Nanog regulates *Pou3f1* expression and represses anterior fate at the exit from pluripotency.

Antonio Barral[1*], Isabel Rollan[1*], Hector Sanchez-Iranzo[1,2], Wajid Jawaid[3], Claudio Badia-Careaga[1], Sergio Menchero[1], Manuel Gomez-Rodriguez[1], Carlos Torroja[1], Fatima Sanchez-Cabo[1], Berthold Göttgens[5], Miguel Manzanares[1#], Julio Sainz de Aja[1,4#].

1. Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), 28029 Madrid, Spain.
2. Present address: EMBL, 69117 Heidelberg, Germany.
3. Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust Cambridge Biomedical Campus Cambridge Cambridgeshire CB2 0QQ
4. Present address: Stem Cell Program, Division of Hematology/Oncology, Children's Hospital Boston, Boston, MA 02115, USA; Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.
5. Department of Haematology, Wellcome and MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK

* These authors contributed equally to this work.

\# Corresponding author julio.sainzdeaja@childrens.harvard.edu & mmanzanares@cnic.es

ABSTRACT

Pluripotency is regulated by a network of transcription factors (TF) that maintains early embryonic cells in an undifferentiated state while allowing them to proliferate. NANOG is a vital TF in maintaining pluripotency and its role in primordial germ cells (PGCs) differentiation has been well described. However, *Nanog* is expressed during gastrulation across all the posterior epiblast, and only later in development its expression is restricted to PGCs. In this work we unveiled a previously unknown mechanism by which *Nanog* specifically represses spatially the anterior epiblast lineage. Transcriptional data collection and analysis from both E6.5 embryos and embryonic stem (ES) cells revealed *Pou3f1* to be negatively correlated with NANOG during gastrulation. We have functionally proved *Pou3f1* to be a target of *Nanog* by using a dual transgene system for controlled expression of *Nanog* and *Nanog* null ES cells that further demonstrated a role for *Nanog* in repressing anterior neural genes. Deletion of a NANOG binding site (BS) 9kb downstream *Pou3f1* transcription start site revealed to have a specific role in the regionalization of *Pou3f1* expression. Our results indicate an active role of *Nanog* in inhibiting neural fate at the onset of gastrulation by repressing *Pou3f1*.

INTRODUCTION

Pluripotency is a steady state in which cells can self-renew and remain undifferentiated, retaining the capacity to give rise to any germ layer. This cell state is maintained by an intricate gene regulatory network (GRN) that is tightly regulated by a core set of transcription factors (TF): NANOG, OCT4 and SOX2 (Trott and Martinez Arias, 2013). These three TFs are involved in establishing and maintaining embryonic pluripotency, both in the blastocyst and in cultured embryonic stem (ES) cells (Chambers and Tomlinson, 2009). This GRN regulates pluripotency by repressing genes involved in differentiation and activating other genes important for pluripotency (Navarro et al., 2012; Thomson et al., 2011a). The same GRN also initiates the process of exiting pluripotency by responding to extrinsic and intrinsic signals and changing the regulatory regions and partners these factors bind to (Kalkan and Smith, 2014; Mohammed et al., 2017a; Pfeuty et al., 2018).

ES cells can be maintained in different stages of differentiation. The more studied being naïve and primed (Joo et al., 2014; Nichols and Smith, 2009a). Both can be maintained *in vitro*: the first with Leukemia Inhibitory Factor (LIF) and 2i, and the latter with Activin and FGF (Tesar et al., 2007). Although these two states can be interconverted *in vitro*, the main functional difference is that cells in a primed state cannot give rise to blastocyst chimeras (Festuccia et al., 2013). In naïve ES cells Nanog is highly and homogeneously expressed while in the primed ES cells Nanog expression levels fluctuate. Transition between these two cell states determines the onset of differentiation.

Interestingly, it has been demonstrated that low levels of *Nanog* expression in ES cells are what trigger differentiation and its overexpression is sufficient to maintain the cells in a LIF-independent pluripotent state (Chambers et al., 2007a). In spite of multiple studies that have addressed ES cell differentiation (Mendjan et al., 2014; Radzisheuskaya et al., 2013; Thomson et al., 2011b), the role of NANOG during the exit from pluripotency *in vivo* is still not well understood (Tam and Behringer, 1997).

During implantation, *Nanog* disappears from the epiblast and is re-expressed in the proximal posterior region of the epiblast after implantation, the region in which

gastrulation starts. Thus, we hypothesized that *Nanog* not only has a role in pluripotency maintenance, but also in defining lineage commitment upon gastrulation (Mendjan et al., 2014). Our lab has recently determined that at the onset of gastrulation, *Nanog* has a determinant role in repressing primitive hematopoiesis and Hox genes expression (Lopez-Jimenez et al.; Sainz de Aja et al., 2019).

Since *Nanog* is re-expressed in the posterior embryo, the prevailing hypothesis is that *Nanog* could be blocking anterior fate by repressing specific TFs in the posterior region. One of the first TFs to be expressed during implantation is *Pou3f1.* POU3F1 is a TF that promotes neural fate differentiation. At the onset of gastrulation, when *Nanog* is re-expressed in the embryo, *Pou3f1* expression becomes restricted to the anterior epiblast. It is accepted that POU3F1 antagonizes extrinsic neural inhibitory signals, however, little is known about the transcriptional regulation of this TF in the early stages of embryonic gastrulation (Zhu et al., 2014a). Here we decipher the role of *Nanog* in the regionalization of *Pou3f1,* important for the anteriorization of the embryo.

To gain insight into this mechanism*,* we studied the effects of NANOG in the transcriptome of different mouse ES cell lines and in mouse embryos. By combining the analysis *o*f our own newly generated RNAseq data together with publicly available data sets, we determined that *Pou3f1* expression is regulated by *Nanog*. To test this hypothesis, we use mouse embryos modified by CRISPR gene targeting. By knocking out NANOG binding sites (BS) next to Pou3f1 locus, we observed that *Nanog* prevents the expression of Pou3f1 in the posterior region of the gastrulating embryo right at the exit from pluripotency. Therefore, in this work we present a previously unknow mechanism by which *Nanog* constrains *Pou3f1* expression to the anterior region of the embryo.

RESULTS

Lack of *Nanog* upregulates anterior genes at the exit from naïve pluripotency

To explore the role of Nanog and to identify putative targets during the transition from pluripotency to lineage specification, we analyzed ES cells mutant for Nanog. We used Nanog KO ES cells and the parental wild-type ES cell line as control (Chambers et al., 2007a). Cells were first cultured with 2i/LIF/KOSR and subsequently changed to serum to induce exit from pluripotency (Heo et al., 2005; Martin Gonzalez et al., 2016). To follow the earliest events taking place, we took cell samples at 0, 12, and 24 hours (Fig 1A). Then, we performed RNAseq and selected genes with changed expression dynamics from 0 to 24 hours (RNAseq Table 1&2). We identified genes repressed by Nanog as those with stable expression in control ES cells but increased expression in Nanog KO cells (Fig 1B) and genes that are positively regulated by Nanog are those activated in controls but not changed in mutant cells (Fig 1C).

Principal component analysis (PCA) of the RNA-seq data showed a clear separation of the samples based on the genotype of the cells (PC1, Fig S1A) and timing of differentiation (PC2, Fig S1A). The genotypic difference resulted in close to 68% variability, whereas timing of differentiation explained 23% of the variability. Interestingly, the comparison at time 0 between control and *Nanog* KO ES cells showed minimal differences in the expression of core pluripotency genes (Fig 1D). The similarities between *Nanog* KO and wildtype cells in the pluripotent stage agree with previous observations on the dispensability of *Nanog* at the pluripotent state (Chambers et al., 2007a). We analyzed changes in gene expression, factoring in their expression over time, (Fig 1E, F) and identified two clusters with the predicted pattern of change (Fig 1G). This analysis indicates that *Nanog* might be involved in the repression of genes implicated in the development of the anterior-neural fate while promoting posterior-mesodermal fate. Our results suggest that *Nanog* has a pivotal role in repressing anterior fate at the exit from pluripotency in ES cells.

RNA-seq data reveal *Pou3f1* as a primary target for repression by NANOG in gastrulating mouse embryos

To address the putative role of Nanog in neural anterior fate in vivo, we took advantage of published E6.5 embryo single-cell RNA-sequencing (scRNAseq) data (Mohammed et al., 2017a; Scialdone et al., 2016). E6.5 is the stage at which Nanog is re-expressed in the posterior part of the mouse embryo (Hart et al., 2004) and several genes including Sox2 or Pou3f1 are already restricted to the epiblast. We merged two single cell RNA-seq expression data sets (Fig S2A) and selected those single cells expressing Nanog above a certain threshold (see Methods). Next we established the correlation of all expressed genes to that of Nanog (Fig 2A; RNAseq Table 3&4). These results confirmed our previous observations in cultured cells. Genes that correlate positively with Nanog are related to gastrulation and mesoderm formation, such as Fgf8, Nodal or Eomes (Fig 2B). Genes that negatively correlated with Nanog were Pou3f1 (Fig S2B) and other neural genes such as Nav2 (Fig 2B). Other early anterior genes, such as Sox2, did not show any correlation with Nanog levels (Fig S2C), suggesting that Nanog might not have a broad impact on anterior specification, but rather has a specific effect on certain TFs. Interestingly, among the negatively correlated genes we also found Utf1, a pluripotency associated gene that is restricted to the anterior region of the embryo during gastrulation and to extraembryonic tissues (Okuda et al., 1998). Enrichment (Chen et al., 2013; Kuleshov et al., 2016) analysis of the clustered genes matching the Jansen tissues gene set library (with a coverage of 19586 genes), allowed us to observe that negatively correlated genes were mostly related to neural development tissues (spinal cord, frontal lobe) (Bell et al., 2000; O'Donnell et al., 2006; Zhu et al., 2014b), and with a lower z-score to endodermal tissues (gut, intestine) (Fig S2E). Genes that positively correlate with Nanog expression were mainly represented in mesodermal tissues (monocyte, b lymphoblastoid cell, bone) (Fig S2F) (Reichenbach et al., 2008; Skerjanc et al.).

