

Article

Identification of Genomic Characteristics and Selective Signals in a Du'an Goat Flock

Qiuming Chen ^{1,2}, Zihao Wang ¹, Junli Sun ¹, Yingfei Huang ¹, Quratulain Hanif ^{3,4} ,
Yuying Liao ^{1,*} and Chuzhao Lei ² 

- ¹ Guangxi Key Laboratory of Livestock Genetic Improvement, Animal Husbandry Research Institute of Guangxi Zhuang Autonomous Region, Nanning 530001, China; cqm19860612@126.com (Q.C.); wangzh761006@163.com (Z.W.); sjn313@126.com (J.S.); Huangyingfei2020@outlook.com (Y.H.)
- ² Key Laboratory of Animal Genetics, Breeding and Reproduction of Shaanxi Province, College of Animal Science and Technology, Northwest A&F University, Yangling 712100, China; leichuzhao1118@nwfau.edu.cn
- ³ Computational Biology Laboratory, Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering, Faisalabad 577, Pakistan; micro32uvas@gmail.com
- ⁴ Department of Biotechnology, Pakistan Institute of Engineering and Applied Sciences, Nilore, Islamabad 45650, Pakistan
- * Correspondence: gxxmskyk@gxxmyjs.com; Tel.: +86-135-7299-2159

Received: 24 February 2020; Accepted: 4 June 2020; Published: 6 June 2020



Simple Summary: Goat plays an irreplaceable role as domestic animals by providing milk, meat, and fiber. With the advancement in human civilization, over 500 goat breeds with various adaptive traits have been established. Whole-genome sequencing and subsequent analyses (including selective sweep analysis) are important tools to reveal the genetic basis underlying these adaptive traits. In this study, we investigated the sequenced genomes of 15 Du'an goats from one flock to identify the genomic characteristics and candidate genomic regions which might be involved in the adaptive traits. Selective sweep analysis revealed selected genes and/or pathways related to immune resistance, small body size, and heat tolerance. The identification of genomic characteristics and selective signals will not only help understand the demographic history and the genetic mechanism underlying adaptive traits but can provide novel insights into genetic diversity and development of breeding strategies in Du'an goats.

Abstract: The Du'an goat is one of the most important farm animals in the Guangxi Autonomous Region of China, but the genetic basis underlying its adaptive traits has still not been investigated. Firstly, in this study, the genomes of 15 Du'an goats from a breeding farm were sequenced (mean depth: 9.50X) to analyze the patterns of genetic variation. A comparable diversity (17.3 million single nucleotide polymorphisms and 2.1 million indels) was observed to be associated with a lower runs of homozygosity-based inbreeding coefficient and smaller effective population size in comparison with other breeds. From selective sweep and gene set enrichment analyses, we revealed selective signals related to adaptive traits, including immune resistance (*serpin* cluster, *INFG1*, *TLR2*, and immune-related pathways), body size (*HMGA2*, *LCOR*, *ESR1*, and cancer-related pathways) and heat tolerance (*MTOR*, *ABCG2*, *PDE10A*, and purine metabolism pathway). Our findings uncovered the unique diversity at the genomic level and will provide the opportunities for improvement of productivity in the Du'an goat.

Keywords: Du'an goat; whole genome sequencing; selective signals

1. Introduction

The goat (*Capra hircus*) was one of the oldest livestock species domesticated in the Zagros Mountains in the Fertile Crescent, around 10,000 years ago [1]. Domestic goats play an important role in the transition of human civilization from foraging and hunting to farming and herding by supplying milk, meat, and fiber. Due to natural and anthropogenic factors (e.g., climate changes, human migration, and population expansion), there are now more than 1006 million individuals and 557 registered breeds worldwide [2,3].

The Du'an goat is a small-statured goat breed which is mainly reared in Du'an county of Guangxi Zhuang Autonomous Region in China and is sporadically found in the neighboring counties. The number of Du'an goats raised per year is more than 0.6 million. Du'an goats are featured as a meat-type breed with the following characteristics: (1) small body size with the average body weight of 41 kg in the female; (2) high heat tolerance with average temperatures of 29 Celsius degree in July in the Du'an county; (3) high disease resistance against *Fasciola hepatica*; (4) diverse types of coat color (white, black, etc.); (5) tender meat; and (6) diverse types of horn (Figure 1) [4,5]. However, the genetic basis underlying these characteristics is still not investigated.



Figure 1. Breed characteristics and geographic location of the Du'an goat.

The decreasing cost of high-throughput sequencing has resulted in considerable progress in genome characterization and demographic history of various organism. Moreover, the selection sweep analysis from the genome data set helped in the identification of the candidate genes/regions associated with adaptive traits. For example, a whole-genome sequencing study showed a very high diversity

associated with linkage disequilibrium in Moroccan goat [6]. The selection sweep analysis of Korean indigenous goats revealed selective signals for Salmonella infection and cardiomyopathy [2]. One of the recently-published studies including 164 modern domestic goat, 52 ancient goat, 24 modern bezoar, and 4 ancient bezoar genomes indicated that Asian goats are genetically distinct from European and African samples. More importantly, the study identified a strong signature of selection harboring *MUC6* gene conferred enhanced immune resistance to gastrointestinal pathogens [7]. Furthermore, studies showed selective signals in other important traits, such as hair growth in the Cashmere goat [8], litter size traits in the Laoshan dairy goat [9], and high altitude adaptation in the Tibetan goat [3]. However, there have been scarce or no attempts to analyze the genomic characteristics and selective signals in the Du'an goat.

In the present study, we performed whole-genome resequencing of 15 Du'an goats to identify genomic characteristics and selective signals. The revelation of selective signals related to immune system, body size, and heat resistance will help further investigate the genetic mechanism underlying adaptive traits in the domestic goats of South China. Our findings will also provide valuable resources for breeding programs and managemental strategies regarding the Du'an goat.

2. Materials and Methods

2.1. Animal Sampling

The ear tissue of 15 Du'an goats was collected from a designated breeding farm in Du'an county of Guangxi Zhuang Autonomous Region, China. Genomic DNA was extracted from the ear tissue using the standard phenol–chloroform protocol [10]. The sampling procedure was approved by the Institutional Animal Care and Use Committee of Northwest A&F University (permit number: NWAFA1019).

