Single-cell gene expression of the bovine blastocyst

Verónica M Negrón-Pérez¹, Yanping Zhang² and Peter J Hansen¹

¹Department of Animal Sciences, D. H. Barron Reproductive and Perinatal Biology Research Program and Genetics Institute, University of Florida, Gainesville, Florida, USA and ²Gene Expression and Genotyping Core, Interdisciplinary Center for Biotechnology Research, University of Florida, Gainesville, Florida, USA

Correspondence should be addressed to P J Hansen; Email: pjhansen@ufl.edu

Abstract

The first two differentiation events in the embryo result in three cell types – epiblast, trophectoderm (TE) and hypoblast. The purpose here was to identify molecular markers for each cell type in the bovine and evaluate the differences in gene expression among individual cells of each lineage. The cDNA from 67 individual cells of dissociated blastocysts was used to determine transcript abundance for 93 genes implicated as cell lineage markers in other species or potentially involved in developmental processes. Clustering analysis indicated that the cells belonged to two major populations (clades A and B) with two subpopulations of clade A and four of clade B. Use of lineage-specific markers from other species indicated that the two subpopulations of clade A represented epiblast and hypoblast respectively while the four subpopulations of clade B were TE. Among the genes upregulated in epiblast were AJAP1, DNMT3A, FGFR4, H2AFZ, KDM2B, NANOG, POU5F1, SAV1 and SLIT2. Genes overexpressed in hypoblast included ALPL, FGFR2, FNT, GATA6, GJA1, HDAC1, MBNL3, PDGFRα and SOX17, while genes overexpressed in all four TE populations were ACTA2, CDX2, CYP11A1, GATA2, GATA3, IFNT, KRT8, RAC1 and SFN. The subpopulations of TE varied among each other for multiple genes including the prototypical TE marker IFNT. New markers for each cell type in the bovine blastocyst were identified. Results also indicate heterogeneity in gene expression among TE cells. Further studies are needed to confirm whether subpopulations of TE cells represent different stages in the development of a committed TE phenotype.

Reproduction (2017) 154 627–644

Introduction

Preimplantation embryonic development involves a series of cleavage divisions and cell differentiation processes that lead to the formation of the three primary cell types in the embryo: epiblast (precursor of fetus), hypoblast (precursor of extraembryonic endoderm including yolk sac) and trophectoderm (precursor of placenta) (Arnold & Robertson 2009). Upon compaction of the morula, the outer cells become polarized and mechanisms for tight junction formation are triggered (Levy et al. 1986, Nikas et al. 1996, Van Soom et al. 1997). Consequently, the outer cells begin to differentiate into trophectoderm (TE) cells while the cells of the inner cell mass (ICM) maintain pluripotency and cell-to-cell communication through gap junctions (Ducibella & Anderson 1975, Lo & Gilula 1979, Togashi et al. 2015). In the mouse, this first lineage differentiation event is regulated in part by the Hippo signaling pathway, which maintains pluripotency of the cells when turned on (i.e. in the ICM) or favors differentiation of the cells when turned off (i.e. in the TE cells) (Paramasivam et al. 2011, Hirate et al. 2013). The presence of gap junctions in the ICM allows activation of a phosphorylation cascade that activates NF2, MST1 and LATS1/2. The last molecule in this cascade phosphorylates and inactivates YAP1 to favor transcription of pluripotency factors such as Pou5f1 (Lian et al. 2010) and Sox2 (Lorthongpanich et al. 2012). In the absence of gap junctions in the outer cells of the embryo, the phosphorylation cascade is inactive, YAP1 remains dephosphorylated and the protein undergoes nuclear translocation and interaction with TEAD4 to promote transcription of differentiation promoting factors such as Cdx2 and Gata3 (Nishioka et al. 2008, Ralston et al. 2010). An additional player of the Hippo pathway is AMOT, which inactivates YAP1 by interacting with YAP1 and membrane-bound CDH1 to prevent nuclear translocation of YAP1. In the outer cells of the embryo, AMOT is unable to inactivate YAP1 because it is restricted to the apical domain of cells by binding to F-actin (Hirate et al. 2013, Leung & Zernicka-Goetz 2013). Transcription factors involved in later TE differentiation include ELF5 (Ng et al. 2008, Pearton et al. 2011) and EOMES (Russ et al. 2000).

As the blastocyst continues to develop, the ICM differentiates into the epiblast and hypoblast. This second differentiation event in the mouse is accomplished in part through activation of the FGF4–FGFR2 pathway, which regulates transcription of Gata6 and Nanog (Kang et al. 2013). The first group of cells to enter the
ICM express Fgf4 but lack FGFR2 so that expression of Nanog is maintained (Morris et al. 2010, 2013). The cells from the second round of cell division in the ICM begin to express Fgr2 while FGF4 from neighboring cells activates FGFR4–FGFR2 signaling to induce expression of Gata6 and repression of Nanog. Thus, the first group of cells become precursors of the epiblast and the second group are precursors of the hypoblast (Morris et al. 2010, 2013).

The processes for formation and early differentiation of the blastocyst have not been well described in species other than the mouse. Indeed, mechanisms for development during the preimplantation period are not completely conserved between mammalian species. There are large-scale species differences in orthologous gene expression patterns during the preimplantation period caused in large part by mutations in transcription factor-binding sites and insertion of transposons containing cis-regulatory elements (Xie et al. 2010). Among the resultant differences in regulatory processes between species are the use of the transcription factor CDX2 to downregulate the pluripotency factor OCT4 (i.e. Pou5f1) in ICM (Berg et al. 2011) and the role of Ccl24 in movement of hypoblast cells (Negrón-Pérez et al. 2017).

