Pluripotent state transitions coordinate morphogenesis in mouse and human embryos

Marta N. Shahbazi[1], **Antonio Scialdone**[2,†], **Natalia Skorupska**[1], **Antonia Weberling**[1], **Gaelle Recher**[1,†], **Meng Zhu**[1], **Agnieszka Jedrusik**[1], **Liani G. Devito**[3], **Laila Noli**[3], **Iain C. Macaulay**[4,†], **Christa Buecker**[5,†], **Yakoub Khalaf**[3], **Dusko Ilic**[3], **Thierry Voet**[4,6], **John C. Marioni**[2,4,7], and **Magdalena Zernicka-Goetz**[1]

[1]Mammalian Embryo and Stem Cell Group, University of Cambridge, Department of Physiology, Development and Neuroscience, Downing Street, Cambridge CB2 3EG, UK

[2]EMBL-European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge CB10 1SD, UK

[3]Faculty of Life Sciences and Medicine, King's College London, Women's Health Academic Centre, Assisted Conception Unit, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK

[4]Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge CB10 1SA, UK

[5]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, California 94305, USA

[6]Laboratory of Reproductive Genomics, Department of Human Genetics, KU Leuven, Herestraat 49, Leuven 3000, Belgium

[7]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, UK

Abstract

The foundations of mammalian development lie in a cluster of embryonic epiblast stem cells that, in response to extracellular matrix signalling, undergo epithelialization creating an apical surface

in contact with a cavity1,2, a fundamental event for all subsequent development. Concomitantly, epiblast cells transit through distinct pluripotent states3,4, before lineage commitment at gastrulation. These pluripotent states have been characterized at the molecular level5, but their biological importance remains unclear. Here we show that exit from an unrestricted naive pluripotent state is required for epiblast epithelialization and generation of the pro-amniotic cavity in mouse embryos. Embryonic stem cells locked in the naive state are able to initiate polarization but fail to undergo lumenogenesis. Mechanistically, exit from naive pluripotency activates an Oct4-governed transcriptional program resulting in expression of glycosylated sialomucin proteins and the vesicle tethering and fusion events of lumenogenesis. Similarly, culture of human embryos beyond implantation reveals that exit from naive pluripotency triggers amniotic cavity formation and developmental progression. Our results add tissue-level architecture as a new criterion for the characterization of different pluripotent states, and show the relevance of transitions between these states during development of the mammalian embryo.

Keywords

Pluripotency; Epiblast; Embryonic Stem cells; Epithelial tissue

---

We sought to determine the relationship between morphogenesis of the epiblast (that is, its polarization and lumenogenesis) and the transition between distinct pluripotency states, both of which are initiated upon embryo implantation. When we examined the temporal correlation between these events, we found that as expression of the naive pluripotency factor Nanog was downregulated6, the anti-adhesive sialomucin protein podocalyxin7 (Podxl) first became polarized apically and was then secreted when the lumen appeared (Fig. 1a). Deep-sequencing analysis at successive implantation stages revealed the kinetics of pluripotent transitions. We found that at embryonic day (E)4.5 and E4.75 epiblasts grouped together and expressed naive pluripotency genes at similar levels, whereas gene expression in the epiblast at E5.0 was distinct and naive pluripotency genes were downregulated (although *Nanog* expression was already reduced at E4.5–E4.75; Fig. 1b, c, Extended Data Fig. 1a–g and Supplementary Table 1). When considered together with previous datasets8, our findings reveal two distinct groups, namely pre-implantation (E3.5–E4.75) and post-implantation (E5.0–E5.5) epiblast populations. In the second population, naive genes were downregulated to a similar extent at E5.5 compared to E5.0, whereas post-implantation gene expression was higher at E5.5 compared to E5.0 (Extended Data Fig. 1h, i). Therefore, the naive gene expression network is dismantled at lumenogenesis. To determine whether there is a causal relationship between these events, we cultured E4.5 embryos in IVC1 medium9 supplemented with 2i/LIF (consisting of a MEK inhibitor, GSK3 inhibitor and leukaemia inhibitory factor (LIF)), which maintains mouse embryonic stem (mES) cells in the naive state10. We found that 2i/LIF preserved expression of Nanog and inhibited Podxl expression and lumenogenesis (Fig. 1d–f). We confirmed that the naive state was maintained in mES cells derived from embryos cultured in IVC1 medium containing 2i/LIF (Extended Data Fig. 1j, k). Embryos in IVC1 medium containing 2i/LIF did not resume development upon 2i/LIF removal (Extended Data Fig. 1l–n), indicating that the kinetics of naive pluripotency exit are tightly coordinated with morphogenesis.

We next examined the kinetics of polarization and lumenogenesis in relation to naive pluripotency exit using mES cells cultured in Matrigel as a model system for embryogenesis1 (Extended Data Fig. 2a). Upon 2i/LIF removal and after the first cell division, all polarity markers that we examined exhibited polarized localization. Upon subsequent divisions, mES cells organized into polarized rosettes that opened to form lumens 36 h after 2i/LIF removal (Extended Data Fig. 2b–d). This coincided with the loss of naive pluripotency gene expression (Extended Data Fig. 2e–l and Supplementary Videos 1–3), whereas expression of the core pluripotency markers Oct4 (also known as *Pou5f1*) and Sox2 was maintained. Expression of early post-implantation genes, including *Otx2*, was induced, but lineage priming factors were not expressed (Extended Data Fig. 2m–q) similar to the E5.5 epiblast4,8. When 2i/LIF was removed before the cells were plated in Matrigel, lumens formed after 24 h without a rosette (Extended Data Fig. 3a–c), similar to post-implantation-like primed human embryonic stem (hES) cells11,12.

To address whether naive pluripotency exit is required for polarization, we cultured mES cells in Matrigel with 2i/LIF for 24-36 hours (Fig. 2a). Notably, cells displayed polarization of subcellular components despite the maintenance of Nanog levels (Fig. 2b–e and Extended Data Fig. 3d–g), and the expression of pluripotency genes was indistinguishable from mES cells cultured in gelatin, which does not induce polarization (Extended Data Figs 2g–o, 3h–m). Moreover, mES cells obtained from cultures in Matrigel lost their polarized organization, grew as canonical mES cells and contributed to post-implantation development in chimaeras, indicating that their naive state was intact (Fig. 2f–h and Extended Data Fig. 3n–p). As expected13, mES cells cultured without 2i/LIF progressively lost the ability to generate colonies (Fig. 2g). In agreement with the above results, *Dgcr8* knockout mES cells, which remain in a naive state14, underwent polarization. Furthermore, maintenance of naive pluripotency with a PKC inhibitor15 did not impair polarization (Fig. 2i, j and Extended Data Fig. 3q–v). Therefore, mES cells can reversibly initiate polarization without losing their naive character.

Next, we analysed mES cells 48 hours after plating in Matrigel (Fig. 3a). Whereas mES cells cultured without 2i/LIF formed lumens, mES cells cultured with 2i/LIF remained as closed rosettes and expressed the naive markers Nanog and Rex1 (also known as Zfp42); Oct4 was expressed at similar levels in both conditions (Fig. 3b and Extended Data Fig. 4a–c). Electron microscopy revealed that while mES cells formed expanded lumens in the absence of 2i/LIF, mES cells cultured with 2i/LIF displayed either no, or only rudimentary, lumens despite having apically localized tight junctions and Golgi apparatus (Extended Data Fig. 4d, e). Over time, rosettes of mES cells in 2i/LIF lost polarization and became disorganized (Fig. 3c, d and Extended Data Fig. 4f–h). Similar to what we observed in embryos, mES cells cultured in Matrigel with medium containing 2i/LIF did not resume morphogenesis upon 2i/LIF removal (Extended Data Fig. 4i–l). To test whether the lumenogenesis defect was a direct consequence of the pluripotency status, we cultured mES cells with a combination of two supplements (both the MEK and GSK3 inhibitors together (2i); the MEK inhibitor and LIF; and the GSK3 inhibitor and LIF) since these are sufficient to maintain naive pluripotency16. All combinations inhibited lumenogenesis and preserved Nanog expression (Extended Data Fig. 4m–o). When given a single supplement, mES cells forming lumens showed decreased Nanog expression (Extended Data Fig. 4p, q). *Dgcr8*

knockout mES cells expressing high Nanog levels also showed a lumenogenesis defect (Fig. 3e–g and Extended Data Fig. 5a–d). Similarly, PKC inhibition preserved Rex1 and Nanog expression, and impaired lumenogenesis (Extended Data Fig. 5e–h). Notably, when we first induced naive pluripotency exit and then cultured cells in Matrigel with medium containing 2i/LIF, those cells that expressed Podxl, but lacked Nanog expression, formed lumens (Extended Data Fig. 5i–m). Next, we sought to identify the transcription factors that regulate lumenogenesis. Constitutive expression of Nanog enhances self-renewal17. However, we found that *Nanog* overexpression in the absence of 2i/LIF and serum was insufficient to block lumenogenesis and to preserve the naive network (Extended Data Fig. 5n–r). Similarly, *Nanog* downregulation was insufficient to promote lumenogenesis with 2i/LIF, and Rex1 levels remained constant (Extended Data Fig. 5 s–v). We next focused on Oct4, because, together with Otx2, it drives enhancer activation upon naive pluripotency exit18,19. Decreasing *Oct4* expression led to defective lumenogenesis and reduced Otx2 levels (Fig. 3h, i and Extended Data Fig. 5w, x). These results indicate that the pluripotency network directs lumen formation and epithelialization through Oct4.

For further mechanistic insights, we tested whether the fusion of apical vesicles20 mediates lumenogenesis. Cells expressing a dominant-negative form of *Rab11a (Rab11a^{S25N})*21 failed to form lumens (Fig. 4a, b). Apical vesicles carry components involved in lumenogenesis, such as Podxl22. Because Podxl was not expressed in naive conditions (Fig. 1a and Extended Data Fig. 6a–d), we evaluated its role in lumenogenesis using RNA interference. mES cells lacking Podxl showed a lumen formation defect at 48 h, without alterations in naive pluripotency exit or polarization (Extended Data Fig. 6e–k). To determine the long-term consequences of Podxl deficiency, we generated *Podxl* knockout mES cells. *Podxl* knockout mES cells showed defective lumenogenesis after 48 h, but formed lumens by 72 h (Fig. 4c, d and Extended Data Fig. 6l–m), in agreement with the lack of lethality of *Podxl* knockout embryos23. This lumenogenesis delay was rescued by overexpression of Cd34, a sialomucin expressed at very low levels at E5.58 (Extended Data Fig. 6n, p). Moreover, treatment with protamine sulfate, which neutralises the negative charges of sialomucins24, blocked lumen formation at 72 h (Fig. 4e–f), indicating that sialomucins mediate lumenogenesis via cell repulsion. Next, we analysed whether *Podxl* expression was regulated by Oct4–Otx2. We found that *Otx2* knockout mES cells expressed low levels of *Podxl* and displayed a lumenogenesis delay (Extended Data Fig. 7a–d), which was rescued by Otx2 overexpression using a doxycycline (DOX)-inducible system (Extended Data Fig. 7e, f). In agreement with previous results6,18, *Otx2* overexpression in the presence of 2i/LIF was able to induce naive pluripotency exit in 50% of wild-type and 35% of *Otx2* knockout mES cells. These cells formed lumens despite the presence of 2i/LIF (Extended Data Fig. 7g–j). Although *Otx2* knockout mES cells showed a delay in naive pluripotency exit, the lumenogenesis defect was rescued by overexpression of GFP–Podxl (Extended Data Fig. 7k-m). Analysis of Oct4 and Otx2 data18 from chromatin immunoprecipitation–sequencing experiments revealed the presence of an intronic enhancer approximately 3.5 Kb downstream of the *Podxl* transcription start site. In naive conditions this enhancer was bound by Oct4 and did not have activating H3K27ac marks. Naive pluripotency exit led to Otx2 binding and increased H3K27ac modifications (Extended Data Fig. 7n). However, Podxl overexpression was insufficient to rescue lumen formation in naive

cells; instead exogenous Podxl co-localized with Rab11 (Fig. 4g and Extended Data Fig. 7o, p). Analysis of the expression of genes involved in the fusion of Podxl-containing vesicles25 in E4.5–E5.5 epiblasts8 revealed that the tight junction protein cingulin (Cgn) was induced upon naive pluripotency exit. We confirmed this in embryos and mES cells (Fig. 4h and Extended Data Fig. 8a, b). In Madin-Darby canine kidney (MDCK) cells, Cgn mediates lumenogenesis by tethering Rab11 vesicles through association with the Rab11-interacting protein Fip526. Indeed, we found that Fip5 showed polarized localization and associated with Podxl upon 2i/LIF removal (Extended Data Fig. 8c). Depletion of Cgn led to a lumenogenesis defect and accumulation of Podxl in the cytoplasm (Fig. 4i, j), without affecting naive pluripotency exit (Extended Data Fig. 8d, e). Thus, in mES cells Cgn mediates the apical fusion of Podxl-containing Rab11 vesicles.

The human epiblast also transforms into an epithelium enclosing a lumen (the amniotic cavity) at implantation2. However, whether this occurs at the same time as a pluripotency change remains unknown. We found that at E6–7, epiblast cells (GATA6–) expressed the naive factor KLF1727 and did not express PODXL, whereas 4 out of 9 human embryos cultured until E9–1011,28, had a PODXL-coated lumen (Fig. 5a, b). Upon addition of inhibitors and growth factors (5i/LAF: MEK inhibitor, GSK3 inhibitor, RAF inhibitor, Src inhibitor, ROCK inhibitor, human LIF, activin A and bFGF2) that preserve naive pluripotency in hES cells29, embryos failed to form a lumen, displayed high levels of KLF17, NANOG and the naive marker CD13030, and showed a mild increase in apoptosis (Fig. 5b-e and Extended Data Fig. 9a–f).

Finally, we determined the tissue-level characteristics of distinctive human pluripotent states in different naive cultures. hES cells overexpressing NANOG and KLF2 with 2i/hLIF remain in a naive-like state29,31. We found that presence of DOX upregulated naive markers and abolished lumenogenesis, whereas DOX removal or switching to primed conditions induced naive pluripotency exit and lumenogenesis (Extended Data Fig. 10a–e). In transgene-independent naive conditions (RSet and 5i/LAF29), we found that although *NANOG* and *KLF2* were downregulated, naive markers were highly expressed and lumenogenesis was abolished without alterations in the initial polarization (Fig. 5f, g and Extended Data Fig. 10f–m). Switching to primed conditions led to both rosette and lumen formation (Extended Data Fig. 10n, o). Therefore, in human embryos and hES cells, a pluripotent transition is required for amniotic cavity formation (Fig. 5h).

Here we show an evolutionary conserved mechanism that unifies transcriptional and architectural changes of the epiblast. Failure to exit from naive pluripotency leads to impaired cavity formation and this in turn could contribute to high rates of embryo loss at implantation.

Methods

Human embryos

All human embryo experiments were carried out in accordance with Human Fertilization and Embryology Authority (HFEA) regulations (license reference R0075). Informed consent was obtained from all participants, who were informed about the conditions and regulations

that apply within the HFEA code of practice. The project also had local approval by a Research Ethics Committee (UK National Health Service Research Ethics Committee reference 06/Q0702/90). All experimental work conformed to the principles of the WMA Declaration of Helsinki. All experimental work was performed following the HFEA Codes of Practice and the HFEA Act 1990 practices.

Slow-frozen human embryos (day 5–6) were thawed using Quinn's Advantage Thaw Kit (ART-8016, LifeGlobal Group), and vitrified human embryos (day 5–6) were thawed using Kitazato Thawing Solutions (VT802-0, Hunter Scientific) following the manufacturer's instructions. Embryos were cultured in human embryo culture media SAGE 1-Step (67010060A, LifeGlobal Group) for 24 h before Zona pellucida removal with Acidic Tyrode's solution (T1788, Sigma-Aldrich). Embryos were washed in SAGE 1-Step medium and subsequently cultured using a pre- to post-implantation culture method, as previously described[11]. In brief, embryos were plated in IbiTreat µ-plates (IB-80826, Ibidi GmbH) in IVC1 (control) or IVC1 supplemented with 5i/LAF and cultured at 37 °C in 21% O2, 5% CO2. After 24 h of culture, half of the medium was replaced with fresh IVC1. After 48 h of culture, half of the medium was replaced with fresh IVC2. After 72 h of culture, embryos were fixed with 4% PFA (15710, Electron Microscopy Sciences). The exact composition of the medium was as follows. IVC1 medium: Advanced DMEM F12 (12634010, Thermo Fisher Scientific), 20% v/v heat-inactivated Fetal Bovine Serum (FBS) (Stem Cell Institute), GlutaMAX (35050061, ThermoFisher Scientific), 25 U ml^{-1} penicillin–25 µg ml^{-1} streptomycin (15140122, ThermoFisher Scientific), 1×ITS-X (10 mg ml^{-1} insulin, 5.5 mg l^{-1} transferrin, 0.0067 mg l^{-1} sodium selenite, 2 mg l^{-1} etholamine; 51500056, Thermo Fisher Scientific), 8 nM β-oestradiol (E8875, Sigma-Aldrich), 200 ng ml^{-1} progesterone (P0130, Sigma-Aldrich) and 25 µM *N*-aceyl-ʟ-cysteine (A7250, Sigma-Aldrich). IVC2 medium: 20% FBS is substituted for 30% KnockOut Serum Replacement (KSR) (10828010, Thermo Fisher Scientific). 5i/LAF: IVC1 and IVC2 were supplemented with 20 ng ml^{-1} recombinant human LIF (300-05, PeproTech), 20 ng ml^{-1} activin A (Stem Cell Institute), 8 ng ml^{-1} bFGF2 (Stem Cell Institute), 1 µM MEK inhibitor PD0325901 (Stem Cell Institute), 0.5 µM GSK3 inhibitor IM-12 (BML-WN102-0005, Enzo Life Sciences), 0.5 µM RAF inhibitor SB-590885 (S0459, LKT Labs), 1 µM Src inhibitor WH-4-023 (5413, Tocris) and 10 µM ROCK inhibitor Y-27632 (72304, StemCell Technologies).

Embryos that attached to the plates and preserved the epiblast lineage were considered for statistical analyses. From a total of 87 embryos thawed, 8 were fixed at the blastocyst stage (Fig. 5a). Out of these 8 embryos, 4 had epiblast cells as determined by immunofluorescence. The remaining 79 human embryos were cultured in the IVC system. Out of these 79 embryos, 34 survived the thawing procedure, attached to the plates and survived in the *in vitro* culture system. Out of these 34, the epiblast lineage was preserved in 20 embryos (Fig. 5 and Extended Data Fig. 9).

Mouse embryos

Mice were kept in the animal house in accordance with national and international guidelines. All experiments have been regulated by the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge

Animal Welfare and Ethical Review Body (AWERB). Experiments were approved by the Home Office. Animals were inspected daily and those that showed health concerns were culled by cervical dislocation.

