

ERF deletion rescues RAS deficiency in mouse embryonic stem cells

Cristina Mayor-Ruiz, Teresa Olbrich, Matthias Drosten, Emilio Lecona, Maria Vega-Sendino, Sagrario Ortega, Orlando Dominguez, Mariano Barbacid, Sergio Ruiz and Oscar Fernandez-Capetillo

SUPPLEMENTAL INFORMATION

Supplemental Inventory:

Supplemental Figures and Figure Legends:

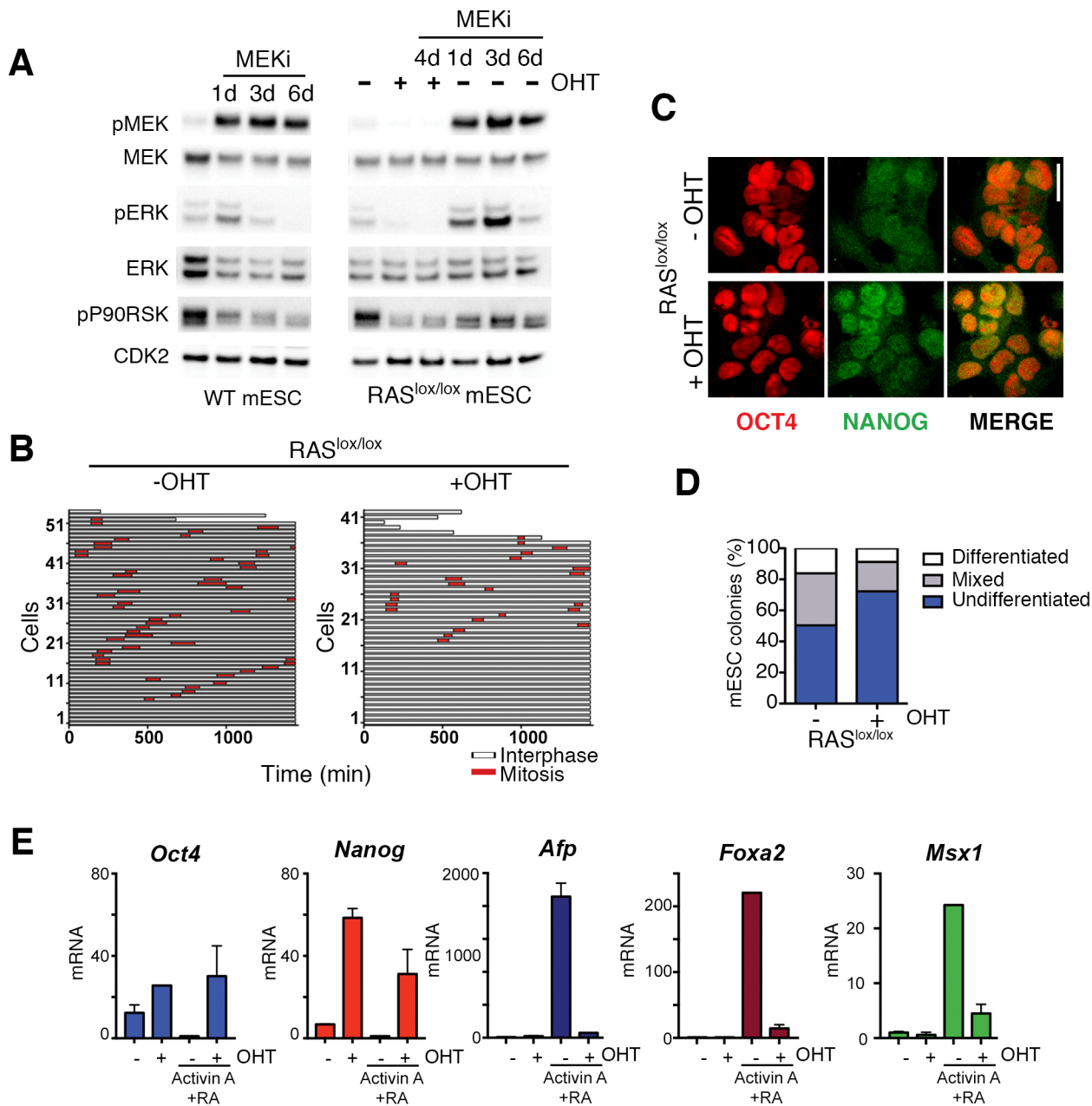
Supplemental Figure S1 – related to Figure 1
Supplemental Figure S2 – related to Figure 1
Supplemental Figure S3 – related to Figure 2
Supplemental Figure S4 – related to Figure 2
Supplemental Figure S5 – related to Figure 3
Supplemental Figure S6 – related to Figure 4

Supplemental Materials and Methods

Supplemental References

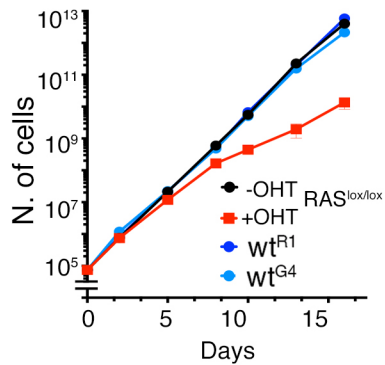
Supplemental Tables S1-7 (separate excel file)

