

Population genomics Reveals PRDM16 underpins the cold tolerance in domestic Cattle

Chunlong Yan¹, Jun Lin², Yuanyuan Huang², Qingshan Gao¹, Zhengyu Piao¹, Shouli Yuan², Li Chen², Xue Ren³, Rongcai Ye², Meng Dong², Hanlin Zhang², Huiqiao Zhou², Xiaoxiao Jiang², Xiangnan Wu², Wanzhu Jin², Xuming Zhou², and Changguo Yan¹

¹Yanbian University

²Affiliation not available

³Annoroad Gene Technology Co. Ltd

July 2, 2020

Abstract

Environmental temperature serves a major driver for adaptive changes in wild organisms, however, its role in domestication has been less characterized. To uncover the mechanisms of cold tolerance in domestic animals, we sequenced genomes of 28 cattle at median coverage from warm and cold areas across China. By characterizing the population structure and demographic history, we identified two genetic clusters, i.e., northern and southern cattle groups, and a common historic population peak at 30 kilo years ago. Genome scan of cold tolerant breeds revealed genes that under selection sweeps enriched in thermogenesis related pathways. Specifically, we determined a substitution of PRDM16 (p.P779L) in north cattle, which maintains the formation of brown adipocytes through boosting expression of thermogenic related genes, indicating a vital role of this gene in cold tolerance. The findings provide a basis of genetic variations in domestic cattle that shaped by temperature environments and highlight a role of reverse mutation in livestock species.

Introduction

Temperature is one of the most important environmental factors that drives evolutionary changes in diverse organisms [1]. Mammals are endotherms and they request a constant body temperature to ensure optimal biological activity [2,3], leading to a strong selection pressure on the heat production system, i.e, shivering and non-shivering thermogenesis [4]. In particular, shivering thermogenesis produces heat in short term [5] and non-shivering thermogenesis is a non-contractile process that able to compensate the defects of shivering thermogenesis and maintain body temperature effectively [4]. Though white adipose tissue (WAT) stores excessive energy as triglycerides, the brown adipose tissue (BAT) which activated by cold exposure has been recognized as a major source of adaptive non-shivering thermogenesis [6-9]. For example, the uncoupling protein-1 (UCP1) in BAT dissipates energy into heat through uncoupled respiration, resulting in increased fatty acid oxidation and heat production [10]. The thermogenic capacity of BAT is particularly effective in maintaining the core body temperature of small mammals and infants [4]. Nevertheless, the thermogenic program in adipose tissue is a complex transcriptional regulation process that has not been fully dissected. The widely reported transcriptional regulators of adipocytes include the peroxisome proliferator-activated receptor-gamma (PPAR γ), peroxisome proliferator activated receptor-gamma coactivator 1 α (PGC1- α), Forkhead box C2 (FoxC2) and PRD1-BF-1-RIZ1 homologous domain-containing protein-16 (PRDM16) [11]. Among these genes, PPAR γ plays a leading role in the differentiation of all adipocytes [12-14]. It is also known that PGC1- α acts together with PPAR γ or the thyroid hormone receptor adaptive thermogenesis [15-16]. FoxC2 can increase the BAT amount to enhance the insulin sensitivity

[17] and PRDM16 can induce the browning of WAT and fibroblasts through driving a brown adipogenesis program while suppressing the white fat adipogenesis program [17].

Cattle are intimately associated with human civilization and culture worldwide. Currently, there are 53 cattle breeds in China, and two species are recognized: *Bos taurus* and *Bos indicus* [18-19]. Archaeological studies supports the claim that *B. taurus* was imported into northern China and north-east Asia from north Eurasia between 5000–4000 BP [20], and *B. indicus* migrated from the Indian subcontinent to East Asia around 3000 BP [21]. Intriguingly, there is a vast difference in the annual average temperature of the habitats among those cattle along with the domestication. Here, to detect the molecular footprints underlying the cold adaptations in domestic cattle, we sequenced genomes of 28 cattle, including 14 cold-tolerant cattle (annual average temperature of habitat: 2–6) and 14 cold-intolerant cattle (annual average temperature of habitat: 20–25). By characterizing the population history and selective sweeps, we identified a candidate gene *PRDM16* that was under selection and responsible for the modification of the BAT function, which underpin the cold-tolerance in northern cattle.

Results and discussion

Genome sequencing and population history

Whole-genome sequencing of 28 cattle with an average depth of 22.44× obtained 17.3 billion clean reads (**Fig. 1a, Additional file 1: Fig. S1 and Table S1**). In total, 45.2 million SNPs were identified and most of the SNPs were located in the intergenic (61.51%) and intron (31.75%) regions (**Additional file 1: Table S2**). Neighbor-joining trees and principal component analysis (PCA) based on the total SNPs clustered the cattle into two main groups: northern group and southern group (**Fig. 1b, c**). The first principal component (PC1), representing 32.41% of the total variation, separated the samples into northern and southern cattle (**Fig. 1c**). ADMIXTURE analyses with different K -values, including $K = 2$, clearly indicated that cattle samples could be classified into northern and southern cattle. Linkage disequilibrium analysis showed that the southern cattle were attenuating compared with northern cattle (**Fig. 1d, Additional file 1: Fig. S2**), suggesting a higher frequency of genetic recombination in southern cattle.

