The asymmetric expression of HSPA2 in blastomeres governs the first embryonic cell-fate decision

Jiayin Gao [1,2,3,4,5,6,7,8†], Jiawei Wang[1,2,3,4,5,6,7,8†], Shiyu Liu[1,2,3,4,5,6,7,8], Jinzhu Song[1,2,3,4,5,6,7,8], Chuanxin Zhang[1,2,3,4,5,6,7,8], Boyang Liu[1,2,3,4,5,6,7,8*], Keliang Wu [1,2,3,4,5,6,7,8*]

Address: [1] Institute of Women, Children and Reproductive Health, Shandong University, 250012, China. [2] State Key Laboratory of Reproductive Medicine and Offspring Health, Shandong University, Jinan, Shandong, 250012, China. [3]National Research Center for Assisted Reproductive Technology and Reproductive Genetics, Shandong University, Jinan, Shandong, 250012, China. [4]Key Laboratory of Reproductive Endocrinology (Shandong University), Ministry of Education, Jinan, Shandong, 250012, China. [5]Shandong Technology Innovation Center for Reproductive Health, Jinan, Shandong, 250012, China. [6]Shandong Provincial Clinical Research Center for Reproductive Health, Jinan, Shandong, 250012, China. [7]Shandong Key Laboratory of Reproductive Medicine, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, 250012, China.[8]Research Unit of Gametogenesis and Health of ART-Offspring, Chinese Academy of Medical Sciences (No.2021RU001), Jinan, Shandong, 250012, China.

[†] These authors contributed equally to this study.

*Corresponding author: boyang Liu and keliang Wu.

Correspondence address: byliu@sdu.edu.cn (B.L.), wukeliang@sdu.edu.cn (K.W.).

Abstract: The first cell-fate decision is the process by which cells of an embryo take on distinct lineage identities for the first time, thus representing the beginning of developmental patterning. Here, we demonstrate that the molecular chaperone heat shock protein A2 (HSPA2), a member of the 70 kDa heat shock protein (HSP70) family, is asymmetrically expressed in the late 2-cell stage of mouse embryos. The knockdown of *Hspa2* in one of the two-cell blastomeres prevented its progeny predominantly toward the inner cell mass (ICM) fate, thus indicating that the differential distribution of HSPA2 in the blastomeres of two-cell embryos can influence the selection of embryonic cell lineages. In contrast, the overexpression of *Hspa2* in one of the two-cell blastomeres did not induce blastomeres to differentiate towards the ICM fate. Furthermore, we demonstrated that HSPA2 forms a complex with CARM1 and activates ICM-specific gene expression. Collectively, our results identify HSPA2 as a critical regulator of the first cell-fate decision which specifies the ICM via the execution of commitment and differentiation phases.

Key words: first cell-fate decision, inner cell mass, HSPA2, CARM1, embryo

1. INTRODUTION

The preimplantation development of mammalian embryos is a complex and orderly process. The first cell-fate decision is a key event in the preimplantation development in mammalian embryos. Specifically, the first cell-fate decision occurs at the morula stage following embryonic compaction and results in the division of embryonic cells into two distinct lineages: the trophectoderm (TE) and the inner cell mass (ICM). The TE is a single layer surrounding the fluid-filled cavity known as the

blastocoel that provides extraembryonic structures such as the placenta, whereas the ICM, which is attached to the inside of the TE, contains pluripotent cells that give rise to the fetus and extraembryonic tissues [1-3]. However, it remains an open question that how these cells shape the first cell-fate decisions of preimplantation development, and when they begin to trigger the initial differentiation signals.

The asymmetry of the blastocyst may be traced back to the asymmetric nuclear abundance of coactivator-associated arginine methyltransferase 1(CARM1) at the 4-cell stage [4-6]. The heterogeneous activity of CARM1 leads to the differential methylation of H3 methylation of arginine 26 (H3R26me). More extensive methylation of H3R26me can result in a higher increase in the expression levels of ICM-specific genes (e.g., *Oct4*, *Nanog* and *Sox2*) and global chromatin accessibility, which directs their progeny into the ICM fate [7]. Recent study has further suggested that asymmetry of the blastomere can even be traced back to the 2-cell stage. For example, a long non-coding RNA, *LincGET*, is known to be asymmetrically expressed in the nuclei of blastomeres in 2 and 4-cell embryos. Specifically, *LincGET* and CARM1 are known to form a complex that promotes the H3 methylation of arginine 26 (H3R26me) and activates the expression of genes in the ICM [8]. In addition, HMGA1, a member of the high mobility group (HMG) protein family, is also known to exhibit heterogeneity as early as the two-cell stage and involves in cell fate differentiation [9]. Collectively, these studies suggest that molecular heterogeneity takes place in the early stage of embryonic development and is associated with cell fate differentiation.

Despite these findings, the specific factors that exhibit heterogeneity at the two-cell stage and the mechanisms by which the cell fate and differentiation can be decided still require further investigation. The molecular chaperone heat shock protein A2 (HSPA2) is a member of the 70 kDa heat shock protein (HSP70) family that is evolutionarily conserved in human and metazoan. It is originally identified in male germ cells and described as a testis-specific protein that is essential for spermatogenesis [10, 11]. The *Hspa2* gene is located on chromosome 14 (14q24.1)[12], and the aberrant expression of HSPA2 in testes was shown to induces primary spermatocytes to arrest in meiosis I and undergo apoptosis, thus leading to male infertility [13, 14]. HSPA2 expression in somatic tissues and during mouse embryogenesis [15, 16], and high levels of HSPA2 expression play a critical role in the genesis and progression of carcinoma [17, 18]. However, the role of HSPA2 in lineage differentiation of embryonic cells is yet to be unveiled.

In this study, we separated blastomeres from the 4-cell stage and cultured to the morula stage, analyzed the gene expression pattern by using Smart-seq2 to identify differentially expressed genes. *Hspa2* was identified and was shown to be asymmetrically distributed in two- to four-cell blastomeres. The knockdown of *Hspa2* in one of the two cell blastomeres at the 2-cell embryo preventedits progeny predominantly towards the ICM fate. Furthermore, HSPA2 and CARM1 were found to form a protein complex and regulate ICM-specific gene expression. Collectively, these results suggest that heterogeneity in *Hspa2* expression patterns can affect cell fate bias in the mouse embryo as early as the 2-cell stage.

2. MATERIALS AND METHODS

2.1 Antibodies

In this study, we utilized CDX2 (ab76541), NANOG (ab17338) and H3R26me2 (ab127095) antibodies from Abcam;a CARM1 (12495s) antibody from Cell Signaling Technology; SOX2 (11064-1-AP), GAPDH (10494-1-AP), CARM1 (55246-1-AP), HSPA2 (12797-1-AP) and ACTIN (66009-1-Ig) antibodies from Proteintech; and an OCT4 (sc-5279) antibody from Santa Cruz Biotech.

2.2 Mouse embryo collection

Animal experiments were approved by the Ethics Committee of Shandong University. ICR mice (6–8 weeks-of-age) were purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. and bred under standard SPF conditions with a 12 h dark and 12 h light cycle with food and water provided ad libitum). Superovulation was induced by an intraperitoneal injection of 5 IU of pregnant mare's serum gonadotrophin (PMSG) (110914564, Ningbo Sansheng Biological Technology Co., Ltd) followed 46–48 h later by 5 IU of human chorionic gonadotrophin (HCG) (110911282, Ningbo Sansheng Biological Technology Co., Ltd.). Oocytes were collected 15 h after hCG injection. First, the collected oocytes were fertilized in G-IVF (#10,136, Vitrolife) with epididymal sperm from adult males that had been capacitated in G-IVF for 1 h. All zygotes were cultured in G1 medium (Vitrolife, Sweden) at 37 ℃ in the presence of 6% CO_2 and 5% O_2. Early 2-, late 2-, 4-, and 8-cell embryos, and morula- and blastocyst stage embryos were collected after 22–26, 34-38, 48–50, 60-65, 70–75, and 96–100 h of culture, respectively. Single

blastomeres were isolated from 2- and 4-cell embryos and cultured to the morula stage as described previously[19].

2.3 Cell culture and transfection

Mouse ESCs (mESCs) were cultured on plates coated with 0.1% gelatin (Oricell) under feeder-free conditions with 5% CO2 at 37 ℃. For cell transfection, the mESCs were transfected with si-*Hspa2* using Lipofectamine 3000 (Invitrogen) according to the manufacturer’s instructions. After transfection for 36 hours, cells were collected for further analysis.

2.4 RT-qPCR analysis

For RT-PCR analysis, we used the Single Cell Sequence Specific Amplification Kit (Vazyme, P621-01) and ChamQ Universal SYBR qPCR Master Mix (Vazyme, Q711-03) in accordance with the manufacturer's instructions. Primers for the candidate genes were designed by online Primer 3.0 software based on the sequence data obtained from the NCBI database. Primer sequences are listed in Table 1. Gene expression was quantified by the comparative CT method. Values are expressed as the mean ± standard deviation (SD). Expression was normalized to Actin. The relative amount of transcript present in each cDNA sample was calculated using the $2^{-\Delta\Delta CT}$ method. All experiments were repeated at least three times, and statistical analysis was performed on individual experimental sets. All measurements were performed in triplicate for each experiment.

Table 1. List of mouse primer pairs used for RT-qPCR analysis

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Hspa2	AAGATTTCTTCAACGGCAAGGAG	GGGATCGTGGTGTTTCTCTTGAT
Oct4	GGGTGATGGGTCAGCAG	TCCGCAGAACTCGTATGC
Sox2	AACGCCTTCATGGTATGG	CTCGGTCTCGGACAAAAG
Nanog	CTGCTCCGCTCCATAACT	GGCTTTCCCTAGTGGCTT
Cdx2	GAAGGGGTGGTGGGTTC	AGGTTGGCTCTGGCATTT
Actin	GCTCTTTTCCAGCCTTCCTT	GTACTTGCGCTCAGGAGGAG
Hprt	CGAGATGTGATGAAGGAAATGG	AGCAGGTCAGCAAAGAATTTATAGC

2.5 Automated western immunoblotting

Western immunoblots were performed on a PeggySue (ProteinSimple) system using a Size Separation Master Kit with Split Buffer (12–230 kDa) in accordance with the manufacturer's instructions. Capillary-based immunoassays were performed using the Wes-Simple Western method (Proteintech: HSPA2 (12797-1-AP), 1:15 dilution; GAPDH (10494-1-AP),1:35 dilution; SOX2 (11064-1-AP), 1:20 dilution; Abcam: CDX2 (ab76541), 1:10 dilution; Santa Cruz Biotech: OCT4 (sc-5279), 1:15 dilution) and an anti-rabbit or an-mouse detection module (ProteinSimple) [20]. Protein expression was detected by chemiluminescence and quantified by area under the curve (AUC) analysis using the Compass for Simple Western program (ProteinSimple).

2.6 Microinjection and HSPA2 knockdown

Aliquots of 2–5 pl of siRNA at 20 μM concentration were microinjected into the nucleus of zygote or one blastomere of the two-cell embryos using an Eppendorf (Hamburg, Germany) micromanipulator under a Leica inverted microscope, and then

cultured in fresh G1 medium.

To knockdown Hspa2, two siRNAs were designed for each gene. The siRNA of *Hspa2* are GGTGCAACAACTTAGTTTA, GCACTGCAGTGATATTAAA. For microinjection in each KD group,20 μM of each siRNA was used. siRNAs were ordered from RINOBIO.

2.7 Western blot analysis and co-immunoprecipitation (Co-IP)

Protein concentration was determined with a BCA protein assay (Bio-Rad, Hemel Hempstead, UK). Proteins were then separated by 12% SDS-PAGE and transferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA, USA) at 300mA for 90min. First, the membranes were incubated in TBST containing 5% non-fat milk for 2h; then, the membranes were incubated with primary antibodies (Abcam: CDX2 (ab76541), 1:1000 dilution; NANOG (ab17338) 1:800 dilution; Cell Signaling Technology: CARM1 (12495s), 1:1000 dilution; Proteintech: SOX2 (11064-1-AP), 1:800 dilution; GAPDH (10494-1-AP), 1:1500 dilution; HSPA2 (12797-1-AP), 1:1000 dilution; ACTIN (66009-1-Ig), 1:700 dilution; Santa Cruz Biotech: OCT4 (sc-5279), 1:800 dilution) overnight at 4°C. On the second day, the membranes were washed three times with TBST and then incubated with horseradish peroxidase-conjugated secondary antibody (1:2000 dilution; Proteintech) for 1 h.

