



Article Genomic Landscape of Copy Number Variations and Their Associations with Climatic Variables in the World's Sheep

Hosein Salehian-Dehkordi ^{1,2,†}[®], Jia-Hui Huang ^{1,†}, Nasrollah Pirany ²[®], Hossein Mehrban ²[®], Xiao-Yang Lv ^{3,4}, Wei Sun ^{3,4}, Ali Esmailizadeh ^{5,*}[®] and Feng-Hua Lv ^{1,*}[®]

- ¹ College of Animal Science and Technology, China Agricultural University, Beijing 100193, China; hosein_salehi6@yahoo.com (H.S.-D.); s20213040574@cau.edu.cn (J.-H.H.)
- ² Department of Animal Science, Faculty of Agriculture, Shahrekord University, Shahrekord 88186-34141, Iran; napirany@yahoo.com (N.P.); hosseinmehrban@gmail.com (H.M.)
- ³ Joint International Research Laboratory of Agriculture and Agri-Product Safety of Ministry of Education of China, Yangzhou University, Yangzhou 225009, China; dx120170085@yzu.edu.cn (X.-Y.L.); dkxmsunwei@163.com (W.S.)
- ⁴ International Joint Research Laboratory in Universities of Jiangsu Province of China for Domestic Animal Germplasm Resources and Genetic Improvement, Yangzhou University, Yangzhou 225009, China
- ⁵ Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran
- * Correspondence: aliesmaili@uk.ac.ir (A.E.); lvfenghua@cau.edu.cn (F.-H.L.)
- + These authors contributed equally to this work.

Abstract: Sheep show characteristics of phenotypic diversity and adaptation to diverse climatic regions. Previous studies indicated associations between copy number variations (CNVs) and climatedriven adaptive evolution in humans and other domestic animals. Here, we constructed a genomic landscape of CNVs (n = 39,145) in 47 old autochthonous populations genotyped at a set of high-density (600 K) SNPs to detect environment-driven signatures of CNVs using a multivariate regression model. We found 136 deletions and 52 duplications that were significantly (P_{adi} < 0.05) associated with climatic variables. These climate-mediated selective CNVs are involved in functional candidate genes for heat stress and cold climate adaptation (e.g., B3GNTL1, UBE2L3, and TRAF2), coat and woolrelated traits (e.g., TMEM9, STRA6, RASGRP2, and PLA2G3), repairing damaged DNA (e.g., HTT), GTPase activity (e.g., COPG), fast metabolism (e.g., LMF2 and LPIN3), fertility and reproduction (e.g., SLC19A1 and CCDC155), growth-related traits (e.g., ADRM1 and IGFALS), and immune response (e.g., *BEGAIN* and *RNF121*) in sheep. In particular, we identified significant ($P_{adj.} < 0.05$) associations between probes in deleted/duplicated CNVs and solar radiation. Enrichment analysis of the gene sets among all the CNVs revealed significant ($P_{adj.} < 0.05$) enriched gene ontology terms and pathways related to functions such as nucleotide, protein complex, and GTPase activity. Additionally, we observed overlapping between the CNVs and 140 known sheep QTLs. Our findings imply that CNVs can serve as genomic markers for the selection of sheep adapted to specific climatic conditions.

Keywords: sheep; CNVs; climate adaptation; association tests; solar radiation

1. Introduction

The earliest domestic sheep were domesticated in the Fertile Crescent [1,2]. Following domestication, they diffused worldwide, and various breeds with diverse phenotypes have been developed under long-term artificial and natural selection in response to human demands and climate change [3–6]. Climate-mediated genetic variations in the genome indicate that environmental factors, such as sunshine, temperature, elevation, and humidity, have influenced the spatial distribution of phenotypic and genetic variation across populations [5,7–10].

Copy number variations (CNVs), complementary to single nucleotide polymorphisms (SNPs) [11], are a major source of variations caused by deletions and duplications. They



Citation: Salehian-Dehkordi, H.; Huang, J.-H.; Pirany, N.; Mehrban, H.; Lv, X.-Y.; Sun, W.; Esmailizadeh, A.; Lv, F.-H. Genomic Landscape of Copy Number Variations and Their Associations with Climatic Variables in the World's Sheep. *Genes* **2023**, *14*, 1256. https://doi.org/10.3390/ genes14061256

Academic Editor: Shaobin Li

Received: 21 April 2023 Revised: 6 June 2023 Accepted: 8 June 2023 Published: 13 June 2023



Copyright: © 2023 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 (https:// creativecommons.org/licenses/by/ 4.0/). account for 4.8–9.5% of the whole genome and contribute to the variability in the genome among individuals [12,13]. Copy number variants (CNVs) can generate meiotically and somatically [14] and play an essential role in the rapid evolution process by changing the expression levels of genes with variable copy numbers [15,16] and dosage [17,18]. CNV of the *ASIP* gene regulates goats' white and grey coat phenotypes [19]. Duplication in the flanking of *KIT* gene is associated with a white coat color in pigs [20]. CNVs identified by genome-wide association studies and selective sweep analyses in sheep have been related to complex traits, such as follicular development and fertility, milk production, wool production, adipogenesis, spleen size, and oxygenated red blood cells [21,22].

Climate-mediated selective pressures impact species distributions, phenotypic variation, and allele frequencies [23]. However, various climate factors, such as temperature, precipitation, altitude, and sunlight, may have distinct impacts on the level of genetic variation among species, influencing their potential for genetic adaptation [5,24]. For example, signatures of local genetic adaptation in genes are related to GTPase activities and energy metabolism caused by mutagenic factors sunlight [5], and positive selection genes with hypoxia responses were enriched in O_2/CO_2 exchange and HIF-1 signaling pathways by increasing the supply of O_2 or reducing the harm of low oxygen saturation level in sheep [8]. Recently, solar ultraviolet (UV) radiation, as a highly mutagenic factor, can hinder DNA replication and transcription with photo-dimers, oxidative DNA lesions, and DNA singlestrand breaks [24], and has been uncovered to influence variations of CNVs in frequency and structure due to sequence permutations in the molecular evolution of DNA [24–26]. Environmental factors, such as vinclozolin, contribute to the transgenerational inheritance of epigenetic modifications in sperm, leading to genome instability. This instability further promotes the acquisition of genetic copy number variations (CNV) in later generations [27].

It was demonstrated that CNVs are involved in adaptive evolution under various climates within diverse species, such as humans [28], cattle [29], goats [30], horses [31], and dogs [32]. CNVs of AMY1 and AMY2B genes in humans and dogs are associated with adapting to the digestion of starchy foods [33,34]. TAS2R16 is involved in gustation, and CNVS of the genes in cattle may be related to adapting to food resources and vegetation diversity [35]. CNVs in goats enriched in the IMAP family genes may play a role in adaptation to hash climates by regulating metabolism [36]. CNVs identified in Tibetan horses overlapped with genes (e.g., CYP4A11, CYP4X1, EIF2AK1, CYP2C18, CYP4F22, NOS2, and *CYP4B1*), which may account for adaptation to high altitude environment [37]. Climate change poses a significant threat to animal husbandry and food safety. A great deal of evidence illustrates that organisms can respond to climate change through phenotypic plasticity and evolutionary adaptation [38]. However, the challenge is identifying variations contributing to climate adaptation across the genome million variations. Nevertheless, a study has yet to be conducted to investigate the associations between CNVs and climatic variables during the post-domestication expansion of sheep. In recent years, the availability of the sheep reference genome of *Ovis aries* (Oar_v4.0) and a comprehensive set of high-density (600 K) single nucleotide polymorphisms (SNPs) provided the opportunity to investigate the contribution of CNVs to rapid local climatic adaptation in autochthonous sheep populations [5,9,39]. Here, we performed unified co-analyses of CNVs, and environmental variables were generated from a global climatic data set of 117 parameters over 40 years in 47 autochthonous sheep populations. We examined the impact of solar radiation on the genomic distribution of CNVs in sheep of different geographic origins.

