

RESEARCH ARTICLE

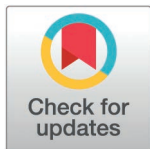
Genome-wide characterization of copy number variants and their functional relevance in indigenous draught cattle of South Asia

Tafara Kundai Mavunga^{1,2*}, Johann Sölkner², Gábor Mészáros², Maano Bryton Malima^{3,4}, Rudolf Pichler¹, Saravanan Ramasamy⁵, Loku Galappattige Sampath Lokugalappatti⁶, Ei Thandar⁷, Menghak Phem⁸, Farai Catherine Muchadeyi¹, Yosra Ressaissi¹, Mario Barbato⁹, Kathiravan Periasamy^{1*}

1 Animal Production and Health Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, Department of Nuclear Sciences and Applications, International Atomic Energy Agency, Vienna, Austria, **2** Institute of Livestock Sciences, Department of Agricultural Sciences, Universität für Bodenkultur Wien, Vienna, Austria, **3** Agriculture Research Council, Onderstepoort, Pretoria, South Africa, **4** South African Medical Research Council Genomics Platform, Tygerberg, South Africa, **5** Tamil Nadu Veterinary and Animal Sciences University, Chennai, India, **6** Faculty of Veterinary Medicine and Animal Science, University of Peradeniya, Peradeniya, Sri Lanka, **7** Livestock Breeding and Veterinary Department, Ministry of Agriculture, Livestock and Irrigation, Yangon, Myanmar, **8** National Animal Health and Production Research Institute, General Directorate of Animal Health and Production, Ministry of Agriculture, Forestry and Fisheries, Phnom Penh, Cambodia, **9** Dipartimento di Scienze Veterinarie, Università degli Studi di Messina, Messina, Italy

* Current address: Indian Council of Agricultural Research-National Research Center on Mithun, Medziphema, Nagaland, India

* Ta.Ku.Mavunga@iaea.org (TKM); kathirvet@yahoo.co.in (KP)



OPEN ACCESS

Citation: Mavunga TK, Sölkner J, Mészáros G, Malima MB, Pichler R, Ramasamy S, et al. (2026) Genome-wide characterization of copy number variants and their functional relevance in indigenous draught cattle of South Asia. PLoS One 21(7): e0353468. <https://doi.org/10.1371/journal.pone.0353468>

Editor: Tofazzal Md Rakib, Chattogram Veterinary and Animal Sciences University, BANGLADESH

Received: January 21, 2026

Accepted: June 23, 2026

Published: July 13, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0353468>

Copyright: © 2026 Mavunga et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/),

Abstract

Copy number variations (CNVs) are an important source of structural genomic variation and contribute to phenotypic diversity in cattle, including traits related to production, reproduction, and adaptation. In this study, we performed a genome-wide characterization of copy number variation regions (CNVRs) in Asian zebu cattle representing eight indigenous draught breeds traditionally used for ploughing, wet-field agriculture, and carting. CNVs were detected using two read-depth-based approaches, CNVnator and CNVcaller, which identified 7,705 and 5,640 CNVs, respectively. Integration of results from both methods using a 50% reciprocal overlap criterion yielded 6,143 CNVRs. The average number of CNVRs per breed ranged from 4,607–5,005, collectively covering approximately 6.23% of the autosomal genome. Of these, 2,697 CNVRs were shared across all breeds, whereas 190 CNVRs were breed specific. Population differentiation based on CNVs, estimated using pairwise V_{ST} statistics, indicated that the Hallikar breed exhibited the highest average differentiation (0.07) relative to other breeds. A total of 4,868 genes overlapped with the identified CNVRs and were enriched for biological processes associated with immune regulation, metabolic function, and adaptive responses. Further comparison with cattle quantitative trait loci identified 65 unique QTLs, predominantly linked to carcass, fertility, reproduction, and growth traits. Overall, this study

which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: The whole genome sequence data generated in the study are available in the form of paired end raw sequences (fastq.gz format) at NCBI under the bio-project accession number PRJNA1358578 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1358578>). The custom scripts used for generating the CNVcaller duplicated window files and reference database are available in the GitHub (<https://github.com/tafarakundai/cnvcaller-custom-reference-db>).

Funding: International Atomic Energy Agency Technical cooperation projects KAM5009, SRL5046 and MYA5022.

Competing interests: The authors have declared that no competing interests exist.

describes the genome-wide distribution and diversity of CNVRs in South Asian indigenous draught cattle and provides baseline genomic information for further investigations into structural variations relevant to breeding, management, and conservation.

1. Introduction

Copy number variants (CNV) in livestock have gained increasing attention in the recent years due to their role in genomic diversity and phenotypic expressions of economically important traits [1]. CNVs are heritable structural variants characterized by deletions or duplications of DNA segments ranging in size from ~1 kilobase (Kb) to several megabases (Mb) [2]. CNVs span a larger proportion of the genome and arise from mechanisms such as unequal recombination, replication errors or genomic rearrangements [3,4]. These repeated genomic segments vary between individuals in a population and contribute to genetic and phenotypic variation. The differences in the copy number of genes result in changes in the gene expression either due to alterations in gene dosage or disruption effect by deletion, duplication, inversion and translocation of DNA sequence [5–7]. CNVs are an important source of variation, and their position largely determines the effect; those located in coding regions alter the protein function, whilst those in regulatory regions alter the level of gene expression [8–10]. Single nucleotide polymorphisms (SNPs) have been routinely used to assess population genetic diversity and genomic predictions for production related traits in cattle. Recent studies [11,12] have indicated SNP markers alone are not sufficient to predict complex traits efficiently and CNVs provide significant additional information to improve the accuracy.

In livestock, several studies have used array-based methods for detection of CNVs [13–17]. However, these approaches have methodological limitations such as limited genomic coverage, low resolution, hybridization noise, limitations in detection of CNVs less than 1kb size and incorrect identification of CNV boundaries [18,19]. Advancement in high throughput sequencing and increased access to next generation sequencing technology has led to improved resolution, sensitivity and the development of new strategies to identify CNVs [20,21]. Whole genome sequencing (WGS) enables the detection of structural variants by identifying abnormal alignment patterns that indicate potential genomic rearrangement breakpoints. Four main approaches are commonly used namely, read pair (RP), split reads (SR), read depth (RD) and assembly based (AS) [18,22–25]. Although each of these methods have been reported to be efficient in detecting CNVs [18], combining one or more approaches can be advantageous in reducing the false positives and improving the accuracy. Various studies have reported the use of these methods in investigating the association between CNVs in specific genes and various phenotypic traits in livestock. For instance, CNVs have been linked to adaptive traits in cattle [26,27], growth performance in sheep [28–30], goats [31,32] and buffaloes [21], indicating their functional significance and impact on economic traits in livestock production.

The zebu cattle breeds of South Asia exhibit extensive genetic diversity in terms of phenotypes and varied utilities such as milk, draught, meat, dung, etc., apart from their adaptability to hot and arid production environments, resistance or tolerance to tropical vectors and diseases [33,34]. Such extensive genetic variations have been shaped by years of adaptation to a wide range of environmental and management conditions and for specific purposes such as draught power [35]. Draught breeds such as Hallikar, Kangayam, Bargur and Deoni in India, Sri Lankan White Cattle, Pyar Zein and Shwe Ni from Myanmar, and Kdarm Red from Cambodia are valued for their strength and draught power, making them ideal for agricultural activities such as ploughing and transportation [36–41]. Hallikar cattle (Karnataka, South India) are known as sturdy and powerful draught animals [36], Kangayam cattle (Tamil Nadu, South India) are known for their ability to work for long hours [42] and Bargur cattle (Tamil Nadu, South India) are adapted to hilly terrain but are relatively temperamental in nature and difficult to train [43,44]. Deoni cattle (Maharashtra, Western India) are a dual-purpose breed used for both draught and milk, and are medium heavy animals found in three-colour variations viz Wannera, Balankaya and Shevera [45]. White Cattle of Sri Lanka are reared inside the forest areas of the Eastern province and are primarily used for draught purposes [46]. Additionally, draught breeds such as Pyar Zein (central and southern regions of Myanmar) and Shwe Ni (central dry zone of Myanmar) are well-suited for wet-field operations, particularly in paddy cultivated areas. Lastly, Kdarm Red from Cambodia (also known as Kor Khmer or Gor Srok) is a dual-purpose breed, primarily used for draught power, but with moderate potential for growth and meat production. The specialized roles of zebu cattle reflect their development in diverse agro-ecological systems, emphasizing the necessity to study the underlying genomic variations that contribute to their adaptation and production potential.

CNVs in cattle have been identified using a range of approaches, including early array-based techniques such as array comparative genomic hybridization (array CGH), quantitative PCR, and fluorescence in situ hybridization [47], as well as more recent high-resolution methods based on SNP arrays and next-generation sequencing (NGS) [48,49]. NGS-based approaches, including read depth, paired-end mapping, and split-read strategies, have improved genome-wide detection and resolution of CNVs [50], enabling population scale analyses across multiple breeds. Despite these advances, the current understanding of CNVs in Asian zebu cattle remains limited. Many studies have focused on a relatively small number of breeds [51–53] or have been biased towards taurine or globally distributed populations such as Nellore, Brahman and Gir [54,55], leading to underrepresentation of indigenous South and Southeast Asian zebu cattle [56]. In addition, inconsistencies in CNV detection methods, reliance on taurine reference genomes, and variability in sequencing depth can affect the accuracy and comparability of results across studies. Consequently, there is a lack of comprehensive, comparative analyses capturing CNV diversity across genetically distinct zebu populations in the region. To address this gap, the present study investigates genome-wide CNV patterns in selected breeds representing South and Southeast Asian zebu cattle. Investigating the role of CNVs in breed differentiation, and adaptation to local agro-ecological conditions may reveal valuable insights for developing sustainable breeding strategies and promoting effective utilization of these native cattle populations. The present study aims to provide new insights into structural genomic variation by conducting a comprehensive whole genome characterization of CNVs in eight South and Southeast Asian zebu cattle populations using the read depth approach and to identify structural variants intersecting key genomic regions associated with production and adaptation traits.

2. Materials and methods

2.1. Animal ethics statement

No animals were specifically sampled for the present study. All genome sequences reported in the study were derived from existing DNA/sample repositories available at Animal Production and Health Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna, Austria. As no live animals were handled and no new samples were collected, ethical approval was not required for this study.

2.2. Inclusivity in global research

Additional information regarding the ethical, cultural, and scientific considerations specific to inclusivity in global research is included in the [S1 File](#).

2.3. Genome resequencing

Whole genome resequencing (WGS) of 40 cattle from eight distinct breeds was performed with five samples per breed, the details of which are shown in [Table 1](#). All WGS data generated in the study are available in the form of paired end raw sequences (fastq.gz format) at NCBI under the BioProject accession number PRJNA1358578 (<https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA1358578>). As indicated under the Ethics Statement, all the DNA samples used in the present study were derived from the existing Genetic Repository at Animal Production and Health Laboratory, Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture, International Atomic Energy Agency, Vienna, Austria. The DNA available at the FAO/IAEA repository were extracted from blood samples collected by jugular venipuncture into EDTA vacutainer tubes and isolated using MasterPure DNA Purification Kit (Biozym, Illumina Inc, USA). It was ensured that the samples included in the study were collected from unrelated cattle using the information available in the Genetic Repository module of FAO/IAEA Genetic Laboratory Information and Data Management System (GLIDMaS). The DNA samples were subjected to quality control using Nanodrop 2000 spectrophotometer (ThermoScientific, USA) initially to estimate concentration, followed by Quant-iT™ PicoGreen™ assay to ascertain the quantity of good quality double strand DNA. Whole genome resequencing was performed using Illumina Novaseq 6000 using paired end 150 bp reads with a target coverage of ~20x per sample at Neogen Europe Ltd., Ayrshire, Scotland.

