animals

MDPI

Article

Genome-Wide Scan for Copy Number Variations in Chinese Merino Sheep Based on Ovine High-Density 600K SNP Arrays

Yuezhen Tian [1], Jing An [1,2], Xinning Zhang [3], Jiang Di [1], Junmin He [4], Ayinuer Yasen [1], Yanpin Ma [1], Gaohaer Sailikehan [3], Xixia Huang [3,*] and Kechuan Tian [4,*]

[1] Key Laboratory of Genetics Breeding and Reproduction of Xinjiang Cashmere and Wool Sheep, Institute of Animal Science, Xinjiang Academy of Animal Science, Urumqi 830011, China
[2] College of Animal Science and Technology, Northwest Agriculture and Forest University, Yangling, Xianyang 712100, China
[3] College of Animal Science, Xinjiang Agricultural University, Urumqi 830052, China
[4] Institute of Animal Science and Veterinary Medicine, Shandong Academy of Agricultural Sciences, Jinan 250100, China
* Correspondence: au-huangxixia@163.com (X.H.); woolcashmere@163.com or tiankechuan@163.com (K.T.)

Simple Summary: A genome-wide copy number variations (CNVs) analysis using high-density Ovine BeadChip array data in 288 Chinese Merino sheep was conducted. A total of 656 CNV regions (CNVRs) with an average size of 66.88 Kbps were identified. These CNVRs contain 519 losses, 60 gains, and 77 gain–losses ranging from 1.5 Kbps to 646 Kbps and covering 43 Mbps (1.58%) of the whole Ovine genome. To validate these results, we performed a quantitative PCR to detect 11 randomly selected CNVRs and successfully verified 8 (72.7%). The gene functional enrichment analysis revealed that a total of 1592 genes in 465 CNVRs were identified as related to biological functions such as ATP binding activity. Our results expand the current CNV map of the sheep genome and provide preliminary fundamental information for carrying out CNV studies of fine wool sheep in the future.

Abstract: Sheep are a vital species in the global agricultural economy, providing essential resources such as meat, milk, and wool. Merino sheep (Junken type) are a key breed of fine wool sheep in China. However, research on fine wool traits has largely overlooked the role of SNPs and their association with phenotypes. Copy number variations (CNVs) have emerged as one of the most important sources of genetic variation, influencing phenotypic traits by altering gene expression and dosage. To generate a comprehensive CNVR map of the ovine genome, we conducted genome-wide CNV detection using genotyping data from 285 fine wool sheep. This analysis revealed 656 CNVRs, including 628 on autosomes and 28 on the X chromosome, covering a total of 43.9 Mbs of the sheep genome. The proportion of CNVRs varied across chromosomes, from 0.45% on chromosome 26 to 3.72% on chromosome 10. Functional annotation through Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses highlighted significantly enriched GO terms, including odorant binding, ATP binding, and sulfuric ester hydrolase activity. The KEGG analysis identified involvement in pathways such as neuroactive ligand–receptor interaction, axon guidance, ECM–receptor interaction, the one-carbon pool by folate, and focal adhesion ($p < 0.05$). To validate these CNVRs, we performed quantitative real-time PCR experiments to verify copy number predictions made by PennCNV software (v1.0.5). Out of 11 selected CNVRs with predicted gain, loss, or gain–loss statuses, 8 (IDs 68, 156, 201, 284, 307, 352, 411, 601) were successfully confirmed. This study marks a significant step forward in mapping CNVs in the ovine genome and offers a valuable resource for future research on genetic variation in sheep.

Keywords: copy number variations; fine wool sheep; Ovine HD BeadChip

check for updates

Citation: Tian, Y.; An, J.; Zhang, X.; Di, J.; He, J.; Yasen, A.; Ma, Y.; Sailikehan, G.; Huang, X.; Tian, K. Genome-Wide Scan for Copy Number Variations in Chinese Merino Sheep Based on Ovine High-Density 600K SNP Arrays. *Animals* **2024**, *14*, 2897. https://doi.org/10.3390/ani14192897

Academic Editor: Huitong Zhou

Received: 29 July 2024
Revised: 1 October 2024
Accepted: 2 October 2024
Published: 8 October 2024

Copyright: © 2024 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/).

1. Introduction

Copy number variations (CNVs) refer to structural variants where specific DNA segments have an altered number of copies, ranging from 1 kilobase (Kb) to several megabases (Mbs) [1,2]. Compared to single nucleotide polymorphisms (SNPs), CNVs can have a greater impact on phenotypic variation by altering gene expression and dosage. CNVs can be either deleted or duplicated through translocation, insertion, or deletion [2,3], making them a significant source of genetic variation between generations and phenotypic diversity among individuals and populations [4–6]. Numerous CNV studies have been conducted on humans [7–9], mice [10–13], dogs [14–16], cattle [17–20], chicken [21], pigs [22], goats [23], and sheep [24]. For example, in the human population, alterations in the number of copies of the *CCL3L1* gene have been found to be linked with susceptibility to HIV-1/AIDS [25,26] and *FCGR3B* copy number variation has a significant effect on systemic lupus erythematosus (SLE) [27]. Additionally, CNVs have been observed in domestic animals, leading to variations in phenotype; for instance, one copy of a 450 kb duplication in the *KIT* gene associated with white coat phenotype in pigs [28]. Furthermore, a 4.6 Kbp duplication in intron 6 of *STX17* influences the phenotype of hair greying and melanoma in horses [29]. In sheep, the gain-type CNV of *PIGY* has shown better growth traits than other types [30]. Above all, all these studies have demonstrated that CNVs are valuable genetic variant resources for understanding complex phenotypes and economically important trait variations in domestic animals.

CNV detection is primarily done using four platforms: comparative genomic hybridization (CGH) arrays, next-generation sequencing (NGS) [31], long-read sequencing [32], and SNP arrays [6]. Although CGH arrays have excellent signal-to-noise ratios [33,34], they are not suitable for large-scale CNV detection, especially on a genome-wide scale. While NGS provides a more detailed characterization of CNVs [2,3,19,35], it remains too expensive for CNV detection in large populations. On the other hand, SNP arrays offer a more desirable performance in genome-wide CNV detection and are more cost-effective, enabling users to increase the number of samples at a reasonable cost. However, the primary limitation of SNP arrays for CNV detection is the low density of SNP coverage, making it difficult to detect small CNVs [36]. This limitation can be mitigated to some extent by using high-density SNP arrays. Therefore, high-density SNPs have been widely and successfully used in detecting CNVRs in human and various animal genomes.

Sheep (*Ovis aries*) are among the earliest domesticated livestock species, providing resources like meat, milk, and wool. These contributions have made sheep a vital part of the global agricultural economy since the Neolithic era [37,38]. Merino sheep (Junken type) is a major breed of fine wool sheep in China, bred since 1972 [33]. In the Xinjiang Uyghur Autonomous Region, where Chinese Merino sheep are raised, wool typically contributes to 21% of the total flock income [39]. They are known for their soft, high-quality wool, highly valued in the textile industry. Over the past few decades, improving wool economic traits in Merino sheep has been a major focus in many countries, utilizing various quantitative and marker-assisted methods to identify quantitative trait loci (QTL) associated with wool traits [40]. Wool keratin-associated proteins (KAPs) are structural elements of wool and hair fibres, creating a matrix that encases the keratin intermediate filaments (KIFs). They are thought to be crucial in determining the physical and mechanical properties of wool fibres [41,42]. The *KAP6* gene has been reported to be linked to Mean Fibre Diameter (MFD) in medium wool Peppin Merino populations [43,44]. A 57 bp deletion in *KRTAP6-1* was linked to the production of coarser wool, characterized by a higher Fibre Diameter Standard Deviation (FDSD), an increased Coefficient of Variation of Fibre Diameter (CVFD), and a greater Prickle Factor (percentage of fibres exceeding 30 microns; PF) [45]. It has been reported that *KRTAP1-2* influences clean fleece weight (CFW), greasy fleece weight (GFW), and yield in Merino cross lambs [46]. Variation in *KRT19-5* was linked to the mean fibre curvature of wool in Chinese Tan Sheep [47]. *KRT31* has been reported to be associated with increases in GFW, CFW, and mean staple length (MSL) [48]. The *KRT34* promoter variation may influence MFD, FDSD, and MSL [49]. *KRT81* may impact wool growth by

modifying the density of wool follicles in the skin, the density of individual fibres, or the skin area producing fibres, rather than by changing the rate of fibre extrusion or their diameter [50]. *KRT84* showed high expression in hair follicles, particularly in the inner root sheath, outer root sheath, and hair medulla, and was consistently expressed across all six lamb ages studied, from 1 to 270 days [51]. However, most of these studies are based on polymorphisms and their association with wool quality traits, and the impact of genome-wide CNV landscapes on Merino sheep and their influence on phenotypic variation remains largely unknown.

