Students t-test, two tailed, data is presented as mean and standard error). n = 6 independent experiments. *b,* A scatter plot showing the changes in RNAP binding and eRNA expression at regulated super enhancers (padj <=0.01), showing a high level of correlation between RNAPII binding and eRNA expression in response to ERK (two-sided Spearman's correlation coefficient). n = 3 biologically independent samples. *c,* A scatter plot showing the changes in H3K27ac and EP300 enrichment at putative enhancers in response to ERK (see materials and methods for further details). ERK activation results in the correlated changes in EP300 binding and H3K27ac deposition at both SEs and TEs (two-sided Spearman's correlation coefficient). n = 636; TE induced, n = 1726; TE repressed, n = 30; repressed SEs, n = 27; activated. Data is derived from 2 biologically independent samples. *d,* A volcano plot showing the expression pattern of ERK-regulated genes associated with ERK-regulated enhancers. n = 3 biologically independent samples. *e,* A stacked bar chart showing the percentage of genes repressed or activated associated with repressed (acDOWN) or activated (acUP) enhancers. 85% of genes associated with repressed enhancers are repressed themselves and similarly 83% of genes associated with activated enhancers are themselves activated. Gene association was performed using GREAT (http://great.stanford.edu/), using single nearest gene cut-off. *f-i,* ATAC profiles at ERK regulated TEs (*f, g*) and SEs (*h, i*) following 2hrs stimulation showing that changes in chromatin accessibility at TEs, but not SEs, is reflective of their activity. n = 3 biologically independent samples. *j,* A scatter plot showing the changes in mRNA and protein levels for a panel of ESC TFs in response to 6hrs ERK signalling. All but Nanog show a degree of persistence into early differentiation. n = 3 biologically independent samples. *k,* A table showing both the mRNA and protein half-lives of a panel of ESC TFs measured by actinomycin d and cycloheximide treatment, respectively. Briefly, cells were treated with actinomycin d (1 μM), or cycloheximide (20 μg/ml) for time periods ranging from 3-6hrs. Samples were collected and processed whereupon the respective half-lives were calculated by $t_x (\log(0.5))/\log(N(t_x)/N(t_0))$, where $t_x$ is the duration of treatment, $N(t_x)$ is protein concentration at time $t_x$ and $N(t_0)$ is initial protein concentration, mean and standard error are presented. n = 3 biologically independent samples. *l,* A bar chart showing the overlap of both Sox2 and Esrrb binding with the list of enhancers defined in (c). Overlap was determined
using the UpSet module from Intervene. Data is derived from two biologically independent samples. **m,** Chow Ruskey diagrams showing overlap between enhancers with changing TF binding. Graphs are scaled to maximal changes at 8hrs for each TF. Data is derived from 2 biologically independent samples.
Extended Data Fig. 4 | Pluripotency TFs remain stably bound to super enhancers, are lost from low affinity binding sites, but are recruited to regions of highly active transcription. a, b, Box and whisker plots showing the changes in TF binding at ERK-regulated SEs at the indicated times. EP300 shows a higher degree of change at 2hrs than do TFs at 2 or 4hrs. * = p-value < 0.01 compared to EP300 (two-sided Wilcoxon test). n = 30; repressed SEs, n = 27; activated SEs. Data is derived from 2 biologically independent samples. c, ChIP-seq tracks of Sox2 and Esrrb at one of the three SE (shown in red) with decreasing binding. Data is representative of 2 biologically independent samples. d, Motif analysis of ERK-regulated enhancers with decreasing TF binding showing no enrichment for any known TF binding motif. n = 334; Sox2, n = 184; Esrrb. Data is derived from 2 biologically independent samples. e, GO analysis of enhancers with increased Esrrb (top) and Sox2 (bottom) binding, GO analysis was performed using GREAT, binomial p values are presented. Data is derived from 2 biologically independent samples. f, A Venn diagram showing a high degree of overlap in the peaks analysed in Figure 3d. g, Box and whisker plots comparing the expression levels (probe intensity) of all ERK-induced genes with genes associated with Esrrb peaks exhibiting increasing binding at 8hrs. Genes that show increased TF binding in response to ERK are the most robustly induced in response to ERK, and the magnitude of this expression is not dependent on Esrrb (two-sided Wilcoxon test [KO:Rescue] p>0.9). n = 3 biologically independent samples.
