Kafkas Univ Vet Fak Derg
24 (6): 835-844, 2018
DOI: 10.9775/kvfd.2018.20022

Kafkas Universitesi Veteriner Fakultesi Dergisi
Journal Home-Page: http://vetdergi.kafkas.edu.tr
Online Submission: http://submit.vetdergikafkas.org

Research Article

Whole Genome Sequencing of the Dzo: Genetic Implications for High Altitude Adaptation, Sterility, and Milk and Meat Production

Jingcheng ZHONG [1,2,†] Zhijie MA [3,†] Zhixin CHAI [1,2] Hui WANG [1,2]
Chengfu ZHANG [4] Qiumei JI [4] Jinwei XIN [4]

[†] *These authors contributed equally to this work*
[1] Institute of Qinghai-Tibetan Plateau, Southwest Minzu University, Chengdu, CHINA
[2] Key Laboratory of Qinghai-Tibetan Plateau Animal Genetic Resource Reservation and Utilization, Ministry of Education, Southwest Minzu University, Chengdu, CHINA
[3] Qinghai Academy of Animal Science and Veterinary Medicine, Qinghai University, Xining, CHINA
[4] Institute of Animal Science and Veterinary, Tibet Academy of Agricultural and Animal Husbandry Sciences, Lhasa, CHINA

Article Code: KVFD-2018-20022 Received: 29.04.2018 Accepted: 09.08.2018 Published Online: 10.08.2018

How to Cite This Article

Zhong J, Ma Z, Chai Z, Wang H, Zhang C, Ji Q, Xin J: Whole genome sequencing of the Dzo: Genetic implications for high altitude adaptation, sterility, and milk and meat production. *Kafkas Univ Vet Fak Derg,* 24 (6): 835-844, 2018. DOI: 10.9775/kvfd.2018.20022

Abstract

The Dzo, a hybrid between the domestic Yak and cattle found on the Qinghai-Tibetan plateau are commercially important owing to increased meat and milk production, as well as their adaptation to extremely high-altitude environments. To better understand the genomic architecture and adaptive capabilities of this unique domesticated hybrid, we performed whole genome resequencing and compared the genomic architecture and variation of the Dzo to its progenitors: the cattle and Yak. In total, 33.17 M single nucleotide variations (SNVs) were detected between the Dzo and cattle reference genomes. Even though the Dzo is known to be sterile, no genetic signatures associated with sterility were found on the Dzo Y chromosome. On the contrary, our results suggest that the autosomal *DMC1* locus (Chromosome 5. 110729098. C>CT) plays a role in the sterility of Dzo, which warrants further exploration of its functions. We integrated the whole genome resequencing data of cattle and Yak to obtain candidate genes with a high degree of variation that might be associated with altitudinal adaptation. We found that the *EPAS1* gene, which encodes hypoxia-inducible factor 2α, exhibited significant variation (Chromosome 11.28664187 C>T) between Yak and cattle, and may play a key role in the genetic basis of altitudinal adaptation. In addition, in analyzing differences between the Dzo, Yak, and cattle genomes, we uncovered several additional genomic signatures relevant to high altitude adaptation and meat and milk production. These findings underscore the need for further studies to improve ruminant stock for sustainable agriculture on the Qinghai-Tibet plateau..

Keywords: *Dzo: Genome sequencing, Dzo, Altitude adaptation, Sterility, Milk and meat production*

Dzo Sığırlarında Tüm Genom Sekanslaması: Yüksek Rakıma Adaptasyon, Sterilite İle Süt ve Et Verimine Etkisi

Öz

Evcil Yak ile Qinghai-Tibet platosunda bulunan sığırın bir hibriti olan Dzo artmış et ve süt üretimi ile birlikte aynı zamanda oldukça yüksek rakımlı bölgelere adaptasyonu nedeni ile ticari olarak önem arz etmektedir. Bu özgün evcil hibritin genomik yapısını ve uyum kapasitelerini daha iyi anlamak amacıyla tüm genom sekanslaması yapılarak genomik yapısı ve varyasyonlar Dzonun progenitörleri olan sığır ve Yak ile karşılaştırıldı. Dzo ile sığır referans genomları arasında toplam 33.17 M tek nükleotid varyasyonu belirlendi. Dzo steril olarak bilinmesine rağmen Dzo Y kromozomunda sterilite ile ilgili hiç bir genetik işarete rastlanmadı. Aksine, elde edilen bulgular otozomal *DMC1* lokusunun (Kromozom 5. 110729098. C>CT) Dzonun sterilitesinde rol oynadığına işaret etmekte ve bu nedenle de daha ileri araştırılmaların yapılması gerekmektedir. Yüksek rakıma adaptasyon ile ilişkili yüksek oranda varyasyona sahip genlerin araştırılması amacıyla sığır ve Yakın tüm genom sekanslama verilerinin entegrasyonu gerçekleştirildi. Hipoksi ile indüklenebilir faktör 2α'yı kodlayan *EPAS1* geni Yak ve sığır arasında yüksek varyasyon (Kromozom 11.28664187 C>T) göstermekteydi ve bu nedenle yüksek rakıma adaptasyonun genetik temelinde rol oynayabileceği düşünüldü. Ayrıca, Dzo, Yak ve sığır genomları arasındaki farklılıklar analiz edildiğinde, yüksek rakıma adaptasyon ile birlikte et ve süt üretimine etki eden bazı ilave genomik işaretlere rastlandı. Bu bulgular Qinghai-Tibet platosunda daha ileri ve devam ettirilebilir ruminant yetiştiriciliği için daha fazla çalışmalara ihtiyaç olduğunu göstermektedir.

Anahtar sözcükler: *Dzo: Genom sekanslama, Dzo, Yüksek rakım adaptasyonu, Sterilite, Süt ve et üretimi*

İletişim (Correspondence)
+86 21 85522285 Fax: +86 21 85522285
zhongjincheng518@126.com

INTRODUCTION

The Yak *(Bos grunniens)* inhabits the Qinghai-Tibetan plateau of China and the adjacent highland regions of Central and East Asia, at altitudes of 2000-5500 m. These animals possess numerous anatomical features and physiological traits that facilitate life at high altitudes, including enlarged lungs and hearts, and the lack of hypoxic pulmonary vasoconstriction. In addition, these animals are efficient foragers in the extreme environment that characterizes livestock production in this challenging landscape. Currently, more than 14 million domestic Yak provide meat, milk, transportation, dung for fuel, and hides for Tibetans and other nomadic pastoralists living in Central Asia, while approximately 15.000 wild Yak also inhabit this region. In addition, the Yak of the Qinghai-Tibetan plateau exhibits high degree of genetic and phenotypic diversity, where this diversification has driven their development as a genetic resource, including the formation of various locally-adapted phenotypes and hybrid forms. In regions of mixed pastoralism and agriculture at lower altitudes in Central Asia, the deliberate hybridization of Yak and cattle is now widespread [1], where the Dzo *(Bos taurus x Bos grunniens)*-the first-generation male hybrid of male cattle and female Yak-has become an economically important resource. The Dzo adapts well to local environmental conditions and retains important production characteristics from both parental species: the ability to adapt to a harsh environment from the Yak and increased productivity from cattle. As a result, the Dzo exhibits higher productivity than the Yak in terms of both milk and meat yields [2,3], yet possesses better physiological adaptation to high altitude environments than cattle.

