The Pharma Innovation Journal 2022; SP-11(11): 2289-2297

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2022; SP-11(11): 2289-2297

www.thepharmajournal.com
Received: 10-08-2022
Accepted: 14-09-2022

Aishwarya Dash
ICAR-National Dairy Research Institute, Karnal, Haryana, India

Dhandapani S
ICAR-National Dairy Research Institute, Karnal, Haryana, India

ID Gupta
ICAR-National Dairy Research Institute, Karnal, Haryana, India

Jayakumar Sivalingam
ICAR-Directorate of Poultry Research, Hyderabad, Telangana, India

Corresponding Author:
Aishwarya Dash
ICAR-National Dairy Research Institute, Karnal, Haryana, India

Delineating copy number variations in livestock animals

Aishwarya Dash, Dhandapani S, ID Gupta and Jayakumar Sivalingam

Abstract
Genetic variation is the difference between individuals detected at molecular level and it is classified into different kinds based on number of nucleotides involved. Structural variations are variations observed due to mutation and genetic recombination in DNA segments (polynucleotide). Copy number variation is an unbalanced structural variation due to loss or gain of DNA fragments varying from 50bp to several mega base pair length. It covers a higher portion of genomic sequence and has higher mutation rate than SNPs. CNVs involve one or more genes and are accountable for change in their structures and quantity and may create new genes. There are several mechanisms such as, non-allelic homologous recombination, non-homologous end-joining, fork stalling and template switching and L1-mediated retro transposition for CNV genesis due to like deletion, duplication, inversions, and translocations of the genes. With the advancement of technology, different techniques (Conventional methods, Array based methods, Next generation Sequencing) are used to detect copy number variations in various livestock species associated with several traits. Hence, CNVs are potentially greater effect in variation and can be considered to be promising causal genetic markers of economic importance.

Keywords: CNV, deletion, duplication, structural variation

Introduction
Individuals show variation among each other phenotypically and genotypically. Phenotypic variation is the difference which is observed with naked eyes, whereas, genetic variation is detected at molecular level. Mutation and genetic recombination result genetic variability due to nucleotide sequence divergences. Single nucleotide Polymorphisms (SNPs) in genome are considered as the most prevalent and vital form of genetic variation and are extensively used in genomic prediction and genome-wide association studies. Gradually large DNA segment mutations were detected and verified by several molecular and cytogenetic analysis studies [1]. These mutated DNA segments are described as structural variations (SV) [2] and may reorganise genes on chromosome. Chromosomal variations are resulted due to several events like insertions, deletions, inversions, duplications and translocations and contribute to genetic and phenotypic variation [3, 4]. Currently, the focus of genetic variation is shifting from specific nucleotide sequences to large-scale changes throughout the whole genome. Copy number variation is a form of structural variation that comprises large number of base pairs.

Copy Number Variation
Copy number variation (CNV) is a phenomenon which affects DNA segments and causes genomic alterations resulting an abnormal number of copies of one or more segments. Precisely it is a duplication or deletion type occurrence [5] and inherits to next generation. [6] Copy number variations (CNVs) are considered as major class of genetic polymorphisms showing large-scale losses and gains of base-pairs [7, 8]. These involve a sequence of nucleotides varying >1 kb in size to several mega base pairs (Mb) [2, 9]. and now these encompass events as small as 50 bp [9, 10]. Hence, CNVs are defined as a class of unbalanced structural variations including variable copy number in comparison with a reference genome. [2,10] Smaller CNVs with <50kb length are much more common than larger ones that are explored with high-resolution studies [11, 12]. The CNV regions (CNVRs) were determined by combining overlapping CNVs identified in different individuals [13, 14].

Effect of CNVs on gene expression
CNVs are found in all chromosomes and are dispersed throughout the genome in a non-random manner demonstrating heterogeneous distribution [12, 13].

These are present in repeated regions like telomere, centromere and heterochromatin and identified by FISH and microsatellite analysis [15]. Structural variations can involve one or more genes and are responsible for alteration of their structures and quantity. As a result of several mechanisms like deletion, duplication, inversions, and translocations of the genes, these can be present as a recessive or dominant allele by change of gene dosage and gene disruption effect. Consequently, there is change in gene regulation and expose of recessive alleles [7, 16]. It also provides materials and mechanisms for creating new genes [17, 18]. When CNVs exist in the protein coding region, they affect transcriptional level of genes within or near the CNVR and subsequently alter translational process and alter the protein function. Interaction of environmental factors and genetic factors influences CNVs for significant phenotypic effect [15]. CNV is therefore considered as a chief source of genetic variation [7]. Studies revealed that CNVs capture 18 to 30% of the genetic variation in terms of gene expression in human and mouse [19, 20]. It has been seen that CNVs occupy approximately 7% of the mouse genome and cattle genome [21]. CNVs are strongly associated with segmental duplications (SDs), which can be considered as catalysts and hotspots for CNV formation due to their fragile structural architecture that prompts frequent rearrangements [12, 22]. Deletion events are more common and have a potentially much higher effect than duplications [10, 23]. This ubiquitous copy number polymorphisms in the genome can be considered to be significant for inherited population variation [2, 24]. CNVs may be associated with phenotypic variations including disease susceptibility [25] and may act as a source for evolutionary mechanisms [17]. Rates of molecular evolution are determined by selection processes acting on genes. Specifically, the type of selection on the individual genes affects the rate of protein-changing (non-synonymous) substitution and the rate of silent (synonymous) substitution at the nucleotide level. Copy number variations are related with normal variation, disease, evolution, and adaptive traits in various species [9, 21].