The main objective of this research is to identify CNVs in Chinese merino sheep. The study analysed data from 288 animals using high-density 680K SNP genotype arrays and the PennCNV program [52]. Furthermore, the CNVRs were verified using the qPCR strategy. This is the first CNVR map of fine wool sheep constructed using high-density SNP genotype data. This study presents a detailed CNVR map of the ovine genome, laying the groundwork for future association studies linking CNVs to economically significant traits in fine wool sheep.

2. Materials and Methods

2.1. Ethics Statement

Blood samples were collected from 288 fine wool sheep in strict accordance with the guidelines approved by the Biological Studies Animal Care and Use Committee of Xinjiang Province for animal experiments in the People's Republic of China. All efforts were made to minimize any discomfort during blood collection. Sample collection and experimental protocols were performed in strict accordance with the guidelines approved by the Biological Studies Animal Care and Use Committee of Xinjiang Province for animal experiments in the People's Republic of China (Approval Number XAASA-20180110).

2.2. Sample Collection and Genotyping

A total of 288 fine wool sheep were randomly selected from the original breeding stock of Chinese Merino sheep (Xinjiang type) in the eastern Yili River Valley, Xinyuan County (43°03′–43°41′ N, 82°28′–84°56′ E), Xinjiang Province, China. A volume of 2–3 mL of whole blood was collected per animal in EDTA anticoagulant tubes from the sheep jugular vein, aliquoted into 2 mL centrifuge tubes and frozen at −20 °C for genomic DNA extraction. Genomic DNA was extracted using a TIANamp Blood DNA Kit. The 288 samples were genotyped using the Illumina Ovine SNP680K BeadChip, which includes 606,006 SNPs across the ovine genome, with an average spacing of 4.28 Kbs. The genotyping platform used was Illumina's Infinium II Multi-Sample Assay and the SNP chips were scanned using iScan and analysed using Illumina's INFINIUM GENOMESTUDIO software (Version 2009.1). To increase the accuracy of CNV inference, stringent quality control criteria were applied to filter out low-quality SNPs (individual call rate > 90% and SNP call frequency > 90%).

2.3. Data Quality Control and CNV Calling

The genomic DNA of 288 individual samples from a fine wool sheep breed (Chinese merino sheep Xinjiang type) was genotyped using the Illumina high-density 680K SNP BeadChip following the manufacturer's protocol. For each SNP marker, log R ratio (LRR) and allelic intensity (B allele frequency: BAF) ratio were exported using Illumina GenomeStudio software (v2.0.5). The population frequency of the B allele (PFB) file was then calculated based on the BAF for each marker. To identify CNVs in the fine wool sheep genome, we employed the PennCNV software [52] http://www.openbioinformatics.org/penncnv (accessed on 9 January 2023) from SNP chip data on both 26 autosomes and chromosome X. To overcome the status of 'genomic waves', we generated the sheep GC model file by calculating the GC content of the 1 Mb genomic region surrounding each marker (500 Kbs on each side) and then used the '-gcmodel' option of PennCNV to adjust the results. PennCNV used a hidden Markov model (HMM)-based approach to model the dependence structures

between copy numbers. It then optimized the HMM parameters using the Baum–Welch algorithm, and we selected the final CNV list based on the posterior probabilities. To reduce the false discovery rate in CNV calling, we adopted the following criteria: (1) the standard deviation (SD) of LRR should be less than 0.30; (2) the BAF drift should be 0.01, and the waviness factor value should be between −0.05 and 0.05; and (3) the CNV should contain ten or more consecutive SNPs. Finally, the CNV regions (CNVRs) were determined by aggregating the overlapping CNVs identified across all samples according to previously published protocols [5]. Following the requirements of PennCNV, we defined CNV call filtering criteria and excluded samples with low-quality signal intensity data. Ultimately, a total of 285 samples remained for further CNV analyses.

2.4. qPCR Confirmation of CNVRs

Quantitative PCR (qPCR) was performed to validate 11 randomly selected CNVRs identified by PennCNV. Following previous methodologies [53], the glyceraldehyde-3-phosphate dehydrogenase gene (GAPDH), known to be present in a single copy across various species, was chosen as the reference gene. Primers were designed using National Center for Biotechnology Information (NCBI) reference sequences via the Primer 3.0 web tool (https://bioinfo.ut.ee/primer3/), as shown in Supplementary Table S3. qPCR reactions were carried out on a BIO-RAD CFX98™ Real-Time System following the manufacturer's instructions and recommended cycling conditions. Each sample was subjected to triplicate reactions using a 20 µL volume containing 1 µL of DNA (approximately 50 ng), 2 µL of forward and reverse primers (concentration 0.1–0.75 µM), 10 µL of 2× Master Mix, and 15 µL of water. All qPCRs were carried out in 96-well reaction plates under the following conditions: initial denaturation at 95 °C for 3 min followed by 40 cycles of denaturation at 94 °C for 10 s and combined annealing and extension at 60 °C for 30 s. The $2^{-\Delta\Delta Ct}$ method was employed to calculate the copy number [54], where Ct represents the threshold cycle. The corresponding equation follows the method reported by Zhu et al. [24]. Analyses of 3 technical replicates of the same DNA samples and 2 negative control samples in the same batch were performed, and then the coefficient of variation (CV%) between reaction tubes of each sample in the same group was calculated to exclude the time effect.

2.5. Functional Enrichment Analysis Procedures for CNVRs

Gene information was retrieved from the BioMart data management system (http://asia.ensembl.org/biomart/martview/) based on the *Ovis aries* (Oar_v3.1) genome assembly. To gain insight into the functional implications of genes overlapping or within CNVRs, Gene Ontology (GO) annotation [55] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses [56] were conducted using the online software DAVID (http://david.abcc.ncifcrf.gov/) (v2023q4) [57]. Due to the limited annotation of the sheep genome, *Ovis aries* homologous Ensembl gene IDs were converted into human orthologue Ensembl gene IDs for analysis. BH-corrected FDR *p*-value of 0.05 was used as the criterion for determining significance levels.

3. Results

3.1. CNV Genotyping

A total of 5101 CNV events were identified after excluding unreliable CNV calls in 285 Chinese fine wool sheep, consisting of 3760 losses, 1214 gains, and 127 gain–loss events (Additional file: Table S1). The CNVs had an average length of 76.1 Kbs and a median length of 55.2 Kbs (Table 1). Additionally, around 41.44% of the CNVs were between 10 and 50 Kbs in length, while 33.93% ranged from 50 to 100 Kbs, with only approximately 3% being small fragment CNVs (<10 Kbs) (Figure 1a). After consolidating overlapping CNVs, a total of 656 CNVRs were identified, including 628 autosomal and 28 X chromosome CNVRs, covering 43.9 Mbs of the sheep genome and accounting for 1.58% of the entire genome (Additional file: Table S2). These CNVRs ranged in size from 1.6 Kbs to 646.2 Kbs, with an average length of 66.9 Kbs and a median of 51.1 Kbs (Figure 1b). Within these 656 CNVRs,

519 loss events, 60 gain events, and 77 gain–loss events were observed. The number of loss events was approximately 8.65 times greater than that of gain events, consistent with findings in other species [22,53,58]. Additionally, a total of 653 CNVRs were found in at least two samples, 11 were identified in more than 50 samples, and 125 CNVRs were identified in more than 10 sheep.

Table 1. Genomic features of copy number variations in fine wool sheep (n = 285).