Mouse embryos were recovered at embryonic days E4.5 (09:00) and E4.75 (16:00) from naturally mated F1 (C57BL/6 × CBA) or MF1 females by flushing the uterus with M2 medium. Embryos recovered at E5.0 (00:00) and E5.5 were manually dissected from the decidua. Females were used for natural matings at 3 ± 1 month of age.

For mES cell–embryo chimaera experiments, embryos were obtained from F1 females superovulated by injection of 7.5 IU of pregnant mares' serum gonadotropin (PMSG, Intervet), followed by injection of 7.5 IU of human chorionic gonadotropin (hCG, Intervet) and mating with F1 males. Chimeric embryos were transferred to F1 females mated with vasectomized CD1 males. Superovulated females were used at 6 ± 1 week of age.

To culture embryos through the pre- to post-implantation transition, mural trophectoderm of E4.5 mouse embryos was removed as previously described[9]. Embryos were subsequently plated in 8-well μ-Slides (80826, Ibidi) for 24 h in IVC1 medium.

Epiblast dissection at peri-implantation stages

CAG–GFP-expressing mouse embryos[32] were recovered at E4.5, E4.75 and E5.0. The mural trophectoderm (E4.5 and E4.75 embryos) and the extraembryonic ectoderm (E5.0 embryos) were removed with a finely pulled glass pipette. Embryos were then treated for 15 min at 4 °C with Cell Dissociation Buffer (enzyme-free) (13150016, Thermo Fisher Scientific) followed by pipetting with a narrow glass pipette. This treatment removed most of the primitive endoderm, visceral endoderm and polar trophectoderm cells. If remaining GFP⁻ cells were observed, these were separated with a finely pulled glass pipette.

mES cell culture

mES cells were routinely cultured in gelatin-coated plates in N2B27 medium with 2i/LIF at 37 °C, 5% CO_2, 21% O_2. N2B27 medium comprised a 1:1 mix of DMEM F12 (21331-020, Thermo Fisher Scientific) and neurobasal A (10888-022, Thermo Fisher Scientific) supplemented with 1% v/v B27 (10889-038, Thermo Fisher Scientific), 0.5% v/v N2 (homemade), 100 μM β-mercaptoethanol (31350-010, Thermo Fisher Scientific), penicillin–streptomycin (15140122, Thermo Fisher Scientific) and GlutaMAX (35050061, Thermo Fisher Scientific). N2 supplement contained DMEM F12 medium (21331-020, Thermo Fisher Scientific), 2.5 mg ml⁻¹ insulin (I9287, Sigma-Aldrich), 10 mg ml⁻¹ Apo-transferrin (T1147, Sigma-Aldrich), 0.75% bovine albumin fraction V (15260037, Thermo Fisher Scientific), 20 μg ml⁻¹ progesterone (p8783, Sigma-Aldrich), 1.6 mg ml⁻¹ putrescine dihydrochloride (P5780, Sigma-Aldrich) and 6 μg ml⁻¹ sodium selenite (S5261, Sigma-Aldrich). 2i/LIF (1 μM MEK inhibitor PD0325901 (Stem Cell Institute), 3 μM GSK3 inhibitor CHIR99021 (Stem Cell Institute) and 10 ng ml⁻¹ LIF (Stem Cell Institute)) was added to N2B27 medium to preserve naive pluripotency. Exit from naive pluripotency was triggered by removal of 2i/LIF. Where indicated, mES cells were cultured for a minimum of four passages using different combinations of inhibitors (MEK inhibitor (MEKi)–LIF, GSK3

inhibitor (GSK3i)–LIF, MEKi–GSK3i (2i)). Cells were routinely tested for mycoplasma contamination by PCR.

As an alternative way to preserve naive pluripotency, mES cells were cultured on mitomycin C (M4287, Sigma-Aldrich)-treated mouse embryonic fibroblasts (MEFs) in Fc medium supplemented with 5 μM Gö6983 (G1918, Sigma-Aldrich). Fc medium contained DMEM (41966, Thermo Fisher Scientific), 15% FBS (Stem Cell Institute), penicillin–streptomycin (15140122, Thermo Fisher Scientific), GlutaMAX (35050061, Thermo Fisher Scientific), MEM non-essential amino acids (11140035, Thermo Fisher Scientific), sodium pyruvate (11360070, Thermo Fisher Scientific) and 100 μM β-mercaptoethanol (31350-010, Thermo Fisher Scientific).

Cells were routinely passaged with trypsin-EDTA (25300054, Thermo Fisher Scientific). Fc medium was subsequently added to neutralize the trypsin and cells were centrifuged at 1,000 r.p.m. for 5 min.

The following mES cell lines were used: E14 wild-type mES cells (provided by A. Smith, Stem Cell Institute), 129 wild-type mES cells (provided by J. Hanna, Weizmann Institute of Science), PE-*Oct4*–GFP mES cells[33] (provided by A. Surani, The Gurdon Institute), *Rex1*::GFPd2 mES cells[34] (provided by A. Smith, Stem Cell Institute), Nanog–YFP mES cells[35], Dgcr8-knockout mES cells[14] (provided by J. Hanna, Weizmann Institute of Science), Otx2-knockout mES cells[18], and KH2 DOX-inducible Nanog mES cells (provided by K. Hochedlinger, Harvard Stem Cell Institute).

mES cell–embryo chimaeras

Chimaeras were generated by combining H2B–GFP-expressing mES cells with eight-cell stage mouse embryos. In brief, four-cell stage embryos were recovered by flushing the oviducts with M2 medium. Embryos were cultured in KSOM medium (MR-020P-5F, Millipore) until the eight-cell stage, at which point the zona pellucida was removed by brief treatment with Acidic Tyrode's solution (T1788, Sigma-Aldrich). H2B–GFP-expressing mES cells were cultured in a 3D Matrigel for 24 or 72 h. To obtain a single-cell suspension, the 3D cultures were treated with Cell Recovery Solution (354253, Corning) for 15 min at 4 °C. The Matrigel was broken into small pieces with a scrapper, the resulting suspension was centrifuged and the cell pellet was gently resuspended in trypsin-EDTA (25300054, Thermo Fisher Scientific). After trypsin inactivation, cells were plated in gelatin-coated plates. The resulting mES cell colonies were gently dissociated into small clumps and aggregated with the eight-cell stage embryos in M2 medium. The resulting chimaeras were cultured in KSOM medium until E3.5, when they were transferred into pseudo-pregnant females. Chimaeras were recovered at E6.5.

mES cell derivation

For Nanog–YFP mES cell derivation, eight-cell stage mouse embryos were recovered from the oviducts of pregnant females and cultured in KSOM (MR-020P-5F, Millipore) in the presence of 2i. After 24 h, the medium was changed to N2B27 with 2i/LIF for 48 h. Hatched blastocysts were subsequently plated on mitomycin C-treated MEFs in Fc medium containing 2i/LIF. Two days later, blastocysts outgrowths were trypsinized and plated to

obtain mES cell colonies. Positive colonies were identified by immunofluorescence and PCR using two different sets of primers:

Nanog–YFP PCR1 FW1: AGCCTTGGAATGCTGCTCCGC;   Nanog–YFP PCR1 RV1: CAATTAGAGCTATGCAGAGAA;   Nanog–YFP PCR1 RV2: CCACCCCGGTGAACAGCTCC;   Nanog–YFP PCR2 FW1: AGCCTTGGAATGCTGCTCCGC;   Nanog–YFP PCR2 RV1: CTGGTCTGCAGAGCTAGTTC;   Nanog–YFP PCR2 RV2: CCACCCCGGTGAACAGCTCC.

For derivation of mES cells from implantation cultures, 24 h after culture in IVC1 or IVC1 +2i/LIF embryos were plated on mitomycin C-treated MEFs in Fc medium containing 2i/ LIF. Two days later, the outgrowths were trypsinized and re-plated and the appearance of dome-shaped mES cell colonies was monitored.

hES cell culture

All hES cell experiments were approved by the UK Stem Cell Bank Steering Committee and comply with the regulations of the UK Code of Practice for the Use of Human Stem Cell Lines. The WIBR3   PE-*OCT4–GFP* DOX-inducible *NANOG KLF2* hES cell line29 (provided by R. Jaenisch, MIT) was routinely cultured on irradiated CF-1 MEFs (GSC-6101G, Amsbio) in N2B27 medium with 2i/hLIF and DOX at 37 °C, 5% CO_2, 21% O_2. N2B27 medium comprised a 1:1 mix of DMEM F12 (21331-020, Thermo Fisher Scientific) and Neurobasal A (10888-022, Thermo Fisher Scientific) supplemented with 2% v/v B27 (10889-038, Thermo Fisher Scientific), 1% v/v N2 (17502048, Thermo Fisher Scientific), 100 μM β-mercaptoethanol (31350-010, Thermo Fisher Scientific), penicillin– streptomycin (15140122, Thermo Fisher Scientific), GlutaMAX (35050061, Thermo Fisher Scientific), MEM non-essential amino acids (11140035, Thermo Fisher Scientific), 50 μg ml^{-1} BSA (A3311, Sigma-Aldrich) and 0.5% v/v KSR (10828010, Thermo Fisher Scientific). To preserve naive pluripotency 2i/hLIF and DOX were added to the medium (1 μM MEK inhibitor PD0325901 (Stem Cell Institute), 3 μM GSK3 inhibitor CHIR99021 (Stem Cell Institute), 20 ng ml^{-1} recombinant human LIF (300-05, PeproTech) and 1 μg ml^{-1} doxycycline hyclate (D9891, Sigma-Aldrich). Cells were routinely passaged once or twice per week by treatment with StemPro Accutase reagent (A1110501, Thermo Fisher Scientific) for 3 min, followed by addition of Fc medium and centrifugation for 3 min at 1,000 r.p.m. Cells were routinely tested for mycoplasma contamination by PCR.

To convert naive hES cells to primed conditions, two days after plating the medium was changed to DMEM F12 (21331-020, Thermo Fisher Scientific) supplemented with 15% FBS (10082139, Thermo Fisher Scientific), 5% KSR (10828010, Thermo Fisher Scientific), GlutaMAX (35050061, Thermo Fisher Scientific), MEM non-essential amino acids (11140035, Thermo Fisher Scientific), penicillin–streptomycin (15140122, Thermo Fisher Scientific), 100 μM β-mercaptoethanol (31350-010, Thermo Fisher Scientific) and 4 ng ml^{-1} bFGF2 (Stem Cell Insitute). Four to five days later, and when clear primed colonies were identified, cells were passaged using collagenase–dispase (000000010269638001, Sigma- Aldrich) for 15 min, followed by addition of Fc medium, centrifugation for 3 min at 1,000 r.p.m. and plating in the presence of 10 μM ROCK inhibitor Y-27632 (72304, StemCell

Technologies) for 24 h. Primed hES cells were switched to a MEF-independent culture by using Geltrex (A1569601, Thermo Fisher Scientific)-coated plates and mTESR1 medium (05850, StemCell Technologies). Primed hES cells cultured in mTESR1 were routinely passaged using StemPro Accutase reagent (A1110501, Thermo Fisher Scientific). During the first 24 h after passaging, 10 µM ROCK inhibitor Y-27632 (72304, StemCell Technologies) was added to the medium.

Naive WIBR3 PE-*OCT4*–*GFP* DOX-inducible *NANOG KLF2* hES cells were plated on CF-1 MEFs and two days after plating the medium was switched to transgene-independent naive hES cell medium. Two different formulations were used: 5i/LAF29 consisted of N2B27 medium supplemented with 20 ng ml^{-1} recombinant human LIF (300-05, PeproTech), 20 ng ml^{-1} activin A (Stem Cell Institute), 8 ng ml^{-1} bFGF2 (Stem Cell Institute), 1 µM MEK inhibitor PD0325901 (Stem Cell Institute), 0.5 µM GSK3 inhibitor IM-12 (BML-WN102-0005, Enzo Life Sciences), 0.5 µM RAF inhibitor SB-590885 (S0459, LKT Labs), 1 µM Src inhibitor WH-4-023 (5413, Tocris) and 10 µM ROCK inhibitor Y-27632. In addition, RSet medium (05970, StemCell Technologies) was used.

Transgene-independent hES cells were routinely cultured on CF-1 MEFs (GSC-6101G, Amsbio) and passaged using StemPro Accutase reagent (A1110501, Thermo Fisher Scientific).

mES cell spheroid formation

To induce polarization and lumenogenesis of mES cells, we used Matrigel, because it can mimic the basement membrane *in vitro* that surrounds the epiblast at implantation *in vivo*1. Two different protocols were used36.

Embedded—Pellets containing 20,000 mES cells were resuspended in 20 µl of ice-cold growth factor-reduced Matrigel (356230, BD Biosciences). The solution was placed as a drop in a well of a µ-Slide 8-well ibiTreat (IB-80826, Ibidi) dish and incubated for 5 min at 37 °C to allow the Matrigel to solidify. Next, 300 µl of medium (N2B27 or Fc depending on the experiment) was added to the well.

3D on top—A well of a µ-Slide 8-well ibiTreat (IB-80826, Ibidi) dish was covered with 35 µl of ice-cold growth factor-reduced Matrigel (356230, BD Biosciences) and incubated for 5 min at 37 °C to allow the Matrigel to solidify. In the meantime, 20,000 mES cells were resuspended in N2B27 and the cell suspension was carefully plated on the Matrigel-coated well. When approximately 80% of the cells had attached to the Matrigel (5–10 min after plating), the medium was removed and replaced with N2B27 containing 5% Matrigel. Where indicated mES cells were treated with 50 µg ml^{-1} of protamine sulfate (1101230005, Merck Millipore).

hES cell spheroid formation

hES cells were induced to polarize and form lumens using the 3D on top protocol as previously described11,32. In brief, a single-cell suspension of hES cells was plated on Matrigel-coated µ-Slide 8-well ibiTreat (IB-80826, Ibidi) dishes as described in 'mES cell

spheroid formation'. Upon attachment to the Matrigel-coated surface, the medium was replaced with hES cell medium (mTESR or naive hES cell medium depending on the experiment) containing 5% Matrigel. To avoid single-cell-induced primed hES cell death, 10 μM ROCK inhibitor Y-27632 (72304, StemCell Technologies) was added to the medium for 24 h. To monitor the dynamics of naive pluripotency exit and morphogenesis, naive hES cells were plated in 3D Matrigel in mTESR medium.

siRNA treatments

mES cells were transfected with Lipofectamine RNAiMAX transfection reagent (13778030, Thermo Fisher Scientific) using a reverse transfection protocol. In brief, 100,000 cells were plated on gelatin-coated 12-well plates together with the Lipofectamine RNAiMAX–siRNA mix. In total, 0.5 μl of a 20-μM siRNA solution was used per well. AllStars negative control siRNAs (1027280, Qiagen) were used as control.

For downregulation of *Podxl* expression an equimolar mix of the following siRNAs (Qiagen) was used: SI01383123: AAGAAUGUAAAUGUCUAUUUA; SI01383130: CUGGAAUUUAUUGAGAGAUUA; SI01383137: CCCAAUUUCCAUCUCCUAUAA.

For downregulation of *Cgn* expression an equimolar mix of the following stealth siRNAs (Thermo Fisher Scientific) was used: CgnMSS230393: CCCUCAUCCAUUGCAUCACUGCUUA; CgnMSS230394: GAAGACAGUUCUGCAGUCCACCAAU; CgnMSS230395: GGCUUGCCUUUAUGAGUCCUAGUAA.

For downregulation of *Nanog* expression an equimolar mix of the following siRNAs (Qiagen) was used: SI01323357: AGCCTTGGAATTATTCCTGAA; SI04460869: TGCCAGTGATTTGGAGGTGAA; SI04460883: CAGGTTTCAGAAGCAGAAGTA.

For downregulation of *Oct4* expression the following stealth siRNA (Thermo Fisher Scientific) was used: Pou5f1MSS237605: ACCUUCUCCAACUUCACGGCAUUGG.

Cloning

Cloning procedures were carried out using Gateway technology (Thermo Fisher Scientific). In brief, the fragment of interest was amplified by PCR to introduce attB sites. This was cloned into a pDONR221 vector (gift of J. Silva, Stem Cell Institute) using the BP clonase II (11789020, Thermo Fisher Scientific). The fragment of interest was further subcloned into a pHygro or pBlasticidin vector containing a hygromycin B- or blasticidin-resistance cassette, respectively (gift of J. Silva, Stem Cell Institute) for expression in mammalian cells. The recombination reaction was carried out using the LR Clonase II (11791100, Thermo Fisher Scientific).

GFP–Rab11a^{S25N} (dominant negative): a *Rab11a*^{S25N}–pEGFP plasmid was used as a template for cloning (provided by J. Clarke, King's College London). The fragment corresponding to *GFP–Rab11a*^{S25N} was amplified by PCR using the following primers: *Rab11a* FW: GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGTGAGCAAGGGCGAGGA

G; *Rab11a* RV: GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGATGTTCTGACAGCACTGC.

GFP–Podxl: a *GFP–Podxl*–pEGFP plasmid22 was used as a template for cloning (provided by A. Echard, Institut Pasteur). The fragment corresponding to CD8 tag–VSV-g tag–EGFP–Rhodopsin–rabbit *Podxl* was amplified by PCR using the following primers: *Podxl* FW:

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGCCTTACCAGTGACCGC; *Podxl* RV:

GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGAGGTGCGTGTCTTCCTC.

GFP–Cd34: Mouse *Cd34* was amplified from cDNA using the following primers: *Cd34* FW: ACCACGGAGACTTCTACACAAGG;

Cd34 RV: TCACAGTTCTGTGTCAGCCAC.

Using the GFP–*Podxl*–pEGFP plasmid as a template, the N-terminal region corresponding to CD8–VSV–G–EGFP–rhodopsin was amplified by PCR using the following primers:

GFP FW: ATGGCCTTACCAGTGACCGC; *GFP* RV: GAATTCCGTCGCATTGGAGAA.

The two fragments were joined by overlap PCR and attB sites were added for cloning into the pBlastidicin vector.

H2B–GFP: an *H2B–GFP*-expressing plasmid was used as template for cloning (provided by M. Perez-Moreno, CNIO). The fragment corresponding to *H2B–GFP* was amplified by PCR using the following primers:

H2B-GFP FW:

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGCCAGAGCCAGCGAAGT; *H2B-GFP* RV:

GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACTTGTACAGCTCGTCCATGC.

To generate a wild-type *GFP–Rab11a* construct we performed site-directed mutagenesis using the *GFP–Rab11a*^{S25N}–pEGFP plasmid as a template. A four-primer PCR strategy was designed to correct the mutated site. The following primers were used:

Rab11a FW (external):

GGGGACAAGTTTGTACAAAAAAGCAGGCTTCACCATGGTGAGCAAGGGCGAGGAG;

Rab11a RV (external):

GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAGATGTTCTGACAGCACTGC; *Rab11a* wild type FW (internal):

CTGGTGTTGGAAAGTCTAATCTCTTGTCTCGATTTACTCGAAATGAGTTTAATCTC GAAAG;

Rab11a wild type RV (internal):

GAGACAAGAGATTAGACTTTCCAACACCAGAATCTCCAATAAGGACAACTTTG.

