Genomics 116 (2024) 110780

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Cell specification and functional interactions in the pig blastocyst inferred from single-cell transcriptomics and uterine fluids proteomics

Adrien Dufour [a], Cyril Kurylo [b], Jan B. Stöckl [c], Denis Laloë [a], Yoann Bailly [d], Patrick Manceau [d], Frédéric Martins [e,f], Ali G. Turhan [g], Stéphane Ferchaud [d], Bertrand Pain [h], Thomas Fröhlich [c], Sylvain Foissac [b], Jérôme Artus [g], Hervé Acloque [a,*]

[a] Université Paris Saclay, INRAE, AgroParisTech, GABI, Domaine de Vilvert, 78350 Jouy en Josas, France
[b] Université de Toulouse, INRAE, ENVT, GenPhySE, Chemin de Borde Rouge, 31326 Castanet-Tolosan, France
[c] Ludwig-Maximilians-Universität München, Genzentrum, Feodor-Lynen-Str. 25, 81377 München, Germany
[d] INRAE, GenESI, La Gouvanière, 86480 Rouillé, France
[e] Plateforme Genome et Transcriptome (GeT-Santé), GenoToul, Toulouse University, CNRS, INRAE, INSA, Toulouse, France
[f] I2MC - Institut des Maladies Métaboliques et Cardiovasculaires, Inserm, Université de Toulouse, Université Paul Sabatier, Toulouse, France
[g] Université Paris Saclay, Inserm, UMRS1310, 7 rue Guy Moquet, 94800 Villejuif, France
[h] Université de Lyon, Inserm, INRAE, SBRI, 18 Av. du Doyen Jean Lépine, 69500 Bron, France

ARTICLE INFO

Keywords:
Pig
Blastocyst
Trophectoderm
Epiblast
Hypoblast
Single-cell

ABSTRACT

The embryonic development of the pig comprises a long in utero pre- and peri-implantation development, which dramatically differs from mice and humans. During this peri-implantation period, a complex series of paracrine signals establishes an intimate dialogue between the embryo and the uterus. To better understand the biology of the pig blastocyst during this period, we generated a large dataset of single-cell RNAseq from early and hatched blastocysts, spheroid and ovoid conceptus and proteomic datasets from corresponding uterine fluids. Our results confirm the molecular specificity and functionality of the three main cell populations. We also discovered two previously unknown subpopulations of the trophectoderm, one characterised by the expression of *LRP2*, which could represent progenitor cells, and the other, expressing pro-apoptotic markers, which could correspond to the Rauber's layer. Our work provides new insights into the biology of these populations, their reciprocal functional interactions, and the molecular dialogue with the maternal uterine environment.

1. Introduction

The pig is a species of increasing interest both as a biomedical model for human pathologies [1], as one of the main sources of animal protein for human nutrition, and as an alternative to rodent animal models for studying early mammalian development [2]. The early embryonic development of the pig comprises a long in utero pre- and peri-implantation process, which dramatically differs from mouse and human. During this protracted peri-implantation period, a complex series of paracrine and exocrine signals establishes an intimate dialogue between the embryo and the uterus [2]. This dialogue leads to essential changes in the uterine receptivity to implantation and embryonic morphology. After fertilization, the porcine embryo undergoes a series of cleavages. At 4 days post-fertilization (dpf), the embryo undergoes the process of compaction, which is associated with an increase in intercellular adhesion and the acquisition of cell polarity, giving it the appearance of a mulberry named morula. At 5 dpf, it undergoes the second major morphogenetic event, the blastocyst formation, characterised by a fluid-filled cavity known as the blastocoel. This process is closely linked to the inner cell mass (ICM) and the trophectoderm (TE) specification. At 6 dpf, the hypoblast (HYPO or primitive endoderm) and the epiblast (EPI) are specified from the ICM. From 7 dpf, the spherical blastocyst elongates as an ovoid, tubular, and filamentous blastocyst, transforming from a 0.5–1 mm diameter sphere to a 1000 mm long filamentous blastocyst at 16 dpf. At the same time, the Rauber's layer corresponding to the polar TE covering the ICM disappears. The underlying EPI becomes directly exposed to molecules present in the uterine fluids. These drastic changes occur before implantation and are likely controlled and coordinated by key functional interactions between cells and tissues.

* Corresponding author.
E-mail address: herve.acloque@inrae.fr (H. Acloque).

https://doi.org/10.1016/j.ygeno.2023.110780
Received 29 June 2023; Received in revised form 8 December 2023; Accepted 30 December 2023
Available online 9 January 2024

Until recently, the study and interpretation of these interactions have been hampered by the lack of molecular tools and datasets that allow the implementation of systems biology approaches. Most of the knowledge about these interactions was inferred from observations made in other mammalian species, particularly in mice, humans, or marsupial [3–5]. Classical genes and pathways known to control early embryonic and extraembryonic cell specification have been tested in cultured pig embryos or explants [6–9]. However, gaining a deeper understanding of such mechanisms has been hampered by the difficulty of accessing pig embryos at late stages of pre-implantation development, the lack of bona fide embryonic and extraembryonic stem cells and the difficulty of performing functional genomics.

The recent development of single-cell RNA sequencing (RNASeq) has enabled remarkable advances in understanding the first steps of mammalian embryogenesis, from the definition of embryonic and extraembryonic populations to the gene regulatory networks controlling cell fates. Indeed, scRNAseq studies of mouse pre-implantation embryos highlight the switch from the naive pluripotent state observed in the ICM of the blastocyst to the formative and primed pluripotent states observed in the epiblast, coinciding with the sequential specification of the trophectoderm and the hypoblast cell populations [4]. Recently, scRNAseq studies of porcine embryos have also been published, describing the switch from a naive epiblast (from 4 dpf to 6 dpf) to a primed epiblast (from 7 dpf) [10,11], which is associated with a switch in signalling pathways. In the ICM, the IL6-STAT3 and PI3K-AKT pathways are mainly activated and regulate the expression of markers of naive embryonic pluripotency (KLF4, ESRRB, STAT3). When the EPI forms, these two signalling pathways decrease in favour of the TGFβ-SMAD2/3 pathway, which in turn regulates the expression of markers of primed pluripotency (NANOG, DNMT3B, OTX2) and is associated with a metabolic switch between OXPHOS and glycolysis. From 10 dpf, these studies described a primed state of pluripotency in the epiblast where canonical Wnt signalling activity increases and primes the pluripotent epiblast for gastrulation [10,11]. These studies also suggest a potentially conserved role of IL1B genes between pigs and humans/monkeys for implantation and rapid trophectoderm expansion in pigs.

However, as these studies have mainly focused on EPI, information on other extraembryonic cell populations still needs to be provided. It is mainly due to the difficulty of obtaining a large number of cells representative of all embryonic and extraembryonic subpopulations using the smart-seq2 approach [12]. Here we used the Chromium 10× Genomics technology to provide a scRNAseq analysis of pig embryos at four different embryonic stages: (1) early blastocyst (Embryonic day 5, E5), (2) hatched blastocyst (E7), (3) spheroid/early ovoid blastocyst (E9) and (4) late ovoid blastocyst (E11). Proteomic datasets were also generated from the uterine fluids of the sows used for embryo production. We characterised a panel of 34,888 cells, from which we first described embryonic and extraembryonic cell populations and their evolution and identified population-specific markers of the three main populations (epiblast, trophectoderm, and hypoblast). We also identified known and novel specific functions associated with the biology of these subpopulations. We then inferred gene regulatory networks working on modules of gene regulation (regulon) and selected those specifically active in each embryonic population. We then linked these regulons to signalling pathways and biological processes. To do this, we constructed signalling networks from ligands (expressed by cells or present in the uterine fluids), receptors, intermediaries, and transcription factors. Our results confirm the molecular specificity and functionality of the three main cell populations and identify novel stage-specific subpopulations. We also provide new insights into the biology of these populations, their reciprocal functional interactions, and the molecular dialogue established with the maternal organism through the uterine fluids.

2. Results and discussion

2.1. Identification of embryonic and extraembryonic cell populations and their associated biological functions

Using the Chromium 10× Genomics technology, we generated a large dataset of scRNAseq at four different embryonic stages corresponding to (1) early blastocyst (E5), (2) hatched blastocyst (E7), (3) spheroid/early ovoid blastocyst (E9) and (4) late ovoid blastocyst (E11) (Fig. 1A and Supplementary Table 1). Raw reads were mapped, reattributed to each cell, and counted. We defined a minimum number of UMIs per cell, a maximum percentage of mitochondrial transcripts per cell, and a maximum number of features per cell to exclude the cells that did not reach a sufficient quality level (Supplementary Table 2). We validated 34,753 cells, distributed as follows: 1226 cells at E5, 4228 cells at E7, 12,727 cells at E9, and 16,572 cells at E11 (Fig. 1B and Supplementary Table 2).

From this dataset, to identify transcriptionally distinct cell populations, we performed a dimensionality reduction and clustering approach using Principal Component Analysis (PCA) and Uniform Manifold Approximation and Projection (UMAP) following the Harmony and Seurat workflows. To identify the distinct cell populations, we then visualised known population marker genes from the literature for ICM (*AICDA*, *KLF4, SPIC*), EPI (*POU5F1*, *ID1*, *DNMT3B*), TE (*DAB2, KRT8, PTGES*), HYPO (*COL18A1*, *GATA6*, *SERPINH1)* (Fig. 1C and D) [13–15].

We then searched for enriched functions in these five cell populations by performing a gene set variation analysis (GSVA) (Supplementary Table 3c) using the expressed genes within each lineage. The most significantly enriched pathways and those selected from the literature are shown in Fig. 1E. We found enrichment for the pools of genes corresponding to Notch, JAK-STAT, Hippo, Hedgehog, and Wnt signalling pathways in ICM cells compared to other cell populations. Similar pathways (except JAK-STAT) were also enriched in EPI. In TE, we found an enrichment of genes associated with ovarian steroidogenesis and estrogen signalling pathway, fatty acid elongation and unsaturated fatty acids and folate biosynthesis, one-carbon, retinol, and thiamine metabolism, reflecting known TE biological functions. In the HYPO, we observed an enrichment of genes associated with the oxytocin signalling pathway, fat digestion and absorption, fatty acid elongation, vitamin digestion and absorption, focal adhesion, and ECM-receptor interaction.

Next, we used SCENIC to identify regulons, defined as functional modules of gene regulation. Each regulon associates a transcription factor (TF) and its direct target genes, defined by their co-expression with the TF and by sharing a common binding motif for this TF in their promoters [16]. For each regulon, activity scores were calculated for each cell using the relative expression of the gene that makes up the regulon. 297 regulons were identified (Supplementary Fig. 1), and their activity score was summarised by their mean expression in each cluster at each state (Supplementary Table 4a). In parallel, we performed the same analysis on two publicly available scRNAseq datasets from pig and human pre-implantation embryos [3,11] (Supplementary Figs. 12, 13), and we looked for common regulons across these scRNAseq analyses (Supplementary Fig. 1). We then selected the most specific regulons for each cell population based on the Regulon Specificity Score (RSS) (Supplementary Table 4b).

2.2. Timely diversification of TE cell population with distinct molecular functions

To further characterise the TE cell population, we selected 18,239 cells from our dataset corresponding to the TE population and performed dimensionality reduction followed by a new clustering (Fig. 2A). It led to the identification of eight TE subpopulations (Fig. 2B). These subpopulations show a clear TE signature expressing various levels of *GATA2, GATA3, DAB2*, and *PTGES* (Fig. 2C), but each has specific