Mechanism for CNV Formation

The CNVs are formed by rearrangements of segments of DNA in the genome and the following four mechanisms account for the majority of CNV formation such as non-allelic homologous recombination (NAHR); non-homologous end-joining (NHEJ); fork stalling and template switching (FoSTeS) and L1-mediated retro transposition [26, 27].

1. Non-allelic homologous recombination (NAHR)

NAHR occurs in meiosis and mitosis due to recombination between the two regions of similar sequence between two loci of chromosome(s). It is of 3 types including inter-chromosomal, inter-chromatid and intra-chromatid. In inter-chromosomal and intra-chromosomal recombination, there is increase in the segment of DNA at the expense of another which may result in duplication and deletion due to crossing over between two non-sister chromatids and sister chromatids respectively. There is inversion of the segment of chromosome in intra-chromatid recombination causing deletion. A strong correlation of large CNVs and SDs in mammals supported the hypothesis that their formation mechanisms were mainly due to NAHR [29]. It has already been shown that SDs provided substrate for NAHR, which in turn, produces novel chromosomal rearrangements and copy number changes. Therefore, CNVRs that overlap with SDs typically display high frequencies as compared to the CNVRs that do not overlap SDs [8]. A possible biological explanation provided that a nonallelic homologous recombination, one of the major sources of CNV, generated more deleted than duplicated regions [10, 12].

2. Non-homologous end-joining (NHEJ)

NHEJ mechanism is utilized by cells to repair DNA double-strand breaks (DSBs) caused by ionizing radiation or reactive oxygen species and physiological forms of DSBs such as variable (diversity) joining [V(D)J] recombination [29, 30]. Small CNVs (<18 kb) were discovered through high-density array CGH or sequence mapping analyses [31, 32]. The overlaps between small CNVs and SDs in human and mouse were significantly lower, suggesting that SDs and NAHR were less involved and other mechanisms, such as NHEJ could be more responsible [27, 28].

3. Fork stalling and template switching (FoSTeS)

FoSTeS is a DNA replication-based mechanism which can account for Complex Genomic Rearrangements and CNVs [7].

4. L1 mediated retro-transposition

It occurs through reverse transcription and integration [33]. SDs are one of the catalysts and hotspots for CNV formation [1, 5]. L1-mediated retrotransposons were associated with various forms of SV and with human genetic diseases [34], which suggested that they might be a major source of genetic structural variation and evolution. [35] More deletions in the L1 regions were detected than in the other non-exonic regions [2].

SNPs vs CNVs

Single nucleotide polymorphism (SNP) is a common form of variation involving change at the specific nucleotide, whereas, CNVs include DNA segments (polynucleotide). In terms of number of base pair, CNV affects a larger fraction of genome compared to Single Nucleotide Polymorphisms (SNPs) [8, 13]. and are widely distributed in the genome. Genomic structural variation covers more base pairs i.e., approximately 1% of a genome, which is much higher than the SNPs comprising 0.1% [11, 36]. Analysis of several autosomal dominant diseases explored that these structural variations show a higher per-locus mutation rate than SNPs do [8, 12]. Mutation rate of CNVs ranges from 1.1×10^{-2} [37] to 1×10^{-8} per locus per generation, [38] by which diverse processes for CNV formation are understood. Stranger *et al*. (2007) showed that CNVs and SNPs contribute 83.6% and 17.7% to complex phenotypes, respectively [20]. Also, it has been reported that CNVs are evolved 2.5 folds faster than SNPs and promote a better adaptation in various habitats [39]. Hence, CNVs involve a higher portion of genomic sequence and have potentially greater effect in variation than SNPs [18, 19] So, CNVs are generally accepted as a major source for heritable variation and can be considered to be promising causal genetic markers for some traits [11, 41].

Methods for Identification of CNVs

1. Conventional Methods

Initially CNVs were studied in terms of specific loci using cytogenetic techniques. Light microscopes were able to find out somatic changes associated with structural variations, but gradually sub-microscopic structural changes can be observed due to the development of new technologies [15]. Fluorescent in situ hybridization (FISH) techniques are used to detect

structural changes due to hybridization of fluorescent probes with the complementary genome. Comparative genomic hybridization method detects CNVs by fluorescent dye visualization and compares length of the chromosomes [41]. These techniques are with constraints of low genomic resolution and detection of only long repeats. Then Bacterial artificial chromosome (BAC) array is used with high genomic resolution throughout the entire gene. High throughput genome sequencing and *in silico* techniques are advance technologies to study whole genome [5]. Sequencing data are aligned with reference sequences are compared using fosmid clones to explore misaligned reads as CNVs [42]. By this method, repetitive regions, inversion induced structural variations can be detected with a high genomic resolution [41] With the advancement of science, Next-generation sequencing techniques replaced array-based techniques to detect copy number variations. These include short and long read sequences of a genome, which can readily recognise structural variations by inversions and translocations processes [1, 3].

2. Array based Methods

Now-a-days CNV identification in livestock is achieved by three main approaches such as CGH array, SNP array, and DNA sequencing. Pros and cons of these three systems are reviewed after different studies [43, 44]. Comparative Genomic Hybridisation Array method follows hybridization technique of a fluorescently labelled target and reference DNA sample. It measures loss and gain of copy numbers based on hybridization intensity and detects small variations in copy number showing higher sensitivity using a single reference sample. Long probe with combination of various coverage and type of reference sample in aCGH method is helpful for finding segmental duplication regions. It has high signal-to-ratio for structural variations [43]. Dense and uniform CGH Array can be produced easily to target a region of interest including repeated regions [45]. Different reference samples can create different relative copy numbers in the target samples. This can be corrected by using same sample as reference within a study. Association study of certain traits is less reliable due to less sample size with high cost [46]. SNP data are aggregated to produce 50K and 777K Bead chip for detecting Copy Number Variations. SNP Array uses population reference unlike CGH array of a single reference sample. These consider LRR (Log R Ratio) and BAF (B Allele Frequency) information. There is inherent bias coverage against the genomic areas commonly having CNVs in this method [44]. SNP probes in bead chip are not that much uniform and dense to get unbiased and high-resolution CNV maps [47]. SNP Array approach can miss small and rare structural variations and exact breakpoints of CNVs due to limited density and high MAF (Minor Allele Frequency) of SNPs. Balanced SV, inversions and translocations are difficult to be captured by SNP chip technique causing incomplete detection of CNVs [2].