2.2. Genome Sequencing and Variant Calling

Each DNA sample was used to construct paired-end sequencing libraries with the insert size of 350 bp. The constructed libraries were sequenced using the Illumina NovaSeq 6000 platform at the Novogene Bioinformatics Institute, Beijing, China. The length of reads was 150 bp. The raw reads were filtered using Trimmomatic [11] with the following parameters: LEADING:20, TRAILING:20, SLIDINGWINDOW:3:15, AVGQUAL:20, MINLEN:35, and TOPHRED33. The clean reads were aligned to the goat reference assembly (ARS1) [12] using Burrows-Wheeler Alignment Maximal Exact Matches algorithm [13] with default parameters. After alignment, the short reads in the BAM file were sorted and duplicated reads were marked (Supplementary Note).

For SNP (single nucleotide polymorphism) and indel detection, we orderly used the modules (HaplotypeCaller, CombineGVCFs, GenotypeGVCFs, and SelectVariants) of Genome Analysis Toolkit 3.8 (GATK) [14] to call the raw variants (Supplementary Note). The raw SNPs were filtered using the VariantFiltration module of GATK with the following parameters recommended by GATK: SOR > 3.0, QD < 2.0, FS > 60.0, MQ < 40.0, MQRankSum < -12.5, ReadPosRankSum < -8.0, DP < 47, and DP > 427. Similarly, raw indels were filtered using the following parameters recommended by GATK: QD < 2.0, FS > 200, and ReadPosRankSum < -20.0. The package ANNOVAR [15] was used to annotate the variants to identify the protein coding mutations caused.

2.3. Population Genetic Analysis

To check for the relatedness among the samples, identity-by-descent testing was performed using PLINK (-genome) [16]. Nucleotide diversity (-window-pi 50,000 -window-pi-step 50,000) and inbreeding coefficient (-het) were also calculated to detect the genomic characteristics in the Du'an goat using VCFtools [17]. The linkage disequilibrium (LD) decay in Du'an goats was estimated using PopLDdecay [18].

Runs of homozygosity (ROH) in the Du'an goat was detected using PLINK [16] (Supplementary Note). The ROH-based inbreeding coefficient (F_{ROH}) was calculated as the average fraction of the autosome covered by ROH. In addition, effective population size (N_e) was calculated using SNeP version 1.1 with default parameters [19].

2.4. Detecting Positive Selection

Before detecting the selective signals, one of the samples was removed due to relatedness (identity-by-descent $PI_HAT = 0.29$). Evidence of positive selection was investigated through three statistics. First, the SNPs were filtered with minor allele frequency < 0.05 and missing rate > 0.10 using VCFtools [16], and missing alleles imputation and haplotype inference were performed using Beagle version 4.1 (Supplementary Note) [20]. The integrated haplotype score (iHS) was estimated from haplotype information to find the selected alleles segregating at intermediate frequency in the Du'an goat using selscan [21]. We also normalized the iHS score using the norm module of selscan. For the iHS statistic, the fraction of SNPs with $|normalized\ iHS\ score| > 2$ in 50 kb windows with 20 kb increments was calculated across the autosomes.

Second, based on the empirical frequency spectrum with all allele frequencies across the autosomes, sweepfinder2 [22] was used to calculate the composite likelihood ratio (CLR) statistics for sites every 1 kb ($-\lg 1000\ FreqFile\ SpectFile\ OutFile$) using allele frequency information to detect the completed sweep. In order to define candidate regions, the genome was divided into 50 kb windows with 20 kb increments. In each window, the maximum CLR was defined as the test statistic.

Third, the nucleotide diversity ($\theta\pi$) was calculated ($-\text{window-pi } 50,000\ -\text{window-pi-step } 20,000$) using VCFtools [16]. Outlier regions supported by two or three methods (higher iHS: top 1%, higher CLR: top 1%, and lower $\theta\pi$: bottom 1%) were defined as the candidate regions under positive selection. In addition, the Tajima'D statistic was also calculated ($-\text{TajimaD } 20,000$) across the autosomes to consolidate our results using VCFtools [16].

2.5. Candidate Genes Analysis

Based on the goat reference genome (ARS1), a custom perl script was used to annotate the regions under positive selection. The protein-coding genes overlapped with the regions under positive selection were defined as candidate genes. Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis was performed on the candidate genes using KOBAS 3.0 [23]. Fisher's exact test with false discovery rate testing was executed and a corrected p -value of less than 0.05 was chosen as an inclusion criterion for functional categories.

3. Results

3.1. Sequencing and Identification of SNPs and Indels

To identify the SNPs and indels in the Du'an goat, we sequenced 15 genomes. In total, ~2.81 billion clean reads were generated and aligned to the goat reference genome (ARS1). The average genome coverage was 99.68% (ranging from 99.24% to 99.79%) with an in-depth mapping coverage of 9.50 folds (ranging from 7.74 to 11.47) (Table S1).

After quality filtering, a total of 17,317,364 SNPs was retained, while the average number of singletons, observed homozygosity, and observed heterozygosity were 163,581 (ranging from 96,181 to 373,907), 11,589,894 (ranging from 10,995,608 to 12,573,915) and 4,433,699 (ranging from 3,414,574 to 5,114,509), respectively. The transitions/transversions ratio was 2.3359, which was comparable to that in the Dazu black goat (2.3347) [24], and lower than the Moroccan goat (2.3552) [6], suggesting that the variant calling of the samples was done correctly. The majority of SNPs were identified in the intergenic (61.6%) and intronic (35.6%) regions, while, only a fraction of SNPs were detected in the genic regions including exonic (0.8%) and untranslated regions (0.7%). Among the exonic SNPs, 45,410 SNPs caused amino acid change, while 4009 and 83 SNPs caused the creation and elimination of a stop

codon, respectively (Figure 2a). Similarly, 2,111,567 indels were detected, where, 0.3% were detected in exonic regions, 37.6% in intronic regions, 59.5% in intergenic regions, and 1% in untranslated regions. Among exonic indels, 1265 and 1020 indels led to frameshift deletion and insertion, respectively, and 58 and 4 indels led to the creation and elimination of a stop codon, respectively (Figure 2b).

