Supplementary Information

Early transcriptional divergence underlies cell fate bias in bovine embryos

Koyama Hinata[1], Daisuke Mashiko[2], Pilar Ferré-Pujol[1], Utano Suzuki[1], Rei Morita[1], Masahiro Kaneda[3], Atchalalt Khurchabilig[1,4], Haruhisa Tsuji[1,5], Tatsuma Yao[6], Satoshi Sugimura[1*]

[1] Biological Production, Institute of Agriculture, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan

[2] Immunology Frontier Research Center, Osaka University, Osaka 565-0871, Japan

[3] Department of Veterinary Medicine, Faculty of Agriculture, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan

[4] Yazhouwan National Laboratory, Sanya 572024, China

[5] Asada Ladies Clinic, Aichi, 450-0002, Japan

[6] Research and Development Center, Fuso Pharmaceutical Industries, Ltd., Osaka 536-8523, Japan*

***Corresponding author:**

Satoshi Sugimura

Biological Production, Institute of Agriculture, Tokyo University of Agriculture and Technology, Tokyo 183-8509, Japan

E-mail: satoshis@cc.tuat.ac.jp

This PDF file includes:

Supplementary Materials and Methods

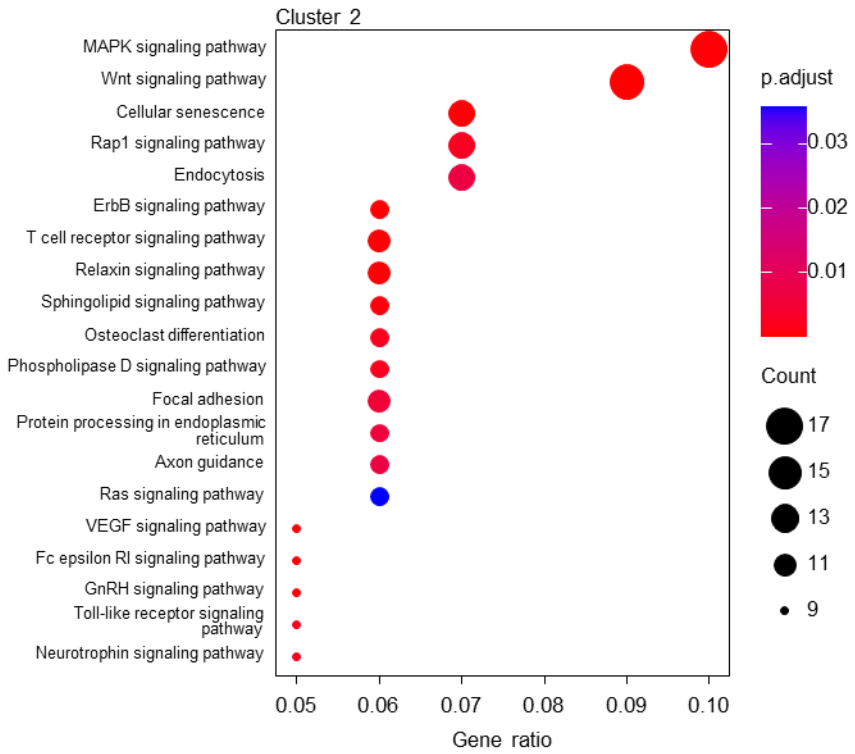
Supplementary Fig. 1-6

Supplementary Table 1-2

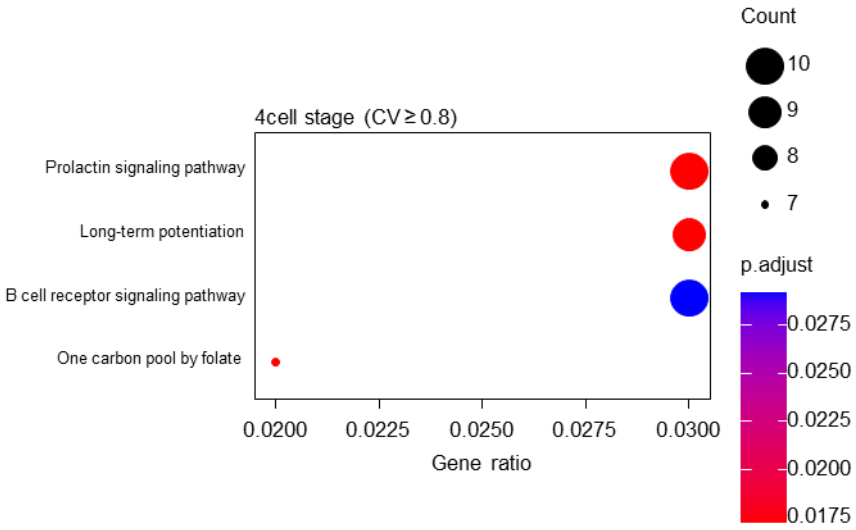