2.4. Quality control of raw sequence reads and alignment

Raw sequencing reads in FASTQ format were initially assessed using FastQC v0.11.9 [57] to evaluate per-base quality scores, GC content and adapter contamination. Reads were then pre-processed with fastp v0.23.2 [59] to remove adapter sequences, trim low-quality bases (Phred <20), and discard short reads (<50 bp), with all other parameters set to default. Post-trimming quality was reassessed using FastQC to confirm improvement in read quality and removal of adapter contamination. High-quality reads were aligned to the indexed Zebu cattle reference genome NIAB_ARS_BosIndicus_Tharparkar_1.0 (NCBI RefSeq assembly, GCF_029378745.1) [58] using BWA-MEM v0.7.17 [59] with default parameters. The average mapping rate was 99% and the average sequencing coverage was 21.2×. The alignment output files in sequence alignment map (SAM) format were sorted, indexed, and converted to binary alignment map (BAM) format using sequence alignment/map tools (SAMtools v1.9) [60]. The potential duplicate reads were marked using Picard tools (<https://broadinstitute.github.io/picard>). The resulting BAM files with marked duplicates were used as input for CNV calling.

Table 1. Description of samples used in the study.

Country	Breed	Number of samples	Type
India	Bargur	5	draught
	Deoni	5	draught/milk
	Hallikar	5	draught
	Kangayam	5	draught
Sri Lanka	White Cattle	5	draught
Cambodia	Kdarm Red	5	draught/meat
Myanmar	Pyar Sein	5	draught
	Shwe Ni	5	draught

<https://doi.org/10.1371/journal.pone.0353468.t001>

2.5. CNV and CNVR detection

A read depth approach was used to detect copy number variations. To increase the reliability of CNV detection, CNVnator [61] and CNVcaller [62] were used. CNVnator utilizes fixed-bin read-depth histograms for sensitive detection of CNVs whereas CNVcaller applies population-based normalization and segmentation models to improve breakpoint resolution and reduce false positives. CNV detection using CNVnator was carried out for the autosomal chromosomes only, using a bin size of 100 bp, following the recommendations of Abyzov et al., [61]. The following steps were run; CNVnator -tree, CNVnator -his, CNVnator -stat, CNVnator -partition and CNVnator -call and identification of CNV occurrence in the preset window was done. The software calls CNVs per individual and to enhance the CNV prediction quality control, raw CNVs were filtered by retaining only those with $q_0 < 0.5$, size $> 1\text{kb}$ and $p\text{-value} < 0.01$ based on t-test statistics. A q_0 filter removes any CNV calls with $q_0 > 0.5$ and the metric represents the fraction of reads within a region that have zero mapping quality which means reads with multiple mapping locations in the genome. This is crucial because regions with $q_0 > 0.5$ are likely enriched for ambiguously mapped reads that potentially lead to false positive CNV detection. To compare the results from CNVcaller, individual CNVs were merged using 50% reciprocal overlap to create population based CNVRs.

CNVcaller is a population-based CNV detection software, that analyses population sequencing data across individuals simultaneously rather than independently. This approach leverages on using population-level read depth information to perform local depth normalization and segmentation across samples (adaptive binning instead of fixed bin size used in CNVnator), resulting in reduced false positives and improved breakpoint accuracy. Initially, the reference genome used in the present study (NIAB_ARS_BosIndicus_Tharparkar_1.0) lacked the duplicated window record files required for CNVcaller's absolute number correction. Hence, custom duplicated window record files were generated following the official CNVcaller pipeline (<https://github.com/JiangYuLab/CNVcaller>). Briefly, the reference database for *Bos indicus* was created using NIAB_ARS_BosIndicus_Tharparkar_1.0 which was segmented into k-mer sequences (overlapping windows) using the CNVcaller script 0.1.Kmer_Generate.py with a recommended window size of 800 bp [62] to optimize CNV detection and boundary resolution. The resulting k-mer FASTA file was then aligned back to the reference genome using BLASR after constructing the suffix array index with sawriter. BLASR alignment was performed using the parameters recommended in the CNVcaller documentation for sensitive k-mer mapping. Finally, duplicated window record files were generated from the BLASR alignment [63] output using 0.2.Kmer_Link.py, producing the genome-wide window linkage file required for downstream CNV analysis. The resulting reference database and duplicated window files were then used for CNV detection in CNVcaller. The custom scripts used for generating the CNVcaller duplicated window files and reference database are available in the GitHub (<https://github.com/tafarakundai/cnvcaller-custom-reference-db>). CNVcaller calculated the GC-corrected normalized read depth for each sample to standardize the copy number in each window and classify different genotypes. The following CNVcaller parameters were used for filtering: -f 0.1 to set the frequency threshold of 10% for downstream analysis, -h 3 to retain CNVs shared in at least three samples, and -r 0.5 to set the minimum reciprocal overlap to be 50% between CNVs for merging or creating population based CNVRs. Additionally, CNVs with a Silhouette score of > 0.5 and those located in unplaced scaffolds were excluded from further analysis. It is important to note that the CNVcaller filtering thresholds (-f 0.1, -h 3) used in this study prioritize high confidence CNVRs and might have reduced sensitivity to rare or low frequency variants.

2.6. Population differentiation based on CNVR

V_{ST} of CNVRs between a pair of populations was calculated as $V_{ST} = V_T - V_S / V_T$ where V_T is the total variance of copy numbers among two breeds and V_S is the average variance within each population, weighted by the number of individuals in the population. The pairwise differentiation among breeds was visualized using a heatmap generated by a custom python script.

2.7. Functional annotation and enrichment analysis of genes overlapping CNVRs

CNVRs identified by both CNVcaller and CNVnator were merged using 50% reciprocal overlap between regions. The merged CNVRs were annotated using ANNOVAR [64] with a custom reference gene database derived from NIAB_ARS_BosIndicus_Tharparkar_1.0 gene transfer format (GTF) file. This GTF file was converted into ANNOVAR refGene format using a custom Python script. ANNOVAR then produced gene-based annotations to identify overlapping transcripts with the CNVRs. Functional analysis of genes overlapping CNVRs for enrichment of Gene Ontology-Biological Process (GOBP) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway was conducted using ShinyGO v0.82 [65]. The list of background gene set used to perform enrichment analysis is provided in the [S2 File](#). A threshold of FDR < 0.01 and Fold Enrichment > 2.5 was used to identify significantly enriched GOBP and KEGG pathways.

2.8. Identification of QTLs associated with CNVRs

The CNVRs detected across all populations as well as those shared among breeds were searched for overlap with Quantitative Trait Loci (QTL) reported in cattle using Animal QTL database, that consisted of 194,095 cattle QTLs or associations [66]. Regions overlapping with known QTLs, and their function associated with QTLs were identified.

3. Results

An average of 421.6 million paired end reads per sample was generated across different breeds of cattle. Evaluation of the quality of sequence data confirmed that all the investigated samples possessed sufficient depth and integrity for the detection of copy number variations. Following quality control and adapter trimming, approximately 99% of reads were successfully aligned to the *Bos indicus* reference genome (NIAB_ARS_BosIndicus_Tharparkar_1.0), indicating minimal alignment bias and high sequence quality. The mean genome-wide coverage was $21.2\times$ providing reliable CNV detection using read depth-based methods. The uniform sequencing depth across the samples is critical for precise CNV calling, and the narrow variation observed across the dataset reflects consistency in sequencing and data quality. A detailed summary of sequencing and alignment statistics for each sample is presented in [Table 2](#).

3.1. Number, distribution and genomic landscape of identified CNVRs

Genome wide CNV detection using CNVnator identified an average of 4627 CNVs per animal (min: 3681, max: 9254, median: 4245). After post-processing, the fraction of mapped reads with zero quality ($q0 < 0.5$), size > 1kb and p-value < 0.01 based on t-test statistics (e-val1 generated by CNVnator, a statistical t-test significance of the read depth difference between the candidate CNV region and the genome-wide background read depth distribution), we retained 7705 autosomal CNVRs. Similarly quality control filtering of CNVcaller output using the silhouette score of > 0.50 retained a total 5640 CNVRs. A sensitivity analysis was performed to assess the effect of different overlap thresholds on CNVR detection while merging the datasets ([S3 File](#)). The total number of detected CNVRs varied only modestly across thresholds, ranging from 5,992 at 30% overlap to 6,536 at 90% overlap. Hence, the 50% reciprocal overlap threshold was retained for the subsequent analysis as it provided a balance between over-merging distinct CNV events at lower thresholds and excessive fragmentation of CNVRs at higher thresholds. After merging both datasets using, a > 50% reciprocal overlap between regions, a total of 6143 high-confidence CNVRs were identified, of which 3048 were deletions, 1480 were duplications and 1615 were mixed events (both). The highest number of CNVRs was observed in Hallikar cattle breed ($n = 5005$) while the lowest ($n = 4607$) was observed in Kangayam cattle ([Table 3](#)). The total length of detected CNVRs across breeds ranged from 114.99 Mb to 163.27 Mb with average coverage of 6.23% of autosomal genome per breed, consistent with previous cattle CNV reports that ranged from 4–12% [26,27,53]. The overall mean length of identified CNVRs was 27 kb and the overall median length was 42 kb. CNVR sizes showed a right-skewed, indicating the presence of both small and moderately large CNVR events. The distribution of CNVRs based on size ([Fig 1a](#)) shows a greater number of CNVRs

Table 2. Sample-wise sequencing metrics of cattle investigated in the study.