Supplemental Movie S1 (separate movie file)

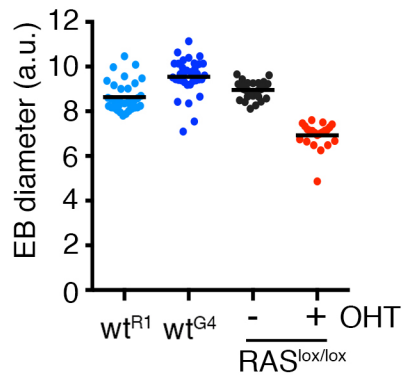


Supplemental Figure 1, related to Figure 1. (A) WB analysis of MEK, ERK and P90RSK phosphorylation levels in WT (G4, left) or RAS^{lox/lox} mESC (right) untreated or treated with OHT or MEKi (PD0325901) for the indicated times. CDK2 is added as a loading control. Note that RAS deficiency has a more profound impact than MEK inhibition in reducing the phosphorylation of these targets. (B) Time spent in mitosis (red) or interphase (white) of individual untreated or OHT-treated RAS^{lox/lox} mESC 7 days after the addition of OHT. Aynschronously growing cells infected with an H2B-EGFP expression virus were imaged every 10 minutes for a total of 24 hours. Time in mitosis was defined as the time between chromosome condensation and cytokinesis. (C) OCT4 (red) and NANOG (green) immunofluorescence signal in untreated or OHT-treated RAS^{lox/lox} mESC 7 days after the addition of OHT. Scale bar (white) indicates 5 μ m. (D) Percentage of Alkaline Phosphatase (AP) positive mESC colonies classified as differentiated, mixed or undifferentiated based on the level of AP activity. 500 individual untreated or OHT-treated RAS^{lox/lox} mESC were plated on gelatin-coated plates, fixed and stained for AP activity. At least, 150 individual colonies were scored 7 days after plating. Two independent experiments were performed using two different mESC lines and analyzed by triplicate. (E) mRNA levels of pluripotent (*Oct4* and *Nanog*) and differentiation (*Afp*, *Foxa2* and *Msx1*) markers in untreated or OHT-treated RAS^{lox/lox} mESC that were induced to differentiate into the mesendoderm lineage with Activin A and Retinoic Acid (RA), as assessed by real-time PCR. *Gapdh* levels were used to normalize gene expression. Two independent experiments were performed and data is shown as the averaged level of triplicates. Error bars indicate SD.

