Rapid production of modified cattle and sheep from spermatid-like stem cells

Lei Yang[1, 2, 4, 5, *], Anqi Di[1, 4], Lishuang Song[1, 4], Xuefei Liu[1], Di Wu[1], Song Wang[1, 3], Zhenting Hao[1], Lige Bu[1], Chunling Bai[1], Guanghua Su[1], Zhuying Wei[1], Li Zhang[1], Zhonghua Liu[3], Shaorong Gao[2, 5, *], and Guangpeng Li[1, 5, *]

1 State Key Laboratory of Reproductive Regulation and Breeding of Grassland Livestock, College of Life Sciences, Inner Mongolia University, Hohhot, China.

2 Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai, China.

3 Key Laboratory of Animal Cellular and Genetics Engineering of Heilongjiang Province, College of Life Science, Northeast Agricultural University, Harbin, China

4 These authors contributed equally: Lei Yang, Anqi Di, Lishuang Song

5 These authors jointly supervised this work: Lei Yang, Shaorong Gao, Guangpeng Li

* e-mail: leiyang@imu.edu.cn; gaoshaorong@tongji.edu.cn; gpengli@imu.edu.cn

Abstract

Haploid androgenetic stem cells (haSCs) are revolutionary resources for studying animal traits and greatly impact livestock breeding, but livestock haSCs have yet to be obtained. Here, we report the derivation of cattle and sheep haSCs using a novel FACE medium, and these cells harbor formative pluripotency features. Notably, knockout cattle and sheep can be efficiently obtained upon injection of the protaminized haSCs with targeted gene mutations into oocytes.

Main

Haploid androgenetic stem cells (haSCs) can support the production of viable transgenic rodents (mice and rats) by oocyte intracytoplasmic haSCs injection (iCHI) [1-3]. Currently, somatic cell nuclear transfer (SCNT) allows CRISPR/Cas-edited cells to give rise to produce genetically modified livestock. However, the low efficiency of CRISPR-SCNT dramatically limits its application. Thus, CRISPR-iCHI holds excellent promise in agrobiotechnology, but livestock haSCs have not been achieved.

To this end, we generated bovine haploid androgenetic (ha) embryos by two different methods: (i) injecting sperm into enucleated oocytes; (ii) removing the female pronucleus from zygotes (Fig.1a, Supplementary Fig.1a). Expectedly, sperm underwent actively DNA demethylate in ha embryos [2] (Supplementary Fig.1b). The blastocyst rate of the injecting and removing approach was comparable (31.16% vs. 30.21%; Supplementary Fig.1d,e, Table 1), indicating that the reprogramming of sperm by ooplasm was not affected by the manipulations. The ha-blastocysts were further cultivated feeder cells under “naïve” or “primed” conditions with mouse, human, and bovine diploid SCs to derive bovine haSCs (b-haSCs; Supplementary Fig.2a). From 261 ha-blastocyst, we only obtained outgrowths consisting of “master-regulators” $OCT4^{+}SOX2^{+}$ in LCDM, CTFR, and bEPSC medium, but couldn’t establish stable cell lines (no more than eight passages; Supplementary Fig.2b-d).

To long-term maintenance of the b-haSCs, we screened the different combinations of inhibitor and cytokine in the above successful culturing medium (11 hits that target ten signal-pathways; Supplementary Fig.3a). Cells cultured in FACE [FGF2::Activin-A::CHIR99021::IWR-1-endo] and FACX [FAC::XAV939] combinations formed colonies, while in others appeared differentiated or did not form colonies (Supplementary Fig.3b,c). Both IWR-1-endo and XAV939 act as Wnt/β-catenin inhibitors and could be used interchangeably [4, 5], thus [FACX] and [FACE] are essentially the same medium. We further titrated Activin-A and obtained cell lines that could be propagated for more than 80 passages in low Activin-A, and haploid cells could be enriched by fluorescence-activated cell sorting (FACS; Fig.1b-d, Supplementary Fig.2e, 3d, 4a-c).

We also succeeded in establishing ovine haSCs (o-haSCs) by FACE medium, which expressed pluripotency marks similar to that of b-haSCs (Fig.1b-d, Supplementary Fig.1c-e, 2e, 4b,c, Table 1). Consistent with the rodent haSCs [1-3], no ruminant haSCs with the Y chromosome (Supplementary Fig.4b,d). Strikingly, originally obtained b- and o-haSCs contain nearly 20% haploid cells, which was higher than that of murine haSCs [1-3] (< 5%; Fig.1c). Overall, the novel FACE medium can be used for both b- and o-haSCs derivation.

Inner cell mass (ICM) of the blastocyst, during gastrulation, is through three pluripotency states known as naïve, formative, and primed states [6] (Supplementary Fig.6a). It is well known that, under the naïve state, mouse haSCs (m-haSCs) tend to rapidly diploidization [1, 2], whereas the ruminant haSCs do not. These results prompted us to determine the pluripotency state of ruminant haESCs. To this end, we analysed the transcriptome by bulk RNA sequencing (buRNA-seq) and found that the b-haSCs couldn't belong to any stage of pre-implantation embryos [7, 8] (Supplementary Fig.5a,b). Meanwhile, b-haSCs also segregated from the recently reported diploid primed and expanded SCs [9, 10] (Supplementary Fig.5c,d), implying that b-haSCs enriches the features of formative pluripotency (Supplementary Fig.5e).

To test this hypothesis, we performed single cell (sc) RNA-seq at the gastrulation period of bovine epiblast from *in vivo* diploid embryos (E8~14; Supplementary Fig.6b,c). Expectedly, b-haSCs were clustered with the embryos at E10 (Supplementary Fig.6b,c). Indeed, we identified three types of cells in b-haSCs: 83.8% formative, 12.1% primed, and a few 4.1% naïve (Fig.1e,f, Supplementary Fig.7a,b). To further dissect the epigenetic features, we performed CUT&Tag assay (Supplementary Fig.8a). The H3K4me3 and H3K27me3 at the promoter of formative genes showed higher and lower in b-haSCs than primed b-SCs, respectively. Also, these modifications of primed genes exhibited the opposite trend (Fig.1g, Supplementary Fig.8b). In agreement, the H3K4me3 and H3K27me3 clustered profiles segregated from primed b-SCs (Supplementary Fig.8c). The presence of H3K4me3 but not H3K27me3 in the promoter of basic-pluripotency genes, including OCT4, SOX2, and SALL4, is common to both b-haSCs and primed diploid b-SCs (Supplementary Fig.8d).

We further undertook a cross-species comparison, and found that o-haSCs clustered closer with b-haSCs than with primed o-SCs (Supplementary Fig.7c). Formative-related regulators were both present in bovine, ovine, and murine formative cells (Fig.1h). The mRNA and protein levels of the key formative markers, including DNMT3A, DNMT3B, and OTX2, were comparable among each species (Supplementary Fig.7d,e). We thus conclude that b- and o-haSCs harbor molecular features of formative pluripotency.