However, although these F_1 interspecies hybrids exhibit clear production advantages, male hybrids are sterile, and thus phenotypic quality and genetic variation in the Dzo is currently a result of parental genotypic combinations and their segregation patterns alone [4]. In recent years, although a number of studies have been carried out to investigate the sterility of the Dzo-such as studies involving breeding, cytogenetics, tissue morphology, endocrinology, biochemistry, and molecular biology-the underlying mechanisms of male sterility remain unclear, despite the potential importance of solving this problem for future agricultural development. Furthermore, understanding the genetic architecture of cattle, Yak, and Dzo from a genomic standpoint could provide important insights into understanding the quantitative trait loci involved in the survival and performance of hybrids, as well as the interactions between genes and pathways that allow hybrids to maintain high levels of productivity in the face of multiple environmental challenges [5,6].

In the present study, we sequenced the genomes of both the Dzo and wild Yak and compared these sequences to previously published domestic Yak genome data and bovine single nucleotide polymorphism (SNP) data from the 1.000 bull genomes project [7,8]. Thus, the objectives of this investigation were to 1) improve our understanding of the mechanisms of adaptation to the plateau environment in the Yak, and 2) to analyze how loci associated with adaptation to this extreme environment are configured in the Dzo, with the goal of elucidating both the basis of trait differences (milk and meat production) in cattle, Yak, and Dzo, and mechanisms of male sterility in the Dzo. As a result, by characterizing the genetic architecture of the Dzo, we have highlighted genetic differences between these regionally important livestock and have begun to explore roles of these differences in local adaptation, and their importance for genetic resource management.

MATERIAL and METHODS

Ethics Statement: Methods of animal care and use were approved by the Institution of Animal Care and Use Committee of Southwest Minzu University and Qinghai University.

Samples, Sequencing, and Library Construction: All Dzo and Yak used in the present study were selected at the age of three from a slaughterhouse in Tongde County (100°63′E; 35°24′N) and the Datong Yak breeding farm (101°67′E; 36°92′N), both in Qinghai province, China. After injecting Dzo with 2×10^{-3} mL xylazine hydrochloride per kg body weight, Dzo were slaughtered and pooled Dzo (PI) samples of testicular tissue from six F1 individuals were obtained. Wild Yak (YA) blood was taken from the jugular veins of six males, and DNA subsequently isolated using the Aidlab Genomic DNA Extraction Kit (Aidlab Biotechnologies Co., Ltd, Beijing, China). For the purposes of the present study, Yak was considered to be wild based on the following taxonomic characters: long hair and large skeleton. Genomic DNA was extracted (TaKaRa, Dalian, China) for library construction, and Illumina paired-end sequencing libraries (insert sizes of 500 bp) were constructed according to the manufacturer's instructions (Illumina, San Diego, California, USA). Sequencing was performed on an Illumina HiSeq 2000 (carried out at BGI, Shenzhen, China). In addition, sequence data from three wild Yak (W1, W2, and W3 from Hoh Xil National Nature Reserve) and three domestic Yaks (D1, D2, and D3, collected from Gansu, Sichuan, and Tibet, respectively) were downloaded from the *Bos grunniens* Genome sequencing BioProject (NCBI Accession No.: PRJNA217895) for comparative analyses.

Read Filtering and Alignment: Raw sequencing reads from a total of eight samples (six Yaks from NCBI and two pooled samples from the present study) were filtered on the basis of chastity score, and trimmed on the basis of quality score, using Fast QC (v0.10.1) and Trimmomatic (v0.32) [9,10] in four steps: 1) removal of adapters; 2) removal of bases from the start or end of a read, if the quality threshold fell below 3; 3) scanning the read with a 4-base pair sliding

window, and removal when the average Phred quality per base fell below 15; and 4) removal of reads shorter than 50 bases or average Phred quality below 20 for the whole read. Clean reads were mapped to the *B. taurus* reference genome (assembly UMD3.1) [11] using BWA (0.7.10-r789) [12] using the BWA-MEM algorithm. Detailed parameters were as follows: java-jar trimmomatic-0.32.jar PE ILLUMINACLIP: TruSeq3-PE. fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW: 4:15 MINLEN:50./bwa mem -t 4 -k 32 -M.

Variant Identification and Annotation: After sequences were aligned, files were converted to BAM format using SAMtools (v1.0) [13]. BAM files were sorted, and duplicate reads were filtered using Picard. Single nucleotide variations (SNVs) and the detection of insertions or deletions (INDELs) were performed using the Genome Analysis Toolkit (GATK, version 2.4-9) [14-16]. SNVs data were downloaded from the 1000 bull genomes project [7] (http://www.1000bullgenomes.com/doco/all_snps_annotated_2013_04_25.tab.gz). Structural variants (SVs) with a confidence score ≥ 40 (Phred quality score) in the YA and PI genomes were detected using Breakdancer (v1.1) [17]. The annotation of SNVs, INDELs, and SVs were conducted using an in-house perl script. Sequence alignments were constructed between proteins that were altered by missense SNVs and proteins encoded by orthologous genes and were scored using the SIFT algorithm to predict the functional impact of protein substitutions [18]. A Y chromosome from cattle *(B. taurus)* was downloaded from NCBI (GenBank No.: CM001061) as a reference sequence. Clean reads were mapped to the Y chromosome reference sequence using BWA (0.7.10-r789) with a BWA-MEM algorithm. The identification of variation and steps for annotation were performed as previously mentioned.

Population Structure: We constructed a phenetic tree using all SNVs with genetic distance matrices between individuals. A neighbor-joining algorithm was used in TreeBeST (https://github.com/Ensembl/treebest) with uncorrected p-distances and bootstrapping (1.000). MEGA5 [19] was used to visualize the tree, and FRAPPE (v1.1) [20] was utilized to infer population structure and ancestry. This analysis was based on 33 M SNVs, and no prior information regarding ancestry was assumed. We ran 10.000 iterations on a pre-defined number of clusters (K; 2-7).

RESULTS

We obtained a total of 230.43 Gb of data from the pooled genome sequences of eight samples, including the Dzo (PI) and wild Yak (YA) collected for this study. The PI and YA samples were sequenced to a depth of approximately 24X and 26X respectively, while the other six Yaks were sequenced to a depth of 6X. In total, 33.17 M single nucleotide variations (SNVs) were identified between the genome sequences of the eight samples and the cattle reference genome (UMD3.1), while 238,188 SNVs were found in coding sequence (CDS) regions, where 99.409 non-synonymous SNVs were annotated. In addition, 13.04 M SNVs were observed in the seven Yak genome sequences, and 30.95 M SNVs were found when comparing the seven Yak genomes to the bovine reference genome. When comparing the Yak data to the 1.000 bull SNP data (26.72 M), we detected 7.58 M SNV sites, where 221.42 K sites exhibited different allelic types (2.92%). As expected, a very high percentage (91.23%) of the 27.50 M SNVs were heterozygous in the Dzo when compared to the bovine reference genome, while 27.63% of the 27.61M SNVs were heterozygous between the YA and bovine reference genome. Furthermore, 3.57 M insertions and deletions (INDELs) were detected between the bovine reference genome and the eight samples (PI, YA, and the six Yak sequences from NCBI [W1, W2, and W3; and D1, D2, and D3]), 7.323 of which were in CDS regions. In total, 58.514 structural variants (SV) were detected in comparison PI and the reference genome, while 73.221 SVs were identified between the YA and the reference genome.