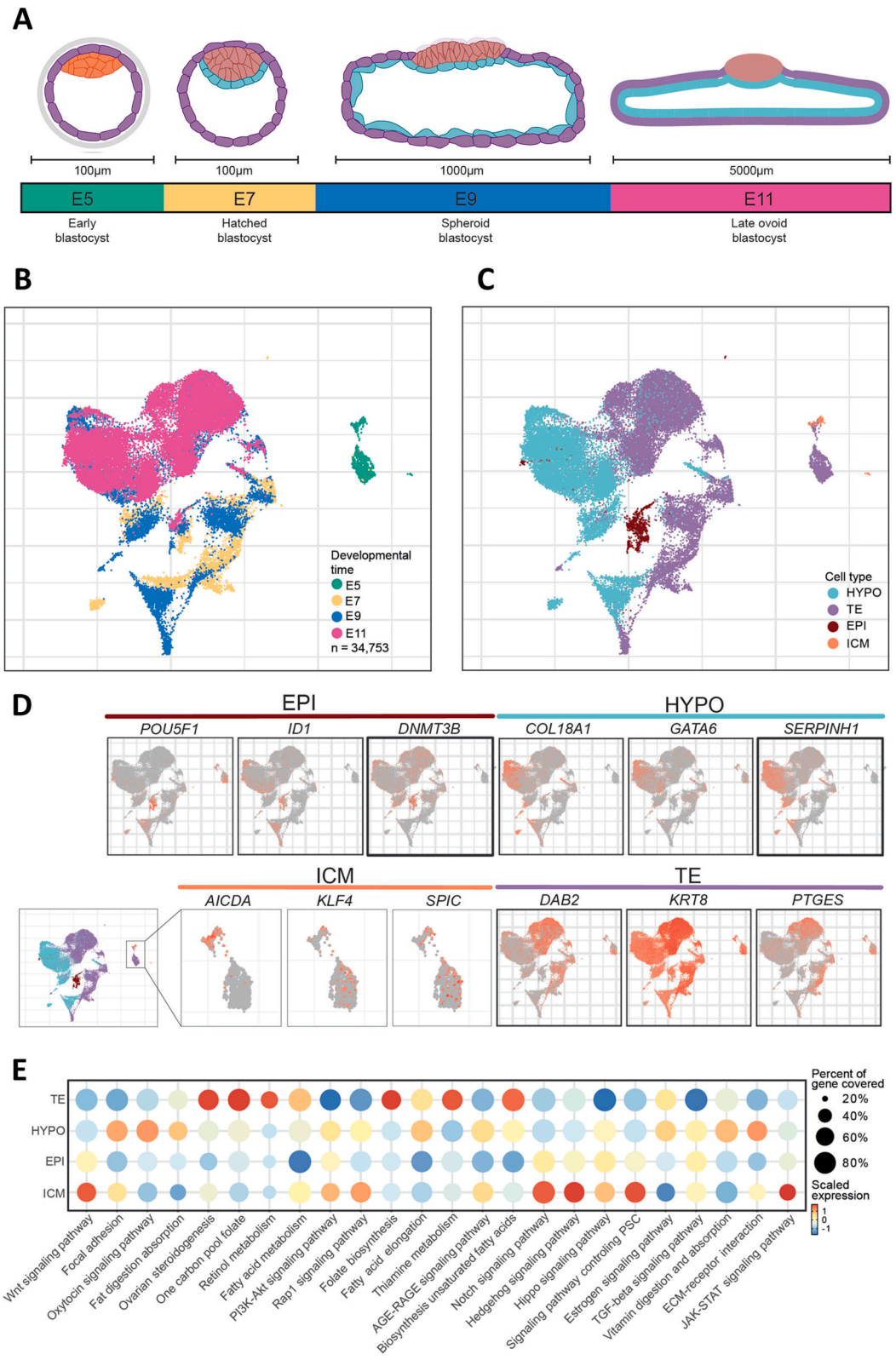
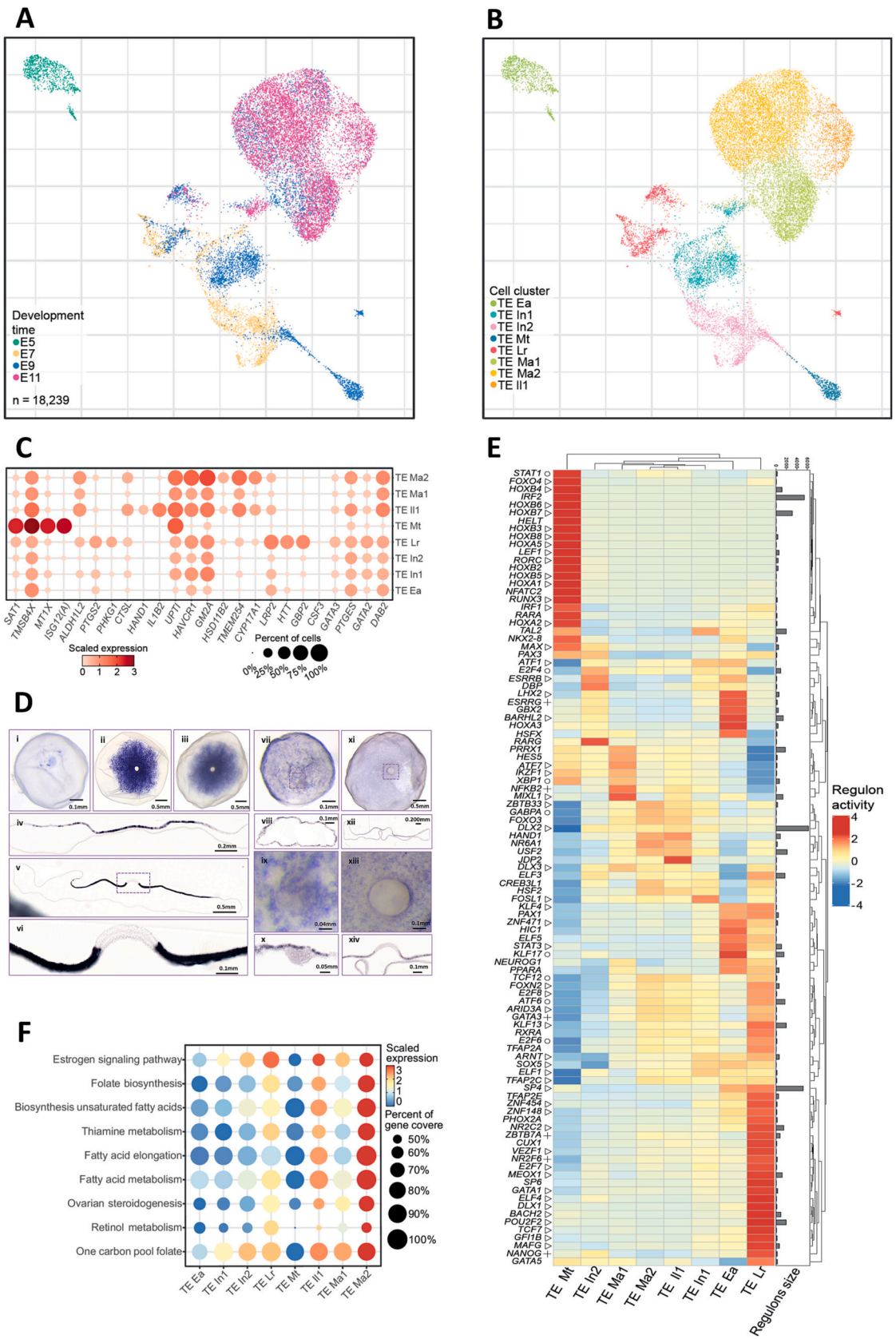


Fig. 1. Dynamic evolution of cell lineages from the early to the ovoid blastocyst stages (E5 to E11).
(A) Schematic view of pig embryo morphology from embryonic day (E)5 to E11. Cells from the ICM and the EPI are represented in orange/red, TE in purple and HYPO in turquoise. (B) Visualisation of cells coloured by developmental time via UMAP: E5 (green), E7 (yellow), E9 (blue), and E11 (pink). (C) Identification of five clusters coloured by population via UMAP: ICM (orange), EPI (red), HYPO (turquoise) and TE (purple). (D) UMAP plot of gene markers for each population. (E) Dot plot visualisation of selected KEGG signalling pathways. The circle size represents the percentage of genes out of all the genes in the pathways that are expressed by the cell populations. The red gradient represents the mean scaled expression of the pathways within the cell populations. The AGE-RAGE signalling pathway in diabetic complications has been abbreviated to the AGE-RAGE signalling pathway. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



(caption on next page)

Fig. 2. Identification and characterisation of different populations in the trophectoderm.
(A) Visualisation of selected TE cells coloured by developmental day via UMAP: E5 (green), E7 (yellow), E9 (blue), E11 (pink). (B) Visualisation of TE populations coloured by cluster via UMAP: TE Ea (green), TE In1 (light blue), TE In2 (pink), TE Lr (red), TE Mt. (dark blue), TE Ma1 (green), TE Ma2 (yellow), TE Il1 (orange). (C) Dot plot visualisation of selected DEG genes, the circle size represents the percentage of cells within the cluster that express the gene. The red gradient represents the mean scaled expression of the genes within the cluster. (D) Expression of *IL1B2 (i-vi) and LRP2 (vii-xiv)* in the TE of spheroid (E9) and ovoid (E11) pig blastocysts. Whole-mount in situ hybridization for *IL1B2* at the spheroid (i) and ovoid stage, dorsal (ii) and ventral (iii) views and *LRP2* at the spheroid (vii) and ovoid (xi) stages. The dotted squares (vii, xi) highlight the area of the embryonic discs that are enlarged (ix, xiii) and show the salt and pepper expression of *LRP2* in the TE (viii, xii). Sections confirmed the localised expression of *IL1B2* (iv-vi) and *LRP2* (viii, x, xii, xiv) in the TE. (E) Heatmap showing scaled values of Regulon Activity Score for the 20 most specific regulons for each cluster, identified by Regulon Specificity Score (RSS). (△): common regulons with another pig study; (+): common regulons with another human study; (O): common regulons within the three studies. Right rows (heatmap): histograms distribution of regulon size (number of genes regulated by the TF in the regulons. (F) Dot plot visualisation of selected KEGG signalling pathways. The circle size represents the percentage of genes out of all the genes in the pathways that are expressed by the cell cluster. The red gradient represents the mean scaled expression of the pathways within the cell cluster. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

characteristics such as TE Mt. subpopulation, in which *GATA3* is not detected in most of the cells.

At the early blastocyst stage (E5), the first cell lineage decision leads to the formation of the early trophectoderm (TE Ea, light green dots in Fig. 2B). This subpopulation is characterised by the expression of early TE marker genes such as *DAB2*, *GATA2*, and *PTGES* [11,17,18] (Fig. 2C and Supplementary Table 3a). Early TE appeared quite distinct from other TE subpopulations regarding gene expression (Fig. 2B, Supplementary Table 3a), with functional enrichment of genes related to aerobic respiration and cell metabolism (Supplementary Table 3d and Supplementary Fig. 2A). We also highlighted specifically active regulons at this stage: *LHX2* and *BARHL2* (identified in our study and from data in [11]), *ESRRG* (in our study and from data in [3]), *GBX2* and *HSFX* (our study) (Fig. 2E). We also observed active *TFAP2C* and *TFAP2E* regulons, in which TFs are known to be involved in ICM/TE segregation in mice [19].

Then, at the subsequent E7 hatched blastocyst stage, E9 early ovoid and E11 late ovoid stages, we identified two subpopulations that are observed at these three stages: intermediate trophectoderm 1 (TE In1, turquoise blue dots in Fig. 2B) a subpopulation that expresses classical TE markers (e.g., *DAB2*, *GATA2*, *PTGES* and *GATA3*) (Fig. 2C and Supplementary Table 3a) [11,17,18,20] and another small TE subpopulation that emerges at E7 and expands in subsequent stages, which we named LRP2 TE (TE Lr, red dots in Fig. 2B) due to its high expression of *LRP2, HTT* and *GBP2* (Fig. 2C, Supplementary Fig. 4). We confirmed the specific expression of *LRP2* in the TE by in situ hybridisation. At the blastocyst spheroid stage, we observed a salt-and-pepper distribution of cells expressing *LRP2* in the TE. At the ovoid stage, *LRP2* expression is more intense in the TE around the embryonic disc than in other regions (Fig. 2D). At E7 and E9, we also identified another subpopulation, which we named intermediate trophectoderm 2 (TE In2). Gene expression profiles were highly similar between TE In1 and TE In2, with notable differences in mitochondrial and ribosomal genes (Supplementary Table 3a), suggesting differences in their cellular metabolism: TE In2 relies on OXPHOS (together with TE Ea and TE Lr), whereas TE In1 relies on glycolysis (together with more mature TE subpopulations) (Supplementary Fig. 2). Slight differences in regulon activity were also observed (Fig. 2E): *ARNT* and *SOX5* regulons (identified in our study and from data in [11]) are less active in TE In2 compared to TE In1, whereas *ESRRB*, *RARA/G*, *E2F4* [21] and *DBP* are more active in TE In2. *ESRRB* is known to play a role in the differentiation of stem cells into TE [22] and *RARA/G* to be involved in cell reprogramming [23]. Regulons specific to TE In1 include TFs described in cell survival, such as *ATF1* [24], differentiation towards TE for *FOSL1* [25], and an unknown role for *TAL2*.

In contrast, TE Lr stands out from the other subpopulations and is characterised by several differentially expressed genes (DEG), including *IGFR*, which controls proliferation, differentiation, growth, and cell survival [26], *CSF3*, which improves embryonic pig development [27] and *GBP2*, *HTT*, and *LRP2* (Figs. 2C, 5B, Supplementary Fig. 4, Supplementary Table 3a). TE Lr cells are cycling with >50% of the cells in S and G2/M phases (Supplementary Fig. 3) and rely on OXPHOS metabolism (together with TE Ea and TE In2) (Supplementary Fig. 2). Strikingly, a large set of specific regulons are often associated with a stem cell signature, some of which were also detected in the ICM (*TCF7*) or the EPI (*NANOG*), which could reflect some cell fate plasticity. It also includes *ARID3A*, which has been described to be required for TE cell maintenance [28], and known TE regulators (i.e., *GATA* and *ELF*-related factors [21]) (Fig. 2E). Taken together, our data suggest that TE Lr may be a population of TE progenitors that emerges around E7 and is maintained until at least E11.

Strikingly, we also observed a new TE cell population at E9, which we named Metallothionein Trophectoderm (TE Mt., blue dots in Fig. 2B). This TE Mt. population is only detected at E9, and its expression profile shows specific DEGs (Fig. 2C and Supplementary Fig. 4). It is characterised by an increased expression of the metallothionein-related gene family (*MT1X*, *MT1A*), but also of *SAT1,* a gene described in human trophoblastic cell apoptosis [29], and *ISG12(A),* also known to have pro-apoptotic activity in human cells [30] (Supplementary Fig. 4). The TE Mt. regulons *IRF2* and *SAT1* have been described to have joint promoters with *ISG12(A)* [31,32]. This population also shows a specific enrichment for regulons of the *HOXB* gene family, *RORC* and *HELT* (Fig. 2E). We did not observe this cluster at later stages, consistent with the fact that these cells are not cycling (Supplementary Fig. 3) and appear to be entering apoptosis as they strongly express pro-apoptotic genes. Taken together, this suggests that TE Mt. cells may correspond to the Rauber's layer that disappears around these stages.

At the ovoid stage (E9 and E11), we inferred three TE subpopulations that are specific to this stage and characterised by a more mature state of differentiation: Mature1 TE (TE Ma1, light green dots in Fig. 2B), Mature2 TE (TE Ma2, yellow dots in Fig. 2B) and Interleukin-1 TE (TE Il1, orange dots in Fig. 2B). These three populations are quite similar regarding gene expression with shared expression of TE markers and TE differentiation and functions (*ALDH1L2, TMEM254, CYP17A1, CTSL*, Fig. 2C, Supplementary Fig. 4), some of these genes (e.g., *CYP17A1*, *TMEM254*, *HSD11B2*, *GM2A* and *HAVCR1*, Supplementary Fig. 4) being described to play a role in elongation [33–35]. TE Il1 is characterised by its elevated expression level of *IL1B2* and of various RNA coding for interleukin beta-like (e.g., *ENSSSCG00000008088*, *ENSSSCG00000033667* and *ENSSSCG00000039214,* Supplementary Fig. 4). These genes are necessary for the rapid elongation of the porcine conceptus [36]. Our results support that only a subset of TE cells expresses *IL1B* at the ovoid stage. We confirmed the specific expression of *IL1B2* in a TE subpopulation by in situ hybridisation. At the spheroid stage, we observed only a few cells expressing *IL1B2*, distributed around the embryonic disc, whereas at the ovoid stage *IL1B2* expression increased and is restricted to the dorsal region of the embryo, with a gradient of expression around the embryonic disc (Fig. 2D). Regulons may drive the regionalised expression of IL1-related genes and for instance, *JDP2* displays a high activity in TE Il1 cells only. It can act together with *HAND1,* which is also highly active in this population (Fig. 2E) and plays a role in differentiation into giant trophoblastic cells [37]. Commonly expressed genes between TE Il1 and TE Ma include *CTSL* and *PTGS2*, which have also been found in extracellular vesicles extracted from the uterine fluid of pregnant ewes [38], suggesting a TE

origin of these secreted proteins in the uterine fluids. Other DEGs (e.g., *TMSB4X*, *ALDH1L2* and *UPTI,* Supplementary Table 3a, Supplementary Fig. 4) have been described in TE during placentation and conceptus elongation [39–41].TE Ma1 and Ma2 differ slightly in their regulons' activity. TE Ma1 specifically activates the regulons *PRRX1*, *HES5*, *NFKB2*, *MIXL1*, *ATF7*, *IKZF1* and *XBP1* while TE Ma2 is more similar to TE Il1 and specifically activates the regulon *ZBTB33*, *GABPA*, *FOXO3*, *HAND1*, *NR6A1* and *USF2* (Fig. 2E).

We then searched for enriched functions in TE populations by performing a GSEA (Supplementary Table 3d) and we observed that fatty acid anabolism and ovarian steroidogenesis increased during TE differentiation (Fig. 2F).

2.3. Hypoblast specification and differentiation

To further characterise the HYPO cell population, we selected 15,335 cells from our dataset corresponding to this population and performed dimensionality reductions followed by a new clustering (Fig. 3A). It led to the identification of four HYPO subpopulations (Fig. 3B). All of these subpopulations show a clear HYPO signature by expressing *APOE, COL18A1*, *GATA4* and *SOX17* (Fig. 3C and Supplementary Table 3a) [11,42], but each one presents distinct features.

