PNAS Nexus, 2023, 2, 1–14
https://doi.org/10.1093/pnasnexus/pgad343
Advance access publication 20 October 2023
Research Report

Ablation of *OCT4* function in cattle embryos by double electroporation of CRISPR-Cas for DNA and RNA targeting (CRISPR-DART)

Jada L. Nix, Gustavo P. Schettini, Savannah L. Speckhart, Alan D. Ealy and Fernando H. Biase*

School of Animal Sciences, Virginia Polytechnic Institute and State University, 175 W Campus dr, Blacksburg, VA 24061, USA
*To whom correspondence should be addressed: Email: fbiase@vt.edu
Edited By: Patrick Stover

Abstract

CRISPR-Cas ribonucleoproteins (RNPs) are important tools for gene editing in preimplantation embryos. However, the inefficient production of biallelic deletions in cattle zygotes has hindered mechanistic studies of gene function. In addition, the presence of maternal RNAs that support embryo development until embryonic genome activation may cause confounding phenotypes. Here, we aimed to improve the efficiency of biallelic deletions and deplete specific maternal RNAs in cattle zygotes using CRISPR-Cas editing technology. Two electroporation sessions with Cas9D10A RNPs targeting exon 1 and the promoter of *OCT4* produced biallelic deletions in 91% of the embryos tested. In most cases, the deletions were longer than 1,000 nucleotides long. Electroporation of Cas13a RNPs prevents the production of the corresponding proteins. We electroporated Cas9D10A RNPs targeting exon 1, including the promoter region, of *OCT4* in two sessions with inclusion of Cas13a RNPs targeting *OCT4* mRNAs in the second session to ablate *OCT4* function in cattle embryos. A lack of *OCT4* resulted in embryos arresting development prior to blastocyst formation at a greater proportion (13%) than controls (31.6%, $P < 0.001$). The few embryos that developed past the morula stage did not form a normal inner cell mass. Transcriptome analysis of single blastocysts, confirmed to lack exon 1 and promoter region of *OCT4*, revealed a significant (False Discovery Rate, FDR < 0.1) reduction in transcript abundance of many genes functionally connected to stemness, including markers of pluripotency (*CADHD1*, *DPPA4*, *GNL3*, *RRM2*). The results confirm that *OCT4* is a key regulator of genes that modulate pluripotency and is required to form a functional blastocyst in cattle.

Keywords: CRISPR, electroporation, gene editing, preimplantation development, preimplantation development

Significance Statement

CRISPR-Cas-mediated DNA editing can revolutionize agriculture and biomedicine due to its simplicity of design and use. Modifications induced in embryos, though challenging to accomplish, are beneficial for the advancement of livestock production and the study of biological function. Here, we developed an approach using CRISPR-Cas enzymes to remove DNA segments of the cattle genome in one-cell embryos. Our results show major advancement in the efficiency of producing large deletions in the genome of cattle embryos. Using our approach, we removed the function of the *OCT4* gene. Our results confirmed *OCT4* as a major regulator of pluripotency genes during embryo development and its requirement for the formation of an inner cell mass in cattle.

Introduction

The driving force behind gene functionality studies is the targeted alteration of genomic sequences followed by observation of phenotypic deviations. The deletion of functional sequences in the genome, also called knockouts (KOs), can be used to study the roles of genes during preimplantation embryonic development (1). Mechanistic studies of gene function provide information connecting genome and phenotype during early embryogenesis, and the data may be used to better understand biological function (2–5) or disease (6–8). The CRISPR-Cas system has been the method of choice for most researchers wishing to alter genome sequences in somatic (9–11), germ (12–15), or embryonic cells (16–25). CRISPR-Cas systems have gained traction due to the simplicity of design and synthesis of gRNAs with sequence complementarity to the target region (26) and improved efficiency when compared with other common methods for sequence alterations (27–30).

Despite recent advancements in protein engineering giving rise to CRISPR-Cas RNPs of greater efficiency and specificity (31), biallelic deletion efficiency, or the deletion of targeted sequences in both chromosomes, remains low in CRISPR-Cas-treated zygotes

Competing Interest: F.H.B. is an inventor on US Provisional patent application no. 63/538,181, which is related to the CRISPR-DART method described in this article. Other authors declare no competing interest.
Received: July 18, 2023. **Accepted:** October 11, 2023

across many species, including cattle (32–37). Interestingly, only four reports provide data on biallelic deletion efficiency in studies utilizing CRISPR-Cas introduced through electroporation of cattle zygotes (33, 35, 36, 38). These studies averaged 75% of sampled embryos containing partial deletions, with the presence of at least one wild-type allele, and 59% containing full deletions with no wild-type alleles. Some intrinsic factors of zygote biology, such as chromatin compaction and the timing of DNA replication, may impair deletion efficiency due to sequence inaccessibility for CRISPR-Cas binding or the increased number of target sites requiring DNA cleavage. Though the introduction of increased amounts of CRISPR-Cas by more intense electroporation conditions is shown to improve editing efficiencies in cattle zygotes, embryonic mortality increases in tandem (38). Alternate methods for increasing CRISPR-Cas content in the zygote have been used, such as zona pellucida drilling prior to electroporation in cattle (35) or zona removal in swine (39). These methods may improve CRISPR-Cas delivery but do not mitigate the setback of embryo mortality. Additionally, it has been suggested that maternally inherited mRNA, present in mammalian zygotes (40–43), may support sufficient protein production in the absence of a functional gene. The presence of mRNA resulting from the gene of interest likely hinders gene functionality studies in preimplantation embryos and may be responsible for inconsistent KO phenotypes. To that end, Cas13a (44), a type of Cas that binds and cleaves single-stranded RNAs (45), may be used to knockdown maternal or nascent mRNA and further obstruct protein production, but this element has not been accounted for in previous cattle studies. Altogether, many factors can influence the efficiency of CRISPR-Cas systems in preimplantation embryos.

The gene *OCT4*, or octamer transcription factor 4, is thought to maintain pluripotency in early cattle (46, 47) and human (48, 49) embryos through its role as a transcription factor for many pluripotency-related genes (48, 50). Additionally, it functions in the HIPPO signaling pathway (51) and is thought to be a key regulator of the first cell-lineage differentiation event in cattle. The function of OCT4 has been studied in murine preimplantation embryogenesis models, and these studies show that normal blastocyst development and first cell-lineage differentiation are possible in the absence of an *OCT4* gene (52, 53), but one murine model results in the development of blastocysts with absent inner cell mass (48). As HIPPO signaling processes vary between bovine and murine preimplantation development (51, 54), these results may not provide adequate translation of information regarding human cell-lineage differentiation. Studies to determine the role of *OCT4* have been completed by CRISPR-Cas-mediated KOs in cattle zygotes, but these studies produced varying outcomes and inconsistent phenotypes (35, 36, 55, 56). Most studies report *OCT4* KO cattle embryos maintaining the ability to reach the blastocyst stage and effectively completing the first cell-lineage differentiation event in the absence of this gene (36, 55, 56). Conversely, one report showed developmental arrest at the morula stage, prior to cell-lineage specification (35). This variability may be due to unaccounted factors, such as maternal or preexisting mRNA, the common presence of wild-type alleles in CRISPR-Cas genome-edited cattle zygotes, and how zygotes were generated.

Here, we aimed to improve the efficiency of CRISPR-Cas-mediated biallelic deletions in cattle zygotes while degrading preexisting RNAs transcribed from the target gene. We targeted the *OCT4* gene, given the inconsistency of results from previous reports. We hypothesized that paired RNPs formed with CRISPR-Cas9D10A, a type of Cas9 that cleaves one strand on the targeted DNA sequence and must be used in pairs to produce deletions (57), produce larger deletions at greater consistency and efficiency than CRISPR-Cas9, a type of Cas9 that cleaves both strands of the targeted DNA sequence, and CRISPR-Cas13a can efficiently knockdown mRNA in cattle zygotes. We also hypothesized that simultaneous targeting of DNA and RNA could ablate gene function in cattle zygotes in vitro. In this study, we have mitigated the barriers of poor deletion efficiency and the presence of preexisting mRNA while maintaining embryo survival. The dual delivery of CRISPR-Cas9D10A, 6 h apart, increases the incidence of gene editing and full deletions. Additionally, we targeted maternally inherited transcripts with CRISPR-Cas13a while simultaneously removing a targeted sequence of the genome. Altogether, we have developed a method for high-efficiency genome and transcriptome editing in bovine zygotes using CRISPR-Cas editing technology. Our approach overcomes many limitations of gene editing for mechanistic studies of gene function in preimplantation embryos. Although cattle blastocyst formation is possible in the absence of OCT4, these embryos lack an inner cell mass and present severe transcriptional dysregulation of several genes related to stemness.

Results

First, we assessed the efficacy of electroporation and the cleavage function of the RNPs. Here, we used electroporation conditions modified from a previous publication (35), as follows: 6 poring pulses of 15 V, with 10% decay, for 2 ms with a 50-ms interval, immediately followed by 5 transfer pulses of 3 V, 40% decay, for 50 ms with a 50 ms interval, alternating the polarity. Fluorescence imaging showed that the RNP formed by Cas9-RFP + scramble guide RNAs (gRNAs) bypassed the zona pellucida in nearly all putative zygotes (PZs) electroporated (Fig. S1A, Appendix). Next, we confirmed that the RNPs formed with Cas9 + OCT4 single-guide RNA (sgRNA) 1 or Cas9 + OCT4sgRNA2 were able to cleave the targeted DNA in vitro (Fig. S1B, Appendix).

Both CRISPR-Cas9 and CRISPR-Cas9D10A produce deletions in cattle zygotes

First, we asked whether Cas9 and Cas9D10A would result in similar editing efficiencies and deletion patterns. High-throughput-targeted sequencing revealed that 73.1 and 81.5% of embryos presented at least 1 segment of DNA deleted when we used Cas9 (N embryos = 26) or Cas9D10A (N embryos = 27), respectively. We observed that 15.4 and 25.9% of the embryos electroporated with Cas9 or Cas9D10A, respectively, and genotyped by sequencing, did not have a wild-type copy of the DNA in the targeted region. The deletions resultant from Cas9 or Cas9D10A varied in their location and length. We observed that Cas9D10A RNPs produced longer deletions and removed the segment of DNA that included both sgRNAs, whereas Cas9 mostly produced small deletions in the region surrounding the sgRNAs but did not cause many deletions spanning both sgRNAs (Fig. 1).

