

Copy number variation in livestock: A mini review

V. Bhanuprakash, Supriya Chhotaray, D. R. Pruthviraj, Chandrakanta Rawat, A. Karthikeyan and Manjit Panigrahi

Division of Animal Genetics, ICAR - Indian Veterinary Research Institute, Izatnagar, Bareilly - 243122, Uttar Pradesh, India.

Corresponding author: V. Bhanuprakash, e-mail: bhanupvhn24@gmail.com

Co-authors: SC: supriya.chhotaray@gmail.com, DRP: saadhane.pdrs@gmail.com, CR: rawatchandrakanta@gmail.com, KA: karthikeyan0318@gmail.com, MP: manjit707@gmail.com

Received: 10-02-2018, **Accepted:** 31-03-2018, **Published online:** 26-04-2018

doi: 10.14202/vetworld.2018.535-541 **How to cite this article:** Bhanuprakash V, Chhotaray S, Pruthviraj DR, Rawat C, Karthikeyan A, Panigrahi M (2018) Copy number variation in livestock: A mini review, *Veterinary World*, 11(4): 535-541.

Abstract

Copy number variation (CNV) is a phenomenon in which sections of the genome, ranging from one kilo base pair (Kb) to several million base pairs (Mb), are repeated and the number of repeats vary between the individuals in a population. It is an important source of genetic variation in an individual which is now being utilized rather than single nucleotide polymorphisms (SNPs), as it covers the more genomic region. CNVs alter the gene expression and change the phenotype of an individual due to deletion and duplication of genes in the copy number variation regions (CNVRs). Earlier, researchers extensively utilized SNPs as the main source of genetic variation. But now, the focus is on identification of CNVs associated with complex traits. With the recent advances and reduction in the cost of sequencing, arrays are developed for genotyping which cover the maximum number of SNPs at a time that can be used for detection of CNVRs and underlying quantitative trait loci (QTL) for the complex traits to accelerate genetic improvement. CNV studies are also being carried out to understand the evolutionary mechanism in the domestication of livestock and their adaptation to the different environmental conditions. The main aim of the study is to review the available data on CNV and its role in genetic variation among the livestock.

Keywords: copy number variation, copy number variation regions, livestock, single nucleotide polymorphisms, quantitative trait loci.

Introduction

The copy number variants (CNVs) are a structural variation in the genome of an individual in the form of losses or gains of DNA fragments. CNV is an important source of genetic and phenotypic variation [1]. Union of overlapping CNVs detected in two different samples are called copy number variation regions (CNVRs) [2]. The difference in the copy number of CNVR genes results in changes in the gene expression and phenotypic variation due to altering gene dosage and gene disruption effect by the deletion, duplication, inversions, and translocations of the gene. It is a source for evolutionary mechanisms [3]. If CNV exists in the protein coding region, it alters the protein function, whereas in the regulatory region, it alters the gene expression level [4]. The current review helps in understanding the CNV and its role in the improvement of economic traits in livestock.

Mechanisms of CNV Formation

Non-allelic homologous recombination (NAHR), non-homologous end-joining (NHEJ), fork stalling and template switching (FoSTeS), and L1-mediated retro transposition are some of the mechanisms which

generate rearrangements in the genome and possibly account for the majority of CNV formation [5,6]. NAHR occurs in meiosis and mitosis due to recombination between the two regions of similar sequence between the non-homologous chromosomes. If crossing over occurs between the sister chromatids, it can increase the segment of DNA at the expense of another which may result in duplication, deletion, and inversion of the segment of chromosome. 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 [7,8]. FoSTeS is a DNA replication-based mechanism which can account for Complex Genomic Rearrangements and CNVs [9]. L1 transposition occurs through reverse transcription and integration [10]. A number of studies have been carried out to identify the CNV in different species such as cattle [11-14], sheep [15,16], goat [17], pig [18-21] and chicken [22].

Algorithm used for Identification of CNV

SNP arrays are being used normally for CNV detection and analysis in humans because of its availability and economic feasibility [2]. In general, most of the studies reported in literature for CNV detection in the population study used comparative genomic hybridization (CGH) arrays and SNP genotyping arrays [23]. Now a days, CNV detection and analysis by whole genome sequencing is practically possible due to decreased cost for next-generation sequencing

Copyright: Bhanuprakash, *et al.* Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

(NGS) techniques. Relatively, sequencing has high resolution over genotyping as it covers the entire genome [24]. SNP arrays utilize a Log R ratio (LRR) and B allele frequency (BAF) which represents the copy numbers and allelic status of the population [25]. Large CNVRs are mostly identified with the SNP50 array since it lacks non-polymorphic probes. Multiple algorithms have been used to identify CNVs and CNVRs [26-29].

PennCNV Software

It is a Hidden Markov Model (HMM) algorithm which integrates multiple parameters such as LRR, BAF, the population frequency of the B allele (PFB) of SNPs, the distance between neighboring SNPs and the allele frequency of SNPs [25-27]. It is based on fitting regression models with GC content to overcome genomic waves [30]. It improves the call rate and accuracy of boundary mapping by considering the pedigree information [12].