Expectedly, both b- and o-haSCs could differentiate into three germ-line with around 16% haploid cells both *in vitro* and *in vivo*, which is much higher than that of m-haSCs [1] (Supplementary Fig.9a-d, 10a-d). Both b- and o-haSCs could localised to ICM positions of the intra-/inter-species chimeric blastocysts (Supplementary Fig.11a-d). Following embryo transfer into the foster mouse, we detected b-haSCs (but not o-haSCs) gave rise to live inter-species chimeras at E18.5 (albino mouse for the host; Fig.1i, Supplementary Fig.12a). However, cattle-mouse chimera only reached 2-day-

105 old with colored eyes (Supplementary Fig.12b,c), suggesting that xenogeneic cells
106 cannot maintain the metabolic environment for each other.
107
108 We next performed the iCHI experiment to replace sperm, by injecting the b- and o-
109 haSCs into bovine and ovine oocytes, respectively [2] (Supplementary Fig.13a,b). The
110 injected b- and o-haSCs formed a pseudo-pronucleus and underwent DNA
111 demethylation, which is consistent with *in vitro* fertilization (IVF) embryos
112 (Supplementary Fig.13c,d). Similar to the embryos derived by SCNT, most b- and o-
113 iCHI embryos could not develop to the blastocyst stage (Fig.2a, Supplementary
114 Fig.14a, Table 2). We and other previous studies revealed aberrant DNA and histone
115 methylation is a barrier for iCHI and SCNT [11, 12]. We thus performed pyrosequencing
116 and whole genome bisulfite sequencing (WGBS) on b- and o-haSCs, and detected
117 that the differentially methylated regions (DMRs) of parental imprints maintained
118 proper status as in sperm, including H19, IGF2, and DIO3 (Fig.2b, Supplementary
119 Fig.15a-c).
120
121 Unlike the asymmetric in IVF zygote, H3K4/K9/K27me3 in all pronuclei of b- and o-
122 iCHI embryos (Fig.2c, Supplementary Fig.14b). Injection with relative demethylases
123 KDM5B/4D/6A mRNA into b- and o-iCHI zygote, could not correct symmetric errors
124 (Supplementary Fig.16a-c, Table 2). Moreover, KDM5B/4D/6A overexpression
125 accelerated self-diploidization of ruminant haSCs (Supplementary Fig.16d). These
126 results prompted us to find another method to rescue the ruminant iCHI embryos.
127
128 As protamine (Pro) can exclude histone modifications during spermatogenesis [13], we
129 used the non-integrated episomal vector to express the PRM1 gene transiently. After
130 Pro-induction in b- and o-haSCs, we observed their nuclei compacted into spermatid-
131 like structures concomitant with H3K4/9/27me3 removal (Supplementary Fig.17a,b).
132 Importantly, the symmetric errors in b- and o-iCHI embryos can be amended by using
133 protaminized b- and o-haSCs as donors, leading iCHI blastocyst rate was comparable
134 to that of IVF embryos (Fig.2a,c, Supplementary Fig.14a,b, Table 2).

To maximize the utility of the Pro-iCHI approach, we chose to edit the Myostatin (MSTN) gene, whose mutation in animals causes a much greater muscle mass known as "double-muscling" [14]. Although the CRISPR prime editor (PE) has been applied in mice, it remains to be tested whether PE could be applied in large animals [15]. Genomic integration of exogenous DNA increases the risk of biosafety. We thus developed an episomal plasmid-based PE (ePE) to generate the gene-edited livestock with integration-free. Indeed, the ePE enables the production of nearly 100% short deletion (11 bp) and nucleotide substitution (G>A) in MSTN loci of the b- and o-haSCs, respectively (Supplementary Fig.18a-e).

We further tested whether the Pro-iCHI approach can produce live animals. Consequently, one MSTN-edited premature lamb with a much smaller size was obtained at 14 weeks (0.39%, 1/256; Fig.2d, Supplementary Table 3). Of the bovine Pro-iCHI embryos, 14.3% (1/7) were born as full-term (287 days) calf using WT b-haSCs as donor cells. Remarkably, we obtained two calves (13.3%, 2/15) with "double-muscle" phenotype using MSTN-edited b-haSCs (Fig.2e, Supplementary Fig.19a-e, Table 3).

Here, we show that FACE medium supports the establishment of b- and o-haSCs. As a proof-of-concept, we also developed a novel Pro-iCHI::ePE system to rapidly produce trait-improved cattle with exogene-free effects.

References

1. Li, W. et al. Androgenetic haploid embryonic stem cells produce live transgenic mice. *Nature* **490**, 407-411 (2012).
2. Yang, H. et al. Generation of genetically modified mice by oocyte injection of androgenetic haploid embryonic stem cells. *Cell* **149**, 605-617 (2012).
3. Li, W. et al. Genetic modification and screening in rat using haploid embryonic stem cells. *Cell Stem Cell* **14**, 404-414 (2014).
4. Lu, J. et al. Structure-activity relationship studies of small-molecule inhibitors of Wnt response. *Bioorg Med Chem Lett* **19**, 3825-3827 (2009).
5. Huang, S.M. et al. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614-620 (2009).
6. Smith, A. Formative pluripotency: the executive phase in a developmental continuum. *Development* **144**, 365-373 (2017).
7. Graf, A. et al. Fine mapping of genome activation in bovine embryos by RNA sequencing. *Proc Natl Acad Sci U S A* **111**, 4139-4144 (2014).
8. Jiang, Z. et al. Transcriptional profiles of bovine in vivo pre-implantation development. *BMC Genomics* **15**, 756 (2014).
9. Bogliotti, Y.S. et al. Efficient derivation of stable primed pluripotent embryonic stem cells from bovine blastocysts. *Proc Natl Acad Sci U S A* **115**, 2090-2095 (2018).
10. Zhao, L. et al. Establishment of bovine expanded potential stem cells. *Proc Natl Acad Sci U S A* **118** (2021).
11. Zhong, C. et al. CRISPR-Cas9-Mediated Genetic Screening in Mice with Haploid Embryonic Stem Cells Carrying a Guide RNA Library. *Cell Stem Cell* **17**, 221-232 (2015).
12. Yang, L., Song, L., Liu, X., Bai, L. & Li, G. KDM6A and KDM6B play contrasting roles in nuclear transfer embryos revealed by MERVL reporter system. *EMBO Rep* **19** (2018).
13. Czernik, M., Iuso, D., Toschi, P., Khochbin, S. & Loi, P. Remodeling somatic nuclei via exogenous expression of protamine 1 to create spermatid-like structures for somatic nuclear transfer. *Nat Protoc* **11**, 2170-2188 (2016).
14. Rodgers, B.D. & Garikipati, D.K. Clinical, agricultural, and evolutionary biology of myostatin: a comparative review. *Endocr Rev* **29**, 513-534 (2008).
15. Anzalone, A.V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149-157 (2019).

Methods

Animals and chemicals

All animal procedures were performed under the ethical guidelines of the State Key Laboratory of Reproductive Regulation and Breeding of Grassland Livestock, College of Life Sciences, Inner Mongolia University (IMU). Specific pathogen-free-grade B6D2F1 (C57BL/6 × DBA/2), CD1, and Kun-Ming (KM) mice were purchased from Laboratory Animal Research Center (Inner Mongolia University) or Vital River Laboratories Co., Ltd. Livestock are raised at the experimental farm of IMU and processed in commercial slaughterhouses. Mice were housed in specific pathogen-free conditions with a 12-h dark/12-h light cycle. Unless otherwise indicated, all chemicals were purchased from Sigma-Aldrich Chemical Reagent Co., Ltd.