Understanding the mechanisms for differentiation and function of cell lineages in the bovine embryo has been hampered by the limited repertoire of markers for each cell type. The only definitive marker identified for epiblast is NANOG (Kuijk et al. 2012, Denicol et al. 2014). One marker for ICM in mouse, Pou5f1, is expressed by both ICM and TE in bovine (Berg et al. 2011, Ozawa et al. 2012). GATA6 can serve as a marker for hypoblast in bovine (Kuijk et al. 2012, Denicol et al. 2014), but interpretation of differences in gene expression for this transcription factor is complicated by the fact that expression can also occur for TE (Kuijk et al. 2012, Denicol et al. 2014). The most commonly used marker for TE in bovine is CDX2 (Dobbs et al. 2013, Schifflmacher & Keeter 2013, Denicol et al. 2014), but the observation that outer cells of the blastocyst only gradually become committed to the TE lineage (Berg et al. 2011) is suggestive that there is heterogeneity within this cell population that could be elucidated if additional markers were available.

Here, we took advantage of recently developed microfluidics-based technology for isolation of RNA from individual cells to identify markers of epiblast, hypoblast and TE in the bovine blastocyst. The approach was to assess the expression of 93 genes including those that are used as markers for epiblast, hypoblast and TE in other species (Chazaud et al. 2006, Guo et al. 2010, Yan et al. 2013, Hermitte & Chazaud 2014, Blakeley et al. 2015, Boroviak et al. 2015), genes differentially expressed between ICM and TE in the bovine (Ozawa et al. 2012, Nagatomo et al. 2013, Brinkhof et al. 2015, Hosseini et al. 2015, Zhao et al. 2016) and genes involved in important processes for the developing embryo including the Hippo signaling pathway, epigenetic regulation, tight junction formation, cell polarity and chemokine signaling. The result was identification of markers for each cell lineage of interest, evidence that the TE is a heterogeneous tissue in the developing blastocyst and identification of markers that distinguish between subtypes of TE.

Materials and methods

In vitro production of embryos

Production of embryos involved use of culture media described elsewhere (Ortega et al. 2017). Oocytes and sperm were from cattle representing B. taurus and admixtures of B. taurus and B. indicus. Oocytes were obtained from ovaries collected at an abattoir by bisecting follicles 3–8 mm in diameter using a scalpel and, after follicles on an ovary were bisected, vigorously washing the ovary in oocyte collection medium (BoviPRO, MOFA Global, Verona, WI, USA) to dislodge cumulus-oocyte complexes (COC). Fluid containing the COC was then filtered and COC were rinsed with fresh collection medium. The COC were then retrieved using a wiretrol pipette (Drummond Scientific Company, Broomall, PA, USA) while visualizing under a dissecting microscope. The COCs were washed another three times in fresh medium. An average of 12–15 COCs (having at least one layer of cumulus cells and containing homogeneous cytoplasm) were collected from each ovary. The COCs were pooled in groups of 10 and matured for 22–24 h in oocyte maturation medium (Tissue Culture Medium-199 with Earle’s salts supplemented with 2% (v/v) bovine steer serum, 100 U/mL penicillin-G, 0.1 mg/mL streptomycin, and 1 mM glutamine) in 50 µL microdrops covered with mineral oil (Sigma-Aldrich) and in an atmosphere of 5% (v/v) CO2 in humidified air at 38.5°C.

Matured COC were pooled, washed three times in HEPES-buffereed synthetic oviduct fluid (HEPES-SOF) and placed in 1.7 mL synthetic oviduct fluid for fertilization (SOF-FERT) in groups of up to 300 COCs. For fertilization, frozen-thawed semen from three bulls was pooled, purified by centrifugation, at 1000 g for 10 min, over a bilayer of Isolate (Irvine Scientific, Santa Ana, CA, USA; 50% (vol/vol) and 90% (vol/vol) Isolate), washed in HEPES-SOF and diluted with SOF-FERT. Then, 120 µL of semen plus 80 µL of 0.5 mM penicillamine, 0.25 mM hypotaurine and 25 µM epinephrine solution were added to the COC for a final sperm concentration of 1 × 106 sperm/mL. Fertilization dishes were incubated in an atmosphere of 5% (v/v) CO2 in humidified air at 38.5°C for 16–18 h.

Presumptive zygotes were collected and exposed to hyaluronidase (1000 U/mL in approximately 0.5 mL HEPES-SOF) to remove cumulus cells and washed three times in HEPES-SOF prior to culture. Embryos were pooled in groups of 25–30 and cultured at 38.5°C in 50 µL microdrops of BBH7 (Cooley Biotech, Gainesville, Florida, USA) covered with mineral oil in a humidified environment consisting of 5% (v/v) O2, 5% (v/v) CO2 and the balance nitrogen.

The proportion of embryos that was cleaved was assessed at day 3 after fertilization and the proportion that became
blastocysts was assessed at day 7. At day 8.75 (207–209 h post insemination), all blastocysts (including non-expanded, expanded, hatching and hatched blastocysts) were collected and subjected to blastomere dissociation. Embryos were collected in a total of three replicates. A replicate was defined as a single in vitro fertilization procedure consisting of insemination of 900–1200 COC. A total of eight bulls were used in the three replicates. The cleavage rate for the three replicates averaged 80% and the percent of inseminated oocytes becoming blastocysts averaged 20% on day 7 and 29.8% on day 8.75.

Preparation of cDNA from single blastomeres

cDNA was prepared from individual blastomeres using the C1 Single-Cell Auto Prep IFC (integrated fluidic circuit) system from Fluidigm (South San Francisco, CA, USA) using manufacturer instructions.

For each replicate, single-cell suspensions were prepared from the blastocysts collected at 207–209 h post insemination. The number of blastocysts processed for each replicate ranged from 227 to 336. Blastocysts were washed three times in Dulbecco’s phosphate-buffered saline (DPBS) containing 0.1% (w/w) polyvinylpyrrolidone (DPBS-PVP; Kodak, Rochester, NY, USA), incubated in 0.1% (w/v) protease from Streptococcus griseus (Sigma-Aldrich) in DPBS until the zonae dissolved, and then washed another three times in fresh DPBS-PVP. Embryos were then incubated in 50 µL drop of TrypLE Select Enzyme 10x (ThermoFisher Scientific) for 15 min at 38.5°C to disaggregate the cells. Finally, blastomeres were transferred to a 1.7 mL tube, vortexed for 2 min, resuspended in 500 µL DPBS-PVP and centrifuged for 5 min at 6000 × g. The supernatant was removed and the cells were resuspended in 10 µL DPBS-PVP.