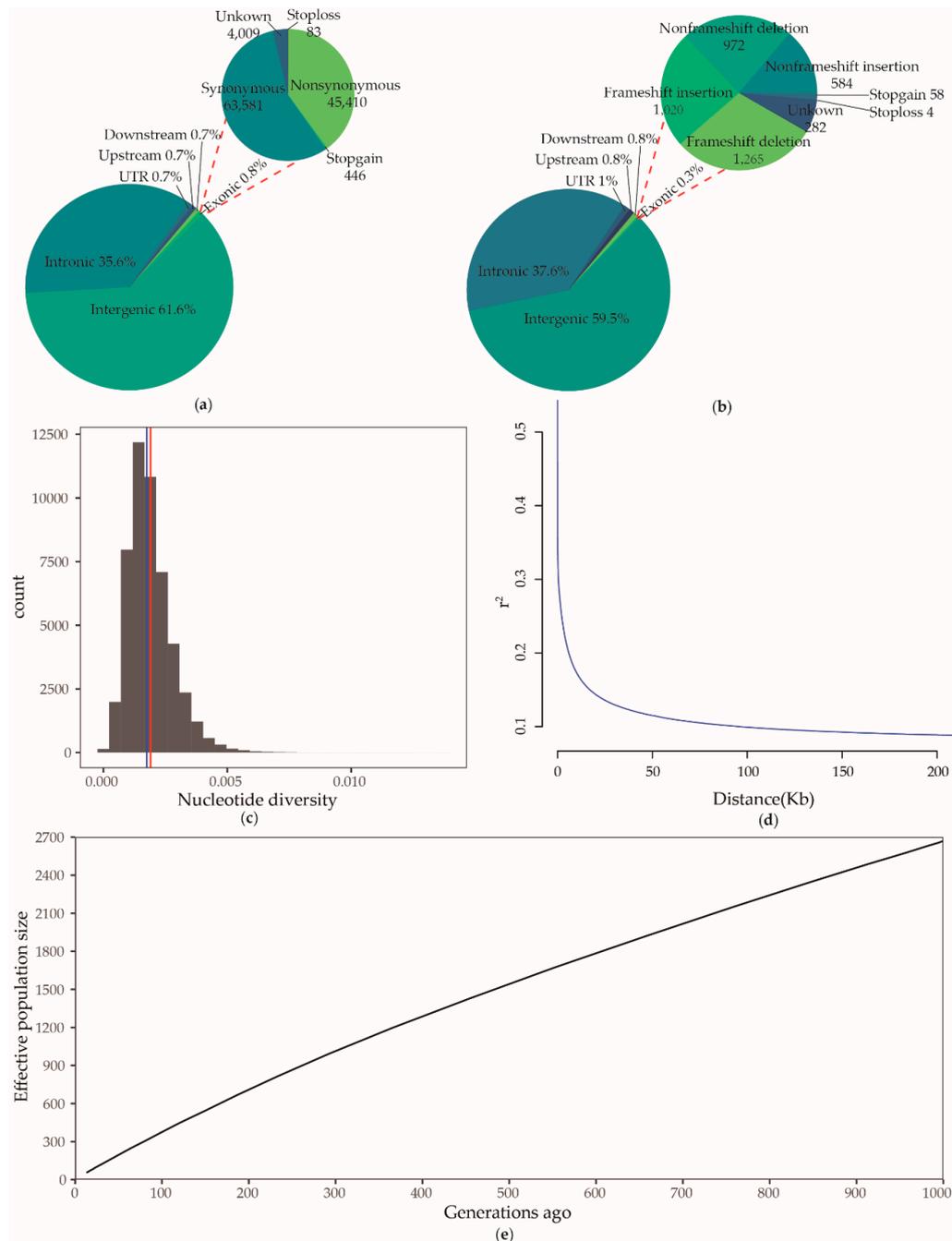


Figure 2. Identification of genomic characteristics. **(a)** Functional classification of the detected single nucleotide polymorphisms. **(b)** Functional classification of the detected indels. **(c)** Distribution of nucleotide diversity in non-overlapping windows of 50 kb across the autosomes. The red line and blue line indicate mean and median, respectively. **(d)** Genome-wide average linkage disequilibrium (LD) decay in the Du'an goat. **(e)** Effective population size over the past 1000 generations.

We also investigated the genomic characteristics in the Du'an goat. According to identity-by-descent analysis, only one pair from the individuals showed PI_HAT value of greater than 0.25 (which was

later excluded from the selection sweep analysis). The mean and median of nucleotide diversity were 0.0019 and 0.0017, respectively, in non-overlapping windows of 50 kb across the autosomes (Figure 2c). The average SNP-based inbreeding coefficient was 0.06613 in the Du'an goat (Table S2). Linkage disequilibrium (LD) analysis indicated the physical distance between SNPs to be 77 kb (reported as half of its maximum) (mean $r^2 = 0.35$) (Figure 2d), while, the average ROH coverage was 6.217 Mb (ranging from ~1–44.312 Mb). Accordingly, the average F_{ROH} was 0.0025 (ranging from ~0.0004–0.0179) (Table S3). In addition, an estimated decrease in the N_e was also observed in the Du'an goat (from 2668 at 999 generations ago to 55 at 13 generations ago) (Figure 2e) (Table S4).

3.2. Identification of Selective Sweep

To identify candidate regions and genes under selection in the Du'an goat, we used iHS statistic, CLR statistic, and nucleotide diversity. The regions with support from two or three methods (lower nucleotide diversity, higher iHS statistic, and CLR statistic) were defined as the candidate regions under selection (Figure 3a–c). After merging the overlapping regions, 124 candidate regions with 245 genes were identified under selection in the Du'an goat (Table S5). Based on annotation of the ANNOVAR software, we also identified 131 nonsynonymous SNPs with the high frequency (>0.8) in 68 selected genes (Table S6) and eight exonic indels with the high frequency (>0.8) in six selected genes (Table S7).

An extreme iHS score was located in the serpin cluster. The selective signals in the serpin cluster were further confirmed by a higher value of CLR, local reduced nucleotide diversity, and lower value of Tajima's D (Figure 3d). *HMG2* was identified as a candidate gene, which showed a strong CLR signal and reduced nucleotide diversity. The region also showed a negative value of Tajima's D indicative of an increase of rare alleles, such as recent bottleneck followed by expansion (Figure 3e). In addition, we found *MTOR* fell in a region that contained some of the most extreme CLR statistic and nucleotide diversity. The negative value of Tajima's D also supported the selection in the region (Figure 3f). Other interesting candidate genes overlapped with selective signals were *IFNGR1*, *TLR2*, *ESR1*, *LCOR*, *ABCG2*, and *PDE10A*.