Type	CNV	CNVR
Total number	5101	656
Average no. of CNVs per individual	17.9	2.3
CNV size per individual	1.4 Mb	0.2 Mb
Loss	3760	519
Gain	1214	60
Gain–loss	127	77
Total length	388.1 Mbs	43.9 Mbs
Average size	76.1 Kbs	66.9 Kbs
Median size	55.2 Kbs	51.1 Kbs

Note: CNVRs come from merging overlapping CNVs.

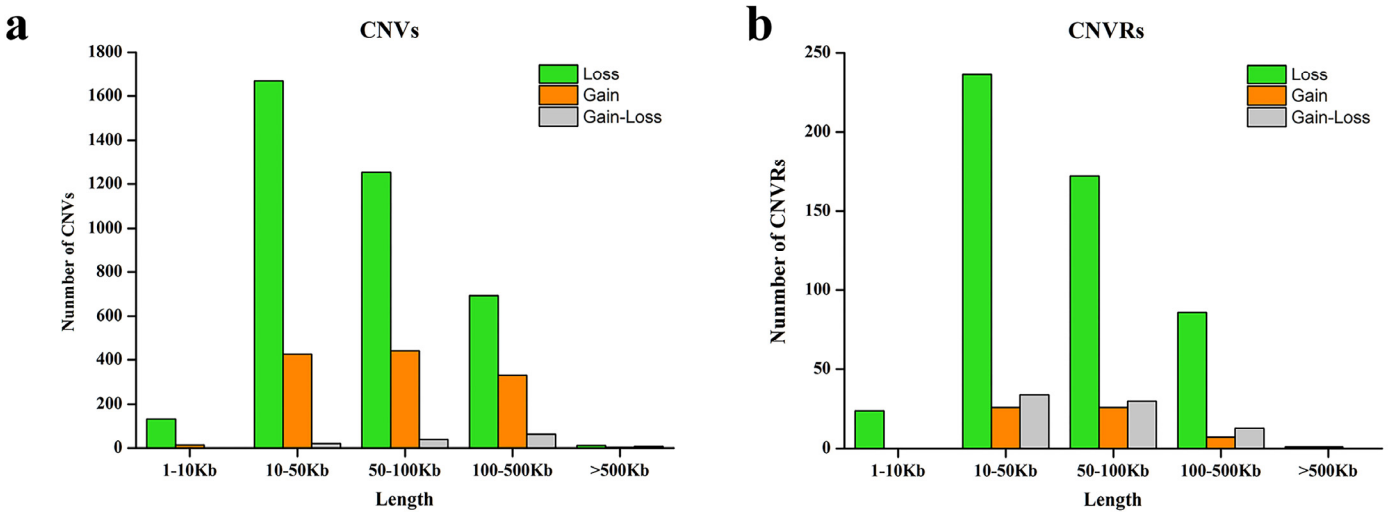


Figure 1. Size distribution of copy number variations: (**a**) CNV size distribution in fine wool sheep and (**b**) CNVR size distribution in fine wool sheep.

3.2. Genome-Wide Surveys of CNVs and CNVRs in Fine Wool Sheep

We produced a map illustrating the distribution of CNVRs across the chromosomes in fine wool sheep (Figure 2). Our findings revealed a non-random distribution of CNVs across the various chromosomes. The percentage of CNVRs on the chromosomes varied, ranging from 0.45% on chromosome 26 to 3.72% on chromosome 10, with the number of CNVRs ranging from 2 on chromosome 26 to 66 on chromosome 3.

3.3. CNV Validation by Quantitative PCR

We performed quantitative real-time PCR experiments to evaluate the accuracy of the copy number designations identified by PennCNV [52]. We chose 11 potential CNVRs that represented various predicted copy number statuses (gain, loss, or gain–loss) for CNV validation. Our findings indicated that eight of these CNVRs (CNVR IDs = 68, 156, 201, 284, 307, 352, 411, 601) were successfully validated (see Figure 3). Further details regarding the primers and results can be found in the Supplementary Information (Additional files: Tables S3 and S4).

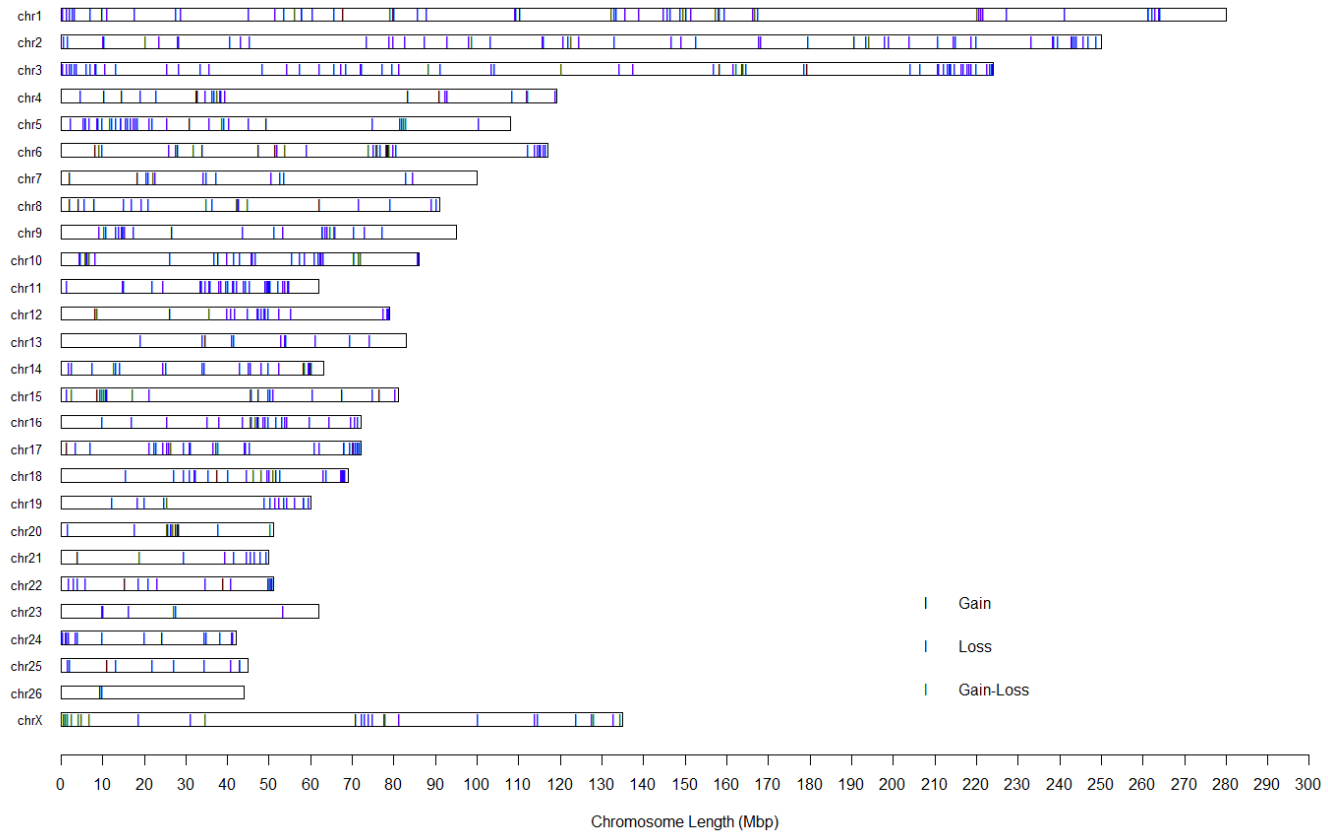


Figure 2. Map of CNVRs in fine wool sheep. Black indicates gains, blue represents losses, and dark green denotes regions with both gain and loss events.