As shown in *Fig. 1A,* the neighbor-joining tree demonstrated that the Dzo (PI) is positioned between the cattle and the Yak. For Bayesian clustering, when K = 2, the cattle and Yak were identified as separate clusters *(Fig 1B);* however, when K = 3, the domestic and wild Yak cluster was divided into two groups. Interestingly, the wild Yak (YA) sample in this study featured approximately 24% of the domestic Yak genomic proportions, although the observed phenotypes of domestic Yak could not be distinguished from those of typical wild Yak. For the three wild Yak sequenced from the Hoh Xil National Nature Reserve, no genomic signals of the domestic Yak could be detected. The finding that there

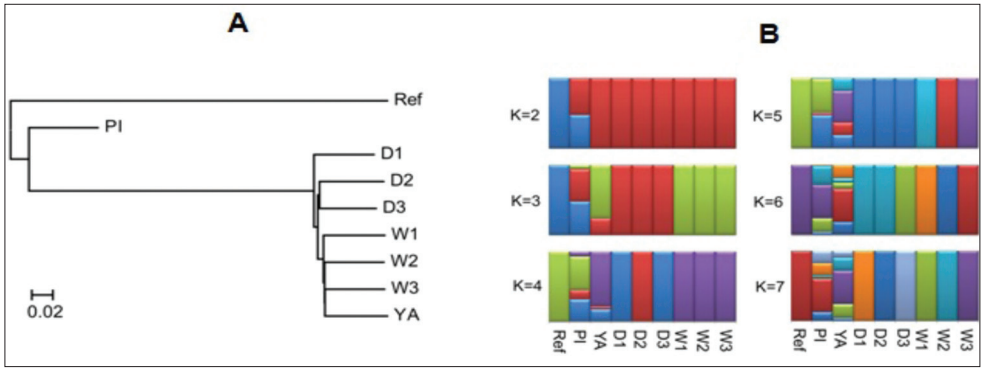

Fig 1. Neighbor-joining tree and population structure for Yak and Dzo. (A) The neighbor-joining tree for Yak and Dzo; (B)The population structures for Yak and Dzo. Ref: Cattle; PI: Dzo; YA: wild yak; W1 to W3 and D1 to D3: The three wild Yaks and domestic Yaks which sequence data were downloaded from *Bos grunniens* Genome Sequencing Bio-project, respectively

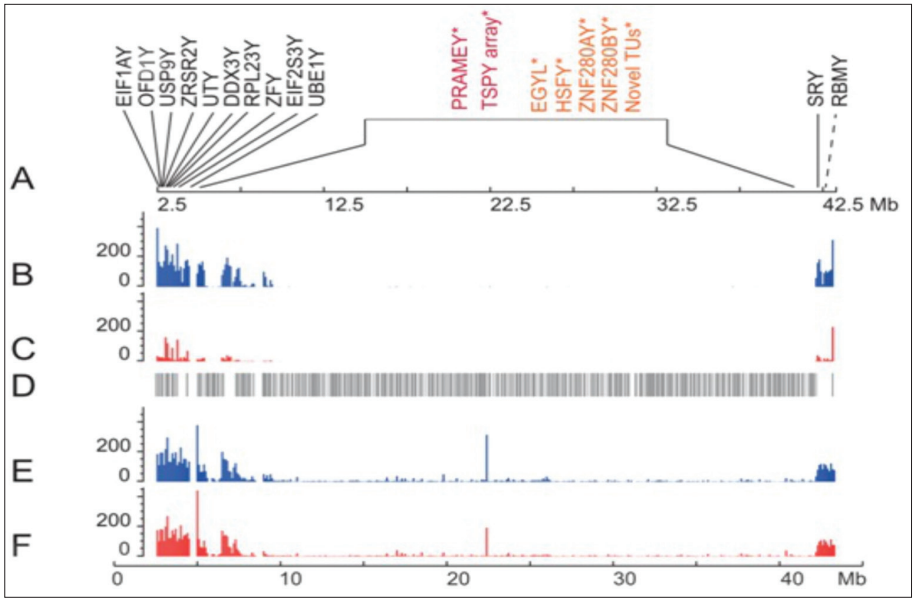

Fig 2. Genome landscape of the bovine male-specific region of the Y chromosome (bMSY). (A) Gene map of bMSY. X-degenerate single-copy genes or transcripts (black) are clustered at either end of the bMSY, whereas multicopy genes/transcripts (red and orange) are present in the majority of the bMSYs. RBMY was missing from the draft assembly. The relative position of RBMY (dashed line) was determined based on a RH-mapping analysis. (B) The SNV distribution in YA chromosome. (C) The SNV distribution in the PI chromosome. (D) A gene map of bMSY. Gray lines indicate the gene region distributed on bMSY. (E) The read depth of YA. (F) The read depth of PI. Window size was set to 100 Kb. The scale on the bottom of the ideogram is based on the bovine Y chromosome assembly

Table 1. *Alignment results of nine X-degradation genes of Y chromosomes between PI and cattle*

Gene ID	Gene Start	Gene End	Gene Covered Base	Gene Depth	Exon Depth
EIF1AY	2678121	2664668	13450	46.56	11.35
OFD1Y	2760114	2824003	63890	313.92	N/A
USP9Y	2907549	3033182	125615	209.99	15.38
ZRSR2Y	3112384	3172379	59968	543.71	18.91
UTY	3224590	3409324	184678	233.05	11.86
DDX3Y	3458510	3468728	10219	33.19	17.31
ZFY	3727723	3748502	20780	125.14	23.55
EIF2S3Y	3815112	3862568	47444	235.17	32.32
SRY	42225210	42225899	690	11.04	11.04

N/A represents no exon information for OFD1Y in NCBI

was a substantial degree of overlap between the genomic sequences of the supposedly wild Yak (YA) sequenced in the present study and the accessed domestic Yak sequences indicates that the wild Yak used in the present study were likely wild/domestic Yak hybrids.

Although the bovine genome project included the Y chromosome, resequencing data for the Y chromosome has largely been ignored owing to its abundance of repetitive and palindromic sequences. In this study, we mapped the sequence data from PI and YA to the cattle Y chromosome (GenBank No.: CM001061). Based on this alignment, we inferred that domestic Yak D1 and D2 are females, which was not elucidated in the original study [8] owing to the lack of alignments involving the key sex determination (*SRY*) gene used in determining the sex of most mammals. The alignment performed in the present study also highlighted the abundance of repetitive sequences on the Y chromosomes, especially at the telomeres *(Fig. 2),* where at the ends of the Y chromosome, genes mainly originate via X chromosomal degradation [20]. The *SPY* gene is important in sex determination, while the highly repetitive sequences around these single copy genes may provide protection against degradation [20,21]. The average sequencing depth of the nine single copy genes analyzed in PI and YA in the present study (Eukaryotic translation initiation factor 1A Y-linked *(EIF1AY)* gene, the Y-linked oral-facial-digital, type 1 *(OFD1Y)* gene, the ubiquitin-specific protease 9, Y chromosome *(USP9Y)* gene, the RNA-binding motif and serine/arginine rich 2 *(ZRSR2Y)* gene, the ubiquitously transcribed tetratricopeptide repeat gene on Y chromosome *(UTY) gene*, the DEAD box helicase 3 Y-linked *(DDX3Y)* gene, zinc-finger protein gene on Y-chromosome *(ZFY)* gene, the eukaryotic translation initiation factor 2, submit 3 and structural gene Y-linked *(EIF2S3Y)* gene and *SRY)*, including introns and intergenic regions, were 180X and 197X, respectively *(Table 1* and *Table 2)*. However, it is important to note that except for STY, some sequences are likely to have been erroneously attributed to these genes owing to the abundant repetitive elements in the introns, which could not be filtered. In contrast, the sequencing depths for the exonic regions in PI and YA were 17X and 20X, respectively. Here, we detected 3.219 SNVs between the Dzo (PI) and cattle Y chromosome,