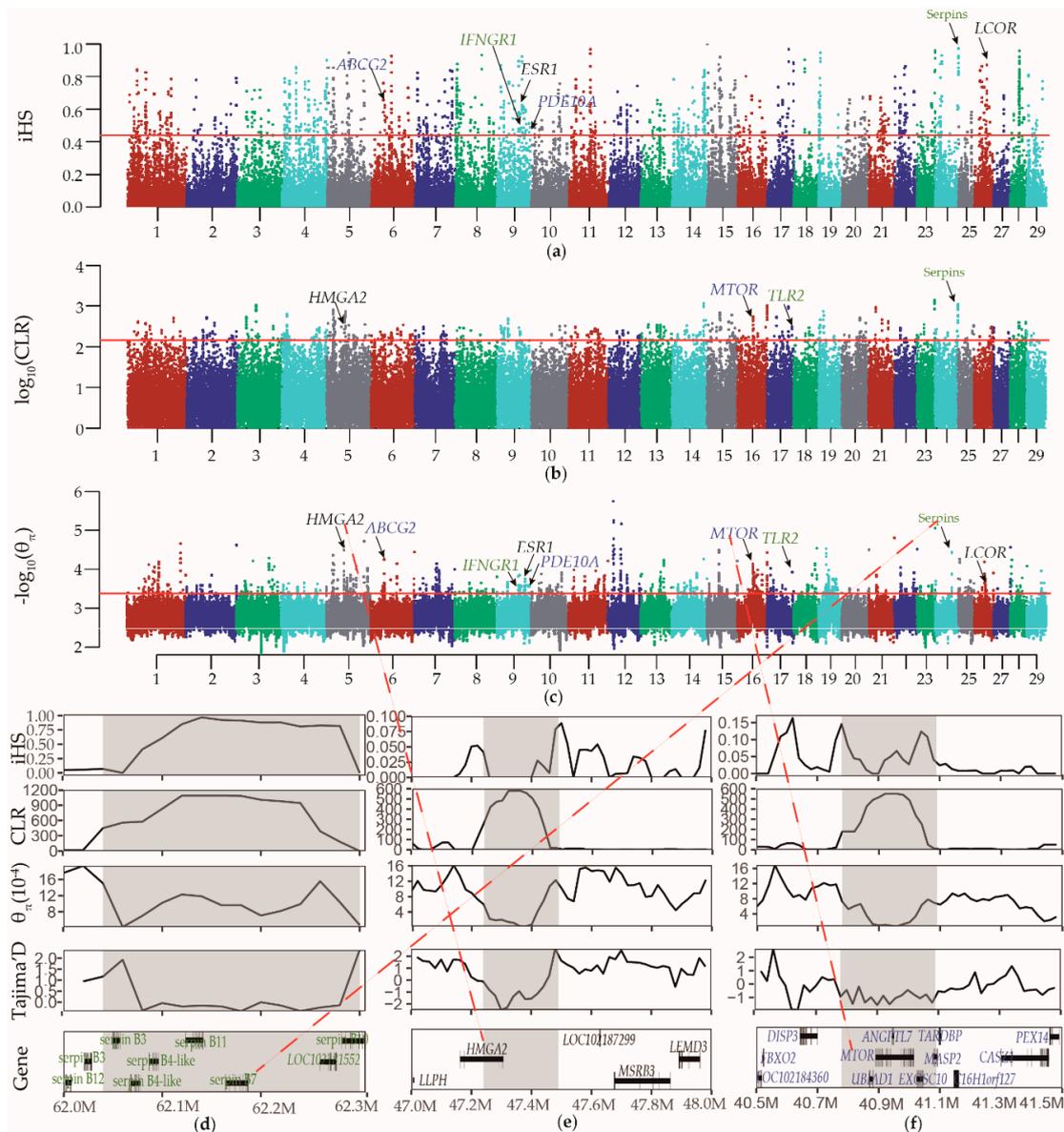


Figure 3. Identification of selective signals in the Du’an goat. Distribution of iHS statistic (a), composite likelihood ratio (CLR) statistic (b), and nucleotide diversity (c) in a 50 kb sliding window with a 20 kb step across all autosomes. The red lines indicate the corresponding threshold (iHS statistic > 0.4424167, CLR statistic > 148.2987 and nucleotide diversity < 0.0004230301). (d) Selective sweep at the serpin cluster on chromosome 24 (62.04–62.33 Mb) (gray area). (e) Selective sweep at the *HMGA2* on chromosome 5 (47.24–47.49 Mb) (gray area). (annotation, black rectangles and black lines represent exons and introns, respectively). (f) Selective sweep at the *MTOR* on chromosome 16 (40.78–41.09 Mb) (gray area). In gene.

3.3. Kegg Pathway Analyses of the Candidate Genes under Selection

In order to provide more functional information about the candidate genes under selection, gene set enrichment analysis of the KEGG pathway was performed, resulting in nine significant pathways. The most significant pathway was olfactory transduction. Other pathways included four immune-related categories (amoebiasis, acute myeloid leukemia, staphylococcus aureus infection, and tuberculosis), two cancer-related categories (PD-L1 expression and PD-1 checkpoint pathway in cancer, and proteoglycans in cancer), and one metabolism-related category (purine metabolism) (Table 1).

Table 1. Enriched gene ontology terms among the candidate genes under selection ¹.

KEGG Term	Corrected <i>p</i>	Gene Count
Olfactory transduction	2.22×10^{34}	44 (448)
Amoebiasis	0.00194	6 (95)
PD-L1 expression and PD-1 checkpoint pathway in cancer	0.0097	5 (89)
Acute myeloid leukemia	0.021178	4 (66)
Staphylococcus aureus infection	0.003584	5 (68)
Tuberculosis	0.023397	6 (179)
Purine metabolism	0.029284	5 (130)
Proteoglycans in cancer	0.032434	6 (203)
Salivary secretion	0.040032	4 (90)

¹ Enriched terms are color-coded to reflect relatedness in functional proximity. Blue, immune system; green, cancer; and orange, metabolism. For each term, gene count shows number of candidate genes (total number of annotated genes).