Collection and *in vitro* maturation of oocytes

Collection and *in vitro* maturation (IVM) of ruminant oocytes were performed following the published protocols with slight modifications [13, 16]. Briefly, bovine or ovine ovaries from the slaughterhouse were transported to the laboratory. Collect cumulus-oocyte complexes (COCs) and transfer them to the IVM medium for 24 h in 5% CO_2 with humidified atmosphere at 38.5 °C [basic M199 medium supplemented with 2 mM glutamine, 300 μM, sodium pyruvate, 100 μM cysteamine, follicle-stimulating hormone (FSH; 50 ng/mL), luteinizing hormone (LH; 3 mg/mL), 1 μg/mL estradiol, and 0.04% gentamicin solution].

Generation of haploid androgenetic embryos

Sperm of Angus cattle or Dorper sheep were prepared for manipulation according to the published protocols [17-19]. The oocytes or zygotes were stained with 5 μg/mL Hoechst 33342 for 10 min to visualize spindle or pronuclei by a short 1-2 sec exposure to the ultraviolet (UV) light. Option A for injecting strategy, one-step SCNT was performed [20, 21], in which somatic cells were replaced by sperms. Briefly, the MII oocyte spindle was adjusted to 8 to 10 o'clock, and one donor sperm was injected into the nearby plasma. Immediately, the spindle was aspirated into the injection pipette and removed from the MII oocyte using the Piezo micromanipulator (Primetech) in the M2 medium containing 5 μg/mL cytochalasin B (CB) on a 37 °C heating stage of an inverted microscope (Nikon). After manipulation, the androgenetic embryos were activated in 5 μM ionomycin for 5 min and then incubated in 2 mM 6-dimethylaminopurine (6-DMAP) and 7.5 μg/mL CB in Ca^{2+}-free KSOM-AA for 5 h. Option B for removing strategy, the zygotes were produced by *in vitro* fertilization (IVF) as previously described [22]. Briefly, the zygotes were cultured in G1 medium for 7-8 h after IVF. Female pronuclei, which were near the second polar body, were removed

from zygotes in the M2 medium containing 5 μg/mL CB using a blunt Piezo-driven pipette. Following activation or removing, the androgenetic embryos were cultured in G1 medium (Vitrolife) for 72 h, and then transferred into G2 medium (Vitrolife) for further culturing at 38.5 °C, humidity to saturation, 5% CO_2, 5% O_2, and 90% N_2. The androgenetic embryos were allowed to develop into blastocysts that were used for haSCs derivation.

Further culture of blastocysts using different SCs medium

The blastocysts with good morphology were selected to generate SCs. In brief, zona pellucida (ZP) were removed from blastocysts by treatment with 5 mg/mL of pronase for 3-4 min under an inverted microscope with a warming stage. After removing ZP, blastocysts were plated on mitomycin C (MMC) treated mouse embryonic fibroblasts (MEFs) in the different mediums for SCs derivation, including 2i/L [23], 4i [24], t2iL+Gö [25], 5i/L/A [26], ABCL [27], LCDM [28], EPSC [29, 30], CTFR [9, 31], bEPSC [10]. After 10 days in culture, the outgrowths were manually picked in separate wells of a 12-well dish seeded with MMC-treated MEFs for further cultivation to avoid contamination of trophoblasts.

Further culture of blastocysts using different small-molecules

Individual denuded blastocysts were placed in separate wells of a 12-well dish seeded with a monolayer of MMC-treated MEFs, and cultured in basal N2B27 [32] or mTeSR [9] medium supplemented with 15% knockout serum replacement (KSR; Thermo), and different small-molecule compounds (Supplementary information, Fig.S3a). The small-molecules are used at the following concentrations: 3 μM CHIR99021 (C; Selleckchem), 2 μM Dimethindene maleate (D; Tocris), 2 μM Minocycline hydrochloride (M; Santacruz), 10 ng/mL LIF (L; Millipore), 2.5 μM IWR-1-endo (E; Selleckchem), 10 ng/mL FGF2 (F; Peprotech), 5 μM XAV939 (X; Selleckchem), 0.3 μM WH-4-023 (W; Selleckchem), 1 ng/mL TGFβ (T; Peprotech), 20 ng/mL Activin-A (A; Peprotech), 50 ng/mL BMP4 (B; R&D systems), 50 μg/mL Vitamin C (V; Selleckchem). All small-molecules are dissolved in dimethyl sulfoxide (DMSO). After 10 days in culture, the outgrowths were manually picked to avoid contamination of trophoblasts.

Derivation and routine culture of haSCs

After removing ZP, each blastocyst was plated into one well of a 96-well plate with the MMC-treated MEFs feeder in FACE medium at 37 °C, humidity to saturation, 5% CO_2, 5% O_2, and 90% N_2. The outgrowths were formed after 10-14 days, the medium was changed by half every 1-2 days, and passaged on the MEFs feeder followed by mechanically separating into several pieces. Cell expansion was carried out from 48-well to 6-well plates with feeder cells and then to 6-well plates for routine culture. Established haSCs were cultured on the MMC-treated MEFs feeder in FACE medium, with daily media change. For passaging, the cells were disaggregated using TrypLE

(Thermo) or Accutase (Stemcell) and passaged onto new feeder layers at a split ratio of 1:10 or 1:20 every 3-5 days. To enhance the survival of haSCs, 10 μM Y-27632 (ROCK inhibitor, Stemcell) was recommended to be added into the FACE medium for 24 h after the passage.

The basal medium for FACE is mTeSR1 (Stemcell, 85850 based media) [33] or N2B27 [32] [total 500 mL was prepared as follows: 240 mL DMEM/F12 (Thermo), 240 mL neurobasal (Thermo), 2.5 mL N2 supplement (Thermo), 5 mL B27 supplement (Thermo), 1% GlutaMAX (Thermo), 1% nonessential amino acids (Thermo), 0.1 mM β-mercaptoethanol (Thermo), 1% penicillin-streptomycin (Thermo), and 15% KSR (Thermo)]. To prepare the FACE medium, small molecules and cytokines were added in the basal medium at the following final concentrations: 10 ng/mL FGF2 (F; Peprotech), 10 ng/mL Activin-A (A; Peprotech), 3 μM CHIR99021 (C; Selleckchem), and 2.5 μM IWR-1-endo (E; Selleckchem).