We next addressed the effect of expressing *Nanog* throughout the early embryo by using an inducible tetON transgenic model (*Nanog*^{tg}) in which *Nanog* expression is induced by the administration of doxycycline (dox) (Piazzolla et al., 2014). We used bulk RNAseq data of induced *Nanog* by administering dox in the

drinking water of pregnant female mice from E4.5 to E7.5 and examined changes in gene expression by RNA-seq using untreated females of the same genotype as controls (Lopez-Jimenez et al.). In this dataset, many genes involved in the early aspects of embryo pattering, such as Hox genes, were downregulated, but the most strongly downregulated gene when *Nanog* was expressed throughout the early embryo was *Pou3f1* (Fig 2C, RNAseq Table 5&6). The expression of other anterior neural genes, for example *Sox2*, *Hesx1* or *Zic3*, was not changed. We confirmed these observations by whole mount *in situ* hybridization of treated and untreated E7.5 *Nanog*^{tg} embryos. Induction of *Nanog* led to a downregulation of *Pou3f1* in the anterior epiblast of treated embryos, while expression of *Sox2* was unchanged (Fig 2D). Interestingly, when *Nanog*^{tg} embryos were recovered at E9.5 treated with dox from E6.5, they presented craniofacial defects that might be a direct consequence of the anterior repression (Fig S2H).

We merged the data from all previous transcriptomic studies, finding only three genes shared by those upregulated in *Nanog* KO ES cells in early differentiation: *Pou3f1*, *Lrp2* and *Clic6.* These genes had a significant negative correlation with *Nanog* in E6.5 single cell transcriptomics and were downregulated in E7.5 *Nanog* gain-of-function embryos (Fig 2E). Genes whose expression positively correlated with that of *Nanog* in E6.5 single cells and that were upregulated in dox-treated *Nanog*^{tg} embryos were mostly genes related to early gastrulation and mesoderm specification, such as *Eomes*, *Fgf8*, *Tdgf1* (*Cripto*) or *Mixl1* (Fig S2E). *Lrp2* and *Clic6* are expressed in primitive endoderm and late derivatives (Gerbe et al., 2008; Sherwood et al., 2007), which are lineages in which *Nanog* has a well-defined negative regulatory role. Therefore, *Pou3f1* is the prime candidate to be a direct target of *Nanog,* mediating its role in suppressing anterior epiblast fate.

Nanog expression impairs neural differentiation *in vitro*

Because we observed that *Pou3f1* is a key factor in early neural differentiation, we wanted to know whether *Nanog* has a direct role in negatively regulating the differentiation towards the early anterior-neural lineage. To confirm whether *Nanog* is blocking anterior fate progression, we derived ES cells from the *Nanog*^{tg} line and differentiated them towards anterior neural fate (Gouti et al., 2014, 2017),

culturing them with or without dox for up to six days. Analysis of gene expression by RT-qPCR showed that upon *Nanog* overexpression, pluripotency related genes, such as *Rex1*, followed the same pattern in untreated or treated cells, despite a slight increase in *Nanog* gain-of-function cells along the time course. On the other hand, anterior neural specific genes were downregulated during the differentiation process, among them Sox1, *Pou3f1, Otx2,* and *Pax6*. *Sox2*, which qualifies as both a pluripotency and an early neural TF, showed a similar pattern of expression by qPCR in both *Nanog*^{tg} +dox and -dox (Fig 3A), but markedly higher protein levels in the differentiation of wildtype cells (Fig 3B). When cells were differentiated towards hindbrain (Gouti et al., 2014, 2017), differences in the expression of neural TFs were reduced, although differences in *Pou3f1* expression still existed (Fig S3). These results indicate that during neural differentiation, *Nanog* downregulates genes important for neural specification.

A distal NANOG-binding element represses *Pou3f1* expression in the posterior epiblast

All evidence indicates that *Pou3f1* is likely a direct transcriptional target of NANOG during anterior-posterior axis specification in the epiblast. To explore possible mechanisms of *Nanog* and *Pou3f1* interaction, we analyzed published ChIP-seq data for NANOG binding in ES and epiblast-like cells (EpiLCs) (Murakami et al., 2016). EpiLCs represent the *in vitro* counterpart of the E6.0-6.5 epiblast of mouse embryos. Murakami *et al.* 2016 describes a broad resetting of NANOG-occupied genomic regions in the transition from ES cells to EpiLCs, resembling the developmental progress from the naïve inner cell mass of the blastocyst to the primed epiblast at gastrulation. Therefore, we examined the *Pou3f1* locus and identified three prominent regions of NANOG binding at 11.5 and 9 kilobases (kb) upstream and 9 kb downstream from the transcription start site. Interestingly, NANOG binds these regions in EpiLC but not in ES cells, suggestive of a specific input of *Nanog* on *Pou3f1* in the epiblast but not at earlier pluripotent stages (Fig 4A).

We hypothesized that the deletion of the NANOG BS would avoid the repression of *Pou3f1* expression and therefore, at E6.5, expanding the domain of expression

of *Pou3f1* towards the posterior region of the embryo. To investigate this hypothesis, we sequentially deleted each of the three NANOG binding sites near the *Pou3f1* locus by CRISPR/Cas9 genome editing in a transient transgenic embryo assay. One by one, we deleted these BS in embryos microinjected with Cas9-gRNAs ribonucleoproteins were recovered at E6.5, processed for whole mount in situ hybridization, and subsequently genotyped the locus of the expected deletion. This assay showed that only deletion of the +9 kb downstream region caused a reproducible change in *Pou3f1* expression. We were expecting an expansion of Pou3f1 expression from the anterior region to the rest of the epiblast, and we only observed this expansion consistently in the +9Kb BS deletion (Fig 4B,C and S4A,B,C). To determine whether the pattern of Pou3f1 expression was expanded towards the posterior epiblast, we generated a stable mouse line carrying this deletion. Mice homozygous for the deletion (Fig S4D) were viable and fertile. This was not completely unexpected, as homozygous null *Pou3f1* mice survive up to birth (Bermingham et al., 1996). We crossed mice heterozygous and homozygous for the deletion and compared littermates for the expression of *Pou3f1* by whole mount *in situ* hybridization. We observed that 3 out of 5 homozygous embryos presented a phenotype of posterior expansion while none of the heterozygous embryos did (Fig 4D). These results demonstrated that the +9 kb NANOG bound region is important for the restriction of Pou3f1 expression to the anterior epiblast.

DISCUSSION

In this work, we sought to capture the first step in the differentiation of naïve ES cells to assess the role of *Nanog* in the exit from pluripotency. We observed that *Nanog* represses differentiation of anterior fate while promoting posterior differentiation. As the pluripotency gene regulatory network develops towards a more differentiatied state, the absence of *Nanog* generates an imbalance in the lineage fate of the cells, which reflects the equilibrium of *Nanog, Oct4,* and *Sox2* during this state transition (Pfeuty et al., 2018). It might be expected that in the absence of *Nanog*, *Sox2* could drive the cells towards a more anterior fate. Moreover, it has already been shown *in vitro* that *Nanog* antagonizes other

anterior genes, such as *Otx2 (Su et al.)*. However, what we have observed instead is that neither *Sox2* nor *Otx2* regulates this transition but rather *Pou3f1*.

Pou3f1 has been reported to be the POUIII family member first expressed in the mouse embryo. It is one of the main TFs driving the progression toward the neural differentiation both in ES cells and in epiblast like stem cells (EpiLCs) through the activation of intrinsic neural lineage genes such as *Sox2* or *Pax6* (Song et al., 2015; Zhu et al., 2014b).

We performed an RNAseq on the first 24h of exit from pluripotency of naïve ES cells from both wildtype and *Nanog*KO. When merging our RNAseq data with the previously published scRNAseq data (Mohammed et al., 2017a; Scialdone et al., 2016) of the epiblast of E6.5 embryos we compared three set of genes: 1) Genes that are upregulated in the *Nanog* KO ES cells at the exit from pluripotency; 2) Genes that are down regulated upon *Nanog* forced expression in the whole gastrulating epiblast; 3) Genes whose expression correlates negatively with the expression of *Nanog* at the onset of gastrulation (Fig 2D,S2G).

Three genes met all three conditions: *Pou3f1*, *Lrd2,* and *Clic6*. These three genes have anterior neural expression and a role in neural differentiation, but only *Pou3f1* regulates the differentiation of the anterior epiblast (Lewis and Tam, 2006; Sherwood et al., 2007). Furthermore, the results of the *Nanog* ectopic expression in the ESC differentiation towards neural fate corroborates the effects observed in the embryo and demonstrates that specific differentiation towards neural is impaired when *Nanog* is expressed during anterior neural fate acquisition (Fig 3A). *Nanog* overexpression still represses *Pou3f1* expression in posterior neural (hindbrain) differentiation (Fig S3A). However, differences in the levels of expression between +dox and –dox are minimal compared with the forebrain differentiation. This could be due to the effect of retinoic acid (RA) on the hindbrain differentiation, which somehow could be bypassing *Nanog* effect in neural differentiation. Sox2 seems to be downregulated in the +dox condition in the immunofluorescence while at the mRNA level does not change. These differences could be due to NANOG regulating SOX2 at a protein level. However, this effect has not been reported before. More experiments will be required to test this hypothesis.

Little is known about the transcriptomic regulation of *Pou3f1*. When we looked at the NANOG-BSs in ES cells and EpiLCs at 1 and 2 days of differentiation (D1 and D2 respectively), we found three binding sites surrounding the *Pou3f1* locus. These BSs are specific to EpiLCs at both D1 and D2, and were absent from ES cells, again pointing towards the transcriptional differences between the naïve pluripotent state and the primed state. The EpiLC (primed) state represents a pre-gastrulating epiblast *in vitro* (Murakami et al., 2016; Nichols and Smith, 2009b; Tesar et al., 2007). Deletion of each of the three NANOG-BS on the *Pou3f1* locus in transient transgenic mouse embryos, yielded diverse results with high variability in the expression of *Pou3f1* depending on the target site and on slight variations in the staging of the embryos. Using the expansion of the *Pou3f1* pattern of expression towards the posterior epiblast in E6.5 as a phenotypic standard, we were able to narrow the site at which NANOG binding apparently had the greatest effect on *Pou3f1* expression. After analyzing the results of the deletion of the different binding sites, we decided to generate a mouse line with the +9Kb BS deleted, which had the highest effect. We observed that 3 out of the 5 homozygous embryos had an expansion of *Pou3f1* expression towards the posterior region of the epiblast, whereas in the heterozygous embryos we did not observe any phenotype. Together these results indicate that +9kb binding site regulates *Pou3f1* expression in the posterior epiblast of the gastrulating embryo, and its expression is probably regulated by *Nanog*.