Although not significant ($P = 0.27$, Fisher's exact test), Cas9D10A produced 10.5% points more full deletions, with no wild-type alleles, when compared with Cas9. Thus, we carried out the next experiments with Cas9D10A and *OCT4*-targeting sgRNAs. Also, considering that many of the Cas9-RFP RNPs remained in the membrane or perivitelline space (Fig. S1A, Appendix), we reasoned that a second electroporation would increase the efficiency of full deletions in cattle PZs. A second electroporation of PZ (~6 h after the first electroporation; see Materials and methods for details) with RNPs composed of

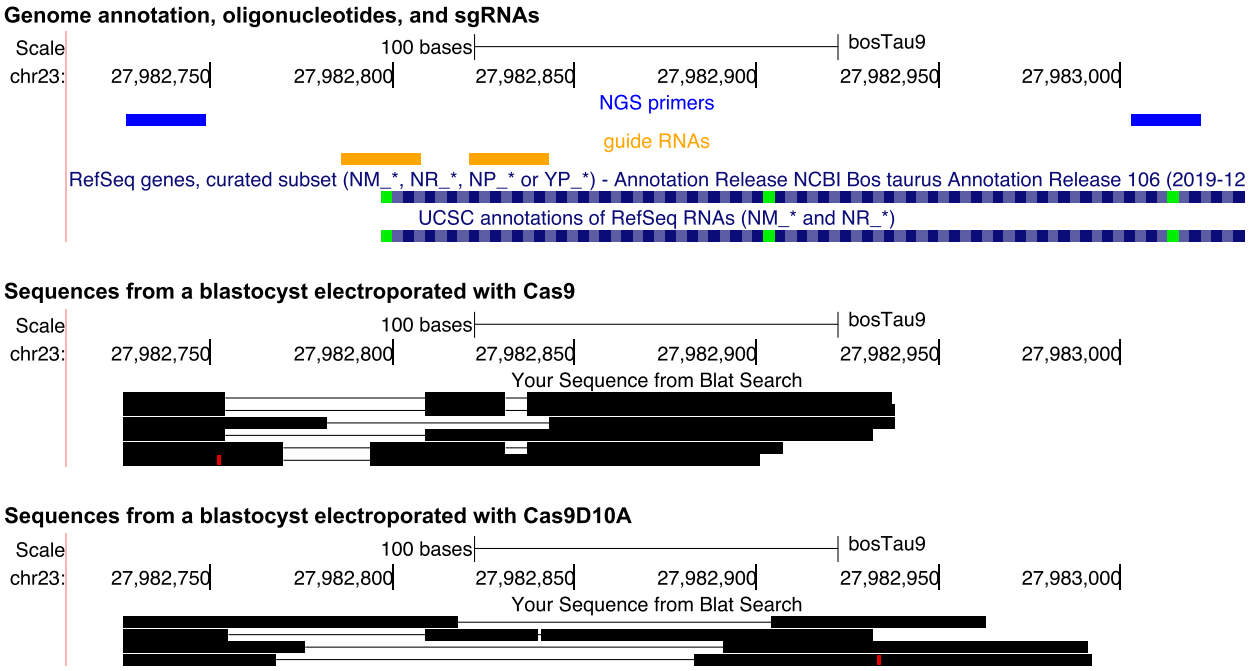

Fig. 1. Targeted DNA deletions using CRISPR-Cas9. Representative images of the DNA mapping of sequences resultant from high-throughput sequencing of embryos electroporated with either Cas9 or Cas9D10A aligned to the cattle genome. Annotation obtained from the Genome Browser.

Cas9D10A and associated sgRNAs resulted in no PCR amplification for most blastocysts when using the oligonucleotides designed for high-throughput short-read sequencing (Fig. S2A, B, Appendix). This outcome, and prior reports that CRISPR-Cas9 can produce unexpected large deletions (58, 59), prompted us to design oligonucleotides to flank a wider region of the DNA surrounding our sgRNA target sequences. Approximately, 19% of the blastocysts tested with this long-range pair of oligonucleotides produced an amplicon (Fig. S2C, Appendix). All blastocysts that had no amplicon produced with oligos surrounding our targeting sgRNAs were tested for amplification of a nontargeted autosomal region of the genome to confirm that an embryo was present in the tube (Fig. S2D, E, Appendix, see Text S1, Appendix for methods).

We sequenced the PCR products from seven blastocysts using the Sanger procedure, and three of these samples produced electropherograms from only one fragment (Fig. 2A). The long-range PCR produced multiple amplicons in the other three samples, which is unsuitable for Sanger sequencing. Therefore, we decided to proceed with Nanopore sequencing for multiple allele detection. Twenty-four blastocysts produced amplicons with long-range oligos and were sequenced by Sanger (Fig. 2B) or Nanopore (Fig. 2C) methods. Sequencing results showed that 95.8% (23/24) of the blastocysts had at least 1 chromosome with a deleted segment on the targeted DNA sequence (see an example in Fig. S3, Appendix). In addition, 70% (17/24) of the blastocysts sequenced did not present a wild-type sequence in the targeted region (Fig. 2B, other two examples in Fig. S4, Appendix). We note that 72 out of 89 blastocysts tested with our long-range oligos did not produce an amplicon, though the presence of DNA was confirmed by amplification of a nontargeted autosomal region in each sample. Therefore, we can reason that these 72 blastocysts had unexpectedly larger DNA deletions (58, 59) that eliminated at least one oligonucleotide pairing site on all chromosomes. Under such reasoning, we can estimate that 91% (81/89) of the blastocysts were edited with no wild-type sequence of the targeted DNA.

Embryo survival following one or two Cas9D10A electroporation sessions

We tested whether the electroporation of Cas9D10A with scramble gRNAs would impact development to the blastocyst stage. One electroporation session with scramble gRNAs produced similar results to controls (164–166 h postfertilization [hpf]—Cas9D10A and scramble gRNAs: 17.1% ± 3.1, controls: 25.3% ± 3.2; 188–190 hpf—Cas9D10A and scramble gRNAs: 31.5% ± 3.8, controls: 30.8% ± 3.4, $P > 0.05$, Tables S1–S3, Appendix). Two electroporation sessions with scramble gRNAs also produced similar results to controls (164–166 hpf—Cas9D10A and scramble gRNAs: 17.7% ± 2.6, controls: 25.3% ± 3.2; 188–190 hpf—Cas9D10A and scramble gRNAs: 28.2 ± 3.1, controls: 30.8% ± 3.4, $P > 0.05$, Tables S1–S3, Appendix). Therefore, one or two electroporation sessions with Cas9D10A and scramble gRNAs did not reduce blastocyst yield and maintained survival like that seen in nonelectroporated embryos.

One electroporation session with Cas9D10A and *OCT4*-targeting sgRNAs reduced the blastocyst yield relative to scramble or control groups (164–166 hpf—Cas9D10A and targeting sgRNAs: 6.8% ± 1.4, controls: 25.3% ± 3.2; 188–190 hpf—Cas9D10A and targeting sgRNAs: 11.6% ± 1.7, controls: 30.8% ± 3.4, $P < 0.001$, Tables S1–S3, Appendix). Two electroporation sessions with Cas9D10A and *OCT4*-targeting sgRNAs also reduced blastocyst development (164–166 hpf—Cas9D10A and targeting sgRNAs: 3.2% ± 0.7, controls: 25.3% ± 3.2; 188–190 hpf—Cas9D10A and targeting sgRNAs: 7.9% ± 1.1, controls: 30.8% ± 3.4, $P < 0.001$, Tables S1–S3, Appendix).

We also evaluated zygotes electroporated twice with Cas9D10A and *OCT4*-targeting sgRNAs ($N = 56$), transferred into individual drops of media at the 8-cell stage and placed in a time-lapse incubator, along with controls that were not electroporated ($N = 28$). A greater number of electroporated embryos arrested their development at the 8-cell (35.5 vs. 17.8% controls, $P = 0.0013$) and morula (51.8 vs. 35.7% controls, $P = 0.0013$) stages. Additionally, a lower proportion of the electroporated embryos developed to the

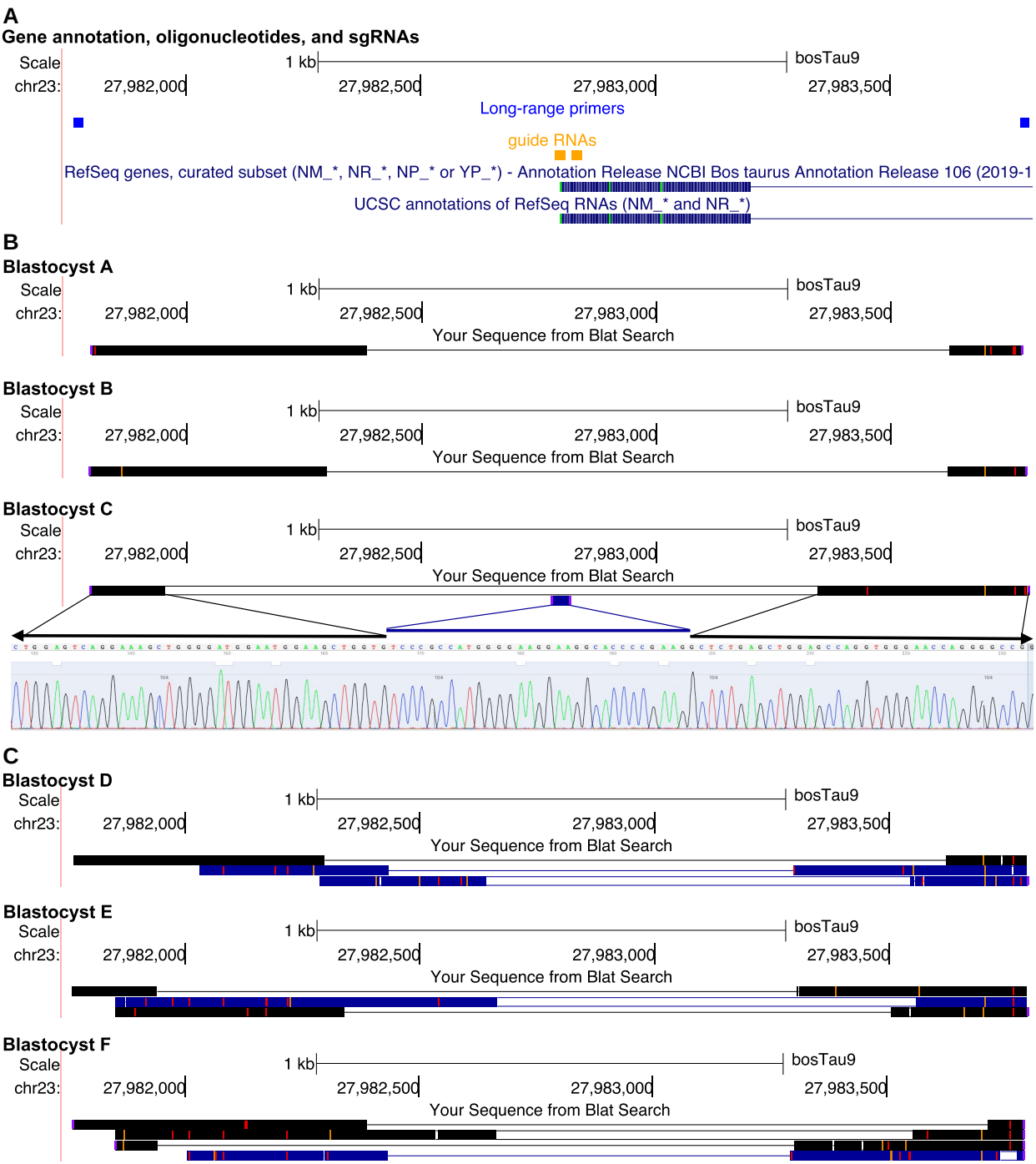

Fig. 2. Representative schematics of DNA mapping from fully edited blastocysts produced by two sessions of electroporation with Cas9D10A and *OCT4*-targeting sgRNAs. (A) Genome annotation identifying sgRNA targets and sequencing primers. (B) Sanger and (C) Nanopore targeted sequencing. Annotation obtained from the Genome Browser.