Following blastocyst disaggregation, cell viability and concentration were measured using the Countess Automated Cell Counter (Life Technologies) and the concentration was adjusted to 300–400 cells/µL. Viability in each replicate ranged from 33% to 82%. A subset of the cells was subjected to a buoyancy test as part of the quality control process for the C1 system to optimize probability of cell capture probability. The cells were mixed with C1 Cell Suspension Reagent at a ratio of 7:1, 7:2 and 7:3 (v:v) and placed in a well of the C1 IFC plate. The buoyancy check was performed under 10x magnification immediately and after 10-min waiting time. Buoyancy was determined by looking at three planes of the well (top, middle and bottom) under the microscope and noting where the majority of the cells were at time 0. This was repeated after the 10-min incubation period. The goal was to have the majority of cells suspended in the middle plane for better capture. The best ratio of suspension reagent to cell suspension was found to be 7:2; at this ratio, 78% of cells were in the middle plane.

The C1 Single-Cell Auto Prep IFC for Preamp (10–17 µm) was primed following the manufacturer’s instructions. Because the cells were 13 µm in diameter on average, the plate and protocol used for priming was C1 DNA Seq IFC: Prime (1773x). The cell suspension was loaded into the C1 plate for individual cell capture, lysis for RNA extraction and cDNA pre-amplification. A viability cell staining was performed for one of the replicates using the LIVE/DEAD Viability/Cytotoxicity kit from Fluidigm to assess the proportion of live and dead cells that had been captured. In this replicate, 54 cells were captured including 11 dead cells and 43 live cells. Only the cDNA from the 43 viable cells was used for further analysis (one of these 43 cells failed to yield cDNA). Overall, the success of the system for capturing individual cells was low. The number of cells yielding cDNA collected for each replicate was 6, 19 and 42 for a total of 67 cells analyzed. The C1 plates containing single-cell cDNA were stored at -20°C until gene expression analysis.

Gene expression analysis

The Fluidigm qPCR microfluidic device Biomark HD system was used for gene expression assays. Primers were designed by Fluidigm Delta Gene assays (Fluidigm Co., San Francisco, CA, USA) and validated by Miami Center for AIDS Research (CFAR) at the University of Miami Miller School of Medicine. A description of the genes analyzed, the primers used and the rationale for including in the analysis are in Supplementary Data 1 (see section on Supplementary Data given at the end of this article). Primers were designed for 2 housekeeping genes (ACTB and GAPDH), 9 epiblast markers from one or more of 3 species; 9 hypoblast markers; 11 trophectoderm markers, 16 genes involved in chemokine signaling, 9 genes involved in Hippo signaling, 13 genes involved in epigenetic modification, 11 genes involved in tight junction formation and cell polarity, and, another 16 genes of interest. Initially, three genes, SDHA, H2AFZ and HPRT1, were also selected as housekeeping genes but they were not used as such because of non-uniform expression.

The procedure for quantitative PCR (qPCR) using the Biomark system (Dominguez et al. 2013) was as follows. Primer-probe sets and samples were loaded on an IFC plate and placed into a controller that prepares the nanolitre reactions. The plate was then transferred into the Biomark machine, which includes a thermocycler for real-time qPCR. A total of 40 PCR cycles were performed using the 96.96 dynamic array IFC developed by the manufacturer. Cycle threshold (Ct) values were calculated by the Fluidigm Real-Time PCR analysis software. The cutoff for detectable genes was those with Ct >27. The geometric mean of the two housekeeping genes was calculated and used to obtain the delta Ct (dCt) values of the other 93 genes of interest. Fold changes were calculated as 2-dCt relative to housekeeping genes.

The parameters for the validation of primers for qPCR using the Biomark HD system were as follows. Validation was achieved using two pools of 40 embryos each (day 5 morulae and day 7 and day 8 blastocysts), endometrium, fetal liver and leukocytes. RNA was extracted from all samples using the RNeasy micro kit (Qiagen); a DNase treatment was included as part of the protocol. Reverse transcription was performed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) following manufacturer’s instructions. The cDNA was pre-amplified following the guidelines for the Ambion Single Cell-to-CTTM kit (ThermoFisher Scientific) and diluted serially in 2-fold dilutions to a single-cell equivalent. Primer-probe sets and samples were loaded on an IFC plate and placed into a controller that prepares the nanolitre reactions. The plate was then transferred into the Biomark machine for real-time qPCR. Standard curves and cycle threshold
Principal component analysis also resulted in clustering of cells into two major populations, with one population corresponding to clade A of the unsupervised clustering analysis and the other population corresponding to clade B (Fig. 2). Principal component analysis also indicated discrimination between subclades identified by the clustering analysis.

**Identification of cell subpopulations as epiblast, hypoblast and TE**

Expression patterns of genes known to be markers for specific cell lineages in the bovine or other species were used to identify each subclade of cells. Expression patterns of nine genes considered as markers of epiblast in bovine (Kuijk et al. 2012, Denicol et al. 2014), mouse (Guo et al. 2010, Boroviak et al. 2015), or human (Yan et al. 2013, Blakeley et al. 2015) are shown in Fig. 3. Three of these genes, FGF4, NANO2 and POUF51, were more highly expressed in subclade A2 than other subclades. In addition, another marker, HNF4A, was more highly expressed in subclades A1 and A2 than for cells in clade B; FN1 was upregulated in clade A1 as compared to the other five subclades. Based on these patterns of gene expression, subclade A2 was considered to represent epiblast. Note that several markers of epiblast in the mouse (Guo et al. 2010, Boroviak et al. 2015) were not preferentially expressed in subclade A2 including BMP4, ESRRB and SOX2.