4. Discussion

4.1. Genetic Diversity in the Du'an Goat

Study of genetic diversity plays an important role in the conservation and utilization of germplasm resources, revelation of evolutionary history, and investigation of phylogenetic relationships. The number of the detected SNPs in the Du'an goat (~17M) was lower than that in the Moroccan (~33M) and Korean indigenous (~37M), and higher than that in the Tibetan Cashmere (~12M), Chengdu Brown (~10M), and Jintang Black goat (~12M) [8], which could be explained by the difference in the number of individuals, reference genome, detection method, filtering criteria, or breed. The average nucleotide diversity in the Du'an goat was lower than that in the Iranian indigenous goat, while it was found to be higher than that in the Nubian and Korean indigenous goat. The average SNP-based inbreeding coefficient in the Du'an goat was higher than the Nubian and Korean indigenous goat. The average r^2 in 50 Kb in the Du'an goat was higher than that in the Iranian indigenous, Moroccan indigenous, and Boer goat, while it was lower than that in the Nubian goat (Table 2). However, ROH-based inbreeding coefficient (0.0025) was lower in the Du'an goat than the Chengdu Brown (0.194), Tibetan Cashmere (0.068), Moroccan indigenous (0.085), and Jintang Black goat (0.119) [8]. The revelation of individual ROH may help the conservation of this breed, since animals with high levels of ROH could be excluded or assigned lower priority for mating to minimize the loss in genetic diversity. In addition, a decreasing N_e was observed in the Du'an goat, and the estimated N_e value was 55 at 13 generations ago, which was lower than that in the Arbas Cashmere (95) and Guangfeng goat (64) [25]. Similarly, a previous study also found higher N_e values in the Argentinian (67), French (57) and South African goat (93) at 10 generations ago [26], which supports our results. A recent decrease in N_e can be due to a stronger selective pressure in recent generations or to a recent bottleneck. From the above results, we can conclude that the Du'an goat has higher SNP-based inbreeding, lower ROH-based inbreeding, fewer variants, higher linkage decay, and smaller effective population size and nucleotide diversity comparable to the other locally-adapted populations, suggesting the unique genetic characteristics.

Table 2. Difference in genomic characteristics between the Du'an goat and seven other goat breeds.

Breed	Number of SNPs	Nucleotide Diversity	Inbreeding Coefficient ¹	Average r ² (50 kb)
Du'an goat	17,317,364	0.001905	0.06613	0.3674
Iranian indigenous goat ²	35,742,191	0.001998	0.06229	0.087908
Moroccan indigenous goat ²	32,914,220	0.001859	0.06143	0.079364
Saanen goat ²	36,845,217	0.001783	0.00207	0.063634
Nubian goat ²	23,726,534	0.001117	−0.0234	0.487128
Boer goat ²	32,384,827	0.001724	−0.04455	0.091387
Korean indigenous goat ²	37,715,208	0.001472	0.01661	0.161031
<i>Capra aegagrus</i> ²	39,222,625	0.001804	0.06821	0.087908

¹ Inbreeding coefficient was estimated using a method of moments. ² The genomic characteristics in these breeds was identified by Badr Benjelloun et al. in 2019 [2].

4.2. Putatively-Selected Genes

4.2.1. Immune System

High disease resistance is one of the key features in the Du'an goat. A previous study reported that *Fasciola hepatica* is endemic in Guangxi Zhuang Autonomous Region inhabited by the Du'an goat [5]. Among the putatively-selected genes, serpins are serine proteinase inhibitors involved in host defense pathways [27]. An in-depth proteomic analysis of *Fasciola hepatica* intra-mammalian stages showed that five serpins were identified in adult excretory/secretory product and somatic soluble newly excysted juveniles, suggesting that serpins could be important for *Fasciola hepatica* establishment and survival in the host [28]. Other noteworthy genes in our candidate list were *IFNGR1* involved in hepatitis B virus and mycobacteria infection [29,30] and *TLR2* implicated in pathogen recognition and activation of innate immunity [31]. Interestingly, the ADAPTmap consortium found that the *IFNGR1* was one of candidate genes involved in environmental adaptation; another 10 candidate genes (*ARL8A*, *CDC25A*, *GPR37L1*, *IL22RA2*, *IQCE*, *KIT*, *LOC102181444*, *MAP28*, *NAV1*, and *PTPN71*) identified by the ADAPTmap consortium were also within the region under selection in the Du'an goat [32]. At the same time, we found four nonsynonymous SNPs with the high frequency in three immune-related genes in the Du'an goat (*IFNGR1*: c.A733G, p.I245V; *SERPINB7*: c.A82G, p.M28V; *TLR2*: c.T47C, p.V16A; and *TLR2*: c.A739G, p.I247V). In addition, gene set enrichment analysis showed four immune-related KEGG pathways. We could speculate that these genes, nonsynonymous SNPs, or pathways could have been affected by selection targeting at immune traits such as resistance to viruses, bacteria, or parasites.

4.2.2. Body Size

Another important Du'an goat characteristic is its small body size. The average body weight is 48 kg in the adult male and 41 kg in the adult female. One interesting observation was the presence of *HMGA2* among the genes with strong signal of selection. *HMGA2* is a widely-studied gene explaining the difference in body size in many organisms including humans [33], cattle [34], dog [35], and horse [36]. Other putatively-selected candidates included *LCOR* associated with cattle stature [34] and *ESR1* associated with body height in two Swedish populations [37]. It is worth mentioning that there were two nonsynonymous SNPs with the high *LCOR* frequency in the Du'an goat (c.G3029A:p.G1010E and c.G2513C:p.S838T). In addition, we observed that two cancer-related KEGG pathways were detected in gene set enrichment analysis. In fact, the relationship between cancer and body size or obesity has been established by numerous studies [38–40]. From the above results, we can conclude that these genes or pathways were strong candidates contributing to the small body size in the Du'an goat.