blastocyst stage (12.5 vs. 46.4% controls, $P = 1.06 \times 10^{-7}$, exact binomial test). Thus, targeting the gene *OCT4* by two electroporation sessions of Cas9D10A and sgRNAs caused partial developmental arrest at the 8-cell and morula stages with a sharp decline in the development to the blastocyst stage but did not eliminate embryo survival.

mRNA knockdown in cattle zygotes by Cas13a

To test whether Cas13a can target mRNAs in zygotes, first, we electroporated PZ with exogenous mRNAs of fluorescent proteins (red [RFP] or green [GFP]). Fluorescence imaging of embryos ~70 hpf showed successful introgression of exogenous mRNAs (GFP and RFP mRNAs) into PZ and expression of the corresponding proteins in cleavage embryos (Fig. 3). In contrast, we quantified a significant reduction of fluorescence (1.37-fold for GFP, and 1.34-fold for RFP, $P < 0.001$, Fig. 3) when we electroporated PZ with the exogenous mRNA and Cas13a + targeting sgRNAs simultaneously. As those PZ treated with Cas13a + targeting sgRNAs did not target an endogenous RNA, we tested whether Cas13a RNPs would impact embryo development. There were no statistical differences in the development to blastocyst stage at 188–190 hpf

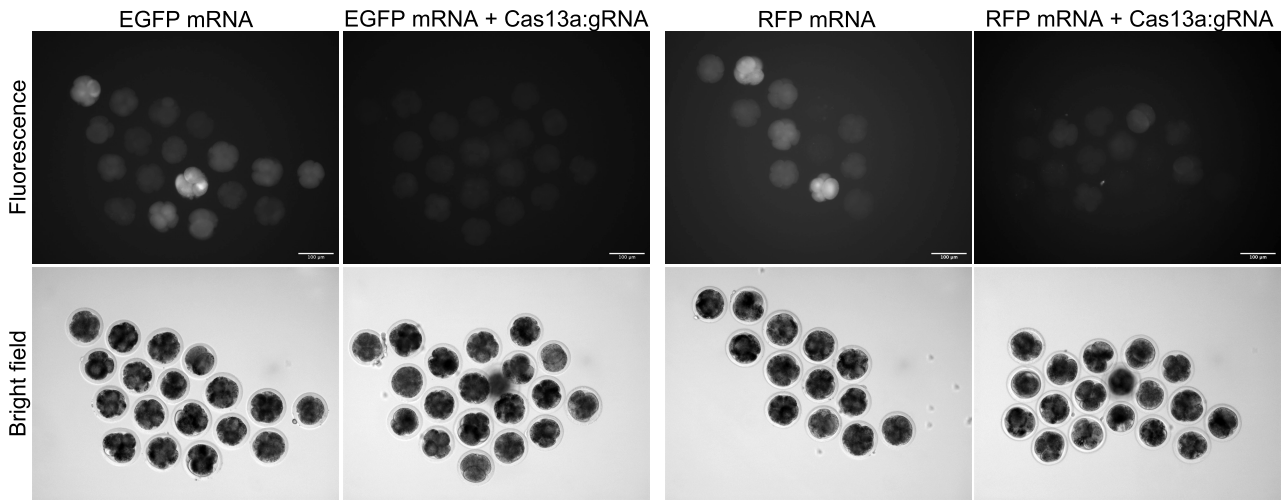

Fig. 3. Knockdown activity of Cas13a in cleavage cattle embryos. Scale bar: 100μm.

(30.8, 37.7, 32.1, and 33.3% for Cas13a + GFP mRNA sgRNAs, Cas13a + RFP mRNA sgRNAs, Cas13a + scramble gRNAs, controls, respectively, $P > 0.8$). Thus, Cas13a targets mRNAs efficiently in cattle zygotes with no alteration in their developmental potential.

Ablation of OCT4 function in cattle preimplantation embryos by CRISPR-DART

We used CRISPR-DART (Fig. 4) to target the promoter (based on orthology with the human genome) and exon 1 of *OCT4*. The induced deletions significantly reduced embryo survival (164–166 hpf—CRISPR-DART: 6.1% ± 0.8, controls: 23% ± 2.3; 188–190 hpf—CRISPR-DART: 13% ± 1.2, controls: 31.6% ± 2.5, $P < 0.001$, Tables S4 and S5, Appendix). Using immunofluorescent staining, we determined that the putatively edited blastocysts (we estimated 91% editing success) did not produce OCT4 protein. Additionally, we detected a decrease of NANOG in the edited blastocysts (Fig. 5A, see Text S1, Appendix for methods). Thus, we confirmed that the deletion of the promoter and exon 1 of *OCT4* resulted in absence of OCT4 protein.

Morphological examination showed an absence of a well-defined inner cell mass in blastocysts deemed *OCT4*$^{-/-}$, whereas a well-defined inner cell mass is clearly visible in the control embryos (Fig. 5B). Time-lapse image analysis of the development of putative *OCT4*$^{-/-}$ embryos confirmed the formation of a blastocoel cavity and absence of a normal inner cell mass (Movies S1 and S2). Next, we interrogated the transcriptome of *OCT4*$^{-/-}$ blastocysts. To confirm that the blastocysts collected were *OCT4*$^{-/-}$, we obtained genomic DNA and total RNA from single embryos. We used the genomic DNA to confirm the absence of the targeted DNA sequence (Fig. 5C) and evaluated the transcript abundance of 14,156 protein-coding or long noncoding genes from 5 *OCT4*$^{-/-}$ blastocysts. Comparative analyses revealed 125 genes with differential transcript abundance between *OCT4*$^{-/-}$ blastocysts and controls (Fig. 5D, False Discovery Rate [FDR] < 0.1). Eighty-three and 42 genes had lower and greater transcript abundance in *OCT4*$^{-/-}$ blastocysts, respectively. Notably, 17 genes with differential transcript abundance were functionally related to the maintenance of pluripotency (see Fig. 5E, for examples). These results indicate that a blastocoel cavity may form in the absence of *OCT4*, but the formation of the inner cell mass and normal gene expression is severely impaired in cattle *OCT4*$^{-/-}$ embryos.

Discussion

We developed an approach using Cas9D10A to delete targeted regions of the DNA and Cas13a to cleave targeted RNA for complete disruption of gene function in cattle zygotes at high efficiency. We used CRISPR-DART to target *OCT4* mRNAs and exon 1, including the promoter region. Our data provide several insights into the function of *OCT4* in cattle preimplantation development. First, most *OCT4*$^{-/-}$ embryos arrest development before the blastocyst stage, but a minor proportion of edited zygotes do still survive. Second, *OCT4*$^{-/-}$ embryos that progress their development are able to form a blastocoel cavity with an outer layer of cells resembling trophectoderm but do not form an inner cell mass with similar morphology observed in control embryos. Finally, the ablation of *OCT4* significantly alters the transcript abundance of genes involved in pluripotency. Our results show that *OCT4* is necessary for the development of a cattle blastocyst with a morphologically normal inner cell mass.

Simultaneous deletion of DNA segments and cleavage of RNA in zygotes

Previous research has reported the use of CRISPR-Cas9 to delete DNA segments in cattle zygotes (32, 35–38). To build on this, we tested the efficacy of Cas9 and Cas9D10A with two sgRNAs targeting the exon 1 of *OCT4*. Although we did not test for off-targets, Cas9D10A produces single-strand DNA breaks and requires two sgRNAs targeting opposite strands to nick the DNA and induce faulty DNA repair (60). This combination of factors significantly reduces mutation elsewhere in the genome. Our results confirmed that Cas9D10A RNPs produce large deletions beyond the region flanked by the sgRNAs (58). However, we only detected deletions larger than 500 nucleotides when we electroporated the zygotes twice in an interval of 6 h between sessions. We expected that two electroporation sessions allow for the introduction of greater quantities of RNPs in the zygote without causing toxicity (61). The combination of Cas9D10A targeting two sequences in the genome, and likely a higher quantity of RNPs entering the cell in two sessions of electroporation, increased the efficiency in producing full edits from 25.9 to 91%, which is higher than previous reports in cattle zygotes (32, 35–38). The timing of electroporation sessions was likely another factor that aided the increased efficiency. The influx of RNPs happened when most PZs were undergoing the

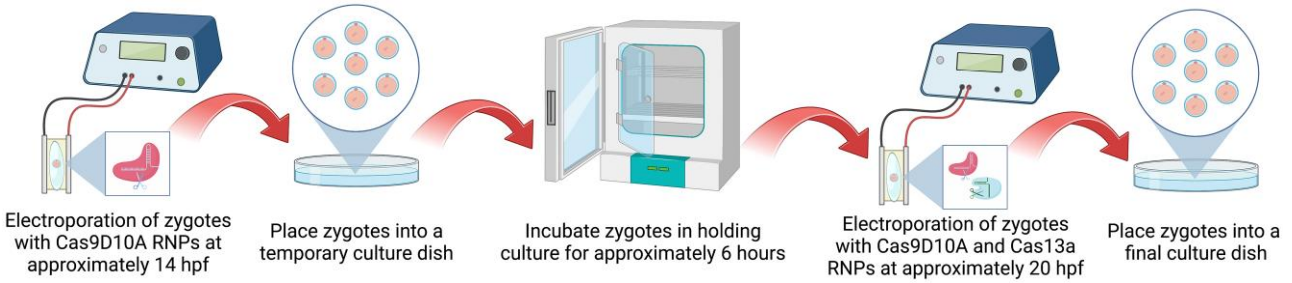

Fig. 4. Schematic of CRISPR-DART procedure. Created with BioRender.com.