The two fragments obtained were joined by overlap PCR.

CRISPR–Cas9

Genome editing was performed using the CRISPR–Cas9 system as previously described[37]. gRNAs were designed using the MIT CRISPR Design Tool (http://crispr.mit.edu/). A two-gRNA strategy was designed in order to remove the initiation ATG. The following gRNAs were used: gRNA1:CGAACTCCGGAGTCGCGATC; gRNA2: CTGCAAGCGGTCCGACCACG.

The gRNA templates were assembled and ligated into the PX459 vector. The following primers were used: gRNA1 FW: CACCGCGAACTCCGGAGTCGCGATC

gRNA1 RV: AAACGATCGCGACTCCGGAGTTCGC; gRNA2 FW: CACCGCTGCAAGCGGTCCGACCACG; gRNA2 RV: AAACCGTGGTCGGACCGCTTGCAGC.

The two gRNA-expressing constructs were co-transfected into E14 mES cells. After 24 h of transfection, 2 µg ml^{-1} of puromycin was added to the medium for 48 h to select for transiently transfected cells. Single cells were subsequently plated on 96-well plates with mitomycin C (M4287, Sigma-Aldrich)-treated MEFs in Fc 2i/LIF medium. Efficient gene editing was assayed by PCR using the following primers: *Podxl* FW: ACACCTTCGCTCGTCGCTGC; *Podxl* RV: GACTAGGAATAAATCCACGATCTGGCC. This generated a 485-bp wild-type band and a 200-bp knockout band.

Three knockout clones were selected for further analyses by RT–PCR (data not shown) and immunofluorescence.

mES cell transfection

mES cells were transfected using Lipofectamine 3000 (L3000001, Thermo Fisher Scientific) following the manufacturer's instructions. In brief, the day before transfection 50,000 cells were plated on 24-well gelatin-coated plates in N2B27 2i/LIF medium without antibiotics. For transfection, 0.5 µg of a PiggyBac transposon vector (gift of J. Silva, Stem Cell Institute) and 0.5 µg of *GFP–Podxl*–pHygro, *GFP–Cd34*–pBlasticidin, tetO–*Otx2*–puro18, *H2B–GFP*–pHygro, *GFP–Rab11a*–pHygro or *GFP–Rab11a^{S25N}*–pHygro were used. Transfected cells were selected with 200 µg ml^{-1} hygromycin B (10687010, Thermo Fisher Scientific), 10 µg ml^{-1} blasticidin (R210-01, Thermo Fisher Scientific) or 2 µg ml^{-1} puromycin (ant-pr-1, Invivogen) and the resulting colonies were manually picked and expanded.

Immunofluorescence

Cells and embryos were fixed in 4% PBS–paraformaldehyde (PFA) (15710, Electron Microscopy Sciences) for 20 min at room temperature and subsequently washed twice with 0.1% Tween–PBS. To stain centrosomes, cells and embryos were fixed in ice-cold methanol for 5 min at 4 °C. Peremeabilization was done in PBS containing 0.3% Triton X-100 and 0.1 M glycine for 30 min at room temperature. Cells were incubated with primary antibodies (Supplementary Table 2) at 4 °C overnight, followed by incubation with fluorescently conjugated Alexa Fluor secondary antibodies (Thermo Fisher Scientific) for 2 h at room temperature (Supplementary Table 2). Both primary and secondary antibodies were diluted in 1% BSA, 0.1% Tween–PBS.

Images were acquired on an inverted SP5 confocal microscope (Leica Microsystems) with a Leica HC PL APO 1.4 NA 63× oil objective or on an inverted SP8 confocal microscope (Leica Microsystems) with a Leica HC PL APO CS2 1.4 NA 63× oil objective or a Leica Fluotar VISIR 0.95 NA 25 x water objective.

Time-lapse microscopy

Time-lapse images of mES cells were acquired on an inverted SP5 confocal microscope (Leica Microsystems) with a Leica HC PL FLUOTAR 0.5 NA 20.0× dry objective. Cells were imaged in a humidified chamber with 21% O_2 and 5% CO_2. Images were taken at intervals of 30 min with a z step of 4 μm.

Fixation and staining for serial block face imaging

Samples were fixed by immersion with 2% formaldehyde (made from PFA) and 2% vacuum-distilled glutaraldehyde, containing 2 mmol l^{-1} $CaCl_2$, in 0.05 M sodium cacodylate buffer at 4 °C and pH 7.4 for 8 h at 4 °C. Samples were rinsed five times for 3 min in cold cacodylate buffer containing 2 mM $CaCl_2$. After buffer washes, samples were incubated in 1% osmium ferricyanide for 18 h at 4 °C and rinsed five times in distilled-deionized water (DIW). This was followed by a 30-min incubation in 1% thiocarbohydrazide at room temperature and five more rinses in DIW. Samples were then incubated in 1% uranyl acetate in 0.05 maleate buffer pH 5.5 at 4 °C for 48 h. The samples were then rinsed five times for 3 min with DIW at room temperature and dehydrated by incubating twice with 50%, 70%, 90% and 100% ethanol, dry ethanol, dry acetone and dry acetonitrile. The dehydrated samples were infiltrated with Quetol 651 epoxy resin over a period of five days. The resin was cured for 48 h at 65 °C. Thin sections were prepared with a Leica Ultracut S mounted on 200-mesh copper grids and viewed with a Tecnia G2 operated at 200 kV.

Image analysis

All microscopy data were analysed using Fiji software38 (http://fiji.sc). For 3D reconstructions, images were cropped and filtered in Fiji and the rendering was done using Chimera (v.1.8.1) (https://www.cgl.ucsf.edu/chimera/). For all quantifications, laser power and detector gain were maintained constant to quantitatively compare different experimental conditions within a single experiment.

Quantification of nuclear fluorescence intensity in ES cells—DAPI images were binarized and a mask was created to segment the nuclei in 2D. These were saved as regions of interest (ROIs), which were used to delineate nuclei in the channel of interest, and nuclear fluorescence intensity was measured within each ROI. For every ES cell spheroid, a unique averaged fluorescence intensity value was generated. Error bars represent the variability (s.e.m.) across different spheroids from different independent experiments. To compare fluorescence intensity values across different experiments, raw fluorescence intensity values were normalized to the average fluorescence intensity of the corresponding control condition. Using this approach, mES cells were considered to be negative for Oct4 expression (in the *Oct4* siRNA group) or negative for Nanog expression (in the *Nanog* siRNA group) when the average fluorescence of the spheroid was below the 25% percentile of the corresponding control group.

Quantification of Podxl and Cgn—To determine whether a spheroid was negative for Podxl expression (in the *Podxl* siRNA group) or negative for Cgn expression (in the *Cgn* siRNA group), the Podxl or Cgn channels in the control group were used to set a threshold value of intensity. This treshold was applied to the experimental group and the images were binarized. mES cell spheroids that did not show any signal based on the threshold value of intensity were considered to be negative.

Quantification of nuclear fluorescence intensity in embryos—A representative *z* plane was chosen for each embryo and the fluorescence intensity of epiblast pluripotency genes was measured using the DAPI channel to segment the nuclei as mentioned above. To account for changes in fluorescence in the *z* axis, the fluorescence intensity of the gene of interest in the epiblast was normalized to the fluorescence intensity of DAPI in the epiblast. Error bars represent the variability (s.e.m.) across different embryos.

Quantification of fluorescence intensity in time-lapse experiments—To quantify the dynamics of naive pluripotency gene expression, the bright field channel was used to create a mask, which was applied to the channel of interest as mentioned above. For each time step, a single fluorescence value corresponding to the average population was measured. Error bars represent the variability (s.e.m.) across different fields of view. To account for photobleaching during imaging in Extended Data Fig. 2e, the raw fluorescence intensity values of the –2i/LIF experimental group were normalized to the raw fluorescence intensity values of the +2i/LIF control group (Extended Data Fig. 3k–m) throughout the time lapse. All experimental conditions shown in Extended Data Figs 2e, 3k–m were imaged simultaneously.

Analysis of centrosome positions—To analyse the position of centrosomes, the nucleus–nucleus axis and the nucleus–centrosome axis were determined. The angle between both vectors was calculated and the *x*–*y* positions were normalized to the internuclear distance.

Analysis of lumen and rosette formation—To quantitatively determine whether rosettes and lumens form in embryos and ES cell spheroids, we immunostained components of the Par polarity complex (either atypical protein kinase C (aPKC), Par6 and/or Podxl), as

these proteins are known to localize to the lumen in 3D *in vitro* cultures[20] and are used as bona fide luminal markers in different model systems[39,40]. Phalloidin staining was used to reveal cell shapes. Lumens and rosettes displayed a polarized organization, whereas disorganized structures did not show a polarized localization of Par proteins and/or Podxl. In addition, the fluorescence intensity profile was plotted from a line drawn through the location of interest (rosettes or lumens). One peak of fluorescence intensity corresponds to a rosette (a closed spot of fluorescence), whereas two peaks depict a lumen (two walls and a central depression) (Figs 1a, 5d and Extended Data Fig. 2b are shown as examples).

In Extended Data Fig. 9a, the F-actin channel was separated from the Nanog channel by applying a subtract function in Fiji.

RNA extraction and RT–PCR

For RNA extraction of mES cells cultured in Matrigel (3D on top protocol), the Matrigel was first removed by treatment with Cell Recovery Solution (354253, Corning) for 15 min at 4 °C. To facilitate the depolimerization, the Matrigel was manually broken into small pieces. Cells were subsequently centrifuged and pellets were washed once with PBS.

RNA was extracted using TRIzol reagent (15596010, Thermo Fisher Scientific) following the manufacturer's instructions. Subsequently, 1 μg of RNA was used to perform a reverse transcriptase reaction in the presence of random primers (C1181, Promega), dNTPs (N0447S, New England BioLabs), RNase inhibitor (M0314L, New England BioLabs) and M-MuLV reverse transcriptase (M0253L, New England BioLabs). RT–PCR reactions were carried out using Power SYBR Green PCR Master Mix (4368708, ThermoFisher Scientific) on a Step One Plus Real-Time PCR machine (Applied Biosystems). The following program was used: 10 min 95 °C followed by 40 cycles of 15 s 95 °C (denaturing) and 1 min 60 °C (annealing and extension). The primers used are listed in Supplementary Table 3.

Mouse gene expression data were normalized to *Gapdh*; human gene expression data were normalized to *HPRT1*. Error bars represent variability (s.e.m.) across multiple independent experiments.

Library preparation, RNA sequencing and mapping of reads

The SMARTSeq2 protocol[41] was used to amplify mRNA with the addition of ERCC spike-in control (1 μl of 1:1,000,000 dilution of mix 1 (4456740, Ambion) per sample).

Nextera XT (Illumina) was used to generate multiplex sequencing libraries from amplified cDNA. Libraries were sequenced on a HiSeq 2500 running in rapid mode. The paired-end reads were then mapped to both the *M. musculus* genome (Ensembl v.38.77) and ERCC sequences using the default settings in GSNAP (v.2014-10-07). Htseq count[42] (v.0.6.1p1 with default options) was used to count the number of reads mapped to each gene. All samples were analysed in a single experiment to avoid batch effects.

Sample quality assessment

Three metrics were used to assess data quality: the fraction of reads that were mapped, the number of genes that had more than 10 reads per million and the fraction of reads that were

mapped to mitochondrial genes. Extended Data Figure 1b–e displays these metrics, with each being shown as a function of the total read number per sample. A principal component analysis (PCA) was run and outliers were defined as performing worse than average for all three metrics (Extended Data Fig. 1f). These analyses uncovered one clear outlier (sample name 'E4.5_Sample1', highlighted in grey), which was then excluded from all further downstream analyses.

We also controlled for possible contamination with cells from the primitive endoderm. We verified that the E5.0 samples did not show any expression of primitive endoderm markers, such as *Gata6*, *Gata4*, *Sox17*, *Sox7* and *Pdgfra.* As for samples from E4.5 and E4.75 embryos, they all cluster with the epiblast samples previously published in ref. 8 (see Extended Data Fig. 1g), except for one sample (E4.5_Sample3), which was found in the primitive endoderm cluster and was therefore excluded from further analysis (see 'PCA and hierarchical clustering' for details on the clustering algorithm).

The samples that passed the quality check were normalized for sequencing depth using size factors43 calculated on endogenous genes.

PCA and hierarchical clustering

The PCA plots shown in Fig. 1b and Extended Data Fig. 1f, h were generated using the $\log_{10}$-transformed read counts (by summing 1 to avoid infinities) of the union of the 5,000 most highly expressed genes in each sample.

The hierarchical clustering shown in Fig. 1c and Extended Data Fig. 1g was based on the dissimilarity matrix defined as $(1 - \rho) / 2$, where ρ is the Spearman correlation coefficient between pairs of samples or genes. Once the dissimilarity matrix was calculated, the hierarchical tree was generated with the R function 'hclust' with the 'average' agglomeration method.

Differential expression analysis

The Bioconductor package DESeq244 was used to find differentially expressed genes between E5.0 samples and E4.5–E4.75 samples at a 0.1 false discovery rate. Before running DESeq2, genes with an average expression level of less than 1 normalized read counts across all samples were excluded.

Comparing our dataset to data in ref. 8

In Extended Data Fig. 1h, before selecting the genes for the PCA, we pooled our data with the epiblast samples previously published in ref. 8 (FPKM normalized read-counts available from ref. 8) and performed a quantile normalization.

The log fold changes plotted in Extended Data Fig. 1i were computed after adding a pseudo-count of 0.1. The genes used for the clustering in Extended Data Fig. 1g are those differentially expressed between epiblast and primitive endoderm at E4.5 as reported in ref. 8.

Statistical analyses

Statistical analyses were performed using GraphPad Prism (excluding the analysis of the sequencing data). Embryos were randomly allocated to control and experimental groups. Sample size was determined based on previous experimental experience. Investigators were not blinded to group allocation. Qualitative data are presented as a contingency table and were analysed with a χ^2 test. Quantitative data are presented as mean ± s.e.m. and were analysed for normality using a D'Agostino–Pearson omnibus normality test. Data with a Gaussian distribution were analysed using a two-tailed unpaired Student's *t*-test (two groups) or ANOVA (multiple groups) with a Tukey's multiple comparison test. Significant differences in the variance were taken into account using a Welch's correction. Data that did not have a Gaussian distribution were analysed using a Mann–Whitney *U*-test (two groups) or Kruskal–Wallis test (multiple groups) with a Dunn's multiple comparison test. For all quantifications a minimum of two independent experiments were performed. Sequencing data were analysed with standard programs and packages.

Code availability

Code used for this study is available from the corresponding author upon reasonable request.