These four clusters divide into two distinct groups according to developmental time and expression of specific genes (Fig. 3A). The first group includes cells from E7 and E9 embryos and comprises early hypoblast (HYPO Ea, pink dots in Fig. 3B) and intermediate hypoblast subpopulations (Hypo In, green dots in Fig. 3B). It is characterised by the expression of *APOC3* and *HIGD1A* (Fig. 3C and Supplementary Fig. 5) and shares many active regulons (Fig. 3D). The second group includes cells from E7 to E11 embryos. It corresponds to mature hypoblast (HYPO Ma, yellow dots in Fig. 3B) and visceral hypoblast (HYPO Ve, turquoise blue dots in Fig. 3B). It is characterised by the expression of higher levels of fibronectin (*FN1)* and *COL4A1* (Supplementary Fig. 5) and of genes associated with focal adhesion and ECM-receptor interaction (Fig. 3F).

In the first group, HYPO is characterised by increased activity of the *IRF2*, *LEF1*, *NFYB*, and *HELT* regulons (Fig. 3E), and some others common to the previously described TE Mt. population, including regulons driven by TFs of the *HOXB* gene family and *RORC.* HYPO Ea is characterised by increased activity of regulons corresponding to TFs known to be early hypoblast marker genes (e.g., *FOXA2*, *THAP1*, and *GATA4*) [43], but also regulons whose TFs are associated with the patterning of the anteroposterior axis (*CDX1*, *HES6*, *GBX2, MEOX1, HOXA3*) [44–46]. This first group could correspond to immature cell populations necessary for patterning the forming hypoblast.

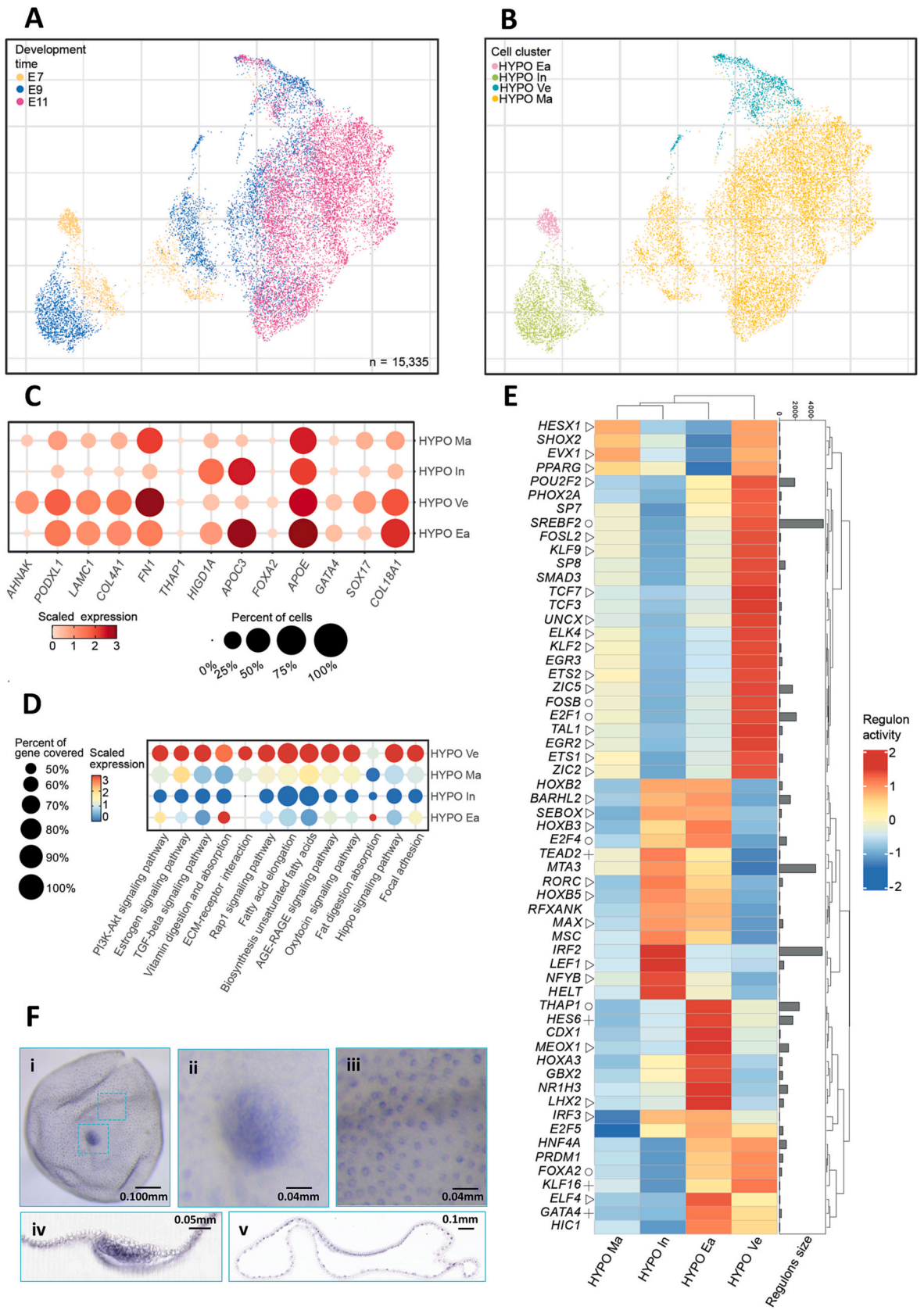
The second group of cells consists of more differentiated hypoblast cells with higher expression levels of genes associated with a mesenchymal phenotype (Fig. 3D). These cells are also more cycling (Supplementary Fig. 3), with >60% and 70% of cells in G2/M and S phases for HYPO Ma and HYPO Ve, respectively, compared to <50% for the HYPO Ea. We also observed an increased expression of genes associated with biological functions related to the biosynthesis of fatty acids and estrogen/oxytocin signalling pathways (Fig. 3D). HYPO Ma represents the majority of cells within this group and should correspond to the parietal hypoblast, underlying the trophectoderm. The second HYPO population, which was named Visceral Hypoblast (HYPO Ve), has DEG markers implicated in hypoblast differentiation (e.g., *GATA6*, *COL4 A1*, *LAMC1*, *PODXL* and *AHNAK,* Supplementary Fig. 5), which have been described in the derivation of extraembryonic endoderm from pluripotent stem cells and the regulation of hypoblast differentiation [42,47–49]. We confirmed the expression of *PODXL1* mainly in HYPO Ve by in situ hybridisation but also in the peri-nuclear region of HYPO Ma cells (Fig. 3F). This population presents a distinct regulon activity profile with HYPO-associated TFs (e.g., *FOSB*, *FOSL2*, *FOXA2*, *GATA4*) [50] but also with known targets of the Wnt signalling pathway (*TCF3*, *TCF7*), BMP/Nodal pathway (*SMAD3*) and the activation of members of the *KLF*, *ETS*, *ZIC*, and *EGR* family (*KLF9*, *KLF2*, *ETS1*, *ETS2*, *ZIC2*, *ZIC5*, *EGR2*, *EGR3*) known for their mitogenic and patterning activity (Fig. 3E). We postulate that this population could be underlying the embryonic disc and should correspond to the visceral endoderm described in mice. This hypothesis is supported by the expression of the Wnt signalling pathway inhibitors *DKK1* [51], which has been described in the mouse anterior visceral endoderm, and *BMP2,* which has been shown to induce visceral endoderm differentiation from XEN cells [52,53] (Supplementary Fig. 5).

2.4. Pluripotency states follow pig epiblast development

To further characterise the ICM and EPI cell populations, we selected 1155 cells from our dataset corresponding to these populations and performed dimensionality reductions followed by a new clustering (Fig. 4A). It revealed two distinct subpopulations (Fig. 4B), which differ in cell proliferation: more cycling cells are present in the EPI compared to the ICM (84% vs 60% in G2/M and S phases, Supplementary Fig. 3). The first subpopulation mainly comprises cells from the early blastocyst stage (E5) (Fig. 4A). It corresponds to the ICM based on the expression profiles of genes associated with naive pluripotency (e.g., *ESRRB*, *KLF4*, *PDGFRA* and *STAT3* [11,17,18]) (Fig. 4B-C and Supplementary Table 3a). This population shares a high degree of transcriptional similarity with TE Ea, as indicated by their high similarity score (>0.97) (Supplementary Fig. 7). This high similarity can also be explained by the ongoing active population segregation between ICM and TE at this stage (Supplementary Fig. 9). In the regulon heatmap (Fig. 4E), we identified early-stage regulons associated with ICM (e.g., *KLF17*, *STAT3*, *TCF7, NR2C2*) and reflecting the activity of known pathways (IL6/STAT3 and Wnt) associated with naive pluripotency in mammalian embryos [10,11]. We also found *ZNF148*, a TF known to suppress Notch signalling in induced pluripotent stem cells [54,55] and *ZNF471*, a TF described to affect stemness by down-regulating pluripotency markers (*NANOG*, *OCT4*, *SOX2*) [56] (Supplementary Fig. 6). The second subpopulation, from E7 to E11, is distinct from the ICM and expresses known EPI population markers associated with formative pluripotency genes (e.g., *NANOG*, *NODAL, DNMT3B*, *POU5F1*, *OTX2* and *SOX2,* Fig. 4C). While this subpopulation appears to be quite stable over time, sharing a high degree of transcriptional similarity (Supplementary Fig. 7), we observed a slight and graduate change in gene expression from E7 to E11, with an increase in the expression of *DNMT3*A, *LIN28A*, *NODAL* together with a decrease in the expression of *NANOG* and *STAT3* (Fig. 4C and Supplementary Table 3a). This dynamic is also observed in terms of regulon activity with a decrease in activity for *NANOG*, *OTX2*, *GBX2*, *FOXD3*, *and ETV1* and an increase in activity for *LHX1*, *LHX4*, *SMAD1*, *PROX1*, and *PATZ1* (Fig. 4E). Our study also highlights novel regulons, whose functions regarding the biology of pluripotent stem cells are poorly understood. It includes *TFDP2, DBP, EN1, RFXANK.* In situ hybridisation for *ETV1* and *TFDP2* confirmed their specific expression in the EPI (Fig. 4D).

2.5. Linking ligand-receptor interactions to regulon activation highlights potential functional interactions between the epiblast and the extraembryonic populations

To link regulons to possible signalling pathways, we used the CellComm pipeline to create pathways by connecting ligands, receptors, and TF based on their expression profiles. The network is based on known protein-protein interactions (Supplementary Table 5a), and the activity of each predicted pathway is scored using the average expression of TF, receptor, and intermediates (Supplementary Table 5b). The main results are shown in Fig. 5A and Supplementary Fig. 8. For the early blastocyst stage (E5), we found known pathways and players associated with naive pluripotency, acting either through a paracrine loop from the TE to the ICM (*ERBB2*, *FGFR4*, *PDGFRA*) or a paracrine/autocrine loop (TE to ICM or ICM to ICM) (*KIT*, *IL6ST*, *EGFR,* and *ITGB1*) (Fig. 5). For the most active ones, these pathways converge to activate the downstream



(caption on next page)

Fig. 3. Identification and characterisation of different populations in the hypoblast.
(A) Visualisation of selected HYPO cells coloured by developmental day via UMAP: E7 (yellow), E9 (blue), and E11 (pink). (B) Visualisation of HYPO populations coloured by cluster via UMAP: HYPO Ea (pink), HYPO In (green), HYPO Ma (yellow), HYPO Ve (blue). (C) Dot plot visualisation of selected DEG genes, the circle size represents the percentage of cells within the cluster that express the gene. The red gradient represents the mean scaled expression of the genes within the cluster. (D) Dot plot visualisation of selected KEGG signalling pathways. The circle size represents the percentage of genes out of all the genes in the pathways that are expressed by the cell cluster. The red gradient represents the mean scaled expression of the pathways within the cell cluster. The AGE-RAGE signalling pathway in diabetic complications has been truncated as the AGE-RAGE signalling pathway. (E) Heatmap showing scaled values of Regulon Activity Score for the 20 most specific regulons for each cluster, identified by Regulon Specificity Score (RSS). (△): common regulons with another pig study; (+): common regulons with another human study; (O): common regulons within the three studies. Right rows (heatmap): histograms distribution of regulon size (number of genes regulated by the TF in the regulons). (F) Expression of *PODXL1* (i) in the HYPO of ovoid (E11) pig blastocysts revealed by whole-mount in situ hybridization. The first dotted square highlights the area of the embryonic disc (ii), which is enlarged to show the expression of *PODXL1* in the visceral hypoblast. The second dotted square highlights extra-embryonic area (iii), which is enlarged to show the peri-nuclear expression of *PODXL1* in HYPO cells. Sections confirmed the specific expression of *PODXL1* in the HYPO underlying the epiblast (iv) and in the parietal hypoblast (v). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

regulator *STAT3,* which is known to play a key role in the ICM and naive pluripotent stem cells [10,46]. We also observed the activation of *TCF7,* which is a direct target of canonical Wnt signalling and new pathways of interest, linking either *EGFR*, *IL6ST*, or *ERBB2* to *ZNF148* or *EGFR* to *UBTF* (Fig. 5A-B and Supplementary Fig. 6). *UBTF* has been described as a regulator of human ESC differentiation by regulating rRNA synthesis. Activin A treatment has also been shown to reduce the binding effects of *UBTF* [57].

At the subsequent E7 hatched blastocyst stage, the pathway activity profile shows no particular autocrine signalling for the EPI (Fig. 5C), while signalling from TE and HYPO to EPI or TE occurs via *ITGB1*. TE also signals to EPI via *LRP2*, and HYPO shows autocrine signalling occurring via *ITGA5,* particularly in HYPO Ve (Fig. 5B and Supplementary Fig. 8).

At the subsequent E9 early ovoid blastocyst stage, paracrine and autocrine signalling by *ITGB1* and *ITGA5* are still predicted in HYPO and EPI but not between TE and EPI, where *CDH1* is mainly involved. We can also observe weak signalling from EPI to TE via *PTPRF*, which is confirmed by the expression of *PTPRF* by TE Lr (Fig. 5A and B). Interestingly *PTPRF* has been found in uterine extracellular vesicles of pregnant sows [58].

At the subsequent E11 late ovoid blastocyst stage, signalling from TE and HYPO converge on EPI, either through the activation of *ERBB3* or *ITGB1*. The two signalling pathways converge to activate *HNF4A*, *STAT3*, *HAND1*, and *SMAD1/3,* suggesting an important convergence of many signalling pathways towards EPI by ITGB1 supports the importance of the extracellular matrix and cellular contacts in transmitting the information necessary for the biology and survival of pluripotent cells. This may be a promising avenue for the recent reactivation of signalling pathways crucial for the early patterning of the embryonic disc and linked to the early steps of gastrulation.