The Pierce Co-IP kit (Thermo Fisher Scientific, USA) was used for co-IP assays. In brief, mESCs were lysed using IP lysis buffer on ice. Next, the lysate was treated with agarose resin at 4°C for 1 h followed by incubation with antibodies (Proteintech: HSPA2(12797-1-AP), 1:100 dilution; CARM1(55246-1-AP), 1:60 dilution) at 4°C

overnight. The captured antigen was then eluted and SDS-PAGE was performed.

2.8 RNA FISH and immunofluorescence

Fluorescence *in situ* hybridization (FISH) was carried out using an RNA-FISH Probe kit (GenePharma, Shanghai), according to the manufacturer's instruction. In brief, embryos were collected and fixed in 4% paraformaldehyde in PBS-DEPC for 15 min at room temperature, permeabilized in 0.1%BufferA (permeabilizing solution) for 15 min at room temperature, and incubated at 37°C in 2× Buffer C for 30 min. The embryos were then dehydrated in 70%, 85%, and 100% EtOH and dried at room temperature for 5 min each time. The embryos were denatured in buffer E (hybridization buffer) containing 10 μM probe at 73°C for 5 min, incubated at 37°C overnight, and washed in 0.1% Buffer F for 5 min to remove non-specific binding. After washing in 2× Buffer C and 1× Buffer C (neutral solution) for 5 min, the nuclei were stained with DAPI for 5 min

To perform immunofluorescence, we removed the zona pellucida with acid Tyrode's solution (Sigma). After that, embryos were fixed in 4% paraformaldehyde in PBS (KGB5001, KeyGEN BioTECH) at room temperature (RT). Embryos were then permeabilized with 1% Triton-100 (T8787, SigmaAldrich) in PBS for 20 min. Embryos were blocked for 30 min at RT on a shaker with a blocking buffer consisting of 1% BSA in PBS containing 0.1% Tween-20 (P9416, Sigma-Aldrich) and 0.01% Triton-100. Next, embryos were incubated with primary antibodies (Abcam: CDX2 (ab76541), 1:150 dilution; H3R26me2 (ab127095), 1:150 dilution; Santa Cruz Biotech: OCT4 (sc-5279), 1:100 dilution) diluted in blocking buffer on a shaker at 4-℃

overnight. On the second day, embryos were washed three times with TBST and incubated with secondary antibody (1:400 dilution; Sigma) and DAPI (D3571, 1:400 dilution; Life Technologies) for 30 min. First, embryos were rinsed three times for 10 min in washing buffer on a shaker at RT; then, the embryos were mounted on slides in washing buffer and examined under a laser scanning confocal microscope (Dragonfly, Andor Technology, UK). Immunofluorescence images were analyzed by ImageJ Software (https://imagej.nih.gov/ij/). To analyze the ICM/TE cell numbers, OCT4 and CDX2 signal at inner layers and outer layers were used to distinguish the ICM and TE cells.

2.9 Molecular docking analysis

For molecular docking analysis, we first generated a PPI network and identified core targets. The proposed protein model was based on the crystal structure downloaded from the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/) and the two- and three-dimensional (2D/3D) structures of potential active compounds were acquired from PubChem (https://pubchem.ncbi.nlm.nih.gov/). Then, molecular docking analysis was performed and visualized in Maestro11.9; the resultant docking score was saved and then compared.

2.10 Proteomics and bioinformatic analysis

For proteomics and bioinformatic analysis, 4 embryos per sample were collected. Four replicates were assessed for NC group and two replicates were assessed for KD group. Protein samples were sonicated three times on ice using a high intensity

ultrasonic processor (Scientz) in lysis buffer (8 M urea, 1% protease inhibitor cocktail). Inhibitors were also added to lysis buffer for PTM experiments. The remaining debris was removed by centrifugation at 12,000 g at 4°C for 10 min. The protein samples were then diluted by adding 200 mM of TEAB to urea concentrations < 2 M. Finally, trypsin was added at a 1:50 trypsin-to-protein mass ratio for the first digestion overnight and a 1:100 trypsin-to-protein mass ratio for a second 4 h-digestion. Finally, the peptides were desalted with a Strata X SPE column. The eluted peptides were then cleaned with C18 ZipTips (Millipore) in accordance with the manufacturer's instructions, followed by analysis with LC–MS/MS. The resulting MS/MS data were processed by MaxQuant with an integrated Andromeda search engine (version 1.4.1.2). False discovery rate thresholds for proteins, peptides and modification sites were specified at 1%.

COG, full name is Clusters of Orthologous Groups of proteins. The proteins that make up each COG are assumed to be derived from an ancestral protein, Orthologs are proteins that are derived from different species, evolved from vertical lineages (species formation), and typically retain the same function as the original protein. COG means "cluster of homologous proteins" in Chinese. COGs are divided into two categories, one for prokaryotes and the other for eukaryotes. Those for prokaryotes are generally referred to as COG databases; those for eukaryotes are generally referred to as KOG databases. Compared to other databases, such as NCBI's COG database, EggNOG provides a more comprehensive classification of species and more homologous protein sequences, with phylogenetic tree construction and functional

annotation for each homologous gene cluster.

2.11 RNA-seq analysis

For RNA sequencing, 3 embryos per sample were collected. Three replicates were assessed for NC group and three replicates were assessed for KD group. After being denuded of the zona pellucida using EmbryoMax Acidic Tyrode's Solution (MR-004-D, Sigma-Aldrich), RNA-seq libraries were generated following the Smart-seq2 protocol as described previously[21]. Sequencing libraries were generated using TruePep™DNA Library Prep Kit V2 for Illumina® (TD502, Vazyme) according to manufacturer's recommendations. The libraries were performed on an Illumina Hiseq X-ten platform with 150 bp paired-end.

Filtering out low quality and joint sequences, the clean reads of each sample obtained were mapped to the reference genome downloaded from the genome website (https://ftp.ensembl.org/pub/release-105/fasta/mus_musculus/dna/Mus_musculus.GRCm39.dna.primary_assembly.fa.gz) using Hisat2. The expression values for each gene were calculated as Fragments Per Kilobase of exon model per Million mapped fragments (TPM+1). DEG analysis was conducted with edgeR package. In this project, the gene were identified as differentially expressed genes (DEGs) if |log2FC| >1 and P<0.05 (FC: fold change). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were s carried out for all the DEGs, utilizing the "cluster Profiler" package.

2.12 Statistical analysis

Statistical analyzes were performed by GraphPad Prism software (version 9.3, GraphPad Prism Software Inc., San Diego, CA) and the Statistics Package for Social Science (SPSS 11.5; SPSS Inc., Chicago, IL, USA). One-way analysis of variance (ANOVA) was used to analyze data. At least three biological replicates were performed for each experiment and results are represented as mean ± standard error of mean (SEM). $p < 0.05$ was considered statistically significant.

2.13 Data access

The RNA-sequencing data used in this study were obtained from the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession numbe (GSE266025). Proteomic data were available via ProteomeX change with identifier PXD052193.

3. RESULTS

3.1 *Hspa*2 is asymmetrically expressed between mouse two- and four-cell blastomeres

In our previous study, we established a culture system that allows blastomeres without their zona pellucida to develop in vitro until the blastocyst stage [19]. Here, in order to expand the heterogeneity between blastomeres and identify factors that may involve in cell lineage differentiation, single blastomeres from the mouse 4-cell embryo were separated and cultured to the morula stage. By using Smart-seq2 technology, we analyzed the gene expression pattern of the morulae, which developed from the 4-cell blastomeres (Figure 1A). Hspa2 was identified by differential gene expression analysis and showed distinct mRNA levels among the four morulae

(Figure 1B). In addition, the level of *Nanog* and *Hspa2* mRNA showed stronger positive correlation (R=0.624) in gene expression than random gene pairs (R=0.029), indicating that *Nanog* expression levels increased with the increase of *Hspa2* (Figure 1C). When the late 2-cell embryo was divided into single blastomeres and cultured to the morula stage, we obtained similar results that *Hspa2* mRNA distinctly distributed between the two morulae (Figure S1A, B). Next, we investigated the expression of *Hspa2* during preimplantation embryo development in whole embryos from meiosis II to the blastocyst stage. qRT-PCR analysis showed that the *Hspa2* mRNA was stored maternally at the MII oocyte and maintained at high level at the zygote stage (Figure S1C).

To directly examine the heterogeneity among individual blastomeres, we separated 4-cell embryos into individual four blastomeres which were then analyzed by single-cell qRT-PCR and Jess TM Simple Western System based automated western immunoblotting with high sensitivity. The results showed that the mRNA and protein levels of HSPA2 differed among the four blastomeres (Figure 1D, E), and confirmed by fluorescence *in situ* hybridization (FISH) analysis (Figure 1F). We also identified heterogeneous HSPA2 expression during the late 2-cell stage (Figure 1G-I). For the early 2-cell stage, however, HSPA2 was symmetrically distributed between the two blastomeres (Figure S1C-E). These results indicating that the heterogeneity occurred as early as the late 2-cell stage. Thus, we hypothesized that the heterogeneous distribution of HSPA2 in late 2-cell plays a direct or indirect role in the lineage commitment of early embryonic stages.

3.2 The reduced expression of *Hspa2* leads to the reduced expression of ICM-marker

To test our hypothesis, we knocked down the *Hspa2* mRNA level by injecting *Hspa2*-specific small interfering RNA (siRNA) into the cytoplasm of zygote, and NC-FAM (negative control with FAM label) siRNA-injected embryos served as the control (Figure 2A). The differentially expressed genes were detected by RNA-seq for *Hspa2* siRNA (Knockdown; KD) and control (NC) groups at the 8-cell stage (Figure 2B). In total, 96 up-regulated genes and 410 down-regulated genes were identified in the *Hspa2*-KD group when compared to the NC group (Figure 2C, D). To investigate the potential biological functions of the DEGs, we conducted GO and KEGG enrichment analysis. The most significant GO enriched terms were closely related to cell fate commitment, stem cell population maintenance, stem cell differentiation and blastocyst development (Figure 2E). According to KEGG analysis, the down-regulated genes were primarily associated with Hippo and Wnt signaling pathways (Figure 2F), which are known to play an important role in cell proliferation, differentiation, and fate decisions in embryonic development[22-25]. Furthermore, the blastocyst formation rate in the *Hspa2-KD* group (52.96±2.489%) was significantly lower than that in the NC group (81.63±2.000%) (Figure 2G, H), thus suggesting that HSPA2 plays an essential role in preimplantation embryonic development. In addition, we found that the down-regulation of *Hspa2* resulted in a remarkable down-regulation of specific ICM-marker genes (*Oct4*, *Sox2* and *Nanog*) at the 4-cell and blastocyst stages. Meanwhile, the mRNA levels of a TE-marker gene, Cdx2, did not

significantly change between the groups (Figure 2I, J). We then used automated Western immunoblotting to detect the protein levels of ICM and TE markers in Hspa2-KD embryos, and found that the expression levels of ICM-marker proteins (OCT4 and SOX2) were inhibited when HSPA2 was down-regulated at the 4-cell and blastocyst stages, while the TE-marker protein CDX2 was not significantly different between the groups (Figure 2K, L). Similarly, for mouse embryonic stem cells (mESCs) the knockdown of HSPA2 resulted in a significant reduction for the ICM-marker proteins, and no change for CDX2 (Figure 2M). These results suggest that HSPA2 plays a role in the regulation of ICM-specific gene expression.