FACS-enrichment of haploid cells

To sort haploid cells, we used fluorescence activating cell sorter (FACS) according to the published protocol with slight modifications [34-36]. Briefly, haSCs were dissociated into single cells with TrypLE (Thermo) or Accutase (Stemcell) and then incubated with 15 μg/mL Hoechst 33342 for 15 min in a 37 °C water bath. Following centrifugation, cells were rinsed with Dulbecco's phosphate-buffered saline (PBS) for removal of the Hoechst 33342, and resuspended in FACE medium with 10 μM Y-27632. Subsequently, the stained cells were through 40 μm cell strainer and the haploid cells were sorted using FACSAria II (BD biosciences). After enrichment, the haploid cells were plated on fresh MEFs feeder in FACE medium containing 10 μM Y-27632. After 24 h, haploid cells were cultured in FACE medium without Y-27632. Usually, haESC clones were formed within 10 days. For cell analysis, after fixation in 70% ethanol, cells were digested by 20 mg/ml RNase A and stained with 50 mg/ml propidium iodide (PI). Then, the haploid cells were analyzed using CytoFLEX LX (Beckman). Analytic flow profiles were performed using FlowJo software (Ashland).

Immunofluorescence staining

Immunofluorescence (IF) was performed as described previously [20]. Briefly, the cells were fixed with 4% paraformaldehyde (PFA) in PBS at RT for 1 h, and permeabilised in PBS supplemented with 0.2% Triton X-100 at RT for 45 min. The cells were incubated with primary antibodies in PBS that contained 3% blocker BSA (Thermo) at 4 °C overnight, then incubated with secondary antibodies for 1 h at RT in the dark. For 5mC and 5hmC staining, permeabilised embryos were incubated in 4 N HCl treatment for 10 min at room temperature, then the oocytes were immediately transferred to Tris-HCl (pH 8.5) for 10 min to neutralise the residual HCl. Finally, the cells were mounted in the anti-fade solution with DAPI (Thermo) and compressed with a cover-slip. All samples were observed by a laser scanning microscope (Nikon). The antibody

information was presented in Supplementary Table 4.

Western blot

Western blot (WB) was performed as described previously [12]. In brief, cells were lysed in RIPA lysis buffer (Beyotime) with protease inhibitor (Thermo) and phosphatase inhibitor (Thermo). Whole-cell lysates were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Following electrophoresis (90 V for 30 min, 30–50 V overnight), the proteins were transferred from the gel to 0.2 mm PVDF membranes (Bio-sharp), and blocked for 2 h at RT in TBST buffer containing 5% BSA or 5% fat-free milk. Films were incubated with the primary antibodies in TBST buffer containing 3% BSA at 4 °C overnight. After washing three times in TBST, PVDF membranes were incubated with secondary antibodies for 1 h at RT. Then, bound antibodies were detected with SuperSignal West Femto Substrate Trial Kit (Thermo). The antibody information was presented in Supplementary Table 4.

Karyotype analysis

Before karyotype analysis, haSCs were incubated with 0.1 μg/mL KaryoMAX Colcemid (Thermo) in FACE medium for 3-4 h, and mechanically separated from the MEFs feeder using pipette tips. Then, haSCs were disaggregated into single cells by 0.05% trypsin-EDTA (Thermo) at 37 °C for 10 min. After centrifugation, the pellet was re-suspended in hypotonic solution (0.56% KCL) at room temperature (RT) for 10 min. Cells were then fixed with ice-cold fixative (3:1 methanol: acetic acid) at RT for 5 min three times. Finally, haSCs suspension was dropped onto an ice-cold fixative washed slide and stained with 4′,6-diamidino-2-phenylindole (DAPI; Thermo). For each cell line, more than 30 metaphases were examined.

Chimeric assay

The chimeric assay was performed as previously described with slight modification [20]. In brief, haSCs were used 1 day before passaging, which showed an optimal undifferentiated morphology. For chimeric mice, ten haSCs were microinjected into CD1 or KM blastocysts using a blunt Piezo-driven pipette. Then, the reconstructed blastocysts were transplanted into the uterus of pseudo-pregnant mice at 3.5 days post-coitum (dpc). For chimeric blastocyst, six haSCs labelled by EGFP fluorescence were microinjected into an 8-cell stage B6D2F1 embryo. The reconstructed embryos were cultured in FACE medium for 4-5 h, then changed into the G1/G2 medium (1:1; Vitrolife) to obtain the chimeric blastocysts.

Embryoid body formation assay

Embryoid body (EB) formation was performed as previously described with slight modification [34]. Briefly, haSCs were dissociated into single cells with TrypLE (Thermo) and cultured for 7 days on ultra-low attachment plates in Iscove's modified Dulbecco's medium (IMDM; Thermo) supplemented with 1% nonessential amino acids (Thermo), 1% penicillin-streptomycin (Thermo), and 15% FBS (Gibco). Subsequently, Ebs were collected and dissociated by 0.05% trypsin-EDTA (Thermo) at 37 °C for 10 min. After

enrichment, cells were plated on atrigel-coated slides for an additional 3 days before IF staining.

Teratoma formation assay

Teratoma formation was performed as previously described [20]. In brief, nearly 10^9 haSCs were subcutaneously injected into the hind limbs of 8-week-old nude mice. After 5 weeks, fully formed teratomas were dissected and fixed with PBS containing 4% PFA, then embedded in paraffin, sectioned, and stained with hematoxylin and eosin (HE) for histological analysis.

Intracytoplasmic haSCs injection

Angus cattle and Dorper sheep ovaries were obtained from commercial slaughterhouses, and transported to the laboratory within 2 h at 28 °C in sterile saline solution (8.5 mg/mL NaCl) with a thermal container. The oocytes, which possessed at least three layers of cumulus, were selected for IVM according to the methods described above. Intracytoplasmic haSCs injection (iCHI) was performed following the published protocols with slight modifications [1, 2, 34]. Briefly, haSCs were arrested metaphase by culturing in FACE medium containing 0.05 μg/mL demecolcine for 12 h. Following centrifugation, cells were rinsed three times with PBS for removal of the demecolcine, and resuspended in the M2 medium containing 10 μM Y-27632 and 1.5% (w/v) polyvinylpyrrolidone (PVP). Each nucleus from haSCs was directly injected into an IVM oocyte using the Piezo micromanipulator (Primetech) in the M2 medium containing 5 μg/mL cytochalasin B (CB) on a 37 °C heating stage of an inverted microscope (Nikon). The reconstructed embryos were cultured in KSOM-AA medium for 1 hr and then activated using ionomycin followed by exposing them to 10 μg/mL cycloheximide (CHX) according to the methods described above.

Embryo Transfer

For bovine embryo transfer, blastocysts were stored in the TCM-199 medium (Thermo) supplemented with 15% FBS (Gibco). The iCHI blastocysts were transferred via the non-surgical transcervical method as previously described [37]. Briefly, a single blastocyst was transferred into the uterine horn of hormonally synchronised (CIDR progesterone + cloprostenol + gonadorelin) nulliparous heifer at day 7, estrus day 0 = day of micromanipulation). For ovine embryo transfer, 2-cell or 4-cell stage embryos were stored in TCM-199 medium (Thermo) supplemented with 10% BSA. The ovine embryos were transferred via surgical method using a TomCat cathether (Sovereign). Eight iCHI embryos were transferred into the oviduct of hormonally synchronized (CIDR progesterone + FSH + cloprostenol) recipient ewes at day 2 (estrus day 0 = day of micromanipulation). Pregnancy detection was diagnosed on day 35 and 50 of bovine and ovine embryonic development, respectively, using transrectal ultrasound EVO scanner (5.0MHz linear probe, E.I. Medical Imaging).