Breed (Country)	Sample	Total Reads	Mapped Reads	Coverage
Bargur (India)	IBR1	434,047,178	99.30%	22.9
	IBR2	408,639,807	99.85%	21.0
	IBR3	399,606,295	99.89%	20.4
	IBR4	419,549,818	99.86%	21.2
	IBR5	425,540,206	99.88%	23.1
Deoni (India)	IDN1	434,140,502	99.90%	22.3
	IDN2	405,129,807	99.90%	20.6
	IDN3	478,578,091	99.75%	25.0
	IDN4	393,221,955	99.92%	19.7
	IDN5	379,314,501	99.87%	19.8
Hallikar (India)	IHL1	446,802,742	99.86%	21.9
	IHL2	436,874,056	99.67%	20.5
	IHL3	385,803,302	99.87%	18.5
	IHL4	434,986,280	99.52%	20.4
	IHL5	418,039,922	99.82%	20.2
Kangayam (India)	IKA1	433,903,903	99.86%	22.5
	IKA2	393,250,124	99.67%	20.2
	IKA3	397,852,521	99.87%	20.4
	IKA4	395,321,367	99.52%	20.4
	IKA5	463,612,943	99.82%	24.4
Kdarm Red (Cambodia)	KDR1	380,089,348	99.94%	20.3
	KDR2	421,341,141	99.93%	21.8
	KDR3	442,710,443	99.69%	24.9
	KDR4	439,812,679	99.94%	22.9
	KDR5	346,874,615	99.93%	17.9
White Cattle (Sri Lanka)	LWC1	502,173,090	99.57%	27.0
	LWC2	419,534,182	99.65%	22.7
	LWC3	398,183,637	99.93%	20.7
	LWC4	496,578,376	99.87%	25.4
	LWC5	459,735,216	99.70%	24.9
Pyar Zein (Myanmar)	YPZ1	392,433,222	99.94%	20.2
	YPZ2	426,388,364	99.92%	22.0
	YPZ3	413,520,511	99.63%	21.1
	YPZ4	448,883,061	99.93%	23.3
	YPZ5	426,713,619	99.94%	22.1
Shwe Ni (Myanmar)	YSN1	422,575,653	99.67%	21.8
	YSN2	407,604,203	99.93%	21.1
	YSN3	417,338,676	99.94%	21.4
	YSN4	407,229,133	99.69%	20.9
	YSN5	396,198,340	99.64%	20.3
Overall Mean $\pm$ SE	—	426,141,071 $\pm$ 2,019,830	99.71 $\pm$ 0.18%	21.63 $\pm$ 0.36

<https://doi.org/10.1371/journal.pone.0353468.t002>

Table 3. Details of Copy Number Variation regions (CNVRs) identified in draught cattle breeds.

Breed	Samples	CNVRs	Total length	Median length	DUP	DEL	Mixed
Bargur	5	4802	127,973,498	4199.0	1336	2947	519
Deoni	5	4692	120,732,008	4199.0	1249	2964	479
Hallikar	5	5005	163,271,395	4599.0	1638	2899	468
Kangayam	5	4607	115,964,793	4099.0	1228	2937	442
Kdarm Red	5	4755	117,829,545	3999.0	1263	2964	528
White Cattle	5	4707	114,987,393	4299.0	1194	3031	482
Pyar Zein	5	4735	118,022,065	4199.0	1217	2992	526
Shwe Ni	5	4660	118,606,940	4199.0	1239	2904	517

<https://doi.org/10.1371/journal.pone.0353468.t003>

($n=3259$) to be under 5 kb. The distribution of CNVRs across autosomes (Fig 1b) showed chromosome 1 had the highest number of CNVRs ($n=350$) covering a length of 4.7 Mb and the lowest number of CNVRs were found in chromosome 27 ($n=83$) covering a length of 2.4 Mb. After normalizing by chromosome length, CNVR density ranged from 1.82 CNVRs per Mb (chromosome 27) to 3.95 CNVRs per Mb (chr 29), and a mean CNVR density revealing that smaller autosomes harboured proportionally more CNVRs than the longer ones (Fig 1c). The results of annotation analysis (Fig 1d,e) revealed 49.9% of identified CNVRs were in intergenic regions, 22.2% in the exonic regions, and 17.2% in intronic regions.

3.2. Population differentiation based on CNVRs

Population differentiation among breeds was evaluated using the V_{ST} statistic derived from autosomal CNVRs (Fig 2). Pairwise V_{ST} values ranged from 0.002 to 0.084, reflecting generally low to moderate levels of population differentiation (Fig 2). The Hallikar (IHL) breed from India showed relatively high overall differentiation from other investigated breeds, with a mean pairwise V_{ST} of 0.072 across all comparisons. Among all the breed-pair combinations, the largest divergence was observed between Hallikar and Kangayam (IKA) ($V_{ST}=0.084$), indicating notable divergence in copy number variation between these two indigenous South Indian draught type breeds. It is also noteworthy to mention that Kangayam exhibited relatively low levels of differentiation, with V_{ST} values ranging from 0.021 to 0.030 across different pairwise comparisons. Conversely, the Southeast Asian breeds, Kdarm Red (KDR) from Cambodia and Shwe Ni (YSN) and Pyar Zein (YPZ) breeds from Myanmar showed lowest differentiation ($V_{ST}=0.002$ to 0.003), suggesting CNVR similarity likely arising from shared ancestry or continuous gene flow. Most other breed pairs exhibited low differentiation ($V_{ST}=0.008$ to 0.012), indicating broadly conserved CNVR patterns among them. Overall, the CNV-based population structure was modest, reflecting limited differentiation in CNVR profiles among South Asian draught cattle populations.

3.3. Breed-specific CNVRs

The breeds analysed in the present study are predominantly draught type cattle and hence CNVRs detected across all 40 individuals were filtered to identify both breed-specific and shared variants. A total of 6143 CNVRs were identified across the autosomes, of which 2697 (43.9%) were shared among all the eight breeds representing the core CNVR set (Fig 3). The high proportion of shared CNVRs indicates significant genomic overlap and a likely history of admixture or common ancestral origin among these indigenous populations. Among the core CNVRs, 1438 were deletions, 654 were duplications, and 605 were mixed events. Breed-specific CNVRs accounted for a small fraction of the total variation, with 190 CNVRs uniquely detected in individual breeds (Fig 3). Among these, Hallikar had the highest number of breed-specific CNVRs ($n=170$). This enrichment may reflect differences in breed history or genomic diversity, however, it may also be influenced by factors such as sample representation or detection sensitivity and therefore should be interpreted with caution. In contrast, the remaining breeds exhibited very few unique CNVRs: Kdarm Red ($n=8$), Deoni ($n=4$),

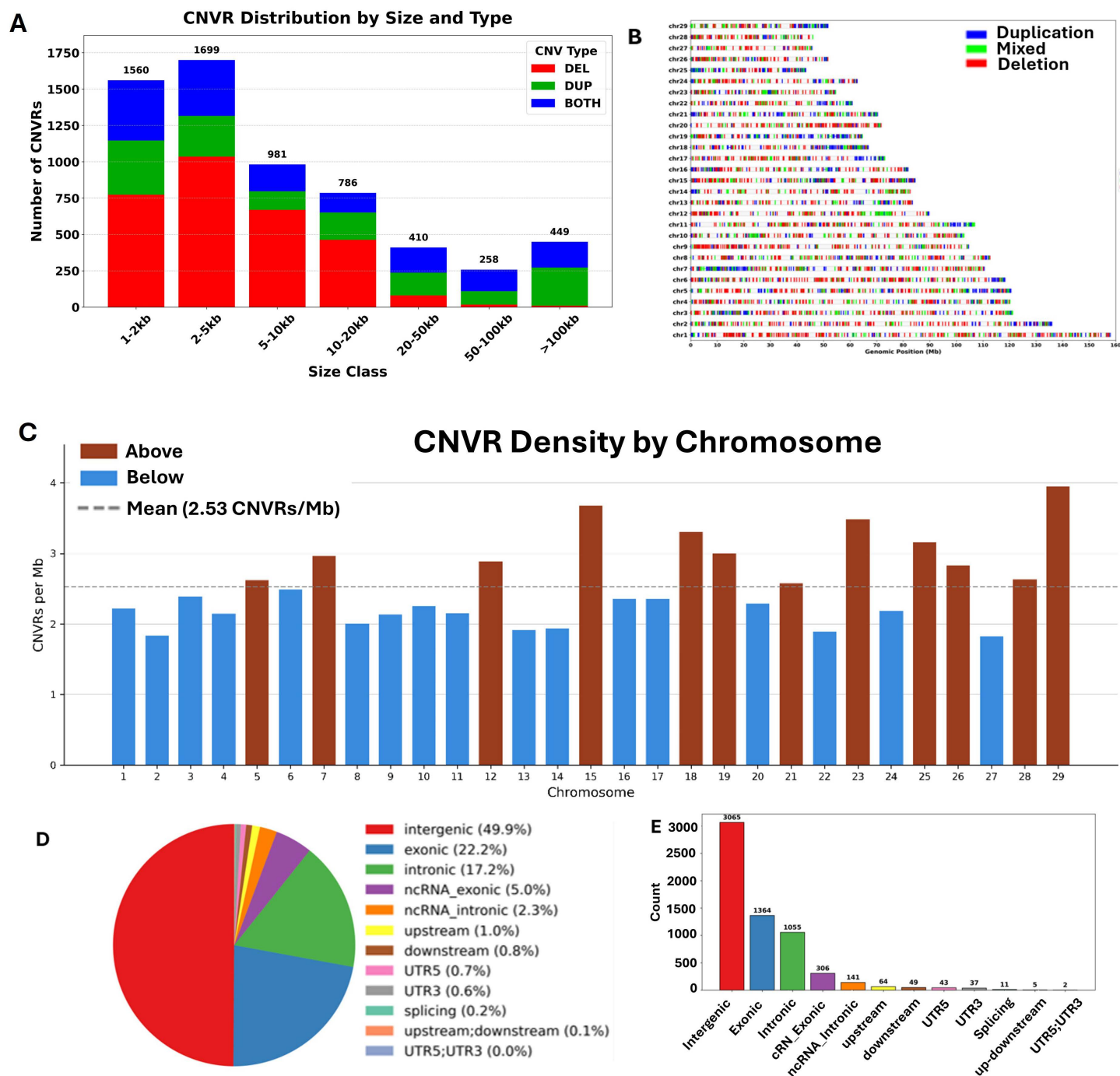


Fig 1. Genomic distribution of Copy Number Variation Regions (CNVRs) (a) the count and length categories of detected CNVRs (b) The autosomal distribution of CNVRs across all breeds (c) CNVR counts by chromosome length reported as CNVR density (number of CNVRs per Mb) (d) CNVRs annotations with various genomic features (e) count distribution of annotated CNVRs.

<https://doi.org/10.1371/journal.pone.0353468.g001>

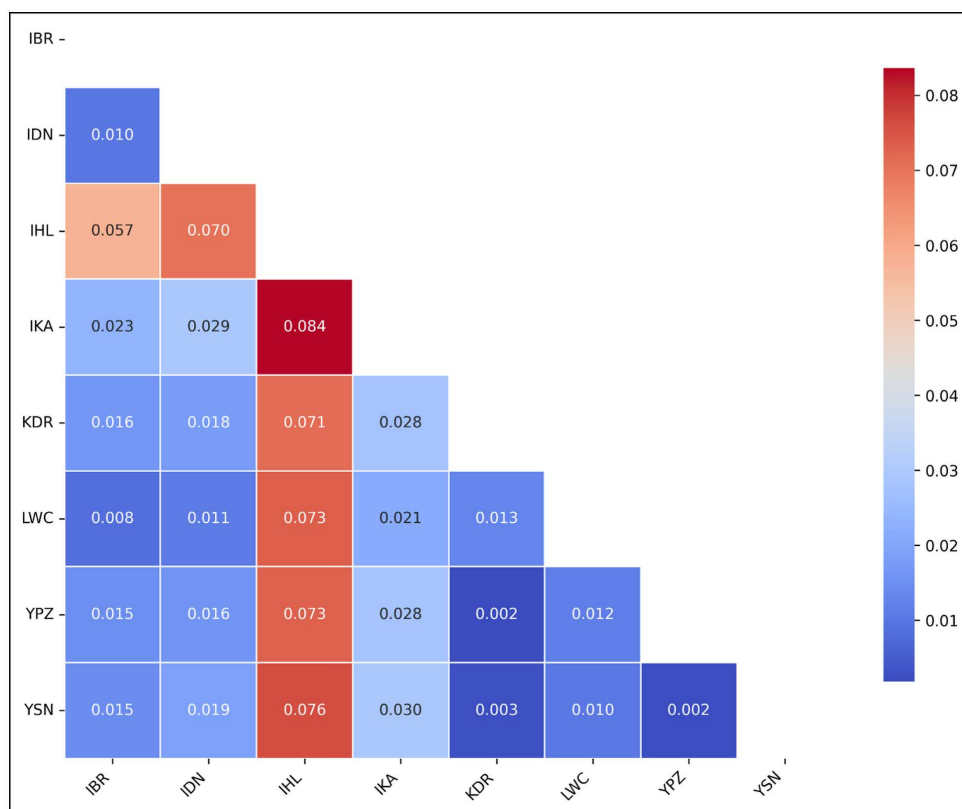


Fig 2. Mean pairwise V_{ST} estimates among cattle breeds based on autosomal CNVRs.