3.3 The knockdown of *Hspa2* expression prevents blastomeres from an ICM fate

To investigate the functional role of HSPA2 in the first lineage segregation during early embryonic development, we examined the proportion of ICM and TE after Hspa2-KD. It was showed that the number of ICM and total cells in the blastocyst were significantly decreased, but the number of TE cells did not change significantly (Figure 3A, B). We then randomly co-injected green fluorescent protein (*Gfp*) mRNA with either *Hspa2*-siRNA or NC-FAM into one blastomere of the two-cell embryos to trace the fate of its descendent cells into the blastocyst (Figure 3C). Compared with the GFP+NC group (48.23%±7.415%), the GFP+*Hspa2*-KD group (33.75%±4.303%) exhibited a significantly reduced ratio of progeny cells with an ICM fate. Moreover, the number of ICM-marker OCT4 positive cells was found to be significantly reduced in the GFP+*Hspa2*-KD group (27.09±3.811%) than in the GFP+NC group (36.95%±2.524%) (Figure 3D, E). This is consistent with our GO

enriched terms of RNA-Seq in that the negative regulation of cell cycle, G1/S transition of mitotic cell cycle, mitotic cell cycle phase transition and regulation of mitotic cell cycle phase transition was found in *Hspa2*-KD group (Figure 2E). Collectively, these results suggested that HSPA2 has a role in ICM lineage establishment and its deficiency leads to partial defects in the ICM cells.

3.4 The overexpression of *Hspa2* does not induce blastomere cells to bias an ICM fate

To determine whether the redundant level of HSPA2 protein also affected first lineage segregation, we overexpressed *Hspa2* by microinjecting *Hspa2* mRNA (approximately 400ng/μL) into the zygote. We observed that the blastocyst formation rate was not affected significantly (Figure 4A, B), thus showing that the overexpression of *Hspa2* did not exert adverse effects on embryonic development. In addition, we verified that the mRNA levels of the ICM-marker genes (*Oct4*, *Sox2* and *Nanog*) and the TE-marker gene *Cdx2* at the 4-cell and blastocyst stage did not differ significantly (Figure 4C, D). We also examined the number of ICM cells and TE cells by immunostaining OCT4 and CDX2 in blastocysts from the overexpression and control groups and found they were not altered significantly (Figure 4E, F). To trace the fate of the descendent, we co-injected GFP mRNA along with either *Hspa2* or NC mRNA randomly into one blastomere of 2-cell embryos and counted the proportion of ICM and TE cells of blastocysts. We found that the percentage of GFP cells that were present in the ICM at the blastocyst stage was similar between the two groups (48.23±7.415 vs. 47.25±4.432), and the percentage of ICM cells also did not differ

significantly (36.95±2.524 vs. 36.24±1.874) (Figure 4G, H). These observations indicated that the redundant level of HSPA2 were not sufficient to induce blastomere cells to bias an ICM fate.

3.5 HSPA2 interacts with CARM1 in the segregation of embryonic cell lineage

To further investigate the mechanism underlying the function of HSPA2 in cell fate decision during early embryonic development, we next analyzed differentially expressed proteins (DEPs) at the blastocyst stage in the *Hspa2*-KD and NC groups by applying untargeted proteomics. Volcano map screening identified a total of 43 annotated DEPs, of which 15 proteins were significantly up-regulated and 28 proteins were downregulated (Figure S2A). These DEPs were mainly involved in post-translational modification by using Clusters of Orthologous Groups (COG)/Karyotic Orthologous Groups (KOG) functional annotation statistics (Figure S2B). In order to compare the functional similarities and differences between proteins with different differential multiples, we divided the DEPs into four classifications: Q1 (<0.5), Q2 (0.5-0.667), Q3 (1.5-2.0), and Q4 (>2.0) according to differential expression multiples (Figure S2C). Biological processes analysis showed that DEPs were associated with positive regulation of the Wnt signaling pathway and cellular development (Figure S2D). In terms of the cellular component, the DEPs were related to the methylome and methyltransferase complex (Figure S2E). These results indicated that the DEPs were mainly associated with the post-translational modification of proteins and cell fate decision.

Coactivator-associated arginine methyltransferase 1 (CARM1), also known as

PRMT4, was the first protein arginine methyltransferase which was proved to be associated with transcriptional activation by methylating and/or regulating histone H3 and non-histone proteins, such as coactivators and transcription factors [26, 27]. Recent studies have demonstrated that CARM1 regulates early mouse development and plays a role in the determination of cell fate [28, 29]. CARM1 levels, along with its specific histone marks (H3R17me2a and H3R26me2a), were found to vary among the four-cell blastomeres and that high levels of CARM1 led to increased levels of H3R26me. This varied distribution led to biased subsequent fate of the blastomeres towards the ICM [5, 30]. Next, we considered whether HSPA2 regulates the decision of the first cell-fate through CARM1. Molecular docking experiments revealed that HSPA2 docked efficiently with CARM1 via several H-bonds and pi-interactions, with strong binding affinity (Figure 5A). The Search Tool for the Retrieval of Interacting Genes (STRING) database was used to predict interactions between HSPA2 and CARM1. The results showed that HSPA2, CARM1 and ICM-specific proteins interacted with each other (Figure S2F). In addition, we found that the *Hspa2*-KD led to reduced CARM1 mRNA as well as protein levels at the 4-cell and blastocyst stages (Figure 5B-E). The knockdown of HSPA2 in mESCs also caused a significant reduction in the levels of CARM1 (Figure 5G).

To further investigate whether HSPA2 could interact with CARM1, we performed co-immunoprecipitation (co-IP) assays in mESCs using HSPA2 and CARM1 as bait. Co-IP analysis showed that HSPA2 and CARM1 proteins interacted effectively (Figure 5F, G). Moreover, we investigated whether HSPA2 could regulate

H3R26me2 modification, a known event that occurs downstream of CARM1 [5]. Immunofluorescence staining confirmed that CARM1-mediated H3R26me2 levels were significantly reduced in the interference group at the 4-cell stage (Figure 5H). Collectively, our data demonstrated that HSPA2 involved in the segregation of embryonic cell lineage by forming a complex with CARM1 and through H3R26me2 modification.

DISCUSSION

During the development of the fertilized eggs to blastocyst stage, the fundamental questions are to explain how the totipotent fertilized egg develops into different cell types and when cellular heterogeneity occurs. Previous studies demonstrated that molecular polarity and differential gene regulation by signaling pathways drive the first cell-fate decision in mammals [25, 31-33]. The tracing of embryonic lineages from the 2-cell stage onwards by live embryo labeling demonstrated that the cellular heterogeneity is initiated early in the embryo, and thus was proposed as cellular heterogeneity hypothesis [34]. H3R26me and its methyltransferase CARM1exhibit asymmetric expression patterns in the 4-cell stage of mouse embryos and plays an important role in regulating the first cell-fate decision. [35, 36]. PRDM14 is also heterogeneously expressed in 4-cellstage mouse embryos and interacts with CARM1 to drive progenies towards pluripotent ICM fate by increasing the expression levels of H3R26me in 4-cell embryos [37]. Despite the progress in our understanding of the mechanisms governing early cell fate decisions, many important questions still remain unanswered. We do not yet fully comprehend

when the initial heterogeneities among cells arise and how are they translated into divergent patterns of gene expression [38, 39].

Although two-cell blastomeres are generally considered to be totipotent, the previous reports showing that when two-cell blastomeres are separated, in the majority of cases, only one of the two-cell blastomeres develops into a mouse [40]. In this study, we provide evidence that HSPA2 is asymmetrically expressed already in the late 2-cell stage and acts as an upstream factor of CARM1 to drive first cell-fate decision. We found that the expression levels of HSPA2 were symmetrically distributed between the two blastomeres at the early 2-cell stage, but showed heterogeneous distribution at the late 2-cell embryos. This indicates that the heterogeneity of HSPA2 at the late 2-cell embryo may be initiated concomitant with the zygotic genome activation (ZGA) [6]. Alternatively, the asymmetry may be triggered by unequally distributed unknown factors that already existed in the early 2-cell embryos or the heterogeneous spatial distribution of maternal genes in the zygote.

The reduced expression of HSPA2 led to the reduced expression levels of ICM-marker genes at both the mRNA and protein levels. Interfering with HSPA2 expression in one blastomere of the 2-cell embryos reduced the contribution of that blastomere to ICM differentiation. In addition, the number of ICM and total cells in the blastocyst was significantly reduced when compared to the control group (Figure 3B, E). We observed that the GO enriched terms were also closely related to negative regulation of cell cycle, G1/S transition of mitotic cell cycle, mitotic cell cycle phase transition and regulation of mitotic cell cycle phase transition (Figure 2E). HSPA2

knockdown via siRNA reduced cell proliferation and led to G1/S phase cell cycle arrest [41]. It is possible that knockdown of HSPA2 results in lower cell number in ICM by interfering the cell cycle. Thus, HSPA2 has a role in ICM lineage establishment, although half of the ICM cells were able to survive with HSPA2 deficiency (Figure 3B).

However, the overexpression of HSPA2 did neither adversely affect embryo development and nor induce blastomere bias towards an ICM fate. It may be because that once the intracellular concentration of *Hspa2* reaches a threshold level, it was not able to control the target gene expression. Furthermore, we also found that HSPA2 interacted with CARM1, and reduced HSPA2 level, led to the reduced expression of CARM1, and thus reduced the levels of H3R26me2 modification at the 4-cell stage. Collectively, these results suggest that HSPA2 physically binds to CARM1 to activate ICM-specific gene expression through H3R26me2 modification (Figure 5I).

Several probable mechanisms are compatible with the asymmetric hypothesis, including polarized cell division and differential gene regulation by signaling [42]. In fact, knock-down of HSPA2 also affected the Hippo and Wnt signaling pathways (Figure 2E). The Hippo signaling pathway has been shown to be responsible for polar cells being specified as ICM, whilst polar cells are specified as TE [25]. Besides, the Wnt signaling pathway directs cell proliferation, cell polarity, and cell fate determination during embryonic development [43].

In summary, our findings provide an exciting step forward for the first time that the role of HSPA2 in governing the first cell-fate decision in the mouse embryo

lineage allocation. It may be beneficial for achieving a better understanding of embryogenesis patterns and cell differentiations.

Compliance with Ethical Standards

The current study was carried out with the approval of the Ethic Committee of Reproductive Medicine of Reproductive Hospital of Shandong University (2021-15), and all experiments performed were abided by relevant regulations and guidelines of the committees.

Acknowledgements

This work was supported by National Natural Science Foundation of China (32170817); The National Key R&D Program of China (2021YFC27003); The Fundamental Research Funds for the Central Universities (2022JC006); Taishan Scholars Program of Shandong Province (tsqn201909194, tsqn202211373); Innovative research team of high-level local universities in Shanghai (SHSMU-ZLCX20210201); The Fundamental Research Funds of Shandong University (2023QNTD004); Shandong Provincial Natural Science Foundation (2022HWYQ-034, ZR2021QC106)

Conflict of interest

The authors declare that they have no conflict of interests.

Author contribution

Jiayin Gao wrote the first draft of the manuscript. Jiawei Wang modified the language. All authors participated in literature research, conception, design, and discussion of the manuscript, and approved its final version.

References

[1] K. Cockburn, J. Rossant, Making the blastocyst: lessons from the mouse, J Clin Invest

492 120(4) (2010) 995-1003.

493 [2] J. Rossant, P.P. Tam, Blastocyst lineage formation, early embryonic asymmetries and axis

494 patterning in the mouse, Development 136(5) (2009) 701-13.

495 [3] B. Liu, H. Zhao, K. Wu, J. Grosshans, Temporal Gradients Controlling Embryonic Cell

496 Cycle, Biology (Basel) 10(6) (2021).

497 [4] M. Goolam, A. Scialdone, S.J.L. Graham, I.C. Macaulay, A. Jedrusik, A. Hupalowska, T.

498 Voet, J.C. Marioni, M. Zernicka-Goetz, Heterogeneity in Oct4 and Sox2 Targets Biases Cell

499 Fate in 4-Cell Mouse Embryos, Cell 165(1) (2016) 61-74.

500 [5] M.E. Torres-Padilla, D.E. Parfitt, T. Kouzarides, M. Zernicka-Goetz, Histone arginine

501 methylation regulates pluripotency in the early mouse embryo, Nature 445(7124) (2007) 214-

502 8.

503 [6] J. Shi, Q. Chen, X. Li, X. Zheng, Y. Zhang, J. Qiao, F. Tang, Y. Tao, Q. Zhou, E. Duan,

504 Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo

505 development revealed by single-cell RNA-seq, Development 142(20) (2015) 3468-77.