401
402 **Reverse transcription-quantitative PCR**
403 The reverse transcription PCR and real-time quantitative PCR reaction were
404 performed using HiScript II One Step qRT-PCR SYBR Green Kit (Vazyme), and signals
405 were detected with Applied Biosystems QuantStudio 3 (Thermo). Analysis of relative
406 gene expression was normalized to the internal GAPDH gene, and measured using
407 the $2^{(\Delta\Delta Ct)}$ method. The primer information was presented in Supplementary Table 5.
408
409 **Bisulphite sequencing**
410 Genomic DNA was extracted using the tissue and blood DNA extraction kit (Qiagen)
411 and treated with the Methylamp DNA modification kit (EpiGentek). To obtain sperm
412 genomic DNA, samples were pretreated with dithiothreitol (DTT) for 3 hr, following
413 proteinase K lysis and phenol-chloroform extraction. The bisulfite conversion was
414 performed as previously published studies [38]. The PCR products were cloned into
415 vectors using TA/Blunt-Zero Cloning Kit (Vazyme), and individual clones were
416 sequenced. The primer information was presented in Supplementary Table 5.
417
418 **Plasmid construction**
419 The coding sequence of Cas9 (H840A) :: M-MLVrt (D200N, T306K, W313F, T330P,
420 L603W), derived from Prime editor 2 system (PE2, Addgene plasmid #132775) [15], was
421 codon-optimized based on the general codon usage in cattle and sheep then
422 synthesized by commercial company. The above vector was designated boPE2 and
423 ovPE2. Indeed, “all-in-one” episomal PE plasmid is the same as epiCRISPR [39] except
424 that the Cas9 was replaced by boPE2 or ovPE2. All CRISPR editor plasmids used in
425 this work were assembled using the USER cloning method as previously described [15].
426 DNA amplification was conducted by Phanta Max Super-Fidelity DNA Polymerase
427 (Vazyme). Integrate the synthesised DNA into the designated positions on the vector
428 using ClonExpress Ultra One Step Cloning Kit V2 (Vazyme). The pegRNA targeting
429 gene was designed by the online program (http://pegfinder.sidichenlab.org) [40] based
430 on GenBank sequence.
431
432 **Genome editing**
433 Genome editing was performed as previously described with slight modification [39]. In
434 brief, approximately 10^4 haSCs were disassociated using TrypLE (Thermo), and
435 cultured in a 6-well plate with FACE medium for 24 h. A total of 2 μg ePE plasmid were
436 transfected into haSCs using Lipomaster 3000 Transfection Reagent (Vazyme)
437 following the manufacturer’s protocol. After transfection 72 h, haSCs were selected by
438 puromycin (0.2-0.5 μg/mL). For analysis of single cell-derived clones, the haSCs were
439 disaggregated into single cells at day 15 post-transfection, and plated them onto new
440 feeder layers with puromycin-free FACE medium for 14 days. Finally, individual

colonies were picked and genotyped. The pegRNA information was presented in Supplementary Table 6.

WGS and data processing

The whole genome sequencing (WGS) was performed as described previously [34,41]. Briefly, libraries were generated from 1 mg of genomic DNA using the TruSeq DNA Sample Preparation Kit (Illumina), and sequenced on the HiSeq 4000 system (Illumina) by Lc-Bio technologies Co., Ltd. For data processing, the copy number variation (CNV) detection was performed as previously described [42].

The template genome was downloaded from NCBI RefSeq, including cattle (ARS-UCD2.0, https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_002263795.4/) and sheep genome (ARS-UI_Ramb_v3.0, https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_016772045.2/).

WGBS and data processing

The whole genome bisulfite sequencing (WGBS) was performed according to previously [34, 43]. Briefly, genomic DNA was extracted from haSCs using the DNeasy B lood & Tissue extraction kit (Qiagen) following end repair and A-tailing. The DNA was fragmented into around 200-300 bp by ultrasonicator (Covaris). Bisulfite-converted DNA was captured using EZ-96 DNA Methylation-DirectTM MagPrep (Zymo) according to the manufacturer's instructions. Library sequencing was performed on the HiSeq 4000 system (Illumina) by Lc-Bio Technologies Co., Ltd. The cleaned data were aligned to the cattle (ARS-UCD2.0) reference with Bismark/Bowtie2 mode. Bismark2summary command was used to calculate the genome CpG methylation level. RnBeads R package was used to calculate the percentage of methylation.

CUT&Tag and data processing

The CUT&Tag was performed using the Hyperactive Universal CUT&Tag Assay Kit for Illumina (Vazyme) according to the manufacturer's instructions. Paired-end sequencing was further performed on the HiSeq 4000 system (Illumina) by Frasergen Co., Ltd. The data processing was performed as described previously [44]. Briefly, paired-end reads were aligned using Bowtie2 (2.2.5) with options: -local-very-sensitive-local–no-unal–no-mixed–no-discordant-phred33 -I 10 -X 700. For peak calling, the parameters used were macs2 callpeak -t input_file -p 1e-5 -f BEDPE/BED -keep-dup all -n out_name. The antibody information was presented in Supplementary Table 4.

buRNA-seq and data processing

The bulk RNA sequencing (buRNA-seq) was followed by previously published studies [20, 35]. Briefly, total RNA was extracted from haSCs using TRIzol Plus RNA Purification Kit (Thermo) following the manufacturer's instructions, and the RNA integrity was checked using Bioanalyzer 2100 (Agilent). The total cDNA library was then amplified by 18–20 cycles for library construction, which was performed following the

manufacturer's instructions (Illumina, USA). Sequencing was further performed on the HiSeq 4000 system (Illumina) by Frasergen Co., Ltd. After removing low-quality reads and adapters, the raw reads were mapped to the cattle (ARS-UCD2.0) or sheep (ARS-UI_Ramb_v3.0) genome using Tophat (v1.3.3) with default parameters [45]. Expression levels for all RefSeq transcripts were quantified to fragments per kilobase of exon model per million mapped reads (FPKM) using Cufflinks (v1.2.0) [46]. Gene Ontology (GO) analyses were performed using R package GO.db and Database for Annotation, Visualization, and Integrated Discovery (DAVID) [47].

scRNA-seq and data processing

The single-cell RNA sequencing (scRNA-seq) was followed by previously published studies [43, 48]. Briefly, haSCs or E8~E14 embryos were dissociated into single cells with TrypLE (Thermo) and resuspended in 1 Ca^{2+}-Mg^{+} free PBS containing 0.04% BSA. Next, 300-1000 living cells were loaded on a Chromium Single Cell Controller (10x Genomics) to generate single-cell gel beads in emulsion (GEMs) cDNAs. Then, recovered cDNAs were used to construct a library using Single Cell 3' Library and Gel Bead Kit V2 (10x Genomics) according to the manufacturer's instructions. Library sequencing was performed on the HiSeq 4000 system (Illumina) with pair-end 150 bp (PE150) by Lc-Bio technologies Co., Ltd. The raw sequencing data were aligned and quantified using the Cell Ranger (10x Genomics) pipeline against the cattle (ARS-UCD2.0) reference genome. The Seurat R package was used to merge the scRNA-seq and cell clustering. Gene expression was calculated by the LogNormalize method in the NormalizeData function. RunPCA function was used in principal component analysis (PCA) with the most variable genes. Cell clusters were identified by a shared nearest neighbor (SNN) modularity optimisation based on a clustering algorithm with the FindClusters function. The *t*-SNE (*t*-distributed stochastic neighbor embedding) dimensionality reduction was generated using the RunTSNE function and the UMAP (uniform manifold approximation and projection) was performed using the RunUMAP function. Differentially expressed genes were identified using FindAllMarkers, and the top 30 differentially expressed genes (DEGs) were marked as signature genes. Heatmaps were plotted by the DoHeatmap function.

