DOI : 10.5958/2277-8934.2020.00010.7 *Vol 27 No 1, p 69-75*

SNPs AND InDels DETECTION AND SELECTION SIGNALS IDENTIFICATION OF ALXA BACTRIAN CAMEL BY WHOLE-GENOME RESEQUENCING

Fenghong Wang[1,2], Huar Bao[1], BAYIN Jiragara[1], Bin Yang[1], Odongerle[1], Serguleng XU[1], Narenbatu[2*] and Demtu Er[1*]

[1]College of Veterinary Medicine, [2]College of Animal Science, Inner Mongolia Agricultural University, Hohhot, Key Laboratory of Clinical Diagnosis and Treatment Technology in Animal Disease, Ministry of Agriculture and Rural Affairs of P.R. China, Hohhot, Inner Mongolia Autonomous Region P.R. China, 010018

ABSTRACT

Alxa Bactrian camel has been raised as an economic and transportation animal in Alxa, Inner Mongolia, China. In a long period of evolution and artificial selection, it has formed many specific physiological characteristics that enable the production of beneficial items such as milk, meat and fur despite in the harsh environment. At the present the Bactrian camel genome is still a draft genome assembled only at scaffold level but the genetic diversity in multiple re-sequenced Bactrian camels has not been investigated yet. In this study, we evaluated the genetic features of SNPs and InDels and conducted selection signals identification to detect specific genes associated with adaptation to the environment and muscle growth of Alxa Bactrian camel by whole-genome resequencing. By next-generation sequencing technology, a total of 367.98 Gb high-quality paired-end reads were generated. On an average, 81.65% of the reference genome sequence was covered with mapping depth of 10-fold. From these data, 6,759,073 SNPs and 976,715 InDels were identified and 15,037 genes were detected, which revealed wide genetic variations and complex genetic features related to adaptation mechanism of Alxa Bactrian camel in harsh environments. By selective signals analysis, we identified 111 genomic regions, including 70 candidate genes such as BCO1, AKR1D1, SVOPL, SMS, PHEX, PCYT1B, POLA1, and MEGF10, which are potentially involved in environment adaptability and muscle growth. Generally, these results provide a framework and comprehensive insights for further genetic studies in the Bactrian camel population and research on genes, which would provide a better understanding of economically important traits and environment adaptability in Bactrian camel and further provide a scientific basis for the selective breeding of Alxa Bactrian camel.

Key words: Alxa Bactrian camel, InDels, selective signals, SNPs, whole-genome resequencing

At present, there are mainly 7 varieties of Bactrian camel in the world, distributed in the cold and desert regions, and 4 varieties are in China. They are called Sunite Bactrian camel, Qinghai Bactrian camel, Xinjiang Bactrian camel and Alxa Bactrian camel according to their breeding regions. China is the largest and most widely distributed country of Bactrian camels in the world, according to the statistics of 2016, there are about 360,000 Bactrian camels in China and 160,000 (44.92%) Bactrian camel are in Alxa, Inner Mongolia. Alxa is located in northwest China, geographically contains the vast Gobi (such as Wuliji, Yinggen) and deserts (such as Badain Jaran Desert, Tengger Desert). Through a long period of evolution and artificial selection, the appearances of the Alxa Bactrian camel living in the Gobi and the desert areas have formed their own characteristics and Alxa Bactrian camel is divided into Gobi camel and desert camel by its geographical distribution. Gobi camel is characterised with two big and erect humps, long and flat wide back waist and brown color. Desert camel is characterised with short trunk and long limb, big and round abdomen and apricot colour. At present, studies on Alxa Bactrian camel are mainly focused on its milk, meat and fat (Moshaverinia and Moghaddas, 2013). Up to date, there is no systematic study on the genome and a series of genome changes in a long period of evolution and artificial selection are still largely unknown.