CnvPartition

CnvPartition is based on a different proprietary sliding window approach which detects CNVs by processing LRR and BAF. Only those homozygous deletion events segregating in different animals were reported by this algorithm due to concern quality calls [31].

cn.MOPS Algorithm

The cn.MOPS (Mixture of PoissonS) algorithm is based on the Bayesian approach for the detection of CNV in multiple samples for NGS data. It decomposes read variations across multiple samples into integer copy numbers and noise by its mixture components and Poisson distributions, respectively. The advantages of using this method are- it identifies overlapping sequences and estimates allele-specific copy numbers [32].

QuantiSNP

QuantiSNP uses different HMMs unlike PennCNV. QuantiSNP uses both LRR and BAF frequency independently whereas in pennCNV treat them as combined. It uses a fixed rate of heterozygosity for each SNP [26].

CNVFinder

It is a python package for CNV detection on whole exome sequencing data from amplicon-based enrichment technologies. This program uses SDe termed as experimental variability, in the LRR distribution [33].

CNV Identification Studies in Domestic Species

Cattle

Upadhyay *et al.* [34] using the Illumina BovineHD Genotyping in cattle identified 9944 CNVs and 923 CNVRs with a length of 61.06 Mb covering

2.5% of the bovine autosomes. These CNVRs were found to be associated with the quantitative trait loci which affect production traits, body measurements, and parasite resistance [23,35,36]. Incidence of overlap of CNVs reported among *taurine* cattle is higher than the overlap between *taurine* and *indicine* cattle. Largest CNV diversity was reported among the zebu cattle [37].

Recent studies reported that CNVs evolved 2.5 folds faster than SNPs and helped to promote a better adaptation in different environments [21]. Liu *et al.* [4] reported the high CNV abundance in *indicine* and African *taurine* cattle breeds than in European *taurine* using Vst for population differentiation which indicates the breed divergence and population history. Pezer *et al.* [38] suggested the variation in the CNV number may be due to the difference in effective population size, gene flow, and selection process among different populations. Upadhyay *et al.* [34] in their study reported that small populations might cause an increase in the CNVs, particularly deletion in CNVs. The discrepancy between the studies observed is due to the small sample size within the breed, large samples from multiple breeds and different SNP arrays used in the study [39]. Different studies using the same method and algorithm for the detection of CNVs, reported varying overlaps. The inconsistency of this overlaps between the studies is due to the platforms and algorithms of CNV calling, differences in size, and population structure under investigation.

Hou *et al.* [36] reported that the more CNV events were detected in *indicine* than in African groups and *taurine* breeds. This observation may suggest the independent domestication events of cattle in Europe, Africa, and Southeast Asia [40,41]. Hou *et al.* [36] using array CGH technique identified 25 germline CNVs in three Holstein bulls. The same group identified over 200 CNVRs from diverse cattle breeds. Fadista *et al.* [42] identified 304 CNVRs with a length of 15.8 Mb of the genome from 20 animals of 4 cattle breeds [36]. Bae *et al.* [11] also identified 368 unique CNV regions from 265 Korean Hanwoo cattle based on 50K SNP array covering 15.8Mb of the genome using PennCNV algorithm without considering the pedigree information and genomic waves. Recent studies on CNV detection using different approaches are given in the Table-1 [11-14,23,30,34-37,42-46] for cattle and Table-2 [14,16,47-52] for other species.

Sheep

Yang *et al.* [16] identified 619 CNV regions, covering 197 Mb, corresponding to ~6.9% of the sheep genome and found several important CNV overlapping genes (*BTG3*, *PTGSI*, and *PSPH*) which are involved in the fetal muscle development, prostaglandin (PG) synthesis, and bone color. Ma *et al.* [48] identified 111 CNVRs from 160 Chinese sheep breeds covering 13.75 Mb of the sheep genome sequence. Fontanesi *et al.* [15] using a cattle-sheep 385K aCGH

Table-1: CNVs detection studies in cattle using different approaches.

Author	Sample size	Breed	Method	CNVs	CNVRs	Total length (Mb)
Bae <i>et al.</i> [11]	265	1	50K	855	368	63.1
Hou <i>et al.</i> [12]	539	21	50K	3666	743	15.8
Jiang <i>et al.</i> [13]	2047	1	50K	219 ^a 169 ^b 140 ^c	101	23.8
Wang <i>et al.</i> [14]	492	1	50K	-	389	70.4
Zhang <i>et al.</i> [23]	14	29	CGH	-	605	2.45
Bickhart <i>et al.</i> [30]	6	3	WGS	-	1265	55.6
Upadhyay <i>et al.</i> [34]		38	770K	9944	196	61.1
Xu <i>et al.</i> [35]	300	8	770K		257	12.4
Hou <i>et al.</i> [36]	674	27	770K	34311	3438	147
Da Silva <i>et al.</i> [37]	1717	1	770K	68007	7319	15.6
Fadista <i>et al.</i> [42]	20	4	CGH	-	254	15.8
Jiang <i>et al.</i> [43]	96	1	770K	1733	357	34.4
Sasaki <i>et al.</i> [44]	1481	1	770K	55593	861	43.6
Liu <i>et al.</i> [45]	20	17	CGH	-	200	36.2
Stothard <i>et al.</i> [46]	2	2	WGS		790	3.3

CNVRs identified by the different algorithm: Superscript a-PennCNV, b-GADA (Genome Alteration Detection Algorithm) and c-cnVPartition. CNV=Copy number variation, CGH=Comparative genomic hybridization, CNVRs=Copy number variation regions

Table-2: CNVs detection studies in domestic animals.