The pluripotency GRN in ES cells and in the inner cell mass of the blastocyst is crucial for the activation and maintenance of pluripotency. However, this GRN is important not only at these stages, but also at the exit from pluripotency, whereby it acquires different roles during differentiation (Iwafuchi-Doi et al., 2012). Our results show that *Nanog* is able to repress *Pou3f1* in the posterior embryo without affecting *Sox2* expression. This work reveals a new role for *Nanog* in a stage in which the pluripotency GRN is untangling, allowing the pluripotency TFs to acquire novel roles past embryonic implantation. Thus, our study offers new insights into the early differentiation of cells that will further contribute to the understanding of the involvement of gene networks, as well as specific genes in pluripotency and early specification in the embryo.

MATERIALS AND METHODS

Animal model

We obtained the *Nanog/rtTA* mouse line (*R26-M2rtTA;Col1a1-tetO-Nanog*) (Piazzolla et al., 2014) from Manuel Serrano (CNIO, Madrid) and Konrad Hochedlinger (Harvard Stem Cell Institute). This is a double transgenic line that carries the *M2-rtTA* gene inserted at the *Rosa26* locus and a cassette containing *Nanog* cDNA under the control of a doxycycline-responsive promoter (tetO) inserted downstream of the *Col1a1* locus. Mice were genotyped by PCR of tail-tip DNA as previously described (Hochedlinger et al., 2005; Piazzolla et al., 2014). Mice were housed and maintained in the animal facility at the Centro Nacional de Investigaciones Cardiovasculares (Madrid, Spain) in accordance with national and European Legislation. Procedures were approved by the CNIC Animal Welfare Ethics Committee and by the Area of Animal Protection of the Regional Government of Madrid (ref. PROEX 196/14).

For the generation of transgenic embryos, 7 week old F1 (C57Bl/6xCBA) females were superovulated to obtain fertilized oocytes as described (Nagy et al., 2014).

Transient transgenic analysis and transgenic line generation

Females were injected with 5 units of pregnant mare's serum gonadotropin (PMSG, Foligon 5000) and 5 units of chorionic gonadotropin (Sigma) two days later, followed by embryo collection the next day. After 1-hour incubation, viable zygotes were microinjected into the pronucleus with commercially available Cas9 protein (30ng/ul; PNABio) and guide RNAs (sgRNA; 25ng/ul; Sigma). All those components were previously hybridized to generate ribonucleoprotein complexes. First, we incubated 100ng/ul of Trans-activating crRNA (tracrRNA) and sgRNA for 5 minutes at 95ºC and then for 10 minutes at room temperature (RT). We then incubated the sgRNAs with the Cas9 for 15 minutes at RT and stored at 4ºC. Injection buffer consisted of Tris 50nM pH7.4, EDTA 1nM, H_2O embryo tested and was filtered through a 0.22um filter. sgRNA were designed with an online tool (http://crispr.mit.edu/). Details of the sequences are shown in Table S1. Primers for genotyping the deletions are detailed in Table S2. PCR products were run on an agarose gel; the specific band was excised and DNA was extracted with the DNA extraction kit (Qiagen). Purified DNA was transcribed

for RNA collection/isolation. After injection, embryos were cultured in M16 (Sigma) covered with mineral oil (Nid Oil, EVB) up to the two-cell stage. Living embryos were then transferred into a pseudopregnant CD1 female, previously crossed with a vasectomized male.

E6.5 embryos were recovered for transient experiments. We also generated stable lines for maintenance. +9Kb Pou3f1 deletion line is maintained in an outbred background (CD1). Control mice for the line experiments are embryos from CD1 females crossed with F1 males.

RT-qPCR assays

RNA was isolated from ESCs using the RNeasy Mini Kit (Qiagen) and then reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). cDNA was used for quantitative-PCR (qPCR) with Power SYBR® Green (Applied Biosystems) in a 7900HT Fast Real-Time PCR System (Applied Biosystems). Primers for qPCR detailed in Table S3.

***In situ* hybridization**

Embryos were collected in cold PBS, transferred to 4% PFA, and fixed overnight at 4 ºC. After washing, embryos were dehydrated in increasing concentrations of PBS-diluted methanol (25%, 50%, 75%, and 2X 100%). *In situ* hybridization in whole mount embryos was performed as previously described (Acloque et al., 2008; Ariza-McNaughton and Krumlauf, 2002). Signal was developed with anti-dioxigenin-AP (Roche) and BM-Purple (Roche). Images were acquired with a Leica MZ-12 dissecting microscope. Probes primers are detailed in Table S4.

Cell culture

ESCs were maintained in serum-free conditions with Knock out serum replacement (Thermo Fisher), LIF (produced in-house), and 2i (CHIR-99021, Selleckchem; and PD0325901, Axon) over inactive mouse embryonic fibroblast (MEFs). The *Nanog*^{KO} (BT12) and their parental wild-type control (E14Tg2a) ES cells were kindly provided by Ian Chambers and Austin Smith (Chambers et al., 2007b). The *Nanog* gain-of-function ES cells (NOES) were derived in the lab from the *Nanog/rtTA* mouse line were derived following standard procedures (Nagy et

al., 2014). Karyotyping of the obtained lines was performed by the pluripotent cell technology unit at CNIC.

Anterior neural differentiation (forebrain and hindbrain)

NOES cells were differentiated to either forebrain or hindbrain lineage as described (Gouti et al., 2014, 2017), in monolayer using Corning p24 plates with 0.1% gelatin (Sigma) added 30 min before passaging. Cells were grown in N2B27 media supplemented with 10 ng/mL bFgf (R&D) for the first 3 days (d1–d3), and then from day 3 to 6 in N2B27 without growth factors for the forebrain fate, or N2B27 supplemented with 10nM retinoic acid for the hindbrain fate.

Statistical analysis

Statistical analysis was performed with the use of two-tailed Student's unpaired t-test analysis (when the statistical significance of differences between two groups was assessed). Fisher exact test was performed for analysis of contingency tables containing the data of the deleted genotypes and expanded phenotypes. Prism software version 7.0 (Graphpad Inc.) was used for representation and statistical analysis. Enriched functional categories in the mouse gene atlas score was calculated using Enrichr (Chen et al., 2013; Kuleshov et al., 2016).

***Nanog^{KO}* ES cell RNAseq**

RNA from *Nanog^{KO}* ES cells and their parental line was extracted as indicated above. Next generation sequencing single read (Illumina HiSeq 2500) and library preparation (New England Biolabs Nest Ultra RNA library prep Kit) were performed in the genomics unit at Centro Nacional de Investigaciones Cardiovasculares (CNIC).

Sequencing reads were inspected by means of a pipeline that used FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) to asses read quality, and Cutadapt v1.3 (Martin, 2011) to trim sequencing reads, eliminating Illumina adaptor remains, and to discard reads that were shorter than 30bp. For libraries amplified with the NuGen Ovation Single Cell kit, the first 8 nucleotides of each read were also eliminated with fastx_trimmer (http://hannonlab.cshl.edu/fastx). The resulting reads were mapped against the mouse transcriptome (GRCm38

assembly, Ensembl release 76) and quantified using RSEM v1.2.20 (Li and Dewey, 2011). Raw expression counts were then processed with an analysis pipeline that used Bioconductor packages EdgeR (Robinson et al., 2010) for normalization (using TMM method) and differential expression testing, and ComBat (Johnson et al., 2007) for batch correction. Only genes expressed at a minimal level of 1 count per million, in at least 3 samples, were considered for differential expression analysis. Changes in gene expression were considered significant if their Benjamini and Hochberg adjusted p-value (FDR) was lower than 0.05.

Analysis of scRNAseq of E6.5 embryos

Two data sets from different experiments (Mohammed et al., 2017b; Scialdone et al., 2016) were normalized by quantiles and batch corrected. To plot the cells in a tSNE, genes with 0 variance were eliminated.

For correlation of genes with *Nanog* we used the slope of the line adjusted to the points per sample. For plotting we used ggPlot package from R. We separated the plots by sample. Statistical analysis was developed in R.

Intersection analysis

The intersection analysis of the genes coming from different RNAseq datasets was performed with the web tool from Bioinformatics and Evolutionary Genomics (http://bioinformatics.psb.ugent.be/webtools/Venn/)

ACKNOWLEDGEMENTS

We thank Manuel Serrano and Konrad Hochedlinger for the *Nanog*^{tg} mouse line; Austin Smith and Ian Chambers for *Nanog*$^{-/-}$ ES cell lines; Cristina Gutierrez-Vazquez, Rocio Nieto-Arellano and Teresa Rayon for comments on the manuscript; Briane Larui for English editing and Karen Pepper for input in the writing; Jesus Victorino for discussion and members of Manzanares lab for continued support. This work was supported by the Spanish government (grant BFU2014-54608-P and BFU2017-84914-P to MM). The Gottgens laboratory is supported by core funding from the Wellcome Trust and MRC to the Wellcome and MRC Cambridge Stem Cell Institute. The CNIC is supported by the Spanish

Ministry of Science, Innovation and Universities (MINECO) and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

AUTHOR CONTRIBUTION

JSdA and MM envisioned the project, JSdA design the experiments with input from MM and BG. AB and IR performed the majority of the experiments with the supervision of JSdA and the help of SM for embryo work and tissue culture; HSI and WJ did RNAseq analysis; MGR, CT and FS analyzed single cell data and designed statistical analysis; and CBC performed in situs. JSdA wrote the manuscript with input from all authors.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Acloque, H., Wilkinson, D.G., and Nieto, M.A. (2008). Chapter 9 In Situ Hybridization Analysis of Chick Embryos in Whole-Mount and Tissue Sections. In Methods in Cell Biology, pp. 169–185.

Ariza-McNaughton, L., and Krumlauf, R. (2002). Non-radioactive in situ hybridization: simplified procedures for use in whole-mounts of mouse and chick embryos. Int. Rev. Neurobiol. *47*, 239–250.

Bell, K.M., Western, P.S., and Sinclair, A.H. (2000). SOX8 expression during chick embryogenesis. Mech. Dev. *94*, 257–260.

Bermingham, J.R., Scherer, S.S., O'Connell, S., Arroyo, E., Kalla, K.A., Powell, F.L., and Rosenfeld, M.G. (1996). Tst-1/Oct-6/SCIP regulates a unique step in peripheral myelination and is required for normal respiration. Genes Dev. *10*, 1751–1762.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007a). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007b). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Festuccia, N., Osorno, R., Wilson, V., and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. Curr. Opin. Genet. Dev. *23*, 504–511.