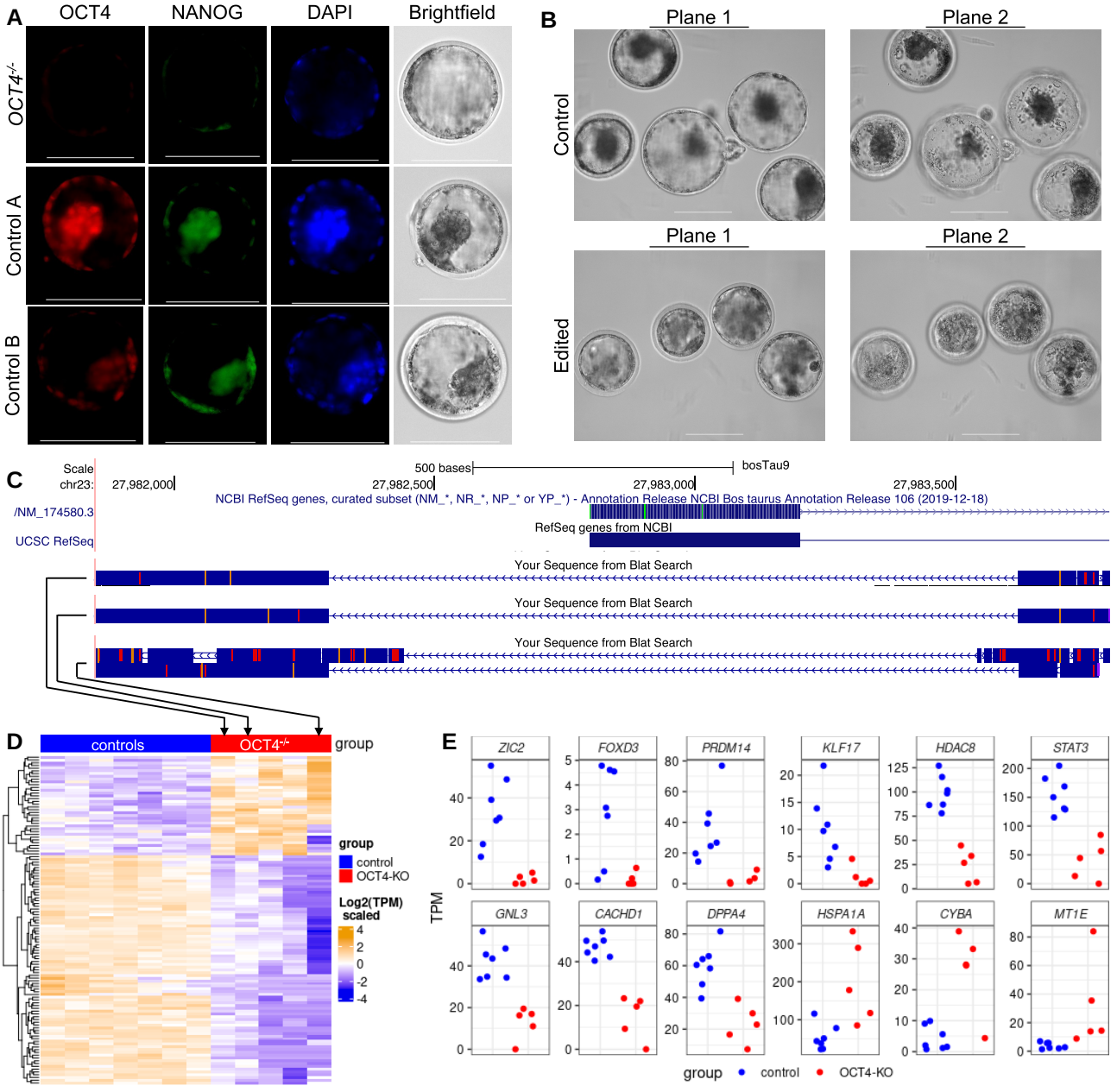

Fig. 5. Impact of *OCT4* KO in cattle preimplantation embryos. (A) Immunofluorescence assay of OCT4 and NANOG in cattle preimplantation embryos. Scale bar: 100 µm. (B) In vitro produced blastocysts 188–190 hpf. Images are presented in two focal planes for the visualization of the inner cell mass and blastocoel cavity. Scale bar: 100 µm. (C) Schematics of the DNA sequence mapping from three of five blastocysts used for RNA-Sequencing. (D) Heatmap depicting the relative differential transcript abundance of 125 genes in *OCT4* KO blastocysts. (E) Transcript abundance of 12 genes functionally associated with the maintenance of pluripotency.

S-phase (62), when the chromatids had to be less condensed (63), thus allowing more accessibility to RPNs.

The RNPs produced by the combination of CRISPR-Cas13a and an sgRNA can target and cleave single-stranded RNAs (64, 65). These RNPs have been used in animal embryos including in mouse (66) and pig (66) to target specific mRNAs. Here, we tested the efficacy of Cas13a in cattle zygotes by introducing and targeting mRNAs for either GFP or RFP. Our experiments showed that Cas13a could efficiently prevent protein synthesis from targeted mRNAs in cattle zygotes. One concern related to Cas13a is that it may cleave unintended mRNAs in the vicinities of targeted RNAs in a cell-dependent manner (67). The introduction of Cas13a + sgRNAs targeting exogenous mRNAs (GFP or RFP) and the corresponding mRNAs into cattle zygotes did not reduce embryo survival, thus, if there are off-target effects, they are negligible in cattle preimplantation embryos. Cas13a can knockdown specific mRNAs in zygotes in conjunction with Cas9D10A to target genomic DNA.

Effects of ablation of *OCT4* in cattle embryos

Our CRISPR-DART approach efficiently deleted exon 1 and the promoter from the *OCT4* in most embryos. We expected that removing the promoter and transcript starting site would impair the production of *OCT4* mRNAs and proteins. Indeed, we confirmed that most of the electroporated embryos tested by immunofluorescence assays did not have detectable OCT4. However, our RNA-sequencing data, produced from confirmed *OCT4*$^{-/-}$ embryos, showed sequences aligning with all five exons of *OCT4* (Fig. S5A, Appendix). The alignment of RNA sequencing reads to the cattle reference genome on the exon 1 of *OCT4* is conflicting with a confirmed deletion on that region, and we reasoned that a processed pseudogene (68, 69) produced RNAs that were sequenced and mapped to the reference genome on the *OCT4* gene. Few lines of evidence support our rationale. We only selected sequences that mapped to the reference genome once to quantify transcript abundance. Messenger RNAs produced by this *OCT4* pseudogene would have mapped to both genomic regions, but this region is not present in the current cattle reference genome, as indicated by the comparative mapping in the UCSC genome browser (Fig. S5B, Appendix). Annotated *OCT4* pseudogenes in human and mice are long noncoding RNAs (70–75) and do not contain introns (Fig. S5B, Appendix). Lastly, one hallmark observation in our data is that no RNA sequencing read produced from edited embryos mapped to *OCT4* intronic regions in the reference genome (Fig. S5A, Appendix), whereas several intronic sequences from all introns were evident in the control embryos (Fig. S5C, Appendix). Thus, we concluded that the RNA sequences mapping to the annotated *OCT4* cattle gene for *OCT4*$^{-/-}$ embryos were transcribed from a pseudogene with no intron. This confounding factor between functional *OCT4* and pseudogenes has been long observed in stem-cell research (72).

The ablation of *OCT4* function in cattle preimplantation embryos severely reduced blastocyst development, but a majority of the blastocysts were confirmed to be fully edited. This finding aligns with reports that produced embryos from *OCT4*$^{-/-}$ somatic cells (55, 56) or produced putative *OCT4*$^{-/-}$ embryos by introducing RNPs into zygotes (35, 36). The major blastocyst phenotype we observed was the absence of an inner cell mass, a phenotype previously reported in KO mice (Pou5f1^{tm1Cgre}/Pou5f1^{tm1Cgre}, genotype id MGI:3040797 (48)). By comparison, Simmet et al. (56) showed that *OCT4* is necessary only for the second-lineage differentiation, and it is possible that dysregulated genomic reprogramming due to somatic-cell cloning could be the cause of the minor discrepancy in the phenotypes. Our results show that *OCT4* is required for the differentiation of inner cell mass in cattle embryos.

We also evaluated the transcriptomic profile of *OCT4*$^{-/-}$ blastocysts. Our results, contrasting the transcriptome of *OCT4*$^{-/-}$ in vitro produced blastocysts and controls, showed 125 genes with significant alteration in transcript abundance. Only seven genes overlapped with the dataset reported elsewhere (*ARF5*, *MT1E*, *NUPR1*, *PLA2G15*, *RRM2*, *STAT3*, and *SWAP70* (56)), but all seven genes had the same direction of altered transcript abundance. Among the genes up-regulated in *OCT4*$^{-/-}$ blastocysts, *MT1E* potentiates cell differentiation (76, 77), and *NUPR1* is activated when the cells are under oxidative stress (78). Conversely, among the genes with lower transcript abundance in *OCT4*$^{-/-}$ blastocysts, *RRM2* is a known marker of pluripotent stem cells (79), *STAT3* is required for embryonic stem-cell pluripotency (80), and the absence of *SWAP70* impairs the self-renewal of mouse hematopoietic stem cells (81).

Our results also highlighted a dysregulation in other genes with roles in regulating stem-cell function. For example, *DPPA4* and *CADHD1* are pluripotent stem-cell markers (79) and have a lower abundance of transcripts in *OCT4*$^{-/-}$ blastocysts. Also, with a lower abundance of transcripts in edited embryos, *ZIC2* (82, 83), *KLF17* (84, 85), *FOXD3* (86), *HDAC8* (86), and *GNL3* (87) are directly involved in the regulation of pluripotency. In contrast, *ADAM9* (88), *ANXA8L1* (88), *CDKN1A* (89), *DDAH2* (90), *ELF3* (90), *HOXC10* (91), and *LGI1* (91) are all up-regulated in *OCT4*$^{-/-}$ blastocysts and promote cellular differentiation. Collectively, these results show a severe imbalance in the regulation of genes associated with stemness or cell differentiation, which is coherent with the absence of an inner cell mass in *OCT4*$^{-/-}$ blastocysts.

Limitations of the study

This study has limitations, mostly related to the editing efficiency and the inherent biology of DNA synthesis in zygotes. First, several technical factors, such as the concentration of RNP and electroporation conditions, may be fine-tuned to improve the efficiency of producing fully edited embryos. Here, we based our technical conditions on previous literature (34, 35), but we are confident there is room for improvement. Second, although matured oocytes are exposed to sperm in the fertilization media simultaneously, there is a window of opportunity for oocytes to be fertilized. Thus, among hundreds of oocytes, the timing of fertilization is heterogeneous. Third, following fertilization, both pronuclei are formed of compacted chromatin (92, 93), in which the RNPs are likely not accessible to the targeted sequence. Fourth, the unwinding of the DNA for synthesis is an asynchronous process across zygotes, and DNA synthesis can happen in a window of ~10 h (32, 62). Therefore, there is tremendous variability in the accessibility of the targeted DNA across zygotes. Fifth, we did not sequence the DNA of embryos that had deletions larger than the DNA sequence flanked by our oligonucleotides. We made several attempts to amplify very large fragments of 6 and 8 kb long, but nonspecific amplification hindered our ability to genotype the embryos accurately. Lastly, we had a relatively small sample size for the transcriptome analysis (five *OCT4*$^{-/-}$ and seven control blastocysts), but we countered this limitation with stringent analysis and scrutiny of the results based on the literature. Despite these limitations, our results show improvement in the efficiency of producing fully edited embryos, and the phenotype observed is coherent with the literature in mice and cattle.