The demographic history of cattle conducted by the pairwise sequentially markovian coalescent (PSMC) [22] model showed two bottlenecks and two expansions, with population peaks at ~50 and ~700 Kya (kilo years ago) and population bottlenecks at ~30 and 400 Kya, respectively (**Fig. 1f**). There were two sharp declines in population both occurred during the glacial period, consistent with the idea that environmental temperature has a determinable impact on population size. Moreover, similar historical patterns appeared in many other mammals, such as the giant panda, yak and sub-nosed monkey [23-25]. Global glaciations are the most probable cause for a sudden change in the global climate, and this extremely cold climate directly affected the decline of species populations. In addition, after the Naynayxungla Glaciation, northern cattle experienced a long-term bottleneck until 70 Kya. In contrast, the effective population size (N_e) of southern cattle recovered rapidly (**Fig. 1f**), which is consistent with previous studies [26-28], and might be explained by the improved living environment in the south during the glaciation [29]. At ~60 Kya, HN and YN cattle showed different trends of N_e . Specifically, the N_e of HN cattle increased rapidly (**Fig. 1f**), probably because of the geographical location of Hainan, a small and comparatively isolated island, which lead to a better environment that, lacked of predators, thereby encouraging the survival and reproduction of HN cattle. According to mitochondrial DNA haplotypes, *B. taurus* (northern cattle) and *B. indicus* (southern cattle) are both derived from the extinct wild aurochs (*Bos primigenius*) and the divergence between the two species dates back to 250,000 years ago [30].

Genome scan of selective sweeps

To determine genetic modifications that occurred under different temperatures, we analyzed selective sweeps between the two cattle groups: northern (MG and YB) and southern (YN and HN). Selective sweep analyses were performed over the whole genomes based on the distribution of F_{ST} values, with the top 5% calculated in 50-kb windows with 25-kb steps using a Fisher's exact test to determine the frequency spectrum in southern versus northern cattle. A total of 269 candidate genes were found to be with strong selective

sweep signals in Chinese cattle (**Fig. 2a**, **Additional file 1: Fig. S3** and **Table S3**). The top enriched KEGG pathway of the candidate genes was RNA degradation ($P = 5.5 \times 10^{-3}$) (**Additional file 1: Table S4**), which mainly acts on gene expression and regulation and plays an important role in cellular mechanisms [31]. Several candidate genes (e.g., *AOC3* and *FASN*) were involved in pathways related to the activation of thermogenic metabolism, such as tyrosine metabolism ($P = 5.3 \times 10^{-2}$) and fatty acid biosynthesis ($P = 5.8 \times 10^{-2}$) (**Additional file 1: Table S5**). This is expected as cold exposure could increase tyrosine hydroxylase, leading to the increased secretion of norepinephrine and enhanced non-shivering thermogenesis in BAT [32]. In addition, *AOC3* is related to diabetes and membrane-bound WAT [33-34]. *FASN* plays a role in the *de novo* biosynthesis of fatty acids, and the down-regulation of *FASN* can induce browning in WAT [35]. Many fatty acids exert a positive effect on thermogenesis by activating BAT [36-38]. We also found many candidate genes (e.g., *APOA5*, *ACOX3* and *ALDOB*) involved in fatty acid, fructose and mannose metabolism and associated with signalling pathways, such as the PPAR signalling pathway (**Additional file 1: Table S5**). *APOA5* treatments can increase gene expression levels of *UCP1* in adipocytes, and fat accumulation can cause *ACOX3* to increase in blood [39]. *ALDOB* is involved in insulin biosynthesis and secretion, as well as insulin receptor signalling [40]. It is well-known that insulin pathways and fat metabolism are inseparable and could affect the development of BAT, leading to obesity and insulin resistance, for example [41-42]. Consistently, enrichment analysis of GO terms revealed that several candidate genes (*LOC100337001*, *PRDM16*, *SH3PXD2B* and *ZFPM1*) are related to fat cell differentiation (GO:0045598), brown fat differentiation (GO:0050873) and white fat cell differentiation (GO:0050872) (**Additional file 1: Fig. S4**, **Table S6** and **S7**).

Among those genes with signals of selective sweep, *PRDM16* is of most interest because it was reported to increase thermogenesis by promoting the expression of the key gene *UCP1* (**Additional file 1: Fig. S5**). In addition, though the F_{ST} (0.26) and θ_{π} (0.2×10^{-3}) values of *PRDM16* were similar to that of three other significant genes (*LOC100337001*, *SH3PXD2B* and *ZFPM1*), the lowest P -value was observed for *PRDM16* ($P = 3.8 \times 10^{-8}$) (**Additional file 1: Table S3** and **S8**). In addition, Tajima's D analysis also supported the idea that *PRDM16* was under selection ($D = -0.34$) (**Fig. 2b**, **Additional file 1: Table S8**). The *PRDM16* genotypes found in northern and southern cattle were well-distinguished from each other and consistent with the phylogenetic tree created using the SNPs of this gene (**Fig. 2c**). We discovered five non-synonymous SNVs of which one (c.2336 T > C, p.L779P) was found at a higher level (93%) in southern cattle than northern cattle (**Additional file 1: Fig. S6** and **Table S8**).