3. Next Generation Sequencing (NGS) method

Array based approaches have some limitations though they have facilitated advancement of CNV studies [44, 48]. They are affected by low probe density and cross hybridization repetitive sequences. Low resolution and limited probe numbers and locations do not cover whole genome. Detection of small CNVs is difficult in array methods [49]. Next Generation Sequencing recognizes CNV regions at a base-pair resolution and explore small variations in DNA fragments [22]. Development of DNA sequencing technologies and complementary analysis programmes make a path to identify genome-wide CNVs systematically. Both common and rare CNVs are constructed at a high rate of resolution showing accuracy and effectiveness of genome sequencing method. Ongoing progresses and cost effective NGS techniques are gaining popularity with more sensitivity. Traditional methods for CNV discovery including hybridization-based microarray approaches like array comparative genomic hybridization (aCGH) and SNP microarrays are now being replaced by powerful sequencing-based computational approaches. In NGS method, comparison between long read and short read sequences revealed the limitation of long read on high cost and short reads on CNV quality [4].

Basic strategies of NGS method

NGS data can recover whole spectrum of SVs with four basic strategies [48, 50] such as:

1) Paired-end mapping (PEM) or Read Pair (RP)
2) Split Read (SR)
3) Read Depth (RD)
4) Sequence Assembly or de novo Assembly of Genome (AS)

Each of the four basic strategies under NGS methods has its own advantages and disadvantages

- The Paired end mapping (PEM) is also known as Read Pair (RP). The PEM method was first to indicate efficacy of NGS data for CNV identification. There is a specific distribution of DNA fragments around insert size in paired-end sequencing [51]. In RP method, CNVs are detected analysing the discordance between mapped paired-reads whose distances are different from normal average insert size. Tools based on this method are PEMer, Hydra, Ulysses and Break Dancer [48, 50].
- In Split Read method one read is split into various fragments randomly. First and last fragments are paired respectively around the reference genome. Copy Number Variations are identified based on alignment and its location and direction. Gapped sequence alignments are also assessed in this method. Split-read based methods include Pindel, Gustaf, SV seq2 and Prism tools [48, 50].
- Read Depth based assessment focuses on hypothesis that there is correlation between depth of coverage and copy number of one genomic region. Duplicated CNVs are found in high coverage region, whereas, deletion regions have low coverage. Alignment of reads with a reference genome, counting of read depth with a bin size, normalization of counts to remove biases due to GC content and repeat regions and applying of segmentation algorithm are followed to call CNVs. Then CNVs are filtered to get significant ones. RP and SR methods reports position of potential CNVs without the counts, but in RD analysis density of aligned reads can discover exact copy number of events. RD can perform better in finding CNVs of large size than RP and SR. CNV-seq, BIC-seq, Cn.MOPS, CNVnator, ERDS, RDXplorer, Read Depth, SegSeq and CNVrd2 utilize RD method for CNV calling [48, 50].
- Sequence Assembly method generates contigs or scaffolds first and then by comparing with reference genome structural polymorphisms can be found. This

method needs complex computational resources. Magnolya tool involve AS method [48, 50].

Copy Number Variation studies in various livestock species

CNVs are genetic variants in the genome that show variations between and within species. CNVs have been studied extensively in human and several livestock species as well. In livestock species, CNVs have been identified to analyse the differences in domesticated animals and for detecting their association with economically important traits. There are several studies on CNVs in the livestock genome. CNVs were investigated for association with economically important traits such as milk production and fertility [8, 52]. The details on CNV investigations by various researchers have been summarised below (Table 1).

Table 1: The details on CNV investigations by various researchers have been summarised below