Extended Data Fig. 5 | AKT/AGC kinase activity is dispensable for transcriptional regulation by ERK. a, A cartoon outlining the experimental setup for SILAC phosphoproteome profiling. See methods for further information on SILAC labelling conditions. b, GO analysis of cluster 2 showing an enrichment in terms associated with transcriptional regulation. FDR values are presented and significance was determined using http://geneontology.org/. Data is derived from 2 biologically independent samples. c, A schematic depicting the PDK1 rescue construct introduced into PDK1−/− ESCs. Stable lines were generated through random integration and all clones were tested for homogenous ERK induction and FLAG-PDK1 expression. d, Western blot analysis of a time course of ERK induction and subsequent inhibition showing ERK-regulated phosphorylation of sites, also co-regulated by AKT, are PDK1 dependent. ERK-mediated phosphorylation of both Gsk3 (S9) and RPS6 (S235/236) are rescued by the re-expression of PDK1. Data is representative of 3 biologically independent samples. e, Correlation (two-tailed Pearson correlation) of gene-expression changes in either PDK1 rescue or KO cells under the indicated conditions showing a high correlation between KO and rescue lines. n = 3 biologically independent samples. f, Scatter plot showing the expression of genes significantly regulated by ERK (abs(log2fc) >1, FDR <=0.01) in PDK1 rescued cells compared with KO. No significant difference could be detected between both samples. n = 3 biologically independent samples. g, The experimental setup to determine the MEK/ERK dependent ESC phosphoproteome. FGF4−/− cells were treated with rFGF4 (40 ng/ml) for 5mins, whereas ERK2AS cells were stimulated for 2hrs. h, An outline of the strategy to generate ERK2-AS cells. ERK2−/− cells were rescued by expressing FLAG-ERK2Q103A from a stably integrated cassette similar to that used for PDK1 rescue depicted in (c). Subclones were screened and subjected to ERK1 mutation by Crispr/CAS9 mutagenesis.
Extended Data Fig. 6 | The kinase activity of ERK, targeting proline-directed serine’s, is essential for its capacity to regulate transcription. a, A cartoon depicting the various ERK alleles in ERK2-AS ESCs. b, A bar chart showing the suppression of Rps6 phosphorylation (S235/236) by treatment with 2 μM PP1 (data is presented as the mean and standard deviation). n = 6 biologically independent samples. c, Western blot analysis for ERK1 in ERK2AS ESCs confirming loss of ERK1 protein. Data is representative of two independent experiments. d, e, Motif analysis of both FGF4 and 4OHT induced sites indicating which are sensitive to inhibition of MEK and ERK, showing the primary motif in sites sensitive to both is, again, a proline directed serine. n = 3; (FGF4)-DMSO, n = 3; (FGF4)-FGF4, n = 3; (FGF4)-MEKi, n = 2; (FGF4)-FGF4/MEKi, n = 3; (ERK2AS)-DMSO, n = 3; (ERK2AS)-4OHT, n = 3; (ERK2AS)-4OHT/PP1, n = 2; (ERK2AS)-4OHT/MEKi, biologically independent samples. f, Box and whisker plots showing the extent to which either MEKi or PP1 treatment can block ERK-dependent transcriptional changes. All sample permutations are significantly different (two-sided Wilcoxon test, p < 2.2e-16). n = 2; 0h DMSO, n = 3; 0h PP1, n = 3; 0h MEKi, n = 3; 2h DMSO, n = 3; 2h PP1, n = 3; 2h MEKi biologically independent samples. g, Quantification of the phosphorylation of the indicated proteins in response to 1hr ERK stimulation in either ERK2+/− cells (EV), or knockout cells rescued with a FLAG-ERK2 construct (data is presented as the mean and standard error). n = 3 biologically independent samples. h, QRTPCR analysis of the expression of the indicated genes following 2hrs ERK stimulation in either ERK2+/− cells (EV), or knockout cells rescued with a FLAG-ERK2 construct (data is presented as the mean and standard error). n = 3 biologically independent samples. i, Venn diagram showing the overlap between Esrrb interacting proteins, with cluster 2.
**Extended Data Fig. 7 | MED24 is a key target of ERK required for normal transcriptional responses.**