Conclusion

The production of KOs is essential for mechanistic studies of gene function in preimplantation embryos. We showed that Cas9D10A is more efficient than Cas9 at producing biallelic deletions in zygotes. Two sessions of electroporation introduce greater quantities of Cas9D10A RNPs and increase the frequency of large biallelic deletions. The sequential introduction of RNPs does not impair embryo development as long sgRNAs targeting proximal sequences in the genome are not used. RNPs consisting of Cas13a prevent protein production from targeted mRNAs in cattle zygotes. Our CRISPR-DART approach increased the efficiency of producing KO zygotes. Lastly, we show that *OCT4* is required for the regulation of several genes that control pluripotency and the formation of an inner cell mass in cattle blastocysts.

Materials and methods

No live animal was handled for this study, thus no approval was required from the Institutional Animal Care and use Committee.

In vitro production of embryos

Unless otherwise specified, all reagents were purchased from Sigma-Aldrich.

All procedures and culture media composition for in vitro production of embryos are described in detail elsewhere (94, 95). Briefly, we purchased cattle ovaries from an abattoir (Brown Packing, Gaffney, SC) and washed them with antibiotic antimycotic (Antibiotic-Antimycotic 100×; Thermo Fisher Scientific, Waltham, MA, USA) and 0.9% saline solution. For the collection of cumulus–oocyte complexes (COCs), we aspirated ovarian follicles 3–8 mm in diameter using an 18G needle (Single-Use Needles BD Medical; VWR, Philadelphia, PA, USA) connected to a regulated vacuum system and collection bottle containing oocyte collection medium (OCM; BoviPlus Oocyte Collection Medium, Minitube, Verona, WI, USA) supplemented with gentamicin (50 µg/µL) and heparin (2 U/mL). We washed COCs twice in OCM, followed by three washes in oocyte maturation medium (OMM). Then we selected COCs with homogeneous, nongranular oocyte cytoplasm and three or more compact layers of cumulus for in vitro maturation. COCs were placed in groups of 10 in 50 µL of OMM covered by light mineral oil. In vitro maturation plates were incubated for 22–24 h at 38.5°C and 5% CO_2 in humidified atmosphere. Following the incubation, we washed the mature COCs in synthetic oviductal fluid (SOF) medium, containing N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid (HEPES; Thermo Fisher Scientific) and SOF for fertilization before transferring into a final fertilization plate (100 COCs/mL). We thawed frozen semen straws and processed sperm prior to transfer into fertilization plates at a concentration of 1,000,000 spermatozoa/mL. COCs and spermatozoa were co-incubated for 12–13 h under the same conditions described for in vitro maturation.

We removed PZs from fertilization medium at ∼14 hpf and denuded the cumulus cells by vortexing in 1% hyaluronidase for 5 min. Next, we moved PZ through three washes of SOF-HEPES and SOF culture medium (SOF-BE1). The PZs used for control groups were placed in their final culture dish immediately after the washes. Alternatively, the PZs used for electroporation were placed in temporary culture dishes, containing 50 µL SOF-BE1 covered with light mineral oil. After electroporation, we washed the PZs in SOF-BE1 before placing them in culture. PZs were cultured in groups of 25–30 in 50 µL SOF-BE1 covered by light mineral oil, incubated at 38.5°C with 5% CO_2, 5% O_2, and 90% N_2 in a humidified Eve Benchtop Incubator (WTA, College Station, TX, USA).

For time-lapse image analysis, we cultured 8-cell embryos individually in 15 µL SOF-BE1 covered by light mineral oil, incubated at 38.5°C with 5% CO_2, 5% O_2, and 90% N_2 in a MIRI Time-Lapse Incubator (Esco Medical, Egaa, Denmark).

Guide RNA design

We designed sgRNAs to target the genomic DNA of the transcriptional start site and exon 1 of *OCT4* using the CRISPOR web service (96). We designed the sgRNAs for Cas13a using New York Genome's cas13designtool (97, 98) to target the fourth exon of the *OCT4* mRNA. As an independent layer of in silico validation, we aligned all sgRNAs targeting the *OCT4* gene or transcript to the bovine genome with the BLAT software in the UCSC Genome Browser (99). Additionally, Cas13a sgRNAs were designed to target CleanCap EGFP and mCherry mRNAs (5moU, TriLink Biotechnologies, San Diego, CA, USA). The targeting sgRNAs used in this study were OCT4 sgRNA1: CTTCGCCTTCTCGCCCCCGCCGG, OCT4 sgRNA2: TGTCCCGCCATGGGGAAGGAAGG, OCT4 mRNA sgRNA: ATGCTCTCCAGGTTGCCTCT, mCherry mRNA sgRNA: TCCTCGAAGTTCATCACCCG, EGFP mRNA sgRNA: CATGATATAGACGTTGTGG. We purchased all sgRNAs as a single RNA molecule comprised of both crRNA and tracrRNA sequences (Integrated DNA Technologies [IDT], Research Triangle Park, NC, USA). We also purchased a scramble gRNA (Alt-R CRISPR-Cas9 Negative Control crRNA #1) and tracrRNA (Alt-R CRISPR-Cas9 tracrRNA) from IDT and combined them following the manufacturer's instructions.

Preparation of RNP and procedures for electroporation

We mixed Cas9 and sgRNAs for the formation of RNPs in Opti-MEM reduced serum medium (Thermo Fisher Scientific), and maintained the solution at room temperature for at least 30 min prior to electroporation. The specific concentrations and enzymes are detailed below.

As detailed above for control cultures, we removed the cumulus cells from the PZ and placed them in holding SOF-BE1 at 38.5°C, 5% CO_2, 5% O_2, and 90% N_2. We removed PZs in groups of 30–40 from a holding culture and briefly washed them in Opti-MEM (previously equilibrated in the incubator at 38.5°C and 5% CO_2). Next, we mixed 3 µL of the solution containing RNPs with 3 µL of Opti-MEM containing PZs. We carried out the electroporation using a BTX oocyte petri dish with platinum electrodes (Harvard Apparatus; VWR). We transferred the final 6 µL to the electroporation chamber. Impedance was checked and, if necessary, adjusted to measure between 0.19 and 0.20 by the addition of Opti-MEM or removal of the electroporation solution. The electroporation parameters were as follows: 6 poring pulses of 15 V, with 10% decay, for 2 ms with a 50-ms interval, immediately followed by 5 transfer pulses of 3 V, 40% decay, for 50 ms with a 50-ms interval, alternating the polarity. Following the electroporation, we washed the PZ with Opti-MEM and SOF-BE1.

Cleavage assay of the targeted DNA

We carried out a cleavage assay to assess the formation and cleavage of DNA by RNPs (100). We amplified a segment of genomic cattle DNA to be targeted by the sgRNAs by assaying a PCR using the following oligonucleotides (forward: GGCAAGGAACTTGATGC ACG and reverse: TGGCCAACCCACTGTTTGAT). The PCR reaction mix consisted of 0.2 IU/µL Phusion Hot Start II DNA Polymerase

(Thermo Fisher Scientific), 1× Phusion HF Buffer, 200 μM dNTPs (Promega, Madison, WI, USA), and forward and reverse oligonucleotides (IDT, Coralville, IA, USA) at 0.10 μM each, in a final volume of 20 μL in 0.2 mL clear PCR tubes. The cycling conditions for this reaction were: 98°C for 1 min, followed by 40 cycles of 98°C for 15 s, 55°C for 45 s, and 72°C for 1 min, followed by a final extension of 4 min at 72°C.

We incubated Cas9 (1 μM; IDT) with either sgRNA1 or sgRNA2 300 nM in Opti-MEM for 30 min at room temperature to form the RNPs. Next, we incubated RNPs with DNA fragments containing the targeted sequence (1:10 [v:v] Cas9 + sgRNA, 3 nM DNA, 1× NEB buffer 3.1) at 37°C for 3 h. Fragments were assessed by electrophoresis on a 1.5% Agarose I gel followed by staining with Diamond Nucleic Acid Dye and imaging.

Evaluation of electroporation efficiency

We evaluated the electroporation efficiency with RNPs formed by Cas9-RFP (Alt-R S.p.Cas9-RFP V3; IDT) at 800 ng/μL and scramble gRNAs at 800 ng/μL. After washing the PZs in Opti-MEM, we imaged them using a fluorescent microscope (details below).

Assessment of sequence deletions by Cas9 or Cas9D10A

To test the pattern of deletions with either a double-cutting enzyme or a nickase, we carried out a single electroporation at ~15 hpf with RNPs formed by either Cas9 or Cas9D10A (IDT) at 800 ng/μL and sgRNAs at 800 ng/μL each. After washing the PZ in SOF-BE1, we placed them in culture as described for control PZs.

Assessment of mRNA cleavage by Cas13a

We carried out a single electroporation of PZs with one of the following solutions: (i) mRNA of either mCherry or GFP at 400 ng/μL; or (ii) mRNA of either mCherry or GFP at 400 ng/μL and RNP formed by Cas13a (GenScript, Piscataway, NJ, USA) at 400 ng/μL and the corresponding targeting sgRNA at 400 ng/μL. After washing the PZ in Opti-MEM, we imaged them using a fluorescent microscope (details below) in SOF-HEPES.

CRISPR-DART

For CRISPR-DART, we carried out the first electroporation at ~14 hpf with 3 μL of RNPs formed by Cas9D10A at 600 ng/μL and sgRNAs at 800 ng/μL each mixed with 3 μL of Opti-MEM. The PZs were maintained in SOF-BE1 media in the incubator. Then, we electroporated them again at ~20 hpf with two solutions of RNP complexes prepared separately. One solution contained Cas9D10A at 600 ng/μL and each sgRNA at 800 ng/μL and the other contained Cas13a at 1,600 ng/μL and sgRNA at 800 ng/μL. At the time of electroporation, we mixed 1.5 μL of each RNP with 3 μL of Opti-MEM containing the PZ. After washing the PZ in SOF-BE1, we placed them in culture as described for control PZs.

Targeted DNA sequencing

All embryos collected for DNA sequencing were washed in Phosphate Buffered Saline (PBS) 0.1% Bovine Serun Albumin (BSA) fraction V, followed by removal of the zona pellucida by exposure to EmbryoMax Acidic Tyrode's Solution and gentle pipetting. Once the zona pellucida was removed, we washed the embryos in PBS 0.1% BSA fraction V twice and collected them individually in microtubes in ~1 μL PBS 0.1% BSA fraction V. We exposed the nucleic acids of each embryo with 5 μL of QuickExtract DNA Extraction Solution (Lucigen; VWR), and incubated at 65°C for 15 min followed by 2 min at 98°C.