2. Materials and Methods

2.1. Data Collecting, CNV Calling, and Quality Controls

We collected the Ovine Infinium HD (600 K) SNP Bead Chip data from previous studies [9,39]. We upgraded all the positions in the SNP BeadChip based on the *O. aries* reference genome Oar_v.4.0 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000298735.2, last accessed on 24 September 2020). To exclude poor-quality DNA samples and SNPs, we only used SNPs that passed the filtering procedure described in Salehian-Dehkordi et al. [39]. Additionally, any SNPs on the X and Y chromosomes were excluded. After quality control and filtering, 47 worldwide autochthonous populations, including 695 individuals (Supplementary Table S1 and Figure 1A), were selected for the following analyses. We detected CNVs and CNVRs (copy number variable regions) using the approach described previously [39]. In summary, we first retrieved each sample's signal intensity values (log R ratio: LRR; B allele frequency: BAF) by the Illumina GenomeStudio v1.0. Second, the B allele matrix population frequency (PFB) was calculated with the script compile_pfb.pl using PennCNV v1.0.53 [40]. Third, we detected CNVs with the same parameters as Salehian-Dehkordi et al. [39]. Finally, the raw CNVs were filtered with the following criteria: (i) individual with a standard deviation of LRR < 0.3; (ii) BAF drift of individual < 0.01; (iii) waviness factor of individual < 0.05; (iv) CNVs with more than 3 consecutive SNPs and at least 1 kb length; (v) CNVs with call count > 100 per sample, which represents poor DNA quality, and relevant individuals were removed. We merged all CNVs into CNVRs using bedtools version 2.30.0 (https://bedtools.readthedocs.io/en/latest/index.html, accessed on 23 January 2021) with the "bedtools merge" option [41].



Figure 1. Population genetic structure. (**A**) Geographic origin of 47 worldwide autochthonous sheep populations. (**B**) Population genetic structure of 47 autochthonous sheep populations (K = 2-10). (**C**) Principal component analysis (PCA) of 47 autochthonous sheep populations. The brown, red, gray, and black symbols represent Europe, Africa, Eastern-Central Asian, and Western Asian populations.

2.2. Environmental Data

We accessed climatic data covering a period of 40 years (1961 to 2001) from the global climate data set (http://www.cru.uea.ac.uk/data, last accessed on 24 September 2021) (The Climatic Research Unit, Norwich [42]). The climatic data contained yearly and monthly means of nine parameters, including (i) average daily duration of bright sunshine in percent

(sunshine fraction, SUN); (ii) mean diurnal temperature range in °C (DTR); (iii) relative humidity in percentage (REH); (iv) precipitation in mm/month (PR); (v) mean temperature in °C (TMP); (vi) the coefficient of variation of monthly precipitation in percent (PRCV); (vii) number of days with ground frost (FRS); (viii) 10 m wind speed in m/s (WND); and (ix) number of days with >0.1 mm rain per month (RDO) (Supplementary Table S1). In total, 117 environmental variables and elevation for the geographic origins of 47 autochthonous populations were obtained based on the Latitude (°N) and Longitude (°E) coordinates (Supplementary Table S1, Supplementary Material online, in Cao et al. 2021 [9]).

2.3. Estimation of Solar Irradiation

We have extracted the solar radiation from the sunshine fraction (SUN variable) using the relationship between sunshine and radiation described in Suehrcke et al. [43]. We first obtained annual average sunshine fractions for the locations of interest from Cao et al. [9]. Next, the annual average extraterrestrial radiation was calculated from average monthly values for each location (see Supplementary Tables S2 and S3) based on the following equation [44]:

$$\overline{H}_{o} = \frac{86,400 G_{sc}}{\pi} \left[1 + 0.033 \cos\left(2\pi \frac{n}{365}\right) \right] \times \left(\cos\Phi\cos\Delta\sin\omega_{s} + \omega_{s}\sin\Phi\sin\Delta\right)$$

where \overline{H}_o is the monthly average of extraterrestrial daily solar radiation on the earth's surface without atmosphere (extra-terrestrial radiation) [MJ/m²/day], G_{sc} is the solar constant (the latest suggested value is 1361 W/m² [45]), π is approximately 3.14159, n is an average day of month, Φ is latitude (North +, South –), Δ is declination: $\Delta = (23.45 \frac{\pi}{180}) \sin(2\pi \frac{284+n}{365})$ [44], and ω_s is the solar hour angle at sunset: $\omega_s = \arccos(-\tan \Phi \times \tan \Delta)$.

After determining the annual sunshine fraction and extraterrestrial radiation, the annual average clearness index \overline{K} was computed to estimate the annual average solar radiation in geographic origins of autochthonous sheep breeds. \overline{K} can be expressed as:

$$\overline{K} = \overline{K}_{clear}(\beta + (1 - \beta) \times S^{\gamma})$$

where \overline{K} is the annual clearness index, \overline{K}_{clear} is the approximate world average clear sky clearness index value (0.7191), β and γ are constants 0.1930 and 0.7283, respectively, and S is the sunshine fraction for the different sampling locations from Cao et al. [9] (see Supplementary Table S1). Finally, we calculated the annual average daily solar radiation (\overline{H}) for each sampling site using the equation $\overline{H} = \overline{K} \times \overline{H}_o$ (see Supplementary Table S3).

2.4. Testing for Genomic Signatures Associated with Climate Variables

Two approaches were used to detect the signatures of local adaptation. We first used multiple univariate logistic regression models [46] to detect genotype–environment associations under individual-based analysis. We analyzed using the Samβada program (https://lasig.epfl.ch/sambada, last accessed 24 September 2020) by incorporating population structure to decrease the occurrences of false genotype–climate relationships under the "*BEST*" option. For the analysis, a specific CNV state of all related samples was marked as either present or absent (i.e., binary information: 1 or 0). Correlations between all possible copies of CNV events and climatic variables were estimated across the sampling locations, and only significant models by Bonferroni correction based on the Wald statistic (p < 0.01) were considered.

We further applied latent factor mixed models to detect associations between CNVs and environmental variables based on population genomic statistics [47]. The methods incorporated fixed, random, and latent effects (i.e., demographic history of autochthonous breeds and isolation-by-distance patterns) to lower the risk of false-positive associations in landscape genomics [47]. We first summarized all the climatic factors and elevation using the first axis of the PCA to reduce the dimensionality of multivariate data into two columns

of principal components with minimal loss of information. We used the ade4 package to perform the PCA analysis for individuals/variables. Following that, the latent factor (K = 4) was identified based on the distribution of populations in structure analyses using Structure v2.3.4 [48] and SmartPCA [49]. For each CNV, we calculated the latent score (Z scores) based on the least-squares estimation approach [50]. A threshold (|Z| scores \geq 10) was used to identify significant CNV effects at a level of $p < 10^{-6}$ after Benjamini-Hochberg correction for multiple testing with 5% type I error [47,50]. We conducted the analysis using latent factor mixed models version 2 (LFMM2) (https://rdrr.io/bioc/LEA/man/lfmm2.html, accessed on 13 September 2022), as implemented in the R package LEA [50].

2.5. Testing for Genomic Signatures Associated with Solar Radiation-Mediated Selective Pressure

Using probe-based statistics, we performed genome-wide association analyses of CNVs with solar radiation variables. CNVs and quality measures were converted to matrices containing probes for deletions and duplications separately using ParseCNV v2 [51] under the parameter "*-includePed*". We implemented two approaches to perform probe-based association regression tests. First, we calculated the probe-based statistical significance of neighboring SNPs using the EMMAX program (http://csg.sph.umich.edu//kang/emmax/download/index.html, accessed on 7 March 2010) [52]. We constructed kinship matrices of BN (Balding-Nichols) using the EMMAX algorithm and the script "*emmax-kin-intel64*". The association between probe-based statistics for CNV occurrence and solar radiation was computed using mixed linear models with the "*emmax-intel64*" argument. To correct for population variation, the PCs as covariates were implemented simultaneously in the model using "*-c*" and "*cov_file*" options. A suggestive threshold ($p < 5 \times 10^{-4}$) was considered for probe-based statistical significance of CNV occurrence [51], and *p* values were computed using the EMMAX method.