Species/Breed (Nos.)	Details of CNVs			Method	Author(s)
	No	Size	Genome coverage		
Cattle					
Holsteins (14) Red Danish (2) Simmental (3) Hereford (1)	304 CNVRs	22 Mb	0.68%	aCGH	Fadista *et al*., 2010 [12]
Bos taurus coreanae (265)	855 CNVs 368 CNVRs			BovineSNP50K	Bae *et al*., 2010 [53]
Taurine, Indicine, Composite (90)	177 CNVRs	28.1Mb	1.07%	aCGH	Liu *et al*., 2010 [54]
Black Angus (1) Holstein (1)	790 CNVs	3.3 Mb	0.13%	WGS	Stothard *et al*., 2011 [55]
Holstein Friesian (1)	196 CNVs 30 CNVs 520 CNVs	6.11 Mb 2.57 Mb 3.63 Mb		aCGH BovineSNP770K WGS	Zhan *et al*., 2011 [56]
Taurine (366) Indicine (70) Composite (46) African Breeds (39)	3666 CNVs 682 CNVRs	139.8 Mb	4.6%	BovineSNP50K	Hou *et al*., 2011 [57]
Angus (3) Brahman (5) Composite (1)	116 CNVs 51 CNVRs	13.5 Mb	0.45%	aCGH	Kijas *et al*., 2011 [58]
Chinese Holstein (2047)	99 CNVRs	23.24 Mb	0.91%	BovineSNP50K	Jiang *et al*., 2012 [59]
Nelore (1) Angus (3) Holstein (1) Hereford (1)	1265 CNVRs	55.6 Mb	2.1%	WGS (RD)	Bickchart *et al*., 2012 [22]
Taurine (447) Indicine (113) Composite (67) African (47)	34311 CNVs 3346 CNVRs	142.7 Mb	4.89%	Bovine HD chip	Hou *et al*., 2012 [60]
Hanwoo (1) vs Black Angus (1)	1173 CNVRs	16.7 Mb	0.63%	WGS	Choi *et al*., 2013 [61]
Hanwoo (1) vs Holstein (1)	963 CNVRs	7.8 Mb	0.29%	WGS	Choi *et al*., 2013 [61]
Chinese Holstein (96)	1733 CNVs 367 CNVRs	42.74 Mb	1.61%	BovineSNP777K	Jiang *et al*., 2013 [62]
Holstein (26,362)	2626669 CNVs			BovineSNP50K	Xu *et al*., 2014 [52]
Holsteins (10) Hanwoo (22)	6811 deleted CNVs	18.6 Mb	0.74%	WGS	Shin *et al*., 2014 [23]
Chinese cattle (129)	370 CNVRs	47 Mb	1.78%	aCGH	Zhang *et al*., 2014 [18]
Holstein (27) Montbéliarde (17) Normande (18)	6426 putative structural variants			WGS	Boussaha *et al*., 2015 [63]
Nguni cattle (492)	433 CNVs 334 CNVRs	(Environmental responses and adaptation)		BovineSNP50K	Wang *et al*., 2015 [64]
Cattle (175)	57 CNVRs	5.27 Mb	0.19%	WES (RD)	Keel *et al*., 2016 [21]
Japanese Black Cattle (1481)	861CNVRs	43.65 Mb	1.74 %	BovineSNP50K	Sasaki *et al*., 2016 [65]
Nellore (723)	49997 CNVs 2600 CNVRs	170.6 Mb	6.5%	BovineSNP777K	da Silva *et al*., 2016 [66]
Polish HF (29)	435594 CNVs			WGS (RD)	Mielczarek *et al*., 2017 [10]
French Dairy, Beef (200)	4178 CNVs		6%	WGS (RD)	Letaief *et al*., 2017 [401]
New Zealand Dairy (556)	43708 CNVs			WGS (RD)	Couldrey *et al*., 2017 [4]
European cattle (149)	9944 CNVs 923 CNVRs	61.06 Mb	2.5%	Bovine HD	Upadhyay *et al*., 2017 [8]
Chinese cattle (188)	13225 CNV 3356 CNVRs			Bovine HD SNP Array	Yang *et al*., 2017 [16]
Holstein (308) Jersey (64)	17518 SV	27.36 Mb		WGS (RP & SR)	Chen *et al*., 2017 [10]

Holstein (8)	14821 CNVs 487CNVRs	8.23 Mb		WGS (RD)	Gao *et al*., 2017 [48]
Taurine Dairy, Beef (553)	6223 CNVRs	107.75 Mb	4.05%	WGS (RD)	Kommadath *et al*., 2019 [9]
Holstein (47)	1758 CNVs 1043 CNVRs	46.8 Mb	2.06%	aCGH	Liu *et al*., 2019 [28]
Species/Breed (Nos.)	**Details of CNVs**			**Method**	**Author(s)**
	No	**Size**	**Genome coverage**		
Buffalo					
Chinese Riverine (2) Swamp (1)	163 CNVRs		1.44%	aCGH	Zhang *et al*., 2014 [48]
Riverine buffalo (14)	13444 deletion CNVRs	*GLYAT* gene as a CNVR adaptation to tropical environments		WGS (RD)	Li *et al*., 2019 [67]
Riverine buffalo (15)	1344 CNVRs	59.8Mb	2.2%	WGS (RD)	Liu *et al*., 2019 [46]
Species/Breed (Nos.)	**Details of CNVs**			**Method**	**Author(s)**
	No	**Size**	**Genome coverage**		
Sheep					
Sheep	135 CNVRs	10.5 Mb	0.4%	aCGH (cattle-sheep)	Fontanesi *et al*., 2011 [68]
Sheep (329)	3624 CNVs 238 CNVRs	60.35 Mb	2.17%	Ovine SNP50K	Liu *et al*., 2013 [69]
Chinese sheep (5)	245 CNVs 51 CNVRs			aCGH	Hou *et al*., 2015 [14]
Sheep (30)	9789 CNVs 3488 CNVRs	67.6 Mb	2.7%	aCGH	Jenkins *et al*., 2016 [70]
Sheep (120) with three types of tails	371CNVRs 301CNVRs 66 CNVRs	71.35 Mb 51.65 Mb 10.56 Mb		OvineSNP600K	Zhu *et al*., 2016 [71]
Chinese sheep (48)	5190 CNVs (Autosomes) 1296 CNVRs		4.7%	OvineSNP600K	Ma *et al*., 2017 [72]
Sheep (2254)	24558 CNVs 619 CNVRs	197 Mb	6.9%	OvineSNP50K	Yang *et al*., 2017 [16]
	BTG3, PTGS1, PSPH genes involved in Foetal muscle development, prostaglandin synthesis and bone colour				
Sheep (4)	1 CNV	2000 bp length			Jiang *et al*., 2019 [73]
	SHE gene for milk fat percentage and bone density				
Sheep (468)	7208 CNVs 365 CNVRs	118.36 Mb	4.05%	OvineSNP50K	Gerlando *et al*., 2019 [74]
Species/Breed (Nos.)	**Details of CNVs**			**Method**	**Author(s)**
	No	**Size**	**Genome coverage**		
Goat					
Goat	161 CNVs 127 CNVRs	11.47 Mb	0.44%	aCGH (bovine-caprine)	Fontanesi *et al*., 2009 [75]
Goats (8)	13347 CNVs	*ASIP* gene duplication for light colour coat in goats			Dong *et al*., 2015 [76]
Goat (1023)	6286 CNVs 978 CNVRs	262 Mb	8.96%	Caprine SNP50K	Liu *et al*., 2019b [47]
Goat (20)	6 CNVs	*KIT* and *ASIP* genes involved in skin pigmentation		WGS	Henkel *et al*., 2019 [77]
Laoshan Dairy Goat High and low fecundity group	13 CNVs	Three times copy numbers duplication in *PRP1* gene and 6 times in *PRP6* gene was associated with high fecundity		WGS (RD)	Zhang *et al*., 2019 [78]
African goat (126)	30 CNVs			Caprine SNP50K	Liu *et al*., 2020 [79]
Species/Breed (Nos.)	**Details of CNVs**			**Method**	**Author(s)**
	No	**Size**	**Genome coverage**		
Pig					
Duroc pig (12)	165 CNVs 37 CNVRs			aCGH	Fadista *et al*., 2008 [80]
Pig (1693)	1315 CNVs 565 CNVRs	143.03 Mb	5.84%	Porcine SNP60K	Chen *et al*., 2012 [81]
Pig (12)	1344 CNVRs	47.79 Mb	1.7%	aCGH	Wang *et al*., 2014 [82]
Pig (678) Belong three generations	48 CNVs	8 CNVs in 6 chromosomes for meat quality trait		Porcine SNP60K	Wang *et al*., 2015 [83]
Pig (16)	1408 CNVRs	Olfactory receptors for food		WGS (RD)	Paudel *et al*., 2015 [84]

