Abstract

Background: Mongolian cattle (MC) is one of an ancient livestock breeds with good economic traits such as adaptation to Mongolian Plateau extreme low temperature in winter, resistant to pathogenic organisms infection and high quality meat. To reveal the molecular mechanism underlying these, the whole genome sequencing and comparative transcriptome sequencing of MC were performed.

Results: By genome sequencing and structure variation analysis, 8 genes related to pathogenic organisms infection, including 4 members of beta-defensins gene family (LAP, DEFB1, DEFB2, and DEFB5), 2 members of interferon (IFN) gene family (IFNW1 and IFNT2) and 2 genes coding for BoLA proteins (Mongolian_cattle_21532 and Mongolian_cattle_19448) were found in MC genome inversion region. By
transcriptome-sequencing, it was elucidated that 8 genes (FATP, FABP, PEPCK, SCP-X, ADIPO, FABP1, SCD-1, APO) related to PPARα pathway, 23 genes involved in oxidative phosphorylation and 10 P450 genes (CYP7B1, CYP4V2, CYP3A5, CYP11A1, CYP2C18, CYP2B6, CYP7A1, CYP2R1, CYP2E1, CYP27B1) were significantly up-regulated in winter MC transcriptomes, comparative to summer MC transcriptomes.

Conclusions: Here, we characterized 41 genes implicated in fatty acid metabolism were up-regulated in winter MC transcriptomes. These genes probably account for the MC adaptation to extreme low temperature. At the same time, 8 genes in MC genome inversion region were discovered. And these genes gave clue to the MC resistant to pathogenic organisms infection. In sum, our revealed genes are of important for us to understand the molecular mechanisms of MC to adapt to their environments and valuable for cattle breeding.

Key words: Mongolian cattle, Genome-sequencing, transcriptome-sequencing, extreme-environment, adaptation.

1. Background

Environmental stress has been a major driving force in the evolution of living organisms (Parsons 2005). Environmental stress may affect species abundance and boundaries, hence affecting biodiversity at all levels, of genes, genomes, individuals, population, species, communities and biota (Nevo 1998).

The Mongolian Plateau (MP) is a highland of the Eurasian continent lying between 87°40′-122°15′N and 37°46′-53°08′E. The varieties of climate and geographical condition in the MP endowed this ancient plateau with high biodiversity. Nevertheless, the harsh environmental condition of this plateau is a big challenge for all organisms lived here. And the animals live here provide optimal models for investigation of how psychrophiles adaptation to its’ extreme cold. Polar bears are typical mammals uniquely adapted to life in the High Arctic. Population genomic sequencing of polar bears along with brown bears revealed that in response to Arctic climates and hyperlipid diet, genes on the polar bear lineage have been under stronger positive selection than in brown bears; nine of the top 16 genes under strong positive selection are associated with cardiomyopathy and vascular disease, implying important reorganization of the cardio-vascular system (Liu et al 2014). Arctic fox is another mammal able to survive in Arctic harsh conditions. Several evolutionary adaptations such as insulating and camouflage colored fur, a compact body, and reduced metabolism during starvation or extremely cold weather enable Arctic foxes such capacity (Audet et al. 2002).
Transcriptome sequencing of Arctic foxes and their close relative, red foxes, revealed that several genes involved in various metabolic and molecular processes such as energy metabolism, cardiac gene regulation, apoptosis and blood coagulation were under positive selection in both species, whereas four of genes are under positive selection only in Arctic foxes transcriptome, two of which are fat metabolism genes (Kumar et al. 2015). In addition to extreme cold, hypoxia and drought are the other two stresses may affect animals living in MP. The Qinghai-Tibetan Plateau is known for its extreme environment with low atmospheric oxygen pressure, cold climate, and limited resources (Wu et al. 2001). And genome sequencing of Tibetans (Simonson et al. 2010), yak (Qiu et al. 2012), Tibetan antelope (Ge et al. 2013), snow leopard (Cho et al. 2013) and Tibetan mastiff (Gou et al. 2014) gave the insights of these extreme high altitude mammals how to adapt to Tibetan hypoxic and extreme cold environments. Drought is another stress with which animals living in MP must be confronted. A few works were reported of mammals lived in desert environments how to adapt to drought. Genomes of sheep in contrasting environments were sequenced and further comparative analyses were preformed. It was revealed that to adapt to desert environments, 11 genes which localized in three pathways, the arachidonic acid metabolism pathway (ANXA6, GPX3, GPX7, and PTGS2), the renin-angiotensin system pathway (CPA3, CPVL, and ECE1) and the oxytocin signaling pathway (CALM2, CACNA2D1, KCNJ5, and COX2) were positively selected in the Taklimakan desert group sheep. All of these positively selective genes were functionally related to regulating water retention and reabsorption in renal cells and blood vessels in the kidney (Yang et al. 2016). Camelus are uniquely evolved to fit desert harsh environments. Genome sequencing and comparative transcriptome sequencing revealed Na⁺/K⁺–ATPase, the epithelial Na⁺ channel (ENaC), aquaporin family genes (AQP1, AQP2 and AQP3) were involved in water and salt reabsorption which are evolved for camel adaptation to desert extreme drought environments (Wu et al. 2014).