We then compared the *PRDM16* protein sequences to other species whose BAT had been found and reported (**Fig. 2d**, **Additional file 1: Fig. S7**) and it is interesting to find that the substitution that occurred at position Leu₇₇₉ of the *PRDM16* gene in northern cattle is the same as that in species that have complete BAT function (e.g., mouse, rat and hamster) (**Fig. 2d**). In rodents, BAT is intact and persists throughout their lifetimes, and its thermogenesis activity is complete [43-44]. However, in many large mammals, BAT is incomplete or even non-existent. For example, the BAT is incomplete in human adults and sheep in which it exists only during infancy and degenerates as they grow [45-46]. Conversely, the substitution in southern cattle, which is proline, is the same as that in species that have an incomplete or null BAT function (sheep, pig, whale, horse, platypus, elephant, sirenian, marsupial, human and rabbit (**Fig. 2d**)). Thus, we hypothesized that this substitution of residue 779 in the *PRDM16* gene is probably related to the BAT function, and this locus is likely to play a role in cold tolerance.

The Mutation Effect of *PRDM16*

To determine the biochemical function of the substitution in *PRDM16*, 3T3-L1 cells (preadipocyte cell line) that ectopically expressed cattle *PRDM16* and *PRDM16* MU (L779P mutation of *PRDM16*) coding sequences were generated and induced to differentiate towards beige adipocytes (**Additional file 1: Fig. S8**). The overexpression efficiency was kept at equivalent levels (**Fig. 3a**). After full differentiation, no differences in morphological characteristics between the *PRDM16* and *PRDM16* MU groups was observed (**Additional file 1: Fig. S9**). In addition, no significant differences in the mRNA and protein expression levels of PPAR γ , a key adipogenesis-regulating gene (**Fig. 3b** and **3c**). Notably, the differentiation efficiency

was lower for the control group (cells infected with an empty vector) than the two other groups, supporting the idea that a lack of PRDM16 significantly impeded the differentiation of brown adipocytes, and the overexpression of PRDM16 would significantly increase the number of brown adipocytes [17]. Despite the similar efficiency in differentiation between the two ectopic *PRDM16* overexpressing groups, the mRNA expression levels of four BAT-selective genes (*UCP1*, β -*EBIIP*, *III α -1a* and *IDEA*) were significantly lower in the PRDM16 MU group than in the PRDM16 group (**Fig. 3d**). Moreover, overexpression of PRDM16 MU largely increased *UCP1* expression, which was far inferior to PRDM16 overexpression (**Fig. 3d, 3e**). The mRNA levels of other BAT-selective genes (β -*EBIIP*, *III α -1a* and *IDEA*) were significantly lower in PRDM16 MU-overexpressing adipocytes than in *PRDM16*-overexpressing adipocytes (**Fig. 3d**). The above results indicate that L779P mutation significantly impaired the normal function of PRDM16 in the formation of brown adipocytes in southern cattle, which live in a warmer area relative to northern cattle.

Conclusions

PRDM16 is a key transcriptional regulator in beige adipocyte formation and can stimulate nearly all key characteristics of authentic brown fat cells²⁶. In addition, although *PRDM16* is introduced before cell differentiation, nearly all adipocytes can be activated to express BAT-selective genes [17]. In this study, we found that BAT-selective genes were up-regulated in PRDM16-overexpressing 3T3-L1 cells compared with the control group, indicating that the mutation in *PRDM16* influences the function of this gene in the brown adipogenesis program. PRDM16 regulates thermogenic genes by forming a complex with various transcription factors, including β -*EBIIP*, *III α -1a*, *IIIAPa* and *IIIAP γ* . Although the same differentiation efficiency was induced, the suppression of β -*EBIIP* and *III α -1a* mRNA expression levels in the PRDM16 MU group indicated a decreased formation of the transcription complex and finally, the suppression of the expression of thermogenic genes, such as *UCP1*, compared with the PRDM16 group. Functional differences in PRDM16 caused by sequence variations could explain why northern cattle are more cold-tolerant than southern cattle. For example, mortality for *B. indicus* is suggested to be higher than that for *B. taurus* in the cold [47], which may be a result of exhausting their post-natal BAT lipids [48]. Therefore, on one hand, *PRDM16* functioned well in northern cattle is required to resist extreme cold, and on the another hand, the functional inactivation of PRDM16 impairs beige adipocyte formation, which is also beneficial for the environment adaptability of southern cattle. In summary, we compared the whole genomes of northern cattle living in extremely cold environments with those of southern cattle living in warm environments to reveal a new mechanism behind northern cattle's adaptability to cold environments. This study is the first to propose a molecular mechanism for cattle's cold-tolerance ability. In particular, we proved that *PRDM16* is a forceful genome effector that facilitates the formation of diverse cold-adaptations using functional experiments. These results could improve the comprehension of adaptive genetic variations in cattle living in different temperatures and other livestock species.

Methods

Genome sequencing

We sampled a total of 28 cattle, from four different regions of China (Mongolia, Yanbian, Hainan and Yunnan). DNA was extracted from the blood of each individual and degradation was monitored based on its concentration by spectrometry, fluorometry, and 1% agarose gel electrophoresis. Paired-end libraries with insert size of 150 bp were constructed for each individual and sequenced using the a HiSeq X Ten Sequencing System (Illumina Inc., San Diego, CA, USA). We mapped the clean reads after filtering sequencing data to the *Bos Taurus* genome assembly (version UMD_3.1.1) using BWA [49]. All the potential single-nucleotide polymorphisms site (SNPs) were extracted and filtered by GATK [50]. ANNOVAR [51] and existing Genome annotation file (GFF/GTF) made corresponding annotations on the detected SNPs. All experimental procedures were performed in accordance with the Regulations for the Administration of Affairs Concerning Experimental Animals approved by the State Council of People's Republic of China.