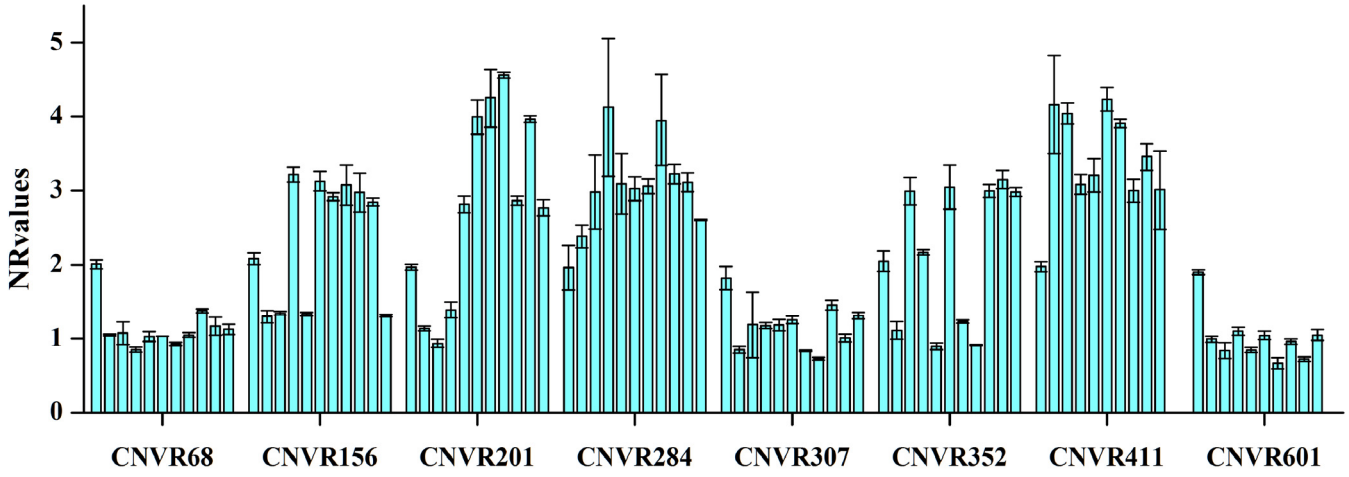


Figure 3. Normalized ratios (NRs) obtained by qPCR for 8 CNVRs. The *y*-axis displays NR values from qPCR, while the *x*-axis represents the different CNV regions. Samples with NRs of around 1 indicate individuals with a single copy (copy loss), NRs of around 2 represent normal individuals (two copies), and NRs of 3 or higher signify individuals with copy number gains.

3.4. Functional Enrichment Analysis of CNVRs

The Ensembl Genes Database from the Biomart data management system was utilized to retrieve the gene content of 656 CNVRs. A total of 1592 annotated genes were identified within 467 CNVRs in fine wool sheep. Among these genes, 1530 are protein-coding genes, 32 are snRNAs, and the remainder include pseudogenes, small nucleolar (snoRNA) genes, and rRNAs. Detailed information on these genes is provided in Additional file Table S5.

To assess the functional annotation of these CNVRs, we conducted GO and KEGG pathways analyses using the DAVID tool. A total of 31 GO terms and five KEGG pathway terms were statistically significant after Benjamin correction (Additional file: Tables S6 and S7). The significantly enriched GO terms were associated with specific biological functions such as odorant binding, ATP binding, and sulfuric ester hydrolase activity. The KEGG analysis confirmed that these genes are primarily involved in neuroactive ligand–receptor interactions, axon guidance, ECM–receptor interactions, one-carbon pool by folate, and focal adhesion ($p < 0.05$). Moreover, we identified genes associated with hair follicle development and fibre diameter traits, including Keratin, Type I Cytoskeletal 17 (*KRT17*), Mitogen-Activated Protein Kinase 1 (*MAPK1*), and EPH receptor A5 (*EPHA5*), which overlapped with CNVRs identified in this study.

4. Discussion

In recent years, CNVs have become increasingly recognized as a significant source of genetic variation, playing a crucial role in phenotypic diversity related to complex diseases in humans and economically important traits in livestock. Additionally, in the field of cancer, CNVs can be used for early diagnosis and detection because they are often involved in specific chromosome segments with relatively strong intensity [59]. For example, an algorithm has been developed to determine cancer types based on the CNVs, aiding in early cancer diagnosis [60].

With the advancement of high-density genotyping arrays, CNV detection using SNP genotyping arrays has become an efficient approach. Several algorithms for CNV inference based on SNP chip data have been developed, such as PennCNV, cnvPartition, GADA, and QuantiSNP. Each algorithm has its strengths and weaknesses [36]. Despite the potential for more reliable results by using multiple algorithms, many studies on CNV detection using SNP arrays in humans and animals have only utilized the PennCNV software (v1.0.5) [22,23,61–63]. In our study, we chose to use PennCNV to detect CNVR for two reasons. Firstly, the PennCNV algorithm incorporates various sets of information, including the total signal intensity of LRR, the distance between neighbouring SNPs, the BAF, and the population frequency of PFB [64]. Additionally, PennCNV employs a computational method that incorporates regression models with GC content to mitigate genomic waves, ensuring greater reliability compared to other algorithms. Secondly, using multiple algorithms to detect CNVs can make it difficult to determine the appropriate number of CNVs. Relying solely on commonly detected CNVs can lead to missing true CNVs, while accepting all CNVs identified by different software may result in numerous false positives. To minimize the risk of false positives associated with using a single algorithm, we applied two stringent criteria (SD of LRR < 0.30 and BAF drift = 0.01). These rigorous standards contributed to a high qPCR validation rate of 72.7% in our study.

In this study, we used the Ovine high density (HD) SNP and identified a total of 653 CNVRs, including 28 CNVRs on the X chromosome. The frequencies of these detected CNVRs are not high, only 11 CNVRs are present in more than 50 samples, accounting for 17.3% of the total samples. Similar results have been identified in other CNV detection studies. Yang et al. (2018) conducted a CNV detection study using ovine SNP50 Beadchip data on 2254 geographically dispersed sheep and identified a total of 619 CNVRs [65]. Wu et al. (2015) performed a CNV detection study using Illumina Bovine HD BeadChip (770k) data on 792 Simmental cattle and identified a total of 263 CNVRs, with only 14 of these CNVRs identified in more than 100 individuals (12.6%) [20]. Di Gerlando et al. (2019) performed a CNV detection study using ovine SNP50K Beadchip data on 468 Valle

del Belice sheep and identified a total of 365 CNVRs [66]. Taghizadeh et al. (2022) have conducted a CNV detection study using ovine SNP50K BeadChip data on 192 genomic samples from three breeds, including 96 Baluchi sheep, 47 Lori-Bakhtiari sheep (fat-tailed), and 47 Zel sheep (thin-tailed), resulting in the identification of 515 CNVRs [67]. Ma et al. (2017) used the Illumina Ovine SNP 600 K BeadChip array for genome-wide detection of CNVs in 48 Chinese Tan sheep and identified a total of 1296 CNVRs, with 553 (42%) detected in only one animal [68]. Wang et al. (2020) performed a CNV detection study using ovine SNP600K Beadchip data on 40 Hu sheep (a highly prolific breed native to China), 165 DS sheep (a synthetic line developed from hybridizing Australian Suffolk sheep and Chinese Hu sheep), and 65 SHH sheep (a crossbreed between DS sheep and Chinese Kazakh sheep) and identified a total of 919 CNVRs [69]. Salehian-Dehkordi et al. (2021) performed a CNV detection study using ovine SNP600K Beadchip data on 2059 sheep from 67 populations worldwide and identified a total of 1217 CNVRs [70]. Many factors may result in the low frequencies of the detected CNVRs in these studies, such as reference genome completeness and variant calling algorithms.

We evaluated our results by comparing them with previously published sheep CNVs. Since some earlier studies are based on the Ovis_aris_1.0 genome build, we first converted these results from Ovis_aris_1.0 to Ovis_v3.1 using the UCSC LIFTOVER tool [71] and then compared them with our results. In this comparison, we only considered the CNVRs on autosomes, as the CNVRs on the X chromosome were excluded in some studies. The detailed comparison results are shown in Table 2.