<https://doi.org/10.1371/journal.pone.0353468.g002>

Kangayam (n=4), Pyar Zein (n=2), Bargur (n=1), White Cattle (n=1), and Shwe Ni (n=0). The relatively low number of unique CNVRs suggests limited detectable divergence in CNVR profiles among the zebu populations. However, given the limitations of read depth-based CNV detection, the absence of CNVRs in some breeds should not be interpreted as true biological absence, and the observed pattern may reflect both their shared utility as draught animals under similar agro-ecological conditions and methodological constraints.

3.4. Functional annotation of CNVR overlapping genes/QTLs

Functional annotation of all 6143 CNVRs (detected across all eight breeds) using *Bos indicus* reference genome resulted in identification of 4868 genes overlapping the structural variants, where a gene was considered to overlap a CNVR if any portion of its genomic coordinates intersected with the CNVR boundaries. The core set of 2697 shared CNVRs overlapped with 2381 genes while the 170 Hallikar specific CNVRs annotated to 618 genes. Functional analysis for detection of enriched Gene Ontology Biological Processes (GOBP) and KEGG pathways using 4868 genes spanning across all the CNVRs were conducted, the results of which are summarized in Figs 4, 5 and (S4 File).

For the complete dataset comprising 4,868 genes overlapping 6,143 CNVRs, enrichment analysis identified 19 GOBP terms with fold enrichment greater than 2.5 (Fig 4A). The enriched categories included processes related to immune regulation, cytokine production, lipid metabolism, cellular stress response, cell adhesion, and developmental regulation. The representative GOBP terms included positive regulation of macrophage cytokine production, regulation of somitogenesis, positive regulation of xenophagy, postsynaptic density organization, cell–cell adhesion mediated by cadherin, and

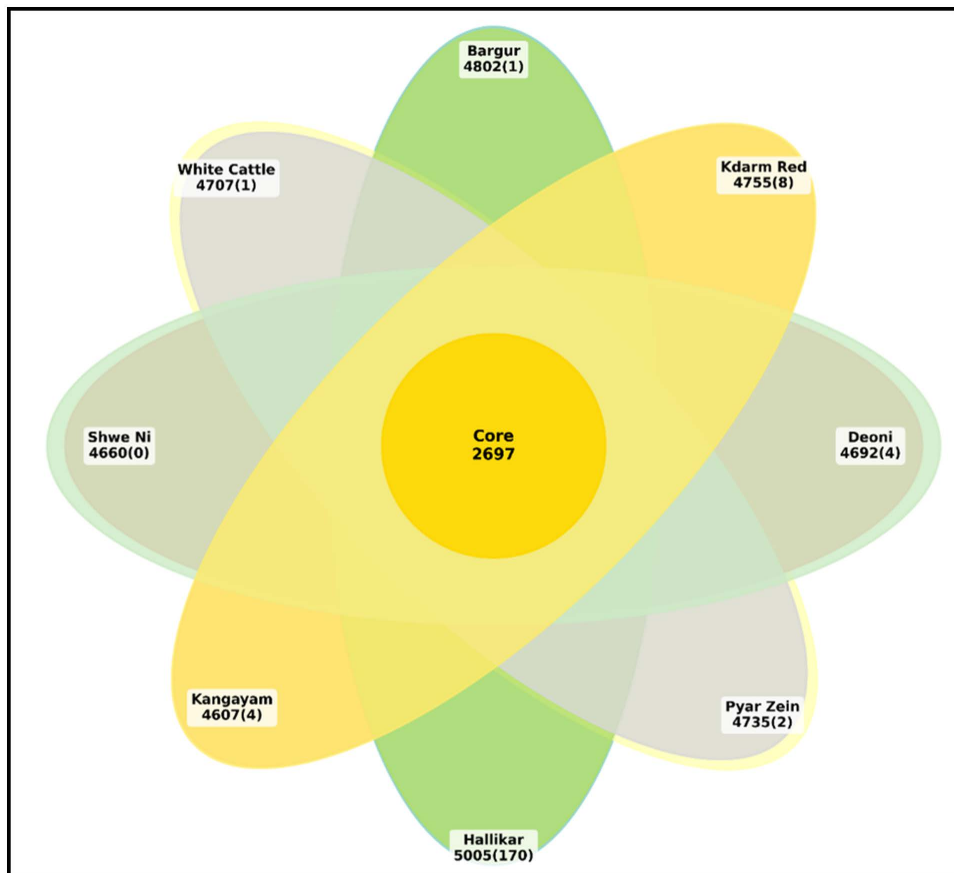


Fig 3. Flower plot diagram showing shared and breed-specific copy number variation regions (CNVRs) across draught cattle breeds of South Asia.

<https://doi.org/10.1371/journal.pone.0353468.g003>

response to UV-B radiation. The 2,381 genes overlapping the core CNVR set revealed 24 significantly enriched GOBP terms (Fig 4B), including categories associated with neuronal development, synaptic organization, vesicle-mediated transport, and signal transduction. Enriched terms included neuron projection guidance, axon guidance, synaptic vesicle cycling, and calcium-ion regulated exocytosis. In the Hallikar breed-specific CNVRs, eight GOBP terms were significantly enriched (Fig 4C), mainly involving morphogenesis, intracellular signalling, and metabolic regulation. The prominent GOBP terms included anatomical structure morphogenesis, regulation of intracellular signal transduction, and positive regulation of cellular metabolic process.

KEGG pathway enrichment analysis identified several pathways associated with signalling, neuroendocrine regulation, and metabolism. In the complete dataset, seven pathways were significantly enriched (Fig 5A), including glutamatergic synapse, parathyroid hormone synthesis and secretion, long-term depression, gonadotropin-releasing hormone (GnRH) secretion, insulin secretion, VEGF signalling, and mucin-type O-glycan biosynthesis. The core CNVR set showed enrichment of two pathways, namely neuroactive ligand–receptor interaction and MAPK signalling (Fig 5B). Similarly, Hallikar-specific CNVRs showed enrichment for MAPK signalling and aldosterone synthesis and secretion pathways (Fig 5C).

To examine the potential relationship between CNVRs and economically important traits, CNVR coordinates from the complete, core, and Hallikar-specific datasets were intersected with cattle QTLs available in the Animal QTLdb [66]. With the complete set of 6,143 CNVRs, 65 unique QTLs were identified (Table 4), including QTLs related to carcass traits (33

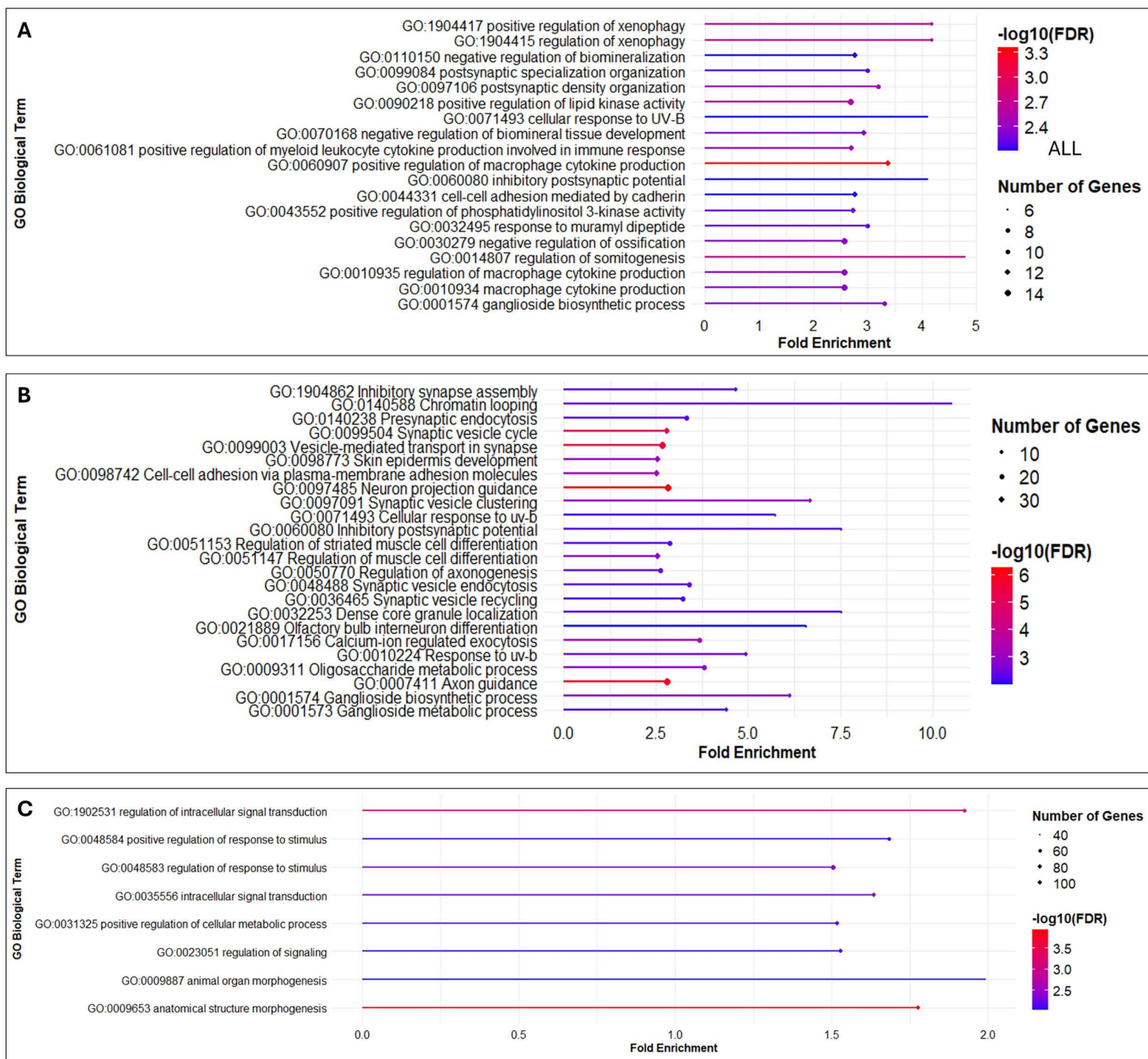


Fig 4. Top GO biological processes significantly enriched ($p < 0.05$) for (A) all CNVRs (B) core CNVRs shared across all breeds and (C) Hallikar-specific CNVRs.