Expression of nine genes that are characteristic of hypoblast in either bovine (Kuijk et al. 2012, Denicol et al. 2014), mouse (Chazaud et al. 2006, Guo et al. 2010, Boroviak et al. 2015) or human (Yan et al. 2013, Blakeley et al. 2015) are shown in Fig. 4. Four of these genes (FGFR2, GATA6, PDGFRA and SOX17) were more expressed in subclade A1 than other populations. Accordingly, subclade A1 was labeled as hypoblast. Note that this subclade, like the epiblast subclade, also expresses high amounts of HNF4A (Fig. 3) and RUNX1, which is a hypoblast marker in mouse (Guo et al. 2010). Other markers of hypoblast in mouse, including CREB3L2, GPR2, SNAI1 and TCF23 (Chazaud et al. 2006, Guo et al. 2010) were not overexpressed in the hypoblast clade.

Results for expression of 11 genes considered as TE markers in bovine (Ozawa et al. 2012, Nagamoto et al. 2013, Brinkhof et al. 2015, Hosseini et al. 2015, Zhao et al. 2016), mouse (Guo et al. 2010, Boroviak et al. 2015), or human (Yan et al. 2013, Blakeley et al. 2015) are shown in Fig. 5. The genes CDX2, GATA3, IFNT and KRT8 exhibited, or tended to exhibit, higher expression in the subpopulations of subclade B than for cells of subclade A. Accordingly, clade B was considered to represent TE and subclades were renamed as follows: B1 = TE1, B2.1 = TE2; B2.2 = TE3 and B2.3 = TE4. Expression of GATA3 tended to be higher for TE1 and TE2 than TE3 and TE4, IFNT was highest in TE3 and
Blastomere-specific gene expression

lowest in TE1, and KRT8 was lowest for TE4. Also, the mouse TE marker, EOMES, was more highly expressed in TE4 than in TE1 and tended to be more expressed than in other TE subpopulations or in epiblast or hypoblast. There was no difference in gene expression between subpopulations for DAB2, MBNL3, PECAM1, TEAD4 or

Figure 1 Identification of embryonic cell populations by unsupervised clustering analysis of gene expression. Heatmaps were generated for expression levels of 90 genes from 67 individual blastomeres after normalizing to housekeeping genes, ACTB and GAPDH. The output yielded two major clades (A and B) that were subdivided into six smaller clusters (highlighted with orange lines): A1, A2, B1, B2.1, B2.2 and B2.3. The numbers in parentheses are the number of cells in each group. The color key is at the bottom right (−3 = bright green, 0 = black, 3 = bright red).
TFAP2C. Also, the mouse TE marker, MBNL3, was most highly expressed in hypoblast (significantly different from epiblast and TE2 and TE4).

Note that expression of GATA6, which is a marker of hypoblast in bovine, but which also shows moderate immunolabeling in TE and low immunolabeling in epiblast (Kuijk et al. 2012, Denicol et al. 2014), exhibited a similar pattern of transcript abundance – high in hypoblast, low in epiblast and moderate in TE populations (Fig. 4).

Other genes overexpressed in epiblast and hypoblast

In addition to genes used to define epiblast (Fig. 3) and hypoblast (Fig. 4), there were six additional genes that were most highly expressed in epiblast (AJAP1, DNMT3A, H2AFZ, KDM2B, SAV1 and SLIT2), two genes most highly expressed in hypoblast (ALPL and GJA1) and three genes (CDH2, HDAC1 and HDAC8) that were more expressed in both epiblast and hypoblast than the TE populations (Fig. 6). Of these latter three genes, CDH2 was expressed equally in epiblast and hypoblast, HDAC1 was more expressed in hypoblast than epiblast, and HDAC8 was more expressed in epiblast than hypoblast (Fig. 6). One gene, ID2, was expressed more in hypoblast than epiblast, with amounts in TE being intermediate (Fig. 6).

Other genes overexpressed in all subpopulations of TE

Besides the genes used to define TE populations, there were an additional five genes that were expressed in all four TE subpopulations (Fig. 7). Of these, expression of GATA2, RAC1 and SFN was relatively uniform between TE subpopulations, AMOT was higher for TE1 than TE4 (with amounts for TE2 and TE3 being intermediate) and expression of CYP11A1 tended to be higher for TE3 than other TE populations.

Figure 2 Principal component analysis of 67 cells subjected to gene expression analysis. Note that cells labeled as Hypo and Epi correspond to cells in Clade 1 from the unsupervised clustering analysis while cells labeled as TE1, TE2, TE3 and TE4 correspond to Clade 2.

Figure 3 Expression of genes considered as prototypical epiblast markers. Data were normalized to the geometric mean of ACTB and GAPDH and are presented as least-squares means ± S.E.M. The number of cells within each subgroup were as follows: epiblast (subclade A2) n = 4, hypoblast (subclade A1) n = 7, TE1 (subclade B1) n = 13, TE2 (subclade B2.1) n = 8, TE3 (subclade B2.2) n = 19 and TE4 (Subclade B2.3) n = 16. Bars labeled with different letters are different from each other (P < 0.05). Species symbols are used to demote upregulation of the gene in epiblast of mouse (Guo et al. 2010, Boroviak et al. 2015), human (Yan et al. 2013, Blakeley et al. 2015) or bovine (Denicol et al. 2014).
There were several genes that were differentially expressed in one or more TE subpopulations as compared to other cell subpopulations. Cells in TE1 expressed higher amounts of transcripts for \textit{CCL26}, \textit{LATS1} and \textit{LATS2} than for all or some subpopulations of cells (Fig. 8). Expression of \textit{LATS1} was lowest for TE3 and TE4. Expression of \textit{TET2} was also higher than for other subpopulations except TE2 (Fig. 8). Cells in TE2 expressed higher amounts of transcript for \textit{ACKR4}, \textit{CCL11}, \textit{CCR2}, \textit{CCR3}, \textit{CRB2}, \textit{DNMT1}, \textit{EPHA4}, \textit{HPRT1}, \textit{ID1}, \textit{MST1}, \textit{PPBP}, \textit{TAZ} and \textit{TJAP1} than other TE subpopulations and, except for \textit{HPRT1} (high expression also in hypoblast), higher than epiblast (Fig. 8). Expression of \textit{INADL} was higher for TE2 than for all cell subpopulations except TE1, \textit{KAT6B} was higher for TE2 than TE4, \textit{STAT3} was higher for TE2 than all populations except hypoblast and \textit{STAT4} (Fig. 8) was higher for TE2 than all populations except epiblast (Fig. 8). \textit{SOX2} was highly expressed in epiblast, hypoblast, TE2 and TE4 but lowly expressed in TE1 and TE3 (Fig. 3).