4.2.3. Heat Tolerance

Guangxi Zhuang Autonomous Region, inhabited by the Du'an goat, is in the south subtropical zone, where the average annual temperature is 17–23 Celsius degree, the average maximum temperature is about 29 Celsius degree in July [41]. In the genome-wide screen, we found a very strong selective

signal with a high-frequent nonsynonymous SNPs in *MTOR* (c.G5507C:p.G1836A). A previous study has shown that *MTOR* is essential for the proteotoxic stress response, *HSF1* activation, and heat shock protein synthesis [42]. Moreover, heat stress experiments showed enhanced *MTOR* signaling in human and rat skeletal muscle [43,44]. A prominent study showed that *MTOR* could regulate phase separation of PGL granules to modulate their autophagic degradation and heat stress adaptation during *Caenorhabditis elegans* embryogenesis [45]. In fact, another study has proved that mutation of *MTOR* gene could be associated with heat tolerance in Chinese cattle [46]. Other strong selection candidates included *ABCG2* implicated in oxidative stress [47], and *PDE10A* implicated in a thermoregulatory role [48]. It has been demonstrated that the duration of thermal stress affects the expression of *ABCG2* in broilers [49]. In addition, we found that purine metabolism was significantly enriched in putatively-selected genes using KEGG pathways. A review has presented the function of heat shock protein and metabolic responses under heat stress or high ambient temperature [50]. Some studies have shown that purine metabolism is involved in heat stress [51,52]. All the above findings indicate that these genes or pathways may be important factors in Du'an goat adaptation to a subtropical environment.

5. Conclusions

Our study used resequencing data from the Du'an goat to identify genomic characteristics and selective signals. The descriptions of genomic characteristics (nucleotide diversity, SNPs, indels, inbreeding coefficient, linkage disequilibrium, ROH, and effective population size) are not only crucial for maintaining herd genetic diversity and development of breeding programs, but also help understand the demographic history in the Du'an goat. The revelation of selective signals related to immune system, body size, and heat tolerance will provide an important opportunity for further investigation of the genetic mechanisms and underlying adaptive traits in the Du'an goat. The identification of some candidate variants will facilitate investigation into related phenotypical data to detect causative variants. Further studies are hereby desired with broader and thorough investigation, leading towards more elaborative genomic characterization and demographic history based on larger datasets, augmenting the current study for better managemental and breeding policies.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2076-2615/10/6/994/s1>, Table S1: Summary of sequencing and mapping results in 15 Du'an goats, Table S2: Inbreeding coefficient (F) for each individual, Table S3: Runs of homozygosity for each individual, Table S4: Historic effective population size over 1000 generations ago, Table S5: Putatively-selected regions in the Du'an goat, Table S6: Nonsynonymous SNPs under selection in Du'an goats, Table S7: Exonic indels under selection in the Du'an goat, Supplementary Note: key parameters for software.

Author Contributions: Conceptualization, Y.L. and C.L.; methodology, Q.C. and Z.W.; software, Q.C.; validation, Z.W., J.S., and Y.H.; formal analysis, Q.C. and Q.H.; investigation, Y.L.; resources, Y.L.; data curation, Q.C.; writing—original draft preparation, Q.C.; writing—review and editing, Q.H. and C.L.; visualization, Q.C.; supervision, C.L. and Y.L.; project administration, C.L.; funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Guangxi special project for innovation-driven development (AA17204024).

Data Availability: The sequencing reads of each sample have been deposited at GenBank with the Bioproject accession number PRJNA607828.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Zeder, M.A.; Hesse, B. The initial domestication of goats (*Capra hircus*) in the Zagros Mountains 10,000 years ago. *Science* **2000**, *287*, 2254–2257. [[CrossRef](#)]
2. Kim, J.-Y.; Jeong, S.; Kim, K.H.; Lim, W.-J.; Lee, H.-Y.; Kim, N. Discovery of Genomic Characteristics and Selection Signatures in Korean Indigenous Goats Through Comparison of 10 Goat Breeds. *Front. Genet.* **2019**, *10*, 699. [[CrossRef](#)] [[PubMed](#)]

3. Wang, X.; Liu, J.; Zhou, G.; Guo, J.; Yan, H.; Niu, Y.; Li, Y.; Yuan, C.; Geng, R.; Lan, X. Whole-genome sequencing of eight goat populations for the detection of selection signatures underlying production and adaptive traits. *Sci. Rep.* **2016**, *6*, 38932. [[CrossRef](#)] [[PubMed](#)]
4. Du, L.; Li, J.; Ma, N.; Ma, Y.; Wang, J.; Yin, C.; Luo, J.; Liu, N.; Jia, Z.; Fu, C. *Animal Genetic Resources in China: Sheep and Goats*; China Agriculture Press: Beijing, China, 2011.
5. Yi, S.; He, B.; Ma, J.; Su, W.; Li, G. Investigation of liver function of Du'an goat in Guangxi. *Guangxi Agric. Sci.* **2008**, *3*, 374–376.
6. Benjelloun, B.; Alberto, F.J.; Streeter, I.; Boyer, F.; Coissac, E.; Stucki, S.; BenBati, M.; Ibnelbachyr, M.; Chentouf, M.; Bechchari, A. Characterizing neutral genomic diversity and selection signatures in indigenous populations of Moroccan goats (*Capra hircus*) using WGS data. *Front. Genet.* **2015**, *6*, 107. [[CrossRef](#)]
7. Zheng, Z.; Wang, X.; Li, M.; Li, Y.; Yang, Z.; Wang, X.; Pan, X.; Gong, M.; Zhang, Y.; Guo, Y. The origin of domestication genes in goats. *Sci. Adv.* **2020**, *6*, eaaz5216. [[CrossRef](#)]
8. Guo, J.; Zhong, J.; Li, L.; Zhong, T.; Wang, L.; Song, T.; Zhang, H. Comparative genome analyses reveal the unique genetic composition and selection signals underlying the phenotypic characteristics of three Chinese domestic goat breeds. *Genet. Sel. Evol.* **2019**, *51*, 70. [[CrossRef](#)]
9. Lai, F.-N.; Zhai, H.-L.; Cheng, M.; Ma, J.-Y.; Cheng, S.-F.; Ge, W.; Zhang, G.-L.; Wang, J.-J.; Zhang, R.-Q.; Wang, X. Whole-genome scanning for the litter size trait associated genes and SNPs under selection in dairy goat (*Capra hircus*). *Sci. Rep.* **2016**, *6*, 38096. [[CrossRef](#)]
10. Green, M.R.; Sambrook, J. *Molecular cloning: A Laboratory Manual*, 4th ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2012.
11. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
12. Bickhart, D.M.; Rosen, B.D.; Koren, S.; Sayre, B.L.; Hastie, A.R.; Chan, S.; Lee, J.; Lam, E.T.; Liachko, I.; Sullivan, S.T. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nat. Genet.* **2017**, *49*, 643–650. [[CrossRef](#)]
13. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)] [[PubMed](#)]
14. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernysky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [[CrossRef](#)] [[PubMed](#)]
15. Wang, K.; Li, M.; Hakonarson, H. ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **2010**, *38*, e164. [[CrossRef](#)] [[PubMed](#)]
16. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.; Daly, M.J. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
17. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [[CrossRef](#)]
18. Zhang, C.; Dong, S.-S.; Xu, J.-Y.; He, W.-M.; Yang, T.-L. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **2019**, *35*, 1786–1788. [[CrossRef](#)]
19. Barbato, M.; Orozco-terWengel, P.; Tapio, M.; Bruford, M.W. SNeP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* **2015**, *6*, 109. [[CrossRef](#)]
20. Browning, S.R.; Browning, B.L. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *Am. J. Hum. Genet.* **2007**, *81*, 1084–1097. [[CrossRef](#)]
21. Szpiech, Z.A.; Hernandez, R.D. Selscan: An efficient multithreaded program to perform EHH-based scans for positive selection. *Mol. Biol. Evol.* **2014**, *31*, 2824–2827. [[CrossRef](#)]
22. DeGiorgio, M.; Huber, C.D.; Hubisz, M.J.; Hellmann, I.; Nielsen, R. SweepFinder2: Increased sensitivity, robustness and flexibility. *Bioinformatics* **2016**, *32*, 1895–1897. [[CrossRef](#)]
23. Xie, C.; Mao, X.; Huang, J.; Ding, Y.; Wu, J.; Dong, S.; Kong, L.; Gao, G.; Li, C.-Y.; Wei, L. KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res.* **2011**, *39*, W316–W322. [[CrossRef](#)] [[PubMed](#)]

