Insights into the function of n-3 PUFAs in *fat-1* transgenic cattle

Xin-feng Liu[1,2], Zhu-ying Wei[1], Chun-ling Bai[1], Xiang-bin Ding[2], Xin Li[2], Guang-hua Su[1], Lei Cheng[1], Li Zhang[1], Hong Guo[2,*] and Guang-peng Li[1,*]

1. The Key Laboratory of Mammalian Reproductive Biology and Biotechnology of the Ministry of Education, Inner Mongolia University, 24 Zhaojun Road, Hohhot 010070, China.

2. College of Animal Science and Animal Medicine, Tianjin Agriculture University, 22 Jinjing Road, Tianjin 300384, China.

*Correspondence authors: Hong Guo (E-mail: guohong64@163.com; Telephone: 86-22-23784590; Fax: 86-22-23784590) and GuangPeng Li (E-mail: imudwzx@163.com; Telephone: 86-22-23784590; Fax: 86-471-4994329).

Runing title: Function analysis of n-3 PUFAs

Supplemental materials and methods

Whole genome sequencing and data analysis

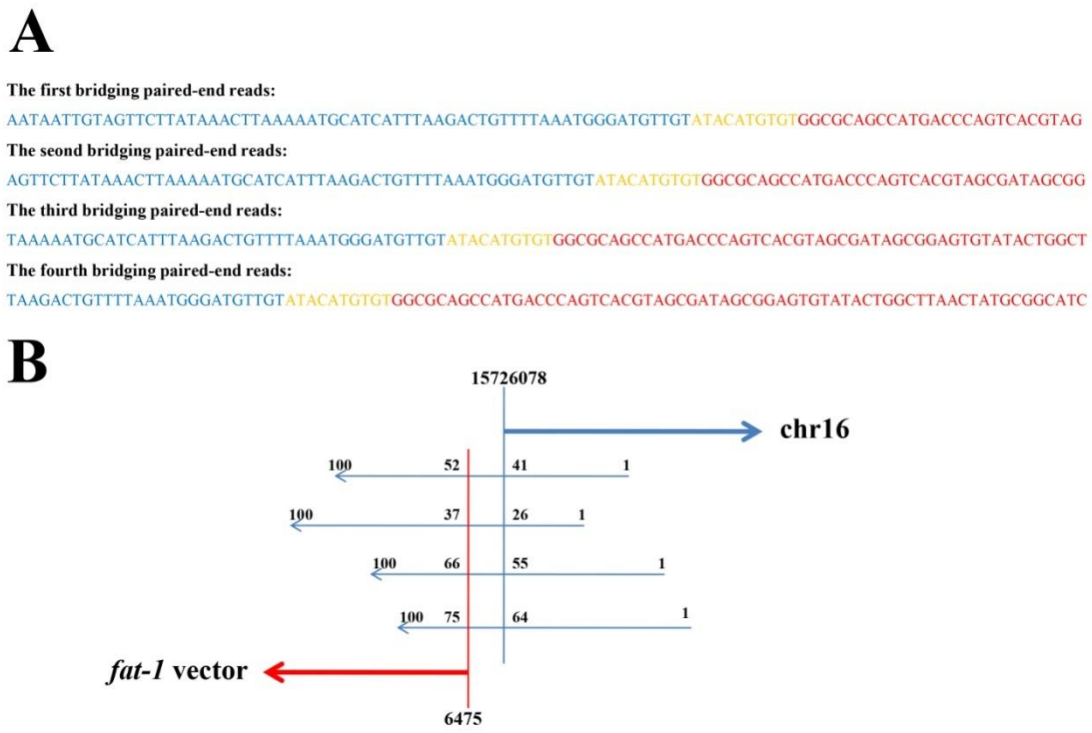
Blood was collected from the jugular vein of the FD006 transgenic cow, and DNA was extracted from the blood with a QIAGEN DNA extraction kit (Qiagen Sciences, Germantown, MD). A total of 1.5 μg whole genomic DNA was sonicated with a Bioruptor sonication system (Diagenode, Inc.) to construct paired-end libraries. The average insert size for the paired-end libraries was 500 bp. These libraries were then sequenced on the Illumina HiSeq 2000 platform according to the manufacturer's protocols (Illumina, San Diego CA). The sequencing reads for FD006 were collected, and low-quality reads were filtered out using custom Perl scripts with a threshold of Q = 20. All of the filtered sequencing reads were then mapped to the reference *Bos taurus* genome sequence (Bos_taurus_UMD_3.1, build 6.1) and the *fat-1* vector sequence to find bridging reads between the host genome and the foreign fragments. To identify the region of transgene integration, abnormal read pairs with one end mapping to the reference and the other end to the transgene or vector were selected for further security. The exact integration breakpoints were identified by split-read analysis that spanned the transgene insertion junctions.