Species	Breed	CNVRs	Length (Mb)	Method	References
Pig	13	49	3131	WGS	Wang <i>et al.</i> [14]
Sheep	68	619	197	OvineSNP50 assay	Yang <i>et al.</i> [16]
Sheep	11	135	77.6	Bovine 385KaCGH arrays	Iafra <i>et al.</i> [47]
Sheep	48	1296	121.8	OvineHD 600K SNP array	Ma <i>et al.</i> [48]
Sheep	3	238	60.35	OvineSNP50 assay	Liu <i>et al.</i> [49]
Goat	10	127	90.3	Bovine 385KaCGH arrays	Hutt <i>et al.</i> [50]
Pig	55	49	754.6	Porcine SNP60 Beadchip	Kijas <i>et al.</i> [51]
Pig	2	172	80.41	PorcineSNP60	Xie <i>et al.</i> [52]

CNV=Copy number variation, CGH=Comparative genomic hybridization, CNVRs=Copy number variation regions, SNP=Single nucleotide polymorphisms

identified 135 CNV regions covering ~10.5 Mb of the sheep genome with reference to the bovine genome. It may suggest that several chromosome regions have a repetitive sequence of CNVRs between the species. CNVR overlapping with the homeobox transcription factor DLX3 was found to be associated with curly hair in sheep [48].

Goat

Fontanesi *et al.* [17] identified 127 CNVRs covering about 11.47 Mb of the goat genome with reference to the bovine genome. Genes with environmental functions were over-represented in goat CNVRs as reported in other mammals [17]. Difference in the copy number at Agouti locus in sheep and goats contributes to the variability of coat color [53].

Horse

Copy number variants account for about 1-3% of the horse genome and mostly of intra-genic than those located in intergenic regions [54]. Ghosh *et al.* [55] using 400K WG tiling oligo array identified 258 CNV regions (CNVRs) comprised of 1.3% of the horse genome across all chromosome except chrY in 16 diverse breeds of horse and also found 20% of the identified CNVRs were located in intergenic regions.

Chicken

Chicken has a unique genome arrangement due to the presence of micro- and macro-chromosome [56]. Griffin *et al.* [57] first studied the chicken CNV with aCGH to establish interspecies genomic rearrangement and they showed that there are more CNVs that involve coding genes than the non-coding sequences. Studies reported the phenotypic association of CNV in chickens, which include a pea-comb, late-feathering, dark brown plumage color, and dermal hyperpigmentation [16,58]. Recent studies for CNVs detection in chicken by different approaches are given in Table-3 [22, 59-66].

CNVRs genes and gene ontology (GO)

Bickhart *et al.* [30] reported the duplication of cathelicidin genes (*CATHL4*) in the Nellore cattle sample, but these genes were found only in a single copy in human and mice. CNV overlapping with *KIT* gene was found to be associated with color-sidedness in English Longhorn cattle [34]. Several CNVs have been identified in cattle for association with milk production traits [35]. Upadhyay *et al.* [34] found genes related to economically important traits of livestock such as *MTHFSD* and *GTF2I* in the CNVRs. Yoshida *et al.* [67] found that the complex

Table-3: CNVs detection studies in chicken using different approaches

Author	Sample size	Method	CNVs	CNVRs	Percentage coverage	Total length (Mb)
Crooijmans <i>et al.</i> [22]	64	aCGH	3154	1556	5.4	60
Zhang <i>et al.</i> [59]	475	60K SNP array	438 ^a 291 ^b	271 ^a 188 ^b	3.92 ^a 2.98 ^b	40.26 ^a 30.60 ^b
Jia <i>et al.</i> [60]	746	60K SNP array	818	209	1.42	13.55
Yi <i>et al.</i> [61]	12	WGS	-	8840	9.4	98.2
Han <i>et al.</i> [62]	10	385 (aCGH) Genome array	-	281	1.07	12
Rao <i>et al.</i> [63]	554	60K SNP array	1875	383	3.97	41
Gorla <i>et al.</i> [64]	256	600K SNP Array	1924	1216	5.12	47
Strillacci <i>et al.</i> [65]	96	580K SNP array	1003	564	1.03	9.43
Fan <i>et al.</i> [66]	2	WGS	-	8839	24.6	

^aLean lines, ^bFat lines in chicken. CNV=Copy number variation, CGH=Comparative genomic hybridization, CNVRs=Copy number variation regions, SNP=Single nucleotide polymorphisms

BoLA-DRB3 lies in the CNVR associated with Mastitis and Bovine leukemia virus infection in various cattle breeds. Liu *et al.* [45] found the duplication in the *CIITA* gene that showed nematode resistance in Angus cattle. MTHFSD gene covering the CNVR1206 found to be associated with milk protein yield in Spanish HF cattle [68]. Reyer *et al.* [69] found GTF2I in the CNVR1703 region which is associated with feed conversion efficiency in chicken. The *MSH4* gene was found to be associated with impaired gamete formations in laboratory mice and recombination rate in cattle [48, 70, 71]. Brenig *et al.* [72] suggested a dose-dependent effect of Belgian-blue type allele in White Park and Galloway cattle in which uneven pigmented spots were seen in the heterozygous condition, whereas in homozygous condition, there was no pigmentation on the body. Gene Ontology (GO) study revealed that CNVRs are particularly enriched in genes related to immunity, sensory perception, response to external stimuli, and neurodevelopmental processes.