24. Guan, D.; Luo, N.; Tan, X.; Zhao, Z.; Huang, Y.; Na, R.; Zhang, J.; Zhao, Y. Scanning of selection signature provides a glimpse into important economic traits in goats (*Capra hircus*). *Sci. Rep.* **2016**, *6*, 36372. [[CrossRef](#)] [[PubMed](#)]
25. Berihulay, H.; Islam, R.; Jiang, L.; Ma, Y. Genome-Wide Linkage Disequilibrium and the Extent of Effective Population Sizes in Six Chinese Goat Populations Using a 50K Single Nucleotide Polymorphism Panel. *Animals* **2019**, *9*, 350. [[CrossRef](#)] [[PubMed](#)]
26. Visser, C.; Lashmar, S.F.; Van Marle-Köster, E.; Poli, M.A.; Allain, D. Genetic diversity and population structure in South African, French and Argentinian Angora goats from genome-wide SNP data. *PLoS ONE* **2016**, *11*. [[CrossRef](#)] [[PubMed](#)]
27. Gettins, P.G. Serpin structure, mechanism, and function. *Chem. Rev.* **2002**, *102*, 4751–4804. [[CrossRef](#)] [[PubMed](#)]
28. Di Maggio, L.S.; Tirloni, L.; Pinto, A.F.; Diedrich, J.K.; Yates, J.R., III; Benavides, U.; Carmona, C.; da Silva Vaz, I., Jr.; Berasain, P. Across intra-mammalian stages of the liver fluke *Fasciola hepatica*: A proteomic study. *Sci. Rep.* **2016**, *6*, 32796. [[CrossRef](#)]
29. Khanizadeh, S.; Mehrdad, R.; Mohebbi, S.; Naghoosi, H.; Mousavinasab, S.; Romnani, S.; Azimzadeh, P.; Sharifian, A.; Zali, M. Correlation between Polymorphism of-56 SNP (T/C) Interferon- γ Receptor 1 Gene and Chronic HBV Infection. *Iran. J. Virol.* **2011**, *2*, 19–24. [[CrossRef](#)]
30. Muszlak, M.; Chappier, A.; Barry, R.H.; Castella, C.; Cremades, F.; Goulois, E.; Laporte, R.; Casanova, J.; Ranaivoarivony, V.; Hebert, J. Multifocal infection due to *Mycobacterium intracellulare*: First case of interferon gamma receptor partial dominant deficiency in tropical French territory. *Arch. De Pediatrie Organe Off. De La Soc. Fr. De Pediatrie* **2007**, *14*, 270–272. [[CrossRef](#)]
31. Zhu, J.; Martinez, J.; Huang, X.; Yang, Y. Innate immunity against vaccinia virus is mediated by TLR2 and requires TLR-independent production of IFN- β . *Blood* **2007**, *109*, 619–625. [[CrossRef](#)]
32. Bertolini, F.; Servin, B.; Talenti, A.; Rochat, E.; Kim, E.; Oget, C.; Palhiere, I.; Crisa, A.; Catillo, G.; Steri, R. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genet. Sel. Evol.* **2018**, *50*, 57. [[CrossRef](#)]
33. Weedon, M.N.; Lettre, G.; Freathy, R.M.; Lindgren, C.M.; Voight, B.F.; Perry, J.R.; Elliott, K.S.; Hackett, R.; Guiducci, C.; Shields, B. A common variant of HMG2 is associated with adult and childhood height in the general population. *Nat. Genet.* **2007**, *39*, 1245–1250. [[CrossRef](#)] [[PubMed](#)]
34. Bouwman, A.C.; Daetwyler, H.D.; Chamberlain, A.J.; Ponce, C.H.; Sargolzaei, M.; Schenkel, F.S.; Sahana, G.; Govignon-Gion, A.; Boitard, S.; Dolezal, M. Meta-analysis of genome-wide association studies for cattle stature identifies common genes that regulate body size in mammals. *Nat. Genet.* **2018**, *50*, 362–367. [[CrossRef](#)] [[PubMed](#)]
35. Plassais, J.; Kim, J.; Davis, B.W.; Karyadi, D.M.; Hogan, A.N.; Harris, A.C.; Decker, B.; Parker, H.G.; Ostrander, E.A. Whole genome sequencing of canids reveals genomic regions under selection and variants influencing morphology. *Nat. Commun.* **2019**, *10*, 1–14. [[CrossRef](#)] [[PubMed](#)]
36. Makvandi-Nejad, S.; Hoffman, G.E.; Allen, J.J.; Chu, E.; Gu, E.; Chandler, A.M.; Lored, A.I.; Bellone, R.R.; Mezey, J.G.; Brooks, S.A. Four loci explain 83% of size variation in the horse. *PLoS ONE* **2012**, *7*. [[CrossRef](#)] [[PubMed](#)]
37. Dahlgren, A.; Lundmark, P.; Axelsson, T.; Lind, L.; Syvänen, A.-C. Association of the estrogen receptor 1 (ESR1) gene with body height in adult males from two Swedish population cohorts. *PLoS ONE* **2008**, *3*, e1807. [[CrossRef](#)] [[PubMed](#)]
38. Gunnell, D.; Okasha, M.; Davey Smith, G.; Oliver, S.; Sandhu, J.; Holly, J. Height, leg length, and cancer risk: A systematic review. *Epidemiol. Rev.* **2001**, *23*, 313–342. [[CrossRef](#)]
39. Khankari, N.K.; Shu, X.-O.; Wen, W.; Kraft, P.; Lindström, S.; Peters, U.; Schildkraut, J.; Schumacher, F.; Boffetta, P.; Risch, A. Association between adult height and risk of colorectal, lung, and prostate cancer: Results from meta-analyses of prospective studies and Mendelian randomization analyses. *PLoS Med.* **2016**, *13*, e1002118. [[CrossRef](#)]
40. Brisson, J.; Morrison, A.S.; Kopans, D.B.; Sadowsky, N.L.; Kalisher, L.; Twaddle, J.A.; Meyer, J.E.; Henschke, C.I.; Cole, P. Height and weight, mammographic features of breast tissue, and breast cancer risk. *Am. J. Epidemiol.* **1984**, *119*, 371–381. [[CrossRef](#)]
41. Chen, K.; Liao, Y.; Zhang, J. The major aeroallergens in Guangxi, China. *Clin. Exp. Allergy* **1988**, *18*, 589–596. [[CrossRef](#)]