Gerbe, F., Cox, B., Rossant, J., and Chazaud, C. (2008). Dynamic expression of Lrp2 pathway members reveals progressive epithelial differentiation of primitive endoderm in mouse blastocyst. Dev. Biol. *313*, 594–602.

Gouti, M., Tsakiridis, A., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2014). In Vitro Generation of Neuromesodermal Progenitors Reveals Distinct Roles for Wnt Signalling in the Specification of Spinal Cord and Paraxial Mesoderm Identity. PLoS Biol. *12*, e1001937.

Gouti, M., Delile, J., Stamataki, D., Wymeersch, F.J., Huang, Y., Kleinjung, J., Wilson, V., and Briscoe, J. (2017). A Gene Regulatory Network Balances Neural and Mesoderm Specification during Vertebrate Trunk Development. Dev. Cell *41*, 243–261.e7.

Hart, A.H., Hartley, L., Ibrahim, M., and Robb, L. (2004). Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. Dev. Dyn. *230*, 187–198.

Heo, J., Lee, J.-S., Chu, I.-S., Takahama, Y., and Thorgeirsson, S.S. (2005). Spontaneous differentiation of mouse embryonic stem cells in vitro: Characterization by global gene expression profiles. Biochem. Biophys. Res. Commun. *332*, 1061–1069.

Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. Cell *121*, 465–477.

Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P.J., Aruga, J., Matsuo, I., and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. Development *139*, 3926–3937.

Johnson, W.E., Li, C., and Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics *8*, 118–127.

Joo, J.Y., Choi, H.W., Kim, M.J., Zaehres, H., Tapia, N., Stehling, M., Jung, K.S., Do, J.T., and Schöler, H.R. (2014). Establishment of a primed pluripotent epiblast stem cell in FGF4-based conditions. Sci. Rep. *4*.

Kalkan, T., and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. Philos. Trans. R. Soc. Lond. B. Biol. Sci. *369*, 20130540-.

Kuleshov, M. V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44*, W90–W97.

Lewis, S.L., and Tam, P.P.L. (2006). Definitive endoderm of the mouse embryo: Formation, cell fates, and morphogenetic function. Dev. Dyn. *235*, 2315–2329.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-

Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Lopez-Jimenez, E., Sainz de Aja, J., Rouco, R., Victorino, J., Santos, E., Badia-Careaga, C., Rollan, I., Acemel, R.D., Torroja, C., Andres-Leon, E., et al. Pluripotency factors control the onset of Hox cluster activation in the early embryo.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.Journal *17*, 10.

Martin Gonzalez, J., Morgani, S.M., Bone, R.A., Bonderup, K., Abelchian, S., Brakebusch, C., and Brickman, J.M. (2016). Embryonic Stem Cell Culture Conditions Support Distinct States Associated with Different Developmental Stages and Potency. Stem Cell Reports *7*, 177–191.

Mendjan, S., Mascetti, V.L., Ortmann, D., Ortiz, M., Karjosukarso, D.W., Ng, Y., Moreau, T., and Pedersen, R.A. (2014). Cell Stem Cell Article NANOG and CDX2 Pattern Distinct Subtypes of Human Mesoderm during Exit from Pluripotency. Stem Cell *15*, 310–325.

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017a). Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. Cell Rep. *20*, 1215–1228.

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017b). Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. Cell Rep. *20*, 1215–1228.

Murakami, K., Günesdogan, U., Zylicz, J.J., Tang, W.W.C., Sengupta, R., Kobayashi, T., Kim, S., Butler, R., Dietmann, S., and Surani, M.A. (2016). NANOG alone induces germ cells in primed epiblast in vitro by activation of enhancers. Nature *529*, 403–407.

Nagy, A., Vintersten, K., Gertsenetein, M., and Behringer, R. (2014). Manipulating the Mouse Embryo: A Laboratory Manual, Fourth Edition.

Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N.P., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M., et al. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. EMBO J. *31*, 4547–4562.

Nichols, J., and Smith, A. (2009a). Naive and Primed Pluripotent States. Cell Stem Cell *4*, 487–492.

Nichols, J., and Smith, A. (2009b). Naive and Primed Pluripotent States. Cell Stem Cell *4*, 487–492.

O'Donnell, M., Hong, C.-S., Huang, X., Delnicki, R.J., and Saint-Jeannet, J.-P. (2006). Functional analysis of Sox8 during neural crest development in Xenopus. Development *133*, 3817–3826.

Okuda, A., Fukushima, A., Nishimoto, M., Orimo, A., Yamagishi, T., Nabeshima, Y., Kuro-o, M., Nabeshima, Y., Boon, K., Keaveney, M., et al. (1998). UTF1, a novel transcriptional coactivator expressed in pluripotent embryonic stem cells and extra-embryonic cells. EMBO J. *17*, 2019–2032.

Pfeuty, B., Kress, C., and Pain, B. (2018). Network Features and Dynamical Landscape of Naive and Primed Pluripotency. Biophys. J. *114*, 237–248.

Piazzolla, D., Palla, A.R., Pantoja, C., Cañamero, M., de Castro, I.P., Ortega, S., Gómez-López, G., Dominguez, O., Megías, D., Roncador, G., et al. (2014). Lineage-restricted function of the pluripotency factor NANOG in stratified epithelia. Nat. Commun. *5*, 4226.

Radzisheuskaya, A., Chia, G.L. Bin, dos Santos, R.L., Theunissen, T.W., Castro, L.F.C., Nichols, J., and Silva, J.C.R. (2013). A defined Oct4 level governs cell state transitions of pluripotency entry and differentiation into all embryonic lineages. Nat. Cell Biol. *15*.

Reichenbach, B., Delalande, J.-M., Kolmogorova, E., Prier, A., Nguyen, T., Smith, C.M., Holzschuh, J., and Shepherd, I.T. (2008). Endoderm-derived Sonic hedgehog and mesoderm Hand2 expression are required for enteric nervous system development in zebrafish. Dev. Biol. *318*, 52–64.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sainz de Aja, J., Menchero, S., Rollan, I., Barral, A., Tiana, M., Jawaid, W., Cossio, I., Alvarez, A., Carreño-Tarragona, G., Badia-Careaga, C., et al. (2019). The pluripotency factor NANOG controls primitive hematopoiesis and directly regulates *Tal1*. EMBO J. e99122.

Scialdone, A., Tanaka, Y., Jawaid, W., Moignard, V., Wilson, N.K., Macaulay, I.C., Marioni, J.C., and Göttgens, B. (2016). Resolving early mesoderm diversification through single-cell expression profiling. Nature *535*, 289–293.

Sherwood, R.I., Jitianu, C., Cleaver, O., Shaywitz, D.A., Lamenzo, J.O., Chen, A.E., Golub, T.R., and Melton, D.A. (2007). Prospective isolation and global gene expression analysis of definitive and visceral endoderm. Dev. Biol. *304*, 541–555.

Skerjanc, I.S., Petropoulos, H., Ridgeway, A.G., and Wilton, S. Myocyte Enhancer Factor 2C and Nkx2–5 Up-regulate Each Other's Expression and Initiate Cardiomyogenesis in P19 Cells*.

Song, L., Sun, N., Peng, G., Chen, J., Han, J.-D.J., and Jing, N. (2015). Genome-wide ChIP-seq and RNA-seq analyses of Pou3f1 during mouse pluripotent stem cell neural fate commitment. Genomics Data *5*, 375–377.

Su, Z., Zhang, Y., Liao, B., Zhong, X., Chen, X., Wang, H., Guo, Y., Shan, Y., Wang, L., and Pan, G. Antagonism between the transcription factors NANOG and OTX2 specifies rostral or caudal cell fate during neural patterning transition.

Tam, P.P.., and Behringer, R.R. (1997). Mouse gastrulation: the formation of a mammalian body plan. Mech. Dev. *68*, 3–25.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

Thomson, M., Liu, S., Zou, L., Smith, Z., and Meissner, A. (2011a). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell *145*.

Thomson, M., Liu, S.J., Zou, L.-N., Smith, Z., Meissner, A., and Ramanathan, S. (2011b). Pluripotency Factors in Embryonic Stem Cells Regulate Differentiation into Germ Layers. Cell *145*, 875–889.

Trott, J., and Martinez Arias, A. (2013). Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency. Biol. Open *2*, 1049–1056.

Zhu, Q., Song, L., Peng, G., Sun, N., Chen, J., Zhang, T., Sheng, N., Tang, W., Qian, C., Qiao, Y., et al. (2014a). The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. Elife *3*.

Zhu, Q., Song, L., Peng, G., Sun, N., Chen, J., Zhang, T., Sheng, N., Tang, W., Qian, C., Qiao, Y., et al. (2014b). The transcription factor Pou3f1 promotes neural fate commitment via activation of neural lineage genes and inhibition of external signaling pathways. Elife *3*.