Supplementary Materials and Methods

Quantitative real-time PCR

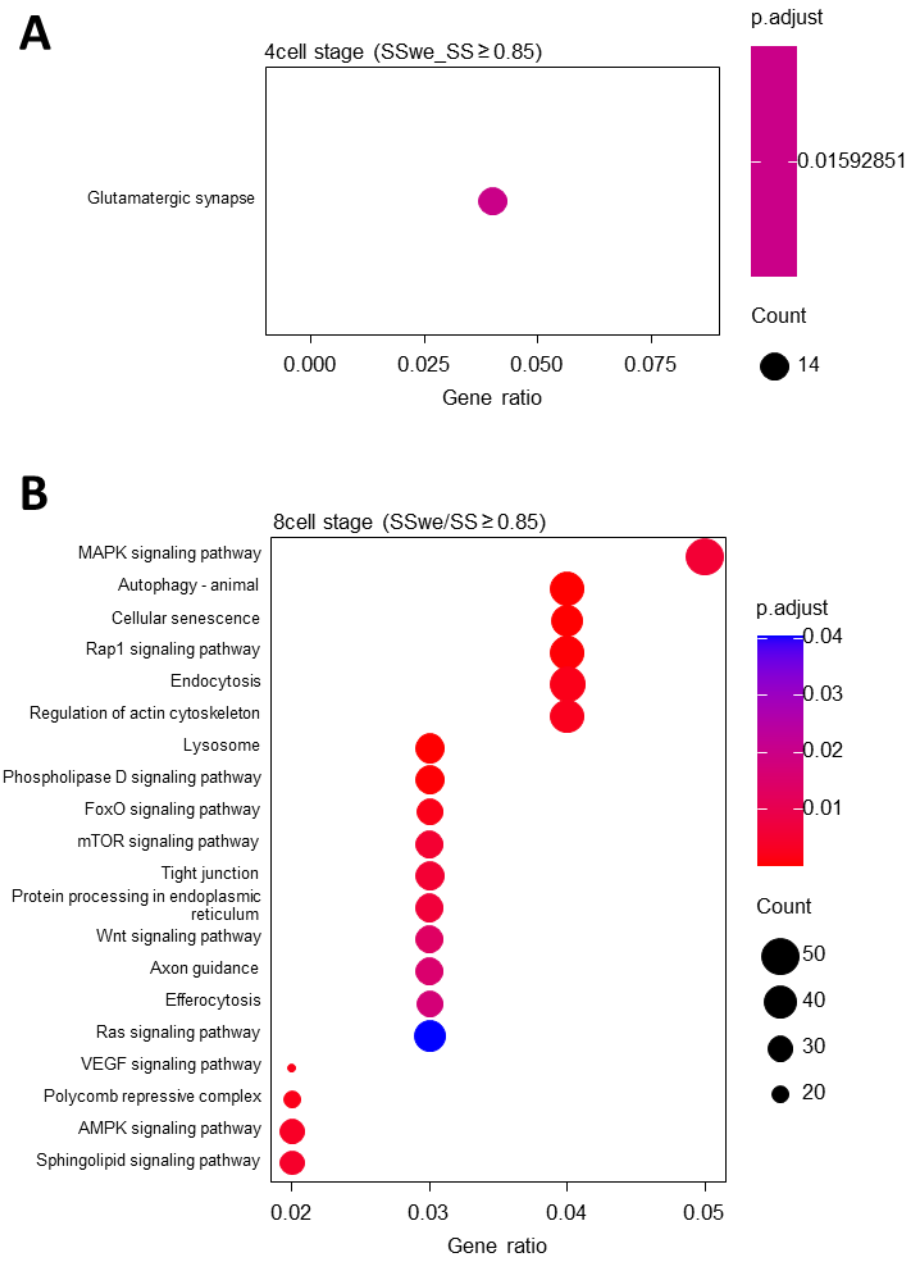
Quantitative real-time PCR (qPCR) was performed on cDNA synthesized from individual blastomeres. Total RNA was reverse-transcribed using the RT-RamDA® cDNA Synthesis Kit (Toyobo, Osaka, Japan) according to the manufacturer's instructions. qPCR was performed on a StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) in a 20-µl reaction volume containing 2 µl of cDNA, 0.5 µl each of the forward and reverse primers (primer sequences listed in Supplementary Table 1), 10 µl of Fast SYBR™ Green Master Mix (Thermo Fisher Scientific), and nuclease-free water. Thermal cycling conditions consisted of an initial denaturation at 95 °C for 20 s, followed by 45 cycles of 95 °C for 3 s, 60 °C for 10 s, and 72 °C for 20 s, followed by melting-curve analysis to confirm product specificity.Absolute transcript levels were quantified using standard curves generated from four serial 5-fold dilutions of purified PCR products. Each reaction was run in duplicate to assess reproducibility. Expression values were normalized using the geometric mean of the housekeeping genes *GAPD* and *YWHAZ*.