2.6. Changes in uterine fluid composition are associated with the transition between early and late blastocysts

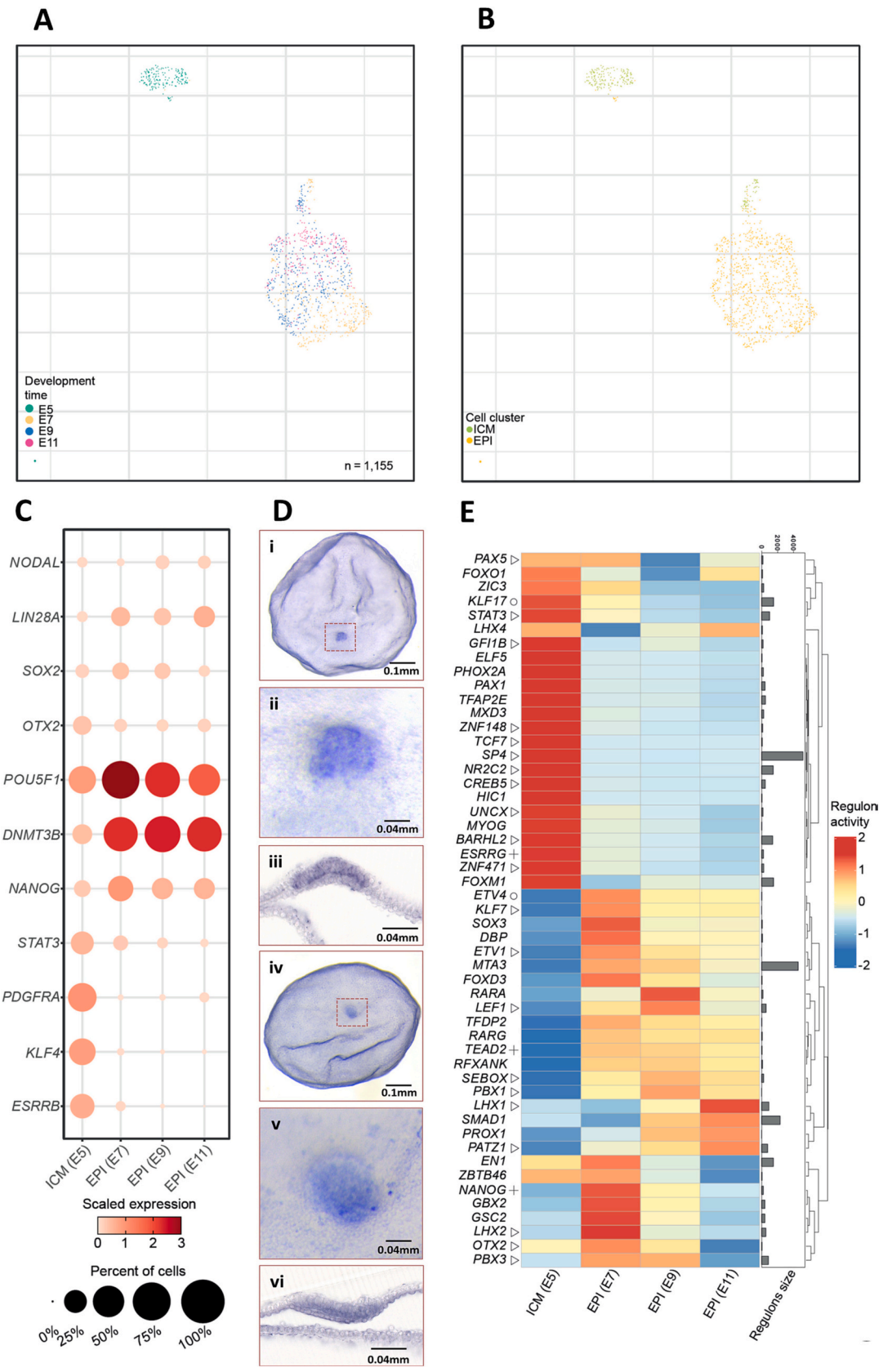
To investigate potential functional interactions between the embryo and its surrounding uterine fluids (UFs), we sampled the UFs from the same sows used to produce the embryos. Uteri were flushed to recover the embryos and the uterine fluids. Uterine fluids were analysed by liquid chromatography-mass spectrometry (LC-MS/MS). A total of 1239 proteins were identified, from which 277 were quantified in the 18 samples (Supplementary Table 6). Expression levels of the identified proteins show a clear discrimination of the proteins based on their LFQ intensities between early (5 dpf) and late (9–11 dpf) UFs, clearly seen in the PCA (Fig. 6A) as well as in the heatmap (Fig. 6B). The early stage shows a protein intensity profile with functions associated with cell metabolism, such as those involved in glycolysis GAPDH, ENO1, AKR1A1, PKM, IDH1 (Fig. 6B) [59–63] pyruvate mechanism LDHA/B (Fig. 6B) and proteins with pleiotropic functions such as proteins of 14–3-3 and YWHAQ/Z/E families, recently identified as key players during the maternal-to-zygotic transition in pigs (Fig. 6B) [64].

In the later stages (E9-E11), classical markers of maternal-embryo recognition are detected, such as the interleukin complex with IL1RAP, IL24 and IL1R1 (Fig. 6B) [65,66]. We also detected receptors identified in the TE, including UPTI, DAG1, PTPRF, DSC2, and LRP2 (Fig. 5B, Supplementary Fig. 4). Surprisingly, the JAK/STAT activator IL6ST was also identified (Supplementary Table 6). UPTI (uterine plasmin/trypsin inhibitor), DAG1, and LRP2 are also known to be expressed by endometrial cells [67–69], supporting the idea of a reciprocal loop of activation for similar signalling pathways between the endometrium and the conceptus to synchronize both for implantation. This process, which begins 7 days before the time window for implantation, seems to be concomitant with the start of embryonic elongation.

By using ligand-receptor analysis, we then associated proteins detected in the UFs (in yellow) with previously identified receptors (in blue) expressed by EPI or TE cells from E9 to E11 embryos in the scRNAseq dataset (Fig. 6C). We first highlighted the importance of extracellular matrix proteins, including FN1 and COL28A1 from UFs, which can activate integrin signalling, like ITGB1, in EPI cells (Figs. 6C and 5A). Regarding TE cells, we linked the expression of DSG2 to that of DSC2 in UFs. DSC2 has been found to interact with several desmoglein receptors to stabilize desmosomes in epithelial cells and is required to form the blastocoel in bovine embryos [70]. We also detected the EPHA-Ephrin-A signalling pathway, with Ephrin receptors expressed by EPI and TE cells and Ephrin A1 (EFNA1) by UFs. EPHA-Ephrin-A is important in the switch between pluripotency and differentiation in murine ESCs [71]. The presence of EFNA1 in uterine fluids at the onset of gastrulation also suggests that segregated EPH-EFN expression could coordinate cell fate and early differentiation before implantation. We also identified proteins annotated as negative regulators of the BMP signalling pathway: NBL1, CHRDL1, SMOC1, CRIM1, and COCH (Fig. 6B) [72] that may also be important to synchronize embryonic development and implantation. EGF is also enriched in UFs at E9-E11 stages and may interact with EPI and TE through different receptors (such as ERRB2 and LRP2, respectively) for self-renewal of pluripotent cells [73] and to prepare TE and endometrial cells for implantation [74]. For proteins identified in endometrial development, we found changes between early stages with S100A6, CAP1 (Fig. 6B) [75] and late stages with QSOX1, COCH, SMPD1 (Fig. 6B) [76–78]. At the two latest stages (E9 and E11), the detection of numerous ligands associated with receptors located on the EPI and TE strongly supports the existence of an intimate dialogue between the mother and the embryos, in order to synchronize embryonic and extraembryonic development, endometrial receptivity and, eventually, implantation.

3. Conclusions

Our study provides new insights into the formation of the first cell lineages of the early pig embryo from single-cell gene expression datasets from E5 to E11 embryos. In particular, our data reveal unsuspected dynamic evolution and heterogeneity within extraembryonic cell populations.



(caption on next page)

Fig. 4. Characterisation of transcriptional change in epiblast populations.
(A) Visualisation of selected ICM/EPI cells coloured by developmental day via UMAP: E5 (green), E7 (yellow), E9 (blue), and E11 (pink). (B) Identification of two cell populations coloured by cluster via UMAP: ICM (green) and EPI (yellow). (C) Dot plot visualisation of selected DEG genes. The circle size represents the percentage of cells within the cluster that express the gene. The red gradient represents the mean scaled expression of the genes within the cluster. (D) Expression of transcription factors *TFDP2* (i) and *ETV1* (iv) in the EPI of ovoid (E11) pig blastocysts revealed by whole-mount in situ hybridization. The dotted square highlights the area of the embryonic disc, which is enlarged (ii, v) to show the expression of *TFDP2* (ii) and *ETV1* (v) in the EPI. Sections confirmed the expression of *TFDP2* (iii) and *ETV1* (vi) in the EPI. (E) Heatmap showing scaled values of Regulon Activity Score for the 20 most specific regulons for each cluster, identified by Regulon Specificity Score (RSS). (△): common regulons with another pig study; (+): common regulons with another human study; (O): common regulons within the three studies. Right rows (heatmap): histograms distribution of regulon size (number of genes regulated by the TF in the regulons. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

First, we confirmed the major difference between the early blastocyst at E5 and the later blastocysts from E7 onwards, both for ICM/EPI and early TE cells, whose transcriptional profiles differ significantly from the later lineages. It also suggests that the time window between E5 and E7 is crucial for the segregation of the three major lineages in the pig species and would deserve a more advanced and detailed analysis to understand the molecular mechanisms involved.

For the later blastocysts, whose biology differs from that of primates and rodents, we have highlighted previously unknown subpopulations, notably from E9 onwards, at the ovoid blastocyst stage. In the trophectoderm, we confirmed the existence of interleukin-1B secreting cells belonging to a specific subpopulation localised around the embryonic disc, as well as known TE functions, such as lipid metabolism and catabolism. Above all, we discovered two previously unknown subpopulations of the TE. The first is characterised by the expression of LRP2, which could represent a subpopulation of TE progenitor or stem cells, and the second is characterised by the expression of numerous pro-apoptotic markers and disappears between E9 and E11. It could correspond to the cells of the Rauber's layer.

Concerning the hypoblast, which we detected from E7 onwards, we observed two main populations: some relatively immature and present at E7 and E9, and others more mature and observed from E7 to E11. Among the latter, we find a population that could correspond to the visceral hypoblast and another to the parietal hypoblast.

We confirmed a quite stable pluripotent state from E7 to E11 for the epiblast, with a graduate priming towards cell differentiation and gastrulation.

An original aspect of our study was to highlight regulatory modules (regulons) specific to each sub-population and potentially conserved between pigs and humans. Experimental validation will be required to confirm the relevance of such regulons in controlling the biology of pluripotent and extraembryonic cells.

Finally, combined with the analysis of the uterine fluid composition, we infer complex dialogues between the maternal environment and the cell lineages of the embryo and identified modules of cell signalling linked to TF-regulated genetic modules.

In the context of the Functional Annotation of ANimal Genomes (FAANG) international action, our work contributes to the $FAANG_{SingleCell}$ objectives by providing single-cell atlases for key tissues of farm animal species [79]. It will help to highlight key regulatory networks at play during early developmental phases and to pinpoint genetic variants involved in these regulations with potential impacts on phenotypes of interest.

4. Methods

4.1. Production of pig embryos

All the embryos were produced at the INRAE experimental unit GenESI (Rouillé, France). All the metadata associated with the biological samples used in this study have been submitted to FAANG Data Portal (https://data.faang.org/home) and BioSamples (https://www.ebi.ac.uk/biosamples/) and are summarised in the Supplementary Table 1.

Two distinct protocols were used for the production of pig embryos. A first batch of embryos (7 and 9 days after artificial insemination) was produced following superovulation (Supplementary information, Table 1). The oestrous cycle was synchronised for each sow using Altrenogest (Regumate), a synthetic progestin, for 18 days. The day after the end of the Altrenogest treatment, sows were superovulated using a first injection of gonadotrophin (1200 UI PMSG), followed 72 h later by an injection of 500 UI hCG. The day after, sows were artificially inseminated, and the insemination was repeated the following day. When the gestational time matched the embryonic stage to be sampled, surgery was performed as follows. The sow was showered the morning of the surgery and received an intramuscular anaesthetic (Ketamine, 10 mg/kg) and analgesic (Xylazine, 2 mg/kg) to calm her down. Then, the anaesthesia mask was placed on her snout; the evaporator was put into operation to diffuse the volatile anaesthetic (Isoflurane 2%). Once the effectiveness of the general anaesthesia was noted, a laparotomy was performed, and the uterus was extracted from the abdominal cavity, clamped, and the embryos were collected by retro-flushing of the uterine horns from the bottom of the horn (uterus) upwards (ovary) in 100 ml of physiological saline solution. The uterus was then replaced in the abdominal cavity, the wound was sutured and the animal was placed back into the recovery room. This procedure was authorised by the French Ministry of higher education, research, and Innovation under the authorisation number: Apafis#10376–20,170,623,130,698. The full protocol has been submitted on the FAANG Data Portal and is publicly accessible using the following link: https://api.faang.org/files/protocols/samples/INRAE_SOP_PLUS4PIGS_EMBRYOS_SAMPLING_PROTO1_20230131.pdf

A second batch of embryos was produced without superovulation (Supplementary information, Table 1). Sows were synchronised and inseminated as previously described. When the gestational time matched the embryonic stage to be sampled (5, 7, 9, and 11 days after artificial insemination), the sows were transported from the breeding unit (Rouillé, France) to the slaughterhouse (Nouzilly, France). They were stunned by electronarcosis and bled. The uterus was clamped and rapidly extracted from the abdominal cavity. Then, the embryos were collected into two tubes of 50 ml by retro-flushing of the uterine horns from the bottom of the horn (uterus) upwards (ovary) in 100 ml of physiological saline solution. The full protocol has been submitted on the FAANG Data Portal and is publicly accessible using the following link: https://api.faang.org/files/protocols/samples/INRAE_SOP_PLUS4PIGS_EMBRYOS_SAMPLING_PROTO2_20230131.pdf

4.2. Preparation of single-cell suspensions

Once recovered, embryos were staged, pooled and transported in embryo holding Media (IMV Technologies) for E5 and E7 embryos or DMEM/F12 for E9 and E11 embryos to the molecular biology laboratory in a thermostatically controlled chamber at 38 °C. At arrival, embryos were transferred to a 4-well dish and staged again under a stereomicroscope. When necessary, embryos of the same stage were pooled together into a drop of DMEM/F12 or IMV Embryo holding media and processed for cell dissociation. The full protocol is publicly accessible on the FAANG Data Portal https://api.faang.org/files/protocols/samples/INRAE_SOP_PLUS4PIGS_EMBRYOS_DISSOCIATION_PROTO3_20230131.pdf. Briefly, the zona pellucida (ZP) of E5 embryos was removed by transferring embryos into drops of 0.5% of pronase for no >5 min.