Verification of transgene integration breakpoint by PCR

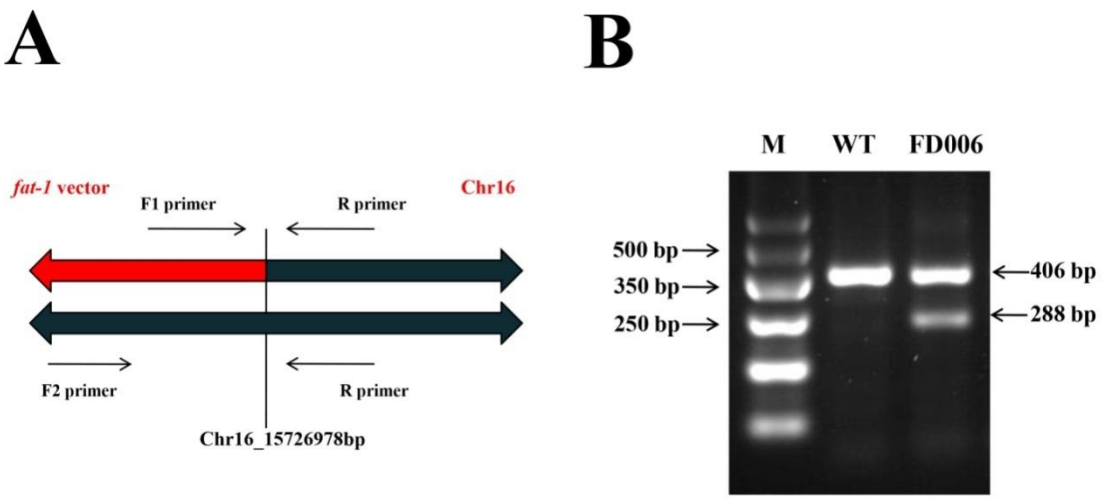
To verify the integration breakpoint, PCR was performed on genomic DNA samples from FD006 and wild-type cattle. Three primers were designed for verifying the transgene integration breakpoints, F1 (GGTAGCTACTAGAAAGGTAAGCA), R (TGAGAAATGACAGAGATTGGTT) and F2 (GCATCTGTGCGGTATTTCAC). The products were amplified using a Bio-rad PCR System T100 (USA) with cycling parameters of

94℃ for 5 minutes, 35 cycles of 94℃ for 30 seconds, 58℃ for 30 seconds and 72℃ for 1 minute, followed by a final extension at 72℃ for 5 minutes. If the FD006 transgenic cow were heterozygous, 288-bp and 406-bp products would be generated by primer pairs F1 + R and F2 + R, respectively. On the other hand, if FD006 were a homozygote, only 288-bp products would be generated. Accordingly, only 406-bp products would be generated in wild-type cattle.

