SUPPLEMENTARY MATERIAL

Production of Hornless Dairy Cattle by Precision Breeding

Daniel F. Carlson[1], Cheryl A. Lancto[1], Bin Zang[2], Eui-Soo Kim[1], Mark Walton[1], David Oldeschulte[3], Christopher Seabury[3], Tad S. Sonstegard[1], and Scott C. Fahrenkrug[1]

1 Recombinetics, St. Paul, Minnesota.
2 Program in Scientific Computation, University of Minnesota, Minneapolis Minnesota.
3 Department of Veterinary Pathobiology, College of Veterinary Medicine, Texas A&M University, College Station, Texas.

Supplementary Table 1: Degenerate TALEN sites in the bovine genome

		# de novo indels	
		RCI-001	RCI-002
# degeneracy in a TALEN pair	# off-target genome wide	within 20bp of an indel	within 20bp of an indel
0_0*	1	1	1
1_4	2	0	0
1_5	5	0	0
2_4	18	0	0
3_3	10	0	0
2_5	101	0	0
3_4	190	0	0
3_5	1449	0	0
4_4	1057	0	0
4_5	13147	0	0
5_5	45771	0	0
total	61751	1	1
* On-Target site			

Supplementary Methods

Experimental procedures involving animals were performed according to established standard operating procedures and protocols approved by the Trans Ova Institutional Animal Care and

Nature Biotechnology doi:10.1038/nbt.3560

Use Committee. The polled phenotype was confirmed by a non-blinded licensed veterinarian at birth, 3, 6 and 10 months of age.

Cell production:

Cells used for somatic cell nuclear transfer and genotyping methods were described in Tan, 2013[1]. Each clone and associated methods, other than HP-24.8 were described previously[1]. Briefly, for HP-24.8, btHP1.3 TALENs mRNA (2 µg) with the following RVD codes (Left monomer- NG NG NG HD NG NG NN NN NG NI NN NN HD NG NN) and (Right monomer- NN NI NI NI NI NN NI NN NI NN NG NG NG NG NN NI NG) and 1 µg of the Pc HDR template (see below) was co-transfected into 500,000 early passage fibroblasts from 2120. For transfection, cells and nucleic acids were resuspended in buffer "R" and electroporated using the 100µl tips by the following parameters: input Voltage; 1800V; Pulse Width; 20ms; and Pulse Number; 1, using the NEON electroporation system (Life Technologies). Individual colonies were propagated and evaluated for introgression of the Pc as previously reported[1].

Somatic Cell Nuclear Transfer:

Nuclear transfer and embryo transfers were performed by Trans Ova Genetics, Sioux Center IA according to their established procedures[2]. The total number of embryo transfers and pregnancy rates are reported in Table 1.

Genotypic analysis of cloned calves:

A first PCR assay was performed using (btHP-F1: 5'- GAAGGCGGCACTATCTTGATGGAA; btHP-R2- 5'- GGCAGAGATGTTGGTCTTGGGTGT) using 1X Gel-track mix (Quanta) and the following parameters: 95°C initial melt, 38 cycles of (95°C for 25 s; 62°C for 25 s; 72°C for 60 s) followed by a 5 min extension. The PCR creates a 591 bp product for Pc compared to the 389

bp product from the horned allele. Secondly, clones were analyzed by PCR using the flanking F1 and R1 primers (HP1748-F1- 5'- GGGCAAGTTGCTCAGCTGTTTTTG; HP1594_1748-R1- 5'-TCCGCATGGTTTAGCAGGATTCA) using 1X Gel-track mix (Quanta) and the following parameters: 95°C initial melt, 38 cycles of (95°C for 25 s; 62°C for 30 s; 72°C for 2 min) followed by a 5 min extension. The PCR creates a 1,748 bp product for Pc compared to the 1,546 bp product from the horned allele. All PCR products were TOPO cloned and sequenced.

Off-Target analysis.