Data availability

RNA sequencing data are available at Array Express (https://www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-5147. The immunofluorescence and time-lapse data that support the findings of this study are available from the corresponding author upon reasonable request. Source Data for RT–PCR experiments and quantifications of the immunofluorescence data are provided with the paper.
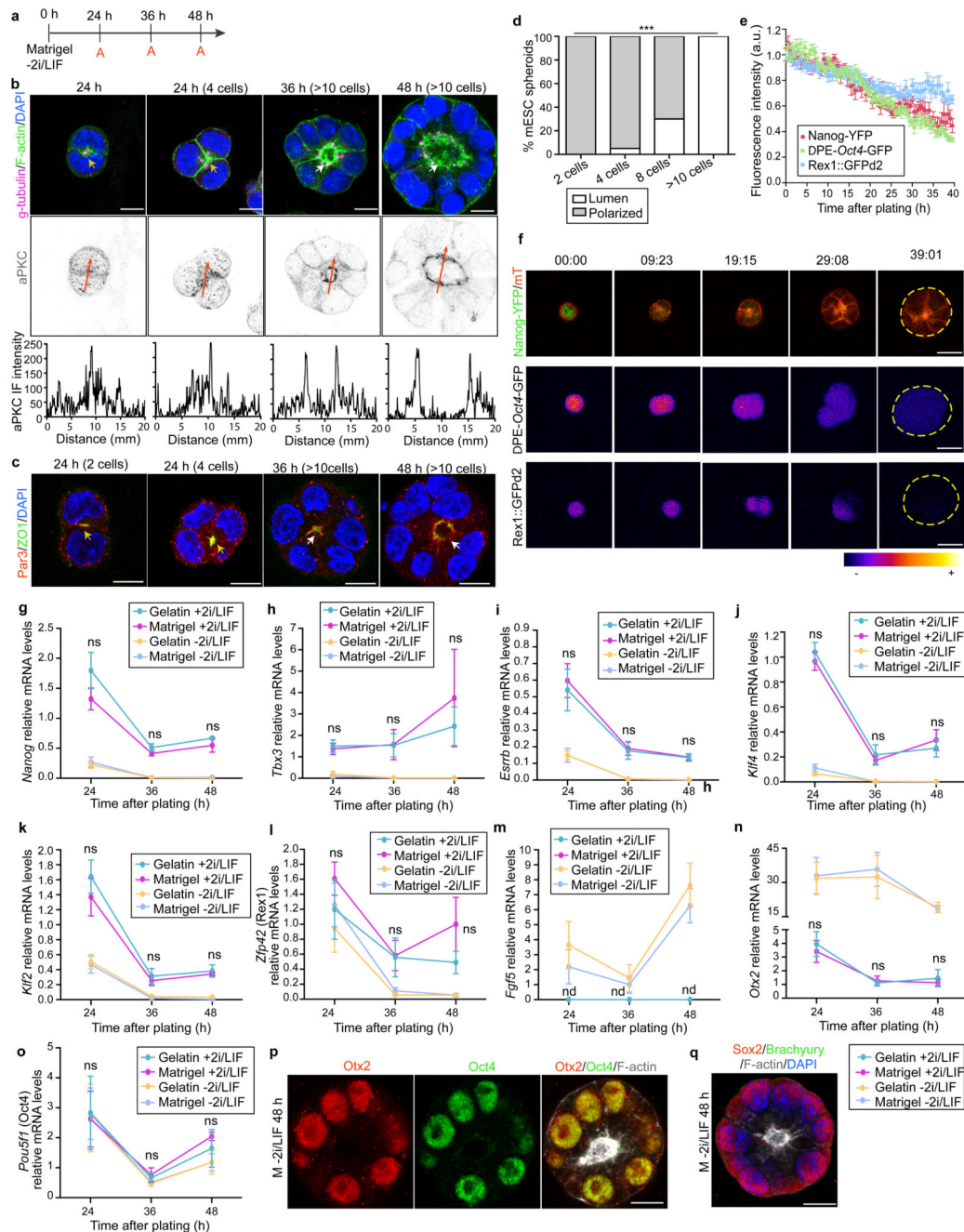
Extended Data
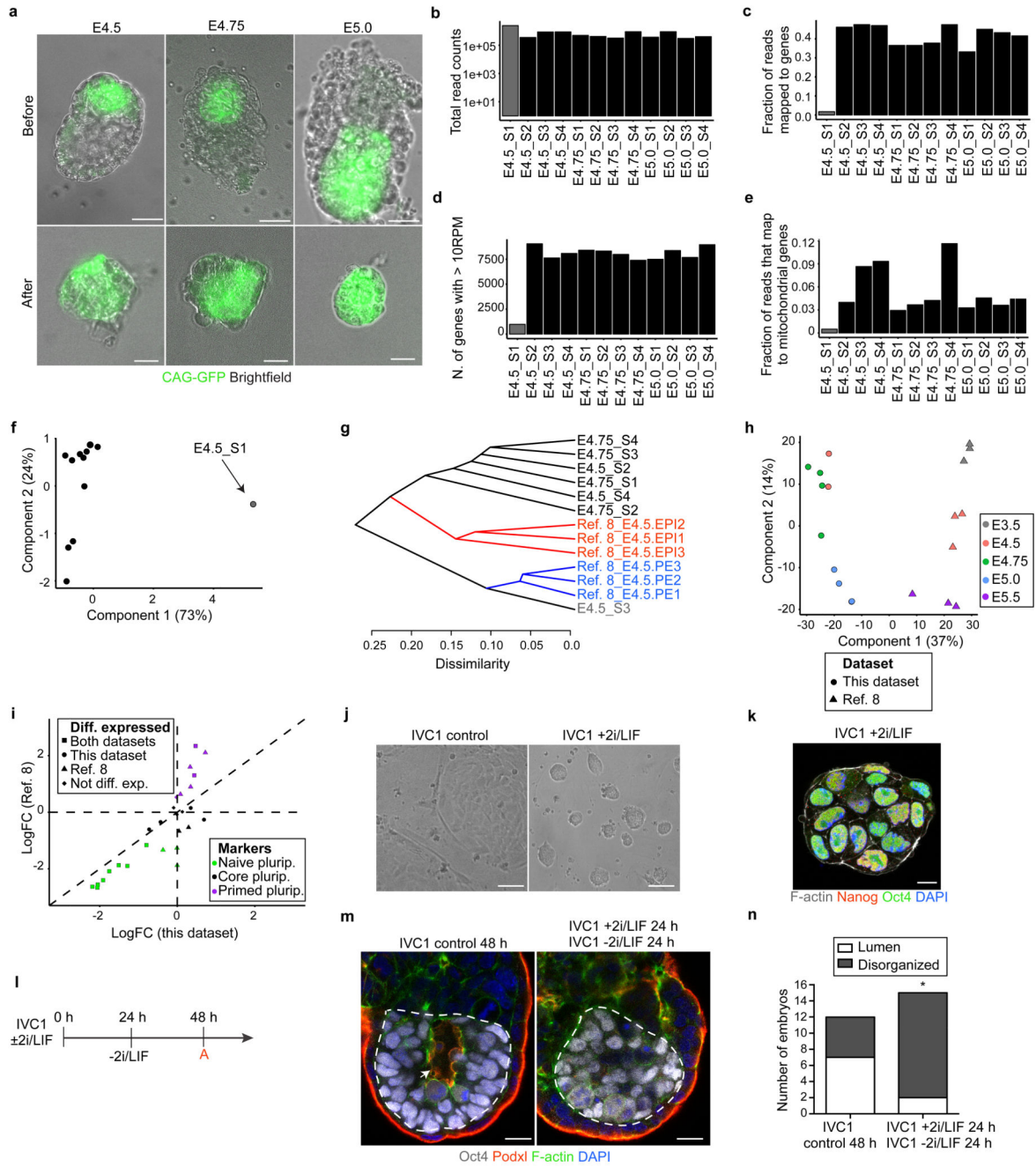
Extended Data Figure 1. Epiblast gene expression patterns at peri-implantation stages.
a, CAG–GFP+ embryos recovered at peri-implantation stages, before and after dissection of
the epiblast. n = 4 E4.5, 4 E4.75 and 4 E5.0 embryos. **b–e,** For all the collected samples we
computed the total number of reads (**b**), the fraction of reads mapped to endogenous genes
(**c**), the number of genes with more than 10 reads per million (RPM) (**d**) and the fraction of
reads mapped to mitochondrial genes (**e**). **f,** A PCA of the metrics shown in b–e. The sample

coloured in grey (E4.5_S1) is characterized by a very low fraction of mapped reads and number of genes detected and was therefore removed from all downstream analysis. **g,** Hierarchical clustering of the E4.5–E4.75 samples that passed the quality check (black) and samples from the epiblast (red) and primitive endoderm (blue) at E4.5 from the dataset of ref. 8. One sample (E4.5_S3, grey) clusters with the primitive endoderm, whereas all others cluster with the epiblast. E4.5_S3 was therefore excluded from further analysis. **h,** PCA of our samples and samples from ref. 8. **i,** log fold change in pluripotency marker genes between E5.0 and E4.5–E4.75 samples from our dataset (x axis) and log fold change between E5.5 and E4.5 epiblast samples from ref. 8 (y axis). Squares, circles and triangles mark genes differentially expressed in both datasets, only in our dataset or only in the dataset of ref. 8 respectively. The naive pluripotency genes significantly downregulated at E5.5 were highly enriched among genes downregulated at E5.0 ($P = 7 \times 10^{-8}$, Fisher's exact test), whereas the enrichment of upregulated early post-implantation factors was only marginally significant ($P = 0.03$, Fisher's exact test). **j,** Bright field images of cells derived from embryos cultured in IVC1 (0/6 mES cell colonies from IVC1 embryos and 4/6 mES cell colonies from IVC1 + 2i/LIF embryos). Scale bars, 50μ m. **k,** Immunostaining of mES cells derived from embryos cultured in IVC1 + 2i/LIF for 24 h. Scale bar, 10 μ m. **l,** Experimental set-up. A, analysis. **m,** Immunostaining of mouse embryos cultured as indicated in l. Dotted lines indicate the epiblast and the arrow points to the pro-amniotic cavity. Scale bars, 20 μm. **n,** Lumen formation in embryos from **m**. n = 12 IVC1 control and 15 IVC1 experimental embryos. $\chi 2$ test, * P = 0.0137.

Extended Data Figure 2. Dynamics of polarization, lumenogenesis and naive pluripotency exit in mES cells.

a, Experimental set-up. A, analysis. **b, c,** Immunostaining of mES cells in 3D Matrigel (**b,** top, **c**). Red long arrows (**b**) indicate the position used to plot intensity profiles (**b,** bottom). Yellow arrows indicate polarization and white arrows lumen formation. Scale bars, 10 μm. **d,** Quantification of lumen formation in cells from **b** and **c**. n = 21 (2 cells), 22 (4 cells), 20 (8 cells) and 21 (> 10 cells) spheroids. χ2 test, * * * P < 0.0001. **e,** Fluorescence intensity in cells from **f**. Raw fluorescence intensity values were normalized to the fluorescence intensity

of reporter cells cultured in gelatin + 2i/LIF throughout the timelapse video. n = 55 Nanog–YFP, 92   PE–Oct4–GFP and 63 Rex1::GFPd2 spheroids. AU, arbitrary units. **f,** Time-lapse frames of naive reporter mES cells cultured in Matrigel without 2i/LIF. mT: membrane Tomato. Time is indicated as h:min. Dotted lines indicate the outline of the spheroids at the end of the experiment. Scale bars, 20 μm. **g–o,** Expression of pluripotency genes in mES cells cultured in gelatin or Matrigel, with or without 2i/LIF. n = 5 (24 h), 3 (36 h), 3 (48 h) independent samples for Nanog, Esrrb and Otx2; n = 5 (24 h), 3 (36 h), 4 (48 h) independent samples for Fgf5; n = 4 (24 h), 3 (36 h), 3 (48 h) independent samples for Klf4 and Klf2; n = 4 (24 h), 3 (36 h), 4 (48 h) independent samples for Zfp42; and n = 3 (24 h), 3 (36 h), 3 (48 h) independent samples for Tbx3 and Pou5f1. Unpaired Student's t-test, NS, not significant. ND, not detected. **p, q,** Immunostaining of mES cells cultured in 3D Matrigel. Scale bars, 10 μm. M, Matrigel.
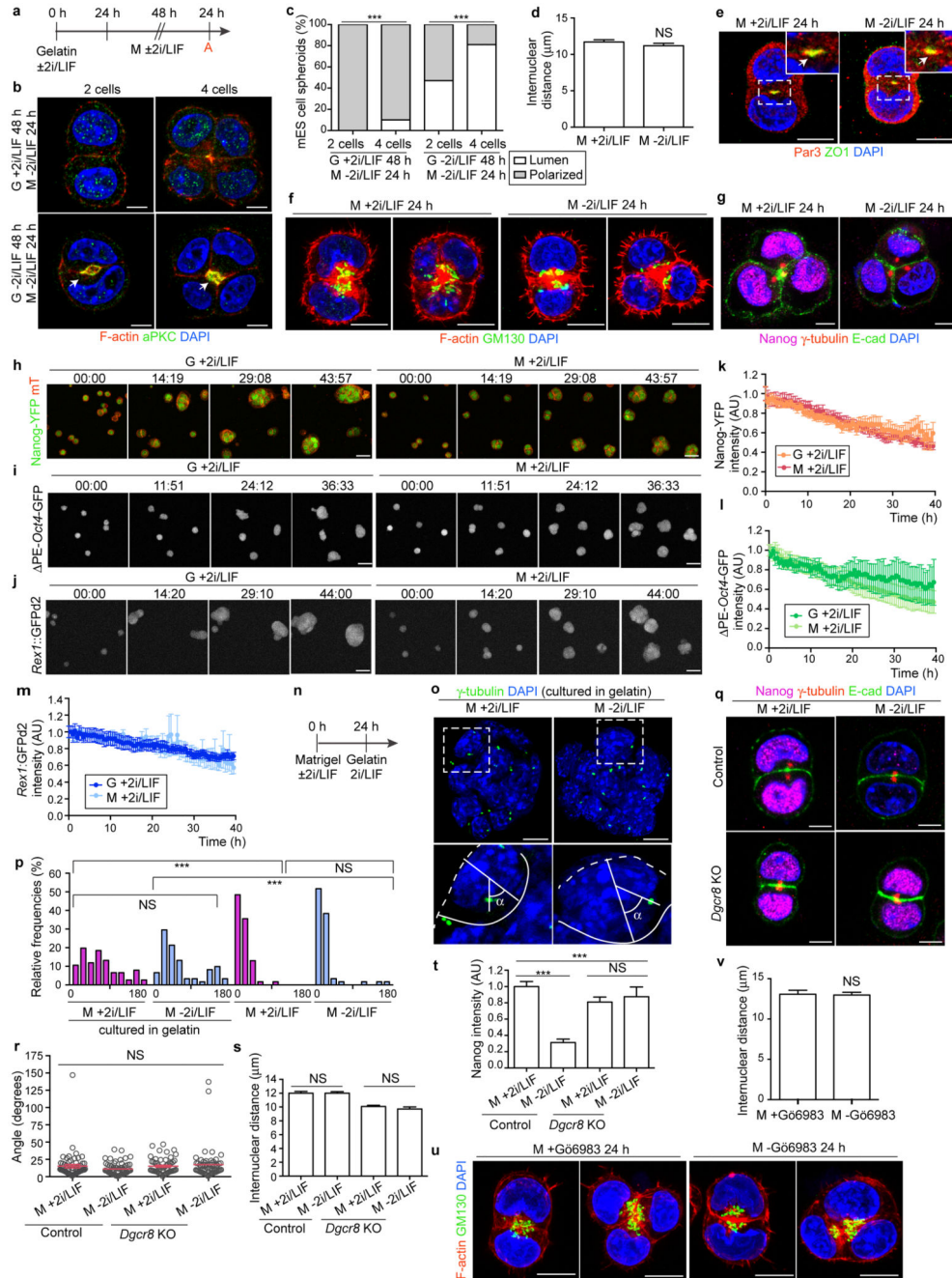
Extended Data Figure 3. mES cells initiate polarization in naïve conditions in a 3D Matrigel culture.

a, Experimental set-up. A, analysis. **b,** Immunostaining of mES cells cultured as indicated in a. Arrows indicate lumens. Scale bars, 5 μm. **c,** Lumen formation in cells from b. n = 31 (2 cells, gelatin + 2i/LIF 48 h, Matrigel – 2i/LIF 24 h), 30 (4 cells, gelatin + 2i/LIF 48 h, Matrigel – 2i/LIF 24 h), 31 (2 cells, gelatin – 2i/LIF 48 h, Matrigel – 2i/LIF 24 h) and 31 (4 cells, gelatin – 2i/LIF 48 h, Matrigel – 2i/LIF 24 h) spheroids. $\chi 2$ test, * * * P < 0.0001. **d,** Internuclear distance in cells from Fig. 2b. n = 31 (+ 2i/LIF) and 30 (– 2i/LIF) spheroids.

Mann–Whitney U-test; NS, not significant. **e–g,** Immunostaining of mES cells cultured in Matrigel with or without 2i/LIF. Squares indicate magnified regions (**e**). Scale bars, 10 μm (**e** and **f**) and 5 μm (**g**). **h–j,** Representative frames from time-lapse experiments using different naive reporter mES cell lines, cultured in either gelatin or Matrigel in the presence of 2i/LIF. Time is indicated as h:min. Scale bars, 50 μ m. **k–m,** Intensity of Nanog–YFP, PE–Oct4–GFP and Rex1::GFPd2 mES cells analysed by time-lapse microscopy from h–j. AU, arbitrary units. n = 65 Nanog–YFP, 55 PE–Oct4–GFP and 43 Rex1::GFPd2 mES cell colonies (gelatin); and n = 61 Nanog–YFP, 84 PE–Oct4–GFP and 61 Rex1::GFPd2 mES cell spheroids (Matrigel). **n,** Experimental set-up. **o,** Immunostaining of mES cells cultured as indicated in n. Squares indicate magnified regions. Dotted lines demarcate the border of the colonies. The nucleus–centrosome angle (α) measured in p is indicated. Scale bars, 10 μm. **p,** Histogram of the angle between the centrosome–nucleus axis and the basal–apical axis of cells in the border of the colonies shown in **o**. Data are shown as relative frequencies (%). n = 76 (Matrigel + 2i/LIF cultured in gelatin), 61 (Matrigel − 2i/LIF cultured in gelatin), 62 (Matrigel + 2i/LIF) and 60 (Matrigel − 2i/LIF) centrosomes. Kruskal–Wallis test; * * * P < 0.0001; NS, not significant. **q,** Immunostaining of control and Dgcr8 knockout mES cells (with or without 2i/LIF). Scale bars, 5 μm. **r,** Angle between the nucleus–centrosome axis and the nuclear–nuclear axis in cells from q. Each dot represents an individual centrosome. n = 60 control + 2i/LIF, 62 control − 2i/LIF, 56 Dgcr8 knockout + 2i/LIF and 58 Dgcr8 knockout − 2i/LIF centrosomes. Kruskal–Wallis test; NS, not significant. **s,** Internuclear distance in cells from q. n = 30 control + 2i/LIF, 31 control − 2i/LIF, 28 Dgcr8 knockout + 2i/LIF and 29 Dgcr8 knockout −2i/LIF spheroids. Kruskal–Wallis test. **t,** Nanog intensity in cells from **q**. n = 30 control + 2i/LIF, 31 control − 2i/LIF, 28 Dgcr8 knockout + 2i/LIF and 30 Dgcr8 knockout − 2i/LIF spheroids. Kruskal–Wallis test; * * * P < 0.0001. **u,** Immunostaining of mES cells cultured in Matrigel with or without Gö6983. Scale bars, 10 μ m. **v,** Internuclear distance in cells from Fig. 2i. n = 26 (+ Gö6983) and 28 (− Gö6983) spheroids. Mann–Whitney U-test; NS, not significant. G, gelatin; M, Matrigel.
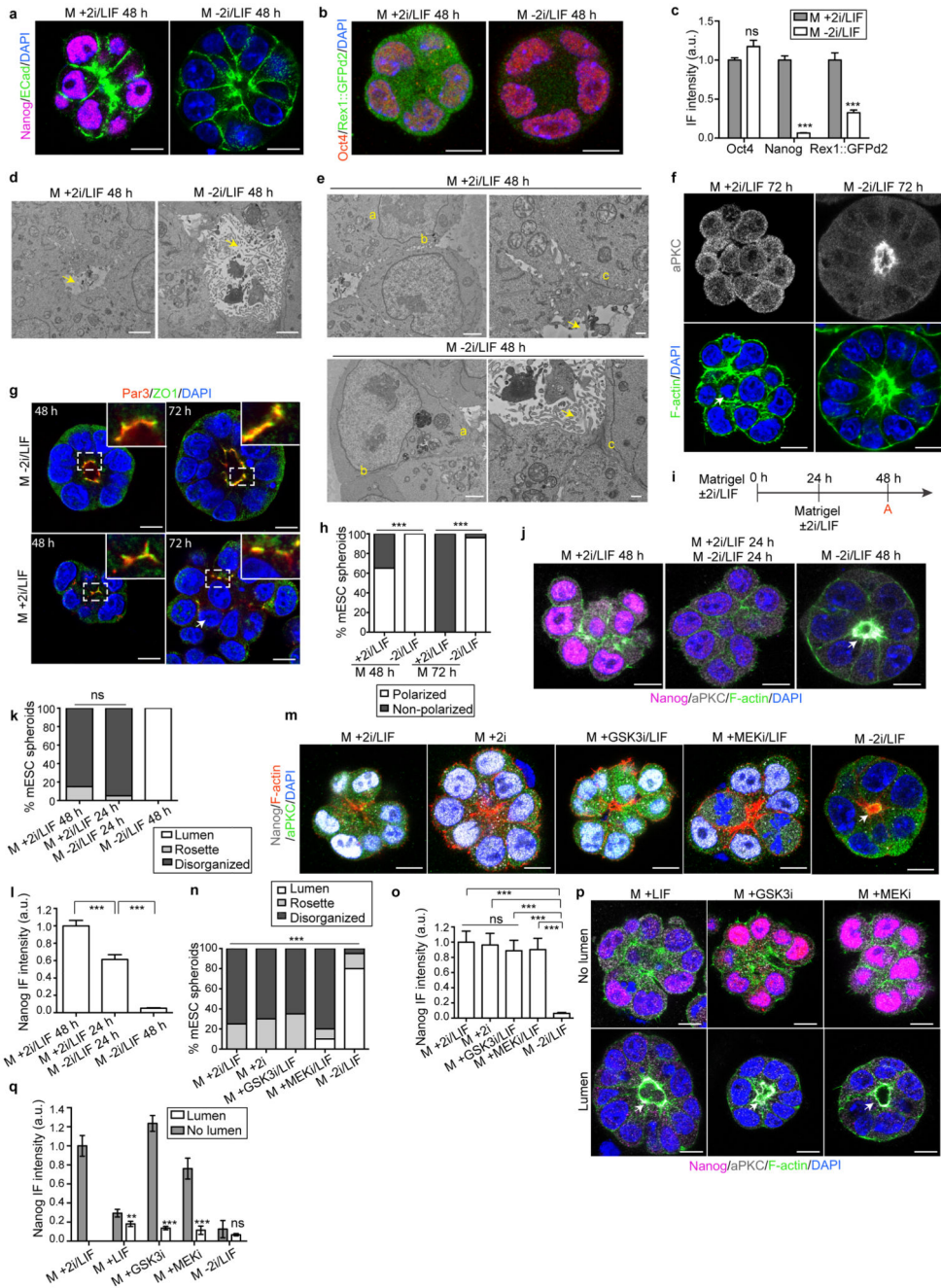
Extended Data Figure 4. 2i/LIF inhibits lumen formation in mES cells.
a, Immunostaining of mES cells cultured in Matrigel. Scale bars, 10 μm. **b,** Immunostaining of Rex1::GFPd2 mES cells cultured in Matrigel. Scale bars, 10 μm. **c,** Oct4, Nanog and Rex1::GFPd2 fluorescence intensity in cells from a and b. n = 30 spheroids per condition (Nanog); n = 21 (+ 2i/LIF) and 20 (− 2i/LIF) spheroids (Oct4); and 21 (+ 2i/LIF) and 20 (− 2i/LIF) spheroids (Rex1::GFPd2). Unpaired Student's t-test; * * * P < 0.0001; NS, not significant. **d, e,** Electron microscopy images of mES cells cultured in Matrigel with or without 2i/LIF. Arrows indicate lumens. a, Golgi apparatus; b, basolateral cell–cell adhesion