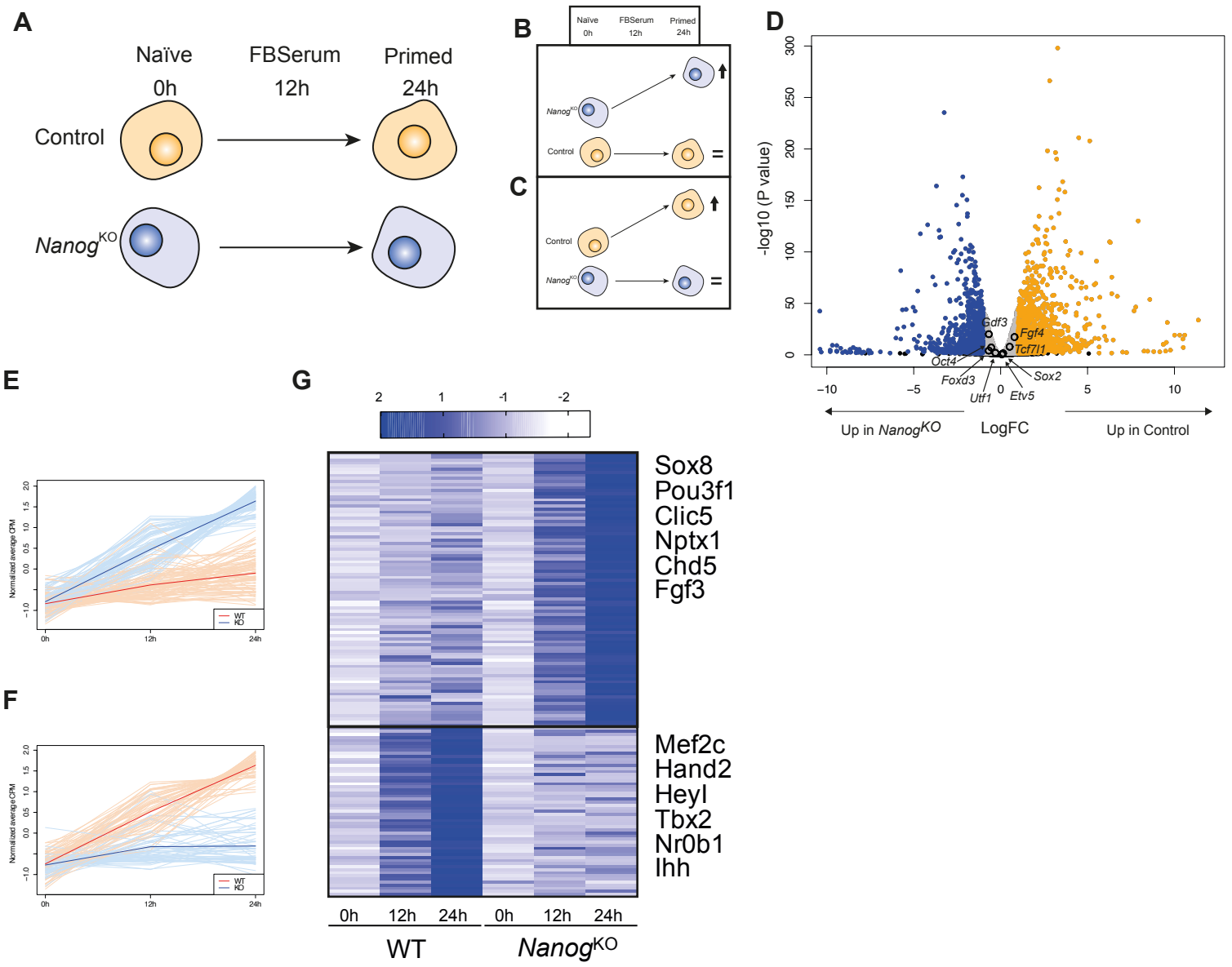

A
Naïve
0h
FBSerum
12h
Primed
24h
Control
Nanog^KO
B
C
D
-log10 (P value)
Gdf3
Fgf4
Tcf7l1
Oct4
Foxd3
Utf1
Etv5
Sox2
Up in Nanog^KO
LogFC
Up in Control
E
F
G
Sox8
Pou3f1
Clic5
Nptx1
Chd5
Fgf3
Mef2c
Hand2
Heyl
Tbx2
Nr0b1
Ihh
0h
12h
24h
WT
Nanog^KO

Figure1

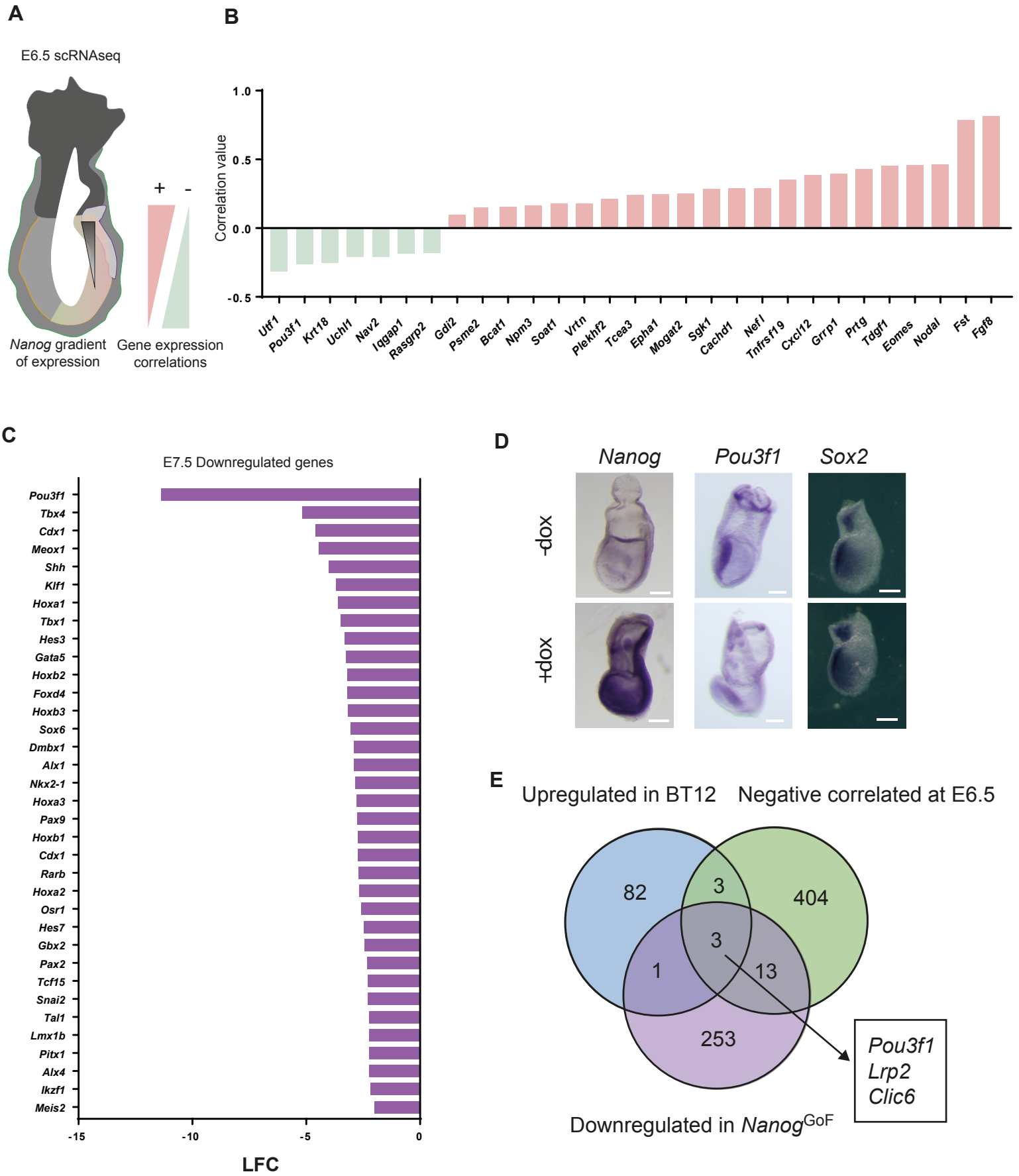
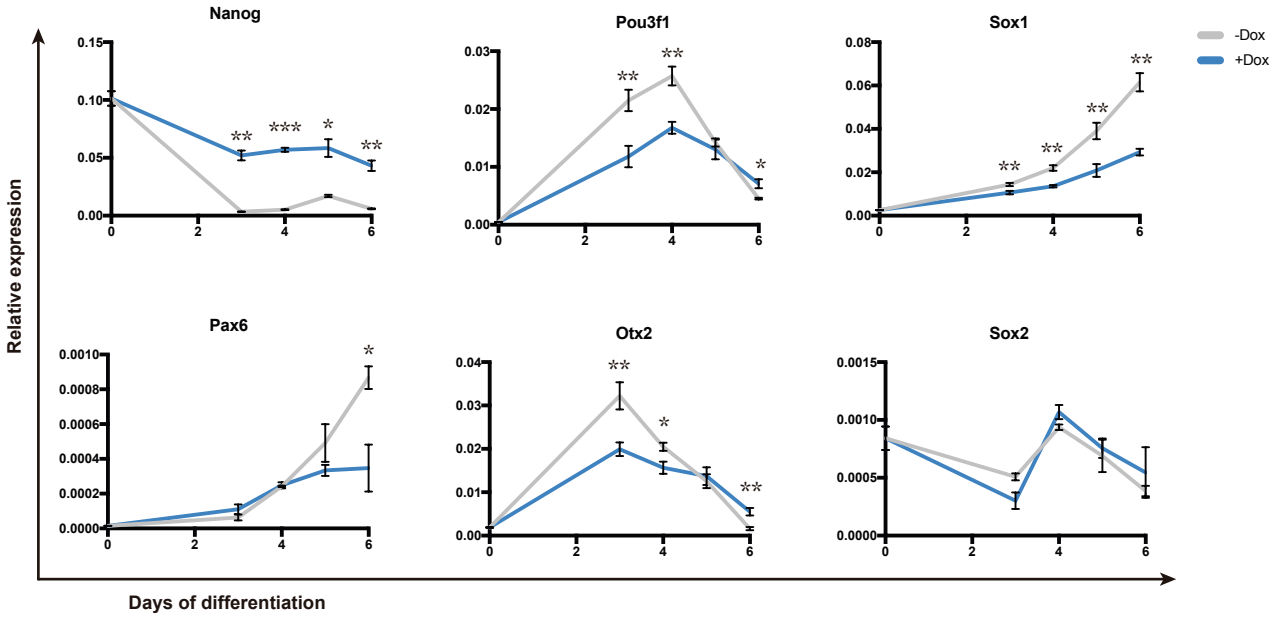

A
E6.5 scRNAseq
+
-
Nanog gradient of expression
Gene expression correlations
B
Correlation value
1.0
0.5
0.0
-0.5
Utf1
Pou3f1
Krt18
Uchl1
Nav2
Iqgap1
Rasgrp2
Gdi2
Psme2
Bcat1
Npm3
Soat1
Vrtn
Plekhf2
Tcea3
Epha1
Mogat2
Sgk1
Cachd1
Nef1
Tnfrsf19
Cxcl12
Grrp1
Prtg
Tdgf1
Eomes
Nodal
Fst
Fgf8
C
E7.5 Downregulated genes
Pou3f1
Tbx4
Cdx1
Meox1
Shh
Klf1
Hoxa1
Tbx1
Hes3
Gata5
Hoxb2
Foxd4
Hoxb3
Sox6
Dmbx1
Alx1
Nkx2-1
Hoxa3
Pax9
Hoxb1
Cdx1
Rarb
Hoxa2
Osr1
Hes7
Gbx2
Pax2
Tcf15
Snai2
Tal1
Lmx1b
Pitx1
Alx4
Ikzf1
Meis2
-15
-10
-5
0
LFC
D
Nanog
Pou3f1
Sox2
-dox
+dox
E
Upregulated in BT12
Negative correlated at E6.5
82
3
404
3
1
13
253
Pou3f1
Lrp2
Clic6
Downregulated in NanogGoF