Phylogenetic and Population Structure

PCA (Principal Components Analysis) was carried out using the EIGENSOFT [52]. A phylogenetic tree

was constructed from the SNPs data by using the neighbor-joining method in the program PHYLIP [53] and graphical demonstration was provided by using the Newick Utilities [54]. Population structure was further inferred using Admixture [55] with kinship (K) set from 2 to 10, according to the K -values with the smallest CV (Cross-Validation) error is the optimal K values.

Linkage Disequilibrium

The LD pattern was calculated for different breeds via the squared correlation coefficient (r^2) between pairwise SNPs using the script PopLDdecay (<https://github.com/BGI-shenzhen/PopLDdecay>).

PSMC Analysis

The PSMC model (<https://github.com/lh3/psmc>) parameters were set as follow: -N25 -t15 -r5 -p "4+25*2+4+6". The mutation rate: $\mu = 1.1 \times 10^{-8}$ and generation time: $g = 5$. The mutation rate is estimated by *baseml* in PAML package.

Selective Sweep Analysis

The population-differentiation statistic (F_{ST}) with Vcftools [56], nucleotide diversity (θ_π) and Tajima's D with Variscan 2.0, were estimated for 50-kb sliding windows with a 25-kb step size along each chromosome and Fisher's exact test was performed on synonymous and non-synonymous SNPs in the exon region by Plink [57], the false discovery rate (FDR) was used to correct the P values. Windows at the top 5% of F_{ST} and $q < 0.01$ of Fisher's exact test were selected, and the corresponding genes were considered as candidates. Enrichment analysis was conducted according to Uniport and KEGG databases.

Cell Culture

Lentiviruses were produced by transfecting the HEK293T cells with the core plasmids and two helper plasmids (psPAX2 and pMD2G). The transfections were implemented using the Polyethylenimine (PEI) method at the ratio of PEI : core plasmid : psPAX2 : pMD2G = 27 : 4 : 3 : 2. The medium was changed 4–6 h after transfection. After 48h, the virus-containing medium was harvested and filtered. Then, the 3T3-L1 cells were incubated overnight with the viral supernatant and 8 $\mu\text{g}/\text{mL}$ polybrene. For browning differentiation, confluent 3T3-L1 cells were incubated for 2 days in a brown adipogenic induction cocktail (DMEM containing 10% FBS, 20 nM insulin, 1 nM 3,3,5-triiodo-L-thyronine (T3), 0.5 mM isobutylmethylxanthine, 0.125 μM indomethacin, and 1 mM dexamethasone). Then, the cells were kept in differentiation medium (DMEM containing 10% FBS, 20 nM insulin, 1 nM T3) for 6 days. The induction medium was changed every 2 days. At day 8, fully differentiated brown adipocytes were applied for all following experiments in this study.

RNA Isolation and Real-time Quantitative PCR

Total RNAs from tissues or cells were extracted with Trizol reagent (Thermo Fisher Scientific Inc, MA, USA). Reverse transcription of 2 μg total RNA was performed with a high-capacity cDNA reverse transcription kit (Promega, Madison, USA). qPCR reaction was performed with a SYBR Green Master Mix (Promega, Madison, USA) and detected using a Prism VIIA7 real-time PCR system (Applied Biosystems Inc, Foster City, CA, USA). Primers were designed using Primer Quest (Integrated DNA Technologies, Inc, Coralville, IA, USA). See supplementary table 10 for the details of primer sequences.

Western Blot Analysis

Cells were lysed in RIPA buffer containing 150mM sodium chloride, 1.0% TritonX-100, 0.5% sodium deoxycholate, 0.1% SDS, 50mM Tris with freshly added protease and phosphatase inhibitor cocktail (Roche Diagnostics Corp, Pleasanton, CA, USA). Equal amount of protein samples were distributed in 10% SDS-polyacryl-amide gels. After electrophoresis, proteins were transferred to a PVDF membrane, incubated with blocking buffer (5% fat-free milk) for 1 h at room temperature, and blotted with the following antibodies overnight: anti-PRDM16 (R&D), anti-UCP1 (Abcam), anti-PPAR γ (CST) and anti- β -actin (SigmaChemical). The membrane was incubated with HRP-conjugated secondary antibodies for 1 h at room temperature.

Signals were visualized using Mini ChemiTM 580 (Sage Creation Science Co, Beijing, China) with Super Signal West Pico Chemiluminescent Substrate (Pierce, Rockford, IL, USA).

Statistics

Data are expressed as means \pm standard error (SE). Comparisons between groups were performed with one-way ANOVA or Student's t-tests. Statistical significance was set at $P < 0.05$.

Funding

This work was supported by the National Key Research and Development Program of China (2017YFC1001001), the Youth Program of National Natural Science Foundation of China (31900830), the National Natural Science Foundation of China (81770834) and CAS pioneer hundred talents program (2018-085) from the Chinese Academy of Sciences and Jilin Provincial Development and Reform Commission Budget Capital Construction Fund Project (2018M640182) and the 111 project (D20034). J.L. was supported by the China Postdoctoral Science Foundation Funded Project (2018M640182).

Author contributions

C.G.Y., X.M.Z. and W.Z.J. designed the research, analyzed data and revised the manuscript. C.L.Y., J.L. and Y.Y.H. performed experiments, analyzed data and wrote the manuscript. Q.S.G. and Z.Y.P. collected samples, performed experiments and analyzed data. S.L.Y. and X.R. analyzed data. L.C. revised the manuscript and R.C.Y, M.D., H.L.Z., H.Q.Z., X.X.J. and X.N.W. collected samples.