Next, we tested the associations between probes of deleted/duplicated CNVs and solar irradiance using the PLINK v1.90 in the ParseCNV environment. Multi-covariate association analyses were applied for the logistic regression models using the codes "-*logistic*" and "-*adjust*" arguments, also "-*covar*" option was run to implement PCA population stratification components. Finally, we used the script "*InsertPlinkPvalues.pl*" in the ParseCNV program to merge adjacent probe-based CNV occurrence *p*-values into significant CNVRs. The code "-*permuteP* < 10,000 >" was run with 10,000 permutations in the algorithm. We implemented a significance threshold of max (T) adjusted (p < 0.05) after multiple testing corrections to define CNVRs and therefore reported associations. To evaluate significant CNVRs for confidence, the program ParseCNV amassed red flags based on various factors referenced in the University of California Santa Cruz (UCSC) browser. These characteristics help to reduce false positive CNV detection based on the criteria such as segmental duplications, the database of genomic variants, centromere/telomere, GC base content, density and count of probes, length of the CNVs, population frequency, peninsula, and inflated (Supplementary Figure S2) [51].

2.6. Gene Annotation and Overlapping with QTLs

We annotated the gene content associated with the candidate CNV and CNV regions using the sheep reference genome Oar_v.4.0. Only genes that overlapped with a CNVR spanning at least 10% of the CNV length were considered. We implemented Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses using the DAVID (database for annotation, visualization, and integrated discovery) [53]. Categories with adjusted *p* value < 0.05 after the Bonferroni correction and at least 12 genes were defined as significantly enriched GO terms and KEGG pathways.

We further annotated the function of the candidate CNV and CNV regions with quantitative trait loci (QTLs). We identified overlapping regions among sheep QTLs and the candidate CNV and CNV regions. The overlapping regions were detected using Bedtools v 2.30.0 under the setting: (i) QTLs with confidence interval < 5 Mb; (ii) CNVR spanning at least 50% of the CNV length [54].

3. Results

3.1. CNV Detection and Population Differential Analyses

In total, 129,446 raw CNVs were identified from 695 samples by PennCNV v1.0.53. After quality controls, 39,145 high-quality CNVs were obtained and merged into 4769 CNVRs based on the overlapping region among CNVs (Table 1, Supplementary Table S4, Figure 2). To determine the accuracy of CNVs, we validated all CNVs by comparing CNVs identified in this study with CNVs in Salehian-Dehkordi et al. [39]. The genotyping concordance was 57.34% (Supplementary Figure S1). We used two approaches (e.g., Structure and PCA analyses) to uncover population features based on CNV frequencies. We observed four genetic clusters corresponding to four geographic regions (Figure 1B,C). All the 47 native breeds were clustered into four groups based on genetic and geographic features: (i) 11 populations (n = 191) from Eastern and Central Asia; (ii) 11 populations (n = 262) from Europe (Figure 1).

Table 1. Summary of CNVs and CNVRs identified in old autochthonous sheep populations.

Туре	CNVs					CNVRs				Unique CNVs			
	Homozygous Deletion	Hemizygous Deletion	Duplication	Biallelic Triplication	Total	Loss	Gain	Mixed	Total	Deletion	Duplication	Mixed	Total
Count Length (Mb)	4726	23,298	11,018	103	39,145	3468	861	440	4769	11,119	4118	510	15,747
	94.4	801.2	460.9	3.2	1359.7	150	23.4	80.9	254	492.6	213.6	14.2	720.4



Figure 2. CNVR distribution in autochthonous sheep populations on each chromosome.

3.2. Function Annotation of CNVRs

A total number of 2892 genes were annotated within 4769 CNVRs in 47 autochthonous sheep populations (Supplementary Table S5). We further investigated the functions of those genes by two methods. Firstly, we used GO term and KEGG pathway enrichment analyses. We detected nine significant ($P_{adj.} < 0.05$) Gene Ontology categories (Supplementary Table S6), including protein complexes in cytoplasm (GO:0005829, $p = 1.91 \times 10^{-8}$ and GO:0005737, $p = 4.4 \times 10^{-5}$), collagen trimer (GO:0005581, $p = 5.31 \times 10^{-5}$), nucleoplasm (GO:0005654, p = 0.01268), ATP binding (GO:0005524, $p = 3.05 \times 10^{-9}$), GTPase activator activity (GO:0005524, $p = 1.71 \times 10^{-5}$), microtubule binding (GO:0008017, $p = 6.5 \times 10^{-4}$), calcium ion binding (GO:0005509, $p = 9.51 \times 10^{-4}$), and guanyl-nucleotide exchange factor activity (GO:0005085, p = 0.0047). Eleven categories of KEGG with the threshold of adjusted P_{adj} value < 0.05 after the Bonferroni correction

were identified, such as axon guidance (oas04360, $p = 5.07 \times 10^{-6}$), focal adhesion (oas04510, $p = 5.00 \times 10^{-4}$), and calcium signaling pathway (oas04020, p = 0.038) (Supplementary Table S6). Secondly, we annotated the 4769 CNVRs and detected 80 unique CNVRs overlapped with 140 QTLs (Supplementary Tables S7 and S8). Of these QTLs, 25 overlapped QTLs for the disease resistance traits, six for reproductive traits, 38 for body size and meat and growth-related traits, 12 for milk-related traits, and four for wool-related traits.

3.3. Climate-Driven Candidate Selective Signatures Testing

The Biplot showed PC1 versus PC2 for the 47 autochthonous sheep populations and the climatic parameters of their geographic origins. All groups of individuals (Figure 3A) and climatic variables (Figure 2B) were mapped using their contributions. The colors show qualities of representation for variables and populations on the maps.



Figure 3. PCA of environmental variables. **(A)** Projection of the nine climatic variables based on 47 worldwide autochthonous populations. The representation of the variable is quantified by cos2, and a high cos2 indicates a good representation of the variable on the principal component. **(B)** Projection of the nine climatic variables on the first and second factor planes. SUN: percent of maximum possible sunshine; DTR: mean diurnal temperature range in °C; REH: relative humidity in percentage; PR: precipitation in mm/month; TMP: mean temperature in °C; PRCV: the coefficient of variation of monthly precipitation in percent; FRS: number of days with ground frost; WND: 10 m wind speed in m/s; RDO: number of days with >0.1 mm rain per month. Contributions of variables to PCs are shown by the colors bar.

We conducted the analysis by Samβada program and selected the top models according to the Wald statistics (p < 0.01 and Wald score > 34), and 377 out of a total of 4,579,965 univariate models (39,145 CNV genotype × 117 climatic variables) were selected. Thirty-four genes were annotated within 377 CNVs, and eight genes from 33 unique CNVs were associated with environmental variables (Supplementary Table S9), such as *ATHL1*, *NLRP6*, *IFITM5* for PR, and *GUCY1A2* for PRCV (Supplementary Table S10). Next, based on the z score in the latent factor mixed models (LFMMs), we detected 155 CNVs with |z| scores ≥ 10 (Supplementary Table S11) that were highly significant ($3.63 \times 10^{-6} \geq p$ values $\geq 1.34 \times 10^{-103}$), associated with environmental parameters in 47 old sheep populations. We then found genes that overlapped CNVs (e.g., *B3GNTL1*, *UBE2L3*, *TRAF2*, *GTF2F1*, and *IGFALS*), which were significantly associated with climatic variables (Supplementary Tables S10 and S12).

3.4. Solar Radiation-Driven Candidate Signatures Test

We implemented association tests for solar radiation variables using the linear mixed models in the EMMAX program. To detect associated deletions and duplications, we found 54,597 and 13,925 probes among deletion and duplication regions for subsequent association tests, respectively. Among the identified probe-based statistics, we found 86 and 45 probes for deletions and duplications, respectively, based on genome-wide CNV significant threshold ($p < 5 \times 10^{-4}$) (see 5th columns in Supplementary Tables S13 and S14). Next, we detected significant associations through multi-covariate association analyses using the logistic regression models. We obtained the top significant ($p < 5 \times 10^{-5}$) SNP probes for deletions (180) and duplications (125) genomic regions among the full set of probes (see 9th columns in Supplementary Tables S13 and S14). We identified common significant probes detected at least by two approaches (EMMAX and ParseCNV-Plink). After Bonferroni correction in duplicated regions, seven significant probes on chromosome 11 with $p \le 6.08 \times 10^{-5}$ were associated with solar irradiance (Figure 4A). In deleted regions, three and eight probes on chromosomes 15 and 4 ($p \le 6.57 \times 10^{-8}$) were significantly associated with solar radiation, respectively (Figure 4B).