The comparison results indicate that our findings are more aligned with CNVRs based on the Ovine HD SNP array compared to CNVRs identified using other platforms. When comparing our study with Zhu et al. [24] and Ma et al. [68], both of which used HD SNP for CNV detection, we found 184 CNVRs in our study overlapped with those reported by Ma, with a total length of 24.4 Mbps, and 165 CNVRs with a length of 9.31 Mbs overlapped with those reported by Zhu et al. [24]. However, when comparing the CNVR findings based on the 50K SNPs, only three CNVRs, totalling 0.45 Mbs in length, and 15 CNVRs, totalling 1.69 Mbs in length, overlapped with those reported by Ma et al. [68] and Liu et al. [72], respectively. Notably, 66 CNVRs with a total length of 17.85 Mbs, which were reported by Yang et al. [65] based on the 50K SNPs, were also identified in our study. This variation may be attributed to the larger sample size of this study, which far exceeds that of other studies. Additionally, when compared to previously reported CNVR findings using the CGH array, only 65 CNVRs with a total length of 5.82 Mbps were identified, and CNVRs with a length of 0.05 Mbps were identified in our study overlapped with Jenkins et al. [73] and Hou et al. [74]. Notably, there are no overlapping CNVRs between our study and Fontanesi et al. [75]. The inconsistency between the results of different studies illustrated that the CNVRs identified using different populations and different technology platforms could lead to different results. It also suggested there is still a vast number of CNVs in the sheep genome that have not been discovered.

Table 2. Comparison of CNVRs on autosomes between this study and results from other studies of the Ovine genome.

			Findings from Different Studies						CNVRs Overlapping with This Study			
Platform	**Study**	**Species**	**Sample**	**Count**	**Total Length (Mbs)**	**Loss**	**Gain**	**Both**	**Count**	**Percentage of Count**	**Total Length (Mbs)**	**Percentage Length**
CGH	Jenkins et al. (2016) [73]	Texel, Coopworth, Perendale, Romney, and Merino	30	3488	67.6	2023	1325	140	65	9.90%	5.82	13.20%
	Fontanesi et al. (2011) [75]	Bagnolese, Comisana, Laticauda, Massese, Sarda, and Valle del Belice	11	135	10.5	75	59	1	0	0%	0	0%
	Hou et al. (2015) [74]	Mongolian sheep, Kazakh Sheep, Tibetan sheep, Hu sheep	122	51	15.55	21	23	7	1	0.15%	0.05	
SNP50	Yang et al. (2018) [65]	Sixty-eight breeds from Africa, America, Asia, and Europe	2254	619	197	-	-	-	66	10.06%	17.85	40.66%
	Di Gerlando et al. (2019) [66]	Valle del Belice sheep	468	365	118.36	320	43	2				
	Goyache et al. (2021) [76]	Djallonké (West African Dwarf) sheep	184	63	82.5	36	7	20				
	Taghizadeh et al. (2022) [67]	Baluchi sheep Lori-Bakhtiari sheep Zel sheep	192	515	73.85	141	364	10				
	Ma et al. (2015) [77]	Dorper, Poll Dorset, Texel, Suffolk, South American Mutton Merino, Borderdale, Gansu Alpine Merino, and Gansu Morden sheep	160	111	13.75	12	99	0	3	0.45%	0.22	0.50%
	Liu et al. (2013) [72]	German Mutton, Dorper, and Sunite	100	238	60.35	219	13	6	15	2.20%	1.69	3.80%
SNP600	Ma et al. (2017) [68]	Chinese Tan sheep	48	1296	121.8	1173	119	5	184	28.00%	24.4	55.60%
	Zhu et al. (2016) [24]	Chinese indigenous sheep	120	490	81.04	390	93	7	165	25.10%	9.31	21.20%
	Wang et al. (2020) [69]	Hu sheep DS sheep SHH sheep	270	919	48.17	730	102	87				
	Salehian-Dehkordi et al. (2021) [70]	sheep from 67 populations all over the world	2059	1217	245	918	197	102				
	This Study	Chinese Merino sheep	288	656	43.9	519	60	77	-	-	-	-

CGH means comparative genomic hybridization arrays, SNP50 stands for SNP50K beadchip array, SNP600 represents SNP600K beadchip array.

We utilized UCSC gene annotation (http://genome.ucsc.edu) to identify genes either fully or partially overlapping with the CNVRs, resulting in the identification of 1592 genes. Interestingly, we identified the functions of some genes associated with hair follicle development and fibre diameter trait. For example, *KRT17*, located in a CNVR detected in 11 animals, is specifically involved in determining the shape and orientation of hair [78]; *MAPK1*, found within a CNVR identified in three animals, plays a critical role in the hair cycle and maintenance of hair follicle stem cell quiescence [79]; and *EPHA5*, located in a CNVR detected in three animals, encodes a protein that belongs to the ephrin receptor subfamily of the protein tyrosine kinase family [80]. Moreover, to make further use of these annotation genes, we conducted a GO analysis and identified 31 over-represented GO terms. Similar to other GO analysis results [67,81], the functions of these genes are enriched in odorant binding activity and ATP binding activity. The KEGG pathway analysis confirmed that the most significantly enriched pathway was neuroactive ligand–receptor interaction, as reported in other studies [82].

To validate the CNVRs detected by PennCNV, we performed qPCR on 11 selected CNVRs and compared the results to a control region known to be free of CNVs. We found that eight (72.7%) of our qPCR results agreed with CNVR predictions in these regions. The high percentage could be explained by the high-density probe of the Ovine HD BeadChip and the large animal resource; additionally, the stringent CNV detection criteria employed in this study, requiring at least ten consecutive SNPs, ensured that most CNVRs predicted by PennCNV closely matched the qPCR validation results. It is important to note that three CNVRs were not validated by qPCR. This discrepancy may be due to the ambiguous boundaries of CNVs inferred from the SNP array, along with potential influences from SNPs and small indels. Future CNV research using next-generation sequencing (NGS) with larger sample sizes and complementary analytical tools will provide more precise CNV structure definitions and enhance the accuracy of CNV detection.

5. Conclusions

To create a comprehensive CNVR map of the ovine genome, we have conducted genome-wide CNV detection using genotyping data from 285 fine wool sheep and identified 656 CNVRs, including 628 on autosomes and 28 on the X chromosome, spanning a total of 43.9 Mbs of the sheep genome. The proportion of CNVRs varied across chromosomes, ranging from 0.45% on chromosome 26 to 3.72% on chromosome 10. For functional annotation, we conducted GO and KEGG pathway analyses. The significantly enriched GO terms were related to biological functions such as odorant binding, ATP binding, and sulfuric ester hydrolase activity. KEGG analysis confirmed that the associated genes are primarily involved in pathways like neuroactive ligand–receptor interaction, axon guidance, ECM–receptor interaction, one-carbon pool by folate, and focal adhesion ($p < 0.05$). Lastly, to validate the identified CNVRs, we performed quantitative real-time PCR experiments to assess the accuracy of copy number predictions made by PennCNV. We selected 11 potential CNVRs with various predicted statuses (gain, loss, or gain–loss) for validation, and 8 of these CNVRs (IDs 68, 156, 201, 284, 307, 352, 411, 601) were successfully confirmed. This study represents a significant advancement in mapping CNVs within the ovine genome and provides a valuable resource for future research on ovine genetic variations. Moreover, our findings offer important insights into the genomic structural variations associated with wool traits in sheep. This will support the selection of fine wool traits based on CNVRs in China's wool industry, ultimately enhancing economic outcomes and benefiting producers.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ani14192897/s1, Table S1: The results of CNVs identified by PennCNV; Table S2: The results of CNVRs identified by PennCNV; Table S3: Information on primer pairs for selected CNVRs and reference genes; Table S4: Results of calculated Normalized Ratio (NR) values for each CNVR; Table S5: The Gene information overlapped by CNVRs; Table S6: Significant GO terms of the genes in this study; Table S7: Significant KEGG pathways of the genes in this study.

Author Contributions: Conceptualization, X.H. and K.T.; Funding acquisition, X.H. and K.T.; Investigation, X.Z., J.D., J.H., Y.M., A.Y. and G.S.; Methodology, J.A.; Project administration, Y.T.; Writing—original draft, J.A.; Writing—review and editing, Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Xinjiang Uygur Autonomous Region (grant number 2022D01E64); the National Natural Science Foundation of Xinjiang Province (grant number 2024TSYCU0002); the National Modern Agriculture Industrial System Project (grant number CARS-39); and the China Postdoctoral Science Foundation (grant number 2017M623287).