sites; c, tight junctions. Scale bars, 2 μm (**d**) and 500 nm (**e**). **f, g,** Immunostaining of mES cells cultured in Matrigel. Squares indicate magnified regions (**g**). Arrows indicate inner non-polarized cells. Scale bars, 10 μ m. **h,** Polarization in cells from **g** and Fig. 3c. n = 27 (+ 2i/LIF 48 h), 26 (− 2i/LIF 48 h), 24 (+ 2i/LIF 72 h) and 24 (− 2i/LIF 72 h) spheroids. χ2 test; * * * P < 0.0001. **i,** Experimental set-up. A, analysis. **j,** Immunostaining of mES cells cultured as indicated in **i**. Arrow points to lumen. Scale bars, 10 μm. **k,** Lumen formation in cells from **j**. n = 20 spheroids per condition. χ2 test; NS, not significant. **l,** Nanog intensity in cells from **j**. n = 20 spheroids per condition. ANOVA; * * * P < 0.0001. **m,** Immunostaining of mES cells cultured in Matrigel with different combinations of inhibitors. Arrow points to lumen. Scale bars, 10 μm. **n,** Lumen formation in cells from **m**. n = 20 spheroids per condition. χ2 test; * * * P < 0.0001. **o,** Nanog intensity in cells from **m**. n = 20 spheroids per condition. ANOVA; * * * P < 0.0001. **p,** Immunostaining of mES cells cultured in Matrigel with a single inhibitor or supplement. 2i/LIF was replaced by a single inhibitor or supplement 48 h before plating the cells in Matrigel. Arrows indicate lumens. Scale bars, 10 μm. **q,** Nanog intensity in cells from **p** as a function of their ability to undergo lumenogenesis. n = 30 (+ 2i/LIF control), n = 50 (+ LIF), n = 48 (+ GSK3i), n = 51 (+ MEKi) and n = 39 (− 2i/LIF) spheroids. Mann–Whitney U-test; * * P = 0.0055; * * * P < 0.0001; NS, not significant. M, Matrigel.
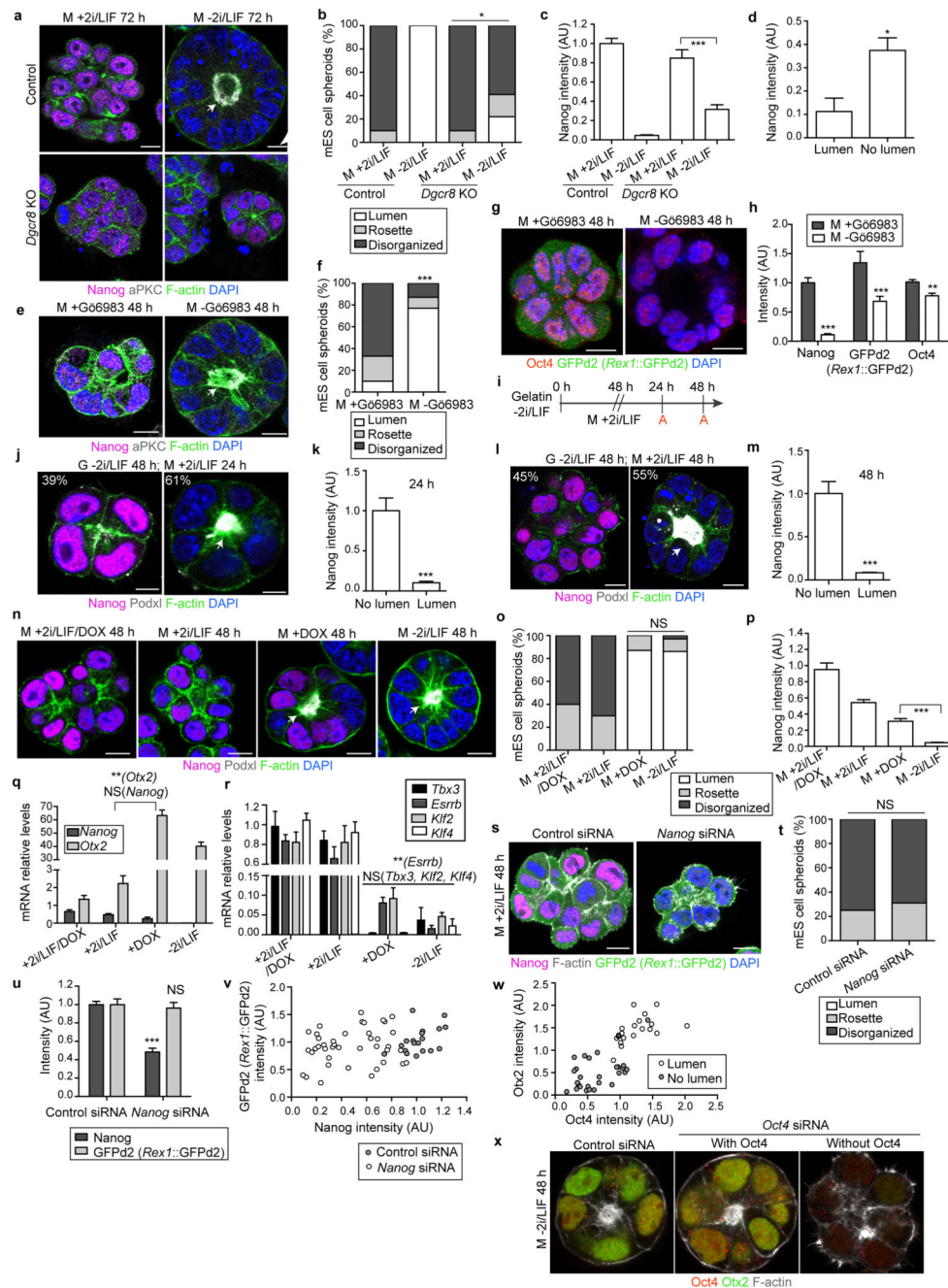
Extended Data Figure 5. Exit from naive pluripotency is required for lumenogenesis in mES cells.

a, Immunostaining of control and Dgcr8 knockout mES cells (with or without 2i/LIF). **b,** Lumen formation in cells from **a**. n = 20 (control + 2i/LIF), 20 (control −2i/LIF), 20 (Dgcr8 knockout + 2i/LIF) and 27 (Dgcr8 knockout − 2i/LIF) spheroids. χ2 test; * P = 0.024. **c,** Nanog intensity in cells from **a**. n = 20 (control + 2i/LIF), 20 (control − 2i/LIF), 20 (Dgcr8 knockout + 2i/LIF) and 27 (Dgcr8 knockout − 2i/LIF) spheroids. ANOVA; * * * P < 0.0001. **d,** Nanog intensity in Dgcr8 knockout mES cells after 72 h of 3D Matrigel culture (**a**), as a

function of their ability to undergo lumenogenesis. n = 6 (lumen) and 21 (no lumen) spheroids. Unpaired Student's t-test; * P = 0.0208. **e,** Immunostaining of mES cells cultured in Matrigel with or without Gö6983. **f,** Lumen formation in cells from **e**. n = 30 spheroids per condition. χ2 test; * * * P < 0.0001. **g,** Immunostaining of Rex1::GFPd2 mES cells cultured in Matrigel with or without Gö6983. **h,** Nanog, Rex1::GFPd2 and Oct4 intensity in cells from **e** and **g**. n = 30 spheroids per condition (Nanog); 41 (+ Gö6983) and 40 (– Gö6983) spheroids (Oct4 and Rex1::GFPd2). Mann–Whitney U-test; * * P = 0.0002; * * * P < 0.0001. **i,** Experimental set-up. A, analysis. **j,** Immunostaining of mES cells cultured as indicated in **i**. The percentage of structures showing the representative phenotype is indicated. **k,** Nanog intensity in cells from j as a function of their ability to undergo lumenogenesis. n = 13 (no lumen) and 20 (lumen) spheroids. Unpaired Student's t-test; * * * P < 0.0001. **l,** Immunostaining of mES cells cultured as indicated in **i**. The percentage of structures showing the representative phenotype is indicated. **m,** Nanog intensity in cells from **l** as a function of their ability to undergo lumenogenesis. n = 14 (no lumen) and 17 (lumen) spheroids. Unpaired Student's t-test; * * * P < 0.0001. **n,** Immunostaining of DOX-inducible Nanog mES cells. **o,** Lumen formation in cells from n. n = 30 (+ 2i/LIF/DOX), 30 (+ 2i/LIF), 31 (+ DOX) and 28 (– 2i/LIF) spheroids. χ2 test, NS, not significant. **p,** Nanog intensity in cells from n. n = 30 (+ 2i/LIF/DOX), 30 (+ 2i/LIF), 31 (+ DOX) and 28 (– 2i/LIF) spheroids. Kruskal–Wallis test; * * * P < 0.0001. **q,** mRNA levels of Nanog and Otx2 in DOX-inducible Nanog mES cells. n = 5 (Nanog) and 3 (Otx2) independent samples per group. Unpaired Student's t-test; * * P = 0.0044; NS, not significant. **r,** mRNA levels of naive pluripotency genes in DOX-inducible Nanog mES cells. n = 4 independent samples per group. Unpaired Student's t-test; * * P = 0.0075; NS, not significant. **s,** Immunostaining of control and Nanog knockdown mES cells. **t,** Lumen formation in cells from s. n = 20 (control siRNA) and 42 (Nanog siRNA) spheroids per condition. χ2 test; NS, not significant. **u,** Nanog and Rex1::GFPd2 intensity in cells from **s**. n = 20 (control siRNA) and 42 (Nanog siRNA) spheroids. **v,** Correlation between Nanog and Rex1::GFPd2 levels in cells from **s**. n = 20 (control siRNA) and 42 (Nanog siRNA) spheroids. **w,** Otx2 and Oct4 intensity in Oct4 knockdown mES cells. n = 19 (lumen) and 28 (no lumen) spheroids. **x,** Immunostaining of control and Oct4 knockdown mES cells. Scale bars, 10 μm (**a, e, g, l, n, s, x**) and 5 μm (**j**); arrows indicate lumens.
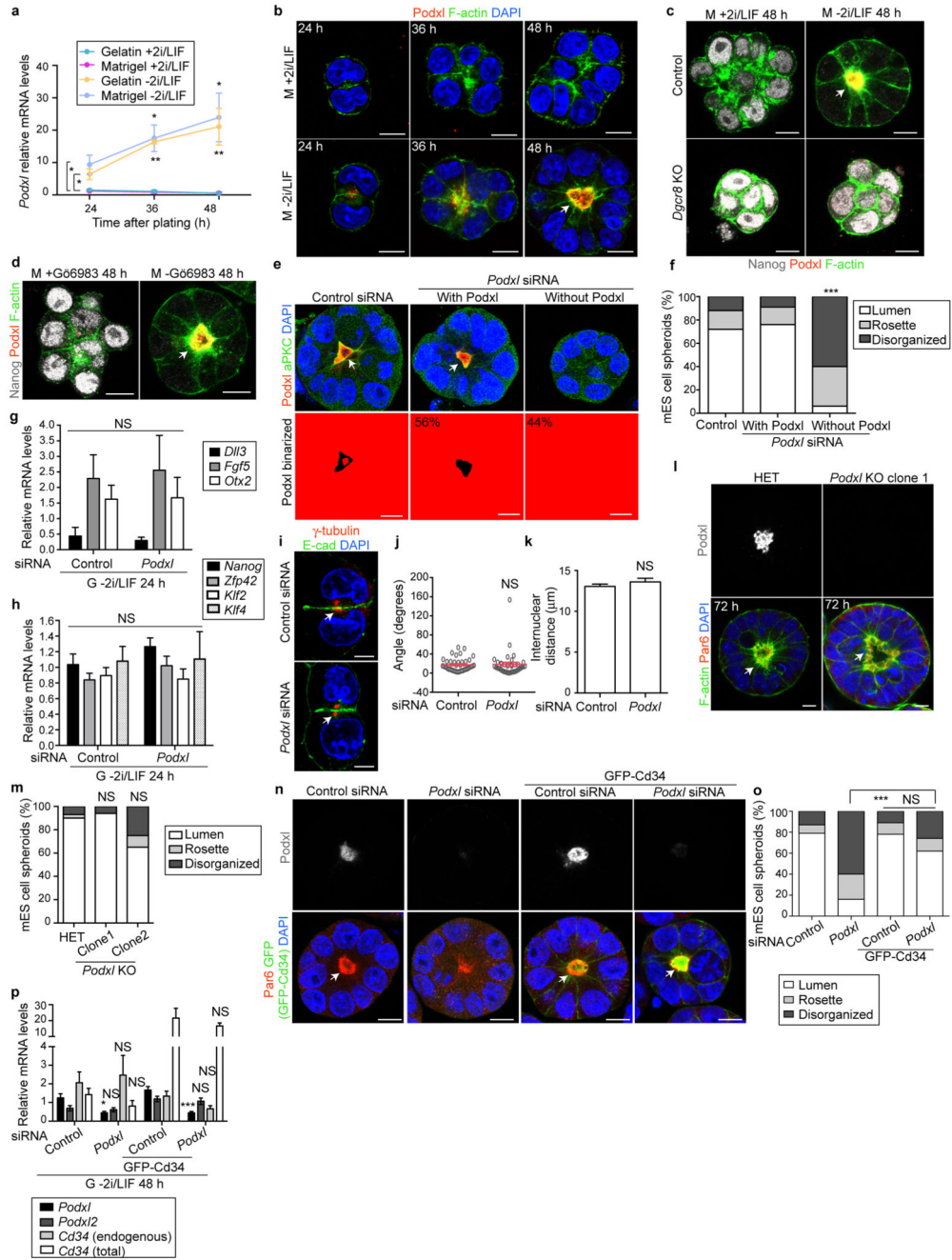
**Extended Data Figure 6. Sialomucins are required for mES cell lumenogenesis.**
a, mRNA levels of Podxl in mES cells. n = 5 (24 h), 3 (36 h) and 4 (48 h) independent samples. Unpaired Student's t-test; * P = 0.0395 (gelatin 24 h); * P = 0.0481 (Matrigel 24 h); * * P = 0.0038 (gelatin 36 h); * P = 0.0126 (Matrigel 36 h); * * P = 0.0075 (gelatin 48 h); * P = 0.0139 (Matrigel 48 h). **b,** Immunostaining of mES cells cultured in 3D Matrigel. Arrow indicates lumen. Scale bars, 10 μm. **c,** Immunostaining of control and Dgcr8 knockout mES cells cultured in Matrigel with or without 2i/LIF. Arrow indicates lumen. Scale bars, 10 μm. **d,** Immunostaining of mES cells cultured in Matrigel with or without