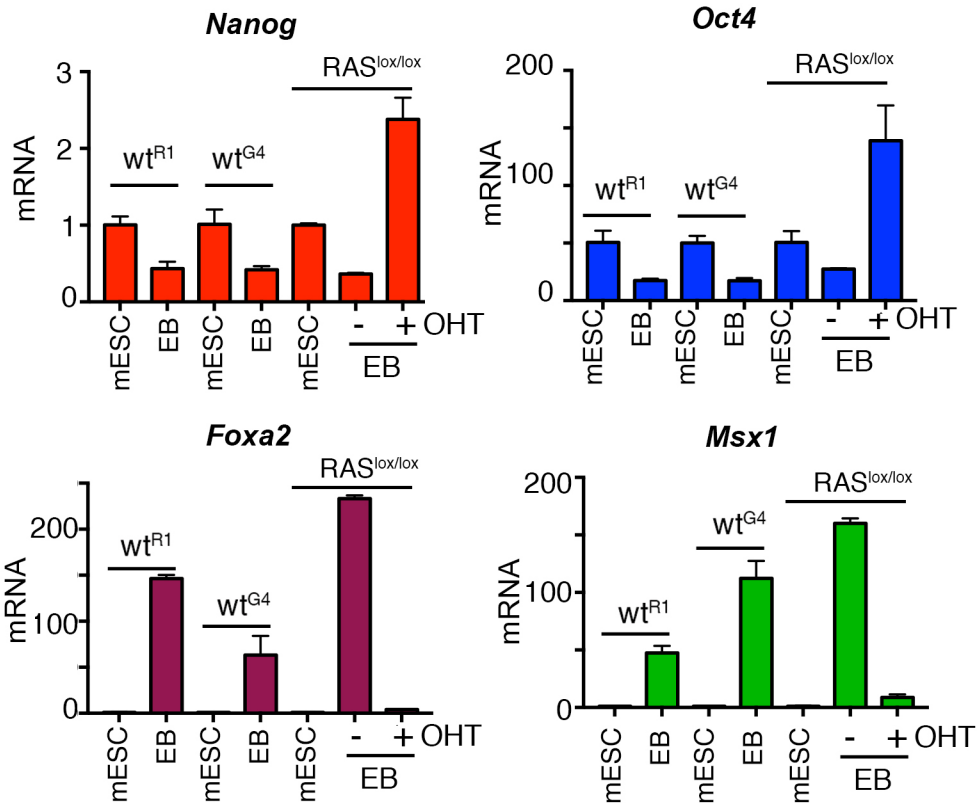
A



B

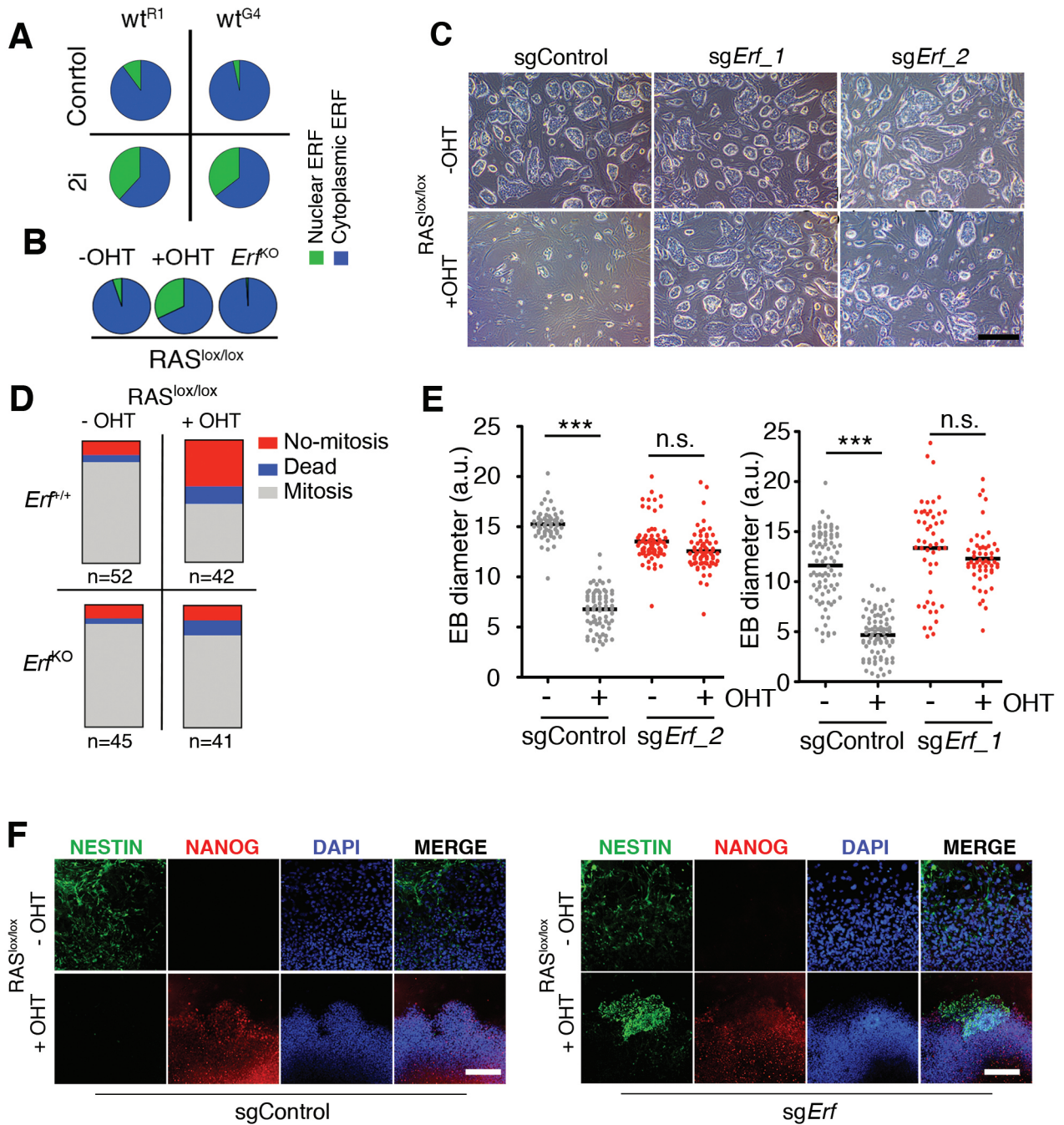


C

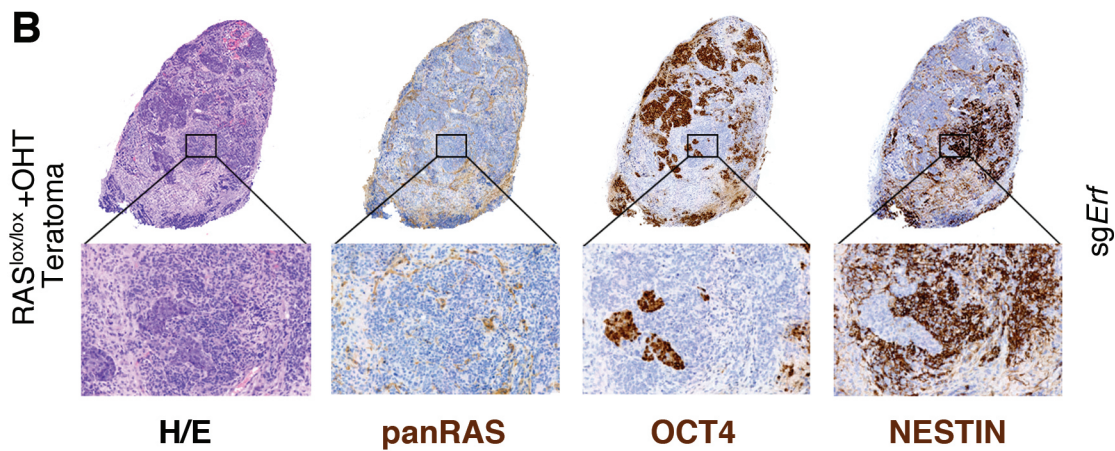
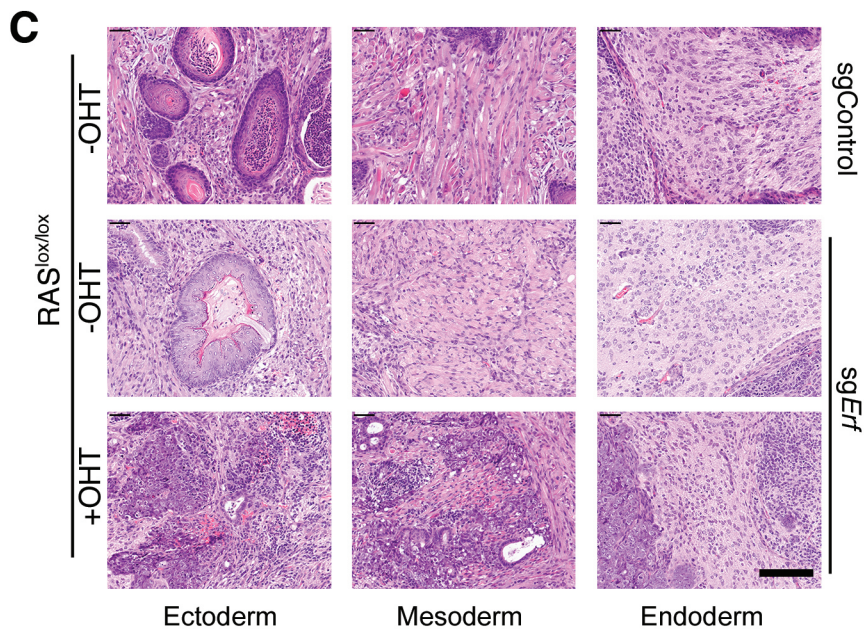
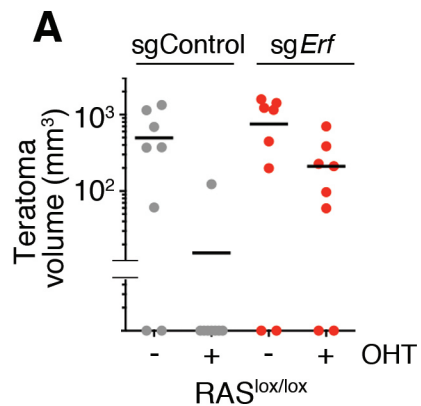


Supplemental Figure 2, related to Figure 1. (A) Cumulative growth curves from cultures of 2 independent wt mESC lines (R1 and G4), as well as from untreated or OHT-treated RAS^{lox/lox} mESC, indicating that the absence of H- and K-Ras in RAS^{lox/lox} mESC does not significantly affect their growth. (B) Diameters of 3-day old EBs derived from 2 independent wt mESC lines (R1 and G4), as well as from untreated or OHT-treated RAS^{lox/lox} mESC. (C) (F) mRNA levels of pluripotent (*Oct4* and *Nanog*) and differentiation (*Foxa2* and *Msx1*) markers in wild type mESC and in EBs derived from 2 independent wt mESC lines (R1 and G4), as well as from untreated or OHT-treated RAS^{lox/lox} mESC as assessed by real-time PCR. *Gapdh* levels were used to normalize gene expression. Error bars indicate SD.

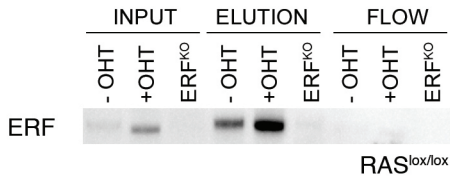
Supplemental Figure 3, related to Figure 2. (A) ERF mRNA expression in mouse tissues obtained from data available at the BioGPS bioportal (probeset: 1435561_at) (<http://ds.biogps.org/?dataset=GSE10246&gene=13875>). (B) ERF mRNA expression obtained from (Boroviak et al. 2015), where the authors performed transcriptome profiling across various stages of early embryonic development (MOR: morula; ICM: inner-cell mass; EPI: epiblast; PrE: Primitive Endoderm).