		foraging and recognition of partner			
Pig (7)	1279 CNVs 540 CNVRs	Fatty acid composition and growth traits		NGS	Revilla *et al*., 2017 [85]
Pig (660)	7097 CNVs 271 CNVRs	Used for eliminating effect of PRRS virus		Porcine SNP60K	Hay *et al*., 2017 [86]
Pigs (240)	39315 CNVs 3538 CNVRs	22.9 Mb	0.94%	WGS (RD)	Keel *et al*., 2019 [87]
Duroc pig (3892)	46118 CNVs 425 CNVRs	197 Mb	7.1%	Porcine SNP80K	Stafuzza *et al*., 2019 [88]
		KIT gene duplication was associated with coat colour			
Large White Pig (857)	312 CNVRs	57.76 Mb	2.36%	Porcine SNP80K	Wang *et al*., 2019 [89]
		GPER1 gene CNV for reproduction			
Meishan and Duroc (61)	12668 CNVRs	3.78 kb	1.71%	WGS (RD)	Zheng *et al*., 2020 [90]
	AHR gene copy number is directly proportional with total number of piglets, number of alive piglets and birth weight				
Large white pigs (857)	4070 CNVs 312 CNVRs	57.76 Mb (2.36%)		PorcineSNP80	Wang *et al*., 2020 [91]
Species/Breed (Nos.)	**Details of CNVs**			**Method**	**Author(s)**
	No	**Size**	**Genome coverage**		
Chicken					
Chicken (64)	3154 CNVs 1556 CNVRs	60 Mb	5.4%	aCGH	Crooijmans *et al*., 2013 [92]
Chicken (12)	8840 CNVRs	98.2 Mb	9.4%	WGS (RD)	Yi *et al*., 2014 [93]
	Associated with disease susceptibility/resistance (*FZD6* and *LIMS1* gene), higher bone mineral density (duplication of *SOCS2* gene)				
RJF (5); CN (20) RIR (20); WL (20)	3079 CNVRs 663 CNVRs	Associated with metabolism and organ development		WGS (RD)	Seol *et al*., 2019 [94]

Conclusions

Copy number variations are unbalanced structural variations, which are responsible for larger genetic variations with several mega base pair length. CNVs can be detected in terms of density of nucleotides in a genomic region and can be described as deletions and duplications. It is observed that CNVs have genomic coverage up to nearly 10% of the genome. These segmental rearrangements act as source of alteration of genetic structure and quantity involving genes. Various traits such as, production, reproduction, growth, morphology, physiology, adaptation, disease resistance etc are associated with segmental variations. Thus, CNVs play vital role for studying genetic variations. Focus should be put on detecting large-scale changes along with specific nucleotide sequences throughout the whole genome in order to get genome level studies more systematic and more significant.

References

1. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. Nature Reviews. Genetics. 2011;12(5):363-376.
2. Chen L, Chamberlain AJ, Reich CM, Daetwyler HD, Hayes BJ. Detection and validation of structural variations in bovine whole-genome sequence data. Genetics Selection Evolution. 2017;49(1):13.
3. 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.
4. Couldrey C, Keehan M, Johnson T, Tiplady K, Winkelman A, Littlejohn M, *et al*. Detection and assessment of copy number variation using PacBio long read and Illumina sequencing in New Zealand dairy cattle. Journal of Dairy Science. 2017;100(7):5472-5478.
5. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, *et al*. Segmental Duplications and Copy-Number Variation in the Human Genome. American Journal of Human Genetics. 2005;77(1):78–88.
6. McCarroll SA, Altshuler DM. Copy-number variation and association studies of human disease. Nature Genetics. 2007;39(7 Suppl):S37-42.
7. Zhang F, Gu W, Hurles ME, Lupski JR. Copy Number Variation in Human Health, Disease and Evolution. Annual Review of Genomics and Human Genetics. 2009;10:451-481.
8. Upadhyay M, da Silva VH, Megens HJ, Visker MHPW, Ajmone-Marsan P, Bâlteanu VA, *et al*. Distribution and Functionality of Copy Number Variation across European Cattle Populations. Frontiers in Genetics. 2017;8:108.
9. Kommadath A, Grant JR, Krivushin K, Butty AM, Baes CF, Carthy TR, *et al*. A large interactive visual database of copy number variants discovered in taurine cattle. Giga Science. 2019;8(6):giz073.
10. Mielczarek M, Frąszczak M, Giannico R, Minozzi G, Williams JL, Wojdak-Maksymiec K, *et al*. Analysis of copy number variations in Holstein- Friesian cow genomes based on whole-genome sequence data. Journal of Dairy Science. 2017;100(7):5515-5525.
11. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, *et al*. Origins and functional impact of copy number variation in the human genome. Nature. 2010;464(7289):704-712.
12. Fadista J, Thomsen B, Holm LE, Bendixen C. Copy number variation in the bovine genome. BMC Genomics. 2010;11(1):284.
13. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, *et al*. Global variation in copy number in