High-throughput short reads

We assayed a PCR using oligonucleotides flanking the targeted deletion site with coupled sequencing adapters on their 5′-end (forward: acactctttccctacacgacgctcttccgatctAGAGGTGTTGAGCA GTCTCTAGG, reverse: gtgactggagttcagacgtgtgctcttccgatctGTAG GCCATCCCTCCACAC; lower case letters indicate adapter, and uppercase letters indicate targeted sequence). The PCR consisted of 0.2 IU/μL Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific), 1× Phusion HF Buffer, 200 μM dNTPs (Promega), and oligonucleotides (IDT) at 0.1 μM each, in a final volume of 20 μL. Reactions were carried out in 0.2 mL clear PCR tubes (VWR), and the cycling conditions were: 98°C for 1 min, followed by 40 cycles of 98°C for 15 s, 61°C for 30 s, and 72°C for 40 s, proceeding a final extension of 4 min at 72°C. We confirmed the amplification using 2% Agarose I (VWR) and gel electrophoresis, followed by DNA staining with Diamond Nucleic Acid Dye (Promega).

Next, we completed the library preparation with a second PCR using oligonucleotides obtained from xGen UDI Primers Plate 1, 8nt (IDT). The reaction mixture consisted of 0.3 IU/μL Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific), 1× Phusion HF Buffer, 200 μM dNTPs (Promega), and 3 μM Illumina adaptors, in a final volume of 25 μL. The reaction was assayed according to the following conditions: 98°C for 30 s, followed by 15 cycles of 98°C for 10 s, 60°C for 30 s, and 72°C for 30 s, proceeding a final extension of 5 min at 72°C. We pooled the amplicons and size-selected the targeted fragments using a 2% Invitrogen UltraPure Low Melting Point Agarose gel (Thermo Fisher Scientific) followed by a purification using the Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA, USA). The libraries were sequenced at the Vanderbilt Technologies for Advanced Genomics using a NovaSeq 6000 System (Illumina, Inc, San Diego, CA, USA) to produce pair-end reads 150 nucleotides long.

We processed the fastq files using an in-house bioinformatic pipeline similar to one published elsewhere (101). We only proceeded with reads #2 because it spanned our targeted DNA region. First, we used trimmomatic v.0.39 (102) to remove the sequencing adapters and filtered reads to retain those with a minimum length of 100 nucleotides and a minimum average quality score of 35. Then, we used clumpify.sh from BBTools (https://sourceforge.net/projects/bbmap/) to remove duplicates. Lastly, we converted the file format from fastq to fasta using seqtk (103).

High-throughput long reads

We produced amplicons by assaying a PCR using the following oligonucleotides (forward: GGCAAGGAACTTGATGCACG and reverse: TGGCCAACCCACTGTTTGAT). The PCR reaction mix consisted of 0.2 IU/μL Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific), 1× Phusion HF Buffer, 200 μM dNTPs (Promega), and forward and reverse oligonucleotides (IDT, Coralville, IA, USA) at 0.10 μM each, in a final volume of 20 μL in 0.2 mL clear PCR tubes. The cycling conditions for this reaction were: 98°C for 1 min, followed by 40 cycles of 98°C for 15 s, 55°C for 45 s, and 72°C for 1 min, followed by a final extension of 4 min at 72°C. We confirmed the amplification by assaying 5 μL of each amplicon by electrophoresis on a 1.5% Agarose I gel before staining with Diamond Nucleic Acid Dye and imaging. When the amplification produced an amplicon, we used the remaining PCR products to prepare sequencing libraries with the Native Barcoding Kit 24 V14 (Oxford Nanopore Technologies, Lexington, MA, USA), following the manufacturer's instructions. We sequenced the libraries on a MinION Mk1C (Oxford Nanopore Technologies).

We processed the fast5 files with Guppy (v 6.4.2) (104) using the configuration file dna_r10.4.1_e8.2_260bps_sup.cfg for super high accuracy base calling. Next, we used porechop (https://github.com/rrwick/Porechop) to remove adapters and used Fitlong (https://github.com/rrwick/Filtlong) to remove sequences smaller than 500 nucleotides long and with a quality of <90%. We aligned the remaining sequences to the cattle reference genome (ARS-UCD1.2) using minimap2 (v 2.24) (105), allowing for spliced alignment (parameters: -ax map-ont –splice -c –cs = long –secondary = no –sam-hit-only -Y –splice-flank = no -G2k). Finally, we used samtools (106) to remove alignments with <500 nucleotides mapped to the genome and supplementary alignments.

Sanger sequencing

We produced PCR amplicons using the procedures described for "high-throughput long reads." When the amplification produced an amplicon, we treated the remaining PCR products with 3 µL ExoSAP-IT Express PCR Product Cleanup Reagent (Thermo Fisher Scientific) and incubated at 37°C for 15 min followed by 80°C for 15 min. The sequencing assay was carried out by the Genomics Sequencing Center at Virginia Tech using the same forward oligonucleotide used for the initial PCR.

DNA and RNA sequencing of single embryos

We collected embryos for DNA and RNA sequencing on stage codes 6 or 7 (107). We washed in PBS 0.1% BSA fraction V, followed by removal of the zona pellucida by exposure to EmbryoMax Acidic Tyrode's Solution and gentle pipetting. Once the zona pellucida was removed, we washed the embryos in PBS 0.1% BSA fraction V twice and collected them individually in microtubes in ~1 µL PBS 0.1% BSA fraction V. We placed the tubes on dry ice and stored the samples at −80°C.

We lysed the embryos by adding 10 µL of lysis solution, consisting of: 8.3 µL Luna Cell Ready Lysis Buffer 2× (New England Biolabs, Ipswich, MA, USA), 0.66 µL Luna Cell Ready RNA Protection Reagent 25× (New England Biolabs), 0.66 µL Luna Cell Ready Protease 25× (New England Biolabs), and 0.33 µL RNasin Plus ribonuclease inhibitor (40 U/µL; Promega). We incubated the solution on ice for 10 min, mixed by pipetting, then split the solution into two tubes. One tube, dedicated to DNA sequencing was further incubated at 37°C for 15 min, followed by the addition of one µl of Luna Cell Ready Stop Solution 10× (New England Biolabs). We extracted DNA (108) by adding a solution of sodium acetate (5 M) for a final concentration of 2.5 M, followed by the addition of 150 µL of Ethanol 100%. We stored the solution −20°C for over 15 h and precipitated the DNA by centrifugation at 15,000xg for 20 min at 4°C. We washed the pellet twice with 150 µL ethanol 70% and eluted it with nuclease-free water. The DNA was used as template for amplification of the targeted region using the oligonucleotides and procedures for high-throughput long reads. Preparation of libraries, sequencing, and processing of sequences were carried out as described above.

We extracted total RNA and from each ½ blastocyst using TRIzol reagent with Phasemaker Tubes for enhanced RNA purity and yield (109–112). RNA was stored in 70% ethanol at −80°C (109) until library preparation. We assessed RNA integrity of samples not used for sequencing with a 2,100 Bioanalyzer Instrument (Agilent, Santa Clara, CA, USA) and RNA 6000 Pico Kit (Agilent). These tests require the total volume of extracted RNA; therefore, only test samples were assessed to ensure the quality and rigor of our procedures. We amplified cDNA using a modified mcSCRB-seq protocol and produced libraries using the Illumina DNA Prep kit (Illumina, Inc) (109, 113). The libraries were sequenced at the Vanderbilt Technologies for Advanced Genomics using a NovaSeq 6000 System (Illumina, Inc) to produce ~30 million pair-end reads 150 nucleotides long per sample.

We aligned the RNA-sequencing data to the cattle reference genome (114) (ARS-UCD1.2/bosTau9) obtained from the Ensembl database (115, 116) using HISAT2 (v2.2.1) (117), followed by filtering with samtools (v1.17) (106, 118) to remove alignment <100 nucleotides long and with >5% mismatch nucleotides, plus removal of duplicates with biobambam2 (v2.0.95) (119). Next, we counted the sequences matching the reference annotation (Bos_taurus.ARS-UCD1.2.105.gtf) using featureCounts (v2.0.1) (120).

Statistical analyses

The analytical procedures used to analyze differences in embryo development and differential transcript abundance, including the supplementary tables and pertinent graphs are available at: https://biase-lab.github.io/crispr_dart/.

Assessment of differences in embryo development

We recorded the number of embryos that developed to the blastocyst stage and the number of PZs with arrested development prior to blastocyst formation at 164–166 hpf and 188–190 hpf for each culture drop. Culture drop was considered biological replicate. We analyzed count data (success of blastocyst development or developmental arrest) using a general linear model with a binomial family, which results in logistic regression analysis, using the "glm" function from the R package "stats." We used the number of blastocysts and the number of PZs that failed to develop into blastocysts as the dependent variable, and the group (control, scramble, or Cas treated) was a fixed effect. The Wald statistical test was conducted with the function "ANOVA" from the R package "car" (121). Finally, we carried out a pairwise comparison using the odds ratio and two-proportion z-test employing the "emmeans" function of the R package "emmeans." The null hypothesis assumed that the odds ratio of the proportion (p) of two groups was not different from 1 (H_0: $p_1/p_2 = 1$). We adjusted the nominal P-value for multiple hypothesis testing with the Bonferroni approach and inferred significance when adjusted P-value <0.05.

We analyzed data obtained from single embryo culture, with each embryo as a biological replicate, using the exact binomial test in R with the function "binom.test" (122). Significance was inferred, if the $P < 0.05$.

Assessment of differences in fluorescence

First, we calculated corrected total cell fluorescence (CTCF) using the standard formula: Integrated density—(Area of selected cell × Mean fluorescence of background readings). We obtained the measurements necessary for the formula using the NIS-elements Imaging Software (v.5.02). Next, we fitted a linear model using the "lm" function of the R package "stats" where Log2(CTCF) was the dependent variable. Replicate and group (fluorescence protein mRNA or fluorescence protein mRNA + Cas13a and targeting sgRNA) were included as fixed effects. We assessed the significance of the variables using the Anova function of the R package "car." Next, we tested the pairwise significance of the two groups by a t-score test employing the "emmeans" function of the R package emmeans. The null hypothesis assumed that the difference between two averages ($\bar{x}$) was not different from zero (H_0: $\bar{x}_1 - \bar{x}_2 = 0$), and significance was inferred at alpha = 0.05.

Differential transcript abundance

In R software (123, 124), we created one matrix with the read counts for all samples and retained genes classified as protein-coding and long noncoding DNA for downstream analysis. We calculated counts per million (CPM) using the function "cpm" from the R package "edgeR" (125) and retained genes if CPM > 1 in 4 or more samples. We also calculated transcript per million as described elsewhere (126, 127) used for plotting the data. We estimated differential transcript abundance between edited and control blastocysts employing the quasi-likelihood negative binomial generalized log-linear model from the R package edgeR (125) and the Wald test from the R package "DESeq2" (128). We inferred statistically significant differences when False discovery rate (129) was <0.1 in both tests.