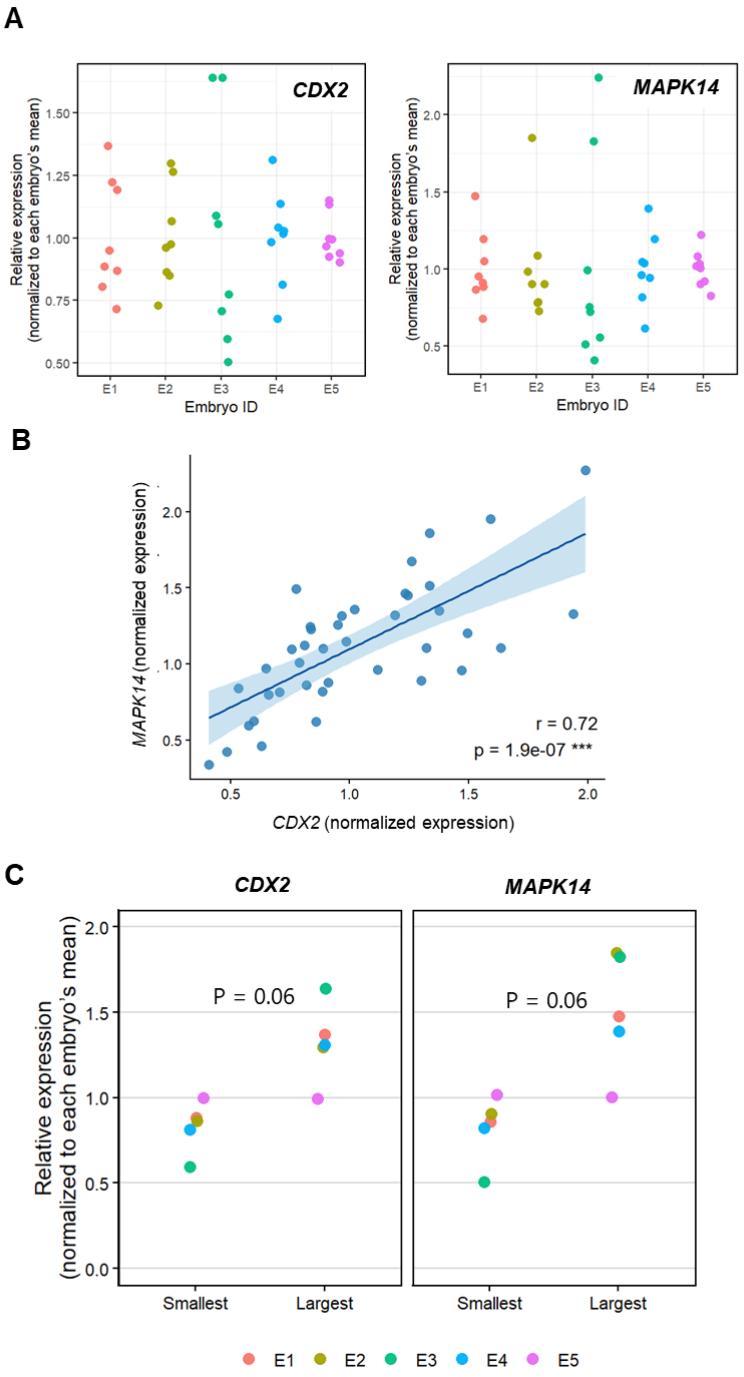
Supplementary Fig. 1. KEGG pathway enrichment analysis of Cluster 2 genes with inter-blastomere variability at 8-cell. This analysis complements the hierarchical clustering shown in Fig. 2A. Dot size and color represent gene count and statistical significance, respectively.