Table 2. *Alignment results of nine X-degradation genes of Y chromosomes between YA and cattle*

Gene ID	Gene Start	Gene End	Gene Covered Base	Gene Depth	Exon Depth
EIF1AY	2678121	2664668	13358	51.26	15.69
OFD1Y	2760114	2824003	63650	340.97	N/A
USP9Y	2907549	3033182	125601	225.72	19.74
ZRSR2Y	3112384	3172379	59995	608.52	19.41
UTY	3224590	3409324	184735	251.36	16.19
DDX3Y	3458510	3468728	10219	36.78	17.09
ZFY	3727723	3748502	20780	134.22	22.56
EIF2S3Y	3815112	3862568	47438	262.80	43.73
SRY	42225210	42225899	690	8.36	8.36

N/A represents no exon information for OFD1Y in NCBI

Table 3. *SNV annotation results of Chromosome Y of PI*

Chr ID	Loci	SNV	Amino Acids	Syn/Non	Gene_ID	Protein_Id
ChrY	3171104	G<->A	R<->H	1	100306950	XP_003584415.1
ChrY	3171086	G<->C	S<->T	1	100306950	XP_003584415.1
ChrY	4394284	G<->A	A<->T	1	100849661	XP_003584416.1
ChrY	4394211	A<->T	E<->V	1	100849661	XP_003584416.1
ChrY	4394169	A<->G	Y<->C	1	100849661	XP_003584416.1
ChrY	3728235	G<->T	V<->L	1	280962	NP_803457.1
ChrY	3506348	A<->G	M<->V	1	790278	XP_001256796.3
ChrY	2544329	G<->A	A<->A	0	100849399	XP_003584412.1
ChrY	5114940	A<->G	T<->T	0	100849792	XP_003584418.1
ChrY	3641828	A<->G	P<->P	0	100849362	XP_003584411.1
ChrY	3641999	C<->T	Y<->Y	0	100849362	XP_003584411.1
ChrY	3642115	C<->T	T<->I	1	100849362	XP_003584411.1
ChrY	3860224	C<->T	Y<->Y	0	100271755	XP_003584417.1
ChrY	3860206	T<->C	S<->S	0	100271755	XP_003584417.1

where only eight SNVs were non-synonymous *(Table 3)*. For wild Yak (YA), 14.630 SNVs were detected, 58 of which were non-synonymous.

We analyzed 52 autosomal genes known to be related to meiotic processes in order to explore male sterility in the Dzo (PI). For example, the positive regulatory domain zinc finger protein 9, encoded by the *PRDM9* gene, is a major determinant of meiotic recombination in humans and mice [22], where variation within this gene strongly influences recombination hotspot activity and meiotic instability in humans [23]. Previous studies have shown that mRNA levels of *PRDM9* are much lower in the testes of sexually immature Yak calves and sterile male cattle-Yaks in comparison to those of normal adult Yaks, suggesting that *PRDM9* might be associated with male fertility in the Dzo. In total, 513 SNVs were detected in these 52 genes, 251 of which were non-synonymous, and 44 of which were found in the SNV set from the 1.000 bull genome project. Of the non-synonymous SNVs, 97 SNVs were fixed, were observed to be different between the Yak and cattle sequences and were found to be heterozygous in the Dzo genome. For the *PRDM9* gene, we found 38 SNVs, where 26 of these were non-synonymous. Interestingly, we also found an INDEL (Chromosome 5.110729098. C>CT) located in the CDS region of the meiotic recombinase *(DMC1)* gene, and seven non-synonymous SNVs in the replication protein A *(RPA)* gene family *(Table 4,* we only present the non-synonymous also in the 1.000 bull project SNV set). The annotated SNVs for these genes may be a starting point for further detailed research into male sterility in the Dzo.

Four genes related to altitudinal adaptation *(EPAS1*, *EGLN1*, *HYOU1*, and *HMBS)* were examined to explore adaptive differences between Yak, Dzo, and cattle. These genes have previously been identified in Tibetan people, and have been shown to help humans adapt to high altitude conditions [24-27], hypoxia, and myocardial infarction [28]. Previous studies of these genes in local domestic livestock were performed either solely within cattle [29] or Yak. In total, 44 SNVs were found in the CDS region of these four genes, including 16 non-synonymous SNVs. Interestingly, only two of these non-synonymous SNVs were in the *EPAS1* gene from the 1.000 bull project SNV set. Moreover, while 14 of the non-synonymous SNVs in the aforementioned genes could be distinguished between cattle (reference

Table 4. *SNP annotation of male-sterile*

ChrID	Location	Pop_SNP	Gene Name	Codon	CDS_ID	Protein_ID	1000BullDB
Chr1	45021322	A G G G R G G G G	PRDM9	AGA<->GGA;	cds276	XP_002683648.2	A/G
Chr1	45028357	T Y Y Y Y Y Y Y Y	PRDM9	TTT<->TCT;	cds276	XP_002683648.2	T/C
Chr1	45028408	C S S S S S S S S	PRDM9	TCC<->TGC;	cds276	XP_002683648.2	C/G
Chr1	45028440	G R R R R R R R R	PRDM9	GGA<->AGA;	cds276	XP_002683648.2	G/A
Chr1	45028459	G R R R R R R R R	PRDM9	GGA<->GAA;	cds276	XP_002683648.2	G/A
Chr1	45029328	T K K K K K K K K	PRDM9	CTC<->CGC;	cds276	XP_002683648.2	T/G
Chr1	45029346	A G G G G G G G G	PRDM9	GAA<->GGA;	cds276	XP_002683648.2	A/G
Chr1	45029661	G A A A R A A R A	PRDM9	AGC<->AAC;	cds276	XP_002683648.2	G/A
Chr1	45033708	G G S G S G S S S	PRDM9	GCA<->CCA;	cds276	XP_002683648.2	G/C
Chr1	45033759	C Y Y Y Y Y Y Y Y	PRDM9	CGA<->TGA;	cds276	XP_002683648.2	C/T
Chr19	23508655	A G G G R G G G G	RPA1	ACC<->GCC;	cds33180	NP_001068644.1	A/G

Table 5. *SNP annotation of EPAS1, EGLN1, HYOU1 and HMBS*

ChrID	Location	Ref D1 D3 D2 PI W1 W2 W3 YA	Gene Name	Codon	Non/Syn	CDS_ID	Protein_ID	1000BullDB
Chr11	28641695	A G G G G G G G G	EPAS1	ACA<->ACG	0	cds20127	NP_777150.1	A/G
Chr11	28650973	C G G G G G G G G	EPAS1	CAA<->GAA	1	cds20127	NP_777150.1	C/G
Chr11	28659125	T C C C Y C C C C	EPAS1	TTT<->TTC	0	cds20127	NP_777150.1	T/C
Chr11	28659176	G G G G R G G G G	EPAS1	GCG<->GCA	0	cds20127	NP_777150.1	/
Chr11	28660443	A G G G R G G G G	EPAS1	ACA<->ACG	0	cds20127	NP_777150.1	A/G
Chr11	28662841	T C C C Y - - C C	EPAS1	CTC<->CCC	1	cds20127	NP_777150.1	T/C
Chr15	30201004	C T T T Y T T T T	HMBS	CCC<->CCT	0	cds25736	NP_001039672.1	C/T
Chr15	30201004	C T T T Y T T T T	HMBS	CCC<->CCT	0	cds25737	XP_005216001.1	C/T

sequence) and Yak, there were no polymorphisms detected within the Yak population. It is notable that the only variant (Chromosome11.28664187C>T) between Yak and cattle is predicted to be deleterious using SIFT and was found to be located in the *EPAS1* gene *(Table 5,* we only present the SNP annotation also in the 1.000 bull project SNV set).