Figure 4. Manhattan plot of GWAS between probes in deleted/duplicated CNVs and solar radiation. (**A**) Manhattan plot based on duplicated CNVs; the horizontal dashed lines correspond to the genome-wide significance thresholds ($P_{adj.} = 6.08 \times 10^{-5}$). (**B**) Manhattan plot based on deleted CNVs; the horizontal dashed lines correspond to the genome-wide significance thresholds ($P_{adj.} = 6.57 \times 10^{-8}$).

Based on the probe-based statistics for CNV occurrence detected by two approaches, we identified associations between CNVRs and solar irradiances. By applying the program ParseCNV, noisy CNVs were removed, and normalized results were implemented in the association test (Supplementary Figure S2B). Afterward, 35 significant CNVRs (adjusted p < 0.05) were associated with solar radiation (Supplementary Table S15), and some of them overlapped with functional genes (Table 2 and Supplementary Table S16). We found that CNVs and CNVRs were commonly selected by statistical tests (Supplementary Table S17).

Significant CNVRs	Count SNPs	Gene ID
chr16:60477356-60501528	9	LOC101104428, LOC101120862
chr16:69228726-70108247	26	IRX1, LOC105602651, LOC105602652, LOC105608090
chr26:399178-402542	4	DLGAP2
		LOC105610048, LOC105616624, LOC105610045, LOC105610042,
		LOC105610043, LOC106991502, RABIF, LOC106991503,
chr12:77533856-78344379	14	LOC106991504, UBE2T, C12H1orf106, LOC105610049, CACNA1S,
		TMEM9, IGFN1, PKP1, LOC101109820, KLHL12, LOC105616599,
		SYT2, LOC105616625, LOC101110351, LOC101110859, KIF21B
		LOC105603604, GATA2, DNAJB8, LOC105603605, LOC105603606,
		SEC61A1, LOC106991760, LOC105603608, LOC105603609,
chr19:58987471-60169432	29	LOC105603631, LOC106991788, RPN1, EEFSEC, RUVBL1 , MGLL,
		ZXDC, SLC41A3 , ALDH1L1, MCM2, KBTBD12, PLXNA1,
		KLF15 , RAB7A
		SLC22A12, MEN1, PPP2R5B, CDC42EP2, GPHA2, ZFPL1, TMEM262,
		TM7SF2, ZNHIT2, FAU, MRPL49, SYVN1, SPDYC, LOC101107147,
		TIGD3, SLC25A45, SLC22A11, LOC105604109, RASGRP2,
chr21:41984078-42773808	4	LOC105604110, MAP4K2, CDC42BPG, EHD1, C21H11orf85, BATF2,
		NAALADL1, CDCA5, VPS51, LOC101105958, LOC105604111,
		SLC22A20, POLA2, DPF2, NRXN2, SF1 , ATG2A, SNX15, SAC3D1,
		ARL2, CAPN1, MIR194, PYGM
		LOC105604138, LOC101118574, PHLDA2 , LOC105604139, DHCR7,
chr21:46919932-47922324	17	NADSYN1, OSBPL5, CARS, NAP1L4, SLC22A18, LOC101118066,
		SHANK2, CDKN1C
chr6:114803668-114941025	22	MSANTD1, RGS12, HTT
chr14:55268932-55284965	7	IZUMO2, LOC101115/29
chr17:70557208-70575800	7	LOC106991688, GST12B, LOC10111139/, LOC101118990
chr13:53886521-53904177	9	SLCO4A1
chr16:31746166-31765064	6	
chr1:262818152-263558472	31	LOC105604794, POFUT2, SLC19A1, PCBP3, ADARB1, COL18A1
chr14:48269630-48282740	4	
chr22:49712254-50377777	25	LUC105604373, PWWP2B, LUC105606161, LUC105604355, JAKMIP3,
-h2.127220004 127274222	0	DP I SL4, LOC101109500, ST K32C, INPP5A, LOC10111028/
cnr5:13/339004-13/3/4322	ð	LUC101123028, LUC10112328/ DEC AIN
CNI18:64006157-64006658	3	BEGAIN

Table 2. The significant CNVRs and candidate genes associated with solar radiation under adaptive selection.

4. Discussion

High-density SNP data have been used to detect CNVs in livestock [21,55,56]. The genome-wide testing of CNVs for phenotypic trait variations was reported previously, but climatic adaptation involving CNVs still needs to be explored. This study showed the results of CNV genotypes called by the high-density SNP. We performed association tests between CNVs and climatic variables and detected candidate genes for environment-driven genetic adaptation. Unlike our previous study, which used only SNPs to find associations between alleles and climates [9], we applied ecological tests to find association signals between CNV states and environmental variables for the first time. We observed a set of novel candidate genes which overlapped with CNVs. They were significantly correlated with local climatic adaptation.

We found 39,145 high-quality CNVs, and the genotype of each CNV was confirmed by being compared with previous results in Salehian-Dehkordi et al. [39] (see Supplementary Tables S3 and S7 in Salehian-Dehkordi et al. [39]). The results showed a high level of genotypic concordance between them (57.34%) (Supplementary Figure S1). We found that the frequency of hemizygous (59.52%) and homozygous deletions (12.07%) was more than duplications (28.15%) and monoallelic triplications (0.26%) (Table 1 and Supplementary Table S4), and the results agree with previous reports [57–60]. As an essential source of polymorphism in genomes, CNVs might account for phenotypic variations [61,62]. Here, we found 2892 genes within CNVRs (Supplementary Table S5). Functional enrichment analysis showed that genes in CNVRs are associated with physiological functions, energy metabolism, nervous system, immunity, and phenotype (Supplementary Table S6). Of the genes, *MLPH* located in chr1:3417932-3436288 encodes the melanophilin carrier protein involved in pigmentation in the hair and skin (Supplementary Table S5) [63]. Interestingly, a hemizygous deletion (chr1:3.112.486-3122600) neared downstream of the mentioned CNVR was detected by the LFMM analysis. We also found other genes related to pigmentation or defense against pathogens, such as *TYR*, *LPO*, and *REN*, which may be involved in local adaptation [64–66].

Our results suggested that autochthonous sheep breeds' adaptation to climates could be partly ascribed to CNVs, as found in organisms such as cattle [67] and balsam poplar [57]. The multiple univariate logistic regression analyses found three critical functional genes on chr21:49691288-49714065 (Supplementary Table S10). CNV harboring functional candidate gene (*OAS2*) with a high Wald score was associated with fat deposition in sheep breeds [68]. The tail fat might benefit adaptation to extreme environments and harsh seasons [69]. These results suggested that the genes with different numbers of copies are under climatemediated selective pressures.

The LFMM analysis further supported the idea that CNVs mediate the genetic response of autochthonous sheep breeds to climates. We explored the involvement of hemizygous deletions resulting from evolutionary climatic adaptation pressure (Supplementary Table S11). Those hemizygous deletions harbored functional genes, such as *B3GNTL1*, *UBE2L3*, *TRAF2*, *TF2F1*, and *IGFALS*. *B3GNTL1* gene identified here has been reported to regulate heat stress adaptation in Egyptian sheep breeds [70]. Additionally, ubiquitin-conjugating enzyme E2-L3 (*UBE2L3*) was involved in response to heat stress in Indian cattle [71]. *TRAF2* and *GTF2F1* (a duplicated gene) genes were identified in Russian sheep breeds, Chinese cattle, and Datong yak populations, indicating their associations with adaptation to high altitudes and cold climates [72–74]. Together, these results suggested the involvement of CNVs in generic and environmental adaptation.

We found two common CNVRs on chromosome 5 identified by LFMM and Samβada analysis (Supplementary Table S18). Four common significant probes were detected in climatic and solar radiation-driven selective pressure testing (Supplementary Table S18). The result indicates potential complex and specific genetic mechanisms responding to environmental factors.