The dominant white coat in sheep is associated with duplication of 190 kb genomic fragment which encompasses three genes viz. the agouti signaling protein (*ASIP*) gene, the itchy E3 ubiquitin protein ligase homolog (mouse) (*ITCH*), and the adenosylhomocysteinase (*AHCY*) loci [73]. Hillbertz *et al.* [74] identified the duplication in fibroblast growth factor genes and the *ORAOVI* gene in Rhodesian and Thai Ridgeback dogs which is responsible for characteristic dorsal hair ridge. A different pattern of white coat color was reported due to the duplication of the *KIT* gene in pig and in some cattle breeds [75].

Sheep and goat

GO analysis and functional studies in sheep reported that many CNVRs are associated with genes related to environmental response and biological functions [48]. Liu *et al.* [76] indicated that ZNF family genes mainly expressed in some sheep breeds are involved in regulating evolutionarily divergent biological traits. Higher expression levels of KIF2A and PHKG2 in Gansu Morden sheep breed compared to other sheep breed indicates the association

of these genes in disease resistance ability [48]. Yang *et al.* [16] found several important CNV-overlapping genes (*BTG3*, *PTGS1*, and *PSPH*) in diverse sheep breeds which were involved in fetal muscle development, PG synthesis, and bone color.

Horse

A homozygous deletion in the *AKRIC* gene may be a possible cause of disorders of sexual development such as male-pseudohermaphroditism due to its role in testicular androgen production and sexual development [55]. GO analysis, and functional studies indicated that the equine CNV genes are mainly involved in biological processes and molecular functions related to transmembrane signal transduction, sensory perception, immune response, reproduction, and steroid metabolism. In horse, *BMPRI1B* gene has been reported for its role in the regulation of the rate of ovulation [77]. Ghosh *et al.* [55] confirmed the role of complex CNVR in chr27 involving *CSMD1* gene which encodes for a transmembrane and a candidate tumor suppressor protein [78].

Chicken

Duplication of segment of DNA at intron 1 non-coding region of the *SOX5* transcription factor interferes with *SOX5* expression, and the regulation of gene expression is critical during cell differentiation for the development of the comb and wattles [79]. Luo *et al.* [80] suggested that CNVs on GGA19 could be a candidate conferring resistance to the Marek's disease. Late feathering locus in the chicken is due to the partial duplication of the *PRLR* and *SPEF2* genes [81]. Lin *et al.* [82] suggested that *SOX6* gene in chicken also has a similar function as reported in many species for the proliferation and differentiation of skeletal muscle cells. *SOX6* gene expression is positively correlated with the number of CNV for CNP13 region in the chicken genome.

Conclusion

Recent studies for CNV detection have enabled the construction of CNV map which in turn helps in identification of CNVs associated with economically important traits. With the advancement

in the techniques and reduced cost of sequencing, researchers are now focusing on the CNV study for detecting the genetic variations, as CNV shows more inclusions and complex genetic variants than SNP sites. Current research for the identification of the CNV regions (CNVRs) throughout the genome in domestic species will change the concept of breeding for genetic improvement. Development of robust and convenient CNV detection techniques could further facilitate unveiling of genetic secrets for molecular breeding of poultry and other farm animals.

Author's Contributions

MP conceptualized and designed the manuscript. BV, SC, and MP prepared manuscript draft and reviewed. CR contributed in literature collection. MP, DRP, and KA edited and made critical comments on the manuscript. MP and BV made critical comments on the revised manuscript and edited for final submission. All authors read and approved the final version.

Acknowledgments

The authors would wish to acknowledge Director and Joint Director (R), IVRI-Bareilly, for their support and providing infrastructural facility to review this study.

Competing Interests

The authors declared that they have no competing interests.