We examined 52 genes related to production traits, including growth, milk, and meat characteristics. Of these genes, 537 SNVs were found in CDS regions, 189 of which were non-synonymous, and 57 of which were found in the 1.000 bull data set. Hormone-sensitive lipase *(LIPE)*, which is considered to be an important candidate gene associated with meat traits, is also involved in free fatty acid mobilization [30]. Here, we found six non-synonymous SNVs in the *LIPE* gene that exhibited fixed differences between Yak and the cattle reference sequence. Of these, three were found in the 1.000 bull project data set (Chromosome18. 51226081G>A; Chromosome18. 51226221A>G; Chromosome18. 51226425A>G), and three were not found in the 1.000 bull SNV set (Chromosome18. 51221047C>T; Chromosome18. 51222967G>A; Chromosome18. 51226270C>A). For fatty acid binding protein *(FABP)* genes (a family of transport proteins for fatty acids and other lipophilic substances, such as eicosanoids and retinoid [31,32], 24 non-synonymous SNVs were found in the CDS region, 11 of which were fixed between Yak and the bovine reference sequence *(Table 6,* we only present the non-synonymous also in the 1.000 bull project SNV set). These fixed SNV differences between Yak and cattle populations could be associated with observed meat quality differences, and thus warrant further investigation. Interestingly, in four genes related to growth and milk production *(GH1*, *GHR*, *PRL*, and *PRLR),* we did not detect any non-synonymous SNVs in *GH1* and *PRL* but did find eight non-synonymous SNVs in *GHR* and *PRL* between cattle and Yak *(Table 6*). These results indicate that the receptor genes for growth hormone and prolactin secretion may play an important role in differences in growth and milk production between the congeners.

DISCUSSION

Although the draft genome sequence of the Yak has been available for four years, and resequencing studies have since been undertaken [6,33,34], genetic variation in the Yak-most notably structural variation and genetic comparisons with other bovid species-Remains poorly understood. Moreover, the genome sequence of the Dzo has not been reported until now. Although the Dzo is the F1 hybrid between cattle and Yak, and the SNPs found in the Dzo indeed derive from that simple cross, it unclear which parental gene combinations produce viable Dzo, or how this variation segregates across individuals. Thus, not only did our study focus on the Dzo, but on the relationships

Table 6. *SNP annotation of milk-meat-growth*

ChrID	Location	Pop_SNP	Gene Name	Codon	CDS_ID	Protein_ID	1000BullDB
Chr6	7105326	G C C C S C C C C	FABP2	CAA<->GAA	cds11066	XP_005207643.1	G/C
Chr6	7105350	T Y Y Y Y C Y Y Y	FABP2	AAA<->GAA	cds11066	XP_005207643.1	T/C
Chr6	7105359	A C C C C C C C C	FABP2	TAA<->GAA	cds11066	XP_005207643.1	A/C
Chr6	7105759	A R A A A A A A A	FABP2	ATA<->ACA	cds11066	XP_005207643.1	A/G
Chr6	7107541	T Y C S S G S S S	FABP2	ATC<->CTC;ATC<->GTC	cds11066	XP_005207643.1	T/G
Chr6	7108785	A W W W W A A W W	FABP2	ATG<->AAG	cds11066	XP_005207643.1	A/T
Chr6	7108842	A R A R R A R R R	FABP2	TTT<->TCT	cds11066	XP_005207643.1	A/G
Chr14	46817895	G C C C C C C C C	FABP9	CTG<->GTG	cds25027	NP_001179339.1	G/C
Chr14	46817903	C T T T Y T T T T	FABP9	TGT<->TAT	cds25027	NP_001179339.1	C/T
Chr14	46819115	G G G G S G G G G	FABP9	GAC<->GAG	cds25027	NP_001179339.1	G/C
Chr14	46835065	T C C C C C C C C	FABP4	ATC<->GTC	cds25028	NP_776739.1	T/C
Chr18	51226081	G A A A R A A A A	LIPE	CGG<->CAG	cds31387	NP_001073689.1	G/A
Chr18	51226081	G A A A R A A A A	LIPE	CGG<->CAG	cds31386	XP_005219231.1	G/A
Chr18	51226221	A G G G R G G G G	LIPE	ATC<->GTC	cds31387	NP_001073689.1	A/G
Chr18	51226221	A G G G R G G G G	LIPE	ATC<->GTC	cds31386	XP_005219231.1	A/G
Chr18	51226425	A G G G R G G G G	LIPE	AAC<->GAC	cds31387	NP_001073689.1	A/G
Chr18	51226425	A G G G R G G G G	LIPE	AAC<->GAC	cds31386	XP_005219231.1	A/G
Chr20	31891025	G A A A A A A A A	GHR	ACC<->ATC	cds35638	NP_788781.1	G/A
Chr20	31891050	T Y T T Y T T T T	GHR	AGC<->GGC	cds35638	NP_788781.1	T/C
Chr20	31891130	T G T T K T T T T	GHR	AAC<->ACC	cds35638	NP_788781.1	T/G
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35722	NP_001034815.1	T/C
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35723	XP_005221632.1	T/C
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35724	XP_005221633.1	T/C
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35725	XP_005221634.1	T/C
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35726	XP_005221635.1	T/C
Chr20	39115345	T C C C Y C C C C	PRLR	AGT<->AAC	cds35727	NP_776580.1	T/C

between the three species, by comparing SNPs in cattle, Yak, and Dzo genomic sequences. In detecting genome-wide variation in this study, very large numbers of SNVs (>30 M) were uncovered between the Yak genomes and the *Bos taurus* reference, as well as between the genomes of eight individuals (Yak and Dzo) and the *Bos taurus* reference, implying that a very large number of variants are segregated within the Yak and between Yak, cattle, and Dzo. When comparing the Dzo to the cattle reference genome, >90% of SNVs were found to be heterozygous in the genome of the F1 Dzo samples. Homo- or heterozygosity in the Dzo genome likely exerts strong effects on gene expression and functionality in this hybrid, which could play an important role in the maintenance of traits of physiological and economic relevance, including adaptations to high altitude. However, this variation may also come at a cost via the incompatibility between alleles [35], a phenomenon that can only be further investigated using single individual sequences that can be phased and subjected to linkage and haplotype analyses.