Supplemental Figure 4, related to Figure 2. (A,B) Pie charts illustrating the percentage of nuclear (green) or cytoplasmic (blue) ERF, as quantified by High-Throughput Microscopy (HTM), that is found in wt mESC (R1 and G4) treated with 2i for 2 hrs (A) or in RAS^{lox/lox} mESC untreated or treated with OHT (96 hrs). An ERF-deficient (*Erf*^{KO}) mESC line was used as an antibody control for these experiments. The experiments were performed three times and a representative one is shown. (C) Representative bright-field images from cultures to untreated or OHT-treated RAS^{lox/lox} mESC infected with lentiviruses expressing Cas9 and sgRNAs against ERF (*sgErf_1* and *sgErf_2*) or a control sgRNA, 7 days after the addition of OHT. Scale bar (black) indicates 20 μ m. This experiment was performed 5 times and a representative example is shown. (D) Quantification of the time spent in mitosis or interphase of individual untreated or OHT-treated *Erf*^{KO}/RAS^{lox/lox} mESC 7 days after the addition of OHT. Aynchronously growing H2B-EGFP transfected cells were imaged every 10 minutes for a total of 24 hours. Time in mitosis was defined as the time between chromosome condensation and cytokinesis. The graphs indicate the percentage of cells that during the course of the 24 hr video underwent mitosis (grey), died (blue), or failed to enter mitosis (red). Note the increase in cells that fail to enter mitosis in the absence of all RAS genes, which is rescued by ERF loss. (E) Quantification of EB diameters from the experiment described in Fig. 2D. Data from two independent experiments using a different *Erf* targeting sgRNA are shown. At least, a total of 30 EBs was scored for each condition. a.u.: arbitrary units. n.s.: non significant; *** $p < 0.001$. (F) Immunofluorescence of NESTIN (green) and NANOG (red) in 11-day old EBs derived from untreated or OHT-treated RAS^{lox/lox} mES cells infected with lentiviruses expressing Cas9 and sgRNAs against *Erf* or a control sgRNA and that were subsequently plated on gelatin to induce spontaneous differentiation (as shown in Figure 2E). DAPI was used to stain DNA. Scale bar (white) indicates 50 μ m.



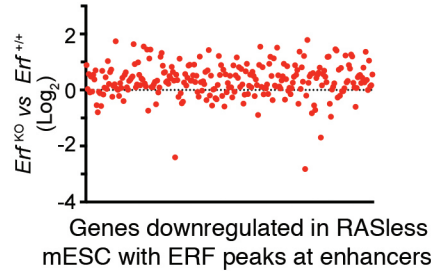
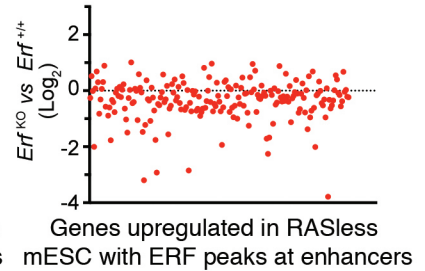
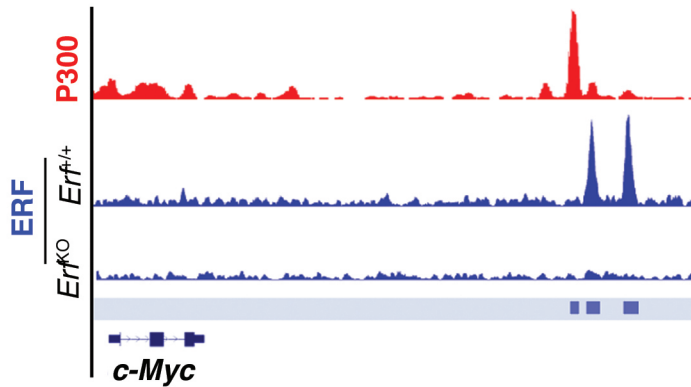
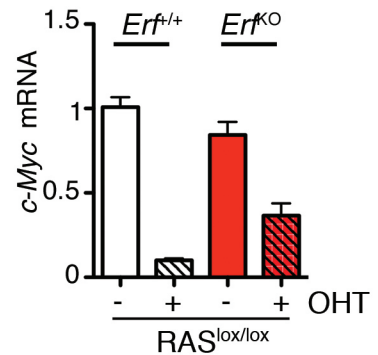
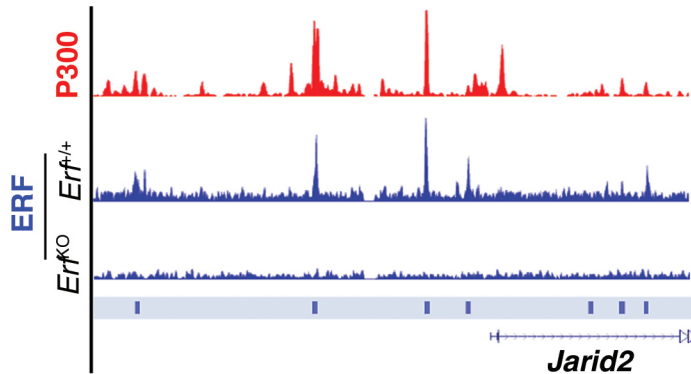
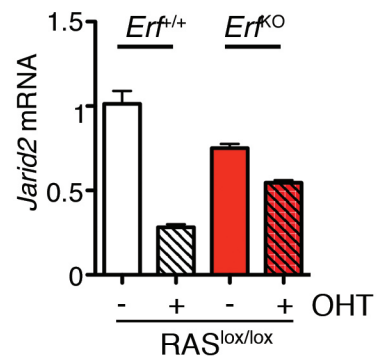
Supplemental Figure 5, related to Figure 3. (A) Teratoma tumor volumes quantified from the experiment shown in Figure 4C. Note that only one teratoma could be obtained from OHT-treated and *Erf*^{+/+} RAS^{lox/lox} mESC, which was a chimera of RAS proficient and deficient cells (see Figure 4E). (B) Immunohistochemistry of panRAS, OCT4 and NESTIN expression in a teratoma obtained from OHT-treated RAS^{lox/lox} mESC that had been infected with a *Erf*-targeting sgRNA. An example region lacking panRAS staining that had lost OCT4 expression and to expressed NESTIN is shown (zoomed-in square). (C) Representative images from a Hematoxylin/Eosin staining of the teratomas of the indicated genotypes, illustrating the presence of areas with differentiation towards ectoderm, mesoderm and endoderm in each case. Scale bar (black) indicates 100 μ m.