Since the first whole-genome assembly of the human genome is completed in 2001, the whole-genome sequencing of mammals has been quickly completed (Venter *et al,* 2001). In recent decades, the next-generation sequencing technology has

SEND REPRINT REQUEST TO NARENBATU ● DEMTU ER email: nmgtrnbt@163.com ● eedmt@imau.edu.cn

significantly facilitated the genetic studies of complex traits in domestic animals (Rosenthal *et al,* 2015) and has revealed natural and artificial selection footprint in many species, such as pig (Li *et al,* 2014, Ai *et al,* 2015), sheep (Yang *et al,* 2016), cattle (Sequencing and Consortium 2009) and dog (Xiao *et al,* 2014). Bactrian camel was present in the third millennium BC and subsequently spread into Central Asia (Peters and Driesch, 1997) and recent studies have focused on the divergence time between the ancestors of Bactrian camel and cattle (Sequencing and Consortium, 2012) genomic Single Nucleotide Polymorphisms (SNPs) rates, (Burger and Palmieri, 2014) the divergence time and demographic history of Bactrian camel, dromedary, alpaca and evidence of camel adaptation to desert environments (Wu *et al,* 2014) protective effect of camel milk on pathogenicity induced by *E. coli.* (Soliman *et al,* 2015). However, the whole-genome resequencing of Bactrian camel had not been utilised in the respective studies. Therefore, here we studied the whole-genome resequencing of Alxa Bactrian camel to examine the genetic variations, obtain genetic features, predict important candidate genes associated with muscle growth and development and understand the adaptation mechanism of Alxa Bactrian camel in harsh environments.

Materials and Methods

In this study, 12 Bactrian camels including 6 Gobi camels and 6 desert camels were randomly chosen to take skeletal muscle sample. Firstly, we extracted DNA from skeletal muscle, and detected the purity, potency and volume of DNA. Secondly, we created DNA library and sequenced the DNA using Illumina HiSeq platform and then filtered data and generated clean reads. Finally, we called Single Nucleotide Polymorphisms (SNPs) and short insertions and deletions (InDels), and then conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis using online software.

Sample collection

Alxa Bactrian camel was sampled from Yingen Gobi and Badain Jaran Desert, where these were raised freely on local pastures. With the assistance of local herdsmen 12 Bactrian camels 8-10 years old including 6 Gobi camels and 6 desert camels, were randomly chosen at slaughter house to take skeletal muscle samples, which were snapped frozen in liquid nitrogen and stored at −80 °C. All experimental procedures were approved by the Animal Care and Use Committee of the Inner Mongolia Agricultural University, and conducted in strict accordance with the animal welfare and ethics guidelines.

Data filtering and generation of clean reads

DNA was extracted from the skeletal muscle samples with the AXYGEN Blood and Tissue Extraction Kit (Corning, USA) according to the manufacturer's instructions. Sequencing was performed on the Illumina HiSeq 2000 platform.

Data filtering and clean reads generation

First, to avoid reads with artificial bias (i.e. low quality paired reads, which mainly resulted from base-calling duplicates and adapter contamination), we removed the following types of reads: (1) Reads of joint contamination (the number of base contaminated by joints of reads > 5 bp; for paired-end sequencing, remove reads if one end is contaminated with the connector); (2) Low quality reads (the base of the mass value Q ≤ 5 in reads accounts for ≥50% of the total base; for paired-end sequencing, if one end is low-quality reads, reads from both ends); (3) Reads with N ratio > 5% (for paired-end sequencing, if one end contains N > 5%, remove the reads at both ends).

SNPs and InDels calling

Clean reads were aligned to the recently released version of the reference Bactrian genome using Burrows-Wheelser Aligner v0.7.10-r789 with default settings (Li and Durbin, 2009). The algorithm MEM was used to find the suffix array coordinates of good matches for each read. Sequence Alignment/Map tools (SAMtools)(Li *et al,* 2009) was used to convert file format from SAM to BAM and to filter the unmapped and non-unique reads. Read pairs with top mapping quality were retained. Local realignment around short InDels was performed with duplication-removed reads using Realigner Target Creator and InDels Realigner in the Genome Analysis Toolkit (GATK, version 3.3-0-g37228af) (McKenna *et al,* 2010; Depristo *et al,* 2011). SNPs and InDels were separated with GATK tool 'Select Variants' and subjected to rigorous processing to exclude false positives.