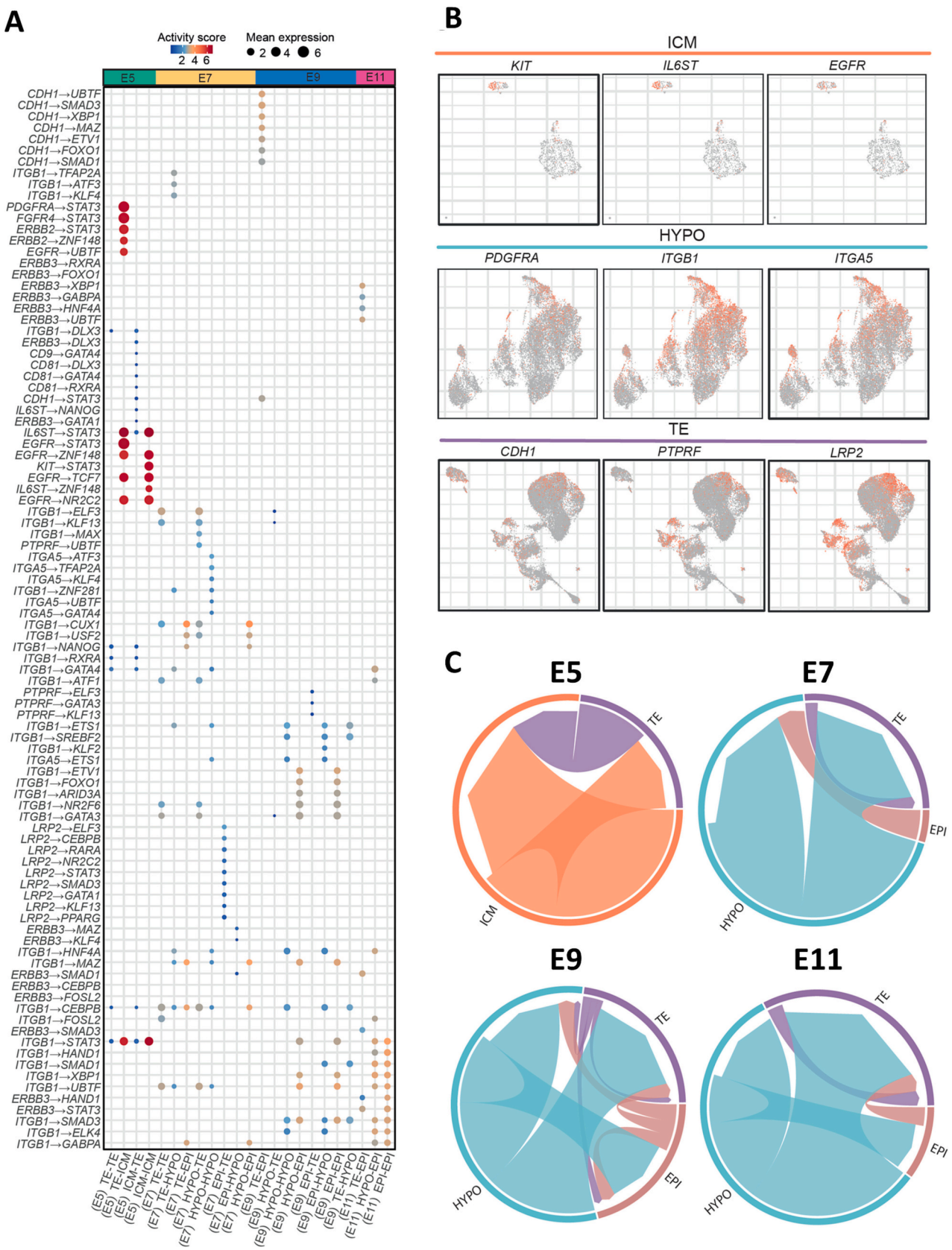


Fig. 5. Cellular crosstalks between cell lineages from E5 to E11.
(A) Dot plot displaying receptor-Transcription-Factor pairs identified for each cluster/cluster interaction coloured by activity score and sized by the cluster mean expression of the receptor. (B) UMAP plot of selected ligands and receptors for the ICM, HYPO and TE. (C) Chord diagram showing cellular interactions between clusters at the four developmental days. Arrow origins represent the sum of the mean for expressed ligands, and arrow ends represent the sum of the mean of expressed receptors.

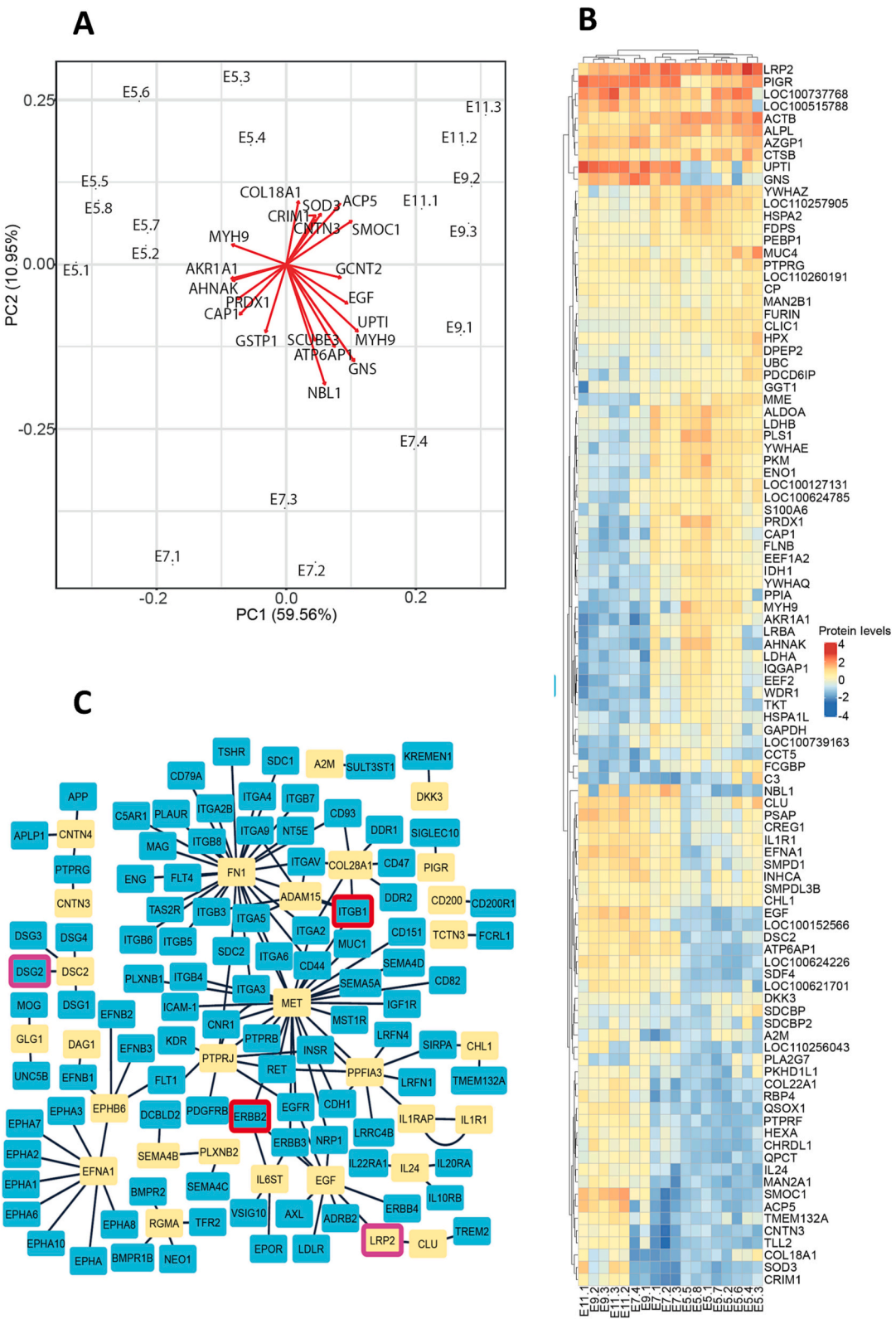


Fig. 6. Change in the uterine fluids between E5 to E11 reflects potential crosstalk between the uterus and the embryo.
(A) Principal Component Analysis (PCA) plot of pig uterine fluids. Each dot represents one uterine fluid sample. Sample IDs consist of embryonic day stages followed by the sample number. Red arrows represent prominently altered (negatively and positively) proteins for each principal component (PC1 and PC2). (B) Heatmap showing scaled protein levels values of the 87 most prominently altered proteins. (C) Visualisation of possible interactions between uterine fluid proteins and receptors on the TE or the EPI. Yellow boxes indicate proteins found in the uterine fluids, blue boxes indicate receptors, red circled boxes indicate receptors preferentially expressed in the EPI, and purple circled boxes indicate receptors preferentially expressed in the TE. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The ZP was removed by aspirating/repulsing the embryo using a pipette tip. Then, the embryos were washed into drops of embryo-holding media. All the embryos (E5 to E11) were incubated in pre-warmed Accutase for 10 min, then pre-warmed TrypLE for 10 min, followed by mechanical dissociation by several rounds of aspiration/repulsing a pipette tip.

4.3. Production of scRNAseq libraries and sequencing

Dissociated cells were washed in DMEM/F12, counted and resuspended in PBS-0.4% BSA according to 10× Genomics protocol: Chromium Next GEM Single Cell 3′ Reagent Kits v3.1 CG000204 Rev. D (for embryos sampled in 2021) or Chromium Single Cell 3′ Reagent Kits v2 CG00052 Rev. B (for embryos sampled in 2017). We loaded 1500 to 10,000 cells depending on the samples and overall, we recovered 2811 cells at E5, 6077 cells at E7, 20,518 cells at E9 and 17,539 cells at E11 (Supplementary Table 2).

After cDNA and library amplification, a Bioanalyzer (or Fragment Analyzer) profile and a Qubit quantification were performed for each sample. The libraries were sequenced on an Illumina HiSeq 3000 (batch 1) and MGI DNBSEQ-G400 (batch 2) to obtain 144 M to 328 M of raw reads per library. The full protocol has been submitted on the FAANG Data Portal and is publicly accessible using the following link: https://api.faang.org/files/protocols/experiments/INRAE_SOP_PLUS4PIGS_scRNASEQ_LIBRARIES_PROTO4_20230228.pdf

4.4. Production of an extended annotation to cover 3′ UTR from porcine transcripts

Raw sequencing files were mapped to the *Sus scrofa* genome assembly version 11.1 (GCA_000003025.6) using the Nextflow pipeline TAGADA v2.2.1 [80]. Gene positions were annotated as per Ensembl build 102, and genes were filtered based on their biotype annotation to only contain genes matching one of these categories: protein-coding, long intergenic non-coding RNA, antisense, immunoglobulin, or T-cell receptor. We then used the quant3p script (github.com/ctlab/quant3p) to extend genes in the 3′ exon for those where reads aligned past annotated genes. We used a genome parameter set to 1,341,049,888 and bam file alignment. Some additional modifications to the annotation were performed, including gene deletion, addition, or position change. These changes are summarised in the project's GitHub repository (https://github.com/Goultard59/pig_embryo_scrnaseq/blob/master/1_generating_matrices/README_annotation_extension.md) and the gtf file is available here: https://github.com/Goultard59/pig_embryo_scrnaseq/blob/master/1_generating_matrices/Sus_scrofa.Sscrofa11.1.102_10_26.filtered.gtf. Then CellRanger (version 6.1.1) was executed on this annotation to produce count matrices.

4.5. Single-cell RNA sequencing analysis

For each developmental stage, raw gene expressions were converted to a Seurat object using the Seurat R package (version 4.3.0) [81]. Cells were removed if they had >10,000 or fewer than 500–300 expressed genes or if over 10 to 25% of their UMIs were derived from the mitochondrial genome (Supplementary Table 2). Mitochondrial genes were identified with AnnotationHub package using *Sus scrofa* ENSEMBL 102 version. Genes where the sequence name corresponds to the mitochondrion were kept (as some mitochondrial genes in pigs use ENSEMBL gene names and are not recognised with Seurat ^MT pattern). Adjustments between samples have been made as described in Supplementary Table 2. After filtering, the gene expression matrices for each sample from the same stages were normalised with NormalizeDATA function from Seurat with default parameters (Log normalisation methods and 10,000 scale factors). The 2000 most variable genes were identified using FindVariablesFeatures and scaled using Seurat's function with the default settings (vst selection methods for variable features and for scaling linear models were used and a maximal value of 10 for scaling). We then performed data integration using Harmony and Seurat Wrappers using embryonic stages as grouping variable with default parameters (theta = 2, lambda = 1, sigma = 0.1, max iterations = 10) [82]. To reduce the dimensionality of this dataset, gene expression matrices were analysed by principal component analysis (PCA), from the function RunPCA from Seurat with default parameters (all variable features, 50 PCA computed, weighted by variance). The first 10–20 principal components were further used as an input with adjustments between stages as described in the Supplementary Table 2 for UMAP dimensionality reduction using the RunUMAP function with current default parameters (30 neighbouring points, uwot methods, learning rate of 1, seed of 42, density lambda of 2, density fraction of 0.3 and minimal distances between points of 0,3). Clustering was conducted using the FindClusters function with stage-adjusted resolution parameters as described in Supplementary Table 2. Cell clusters in the resulting two-dimensional representation were annotated to known biological cell types (ICM, EPI, TE, HYPO) according to curated known cell markers described in pig and other mammals [11] (Supplementary Figs. 9 and 10). At stage E11, one cluster of cells classed as unknown (UNK) was complicated to assign at a specific population. Despite the expression of AVE markers (*LEFTY2*, *OTX2* and *NODAL*), the number of cells (137) was too small to make a definitive conclusion and we decided to remove those cells from further analysis (Supplementary Fig. 10).

Cell cycle assignment was performed using CellCycleScoring Seurat function. The genes used for assignment come from the Seurat 4.0 cc. genes objects. Orthologuous genes between humans and pigs were identified using gprofiler with ENTREZGENE_ACC numeric namespaces (*TACC3* was not detected in any stage, *ANP32E* was only detected at E9). The cell cycle phase with the highest score was assigned to each cell. Quality Controls were performed to ensure a good quality distribution of transcripts, samples and cell cycle phase for each stage (Supplementary Fig. 11).

To produce the UMAP used in Fig. 1, all raw filtered matrices from all stages and samples were merged, then processed with normalisation, variable feature identification, and scaled. Dimension reduction identification and batch correction with Harmony was applied with default parameters as described previously. The first 30 dimensions were used to generate the UMAP. Cells were coloured according to the previously identified population at each stage in Fig. 1C.

To produce the UMAP used in Fig. 2, all raw filtered matrices from all stages and samples with TE-assigned cells were merged, then processed with normalisation, variable feature identification, scaling, dimension reduction identification and batch correction using Harmony with default parameters as described previously. The first 30 dimensions were used to generate the UMAP, and clustering was performed using FindCluster function from Seurat with a resolution of 0.05 and default parameters.

To generate the UMAPs used in Figs. 3 and 4, we used the same method as in Fig. 2 but with the first 20 and 25 dimensions, respectively, and clustering was performed with a resolution of 0.05 and 0.1, respectively, and default parameters.