Data availability

All WGBS, CUT&Tag, buRNA-Seq, scRNA-seq, WGS data are available through GEO under the accession numbers GSE250491, GSE250493, GSE250494, GSE250495, and GSE250496. In addition, the SuperSeries GSE250497 links all the above data sets together.

Acknowledgements

We thank Hong Wang (Inner Mongolia University) for technical help. This study was supported by the National Natural Science Foundation of China (32360837, 32341052), Scientific and Technological Innovation 2030 (2023ZD0404803), Inner Mongolia Open

Competition Projects (2022JBGS0025), Inner Mongolia Science and Technology Leading Team (2022LJRC0006), Inner Mongolia Science and Technology Major Projects (2021ZD0009, 2021ZD0008, 2022ZD0008, 2023KJHZ0028), Inner Mongolia Young Talents Projects (NJYT23138), Inner Mongolia Natural Science Foundation (2023MS03004), Central Government Guides Development (2022ZY0212), National Agricultural Science and Technology Project (NK2022130203), Collaborative Innovation among Universities in Hohhot (XTCX2023-06), Ministry of Education Engineering Centre Project (JYBGCSYS2022).

Author contributions

L.Y., A.D., X.L., L.S., D.W., S.W., Z.H., L.B., C.B., G.S., Z.W., L.Z. performed experiments. L.Y., G.L., S.G. designed experiments, L.Y. and G.L. wrote the manuscript.

Competing interests

The authors declare no competing interests.

543 ## References

544 16. Ross, P.J. & Cibelli, J.B. Bovine somatic cell nuclear transfer. *Methods Mol*
545 *Biol* **636**, 155-177 (2010).
546 17. Czernik, M., Iuso, D., Toschi, P., Khochbin, S. & Loi, P. Remodeling somatic
547 nuclei via exogenous expression of protamine 1 to create spermatid-like
548 structures for somatic nuclear transfer. *Nat Protoc* **11**, 2170-2188 (2016).
549 18. Zhang, J. et al. Dissecting the molecular features of bovine-arrested eight-cell
550 embryos using single-cell multi-omics sequencingdagger. *Biol Reprod* **108**,
551 871-886 (2023).
552 19. Alberio, R., Motlik, J., Stojkovic, M., Wolf, E. & Zakhartchenko, V. Behavior of
553 M-phase synchronized blastomeres after nuclear transfer in cattle. *Mol*
554 *Reprod Dev* **57**, 37-47 (2000).
555 20. German, S.D., Lee, J.H., Campbell, K.H., Sweetman, D. & Alberio, R. Actin
556 depolymerization is associated with meiotic acceleration in cycloheximide-
557 treated ovine oocytes. *Biol Reprod* **92**, 103 (2015).
558 21. Yang, L. et al. Transient Dux expression facilitates nuclear transfer and
559 induced pluripotent stem cell reprogramming. *EMBO Rep* **21**, e50054 (2020).
560 22. Zhou, Q. et al. Generation of fertile cloned rats by regulating oocyte
561 activation. *Science* **302**, 1179 (2003).
562 23. Kinoshita, M. et al. Pluripotent stem cells related to embryonic disc exhibit
563 common self-renewal requirements in diverse livestock species. *Development*
564 **148** (2021).
565 24. Ying, Q.L. et al. The ground state of embryonic stem cell self-renewal. *Nature*
566 **453**, 519-523 (2008).
567 25. Irie, N. et al. SOX17 is a critical specifier of human primordial germ cell fate.
568 *Cell* **160**, 253-268 (2015).
569 26. Takashima, Y. et al. Resetting transcription factor control circuitry toward
570 ground-state pluripotency in human. *Cell* **158**, 1254-1269 (2014).
571 27. Theunissen, T.W. et al. Systematic identification of culture conditions for
572 induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**,
573 471-487 (2014).
574 28. Bao, S. et al. Derivation of hypermethylated pluripotent embryonic stem cells
575 with high potency. *Cell Res* **28**, 22-34 (2018).
576 29. Yang, Y. et al. Derivation of pluripotent stem cells with in vivo embryonic and
577 extraembryonic potency. *Cell* **169**, 243-257 e225 (2017).
578 30. Yang, J. et al. Establishment of mouse expanded potential stem cells. *Nature*
579 **550**, 393-397 (2017).

580 31. Gao, X. et al. Establishment of porcine and human expanded potential stem
581 cells. *Nat Cell Biol* **21**, 687-699 (2019).
582 32. Bogliotti, Y.S. et al. Efficient derivation of stable primed pluripotent embryonic
583 stem cells from bovine blastocysts. *Proc Natl Acad Sci U S A* **115**, 2090-2095
584 (2018).
585 33. Vilarino, M. et al. Derivation of sheep embryonic stem cells under optimized
586 conditions. *Reproduction* **160**, 761-772 (2020).
587 34. Zhao, L. et al. Establishment of bovine expanded potential stem cells. *Proc*
588 *Natl Acad Sci U S A* **118** (2021).
589 35. De Los Angeles, A., Okamura, D. & Wu, J. Highly efficient derivation of
590 pluripotent stem cells from mouse preimplantation and postimplantation
591 embryos in serum-free conditions. *Methods Mol Biol* **2005**, 29-36 (2019).
592 36. Ludwig, T.E. et al. Feeder-independent culture of human embryonic stem
593 cells. *Nat Methods* **3**, 637-646 (2006).
594 37. Zhang, X.M. et al. In vitro expansion of human sperm through nuclear
595 transfer. *Cell Res* **30**, 356-359 (2020).
596 38. Zhong, C. et al. Generation of human haploid embryonic stem cells from
597 parthenogenetic embryos obtained by microsurgical removal of male
598 pronucleus. *Cell Res* **26**, 743-746 (2016).
599 39. Elling, U. et al. Derivation and maintenance of mouse haploid embryonic stem
600 cells. *Nat Protoc* **14**, 1991-2014 (2019).
601 40. Yang, L., Song, L., Liu, X., Bai, L. & Li, G. KDM6A and KDM6B play
602 contrasting roles in nuclear transfer embryos revealed by MERVL reporter
603 system. *EMBO Rep* **19** (2018).
604 41. Yang, H. et al. Generation of genetically modified mice by oocyte injection of
605 androgenetic haploid embryonic stem cells. *Cell* **149**, 605-617 (2012).
606 42. Li, W. et al. Androgenetic haploid embryonic stem cells produce live
607 transgenic mice. *Nature* **490**, 407-411 (2012).
608 43. Owen, J.R. et al. One-step generation of a targeted knock-in calf using the
609 CRISPR-Cas9 system in bovine zygotes. *BMC Genomics* **22**, 118 (2021).
610 44. Hajkova, P. et al. DNA-methylation analysis by the bisulfite-assisted genomic
611 sequencing method. *Methods Mol Biol* **200**, 143-154 (2002).
612 45. Anzalone, A.V. et al. Search-and-replace genome editing without double-
613 strand breaks or donor DNA. *Nature* **576**, 149-157 (2019).
614 46. Xie, Y. et al. An episomal vector-based CRISPR/Cas9 system for highly
615 efficient gene knockout in human pluripotent stem cells. *Sci Rep* **7**, 2320
616 (2017).
617 47. Chow, R.D., Chen, J.S., Shen, J. & Chen, S. A web tool for the design of
618 prime-editing guide RNAs. *Nat Biomed Eng* **5**, 190-194 (2021).