**a,** ERK robustly phosphorylates the Mediator component, MED24 on two serine’s in the C-terminal end of the protein directly adjacent to a hormone-receptor interaction domain, values taken from (Figure 4a). Data is presented as the mean. n = 2 biologically independent samples. **b,** Expression values for Med24 at the indicated developmental stages from ref44. **c,** A schematic outlining the strategy for making Med24 conditional ESCs and where sgRNA guides are targeted in the endogenous locus. Med24 expression is maintained by a doxycycline-inducible promoter, where stable integration of the cassette can be selected for by either neomycin (top), or hygromycin (bottom). **d, e,** Western blot analysis of the expression levels of MED24 in WT ESCs expressing either construct depicted in (c) prior to deletion of the endogenous MED24. Selection of exo-Med24 with neomycin (d) gives slightly higher expression levels than selection with hygromycin (e). Data is representative of two independent experiments. All mutant phenotypes were reproduced in knock out cell lines rescued with either construct. ChIP-seq and transcriptome analysis was performed using cells expressing the neomycin rescue. **f,** Western blot analysis confirming disruption of endogenous MED24 from hygromycin selected cells. C14, 15, 2 and 5 are homozygous mutant cell lines. Data is representative of two independent experiments. **g,** Western blot analysis confirming disruption of endogenous MED24 from neomycin selected cells. WT1, 3, 5 and 6 are homozygous mutant cell lines. Data is representative of two independent experiments. **h,** GO analysis for MED24-bound regions showing that bound regions are associated with genes involved in peri-implantation development. MED24 peaks were located within TEs (from Extended Data Figure 3c) and ontologically annotated with GREAT. n = 3 biologically independent samples. **i,** Motif analysis for MED24 binding regions within repressed and induced enhancers showing enrichment for Esrrb (repressed) and ETS/AP1-factors (activated) motifs. n = 636; activated and 1726; repressed. **j,** Quantification of the levels of NANOG::EGFP expression upon the deletion of MED24 (data is presented as the mean and standard error). n = 6 biologically independent samples. **k,** ChIP-QPCR analysis of RNAPII binding to the Nanog enhancer in either WT or KO cells MED24, enrichment at an unbound region between the proximal enhancer and TSS serves as a negative control (data is presented as the mean and standard error). n = 3 biologically independent samples. **l,** PCA of the nascent transcriptional response in WT and MED 24 KO samples following 2hrs ERK stimulation. n = 3 biologically independent samples.
Extended Data Fig. 8 | Mediator binding is regulated by ERK and is not required for neural specification. a, Expression levels of eRNA expressed from the Klf4 or Spry4 SEs in MED24 WT or KO cells following 2hrs ERK activity (data is presented as the mean and standard error, \(^*p = 0.03, **p = 0.006,\) one-tailed Students t-test). Reads were normalised to housekeeping genes. \(n = 3\) biologically independent samples. b, Co-localisation analysis of ChIP peaks for ESRRB, MED1 and MED24 showing MED1 is rarely found with Esrrb in absence of MED24. c, A stacked bar chart showing the percentage of all potential enhancers with changes in EP300, MED1 or MED24 following 2hrs ERK induction. Changes in co-factor binding range from 10 to 20% of all putative enhancers, indicating that the regulation of binding is selective. Data is derived from 2 biologically independent samples. d, Box and whisker plots showing the binding of Mediator components to repressed and activated TEs and SEs. The magnitude of change relative to EP300 is significantly different at SEs (two-sided Wilcoxon test). \(n = 1726;\) repressed TEs, \(n = 636;\) activated TEs, \(n = 231;\) all SEs, \(n = 30;\) repressed SEs, \(n = 27;\) activated SEs. Data is derived from two biologically independent samples. e, f, Chow-Ruskey diagrams showing the relative loss (e), or gain (f) of co-factor components from ERK-regulated enhancers. Data is derived from 2 biologically