A

B
GREAT 3.0 ontology Analysis

GO Biological Process (Term name)	Binomial FDR P-value
Stem cell maintenance	6.11 x 10 ⁻¹⁰
Regulation of peptidyl-threonine phosphorylation	7.95 x 10 ⁻⁹
Blastocyst formation	1.13 x 10 ⁻⁷
Covalent chromatin modification	1.32 x 10 ⁻⁷
Negative regulation of MAPK cascade	1.59 x 10 ⁻⁷
Peptidyl-lysine modification	2.48 x 10 ⁻⁷
Histone modification	3.41 x 10 ⁻⁷

Functions associated to ERF-bound enhancers
C

Motifs	p-value	Matched TFs
	2.2e-354	ETS (FEV, ELF5, ELF3, FLI-1, ERG, ETS1, ERF...)
	1.1e-78	SOX(3, 2, 11, 10, 4, 12, 6, 9, 14), FOXL1, FOXA2...
	2.9e-51	POU(2F3, 2F2, 2F1, 3F1, 3F4, 5F1P1, 1F1, 3F2...)
	1.7e-25	KLF(4, 5, 1, 7), ASCL2, SP1, EGR(2, 1), FOXO3...
	2.1e-14	ZIC(1, 2, 3, 4), CTCF, SCRT(1, 2), ATOH1...

D

I

E

F

G

H


Supplemental Figure 6, related to Figure 4. (A) Western blot analysis showing the expression levels of ERF in the immunoprecipitated bound regions. A total of 0.3% of the sample was used for the Input and flow-through (Flow) extracts. (B) Genomic Regions Enrichment of Annotations Tool (GREAT) for Gene Ontology Biological Process corresponding to the closest genes associated to the 1954 ERF-bound enhancers in RAS^{less} mESC. (C) Motif analysis of ERF-bound regions in RAS^{less} mESC using MEME. The top sequence motifs are shown together with the corresponding p-value and the putative transcription factors (TFs) that bind to these sequences. (D) Graph illustrating the fold change (FC) in expression between ERF-knockout and wild type mESC for genes that are: (1) downregulated in RAS^{less} mESC and (2) in close proximity to ERF-bound enhancer regions. (E) Representative ChIP-Seq tracks of ERF at the *Myc* locus in RAS^{less} mESC. The ChIP-Seq signal from ERF-deficient mESC is also provided as a specificity control. The X-axis corresponds to the genomic location and the Y-axis corresponds to normalized ChIP-Seq signal-density. The ChIP-Seq track of P300 in mESC (GEO accession number GSM918750) is provided for comparison with a reference enhancer-associated mark. Known mESC enhancers are shown as blue boxes in the bottom track. (F) *Myc* mRNA levels in untreated or OHT-treated wild type (*Erf*^{+/+}) and ERF-deficient (*Erf*^{KO}) RAS^{lox/lox} mESC as assessed by qRT-PCR. *Myc* levels were normalized to *Gapdh* expression. (G) Representative ChIP-Seq tracks of ERF at the *Jarid2* locus in RAS^{less} mESC. The ChIP-Seq signal from ERF-deficient mESC is also provided as a specificity control. The X-axis corresponds to the genomic location and the Y-axis corresponds to normalized ChIP-Seq signal-density. The ChIP-Seq track of P300 in mESC (Heintzman et al. 2007) is provided for comparison with a reference enhancer-associated mark. Known mESC enhancers are shown as blue boxes in the bottom track. (H) *Jarid2* mRNA levels in untreated or OHT-treated wild type (*Erf*^{+/+}) and ERF-deficient (*Erf*^{KO}) RAS^{lox/lox} mESC as assessed by qRT-PCR. *Jarid2* levels were normalized to *Gapdh* expression. (I) Graph illustrating the fold change (FC) in expression between ERF-knockout and wild type mESC for genes that are: (1) upregulated in RAS^{less} mESC and (2) in close proximity to ERF-bound enhancer regions.