Functional enrichment analysis

We performed functional enrichment analysis within the GO and KEGG pathway terms using the Database for Annotation Visualisation and Integrated (DAVID) tool (http://david.abcc.ncifcrf.gov/) for enrichment analysis of the significant overrepresentation of GO biological processes (GO-BP), molecular function (GO-MF) terminologies, and KEGG-pathway. Only the enriched GO terms with

raw p-value <0.05 were used for further interpretation in this study. The functional relationships of the genes of interest were used in the Pathway studio program

Results and Discussion

Whole-genome sequencing, SNPs and InDels detection

General SNPs and InDels detection

Whole-genome sequence data of 12 Bactrian camels were obtained from Hiseq-2000 platform (Illumina, San Diego, CA, USA). Approximately 2,453,332,756 reads were produced (a total of 367.98 Gb paired-end sequence). Read mapping to the reference sequence was performed using BWA. About 2,003,009,380 high-quality paired-end sequences were successfully mapped to the Bactrian camel reference genome (https://www.ncbi.nlm.nih.gov/genome/10741 genome_assembly_id = 212083). On an average, 81.65% of the reference genome sequence was covered with mapping depth of 10-fold (Fig 1). These data yielded 7,735,788 variations from 12 Alxa Bactrian camels including 6,759,073 SNPs and 976,715 InDels using GATK and SAMtools. Of the identified SNPs, 4,220,597 (28.04%) were located in intergenic regions, 4,255,696 (28.27%) were located in introns, 992,658 (6.59%) and 992,658 (6.59%) and 1,007,762 (6.69%) were located in upstream and downstream, respectively. Of the identified InDels, 3,980,120 were located in intergenic, 1,902,612 were located in introns, 3,602,124 were located in upstream and downstream, respectively (Table 1).

Table 1. Summary and annotation of SNPs of Alxa Bactrian camel.

Type	Number	Percentage
UPSTREAM	992,658	6.59%
DOWNSTREAM	1,007,762	6.69%
EXON	121,018	0.80%
INTRON	4,255,699	28.27%
INTERGENIC	4,220,597	28.04%
TRANSCRIPT	4,387,197	29.14%
UTR_3_PRIME	43,408	0.29%
UTR_5_PRIME	14,532	0.10%

Specific SNPs detection

In total, 6,478,201 variations were identified from Gobi camel, including 5,770,567 SNPs and 707,633 short InDels. The number of specific SNPs was 1,494,099 in Gobi camel. In total, 5,800,250 variations were identified from desert camel, including 5,143,589 SNPs and 656,611 short InDels. The number of specific SNPs was 867,120 in desert camel. Compared to desert camel (867,120), Gobi camel (1,494,099) has more specific SNPs, indicating weaker intensive selection of desert camel (Fig 2).

Annotation and enrichment analysis

Annotation and enrichment analysis of general SNPs

In our Alxa Bactrian camel data set, 6,759,073 SNPs were detected in a total of 15,037 genes. Of the 15,037 genes, many of them matched to genes which were potentially associated with camel's system of adaptation to the environment, such as energy and fat metabolism (DGKZ), adaptation of respiratory system (FOXP3, CX3CR1, CYSLTR2 and SEMA4A), adaptation of visual system (OPN1SW, CX3CR1 and CNTFR), salt metabolism (NR3C2 and IRS1), water reservation (AQP1), osmoregulation (NFAT5 and BGTI) (Wu *et al,* 2014). As it is well known, the soil of Alxa is saline-alkali soil, therefore Bactrian camel living in Alxa has acquired specific abilities of salt tolerance. It is reported that camel can tolerate a high dietary intake of salt, and the salt tolerance level is eight times more than cattle and sheep, and the blood glucose level is twice more than other ruminants (Alali *et al,* 1988; Ali, 1994). Salt metabolism plays an important role in salt tolerance and water balance in Bactrian camel. NR3C2 and IRS1 genes play critical roles in sodium re-absorption and water balance in kidney (Sun *et al,* 1991). NFAT5 gene is the only known tonicity regulated transcription factor in mammals (Cheung and Ko, 2013). Also, Alxa is a very cold and dry region and MAPK4, NOX4, IFNGR2, SLC2A4, and PDK1 genes show significant correlations with climate variation (Lv *et al,* 2014; Yang *et al,* 2016). Additionally, we also found that ADCY4, CACNA2D1, AGT and PTGER genes are associated with adaptation of cold and dry environment. For example, ADCY4 and AC stimulated cAMP genes are involved in cAMP induced cell proliferation in cultured adrenal cells and are the key mediators of Na and water transport (Al-Hakim, 2004; Strait *et al,* 2010). These useful findings explains that Alxa Bactrian camel is well-adapted to the harsh environment.