Figure 2

A

Forebrain differentiation of Nanog^{GoF} ESC



B

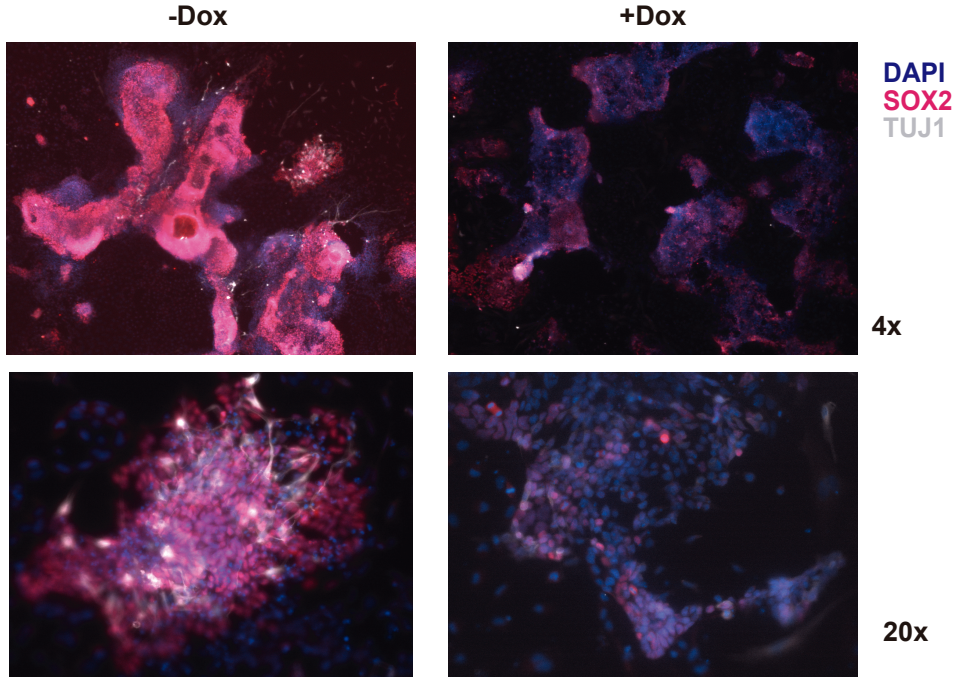
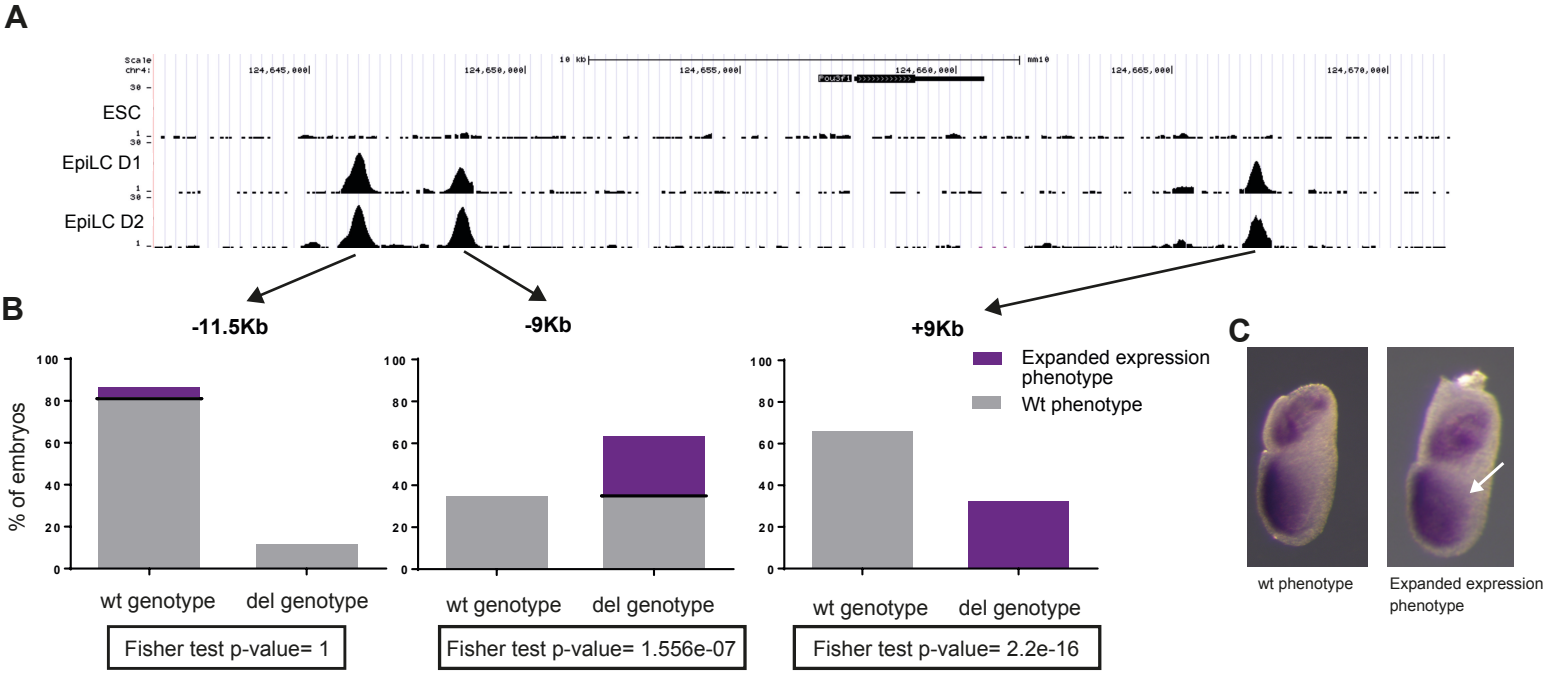


Figure 3



D

GGGGTCACCAAATGGATCATTTGTCACCTTGGACCTCAAACTAGGA
AGTAACCATTCCTCCATTCACTTATCCATCAGTCCGTTCTCCAGCC
AGCTACTCCTCCA**TTCACCGAGCCACCCACCTT**TGCATGCAGCCAT
TCACCTACCTACCTATTCAGCTACCCACCCATCATCTGCCATCCAT
TCATCCGATCTTCTTGCTTAACGAATGTCAAGAAGGCTGATGGCCT
CTAACCTCCACGCTCTTCTAGCTTTGTCTCATTCCCTTCCTGCCCC
TC**AGCCATATAGGGTGTTTGAG**AAAACGGTGTAGAAGGTAATGATG
AAACCTGTTGTCATATCTACGAAGTTGTGTGAGTCCAAGTGCTGGA

NANOG BINDING SITE
DELETED SEQUENCE
GUIDE RNA

E

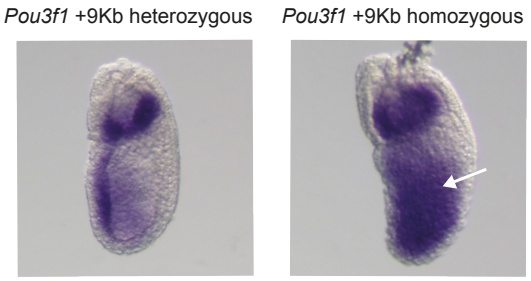


Figure 4

FIGURE LEGENDS

Figure 1: Nanog KO ESC Naïve to Prime differentiation

A. Three time points were taken during ES cells priming: 0h, 12h and 24h per triplicate.

B-C. Graphical model of the putative gene groups: **B** Schematic of genes upregulated in KO conditions during priming of ES cells and **C** genes upregulated in control conditions during priming of ES cells.

D. Volcano plot of wt vs KO cells at time 0h. In blue, we find genes upregulated in *Nanog*^{KO} cells. In orange, genes upregulated in control cells. In grey genes that have less than [0.5] Log Fold Change LogFC. In black genes with more than 0.05 adjusted p-value. Core pluripotency factors are indicated.

E. Genes upregulated in *Nanog*^{KO} across time.

F. Genes upregulated in control across time. Blue lines represent *Nanog*^{KO} ES cells, red lines represent control cells. Thicker lines represent the mean of normalized average CPM of indicated genes.

G. Heatmap comparing both set of genes with representative genes.

Figure 2: *Nanog* RNAseq merging

A. Schematic representation of an E6.5 embryo. Light red indicates *Nanog*'s expression pattern. Black triangle on the embryo points to the diminishing levels of *Nanog* towards the distal region of the embryo. Red and green triangles represent the positive red and negative green correlations between *Nanog* and any other given gene.

B. Correlation values of the genes that show the highest statistical correlation negative in green and positive in red in two different data sets (Mohammed et al., 2017a; Scialdone et al., 2016).

C. List of the most downregulated genes in E7.5 embryos with ectopic expression of *Nanog* when compared with control embryos: Bars mark the log fold change of the differences between *Nanog* induced and control embryos.

D. *In situ* hybridization of *Nanog*^{tg} embryos with and without doxycycline. n=5. Scale bar 300um

E. Venn diagram of the different RNAseq analyzed: in blue are all genes significantly upregulated upon *Nanog* loss of function in ES cells priming, in green genes that are negatively correlated with *Nanog,* and in purple genes downregulated upon *Nanog* ectopic expression. Boxed genes are the only found in all three groups.

Figure 3: RT-qPCR screening of anterior neural differentiation of *Nanog*^{tg} ES cells.

A. RT-qPCR of different genes with timepoints at day 0, 3, 4, 5 and 6 of anterior neural differentiation. In blue differentiation with doxycycline. In gray differentiation without doxycycline (see methods).

B. Immunofluorescence at day 6 of anterior neural differentiation showing DAPI in blue, SOX2 in red and TUJ1 in white.

Figure 4: NANOG-BS deletion in E6.5 by CRISPRCas9 injection in the zygote.

A. Locus of *Pou3f1* containing three NANOG-BS in the EpiLCs after 1 day of differentiation D1 or two days after differentiation D2 but not in the ES cells ESC.