506 [7] F. Sotomayor-Lugo, N. Iglesias-Barrameda, Y.M. Castillo-Aleman, I. Casado-Hernandez,

507 C.A. Villegas-Valverde, A.A. Bencomo-Hernandez, Y. Ventura-Carmenate, R.A. Rivero-

508 Jimenez, The Dynamics of Histone Modifications during Mammalian Zygotic Genome

509 Activation, Int J Mol Sci 25(3) (2024).

510 [8] J. Wang, L. Wang, G. Feng, Y. Wang, Y. Li, X. Li, C. Liu, G. Jiao, C. Huang, J. Shi, T.

511 Zhou, Q. Chen, Z. Liu, W. Li, Q. Zhou, Asymmetric Expression of LincGET Biases Cell Fate in

512 Two-Cell Mouse Embryos, Cell 175(7) (2018) 1887-1901 e18.

513 [9] Y. Wang, X. Zheng, R. Cheng, J. Han, X. Ma, W. Xu, L. Gao, A. Lei, J. Liu, F. Quan, Y.

514 Zhang, X. Liu, Asymmetric expression of maternal mRNA governs first cell-fate decision,
515 FASEB J 35(11) (2021) e22006.
516 [10] D. Scieglinska, Z. Krawczyk, Expression, function, and regulation of the testis-enriched
517 heat shock HSPA2 gene in rodents and humans, Cell Stress Chaperones 20(2) (2015) 221-
518 35.
519 [11] M.J. Gomez-Torres, N. Huerta-Retamal, P. Saez-Espinosa, L. Robles-Gomez, M. Aviles,
520 J. Aizpurua, Molecular Chaperone HSPA2 Distribution During Hyaluronic Acid Selection in
521 Human Sperm, Reprod Sci 30(4) (2023) 1176-1185.
522 [12] H. Zhang, W. Chen, C.J. Duan, C.F. Zhang, Overexpression of HSPA2 is correlated with
523 poor prognosis in esophageal squamous cell carcinoma, World J Surg Oncol 11 (2013) 141.
524 [13] L.L. Bonnycastle, C.E. Yu, C.R. Hunt, B.J. Trask, K.P. Clancy, J.L. Weber, D. Patterson,
525 G.D. Schellenberg, Cloning, sequencing, and mapping of the human chromosome 14 heat
526 shock protein gene (HSPA2), Genomics 23(1) (1994) 85-93.
527 [14] D.J. Dix, J.W. Allen, B.W. Collins, C. Mori, N. Nakamura, P. Poorman-Allen, E.H.
528 Goulding, E.M. Eddy, Targeted gene disruption of Hsp70-2 results in failed meiosis, germ cell
529 apoptosis, and male infertility, Proc Natl Acad Sci U S A 93(8) (1996) 3264-8.
530 [15] A.K. Murashov, D.J. Wolgemuth, Distinct transcripts are recognized by sense and
531 antisense riboprobes for a member of the murine HSP70 gene family, HSP70.2, in various
532 reproductive tissues, Mol Reprod Dev 43(1) (1996) 17-24.
533 [16] W. Rupik, A. Stawierej, I. Stolarczyk, W. Widlak, Promoter of the heat shock testis-
534 specific Hsp70.2/Hst70 gene is active in nervous system during embryonic development of
535 mice, Anat Embryol (Berl) 211(6) (2006) 631-8.

536 [17] M. Garg, D. Kanojia, S. Saini, S. Suri, A. Gupta, A. Surolia, A. Suri, Germ cell-specific
537 heat shock protein 70-2 is expressed in cervical carcinoma and is involved in the growth,
538 migration, and invasion of cervical cells, Cancer 116(16) (2010) 3785-96.
539 [18] M. Garg, D. Kanojia, A. Seth, R. Kumar, A. Gupta, A. Surolia, A. Suri, Heat-shock protein
540 70-2 (HSP70-2) expression in bladder urothelial carcinoma is associated with tumour
541 progression and promotes migration and invasion, Eur J Cancer 46(1) (2010) 207-15.
542 [19] J. Song, J. Zhang, X. Yuan, B. Liu, W. Tao, C. Zhang, K. Wu, Functional substitution of
543 zona pellucida with modified sodium hyaluronate gel in human embryos, J Assist Reprod
544 Genet 39(11) (2022) 2669-2676.
545 [20] G.G. Schiattarella, F. Altamirano, D. Tong, K.M. French, E. Villalobos, S.Y. Kim, X. Luo,
546 N. Jiang, H.I. May, Z.V. Wang, T.M. Hill, P.P.A. Mammen, J. Huang, D.I. Lee, V.S. Hahn, K.
547 Sharma, D.A. Kass, S. Lavandero, T.G. Gillette, J.A. Hill, Nitrosative stress drives heart
548 failure with preserved ejection fraction, Nature 568(7752) (2019) 351-356.
549 [21] S. Picelli, O.R. Faridani, A.K. Bjorklund, G. Winberg, S. Sagasser, R. Sandberg, Full-
550 length RNA-seq from single cells using Smart-seq2, Nat Protoc 9(1) (2014) 171-81.
551 [22] B. Swierczek-Lasek, D. Dudka, D. Bauer, T. Czajkowski, K. Ilach, W. Streminska, A.
552 Kominek, K. Piwocka, M.A. Ciemerych, K. Archacka, Comparison of Differentiation Pattern
553 and WNT/SHH Signaling in Pluripotent Stem Cells Cultured under Different Conditions, Cells
554 10(10) (2021).
555 [23] D. ten Berge, D. Kurek, T. Blauwkamp, W. Koole, A. Maas, E. Eroglu, R.K. Siu, R. Nusse,
556 Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells, Nat
557 Cell Biol 13(9) (2011) 1070-5.

558 [24] A. de Jaime-Soguero, W.A. Abreu de Oliveira, F. Lluis, The Pleiotropic Effects of the
559 Canonical Wnt Pathway in Early Development and Pluripotency, Genes (Basel) 9(2) (2018).
560 [25] N. Nishioka, K. Inoue, K. Adachi, H. Kiyonari, M. Ota, A. Ralston, N. Yabuta, S. Hirahara,
561 R.O. Stephenson, N. Ogonuki, R. Makita, H. Kurihara, E.M. Morin-Kensicki, H. Nojima, J.
562 Rossant, K. Nakao, H. Niwa, H. Sasaki, The Hippo signaling pathway components Lats and
563 Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass, Dev Cell
564 16(3) (2009) 398-410.
565 [26] M.R. Stallcup, J.H. Kim, C. Teyssier, Y.H. Lee, H. Ma, D. Chen, The roles of protein-
566 protein interactions and protein methylation in transcriptional activation by nuclear receptors
567 and their coactivators, J Steroid Biochem Mol Biol 85(2-5) (2003) 139-45.
568 [27] N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, Functional insights
569 from structures of coactivator-associated arginine methyltransferase 1 domains, EMBO J
570 26(20) (2007) 4391-401.
571 [28] A. Hupalowska, A. Jedrusik, M. Zhu, M.T. Bedford, D.M. Glover, M. Zernicka-Goetz,
572 CARM1 and Paraspeckles Regulate Pre-implantation Mouse Embryo Development, Cell
573 175(7) (2018) 1902-1916 e13.
574 [29] M.R. Pawlak, C.A. Scherer, J. Chen, M.J. Roshon, H.E. Ruley, Arginine N-
575 methyltransferase 1 is required for early postimplantation mouse development, but cells
576 deficient in the enzyme are viable, Mol Cell Biol 20(13) (2000) 4859-69.
577 [30] D.E. Parfitt, M. Zernicka-Goetz, Epigenetic modification affecting expression of cell
578 polarity and cell fate genes to regulate lineage specification in the early mouse embryo, Mol
579 Biol Cell 21(15) (2010) 2649-60.

580 [31] M.H. Johnson, C.A. Ziomek, Induction of polarity in mouse 8-cell blastomeres: specificity,
581 geometry, and stability, J Cell Biol 91(1) (1981) 303-8.
582 [32] E. Korotkevich, R. Niwayama, A. Courtois, S. Friese, N. Berger, F. Buchholz, T. Hiiragi,
583 The Apical Domain Is Required and Sufficient for the First Lineage Segregation in the Mouse
584 Embryo, Dev Cell 40(3) (2017) 235-247 e7.
585 [33] Y. Hirate, H. Sasaki, The role of angiomotin phosphorylation in the Hippo pathway during
586 preimplantation mouse development, Tissue Barriers 2(1) (2014) e28127.
587 [34] K. Piotrowska, M. Zernicka-Goetz, Role for sperm in spatial patterning of the early mouse
588 embryo, Nature 409(6819) (2001) 517-21.
589 [35] R. Xu, C. Li, X. Liu, S. Gao, Insights into epigenetic patterns in mammalian early embryos,
590 Protein Cell 12(1) (2021) 7-28.
591 [36] W. Cui, A. Cheong, Y. Wang, Y. Tsuchida, Y. Liu, K.D. Tremblay, J. Mager, MCRS1 is
592 essential for epiblast development during early mouse embryogenesis, Reproduction 159(1)
593 (2020) 1-13.
594 [37] A. Burton, J. Muller, S. Tu, P. Padilla-Longoria, E. Guccione, M.E. Torres-Padilla, Single-
595 cell profiling of epigenetic modifiers identifies PRDM14 as an inducer of cell fate in the
596 mammalian embryo, Cell Rep 5(3) (2013) 687-701.
597 [38] I. Chambers, J. Silva, D. Colby, J. Nichols, B. Nijmeijer, M. Robertson, J. Vrana, K. Jones,
598 L. Grotewold, A. Smith, Nanog safeguards pluripotency and mediates germline development,
599 Nature 450(7173) (2007) 1230-4.
600 [39] A.M. Singh, T. Hamazaki, K.E. Hankowski, N. Terada, A heterogeneous expression
601 pattern for Nanog in embryonic stem cells, Stem Cells 25(10) (2007) 2534-42.

[40] E. Casser, S. Israel, A. Witten, K. Schulte, S. Schlatt, V. Nordhoff, M. Boiani, Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos, Sci Rep 7(1) (2017) 8299.

[41] L. Cao, X. Yuan, F. Bao, W. Lv, Z. He, J. Tang, J. Han, J. Hu, Downregulation of HSPA2 inhibits proliferation via ERK1/2 pathway and endoplasmic reticular stress in lung adenocarcinoma, Ann Transl Med 7(20) (2019) 540.

[42] F.H. Biase, X. Cao, S. Zhong, Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing, Genome Res 24(11) (2014) 1787-96.

[43] S.A. Morris, Y. Guo, M. Zernicka-Goetz, Developmental plasticity is bound by pluripotency and the Fgf and Wnt signaling pathways, Cell Rep 2(4) (2012) 756-65.

Figure caption

Figure 1. *Hspa2* is asymmetrically expressed between mouse 2- and 4-cell blastomeres. (A) Schematic overview. The 4-cell embryo divided into single blastomeres and cultured to the morula stage. Created with biorender.com. **(B)** *Hspa2* expression is heterogeneous in morula developed from 4-cell blastomeres. There experimental replicates were performed. **(C)** *Nanog* and *Hspa2* show stronger positive correlation (R=0.624) in gene expression than random gene pairs (n=100, n means the number of random gene pairs) **(D and G)** Single-cell qRT-PCR results showing *Hspa2* distribution between 4- (D) (n=15), late 2- (G) (n=17) cell blastomere. *Hprt* as the housekeeping gene. CT values are used for level analysis. Two-tailed Student's t tests were used for the statistical analysis. Blastomeres are named B1 to B4 (D) or B1 and B2 (G) according to increasing *Hspa2* concentration. $*p < 0.05$. **(E and H)** Automated western immunoblotting results showing

HSPA2 distribution between 4- (E)and late 2- (H) cell blastomeres. Blastomeres are named B1 to B4 (E) or B1 and B2 (H) according to decreasing *Hspa2* concentration. The band intensity was assessed. There experimental replicates were performed. *p <0.05. **(F and I)** FISH of *Hspa2* mRNA in early 4- (F) (n=17), late 2- (I) (n=22) cell embryos. Blastomeres with low expression as white triangles. RNA-FISH fluorescence quantified and normalized to the nucleus with the strongest staining per individual two- or four-cell embryo. Blastomeres are named B1 and B2 (G) or B1 to B4 (D) according to increasing *Hspa2* concentration. Three experimental replicates were performed. Bar, 25 μm. *p <0.05.