Gö6983. Arrow indicates lumen. Scale bars, 10 μm. **e,** Immunostaining of control and Podxl knockdown mES cells. A binarized image of the Podxl channel was used to determine presence or absence of Podxl. The percentages indicate the proportion of spheroids with or without Podxl in the Podxl siRNA group. Arrows indicate lumens. Scale bars, 10 μm. **f,** Lumen formation in cells from **e**. n = 43 (control), 45 (Podxl siRNA with Podxl (Podxl siRNA-treated cells that did not show a reduction in Podxl expression)) and 35 (Podxl siRNA without Podxl (Podxl siRNA-treated cells that showed a reduction in Podxl expression)) spheroids. χ2 test; * * * P < 0.0001. **g,** mRNA levels of early post-implantation factors in control and Podxl knockdown mES cells 24 h after removal of 2i/LIF. n = 3 independent samples per group. Student's t-test; NS, not significant. **h,** mRNA levels of naive pluripotency genes in control and Podxl knockdown mES cells 24 h after removal of 2i/LIF. n = 3 independent samples per group. Student's t-test; NS, not significant. **i,** Immunostaining of control and Podxl knockdown mES cells cultured in Matrigel. Arrows indicate polarization. Scale bars, 5 μm. **j,** Centrosome positions in cells from **i**. Each dot represents an individual centrosome. n = 40 (control siRNA) and 36 (Podxl siRNA) centrosomes. Mann–Whitney U-test; NS, not significant. **k,** Internuclear distance in cells from **i**. n = 20 (control siRNA) and 18 (Podxl siRNA) spheroids. Unpaired Student's t-test; NS, not significant. **l,** Immunostaining of Podxl heterozygous (HET) and knockout mES cells. One knockout clone is shown as an example. Arrows indicate lumens. Scale bars, 10 μm. **m,** Lumen formation in cells from **l**. n = 30 (HET), 33 (clone 1) and 20 (clone 2) spheroids. χ2 test; NS, not significant. **n,** Immunostaining of control and Podxl knockdown mES cells with or without GFP–Cd34 **o**verexpression. Arrows indicate lumens. Scale bars, 10 μm. **o,** Lumen formation in cells from n. n = 37 (control siRNA), 29 (Podxl siRNA), 33 (control siRNA + GFP–Cd34) and 31 (Podxl siRNA + GFP–Cd34) spheroids. χ2 test; * * * P < 0.0001; NS, not significant. **p,** mRNA levels of sialomucin proteins in control and Podxl knockdown mES cells (with or without GFP–Cd34 overexpression) 48 h after removal of 2i/LIF. n = 4 (without GFP–Podxl) and 6 (with GFP–Cd34) independent samples per group. Unpaired Student's t-test; * P = 0.0145; * * * P = 0.0001. G, gelatin; M, Matrigel.
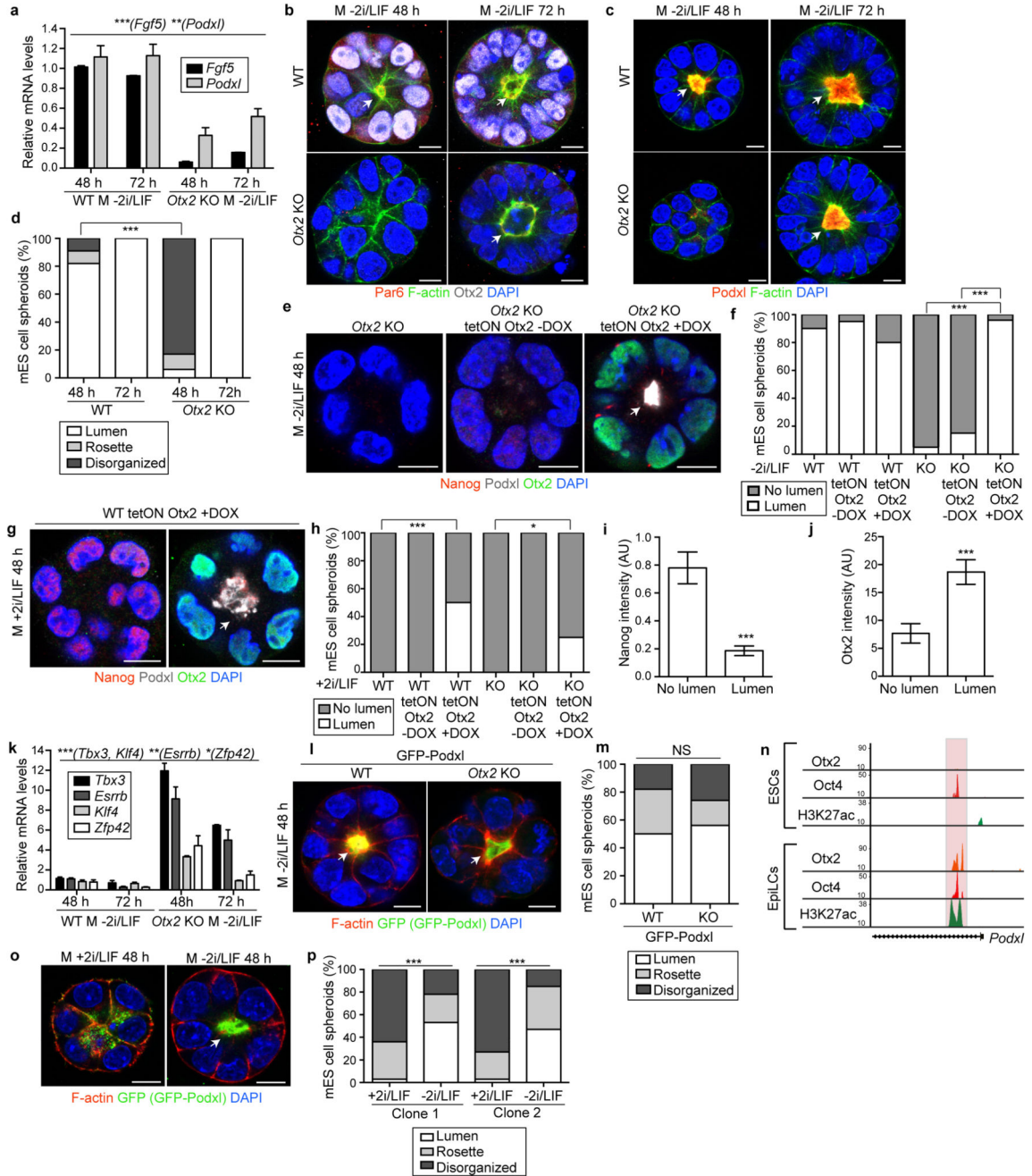
Extended Data Figure 7. Otx2 and Oct4 induce Podxl expression upon naive pluripotency exit.
a, mRNA levels of Fgf5 and Podxl in wild-type and Otx2 knockout mES cells at different time points after removal of 2i/LIF. n = 2 independent samples per group. ANOVA; * * P = 0.0092; * * * P < 0.0001. **b, c,** Immunostaining of wild-type and Otx2 knockout mES cells cultured in Matrigel. **d,** Lumen formation in cells from b and c. n = 46 (wild-type 48 h), 34 (wild-type 72 h), 36 (Otx2 knockout 48 h) and 35 (Otx2 knockout 72 h) spheroids. χ2 test; * * * P < 0.0001. **e,** Immunostaining of Otx2 knockout and Otx2-knockout tetON-Otx2 mES cells in the presence or absence of DOX (DOX addition triggers Otx2 overexpression). **f,**

Lumen formation in cells from e and in wild-type mES cells cultured without 2i/LIF and with or without Otx2 overexpression. n = 19 (wildtype), 20 (wild-type tetON-Otx2 – DOX), 20 (wild-type tetON-Otx2 +DOX), 20 (Otx2 knockout), 20 (Otx2 knockout tetON-Otx2 – DOX) and 23 (Otx2 knockout tetON-Otx2 + DOX) spheroids. $\chi 2$ test; * * * P < 0.0001. **g,** Immunostaining of wild-type tetON-Otx2 mES cells overexpressing Otx2 (addition of DOX). h, Lumen formation in cells from **g** (and their corresponding controls) and in Otx2 knockout mES cells cultured with 2i/LIF and with or without Otx2 overexpression. n = 20 (wild-type), 20 (wild-type tetON-Otx2 – DOX), 50 (wild-type tetON-Otx2 + DOX), 19 (Otx2 knockout), 21 (Otx2 knockout tetON-Otx2 –DOX) and 20 (Otx2 knockout tetON-Otx2 + DOX) spheroids. $\chi 2$ test; * P = 0.0196; * * * P < 0.0001. **i, j,** Nanog (**i**) and Otx2 (**j**) intensity in cells from **g** as a function of their ability to undergo lumogenesis. n = 25 spheroids per condition. Mann–Whitney U-test; * * * P < 0.0001. **k,** mRNA levels of naive pluripotency genes in wild-type and Otx2 knockout mES cells at different time points after removal of 2i/LIF. n = 2 independent samples per group. ANOVA; * P = 0.02; * * P = 0.0045; * * * P = 0.0002 (Klf4); * * * P = 0.0001 (Tbx3). **l,** Immunostaining of wild-type and Otx2 knockout mES cells overexpressing GFP–Podxl. **m,** Lumen formation in cells from **l**. n = 66 (wild-type) and 61 (Otx2 knockout) spheroids. $\chi 2$ test; NS, not significant. n, ChIP–seq analysis of Oct4, Otx2 and H3K27ac in mES cells (top) and after 48 h of transition into epiblast-like cells (EpiLCs, bottom). Shaded area indicates an enhancer within the first intron of Podxl that is activated during cell fate transition. **o,** Immunostaining of GFP–Podxl overexpressing mES cells. **p,** Lumen formation in cells from **o**. n = 30 (+ 2i/LIF clone 1), 32 (– 2i/LIF clone 1), 33 (+ 2i/LIF clone 2) and 34 (– 2i/LIF clone 2) spheroids. $\chi 2$ test; * * * P < 0.0001. Scale bars, 10 μm; arrows indicate lumens; (**b, c, e, g, l, o**). M, Matrigel.
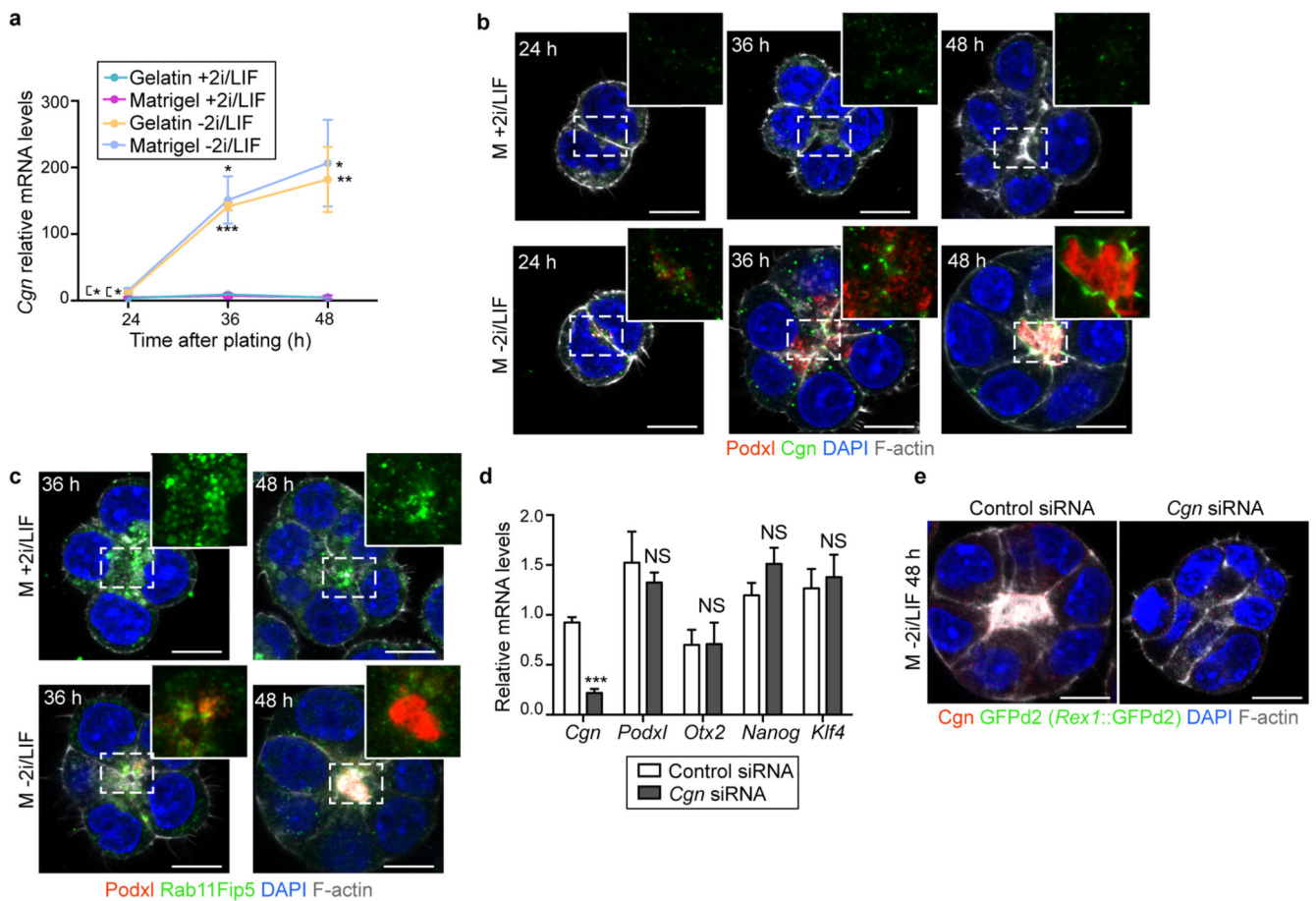
Extended Data Figure 8. Cgn is induced upon naive pluripotency exit and mediates Rab11 vesicle tethering to the apical membrane.

a, mRNA levels of Cgn in mES cells. n = 5 (24 h), 3 (36 h) and 4 (48 h) independent samples per group. Unpaired Student's t-test; * P = 0.0431 (gelatin 24 h); * P = 0.0347 (Matrigel 24 h); * * * P = 0.0003 (gelatin 36 h), * P = 0.0126 (Matrigel 36 h); * * P = 0.0075 (gelatin 48 h); * P = 0.0139 (Matrigel 48 h). **b, c,** Immunostaining of mES cells cultured in Matrigel with or without 2i/LIF. Squares indicate magnified regions. **d,** mRNA levels of naive pluripotency genes and early post-implantation factors in control and Cgn knockdown mES cells. n = 3 independent samples per group. Unpaired Student's t-test; * * * P < 0.0001; NS, not significant. **e,** Immunostaining of Rex1::GFPd2 mES cells in control and Cgn knockdown mES cells. Scale bars, 10 μm (**b, c, e**). Matrigel (M).
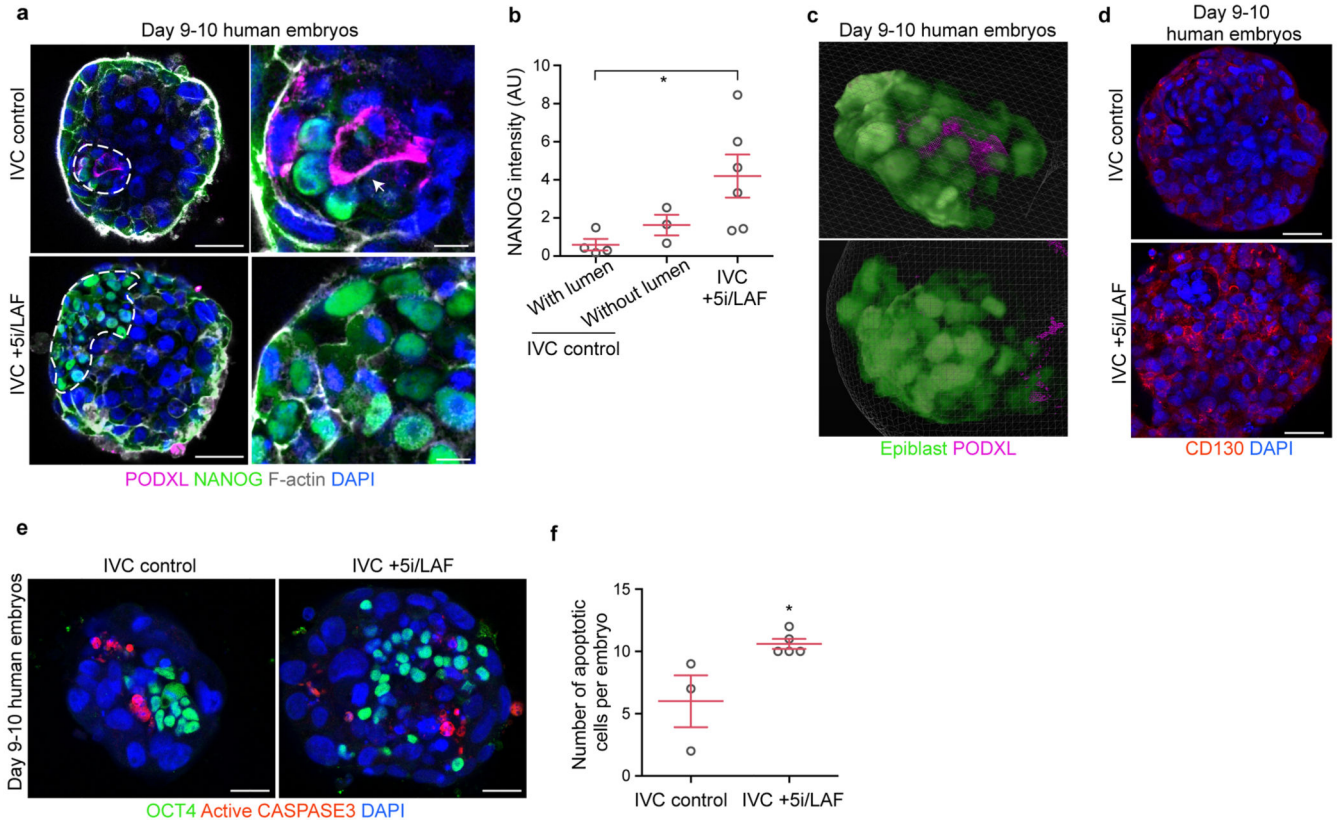
Extended Data Figure 9. Characterization of pluripotency gene expression and epiblast morphogenesis in post-implantation human embryos.

a, Immunostaining of day 9–10 human embryos. Dotted lines indicate the epiblast and the arrow indicates the amniotic cavity. Scale bars, 50 μm and 10 μm (magnified areas). **b,** NANOG intensity in embryos from **a**. n = 4 (IVC control with lumen), 3 (IVC control without lumen) and 6 (IVC +5i/LAF) human embryos. Mann–Whitney U-test; * P = 0.0381. **c,** Three-dimensional reconstruction of the epiblast (based on the KLF17 staining) and the amniotic cavity (based on the PODXL staining) of embryos from Fig. 5b. **d,** Immunostaining of day 9–10 human embryos. n = 2 embryos per group. Scale bars, 50 μm. **e,** Immunostaining of day 9–10 human embryos. Scale bars, 50 μm. **f,** Number of apoptotic cells per embryo in embryos from e. n = 3 (IVC control) and 5 (IVC + 5i/LAF) embryos per condition. Mann–Whitney U-test; * P = 0.0179.
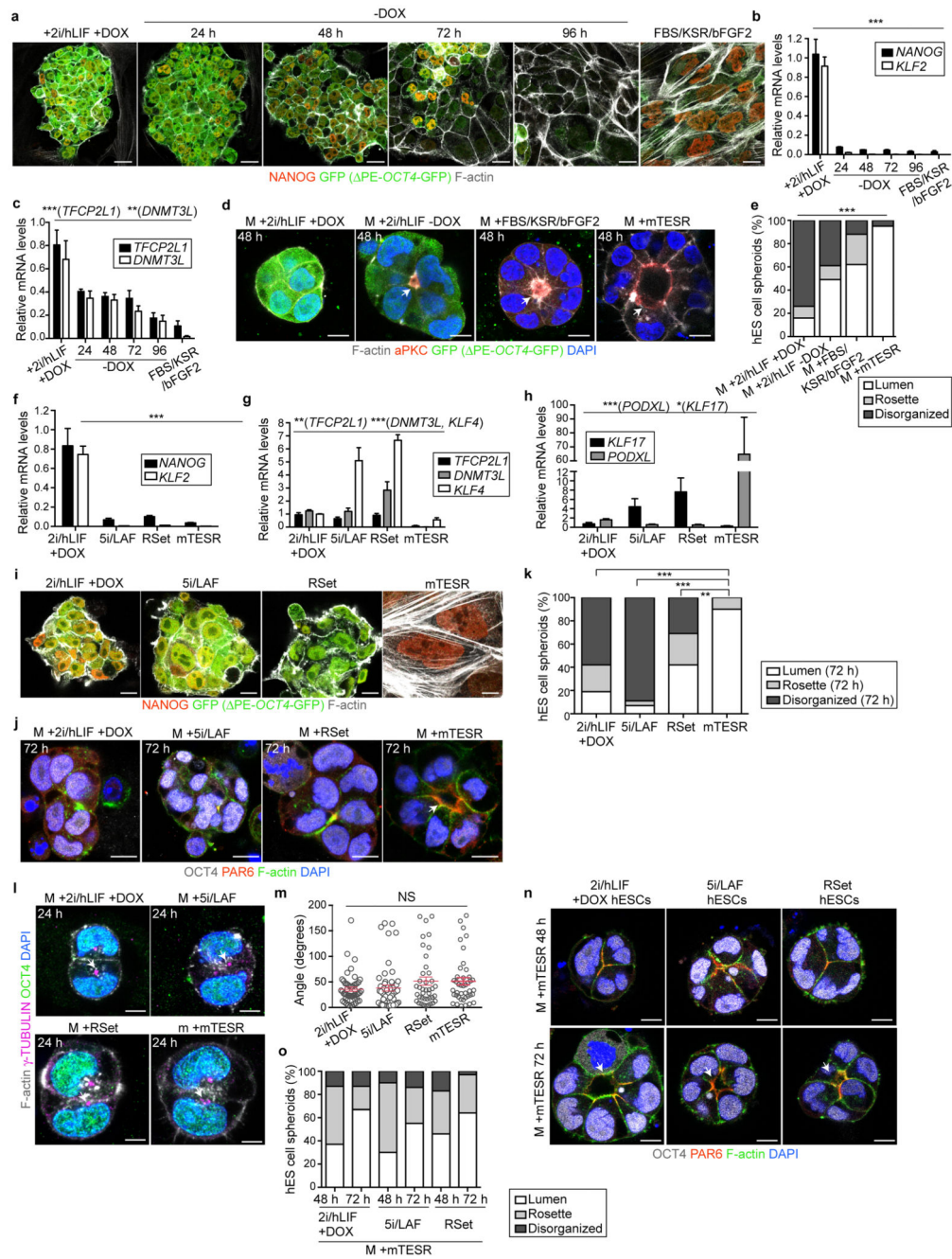
Extended Data Figure 10. Exit from naive pluripotency is required for lumenogenesis in hES cells.
a, Immunostaining of WIBR3 ΔPE–OCT4–GFP cells cultured in primed FBS/KSR/bFGF2 medium or N2B27 2i/hLIF with or without DOX. DOX addition induces the expression of NANOG and KLF2. Scale bars, 20 μm. **b, c,** mRNA levels of pluripotency genes in cells from a. n = 3 independent samples per condition. ANOVA; * * * P < 0.0001 (NANOG and KLF2); * * * P = 0.0001 (TFCP2L1); * * P = 0.0013 (DNMT3L). **d,** Immunostaining of WIBR3 ΔPE–OCT4–GFP cells cultured in Matrigel in primed or naive conditions. Arrows