independent samples. g, h, Motif analysis on regions defined in (e, f) showing regions enriched in Esrrb and ETS/AP1 motifs are regions where both Mediator and EP300 binding changes in response to ERK. A similar, but less dramatic effect is observed for canonical pluripotency binding motifs that are also enriched within these regions. At activated enhancers, the majority of regions with EP300 binding show changes in Mediator, whereas at repressed enhancers, there is an appreciable number of changes in EP300 binding that appear to occur in the absence of Mediator dissociation. \(n = 367;\) Med/EP300 activated, \(n = 73;\) EP300 activated, \(n = 261;\) Med/EP300 repressed, \(n = 903;\) Med/EP300 repressed. i, Flow cytometry analysis of MED24 WT or KO cells at day 5 neural differentiation for the cell surface markers PDGFRA and NANOG::EGFP. Data is representative of 3 biologically independent samples. j, Immunofluorescence of MED24 WT or KO cells following 5 days of neural differentiation for the early neural marker Nestin. Data is representative of 3 biologically independent samples.
Extended Data Fig. 9 | MED24 is specifically required for PrE specification, but not its survival or expansion. a, QRTPCR analysis of ESC and neural markers following 4 days of neural differentiation in WT and MED24 KO cells. Data is presented as the mean and standard error. n = 3 biologically independent samples. b, Flow cytometry analysis of MED24 WT or KO cells at day 5 of PrE differentiation for the cell-surface markers PDGRFA (PrE) and PECAM1 (ESC). Data is representative of 3 biologically independent samples. c, Quantification of flow cytometry analysis for the PrE-cell-surface marker PDGFRA staining of cells following 5 days PrE differentiation (data is presented as the mean and standard error, *p = 0.01, one tailed Students t-test). n = 3 biologically independent samples. d, e, QRTPCR for either ESC or PrE-markers showing both a significant failure to repress pluripotent transcription, and conversely to induce PrE-markers in MED24 mutant (data is presented as the mean and standard error, *p = 0.03, **p = 0.02, ***p = 0.0004, ****p = 0.00002, two-tailed Students t-test). n = 3 biologically independent samples. f, Quantification of the numbers of Ki67+ve nEND cells with and without MED24. Cells were differentiated towards stable and homogenous endoderm culture whereupon doxycycline was removed to deplete cells of MED24. Analysis was preformed after 48hrs and no detectable difference in the number of proliferating cells between both conditions was observed (data is presented as the mean and standard error). n = 3 biologically independent samples. g, As in (f) except analysis of cleaved caspase 3 as a marker of apoptosis (data is presented as the mean and standard error). n = 3 biologically independent samples. h, QRTPCR analysis of the ERK target genes EGR3 and Fosl2 following 2hrs activity in Med24 WT or Med24 SA mutant cells. A slightly attenuated response was observed in the induction of two of the most MED24 dependent genes from Fig. 5a when both serine’s 860 and 871 were mutated to alanine (data is presented as the mean and standard error, *p = 0.04, **p = 0.01, two-tailed Students t-test). n = 3 biologically independent samples. i, Quantification of flow cytometry analysis for the PrE cell-surface marker PDGFRA at day 5 of differentiation in KO cells expressing either WT MED24 or MED24 (S829A, S840A). n = 3; WT, n = 2; SA biologically independent samples.
Extended Data Fig. 10 | A model for PrE-priming involving the stable association of TFs alongside the reversible reallocation of co-factors and the RNAPII complex. A cartoon describing the sequence leading to ERK-dependent, PrE commitment. ERK activation results in the rapid, but reversible, disassociation or association of RNAPII and cofactors from repressed and induced genes, respectively. If the signal is maintained to a point where the activating TFs repressed by ERK (specifically pluripotency factors) drop below a minimum threshold, the ability to reform a functionally active transcription complex is lost and terminal differentiation occurs.
References