DNA from the progenitor cell lines and edited animals were submitted to the Genomics and Bioinformatics Service core at Texas A&M. Samples were sequenced to an average 20X coverage on the Illumina HiSeq 2500 high output mode with paired end 125 bp reads were compared to the bovine reference sequence (UMD3.1). Our goal was to characterize two unique progenitor-clone pairs. RCI-003 was excluded since it is an identical clone to RCI-002 and the carcass of RCI-005 was disposed of prior to sample collection. RCI-001 was the only animal from progenitor 2122 and a coin-flip was used to randomize between RCI-002 and RCI-004. Structural variations were called using CLC probabilistic variant detection tools, and those with >7 reads were further considered even though this coverage provides only a 27.5% probability of accurately detecting heterozygosity[3]. Raw sequences are available in NCBI SRA database, study number SRP072240. Indel calls were groomed by the following methods. Upon indel calls for the original non-edited cell lines and 2 of the edited animals, we screened for de novo indels in edited animal RCI-001, which are not in the progenitor cell-line, 2120. We retrieved those potential de novo indels by subtracting common indels of RCI-001 and 2120 from all indels called for RCI-001. The same procedure was performed for edited RCI-002 against original cell-line 2122. We then applied PROGNOS[4] with reference bovine genome build UMD3.1 to compute all potential off-targets likely caused by the TALENs pair. The spacer allowance was set to 12-25 bp, mismatches, up to 5bp per TALEN monomer, 10 mismatches combined per TALEN pair. Both homo and heterodimers were considered. The distributions of sites with different degeneracy levels are shown in Table S1. For all matching sequences computed, we extract their corresponding information for comparison with de novo indels of RCI-001 and RCI-

002. BEDTools[5] was adopted to find de novo indels within 20 bp distance of predicted potential targets for the edited animal.

Pc HDR template (cloned into PCR 2.1 vector (Life Technologies):

```
atcgaacctgggtcttctgcattggctggcagattctttaccactgagccaccacaccctagag
tgcaaagggggcttagcaccacaggaggttcacaaagatgtaagctgttcttattaaggctgag
gtgggggttgggagaaggggagaaaaagttttgtaagttgttaatttataataaatccccaaa
gaaatggtctttcaagtacatacttatctaaaactttgtcaataggggaaatgttcttaggaga
gaaaaggaatttttttcttttagcataaagctgactttctaatatgggcttccctagtagctcag
ctggtgaagaacccgcctgcaatgtgggaaacctgggtttgagccctgggttgggaagatcccc
tggagaaaagaatggctacccactccagtattctggcctggagaattccatggactgtatacca
tggggttgcaaagagttggacacgactgagtgactttcattttcactttctgactttctaatat
tcgaggaatgcttagaagtgtggccggtagaaaatagtcctttgtgcctgggacttcaagaagg
cggcactatcttgatggaactcagtctcatcacctgtgaaatgaagagtacgtggtaccaacta
ctttctgagctcacgcacagctggacgtctgcgcctttcttgttatactgcagatgaaaacatt
ttatcagatgtttgcctaagtatggattacatttaagatacatatttttctttcttgtctgaaa
gtctttgtagtgagagcaggctggaattatgtctggggtgagatagttttcttggtaggctgtg
aaatgaagagtacgtggtaccaactactttctgagctcacgcacagctggacgtctgcgccttt
cttgttatactgcagatgaaaacattttatcagatgtttgcctaagtatggattacatttaaga
tacatatttttctttcttgtctgaaagtctttgtagtgagagcaggctggaattatgtctgggg
tgagatagttttctttgctctttagatcaaaactctcttttcatttttaagtctatcccaaaag
tgtgggaggtgtccttgatgttgaattataggcagagggtcagtttatcaacacccaagaccaa
catctctgcctttgataagagatagaaatagaagtggagagagaggaggaaaaacatgactcac
gatacattctgggttgtttgtttttgtttttatttttgttttggaaggagcgggtgggggaacg
tgctgattaaagaaagtctagagagaacaagattcttaaaaataaatccacagtgaagcccagc
gggtggggatttcccacagattttcagggctttttttgtgttgtcatggggatattagtcaatgt
tggtgtcttattttggagtcactatgagtgaaccatgtttaaggagctatggctcagctgctaa
actattttcaaaaggaaaatggtgtgttacggtttccgagcagtggggccctggtacaggtaat
atcatactcaaaagcactcttttgtgctacatgaatgaatcctgctaaaccatgcgga
```

References:

1 Tan, W. *et al.* Efficient nonmeiotic allele introgression in livestock using custom endonucleases. *Proc Natl Acad Sci U S A* **110**, 16526-16531, doi:10.1073/pnas.1310478110 (2013).

2 Kasinathan, P. *et al.* Acceleration of genetic gain in cattle by reduction of generation interval. *Sci Rep* **5**, 8674, doi:10.1038/srep08674 (2015).

3 Illumina. *Calling Sequencing SNPs*, <http://www.illumina.com/Documents/products/technotes/technote_snp_caller_sequencing.pdf> (2010).

4 Fine, E. J., Cradick, T. J. & Bao, G. Identification of off-target cleavage sites of zinc finger nucleases and TAL effector nucleases using predictive models. *Methods Mol Biol* **1114**, 371-383, doi:10.1007/978-1-62703-761-7_24 (2014).

5 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842, doi:10.1093/bioinformatics/btq033 (2010).