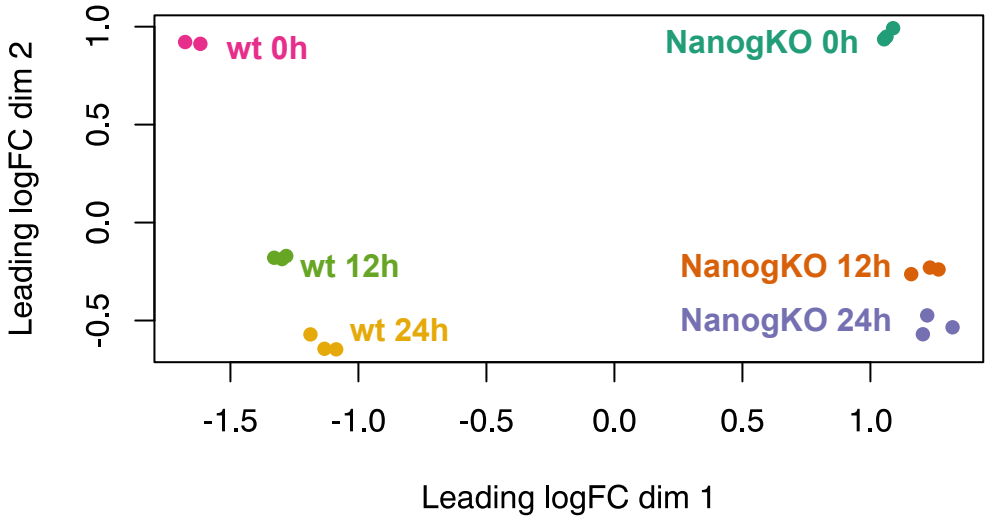
B. Graphical representation of the phenotypic observations in embryos recovered at E6.5 and displaying the fisher test p value below each graph that belongs to the specific deletion.

C. In situ hybridization ISH of E6.5 with wildtype phenotype and expanded deleted phenotype, this presenting the expanded pattern of expression of *Pou3f1* towards the posterior region of the embryo epiblast as indicated by the white arrow.

D. Sequence of the deleted locus. In purple the gRNAs, in grey the deleted region and in black NANOG binding site.

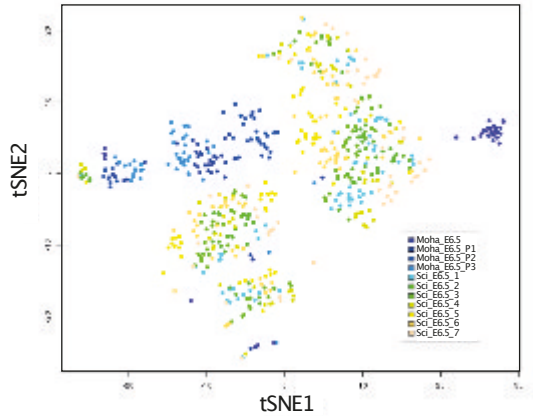
E. *Pou3f1* ISH. White arrows show the expansion of the expression pattern of *Pou3f1* towards the posterior region of the embryo in the embryo homozygous for the deletion.

A

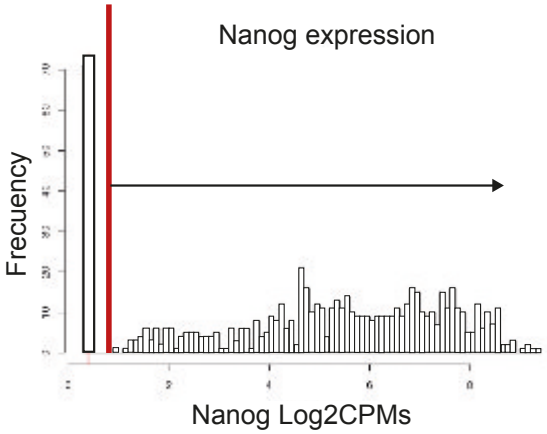


Supplementary Figure 1

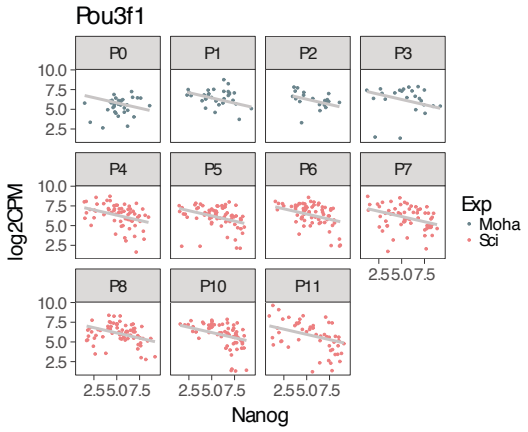
A



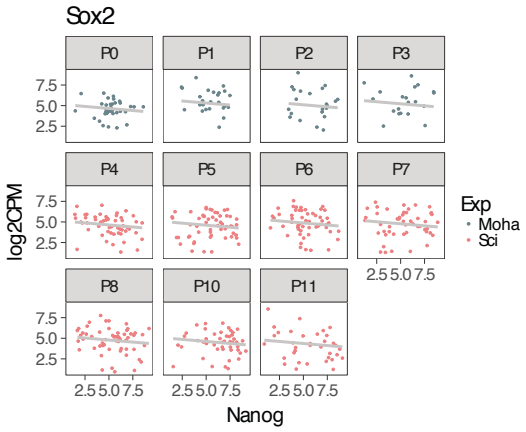
B



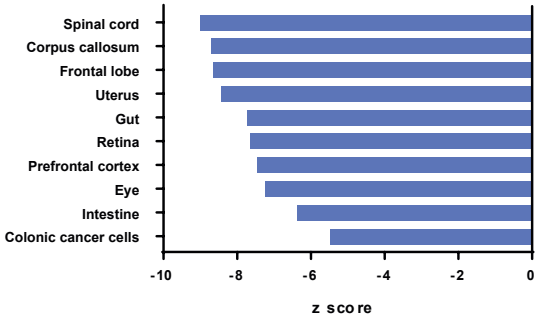
C



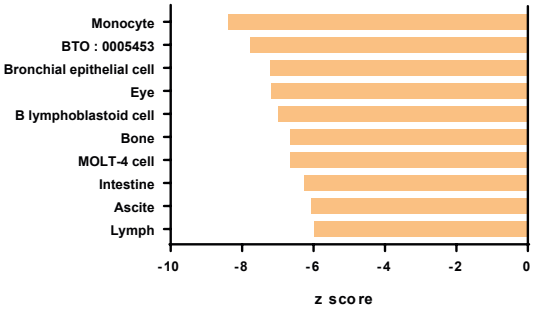
D



E



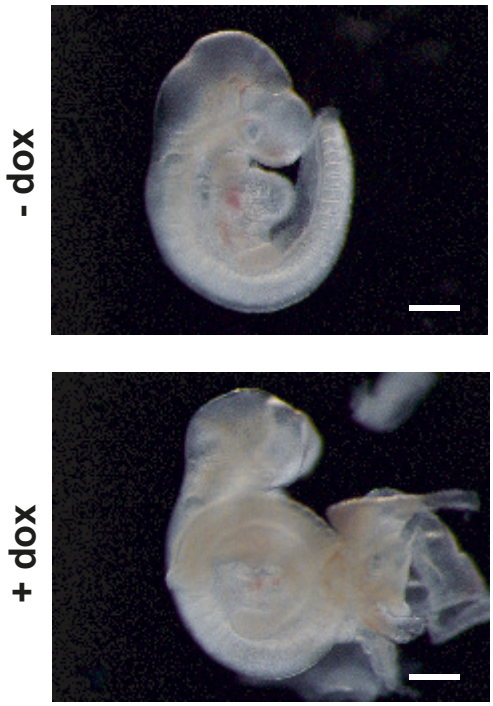
F



G

Intersections	total	Genes	Intersections	total	Genes
down in 7.5 neg corr up in ko	3	Pou3f1	neg corr up in ko	3	Gclm
		Lrp2			Mt3
		Clic6			Dusp4
up in 7.5 up in ko	3	Serpinb9f	down in 7.5 up in ko	1	Ctsh
		Serpinb9e	neg corr up in wt	1	Ccdc64b
		Rassf4	down in 7.5 up in wt	1	Maf
neg corr up in 7.5	9	Nid1	pos corr up in 7.5	16	Dact1
		Krt19			Fes
		Scamp5			Ahcy
		Tmem14c			Eomes
		Nxn			Nefl
		Skil			Dppa5a
		Enc1			Tdgf1
		Car2			Palm
		Lgals1			Rps2
down in 7.5 neg corr	13	Rbp4			Rpl21
		3110002H16Rik			Upp1
		Crabp2			Fgf8
		Tmem181b-ps			Epha2
		Npl			Mixl1
		Gstm1			Socs2
		Gnb4			Pmepa1
		Hnf4a	down in 7.5 pos corr	5	Bbx
		Slc7a8			Wls
		Ddah1			Hmga1-rs1
		Igf2			Tdh
		Cubn			Slc25a36
		Cd59a			