Competing interests

The authors declare that they have no competing interests.

Data Availability Statement:

Data that support the findings of this study are available in the Dryad Data Repository.

Reference

- 1 Parsons, P. A. Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. *Biol Rev Camb Philos Soc* 80, 589-610 (2005).
- 2 Haim, A. & Levi, G. Role of Body-Temperature in Seasonal Acclimatization - Photoperiod-Induced Rhythms and Heat-Production in *Meriones-Crassus*. *J Exp Zool* 256, 237-241 (1990).
- 3 Hayes, J. P. & Garland, T., Jr. The Evolution of Endothermy: Testing the Aerobic Capacity Model. *Evolution* 49, 836-847 (1995).
- 4 Cannon, B. & Nedergaard, J. Brown adipose tissue: function and physiological significance. *Physiol Rev* 84, 277-359 (2004).
- 5 Heldmaier, G., Steinlechner, S., Ruf, T., Wiesinger, H. & Klingenspor, M. Photoperiod and Thermoregulation in Vertebrates - Body-Temperature Rhythms and Thermogenic Acclimation. *J Biol Rhythm* 4, 251-265 (1989).
- 6 Nicholls, D. G. & Locke, R. M. Thermogenic Mechanisms in Brown Fat. *Physiological Reviews* 64, 1-64 (1984).
- 7 Hughes, D. A., Jastroch, M., Stoneking, M. & Klingenspor, M. Molecular evolution of UCP1 and the evolutionary history of mammalian non-shivering thermogenesis. *Bmc Evol Biol* 9, 4 (2009).
- 8 Rowlett, U., Mrosovsky, N. & English, A. A comparative survey of brown fat in the neck and axilla of mammals at birth. *Biol Neonate* 17, 53-83 (1971).
- 9 Saito, S., Saito, C. T. & Shingai, R. Adaptive evolution of the uncoupling protein 1 gene contributed to the acquisition of novel nonshivering thermogenesis in ancestral eutherian mammals. *Gene* 408, 37-44 (2008).

- 10 Klingenberg, M. Uncoupling protein - A useful energy dissipator. *J Bioenerg Biomembr* 31, 419-430 (1999).
- 11 Kajimura, S., Seale, P. & Spiegelman, B. M. Transcriptional control of brown fat development. *Cell Metab* 11, 257-262 (2010).
- 12 Tontonoz, P., Hu, E. & Spiegelman, B. M. Stimulation of adipogenesis in fibroblasts by PPAR gamma 2, a lipid-activated transcription factor. *Cell* 79, 1147-1156 (1994).
- 13 Barak, Y. et al. PPAR gamma is required for placental, cardiac, and adipose tissue development. *Mol Cell* 4, 585-595 (1999).
- 14 Nedergaard, J., Petrovic, N., Lindgren, E. M., Jacobsson, A. & Cannon, B. PPARgamma in the control of brown adipocyte differentiation. *Biochim Biophys Acta* 1740, 293-304 (2005).
- 15 Puigserver, P. et al. A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* 92, 829-839 (1998).
- 16 Handschin, C. & Spiegelman, B. M. Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism. *Endocr Rev* 27, 728-735 (2006).
- 17 Seale, P. et al. Transcriptional control of brown fat determination by PRDM16. *Cell Metab* 6, 38-54 (2007).
- 18 Lai, S. J., Liu, Y. P., Liu, Y. X., Li, X. W. & Yao, Y. G. Genetic diversity and origin of Chinese cattle revealed by mtDNA D-loop sequence variation. *Mol Phylogenet Evol* 38, 146-154 (2006).
- 19 Lei, C. Z. et al. Origin and phylogeographical structure of Chinese cattle. *Anim Genet* 37, 579-582 (2006).
- 20 Cai, D. W. et al. The origins of Chinese domestic cattle as revealed by ancient DNA analysis. *J Archaeol Sci* 41, 423-434 (2014).
- 21 Payne, W. J. A. & Hodges, J. *Tropical cattle : origins, breeds, and breeding policies*. (Blackwell Science, 1997).
- 22 Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* 475, 493-496 (2011).
- 23 Qiu, Q. et al. Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat Commun* 6, 10283 (2015).
- 24 Zhao, S. et al. Whole-genome sequencing of giant pandas provides insights into demographic history and local adaptation. *Nat Genet* 45, 67-71 (2013).
- 25 Zhou, X. et al. Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. *Nat Genet* 46, 1303-1310 (2014).
- 26 Chen, N. et al. Whole-genome resequencing reveals world-wide ancestry and adaptive introgression events of domesticated cattle in East Asia. *Nat Commun* 9, 2337 (2018).
- 27 Mei, C. et al. Genetic Architecture and Selection of Chinese Cattle Revealed by Whole Genome Resequencing. *Mol Biol Evol* 35, 688-699 (2018).
- 28 Lan, D. et al. Genetic Diversity, Molecular Phylogeny, and Selection Evidence of Jinchuan Yak Revealed by Whole-Genome Resequencing. *G3 (Bethesda)* 8, 945-952 (2018).
- 29 Murray, C., Huerta-Sanchez, E., Casey, F. & Bradley, D. G. Cattle demographic history modelled from autosomal sequence variation. *Philos Trans R Soc Lond B Biol Sci* 365, 2531-2539 (2010).
- 30 Bradley, D. G., MacHugh, D. E., Cunningham, P. & Loftus, R. T. Mitochondrial diversity and the origins of African and European cattle. *Proc Natl Acad Sci U S A* 93, 5131-5135 (1996).