References

- Zhang, F., Gu, W., Hurler, M.E. and Lupski, J.R. (2009) Copy number variation in human health, disease, and evolution. *Annu. Rev. Genomics. Hum. Gene.*,10:451-481.
- Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W. and Cho, E.K. (2006) Global variation in copy number in the human genome. *Nature*,444(7118): 444.
- Emerson, J.J., Cardoso-Moreira, M., Borevitz, J.O. and Long, M. (2008) Natural selection shapes genome-wide patterns of copy-number polymorphism in *Drosophila melanogaster*. *Science*,320(5883): 1629-1631.
- Liu, G.E., Ventura, M., Cellamare, A., Chen, L., Cheng, Z., Zhu, B., Li, C., Song, J. and Eichler, E.E. (2009) Analysis of recent segmental duplications in the bovine genome. *BMC Genomics*,10: 571.
- Gu, W., Zhang, F. and Lupski, J.R. (2008) Mechanisms for human genomic rearrangements. *Pathogenetics*,1: 4.
- Hastings, P.J., Lupski, J.R., Rosenberg, S.M. and Ira, G. (2009) Mechanisms of change in gene copy number. *Nat. Rev. Genet.*,10:551-564.
- Lieber, M.R., Ma, Y., Pannicke, U. and Schwarz, K. (2003) Mechanism and regulation of human non-homologous DNA end-joining. *Nat. Rev. Mol. Cell Biol.*,4:712.
- Lieber, M.R., Lu, H., Gu, J. and Schwarz, K. (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: Relevance to cancer, aging, and the immune system. *Cell Res.*,18: 125.
- Lee, J.A., Carvalho, C.M. and Lupski, J.R. (2007) A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell*,131: 1235-1247.
- Ostertag, E.M. and Kazazian H.H. Jr. (2001) Biology of mammalian L1 retrotransposons. *Annu. Rev. Genet.*,35: 501-538.
- Bae, J.S., Cheong, H.S., Kim, L.H., NamGung, S., Park, T.J., Chun, J.Y., Kim, J.Y., Pasaje, C.F.A., Lee, J.S. and Shin, H.D. (2010) Identification of copy number variations and common deletion polymorphisms in cattle. *BMC Genomics*, 11:232.
- Hou, Y., Liu, G.E., Bickhart, D.M., Cardone, M.F., Wang, K., Kim, E.S., Matukumalli, L.K., Ventura, M., Song, J., VanRaden, P.M. and Sonstegard, T.S. (2011) Genomic characteristics of cattle copy number variations. *BMC Genomics*, 12:127.
- Jiang, L., Jiang, J., Wang, J., Ding, X., Liu, J. and Zhang, Q. (2012) Genome-wide identification of copy number variations in Chinese Holstein. *PLoS One*,7: e48732.
- Wang, M.D., Dzama, K., Hefer, C.A. and Muchadeyi, F.C. (2015) Genomic population structure and prevalence of copy number variations in South African Nguni cattle. *BMC Genomics*,16: 894.
- Fontanesi, L., Beretti, F., Martelli, P.L., Colombo, M., Dall'Olio, S., Occidente, M., Portolano, B., Casadio, R., Matassino, D. and Russo, V. (2011) A first comparative map of copy number variations in the sheep genome. *Genomics*,97:158-165.
- Yang, L., Xu, L., Zhou, Y., Liu, M., Wang, L., Kijas, J.W., Zhang, H., Li, L. and Liu, G.E. (2017) Diversity of copy number variation in a worldwide population of sheep. *Genomics*, 110: 143-148.
- Fontanesi, L., Martelli, P.L., Beretti, F., Riggio, V., Dall'Olio, S., Colombo, M., Casadio, R., Russo, V. and Portolano, B. (2010) An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC Genomics*,11:639.
- Fadista, J., Nygaard, M., Holm, L.E., Thomsen, B. and Bendixen, C. (2008) A snapshot of CNVs in the pig genome. *PLoS One*,3:e3916.
- Ramayo-Caldas, Y., Castelló, A., Pena, R.N., Alves, E., Mercadé, A., Souza, C.A., Fernández, A.I., Perez-Enciso, M. and Folch, J.M. (2010) Copy number variation in the porcine genome inferred from a 60 k SNP bead chip. *BMC Genomics*,11:593.
- Chen, C., Qiao, R., Wei, R., Guo, Y., Ai, H., Ma, J., Ren, J. and Huang, L. (2012) 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*,13:733.
- Paudel, Y., Madsen, O., Megens, H.J., Frantz, L.A., Bosse, M., Crooijmans, R.P. and Groenen, M.A. (2015) Copy number variation in the speciation of pigs: A possible prominent role for olfactory receptors. *BMC Genomics*,16:330.
- Crooijmans, R.P., Fife, M.S., Fitzgerald, T.W., Strickland, S., Cheng, H.H., Kaiser, P., Redon, R. and Groenen, M.A. (2013) Large-scale variation in DNA copy number in chicken breeds. *BMC Genomics*,14:398.
- Zhang, L., Jia, S., Yang, M., Xu, Y., Li, C., Sun, J., Huang, Y., Lan, X., Lei, C., Zhou, Y. and Zhang, C. (2014) Detection of copy number variations and their effects in Chinese bulls. *BMC Genomics*,15:480.
- Alkan, C., Coe, B.P., Eichler, E.E. (2011) Applications of next-generation sequencing genome structural variation discovery and genotyping. *Nat. Rev. Genet.*, 12: 363-375.
- Peiffer, D.A., Le, J.M., Steemers, F.J., Chang, W., Jenniges, T., Garcia, F., Haden, K., Li, J., Shaw, C.A., Belmont, J. and Cheung, S.W. (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res.*,16: 1136-1148.
- Colella, S., Yau, C., Taylor, J.M., Mirza, G., Butler, H., Clouston, P., Bassett, A.S., Seller, A., Holmes, C.C. and Ragoussis, J. (2007) QuantiSNP: An objective bayes hidden-markov model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids*