Cattle had a pivotal position in the advancement of human culture and civilization. Ever since its domestication, cattle distributed almost all continents. To fit for their divergent environments, cattle have involved many breeds with different economic traits and environment adaptation traits. For many cattle breeds, the mechanisms underlying their economic traits were revealed by genome-wide association studies (Takeda et al. 2019, Smith et al. 2019, Cruz et al. 2019, Gonzalez-Ruiz et al. 2019, Bedhane et al. 2019). However, few works were reported to address cattle living in different ecosystems how to adapt to their environments, particularly, Mongolian cattle (MC) how to adapt to MP harsh environments. Here, by genome sequencing and comparative transcriptome sequencing, we proved that structure variants play roles in Mongolian cattle adaptation to their extreme environments as well as to apply these knowledge to livestock breeding.
2. METHODS

2.1 DNA sampling, extraction and sequencing

MC living in Khövsgöl Province of Mongolian with pure blood was used in this research. Blood samples collected from target animal’s arteria carotis by veterinarian. And genomic DNA was extracted from blood samples according to the instrument of DNA extraction kits (Qiagen). For sequencing, Pair-end libraries of different insert sizes (from 170bp to 20kb) were prepared, and then they were sequenced on IlluminaHiseq 2000 system.

2.2 Phylogenetic analyses

Previously, including MC, we’ve sequenced 4 MP livestock genome. The other animals are Mongolian sheep (MS), Mongolian horse (MH) and Alxa Bactrain camel. By aligning MC genome with MS, MH genome and sheep, cattle, horse, mouse, human genome extracted from NCBI data, individually, the divergence among them was revealed and the phylogenetic tree established.

2.3 Genome structure variation analyses

By comparative analysis of MC genome with published cattle (*Bos taurus*, Hereford breed) genome (Elsik et al. 2010), MC genome structure variations including SNPs (Single Nucleotide Polymerphism), short InDels, CNVs (Copy Number Variations), SDs (Segmental Duplications) and IVs (Inversion Variants) were analyzed.

2.4 Transcriptome sequencing

For transcriptome sequencing, we choose 3 groups of animals termed B (Cattle living in low altitude), S (MC living in summer) and W (MC living in winter). Each group has 3 replicates. After slaughtered in slaughtering house, four kinds of tissues (skin, liver, back fat, muscle) were carefully collected and quickly put into liquid nitrogen. In lab, the total RNA was extracted from above tissues according to the instrument of RNA extraction kits (Qiagen). Transcriptome sequencing was performed on IlluminaHiseq 2000 system.
2.5 Transcriptome DEGs analysis

By pair-wise comparisons, differentially expressed genes (DEGs) existing in different tissues from different group of animals were revealed. Focused on their functions, DEGs involved in MP harsh environment adaptation were enriched in PPAR signaling pathway, Oxidative phosphorylation pathway, PI3K-Akt signaling pathway and P450 gene family.

3 RESULTS

3.1 Genome summary and evolution

We sequenced the genomes of MC using Illumina Hiseq 2000 system. Pair-end libraries of different insert sizes (from 170bp to 20kb) were prepared and generated 292.67 Gb raw sequence data for MC (Supplementary Table 1 and 2). Using 17-mer analyzing, the size of genome was estimated to be 3.2Gb (Supplementary Table 3, Supplementary Fig. 1). The total size of the assembled MC was ~2.68Gb which close to the 2.66Gb and 2.65Gb of sequences of yak and cattle genome (UMD3.1), respectively, with scaffold N50 of 1.58Mb and contig N50 of 27.16kb (Supplementary Table 4, Supplementary Fig. 2-3). From gene structural and functional annotation using genomic data, we obtained in total of 22,098 protein coding genes for MC, which is in parallel with that of yak (22,282) (Supplementary Tables 6-8, Supplementary Fig. 4-5). We further compared the gene families of MC, MS, MH, cattle, sheep, horse and human by using CAFE to find the expansion and contracted gene families. We identified 577 expansion genes of MC (Fig. 1a) and GO enrichment analysis revealed they are enriched in the housekeeping gene families, such as 5S rRNA binding and nucleosome assembly (Supplementary Table 9-11).

We sequenced the genomes of MC using Illumina Hiseq 2000 system. Pair-end libraries of different insert sizes (from 170bp to 20kb) were prepared and generated 292.67 Gb raw sequence data for MC (Supplementary Table 1 and 2). Using 17-mer analyzing, the size of genome was estimated to be 3.2Gb (Supplementary Table 3, Supplementary Fig. 1). The total size of the assembled MC was ~2.68Gb which close to the 2.66Gb and 2.65Gb of sequences of yak and cattle genome (UMD3.1), respectively, with scaffold N50 of 1.58Mb and contig N50 of 27.16kb (Supplementary Table 4, Supplementary Fig. 2-3). From gene structural and functional annotation using genomic data, we obtained in total of 22,098 protein coding genes for MC, which is in parallel with that of yak (22,282) (Supplementary Tables 6-8, Supplementary Fig. 4-5). We further compared the gene families of MC, MS, MH, cattle, sheep, horse and human by using CAFE to find the expansion and contracted gene families. We identified 577 expansion genes of MC (Fig. 1a) and GO enrichment analysis revealed they are enriched in the housekeeping gene families, such as 5S rRNA binding and nucleosome assembly (Supplementary Table 9-11).

We identified 7,580 one2one orthologous genes by comparing MC, Mongolian sheep (MS), Mongolian horse (MH), sheep, cattle, horse, mouse and human sequences and built the phylogenetic tree. Mongolia mammals and corresponding domestic ones were diverged about 2-3 million years ago (Mya, MC-cattle: 1.8 Mya, MS-sheep: 2.6 Mya and MH-horse: 2.7 Mya) (Fig. 1a), indicating their divergence maybe influenced with the environment and climate at that time. To identify the population size history, PSMC analysis of MC, MS, MH, cattle, sheep and horse also revealed they diverged about 3 Mya (Fig. 1b). We found population size changes of Mongolia and their domestic ruminants are
similar along the history. Interestingly, all of them decreased to the low about ~0.1 Mya which is fitly overlap the peak of atmospheric surface air temperature distribution, suggesting the substantial population decreasing maybe resulted from the temperature sharply increasing (Fig. 1b).