Figure 2. Knockdown of *Hspa2* lead to decreased expression of ICM-mark at both the mRNA and protein level. (A) Schematic diagram of microinjection of zygotic embryos. Created with biorender.com. (B) Bar charts showing the expression of *Hspa2* in the negative control (NC) group (injected with negative control siRNA) and siRNA targeting group based on RNA-seq. The error bars denote the standard deviations of three experimental replicates of RNA-seq. *p <0.05. **(C)** Heat map of the hierarchical clustering analysis of the DEGs between NC and KD groups. **(D)** Volcano plots of significantly differentially expressed genes (DEGs). Red, upregulated; Blue, down-regulated**. (E)** GO enrichment analysis from down-regulated DEGs between NC and KD groups. **(F)** Significantly down-regulated DEGs were analyzed by KEGG enrichment. **(G)** Representative images of preimplantation embryos in the NC and KD groups from 2-cell to blastocyst. Red arrowhead denotes degenerating blastocysts. Bar, 100 μm. **(H)** The blastocyst rate between NC (n=158) and KD (n=155) groups. n represents the number of blastocysts per experiment. Three experimental replicates were performed. *p <0.05. **(I and J)** qRT-PCR of 4-cell embryo and blastocyst showed that ICM-marker gene (*Oct4*, *Sox2* and *Nanog*) mRNA was

significantly lower in the interference compared with the NC group. There were no significant differences in Cdx2 between the two groups. *$p < 0.05$, **$p < 0.01$. **(K and L)** Automated western blotting analysis of ICM-marker protein (OCT4, SOX2 and NANOG) expression in the interference and NC groups. Three experimental replicates were performed. *$p < 0.05$. **(M)** Western blot analysis of ICM-marker protein (OCT4, SOX2 and NANOG) and TE-marker protein (CDX2) expression in mESCs. The band intensity was assessed with ImageJ. Three experimental replicates were performed. *$p < 0.05$.

Figure 3. Knockdown of *Hspa2* prevents blastomeres from undergoing an ICM fate. (A) Immunofluorescence of OCT4 (green) and CDX2 (red) in NC (n=40) and si-*Hspa2* (n=36) groups. OCT4 was used as an ICM marker. Nuclei were visualized with DAPI staining (blue). Bar, 50 μm. **(B)** The number of ICM and TE. OCT4 was used as an ICM marker and CDX2 was used as a TE marker. *$p < 0.05$. **(C)** Schematic diagram of microinjection of early 2-cell embryo single blastomeres. Created with biorender.com. **(D)**OCT4 and GFP fluorescent staining of blastocysts. OCT4 was used as an ICM marker. The results show that less OCT4-positive cells are GFP-positive in embryos injected with GFP and *Hspa2 siRNA* (GFP+*si-Hspa2*). Three experimental replicates were performed. Bar, 50 μm. **(E)** Analysis of the distribution of progeny of injected blastomere at the blastocyst stage based on fluorescence staining. Total cells (n) are the total number of cells in the blastocyst; inner cells (%) is the percentage of inner cells relative to the total number of cells in the blastocyst; GFP cells that are inner (%) is the percentage of GFP-positive inner cells relative to the total number of GFP-positive cells in the blastocyst. Data derived from three independent experiments and presented as mean ± standard deviation. Three experimental replicates were performed. *$p < 0.05$.

Figure 4. Overexpression of *Hspa2* not induces blastocyst bias towards ICM fate. (A) Representative images of preimplantation embryos in the NC and *Hspa2* mRNA groups from 2-cell to blastocyst. Bar, 100 μm. **(B)** The blastocyst rate between NC (n=151) and overexpression (n=150) groups. Three experimental replicates were performed. **(C and D)** qRT-PCR of 4-cell embryo and blastocyst showed that ICM-marker gene (*Oct4*, *Sox2* and *Nanog*) and TE-marker gene (*Cdx2*) were no significant differences in the overexpression compared with the control group. Three experimental replicates were performed. $*p < 0.05$, $****p < 0.001$. **(E)** Immunofluorescence of OCT4 (green) and CDX2 (purple) in NC (n=38) and overexpression (n=40) groups. OCT4 was used as an ICM marker. Nuclei were visualized with DAPI staining (blue). Bar, 50 μm. **(F)** The number of ICM and TE. OCT4 was used as an ICM marker and CDX2 was used as a TE marker. **(G)** OCT4 and GFP fluorescent staining of blastocysts. OCT4 was used as an ICM marker. The results show that the percentage of GFP cells that were present in the ICM was similar between the two groups. Three experimental replicates were performed. Three experimental replicates were performed. Bar, 50 μm. **(H)** Analysis of the distribution of progeny of injected blastomere at the blastocyst stage based on fluorescence staining. Total cells (n) are the total number of cells in the blastocyst; inner cells (%) is the percentage of inner cells relative to the total number of cells in the blastocyst; GFP cells that are inner (%) is the percentage of GFP-positive inner cells relative to the total number of GFP-positive cells in the blastocyst. Data derived from three independent experiments and presented as mean ± standard deviation. Three experimental replicates were performed.

Figure 5. HSPA2 and CARM1 form a complex. (A) Molecular docking analysis for HSPA2 binding to CARM1 proteins. **(B and C)** qRT-PCR of 4-cell embryo and blastocyst showed that

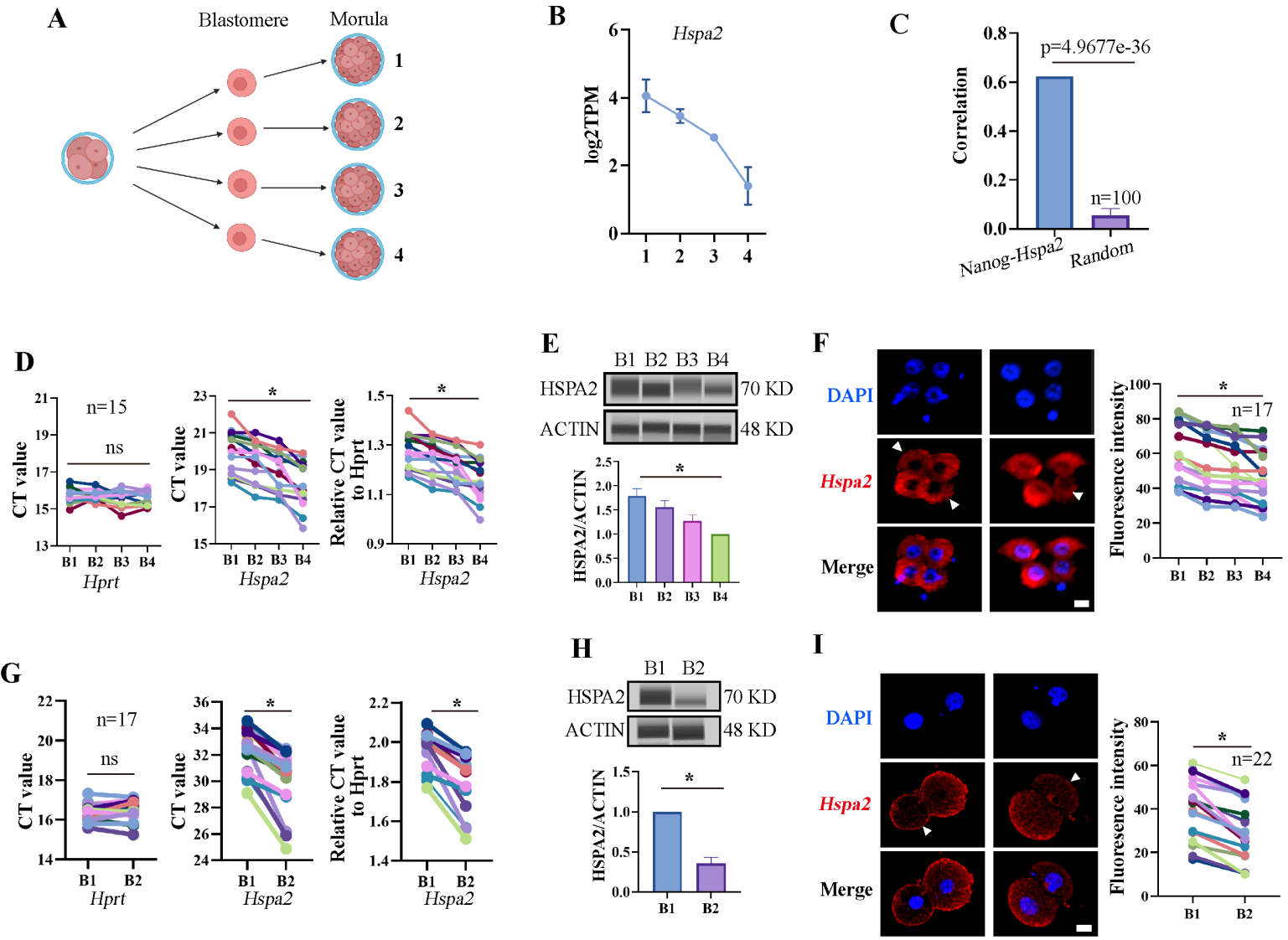
Carm1 was significantly lower in the interference compared with the control group. Three experimental replicates were performed. *$p < 0.05$. **(D and E)** Automated western blotting analysis of CARM1 expression in the interference and control groups. The band intensity was assessed. Three experimental replicates were performed. *$p < 0.05$ **(F and G)** CO-IP analysis of HSPA2 and CARM1 in mESCs. **(H)** H3R26me2 staining of 4-cell embryo shows that HSPA2 knockdown led to a dramatic decrease in H3R26me2 modification. Intensity relative to DNA signal was used. Three experimental replicates were performed. Bar, 50 μm.*$p < 0.05$. (I) The model of HSPA2 physically binds to CARM1 to activate ICM-specific gene expression.