- Res., 35: 2013-2025.
27. Wang, K., Li, M., Hadley, D., Liu, R., Glessner, J., Grant, S.F., Hakonarson, H. and Bucan, M. (2007) PennCNV: An integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.*,17: 1665-1674.
 28. Korn, J.M., Kuruvilla, F.G., McCarroll, S.A., Wysoker, A., Nemesh, J., Cawley, S., Hubbell, E., Veitch, J., Collins, P.J., Darvishi, K. and Lee, C. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genetics*,40:1253.
 29. Yavaş, G., Koyutürk, M., Özsoyoğlu, M., Gould, M.P. and LaFramboise, T. (2009) An optimization framework for unsupervised identification of rare copy number variation from SNP array data. *Genome Biol.*,10:R119.
 30. Bickhart, D.M., Hou, Y., Schroeder, S.G., Alkan, C., Cardone, M.F., Matukumalli, L.K., Song, J., Schnabel, R.D., Ventura, M., Taylor, J.F. and Garcia, J.F. (2012) Copy number variation of individual cattle genomes using next-generation sequencing. *Genome Res.*, 22:778-790.
 31. Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V. and Abdel-Hamid, H. (2011) A copy number variation morbidity map of developmental delay. *Nat. Genet.*,43:838.
 32. Klambauer, G., Schwarzbauer, K., Mayr, A., Clevert, D.A., Mitterecker, A., Bodenhofer, U. and Hochreiter, S. (2012) cn.MOPS: mixture of Poissons for discovering copy number variations in next-generation sequencing data with a low false discovery rate. *Nucleic Acids Res.*,40:e69-e69.
 33. Fiegler, H., Redon, R., Andrews, D., Scott, C., Andrews, R., Carder, C., Clark, R., Dovey, O., Ellis, P., Feuk, L. and French, L. (2006) Accurate and reliable high-throughput detection of copy number variation in the human genome. *Genome Res.*,16:1566-1574.
 34. Upadhyay, M.R., Chen, W., Lenstra, J.A., Goderie, C.R.J., MacHugh, D.E., Park, S.D.E., Magee, D.A., Matassino, D., Ciani, F., Megens, H.J. and Van Arendonk, J.A.M. (2017) Genetic origin, admixture and population history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity*,118:169-176.
 35. Xu, L., Cole, J.B., Bickhart, D.M., Hou, Y., Song, J., VanRaden, P.M., Sonstegard, T.S., Van Tassell, C.P. and Liu, G.E. (2014) Genome wide CNV analysis reveals additional variants associated with milk production traits in Holsteins. *BMC Genomics*,15:683.
 36. Hou, Y., Bickhart, D.M., Hvinden, M.L., Li, C., Song, J., Boichard, D.A., Fritz, S., Eggen, A., DeNise, S., Wiggans, G.R. and Sonstegard, T.S. (2012) Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. *BMC Genomics*,13:376.
 37. da Silva, V.H., de Almeida, R.L.C., Geistlinger, L., Pértille, F., Giachetto, P.F., Brassalotti, R.A., Morosini, N.S., Zimmer, R. and Coutinho, L.L. (2016) Genome-wide detection of CNVs and their association with meat tenderness in Nelore cattle. *PLoS One*,11: e0157711.
 38. Pezer, Ž., Harr, B., Teschke, M., Babiker, H. and Tautz, D. (2015) Divergence patterns of genic copy number variation in natural populations of the house mouse (*Mus musculus domesticus*) reveal three conserved genes with major population-specific expansions. *Genome Res.*,25: 1114-1124.
 39. Pinto, D., Darvishi, K., Shi, X., Rajan, D., Rigler, D., Fitzgerald, T., Lionel, A.C., Thiruvahindrapuram, B., MacDonald, J.R., Mills, R. and Prasad, A. (2011) Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat. Biotechnol.*,29:512.
 40. Troy, C.S., MacHugh, D.E., Bailey, J.F., Magee, D.A., Loftus, R.T., Cunningham, P., Chamberlain, A.T., Sykes, B.C. and Bradley, D.G. (2001) Genetic evidence for near-eastern origins of European cattle. *Nature*,410:1088-1091.
 41. Caramelli, D. (2006) The origins of domesticated cattle. *Human Evol.*, 21:107-122.