<https://doi.org/10.1371/journal.pone.0353468.g004>

QTLs), fertility and reproduction [13], growth [6], disease susceptibility and immune-related traits [4], milk mineral composition [3], and other miscellaneous traits [6]. The overlapping QTLs included those associated with antral follicle number, Anti-Müllerian hormone levels, average daily gain, carcass weight, somatic cell score, subcutaneous fat thickness, muscularity, and mineral composition traits. The core set of 2,697 shared CNVRs overlapped with 40 unique QTLs (Table 4)

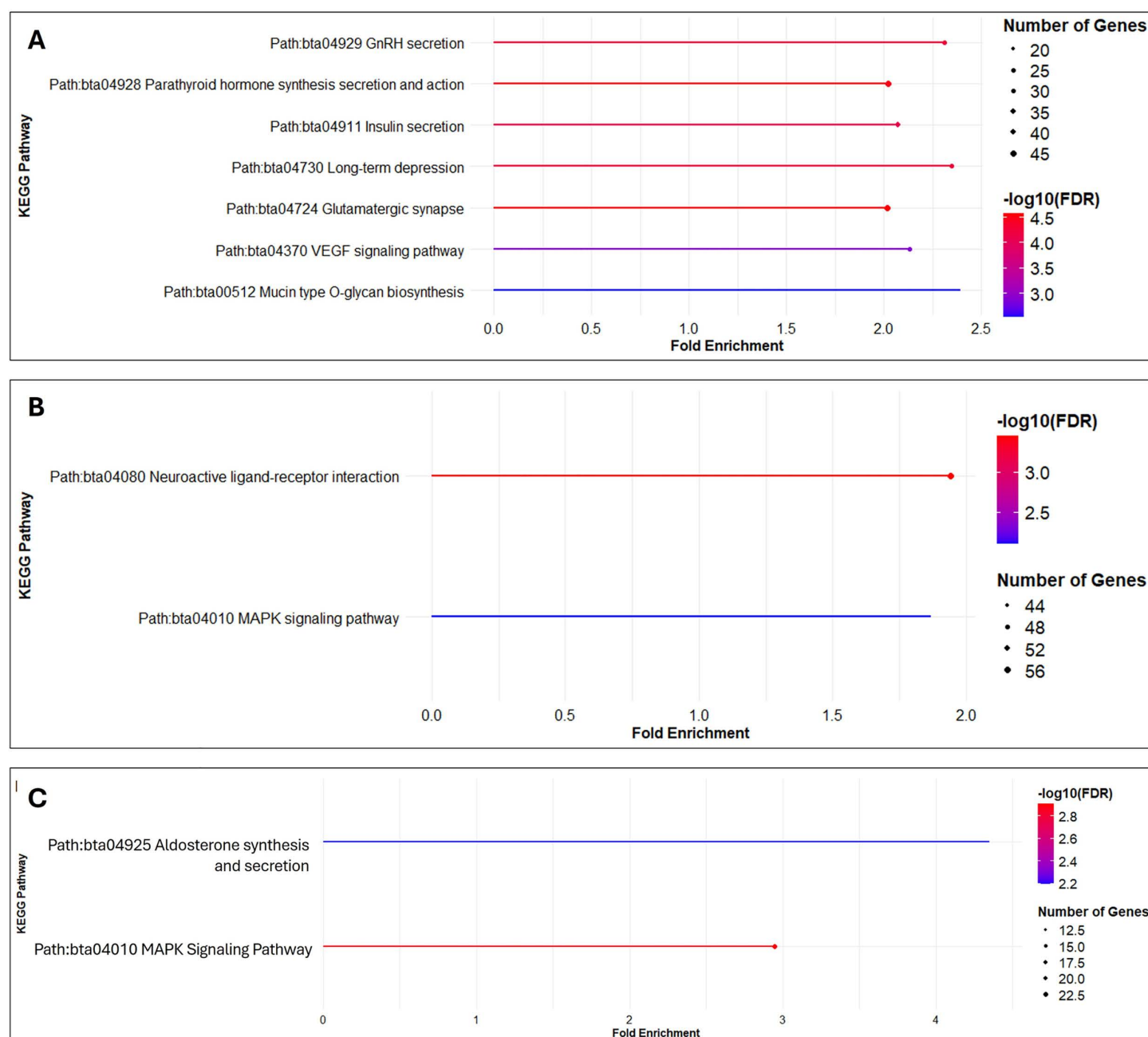


Fig 5. Top KEGG pathways significantly enriched ($p < 0.05$) for (A) all CNVRs (B) core CNVRs shared across all breeds (C) Hallikar-specific CNVRs.

<https://doi.org/10.1371/journal.pone.0353468.g005>

while the 170 Hallikar-specific CNVRs overlapped with nine unique QTLs, comprising five carcass-related QTLs, three QTLs associated with mineral composition traits, and one growth-related QTL. Overall, the QTL overlap analysis identified CNVRs located within genomic intervals previously reported to be associated with economically important traits in cattle.

Table 4. Summary of QTLs overlapping with ALL, core (shared) and Hallikar-specific CNVRs.

Quantitative trait Category	Quantitative trait	No. QTLs					
		ALL CNVRs	Total	Core CNVRs	Total	Hallikar-specific CNVRs	Total
Carcass traits	Carcass weight	1	33	—	21	—	5
	Finishing precocity	1		—		—	
	Juiciness	1		1		—	
	Longissimus muscle area	6		4		2	
	Marbling score	3		3		—	
	Meat color	2		1		—	
	Meat firmness	1		1		—	
	Tenderness score	2		2		—	
	Subcutaneous fat thickness	16		9		3	
Growth	Average daily gain	1	6	—	3	—	1
	Body weight	1		—		—	
	Conformation score	1		1		—	
	Muscularity	3		2		1	
Fertility & Reproductive traits	Anti-Müllerian hormone level	5	13	3	8	—	
	Antral follicle number	8		5		—	
Disease susceptibility/Resistance	M. paratuberculosis susceptibility	3	4	3	4	—	
	Somatic cell score	1		1		—	
Milk composition (Minerals)	Milk calcium content	1	3	—		1	3
	Milk magnesium content	1		—		1	
	Milk phosphorus content	1		—		1	
Others	Blood calcium level	2	6	1	4	—	
	Intestinal atresia	2		1		—	
	Mean corpuscular hemoglobin concentration	1		1		—	
	Ketosis	1		1		—	
Total		65	65	40	40	9	9

<https://doi.org/10.1371/journal.pone.0353468.t004>

4. Discussion

Genomic diversity in zebu cattle of South Asia has been studied extensively using single nucleotide polymorphisms (SNPs) [35,67,68], microsatellite and mitochondrial DNA markers [39,69]. However, comprehensive information on the diversity and distribution of structural variations in Asian zebu cattle genome remains limited. Most of the CNV analysis in zebu breeds relied on high quality taurine reference genome assemblies such as ARS-UCD1.2 [70], UMD3.1 [71], Btau_4 [72], and UOA_Angus_1 [73]. The recent availability of a *Bos indicus* reference genome NIAB_ARS_BosIndicus_Tharparkar_1.0 (GCF_029378745.1) provided an improved framework for CNV detection by enhancing breakpoint resolution and reducing reference bias. Vani et al [53] investigated CNVs using *Bos indicus* Nellore reference genome (GCF_000247795.1_Bos_indicus_1.0) [74] in 15 pooled samples belonging to five zebu draught breeds. In the present study we expanded the dataset to include 40 individuals from eight draught breeds across India, Sri Lanka, Myanmar and Cambodia, thereby broadening the geographic and genetic representation of South and Southeast Asian Zebu populations.

CNVRs were identified with high reliability with an average sequencing depth of $21.2\times$, which is higher than $8\text{--}10\times$ reported for earlier South Asian cattle studies [53,23,49,75]. The number of CNVRs detected in the present study is broadly comparable to previous reports across diverse cattle populations, including Asian and Indian draught breeds (11,065 CNVRs) [53], indigenous Chinese cattle populations (ranging from 1,651–9,349 CNVRs) [76], Hainan cattle

(5,458 CNVRs) [77], Pinan and Nanyang cattle (9,631 CNVRs) [78], Nellore cattle (1,884 CNVRs) [54], Tharparkar cattle (447 CNVRs) [52], and Ethiopian indigenous cattle breeds (3,893 CNVRs) [27]. The CNV-based genomic architecture inferred in this study is consistent with earlier investigations, particularly with respect to autosomal CNV coverage which accounted for approximately 6.3% of the genome. This estimate falls within the range reported for most cattle populations [8,76,78,79], although it is lower than the substantially higher CNV coverage (11–17%) reported in draught cattle populations of Tamil Nadu, India [53]. Furthermore, a higher proportion of deletions relative to duplications was observed, a pattern that has been consistently reported in CNV studies across cattle and other mammalian species [12,27,31,80].

Population differentiation assessed using CNVR derived V_{ST} values revealed low to moderate but significant structural genomic divergence among the investigated draught cattle populations. The observed range of pairwise V_{ST} values (0.002–0.084) indicated a relatively low magnitude of differentiation, consistent with the shared zebu ancestry and common evolutionary pressures associated with draught utility and adaptation to tropical agro-ecological environments. While most breeds share broadly conserved CNVR patterns, certain breeds showed relatively higher differentiation. Specifically, Hallikar cattle from southern India displayed the highest observed value for CNV based differentiation, as evidenced by the highest mean pairwise V_{ST} values across comparisons. The significant divergence observed between Hallikar and Kangayam cattle, despite their geographic proximity, may indicate distinct selective breeding practices resulting in differential accumulation of structural variants. This divergence could reflect long-term selection for functional traits such as endurance, gait efficiency, or workload specialization, which are known to vary subtly among traditional draught breeds. In contrast, the Southeast Asian breeds, Kdarm Red from Cambodia and Shwe Ni and Pyar Zein from Myanmar showed low levels of CNV based differentiation, indicative of highly similar CNVR profiles. This genomic homogeneity likely reflects historical relatedness or continuous gene flow among them, potentially facilitated by transboundary livestock movement and overlapping production systems. Similarly, most Indian indigenous breeds and Sri Lankan cattle exhibited low pairwise V_{ST} values, pointing to broadly conserved structural genomic variations among them. The limited differentiation observed among these populations suggests that adaptive divergence may be concentrated in a small subset of CNVRs rather than being genome-wide, highlighting the role of specific structural variants in mediating local adaptation. However, it is important to mention that CNV detection using read depth-based approaches is influenced by the choice of filtering and detection parameters. These settings may preferentially retain high confidence CNVRs while reducing sensitivity to rare or low frequency variants. As a result, the observed distribution of differentiation and breed specific CNVRs may be partially affected by methodological constraints in addition to underlying biological variation.