Methods

Data Availability

The microarray, ChIP-seq and RNA-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through the GEO Series accession number: GSE132444.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.protemexchange.org) via the PRIDE partner repository with the dataset identifiers PXD008964 and PXD012573.

ESC culture

ESCs were maintained in complete ESC medium: GMEM (Sigma) supplemented with 10% FBS (Gibco), 100 μM 2-mercaptoethanol (Sigma), 1xMEM non-essential amino acids, 2 mM L-glutamine, 1 mM sodium pyruvate (all from Gibco), 1000 units/ml LIF (made in house) and unless otherwise stated 1.6μM Gsk3i (Chir99021: Axon Medchem) on gelatinised tissue culture flasks (Corning). For serum-free culture, ESC lines were maintained in N2B27 supplemented with 3 μM Gsk3i and 1 μM MEKi (PD0325901: Sigma) and LIF, and passaged with Accutase (Sigma). Both FGFR inhibitor (PD173074, FGFRi) and (Z)-4-Hydroxytamoxifen (4OHT), were used at 250 nM and were from Sigma. 1-NM-PP1 (Cayman Chemicals) was used at 2 μM, The EZH2 inhibitor EPZ-6438 was used at 2 μM. Doxycycline (Sigma) was used at 500 ng/ml. Recombinant human Fgf4 (R&D Systems) was used at 40 nM, Activin A (ProproTech) was used at 20 ng/mL. Neural monolayer differentiation was performed as outlined in ref32) and PrE differentiation was performed as outlined in ref25. SILAC labelling was performed as in ref33. For ERK induction, ESCs were plated 24hrs prior to induction in serum and LIF containing medium supplemented with 250 nM PD173074 at a density of 25x10^3 per cm^2. The following day 4OHT was added directly to the medium to a final concentration of 250 nM. For reversion experiments, MEKi (PD0325901) was added directly to the induction medium to a final concentration of 1 μM.

QRTPCR

Total RNA was collected using the RNeasy Mini Kit (Qiagen). Genomic DNA was eliminated by DNase treatment (Qiagen) and 1 μg of total RNA was used for first strand synthesis using SuperScript III reverse transcriptase according to the manufacturer’s instructions. cDNA
corresponding to 10 ng total RNA was used for QRTPCR analysis using the Roche LC480 and target amplification was detected with the Universal Probe Library system. See supplementary table 22 for a list of primers and probes used.

**Western blotting**

Blotting was performed as previously described in ref\(^7\), primary antibodies were detected using fluorescently conjugated secondary antibodies (AlexaFlour, Molecular Probes), visualized using a Chemidoc MP (Biorad) and quantified using Fiji. See supplementary table 23 for a list of antibodies and concentrations used, and supplemental figure 1 for all uncropped western blot images.

**Immunofluorescence**

Cells were washed and fixed in 4% formaldehyde (FA) at room temperature for 15 minutes (Fisher Scientific, PI-28906), followed by a 10 minute incubation in ice cold methanol at -20 °C. Samples were subsequently blocked and permeabilized in 5% donkey serum and 0.3% Triton. Antibodies were incubated overnight in 1% BSA, 0.3% Triton in PBS, and subsequently visualized with the appropriate secondary (AlexaFluor, molecular probes). See supplementary table 23 for a list of antibodies and concentrations used.

**Flow Cytometry**

Cells were dissociated with Accutase and incubated with the appropriate antibody in 10% FCS-PBS for 30 minutes, washed extensively and analysed on an LSRFortessa (BD Biosciences). Dead cells were excluded based on DAPI inclusion. A gating strategy example can be found in supplemental figure 2.

**Microarray Analysis**

Transcriptome analysis was carried out in-house using one colour 8x60K Mouse Gene Expression Microarrays (Agilent Technologies- G4852A) as per manufacturer’s instructions. Briefly, 150 ng of total RNA was labelled using the Low Input Quick Amp Labelling Kit (Agilent Technologies- 5190-2305). Labelled samples were hybridized overnight, and then washed and scanned using the high sensitivity protocol (AgilentG3_HiSen_GX_1color) on a SureScan microarray scanner (Agilent Technologies) and probe intensities were obtained by taking the
gProcessedSignal from the output of Agilent feature extraction software using default settings. Probe intensities were analysed using ExAtlas\textsuperscript{34}. Significant genes were determined using an ANOVA FDR cut-off of 0.01.