As compared to other cell populations, expression of \textit{ACTA2} was higher for TE3 than other TE populations and cells in TE4 expressed higher amounts of ASGR1, CCL24, \textit{CCR5}, CDC42EP4 (difference with epiblast not significant), \textit{IL4}, \textit{ITK}, \textit{PLCB1} and \textit{SOX30} than other cells; cells in TE3 expressed lower amounts of transcripts for \textit{EZH2}, cells in TE3 and TE4 expressed lower amounts of \textit{LATS1} than TE1, cells in TE3 and TE4 expressed lower amounts of \textit{SDHA} than for other populations and cells in TE4 expressed the lowest amount of \textit{HSD3B1} (Fig. 9). Moreover, expression of \textit{TET1} was highest in TE2 and lowest in TE3 (Fig. 9).

There were 14 genes in which there was no significant difference in expression between cell populations (Fig. 10).

In the present experiment, analysis of gene expression at the single-cell level has revealed several heretofore unknown characteristics of the bovine blastocyst. This accomplishment was achieved despite the inefficiency of the procedures for isolation of individual cells using the C1 Single-Cell Auto Prep IFC. Only 67 usable cells were captured in three replicates although the theoretical number of cells that could be captured was 288. The reasons for the low capture rate are not known. Contributing factors could have included cell aggregation or suboptimal buoyancy conditions for cell capture.
Clearly, more work is required to optimize the use of the C1 technology for isolation of individual cells from the bovine blastocyst. Alternative approaches for isolation of individual cells could include micromanipulation to pluck individual cells from disaggregated embryos, as has been performed for the mouse and human embryo (Guo et al. 2010, Yan et al. 2013, Blakely et al. 2015).

One consequence of the low number of cell numbers analyzed is that there could be some differences in gene expression between populations that did not approach statistical significance because of the small degree of replication. It is also possible that the selection of cells was not random but that cells of certain characteristics were more likely to be captured than others. Nonetheless, the experiment was successful in the identification of molecular markers for epiblast, hypoblast and TE lineages in the bovine blastocyst, including some markers that allow discrimination of 4 subpopulations of TE. These are summarized in Fig. 11.

It is proposed that mRNA for these genes, or the protein products encoded by those mRNA, can be used to identify and study the function of specific lineages in the bovine blastocyst. Some of the molecular markers for epiblast, hypoblast and TE were expressed almost exclusively by one cell lineage. Examples include AJAP1, FGF4 and NANOG in epiblast, FGFR2, FN1 and SOX17 in hypoblast and CDX2, GATA2, GATA3, IFNT, KRT8 and SFN in TE. Others, while more expressed in one cell subpopulation than others, were expressed to variable degrees by more than one cell subpopulation. Examples include HDAC1 and POU5F1, which were highest in epiblast but which were also expressed in hypoblast and TE cells, GATA6, in which transcript abundance was highest for hypoblast but which was also expressed in TE cell subpopulations, and SLIT2, most expressed in epiblast but also moderately expressed in the TE4 subpopulation. Gene markers identified in Fig. 11 that allow discrimination between
the four subpopulations of TE also varied in the degree to which expression was limited to one subpopulation. Thus, some genes, like ACKR4, CCL11, CCL26, CCR2, CCR3, CRB2, DNMT1, EOMES, EPHA4, ID1, MST1, PPBP, STAT3, STAT4, TAZ and TJAP1, were much more expressed for one TE subpopulation than for other

Figure 6 Additional genes that were differentially expressed in epiblast and hypoblast. Refer to Fig. 3 for additional information.

Figure 7 Additional genes upregulated in trophectoderm that were expressed in all four TE subpopulations. Refer to Fig. 3 for additional information.
Figure 8 Genes that were differentially expressed in TE1 and TE2 cell populations. Refer to Fig. 3 for additional information.
Blastomere-specific gene expression

subpopulations, whereas others, like AMOT, CYP11A1, LATS1 and TET2, while being more expressed in one or more cell subpopulations than others, were expressed more uniformly among the subpopulations.

Many of the epiblast, hypoblast and TE markers identified in Fig. 11 have been previously identified as cell lineage markers. Examples include Cdx2, Fgf4, Fgfr2, Gata3, Gata6, Hnf4a, Krt8, Nanog, Pdgfra, Pou5f1, Runx1 and Sox17 for mouse (Guo et al. 2010, Boroviak et al. 2015), HNF4A, NANOG, PDGFRA, POU5F1 and SOX17 for human (Yan et al. 2013, Blakeley et al. 2015) and CDX2, GATA3, GATA6, IFNT, KRT8 and NANOG for bovine (Kuijk et al. 2012, Ozawa et al. 2012, Nagatomo et al. 2013, Denicol et al. 2014, Brinkhof et al. 2015, Hosseini et al. 2015, Zhao et al. 2016). Others have not been previously described as cell lineage markers, for example, GATA2, MST1, SFN and TJAP1 for TE.

Several of the cell lineage markers in the mouse or human embryo displayed an expression pattern in the bovine embryo distinct from the expected pattern. Examples include BMP4, CREB3L2, ESRRB, FNT, GRB2, MBNL3, SNAI1, TCF23 and TEAD4. Also, some TE markers in mice were only expressed in a subset of TE cells, most notably ELF5 and EOMES. Differences between the bovine and other species is not surprising because evolutionary analysis indicates that there is a large degree of heterogeneity in the control of gene expression in the early embryo among species (Xie et al. 2010, Berg et al. 2011, Niakan & Eggan 2013, Hosseini et al. 2015).