Supplementary Fig. 2. KEGG pathway enrichment analysis of genes exhibiting high inter-blastomere variability (CV ≥ 0.8) at the 4-cell stage. Dot size and color represent gene count and statistical significance, respectively.

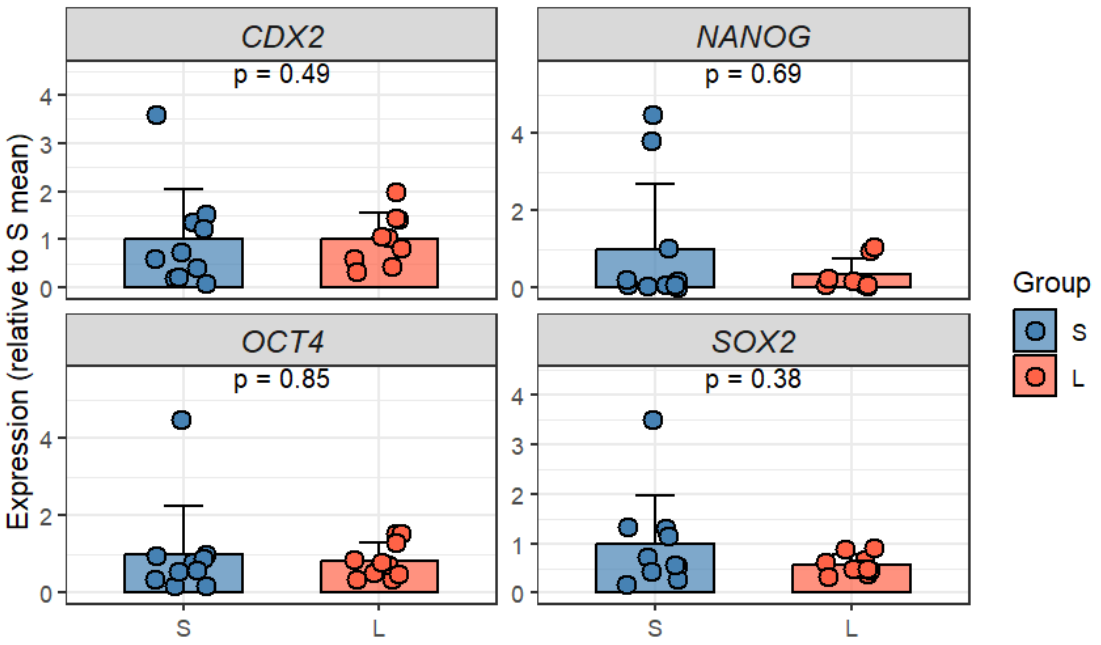


Supplementary Fig. 3. KEGG pathway enrichment analysis of genes with high SSwe/SS ratios (≥ 0.85) at the 4-cell and 8-cell stages. (A) 4-cell stage; (B) 8-cell stage. Dot size and color represent gene count and statistical significance, respectively.

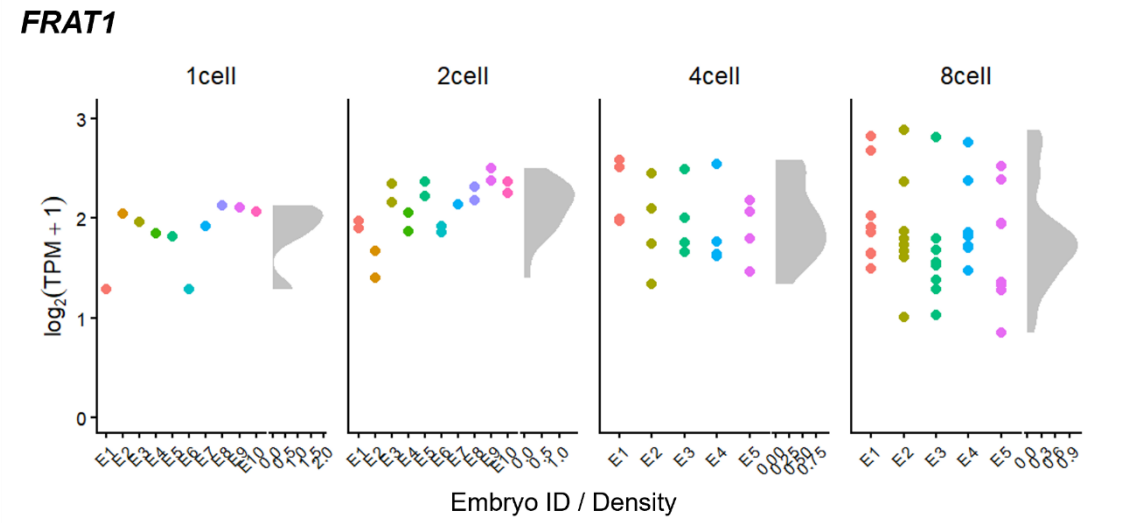


Supplementary Fig. 4. Validation of RNA-seq–identified inter-blastomere heterogeneity of *CDX2* and *MAPK14* expression by real-time qPCR in bovine 8-cell embryos. (A) Relative expression levels of *CDX2* and *MAPK14* in individual blastomeres from five 8-cell embryos (E1–E5), measured to validate RNA-seq results. Expression values were normalized to the mean expression level within each embryo. Each dot represents a single blastomere (eight

per embryo), and colors indicate embryo identity. Consistent with the RNA-seq data, both genes exhibited substantial variability among sister blastomeres within the same embryo. (B) Correlation between *CDX2* and *MAPK14* expression across individual blastomeres in the validation dataset. Each dot represents one blastomere (n = 40). A positive correlation was observed (Pearson's r = 0.72), supporting coordinated expression of *CDX2* and *MAPK14* as identified by RNA-seq. The solid line represents the linear regression fit, and the shaded area indicates the 95% confidence interval. (C) Comparison of relative expression levels between the smallest and largest blastomeres within each embryo for *CDX2* (left) and MAPK14 (right). Each colored dot represents an individual embryo (n = 5). Although the differences did not reach statistical significance, expression levels tended to be higher in the largest blastomeres, consistent with trends observed in the RNA-seq analysis (paired Wilcoxon signed-rank test, one-sided; $p = 0.06$ for both genes; $p < 0.05$ was considered statistically significant).