Acknowledgments

The authors thank Select Sires for the donation of semen straws used in this research.

Preprint

This manuscript was posted on a preprint: https://doi.org/10.1101/2023.07.07.548144

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This project was supported by Grant no. 2018-67015-31936 from the USDA National Institute of Food and Agriculture.

Author Contributions

J.L.N. performed the research, analyzed data, and co-wrote the paper. G.P.S.: analyzed data. S.L.S. and A.D.E. contributed with the methodology and resources. F.H.B. designed the research, performed the research, analyzed data, wrote the paper, and secured funding for the research.

Data Availability

The transcriptome data produced in this research is publicly available in the Gene Expression Omnibus repository under the accession GSE236474.

References

1 Kwon J, Namgoong S, Kim NH. 2015. CRISPR/Cas9 as tool for functional study of genes involved in preimplantation embryo development. *PLoS One*. 10:e0120501.
2 Ikonomov OC, *et al.* 2011. The phosphoinositide kinase PIKfyve is vital in early embryonic development PREIMPLANTATION LETHALITY OF PIKfyve(-/-) EMBRYOS BUT NORMALITY OF PIKfyve(+/-) MICE. *J Biol Chem*. 286:13404–13413.
3 Xiao Y, *et al.* 2021. Regulation of gene expression in the bovine blastocyst by colony-stimulating factor 2 is disrupted by CRISPR/Cas9-mediated deletion of CSF2RA. *Biol Reprod*. 104:995–1007.
4 Stamatiadis P, *et al.* 2022. TEAD4 Regulates trophectoderm differentiation upstream of CDX2 in a GATA3-independent manner in the human preimplantation embryo. *Hum Reprod*. 37:1760–1773.
5 Cosemans G, *et al.* 2022. CRISPR/Cas9 mediated knock-out (KO) reveals a divergent role for trophectoderm markers GATA2/3 in the mouse and human preimplantation embryo. *Hum Reprod*. 37:I128–I128.
6 Tu ZC, Yang WL, Yan S, Guo XY, Li XJ. 2015. CRISPR/Cas9: a powerful genetic engineering tool for establishing large animal models of neurodegenerative diseases. *Mol Neurodegener*. 10:35.
7 Motta BM, Pramstaller PP, Hicks AA, Rossini A. 2017. The impact of CRISPR/Cas9 technology on cardiac research: from disease modelling to therapeutic approaches. *Stem Cells Int*. 2017:1–13.
8 Whitelaw CBA. 2016. Engineering large animal models of human disease. *J Pathol*. 240:5–5.
9 Marr E, Potter CJ. 2021. Base editing of somatic cells using CRISPR-Cas9 in *Drosophila*. *CRISPR J*. 4:836–845.
10 Xie C, *et al.* 2016. Genome editing with CRISPR/Cas9 in postnatal mice corrects PRKAG2 cardiac syndrome. *Cell Res*. 26:1099–1111.
11 Swiech L, *et al.* 2015. In vivo interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat Biotechnol*. 33:102–106.
12 Vilarino M, *et al.* 2017. CRISPR/Cas9 microinjection in oocytes disables pancreas development in sheep. *Sci Rep*. 7:17472.
13 Sergio Navarro-Serna AH, *et al.* 2021. Two-pore channel 2 biallelic knockout pigs in one generation by CRISPR-Cas9 microinjection before oocyte insemination. *CRISPR J*. 4:132–146.
14 Delerue F, Ittner LM. 2017. Generation of genetically modified mice through the microinjection of oocytes. *J Vis Exp*. 124:55765.
15 Gim GM, *et al.* 2022. Germline transmission of MSTN knockout cattle via CRISPR-Cas9. *Theriogenology*. 192:22–27.
16 Morin V, Veron N, Marcelle C. 2017. CRISPR/Cas9 in the chicken embryo. *Methods Mol Biol*. 1650:113–123.
17 Wang B, *et al.* 2015. Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. *Biotechniques*. 59:201–202. 204, 206–208
18 Owen JR, *et al.* 2021. One-step generation of a targeted knock-in calf using the CRISPR-Cas9 system in bovine zygotes. *BMC Genomics*. 22:118.
19 Ryu J, Lee K. 2017. CRISPR/Cas9-mediated gene targeting during embryogenesis in swine. *Methods Mol Biol*. 1605:231–244.
20 Song Y, *et al.* 2016. Efficient dual sgRNA-directed large gene deletion in rabbit with CRISPR/Cas9 system. *Cell Mol Life Sci*. 73:2959–2968.
21 Zhu Z, Verma N, Gonzalez F, Shi ZD, Huangfu D. 2015. A CRISPR/Cas-mediated selection-free knockin strategy in human embryonic stem cells. *Stem Cell Rep*. 4:1103–1111.
22 Wang H, *et al.* 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*. 153:910–918.
23 Hai T, Teng F, Guo R, Li W, Zhou Q. 2014. One-step generation of knockout pigs by zygote injection of CRISPR/Cas system. *Cell Res*. 24:372–375.
24 Wan H, *et al.* 2015. One-step generation of p53 gene biallelic mutant Cynomolgus monkey via the CRISPR/Cas system. *Cell Res*. 25:258–261.
25 Park KE, *et al.* 2020. One-step homology mediated CRISPR-Cas editing in zygotes for generating genome edited cattle. *CRISPR J*. 3:523–534.
26 Jinek M, *et al.* 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 337:816–821.
27 Nerys-Junior A, Braga-Dias LP, Pezzuto P, Cotta-de-Almeida V, Tanuri A. 2018. Amilcar tanuri comparison of the editing

patterns and editing efficiencies of TALEN and CRISPR-Cas9 when targeting the human CCR5 gene. *Genet Mol Biol*. 41: 167–179.
28 He Z, Proudfoot C, Whitelaw CB, Lillico SG. 2016. Comparison of CRISPR/Cas9 and TALENs on editing an integrated EGFP gene in the genome of HEK293FT cells. *Springerplus*. 5:814.
29 Zhang J, *et al*. 2019. Comparison of gene editing efficiencies of CRISPR/Cas9 and TALEN for generation of MSTN knock-out cashmere goats. *Theriogenology*. 132:1–11.
30 Zhang ZB, *et al*. 2018. Comparing successful gene knock-in efficiencies of CRISPR/Cas9 with ZFNs and TALENs gene editing systems in bovine and dairy goat fetal fibroblasts. *J Integr Agric*. 17:1171–1180.
31 Vos PD, Filipovska A, Rackham O. 2022. Frankenstein Cas9: engineering improved gene editing systems. *Biochem Soc Trans*. 50: 1505–1516.
32 Lamas-Toranzo I, *et al*. 2019. Strategies to reduce genetic mosaicism following CRISPR-mediated genome edition in bovine embryos. *Sci Rep*. 9:14900.
33 Miskel D, *et al*. 2022. The cell cycle stage of bovine zygotes electroporated with CRISPR/Cas9-RNP affects frequency of loss-of-heterozygosity editing events. *Sci Rep*. 12:10793.
34 Lin JC, Van Eenennaam AL. 2021. Electroporation-mediated genome editing of livestock zygotes. *Front Genet*. 12:648482.
35 Camargo LSA, Owen JR, Van Eenennaam AL, Ross PJ. 2020. Efficient one-step knockout by electroporation of ribonucleoproteins into zona-intact bovine embryos. *Front Genet*. 11: 570069.
36 Daigneault BW, Rajput S, Smith GW, Ross PJ. 2018. Embryonic POU5F1 is required for expanded bovine blastocyst formation. *Sci Rep*. 8:7753.
37 Miao D, Giassetti MI, Ciccarelli M, Lopez-Biladeau B, Oatley JM. 2019. Simplified pipelines for genetic engineering of mammalian embryos by CRISPR-Cas9 electroporationdagger. *Biol Reprod*. 101:177–187.
38 Namula Z, *et al*. 2019. Genome mutation after the introduction of the gene editing by electroporation of Cas9 protein (GEEP) system into bovine putative zygotes. *In Vitro Cell Dev Biol Anim*. 55:598–603.
39 Namula Z, *et al*. 2022. Zona pellucida treatment before CRISPR/Cas9-mediated genome editing of porcine zygotes. *Veterinary Medicine and Science*. 8:164–169.
40 Mourot M, *et al*. 2006. The influence of follicle size, FSH-enriched maturation medium, and early cleavage on bovine oocyte maternal mRNA levels. *Mol Reprod Dev*. 73: 1367–1379.
41 Fair T, *et al*. 2004. Analysis of differential maternal mRNA expression in developmentally competent and incompetent bovine two-cell embryos. *Mol Reprod Dev*. 67:136–144.
42 Sha Q-Q, Zhang J, Fan H-Y. 2019. A story of birth and death: mRNA translation and clearance at the onset of maternal-to-zygotic transition in mammals†. *Biol Reprod*. 101: 579–590.
43 Zhang K, Smith GW. 2015. Maternal control of early embryogenesis in mammals. *Reprod Fertil Dev*. 27:880–896.
44 Abudayyeh OO, *et al*. 2017. RNA Targeting with CRISPR-Cas13. *Nature*. 550:280–284.
45 Kick LM, von Wrisberg MK, Runtsch LS, Schneider S. 2022. Structure and mechanism of the RNA dependent RNase Cas13a from Rhodobacter capsulatus. *Commun Biol*. 5:71.
46 Kurosaka S, Eckardt S, McLaughlin KJ. 2004. Pluripotent lineage definition in bovine embryos by Oct4 transcript localization. *Biol Reprod*. 71:1578–1582.
47 Khan DR, *et al*. 2012. Expression of pluripotency master regulators during two key developmental transitions: EGA and early lineage specification in the bovine embryo. *PLoS One*. 7:e34110.
48 Nichols J, *et al*. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*. 95:379–391.
49 Fogarty NME, *et al*. 2017. Genome editing reveals a role for OCT4 in human embryogenesis. *Nature*. 550:67–73.
50 Pan GJ, Chang ZY, Scholer HR, Pei D. 2002. Stem cell pluripotency and transcription factor Oct4. *Cell Res*. 12:321–329.
51 Sharma J, Antenos M, Madan P. 2021. A comparative analysis of hippo signaling pathway components during murine and bovine early mammalian embryogenesis. *Genes (Basel)*. 12:281.
52 Frum T, *et al*. 2013. Oct4 cell-autonomously promotes primitive endoderm development in the mouse blastocyst. *Dev Cell*. 25: 610–622.
53 Wu G, *et al*. 2013. Establishment of totipotency does not depend on Oct4A. *Nat Cell Biol*. 15:1089–1097.
54 Gerri C, *et al*. 2023. A conserved role of the Hippo signalling pathway in initiation of the first lineage specification event across mammals. *Development*. 150:dev201112.
55 Simmet K, *et al*. 2018. OCT4/POU5F1 is required for NANOG expression in bovine blastocysts. *Proc Natl Acad Sci U S A*. 115: 2770–2775.
56 Simmet K, *et al*. 2022. OCT4/POU5F1 is indispensable for the lineage differentiation of the inner cell mass in bovine embryos. *FASEB J*. 36:e22337.
57 Gopalappa R, Suresh B, Ramakrishna S, Kim HH. 2018. Paired D10A Cas9 nickases are sometimes more efficient than individual nucleases for gene disruption. *Nucleic Acids Res*. 46:e71.
58 Korablev A, Lukyanchikova V, Serova I, Battulin N. 2020. On-target CRISPR/Cas9 activity can cause undesigned large deletion in mouse zygotes. *Int J Mol Sci*. 21:3604.
59 Owens DDG, *et al*. 2019. Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res*. 47:7402–7417.
60 Ran FA, *et al*. 2013. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 154:1380–1389.
61 Ye S, Enghiad B, Zhao H, Takano E. 2020. Fine-tuning the regulation of Cas9 expression levels for efficient CRISPR-Cas9 mediated recombination in Streptomyces. *J Ind Microbiol Biotechnol*. 47:413–423.
62 Comizzoli P, Marquant-Le Guienne B, Heyman Y, Renard JP. 2000. Onset of the first S-phase is determined by a paternal effect during the G1-phase in bovine zygotes. *Biol Reprod*. 62: 1677–1684.
63 Levchenko SM, Pliss A, Peng X, Prasad PN, Qu J. 2021. Fluorescence lifetime imaging for studying DNA compaction and gene activities. *Light Sci Appl*. 10:224.
64 Abudayyeh OO, *et al*. 2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*. 353:aaf5573.
65 Cox DBT, *et al*. 2017. RNA Editing with CRISPR-Cas13. *Science*. 358:1019–1027.
66 Bi D, *et al*. 2021. CRISPR/Cas13d-mediated efficient KDM5B mRNA knockdown in porcine somatic cells and parthenogenetic embryos. *Reproduction*. 162:149–160.
67 Ai Y, Liang D, Wilusz JE. 2022. CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. *Nucleic Acids Res*. 50:e65.
68 Troskie RL, Faulkner GJ, Cheetham SW. 2021. Processed pseudogenes: a substrate for evolutionary innovation: retrotransposition contributes to genome evolution by propagating