### 3.2 Variation analysis

Using cattle genome as reference sequences, we totally detected ~7.62 million single nucleotide variations and short InDels, 4689 copy number variations (block length >5kb, 3750 deletions and 939 duplication) and 32,944 structural variations (19931 deletions, 4343 insertions and 8670 inversions) of MC, which will provide abundant genetic heterogeneity for use in future breeding application (Fig. 2). We also identified the segmental duplication (SD) in MC cattle, and found 251 specific SDs in its’ genome. GO analysis showed these SD genes were enriched in the categories including defense response (P=1.26E-7), response to stress (P=3.05E-3) and response to virus (P=2.24E-5), which might be relevant to adaption to extreme environment (Fig. 3a). Genes identified in specific SD area in MC were enriched in a GO term called olfactory receptor activity and they belong to different subfamilies of olfactory receptor proteins (Fig. 3b). Olfactory receptors are members of class-A rhodopsin-like gene family of G protein-coupled receptors, which play important roles in signal transduction cascade (Gaillard et al., 2004, Hussain et al. 2009). In cattle, understanding the OR gene repertoire could lead to link the effects of genetic differences in these genes to variation in olfaction. A study conducted by Lee K. etc. finds that the cattle genome contains 1,071 OR-related sequences including 881 functional, 190 pseudo, and 352 partial OR sequences and they also noticed significant copy number variations are present in the OR repertoire of the cattle (Lee et al. 2013).

We also totally identified 749 inversion variants (IVs) in MC-cattle based on the whole genome alignment results. This is challenging because these inverted regions contain many genes, thus we mainly do enrichment analysis for all the genes in the vicinity of breakpoints of inversion (1kb up- and downstream flank). We found seven genes (Mongolian_cattle_16163, Mongolian_cattle_18647, Mongolian_cattle_03540, Mongolian_cattle_11348, Mongolian_cattle_20851, Mongolian_cattle_19494 and Mongolian_cattle_14394) in MC were enriched in a biological process called defense response in which Beta-defensins gene family and Interferon (IFN) gene family are implicated (Supplementary Fig. 6). In our data, the Beta-defensins genes are *LAP, DEFB1, DEFB5*, and *DEFB7*. The lingual antimicrobial peptide (LAP) is involved in the innate system and plays an important roles in killing a large variety of microbes, and immunolocalization of it has revealed its existence in both
infected and non-infected alveolus in cattle (Isobe et al. 2009). Study pointed out that the mRNA abundance of DEFB1 (Roosen et al. 2004) and DEFB5 in the cattle mammary glands increased, suggesting beta-defensins may play a role in local host defense during under infections. IFNs are members of cytokines and the release of them could trigger the protective defenses of the immune system to help eliminate pathogens (Parkin and Cohen 2001). In our analysis, we found Interferon-omega 1 (IFNW1) and Interferon tau-2 (IFN) located in the breakpoints. Mo_sheep_14980 of MS and Mongolian_cattle_14394 of MC are LOC781948, which is interferon Omega-1 (Type I interferon). Studies showed IFNT2 had antiviral activity in different cell lines and antiluteolytic effect in cyclic ewes (Boue et al. 2000). In general, these two IFNs mainly perform antiviral activity to strengthen the immune system.

We also found two genes encoding major histocompatibitily complex (MHC) class II protein complex (Mongolian_cattle_21532 and Mongolian_cattle_19448) in the breakpoints (Supplementary Fig. 7). MHC molecules are transmembrane proteins, which localized in cell surface and involved in antigen recognition, immunologic rejection of homotransplantations, immune response and regulation. Amorena and his colleagues first reported MHC class I antigens in cattle, and named it as the bovine leucocyte antigen, BoLA system (Amorena and Stone, 1978). BoLA genes have tight connections to breeding and disease control. The features of BoLA make it become a hot spot in disease resistance area in breeding.

3.3 Transcriptomes analysis

We sequenced transcriptomes of four tissues (skin, liver, back fat, muscle) in three states (B represents cattle, S is the Mongolian cattle sequenced in summer, W is the Mongolian cattle sequenced in winter) (Supplementary table 12). Based on the differentially expressed genes (DEGs) data, we have conducted pathway analysis. By pair-wise comparisons in four tissues, we identified 710, 523, 403, and 383 DEGs in back fat, skin, muscle and liver, respectively. Based on the DEGs, we performed the enrichment analysis and found significantly enriched DEGs with functions terming in response to physiological adaptation (Supplementary Fig. 6-9).

3.4 Enriched DEGs pathway analysis

3.4.1 PPAR signaling pathway
The first significant enriched pathway of DEGs was PPAR signaling pathway, with 16 DEGs (B-W) and 13 DEGs (S-W) significantly enriched in this pathway, while in B-S, no DEGs were enriched. Peroxisome proliferator-activated receptors (PPARs) are lipid-activated transcription factors belonging to steroid acid receptor superfamily. All of the encoded proteins are involved in lipid homeostasis (Lemberger et al. 1996). PPARα is the principle transcriptional regulator of FAO enzyme genes, which can control the capacity to produce energy via the utilization of fats by mammalian postnatal heart (Barger et al. 2000). And other researches pointed out that the PPARG/RXRA complex triggers activation of target genes involved in fatty acid oxidation resulting in increased triglyceride formation by ATP production (Lim et al. 2015). In B-W group, genes involved in this pathway like FATP, FABP, PEPCK, SCP-X, ADIPO, FABP1, SCD-1, APO were all up-expressed in W state cattle, indicating an up-regulating PPAR pathway in Mongolian cattle skin in winter (Fig. 4a and 4b). As for the S-W group, results were consistent with the B-W group, genes such as SCD, ACBP etc. were up-expressed in W cattle, again showing an enhancement of PPAR pathway in Mongolian cattle skin in winter (Fig. 4a and 4b). Based on the data, we speculated that the synthesis and metabolism of fatty acids in Mongolian cattle skin were greater than normal buffalo, and between Mongolian cattle, the signaling was also more strong in winter, indicating its connection to the cold adaption of Mongolian cattle.