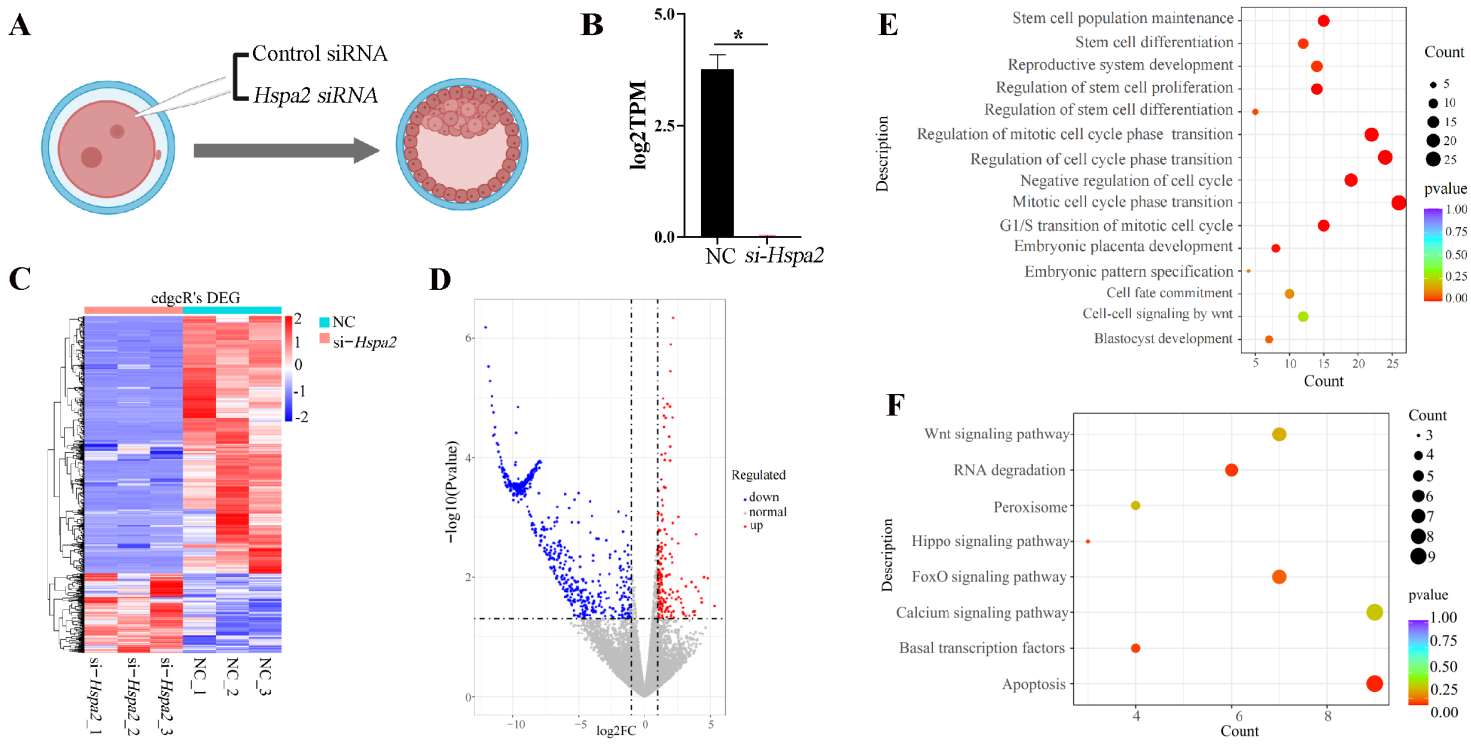
Supplemental Figure caption

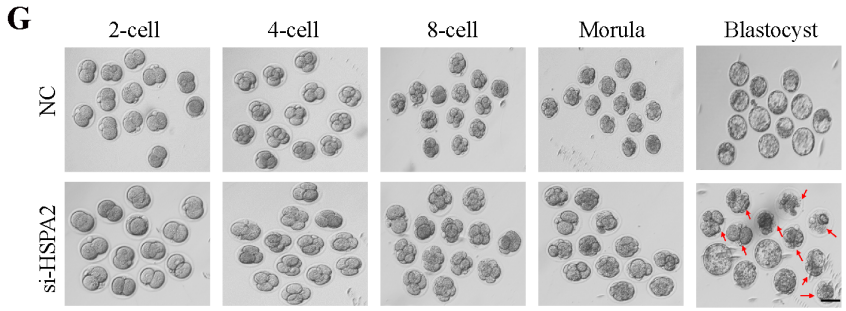
Figure S1. *Hspa2* is asymmetrically expressed between mouse 2- and 4-cell blastomeres, Related to Figure 1. (A) Schematic overview. Late 2-cell divided into single blastomeres and cultured to the morula stage. Created with biorender.com. **(B)** *Hspa2* expression is heterogeneous and the level of *Nanog* is positively correlated with the expression of *Hspa2* in morula developed from late 2-cell blastomeres. Three experimental replicates were performed. **(C)**qPCR analysis of the relative expression of *Hspa2* mRNA form the meiosis II to blastocyst stages. The error bars represent SEM. About 20 embryos of each stage were used and three experimental replicates were performed. **(D)** Single-cell qRT-PCR results showing *Hspa2* distribution between early 2-cell (n=17) blastomere. *Hprt* as the housekeeping gene. CT values are used for level analysis. Two-tailed Student's t tests were used for the statistical analysis. Blastomeres are named B1 and B2 according to increasing *Hspa2* concentration. Three experimental replicates were performed. **(E)** Automated western immunoblotting results showing HSPA2 distribution between early 2-cell blastomeres. Blastomeres are abbreviated named B. The protein level is normalized B1. Three

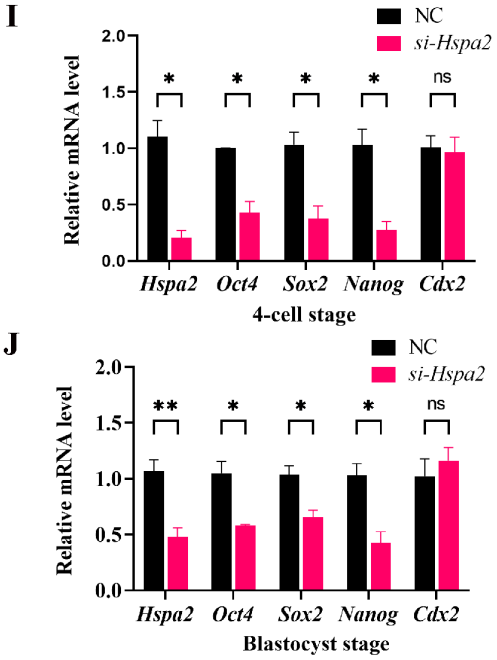
experimental replicates were performed. **(F)** FISH of *Hspa2* mRNA in early 2-cell embryos (n=40). RNA-FISH fluorescence quantified and normalized to the nucleus with the strongest staining per individual early 2-cell embryo. Three experimental replicates were performed. Bar, 25 μm.

Figure S2. HSPA2 and CARM1 form a complex. Related to Figure 5. (A) Volcano plot of the DEPs in *Hspa2* siRNA and control groups at blastocyst stage. **(B)** COG/KOG enrichment analysis for DEPs. **(C)** Quantitative classification and numbers of DEPs in each quantified part. The DEPs were divided into four quantified parts according to their differential expression multiples: Q1 (<0.5), Q2 (0.5-0.667), Q3 (1.5-2.0), and Q4 (>2.0). **(D and E)** Clustering analysis including biological process analysis (D) and cellular component (E) of quantitative DEPs in each Q group. **(F)** The Protein-protein interactions (PPIs) network was mapped by STRING database (https://cn.string-db.org/cgi/network?taskId=bk9UNYyN0p0f&sessionId=bS81YN78aR). **(G)** Western blot analysis of CARM1 expression in the knockdown and control groups. The band intensity was assessed with ImageJ. Three experimental replicates were performed. ***$p < 0.001$.


A
Blastomere
Morula
1
2
3
4
B
Hspa2
log2TPM
C
p=4.9677e-36
Correlation
n=100
Nanog-Hspa2
Random
D
n=15
ns
CT value
Hprt
Hspa2
Relative CT value to Hprt
E
B1 B2 B3 B4
HSPA2
70 KD
ACTIN
48 KD
HSPA2/ACTIN
F
DAPI
Hspa2
Merge
Fluoresence intensity
n=17
G
n=17
ns
H
B1 B2
I
n=22

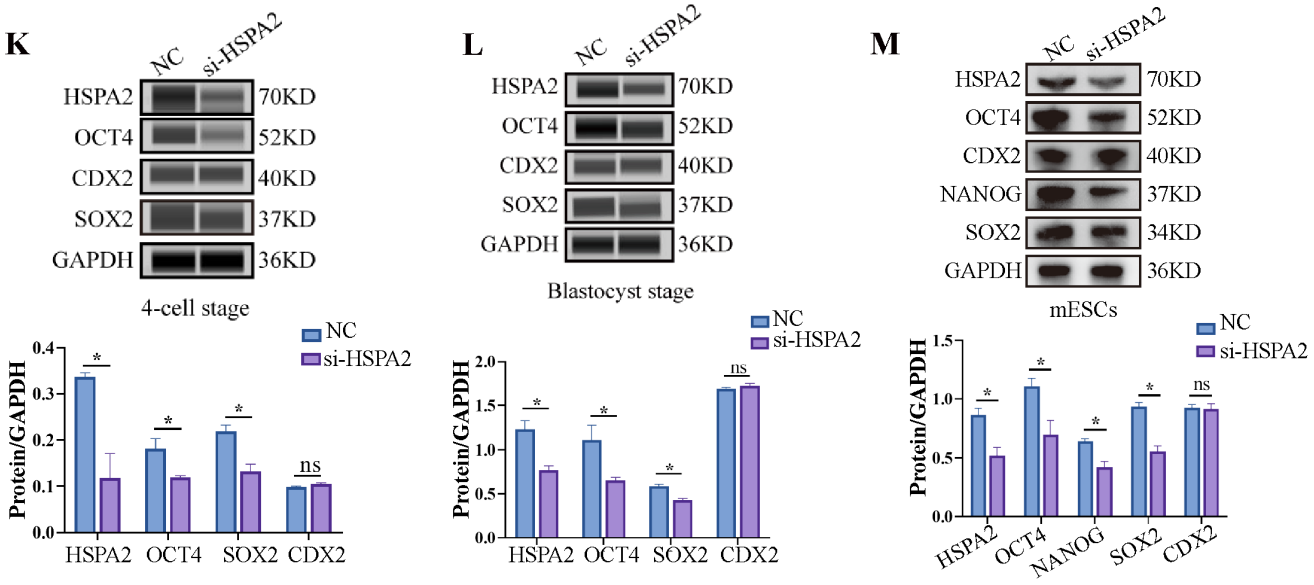




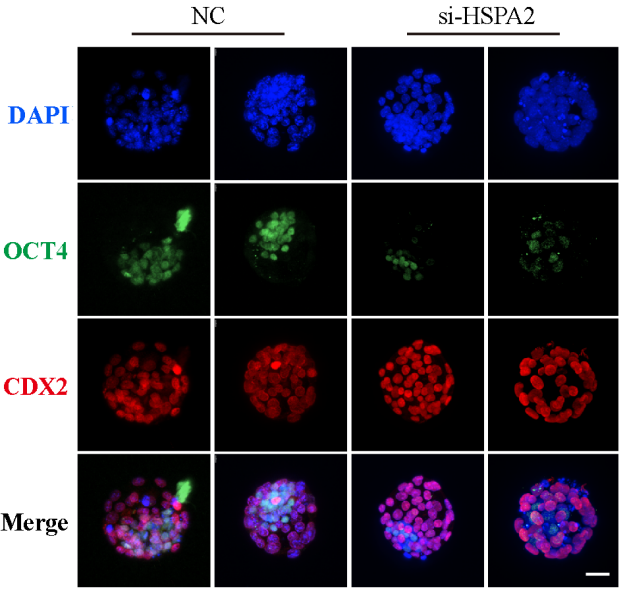


H

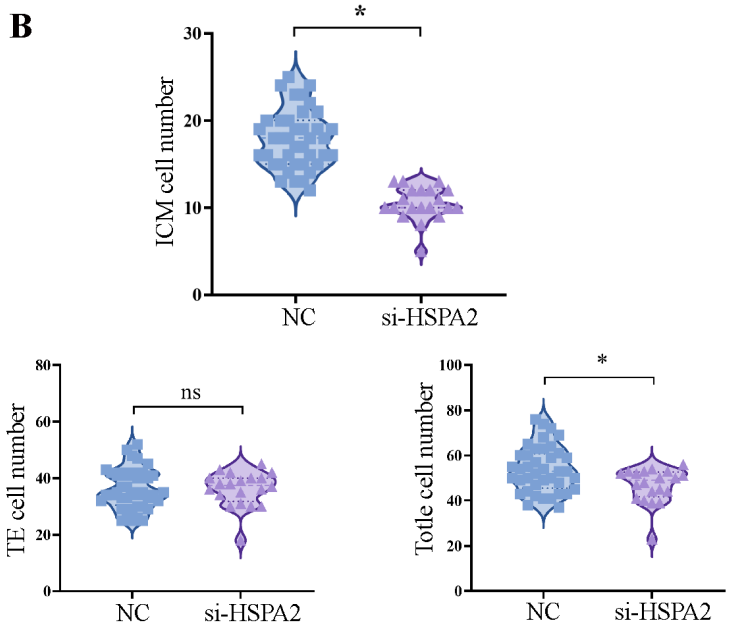
Group	Total Number of Zygote (Repeating count)	Blastocyst (n±S.E)	Blastocyst Rate (%±S.E)
NC	158(3)	129(43±1.155)	81.63±2.000
si-*Hspa2*	155(3)	82(27±1.155)	52.96±2.489*