To find significant associations between CNV-associated genes and solar radiation, we merged probe-based statistics from CNV calls into CNVRs (Supplementary Table S16). Interestingly, we identified candidate genes such as *HTT* that play an important role in repairing damaged DNA [75], *RUVBL1* for heat and parasite stress [76], *SF1* and *PHLDA2* for litter size and reproduction [29], *SLC41A3* for immune responses [77], and *KLF15* for growth-related traits [78] (Supplementary Table S17). *HTT* is important in repairing DNA damage created by solar and UV radiation. In addition, a cascade of physiological events will be initiated when an animal is exposed to solar radiation [79,80]. For example, an endogenous circannual rhythm driven and synchronized by the annual photoperiod cycle regulates the breeding season in sheep [81,82]. UV radiation and climatic factors, including temperature and humidity, can also affect the spread of pathogens in mammals [83,84].

Many selected CNV-overlapping genes via ecological tests were associated with production traits and immunity (Supplementary Table S19), suggesting that these CNV-associated genes might be signatures of natural and artificial selections in their adaptation to extreme climates. Our results supported the hypothesis of the 119 selective functional genes associated with climate-mediated selection—21 genes (e.g., *STRA6*, *TMEM9*, and *PLA2G3*) were responsible for coat and wool, 20 genes (e.g., *B3GNTL1*, *TRAF2*, and *GUCY1A2*) for environmental condition and stress, 15 genes (e.g., *LMF2*, *LPIN3*, *COPG1*, and *OAS2*) for metabolism, including GTPase and lipid metabolism activities, 14 genes (e.g., *NPBWR2*, *OAS1*, and *SLC19A1*) for fertility and reproduction, 18 genes (e.g., *SFTPD*, *RNF121*, and

TMEM154) for immunity, cancer, and disorders, and 28 genes (e.g., *KLF15*, *CACNA1S*, and *CCDC152*) for growth and production traits (Supplementary Table S19). We observed some candidate genes (e.g., *LTN1*, *TRAF2*, *COPG1*, and *SLC19A1*) overlapped with candidate CNVs in sheep detected previously [74,85,86]. Additionally, some functional genes were previously detected as related to climate-mediated livestock adaptation [5,9].

Previous reports showed the role of important CNVs for specific traits such as coat color [87]. A considerable number of selective genes from association tests, by the LFMM approach, were associated with coat and wool traits in livestock (e.g., *MEGF6, STRA6, LTN1, FAM83G*, and *TRAF2*) (Supplementary Table S12). These genes were under natural selection due to the high-altitude adaptation or the intensive UV and solar radiations under harsh environments, such as the Tibetan plateau [88].

Genotype-by-environment interactions for production traits in beef cattle studies indicated the role of the biological system (e.g., vasodilation, metabolism, and nervous) in the genetic sensitivity to environmental stress [89,90]. In agreement with previous studies, we found many significant CNVs that overlapped genes (e.g., *PCDH15*, *MAPK1*, *DNAJB8*, and *SHANK2*) related to heat stress and metabolism (Supplementary Table S19). Of the common CNVs and CNVRs obtained from more than one association test (Supplementary Table S18), *INPP5A* and *KLF15* were related to heat stress and production traits, respectively [78,91]. Our results indicated that genotype-by-environment interactions could have contributed to copy number status during adaptation to different environmental stress.

5. Conclusions

This is one of the first comprehensive CNV studies to reveal novel associations with environmental variables. Overall, 39,145 high-quality CNVs were identified in the world-wide sheep populations, which harbored 2892 genes associated with phenotypic traits and climatic-mediated adaptive variations. We demonstrated evidence for the climate-mediated genes in CNVRs. This study indicated the genes with different numbers of copies retrieved from natural selection for specific traits such as heat stress, cold adaptation, and metabolism. CNV harbors specific functional candidate genes such as *B3GNTL1*, *UBE2L3*, *SHANK2*, *COPG1*, *TRAF2*, and *GTF2F1* for heat stress and cold climate adaptation *LTN1*, *STRA6*, *RASGRP2*, and *HTT* for repairing damaged DNA. This study generated a CNV map and revealed important candidate genes associated with climate adaptation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes14061256/s1.