Annotation and enrichment analysis of specific SNPs

We identified 1,494,099 and 867,120 specific SNPs, 14346 and 13132 genes of Alxa Gobi and desert camel, respectively. By analysing the annotated genes in two species of camel, 12,292 genes were

found in two species of camel at the same time. So, there are 2054 specific genes in Gobi camel and 840 specific genes in desert camel (Fig 3). In Gobi camel, Gene Ontology (GO) terms were associated with 2054 genes (Harris *et al,* 2004; Su and Zhou, 2007). GO analysis of these genes revealed enrichment in 14 GO terms in the biological processes, 14 GO terms in the cellular components and 21 GO terms in the molecular functions. KEGG enrichment analysis of these genes identified 10 pathways. The analysis showed that the genes associated with cellular components, such as nucleus extracellular exosome, cytoplasm, nucleoplasm and cytosol, were significantly enriched in Gobi camel. In desert camel, GO terms are associated with 840 genes. GO analysis of these genes revealed enrichment in 14 GO terms in the biological processes, 6 GO terms in the cellular components and 19 GO terms in the molecular functions. KEGG enrichment analysis of these genes identified 5 pathways. The analysis showed that the genes associated with cellular components, such as metal ion binding, nucleic acid binding, and protein cysteine S-palmitoyl transferase activity were significantly enriched in desert camel. These studies results suggest that the phenotypes associated with these genes may represent specific characteristics of Gobi camel and desert camel.

Selective sweep signals analysis

In order to detect genome selection signals related to evolution and meat production in Alxa Bactrian camel, our study used the sliding window method to find the highly selected genomic regions. By using 20Kb sliding window and 10Kb step to scan the whole genome of Alxa Bactrian camel to find the selected genomic signals in the process of evolution. We identified a total of 111 genomic regions under selective sweep signals analysis containing 70 candidate genes that are associated with Alxa Bactrian camel traits (Fig 4).

The annotation analysis found that these genomic regions contain 70 genes and GO annotation of these genes revealed enrichment in the biological process terms, such as cell development, cell differentiation, single-organism cellular process, single-organism developmental process, myofibrillar assembly, system development, and muscle structure development. KEGG enrichment analysis of these candidate genes identified 1 significant metabolic pathway (Fig 5). Our study results are consistent with Jirimutu *et al* (2012). They reported that these changes may underline the insulin resistance typically (Sequencing and Consortium, 2012). Candidate genes play an important role in regulation of skeletal muscle and those candidate genes associated with the growth and development of skeletal muscle were identified in several genomic regions under selective sweep signals analysis, including SMS, PHEX (NW-011517570.1), PCYT1B, POLA1, and MEGF10 (NW-011515253.1). Previous study showed that MEGF10 plays an important role in muscle stem cells and can regulate the development of skeletal muscle (Park *et al,* 2014). Because these cells are connected with the surface of muscle fibre, which is also called satellite cell, which normally are not active. Once the muscle muscle fibre is damaged the satellite cells are activated, differenced, proliferated and regenerate the muscle fibre fusion to repair the damaged muscle. PHEX also is an important

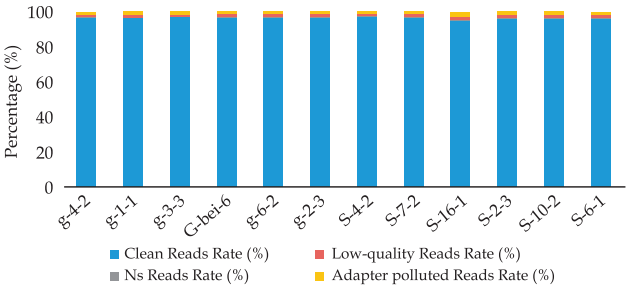

Fig 1. Raw reads statistics of Alxa Bactrian camel by whole-genome resequencing

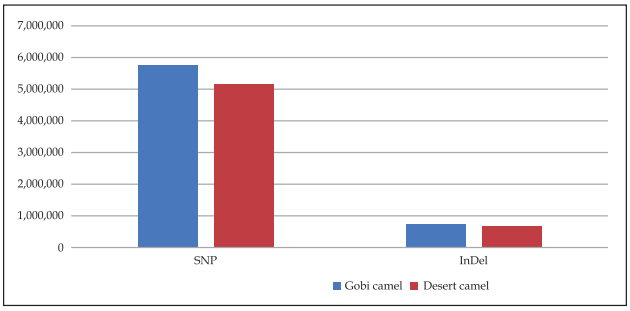

Fig 2. SNPs and InDels in Alxa Gobi and desert camel.