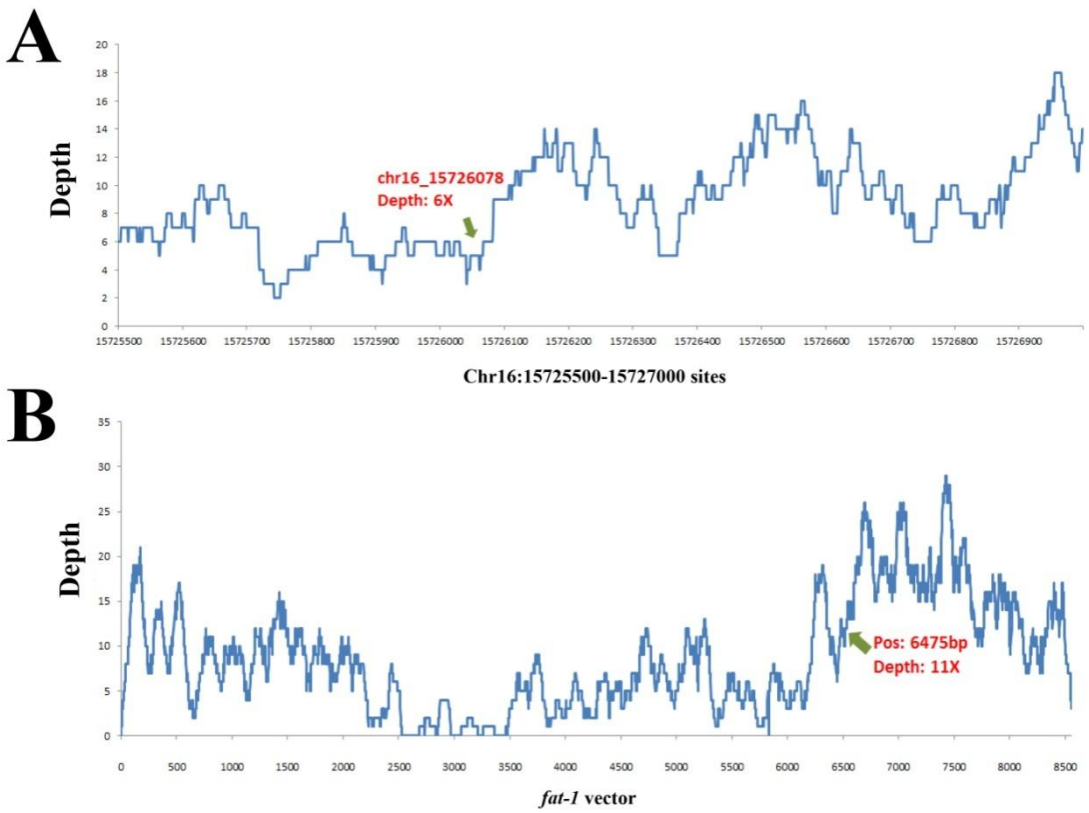
Supplemental Figures



Supplemental Figure S1: Transgene integration sites. (A) Four bridging paired-end reads; the blue part is the *fat-1* vector sequence, the red part is the sequence of chromosome 16, and the yellow part is an introduction of a 10-nucleotide portion of the bovine genome. (B) Schematic of the integration sites for the *fat-1* vector on chromosome 16 in the bovine genome.



Supplemental Figure S2: Verification the transgene integration sites by PCR. (A) Schematic of primer design; F1 is a transgene-specific primer, whereas F2 and R are genome-specific primers. (B) PCR detection of the integration sites in the FD006 transgenic cow and one wild-type cow. The amplified product for the wild-type (WT) sequence was 406 bp, whereas the transgenic (FD006) sequences were 406 bp and 288 bp, respectively.



Supplemental Figure S3: Sequencing depth of the Chr16:15725500–15727000 sites and the *fat-1* vector. After mapping the Illumina reads onto the (A) Chr16:15725500-15727000 sites and (B) *fat-1* vector, the sequencing depth was calculated in 10-bp sliding windows. The X-axis denotes the length of the (A) Chr16:15725500-15727000 sites and (B) *fat-1* vector in bp, and the Y-axis denotes sequencing depth.