SUPPLEMENTARY METHODS

Western blot

Cell pellets were lysed in 50mM Tris pH 7.9, 8M Urea and 1%Chaps and incubated with shaking at 4°C for at least 30 minutes. 20 µg of supernatants were run and transferred for detection by using the corresponding antibodies. See Supplementary Table S7 for a list of the antibodies used in this work.

Chromatin immunoprecipitation

mESC from ERF-knockout, untreated and OHT-treated RAS^{lox/lox} cells were fixed in DMEM containing 10 mM Hepes, pH 7.6, 1% formaldehyde, 15 mM NaCl, 0.15 mM EDTA and 0.075 mM EGTA, for 10 min at room temperature. The reaction was quenched with 0.125 M glycine for 5 min. After washing with PBS, cells were collected and the pellet lysed in 50 mM Hepes, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% Igepal CA630, and 0.25% Triton X-100. After centrifugation, isolated nuclei were washed once in 10 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA and 0.5 mM EGTA, then resuspended in 10 mM Tris pH 8.0, 1 mM EDTA, 0.5 mM EGTA and 0.5% N-Lauryl Sarcosine, and sonicated using a Covaris S2 sonicator to an average chromatin size of 200 bp. Chromatin was diluted with two volumes of 30 mM Tris pH 8.0, 1.5 mM EDTA, 150 mM NaCl and 0.66% NP-40 and pre-cleared by adding ERF antibody and incubated at 4°C over-night. Protein G Dynabeads (Invitrogen) were blocked with BSA, and 50 µl of beads added to each condition. Following one hour of incubation at 4°C, beads were washed two times 20 mM Tris pH 8, 1 mM EDTA, 0.2% N-Lauryl Sarcosine, 1% NP-40, and 0.5 M LiCl. After a final wash with 10 mM Tris pH 8.0, 1 mM EDTA and 50 mM NaCl, beads were resuspended in 20nM Tris pH8, 1mM EDTA, 50nM NaCl and 0.05% NP40 and washed five times. Crosslink reversal was performed in 100 mM sodium bicarbonate, 200 mM NaCl, and 1% SDS. A total of 40 µg of RNase was added to each sample and the chromatin incubated for 1h at 37°C. Then, 1 µg of proteinase K was added to each sample, incubated for 2h at 55°C and the cross-linking reversed at 65°C for 16 hr. 10% of the input was treated in parallel. Finally, DNA was extracted by

using a PCR-fragment purification kit (Qiagen).

Library construction

Fragmented DNA samples were converted into sequencing libraries with "NEBNext Ultra II DNA Library Prep Kit for Illumina" (New England BioLabs). Adapter-ligated libraries were completed by limited-cycle PCR and extracted with a double-sided SPRI size selection. Libraries were sequenced on an Illumina HiSeq2000 instrument by following manufacturer's protocols.

ChIP-Seq

ERF ChIP samples from ERF-knockout, OHT-untreated Ras^{lox/lox} mESC and a pool of all the input DNAs were used as controls. Sequenced reads from ChIP-seq experiments were mapped to the NCBI37/mm9 mouse genome assembly with Bowtie (Langmead et al. 2009). Peak-calling was performed with MACS2 (Model-based Analysis of ChIP-Seq 2) (Zhang et al. 2008). Fragment length estimated on the WT-OHT sample with "macs2 predictd" module was used for "macs2 callpeak" enrichment analyses. Normalized genome-wide read densities were converted with bedGraphToBigWig and visualized on the UCSC genome browser (<http://genome.ucsc.edu>). Gene Ontology analyses were performed with GREAT (McLean et al. 2010) and to identify the genes associated to ERF-bound genomic regions we used the default association rule of the GREAT pipeline (Basal+extension: 5000 bp upstream, 1000 bp downstream, 1000000 bp max extension, curated regulatory domains included). Annotation was complemented with the Nearest Downstream Gene subroutine from PeakAnalyzer (Salmon-Divon et al. 2010), PeakAnnotator, which was used to calculate the distribution of ERF peaks to the closest transcription start site (TSS). The AnnotatePeaks.pl script from the HOMER package (Heinz et al. 2010) was used to analyse ERF peaks overlapping genomic features. Heatmap representations were generated using as a reference a previously described database of all mES cell enhancers (Hnisz et al. 2013) and represented with seqMINER (Ye et al. 2011). The GEO accession IDs for the mESC enhancer

data sets are GSM918750 and GSM1003750 for p300 and H3K4me1, respectively. Finally, IntervalStats (Chikina and Troyanskaya 2012) was used to calculate the statistical significance of the overlap between the ERF binding sites found in RASless mESC and the previously reported dataset of mESC enhancers (Hnisz et al. 2013). This analysis shows that 1324 out of 4297 the ERF peaks (30,8%) have a P-value lower than 0,001 (***) of overlapping with mESC enhancer sequences (vs randomly distributed); while 42,2% of the peaks have a P-value lower than 0,01 (***) and 49,1% a P-value lower than 0,05 (*). Two independent ChIP-Seq experiments were performed and although results from one are shown, the overlap between the peaks obtained in both experiments was over 95%.