Author Contributions: H.S.-D., A.E. and F.-H.L. conceived and designed the project. H.S.-D., N.P., H.M., X.-Y.L. and W.S. collected the datasets. H.S.-D. and J.-H.H. analyzed the data. H.S.-D. and J.-H.H. wrote the manuscript with contributions from F.-H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported by grants from the National Natural Science Foundation of China (Nos. 32061133010, 31972527, and U21A20246), the National Key Research and Development Program-Key Projects (2021YFD1200900 and 2021YFD1300904) and the grant of high quality economic and social development in southern Xinjiang (NFS2101).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Acknowledgments: We thank Saeid Karimi-Dehkordi (Department of Animal Science, Faculty of Agriculture, Shahrekord University, Iran) and vice presidency for science and technology of Iran for their support and providing appropriate environment to perform some analyses. We are grateful to thank Harry Suehrcke (Sunspin Pty Ltd., Australia) for assisting in part of estimation of solar irradiation. We thank Meng-Hua Li for his helpful comments and revisions.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Zeder, M.A. Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 11597–11604. [CrossRef] [PubMed]
- Driscoll, C.A.; Macdonald, D.W.; O'Brien, S.J. From wild animals to domestic pets, an evolutionary view of domestication. *Proc. Natl. Acad. Sci. USA* 2009, 106, 9971–9978. [CrossRef] [PubMed]
- Tapio, M.; Marzanov, N.; Ozerov, M.; Ćinkulov, M.; Gonzarenko, G.; Kiselyova, T.; Murawski, M.; Viinalass, H.; Kantanen, J. Sheep Mitochondrial DNA Variation in European, Caucasian, and Central Asian Areas. *Mol. Biol. Evol.* 2006, 23, 1776–1783. [CrossRef] [PubMed]
- 4. Chessa, B.; Pereira, F.; Arnaud, F.; Amorim, A.; Goyache, F.; Mainland, I.; Kao, R.R.; Pemberton, J.M.; Beraldi, D.; Stear, M.J.; et al. Revealing the History of Sheep Domestication Using Retrovirus Integrations. *Science* **2009**, *324*, 532–536. [CrossRef]
- Lv, F.H.; Agha, S.; Kantanen, J.; Colli, L.; Stucki, S.; Kijas, J.W.; Joost, S.; Li, M.-H.; Marsan, P.A. Adaptations to Climate-Mediated Selective Pressures in Sheep. *Mol. Biol. Evol.* 2014, *31*, 3324–3343. [CrossRef]
- Lv, F.H.; Peng, W.F.; Yang, J.; Zhao, Y.X.; Li, W.R.; Liu, M.J.; Ma, Y.H.; Zhao, Q.J.; Yang, G.L.; Wang, F.; et al. Mitogenomic Meta-Analysis Identifies Two Phases of Migration in the History of Eastern Eurasian Sheep. *Mol. Biol. Evol.* 2015, *32*, 2515–2533. [CrossRef]
- 7. Kim, E.S.; Elbeltagy, A.R.; Aboul-Naga, A.M.; Rischkowsky, B.; Sayre, B.; Mwacharo, J.M.; Rothschild, M.F. Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity* **2016**, *116*, 255–264. [CrossRef]
- Yang, J.; Li, W.R.; Lv, F.H.; He, S.G.; Tian, S.L.; Peng, W.F.; Sun, Y.W.; Zhao, Y.X.; Tu, X.L.; Zhang, M.; et al. Whole-Genome Sequencing of Native Sheep Provides Insights into Rapid Adaptations to Extreme Environments. *Mol. Biol. Evol.* 2016, 33, 2576–2592. [CrossRef]
- Cao, Y.H.; Xu, S.S.; Shen, M.; Chen, Z.H.; Gao, L.; Lv, F.H.; Xie, X.L.; Wang, X.H.; Yang, H.; Liu, C.B.; et al. Historical Introgression from Wild Relatives Enhanced Climatic Adaptation and Resistance to Pneumonia in Sheep. *Mol. Biol. Evol.* 2021, 38, 838–855. [CrossRef]
- 10. Senczuk, G.; Criscione, A.; Mastrangelo, S.; Biscarini, F.; Marletta, D.; Pilla, F.; Laloë, D.; Ciampolini, R. How Geography and Climate Shaped the Genomic Diversity of Italian Local Cattle and Sheep Breeds. *Animals* **2022**, *12*, 2198. [CrossRef]
- 11. Hou, Y.; Liu, G.E.; Bickhart, D.M.; Cardone, M.F.; Wang, K.; Ventura, M.; Song, J.; VanRaden, P.M.; Sonstegard, T.S.; Van Tassell, C.P.; et al. Genomic characteristics of cattle copy number variations. *BMC Genom.* **2011**, *12*, 127. [CrossRef]
- 12. Zarrei, M.; MacDonald, J.R.; Merico, D.; Scherer, S.W. A copy number variation map of the human genome. *Nat. Rev. Genet.* 2015, 16, 172–183. [CrossRef]
- 13. Hastings, P.J.; Lupski, J.R.; Rosenberg, S.M.; Ira, G. Mechanisms of change in gene copy number. *Nat. Rev. Genet.* **2009**, *10*, 551–564. [CrossRef]
- 14. Bruder, C.E.G.; Piotrowski, A.; Gijsbers, A.A.C.J.; Andersson, R.; Erickson, S.; de Stahl, T.D.; Menzel, U.; Sandgren, J.; von Tell, D.; Poplawski, A.; et al. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am. J. Hum. Genet.* **2008**, *82*, 763–771. [CrossRef]
- 15. Nguyen, D.Q.; Webber, C.; Ponting, C.P. Bias of selection on human copy-number variants. PLoS Genet. 2006, 2, e20. [CrossRef]
- 16. Dumas, L.; Kim, Y.H.; Karimpour-Fard, A.; Cox, M.; Hopkins, J.; Pollack, J.R.; Sikela, J.M. Gene copy number variation spanning 60 million years of human and primate evolution. *Genome Res.* **2007**, *17*, 1266–1277. [CrossRef]
- 17. Conrad, D.F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T.D.; Barnes, C.; Campbell, P.; et al. Origins and functional impact of copy number variation in the human genome. *Nature* **2010**, *464*, 704–712. [CrossRef]
- 18. Kibriya, M.G.; Jasmine, F.; Parvez, F.; Argos, M.; Roy, S.; Paul-Brutus, R.; Islam, T.; Ahmed, A.; Rakibuz-Zaman, M.; Shinkle, J.; et al. Association between genome-wide copy number variation and arsenic-induced skin lesions: A prospective study. *Environ. Health* **2017**, *16*, 75. [CrossRef] [PubMed]
- 19. Fontanesi, L.; Beretti, F.; Riggio, V.; Gómez González, E.; Dall'Olio, S.; Davoli, R.; Russo, V.; Portolano, B. Copy number variation and missense mutations of the agouti signaling protein (*ASIP*) gene in goat breeds with different coat colors. *Cytogenet. Genome Res.* **2009**, *126*, 333–347. [CrossRef] [PubMed]
- Giuffra, E.; Törnsten, A.; Marklund, S.; Bongcam-Rudloff, E.; Chardon, P.; Kijas, J.M.; Anderson, S.I.; Archibald, A.L.; Andersson, L. A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT. *Mamm. Genome* 2002, 13, 569–577. [CrossRef] [PubMed]
- Zhu, C.; Fan, H.; Yuan, Z.; Hu, S.; Ma, X.; Xuan, J.; Wang, H.; Zhang, L.; Wei, C.; Zhang, Q.; et al. Genome-wide detection of CNVs in Chinese indigenous sheep with different types of tails using ovine high-density 600K SNP arrays. *Sci. Rep.* 2016, *6*, 27822. [CrossRef]
- 22. Li, X.; Yang, J.; Shen, M.; Xie, X.L.; Xu, Y.X.; Lv, F.H.; Shen, M.; Yang, H.; Yang, Y.L.; Liu, C.B.; et al. Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. *Nat. Commun.* **2020**, *11*, 2815. [CrossRef]
- Hancock, A.M.; Witonsky, D.B.; Alkorta-Aranburu, G.; Beall, C.M.; Gebremedhin, A.; Sukernik, R.; Utermann, G.; Pritchard, J.K.; Coop, G.; Di Rienzo, A. Adaptations to Climate-Mediated Selective Pressures in Humans. *PLoS Genet.* 2011, 7, e1001375. [CrossRef]

- 24. Hovhannisyan, G.; Harutyunyan, T.; Aroutiounian, R.; Liehr, T. DNA Copy Number Variations as Markers of Mutagenic Impact. *Int. J. Mol. Sci.* **2019**, 20, 4723. [CrossRef]
- Tsilimigaki, S.I.; Messini-Nikolaki, N.; Kanariou, M.; Piperakis, S.M. A study on the effects of seasonal solar radiation on exposed populations. *Mutagenesis* 2003, *18*, 139–143. [CrossRef] [PubMed]
- Zitzelsberger, H.; Unger, K. DNA copy number alterations in radiation-induced thyroid cancer. *Clin. Oncol. (R. Coll. Radiol.)* 2011, 23, 289–296. [CrossRef] [PubMed]
- 27. Skinner, M.K.; Guerrero-Bosagna, C.; Haque, M.M. Environmentally induced epigenetic transgenerational inheritance of sperm epimutations promote genetic mutations. *Epigenetics* **2015**, *10*, 762–771. [CrossRef] [PubMed]
- Hsieh, P.; Vollger, M.R.; Dang, V.; Porubsky, D.; Baker, C.; Cantsilieris, S.; Hoekzema, K.; Lewis, A.P.; Munson, K.M.; Sorensen, M.; et al. Adaptive archaic introgression of copy number variants and the discovery of previously unknown human genes. *Science* 2019, *366*, eaax2083. [CrossRef]
- Zhang, Y.; Hu, Y.; Wang, X.; Jiang, Q.; Zhao, H.; Wang, J.; Ju, Z.; Yang, L.; Gao, Y.; Wei, X.; et al. Population Structure, and Selection Signatures Underlying High-Altitude Adaptation Inferred From Genome-Wide Copy Number Variations in Chinese Indigenous Cattle. Front. Genet. 2020, 10, 1404. [CrossRef]
- 30. Liu, M.; Zhou, Y.; Rosen, B.D.; Van Tassell, C.P.; Stella, A.; Tosser-Klopp, G.; Rupp, R.; Palhiere, I.; Colli, L.; Sayre, B.; et al. Diversity of copy number variation in the worldwide goat population. *Heredity* **2019**, *122*, 636–646. [CrossRef] [PubMed]
- Solé, M.; Ablondi, M.; Binzer-Panchal, A.; Velie, B.D.; Hollfelder, N.; Buys, N.; Ducro, B.J.; François, L.; Janssens, S.; Schurink, A.; et al. Inter- and intra-breed genome-wide copy number diversity in a large cohort of European equine breeds. *BMC Genom.* 2019, 20, 759. [CrossRef] [PubMed]
- 32. Berglund, J.; Nevalainen, E.M.; Molin, A.M.; Perloski, M.; André, C.; Zody, M.C.; Sharpe, T.; Hitte, C.; Lindblad-Toh, K.; Lohi, H.; et al. Novel origins of copy number variation in the dog genome. *Genome Biol.* **2012**, *13*, R73. [CrossRef] [PubMed]
- 33. Perry, G.H.; Dominy, N.J.; Claw, K.G.; Lee, A.S.; Fiegler, H.; Redon, R.; Werner, J.; Villanea, F.A.; Mountain, J.L.; Misra, R.; et al. Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* 2007, 39, 1256–1260. [CrossRef]
- 34. Reiter, T.; Jagoda, E.; Capellini, T.D. Dietary Variation and Evolution of Gene Copy Number among Dog Breeds. *PLoS ONE* **2016**, *11*, e0148899. [CrossRef]
- Gautier, M.; Moazami-Goudarzi, K.; Levéziel, H.; Parinello, H.; Grohs, C.; Rialle, S.; Kowalczyk, R.; Flori, L. Deciphering the wisent demographic and adaptive histories from individual whole-genome sequences. *Mol. Biol. Evol.* 2016, 33, 2801–2814. [CrossRef]
- 36. Fontanesi, L.; Martelli, P.L.; Beretti, F.; Riggio, V.; Dall'Olio, S.; Colombo, M.; Casadio, R.; Russo, V.; Portolano, B. An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC Genom.* **2010**, *11*, 639. [CrossRef]
- 37. Wang, W.; Wang, S.; Hou, C.; Xing, Y.; Cao, J.; Wu, K.; Liu, C.; Zhang, D.; Zhang, L.; Zhang, Y.; et al. Genome-wide detection of copy number variations among diverse horse breeds by array CGH. *PLoS ONE* **2014**, *30*, e86860. [CrossRef] [PubMed]
- Scheffers, B.R.; De Meester, L.; Bridge, T.C.; Hoffmann, A.A.; Pandolfi, J.M.; Corlett, R.T.; Butchart, S.H.; Pearce-Kelly, P.; Kovacs, K.M.; Dudgeon, D.; et al. The broad footprint of climate change from genes to biomes to people. *Science* 2016, 354, aaf7671. [CrossRef]
- Salehian-Dehkordi, H.; Xu, Y.X.; Xu, S.S.; Li, X.; Luo, L.Y.; Liu, Y.J.; Wang, D.F.; Cao, Y.H.; Shen, M.; Gao, L.; et al. Genome-Wide Detection of Copy Number Variations and Their Association With Distinct Phenotypes in the World's Sheep. *Front. Genet.* 2021, 12, 670582. [CrossRef]
- Wang, K.; Li, M.; Hadley, D.; Liu, R.; Glessner, J.; Grant, S.F.A.; Hakonarson, H.; Bucan, M. PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* 2007, *17*, 1665–1674. [CrossRef]
- Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J.; et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* 2012, 337, 1190–1195. [CrossRef] [PubMed]
- 42. New, M.; Lister, D.; Hulme, M.; Makin, I. A high-resolution data set of surface climate over global land areas. *Clim. Res.* 2002, 21, 1–25. [CrossRef]
- 43. Suehrcke, H.; Bowden, R.S.; Hollands, K. Relationship between sunshine duration and solar radiation. *Sol. Energy* **2013**, *92*, 160–171. [CrossRef]
- 44. Duffie, J.A.; Beckman, W.A. Solar Engineering of Thermal Processes; John Wiley & Sons: Hoboken, NJ, USA, 2013.
- 45. Gueymard, C.A. Revised composite extraterrestrial spectrum based on recent solar irradiance observations. *Sol. Energy* **2018**, *169*, 434–440. [CrossRef]
- 46. Duruz, S.; Sevane, N.; Selmoni, O.; Vajana, E.; Leempoel, K.; Stucki, S.; Orozco-terWengel, P.; Rochat, E.; Dunner, S.; The NEXTGEN Consortium; et al. Rapid identification and interpretation of gene–environment associations using the new R. SamBada landscape genomics pipeline. *Mol. Ecol. Resour.* 2019, 19, 1355–1365. [CrossRef]
- Frichot, E.; Schoville, S.D.; Bouchard, G.; François, O. Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Mol. Biol. Evol.* 2013, 30, 1687–1699. [CrossRef]
- Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* 2009, *9*, 1322–1332. [CrossRef]
- 49. Patterson, N.; Price, A.L.; Reich, D. Population Structure and Eigenanalysis. PLoS Genet. 2006, 2, e190. [CrossRef]