42. Chou, S.-D.; Prince, T.; Gong, J.; Calderwood, S.K. mTOR is essential for the proteotoxic stress response, HSF1 activation and heat shock protein synthesis. *PLoS ONE* **2012**, *7*, e39679. [[CrossRef](#)]
43. Kakigi, R.; Naito, H.; Ogura, Y.; Kobayashi, H.; Saga, N.; Ichinoseki-Sekine, N.; Yoshihara, T.; Katamoto, S. Heat stress enhances mTOR signaling after resistance exercise in human skeletal muscle. *J. Physiol. Sci.* **2011**, *61*, 131–140. [[CrossRef](#)] [[PubMed](#)]
44. Yoshihara, T.; Naito, H.; Kakigi, R.; Ichinoseki-Sekine, N.; Ogura, Y.; Sugiura, T.; Katamoto, S. Heat stress activates the Akt/mTOR signalling pathway in rat skeletal muscle. *Acta Physiol.* **2013**, *207*, 416–426. [[CrossRef](#)] [[PubMed](#)]
45. Zhang, G.; Wang, Z.; Du, Z.; Zhang, H. mTOR regulates phase separation of PGL granules to modulate their autophagic degradation. *Cell* **2018**, *174*, 1492–1506.e1422. [[CrossRef](#)] [[PubMed](#)]
46. Ning, Q.; Qu, K.; Hanif, Q.; Jia, Y.; Cheng, H.; Zhang, J.; Chen, N.; Chen, H.; Huang, B.; Lei, C. MTOR Variation Related to Heat Resistance of Chinese Cattle. *Animals* **2019**, *9*, 915. [[CrossRef](#)]
47. Adachi, T.; Nakagawa, H.; Chung, I.; Hagiya, Y.; Hoshijima, K.; Noguchi, N.; Kuo, M.T.; Ishikawa, T. Nrf2-dependent and-independent induction of ABC transporters ABCC1, ABCC2, and ABCG2 in HepG2 cells under oxidative stress. *J. Exp. Ther. Oncol.* **2007**, *6*, 335–348.
48. Hankir, M.K.; Kranz, M.; Gnad, T.; Weiner, J.; Wagner, S.; Deuther-Conrad, W.; Bronisch, F.; Steinhoff, K.; Luthardt, J.; Klötting, N. A novel thermoregulatory role for PDE10A in mouse and human adipocytes. *Embo Mol. Med.* **2016**, *8*, 796–812. [[CrossRef](#)]
49. Wang, J.; Xue, X.; Liu, Q.; Zhang, S.; Peng, M.; Zhou, J.; Chen, L.; Fang, F. Effects of duration of thermal stress on growth performance, serum oxidative stress indices, the expression and localization of ABCG2 and mitochondria ROS production of skeletal muscle, small intestine and immune organs in broilers. *J. Therm. Biol.* **2019**, *85*, 102420. [[CrossRef](#)]
50. Belhadj Slimen, I.; Najjar, T.; Ghram, A.; Abdrrabba, M. Heat stress effects on livestock: Molecular, cellular and metabolic aspects, a review. *J. Anim. Physiol. Anim. Nutr.* **2016**, *100*, 401–412. [[CrossRef](#)]
51. Ippolito, D.L.; Lewis, J.A.; Yu, C.; Leon, L.R.; Stallings, J.D. Alteration in circulating metabolites during and after heat stress in the conscious rat: Potential biomarkers of exposure and organ-specific injury. *BMC Physiol.* **2014**, *14*, 14. [[CrossRef](#)]
52. Hou, Y.; Wang, X.; Ping, J.; Lei, Z.; Gao, Y.; Ma, Z.; Jia, C.; Zhang, Z.; Li, X.; Jin, M. Metabonomics approach to assessing the modulatory effects of kisspeptin-10 on liver injury induced by heat stress in rats. *Sci. Rep.* **2017**, *7*, 7020. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).