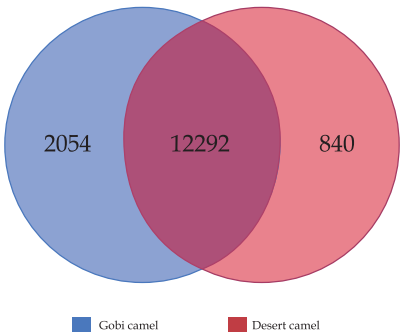

Fig 3. Common and specific genes numbers in Alxa Gobi and desert camel.

gene, which can regulate phosphate homeostasis and skeletal mineralisation (Quarless, 2003) and EZH1 can promote skeletal growth (Lui *et al,* 2016). We also found other candidate genes associated with skeletal muscle development, for example BCO1 (The enzyme β-carotene oxygenase 1) can catalyse the breakdown of provitamin A, which can be converted to vitamin A. Vitamin A plays an important role in vision (Saari, 2012), epithelium maintenance (Takahashi *et al,* 1993; Kumar *et al,* 2017), immune competence (Duriancik *et al,* 2010), reproduction, embryonic growth and development (Clagettdame and Knutson, 2011) and its neuroprotective nature (Ramani *et al,* 2017) is essential for the life of all animals. So, the BCO1 function is irreplaceable and could be efficient to regulate myogenesis and satellite cell activity *in vivo* (Praud *et al,* 2017). We also detected some genes associated with immune system, such as SASH1

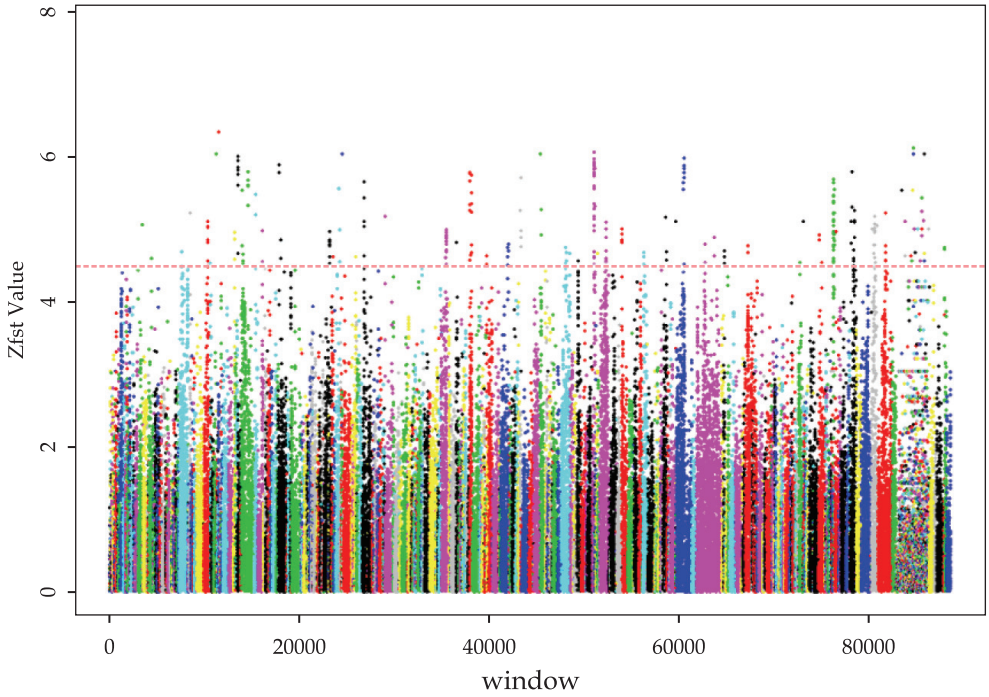

Fig 4. Selective sweep signals of Alxa Bactrian camel.