- 31 Petrillo, E., Godoy Herz, M. A., Barta, A., Kalyna, M. & Kornblihtt, A. R. Let there be light: regulation of gene expression in plants. *RNA Biol* 11, 1215-1220 (2014).
- 32 Nguyen, K. D. et al. Alternatively activated macrophages produce catecholamines to sustain adaptive thermogenesis. *Nature* 480, 104-108 (2011).
- 33 Ke, M. et al. Differential proteomic analysis of white adipose tissues from T2D KKAY mice by LC-ESI-QTOF. *Proteomics* 17 (2017).
- 34 Obata, T. Diabetes and semicarbazide-sensitive amine oxidase (SSAO) activity: a review. *Life Sci* 79, 417-422 (2006).
- 35 Guilherme, A. et al. Adipocyte lipid synthesis coupled to neuronal control of thermogenic programming. *Mol Metab* 6, 781-796 (2017).
- 36 Quan, L. H. et al. Myristoleic acid produced by enterococci reduces obesity through brown adipose tissue activation. *Gut* (2019).
- 37 Heeren, J. & Scheja, L. Brown adipose tissue and lipid metabolism. *Curr Opin Lipidol* 29, 180-185 (2018).
- 38 Li, Z. et al. Butyrate reduces appetite and activates brown adipose tissue via the gut-brain neural circuit. *Gut* 67, 1269-1279 (2018).
- 39 Zheng, X. Y., Yu, B. L., Xie, Y. F., Zhao, S. P. & Wu, C. L. Apolipoprotein A5 regulates intracellular triglyceride metabolism in adipocytes. *Mol Med Rep* 16, 6771-6779 (2017).
- 40 Gerst, F. et al. The Expression of Aldolase B in Islets Is Negatively Associated With Insulin Secretion in Humans. *J Clin Endocrinol Metab* 103, 4373-4383 (2018).
- 41 Lynes, M. D., Schulz, T. J., Pan, A. J. & Tseng, Y. H. Disruption of Insulin Signaling in Myf5-Expressing Progenitors Leads to Marked Paucity of Brown Fat but Normal Muscle Development. *Endocrinology* 156, 1637-1647 (2015).
- 42 Montanari, T., Posic, N. & Colitti, M. Factors involved in white-to-brown adipose tissue conversion and in thermogenesis: a review. *Obes Rev* 18, 495-513 (2017).
- 43 Scarpace, P. J., Matheny, M., Borst, S. & Tumer, N. Thermoregulation with age: role of thermogenesis and uncoupling protein expression in brown adipose tissue. *Proc Soc Exp Biol Med* 205, 154-161 (1994).
- 44 Kirov, S. A., Talan, M. I., Kosheleva, N. A. & Engel, B. T. Nonshivering thermogenesis during acute cold exposure in adult and aged C57BL/6J mice. *Exp Gerontol* 31, 409-419 (1996).
- 45 Alexander, G. & Bell, A. W. Quantity and calculated oxygen consumption during summit metabolism of brown adipose tissue in new-born lambs. *Biol Neonate* 26, 214-220 (1975).
- 46 Lidell, M. E. et al. Evidence for two types of brown adipose tissue in humans. *Nat Med* 19 (2013).
- 47 Carstens, G. E. Cold Thermoregulation in the Newborn Calf. *Vet Clin N Am-Food A* 10, 69-106 (1994).
- 48 Smith, S. B., Carstens, G. E., Randel, R. D., Mersmann, H. J. & Lunt, D. K. Brown adipose tissue development and metabolism in ruminants. *J Anim Sci* 82, 942-954 (2004).
- 49 Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760 (2009).
- 50 McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20, 1297-1303 (2010).
- 51 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38, e164 (2010).

- 52 Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38, 904-909 (2006).
- 53 Plotree, D. & Plotgram, D. PHYLIP-Phylogeny inference package (version 3.2). Vol. 5 (1989).
- 54 Junier, T. & Zdobnov, E. M. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26, 1669-1670 (2010).
- 55 Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19, 1655-1664 (2009).
- 56 Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* 27, 2156-2158 (2011).
- 57 Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81, 559-575 (2007).