**RNA-Seq**

Total RNA was purified by standard methods and rRNA depleted using the Ribo-Zero kit (Illumina, as per manufacturer’s instructions). Libraries were prepared for Illumina sequencing using the NEBNext\textsuperscript{®} Ultra\textsuperscript{™} kit as per manufacturer’s instructions. Barcoded libraries were pooled and run on two lanes (for improved coverage) on a HiSeq 2500 Sequencer (Illumina). Sequencing reads (60 bases) were aligned using the STAR package\textsuperscript{35}. Allocation of reads at introns (and exons) were examined using Table Browser (UCSC) to define the corresponding genomic intervals. Reads per gene per class were counted using HTSeq, using the categorised alignment files as input. To then calculate differences between time points (0h, 1h, 2h, 4h after ERK induction), we utilised the Bioconductor package, DESeq2\textsuperscript{36}. EU-seq was performed with the Click-iT™ Nascent RNA Capture Kit (Thermo Scientific) as per the manufacturer’s instructions. Cells were labelled with 5 mM 5-ethynyl uridine for 15 minutes prior to sample collection. RNA-seq libraries were prepared on-bead NEBNext\textsuperscript{®} Ultra™ kit as per manufacturer’s instructions and subsequently sequenced using a Next-Seq 500 Sequencer (Illumina). RNA-seq was analysed using the DESeq2 package. Genes were considered scientifically regulated if they exhibited an abs(log2fc) of >=1, and an adjusted p value of <=0.01 (Wald’s test p-values corrected by using the Benjamini-Hochberg procedure). eRNA expression was calculated by counting reads per region defined as SEs in ref\textsuperscript{15}, with regions +/- 1kb from a coding gene subtracted from the analysis to avoid confounding nascent gene expression with RNA emanating from intergenic enhancers.

**ChIP-Seq**

Chromatin was crosslinked for 10 minutes in 1% FA (single fixation), or with at subsequent 45-minute fixation with 2mM Di(N-succinimidyl) glutarate (double fixation) and stopped in glycine. Nuclei were extracted in 50mM HEPES KOH, pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, and subsequently sonicated in 10mM Tris pH 8, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% sodium deoxycholate, 0.5% N-lauryl sarcosine. Immunoprecipitation was performed against 50ug of chromatin using 10ug of antibody, in
sonication buffer supplemented with 1% Triton and antibody-chromatin complexes were collected with the appropriate magnetic bead and extensively washed in 50mM Hepes (pH 7.6), 1mM EDTA, 0.5M LiCl, 0.7% sodium deoxycholate, 1% NP-40. For replicate ChIP experiments all antibody incubations, immunoprecipitations and washes were performed in parallel to reduce technical variation. Bound DNA was eluted, purified and subjected to either QRTPCR or prepared for Illumina sequencing using the NEBNext Ultra™ II DNA Library Prep Kit as per manufacturer’s instructions. Libraries were subjected to size-selection and fragments between 400- and 200bp were sequenced in house on a Next-Seq 500 Sequencer (Illumina). Raw sequence reads were aligned to the mm10 reference genome (with Bowtie2\textsuperscript{37}), reads were quality filtered and duplicates removed prior to analysis. Peak calling was performed with \textit{macs2}\textsuperscript{38}, with input samples as controls using q 0.01. Reads per regions were calculated with \textit{coverage} (\textit{bedtools}\textsuperscript{39}). Count tables were analysed using the \textit{DESeq2} package. Genes/genomic regions were considered significantly regulated if they exhibited an abs(log2fc) of >=1, and an adjusted p value of <=0.01 (Wald’s test p-values corrected by using the Benjamini-Hochberg procedure). See supplementary table 23 for a list of antibodies and concentrations used. To generate the ESC enhancer dataset for subsequent analysis raw data from EP300 ChIP\textsuperscript{40} and H3K27ac ChIP-seq from 0h and 2h ERK treatment was processed as described above, with regions containing both EP300 and H3K27ac peaks considered putative enhancers. Motif analysis was performed using HOMER\textsuperscript{41} specifying motif size between 7 and 10 nucleotides.