3.4.2 Oxidative phosphorylation pathway

We also found oxidative phosphorylation pathway was enriched by DEGs. Although in S-W group, no DEGs were enriched in this pathway, we did found 23 DEGs in B-S and 20 B-W enriched in this pathway (Fig. 5a). And most of these DEGs were up-regulated expressed in S and W Mongolian cattle, respectively, suggesting a stronger oxidative phosphorylation process in Mongolian cattle than in other cattle breeds. Oxidation phosphorylation is the metabolic pathway in which mitochondria in cells use energy released by nutrients oxidation to form ATP, therefore stronger oxidation phosphorylation reaction in Mongolian cattle means a larger amount of ATP synthesis in Mongolian cattle skin, further indicating a stronger energy use in Mongolian cattle, which may be connected to cold-resistance.

3.4.3 PI3K-Akt signaling pathway
We also noticed the downstream part of the PI3K-Akt signaling pathway (Fig. 5B). PI3K-Akt transduction pathway is directly related to cellular quiescence, proliferation, cancer, and longevity. And some studies pointed out that the inhibition of PI3K pathway could indirectly decreases the expression of the glucose transporter GLUT1 at the plasma membrane (Riler et al. 2006). In our data, in S-W group, we could found that GYS, PEPCK, G6Pase were up-regulated in W Mongolian cattle (5C), showing a more active Glycolysis and Gluconeogenesis process in W cattle, which may indicate in winter, the energy production was more prosperous.

3.4.4 Complement and coagulation cascades pathway
Besides, we also found a totally differentially expressed pathway called complement and coagulation cascades pathway between the skin tissues and muscle tissues, which is connected to the immunocompetence of the species. On skin vs skin group, there was no DEGs enriched in this pathway between B-S group, while 23 and 13 DEGs in B-W and S-W group were significantly enriched, respectively, and all of these DEGs were up-regulated in W state Mongolian cattle (Fig.6A and 6B). On muscle vs muscle group, no DEGs in S-W group were enriched, while 21 and 25 DEGs in B-S and B-W group were significantly enriched in this pathway, respectively, and all of these DEGs (except C3AR1) were up-regulated in B cattle, which is different from the results of skin group (Fig. 6a and 6b), suggesting a different immune state in skin and muscle tissues in these two kinds of cattle. Among the up-regulated genes, most encode proteins like coagulation factors, complement components and other proteins related to immunity. Kininogen-1 (KNG1), also known as alpha-2-thiol protease inhibitor, Williams-Fitzgerald-Flaujeac factor or the HMWK-kallikrein factor is a protein that in humans is encoded by the KNG1 gene. Kininogen-1 is the precursor protein to high-molecular-weight kininogen (HMWK), low-molecular-weight kininogen (LMWK), and bradykinin. Bradykinin is a potent endothelium-dependent vasodilator, leading to a drop in blood pressure. It also causes contraction of non-vascular smooth muscle in the bronchus and gut, increases vascular permeability and is also involved in the mechanism of pain (Mutschler et al. 1997). Bradykininalso causes natriuresis, contributing to the drop in blood pressure.

3.4.5 P450 gene family
Recent studies have indicated that cytochrome P450 is a key enzyme in drug metabolism, and has important effects on cell factor and body temperature regulation. In MC, we found 10 up-expressed P450 genes including CYP7B1, CYP4V2, CYP3A5, CYP11A1, CYP2C18, CYP2B6,
Cytochromes P450 (CYPs) belong to the superfamily of proteins containing a heme cofactor and, therefore, are hemoproteins. CYPs use a variety of small and large molecules as substrates in enzymatic reactions. It participates in the metabolism of endogenous substances and exogenous substances including drugs and environmental compounds. In the cell, CYPs are mainly distributed in the endoplasmic reticulum and the inner membrane of the mitochondria.

4. DISCUSSION

MC is an ancient cattle breed with valuable economic traits, such as adaptation to harsh environments, resistant to pathogenic organisms infection and high quality meat. By complete mitochondrial DNA D-loop sequences, it’s revealed that the MC has the highest genetic diversity (Xia et al 2018), which may account for their varied economic traits. However, the detail molecular mechanism underlying above economic traits remain to be elucidated.

The notable trait of MC is its’ tolerance of Mongolian Plateau extreme cold weather in winter. Adipose tissue, including Brown Adipose Tissue (BAT), White Adipose Tissue (WAT) and newly revealed adipocyte termed beige adipocyte, play crucial roles in mammal cold adaptation. In polar bear or belugas, blubbers (deep subcutaneous fat) serve as for insulation keeping heats from lose. Small mammals’ BATs are important organs for thermogenesis in cold adaptation. After cold acclimation, a serial of reactions, including releasing of noradrenaline from sympathetic nervous system, binding of it to its’ receptor, activating a transcriptional coactivator PGC-1α, up-regulating genes involved in fatty acid transport, β-oxidation, TCA cycle, electron transport system, glycolysis, lipogenesis and PPARα activation pathways occur in BATs (Kim et al. 2016). Finally, the proton gradient in mitochondria established and such gradient converted into heats by the function of uncoupling protein1 (UCP1) (Nedergaard et al. 2001). In mammals without BAT, such as cold tolerant pig breeds, Tibetan pig and Min pig, when expose to cold climate, WATs first are browning to be beige adipocytes, where similar thermogenesis process occurred in BAT happening (Lin et al. 2017).