Enrichment analysis identified GO terms that were also reported in other Asian cattle from China, and India [53,76]. CNVR overlapping genes involved in immune regulation, such as cytokine and chemokine signalling (for example *CXCL2*, *CCL19*, and other interleukin-related genes), likely affect macrophage activation and host defence pathways. These cytokines and chemokines are central to pro-inflammatory responses and leukocyte recruitment, shaping innate immune response to pathogens common in tropical cattle production environments. Similar enrichment of immune related genes has been reported in previous cattle CNV studies, indicating potential association of structural variants with disease resistance and environmental resilience [27,81,82]. Further, CNVRs intersecting genes linked to cellular stress responses and tissue integrity may indicate functional impact. For example, variants in genes that mediate cellular response to damage or pathogen exposure may influence the threshold and extent of stress responses, potentially affecting heat tolerance or oxidative stress regulation [77,83]. Within the core CNVRs shared across breeds, neuronal and synaptic processes were strongly overrepresented. Enrichment of neuron projection guidance, synaptic vesicle cycling, and calcium-dependent exocytosis points to the involvement of genes such as *RAB3A* and related vesicle-trafficking regulators, which are critical for controlled neurotransmitter release and effective neuromuscular signal transmission. These processes indicate coordinated movement and sustained muscle activity, suggesting a potential role of structural variants in behavioural stability, stamina, and work efficiency. Calcium-dependent exocytosis processes are conserved mechanisms mediating synaptic

vesicle release and are critical for efficient nerve-to-muscle signalling that underlies coordinated movement and stamina [84].

KEGG pathway enrichment analysis revealed that genes overlapping CNVRs are distributed across signalling pathways that regulate neuroendocrine control, reproduction, stress adaptation, growth, and metabolic homeostasis. Among the most significantly enriched pathways were glutamatergic synapse and long-term depression, indicating that structural variation affects genes central to neuronal signalling. These pathways included genes such as *GNAI1*, *GNAI2*, *GNB2*, *GNG2*, *ADCY7*, *ITPR2*, and *RYR1*, which regulate neurotransmitter-mediated calcium signalling and downstream second-messenger responses [85,86]. Glutamatergic and calcium-dependent signalling is essential not only for central nervous system function but also for neuromuscular coordination, behavioural responses, and stress perception [85,87,88]. The enrichment of the GnRH secretion pathway highlights the role of CNVR-associated genes in reproductive endocrinology. This pathway comprised key regulators such as *KISS1*, *KISS1R*, *PLCB1*, *PLCB2*, *AKT1*, *ITPR1*, *ITPR2*, and multiple voltage-gated calcium and potassium channel genes (*CACNA1* and *KCNN* families). These genes collectively regulate pulsatile GnRH release and pituitary gonadotropin secretion, thereby influencing ovulation rate, oestrous cyclicity, and fertility. Structural variation affecting these components may contribute to differences in reproductive efficiency and adaptability under environmental stress [89]. Pathways related to hormonal and stress regulation, including parathyroid hormone synthesis and secretion and aldosterone synthesis and secretion, were also significantly enriched. Genes such as *GNA11*, *ADCY7*, *BCL2*, and *MEF2D* link calcium homeostasis, endocrine signalling, and cellular survival. These pathways play important roles in mineral balance, muscle contraction, and physiological adaptation to nutritional and thermal challenges [90]. Growth- and metabolism-associated pathways, including MAPK signalling, insulin signalling, Rap1 signalling, VEGF signalling, and endocytosis, were also consistently over-represented. Core genes such as *MAP2K1*, *MAP2K5*, *MAPK7*, *PRKCA*, *AKT1*, *HRAS*, and *CDC42* regulate cell proliferation, angiogenesis, glucose uptake, and nutrient partitioning between muscle and adipose tissues. Variation in these pathways has been associated with growth rate, feed efficiency, muscle development, and fat deposition in cattle [91,92]. Overall, the KEGG pathway enrichment profile indicated intersection of CNVRs with highly integrated neuroendocrine, metabolic, and stress-responsive pathways, suggesting the potential influence of structural variation on coordinated regulation of reproduction, growth, and resilience in indigenous cattle populations raised under challenging production environments. Although these functional enrichments provided an indication of the biological functions represented within the CNVR-overlapping genes; it is important to exercise caution while interpreting their association with specific adaptive/productive characteristics in the studied cattle populations.

The overlap of CNVRs with carcass-related QTLs [83,93,94] indicates that copy number changes are located in genomic regions influencing muscle development, fat deposition, and body conformation. The detection of CNVRs in QTL regions associated with fertility and reproduction traits [17,29] suggests that structural variation may influence reproductive efficiency in low-input production systems, where animals must allocate limited resources between basic physiologic/metabolic needs, work and reproduction. Similarly, the overlap of core set of shared CNVRs with QTLs linked to disease susceptibility and immune responsiveness points to a role for copy number variation in tolerance to infection challenges [95,96].

5. Conclusion

The present study provides a comprehensive genome-wide assessment of copy number variation regions in South Asian indigenous draught cattle using whole-genome sequencing data. A large number of CNVRs were identified across eight breeds, with a substantial proportion shared among populations while a smaller number of breed-specific CNVRs were detected. Functional enrichment analyses showed that genes overlapping CNVRs were predominantly involved in immune regulation, metabolic control, neuroendocrine signalling, and cellular responses to stress, indicating biological processes relevant to adaptation, work capacity, and general fitness. Mapping of CNVRs to known quantitative trait loci further

supported associations with carcass characteristics, growth, fertility, and reproduction traits. Overall, the findings expand current knowledge of structural genomic diversity in South Asian draught cattle and provide baseline genomic information that can support future studies on their adaptation, breeding, management and conservation.

Supporting information

S1 File. Declaration on Inclusivity in Global Research.

(DOCX)

S2 File. List of background gene set used for enrichment analysis using ShinyGo.

(TXT)

S3 File. Sensitivity analysis to assess the effect of different overlap thresholds on CNVR detection.

(DOCX)

S4 File. List of enriched GO processes and KEGG pathways involving genes overlapping with ALL, core (shared) and Hallikar-specific CNVRs.

(XLSX)

Acknowledgments

We would like to express our gratitude to local farmers and veterinarians in India, Sri Lanka, Myanmar and Cambodia for their assistance during sample collection and enriching the FAO/IAEA genetic repository.

Author contributions

Conceptualization: Johann Sölkner, Gábor Mészáros.

Data curation: Tafara Kundai Mavunga, Kathiravan Periasamy.

Formal analysis: Tafara Kundai Mavunga, Maano Bryton Malima, Kathiravan Periasamy.

Funding acquisition: Yosra Ressaissi, Farai Catherine Muchadeyi.

Investigation: Saravanan Ramasamy, Loku Galappattige Sampath Lokugalappatti, Ei Thandar, Menghak Phem.

Methodology: Tafara Kundai Mavunga, Rudolf Pichler, Saravanan Ramasamy, Loku Galappattige Sampath Lokugalappatti, Ei Thandar, Menghak Phem.

Project administration: Kathiravan Periasamy.

Writing – original draft: Tafara Kundai Mavunga.

Writing – review & editing: Johann Sölkner, Gábor Mészáros, Yosra Ressaissi, Farai Catherine Muchadeyi, Mario Barbato, Kathiravan Periasamy.

References

1. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, et al. Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci.* 2009;92(1):16–24. <https://doi.org/10.3168/jds.2008-1514> PMID: [19109259](https://pubmed.ncbi.nlm.nih.gov/19109259/)
2. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet.* 2006;7(2):85–97. <https://doi.org/10.1038/nrg1767> PMID: [16418744](https://pubmed.ncbi.nlm.nih.gov/16418744/)
3. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al. Mapping copy number variation by population-scale genome sequencing. *Nature.* 2011;470(7332):59–65. <https://doi.org/10.1038/nature09708> PMID: [21293372](https://pubmed.ncbi.nlm.nih.gov/21293372/)
4. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, et al. An integrated map of structural variation in 2,504 human genomes. *Nature.* 2015;526(7571):75–81. <https://doi.org/10.1038/nature15394> PMID: [26432246](https://pubmed.ncbi.nlm.nih.gov/26432246/)