**Mass Spectrometry**

Cells were grown SILAC DMEM (PAA Laboratories GmbH, Germany) lacking arginine and lysine supplemented with 10% dialyzed fetal bovine serum and all other standard culture components supplemented with either light amino acids (Sigma), with medium and heavy variants from Cambridge Isotope Labs (Massachusetts, US). Samples were lysed in modified RIPA buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1% NP-40, 0.1% sodium deoxycholate, 1 mM EDTA, 5 mM β-glycerolphosphate, 5 mM sodium fluoride, 1 mM sodium orthovanadate) supplemented with 1 complete® protease inhibitor cocktail tablet (Roche, Basel, Switzerland) and N-ethylmaleimide (Sigma). Three cell populations were obtained labelled with light (Lys0, Arg0), medium (L-[13C6] Arg (+6) and L-[2H4]Lys (+4), and heavy(L-[13C6,15N4]Arg (+10) and L- [13C6,15N2]Lys (+8) ) amino acids. 20 mg digested protein was incubated with titanium
dioxide (TiO2) beads (GL Sciences, Japan) washed and bound phosphopeptides were eluted directly into a 96-well plate by 5% NH4OH followed by 10% NH4OH, 25% acetonitrile. Full-scan MS spectra (m/z 300-1,750, resolution 70,000 at m/z 200) were detected in the Orbitrap analyser and raw data was analysed by the MaxQuant software suite with the integrated Andromeda search engine. Only peptides with site localization probability of at least 0.75 (class I) were included in the bioinformatics analyses. Each time point was considered an independent experiment and we defined a site regulated, if its SILAC ratio was higher than 2-fold or lower than 0.5-fold in at least one experimental condition. The resulting data was clustered by the fuzzy c-means algorithm in GProX 1.1.12., by requesting six clusters with a fuzzification parameter of 2 and 100 algorithm iterations. For identification of potential kinase motifs, the sequence window of the regulated phosphorylation sites was compared to the non-regulated sites using the IceLogo resource with default parameters.

For TMT-multiplexed analysis, cells were washed in PBS and lysed in 6 M guanidine-HCl, 100 mM Tris pH 8.5, 5 mM TCEP and 10 mM CAA for 10 min at 99°C and whole cell extracts were sonicated. Cell lysates were digested by Lys-C (Wako) in an enzyme/protein ratio of 1:300 (w/w) for 1 hour, followed by a dilution with 25 mM Tris pH 8.5, to 2 M guanidine-HCl and further digested overnight with trypsin (Sigma-Aldrich) 1:100 (w/w). Protease activity were quenched by acidification with TFA and the resulting peptide mixture was concentrated on C18 Sep-Pak (Waters). Peptides were eluted with 40% ACN, followed by 60% ACN. 300 micrograms of peptide from each sample was labelled with 30 μl of one of 11 different TMT-labels according to the manufacturer’s protocol (ThermoScientific). After labelling, the 11 samples were combined and adjusted to 80% ACN, 6% TFA and phosphopeptides were further enriched by two sequential rounds of titansphere chromatography as described in ref^42. The eluted phosphopeptides were concentrated in a SpeedVac and fractionated with high-pH-reversed-phase fractionation^42 and the resulting 12 fractions of TMT11-plex labelled peptide mixture were analysed by online nanoflow LC-MS/MS on an EASY-nLC system coupled to a Q-Exactive HF-X instrument (Thermo Fisher Scientific) equipped with a nanoelectrospray source. MS raw files were analysed using MaxQuant software version 1.6.0.17 using the Andromeda search engine. Subsequent data analysis was performed using Perseus version 1.6.2.2. For each TMT 11-plex phosphoproteome, all measured peptide intensities were normalized using the “normalizeQuantiles” function from the Bioconductor R package LIMMA^43 and further
normalized by median subtraction across rows within the same experimental replicate, and the data were filtered for contaminants and reverse hits. Only peptides with a phosphorylation site localization probability of at least 0.75 (class 1) were included in the bioinformatic analyses. For each TMT multiplex analysis significant regulation was deemed based on ANOVA testing (FDR: 0.05 and s0=0.1).
Additional references


Klf2

all repressed enhancers (1726)

all activated enhancers (636)

ESRRB

SOX2

Arc

all repressed enhancers (1726) all activated enhancers (636)

ESRRB

SOX2

Arc

0h 2h 4h 8h

all induced genes super enhancer driven

JUN−AP1 FOSL2 FRA2 FRA1 JUNB ATF3 BATF ESRRB SOX2

SOX2 binding (-log P)

ESRRB binding (-log P)

0 10 20 30 40

log2 (norm. probe intensity after reversion)

EP300 (69.1%)

ESRRB (4.8%)

SOX2 (5.9%)

EP300 (49.7%)

ESRRB (0.4%)

SOX2 (0.4%)

2h 4h 8h

% of enhancers with changing binding

activated
driven

repessed

0h 2h 4h 8h

all induced genes super enhancer driven

0 10 20 30 40
Fgf4 induced and MEK inhibited
n = 784
Raf-induced and PP1/MEK inhibited
n = 3,478