Motif analysis

Representative motifs found in the sequences resulting from the ERF-associated peak calling were identified using DREME, which is available as part of MEME Suite of motif-based sequence analysis tools (<http://meme.nbcr.net>). DREME computes the (uncorrected) *P* value of Fisher's exact test to determine the significance of the representation of each motif in the positive set compared with that in the negative set. For each motif, the *E*-value was calculated with the corrected *P* value. Motifs with an *E*-value of <0.05 were considered significantly enriched. We then aligned those enriched motifs to sequences in the JASPAR database using TOMTOM (*P* value cut-off of 5e-14) (Gupta et al. 2007).

Gene Set Enrichment Analysis

Gene set enrichment analysis of the expression changes associated to ERF expression levels was carried out using prerranked GSEA (Barbie et al. 2009).

Microarray Analyses

WT and ERF-deficient RAS^{lox/lox} mESC (4 independent clones per genotype from 2 independent RAS^{lox/lox} lines) were treated with OHT for 2 days. Total RNA samples, obtained with Absolutely RNA Microprep Kit (Stratagene), were

labeled with the "One-Color Microarray-Based Gene Expression Analysis, Low Input Quick Amp Labeling" kit (Agilent) and hybridized to the Mouse Gene Expression G3 60K microarray (Agilent design ID 028005). Washed microarrays were scanned on a G2505C DNA microarray scanner (Agilent) and their images analyzed by Agilent Feature Extraction Software (ver. 11.5). Further information is provided at the NCBI GEO under the accession number GSE99477.

Quantitative RT-PCR

RNA was isolated with Absolutely RNA Microprep Kit (Agilent Technologies) and reverse transcribed into cDNA using the SuperScript III Reverse Transcriptase kit for RT-PCR (Invitrogen). Real-time quantitative PCR was performed with the SYBR-Select Master Mix (Applied Biosystems) on a QuantStudio 6K machine (Thermo Scientific). GAPDH expression level was used to normalize values of gene expression. Data are shown as fold change relative to the sample control and at least two independent experiments in triplicate were performed for each analysis. Primer sequences are available upon request.

Flow cytometry analysis

To analyze cell cycle profiles by flow cytometry, 7-day untreated and OHT-treated RAS^{lox/lox} mESC were incubated with 20 μ M EdU for 30 min. Trypsinized cells were fixed with 4% paraformaldehyde for 10 min and EdU incorporation was detected by the Click-iT™ EdU Alexa Fluor® Imaging kit (Invitrogen/Molecular Probes). DNA content was visualized with propidium iodide (PI). Untreated, OHT-treated or 2i treated RAS^{lox/lox} mESC with a knocked-in mCherry reporter at the Nanog locus (Maza et al. 2015) were trypsinized and the levels of NANOG expression were analyzed by flow cytometry.

Growth curves

mESC were seeded on feeders at a density of 75,000 cells per well in 12-well plates and counted every 2 or 3 days using a Neubauer camera.

Live cell imaging

To evaluate mitosis entry and duration on mESC cells, exponentially growing mESC were infected with lentiviruses encoding a histone H2B-EGFP fusion and seeded on 8 wells μ -Slide (Ibidi, 80826) pre-treated with gelatin 0.1%. The day after, control and RAS^{lox/lox} mESC that had been pretreated for 7 days with OHT were imaged every 10 minutes for a total of 24 hours in a Leica DMI 6000 B system. Mitotic events were scored by chromatin condensation/decondensation. At least 40 individual cells were followed per condition

Immunohistochemistry

Tumor samples were fixed with 4% PFA and embedded in paraffin. Subsequently, 5- μ m sections were prepared and stained with H&E or performed immunohistochemistry analyses by standard procedures to detect the expression of panRAS, OCT4 or NESTIN (See Supplemental Table S7 for a list of antibodies).

SUPPLEMENTARY REFERENCES

- Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C et al. 2009. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**: 108-112.
- Boroviak T, Loos R, Lombard P, Okahara J, Behr R, Sasaki E, Nichols J, Smith A, Bertone P. 2015. Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev Cell* **35**: 366-382.
- Chikina MD, Troyanskaya OG. 2012. An effective statistical evaluation of ChIPseq dataset similarity. *Bioinformatics* **28**: 607-613.
- Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between motifs. *Genome Biol* **8**: R24.
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**: 311-318.
- Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**: 576-589.
- Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA. 2013. Super-enhancers in the control of cell identity and disease. *Cell* **155**: 934-947.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Maza I, Caspi I, Zviran A, Chomsky E, Rais Y, Viukov S, Geula S, Buenrostro JD, Weinberger L, Krupalnik V et al. 2015. Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. *Nat Biotechnol* **33**: 769-774.
- McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* **28**: 495-501.
- Salmon-Divon M, Dvinge H, Tammoja K, Bertone P. 2010. PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics* **11**: 415.
- Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, Davidson I, Tora L. 2011. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* **39**: e35.
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**: R137.