4.6. Secondary analysis

Functional analysis (Gene Ontology, Human Phenotype Ontology, miRBase and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment) was performed to obtain the most significant pathways for each cell population using the GSVA package [83] with parameters set to ssgea method and sz minimum of 1. A linear model was then applied to the output matrix, followed by empirical Bayes statistics for differential enrichment analysis between cell types (Supplementary Table 3c). Then, DEG analyses were performed in a pairwise fashion between the different stages (Supplementary Table 3b) and between cell populations (Supplementary Table 3a), using the FindMarkers function of the Seurat package, with a filter set at 0.05 for *p*-value and at 25% for the minimum

percentage of cells expressing DEG in the given identity. Based on the resulting DEG for each cluster, an enrichment analysis was performed using the most function from the gprofiler2 package [84] with parameters set to *Sus scrofa* (Supplementary Table 3d).

4.7. Multiple factor analysis

Multiple Factor Analysis (MFA) represents an extension of PCA for the case where multiple quantitative data tables are to be simultaneously analysed [85,86]. As such, MFA is a dimension reduction method that decomposes the features from a given gene set into a lower dimension space. In particular, the MFA approach weights each table individually to ensure that tables with more features or those on a different scale do not dominate the analysis; all features within a given table are given the same weight. These weights are chosen such that the first eigenvalue of a PCA performed on each weighted table equals 1, ensuring that all tables play an equal role in the global multi-table analysis.

4.8. SCENIC analysis

A custom database for RcisTarget and GRNboost was built to run SCENIC on the pig genome. Transcription factor lists were generated following the methods used for the AnimalTFDB 3.0 [87] with the pig Ensembl build 102. The motif to Transcription Factor annotation was adapted to pig by gene orthology using OrthoFinder [88]. The motif database was built using aertslab scripts (github.com/aertslab/create_cisTarget_databases) with the best transcript score for each gene of the genomes based on the 10 kb upstream regions. RcisTargetdatabase feather files, motif to transcription factor annotation and transcription factor list are available at (https://doi.org/10.5281/zenodo.8232600). The library of motifs used in this manuscript comprises 10,560 PWMs from several sources [89]. The SCENIC pipeline was run using the VSN pipeline with an aggregation of 10 runs using raw counts matrices filtered from previous quality controls [90] (github.com/vib-singlecell-nf/vsn-pipelines).

For each regulon, RSS was computed based on the specificity results [91] (Supplementary Table 4b). The 20 regulons with the highest RSS were selected for each cell type and each stage to produce the heatmap and the CellComm analyses. For heatmap production, the AUCell regulons activity was averaged for each cell type and scaled before plotting.

4.9. Meta-analysis comparisons

Processed scRNAseq expression matrix from Meistermann et al. (2021) datasets were downloaded from https://gitlab.univ-nantes.fr/E114424Z/meistermannbruneauetalprocessed using raw counts and annotations of cells. Raw sequencing reads (fastq files) from Zhi et al. (2021) [11] datasets were downloaded from Genome Sequence Archive (GSA) using accession number CRA003960. Reads were then split using sabre (https://github.com/najoshi/sabre) with the parameters pairedEnd mode and max mismatch of 2. We used the barcode list from Supplementary Table 1 of [11] as the barcode input list for the sabre. Reads were then trimmed using Trim Galore [92]. Finally, the reads were mapped using the same script used for the human dataset, available at https://gitlab.univ-nantes.fr/E114424Z/SingleCell_Align with a heat index based on pig reference genome assembly version 11.1 (GCA_000003025.6), a pig annotation reference downloaded from ENSEMBL (version 104), and paired-end mode.

For SCENIC analysis, the previous pipeline was applied to Zhi et al. datasets (Supplementary Fig. 12). For Meistermann et al., we applied the previous pipeline with human adaptation as the v9 Motif2TF annotations and the feather file-based hg38 with TSS+/−10kbp was used (Supplementary Fig. 13). Then we compared the transcription factors identified in these two datasets with ours (Supplementary Fig. 14).

4.10. Ligand receptor analysis

Ligand to transcription factor pathways were inferred using CellComm from FUSCA [93]. Transcription factors were retrieved from our previous SCENIC analysis. Ligands and receptors were identified with LIANA packages [94], where gene names were converted from humans to pigs with OmniPath [95]. The package used sca, Naomi, logic, connectome, call_italk, and call_connectome methods. Inference between all populations at all stages was made with default parameters following CellCom tutorials. Protein-protein interaction was obtained from OmniPath with the conversion from human to pig (Supplementary Table 5a and 5b). After interactions inference, the top 20 ranked pathways by possible autocrine or paracrine interactions were kept for visualisation.

We considered the ligand-to-receptor activity score (Supplementary Table 5a) corresponding to the matching criteria to produce a circus plot visualisation of interactions between the different populations. Ligand and receptor were then assigned to a population for which sender cell types show higher expression than the average expression plus the standard deviation. Unassigned ligands and receptors were subsequently discarded. Then, circus plots were produced using the sum of the mean ligand and receptor expressions.

4.11. LC-MS/MS analysis of uterine fluids

Uterine fluids were recovered with porcine embryos by retroflushing the uterine horns from the bottom of the horn (uterus) upwards (ovary) in 100 ml of physiological saline solution. The embryos were removed, and the resulting solution was then centrifugated and the supernatant filtered through a 70 μm cell strainer and stored at −80 °C before being processed.

Samples were concentrated using centrifugal filter devices (Amicon-4 10 K, Merck) at 4000 g for 25 min. Aliquots of concentrate were brought to a concentration of 1 M urea / 50 mM ammonium bicarbonate. The samples were then reduced in 5 mM DTE for 30 min at 37 °C and alkylated in 15 mM iodoacetamide at room temperature in the dark for 30 min. Samples were digested with 70 ng of modified porcine trypsin (Promega) at 37 °C overnight. Peptides were dried using a vacuum concentrator and resuspended in 0.1% formic acid. Peptide samples were analysed using an Ultimate 3000 nano-LC system online coupled to a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific). Peptides were injected on a PepMap 100 C18 trap column (100 μm × 2 cm, 5 μM particles, Thermo Fisher Scientific) and separated with an EASY-Spray analytical column (PepMap RSLC C18, 75 μm × 50 cm, 2 μm particles, Thermo Fisher Scientific). Chromatography was performed at a 250 ml/min flow rate with 0.1% formic acid as solvent A and 0.1% formic acid as solvent B. The chromatographic method consisted of i) a 10 min equilibration step with 3% solvent B, ii) a 90 min gradient from 6% to 20% solvent B, iii) a 10 min gradient from 20% to 40% solvent B, and iv) a 10 min final elution step at 85% solvent B. MS spectra were acquired using a top-15 data-dependent acquisition method on a Q Exactive HF-X mass spectrometer.

Protein identification and quantification were performed with MaxQuant (v.1.6.1.0) and the NCBI RefSeq *Sus scrofa* database based on the Sscrofa11.1 genome assembly and Sus_scrofa_annotation_release_106. For the database search, the following parameters were used: enzyme: Trypsin/P; missed cleavages ≤2; 4.5 ppm mass tolerance for precursor main search; 20 ppm mass tolerance MS/MS search; carbamidomethylation of cysteine as fixed modification and acetyl (Protein N-terminus) as well as oxidised methionine as variable modifications. Label-free quantification was used.

4.12. Proteomics data analysis

The output table proteinGroups.txt from MaxQuant was loaded into Perseus [96] v.1.6.7 for downstream analyses. Data were filtered to

remove contaminants, reverse peptides that match a decoy database, and proteins identified only by modified peptides. The matrix was normalised by log [2], and the samples were categorised into three categories early (E5), intermediate (E7), and late (E10). Proteins were kept if there were at least in 70% of one category. Finally, missing values were replaced by a normal distribution (Supplementary information, Table S6). We used the most identified protein IDs (Majority Protein IDs from MaxQuant output) and converted the RefSeq Protein Accession to gene symbol and ENSEMBL gene pig using profiler to assign protein names. The remaining ambiguous proteins were LOC100521789, assigned to AMY2, LOC110259139, assigned to PCDHA11; and ARF3, assigned to ARF1. PCA was performed using the R function autopilot from the package ggfortify [97]. Differential expression tests between early and late categories were produced using the Perseus Volcano plot function with parameters *t*-test, on both sides, 250 randomisations, an FDR threshold of 0.05, and an S0 of 0.1. Interaction between uterine fluids and receptors from TE and EPI were produced using LIANA: single-cell data from EPI and TE at stages E9 and E11 were pulled together with 200 « artificial cells » made from protein differentially enriched in the late categories. All cells were processed on LIANA using OmniPath resources, and orthologue genes were converted using profiler and methods connectome, logic, Naomi, sca, cell phoned, total, call_squidpy, call_cellchat, call_connectome, call_italk. Results were filtered to keep only interactions between uterine fluids as ligands and TE and EPI cell populations as receptors. The resulting interactions were visualised and plotted using Cytoscape [98]. Receptors were assigned to a cell population based on their appearance on the differential expression testing between EPI and TE.

4.13. Whole-mount in situ hybridization

Whole-mount in situ hybridization was carried out as described previously [99]. Digoxigenin-labelled probes were synthesized by VectorBuilder from the partial pig cDNAs of *TFDP2, ETV1, IL1B2, LRP2, PODXL1* and inserted into VectorBuilder In Vitro Transcription Vector (for In Situ hybridization). Sequences used for the synthesis of the probes are provided in Supplementary Table 7. After in situ hybridization, embryos were photographed under a Leica M80 stereomicroscope and subsequently embedded in gelatin, sectioned at 40 μm using a Leica VT1000S vibratome. All slides were scanned with a Pannoramic Scan 150 (3D Histech) and analysed with a CaseCenter 2.9 viewer (3D Histech).

Contributions

J.S., Y.B., P.M., F.M., T.F., J.A., and H.A. performed experiments. A. D., C.K., J.S., D.L., T.F., S.F., J.A., and H.A. analysed the data and prepared figs. A.G.T., S.F., and B.P. provided reagents, technical expertise, and conceptual inputs. A.D., S.F., J.A., and H.A. wrote the paper.

Funding

This work was supported by the ANR PluS4PiGs (ANR-19-CE20-0019). Adrien Dufour is funded by the DIM-1HEALTH from Région Île-de-France, the Animal Genetics division of INRAE and the IFIP, Institut du porc. Adrien Dufour is also the recipient of an EMBO scientific exchange fellowship.

Author statement

We certify that all *authors* have seen and approved the final version of the manuscript being submitted. We warrant that the article is the *authors*' original work, hasn't received prior publication and isn't under consideration for publication elsewhere.

CRediT authorship contribution statement

Adrien Dufour: Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing, Visualization. **Cyril Kurylo:** Data curation, Software. **Jan B. Stöckl:** Data curation, Formal analysis. **Denis Laloë:** Formal analysis, Methodology. **Yoann Bailly:** Formal analysis, Resources. **Patrick Manceau:** Methodology. **Frédéric Martins:** Investigation, Methodology. **Ali G. Turhan:** Funding acquisition, Resources. **Stéphane Ferchaud:** Formal analysis, Methodology, Resources. **Bertrand Pain:** Conceptualization, Funding acquisition. **Thomas Fröhlich:** Formal analysis, Funding acquisition, Methodology, Resources. **Sylvain Foissac:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Resources, Supervision, Writing – original draft. **Jérôme Artus:** Conceptualization, Funding acquisition, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing. **Hervé Acloque:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the data of this study are publicly available and have been deposited into public repositories. All raw sequencing data and associated metadata are available in the FAANG Data portal (https://data.faang.org/home) and ENA under accession number PRJEB60517 with their associated metadata (Supplementary Table 8). The code used for the analysis is available at (https://doi.org/10.5281/zenodo.10261174). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [100] partner repository with the dataset identifier PXD042421.

Acknowledgements

We are grateful to the Genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) and Claire Kuchly (GeT-PlaGe) for providing help, computing, and storage resources. We are grateful to people from the INRAE experimental farm (https://doi.org/10.15454/1.5572415481185847E12) who took care for the animals. We are grateful to Sylvain Bourgeois and the people from the Unité Expérimentale de Physiologie Animale de l'Orfrasière INRAE. We thank Marie-José Mercat from the IFIP, Institut du Porc, for her help and funding. We also thank Sarah Djebali, Nathalie Beaujean, Nathalie Vialaneix, Camille Berthelot, and Laurent David for sharing data, codes, pipelines, and helpful suggestions. We thank M. Kösters at LMU, Michel Cohen-Tannoudji and Laure Bally-Cuif at Pasteur Institute and Julie Rivière at @BRIDGe facility for excellent technical assistance. We thank Gert Hulselmans and Chris C Flerins for their helps on Github. We thank people from our respective laboratories for their helpful comments and discussion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2023.110780.

References

[1] B.P. Griffith, C.E. Goerlich, A.K. Singh, M. Rothblatt, C.L. Lau, A. Shah, M. Lorber, A. Grazioli, K.K. Saharia, S.N. Hong, et al., Genetically modified porcine-to-human cardiac xenotransplantation, N. Engl. J. Med. 387 (2022) 35–44, https://doi.org/10.1056/NEJMoa2201422.

[2] I.Hue Artus, H. Acloque, Preimplantation Development in Ungulates: a 'Ménage à Quatre' Scenario, Reproduction 159 (2020) R151–R172, https://doi.org/10.1530/REP-19-0348.

[3] D. Meistermann, A. Bruneau, S. Loubersac, A. Reignier, J. Firmin, V. François-Campion, S. Kilens, Y. Lelièvre, J. Lammers, M. Feyeux, et al., Integrated pseudotime analysis of human pre-implantation embryo single-cell transcriptomes reveals the dynamics of lineage specification, Cell Stem Cell 28 (2021) 1625–1640.e6, https://doi.org/10.1016/j.stem.2021.04.027.

[4] T. Boroviak, G.G. Stirparo, S. Dietmann, I. Hernando-Herraez, H. Mohammed, W. Reik, A. Smith, E. Sasaki, J. Nichols, P. Bertone, Single cell transcriptome analysis of human, marmoset and mouse embryos reveals common and divergent features of preimplantation development, Development 145 (2018) dev167833, https://doi.org/10.1242/dev.167833.

[5] S.K. Mahadevaiah, M.N. Sangrithi, T. Hirota, J.M.A. Turner, A single-cell transcriptome atlas of marsupial embryogenesis and X inactivation, Nature. (2020), https://doi.org/10.1038/s41586-020-2629-6.

[6] X. Gao, M. Nowak-Imialek, X. Chen, D. Chen, D. Herrmann, D. Ruan, A.C. H. Chen, M.A. Eckersley-Maslin, S. Ahmad, Y.L. Lee, et al., Establishment of porcine and human expanded potential stem cells, Nat. Cell Biol. 21 (2019) 687–699, https://doi.org/10.1038/s41556-019-0333-2.

[7] Y. Yuan, J. Park, Y. Tian, J. Choi, R. Pasquariello, A.P. Alexenko, A. Dai, S. K. Behura, R.M. Roberts, T. Ezashi, A six-inhibitor culture medium for improving naïve-type pluripotency of porcine pluripotent stem cells, Cell Death Dis. 5 (2019) 104, https://doi.org/10.1038/s41420-019-0184-4.

[8] R. Alberio, N. Croxall, C. Allegrucci, Pig epiblast stem cells depend on Activin/Nodal signaling for pluripotency and self-renewal, Stem Cells Dev. 19 (2010) 1627–1636, https://doi.org/10.1089/scd.2010.0012.