Previously, Qi et al.[36] assessed the impact of cattle admixture on domestic Yak and concluded that admixture between Yak and cattle has impacted the contemporary genetic makeup of the domestic Yak. Recently, Medugorac et al.[1] inferred bovine haplotypes in the genomes of Mongolian Yaks using whole-genome sequencing and showed that these introgressed regions are enriched for genes involved in nervous system development and function (gene ontogenies often associated with domestication), which supports the idea that introgressive hybridization could have facilitated Yak management and breeding. However, until now, no studies have assessed the genetic implications of domestication on wild Yak. In our study, we detected the introgression of domestic Yak in the wild Yak (YA) genome for the first time. Although the phenotypes of the sampled individuals were superficially indistinguishable from wild Yaks, approximately 24% of their genome was inferred as having a domestic Yak origin, leading us to infer second-generation introgression. This may be related to the unique method of grazing in pastoral areas of the Qinghai-Tibetan plateau, where herders often manage domestic Yak adjacent to wild Yak populations. We were able to detect this introgression by showing that population differentiation can be easily and precisely identified between wild and domestic Yak at the genome level, potentially shedding light on the mechanism and spread of Yak domestication and its contemporary impact

on wild Yak populations-a process that clearly needs to be monitored as it potentially threatens their genetic integrity. Based on our results, it is clear that further range-wide studies of domestic/wild Yak introgression are necessary, and that hybridization between wild Yak and domestic Yak should be tightly controlled to maintain the genetic integrity of wild and domestic Yak populations, in addition, the mechanism of male sterility in the Dzo is one of the most pressing management questions in Yak science, and yet it remains unresolved. In combining genomic information from Y chromosomes and autosomal sequence from cattle, Dzo, and Yak, we were able to explore this problem by examining variations across the Y chromosome and in 52 autosomal genes known to be related to meiotic processes. Intriguingly, however, we only detected eight non-synonymous SNVs in coding regions of the Y chromosome between the Dzo and cattle, while 58 were found to distinguish Yak and cattle. As previously described in species (including cattle) [37,38], we found that both ends of the Y chromosomes in Yak and Dzo comprise a large number of repeat sequences, mainly distributed in intergenic regions and introns. These repetitive sequences have played a key role in the formation of the Y chromosome [20,21,37,38] by preventing recombination, thereby protecting important single copy genes at both ends [20,21,39]. However, we could not find strong evidence for the determination of Dzo sterility in genes along the Y chromosome based on our SNV results, suggesting that autosomal genes maybe hold the key to Dzo male sterility. Thus, discrepancies in the Y chromosome between cattle, Yak, and Dzo need to be further explored to elucidate the substantial difference between the expected (approximately 30 non-synonymous substitutions between Dzo and cattle) and observed (eight) non-synonymous SNVs. While this observation suggests that reproductively viable Dzo may require the presence of a strongly cattle-like Y chromosome segregating in the parental population to survive, variations in the Y-chromosomes of domestic Yak populations need to be further characterized.

The meiotic recombination protein DMC1 plays a central role in homologous recombination in meiosis by assembling the sites of programmed DNA double strand breaks and localizing allelic DNA sequences located on homologous chromatids. In the present study, the INDEL (Chromosome 5.110729098. C>CT) discovered in the CDS region of the *DMC1* gene in the Dzo is predicted to have a major impact on gene function. Replication protein A (RPA) not only binds to single-stranded DNA during replication and keeps DNA unwound during replication, it also binds to ssDNA during the initial phase of homologous recombination, including prophase I of meiosis [40]. variation of *PRDM9* also influence male recombination in cattle with variation in *REC8* and *RNF212*, which correlate with genome-wide recombination rates, while variation in *PRDM9* also influence genome-wide hotspot frequency [41]. Here, we detected seven non-synonymous SNVs in the *RPA* gene family, and 26 non-synonymous SNVs in the *PRDM9* gene *(Table 4)*, potentially leading to changes in gene function.

Previous research has indicated that first and second backcross Dzo generations are also sterile, where no spermatogonia were detected in the seminiferous tubules in first generation hybrids and other early backcross generations. By the third generation of backcrossing (with either Yak or cattle), some spermatocytes can be present and the occasional male is found to be fertile, although fertility is not assured until the fourth or fifth generation of backcrossing [4]. Possible causes for this phenomenon have been attributed to X- and Y-chromosome incompatibility. Although we only focused on the Y chromosome here, our results provide informative data for the further study of Dzo male sterility, which should focus on both sex-linked genes and functional studies of the autosomal genes highlighted here.

Linking genomic correlations to environmental parameters is of special importance in Yak science, given the unusual habitat that this species occupies. In humans, a SNV in *EPAS1* has been detected between Tibetan (high-altitude) and Han (low-altitude) populations, and is associated with erythrocyte abundance, and thus supports the role of *EPAS1* in the adaptation to hypoxia necessary for adaptation to high altitudes [24]. Newman et al.[29] found a high degree of association in the oxygen degradation domain of EPAS1 between an *EPAS1 (HIF-2α)* double variant in Angus cattle with high-altitude pulmonary hypertension (HAPH) and demonstrated that the variant appears to be prevalent only in lowland cattle. In Yak, three novel SNPs have also been identified in *EPAS1*, and have been genotyped in three breeds. In the present study, we predicted that one variant (Chromosome11.28664187C>T) in *EPAS1* between Yak and cattle is deleterious. However, we did not identify the same SNPs that have been previously identified in the *EPAS1* gene. This is likely a result of the fact that previous studies have been examining intraspecific variation (within cattle and Yak populations) but have not compared across species and did not specifically investigate interspecific differences in altitudinal adaptation. Overall, in the four genes that are known to be related to altitudinal adaptation and that have been closely examined here (*EPAS1*, *EGLN1*, *HYOU1* and *HMBS*), we found 16 non-synonymous SNVs in CDS regions, and 14 non-synonymous SNVs that distinguished cattle and Yak, suggesting that differences in high plateau adaptation between cattle and Yak could be linked to some of the substitutions detected in these genes. The further investigation of the potential roles of these substitutions will require functional analyses.

For genes related to growth, milk, and meat traits, six non-synonymous SNVs in the *LIPE* gene and 11 non-synonymous SNVs in *FABP* genes possessed fixed differences between Yak and the cattle reference genome. Some of these genes are known to be linked to leanness in

cattle, and therefore phenotypic differences between Yak and cattle may be important in this context. Interestingly, eight non-synonymous SNVs were found in genes for the growth hormone and prolactin receptors *GHR* and *PRLR*, respectively. Previous studies have shown that some homozygous, or compound heterozygous, mutations in *GHR* may induce a partial insensitivity to GH in humans [42]. Prolactin (PRL) is secreted from the anterior pituitary, and plays an extensive role in corpus luteum formation, mammogenesis, and lactogenesis; where the functional activity of PRL is mediated by its receptor (PRLR) in the PRL signal transduction cascade [43]. Moreover, SNVs in PRLR have also been detected in goat, and have been shown to influence milk yield in different goat breeds [44]. Our study indicates that GHR and PRLR may play an important role in differences in growth and milk traits between Yak, Dzo, and cattle.

To better understand genomic differences between cattle (assembly UMD3.1), Yak (GenBank No.: GCA_000298355.1), and Zebu *(B. indicus*; GenBank No.: GCA_000247795.1), we aligned the Zebu and Yak genomes to the cattle reference sequence. The results of this alignment showed that the Yak and Zebu genomic sequences covered 89.99% and 88.76% of the cattle genome, respectively (not present). Thus, a subspecies relationship between cattle and Zebu ancestors that is inferred and likely reflects the quality of the reference sequences [45]. The results of the genome coverage analysis showed large differences in the genomic sequences for Yak, Zebu, and cattle. One potential but important reason for this finding may be the integrity of the genome assembly. Resequencing-based assemblies are limited in establishing differences between bovine genomes alone. In the future, it will be important to complete the genomes for Yak and Zebu. Moreover, it should be noted that our results were based on a pooled sequencing strategy for six wild Yaks and six Dzo individuals. Although the pooling of samples has been successfully used in studies on dog and chicken domestication, the strategy of pooling samples has been shown to affect some statistical analyses, most notably haplotype and linkage analyses [46,47]. Therefore, the sequencing of multiple genomes separately for Yak and Dzo is desirable in the future.