Figures and Legends

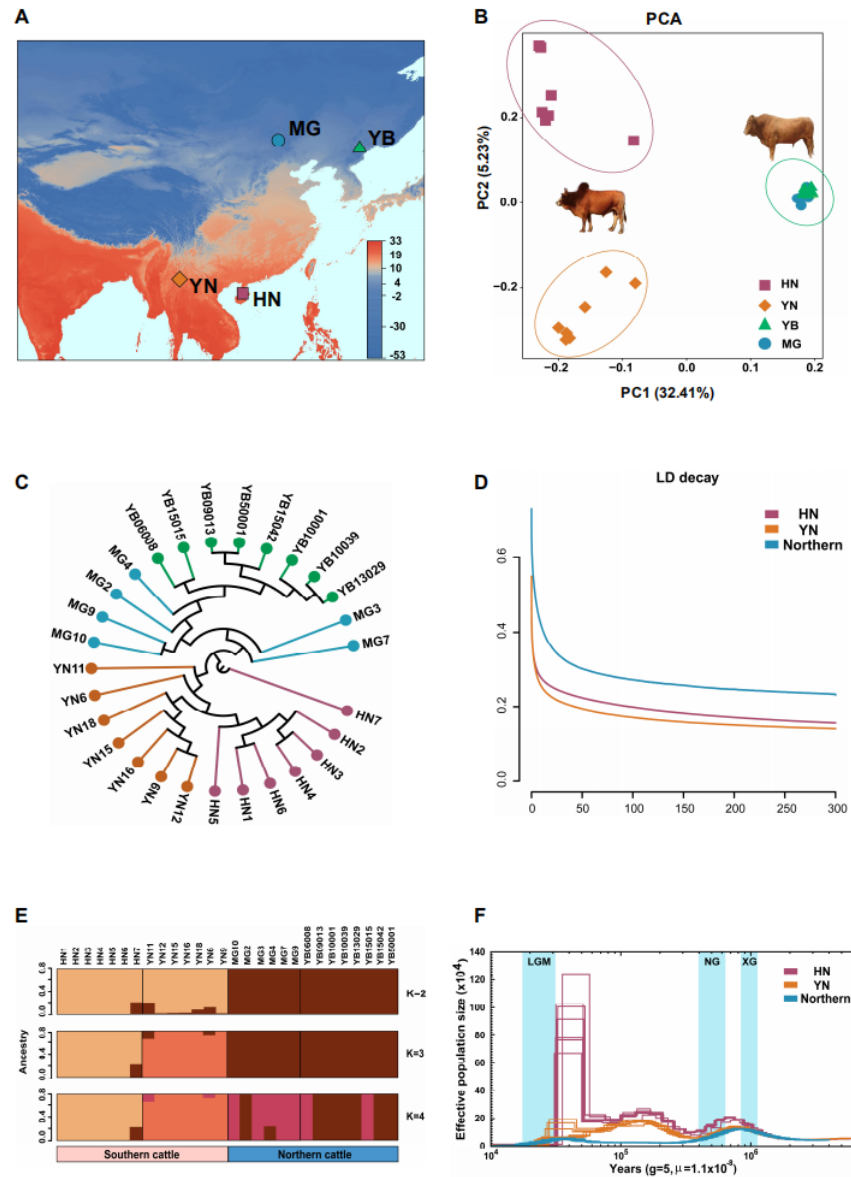


Figure 1. Phylogeny and population history analysis.(A) Geographical distribution of the selected cattle in this study. The map was drawn using annual temperature data (www.worldclim.org). MG, Mongolia cattle; YB, Yanbian cattle; YN, Yunnan cattle; HN, Hainan cattle. (B) Principal component (PC) analysis, PC1 against PC2. (C) Neighbor-joining tree of the relationships between the four cattle breeds. (D) LD decay in HN cattle, YN cattle and northern cattle. (E) Genetic structure of cattle breeds using the ADMIXTURE program. (F) Demographic history inferred by PSMC. The period of the Xixiabangma Glaciation (XG, 1,170–800 kya), Naynayxungla Glaciation (NG, 780–500 kya) and the last glacial maximum (LGM, 20 kya).