the human genome. Nature. 2006;444(7118):444–454.
14. Hou CL, Meng FH, Wang W, Wang SY, Xing YP, Cao JW, *et al*. Genome-wide analysis of copy number variations in Chinese sheep using array comparative genomic hybridization. Small Ruminant Research, 2015;128:19-26.
15. Clancy S. Genetic mutation. Nature Education. 2008;1(1):187.
16. Yang L, Xu L, Zhu B, Niu H, Zhang W, Miao J, *et al*. Genome-wide analysis reveals differential selection involved with copy number variation in diverse Chinese Cattle. Scientific Reports. 2017;7(1):14299.
17. Emerson JJ, Cardoso-Moreira M, Borevitz JO, Long M. Natural Selection Shapes Genome-Wide Patterns of Copy-Number Polymorphism in Drosophila melanogaster. Science. 2008;320(5883):1629-1631.
18. Zhang L, Jia S, Yang M, Xu Y, Li C, Sun J, *et al*. Detection of copy number variations and their effects in Chinese bulls. BMC Genomics. 2014;15(1):480.
19. Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, *et al*. Segmental copy number variation shapes tissue transcriptomes. Nature Genetics. 2009;41(4):424-429.
20. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, *et al*. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science (New York, N.Y.). 2007;315(5813):848–853.
21. Keel BN, Lindholm-Perry AK, Snelling WM. Evolutionary and Functional Features of Copy Number Variation in the Cattle Genome. Frontiers in Genetics. 2016;7:207.
22. 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 Research. 2012;22(4):778-790.
23. Shin DH, Lee HJ, Cho S, Kim HJ, Hwang JY, Lee CK, *et al*. Deleted copy number variation of Hanwoo and Holstein using next generation sequencing at the population level. BMC Genomics. 2014;15(1):240.
24. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, *et al*. Large-scale copy number polymorphism in the human genome. Science (New York, N.Y.). 2004;305(5683):525-528.
25. Stankiewicz P, Lupski JR. Structural Variation in the Human Genome and its Role in Disease. Annual Review of Medicine. 2010;61(1):437-455.
26. Gu W, Lupski JR. CNV and nervous system diseases – what's new? Cytogenetic and Genome Research. 2008;123(1-4):54-64.
27. Hastings P, Lupski JR, Rosenberg SM, Ira G. Mechanisms of change in gene copy number. Nature Reviews. Genetics. 2009;10(8):551-564.
28. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang L, Cellamare A, *et al*. Analysis of copy number variations among diverse cattle breeds. Genome Research. 2010;20(5):693-703.
29. Lieber MR, Ma Y, Pannicke U, Schwarz K. Mechanism and regulation of human non-homologous DNA end-joining. Nature reviews Molecular cell biology. 2003;4(9):712-720.
30. Lieber MR. The mechanism of human nonhomologous DNA end joining. The Journal of Biological Chemistry. 2008;283(1):1-5.
31. Kim PM, Lam HYK, Urban AE, Korbel JO, Affourtit J, Grubert F, *et al*. Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. Genome Research. 2008;18(12):1865-1874.
32. Cahan P, Li Y, Izumi M, Graubert TA. The impact of copy number variation on local gene expression in mouse hematopoietic stem and progenitor cells. Nature Genetics. 2009;41(4):430-437.
33. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, *et al*. A mouse model of human L1 retrotransposition. Nature Genetics. 2002;32(4):655–660.
34. Beck CR, Collier P, Macfarlane C, Malig M, Kidd JM, Eichler EE, *et al*. LINE-1 Retrotransposition Activity in Human Genomes. Cell. 2010;141(7):1159–1170.
35. Lupski JR. Retrotransposition and Structural Variation in the Human Genome. Cell. 2010;141(7):1110–1112.
36. Jenkins GM, Goddard ME, Black MA, Brauning R, Auvray B, Dodds KG, *et al*. Copy number variants in the sheep genome detected using multiple approaches. BMC Genomics. 2016a;17(1):441.
37. Egan CM, Sridhar S, Wigler M, Hall IM. Recurrent DNA copy number variation in the laboratory mouse. Nature Genetics. 2007;39(11):1384-1389.
38. Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, *et al*. Whole Genome Sequencing in Autism Identifies Hotspots for De Novo Germline Mutation. Cell. 2012;151(7):1431-1442.
39. Paudel Y, Madsen O, Megens HJ, Frantz LAF, Bosse M, Crooijmans RPMA, *et al*. Copy number variation in the speciation of pigs: A possible prominent role for olfactory receptors. BMC Genomics. 2015;16(1):330.
40. Letaief R, Rebours E, Grohs C, Meersseman C, Fritz S, Trouilh L, *et al*. Identification of copy number variation in French dairy and beef breeds using next-generation sequencing. Genetics Selection Evolution. 2017;49(1):77.
41. Freeman JL, Perry GH, Feuk L, Redon R, McCarroll SA, Altshuler DM, *et al*. Copy number variation: New insights in genome diversity. Genome Research. 2006;16(8):949-961.
42. Tüzün E, Sharp A, Bailey J, Kaul R, Morrison V, Pertz L, *et al*. Fine-scale structural variation of the human genome. Nature Genetics. 2005;37:727-732.
43. Pinto D, Darvishi K, Shi X, Rajan D, Rigler D, Fitzgerald T, *et al*. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. Nature Biotechnology. 2011;29(6):512-520.
44. Winchester L, Yau C, Ragoussis J. Comparing CNV detection methods for SNP arrays. Briefings in Functional Genomics. 2009;8(5):353-366.
45. Selzer RR, Richmond TA, Pofahl NJ, Green RD, Eis PS, Nair P, *et al*. Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes, Chromosomes and Cancer. 2005;44(3):305-319.
46. Liu S, Kang X, Catacchio CR, Liu M, Fang L, Schroeder SG, *et al*. Computational detection and experimental validation of segmental duplications and associated copy number variations in water buffalo (Bubalus bubalis). Functional and Integrative Genomics. 2019a;19(3):409-419.
47. Liu M, Zhou Y, Rosen BD, Van Tassell CP, Stella A, Tosser-Klopp G, *et al*. Diversity of copy number