In brief, our results provide basic high-value data for researchers studying Yak, cattle, and Dzo breeding. Comparative analyses of the genomic sequences of Dzo, Yak, and cattle using whole genome sequencing strategies, provide a more in-depth understanding of the genetic variations between Yak, cattle, and Dzo. Here, we have provided evidence of variation in genes related to meiosis and gametogenesis, substantial variation in genes related to high-altitude adaptation, as well as genes associated with meat and milk production between the Yak and cattle genomes. In summary, these genetic variations may be helpful in genotype-phenotype association analyses of Qinghai-Tibet plateau ruminants, most notably for Yak and Dzo, and could also help to shape the management practices of wild and domesticated ruminant populations to preserve the genetic integrity of both populations.

Availability of Supporting Data

All raw reads generated in this work have been deposited in the NCBI database under BioProject accessionPRJNA359997 (alias: SRP095965). The whole genome SNP and INDEL data set has been deposited in the following link: https://pan.baidu.com/s/1bzB1t8.

Competing Financial Interests

The authors have declared that no competing interests exist.

Funding

This study was supported by the Program of National Beef Cattle and Yak Industrial Technology System (CARS-37).

REFERENCES

1. Medugorac I, Graf A, Grohs C, Rothammer S, Zagdsuren Y, Gladyr E, Zinovieva N, Barbieri J, Seichter D, Russ I, Eggen A, Hellenthal G Brem G, Blum H, Krebs S, Capitan A: Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian Yaks. *Nat Genet,* 49 (3): 470-475. 2017. DOI: 10.1038/ng.3775

2. Anand IS, Harris E, Ferrari R, Pearce P, Harris P: Pulmonary haemo-dynamics of the Yak, cattle, and cross breeds at high altitude. *Thorax,* 41 (9): 696-700, 1986. DOI: 10.1136/thx.41.9.696

3. Tumennasan K, Tuya T, Hotta Y, Takase H, Speed RM, Chandley AC: Fertility investigations in the F1 hybrid and backcross progeny of cattle (Bos taurus) and yak *(B. grunniens)* in Mongolia. *Niigata J Healt Welfare*, 78 (1): 69-73, 1997. DOI: 10.1159/000134633

4. Luo XL, Song HF, Guan JQ: Investigation on mechanism of sterility of male hybrids between Yak and cattle. *J Appl Anim Res,* 42 (4): 395-399, 2014. DOI: 10.1080/09712119.2013.875907

5. Cheviron ZA, Brumfield RT: Genomic insights into adaptation to high-altitude environments. *Heredity*. 108, 354-361, 2012. DOI: 10.1038/hdy.2011.85

6. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, Cao C, Hu Q, Kim J, Larkin DM, Auvil L, Capitanu B, Ma J, Lewin HA, Qian X, Lang Y, Zhou R, Wang L, Wang K, Xia J, Liao S, Pan S, Lu X, Hou H, Wang Y, Zang X, Yin Y, Ma H, Zhang J, Wang Z, Zhang Y, Zhang D, Yonezawa T, Hasegawa M, Zhong Y, Liu W, Zhang Y, Huang Z, Zhang S, Long R, Yang H, Wang J, Lenstra JA, Cooper DN, Wu Y, Wang J, Shi P, Wang J, Liu J: The Yak genome and adaptation to life at high altitude. *Nat Genet,* 44, 946-949, 2012. DOI: 10.1038/ng.2343

7. Daetwyler HD, Capitan A, Pausch H, Stothard P, van Binsbergen R, Brøndum RF, Liao X, Djari A, Rodriguez SC, Grohs C, Esquerré D, Bouchez O, Rossignol MN, Klopp C, Rocha D, Fritz S, Eggen A, Bowman PJ, Coote D, Chamberlain AJ, Anderson C, VanTassell CP, Hulsegge I, Goddard ME, Guldbrandtsen B, Lund MS, Veerkamp RF, Boichard DA, Fries R, Hayes BJ:: Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nat Genet,* 46, 858-865, 2014. DOI: 10.1038/ng.3034

8. Wang K, Hu Q, Ma H, Wang L, Yang Y, Luo W, Qiu Q: Genome-wide variation within and between wild and domestic Yak. *Mol Ecol Resour,* 14, 794-801, 2014. DOI: 10.1111/1755-0998.12226

9. Bolger AM, Lohse M, Usadel B: Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114-2120, 2014. DOI: 10.1093/bioinformatics/btu170

10. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, Hanrahan F, Pertea G, Van Tassell CP, Sonstegard TS, Marçais G, Roberts M, Subramanian P, Yorke JA, Salzberg SL: A whole-genome assembly of the domestic cow, *Bos taurus. Genom Biol,* 10:R42, 2009. DOI: 10.1186/gb-2009-10-4-r42

11. Li H, Durbin R: Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25, 1754-1760, 2009. DOI: 10.1093/bioinformatics/btp324

12. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J: The sequence alignment/map format and SAMtools. *Bioinformatics*, 25, 2078-2079, 2009. DOI: 10.1093/bioinformatics/btp352

13. Li H1, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup: The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genom Res,* 20, 1297-1303, 2010. DOI: 10.1101/gr.107524.110

14. 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, Daly MJ: A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet,* 43, 491-498, 2011. DOI: 10.1038/ng.806

15. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA: From fastQ data to high confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics,* 43 (11): 11.1-33, 2013. DOI: 10.1002/0471250953.bi1110s43

16. Chen K, Wallis JW, Mclellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, Shi X, Fulton RS, Ley TJ, Wilson RK, Ding L, Mardis ER: BreakDancer: An algorithm for high-resolution mapping of genomic structural variation. *Nat Methods*, 6, 677-681, 2009. DOI: 10.1038/nmeth.1363

17. Kumar P, Henikoff S, Ng PC: Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*, 4, 1073-1081, 2009. DOI: 10.1038/nprot.2009.86

18. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol,* 28, 2731-2739, 2011. DOI: 10.1093/molbev/msr121

19. Tang H, Peng J, Wang P, Risch NJ: Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol,* 28, 289-301, 2005. DOI: 10.1002/gepi.20064

20. Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho TJ, Koutseva N, Zaghlul S, Graves T, Rock S, Kremitzki C, Fulton RS, Dugan S, Ding Y, Morton D, Khan Z, Lewis L, Buhay C, Wang Q, Watt J, Holder M, Lee S, Nazareth L, Alföldi J, Rozen S, Muzny DM, Warren WC, Gibbs RA, Wilson RK, Page DC: Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature,* 508, 494-499, 2014. DOI: 10.1038/nature13206

21. Soh YQ, Alföldi J, Pyntikova T, Brown LG, Graves T, Minx PJ, Fulton RS, Kremitzki C, Koutseva N, Mueller JL, Rozen S, Hughes JF, Owens E, Womack JE, Murphy WJ, Cao Q, de Jong P, Warren WC, Wilson RK, Skaletsky H, Page DC: Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell*, 159, 800-813, 2014. DOI: 10.1016/j.cell.2014.09.052

22. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B: PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science*, 327, 836-840, 2010. DOI: 10.1126/science.1183439

23. Berg IL, Rita N, Lam KG, Shriparna S, Linda OH, May CA, Jeffreys AJ: PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet,* 42, 859-863, 2010. DOI: 10.1038/ng.658