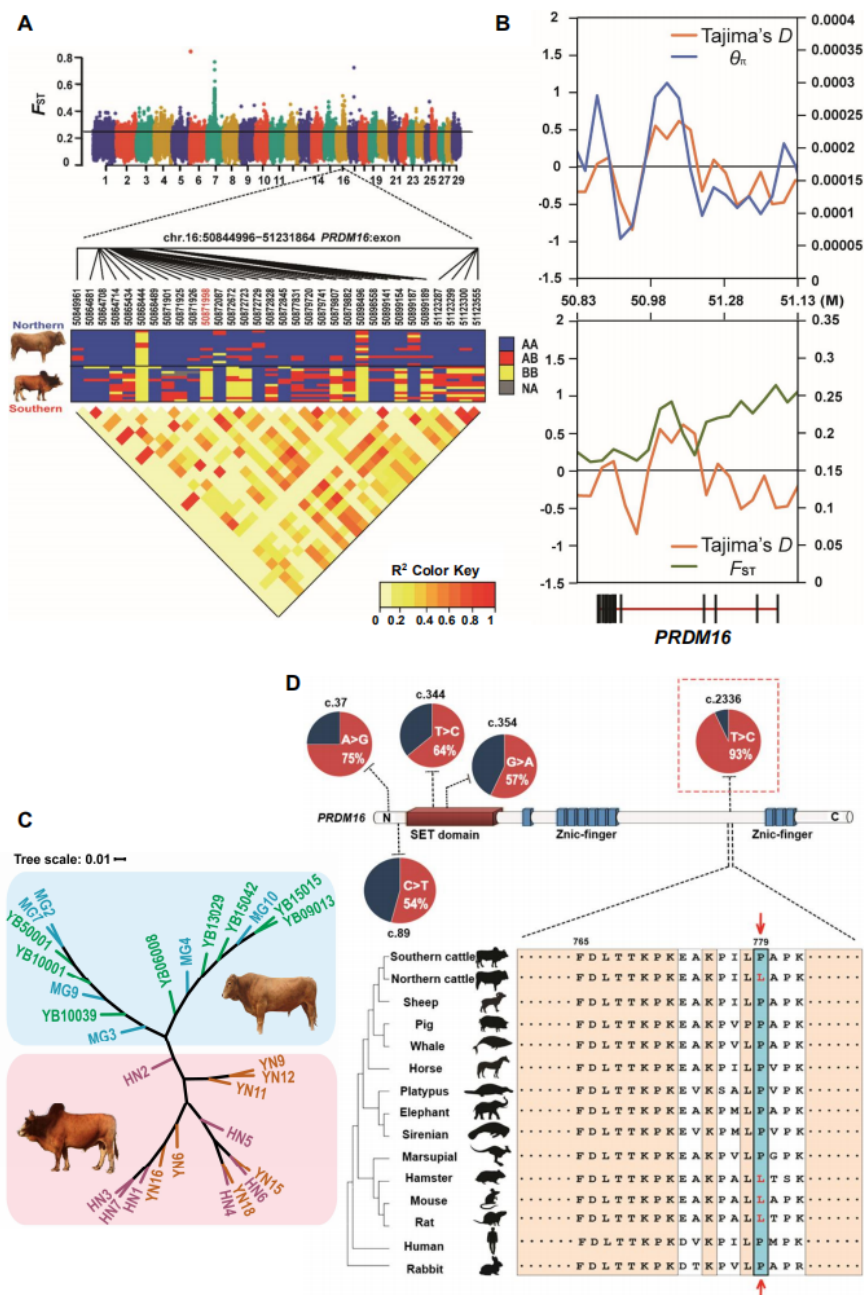


Figure 2. Genome scan of selective sweep. (A) LD analysis and genotype heat map were performed for the non-synonymous and synonymous mutation sites in *PRDM16* exon. (B) The Tajima's D , θ_{π} and F_{ST} values around the *PRDM16* gene. (C) The SNPs tree of *PRDM16*. (D) Mutation rates of the nonsynonymous SNVs in southern cattle and information of mutation site of *PRDM16* in different species.

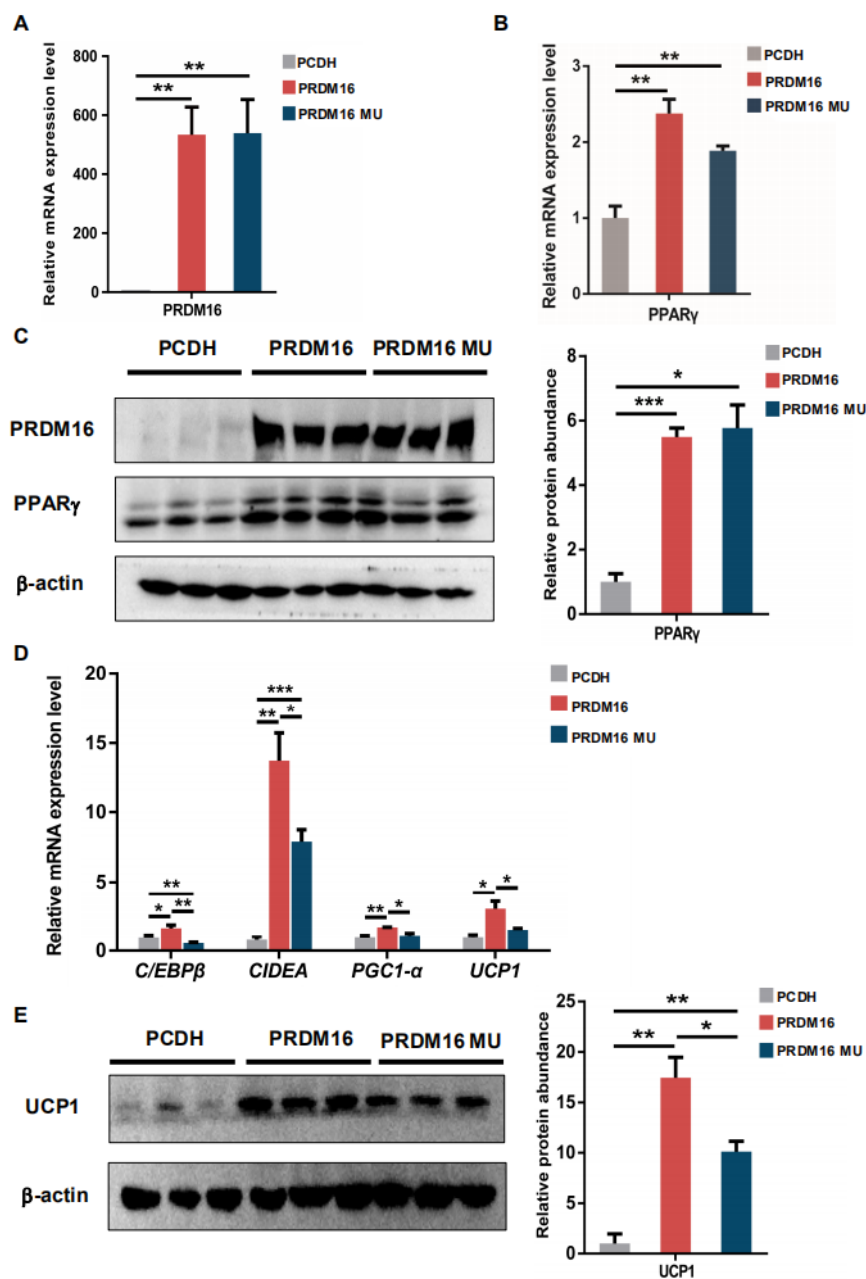


Figure 3. Mutation of *PRDM16* decrease brown adipogenesis. (A) The mRNA level of *PRDM16*. (B) The mRNA level of *PPAR γ* at fully differentiation. (C) Protein expression of *PRDM16* and *PPAR γ* . (D) mRNA levels of brown fat-selective genes analyzed. (E) Protein expression of *UCP1*. Data are means \pm SE. (A, B, E) $n=5-6$ /group; (C, F) $n=3$ /group. One-way ANOVA with Tukey post-hoc test was used to compare groups. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.