 42. Fadista, J., Thomsen, B., Holm, L.E. and Bendixen, C. (2010) Copy number variation in the bovine genome. *BMC Genomics*, 11:284.
 43. Jiang, L., Jiang, J., Yang, J., Liu, X., Wang, J., Wang H., Ding, X., Liu, J., Zhang, Q. (2013) Genome-wide detection of copy number variations using high-density SNP genotyping platforms in Holsteins. *BMC Genomics*,14:131.
 44. Sasaki, S., Watanabe, T., Nishimura, S. and Sugimoto, Y. (2016) Genome-wide identification of copy number variation using high-density single-nucleotide polymorphism array in Japanese Black cattle. *BMC Genetics*,17:26.
 45. Liu, G.E., Hou, Y., Zhu, B., Cardone, M.F., Jiang, L., Cellamare, A., Mitra, A., Alexander, L.J., Coutinho, L.L., Dell'Aquila, M.E. and Gasbarre, L.C. (2010) Analysis of copy number variations among diverse cattle breeds. *Genome Res.*,20: 693-703.
 46. Stothard, P., Choi, J.W., Basu, U., Sumner-Thomson, J.M., Meng, Y., Liao, X. and Moore, S.S. (2011) Whole genome resequencing of black angus and holstein cattle for SNP and CNV discovery. *BMC Genomics*,12: 559.
 47. Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W. and Lee, C. (2004) Detection of large-scale variation in the human genome. *Nat. Genetics*,36:949.
 48. Ma, L., O'Connell, J.R., VanRaden, P.M., Shen, B., Padhi, A., Sun, C., Bickhart, D.M., Cole, J.B., Null, D.J., Liu, G.E. and Da, Y. (2015) Cattle sex-specific recombination and genetic control from a large pedigree analysis. *PLoS Genetics*,11:e1005387.
 49. Liu, B., Morrison, C.D., Johnson, C.S., Trump, D.L., Qin, M., Conroy, J.C., Wang, J. and Liu, S. (2013) Computational methods for detecting copy number variations in cancer genome using next generation sequencing: principles and challenges. *Oncotarget*,4:1868.
 50. Hutt, K.J., McLaughlin, E.A. and Holland, M.K. (2006) Kit ligand and c-Kit have diverse roles during mammalian oogenesis and folliculogenesis. *Mol. Human Reprod.*,12: 61-69.
 51. Kijas, J.W., Barendse, W., Barris, W., Harrison, B., McCulloch, R., McWilliam, S. and Whan, V. (2011) Analysis of copy number variants in the cattle genome. *Gene*,482:73-77.
 52. Xie, J., Li, R., Li, S., Ran, X., Wang, J., Jiang, J. and Zhao P. (2016) Identification of copy number variations in Xiang and Kele pigs. *PLoS One*,11:e0148565.
 53. Lai, W., Choudhary, V. and Park, P.J. (2008) CGHweb: A tool for comparing DNA copy number segmentations from multiple algorithms. *Bioinformatics*,24:1014-1015.
 54. Wang, X. and Byers, S. (2014) Copy number variation in chickens: A review and future prospects. *Microarrays*,3: 24-38.
 55. Ghosh, S., Qu, Z., Das, P.J., Fang, E., Juras, R., Cothran, E.G., McDonnell, S., Kenney, D.G., Lear, T.L., Adelson, D.L. and Chowdhary, B.P. (2014) Copy number variation in the horse genome. *PLoS Genetics*,10: e1004712.
 56. Wang, X., Li, J. and Leung, F.C. (2002) Partially inverted tandem repeat isolated from pericentric region of chicken chromosome 8. *Chromosome Res.*, 10: 73-82.
 57. Griffin, D.K., Robertson, L.B., Tempest, H.G., Vignal, A., Fillon, V., Crooijmans, R.P., Groenen, M.A., Deryusheva, S., Gaginskaya, E., Carré, W. and Waddington, D. (2008) Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics*,9:168.
 58. Elferink, M.G., Vallée, A.A., Jungerius, A.P., Crooijmans, R.P. and Groenen, M.A. (2008) Partial duplication of the PRLR and SPEF2 genes at the late feathering locus in chicken. *BMC Genomics*, 9: 391.
 59. Zhang, H., Du, Z.Q., Dong, J.Q., Wang, H.X., Shi, H.Y., Wang, N., Wang, S.Z. and Li, H. (2014) Detection of