[9] D.-R. Hou, Y. Jin, X.-W. Nie, M.-L. Zhang, N. Ta, L.-H. Zhao, N. Yang, Y. Chen, Z.-Q. Wu, H.-B. Jiang, et al., Derivation of porcine embryonic stem-like cells from in vitro-produced blastocyst-stage embryos, Sci. Rep. 6 (2016) 25838, https://doi.org/10.1038/srep25838.

[10] P. Ramos-Ibeas, F. Sang, Q. Zhu, W.W.C. Tang, S. Withey, D. Klisch, L. Wood, M. Loose, M.A. Surani, R. Alberio, Pluripotency and X chromosome dynamics revealed in pig pre-gastrulating embryos by single cell analysis, Nat. Commun. 10 (2019) 500, https://doi.org/10.1038/s41467-019-08387-8.

[11] M. Zhi, J. Zhang, Q. Tang, D. Yu, S. Gao, D. Gao, P. Liu, J. Guo, T. Hai, J. Gao, et al., Generation and characterization of stable pig pregastrulation epiblast stem cell lines, Cell Res. 32 (2022) 383–400, https://doi.org/10.1038/s41422-021-00592-9.

[12] T. Liu, J. Li, L. Yu, H.-X. Sun, J. Li, G. Dong, Y. Hu, Y. Li, Y. Shen, J. Wu, et al., Cross-species single-cell transcriptomic analysis reveals pre-gastrulation developmental differences among pigs, monkeys, and humans, Cell Discov. 7 (2021) 8, https://doi.org/10.1038/s41421-020-00238-x.

[13] G. Guo, J. Yang, J. Nichols, J.S. Hall, I. Eyres, W. Mansfield, A. Smith, Klf4 reverts developmentally programmed restriction of ground state pluripotency, Development 136 (2009) 1063–1069, https://doi.org/10.1242/dev.030957.

[14] A.S. Bernardo, A. Jouneau, H. Marks, P. Kensche, J. Kobolak, K. Freude, V. Hall, A. Feher, Z. Polgar, C. Sartori, et al., Mammalian embryo comparison identifies novel pluripotency genes associated with the naïve or primed state, Biol. Open (2018), https://doi.org/10.1242/bio.033282 bio.033282.

[15] R. Kumar, T. Evans, Activation-induced cytidine deaminase regulates fibroblast growth factor/extracellular signal-regulated kinases signaling to achieve the Naïve pluripotent state during reprogramming, Stem Cells 37 (2019) 1003–1017, https://doi.org/10.1002/stem.3023.

[16] S. Aibar, C.B. González-Blas, T. Moerman, V.A. Huynh-Thu, H. Imrichova, G. Hulselmans, F. Rambow, J.-C. Marine, P. Geurts, J. Aerts, et al., SCENIC: single-cell regulatory network inference and clustering, Nat. Methods 14 (2017) 1083–1086, https://doi.org/10.1038/nmeth.4463.

[17] S. Petropoulos, D. Edsgärd, B. Reinius, Q. Deng, S.P. Panula, S. Codeluppi, A. Plaza Reyes, S. Linnarsson, R. Sandberg, F. Lanner, Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos, Cell 165 (2016) 1012–1026, https://doi.org/10.1016/j.cell.2016.03.023.

[18] Y.E. Wu, L. Pan, Y. Zuo, X. Li, W. Hong, Detecting activated cell populations using single-cell RNA-Seq, Neuron 96 (2017) 313–329.e6, https://doi.org/10.1016/j.neuron.2017.09.026.

[19] Z. Cao, T.S. Carey, A. Ganguly, C.A. Wilson, S. Paul, J.G. Knott, Transcription factor AP-2γ induces early *Cdx2* expression and represses HIPPO signaling to specify the trophectoderm lineage, Development (2015), https://doi.org/10.1242/dev.120238 dev.120238.

[20] C. Gerri, A. McCarthy, G. Alanis-Lobato, A. Demtschenko, A. Bruneau, S. Loubersac, N.M.E. Fogarty, D. Hampshire, K. Elder, P. Snell, et al., Initiation of a conserved trophectoderm program in human, cow and mouse embryos, Nature 587 (2020) 443–447, https://doi.org/10.1038/s41586-020-2759-x.

[21] C. Dong, M. Beltcheva, P. Gontarz, B. Zhang, P. Popli, L.A. Fischer, S.A. Khan, K. Park, E.-J. Yoon, X. Xing, et al., Derivation of trophoblast stem cells from naïve human pluripotent stem cells, eLife 9 (2020) e52504, https://doi.org/10.7554/eLife.52504.

[22] H. Gao, R. Gao, L. Zhang, W. Xiu, R. Zang, H. Wang, Y. Zhang, J. Chen, Y. Gao, S. Gao, Esrrb plays important roles in maintaining self-renewal of trophoblast stem cells (TSCs) and reprogramming somatic cells to induced TSCs, J. Mol. Cell Biol. 11 (2019) 463–473, https://doi.org/10.1093/jmcb/mjy054.

[23] J. Yang, W. Wang, J. Ooi, L.S. Campos, L. Lu, P. Liu, Signalling through retinoic acid receptors is required for reprogramming of both mouse embryonic fibroblast cells and epiblast stem cells to induced pluripotent stem cells, Stem Cells 33 (2015) 1390–1404, https://doi.org/10.1002/stem.1926.

[24] S.C. Bleckmann, J.A. Blendy, D. Rudolph, A.P. Monaghan, W. Schmid, G. Schütz, Activating transcription factor 1 and CREB are important for cell survival during early mouse development, Mol. Cell. Biol. 22 (2002) 1919–1925, https://doi.org/10.1128/MCB.22.6.1919-1925.2002.

[25] B.-K. Lee, N. Uprety, Y.J. Jang, S.K. Tucker, C. Rhee, L. LeBlanc, S. Beck, J. Kim, Fosl1 overexpression directly activates trophoblast-specific gene expression programs in embryonic stem cells, Stem Cell Res. 26 (2018) 95–102, https://doi.org/10.1016/j.scr.2017.12.004.

[26] I. Bedzhov, E. Liszewska, B. Kanzler, M.P. Stemmler, Igf1r signaling is indispensable for preimplantation development and is activated via a novel function of E-cadherin, PLoS Genet. 8 (2012) e1002609, https://doi.org/10.1371/journal.pgen.1002609.

[27] L. Cai, Y. Jeong, Y. Jin, J. Lee, Y. Jeong, K. Hwang, S. Hyun, W. Hwang, Effects of human recombinant granulocyte-colony stimulating factor treatment during in vitro culture on porcine pre-implantation embryos, PLoS One 15 (2020) e0230247, https://doi.org/10.1371/journal.pone.0230247.

[28] C. Rhee, B.-K. Lee, S. Beck, A. Anjum, K.R. Cook, M. Popowski, H.O. Tucker, J. Kim, Arid3a is essential to execution of the first cell fate decision via direct embryonic and extraembryonic transcriptional regulation, Genes Dev. 28 (2014) 2219–2232, https://doi.org/10.1101/gad.247163.114.

[29] T. Liang, J. Xie, J. Zhao, W. Huang, Z. Xu, P. Tian, C. Mi, M. Dai, S. Zhang, H. Zhang, Novel lnc-HZ03 and miR-hz03 promote BPDE-induced human trophoblastic cell apoptosis and induce miscarriage by upregulating p53/SAT1 pathway, Cell Biol. Toxicol. 37 (2021) 951–970, https://doi.org/10.1007/s10565-021-09583-3.

[30] H. Gytz, M.F. Hansen, S. Skovbjerg, A.C.M. Kristensen, S. Hørlyck, M.B. Jensen, M. Fredborg, L.D. Markert, N.A. McMillan, E.I. Christensen, et al., Apoptotic properties of the type 1 interferon induced family of human mitochondrial membrane ISG12 proteins: ISG12-induced apoptosis, Biol. Cell. 109 (2017) 94–112, https://doi.org/10.1111/boc.201600034.

[31] P.M. Martensen, T.M.M. Søgaard, I.M. Gjermandsen, H.N. Buttenschøn, A. B. Rossing, V. Bonnevie-Nielsen, C. Rosada, J.L. Simonsen, J. Justesen, The interferon alpha induced protein ISG12 is localized to the nuclear membrane: ISG12 is localized to the nuclear membrane, Eur. J. Biochem. 268 (2001) 5947–5954, https://doi.org/10.1046/j.0014-2956.2001.02545.x.

[32] B.C. Mounce, E.Z. Poirier, G. Passoni, E. Simon-Loriere, T. Cesaro, M. Prot, K. A. Stapleford, G. Moratorio, A. Sakuntabhai, J.-P. Levraud, et al., Interferon-induced spermidine-spermine acetyltransferase and polyamine depletion restrict Zika and chikungunya viruses, Cell Host Microbe 20 (2016) 167–177, https://doi.org/10.1016/j.chom.2016.06.011.

[33] X. Chu, C.J. Corbin, M.A. Kaminski, A.J. Conley, Unique Regulation of CYP17 Expression in the Trophectoderm of the Preattachment Porcine Blastocyst, Endocrinology 140 (1999) 632–640, https://doi.org/10.1210/endo.140.2.6472140.

[34] K. Brooks, G. Burns, T.E. Spencer, Biological roles of hydroxysteroid (11-Beta) dehydrogenase 1 (HSD11B1), HSD11B2, and glucocorticoid receptor (NR3C1) in sheep conceptus elongation1, Biol. Reprod. 93 (2015), https://doi.org/10.1095/biolreprod.115.130757.

[35] G. Guo, G.G. Stirparo, S.E. Strawbridge, D. Spindlow, J. Yang, J. Clarke, A. Dattani, A. Yanagida, M.A. Li, S. Myers, et al., Human naive epiblast cells possess unrestricted lineage potential, Cell Stem Cell 28 (2021) 1040–1056.e6, https://doi.org/10.1016/j.stem.2021.02.025.

[36] J.J. Whyte, A.E. Meyer, L.D. Spate, J.A. Benne, R. Cecil, M.S. Samuel, C. N. Murphy, R.S. Prather, R.D. Geisert, Inactivation of porcine interleukin-1β results in failure of rapid conceptus elongation, Proc. Natl. Acad. Sci. 115 (2018) 307–312, https://doi.org/10.1073/pnas.1718004115.

[37] G. Wu, L. Gentile, T. Fuchikami, J. Sutter, K. Psathaki, T.C. Esteves, M.J. Araúzo-Bravo, C. Ortmeier, G. Verberk, K. Abe, et al., Initiation of trophectoderm lineage specification in mouse embryos is independent of Cdx2, Development 137 (2010) 4159–4169, https://doi.org/10.1242/dev.056630.

[38] G. Burns, K. Brooks, M. Wildung, R. Navakanitworakul, L.K. Christenson, T. E. Spencer, Extracellular vesicles in luminal fluid of the ovine uterus, PLoS One 9 (2014) e90913, https://doi.org/10.1371/journal.pone.0090913.

[39] D. Liu, Y. Chen, Y. Ren, P. Yuan, N. Wang, Q. Liu, C. Yang, Z. Yan, M. Yang, J. Wang, et al., Primary specification of blastocyst trophectoderm by scRNA-seq: new insights into embryo implantation, Sci. Adv. 8 (2022) eabj3725, https://doi.org/10.1126/sciadv.abj3725.

[40] L. Aghajanova, S. Shen, A.M. Rojas, S.J. Fisher, J.C. Irwin, L.C. Giudice, Comparative transcriptome analysis of human trophectoderm and embryonic stem cell-derived trophoblasts reveal key participants in early implantation1, Biol. Reprod. 86 (2012), https://doi.org/10.1095/biolreprod.111.092775.

[41] J. Xu, L. Yu, J. Guo, J. Xiang, Z. Zheng, D. Gao, B. Shi, H. Hao, D. Jiao, L. Zhong, et al., Generation of pig induced pluripotent stem cells using an extended pluripotent stem cell culture system, Stem Cell Res Ther 10 (2019) 193, https://doi.org/10.1186/s13287-019-1303-0.

[42] B. Sozen, V. Jorgensen, B.A.T. Weatherbee, S. Chen, M. Zhu, M. Zernicka-Goetz, Reconstructing aspects of human embryogenesis with pluripotent stem cells, Nat. Commun. 12 (2021) 5550, https://doi.org/10.1038/s41467-021-25853-4.

[43] X.A. Wolf, P. Serup, P. Hyttel, Three-dimensional immunohistochemical characterization of lineage commitment by localization of T and FOXA2 in