It remains to be determined whether differences in protein expression recapitulate differences in transcript abundance and whether markers that are useful for distinguishing between cell lineages at the mRNA level will also prove useful at the protein level. It is notable,
however, that for the proteins previously examined, there is concordance with results on mRNA determined here. CDX2 is a widely used protein marker for TE (Ross et al. 2009, Kuijk et al. 2012, Schiffmacher & Keefer 2013, Denicol et al. 2014) and was found to be highly expressed in all four TE populations in the present study. In addition, IFNT, whose transcript was more expressed in TE than epiblast and hypoblast, was immunolocalized to TE but not ICM (Johnson et al. 2006). Similarly, like for present findings regarding gene expression, GATA6 can be localized to both hypoblast and TE, with amounts of immunoreactive protein greater in hypoblast than TE (Kuijk et al. 2012, Denicol et al. 2014). The present finding that POU5F1 was most highly expressed in the epiblast but was also detected in populations of hypoblast and TE is consistent with earlier results that immunoreactive POU5F1 was present in ICM and TE, with more intense labeling in the former (Berg et al. 2011).

There was one case, however, where there is a lack of concordance between gene expression and protein accumulation. Immunoreactive YAP1 was greater in TE compared to ICM in day 7 bovine blastocysts (Tribulo et al. 2017) even though YAP1 mRNA was not different between cell lineages in the current study. Indeed, at least in the mouse, YAP1 is regulated post-translationally through control of phosphorylation and degradation (Basu et al. 2003, Verghese et al. 2012). It is also possible that differences in gene expression between cell lineages, seen here with blastocysts at day

Figure 10 Transcripts whose expression did not differ significantly among cell populations. Refer to Fig. 3 for additional information.
8.75 of development, may be altered at other stages of development. One example is CCL24, whose expression in the blastocyst peaks at about Day 7 of development (Negrón-Pérez et al. 2017). At Days 7–8, CCL24 is more expressed in ICM than TE (Ozawa et al. 2012, Brinkhof et al. 2015, Hosseini et al. 2015, Zhao et al. 2016). In the current experiment, however, CCL24 expression was largely limited to one TE subpopulation (TE4), suggesting a developmental switch in the site of expression of this chemokine gene.

Patterns of gene expression revealed in this study provide some insight into the processes leading to the formation and subsequent development of each cell lineage. As expected, the epiblast, which remains in a pluripotent state after formation of the TE and hypoblast (Nichols et al. 1998, Kirchhof et al. 2000, Avilion et al. 2003, Silva et al. 2009), exhibited upregulation of the pluripotency genes NANOG and POU5F1. The epiblast also overexpressed FGF4, which has been associated with high levels of SOX2 expression in the epiblast of mouse species (Yuan et al. 1995). Expression of two genes involved in epigenetic remodeling in the epiblast, DNMT3A and KDM2B, are also consistent with maintenance of epiblast cells. Studies using embryonic stem cells deficient for DNMT1/DNMT3A/DNMT3B indicated that methyltransferases are important for allowing embryonic stem cells to contribute to embryonic lineages at day E10.5 but is not required for formation of extraembryonic tissues (Sakaue et al. 2010). KDM2B has also been implicated in maintenance of pluripotency by driving noncanonical Polycomb-repressive complex 1 (ncPRC1) to CpG islands (Morey et al. 2015, Schoenfelder et al. 2015) and subsequent repression of differentiation promoting genes (Smith et al. 2016). Another gene upregulated in epiblast, H2AFZ, encodes for a variant of the histone octamer member H2A. The specific function of H2AFZ during early embryonic development is unknown but H2afz deficient mouse embryos die by E5.5 and have an abnormal ICM (Faast et al. 2001). A histone deacetylase, HDAC8, was upregulated in both epiblast and hypoblast, in agreement with a previous finding that the gene was upregulated in the ICM of the bovine embryo (Hosseini et al. 2015). The significance of this finding is unclear since HDAC8 is involved in silencing pluripotency genes (Saha et al. 2013). Another two genes upregulated in the epiblast may be involved in the Hippo signaling pathway that controls formation of the ICM. AJAP1 encodes for an adherens junction-associated protein. During formation of the ICM and TE, adherens junctions play a crucial role in promoting ICM formation by interacting with CDH1 and AMOT (Sasaki 2015). Another gene upregulated in epiblast, SAV1, is an upstream member of the phosphorylation cascade of the Hippo signaling pathway (Lee et al. 2008).

Differentiation of the hypoblast in mice involves actions of FGF4 from neighboring cells on hypoblast precursors (Morris et al. 2010, 2013). The situation with

---

**Figure 11** Model of genes whose expression differs between cell populations in the bovine blastocyst. Genes listed under Epiblast, Hypoblast, Epi & hypo and Trophoderm were upregulated in that cell type as compared to other cell populations: For the four populations of trophoderm cells, two types of genes are listed: those that were upregulated relative to one or more other populations of trophoderm cells and those that were downregulated.
respect to the bovine is unclear. Treatment of bovine embryos with FGF4 and heparin can lead to blastocysts where the ICM is composed entirely of hypoblast cells (Kuijk et al. 2012). In this same paper, inhibition of FGF4 signaling caused by addition of sodium chlorate caused no effect on number of hypoblast cells (Kuijk et al. 2012). Addition of FGF2 to bovine blastocysts increased the outgrowth of hypoblast cells expressing GATA4 and GATA6 (Yang et al. 2011). Present results evaluating gene expression are consistent with a role of FGF4 in the formation of the hypoblast. In particular, the epiblast subpopulation preferentially expressed FGF4 while cells of the hypoblast subpopulation preferentially expressed FGFR2.