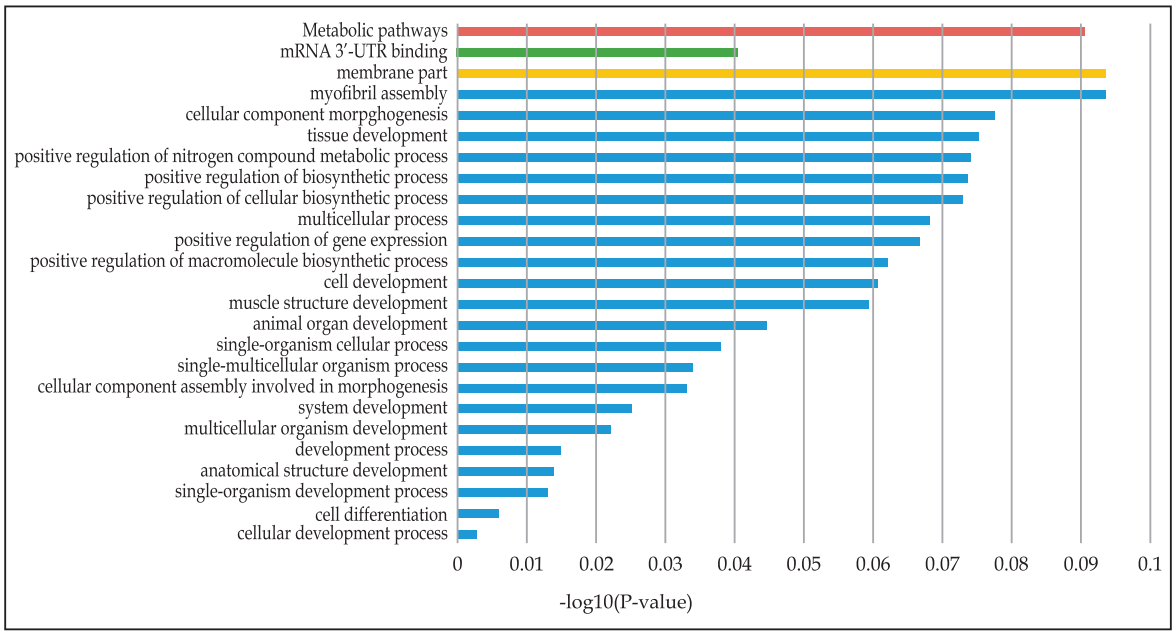

Fig 5. GO and KEGG pathways enrichment analysis of candidate genes selection of Alxa Bactrian camel.

and MS4A genes that play important roles in the development and progression of various diseases by regulating the cell growth, proliferation and apoptosis. (Liang *et al,* 2001, Lin *et al,* 2012) AKR1D1 gene can be repressed in diabetic patients (Valanejad *et al,* 2017). SVOPL gene is one of the SLC22 (solute carrier family 22) members. It is expressed in kidney and liver and regulates the uptake and excretion of environment toxins (Jacobsson *et al,* 2007). These results may explain why Alxa Bactrian camel has special functions such as high drought resistance, crude feed tolerance and adaptation to cold and dry environment compared with other animals.

Conclusion

This study is the first to explore the whole genome resequencing of Alxa Bactrian camel. We found a large number of SNPs and InDels by whole-genome resequencing between Gobi camel and desert camel. We annotated 14,346 and 13,132 genes by specific SNPs of Alxa Gobi and desert camel. These specific SNPs contained genes, such as NR3C2, IRS1 AQP1, FOXP3, CX3CR1 and CYSLTR2 are potentially involved in camel's system of adaptation to the environment. By selective signals analysis, we identified 111 genomic regions associated with environment adaptability and muscle growth. These selected genomic regions contained genes, such as BCO1, AKR1D1, SVOPL, SMS, PHEX, PCYT1B, POLA1, and MEGF10. It means these genes expressed positive selection trend in a long period of evolution and artificial selection of Alxa Bactrian camel. The detected SNPs and InDels are enough to show genetic variations between Alxa Gobi and desert camel.

Acknowledgements

The project (No. 201502069) was supported by Inner Mongolia Science & Technology Plan.

References

Ai H, Fang X, Yang B, Huang Z, Chen H, Mao L, Zhang F, Zhang L, Cui L, He W, Yang J, Yao X, Zhou L, Han L, Li J, Sun S, Xie X, Lai B, Su Y, Lu Y, Yang H, Huang T, Deng W, Nielsen R, Ren J and Huang L (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. Nature Genetics 47(3):217-225.

Al-Hakim A, Rui X, Tsao J, Albert PR and Schimmer BP (2004). Forskolin-resistant Y1 adrenal cell mutants are deficient in adenylyl cyclase type 4. Molecular & Cellular Endocrinology 214(1-2):155-165.