- Caye, K.; Jumentier, B.; Lepeule, J.; François, O. LFMM 2: Fast and Accurate Inference of Gene-Environment Associations in Genome-Wide Studies. *Mol. Biol. Evol.* 2019, *36*, 852–860. [CrossRef]
- 51. Glessner, J.T.; Li, J.; Hakonarson, H. ParseCNV integrative copy number variation association software with quality tracking. *Nucleic Acids Res.* **2013**, *41*, e64. [CrossRef]
- Kang, H.M.; Sul, J.H.; Service, S.K.; Zaitlen, N.A.; Kong, S.y.; Freimer, N.B.; Sabatti, C.; Eskin, E. Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 2010, 42, 348–354. [CrossRef] [PubMed]
- 53. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2008**, *4*, 44–57. [CrossRef] [PubMed]
- 54. Liu, M.; Fang, L.; Liu, S.; Pan, M.G.; Seroussi, E.; Cole, J.B.; Ma, L.; Chen, H.; Liu, G.E. Array CGH-based detection of CNV regions and their potential association with reproduction and other economic traits in Holsteins. *BMC Genom.* **2019**, *20*, 181. [CrossRef] [PubMed]
- 55. Bae, J.S.; Cheong, H.S.; Kim, L.H.; NamGung, S.; Park, T.J.; Chun, J.Y.; Kim, J.Y.; Pasaje, C.F.A.; Lee, J.S.; Shin, H.D. Identification of copy number variations and common deletion polymorphisms in cattle. *BMC Genom.* **2010**, *11*, 232. [CrossRef]
- 56. Wang, M.D.; Dzama, K.; Hefer, C.A.; Muchadeyi, F.C. Genomic population structure and prevalence of copy number variations in South African Nguni cattle. *BMC Genom.* 2015, *16*, 894. [CrossRef]
- 57. Prunier, J.; Giguère, I.; Ryan, N.; Guy, R.; Soolanayakanahally, R.; Isabel, N.; MacKay, J.; Porth, I. Gene copy number variations involved in balsam poplar. *Mol. Ecol.* 2019, *28*, 1476–1490. [CrossRef]
- 58. Mei, C.G.; Junjvlieke, Z.; Raza, S.H.A.; Wang, H.B.; Cheng, G.; Zhao, C.; Zhu, W.; Zan, L. Copy number variation detection in Chinese indigenous cattle by whole genome sequencing. *Genomics* **2019**, *112*, 831–836. [CrossRef]
- 59. Wang, Y.; Song, H.; Wang, W.; Zhang, Z. Generation and characterization of *Megf6* null and Cre knock-in alleles. *Genesis* **2019**, 57, e23262. [CrossRef]
- 60. Xu, L.Y.; Yang, L.; Wang, L.; Zhu, B.; Chen, Y.; Gao, H.; Gao, X.; Zhang, L.; Liu, G.E.; Li, J. Probe-based association analysis identifies several deletions associated with average daily gain in beef cattle. *BMC Genom.* **2019**, *20*, 31. [CrossRef]
- 61. Beckmann, J.S.; Estivill, X.; Antonarakis, S.E. Copy number variants and genetic traits: Closer to the resolution of phenotypic to genotypic variability. *Nat. Rev. Genet.* **2007**, *8*, 639–646. [CrossRef]
- 62. O'Donovan, M.C.; Kirov, G.; Owen, M.J. Phenotypic variations on the theme of CNVs. *Nat. Genet.* 2008, 40, 1392–1393. [CrossRef] [PubMed]
- 63. Fukuda, M.; Kuroda, T.S.; Mikoshiba, K. Slac2-a/Melanophilin, the Missing Link between Rab27 and Myosin Va. J. Biol. Chem. 2002, 277, 12432–12436. [CrossRef]
- 64. Gudbjartsson, D.F.; Sulem, P.; Stacey, S.N.; Goldstein, A.M.; Rafnar, T.; Sigurgeirsson, B.; Benediktsdottir, K.R.; Thorisdottir, K.; Ragnarsson, R.; Sveinsdottir, S.G.; et al. *ASIP* and *TYR* pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat. Genet.* **2008**, *40*, 886–891. [CrossRef]
- 65. Yamakaze, J.; Lu, Z. Deletion of the lactoperoxidase gene causes multisystem inflammation and tumors in mice. *Sci. Rep.* **2021**, 11, 12429. [CrossRef]
- 66. Zhu, X.; Chang, Y.P.C.; Yan, D.; Weder, A.; Cooper, R.; Luke, A.; Kan, D.; Chakravarti, A. Associations Between Hypertension and Genes in the Renin-Angiotensin System. *Hypertension* **2003**, *41*, 1027–1034. [CrossRef] [PubMed]
- 67. Singh, V.K.; Singh, S.; Nandhini, P.B.; Bhatia, A.K.; Dixit, S.P.; Ganguly, I. Comparative genomic diversity analysis of copy number variations (CNV) in indicine and taurine cattle thriving in Europe and Indian subcontinent. *Anim. Biotechnol.* **2023**, 1–12. [CrossRef]
- 68. Kalds, P.; Zhou, S.; Gao, Y.; Cai, B.; Huang, S.; Chen, Y.; Wang, X. Genetics of the phenotypic evolution in sheep: A molecular look at diversity-driving genes. *Genet. Sel. Evol.* **2022**, *54*, 61. [CrossRef] [PubMed]
- Xu, Y.X.; Wang, B.; Jing, J.N.; Ma, R.; Luo, Y.H.; Li, X.; Yan, Z.; Liu, Y.J.; Gao, L.; Ren, Y.L.; et al. Whole-body adipose tissue multi-omic analyses in sheep reveal molecular mechanisms underlying local adaptation to extreme environments. *Commun. Biol.* 2023, *6*, 159. [CrossRef]
- Aboul-Naga, A.M.; Alsamman, A.M.; El Allali, A.; Elshafie, M.H.; Abdelal, E.S.; Abdelkhalek, T.M.; Abdelsabour, T.H.; Mohamed, L.G.; Hamwieh, A. Genome-wide analysis identified candidate variants and genes associated with heat stress adaptation in Egyptian sheep breeds. *Front. Genet.* 2022, *3*, 898522. [CrossRef]
- 71. Nabi Khan, R.I.; Sahu, A.R.; Malla, W.A.; Praharaj, M.R.; Hosamani, N.; Kumar, S.; Gupta, S.; Sharma, S.; Saxena, A.; Varshney, A.; et al. Systems biology under heat stress in Indian cattle. *Gene* **2021**, *30*, 145908. [CrossRef]
- Seifert, A.; Schofield, P.; Barton, G.J.; Hay, R.T. Proteotoxic stress reprograms the chromatin landscape of SUMO modification. *Sci. Signal.* 2015, *8*, rs7. [CrossRef] [PubMed]
- 73. Wang, K.; Yang, Y.; Wang, L.; Ma, T.; Shang, H.; Ding, L.; Han, J.; Qiu, Q. Different gene expressions between cattle and yak provide insights into high-altitude adaptation. *Anim. Genet.* **2016**, *47*, 28–35. [CrossRef] [PubMed]
- 74. Igoshin, A.V.; Deniskova, T.E.; Yurchenko, A.A.; Yudin, N.S.; Dotsev, A.V.; Selionova, M.I.; Zinovieva, N.A.; Larkin, D.M. Copy number variants in genomes of local sheep breeds from Russia. *Anim. Genet.* **2022**, *53*, 119–132. [CrossRef]
- 75. Maiuri, T.; Mocle, A.J.; Hung, C.L.; Xia, J.; van Roon-Mom, W.M.; Truant, R. Huntingtin is a scaffolding protein in the ATM oxidative DNA damage response complex. *Hum. Mol. Genet.* **2017**, *26*, 395–406. [CrossRef]