- genome-wide copy number variations in two chicken lines divergently selected for abdominal fat content. *BMC Genomics*,15: 517.
60. Jia, X., Chen, S., Zhou, H., Li, D., Liu, W. and Yang, N. (2013) Copy number variations identified in the chicken using a 60K SNP bead chip. *Anim. Gene.*,44: 276-284.
 61. Yi, G., Qu, L., Liu, J., Yan, Y., Xu, G. and Yang, N. (2014) Genome-wide patterns of copy number variation in the diversified chicken genomes using next-generation sequencing. *BMC Genomics*,15: 962.
 62. Han, R., Yang, P., Tian, Y., Wang, D., Zhang, Z., Wang, L., Li, Z., Jiang, R. and Kang, X. (2014) Identification and functional characterization of copy number variations in diverse chicken breeds. *BMC Genomics*,15: 934.
 63. Rao, Y.S., Li, J., Zhang, R., Lin, X.R., Xu, J.G., Xie, L., Xu, Z.Q., Wang, L., Gan, J.K., Xie, X.J. and He, J. (2016) Copy number variation identification and analysis of the chicken genome using a 60K SNP bead chip. *Poul. Sci.*,95: 1750-1756.
 64. Gorla, E., Cozzi, M.C., Román-Ponce, S.I., López, F.R., Vega-Murillo, V.E., Cerolini, S., Bagnato, A. and Strillacci, M.G. (2017) Genomic variability in Mexican chicken population using copy number variants. *BMC Genetics*,18:61.
 65. Strillacci, M.G., Cozzi, M.C., Gorla, E., Mosca, F., Schiavini, F., Román-Ponce, S.I., López, F.R., Schiavone, A., Marzoni, M., Cerolini, S. and Bagnato, A. (2017) Genomic and genetic variability of six chicken populations using single nucleotide polymorphism and copy number variants as markers. *Animal*,11: 737-745.
 66. Fan, W.L., Ng, C.S., Chen, C.F., Lu, M.Y.J., Chen, Y.H., Liu, C.J., Wu, S.M., Chen, C.K., Chen, J.J., Mao, C.T. and Lai, Y.T. (2013) Genome-wide patterns of genetic variation in two domestic chickens. *Genome Biol. Evol.*,5: 1376-1392.
 67. Yoshida, T., Furuta, H., Kondo, Y. and Mukoyama, H. (2012) Association of BoLA-DRB3 alleles with mastitis resistance and susceptibility in Japanese Holstein cows. *Anim. Sci. J.*,83: 359-366.
 68. Sassi, N.B., González-Recio, Ó., De Paz-Del Río, R., Rodríguez-Ramilo, S.T. and Fernández, A.I. (2016) Associated effects of copy number variants on economically important traits in Spanish Holstein dairy cattle. *J. Dairy Sci.*,99: 6371-6380.
 69. Reyer, H., Hawken, R., Murani, E., Ponsuksili, S. and Wimmers, K. (2015) The genetics of feed conversion efficiency traits in a commercial broiler line. *Sci. Reports*,5: 16387.
 70. Kneitz, B., Cohen, P.E., Avdievich, E., Zhu, L., Kane, M.F., Hou, H., Kolodner, R.D., Kucherlapati, R., Pollard, J.W. and Edlmann, W. (2000) MutS homolog 4 localization to meiotic chromosomes is required for chromosome pairing during meiosis in male and female mice. *GenesDev.*,14: 1085-1097.
 71. Kadri, N.K., Harland, C., Faux, P., Cambisano, N., Karim, L., Coppeters, W., Fritz, S., Mullaart, E., Baurain, D., Boichard, D. and Spelman, R. (2016) Coding and noncoding variants in HFM1, MLH3, MSH4, MSH5, RNF212, and RNF212B affect recombination rate in cattle. *Genome Res.*,26: 1323-1332.
 72. Brenig, B., Beck, J., Floren, C., Bornemann-Kolatzki, K., Wiedemann, I., Hennecke, S., Swalve, H. and Schütz, E. (2013) Molecular genetics of coat colour variations in white galloway and white park cattle. *Anim. Genetics*,44: 450-453.
 73. Norris, B.J. and Whan, V.A. (2008) A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Res.*,18: 1282-1293.
 74. Hillbertz, N.H.S., Isaksson, M., Karlsson, E.K., Hellmen, E., Pielberg, G.R., Savolainen, P., Wade, C.M., Von Euler, H., Gustafson, U., Hedhammar, Å. and Nilsson, M. (2007) Duplication of FGF3, FGF4, FGF19 and ORAOV1 causes hair ridge and predisposition to dermoid sinus in ridgeback dogs. *Nat. Genetics*,39:1318.
 75. Pielberg, G., Olsson, C., Syvänen, A.C. and Andersson, L. (2002) Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics*,160: 305-311.
 76. Liu, J., Zhang, L., Xu, L., Ren, H., Lu, J., Zhang, X., Zhang, S., Zhou, X., Wei, C., Zhao, F. and Du, L. (2013) Analysis of copy number variations in the sheep genome using 50K SNP bead chip array. *BMC Genomics*,14:229.
 77. Doan, R., Cohen, N., Harrington, J., Veazey, K., Juras, R., Cothran, G. and McCue, M.E. (2013) Erratum: Identification of copy number variants in horses (genome research (2012) 22 (899-907). *Genome Res.*,23:1362.
 78. Shull, A.Y., Clendenning, M.L., Ghoshal-Gupta, S., Farrell, C.L., Vangapandu, H.V., Dudas, L., Wilkerson, B.J. and Buckhaults, P.J. (2013) Somatic mutations, allele loss, and DNA methylation of the cub and sushi multiple domains 1 (CSMD1) gene reveals association with early age of diagnosis in colorectal cancer patients. *PLoSOne*,8:e58731.
 79. Wright, D., Boije, H., Meadows, J.R., Bed'Hom, B., Gourichon, D., Vieaud, A., Tixier-Boichard, M., Rubin, C.J., Imstrand, F., Hallböök, F. and Andersson, L. (2009) Copy number variation in intron 1 of SOX5 causes the pea-comb phenotype in chickens. *PLoSGenetics*,5:e1000512.
 80. Luo, C., Shen, X., Rao, Y., Xu, H., Tang, J., Sun, L., Nie, Q. and Zhang, X. (2012) Differences of Z chromosome and genomic expression between early-and late-feathering chickens. *Mol. Biol. Rep.*,39: 6283-6288.
 81. Marenne, G., Real, F.X., Rothman, N., Rodríguez-Santiago, B., Pérez-Jurado, L., Kogevinas, M., García-Closas, M., Silverman, D.T., Chanock, S.J., Génin, E. and Malats, N. (2012) Genome-wide CNV analysis replicates the association between GSTM1 deletion and bladder cancer: A support for using continuous measurement from SNP-array data. *BMC Genomics*,13: 326.
 82. Lin, S., Lin, X., Zhang, Z., Jiang, M., Rao, Y., Nie, Q. and Zhang, X. (2018) Copy number variation in SOX6 contributes to chicken muscle development. *Genes*,9: 42.