Institutional Review Board Statement: Sample collection and experimental protocols were performed in strict accordance with the guidelines approved by the Biological Studies Animal Care and Use Committee of Xinjiang Province for animal experiments in the People's Republic of China (Approval Number XAASA-20180110).

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are provided within the article. Additional datasets used and/or analysed during this study are available from the corresponding author upon reasonable request.

Acknowledgments: I would like to express my gratitude to the authors of the article for their help during the experiment and writing processes.

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

References

1. Sanchez, S.; Juarez, U.; Dominguez, J.; Molina, B.; Barrientos, R.; Martinez-Hernandez, A.; Carnevale, A.; Grether-Gonzalez, P.; Mayen, D.G.; Villarroel, C.; et al. Frequent copy number variants in a cohort of Mexican-Mestizo individuals. *Mol. Cytogenet.* **2023**, *16*, 2. [CrossRef] [PubMed]
2. Gordeeva, V.; Sharova, E.; Arapidi, G. Progress in Methods for Copy Number Variation Profiling. *Int. J. Mol. Sci.* **2022**, *23*, 2143. [CrossRef] [PubMed]
3. 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**, *34*, 3483–3494. [CrossRef] [PubMed]
4. Meng, G.; Bao, Q.; Ma, X.; Chu, M.; Huang, C.; Guo, X.; Liang, C.; Yan, P. Analysis of copy number variation in the whole genome of normal-haired and long-haired tianzhu white yaks. *Genes* **2022**, *13*, 2405. [CrossRef]
5. Wang, Z.; Guo, Y.; Liu, S.; Meng, Q. Genome-Wide Assessment Characteristics of Genes Overlapping Copy Number Variation Regions in Duroc Purebred Population. *Front. Genet.* **2021**, *12*, 753748. [CrossRef]
6. Xu, C.; Zhang, W.; Jiang, Y.; Zhou, M.; Liu, L.; Su, S.; Li, X.; Wang, C. Genome-Wide Detection and Analysis of Copy Number Variation in Anhui Indigenous and Western Commercial Pig Breeds Using Porcine 80K SNP BeadChip. *Genes* **2023**, *14*, 654. [CrossRef]
7. Zamariolli, M.; Auwerx, C.; Sadler, M.C.; van der Graaf, A.; Lepik, K.; Schoeler, T.; Moyses-Oliveira, M.; Dantas, A.G.; Melaragno, M.I.; Kutalik, Z. The impact of 22q11.2 copy-number variants on human traits in the general population. *Am. J. Hum. Genet.* **2023**, *110*, 300–313. [CrossRef]
8. Benedetti, F.; Silvestri, G.; Saadat, S.; Denaro, F.; Latinovic, O.S.; Davis, H.; Williams, S.; Bryant, J.; Ippodrino, R.; Rathinam, C.V.; et al. Mycoplasma DnaK increases DNA copy number variants in vivo. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2219897120. [CrossRef]
9. Liu, G.; Yang, H.; Yuan, X. A shortest path-based approach for copy number variation detection from next-generation sequencing data. *Front. Genet.* **2022**, *13*, 1084974. [CrossRef]
10. Mihola, O.; Kobets, T.; Krivankova, K.; Linhartova, E.; Gasic, S.; Schimenti, J.C.; Trachtulec, Z. Copy-number variation introduced by long transgenes compromises mouse male fertility independently of pachytene checkpoints. *Chromosoma* **2020**, *129*, 69–82. [CrossRef]
11. Ueo, D.; Furuhashi, M.; Sasaki, T.; Kudoh, J.; Parry, D.A.; Winter, D.J.; Sasaki, T.; Hashimoto, T.; Tsuruta, D.; Fujiwara, S. Intragenic copy number variation in mouse epiplakin 1 (Eppk1) and the conservation of the repeat structures in the lower vertebrates. *J. Dermatol. Sci.* **2021**, *103*, 186–189. [CrossRef] [PubMed]
12. White, J.; O'Neill, M.; Sheth, K.; Lamb, D. MP65-12 Murine Rbfox-2 Haploinsufficiency Parallels Congenital Anomalies in Human Patients with Rbfox-2 Copy Number Variants. *J. Urol.* **2020**, *203*, e978. [CrossRef]
13. Keshavarz, M.; Savriama, Y.; Refki, P.; Reeves, R.G.; Tautz, D. Natural copy number variation of tandemly repeated regulatory SNORD RNAs leads to individual phenotypic differences in mice. *Mol. Ecol.* **2021**, *30*, 4708–4722. [CrossRef] [PubMed]