One downstream gene for FGF signaling in human pluripotent stem cells is HNF4A (Twaroski et al. 2015), which is also a hypoblast marker in the mouse (Hermitte & Chazaud 2014) and human (Yan et al. 2013). This gene was overexpressed in the epiblast and hypoblast subpopulations and mRNA was localized to ICM, particularly those lining the blastocoele, in the bovine blastocyst (Nagatomo et al. 2013).

Another gene that is regulated by FGF4 in mouse embryo is the transcription factor Sox17 (Frankenberg et al. 2011), which in turn can act as an enhancer for Fn1 (Shirai et al. 2005). Both of these hypoblast markers were found to be upregulated in the hypoblast subpopulation in the present experiment. Fn1 encodes fibronectin 1, a component of extracellular matrix. In the bovine, FN1 is primarily localized to ICM although a small fraction of immunoreactive FN1 is associated with what could be hypoblast (Goossens et al. 2009).

Another growth factor important for hypoblast survival and expansion in the mouse is PDGF (Artus et al. 2010, 2013). The receptor for this growth factor, PDGFRα, was overexpressed in the hypoblast subpopulation. Hypoblast cells also overexpressed two genes involved in gap junction formation – GJA1, and CDH2, which was also highly expressed in epiblast. Gap junctions play an important role in Hippo signaling in the ICM by regulating localization of CDH1 which is a direct regulator of the Hippo signaling (Sasaki 2015). Perhaps these specialized junctional complexes are important for differentiation or maintenance of the hypoblast.

Although hypoblast cells represent a more differentiated cell type than epiblast, gene expression patterns suggest that this cell line retains some degree of an undifferentiated estate, at least at this stage of development. Thus, hypoblast cells overexpressed ALPL, which is highly expressed in embryonic stem cells in mouse (Tielens et al. 2006) and pluripotent stem cells derived from cloned bovine embryos (Wang et al. 2005). The hypoblast is also apparently undergoing epigenetic regulation as indicated by overexpression of HDAC1 and HDAC8. Hdac1 is a histone deacetylase that is embryonic lethal (Montgomery et al. 2007) and is involved in recruiting silencing complexes to maintain pluripotency of mouse embryonic stem cell (Dovey et al. 2010). In addition, Hdac1 cooperates with Brg1 to silence Nanog in the TE (Carey et al. 2015). Perhaps the same phenomenon occurs in hypoblast. Inhibition of Nanog expression is sufficient to induce hypoblast formation in mouse embryos (Frankenberg et al. 2011).

In the current experiment, a large number of genes involved in chemokine signaling were selected for analysis because of a recent finding that suggests CCL24 participates in positioning of hypoblast precursors to the outside of the ICM (Negrón-Pérez et al. 2017). This chemokine gene reaches maximal expression at Day 7 of development (Negrón-Pérez et al. 2017) and is more highly expressed in the ICM than in the TE at Day 7–9 (Ozawa et al. 2012, Brinkhof et al. 2015, Hosseini et al. 2015, Zhao et al. 2016). Moreover, inhibition of CCL24 signaling by addition of receptor antagonists or knockdown of CCL24 mRNA reduces the proportion of GATA6+ cells in the outer portion of the ICM (Negrón-Pérez et al. 2017). Based on these results, it was expected to observe higher expression of CCL24 in epiblast and of various chemokine receptor genes in hypoblast. In fact, however, CCL24 was overexpressed in one population of TE cells and there was no evidence for overexpression of chemokine receptors in hypoblast. One possibility is that the apparent expression of CCL24 in ICM reflects contamination with TE. This could be possible because, as discussed below, the TE4 population may represent polar TE that is adjacent to ICM. However, immunolocalization of CCL24 indicated that the protein was most commonly localized to the ICM (Negrón-Pérez et al. 2017). It may be more likely, therefore, that the site of CCL24 expression changes as the embryo advances in development so that by day 8.75 of development, the time examined here, CCL24 expression by cells of the ICM becomes inhibited and the TE begins expressing CCL24.

CDX2 is the transcription factor upregulated in TE as compared to epiblast or hypoblast that has been shown to be necessary for development of the trophoblast in the bovine. Zygotic deletion of CDX2 does not prevent blastocyst formation (Berg et al. 2011) but is important for upregulation of GATA3 (Sakurai et al. 2016), and maintenance of tight junctions in the TE (Goissis & Cibelli 2014). Transgenic CDX2-deficient embryos were unable to undergo trophoblast elongation at Day 14 of gestation after transfer to females but expression of IFNT at this stage was not blocked (Berg et al. 2011). A recent paper from the mouse indicates that maternally-derived CDX2 is important for blastocyst formation (Jedrusik et al. 2015) and it remains to be determined whether a similar phenomenon occurs in the bovine.

In the mouse, transcription of Cdx2 depends upon formation of a complex of YAP1 and TEAD4 after inactivation of Hippo signaling in the outer cells of the embryo (reviewed by Lorthongpanich & Issaragrisil 2015). It is not known whether a similar signaling system is involved in the bovine. Expression of TEAD4
and YAP1 were not different between cell types of the bovine blastocyst and other components of Hippo signaling were increased in epiblast (SAV1), all four TE subpopulations (SFN), TE1 (AMOT, LAT51, or LAT52), or TE2 (MST1 or TAZ). Bovine blastocysts can form, appear to show no differences of ICM and TE and have normal CDX2, GATA6, IFNT and POU5F1 levels in the presence of reduced amounts of mRNA for TEAD4 (Sakurai et al. 2016).

Use of single-cell analysis of blastocyst gene expression revealed that the cells of the TE are heterogeneous. There are at least four subpopulations of TE cells at day 8.75 of development with TE1 forming one subclade and TE2, TE3 and TE4 forming a second subclade. The existence of so many populations of TE could reflect variation among individual blastocysts or the presence of more than one population of cells in an individual blastocyst. Blastocysts used for the experiment included a mixture of genetic types and a variety of stages (i.e. early, expanding, hatching and hatched blastocysts) and some of the heterogeneity in TE populations could reflect variation among blastocysts in genetic background or stage of development. Moreover, there are differences between male and female blastocysts in gene expression (Bermejo-Alvarez et al. 2010) and such differences could lead to differences in gene expression among TE cells. However, it is also likely that the TE is not functionally uniform within an individual blastocyst. Indeed, there is evidence that gene expression differs between mural and polar TE cells in the bovine (Nagatomo et al. 2013) and immunolocalization of IFNT in the TE is not uniform (Johnson et al. 2006).