indicate lumens. Scale bars, 20 μ m. **e,** Lumen formation in cells from **d**. n = 38 (2i/hLIF + DOX), 33 (2i/hLIF − DOX), 34 (FBS/KSR/bFGF2) and 20 (mTESR) spheroids. χ2 test; * * * P < 0.0001. **f–h,** mRNA levels of pluripotency genes and early post-implantation factors in cells from **i**. n = 4 (NANOG, KLF2, TFCP2L1, DNMT3L, KLF4 and PODXL) and 3 (KLF17) independent samples per group. Kruskal–Wallis test; * * * P = 0.0002 (NANOG); * * * P < 0.0001 (KLF2, DNMT3L, KLF4, PODXL); * * P = 0.0058 (TFCP2L1); * P = 0.0205 (KLF17). **i,** Immunostaining of WIBR3 PE–OCT4–GFP cells cultured in primed mTESR medium or naive conditions (2i/hLIF + DOX, 5i/LAF and RSet). Scale bars, 10 μm. **j,** Immunostaining of hES cells cultured in 3D Matrigel with different naive and primed conditions. Arrow indicates lumen. Scale bars, 10 μm. **k,** Lumen formation in cells from **j**. n = 31 (2i/hLIF + DOX), 27 (5i/LAF), 33 (RSet) and 20 (mTESR) spheroids. χ2 test; * * P = 0.0019; * * * P < 0.0001. **l,** Immunostaining of hES cells cultured in 3D Matrigel with different naive and primed conditions. Arrows indicate polarized centrosomes. Scale bars, 5 μm. **m,** Angle between the nucleus–centrosome axis and the nuclear–nuclear axis in cells from l. Each dot represents an individual centrosome. n = 53 (2i/LIF + DOX), 51 (5i/LAF), 40 (RSet) and 42 (mTESR) centrosomes. Kruskal–Wallis test; NS, not significant. **n,** Immunostaining of naive hES cells cultured in 3D Matrigel with mTESR. The initial naive conditions in which the cells were cultured are indicated. Arrows indicate lumens. Scale bars, 10 μm. **o,** Lumen formation in cells from **n**. n = 30 (2i/hLIF + DOX 48 h), 30 (2i/hLIF + DOX 72 h), 30 (5i/LAF 48 h), 29 (5i/LAF 72 h), 30 (RSet 48 h) and 30 (RSet 72 h). M, Matrigel.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Bedzhov I, Zernicka-Goetz M. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. Cell. 2014; 156:1032–1044. [PubMed: 24529478]

2. Hertig AT, Rock J, Adams EC. A description of 34 human ova within the first 17 days of development. Am J Anat. 1956; 98:435–493. [PubMed: 13362122]

3. De Los Angeles A, et al. Hallmarks of pluripotency. Nature. 2015; 525:469–478. [PubMed: 26399828]

4. Kalkan T, Smith A. Mapping the route from naive pluripotency to lineage specification. Phil Trans R Soc Lond B. 2014; 369 20130540.

5. Weinberger L, Ayyash M, Novershtern N, Hanna JH. Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat Rev Mol Cell Biol. 2016; 17:155–169. [PubMed: 26860365]

6. Acampora D, Di Giovannantonio LG, Simeone A. Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition. Development. 2013; 140:43–55. [PubMed: 23154415]

7. Nielsen JS, McNagny KM. The role of podocalyxin in health and disease. J Am Soc Nephrol. 2009; 20:1669–1676. [PubMed: 19578008]

8. Boroviak T, et al. Lineage-specific profiling delineates the emergence and progression of naive pluripotency in mammalian embryogenesis. Dev Cell. 2015; 35:366–382. [PubMed: 26555056]

9. Bedzhov I, Leung CY, Bialecka M, Zernicka-Goetz M. In vitro culture of mouse blastocysts beyond the implantation stages. Nat Protoc. 2014; 9:2732–2739. [PubMed: 25356584]

10. Ying QL, et al. The ground state of embryonic stem cell self-renewal. Nature. 2008; 453:519–523. [PubMed: 18497825]

11. Shahbazi MN, et al. Self-organization of the human embryo in the absence of maternal tissues. Nat Cell Biol. 2016; 18:700–708. [PubMed: 27144686]

12. Taniguchi K, et al. Lumen formation is an intrinsic property of isolated human pluripotent stem cells. Stem Cell Reports. 2015; 5:954–962. [PubMed: 26626176]

13. Betschinger J, et al. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. Cell. 2013; 153:335–347. [PubMed: 23582324]

14. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet. 2007; 39:380–385. [PubMed: 17259983]

15. Dutta D, et al. Self-renewal versus lineage commitment of embryonic stem cells: protein kinase C signaling shifts the balance. Stem Cells. 2011; 29:618–628. [PubMed: 21308862]

16. Dunn SJ, Martello G, Yordanov B, Emmott S, Smith AG. Defining an essential transcription factor program for naïve pluripotency. Science. 2014; 344:1156–1160. [PubMed: 24904165]

17. Chambers I, et al. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell. 2003; 113:643–655. [PubMed: 12787505]

18. Buecker C, et al. Reorganization of enhancer patterns in transition from naïve to primed pluripotency. Cell Stem Cell. 2014; 14:838–853. [PubMed: 24905168]

19. Yang SH, et al. Otx2 and Oct4 drive early enhancer activation during embryonic stem cell transition from naive pluripotency. Cell Rep. 2014; 7:1968–1981. [PubMed: 24931607]

20. Bryant DM, et al. A molecular network for de novo generation of the apical surface and lumen. Nat Cell Biol. 2010; 12:1035–1045. [PubMed: 20890297]

21. Ullrich O, Reinsch S, Urbé S, Zerial M, Parton RG. Rab11 regulates recycling through the pericentriolar recycling endosome. J Cell Biol. 1996; 135:913–924. [PubMed: 8922376]

22. Meder D, Shevchenko A, Simons K, Fullekrug J. Gp135/podocalyxin and NHERF-2 participate in the formation of a preapical domain during polarization of MDCK cells. J Cell Biol. 2005; 168:303–313. [PubMed: 15642748]

23. Doyonnas R, et al. Anuria, omphalocele, and perinatal lethality in mice lacking the CD34-related protein podocalyxin. J Exp Med. 2001; 194:13–28. [PubMed: 11435469]

24. Strilic´ B, et al. Electrostatic cell-surface repulsion initiates lumen formation in developing blood vessels. Curr Biol. 2010; 20:2003–2009. [PubMed: 20970336]

25. Blasky AJ, Mangan A, Prekeris R. Polarized protein transport and lumen formation during epithelial tissue morphogenesis. Annu Rev Cell Dev Biol. 2015; 31:575–591. [PubMed: 26359775]

26. Mangan AJ, et al. Cingulin and actin mediate midbody-dependent apical lumen formation during polarization of epithelial cells. Nat Commun. 2016; 7:12426. [PubMed: 27484926]

27. Blakeley P, et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. Development. 2015; 142:3151–3165. [PubMed: 26293300]

28. Deglincerti A, et al. Self-organization of the in vitro attached human embryo. Nature. 2016; 533:251–254. [PubMed: 27144363]

29. Theunissen TW, et al. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell. 2014; 15:471–487. [PubMed: 25090446]

30. Collier AJ, et al. Comprehensive Cell surface protein profiling identifies specific markers of human naive and primed pluripotent states. Cell Stem Cell. 2017; 20:874–890. [PubMed: 28343983]

31. Takashima Y, et al. Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell. 2014; 158:1254–1269. [PubMed: 25215486]

32. Vuoristo S, Jedrusik A, Shahbazi M, Zernicka-Goetz M. Culture of human embryos through implantation stages in vitro. Protoc Exch. 2016; doi: 10.1038/protex.2016.017

33. Rhee JM, et al. In vivo imaging and differential localization of lipid-modified GFP-variant fusions in embryonic stem cells and mice. Genesis. 2006; 44:202–218. [PubMed: 16604528]

34. Yeom YI, et al. Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development. 1996; 122:881–894. [PubMed: 8631266]

35. Kalkan T, et al. Tracking the embryonic stem cell transition from ground state pluripotency. Development. 2017; 144:1221–1234. [PubMed: 28174249]

36. Panamarova M, et al. The BAF chromatin remodelling complex is an epigenetic regulator of lineage specification in the early mouse embryo. Development. 2016; 143:1271–1283. [PubMed: 26952987]

37. Lee GY, Kenny PA, Lee EH, Bissell MJ. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods. 2007; 4:359–365. [PubMed: 17396127]

38. Ran FA, et al. Genome engineering using the CRISPR–Cas9 system. Nat Protoc. 2013; 8:2281–2308. [PubMed: 24157548]

39. Schneider CA, Rasband WS, Eliceiri KW. NIH image to ImageJ: 25 years of image analysis. Nat Methods. 2012; 9:671–675. [PubMed: 22930834]

40. Martín-Belmonte F, et al. Cell-polarity dynamics controls the mechanism of lumen formation in epithelial morphogenesis. Curr Biol. 2008; 18:507–513. [PubMed: 18394894]

41. Yang Z, et al. De novo lumen formation and elongation in the developing nephron: a central role for afadin in apical polarity. Development. 2013; 140:1774–1784. [PubMed: 23487309]

42. Picelli S, et al. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014; 9:171–181. [PubMed: 24385147]

43. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015; 31:166–169. [PubMed: 25260700]

44. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11:R106. [PubMed: 20979621]

45. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15:550. [PubMed: 25516281]
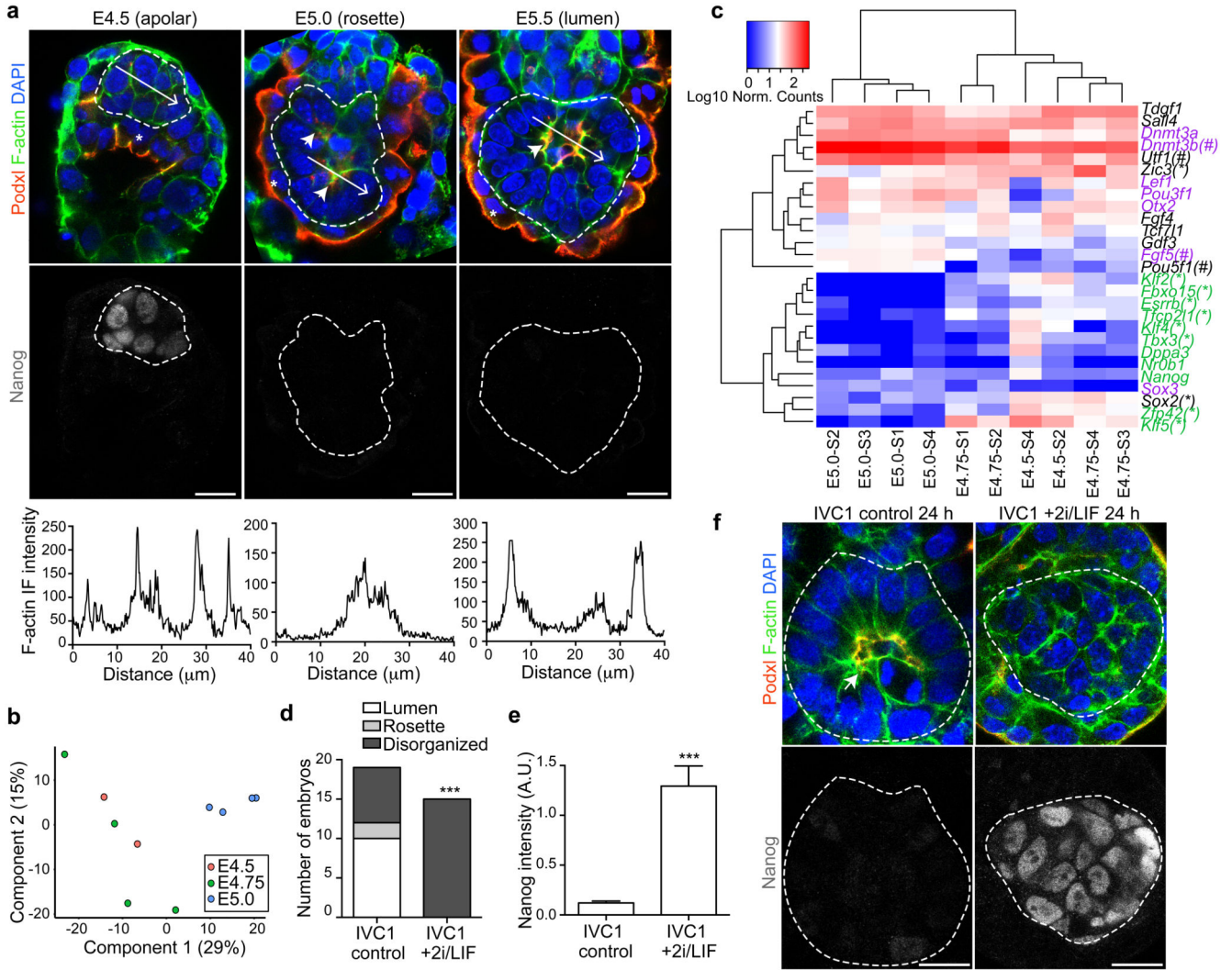
Figure 1. Epiblast gene expression at peri-implantation.