14. Serres-Armero, A.; Davis, B.W.; Povolotskaya, I.S.; Morcillo-Suarez, C.; Plassais, J.; Juan, D.; Ostrander, E.A.; Marques-Bonet, T. Copy number variation underlies complex phenotypes in domestic dog breeds and other canids. *Genome Res.* **2021**, *31*, 762–774. [CrossRef] [PubMed]
15. Cerezo-Echevarria, A.; Kehl, A.; Beitzinger, C.; Müller, T.; Klopfleisch, R.; Aupperle-Lellbach, H. Evaluating the histologic grade of digital squamous cell carcinomas in dogs and copy number variation of KIT Ligand—A correlation study. *Vet. Sci.* **2023**, *10*, 88. [CrossRef]
16. Di Gerlando, R.; Mastrangelo, S.; Sardina, M.T.; Ragatzu, M.; Spaterna, A.; Portolano, B.; Biscarini, F.; Ciampolini, R. A genome-wide detection of copy number variations using SNP genotyping arrays in braque français type pyrénées dogs. *Animals* **2019**, *9*, 77. [CrossRef]
17. Rafter, P.; Gormley, I.C.; Parnell, A.C.; Naderi, S.; Berry, D.P. The Contribution of Copy Number Variants and Single Nucleotide Polymorphisms to the Additive Genetic Variance of Carcass Traits in Cattle. *Front. Genet.* **2021**, *12*, 761503. [CrossRef]
18. Buggiotti, L.; Yudin, N.S.; Larkin, D.M. Copy number variants in two northernmost cattle breeds are related to their adaptive phenotypes. *Genes* **2022**, *13*, 1595. [CrossRef]
19. Kumar, H.; Panigrahi, M.; Saravanan, K.; Rajawat, D.; Parida, S.; Bhushan, B.; Gaur, G.; Dutt, T.; Mishra, B.; Singh, R. Genome-wide detection of copy number variations in Tharparkar cattle. *Anim. Biotechnol.* **2023**, *34*, 448–455. [CrossRef]
20. Wu, Y.; Fan, H.; Jing, S.; Xia, J.; Chen, Y.; Zhang, L.; Gao, X.; Li, J.; Gao, H.; Ren, H. A genome-wide scan for copy number variations using high-density single nucleotide polymorphism array in Simmental cattle. *Anim. Genet.* **2015**, *46*, 289–298. [CrossRef]
21. Yi, G.; Qu, L.; Liu, J.; Yan, Y.; Xu, G.; Yang, N. Genome-wide patterns of copy number variation in the diversified chicken genomes using next-generation sequencing. *BMC Genom.* **2014**, *15*, 962. [CrossRef] [PubMed]
22. Keel, B.N.; Nonneman, D.J.; Lindholm-Perry, A.K.; Oliver, W.T.; Rohrer, G.A. A survey of copy number variation in the porcine genome detected from whole-genome sequence. *Front. Genet.* **2019**, *10*, 737. [CrossRef] [PubMed]
23. Nandolo, W.; Mészáros, G.; Wurzinger, M.; Banda, L.J.; Gondwe, T.N.; Mulindwa, H.A.; Nakimbugwe, H.N.; Clark, E.L.; Woodward-Greene, M.J.; Liu, M. Detection of copy number variants in African goats using whole genome sequence data. *BMC Genom.* **2021**, *22*, 398. [CrossRef] [PubMed]
24. Zhu, C.; Fan, H.; Yuan, Z.; Hu, S.; Ma, X.; Xuan, J.; Wang, H.; Zhang, L.; Wei, C.; Zhang, Q. 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]
25. Saitou, M.; Gokcumen, O. An evolutionary perspective on the impact of genomic copy number variation on human health. *J. Mol. Evol.* **2020**, *88*, 104–119. [CrossRef]
26. Isa, I.I.M.; Jamaluddin, J.; Achim, N.H.; Abubakar, S. Population-specific profiling of CCL3L1 copy number of the three major ethnic groups in Malaysia and the implication on HIV susceptibility. *Gene* **2020**, *754*, 144821.
27. Qi, Y.; Zhou, X.; Bu, D.; Hou, P.; Lv, J.; Zhang, H. Low copy numbers of FCGR3A and FCGR3B associated with Chinese patients with SLE and AASV. *Lupus* **2017**, *26*, 1383–1389. [CrossRef]
28. Liang, X.; Lan, J.; Xu, M.; Qin, K.; Liu, H.; Sun, G.; Liu, X.; Chen, Y.; He, Z. Impact of KIT editing on Coat pigmentation and Fresh Meat Color in Yorkshire pigs. *CRISPR J.* **2022**, *5*, 825–842. [CrossRef]
29. Nowacka-Woszuk, J.; Mackowski, M.; Stefaniuk-Szmukier, M.; Cieslak, J. The equine graying with age mutation of the STX17 gene: A copy number study using droplet digital PCR reveals a new pattern. *Anim. Genet.* **2021**, *52*, 223–227. [CrossRef]
30. Feng, Z.; Li, X.; Cheng, J.; Jiang, R.; Huang, R.; Wang, D.; Huang, Y.; Pi, L.; Hu, L.; Chen, H. Copy number variation of the PIGY gene in sheep and its association analysis with growth traits. *Animals* **2020**, *10*, 688. [CrossRef]
31. Henriksen, R.A.; Zhao, L.; Korneliussen, T.S. NGSNGS: Next-generation simulator for next-generation sequencing data. *Bioinformatics* **2023**, *39*, btad041. [CrossRef] [PubMed]
32. Sedlazeck, F.J.; Rescheneder, P.; Smolka, M.; Fang, H.; Nattestad, M.; Von Haeseler, A.; Schatz, M.C. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **2018**, *15*, 461–468. [CrossRef] [PubMed]
33. Shen, X.; Song, C. Responses of Chinese merino sheep (Junken Type) on copper-deprived natural pasture. *Biol. Trace Elem. Res.* **2021**, *199*, 989–995. [CrossRef] [PubMed]
34. Cheung, S.W.; Bi, W. Novel applications of array comparative genomic hybridization in molecular diagnostics. *Expert Rev. Mol. Diagn.* **2018**, *18*, 531–542. [CrossRef]
35. Zhan, B.; Fadista, J.; Thomsen, B.; Hedegaard, J.; Panitz, F.; Bendixen, C. Global assessment of genomic variation in cattle by genome resequencing and high-throughput genotyping. *BMC Genom.* **2011**, *12*, 557. [CrossRef]
36. Sastre, N.; Mercadé, A.; Casellas, J. SNP+ to predict dropout rates in SNP arrays. *Conserv. Genet. Resour.* **2023**, *15*, 113–116. [CrossRef]
37. Ceccobelli, S.; Landi, V.; Senczuk, G.; Mastrangelo, S.; Sardina, M.T.; Ben-Jemaa, S.; Persichilli, C.; Karsli, T.; Bâlteanu, V.-A.; Raschia, M.A. A comprehensive analysis of the genetic diversity and environmental adaptability in worldwide Merino and Merino-derived sheep breeds. *Genet. Sel. Evol.* **2023**, *55*, 24. [CrossRef]
38. Li, X.; Yang, J.; Shen, M.; Xie, X.-L.; Liu, G.-J.; Xu, Y.-X.; Lv, F.-H.; Yang, H.; Yang, Y.-L.; Liu, C.-B. Whole-genome resequencing of wild and domestic sheep identifies genes associated with morphological and agronomic traits. *Nat. Commun.* **2020**, *11*, 2815. [CrossRef]
39. Mu, F.; Rong, E.; Jing, Y.; Yang, H.; Ma, G.; Yan, X.; Wang, Z.; Li, Y.; Li, H.; Wang, N. Structural characterization and association of ovine Dickkopf-1 gene with wool production and quality traits in Chinese Merino. *Genes* **2017**, *8*, 400. [CrossRef]

40. Roldan, D.; Dodero, A.; Bidinost, F.; Taddeo, H.; Allain, D.; Poli, M.; Elsen, J.M. Merino sheep: A further look at quantitative trait loci for wool production. *Animal* **2010**, *4*, 1330–1340. [CrossRef]
41. Li, S.; Xi, Q.; Zhao, F.; Wang, J.; He, Z.; Hu, J.; Liu, X.; Luo, Y. A highly polymorphic caprine keratin-associated protein gene identified and its effect on cashmere traits. *J. Anim. Sci.* **2021**, *99*, skab233. [CrossRef] [PubMed]
42. Powell, B.a.; Rogers, G. The role of keratin proteins and their genes in the growth, structure and properties of hair. *Exs* **1997**, *78*, 59–148. [PubMed]
43. Parsons, Y.; Piper, L.; Cooper, D. Linkage relationships between keratin-associated protein (KRTAP) genes and growth hormone in sheep. *Genomics* **1994**, *20*, 500–502. [CrossRef] [PubMed]
44. Gong, H.; Zhou, H.; Forrest, R.H.; Li, S.; Wang, J.; Dyer, J.M.; Luo, Y.; Hickford, J.G. Wool keratin-associated protein genes in sheep—A review. *Genes* **2016**, *7*, 24. [CrossRef]
45. Zhou, H.; Gong, H.; Li, S.; Luo, Y.; Hickford, J. A 57-bp deletion in the ovine KAP 6-1 gene affects wool fibre diameter. *J. Anim. Breed. Genet.* **2015**, *132*, 301–307. [CrossRef]
46. Gong, H.; Zhou, H.; Hodge, S.; Dyer, J.M.; Hickford, J.G. Association of wool traits with variation in the ovine KAP1-2 gene in Merino cross lambs. *Small Rumin. Res.* **2015**, *124*, 24–29. [CrossRef]
47. Bai, L.; Zhou, H.; Li, W.; Tao, J.; Hickford, J.G. Exploring variation in ovine KRTAP19-5 and its effect on fine wool fibre curvature in Chinese Tan sheep. *Animals* **2024**, *14*, 2155. [CrossRef]
48. Chai, W.; Zhou, H.; Gong, H.; Wang, J.; Luo, Y.; Hickford, J. Nucleotide variation in the ovine KRT31 promoter region and its association with variation in wool traits in Merino-cross lambs. *J. Agric. Sci.* **2019**, *157*, 182–188. [CrossRef]
49. Chai, W.; Zhou, H.; Gong, H.; Hickford, J.G. Variation in the ovine KRT34 promoter region affects wool traits. *Small Rumin. Res.* **2022**, *206*, 106586. [CrossRef]
50. Li, W.; Bai, L.; Zhou, H.; Zhang, Z.; Ma, Z.; Wu, G.; Luo, Y.; Tanner, J.; Hickford, J.G. Ovine KRT81 Variants and Their Influence on Selected Wool Traits of Commercial Value. *Genes* **2024**, *15*, 681. [CrossRef]
51. Yu, X.; Li, S.; Zhou, H.; Zhao, F.; Hu, J.; Wang, J.; Liu, X.; Li, M.; Zhao, Z.; Hao, Z. Spatiotemporal expression and haplotypes identification of KRT84 gene and their association with wool traits in Gansu Alpine fine-wool sheep. *Genes* **2024**, *15*, 248. [CrossRef] [PubMed]
52. Wang, K.; Li, M.; Hadley, D.; Liu, R.; Glessner, J.; Grant, S.F.; 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] [PubMed]
53. Kim, Y.-M.; Ha, S.-J.; Seong, H.-S.; Choi, J.-Y.; Baek, H.-J.; Yang, B.-C.; Choi, J.-W.; Kim, N.-Y. Identification of copy number variations in four horse breed populations in south korea. *Animals* **2022**, *12*, 3501. [CrossRef] [PubMed]
54. Perry, G.H.; Yang, F.; Marques-Bonet, T.; Murphy, C.; Fitzgerald, T.; Lee, A.S.; Hyland, C.; Stone, A.C.; Hurles, M.E.; Tyler-Smith, C. Copy number variation and evolution in humans and chimpanzees. *Genome Res.* **2008**, *18*, 1698–1710. [CrossRef]
55. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef]
56. Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **2010**, *38*, D355–D360. [CrossRef]
57. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [CrossRef]
58. 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]
59. Molparia, B.; Oliveira, G.; Wagner, J.L.; Spencer, E.G.; Torkamani, A. A feasibility study of colorectal cancer diagnosis via circulating tumor DNA derived CNV detection. *PLoS ONE* **2018**, *13*, e0196826. [CrossRef]
60. Zou, J.; Wang, E. eTumorType, An algorithm of discriminating cancer types for circulating tumor cells or cell-free DNAs in blood. *Genom. Proteom. Bioinform.* **2017**, *15*, 130–140. [CrossRef]
61. Huang, J.; Li, R.; Zhang, X.; Huang, Y.; Dang, R.; Lan, X.; Chen, H.; Lei, C. Copy number veriation regions detection in Qinchuan cattle. *Livest. Sci.* **2017**, *204*, 88–91. [CrossRef]
62. Zhang, C.; Zhao, J.; Guo, Y.; Xu, Q.; Liu, M.; Cheng, M.; Chao, X.; Schinckel, A.P.; Zhou, B. Genome-wide detection of copy number variations and evaluation of candidate copy number polymorphism genes associated with complex traits of pigs. *Front. Vet. Sci.* **2022**, *9*, 909039. [CrossRef] [PubMed]
63. Sun, N.; Zhang, J.; Zhang, C.; Shi, Y.; Zhao, B.; Jiao, A.; Chen, B. Using Illumina Infinium HumanMethylation 450K BeadChip to explore genome-wide DNA methylation profiles in a human hepatocellular carcinoma cell line. *Mol. Med. Rep.* **2018**, *18*, 4446–4456. [CrossRef] [PubMed]
64. Lepamets, M.; Auwerx, C.; Nõukas, M.; Claringbould, A.; Porcu, E.; Kals, M.; Jürgenson, T.; Morris, A.P.; Võsa, U.; Bochud, M. Omics-informed CNV calls reduce false-positive rates and improve power for CNV-trait associations. *Hum. Genet. Genom. Adv.* **2022**, *3*, 100133. [CrossRef]
65. Yang, L.; Xu, L.; Zhou, Y.; Liu, M.; Wang, L.; Kijas, J.W.; Zhang, H.; Li, L.; Liu, G.E. Diversity of copy number variation in a worldwide population of sheep. *Genomics* **2018**, *110*, 143–148. [CrossRef]