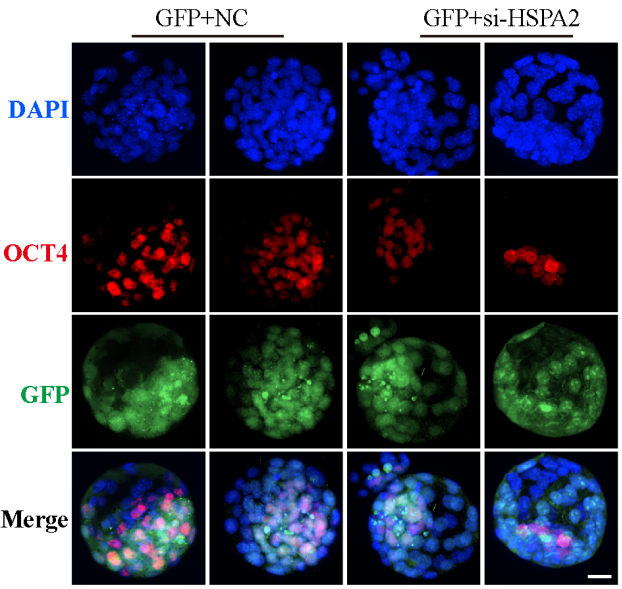
A



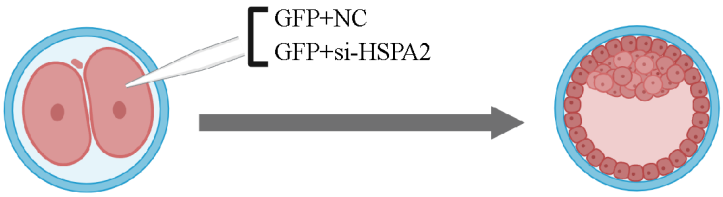
B



D

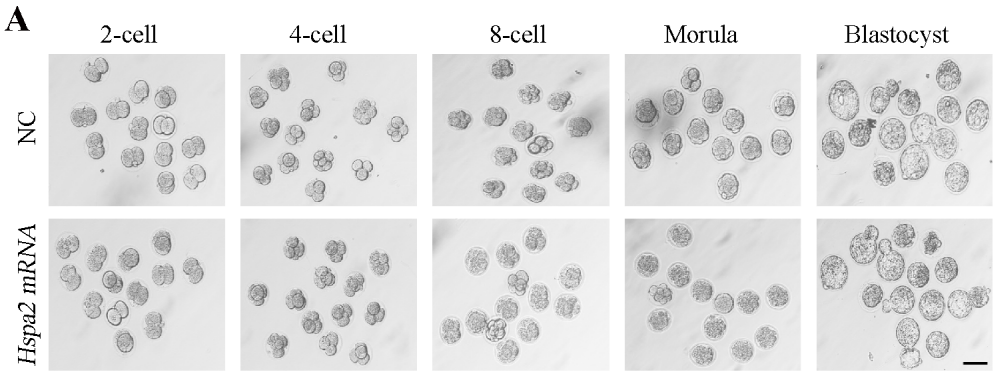


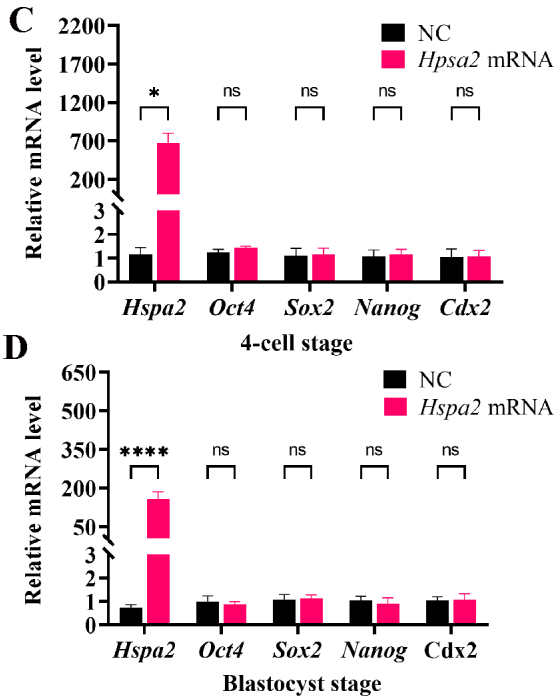
C



E

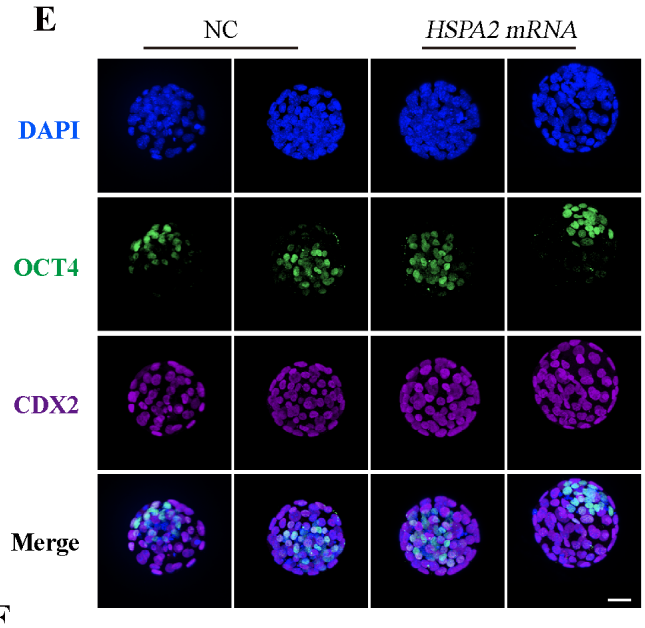
Group	Blastocysts (n)	Total cells (n)	Inner cells (%)	GFP cells that are inner (%)
GFP+NC	49	48.64±7.904	36.95±2.524	48.23±7.415
GFP+*si-Hspa2*	37	46.62±5.139	27.09±3.811*	33.75±4.303*

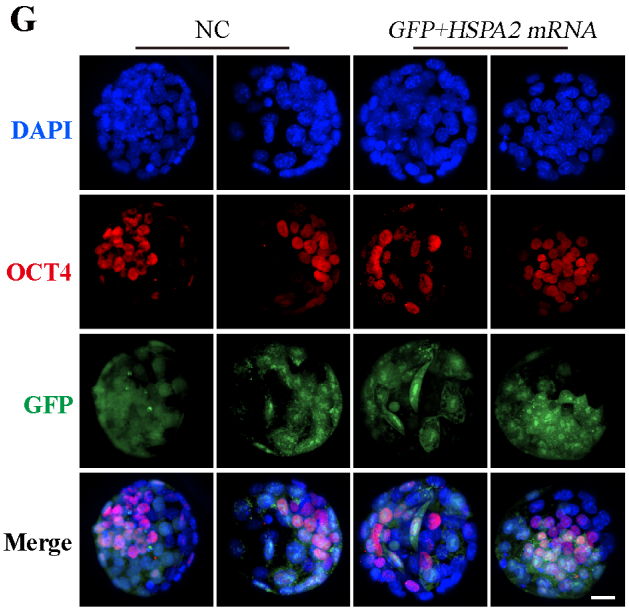


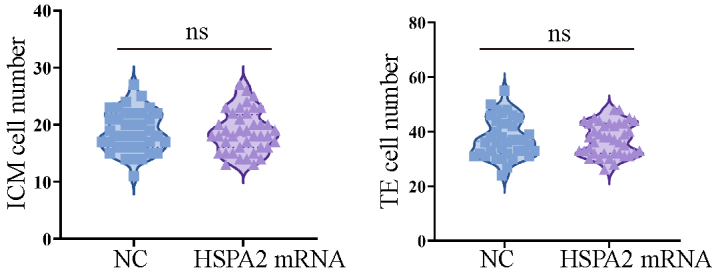


B

Group	Total Number of Zygote (Repeating count)	Blastocyst (n±S.E)	Blastocyst Rate (%±S.E)
NC	151(3)	126(42±1.529)	83.39±1.579
Hspa2 mRNA	150(3)	122(40.67±1.453)	81.49±2.136