porcine peri-implantation embryos, Dev. Dyn. 240 (2011) 890–897, https://doi.org/10.1002/dvdy.22602.
[44] A. Kawamura, H. Ovara, Y. Ooka, H. Kinoshita, M. Hoshikawa, K. Nakajo, D. Yokota, Y. Fujino, S. Higashijima, S. Takada, et al., Posterior–anterior gradient of zebrafish hes6 expression in the presomitic mesoderm is established by the combinatorial functions of the downstream enhancer and 3′UTR, Dev. Biol. 409 (2016) 543–554, https://doi.org/10.1016/j.ydbio.2015.11.010.
[45] E. van den Akker, S. Forlani, K. Chawengsaksophak, W. de Graaff, F. Beck, B. I. Meyer, J. Deschamps, *Cdx1* and *Cdx2* have overlapping functions in anteroposterior patterning and posterior axis elongation, Development 129 (2002) 2181–2193, https://doi.org/10.1242/dev.129.9.2181.
[46] M. Wang, L. Tang, D. Liu, Q.-L. Ying, S. Ye, The transcription factor Gbx2 induces expression of Kruppel-like factor 4 to maintain and induce naïve pluripotency of embryonic stem cells, J. Biol. Chem. 292 (2017) 17121–17128, https://doi.org/10.1074/jbc.M117.803254.
[47] N. Smyth, H.S. Vatansever, P. Murray, M. Meyer, C. Frie, M. Paulsson, D. Edgar, Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation, J. Cell Biol. 144 (1999) 151–160, https://doi.org/10.1083/jcb.144.1.151.
[48] K.M. Mackinlay, B.A. Weatherbee, V. Souza Rosa, C.E. Handford, G. Hudson, T. Coorens, L.V. Pereira, S. Behjati, L. Vallier, M.N. Shahbazi, et al., An in vitro stem cell model of human epiblast and yolk sac interaction, eLife 10 (2021) e63930, https://doi.org/10.7554/eLife.63930.
[49] H.J. Lim, J. Kim, C.-H. Park, S.A. Lee, M.R. Lee, K.-S. Kim, J. Kim, Y.S. Bae, Regulation of c-Myc expression by Ahnak promotes induced pluripotent stem cell generation, J. Biol. Chem. 291 (2016) 752–761, https://doi.org/10.1074/jbc.M115.659276.
[50] E. Corujo-Simon, A.H. Radley, J. Nichols, Evidence implicating sequential commitment of the founder lineages in the human blastocyst by order of hypoblast gene activation, Development 150 (2023) dev201522, https://doi.org/10.1242/dev.201522.
[51] C.D. Stern, K.M. Downs, The hypoblast (visceral endoderm): an evo-devo perspective, Development 139 (2012) 1059–1069, https://doi.org/10.1242/dev.070730.
[52] A. Paca, C.A. Séguin, M. Clements, M. Ryczko, J. Rossant, T.A. Rodriguez, T. Kunath, BMP signaling induces visceral endoderm differentiation of XEN cells and parietal endoderm, Dev. Biol. 361 (2012) 90–102, https://doi.org/10.1016/j.ydbio.2011.10.013.
[53] J. Artus, P. Douvaras, A. Piliszek, J. Isern, M.H. Baron, A.-K. Hadjantonakis, BMP4 signaling directs primitive endoderm-derived XEN cells to an extraembryonic visceral endoderm identity, Dev. Biol. 361 (2012) 245–262, https://doi.org/10.1016/j.ydbio.2011.10.015.
[54] N. Salomonis, P.J. Dexheimer, L. Omberg, R. Schroll, S. Bush, J. Huo, L. Schriml, S. Ho Sui, M. Keddache, C. Mayhew, et al., Integrated genomic analysis of diverse induced pluripotent stem cells from the progenitor cell biology consortium, Stem Cell Rep. 7 (2016) 110–125, https://doi.org/10.1016/j.stemcr.2016.05.006.
[55] N. Wang, M. Li, Y. Liu, J. Yu, J. Ren, Z. Zheng, S. Wang, S. Yang, S. Yang, L. Liu, et al., ZBP-89 negatively regulates self-renewal of liver cancer stem cells via suppression of Notch1 signaling pathway, Cancer Lett. 472 (2020) 70–80, https://doi.org/10.1016/j.canlet.2019.12.026.
[56] C. Tao, J. Luo, J. Tang, D. Zhou, S. Feng, Z. Qiu, T.C. Putti, T. Xiang, Q. Tao, L. Li, et al., The tumor suppressor zinc finger protein 471 suppresses breast cancer growth and metastasis through inhibiting AKT and Wnt/β-catenin signaling, Clin. Epigenetics 12 (2020) 173, https://doi.org/10.1186/s13148-020-00959-6.
[57] J.L. Woolnough, B.L. Atwood, Z. Liu, R. Zhao, K.E. Giles, The regulation of rRNA gene transcription during directed differentiation of human embryonic stem cells, PLoS One 11 (2016) e0157276, https://doi.org/10.1371/journal.pone.0157276.
[58] J. Szuszkiewicz, K. Myszczynski, Z.P. Reliszko, Y. Heifetz, M.M. Kaczmarek, Early steps of embryo implantation are regulated by exchange of extracellular vesicles between the embryo and the endometrium, FASEB J. 36 (2022), https://doi.org/10.1096/fj.202200677R.
[59] G. Peng, S. Suo, G. Cui, F. Yu, R. Wang, J. Chen, S. Chen, Z. Liu, G. Chen, Y. Qian, et al., Molecular architecture of lineage allocation and tissue organization in early mouse embryo, Nature 572 (2019) 528–532, https://doi.org/10.1038/s41586-019-1469-8.
[60] M. Arias-Alvarez, P. Bermejo-Alvarez, A. Gutierrez-Adan, D. Rizos, P.L. Lorenzo, P. Lonergan, Effect of leptin supplementation during in vitro oocyte maturation and embryo culture on bovine embryo development and gene expression patterns, Theriogenology 75 (2011) 887–896, https://doi.org/10.1016/j.theriogenology.2010.10.031.
[61] M. García-Herreros, C.A. Simintiras, P. Lonergan, Temporally differential protein expression of glycolytic and glycogenic enzymes during in vitro preimplantation bovine embryo development, Reprod. Fertil. Dev. 30 (2018) 1245, https://doi.org/10.1071/RD17429.
[62] Y. Xu, Y. Zhang, J.C. García-Cañaveras, L. Guo, M. Kan, S. Yu, I.A. Blair, J. D. Rabinowitz, X. Yang, Chaperone-mediated autophagy regulates the pluripotency of embryonic stem cells, Science 369 (2020) 397–403, https://doi.org/10.1126/science.abb4467.
[63] H. Miyazawa, Y. Yamaguchi, Y. Sugiura, K. Honda, K. Kondo, F. Matsuda, T. Yamamoto, M. Suematsu, M. Miura, Rewiring of embryonic glucose metabolism via suppression of PFK-1 and aldolase during mouse chorioallantoic branching, Development 144 (2017) 63–73, https://doi.org/10.1242/dev.138545.
[64] Z.-Q. Du, H. Liang, X.-M. Liu, Y.-H. Liu, C. Wang, C.-X. Yang, Single cell RNA-seq reveals genes vital to in vitro fertilized embryos and parthenotes in pigs, Sci. Rep. 11 (2021) 14393, https://doi.org/10.1038/s41598-021-93904-3.
[65] H. Ka, H. Seo, Y. Choi, I. Yoo, J. Han, Endometrial response to conceptus-derived estrogen and interleukin-1β at the time of implantation in pigs, J. Anim. Sci. Biotechnol. 9 (2018) 44, https://doi.org/10.1186/s40104-018-0259-8.
[66] H. Seo, Y. Choi, J. Shim, Y. Choi, H. Ka, Regulatory mechanism for expression of IL1B receptors in the uterine endometrium and effects of IL1B on prostaglandin synthetic enzymes during the implantation period in Pigs1, Biol. Reprod. 87 (2012), https://doi.org/10.1095/biolreprod.112.099051.
[67] M.L. Stallings-Mann, M.G. Burke, W.E. Trout, R.M. Roberts, Purification, characterization, and cDNA cloning of a Kunitz-type proteinase inhibitor secreted by the porcine uterus, J. Biol. Chem. 269 (1994) 24090–24094.
[68] M. Santhanakrishnan, K. Ray, K. Oppenheimer, E.A. Bonney, Dynamic regulation of alpha-dystroglycan in mouse placenta, Placenta 29 (2008) 932–936, https://doi.org/10.1016/j.placenta.2008.08.021.
[69] S.J. Oh, T.H. Kim, J.M. Lim, J.-W. Jeong, Progesterone induces expression of Lrp2 in the murine uterus, Biochem. Biophys. Res. Commun. 441 (2013) 175–179, https://doi.org/10.1016/j.bbrc.2013.10.037.
[70] F. Wang, X. Tian, L. Zhang, C. Gao, C. He, Y. Fu, P. Ji, Y. Li, N. Li, G. Liu, Beneficial effects of melatonin on in vitro bovine embryonic development are mediated by melatonin receptor 1, J. Pineal Res. 56 (2014) 333–342, https://doi.org/10.1111/jpi.12126.
[71] R. Fernandez-Alonso, F. Bustos, M. Budzyk, P. Kumar, A.O. Helbig, J. Hukelmann, A.I. Lamond, F. Lanner, H. Zhou, E. Petsalaki, et al., Phosphoproteomics identifies a bimodal EPHA2 receptor switch that promotes embryonic stem cell differentiation, Nat. Commun. 11 (2020) 1357, https://doi.org/10.1038/s41467-020-15173-4.
[72] M. Katoh, M. Katoh, CER1 is a common target of WNT and NODAL signaling pathways in human embryonic stem cells, Int. J. Mol. Med. (2006), https://doi.org/10.3892/ijmm.17.5.795.
[73] L. Wang, T.C. Schulz, E.S. Sherrer, D.S. Dauphin, S. Shin, A.M. Nelson, C.B. Ware, M. Zhan, C.-Z. Song, X. Chen, et al., Self-renewal of human embryonic stem cells requires insulin-like growth factor-1 receptor and ERBB2 receptor signaling, Blood 110 (2007) 4111–4119, https://doi.org/10.1182/blood-2007-03-082586.
[74] W. Jeong, S. Jung, F.W. Bazer, G. Song, J. Kim, Epidermal growth factor: porcine uterine luminal epithelial cell migratory signal during the peri-implantation period of pregnancy, Mol. Cell. Endocrinol. 420 (2016) 66–74, https://doi.org/10.1016/j.mce.2015.11.023.
[75] J. von Hof, N. Sprekeler, G. Schuler, A. Boos, M.P. Kowalewski, Uterine and placental expression of HPGD in cows during pregnancy and release of fetal membranes, Prostaglandins Other Lipid Mediat. 128–129 (2017) 17–26, https://doi.org/10.1016/j.prostaglandins.2016.12.003.
[76] M. Hoelker, D. Salilew-Wondim, M. Drillich, G.-B. Christine, N. Ghanem, L. Goetze, D. Tesfaye, K. Schellander, W. Heuwieser, Transcriptional response of the bovine endometrium and embryo to endometrial polymorphonuclear neutrophil infiltration as an indicator of subclinical inflammation of the uterine environment, Reprod. Fertil. Dev. 24 (2012) 778, https://doi.org/10.1071/RD11171.
[77] T. Kaneko-Tarui, L. Zhang, K.J. Austin, L.E. Henkes, J. Johnson, T.R. Hansen, J. K. Pru, Maternal and embryonic control of uterine sphingolipid-metabolizing enzymes during murine embryo Implantation1, Biol. Reprod. 77 (2007) 658–665, https://doi.org/10.1095/biolreprod.107.061044.
[78] H.-S. Lin, R.K.-K. Lee, T.-H. Yang, H.-W. Fang, S.-H. Li, Spatiotemporal protein expression profiles of QSOX1 in the murine uterus, placenta, and embryo during pregnancy, Appl. Sci. 11 (2021) 10151, https://doi.org/10.3390/app112110151.
[79] E.L. Clark, A.L. Archibald, H.D. Daetwyler, M.A.M. Groenen, P.W. Harrison, R. D. Houston, C. Kühn, S. Lien, D.J. Macqueen, J.M. Reecy, et al., From FAANG to fork: application of highly annotated genomes to improve farmed animal production, Genome Biol. 21 (2020) 285, https://doi.org/10.1186/s13059-020-02197-8.
[80] C. Kurylo, C. Guyomar, S. Foissac, S. Djebali, TAGADA: a scalable pipeline to improve genome annotations with RNA-seq data, NAR Genomics Bioinforma. 5 (2023) lqad089, https://doi.org/10.1093/nargab/lqad089.
[81] A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species, Nat. Biotechnol. 36 (2018) 411–420, https://doi.org/10.1038/nbt.4096.
[82] I. Korsunsky, N. Millard, J. Fan, K. Slowikowski, F. Zhang, K. Wei, Y. Baglaenko, M. Brenner, P. Loh, S. Raychaudhuri, Fast, sensitive and accurate integration of single-cell data with harmony, Nat. Methods 16 (2019) 1289–1296, https://doi.org/10.1038/s41592-019-0619-0.
[83] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-Seq data, BMC Bioinformatics 14 (2013) 7, https://doi.org/10.1186/1471-2105-14-7.
[84] U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, J. Vilo, g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update), Nucleic Acids Res. 47 (2019) W191–W198, https://doi.org/10.1093/nar/gkz369.
[85] B. Escofier, J. Pagès, Multiple factor analysis (AFMULT package), Comput. Stat. Data Anal. 18 (1994) 121–140, https://doi.org/10.1016/0167-9473(94)90135-X.
[86] A. Rau, R. Manansala, M.J. Flister, H. Rui, F. Jaffrézic, D. Laloë, P.L. Auer, Individualized multi-omic pathway deviation scores using multiple factor analysis, Biostatistics. 23 (2022) 362–379, https://doi.org/10.1093/biostatistics/kxaa029.
[87] H. Hu, Y.-R. Miao, L.-H. Jia, Q.-Y. Yu, Q. Zhang, A.-Y. Guo, AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors, Nucleic Acids Res. 47 (2019) D33–D38, https://doi.org/10.1093/nar/gky822.

[88] D.M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative genomics, Genome Biol. 20 (2019) 238, https://doi.org/10.1186/s13059-019-1832-y.
[89] R. Janky, A. Verfaillie, H. Imrichová, B. Van de Sande, L. Standaert, V. Christiaens, G. Hulselmans, K. Herten, M. Naval Sanchez, D. Potier, et al., iRegulon: from a gene list to a gene regulatory network using large motif and track collections, PLoS Comput. Biol. 10 (2014) e1003731, https://doi.org/10.1371/journal.pcbi.1003731.
[90] C.C. Flerin, K. Davie, G. Hulselmans, M.D. Waegeneer, vib-singlecell-nf/vsn-pipelines: v0.27.0. (Zenodo), 2021, https://doi.org/10.5281/zenodo.5751297.
[91] S. Suo, Q. Zhu, A. Saadatpour, L. Fei, G. Guo, G.-C. Yuan, Revealing the critical regulators of cell identity in the mouse cell atlas, Cell Rep. 25 (2018) 1436–1445. e3, https://doi.org/10.1016/j.celrep.2018.10.045.
[92] F. Krueger, F. James, P. Ewels, E. Afyounian, M. Weinstein, B. Schuster-Boeckler, G. Hulselmans, Sclamons, FelixKrueger/TrimGalore: v0.6.10 - add Default Decompression Path. Version 0.6.10 (Zenodo), 2023, https://doi.org/10.5281/ZENODO.7598955.
[93] E. Lummertz da Rocha, C. Kubaczka, W.W. Sugden, M.A. Najia, R. Jing, A. Markel, Z.C. LeBlanc, R. dos Santos Peixoto, M. Falchetti, J.J. Collins, et al., CellComm infers cellular crosstalk that drives haematopoietic stem and progenitor cell development, Nat. Cell Biol. 24 (2022) 579–589, https://doi.org/10.1038/s41556-022-00884-1.
[94] D. Dimitrov, D. Türei, C. Boys, J.S. Nagai, R.O. Ramirez Flores, H. Kim, B. Szalai, I.G. Costa, A. Dugourd, A. Valdeolivas, et al., Comparison of Resources and Methods to infer Cell-Cell Communication from Single-cell RNA Data (Bioinformatics), 2021, https://doi.org/10.1101/2021.05.21.445160.
[95] D. Türei, T. Korcsmáros, J. Saez-Rodriguez, OmniPath: guidelines and gateway for literature-curated signaling pathway resources, Nat. Methods 13 (2016) 966–967, https://doi.org/10.1038/nmeth.4077.
[96] S. Tyanova, T. Temu, P. Sinitcyn, A. Carlson, M.Y. Hein, T. Geiger, M. Mann, J. Cox, The Perseus computational platform for comprehensive analysis of (prote) omics data, Nat. Methods 13 (2016) 731–740, https://doi.org/10.1038/nmeth.3901.
[97] Y. Tang, M. Horikoshi, W. Li, Ggfortify: unified Interface to visualize statistical results of popular R packages, R J. 8 (2016) 474, https://doi.org/10.32614/RJ-2016-060.
[98] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504, https://doi.org/10.1101/gr.1239303.
[99] H. Acloque, O.H. Ocaña, D. Abad, C.D. Stern, M.A. Nieto, *Snail2* and *Zeb2* repress *P-cadherin* to define embryonic territories in the chick embryo, Development 144 (2017) 649–656, https://doi.org/10.1242/dev.142562.
[100] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D.J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, et al., The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic Acids Res. 50 (2022) D543–D552, https://doi.org/10.1093/nar/gkab1038.