66. Di Gerlando, R.; Sutera, A.M.; Mastrangelo, S.; Tolone, M.; Portolano, B.; Sottile, G.; Bagnato, A.; Strillacci, M.G.; Sardina, M.T. Genome-wide association study between CNVs and milk production traits in Valle del Belice sheep. *PLoS ONE* **2019**, *14*, e0215204. [CrossRef]
67. Taghizadeh, S.; Gholizadeh, M.; Rahimi-Mianji, G.; Moradi, M.H.; Costilla, R.; Moore, S.; Di Gerlando, R. Genome-wide identification of copy number variation and association with fat deposition in thin and fat-tailed sheep breeds. *Sci. Rep.* **2022**, *12*, 8834. [CrossRef]
68. Ma, Q.; Liu, X.; Pan, J.; Ma, L.; Ma, Y.; He, X.; Zhao, Q.; Pu, Y.; Li, Y.; Jiang, L. Genome-wide detection of copy number variation in Chinese indigenous sheep using an ovine high-density 600 K SNP array. *Sci. Rep.* **2017**, *7*, 912. [CrossRef]
69. Wang, Z.; Guo, J.; Guo, Y.; Yang, Y.; Teng, T.; Yu, Q.; Wang, T.; Zhou, M.; Zhu, Q.; Wang, W. Genome-wide detection of CNVs and association with body weight in sheep based on 600K SNP arrays. *Front. Genet.* **2020**, *11*, 558. [CrossRef]
70. 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. Genome-wide detection of copy number variations and their association with distinct phenotypes in the world's sheep. *Front. Genet.* **2021**, *12*, 670582. [CrossRef]
71. Navarro Gonzalez, J.; Zweig, A.S.; Speir, M.L.; Schmelter, D.; Rosenbloom, K.R.; Raney, B.J.; Powell, C.C.; Nassar, L.R.; Maulding, N.D.; Lee, C.M. The UCSC genome browser database: 2021 update. *Nucleic Acids Res.* **2021**, *49*, D1046–D1057. [CrossRef] [PubMed]
72. Liu, J.; Zhang, L.; Xu, L.; Ren, H.; Lu, J.; Zhang, X.; Zhang, S.; Zhou, X.; Wei, C.; Zhao, F. Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. *BMC Genom.* **2013**, *14*, 229. [CrossRef] [PubMed]
73. Jenkins, G.M.; Goddard, M.E.; Black, M.A.; Brauning, R.; Auvray, B.; Dodds, K.G.; Kijas, J.W.; Cockett, N.; McEwan, J.C. Copy number variants in the sheep genome detected using multiple approaches. *BMC Genom.* **2016**, *17*, 441. [CrossRef] [PubMed]
74. Hou, Y.; Liu, G.E.; Bickhart, D.M.; Cardone, M.F.; Wang, K.; Kim, E.-s.; Matukumalli, L.K.; Ventura, M.; Song, J.; VanRaden, P.M. Genomic characteristics of cattle copy number variations. *BMC Genom.* **2011**, *12*, 127. [CrossRef]
75. Fontanesi, L.; Beretti, F.; Martelli, P.; Colombo, M.; Dall'Olio, S.; Occidente, M.; Portolano, B.; Casadio, R.; Matassino, D.; Russo, V. A first comparative map of copy number variations in the sheep genome. *Genomics* **2011**, *97*, 158–165. [CrossRef]
76. Goyache, F.; Fernández, I.; Tapsoba, A.S.R.; Traoré, A.; Menéndez-Arias, N.A.; Álvarez, I. Functional characterization of Copy Number Variations regions in Djallonké sheep. *J. Anim. Breed. Genet.* **2021**, *138*, 600–612. [CrossRef]
77. Ma, Y.; Zhang, Q.; Lu, Z.; Zhao, X.; Zhang, Y. Analysis of copy number variations by SNP50 BeadChip array in Chinese sheep. *Genomics* **2015**, *106*, 295–300. [CrossRef]
78. Chen, Y.; Bao, Z.; Liu, M.; Li, J.; Dai, Y.; Wang, F.; Zhang, X.; Zhai, P.; Zhao, B.; Wu, X. Promoter methylation changes in KRT17: A novel epigenetic marker for wool production in angora rabbit. *Int. J. Mol. Sci.* **2022**, *23*, 6077. [CrossRef]
79. Cai, B.; Li, M.; Zheng, Y.; Yin, Y.; Jin, F.; Li, X.; Dong, J.; Jiao, X.; Liu, X.; Zhang, K. MicroRNA-149-Mediated MAPK1/ERK2 Suppression Attenuates Hair Follicle Stem Cell Differentiation. *Hum. Gene Ther.* **2022**, *33*, 625–637. [CrossRef]
80. Liu, C.; He, Y.; Feng, X.; Li, J.; Wang, J. Expression of EPHA5 in lung adenocarcinoma is associated with lymph node metastasis and EGFR mutation. *APMIS* **2022**, *130*, 338–345. [CrossRef]
81. Braga, L.G.; Chud, T.C.; Watanabe, R.N.; Savegnago, R.P.; Sena, T.M.; do Carmo, A.S.; Machado, M.A.; Panetto, J.C.d.C.; da Silva, M.V.G.; Munari, D.P. Identification of copy number variations in the genome of Dairy Gir cattle. *PLoS ONE* **2023**, *18*, e0284085. [CrossRef]
82. Rees, E.; Kirov, G. Copy number variation and neuropsychiatric illness. *Curr. Opin. Genet. Dev.* **2021**, *68*, 57–63. [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.