5. Hastings PJ, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. *Nat Rev Genet.* 2009;10(8):551–64. <https://doi.org/10.1038/nrg2593> PMID: [19597530](#)
6. Harewood L, Fraser P. The impact of chromosomal rearrangements on regulation of gene expression. *Hum Mol Genet.* 2014;23(R1):R76–82. <https://doi.org/10.1093/hmg/ddu278>
7. Yang T-L, Hao R-H, Guo Y, Papasian CJ, Deng H-W. Copy Number Variation. *Genetics of Bone Biology and Skeletal Disease.* Elsevier; 2018. pp. 43–54. <https://doi.org/10.1016/b978-0-12-804182-6.00004-6>
8. Fadista J, Thomsen B, Holm L-E, Bendixen C. Copy number variation in the bovine genome. *BMC Genomics.* 2010;11:284. <https://doi.org/10.1186/1471-2164-11-284> PMID: [20459598](#)
9. Clap A, Vidal O, Amills M. Copy number variation in the genomes of domestic animals. *Anim Genet.* 2012;43(5):503–17. <https://doi.org/10.1111/j.1365-2052.2012.02317.x> PMID: [22497594](#)
10. Gamazon ER, Stranger BE. The impact of human copy number variation on gene expression. *Brief Funct Genomics.* 2015;14(5):352–7. <https://doi.org/10.1093/bfpg/evl017> PMID: [25922366](#)
11. Xu L, Cole JB, Bickhart DM, Hou Y, Song J, VanRaden PM, et al. Genome wide CNV analysis reveals additional variants associated with milk production traits in Holsteins. *BMC Genomics.* 2014;15(1):683. <https://doi.org/10.1186/1471-2164-15-683> PMID: [25128478](#)
12. Hay EHA, Utsunomiya YT, Xu L, Zhou Y, Neves HHR, Carvalheiro R, et al. Genomic predictions combining SNP markers and copy number variations in Nellore cattle. *BMC Genomics.* 2018;19(1):441. <https://doi.org/10.1186/s12864-018-4787-6> PMID: [29871610](#)
13. Wu Y, Fan H, Jing S, Xia J, Chen Y, Zhang L, et al. A genome-wide scan for copy number variations using high-density single nucleotide polymorphism array in Simmental cattle. *Anim Genet.* 2015;46(3):289–98. <https://doi.org/10.1111/age.12288> PMID: [25917301](#)
14. Salomón-Torres R, González-Vizcarra VM, Medina-Basulto GE, Montañón-Gómez MF, Mahadevan P, Yaurima-Basaldúa VH, et al. Genome-wide identification of copy number variations in Holstein cattle from Baja California, Mexico, using high-density SNP genotyping arrays. *Genet Mol Res.* 2015;14(4):11848–59. <https://doi.org/10.4238/2015.October.2.18> PMID: [26436509](#)
15. Prinsen RTMM, Strillacci MG, Schiavini F, Santus E, Rossoni A, Maurer V. A genome-wide scan of copy number variants using high-density SNPs in Brown Swiss dairy cattle. *Livest Sci.* 2016;191:153–60. <https://doi.org/10.1016/j.livsci.2016.08.006>
16. Durán Aguilar M, Román Ponce SI, Ruiz López FJ, González Padilla E, Vásquez Peláez CG, Bagnato A, et al. Genome-wide association study for milk somatic cell score in holstein cattle using copy number variation as markers. *J Anim Breed Genet.* 2017;134(1):49–59. <https://doi.org/10.1111/jbg.12238> PMID: [27578198](#)
17. Liu M, Fang L, Liu S, Pan MG, Seroussi E, Cole JB, et al. Array CGH-based detection of CNV regions and their potential association with reproduction and other economic traits in Holsteins. *BMC Genomics.* 2019;20(1):181. <https://doi.org/10.1186/s12864-019-5552-1> PMID: [30845913](#)
18. Zhao M, Wang Q, Wang Q, Jia P, Zhao Z. Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics.* 2013;14 Suppl 11(Suppl 11):S1. <https://doi.org/10.1186/1471-2105-14-S11-S1> PMID: [24564169](#)
19. 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 Genomics.* 2014;15(1):962. <https://doi.org/10.1186/1471-2164-15-962> PMID: [25378104](#)
20. Ho SS, Urban AE, Mills RE. Structural variation in the sequencing era. *Nat Rev Genet.* 2020;21(3):171–89. <https://doi.org/10.1038/s41576-019-0180-9> PMID: [31729472](#)
21. Ahmad SF, Chandrababu Shailaja C, Vaishnav S, Kumar A, Gaur GK, Janga SC, et al. Read-depth based approach on whole genome resequencing data reveals important insights into the copy number variation (CNV) map of major global buffalo breeds. *BMC Genomics.* 2023;24(1):616. <https://doi.org/10.1186/s12864-023-09720-8> PMID: [37845620](#)
22. Zhang ZD, Du J, Lam H, Abyzov A, Urban AE, Snyder M, et al. Identification of genomic indels and structural variations using split reads. *BMC Genomics.* 2011;12:375. <https://doi.org/10.1186/1471-2164-12-375> PMID: [21787423](#)
23. Pirooznia M, Goes FS, Zandi PP. Whole-genome CNV analysis: advances in computational approaches. *Front Genet.* 2015;6:138. <https://doi.org/10.3389/fgene.2015.00138> PMID: [25918519](#)
24. Whitford W, Lehnert K, Snell RG, Jacobsen JC. Evaluation of the performance of copy number variant prediction tools for the detection of deletions from whole genome sequencing data. *J Biomed Inform.* 2019;94:103174. <https://doi.org/10.1016/j.jbi.2019.103174> PMID: [30965134](#)
25. Yoon S, Xuan Z, Makarov V, Ye K, Sebat J. Sensitive and accurate detection of copy number variants using read depth of coverage. *Genome Res.* 2009;19(9):1586–92. <https://doi.org/10.1101/gr.092981.109> PMID: [19657104](#)
26. Liu Y, Mu Y, Wang W, Ahmed Z, Wei X, Lei C, et al. Analysis of genomic copy number variations through whole-genome scan in Chinese Qaidam cattle. *Front Vet Sci.* 2023;10:1148070. <https://doi.org/10.3389/fvets.2023.1148070> PMID: [37065216](#)
27. Ayalew W, Xiaoyun W, Tarekegn GM, Tessema TS, Chu M, Liang C, et al. Whole-genome sequencing of copy number variation analysis in Ethiopian cattle reveals adaptations to diverse environments. *BMC Genomics.* 2024;25(1):1088. <https://doi.org/10.1186/s12864-024-10936-5> PMID: [39548375](#)
28. Taghizadeh S, Gholizadeh M, Rahimi-Mianji G, Moradi MH, Costilla R, Moore S, et al. Genome-wide identification of copy number variation and association with fat deposition in thin and fat-tailed sheep breeds. *Sci Rep.* 2022;12(1):8834. <https://doi.org/10.1038/s41598-022-12778-1> PMID: [35614300](#)

29. Ladeira GC, Pilonetto F, Fernandes AC, Bóscollo PP, Dauria BD, Titto CG, et al. CNV detection and their association with growth, efficiency and carcass traits in Santa Inês sheep. *J Anim Breeding Genetics*. 2022;139(4):476–87. <https://doi.org/10.1111/jbg.12671>
30. Shi H, Li T, Su M, Wang H, Li Q, Lang X, et al. Identification of copy number variation in Tibetan sheep using whole genome resequencing reveals evidence of genomic selection. *BMC Genomics*. 2023;24(1):555. <https://doi.org/10.1186/s12864-023-09672-z> PMID: 37726692
31. Liu M, Zhou Y, Rosen BD, Van Tassel CP, Stella A, Tosser-Klopp G, et al. Diversity of copy number variation in the worldwide goat population. *Heredity (Edinb)*. 2019;122(5):636–46. <https://doi.org/10.1038/s41437-018-0150-6> PMID: 30401973
32. Nandolo W, Mészáros G, Wurzinger M, Banda LJ, Gondwe TN, Mulindwa HA, et al. Detection of copy number variants in African goats using whole genome sequence data. *BMC Genomics*. 2021;22(1):398. <https://doi.org/10.1186/s12864-021-07703-1> PMID: 34051743
33. Chen N, Xia X, Hanif Q, Zhang F, Dang R, Huang B, et al. Global genetic diversity, introgression, and evolutionary adaptation of indicine cattle revealed by whole genome sequencing. *Nat Commun*. 2023;14(1):7803. <https://doi.org/10.1038/s41467-023-43626-z> PMID: 38016956
34. Chan EKF, Nagaraj SH, Reverter A. The evolution of tropical adaptation: comparing taurine and zebu cattle. *Anim Genet*. 2010;41(5):467–77. <https://doi.org/10.1111/j.1365-2052.2010.02053.x> PMID: 20477791
35. Mavunga TK, Sölkner J, Mészáros G, Pichler R, Ramasamy S, Manomohan V. Genomic diversity and selection signatures in Asian Zebu Cattle: insights into adaptation and genetic erosion. *Sci Rep*. 2025;15(1):33346. <https://doi.org/10.1038/s41598-025-14744-z>
36. Singh PK. Phenotypic characterization and performance evaluation of Hallikar cattle in its native tract. *Indian J Anim Sci*. 2008;78(2):211–4.
37. Kuralkar SV, Bankar PS, Chopade MM, Kuralkar P, Dhaware SA. Phenotypic characteristics, production and reproduction performance of Deoni cattle in its native tract. *Indian J Anim Sci*. 2014;84(1):75–7. <https://doi.org/10.56093/ijans.v84i1.37315>
38. Hattarakihal M, Patil VM, Waghmare PG, Suranagi M, Kulkarni S, Desai AR, et al. Morphometric traits of different Deoni cattle strains. *Int J Vet Sci Anim Husbandry*. 2023;8(4S):151–4. <https://doi.org/10.22271/veterinary.2023.v8.i4sc.663>
39. Lokugalappatti LGS, Wickramasinghe S, Alexander PABD, Abbas K, Hussain T, Ramasamy S, et al. Indigenous cattle of Sri Lanka: Genetic and phylogeographic relationship with Zebu of Indus Valley and South Indian origin. *PLoS One*. 2023;18(8):e0282761. <https://doi.org/10.1371/journal.pone.0282761> PMID: 37585485
40. Lwin M, Mon SLY, Yamanaka H, Nagano Y, Mannen H, Faruque MO, et al. Genetic diversities and population structures of four popular Myanmar local cattle breeds. *Anim Sci J*. 2018;89(12):1648–55. <https://doi.org/10.1111/asj.13112> PMID: 30318818
41. Namikawa T, M H, T K, N K, T Y. Coat-color variations of cattle observed in the fields of Cambodia, and withers-height and other traits in the native cattle subjected to further experimental analyses. *Rep Soc Res Native Livestock*. 2006;23:31–44.
42. Natarajan ASKVK. Productivity performance of Kangayam cattle. *Indian J Anim Sci*. 2012;82(11):1440–1.
43. Ganapathi PR. Distribution and population status of Bargur cattle. *Indian Vet J*. 2009;86(1):971–2.
44. Pundir RK, Kathiravan P, Singh PK, Vamanikhandan VA. Bargur cattle: status, characteristics and performance. *Indian J Anim Sci*. 2024;79(7). <https://doi.org/10.56093/ijans.v79i7.76434>
45. Dongre VB, Gandhi RS, Salunke VM, Kokate LS, Durge SM, Khandait VN, et al. Present status and future prospects of Deoni Cattle. *Indian J Anim Sci*. 2017;87(7). <https://doi.org/10.56093/ijans.v87i7.72112>
46. Shanjayan NLL. A morphometric analysis of indigenous white cattle (Thamankaduwa breed) in the Eastern province of Sri Lanka with a description of a novel character. In: 5th International Symposium-IntSym. SEUSL; 2015.
47. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, et al. Copy number variation: new insights in genome diversity. *Genome Res*. 2006;16(8):949–61. <https://doi.org/10.1101/gr.3677206> PMID: 16809666
48. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang L, Cellamare A, et al. Analysis of copy number variations among diverse cattle breeds. *Genome Res*. 2010;20(5):693–703. <https://doi.org/10.1101/gr.105403.110> PMID: 20212021
49. Bickhart DM, Hou Y, Schroeder SG, Alkan C, Cardone MF, Matukumalli LK, et al. Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res*. 2012;22(4):778–90. <https://doi.org/10.1101/gr.133967.111> PMID: 22300768
50. Mason-Suares H, Landry L, Lebo S. Detecting copy number variation via next generation technology. *Curr Genet Med Rep*. 2016;4(3):74–85. <https://doi.org/10.1007/s40142-016-0091-4>
51. Singh VK, Singh S, Nandhini PB, Bhatia AK, Dixit SP, 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(8):3483–94. <https://doi.org/10.1080/10495398.2022.2162910> PMID: 36592947
52. Kumar H, Panigrahi M, Saravanan KA, Rajawat D, Parida S, Bhushan B, et al. Genome-wide detection of copy number variations in Tharparkar cattle. *Anim Biotechnol*. 2023;34(2):448–55. <https://doi.org/10.1080/10495398.2021.1942027> PMID: 34191685
53. Vani S, Balasubramanyam D, Tirumurugaan KG, Gopinathan A, Karthickeyan SMK. Genome-wide copy number variation regions in indigenous (*Bos indicus*) cattle breeds of Tamil Nadu, India. *Anim Biosci*. 2025;38(3):395–407. <https://doi.org/10.5713/ab.23.0525> PMID: 39210813
54. Benfica LF, Brito LF, do Bem RD, de Oliveira LF, Mulim HA, Braga LG, et al. Detection and characterization of copy number variation in three differentially-selected Nellore cattle populations. *Front Genet*. 2024;15:1377130. <https://doi.org/10.3389/fgene.2024.1377130> PMID: 38694873
55. Braga LG, Chud TCS, Watanabe RN, Savegnago RP, Sena TM, do Carmo AS, et al. Identification of copy number variations in the genome of Dairy Gir cattle. *PLoS One*. 2023;18(4):e0284085. <https://doi.org/10.1371/journal.pone.0284085> PMID: 37036840