Additional studies will be required to identify temporal, spatial and functional differences between the TE subpopulations. However, there are indications that TE1, which forms a separate subclade from TE2, TE3 and TE4, may represent immature TE cells. This population of cells had low expression of the maternal recognition of pregnancy signal IFNT as well as the lowest amounts of the transcription factor EOMES. Secretion of IFNT was higher for expanded blastocysts than for non-expanded blastocysts (Kubisch et al. 2001) and increased as blastocysts were cultured for 24 or 48 h (Kubisch et al. 2004). Eomes is a late TE marker in the mouse blastocyst and its transcription depends on actions of TEAD4 and CDX2 (Probst & Arnold 2017). TE1 cells also exhibited elevated expression of three molecules involved in Hippo signaling – AMOT, LAT51 and LAT52. Moreover, TE1 cells expressed high amounts of TET2 which is involved in inhibition of differentiation of mouse embryonic stem cells to trophoblast (Koh et al. 2011). The fact that TE1 cells highly express CCL26, a marker of polar TE in the bovine (Nagatomo et al. 2013), could mean that these cells are preferentially localized in this region of the TE. While trophoblast stem cells (TSC) are also located in polar TE, at least in the mouse (Probst & Arnold 2017), it is unlikely that TE1 cells represent TSC because of the low expression of EOMES, which is highly expressed in TSC of the mouse (Probst & Arnold 2017).

A better candidate for TSC are cells in the TE4 subpopulation. These cells highly express EOMES and two markers of polar TE (Nagatomo et al. 2013) – ASGR1 and (non-significantly) TMEM232. TE4 cells also express high amounts of SOX2, which has been implicated in maintenance of TSC in mouse (Adachi et al. 2013) as well as the transcription factor SOX30, a member of the SRY family whose function is not well understood.

Perhaps the most differentiated of the TE subpopulations is the TE2 subclade. This subpopulation had the highest number of genes that were upregulated in expression compared to other TE subpopulations. Among the 18 genes upregulated in TE2 are three implicated in placental function in other species: the Hippo signaling molecule MST1, the transcriptional regulator ID1 and the ephrin receptor EPHA4. Mst1 is important for placentation differentiation in the mouse (Du et al. 2014), ID2, which is closely related to ID1, participates in regulation of basic helix-loop-helix transcription factors in human placenta (Liu et al. 2004), and interactions between ephrin A and ephrin receptors have been proposed to block premature binding of blastocysts to endometrium (Fujii et al. 2006). Another two genes that were upregulated were STAT3 and TAZ. STAT3 is an intracellular signal transduction molecule for various growth factors affecting embryonic development including, in the mouse, LIF and EGF (Cheng et al. 2016). TAZ is another Hippo signaling molecule and is directly related to YAP1 (Kanai et al. 2000). In mouse embryos, TAZ works in conjunction with YAP1 to induce Cdx2 and an increase of TAZ is enough to upregulate CDX2 (Nishioka et al. 2009). Another gene upregulated in TE2 was TET1. This gene, which encodes for an enzyme that converts 5-methylcytosine to 5-hydroxymethylcytosine, is not required for successful development in mice but knocking out the gene is associated with increased development of trophoderm (Dawlaty et al. 2011).

Analysis of single cells indicated expression of some genes previously reported as not being expressed by the bovine blastocyst, most notably the chemokine receptors CCR3 and CCR5 (Negrón-Pérez et al. 2017) and EOMES (Berg et al. 2011). One possibility is that these genes are expressed at low levels by specific cell populations in the TE. Alternatively, it may be that the low levels of these transcripts detected in specific cell populations represent non-specific amplification or low amounts of DNA contamination.

We conclude that the bovine late-blastocyst cell population is heterogeneous and composed of epiblast, hypoblast and four TE cell types. Each cell population may be identified with specific markers previously described and with novel markers identified here. Further studies are needed to confirm whether subpopulations of TE cells represent different maturity stages in development of a committed TE phenotype.
Supplementary data

This is linked to the online version of the paper at http://dx.doi.org/10.1530/REP-17-0345.

Declaration of interest

PIH is a co-owner of Cooley Biotech LLC, the manufacturer of BBH7.

Funding

Research was supported by Agriculture and Food Research Initiative Competitive Grant No. 2011-67015-30688 from the USDA National Institute of Food and Agriculture and from funds from the L.E. ‘Red’ Larson Endowment. Verónica Negrón-Pérez was supported by a McKnight Doctoral Fellowship from the Florida Education Fund, Inc.

Acknowledgements

The authors thank owners and employees of Central Beef Packing Co. (Center Hill, FL), Adena Meat Products L.P. (Fort McCoy, FL), and Florida Beef Inc. (Zolfo Springs, FL) for providing ovariies; William Rembert and Eddie Cummings for ovary collection; Mei Zhang from the Gene Expression and Genotyping Core, Interdisciplinary Center for Biotechnology Research, University of Florida for assisting in the sample processing development; Maxine McClain and Pedja Sekaric from the Fluidigm Corporation for assisting in the troubleshooting process, and the Miami Center for AIDS Research (CFAR) at the University of Miami Miller School of Medicine, which is funded by a grant (P30AI073961) from the National Institutes of Health, and which performed qPCR.

References


Dovey OM, Foster CT & Cowley SM 2010 Histone deacetylase 1 (HDAC1), but not HDAC2, controls embryonic stem cell differentiation. PNAS 107 8422–8427. (doi:10.1073/pnas.1000478107)


Received 7 June 2017
First decision 18 July 2017
Revised manuscript received 3 August 2017
Accepted 16 August 2017