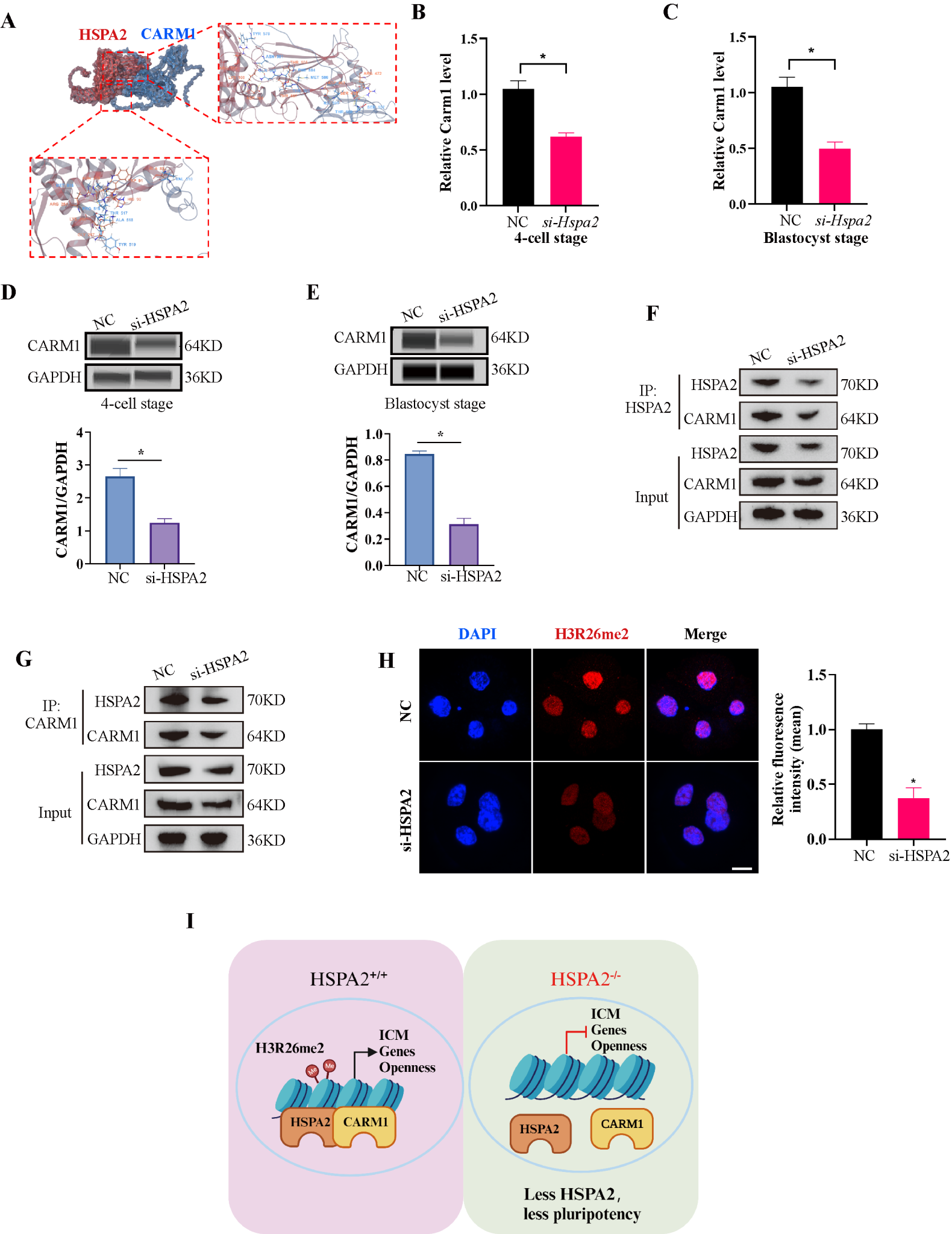


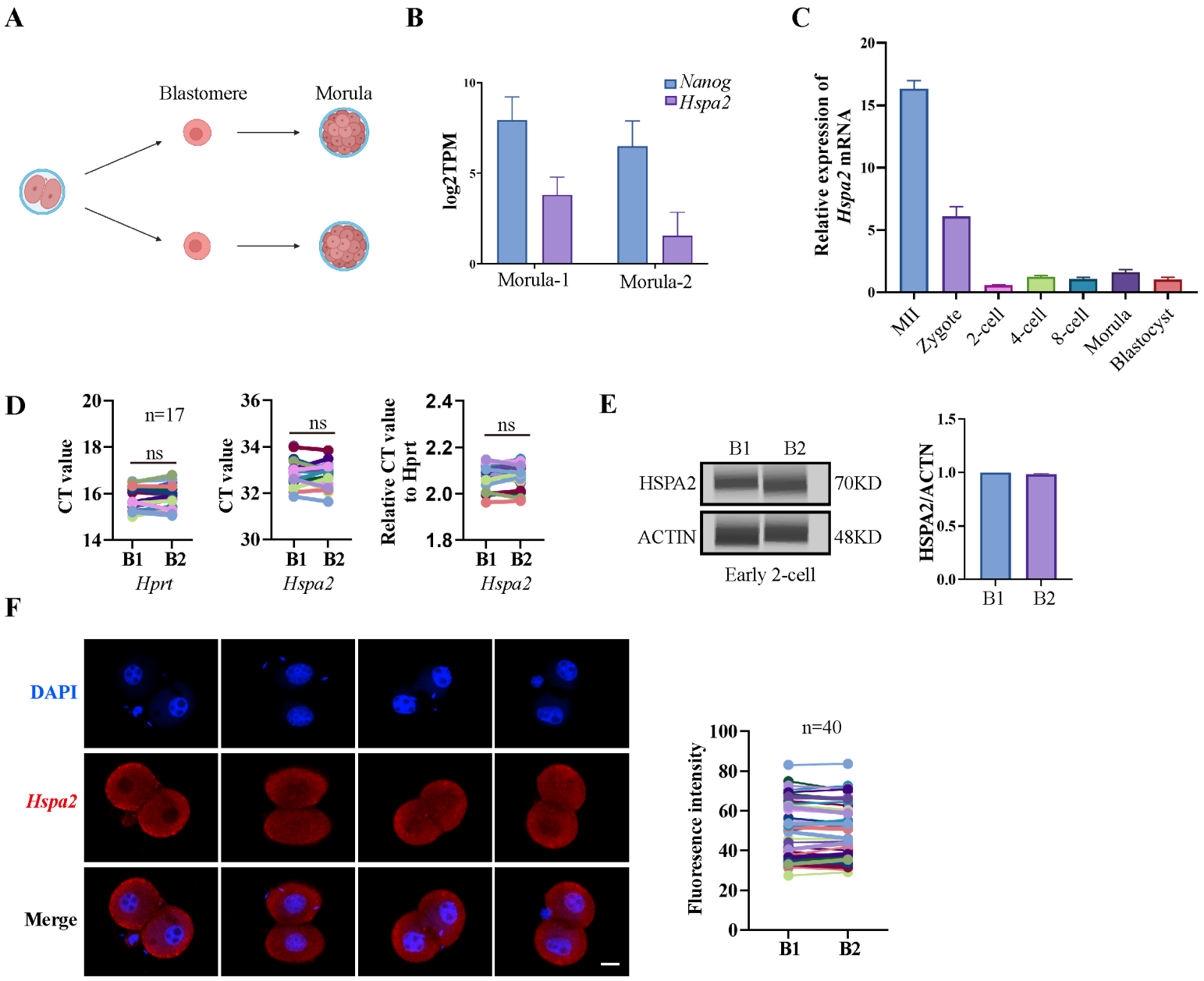


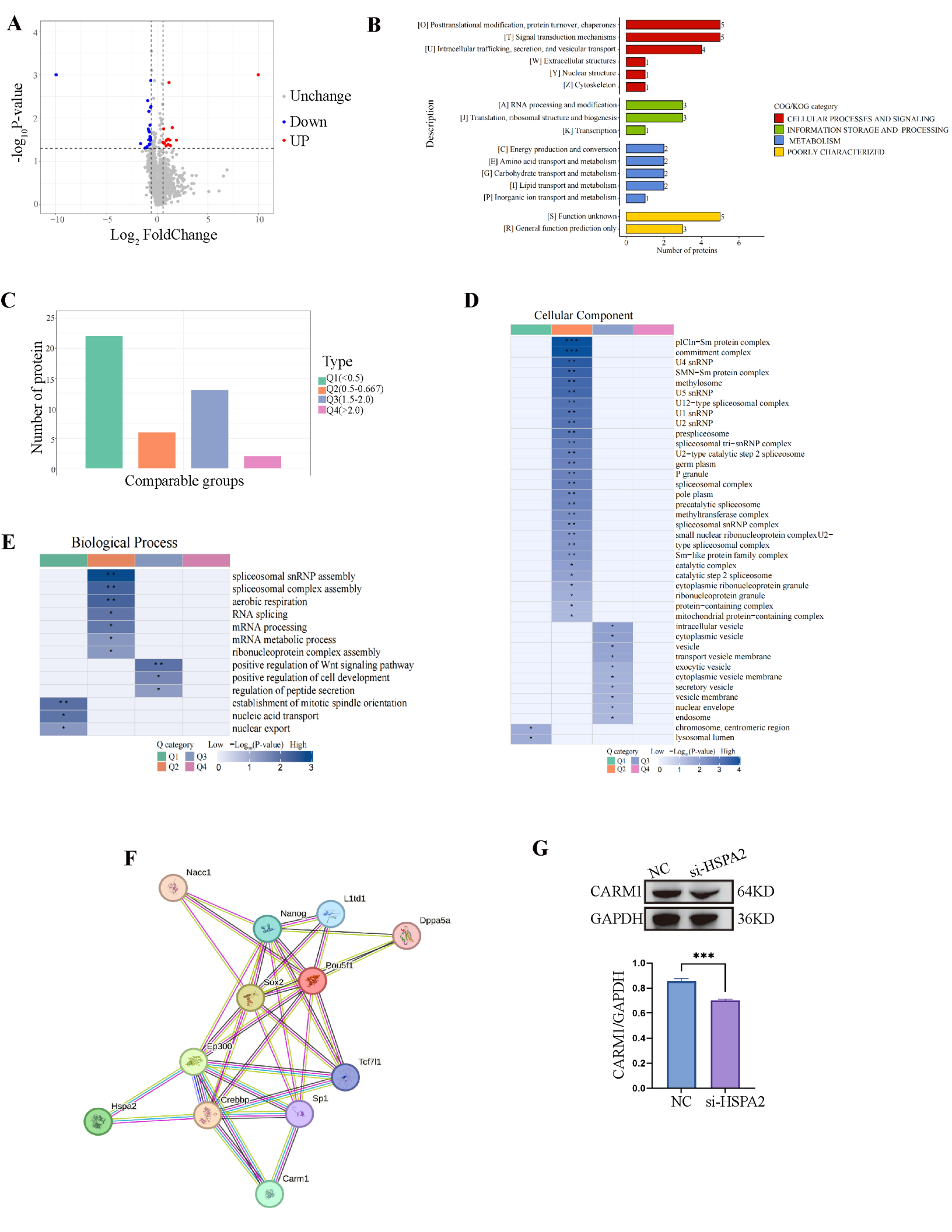


H

Group	Blastocysts (n)	Total cells (n)	Inner cells (%)	GFP cells that are inner (%)
GFP+NC	49	48.64±7.904	36.95±2.524	48.23±7.415
GFP+*Hspa2* mRNA	23	49.13±6.546	36.24±1.874	47.25±4.432


A
HSPA2
CARM1
B
Relative Carm1 level
1.5
1.0
0.5
0.0
*
NC
si-Hspa2
4-cell stage
C
Relative Carm1 level
1.5
1.0
0.5
0.0
*
NC
si-Hspa2
Blastocyst stage
D
NC
si-HSPA2
CARM1
64KD
GAPDH
36KD
4-cell stage
CARM1/GAPDH
4
3
2
1
0
*
NC
si-HSPA2
E
NC
si-HSPA2
CARM1
64KD
GAPDH
36KD
Blastocyst stage
CARM1/GAPDH
1.0
0.8
0.6
0.4
0.2
0.0
*
NC
si-HSPA2
F
NC
si-HSPA2
IP:
HSPA2
HSPA2
70KD
CARM1
64KD
Input
HSPA2
70KD
CARM1
64KD
GAPDH
36KD
G
NC
si-HSPA2
IP:
CARM1
HSPA2
70KD
CARM1
64KD
Input
HSPA2
70KD
CARM1
64KD
GAPDH
36KD
H
DAPI
H3R26me2
Merge
NC
si-HSPA2
Relative fluoresence intensity (mean)
1.5
1.0
0.5
0.0
*
NC
si-HSPA2
I
HSPA2+/+
H3R26me2
Me
Me
ICM
Genes
Openness
HSPA2
CARM1
HSPA2-/-
ICM
Genes
Openness
HSPA2
CARM1
Less HSPA2,
less pluripotency


A
Blastomere
Morula
B
Nanog
Hspa2
log2TPM
10
5
0
Morula-1
Morula-2
C
Relative expression of Hspa2 mRNA
20
15
10
5
0
MII
Zygote
2-cell
4-cell
8-cell
Morula
Blastocyst
D
n=17
ns
CT value
20
18
16
14
B1
B2
Hprt
ns
CT value
36
34
32
30
B1
B2
Hspa2
ns
Relative CT value to Hprt
2.4
2.2
2.0
1.8
B1
B2
Hspa2
E
B1
B2
HSPA2
70KD
ACTIN
48KD
Early 2-cell
HSPA2/ACTN
1.5
1.0
0.5
0.0
B1
B2
F
DAPI
Hspa2
Merge
n=40
Fluoresence intensity
100
80
60
40
20
0
B1
B2


A
-log10 P-value
Log2 FoldChange
Unchange
Down
UP
B
Description
Number of proteins
[O] Posttranslational modification, protein turnover, chaperones
[T] Signal transduction mechanisms
[U] Intracellular trafficking, secretion, and vesicular transport
[W] Extracellular structures
[Y] Nuclear structure
[Z] Cytoskeleton
[A] RNA processing and modification
[J] Translation, ribosomal structure and biogenesis
[K] Transcription
[C] Energy production and conversion
[E] Amino acid transport and metabolism
[G] Carbohydrate transport and metabolism
[I] Lipid transport and metabolism
[P] Inorganic ion transport and metabolism
[S] Function unknown
[R] General function prediction only
COG/KOG category
CELLULAR PROCESSES AND SIGNALING
INFORMATION STORAGE AND PROCESSING
METABOLISM
POORLY CHARACTERIZED
C
Number of protein
Comparable groups
Type
Q1(<0.5)
Q2(0.5-0.667)
Q3(1.5-2.0)
Q4(>2.0)
D
Cellular Component
pICln−Sm protein complex
commitment complex
U4 snRNP
SMN−Sm protein complex
methylosome
U5 snRNP
U12−type spliceosomal complex
U1 snRNP
U2 snRNP
prespliceosome
spliceosomal tri−snRNP complex
U2−type catalytic step 2 spliceosome
germ plasm
P granule
spliceosomal complex
pole plasm
precatalytic spliceosome
methyltransferase complex
spliceosomal snRNP complex
small nuclear ribonucleoprotein complex U2−type spliceosomal complex
Sm−like protein family complex
catalytic complex
catalytic step 2 spliceosome
cytoplasmic ribonucleoprotein granule
ribonucleoprotein granule
protein−containing complex
mitochondrial protein−containing complex
intracellular vesicle
cytoplasmic vesicle
vesicle
transport vesicle membrane
exocytic vesicle
cytoplasmic vesicle membrane
secretory vesicle
vesicle membrane
nuclear envelope
endosome
chromosome, centromeric region
lysosomal lumen
Q category
Low −Log10(P-value) High
E
Biological Process
spliceosomal snRNP assembly
spliceosomal complex assembly
aerobic respiration
RNA splicing
mRNA processing
mRNA metabolic process
ribonucleoprotein complex assembly
positive regulation of Wnt signaling pathway
positive regulation of cell development
regulation of peptide secretion
establishment of mitotic spindle orientation
nucleic acid transport
nuclear export
Q category
Low −Log10(P-value) High
F
Nacc1
L1td1
Nanog
Dppa5a
Pou5f1
Sox2
Ep300
Tcf7l1
Hspa2
Crebbp
Sp1
Carm1
G
NC
si-HSPA2
CARM1
GAPDH
64KD
36KD
***
CARM1/GAPDH
NC
si-HSPA2