619 48. Patch, A.M. et al. Whole-genome characterization of chemoresistant ovarian
620 cancer. *Nature* **521**, 489-494 (2015).
621 49. Xie, C. & Tammi, M.T. CNV-seq, a new method to detect copy number
622 variation using high-throughput sequencing. *BMC Bioinformatics* **10**, 80
623 (2009).
624 50. Shen, H. et al. Mouse totipotent stem cells captured and maintained through
625 spliceosomal repression. *Cell* **184**, 2843-2859 e2820 (2021).
626 51. Kaya-Okur, H.S. et al. CUT&Tag for efficient epigenomic profiling of small
627 samples and single cells. *Nat Commun* **10**, 1930 (2019).
628 52. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions
629 with RNA-seq. *Bioinformatics* **25**, 1105-1111 (2009).
630 53. Trapnell, C. et al. Transcript assembly and quantification by RNA-seq reveals
631 unannotated transcripts and isoform switching during cell differentiation. *Nat*
632 *Biotechnol* **28**, 511-515 (2010).
633 54. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative
634 analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*
635 **4**, 44-57 (2009).
636 55. Zhao, T. et al. Single-cell RNA-seq reveals dynamic early embryonic-like
637 programs during chemical reprogramming. *Cell Stem Cell* **23**, 31-45 e37
638 (2018).
639

Figure legends

Fig.1 | *De novo* derivation of ruminant haSCs with formative features.

a. Schematic overview of haSCs derivation, haploid embryos were generated by injecting sperm into enucleated oocytes.

b. Morphology of ruminant haSCs derived from haploid blastocysts.

c. The effects of FACE medium on haploidy maintenance of ruminant haSCs with FACS enrichment.

d. Immunofluorescence images showing that ruminant haSCs expressed core pluripotency markers (SOX2 and OCT4) and formative pluripotency marker (OTX2).

e. UMAP projection of b-haSCs, colored by pluripotency-associated marker genes.

f. Heatmap and bubble displaying the expression of formative and primed marker genes in different samples.

g. Heatmap of H3K4me3 and H3K27me3 signals in b-haSCs and primed-b-SCs.

h. Ternary plot showing cross-species comparison of orthologous transcription factors, formative-pluripotency related genes.

i. Chimeric mouse generated by EGFP-labeled ruminant haSCs at E18.5.

Scale bars: b, 100 μm; d, 25 μm; n ≥ 3 independent experiments; B, bovine; O, ovine; M, mouse.

Fig.2 | Generation of MSTN-knockout ruminants using the protaminized haSCs.

a. Morphology of ruminant blastocysts derived by iCHI, Pro-iCHI, and IVF.

b. Bisulfite pyrosequencing analysis of paternally imprinted regions. The filled and open squares represent methylated and unmethylated CpG sites, respectively.

c. Immunostaining for H3K4me3, H3K9me3, and H3K27me3 of ruminant blastocysts derived by iCHI, Pro-iCHI, and IVF.

d. Image of ovine fetus at 98 days from MSTN-edited o-haSCs.

e. Representative images of WT and MSTN-edited cattle through the Pro-iCHI approach.

Scale bars: a, 100 μm; c, 25 μm; n ≥ 3 independent experiments; B, bovine; O, ovine; ♀, female; ♂, male.