- Naji, M.M.; Jiang, Y.; Utsunomiya, Y.T.; Rosen, B.D.; Sölkner, J.; Wang, C.; Jiang, L.; Zhang, Q.; Zhang, Y.; Ding, X.; et al. Favored single nucleotide variants identified using whole genome Re-sequencing of Austrian and Chinese cattle breeds. *Front. Genet.* 2022, 13, 974787. [CrossRef] [PubMed]
- Liu, J.; Zhang, S.; Dai, W.; Xie, C.; Li, J.C. A Comprehensive Prognostic and Immune Analysis of *SLC41A3* in Pan-Cancer. *Front.* Oncol. 2021, 10, 586414. [CrossRef] [PubMed]
- 78. Zhai, R.; Wang, W.; Zhang, D.; Li, X.; Zhang, Y.; Zhao, Y.; Zhao, L.; Wang, J.; Xu, D.; Cheng, J.; et al. Novel polymorphism at *KLF15* gene and its association with growth traits in Hu sheep. *Anim. Biotechnol.* **2022**, 1–7. [CrossRef]
- Chen, J.; Okimura, K.; Yoshimura, T. Light and Hormones in Seasonal Regulation of Reproduction and Mood. *Endocrinology* 2020, 161, bqaa130. [CrossRef]
- 80. Brunet, A.G.; Santiago-Moreno, J.; Toledano-Díaz, A.; López-Sebastián, A. Reproductive seasonality and its control in Spanish sheep and goats. *Biol. Trop. Subtrop. Agroecosyst.* **2011**, *15*, S47–S70.
- Gomez-Brunet, A.; Santiago-Moreno, J.; del Campo, A.; Malpaux, B.; Chemineau, P.; Tortonese, D.J. Endogenous Circannual Cycles of Ovarian Activity and Changes in Prolactin and Melatonin Secretion in Wild and Domestic Female Sheep Maintained under a Long-Day Photoperiod. *Biol. Reprod.* 2008, 78, 552–562. [CrossRef]
- Chemineau, P.; Bodin, L.; Migaud, M.; Thiéry, J.C.; Malpaux, B. Neuroendocrine and Genetic Control of Seasonal Reproduction in Sheep and Goats. *Reprod. Domest. Anim.* 2010, 45, 42–49. [CrossRef] [PubMed]
- Bertolini, F.; Servin, B.; Talenti, A.; Rochat, E.; Kim, E.S.; Oget, C.; Palhière, I.; Crisà, A.; Catillo, G.; Steri, R.; et al. Signatures of selection and environmental adaptation across the goat genome post-domestication. *Genet. Sel. Evol.* 2018, 50, 57. [CrossRef] [PubMed]
- Flori, L.; Moazami-Goudarzi, K.; Alary, V.; Araba, A.; Boujenane, I.; Boushaba, N.; Casabianca, F.; Casu, S.; Ciampolini, R.; D'Acier, A.C.; et al. A genomic map of climate adaptation in Mediterranean cattle breeds. *Mol. Ecol.* 2019, 28, 1009–1029. [CrossRef]
- Yuan, C.; Lu, Z.; Guo, T.; Yue, Y.; Wang, X.; Wang, T.; Zhang, Y.; Hou, F.; Niu, C.; Sun, X.; et al. A global analysis of CNVs in Chinese indigenous fine-wool sheep populations using whole-genome resequencing. *BMC Genom.* 2021, 22, 78. [CrossRef] [PubMed]
- Di Gerlando, R.; Mastrangelo, S.; Tolone, M.; Rizzuto, I.; Sutera, A.M.; Moscarelli, A.; Portolano, B.; Sardina, M.T. Identification of Copy Number Variations and Genetic Diversity in Italian Insular Sheep Breeds. *Animals* 2022, 12, 217. [CrossRef]
- 87. Norris, B.J.; Whan, V.A. A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. *Genome Res.* **2008**, *18*, 1282–1293. [CrossRef] [PubMed]
- 88. Guo, J.; Tao, H.; Li, P.; Li, L.; Zhong, T.; Wang, L.; Ma, J.; Chen, X.; Song, T.; Zhang, H. Whole-genome sequencing reveals selection signatures associated with important traits in six goat breeds. *Sci. Rep.* **2018**, *8*, 10405. [CrossRef] [PubMed]
- Braz, C.U.; Rowan, T.N.; Schnabel, R.D.; Decker, J.E. Genome-wide association analyses identify genotype-by-environment interactions of growth traits in Simmental cattle. *Sci. Rep.* 2021, *11*, 13335. [CrossRef]
- 90. Smith, J.L.; Wilson, M.L.; Nilson, S.M.; Rowan, T.N.; Schnabel, R.D.; Decker, J.E.; Seabury, C.M. Genome-wide association and genotype by environment interactions for growth traits in U.S. Red Angus cattle. *BMC Genom.* **2022**, *23*, 517. [CrossRef]
- Corvo, M.D.; Lazzari, B.; Capra, E.; Zavarez, L.; Milanesi, M.; Utsunomiya, Y.T.; Utsunomiya, A.T.H.; Stella, A.; de Paula Nogueira, G.; Garcia, J.F.; et al. Methylome Patterns of Cattle Adaptation to Heat Stress. *Front. Genet.* 2021, 12, 633132. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.