Alali AK, Husayni HA and Power DM (1988). A comprehensive biochemical analysis of the blood of the camel (*Camelus dromedarius*). Comparative Biochemistry and Physiology Part B Comparative Biochemistry 89(1):35-37.

Ali TM (1994). A manual for the Primary Animal Health Care Worker.

Burger PA and Palmieri N (2014). Editor's choice: Estimating the Population Mutation Rate from a de novo Assembled Bactrian Camel Genome and Cross-Species Comparison with Dromedary ESTs. Journal of Heredity 105(6): 839-846.

Cheung Chris YK and Ko Ben Cb (2013). NFAT5 in cellular adaptation to hypertonic stress – regulations and functional significance. Journal of Molecular Signaling 8(1):5.

Clagett-Dame M and Knutson D (2011). Vitamin A in Reproduction and Development. Nutrients 3(4):385-428.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D and Daly MJ (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43(5):491-498.

Duriancik DM, Lackey DE and Hoag KA (2010). Vitamin A as a regulator of antigen presenting cells. Journal of Nutrition 140(8):1395-1399.

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B and Mungall C (2004). The Gene Ontology (GO) database and informatics resource. Nucleic Acids Research 32:258-261.

Jacobsson JA, Haitina T, Lindblom J and Fredriksson R (2007). Identification of six putative human transporters with structural similarity to the drug transporter SLC22 family. Genomics 90(5):95-609.

Jirimutu, Wang Z, Ding G, Chen G, Sun Y, Sun Z, Zhang H, Wang L, Hasi S, Zhang Y, Li J, Shi Y, Xu Z, He C, Yu S, Li S, Zhang W, Batmunkh M, Ts B, Narenbatu, Unierhu, Bat-Ireedui S, Gao H, Baysgalan B, Li Q, Jia Z, Turigenbayila, Subudenggerile, Narenmanduhu, Wang Z, Wang J, Pan L, Chen Y, Ganerdene Y, Dabxilt, Erdemt, Altansha, Altansukh, Liu T, Cao M, Aruuntsever, Bayart, Hosblig, He F, Zha-ti A, Zheng G, Qiu F, Sun Z, Zhao L, Zhao W, Liu B, Li C, Chen Y, Tang X, Guo C, Liu W, Ming L, Temuulen, Cui A, Li Y, Gao J, Li J, Wurentaodi, Niu S, Sun T, Zhai Z, Zhang M, Chen C, Baldan T, Bayaer T, Li Y and Meng H (2012). Genome sequences of wild and domestic bactrian camels. Nature Communications 3:1202.

Kumar S, Dollé P, Ghyselinck NB and Duester G (2017). Endogenous retinoic acid signaling is required for maintenance and regeneration of cornea. Experimental Eye Research 154:190-195.

Li H and Durbin R (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25(14-15):1754-1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G and Durbin R (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078-2079.

Li M, Tian S, Yeung CKL, Meng X, Tang Q, Niu L, Wang X, Jin L, Ma J and Long K (2014). Whole-genome sequencing

of Berkshire (European native pig) provides insights into its origin and domestication. Scientific Reports 4(4):4678.

Liang Y, Buckley TR, Tu L, Langdon SD and Tedder TF (2001). Structural organization of the human MS4A gene cluster on Chromosome 11q12. Immunogenetics 53(5):357-368.

Lin S, Zhang J, Xu J, Wang H, Sang Q, Xing Q and He L (2012). Effects of SASH1 on melanoma cell proliferation and apoptosis *in vitro*. Molecular Medicine Reports 6(6):1243-1248.

Lui J, Garrison P, Nguyen Q, Ad M, Keembiyehetty C, Chen W, Jee Y, Landman E, Nilsson O, Barnes K and Baron J (2016). EZH1 and EZH2 promote skeletal growth by repressing inhibitors of chondrocyte proliferation and hypertrophy. Nature Communications 7:13685.

Lv FH, Agha S, Kantanen J, Colli L, Stucki S, Kijas JW, Joost S, Li MH and Ajmone MP (2014). Adaptations to climate-mediated selective pressures in sheep. Molecular Biology and Evolution 31(12):3324-3343.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo M (2010). The Genome Analysis Toolkit: a MapReduce framework for analysing next-generation DNA sequencing data. Genome Research 20(9):1297-1303.