pseudogene sequences with rich regulatory potential throughout the genome. *Bioessays*. 43:e2100186.
69 Zhang X, Wacker C, Schutz E, Brenig B. 2020. Processed pseudogene confounding the identification of a putative lethal recessive deletion in the bovine 60S ribosomal protein L11 gene (uL5). *Anim Genet*. 51:146–147.
70 Scarola M, *et al.* 2015. Epigenetic silencing of Oct4 by a complex containing SUV39H1 and Oct4 pseudogene lncRNA. *Nat Commun*. 6:7631.
71 Lin H, Shabbir A, Molnar M, Lee T. 2007. Stem cell regulatory function mediated by expression of a novel mouse Oct4 pseudogene. *Biochem Biophys Res Commun*. 355:111–116.
72 Liedtke S, Enczmann J, Waclawczyk S, Wernet P, Kögler G. 2007. Oct4 and its pseudogenes confuse stem cell research. *Blood*. 110: 1081a.
73 Yadav PS, Kues WA, Herrmann D, Carnwath JW, Niemann H. 2005. Bovine ICM derived cells express the Oct4 ortholog. *Mol Reprod Dev*. 72:182–190.
74 Salem NA, *et al.* 2021. A novel Oct4/Pou5f1-like non-coding RNA controls neural maturation and mediates developmental effects of ethanol. *Neurotoxicol Teratol*. 83:106943.
75 Schiffmacher A, Keefer C. 2008. Expression of an OCT4 (POU5F1) pseudogene in the bovine embryo derived CT-1 cell line. *Biol Reprod*. 78:100–100.
76 Sim EZ, *et al.* 2022. Methionine metabolism regulates pluripotent stem cell pluripotency and differentiation through zinc mobilization. *Cell Rep*. 40:111120.
77 Shiraki N, *et al.* 2014. Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. *Cell Metab*. 19:780–794.
78 Huang C, Santofimia-Castano P, Iovanna J. 2021. NUPR1: a critical regulator of the antioxidant system. *Cancers (Basel)*. 13:3670.
79 Mallon BS, *et al.* 2013. StemCellDB: the human pluripotent stem cell database at the National Institutes of Health. *Stem Cell Res*. 10:57–66.
80 Raz R, Lee CK, Cannizzaro LA, d'Eustachio P, Levy DE. 1999. Essential role of STAT3 for embryonic stem cell pluripotency. *Proc Natl Acad Sci U S A*. 96:2846–2851.
81 Ripich T, *et al.* 2016. SWEF Proteins distinctly control maintenance and differentiation of hematopoietic stem cells. *PLoS One*. 11:e0161060.
82 Luo Z, *et al.* 2015. Zic2 is an enhancer-binding factor required for embryonic stem cell specification. *Mol Cell*. 57:685–694.
83 Lim LS, *et al.* 2007. Zic3 is required for maintenance of pluripotency in embryonic stem cells. *Mol Biol Cell*. 18:1348–1358.
84 Lea RA, *et al.* 2021. KLF17 Promotes human naive pluripotency but is not required for its establishment. *Development*. 148: dev199378.
85 Wang SH, *et al.* 2022. KLF17 Promotes human naive pluripotency through repressing MAPK3 and ZIC2. *Sci China Life Sci*. 65:1985–1997.
86 Hua WK, *et al.* 2017. HDAC8 Regulates long-term hematopoietic stem-cell maintenance under stress by modulating p53 activity. *Blood*. 130:2619–2630.
87 Tsai RY, McKay RD. 2002. A nucleolar mechanism controlling cell proliferation in stem cells and cancer cells. *Genes Dev*. 16: 2991–3003.
88 Barbosa CMV, *et al.* 2019. Extracellular annexin-A1 promotes myeloid/granulocytic differentiation of hematopoietic stem/progenitor cells via the Ca(2+)/MAPK signalling transduction pathway. *Cell Death Discov*. 5:135.
89 Kreis NN, Louwen F, Yuan J. 2019. The multifaceted p21 (Cip1/Waf1/CDKN1A) in cell differentiation. Migration and cancer therapy. *Cancers (Basel)*. 11:1220.
90 Tran LL, Dang T, Thomas R, Rowley DR. 2021. ELF3 Mediates IL-1alpha induced differentiation of mesenchymal stem cells to inflammatory iCAFs. *Stem Cells*. 39:1766–1777.
91 Xie YJ, *et al.* 2018. Leucine-rich glioma inactivated 1 promotes oligodendrocyte differentiation and myelination via TSC-mTOR signaling. *Front Mol Neurosci*. 11:231.
92 Laurincik J, Kopecny V, Hyttel P. 1996. Detailed analysis of pronucleus development in bovine zygotes in vivo: ultrastructure and cell cycle chronology. *Mol Reprod Dev*. 43:62–69.
93 Laurincik J, *et al.* 1998. A detailed analysis of pronucleus development in bovine zygotes in vitro: cell-cycle chronology and ultrastructure. *Mol Reprod Dev*. 50:192–199.
94 Jada Nix MAM, Oliver MA, Rhoads M, Ealy AD, Biase FH. 2023. Cleavage kinetics is a better indicator of embryonic developmental competency than brilliant cresyl blue staining of oocytes. *Anim Reprod Sci*. 248:107174.
95 Tribulo P, Rivera RM, Ortega Obando MS, Jannaman EA, Hansen PJ. 2019. Production and culture of the bovine embryo. *Methods Mol Biol*. 2006:115–129.
96 Concordet JP, Haeussler M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res*. 46:W242–W245.
97 Wessels HH, *et al.* 2020. Massively parallel Cas13 screens reveal principles for guide RNA design. *Nat Biotechnol*. 38:722–727.
98 Guo X, *et al.* 2021. Transcriptome-wide Cas13 guide RNA design for model organisms and viral RNA pathogens. *Cell Genom*. 1: 100001.
99 Kent WJ, *et al.* 2002. The human genome browser at UCSC. *Genome Res*. 12:996–1006.
100 Karmakar S, Behera D, Baig MJ, Molla KA. In vitro Cas9 cleavage assay to check guide RNA efficiency. 2021, In: Islam MT, Molla KA, editors. *CRISPR-Cas methods: volume 2*. New York (NY): Springer US. p. 23–39.
101 Clement K, *et al.* 2019. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*. 37:224–226.
102 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*. 30: 2114–2120.
103 Shen W, Le S, Li Y, Hu FQ. 2016. Seqkit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *Plos One*. 11: e0163962.
104 Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol*. 20:129.
105 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 34:3094–3100.
106 Li H, *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.
107 Bó G, Mapletoft R. 2018. Evaluation and classification of bovine embryos. *Anim Reprod*. 10:344–348.
108 Biase FH, Franco MM, Goulart LR, Antunes RC. 2002. Protocol for extraction of genomic DNA from swine solid tissues. *Genet Mol Biol*. 25:313–315.
109 Biase FH. 2021. Isolation of high-quality total RNA and RNA sequencing of single bovine oocytes. *STAR Protoc*. 2:100895.
110 Puissant C, Houdebine LM. 1990. An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Biotechniques*. 8: 148–149.

111 Chomczynski P. 1993. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques*. 15:532.

112 Chomczynski P, Sacchi N. 2006. The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nat Protoc*. 1:581–585.

113 Bagnoli JW, *et al.* 2018. Sensitive and powerful single-cell RNA sequencing using mcSCRB-seq. *Nat Commun*. 9:2937.

114 Elsik CG, *et al.* 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 324:522–528.

115 Kinsella RJ, *et al.* 2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database*. 2011:bar030.

116 Flicek P, *et al.* 2012. Ensembl. *Nucleic Acids Res*. 40:D84–D90.

117 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. 37:907–915.

118 Danecek P, *et al.* 2021. Twelve years of SAMtools and BCFtools. *Gigascience*. 10:giab008.

119 Tischler G, Leonard S. 2014. Biobambam: tools for read pair collation based algorithms on BAM files. *Source Code Biol Med*. 9:1–18.

120 Liao Y, Smyth GK, Shi W. 2014. Featurecounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 30:923–930.

121 Fox J, Weisberg S. 2019. *An R companion to applied regression*. 3rd ed. Thousand Oaks (CA): Sage.

122 Hollander M, Wolfe DA, Chicken E. 2013. *Nonparametric statistical methods*. Hoboken (NJ): John Wiley & Sons.

123 Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. *J Comput Graph Stat*. 5:299–314.

124 RCoreTeam. 2020. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

125 McCarthy DJ, Smyth GK. 2010. Edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26:139–140.

126 Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 12:323.

127 Wagner GP, Kin K, Lynch VJ. 2012. Measurement of mRNA abundance using RNA-Seq data: RPKM measure is inconsistent among samples. *Theory Biosci*. 131:281–285.

128 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol*. 15:550.

129 Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met*. 57:289–300.