Except for WATs, myogenic intermediates are able to differentiate into beige adipocytes (Chen et al. 2019), which implying skeletal muscle implicated in the nonshivering thermogenesis (NST) process, too. In our transcriptome data, we found 8 genes (FATP, FABP, PEPCK, SCP-X, ADIPO, FABP1, SCD-1, APO) related to PPARα pathway, 23 genes involved in oxidative phosphorylation and 10 P450 genes (CYP7B1, CYP4V2, CYP3A5, CYP11A1, CYP2C18, CYP2B6, CYP7A1, CYP2R1, CYP2E1, CYP27B1) were significantly up-regulated in winter MC transcriptomes. Since all of these genes implicated in fatty acid metabolism, we proposed that similar to cold tolerant pig breeds, MC in cold
adaptation could convert the WAT into beige adipocytes, where heats were generated by metabolism of fatty acids and glucose to sustain body core temperature. Unlike polar region climates, where all the year is extreme cold, in summer, the temperature in MP, especially in semiarid regions such as Ordos steppe, is high. Therefore, precious detection of such temperature changes also critical for MC fitting for their environments. Genome structure variation analysis indicated SD area in MC enriched in a GO term called olfactory receptor. The main function of olfactory receptor is perception of chemical and climate factors. Therefore, genome structure variation in MC olfactory receptor probably underlying the molecular mechanism of MC how to precisely detect the temperature change.

Resistant to pathogenic organisms infection is another good trait of MC. Via GWAS (genome-wide association study), molecule based on different cattle breeds such as Holstein, *Bos taurus*×*Bos indicus* crossbred cattle and Angus resistant to mastitis ((Kurz et al. 2019), foot-and-mouth disease (FMD) (Lee et al. 2015), tick (Otto et al. 2018), parasite (gastrointestinal nematodes) (Hou et al. 2012) were found out. Here, by whole genome sequencing and genome structure variation analysis, we found millions (~7.62 million) of SNPs and numerous of structure variations (including CNVs, segmental duplications, SDs, inversions). By critical analysis of vicinity of breakpoints of inversion, we found 8 genes related to pathogenic organisms infection. 4 of them (*LAP, DEFB1, DEFB2, and DEFB5*) belong to Beta-defensins gene family. It’s well characterized that as an antimicrobial peptide, LAP (lingual antimicrobial peptide) is involved in the innate system and has versatile antimicrobial activities (Isobe et al. 2009). The investigation of cDNA from cattle mammary gland showed that many members of beta-defensins gene family were expressed, which implied these robust antimicrobial peptides playing important roles for cattle resistant to mastitis. 2 of them belong to interferon (IFN) gene family. In our data, they are *IFNW1* and *IFNT2*. Interferon is known for its’ anti-virus activities. In this case, *IFNW1* and *IFNT2* are in the vicinity of breakpoints of inversion. Such structure variation probably up-regulated the expression of these anti-virus proteins, which may accounts for MC more resistant to FMD infection than other breeds. Another two genes are coding for MHC class II proteins. In cattle, MHC class I proteins were first found by Amorena and named them BoLA system (For Bovine Leucocyte Antigen). Ever after being found, for their importance in host immunities, BoLAs have been extensively studied and their roles in bovine mastitis (Sharif et al. 1998), bovine leukemia virus infection (Stear et al. 1998) and parasitic load (Martinez et al. 2006) were well documented. Here, we found two genes coding for BoLA proteins are localized in MC genome inversion region. Such structure change mostly will affect their expression and in turn affect their anti-pathogenic organism infection activities.
MC was considered as a good cattle breed with high quality meat. The criterion used to assess beef includes fat percentage (FAR), unsaturated fat acid (oleic acid) percentage and fatty acid composition. Still by GWAS, candidate genes involved in Japan Black Cattle above trait were revealed respectively (Nakajima et al. 2018, Kawaguchi et al. 2019, Sasago et al. 2016). From our MC genome and transcriptome data, we didn’t obtain the direct evidence to support that MC have high quality meat. However, as discussed, FAR is important for mammals living in high altitude to adapt to their environments. To adapt to Arctic extreme cold, half of polar bear mass is composed of fat tissue. And unsaturated fat acid is important for thermogenesis. While COX (Cyclooxygenases)-derived oxylipins (unsaturated fat acid such as omega-3 lipid oxidation products) is required for the formation of beige adipocyte (Vegiopoulos et al. 2010), the P450 expoxyxygenase and epoxide-derived oxylipins, (±) 12, 13-dihydroxy-9Z-octade-cenoic acid (12,13-diHOME) can induct the fatty acid uptake into BAT and skeletal muscle (Lynes et al. 2017). In MC, no BAT were observed, which indicated that in response to extreme cold weather in winter, MC firstly need brown WAT into beige adipocyte, then uptake fatty acid into these nearly formed adipocyte to generate heats. The prerequisite of these two processes is the existence of large amounts of unsaturated fat acid. The indirect evidence is that in winter MC transcriptomes, 10 P450 genes were significantly up-regulated, which implied these up-regulated genes may involved in the oxidation of unsaturated fat acid into oxylipins.

MC is a valuable cattle breed. Nevertheless, so far, few research works performed on this important livestock. By genome sequencing association with comparative transcriptome sequencing, we revealed the molecular mechanisms of MC how to adaptation to MP harsh environments as well as resistant to pathogenic organisms. Undoubtedly, these uncovered genes will become cherished genetic resources for cattle breeding.