Moshaverinia A and Moghaddas E (2013). Prevalence of tick infestation in dromedary camels (*Camelus dromedarius*) brought for slaughter in Mashhad abattoir, Iran. Journal of Parasitic Diseases 39(3).

Park SY, Yun Y, Kim MJ and Kim IS (2014). Myogenin is a positive regulator of MEGF10 expression in skeletal muscle. Biochemical and Biophysical Research Communications 450(4):1631-1637.

Peters J and Driesch AVD (1997). The two-humped camel (*Camelus bactrianus*): new light on its distribution, management and medical treatment in the past. Journal of Zoology 242(4):651-679.

Praud C, Al Ahmadieh S, Voldoire E, Le Vern Y, Godet E, Couroussé N, Graulet B, Le Bihan Duval E, Berri C and Duclos MJ (2017). Beta-carotene preferentially regulates chicken myoblast proliferation withdrawal and differentiation commitment via BCO1 activity and retinoic acid production. Experimental Cell Research 358(2):140-146.

Quarles LD (2003). FGF23, PHEX, and MEPE regulation of phosphate homeostasis and skeletal mineralisation. American Journal of Physiology Endocrinology and Metabolism 285(1):1-9.

Ramani M, Van GT, Kadish I, Ambalavanan Nand Mcmahon LL (2017). Vitamin A and Retinoic acid combination attenuates neonatal hyperoxia-induced neurobehavioral impairment in adult mice. neurobisology of learning and memory 141:209-216.

Rosenthal E, Blue E and Jarvik GP (2015). Next-generation gene discovery for variants of large impact on lipid traits. Current Opinion in Lipidology 26(2):114-119.

Saari JC (2012). Vitamin A metabolism in rod and cone visual cycles. Annual Review of Nutrition 32(1):125.

Sequencing TBCG and A Consortium (The Bactrian Camels Genome Sequencing and Analysis Consortium) (2012). Corrigendum: Genome sequences of wild and domestic Bactrian camels. Nature Communications 3:1202.

Sequencing TBG and A Consortium (2009). The genome sequence of taurine cattle: A window to ruminant biology and evolution. Science 324(5926):522-528.

Soliman MM, Hassan MY, Mostafa SA, Ali HA and Saleh OM (2015). Protective effects of camel milk against pathogenicity induced by *Escherichia coli* and *Staphylococcus aureus* in Wistar rats. Molecular Medicine Reports 12(6):8306-12.

Strait KA, Stricklett PK, Chapman M and Kohan DE (2010). Characterisation of vasopressin-responsive collecting duct adenylyl cyclases in the mouse. American Journal of Physiology-Renal Physiology 298(2):F859-867.

Su Z and Zhou X (2007). EasyGO: Gene Ontology-based annotation and functional enrichment analysis tool for agronomical species. BMC Genomics 8(1):246.

Sun XJ, Rothenberg P, Kahn CR, Backer JM, Araki E, Wilden PA, Cahill DA, Goldstein BJ and White MF (1991). Structure of the insulin receptor substrate IRS-1 defines a unique signal transduction protein. Nature 352(6330):73-77.

Takahashi Y, Miura T and Takahashi K (1993). Vitamin A is involved in maintenance of epithelial cells on the bronchioles and cells in the alveoli of rats. Journal of Nutrition 123(4):634-641.

Valanejad L, Ghareeb M, Shiffka S, Nadolny C, Chen Y, Guo L, Verma R, You S, Akhlaghi F and Deng R (2017). Dysregulation of Δ4-3-oxosteroid 5β-reductase in diabetic patients: Implications and mechanisms. Molecular and Cellular Endocrinology 470:127-141.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA and Holt RA *et al* (2001). The sequence of the human genome. Science 291(5507):1304-1351.

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

Xiao Gou, Wang Z, Li N, Qiu F, Xu Z, Yan D, Yang S, Jia J, Kong X, Wei Z, Lu S, Lian L, Wu C, Wang X, Li G, Ma T, Jiang Q, Zhao X, Yang J, Liu B, Wei D, Li H, Yang J, Yan Y, Zhao G, Dong X, Li M, Deng W, Leng J, Wei C, Wang C, Mao H, Zhang H, Ding G, and Li Y (2014). Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. Genome Research 24(8):1308-1315.

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. Molecular Biology and Evolution 33(10):2576-2592.