a, Immunostaining of mouse embryos (top). Dotted lines indicate the epiblast; arrowheads indicate polarized Podxl; asterisks indicate Podxl in the primitive endoderm; white long arrows show positions used to plot intensity profiles (bottom). **b**, Principal component analysis of all samples by all expressed genes. Numbers in brackets indicate percentage of variance. **c**, Heat map showing expression of core (black), naive (green) and post-implantation (purple) genes. Genes with significantly high (*) or low (#) expression in E4.5–E4.75 compared to E5.0 cells are indicated. $P<0.1$. **d**, Lumen formation in embryos from **f**. n = 19 (IVC1) and 15 (IVC1 +2i/LIF) embryos. χ^2 test; ***P = 0.0007. **e**, Nanog intensity in embryos from **f**. n = 19 (IVC1) and 15 (IVC1 +2i/LIF) embryos. Unpaired Student's t-test; ***$P<0.0001$. **f**, Immunostaining of *in vitro* cultured mouse embryos. Dotted lines indicate the epiblast. Arrow indicates lumen. Scale bars, 20 μm (**a**, **f**).
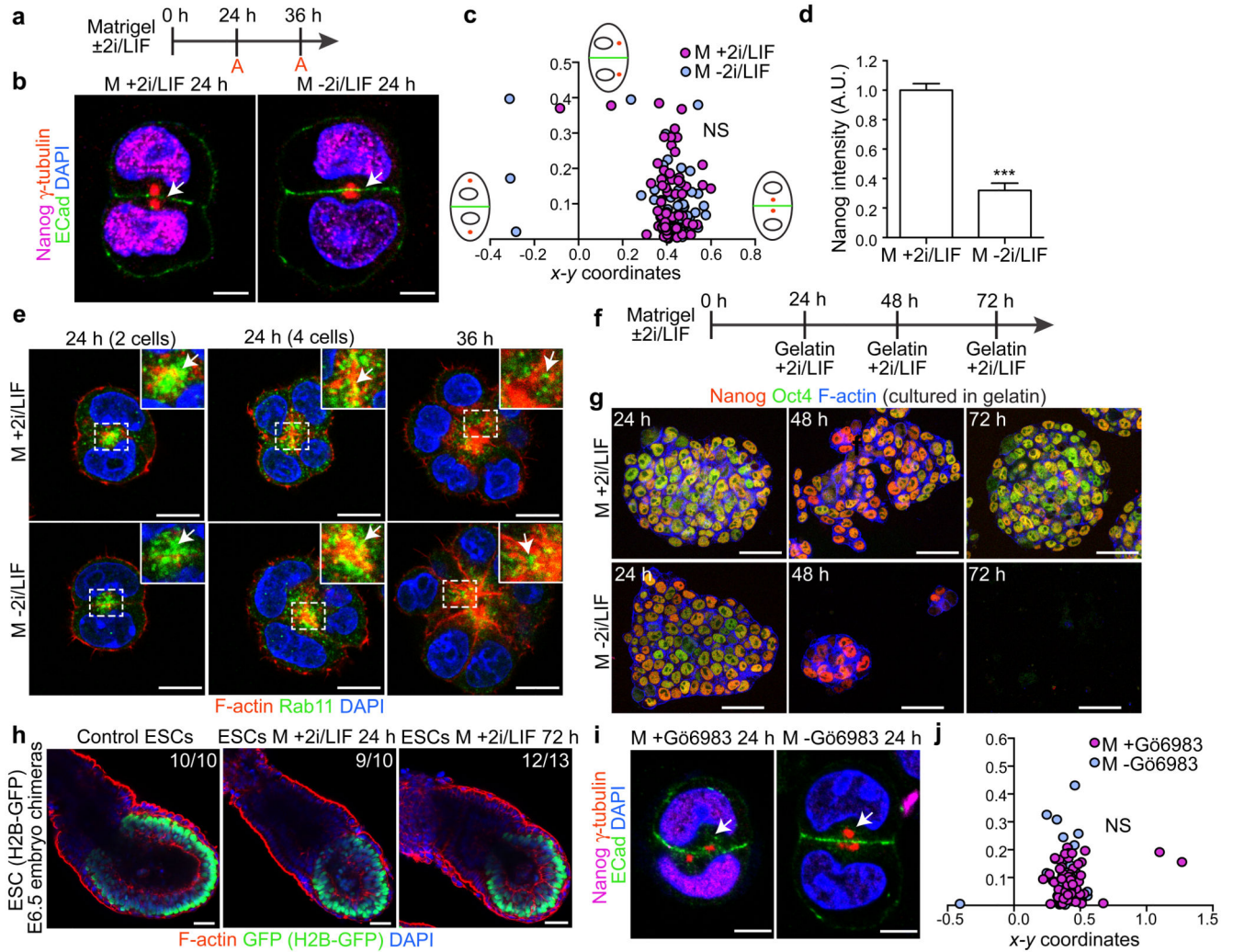
Figure 2. Naive mES cells initiate polarization in Matrigel.
a, Experimental set-up. A, analysis. **b**, Immunostaining of mES cells cultured as indicated in **a**. Scale bars, 5 μm. **c**, Centrosome positions in cells from **b**. *n* = 60 (+2i/LIF) and 62 (−2i/LIF) centrosomes. Unpaired Student's *t*-test; NS, not significant. **d**, Nanog intensity in cells from **b**. *n* = 30 (+2i/LIF) and 31 (−2i/LIF) spheroids. Unpaired Student's *t*-test; ***P< 0.0001. **e**, Immunostaining of mES cells cultured as indicated in **a**. Squares denote magnified regions. Scale bars, 10 μm. **f**, Experimental set-up. **g**, Immunostaining of mES cells cultured as indicated in **f**. Scale bars, 50 μm. **h**, mES cell–embryo chimaeras generated using H2B–GFP mES cells cultured as indicated in **f**. Scale bars, 40 μm. **i**, Immunostaining of mES cells. Scale bars, 5 μm. **j**, Centrosome positions in cells from **i**. *n* = 52 (+Gö6983) and 56 (−Gö6983) centrosomes. Mann–Whitney *U*-test; NS, not significant. Arrows denote polarization. M, Matrigel.
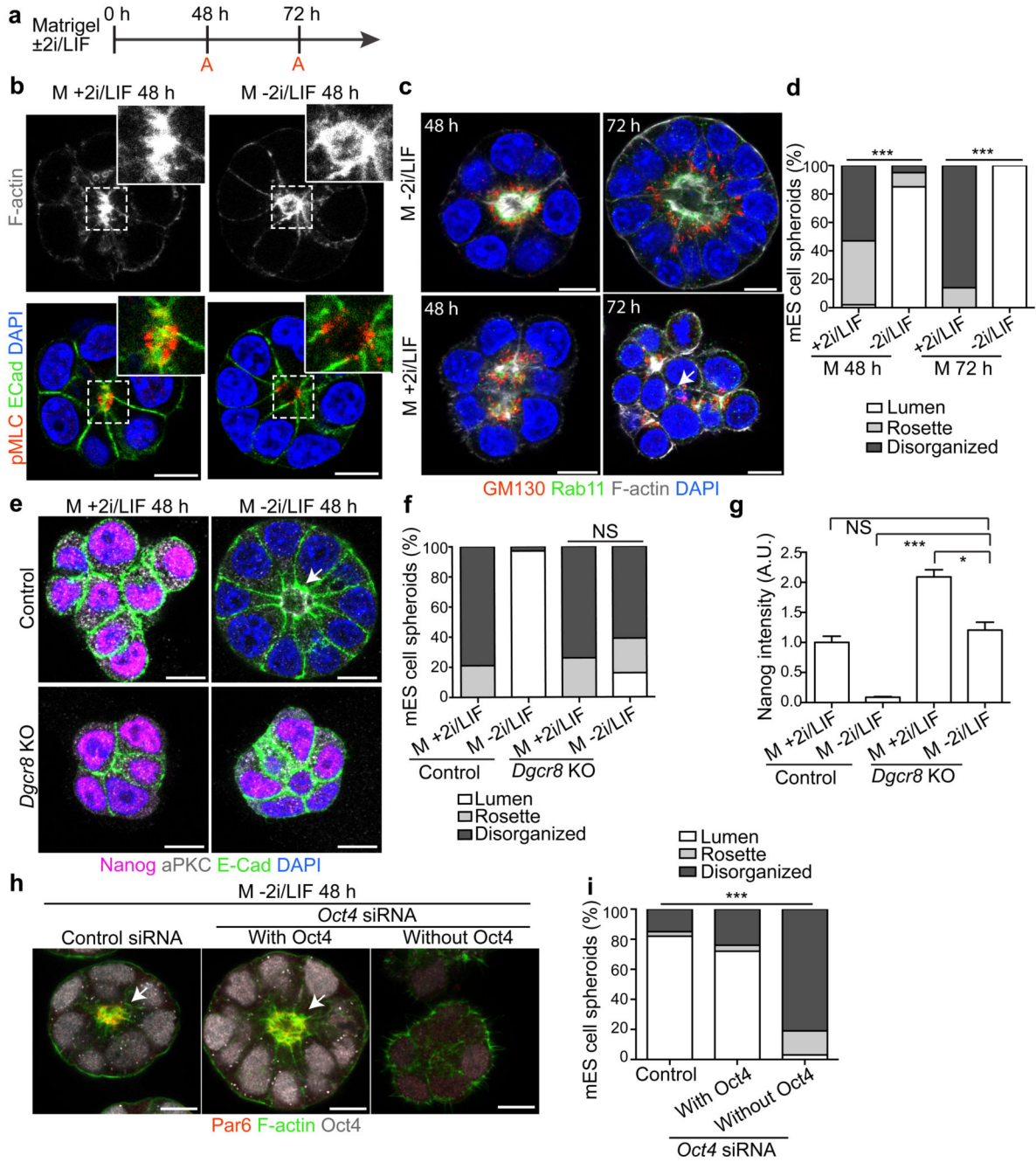
Figure 3. Naive pluripotency exit is required for lumenogenesis.

a, Experimental set-up. A, analysis. **b**, **c**, Immunostaining of mES cells cultured as indicated in **a**. Squares denote magnified regions and the arrow in **c** non-polarized cells. **d**, Lumen formation in cells from **b** and **c**. n = 40 (+2i/LIF 48 h), 40 (−2i/LIF 48 h), 21 (+2i/LIF 72 h) and 22 (−2i/LIF 72 h) spheroids. χ^2 test; ***$P < 0.0001$. **e**, Immunostaining of control and *Dgcr8* knockout (KO) mES cells. Arrow indicates lumen. **f**, Lumen formation in cells from **e**. n = 29 (control +2i/LIF), 31 (control −2i/LIF), 31 (*Dgcr8* knockout +2i/LIF) and 31 (*Dgcr8* knockout −2i/LIF) spheroids. χ^2 test; NS, not significant. **g**, Nanog intensity in cells

from **e**. $n = 20$ (control +2i/LIF), 20 (control −2i/LIF), 19 (*Dgcr8* knockout +2i/LIF) and 21 (*Dgcr8* knockout −2i/LIF) spheroids. Kruskal–Wallis test; ***$P < 0.0001$; *$P < 0.05$, NS, not significant. **h**, Immunostaining of control and *Oct4* knockdown mES cells. Arrows indicate lumens. **i**, Lumen formation in cells from **h**. $n = 39$ (control), 53 (*Oct4* siRNA with Oct4) and 32 (*Oct4* siRNA without Oct4) spheroids. χ test; ***$P < 0.0001$. Scale bars, 10 μm (**b**, **c**, **e**, **h**). M, Matrigel.
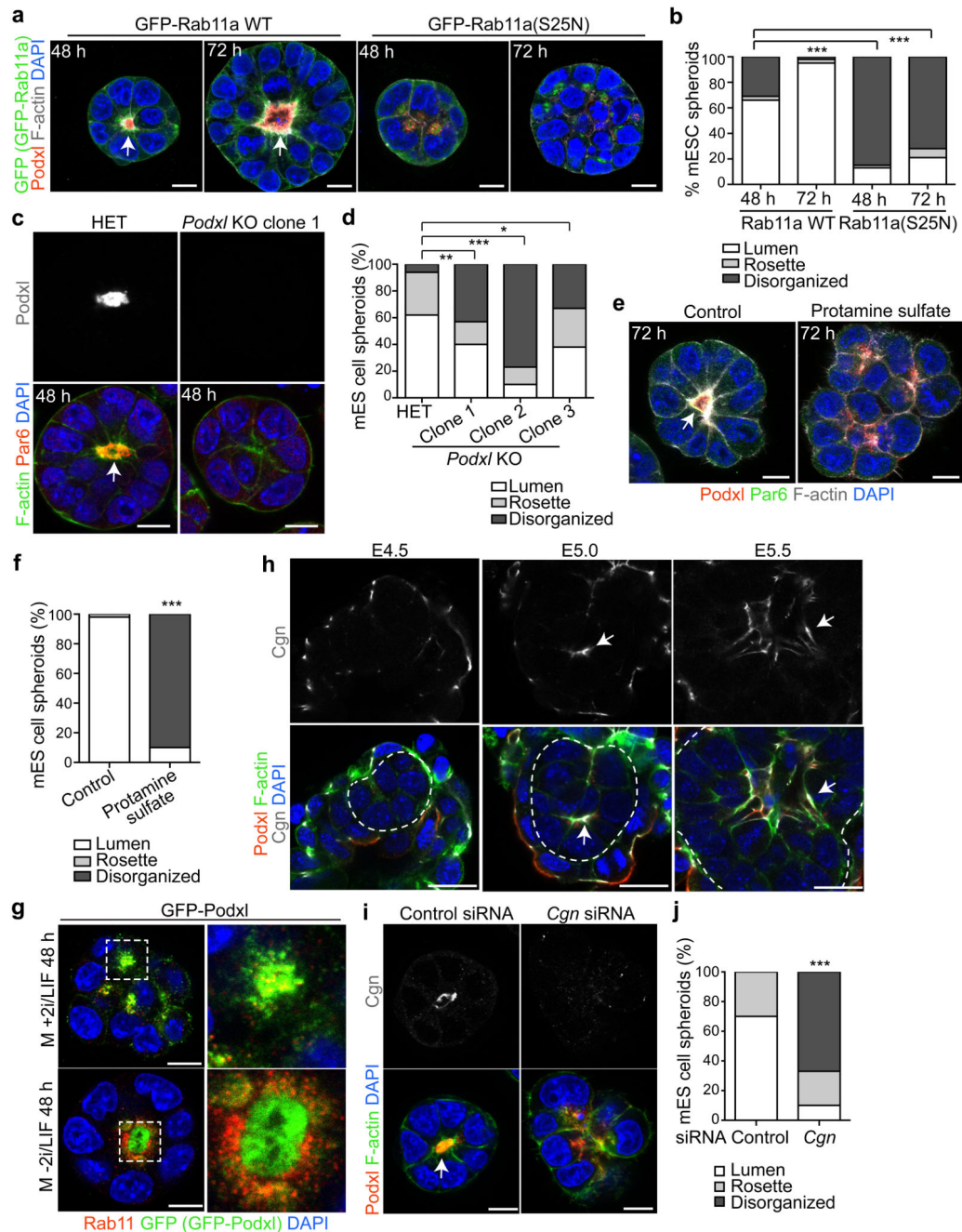
Figure 4. Vesicle fusion downstream of naive pluripotency exit.
a, Immunostaining of GFP–Rab11a and GFP–Rab11a(S25N) overexpressing mES cells.
Arrows indicate lumens. Scale bars, 10 μm. WT, wild-type. **b**, Lumen formation in cells
from **a**. $n = 35$ (Rab11a wild-type 48 h), 60 (Rab11a wild-type 72 h), 39 (Rab11a(S25N) 48
h) and 29 (Rab11a(S25N) 72 h) spheroids. χ^2 test; ***$P < 0.0001$. **c**, Immunostaining of
heterozygous (HET) and *Podxl* knockout (KO) mES cells. Arrow indicates lumen. Scale
bars, 10 μm. **d**, Lumen formation in cells from **c**. $n = 34$ (HET), 32 (clone 1), 30 (clone 2)
and 21 (clone 3) spheroids. χ^2 test; *$P = 0.0245$; **$P = 0.0019$; ***$P < 0.0001$. **e**,

Immunostaining of control and protamine sulfate-treated mES cells. Arrow indicates lumen. Scale bars, 10 μm. **f**, Lumen formation in cells from **e**. $n = 41$ (control) and 40 (protamine sulfate) spheroids. χ^2 test; ***$P < 0.0001$. **g**, Immunostaining of GFP–Podxl overexpressing mES cells. Squares denote magnified regions. Scale bars, 10 μm. M, Matrigel. **h**, Immunostaining of mouse embryos. Dotted lines indicate the epiblast, arrows indicate polarized Cgn. Scale bars, 20 μm. **i**, Immunostaining of control and *Cgn* knockdown mES cells. Arrow indicates lumen. Scale bars, 10 μm. **j**, Lumen formation in cells from **i**. $n = 40$ (control) and 39 (*Cgn* siRNA) spheroids. χ^2 test; ***$P < 0.0001$.
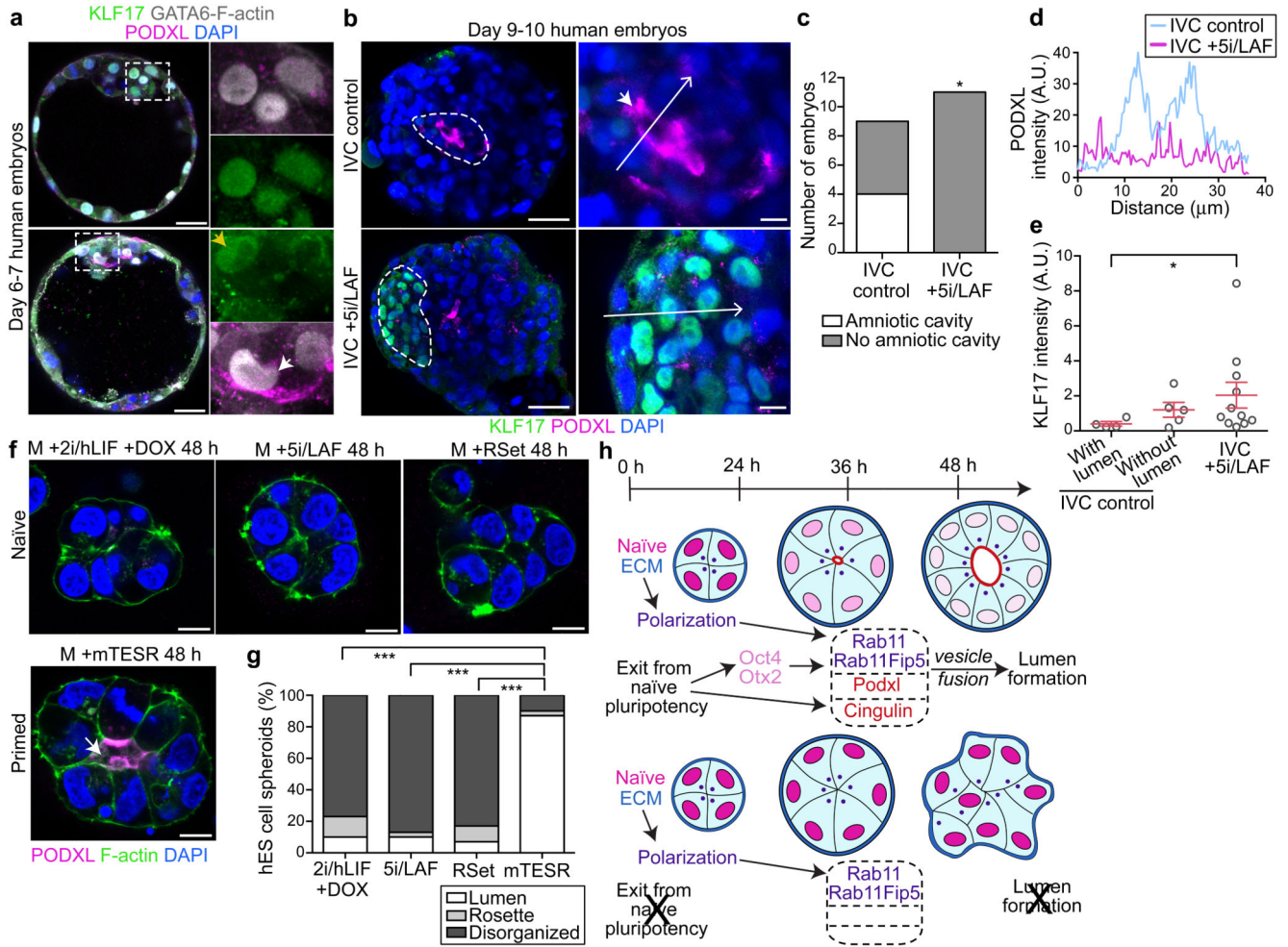
Figure 5. Naive pluripotency exit direct lumenogenesis in human embryos and hES cells.
a, Immunostaining of day 6–7 human embryos. Squares denote magnified regions, yellow arrowhead indicates the epiblast and white arrowhead indicates the primitive endoderm. Scale bars, 30 μm. **b**, Immunostaining of day 9–10 human embryos. Dotted lines indicate the epiblast; arrowhead indicates the amniotic cavity; white long arrows show the position used to plot intensity profiles (**d**). Scale bars, 50 μm and 10 μm (magnified areas). **c**, Lumen formation in embryos from **b**. n=9 (IVC control) and 11 (IVC +5i/LAF). χ^2 test; *$P=$ 0.0134. **d**, PODXL intensity profiles in embryos from **b**. **e**, KLF17 intensity in embryos from **b**. $n = 4$ (control with lumen), 5 (control without lumen) and 11 (IVC +5i/LAF) embryos. Unpaired Student's t-test; *$P = 0.0366$. **f**, Immunostaining of hES cells. Arrow indicates lumen. Scale bars, 10 μm. **g**, Lumen formation in cells from **f**. $n = 31$ (2i/hLIF +DOX), 30 (5i/LAF), 30 (RSet) and 30 (mTESR) spheroids. χ^2 test; ***$P < 0.0001$. **h**, Model summarizing the findings of the study.