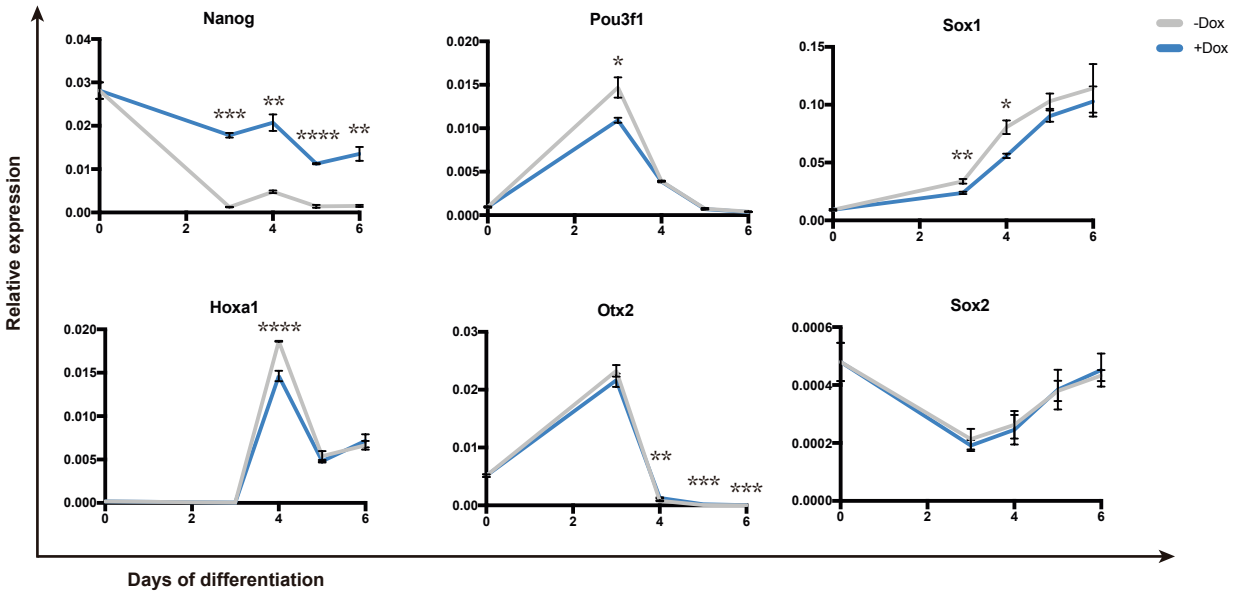
H



Supplementary 2

A

Hindbrain differentiation of Nanog^{GoF} ESC



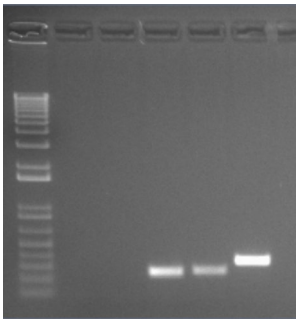
Supplementary 3

A

-9Kb locus

```
AGGACCCCGCAGCGTCCCCATCTAAGCTAAGTCAGGATCTATATTCCCTG
GGAAAGGGTCTGGCCTGTGCTGGAACAGTCATCTCCACTAGATTTAGCCT
GCCCACTACAGAGGCTAGTATGCCCGGGTTTGTGGATGTCCCTGGCCTTC
CCTGATTTTCCCTGCTCCTACTGGCACGAATGATTTCCTCTCCTACCATT
CTATTCATCTTCATTTTGGTCGTGGATCAAGGCTATTAGGCAAAACTGAG
CCCTAATAGGCCCCCCTCCTAGTGCTCTGGGTCCATTAACCAGCCCACTT
TGACGACAAGCAATGACCCAGTGTGAATAAGAAGATCTACCCGCTTCTCT
GAAGCGGAAAGCCAGAAACGGCAATATCCTCTCGTCATCCCCCTTTTCAA
AAGAAAGCACTAAGCCAGGCATGGTGGCGCACACCTGCAATCCCAGAACC
CAGGAAGCTGAGGCAGAAGAGCCAGAAGGACCATTGATTCCAAGCCAGCC
TGATTTACATAGCAAGAGCCGGTCGCCAAAAGCTAAGGAAATAAATAAAA
AAACTTAAAAGAGGAAAGTAAGTGGGCACACCTTTAATCCCAGTATTGGA
```

-9Kb PCR gel



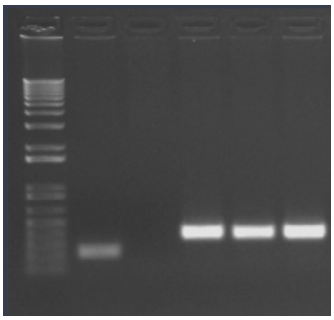
373bp wt
243bp del

B

-11.5Kb locus

```
TGCAGGGCCCTGAAGCCTGACAGAAAAATCACAGACACACTCTGACATTA
ATCAAAGACTCTCATAAATAAATATGTGTAAATCATATGCAAACCAGGTA
GGAATTTACATCTCTCTCTCTCCCCCTGATGAGTCAGCACCTCCAGACAC
CAGGGAGGCTGCCTGCTTGCTGCTCTCTCGGGAGGTTTAGGACAAGAGCC
CTCAACATCCAAGGGAGCCCGCTGAATTCACGATCTCAGATCCTAAATGT
CCATGCCATCAATATAAATCATGCAATTTGCCACAATGCTTTCTCCTACA
CACCAACTACTAAAAGCGTTTAATGAGAAACTATGTGAAGCTCTCAGAAG
GGTCGTTGGTGGTATGCACACAGCTTACCTACACTCTTCTTCATATTATC
TCATCTAGCTTTCACAATAATAGCAATACTCATTACAACAATCATAAAAG
TAATAGGTGTTCTTCTGTCGTGGAGGCTTCCACGACGTCCCTGTGTGCGG
CTCCCGAGGTCCAGTCTTCCACCCCCATCCCTCCAGGCCCTACCATTTCA
GTCCACTCCAGCCTGGCTTCCTACTGCCAGGAACTGCAGCTTTGCCTGAA
GGCTTCTGCAGATCTCCTGGAGGCTGCCCTCCCCTCCTAAGGCAGGCCAC
AAGTGTTGGAGAGTCAGCCCCCACAGCAGAGA
```

-11.5Kb PCR gel



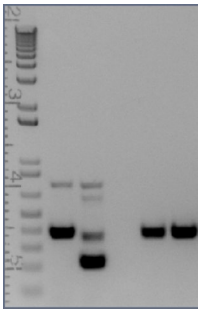
367bp wt
210bp del

C

+9Kb locus

```
AAGTCTGACTGAAGCTTCCTTCAGGTCTCCGTGAGTAGTTTTACTGCAAC
CCTAGCCAGAGGTTAAGCAGAGGTTAACCAGCCTCAGACATTTGGGGTCA
CCAAATGGATCATTTGTCACCTTGGACCTCAAACTAGGAAGTAACCATTC
CTCCATTCACTTATCCATCAGTCCGTTCTCCAGCCAGCTACTCCTCCATT
CACCGAGCCACCCACCTTTGCATGCAGCCATTCACCTACCTACCTATTCA
GCTACCCACCCATCATCTGCCATCCATTCATCCGATCTTCTTGCTTAACG
AATGTCAAGAAGGCTGATGGCCTCTAACCTCCACGCTCTTCTAGCTTTGT
CTCATTCCCTTCCTGCCCCTCAGCCATATAGGGTGTTTGAGAAAACGGTG
TAGAAGGTAATGATGAAACCTGTTGTCATATCTACGAAGTTGTGTGAGTC
CAAGTGCTGGAACTTTCCACACTCCCCTCCTCTCCCCAAAGAAGCTATAG
TTGATACAGCCTGTGGCAAAAGTCTAGATGGCCTTAGTTCTCATTTCTTA
CCCCAGGAATGTCAAGTTCTGGGCTAAATGTAAGGGGTTTCAAAGGTCAC
```

+9Kb PCR gel



386bp wt
200bp del

D

Sequence of deleted +9Kb locus Forward in generated mouse line

```
NNNNNNNNNNNNNTGNNNNNNGGACTNANNCTAGGAAGTAACCATTCCTCC
NTTCACTTATCCATCAGTCCGTTCTCCAGCCAGCTACTCCTCAGCCATAT
AGGGTGTTTGAGAAAACGGTGTAGAAGGTAATGATGAAACCTGTTGTCAT
ATCTACGAAGTTGTGTGNNNNNANNNNNNNNNNNNNANNNNNNNGANNNN
NNNANNTCNNNNNNNNNTNNCNNCNNNNNNNNNNNNNNNNCTATATGGNT
GANGANTANCTGGNNGGANAACGGACTGATGGATAAGTGAATGGANGAAT
GGTTACTTCCTANNNNGNGGNGNNANGNGACNAANGATCCNNNNGGNGAC
CCNNNNNTCNNNNGNNNNNANNNNN
```

Sequence of deleted +9Kb locus Reverse in generated mouse line

```
NNNNNNNNNNNNNGGNTTCNTCNTTACNNTCTACNCCGTTTTCTCAAACAC
CCTATATGGCTGAGGAGTAGCTGGCTGGAGAACGGACTGATGGATAAGTG
AATGGAGGAATGGTTACTTCCTAGTTTGAGGTCCAAGGTGACAAATGATC
CATTTGGTGACCCCAAATGTCTGNNNNNNNNTANANNNNN
```

PCR PRIMERS

GUIDE RNA

SUPPLEMENTARY FIGURE LEGENDS

Supplementary figure 1: PCA showing sample and time distribution.

Two dimensions (dim) are represented: Dim1 genotype (wildtype and NanogKO), dim2 time (from 0h to 24h)

Supplementary figure 2: Analysis of the RNAseq in all samples.

A. tSNE of all the experiments from two different datasets: Moha (Mohammed et al., 2017b) and Sci (Scialdone et al., 2016). Cells from each experiment are represented in different colours according to the legend.

B. Levels of expression of *Nanog* in Log2 counts per million. Red bar and black arrow represent the point from which we start counting cells expressing *Nanog* with >0 CPM.

C-D. Graphic representation of examples of the correlations with *Nanog* on scRNAseq experiments separated by sorted samples P0-P11): On the Y axis we find Log2CPM of Pou3f1 (C) and Sox2 (D). On the X axis Nanog expression in Log2COM is represented. On the right correlation with *Nanog* of each experiment separately. Red cells come from Scialdone experiments (Scialdone et al., 2016) and blue cells come from Mohammed experiments (Mohammed et al., 2017b).

E. Enrichr for negatively correlated genes in E6.5 scRNAseq analysis.

F. Enrichr for positively correlated genes in E6.5 scRNAseq analysis.

G. List of genes included in the different overlap of the venn diagram (Fig 2E). Intersecting columns stablish the gene sets that intersect. Total columns regard the total number of genes that intersect. Gene's columns contain the gene symbols of the gene belonging to a specific intersection.

H. Bright field of E9.5 freshly recovered embryos without doxycycline and with doxycycline presenting cranio facial malformations. Scale bar= 500um

Supplementary figure 3: RT-qPCR screening of hindbrain differentiation of *Nanog*^{tg} ES cells.

A. RT-qPCR of different genes with timepoints at day 0, 3, 4, 5 and 6 of posterior neural differentiation. In blue differentiation with doxycycline. In gray differentiation without doxycycline (see methods).

Supplementary figure 4: NANOG-BS deletion genotyping.

A. -9Kb locus sequence and gel with wildtype and deleted bands. PCR primers are in orange, guide RNA are in purple.

B. -11.5Kb locus sequence and gel with wildtype and deleted bands.

C. +9Kb locus sequence and gel with wildtype and deleted bands. PCR primers are in orange, guide RNA are in purple.

D. Sequence of the deleted +9Kb locus using forward and reverse primers. PCR primers are in orange, guide RNA are in purple.