variation in the worldwide goat population. Heredity. 2019b;122(5):636-646.
48. Gao Y, Jiang J, Yang S, Hou Y, Liu GE, Zhang S, *et al*. CNV discovery for milk composition traits in dairy cattle using whole genome resequencing. BMC Genomics. 2017;18(1):265.
49. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, *et al*. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008;456(7218):53-59.
50. Pirooznia M, Goes FS, Zandi PP. Whole-genome CNV analysis: Advances in computational approaches. Frontiers in Genetics. 2015;6:138.
51. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, *et al*. Paired-End Mapping Reveals Extensive Structural Variation in the Human Genome. Science. 2007;318(5849):420-426.
52. 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.
53. Bae JS, Cheong HS, Kim LH, NamGung S, Park TJ, Chun JY, *et al*. Identification of copy number variations and common deletion polymorphisms in cattle. BMC Genomics. 2010;11(1):232.
54. Liu GE, Hou Y, Zhu B, Cardone MF, Jiang L, Cellamare A, *et al*. Analysis of copy number variations among diverse cattle breeds. Genome Research. 2010;20(5):693-703.
55. Stothard P, Choi JW, Basu U, Sumner-Thomson JM, Meng Y, Liao X, *et al*. Whole genome resequencing of black Angus and Holstein cattle for SNP and CNV discovery. BMC Genomics. 2011;12(1):559.
56. 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 Genomics. 2011;12(1):557.
57. Hou Y, Liu GE, Bickhart DM, Cardone MF, Wang K, Kim E, *et al*. Genomic characteristics of cattle copy number variations. BMC Genomics. 2011;12(1):127.
58. Kijas J, Barendse W, Barris W, Harrison B, McCulloch R, McWilliam S, *et al*. Analysis of copy number variants in the cattle genome. Gene. 2011;482:73-77.
59. Jiang L, Jiang J, Wang J, Ding X, Liu J, Zhang Q. Genome-Wide Identification of Copy Number Variations in Chinese Holstein. Plos One. 2012;7(11):e48732.
60. Hou Y, Bickhart DM, Hvinden ML, Li C, Song J, Boichard DA, *et al*. Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. BMC Genomics. 2012;13(1):376.
61. Choi JW, Lee KT, Liao X, Stothard P, An HS, Ahn S, *et al*. Genome-wide copy number variation in Hanwoo, Black Angus and Holstein cattle. Mammalian Genome: Official Journal of the International Mammalian Genome Society. 2013;24(3-4):151-163.
62. Jiang L, Jiang J, Yang J, Liu X, Wang J, Wang H, *et al*. Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. BMC Genomics. 2013;14(1):131.
63. Boussaha M, Esquerré D, Barbieri J, Djari A, Pinton A, Letaief R, *et al*. Genome-Wide Study of Structural Variants in Bovine Holstein, Montbéliarde and Normande Dairy Breeds. Plos one. 2015;10(8):e0135931.
64. Wang L, Xu L, Liu X, Zhang T, Li N, Hay EH, *et al*. Copy number variation-based genome wide association study reveals additional variants contributing to meat quality in Swine. Scientific Reports. 2015;5:12535.
65. Sasaki S, Watanabe T, Nishimura S, Sugimoto Y. Genome-wide identification of copy number variation using high-density single-nucleotide polymorphism array in Japanese Black cattle. BMC Genetics. 2016;17(1):26.
66. 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(1):454.
67. Li W, Bickhart DM, Ramunno L, Iamartino D, Williams JL, Liu GE. Comparative sequence alignment reveals River Buffalo genomic structural differences compared with cattle. Genomics. 2019;111(3):418-425.
68. Fontanesi L, Beretti F, Martelli PL, Colombo M, Dall'olio S, Occidente M, *et al*. A first comparative map of copy number variations in the sheep genome. Genomics. 2011;97(3):158-165.
69. Liu J, Zhang L, Xu L, Ren H, Lu J, Zhang X, *et al*. Analysis of copy number variations in the sheep genome using 50K SNP BeadChip array. BMC Genomics. 2013;14(1):229.
70. Jenkins GM, Goddard ME, Black MA, Brauning R, Auvray B, Dodds KG, *et al*. Copy number variants in the sheep genome detected using multiple approaches. BMC Genomics. 2016a;17(1):441.
71. Zhu C, Fan H, Yuan Z, Hu S, Ma X, Xuan J, *et al*. Genome-wide detection of CNVs in Chinese indigenous sheep with different types of tails using ovine high-density 600K SNP arrays. Scientific Reports. 2016;6:27822.
72. Ma Q, Liu X, Pan J, Ma L, Ma Y, He X, *et al*. Genome-wide detection of copy number variation in Chinese indigenous sheep using an ovine high-density 600 K SNP array. Scientific Reports. 2017;7(1):912.
73. Jiang R, Cheng J, Cao XK, Ma YL, Chaogetu B, Huang YZ, *et al*. Copy Number Variation of the SHE Gene in Sheep and Its Association with Economic Traits. Animals: An Open Access Journal from MDPI. 2019;9(8).
74. Gerlando RD, Sutera AM, Mastrangelo S, Tolone M, Portolano B, Sottile G, *et al*. Genome-wide association study between CNVs and milk production traits in Valle del Belice sheep. Plos one. 2019;14(4):e0215204.
75. Fontanesi L, Beretti F, Riggio V, Gómez González E, Dall'Olio S, Davoli R, *et al*. Copy number variation and missense mutations of the agouti signaling protein (ASIP) gene in goat breeds with different coat colors. Cytogenetic and Genome Research. 2009;126(4):333-347.
76. Dong Y, Zhang X, Xie M, Arefnezhad B, Wang Z, Wang W, *et al*. Reference genome of wild goat (*Capra aegagrus*) and sequencing of goat breeds provide insight into genic basis of goat domestication. BMC Genomics. 2015;16:431.
77. Henkel J, Saif R, Jagannathan V, Schmocker C, Zeindler F, Bangerter E, *et al*. Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. PLOS Genetics. 2019;15(12):e1008536.
78. Zhang RQ, Wang JJ, Zhang T, Zhai HL, Shen W. Copy-