5. Conclusion

In this paper, by whole genome sequencing and comparative transcriptome sequencing, genes localized in MC genome inversion region and genes up-regulated in winter MC transcriptomes were characterized, which probably account for MC adaptation to MP extreme environments and resistant to pathogenic organisms infection. Therefore, these genes are of valuable for cattle breeding.
Abbreviations:

MC: Mongolian Cattle; MP: Mongolian Plateau.

Declarations

Ethics approval and consent to participate: Procedures involving animals, their care and humane kill were conducted in conformity with Guidelines on the Humane Treatment of Laboratory Animals (HTLA Pub. Chapter2-6, revised 2006 in China).

Consent for publication: Not applicable.

Availability of data and material: All data generated or analyzed during this study are included in this manuscript.

Competing interests: The authors declare that they have no competing interests.

Funding: Not applicable.

Authors’ contributions: YP Xing was a major contributor in writing the manuscript, analyzed and interpreted the genome data. Y Qi analyzed and interpreted the transcriptome data. The collection of tissue samples and extraction of DNA and RNA were performed by C Purev, SY Wang, HB Wang, KF Wu, JW Cao, CX Liu, YY Liu, LU Li. Professor YR Zhang and HM Zhou reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements: We appreciated WK Jing for his kindly help in sample collection.
References

Amorena B and Stone WH. 1978. Serologically defined (SD) locus in cattle. Science. 201:159-160.

Audet AM, Robbins CB, Larivière S, Alopexlagopus. 2002. Mamm Species. 713:1–10.

Barger PM, and Kelly DP. 2000. PPAR signaling in the control of cardiac energy metabolism. Trends in cardiovascular medicine. 10:238-245.

Bedhane M, van der Werf J, Gondro C, Duijvesteijn N, Lim D, Park B, Park MN, Hee RS, Clark S. 2019. Genome-Wide Association Study of Meat Quality Traits in Hanwoo Beef Cattle Using Imputed Whole-Genome Sequence Data. Front Genet. 10:1235.

Boue O, Garcia JL, Villar M, Alazo K, Perez A, Ramos E, Morales C, Morera OL, Redondo M, Montero C, Rodriguez M. 2000. Antiviral and antiluteolytic activity of recombinant bovine IFN-omega1 obtained from Pichia pastoris. Journal of interferon & cytokine research. 20: 677-683.

Chen Y, Ikeda K, Yoneshiro T, Scaramozza A, Tajima K, Wang Q, Kim K, Shinoda K, Sponton CH, Brown Z, et al. 2019. Thermal stress induces glycolytic beige fat formation via a myogenic state. Nature. 565: 180-185.

Cho YS, Hu L, Hou H, Lee H, Xu J, Kwon S, Oh S, Kim HM, Jho S, Kim S. et al. 2013. The tiger genome and comparative analysis with lion and snow leopard genomes. Nat Commun. 4: 2433.

Cruz VAR, Oliveira HR, Brito LF, Fleming A, Larmer S, Miglior F, Schenkel FS. 2019. Genome-Wide Association Study for Milk Fatty Acids in Holstein Cattle Accounting for the DGAT1 Gene Effect. Animals. 9: 997-1009.

Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, Weinstock GM, Adelson DL, Eichler EE, Elnitski L, Guigó R. et al. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science. 324: 522-528.

Gaillard I, Rouquier S, and Giorgi D. 2004. Olfactory receptors. Cell Mol Life Sci. 61: 456-69.
Ge RL, Cai QL, Shen YY, San A, Ma L, Zhang Y, Yi X, Chen Y, Yang LF, Huang Y. et al. 2013. Draft genome sequence of the Tibetan antelope. *Nat Commun*. 4: 1858.

González-Ruiz S, Strillacci MG, Durán-Aguilar M, Cantó-Alarcón GJ, Herrera-Rodriguez SE, Bagnato A, Guzmán LF, Milián-Suazo F, Román-Ponce SI. 2019. Genome-Wide Association Study in Mexican Holstein Cattle Reveals Novel Quantitative Trait Loci Regions and Confirms Mapped Loci for Resistance to Bovine Tuberculosis. *Animals*. 9: 636-651.

Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z. et al. 2014. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Res.* 24: 1308–1315.

Hou YL, Liu GE, Bickhart DM, MatukumalliLK, Li CJ, Song JZ, Gasbarre LC, Van Tassell CP, Sonstegard TS. 2012. Genomic regions showing copy number variations associate with resistance or susceptibility to gastrointestinal nematodes in Angus cattle. *Funct Integr Genomics*. 12: 81-92.

Hussain A, Saraiva LR, and Korschling SI. Positive Darwinian selection and the birth of an olfactory receptor clade in teleosts. 2009. *Proc Natl Acad Sci U S*. 106: 4313-4318.

Isobe N, Hosoda K, Yoshimura Y. 2009. Immunolocalization of lingual antimicrobial peptide (LAP) in the bovine mammary gland. *Animal science journal = Nihon chikusan Gakkaiho*. 80: 446-450.

Isobe N, Nakamura J, Nakano H, Yoshimura Y. 2009. Existence of functional lingual antimicrobial peptide in bovine milk. *Journal of dairy science*. 92: 2691-2695.

Kawaguchi F, Kigoshi H, Fukushima M, Iwamoto E, Kobayashi E, Oyama K, Mannen H, Sasazaki S. 2018. Whole-genome resequencing to
identify candidate genes for the QTL for oleic acid percentage in Japanese Black cattle. *Anim Sci J.* 90: 467-472.