24. Huerta-Sanchez E, Jin X, Bianba Z, Peter BM, Vinckenbosch N, Liang Y, Yi X, He M, Somel M, Ni P, Wang B, Ou X, Huasang, Luosang J, Cuo ZX, Li K, Gao G, Yin Y, Wang W, Zhang X, Xu X, Yang H, Li Y, Wang J, Wang J, Nielsen R: Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature*, 512, 194-197, 2014. DOI: 10.1038/nature13408

25. Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZXP, Pool JE, Xu X, Jiang H, Vinckenbosch N, Korneliussen TS, Zheng H, Liu T, He W, Li K, Luo R, Nie X, Wu H, Zhao M, Cao H, Zou J, Shan Y, Li S, Yang Q, Asan, Ni P, Tian G, Xu J, Liu X, Jiang T, Wu R, Zhou G, Tang M, Qin J, Wang T, Feng S, Li G, Huasang, Luosang J, Wang W, Chen F, Wang Y, Zheng X, Li Z, Bianba Z, Yang G, Wang X, Tang S, Gao G, Chen Y, Luo Z, Gusang L, Cao Z, Zhang Q, Ouyang W, Ren X, Liang H, Zheng H, Huang Y, Li J, Bolund L, Kristiansen K, Li Y, Zhang Y, Zhang X, Li R, Li S, Yang H, Nielsen R, Wang J, Wang J: Sequencing of 50 human exomes reveals adaptation to high altitude. *Science*, 329, 75-78, 2010. DOI: 10.1126/science.1190371

26. Berra E, Benizri E, Ginouvès A, Volmat V, Roux D, Pouysségur J: HIF prolylhydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1alpha in normoxia. *EMBO J,* 22, 4082-4090, 2003. DOI: 10.1093/emboj/cdg392

27. Lorenzo FR, Huff C, Myllymäki M, Olenchock B, Swierczek S, Gordeuk V, Wuren T, Ri-Li G, McClain DA, Khan TM, Koul PA, Guchhait P, Salama ME, Xing J, Semenza GL, Liberzon E, Wilson A, Simonson TS, Jorde LB, Kaelin WG Jr, Koivunen P, Prchal JT: A genetic mechanism for Tibetan high-altitude adaptation. *Nat Genet,* 46, 951-956, 2014. DOI: 10.1038/ng.3067

28. Ikeda J, Kaneda S, Kuwabara K, Ogawa S, Kobayashi T, Matsumoto M, Yura T, Yanagi H: Cloning and expression of cDNA encoding the human 150 kDa oxygen-regulated protein, ORP150. *Biochem Biophysic Res Commun,* 230, 94-99, 1997. DOI: 10.1006/bbrc.1996.5890

29. Newman JH, Holt TN, Cogan JD, Womack B, Phillips JA 3rd, Li C, Kendall Z, Stenmark KR, Thomas MG, Brown RD, Riddle SR, West JD, Hamid R: Increased prevalence of EPAS1 variant in cattle with high-altitude pulmonary hypertension. *Nat Commun,* 6: 6863, 2015. DOI: 10.1038/ncomms7863

30. Zechner R, Kienesberger PC, Haemmerle G, Zimmermann R, Lass A: Adipose triglyceride lipase and the lipolytic catabolism of cellular fat stores. *Jf Lipid Res,* 50, 3-21, 2009. DOI: 10.1194/jlr.R800031-JLR200

31. Chmurzynska A: The multigene family of fatty acid-binding proteins (FABPs): Function, structure and polymorphism. *J Appl Genet,* 47, 39-48, 2006. DOI: 10.1007/BF03194597

32. Smathers RL, Petersen DR: The human fatty acid-binding protein family: Evolutionary divergences and functions. *Human Genom,* 5, 170-191, 2011. DOI: 10.1186/1479-7364-5-3-170

33. Qiu Q, Wang L, Wang K, Yang Y, Ma T, Zhang X, Ni Z, Hou F, Long R, Abbott R, Lenstra J, Liu J: Yak whole-genome resequencing reveals domestication signatures and prehistoric population expansions. *Nat Commun,* 6: 10283, 2015. DOI: 10.1038/ncomms10283

34. Xiao Z, Wang K, Wang L, Yang Y, Ni Z, Xie X, Shao X, Han J, Wan D, Qiu Q: Genome-wide patterns of copy number variation in the Chinese Yak genome. *BMC Genom,* 17: 379, 2016. DOI: 10.1186/s12864-016-2702-6

35. Soulé ME: Heterozygosity and developmental stability: another look. *Evolution*, 33, 396-401, 1979. DOI: 10.1111/j.1558-5646.1979.tb04693.x

36. Qi X, Han J, Wang G, Rege JE, Hanotte O: Assessment of cattle genetic introgression into domestic Yak populations using mitochondrial and microsatellite DNA markers. *Anim Genet,* 41, 242-252, 2010. DOI: 10.1111/j.1365-2052.2009.01989.x

37. Cortez D, Marin R, Toledoflores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H: Origins and functional evolution of Y chromosomes across mammals. *Nature*, 508, 488-493, 2014. DOI: 10.1038/nature13151

38. Chang TC, Yang Y, Retzel EF, Liu WS: Male-specific region of the bovine Y chromosome is gene rich with a high transcriptomic activity in testis development. *Proc Nat Acad Sci USA*, 110, 12373-12378, 2013. DOI: 10.1073/pnas.1221104110

39. Bachtrog D: Signs of genomic battles in mouse sex chromosomes. *Cell*. 159, 716-718, 2014. DOI: 10.1016/j.cell.2014.10.036

40. Binz SK, Sheehan AM, Wold MS: Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair*, 3, 1015-1024, 2004. DOI: 10.1016/j.dnarep.2004.03.028

41. Sandor C, Li W, Coppieters W, Druet T, Charlier C, Georges M: Genetic variants in REC8, RNF212, and PRDM9 influence male recombination in cattle. *PLOS Genet,* 8:e1002854, 2012. DOI: 10.1371/journal.pgen.1002854

42. Sanchez JE, Perera E, Baumbach L, Cleveland WW: Growth hormone receptor mutations in children with idiopathic short stature. *J Clin Endocrinol Metab,* 83, 4079-4083, 1998. DOI: 10.1210/jcem.83.11.5238

43. Wilkanowska A, Mazurowski A, Mroczkowski S, Kokoszynski D: Prolactin (PRL) and prolactin receptor (PRLR) genes and their role in poultry production traits. *Folia Biol,* 62, 1-8, 2014. DOI: 10.3409/fb62_1.1

44. Hou JX, An XP, Song YX, Wang JG, Ma T, Han P, Fang F, Cao BY: Combined effects of four SNPs within goat PRLR gene on milk production traits. *Gene,* 529, 276-281 2013. DOI: 10.1016/j.gene.2013.07.057

45. Decker JE, Mckay SD, Rolf MM, Kim J, Molina AA, Sonstegard TS, Hanotte O, Götherström A, Seabury CM, Praharani L, Babar ME, Correia de Almeida Regitano L, Yildiz MA, Heaton MP, Liu WS, Lei CZ, Reecy JM, Saif-Ur-Rehman M, Schnabel RD, Taylor JF: Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLOS Genet,* 10:e1004254, 2014. DOI: 10.1371/journal.pgen.1004254

46. Rubin CJ, Zody MC, Eriksson J, Meadows JR, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F, Besnier F, Carlborg O, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L: Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 464 (7288): 587-591, 2010. DOI: 10.1038/nature08832

47. Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar A, Lindblad-Toh K: The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature,* 495 (7441): 360-364, 2013. DOI: 10.1038/nature11837