Supplementary Fig. 5. Gene expression patterns in embryos derived from the smallest (S) and largest (L) blastomeres isolated from 8-cell stage embryos. Individual smallest (S) and largest (L) blastomeres were isolated from 8-cell stage embryos and cultured independently. Gene expression of blastocyst-forming embryos on Day 7 post-fertilization was analyzed (n = 10 embryos per group). Bar plots show mean normalized expression (S mean = 1) with SD overlaid with individual data points; p-values were determined using an unpaired Wilcoxon test and $P < 0.05$ was considered statistically significant.



Supplementary Fig. 6. Comparison of expression levels of *FRAT1* between sister blastomeres at the 8-cell stage. *FRAT1* is a Wnt signaling–related gene identified among the differentially expressed genes (DEGs) between the smallest (S) and largest (L) blastomeres. Gene expression was analyzed from the 1-cell to 8-cell stages: 10 embryos at the 1-cell stage (E1–E10), two sister blastomeres from each of 10 embryos at the 2-cell stage (E1–E10), four sister blastomeres from each of five embryos at the 4-cell stage (E1–E5), and eight blastomeres from each of five embryos at the 8-cell stage (E1–E5). Gray shading on the right indicates the density distribution of expression levels.

Supplementary Table 1. Sequences of primers used for quantitative real-time PCR.

Gene	Accession No.	Primer sequence (5'→3')	Product size (bp)
GAPDH	XM_618013	F: TTCAACGGCACAGTCAAGG R: ACATACTCAGCACCAGCATCAC	119
YWHAZ	BM446307	F: GCATCCCACAGACTATTTCC R: GCAAAGACAATGACAGACCA	120
CDX2	XM_871005	F: GCCACCATGTACGTGAGCTAC R: ACATGGTATCCGCCGTAGTC	140
MAPK14	NM_001102	F: GCTGTCGACCTGCTGGAGAAGATG R: TCGTCGTCAGGATCGTGGT ACTGG	110
NANOG	DQ069776	F: GTCCCGGTCAAGAAACAAAA R: TGCATTTGCTGGAGACTGAG	107
SOX2	NM_001105463	F: GGTTGACATCGTTGGTAATTTATAATAGC R: CACAGTAATTTCATGTTGGTTTTTCA	88
OCT4	NM_174580.3	F: ACATGTGTAAGCTGCGGCCC R: CTTTCGGGCCTGCACAAGGG	107

Supplementary Table 2. Bootstrap estimation of expression differences between the smallest and largest blastomeres*

Gene	Stage	Mean difference (TPM)	95% CI (TPM)	p-value
CDX2	4cell	2.04	[0.30, 4.29]	0.017
	8cell	2.19	[0.14, 4.69]	0.028
HRAS	4cell	0.58	[-4.35, 5.72]	0.396
	8cell	7.38	[0.76, 13.9]	0.017
MAPK14	4cell	1.21	[-8.10, 10.32]	0.415
	8cell	11.4	[1.48, 21.32]	0.013
RAC1	4cell	120.82	[28.4, 92.8]	< 0.001
	8cell	59.18	[28.4, 92.8]	< 0.001

* Mean differences in gene expression (TPM) between the smallest and largest blastomeres within the same embryo at the 4-cell and 8-cell stages. Values represent the bootstrap mean difference (Smallest − Largest), 95% confidence interval, and p-values calculated by bootstrap resampling. Statistical significance was defined as $p < 0.05$ and by 95% CI not crossing zero.