56. Jang J, Terefe E, Kim K, Lee YH, Belay G, Tijjani A, et al. Population differentiated copy number variation of *Bos taurus*, *Bos indicus* and their African hybrids. *BMC Genomics*. 2021;22(1):531. <https://doi.org/10.1186/s12864-021-07808-7> PMID: [34253178](#)
57. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online] [Internet]. 2010 [cited 2025 Sep 4]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
58. USDAARS. Genome assembly NIAB-ARS_B.indTharparkar_mat_pri_1.0 [Internet]. 2023 [cited 2025 Oct 28]. Available from: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_029378745.1/
59. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324> PMID: [19451168](#)
60. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO. Twelve years of SAMtools and BCFtools. *Gigascience*. 2021;10(2). <https://doi.org/10.1093/gigascience/giab008>
61. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res*. 2011;21(6):974–84. <https://doi.org/10.1101/gr.114876.110> PMID: [21324876](#)
62. Wang X, Zheng Z, Cai Y, Chen T, Li C, Fu W. CNVcaller: highly efficient and widely applicable software for detecting copy number variations in large populations. *Gigascience*. 2017;6(12). <https://doi.org/10.1093/gigascience/gix115>
63. Chaisson MJ, Tesler G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics*. 2012;13:238. <https://doi.org/10.1186/1471-2105-13-238> PMID: [22988817](#)
64. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. <https://doi.org/10.1093/nar/gkq603> PMID: [20601685](#)
65. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 2019;36(8):2628–9. <https://doi.org/10.1093/bioinformatics/btz931>
66. Hu Z-L, Park CA, Reecy JM. Bringing the Animal QTLdb and CorrDB into the future: meeting new challenges and providing updated services. *Nucleic Acids Res*. 2022;50(D1):D956–61. <https://doi.org/10.1093/nar/gkab1116> PMID: [34850103](#)
67. Strucken EM, Gebrehiwot NZ, Swaminathan M, Joshi S, Al Kalaldehy M, Gibson JP. Genetic diversity and effective population sizes of thirteen Indian cattle breeds. *Genet Sel Evol*. 2021;53(1):47. <https://doi.org/10.1186/s12711-021-00640-3> PMID: [34074236](#)
68. Dixit SP, Bhatia AK, Ganguly I, Singh S, Dash S, Sharma A, et al. Genome analyses revealed genetic admixture and selection signatures in *Bos indicus*. *Sci Rep*. 2021;11(1):21924. <https://doi.org/10.1038/s41598-021-01144-2> PMID: [34753978](#)
69. Manomohan V, Saravanan R, Pichler R, Murali N, Sivakumar K, Sudhakar K, et al. Legacy of draught cattle breeds of South India: Insights into population structure, genetic admixture and maternal origin. *PLoS One*. 2021;16(5):e0246497. <https://doi.org/10.1371/journal.pone.0246497> PMID: [34029341](#)
70. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E. De novo assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020;9(3). <https://doi.org/10.1093/gigascience/giaa021>
71. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, et al. A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome Biol*. 2009;10(4):R42. <https://doi.org/10.1186/gb-2009-10-4-r42> PMID: [19393038](#)
72. Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC, Gibbs RA, Muzny DM, et al. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*. 2009;324(5926):522–8. <https://doi.org/10.1126/science.1169588> PMID: [19390049](#)
73. Low WY, Tearle R, Liu R, Koren S, Rhie A, Bickhart DM, et al. Haplotype-resolved genomes provide insights into structural variation and gene content in Angus and Brahman cattle. *Nat Commun*. 2020;11(1):2071. <https://doi.org/10.1038/s41467-020-15848-y> PMID: [32350247](#)
74. Canavez FC, Luche DD, Stothard P, Leite KRM, Sousa-Canavez JM, Plastow G, et al. Genome sequence and assembly of *Bos indicus*. *J Hered*. 2012;103(3):342–8. <https://doi.org/10.1093/jhered/esr153> PMID: [22315242](#)
75. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A. Diversity of human copy number variation and multicopy genes. *Science*. 2010;330(6004):641–6. <https://doi.org/10.1126/science.1197005>
76. Mei C, Junjvlieke Z, Raza SHA, Wang H, Cheng G, Zhao C, et al. Copy number variation detection in Chinese indigenous cattle by whole genome sequencing. *Genomics*. 2020;112(1):831–6. <https://doi.org/10.1016/j.ygeno.2019.05.023> PMID: [31145994](#)
77. Zhong Z, Wang Z, Xie X, Pan D, Su Z, Fan J. Insights into adaption and growth evolution: genome-wide copy number variation analysis in Chinese Hainan yellow cattle using whole-genome re-sequencing data. *Int J Mol Sci*. 2024;25(22):11919. <https://doi.org/10.3390/ijms252211919>
78. Song X, Zhang Z, Xing S, Liu X, Zhang Y, Wang J, et al. The CNV map construction and ROH analysis of Pinan cattle. *BMC Genomics*. 2025;26(1):480. <https://doi.org/10.1186/s12864-025-11626-6> PMID: [40369451](#)
79. Bhanuprakash V, Chhotaray S, Pruthviraj DR, Rawat C, Karthikeyan A, Panigrahi M. Copy number variation in livestock: a mini review. *Vet World*. 2018;11(4):535–41. <https://doi.org/10.14202/vetworld.2018.535-541> PMID: [29805222](#)
80. Pierce MD, Dzama K, Muchadeyi FC. Corrigendum: genetic diversity of seven cattle breeds inferred using copy number variations. *Front Genet*. 2018;9:252. <https://doi.org/10.3389/fgene.2018.00252> PMID: [30038636](#)
81. Hou Y, Liu GE, Bickhart DM, Cardone MF, Wang K, Kim E-S, et al. Genomic characteristics of cattle copy number variations. *BMC Genomics*. 2011;12:127. <https://doi.org/10.1186/1471-2164-12-127> PMID: [21345189](#)

82. Zhou Y, Utsunomiya YT, Xu L, Hay EHA, Bickhart DM, Sonstegard TS, et al. Comparative analyses across cattle genders and breeds reveal the pitfalls caused by false positive and lineage-differential copy number variations. *Sci Rep*. 2016;6:29219. <https://doi.org/10.1038/srep29219> PMID: [27381368](https://pubmed.ncbi.nlm.nih.gov/27381368/)
83. da Silva JM, Giachetto PF, da Silva LO, Cintra LC, Paiva SR, Yamagishi MEB, et al. Genome-wide copy number variation (CNV) detection in Nelore cattle reveals highly frequent variants in genome regions harboring QTLs affecting production traits. *BMC Genomics*. 2016;17:454. <https://doi.org/10.1186/s12864-016-2752-9> PMID: [27297173](https://pubmed.ncbi.nlm.nih.gov/27297173/)
84. Sudhof TC. Calcium control of neurotransmitter release. *Cold Spring Harb Perspect Biol*. 2012;4(1):a011353. <https://doi.org/10.1101/cshperspect.a011353>
85. Niciu MJ, Kelmendi B, Sanacora G. Overview of glutamatergic neurotransmission in the nervous system. *Pharmacol Biochem Behav*. 2012;100(4):656–64. <https://doi.org/10.1016/j.pbb.2011.08.008> PMID: [21889952](https://pubmed.ncbi.nlm.nih.gov/21889952/)
86. Nanou E, Catterall WA. Calcium channels, synaptic plasticity, and neuropsychiatric disease. *Neuron*. 2018;98(3):466–81. <https://doi.org/10.1016/j.neuron.2018.03.017>
87. Baj A, Moro E, Bistoletti M, Orlandi V, Crema F, Giaroni C. Glutamatergic signaling along the microbiota-gut-brain axis. *Int J Mol Sci*. 2019;20(6):1482. <https://doi.org/10.3390/ijms20061482>
88. Chen S-Y, Yao J, Hu Y-D, Chen H-Y, Liu P-C, Wang W-F, et al. Control of behavioral arousal and defense by a glutamatergic Midbrain-Amygdala pathway in mice. *Front Neurosci*. 2022;16:850193. <https://doi.org/10.3389/fnins.2022.850193> PMID: [35527820](https://pubmed.ncbi.nlm.nih.gov/35527820/)
89. Lehman MN, Coolen LM, Goodman RL. Minireview: kisspeptin/neurokinin B/dynorphin (KNDy) cells of the arcuate nucleus: a central node in the control of gonadotropin-releasing hormone secretion. *Endocrinology*. 2010;151(8):3479–89. <https://doi.org/10.1210/en.2010-0022> PMID: [20501670](https://pubmed.ncbi.nlm.nih.gov/20501670/)
90. Sejjan V, Bhatta R, Gaughan JB, Dunshea FR, Lacetera N. Review: adaptation of animals to heat stress. *Animal*. 2018;12(s2):s431–44. <https://doi.org/10.1017/S1751731118001945> PMID: [30139399](https://pubmed.ncbi.nlm.nih.gov/30139399/)
91. Roudbari Z, Coort SL, Kutmon M, Eijssen L, Melius J, Sadkowski T, et al. Identification of biological pathways contributing to marbling in skeletal muscle to improve beef cattle breeding. *Front Genet*. 2020;10:1370. <https://doi.org/10.3389/fgene.2019.01370> PMID: [32117419](https://pubmed.ncbi.nlm.nih.gov/32117419/)
92. Reis IA, Baldassini WA, Ramírez-Zamudio GD, de Farias IMSC, Chiaratti MR, Pereira Junior S, et al. Muscle tissue transcriptome of F1 Angus-Nellore bulls and steers feedlot finished: impacts on intramuscular fat deposition. *BMC Genomics*. 2024;25(1):1178. <https://doi.org/10.1186/s12864-024-11066-8> PMID: [39633259](https://pubmed.ncbi.nlm.nih.gov/39633259/)
93. Guo S, Wu X, Pei J, Wang X, Bao P, Xiong L, et al. Genome-wide CNV analysis reveals variants associated with high-altitude adaptation and meat traits in Qaidam cattle. *Electron J Biotechnol*. 2021;54:8–16. <https://doi.org/10.1016/j.ejbt.2021.07.006>
94. Marín-Garzón NA, Mota LFM, Vargas G, Arikawa LM, Fonseca LFS, Fernandes Júnior GA, et al. Detection and functional assessment of structural variants using whole-genome re-sequencing data in Nelore cattle. *Sci Rep*. 2025;15(1):30364. <https://doi.org/10.1038/s41598-025-14139-0> PMID: [40830176](https://pubmed.ncbi.nlm.nih.gov/40830176/)
95. Cozzi MC, Martinez-Ruiz CP, Roman-Ponce SI, Murillo VEV, Utrera ÁR, Montaña-Bermúdez MM. Copy number variants reveal genomic diversity in a Mexican Creole cattle population. *Livest Sci*. 2019;229:194–202. <https://doi.org/10.1016/j.livsci.2019.09.030>
96. Badia-Bringué G, Canive M, Vázquez P, Garrido JM, Fernández A, Juste RA, et al. Genome-wide association study reveals quantitative trait Loci and candidate genes associated with high interferon-gamma production in Holstein Cattle naturally infected with *Mycobacterium Bovis*. *Int J Mol Sci*. 2024 Jun 3;25(11):6165. <https://doi.org/10.3390/ijms25116165>