PC1: 82% variance
PC2: 8% variance

SILAC cluster 1-6
Phosphoproteins: n=579
Includes Med24
pS871 (TMT)
pS860 and pS871 (SILAC)

926
116
158
231
11
294
92
FGF4 regulated super enhancers

* * *

TE: R=0.84
SE: R=0.83

Repressed traditional enhancers
Activated traditional enhancers

Repressed super enhancers
Activated super enhancers

No. of Intersections

Enhancers
SOX2
ESRRB

ESRRB
SOX2

Activated
Repressed
Activated
Repressed

Protein
mRNA

<table>
<thead>
<tr>
<th>Protein</th>
<th>mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nanog</td>
<td>1.5(±0.1)</td>
</tr>
<tr>
<td>Esrrb</td>
<td>2.1(±0.4)</td>
</tr>
<tr>
<td>Klf2</td>
<td>4.8(±0.5)</td>
</tr>
<tr>
<td>Tfcp211</td>
<td>2.0(±0.1)</td>
</tr>
<tr>
<td>Klf4</td>
<td>7.7(±1.2)</td>
</tr>
<tr>
<td>Oct4</td>
<td>9.1(±0.7)</td>
</tr>
<tr>
<td>Sox2</td>
<td>10(±2.2)</td>
</tr>
</tbody>
</table>
Activated super enhancers

Repressed super enhancers

Motifs enriched at regions decreasing in TF binding

GO Biological Process

log_{10}(probe intensity)

log_{10}(probe intensity)

log_{10}(probe intensity)
**Lys0, Arg0**

“Light”

**Lys4, Arg6**

“Medium”

**Lys8, Arg10**

“Heavy”

0m

30m

120m

**Nucleic acid binding**

Poly(A) RN

A binding RN

**Heterocyclic compound binding**

Organic cyclic compound binding

**Protein binding**

DN A binding

**Transcription factor activity**, transcription factor binding

**Transcription coactivator activity**

Chromatin binding

**Spearman correlation = 0.956**

Sample 1  
Sample 2  
Sample 3  
Sample 4  
Sample 5  
Sample 6  
Sample 7  
Sample 8  
Sample 9  
Sample 10  
Sample 11

**FGF4+MEKi**

**ERK**

**AS +MEKi/PP1**

Isolate proteins

6 M GndHCl

Reduce

Alkylate Digest

Label

TMT

Combine

Fractionate and perform LC-MS/MS

**Erk2 KO ES cells**

Rescue with **Erk2**

Knock out **Erk1** with CRISPR/Cas9
Cluster 2 (236)

Motif overrepresentation in clusters

FGF4 induced and MEK inhibited
cRAF-ER

T2-induced and PP1/MEK inhibited

P

K

A

G

p

r

o
t

i

c

u

d

r

e

k

i

n

a

s

1

E

R

K

1

, 2

P

r

o

l

i

d

i

d

i

e

r

c

t

e

d

k

i

n

a

s

G

S

-3

E

R

K

1

, E

R

K

2

, C

D

K

5

F

R

P

o

l

i

r

e

d

i

e

n

d

i

d

i

e

r

c

t

e

d

k

i

n

a

s

I

I

E

R

K

1

, 2

P

K

C

ki

n

a

s

G

S

-3

E

R

K

1

, E

R

K

2

, C

D

K

5

P

r

o

l

i

d

i

d

i

e

r

c

t

e

d

k

i

n

a
**Figure Legends**

**Figure a** shows a bar graph representing log2 (rel. expression) of various genes: Pou5F1, Nanog, Klf4, NeuroD1, and NeuroD3. The graph compares WT and KO conditions for days 0 and 5.

**Figure b** illustrates a scatter plot of PDGFRA expression with a significant increase in KO condition compared to WT.

**Figure c** displays a bar graph showing the increase in %PDGFRA +ve in the KO condition compared to WT.

**Figure d** and **e** present bar graphs for Pluripotency and Primitive endoderm, respectively, comparing log2 (rel. expression) of Nanog, Klf2, and Tbx3.

**Figure f** and **g** depict bar graphs for nEnd and Egr3, respectively, indicating cleaved-uncleaved ratio and log2 (rel. expression) over time.

**Figure h** and **i** show bar graphs for Fosl2 and MED24 phospho-mutant, respectively, indicating % PDGFRA+ve with time.