Figure 1

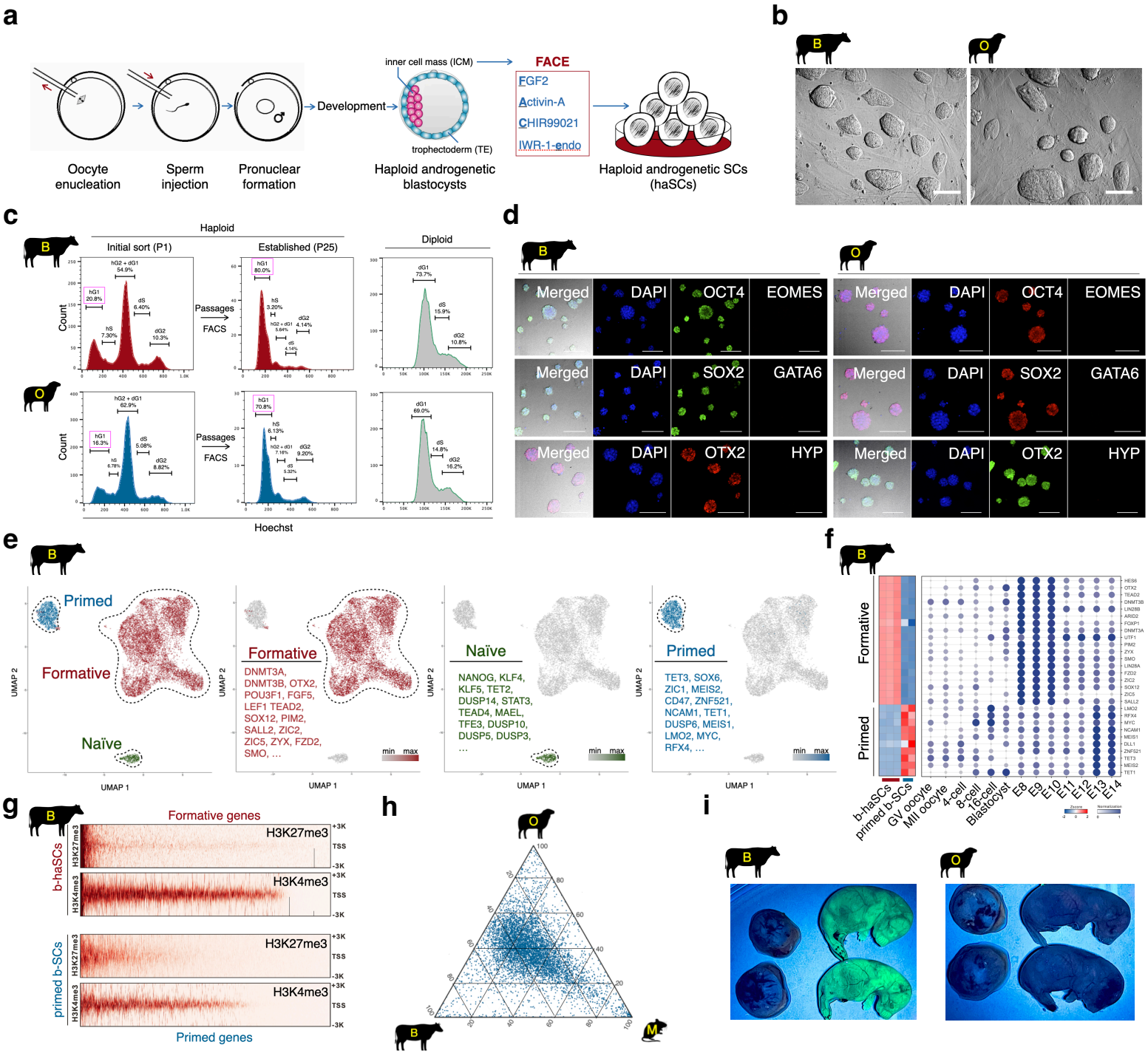
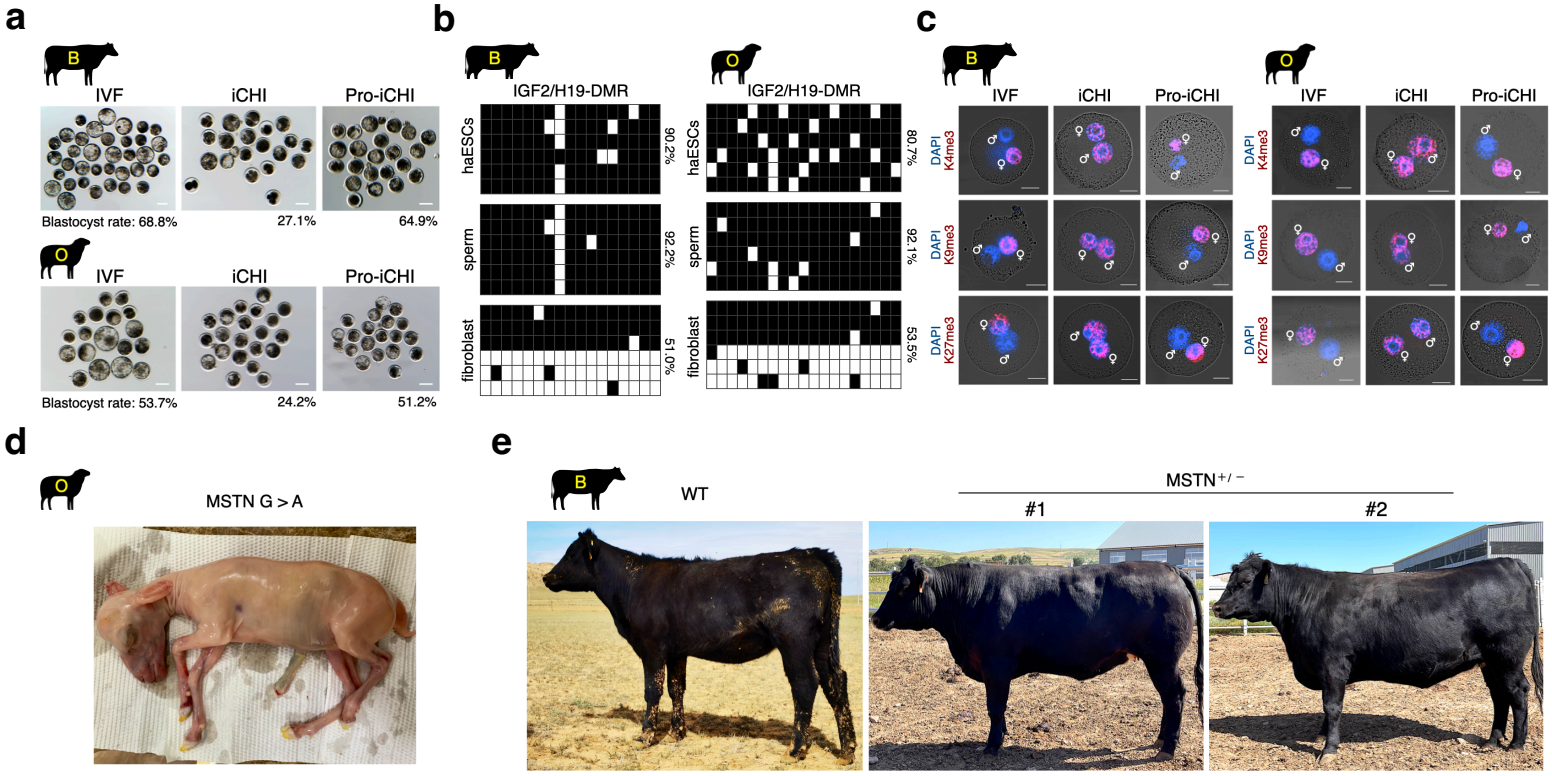

a
Oocyte enucleation
Sperm injection
Pronuclear formation
Development
inner cell mass (ICM)
trophectoderm (TE)
Haploid androgenetic blastocysts
FACE
FGF2
Activin-A
CHIR99021
IWR-1-endo
Haploid androgenetic SCs (haSCs)
b
B
O
c
Haploid
Initial sort (P1)
Established (P25)
Diploid
Passages
FACS
Count
Hoechst
d
Merged
DAPI
OCT4
EOMES
SOX2
GATA6
OTX2
HYP
e
Primed
Formative
Naïve
UMAP 1
UMAP 2
DNMT3A, DNMT3B, OTX2, POU3F1, FGF5, LEF1 TEAD2, SOX12, PIM2, SALL2, ZIC2, ZIC5, ZYX, FZD2, SMO, …
NANOG, KLF4, KLF5, TET2, DUSP14, STAT3, TEAD4, MAEL, TFE3, DUSP10, DUSP5, DUSP3, …
TET3, SOX6, ZIC1, MEIS2, CD47, ZNF521, NCAM1, TET1, DUSP6, MEIS1, LMO2, MYC, RFX4, …
min
max
f
Formative
Primed
b-haSCs
primed b-SCs
GV oocyte
MII oocyte
4-cell
8-cell
16-cell
Blastocyst
E8
E9
E10
E11
E12
E13
E14
g
Formative genes
Primed genes
b-haSCs
primed b-SCs
H3K27me3
H3K4me3
TSS
h
M
i

Figure 2


a
B
IVF
iCHI
Pro-iCHI
Blastocyst rate: 68.8%
27.1%
64.9%
O
IVF
iCHI
Pro-iCHI
Blastocyst rate: 53.7%
24.2%
51.2%
b
B
IGF2/H19-DMR
haESCs
sperm
fibroblast
90.2%
92.2%
51.0%
O
IGF2/H19-DMR
haESCs
sperm
fibroblast
80.7%
92.1%
53.5%
c
B
IVF
iCHI
Pro-iCHI
DAPI
K4me3
DAPI
K9me3
DAPI
K27me3
O
IVF
iCHI
Pro-iCHI
DAPI
K4me3
DAPI
K9me3
DAPI
K27me3
d
O
MSTN G > A
e
B
WT
MSTN+/−
#1
#2