Kim J, Fernand VE, Henagan TM, Shin J, Huypens P, Newman S, Gettys TW, Chang JS. 2016. Regulation of Brown and White Adipocyte Transcriptome by the Transcriptional Coactivator NT-PGC-1α. *PLoS ONE.* 11: e0159990.

Kumar V, Kutschera VE, Nilsson MA, Janke A. 2015. Genetic signatures of adaptation revealed from transcriptome sequencing of Arctic and red foxes. *BMC Genomics.* 16: 585.

Kurz JP, Yang Z, Weiss RB, Wilson DJ, Rood KA, Liu GE, Wang Z. 2019. A genome-wide association study for resistance in phenotypically well-characterized Holstein dairy cattle using a selective genotyping approach. *Immunogenetics.* 71: 35-47.

Lee BY, Lee KN, Lee T, Park JH, Kim SM, Lee HS, Chung DS, Shim HS, Lee HK, Kim H. 2015. Bovine genome-wide association study for genetic elements to resist the infection of Foot-and-mouth disease in the field. *Asian-Australas J Animal Science.* 28: 166-170.

Lee K, Nguyen DT, Choi M, Cha SY, Kim JH, Dadi H, Seo HG, Seo K, Chun T, Park C. 2013. Analysis of cattle olfactory subgenome: the first detail study on the characteristics of the complete olfactory receptor repertoire of a ruminant. *BMC Genomics.* 14: 596-607.

Lemberger T, Desvergne B and Wahli W. 1996. Peroxisome proliferator-activated receptors: a nuclear receptor signaling pathway in lipid physiology. *Annual review of cell and developmental biology.* 12: 335-363.

Lim D, Chai HH, Lee SH, Cho YM, Choi JW, Kim NK. 2015. Gene Expression Patterns Associated with Peroxisome Proliferator-activated Receptor (PPAR) Signaling in the Longissimus dorsi of Hanwoo (Korean Cattle). *Asian-Australasian journal of animal sciences.* 28: 1075-1083.

Lin J, Cao C, Tao C, Ye R, Dong M, Zheng Q, Wang C, Jiang X, Qin G, Yan C et al. 2017. Cold adaptation in pigs depends on UCP3 in beige
adipocytes. *J Mol Cell Biol.* 9: 364-375.

Liu S, Lorenzen ED, Fumagalli M, Li B, Harris K, Xiong Z, Zhou L, Korneliussen TS, Somel M, Babbitt C. et al. 2014. Population genomics reveal recent speciation and rapid evolutionary adaptation in polar bears. *Cell.* 157: 785–794.

Lynes MD, Leiria LO, Lundh M, Bartelt A, Shamsi F, Huang TL, Takahashi H, Hirshman MF, Schlein C, Lee A, et al. 2017. The cold-induced lipokine 12,13-diHOME promotes fatty acid transport into brown adipose tissue. *Nat. Med.* 23: 631–637.

Martinez ML, Machado MA, Nascimento CS, Silva MV, Teodoro RL, Furlong J, Prata MC, Campos AL, Guimarães MF, Azevedo AL, et al. 2006. Association of BoLA-DRB3.2 alleles with tick (Boophilus microplus) resistance in cattle. *Genet. Mol. Res.* 5: 513-524.

Mutschler E, Schäfer-Korting, Monika. 1997. Arzneimittelwirkungen (in German) (7 ed.). Stuttgart: Wissenschaftliche Verlagsgesellschaft. ISBN 3-8047-1377-7.

Nevo E. 1998. Book review. *Heredity.* 81: 591–597.

Nakajima A, Kawaguchi F, Uemoto Y, Fukushima M, Yoshida E, Iwamoto E, Akiyama T, Kohama N, Kobayashi E, Honda T, et al. 2018. A genome-wide association study for fat-related traits computed by image analysis in Japanese Black cattle. *Anim Sci J.* 89:743-751.

Nedergaard J, Golozoubova V, Matthias A, Asadi A, Jacobsson A, Cannon B. 2001. UCP1: the only protein able to mediate adaptive non-shivering thermogenesis and metabolic inefficiency. *Biochim Biophys Acta.* 1504: 82–106.

Otto PI, Guimarães SEF, Verardo LL, Azevedo ALS, Vandenplas J, Soares ACC, Sevillano CA, Veroneze R, de Fatima A Pires M, de Freitas C, et al. 2018. Genome-wide association studies for tick resistance in Bos taurus\* Bos indicus crossbred cattle: A deeper look into this intricate mechanism. *J Dairy Sci.* 101:11020-11032.

Parkin J, Cohen B. 2001. An overview of the immune system. *Lancet.* 357: 1777-1789.
Parsons PA. 2005. Environments and evolution: interactions between stress, resource inadequacy and energetic efficiency. Biol Rev. 80:89–610.

Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM. et al. 2012. The yak genome and adaptation to life at high altitude. Nat Genet. 44: 946–949.

Riley JK, Carayannopoulos MO, Wyman AH, Chi M and Moley KH. 2006. Phosphatidylinositol 3-kinase activity is critical for glucose metabolism and embryo survival in murine blastocysts. The Journal of biological chemistry. 281: 6010-6019.

Roosen S, Exner K, Paul S, Schroder JM, Kalm E, Looft C. 2004. Bovine beta-defensins: identification and characterization of novel bovine beta-defensin genes and their expression in mammary gland tissue. Mammalian genome. 15: 834-842.

Sasago N, Abe T, Sakuma H, Kojima T, Uemoto Y. 2017. Genome-wide association study for carcass traits, fatty acid composition, chemical composition, sugar, and the effects of related candidate genes in Japanese Black cattle. Anim Sci J. 88:33-44.