number variation in goat genome sequence: A comparative analysis of the different litter size trait groups. Gene. 2019;696:40-46.
79. Liu M, Woodward-Greene J, Kang X, Pan MG, Rosen B, Van Tassell CP, *et al*. Genome-wide CNV analysis revealed variants associated with growth traits in African indigenous goats. Genomics. 2020;112(2):1477-1480.
80. Fadista J, Nygaard M, Holm LE, Thomsen B, Bendixen C. A snapshot of CNVs in the pig genome. PloS One. 2008;3(12):e3916.
81. Chen C, Qiao R, Wei R, Guo Y, Ai H, Ma J, *et al*. A comprehensive survey of copy number variation in 18 diverse pig populations and identification of candidate copy number variable genes associated with complex traits. BMC Genomics. 2012;13(1):733.
82. Wang J, Jiang J, Wang H, Kang H, Zhang Q, Liu JF. Enhancing genome- wide copy number variation identification by high density array CGH using diverse resources of pig breeds. PloS One. 2014;9(1):e87571.
83. Wang L, Xu L, Liu X, Zhang T, Li N, Hay EH, *et al*. Copy number variation-based genome wide association study reveals additional variants contributing to meat quality in Swine. Scientific Reports. 2015;5:12535.
84. Paudel Y, Madsen O, Megens HJ, Frantz LAF, Bosse M, Crooijmans RPMA, Groenen MAM. Copy number variation in the speciation of pigs: A possible prominent role for olfactory receptors. BMC Genomics. 2015;16(1):330.
85. Revilla M, Puig-Oliveras A, Castelló A, Crespo-Piazuelo D, Paludo E, Fernández AI, *et al*. A global analysis of CNVs in swine using whole genome sequence data and association analysis with fatty acid composition and growth traits. Plos one. 2017;12(5):e0177014.
86. Hay EH, Choi I, Xu L, Zhou Y, Rowland R, Lunney J, Liu G. CNV Analysis of Host Responses to Porcine Reproductive and Respiratory Syndrome Virus Infection. Journal of Genomics. 2017;5:58-63.
87. Keel BN, Nonneman DJ, Lindholm-Perry AK, Oliver WT, Rohrer GA. A Survey of Copy Number Variation in the Porcine Genome Detected From Whole- Genome Sequence. Frontiers in Genetics. 2019;10:737.
88. Stafuzza NB, Silva RM, de O, Fragomeni B de O, Masuda Y, Huang Y, *et al*. A genome-wide single nucleotide polymorphism and copy number variation analysis for number of piglets born alive. BMC Genomics. 2019;20(1):321.
89. Wang C, Chen H, Wang X, Wu Z, Liu W, Guo Y, *et al*. Identification of copy number variations using high density whole-genome SNP markers in Chinese Dongxiang spotted pigs. Asian-Australasian Journal of Animal Sciences, 2019, 1809-1815.
90. Zheng X, Zhao P, Yang K, Ning C, Wang H, Zhou L, Liu J. CNV analysis of Meishan pig by next-generation sequencing and effects of AHR gene CNV on pig reproductive traits. Journal of Animal Science and Biotechnology. 2020;11(1):42.
91. Wang Y, Zhang T, Wang C. Detection and analysis of genome-wide copy number variation in the pig genome using an 80 K SNP Beadchip. Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie. 2020;137(2):166–176.
92. Crooijmans RP, Fife MS, Fitzgerald TW, Strickland S, Cheng HH, Kaiser P, *et al*. Large scale variation in DNA copy number in chicken breeds. BMC Genomics. 2013;14(1):398.
93. 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.
94. Seol D, Ko BJ, Kim B, Chai HH, Lim D, Kim H. Identification of Copy Number Variation in Domestic Chicken Using Whole-Genome Sequencing Reveals Evidence of Selection in the Genome. Animals: An Open Access Journal from MDPI. 2019;9(10):809.