Sharif S, Mallard BA, Wilkie BN, Sargeant JM, Scott HM, Dekkers JC, Leslie KE. 1998. Associations of the bovine major histocompatibility complex DRB3 (BoLA-DRB3) alleles with occurrence of disease and milk somatic cell score in Canadian dairy cattle. Anim. Genet. 29: 185-193.

Simonson TS, Yang Y, Huff CD, Yun H, Qin G, Witherspoon DJ, Bai Z, Lorenzo FR, Xing J, Jorde LB. et al. 2010. Genetic evidence for high-altitude adaptation in Tibet. Science. 329:72-74.

Smith JL, Wilson ML, Nilson SM, Rowan TN, Oldeschulte DL, Schnabel RD, Decker JE, Seabury CM. 2019. Genome-wide association and genotype by environment interactions for growth traits in U.S. Gelbvieh cattle. BMC Genomics. 20: 926-940.

Stear MJ, Dimmock CK, Newman MJ, Nicholas FW. 1988. BoLA antigens are associated with increased frequency of persistent lymphocytosis.
in bovine leukaemia virus infected cattle and with increased incidence of antibodies to bovine leukaemia virus. *Anim Genet.* 19: 151-158.

Takeda M, Uemoto Y, Inoue K, Ogino A, Nozaki T, Kurogi K, Yasumori T, Satoh M. 2019. Genome-wide association study and genomic evaluation of feed efficiency traits in Japanese Black cattle using single-step genomic best linear unbiased prediction method. *Anim Sci J.* 00:1–11.

Vegiopoulos A, Mueller-Decker K, Strzoda D, Schmitt I, Chichelnitskiy E, Ostertag A, Berriel Diaz M, Rozman J, Hrabe de Angelis M, Nuelsing RM, et al. 2010. Cyclooxygenase-2 controls energy homeostasis in mice by de novo recruitment of brown adipocytes. *Science.* 328: 1158–1161.

Wu H, Guang X, Al-Fageeh MB, Cao J, Pan S, Zhou H, Zhang L, Abutarboush MH, Xing Y, Xie Z, et al. 2014. Camelid genomes reveal evolution and adaptation to desert environments. *Nat Commun.* 5:5188.

Wu T. 2001. The Qinghai-Tibetan plateau: how high do Tibetans live? *High Alt. Med. Bio.* 2:489-499.

Xia X, Qu K, Zhang G, Jia Y, Ma Z, Zhao X, Huang Y, Chen H, Huang B, Lei C. 2019. Comprehensive analysis of the mitochondrial DNA diversity in Chinese cattle. *Anim Genet.* 50:70-73.

Yang J, Li WR, Lv FH, He SG, Tian SL, Peng WF, Sun YW, Zhao YX, Tu XL, Zhang M, Xie XL, Wang YT, Li JQ, Liu YG, Shen ZQ, Wang F, Liu GJ, Lu HF, Kantanen J, Han JL, Li MH, Liu MJ. 2016. Whole-Genome Sequencing of Native Sheep Provides Insights into Rapid Adaptations to Extreme Environments. *Mol Biol Evol.* 10: 2576-2592.
**Figure legends**

**FIG 1** The phylogenetic tree of MP mammals and their divergence.

(A) The phylogenetic tree of Mongolian Cattle (MC), Mongolian sheep (MS), Mongolian horse (MH), sheep, cattle, horse, mouse and human. Mongolia mammals and corresponding domestic ones were diverged about 2-3 million years ago (Mya, MC-cattle: 1.8 Mya, MS-sheep: 2.6 Mya and MH-horse: 2.7Mya).

(B) PSMC analysis of MC, MS, MH, cattle, sheep and horse. PSMC analysis revealed MC, MS, MH diverged from their corresponding domestic ones about 3 Mya.

**FIG 2** Structure variation analysis of MC genome.

(A) SNPs (Single Nucleotide Polymorphisms) detection; (B) Indel detection;

(C) Inversions detection; (D) CNVs (Copy Number Variations).

Using cattle genome as reference sequences, we totally detected ~7.62 million single nucleotide variations and short InDels, 749 inversion variants and 4689 CNVs (block length >5kb, 3750 deletions and 939 duplication). In addition, 32,944 structural variations, including 19931 deletions, 4343 insertions and 8670 inversions were also detected.

**FIG 3** Segmental duplications (SDs) analysis and GO enrichment.

(A) Segmental duplications (SDs) were detected in MC cattle genome and GO enriched in terms related to adaptation to extreme environments.

(B) SDs enriched in a GO term called olfactory receptor activity.
Using cattle genome as reference sequences, 251 specific SDs were identified in MC genome and enriched in GO terms related to adaptation to extreme environments and GO term called olfactory receptor activity.

FIG 4 Differentially Expressed Genes (DEGs) enriched in PPAR signaling pathway.
(A) DEGs enriched in PPAR signaling pathway.
(B) RPKM analysis of DEGs enriched in PPAR signaling pathway.

FIG 5 DEGs enriched in Oxidative phosphorylation pathway and PI3K-Akt signaling pathway.
(A) DEGs enriched in Oxidative phosphorylation pathway.
(B) DEGs enriched in PI3K-Akt signaling pathway.
(C) RPKM analysis of 3 up-regulated genes related to PI3K-Akt signaling pathway.

FIG 6 DEGs enriched in Complement and coagulation cascades pathway.
(A) DEGs enriched in Complement pathway.
(B) DEGs enriched in coagulation cascades pathway.
(C) RPKM analysis of DEGs enriched in Complement and coagulation cascades pathway.
Figures.

Figure 1
Figure 3
Figure 4
Figure 5
Figure 6