527. Copy number variation analysis in Nguni and Bonsmara crossbred cattle

B. Bhika Kooverjee[1*], P. Soma[1], F.W.C Neser[2], M.A. van der Nest[3] and M.M Scholtz[1,2]

[1]*Agricultural Research Council – Animal Production, Irene, South Africa; [2]Department of Animal, Wildlife and Grassland Sciences, University of the Free State, Bloemfontein, South Africa; [3]Agricultural Research Council – Biotechnology Platform, Onderstepoort, South Africa; kooverjeeb@arc.agric.za*

Abstract

Crossbreeding is a useful tool in livestock systems in mitigating the effects of climate change. The aim of this study was to perform a copy number variation analysis in indigenous Nguni and Bonsmara crossbred cattle. Four crossbred individuals were sequenced at 10× coverage. Following quality control, sequence data were interrogated for CNVs using the panelcn.MOPS tool. CNVs detected harboured genes related to important biological processes. Genes relating to fertility (MTERF2), Heat tolerance (LOC109577026), Stress response (HTRA2) and carbohydrate metabolism (LOC109573033) was found in the Bonsmara-sired crossbreds. While genes relating to embryogenesis (*ARMCX1*) and lipid metabolism (*ENHO*) were found in the Nguni-sired crossbreds. Results of this study suggested these crossbred cattle display potential for inclusion in various crossbreeding programs as they harbour genetic resources that may enhance the presence of significant traits such as meat quality, heat tolerance and adaptation, while simultaneously limiting the adverse effect on the environment.

Introduction

Copy number variations (CNVs) is defined as the deletion or repetition of a genome copy number (Goshu *et al.,* 2018), and may have a functional and evolutionary impact (Keel *et al.*, 2016). Changes in gene copy number also influence gene dosage, unjustified gene fusion, gene interruption, and position effects (Liu *et al.,* 2010). Several studies attempted to increase the understanding why crossbred cattle differ in their response to environmental changes such as heat stress. Since CNVs are an important source of genetic variability associated with phenotypic variation, the aim of this study was to examine genome-wide copy number differences between F1 progeny sired from Nguni and Bonsmara dams and with F1 progeny sired from Bonsmara and Nguni dams, which are both indigenous breeds. For this purpose, the genomes of F1 progeny from Bonsmara × Nguni reciprocal crosses were sequenced and panelcn.MOPS, a read depth approach based on next-generation sequencing, was used to identify genes that differ in copy number variation between crossbred progeny from Bonsmara and Nguni sires.

Materials & methods

Ethical approval and animal selection. Ethical approval was obtained from the Agricultural Research Council, Animal Production Ethics Committee (APIEC21/06). Hair samples from two Nguni-sired (N×B) and two Bonsmara-sired (B×N) crossbred animals were collected from a crossbreeding project in the Northern Cape province.

DNA extraction and sequencing. DNA was extracted from hair samples using the Qiagen DNeasy Blood and tissue kit as per manufacturer's protocol. For whole genome sequencing of the four crossbreds, libraries were prepared according to the Illumina TruSeq Nano DNA Library Prep Kit sample preparation guide. Sequence reads were filtered for base quality and adapter trimming using Trimmomatic v0.36. After trimming, only pairs of DNA sequences with reads exceeding 36 bp were retained for analysis. BWA-MEM v0.7.17 software (Li and Durbin, 2009) was used to align sequences with the reference genome (*Bos_indicus*_1.0, Canavez *et al.,* 2012). The *Bos indicus* reference (GCF_000247795.1_*Bos_indicus*_1.0) was

chosen due to African Sanga cattle owning unique indicine ancestry that is highly distinguished from that found in modern cattle populations such as the Brahman and Nellore (van Marle-Koster *et al.,* 2021). The sequencing depth determined with Samtools depth command for the crossbred ranged from 14.7 to 19.8.

CNV analysis and gene annotation. The GUI version of PanelcnMOPS program, CNV Detective (https://www.bioinf.jku.at/software/panelcnmops/) was used to identify copy number variations in the crossbreds. The two Nguni-sired crossbreds were used as the control group while the two Bonsmara-sired crossbreds were used to represent the test group. Through the advanced option setting, the duplication threshold was set at 1.46, the deletion threshold was set at 0.57, along with minimum median RC/ROI ratio = 30 and the sex = male. Quality control (QC) involves the minimum median RC/ROI ratio, where, samples with a median RC across all ROIs that is lower than 0.55 times the median of all samples fail the first step of the sample QC. For each ROI, the ratio between the normalized RCs of each sample and the median across all remaining samples is calculated, when samples show a high variation in RC ratios it fails the second QC step. CNVs labelled with 'LowQual' were regarded as insignificant and removed. The gene content of significant CNV regions were assessed based on the gene annotation of the UMD3.1 genome assembly using Ensembl (Ensembl Genes104). The gene ontology terms (for biological processes, cellular component and molecular function) identified using PANTHER with Bonferoni correction with significance of $P<0.05$.

Results

The results are indicated in Figure 1 and Table 1. Figure 1 displays the contribution of each copy number class to the total number of CNV calls per chromosome. In Table 1 the genes associated with different biological traits are listed.
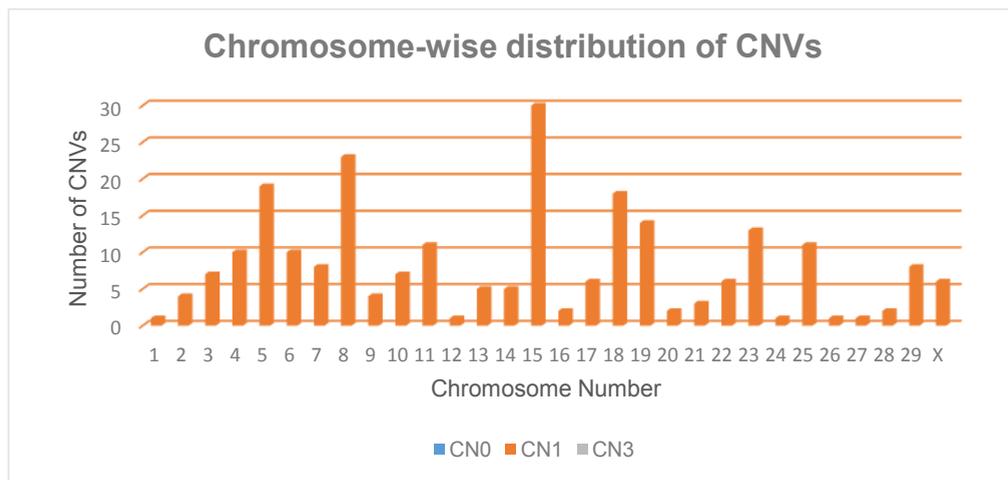

Figure 1. Bar plot displaying the contribution of each copy number class to the total number of CNV calls per chromosome. CN0 (blue) and CN1 (orange) represents the double and single copy number deletion events, respectively. CN3 (grey) represents the single duplication copy number events.

Table 1. CNV genes identified in the Nguni-sired (N×B) and Bonsmara-sired (B×N) crossbreds detected by panelcn. MOPS.

Breed	Trait	Chr	Start	End	Gene
N×B	Embryogenesis	X	30880957	30882318	ARMCX1
N×B	Lipid metabolism	8	80116078	80116308	ENHO
B×N	Female fertility	5	75275935	75277092	MTERF2
B×N	Heat tolerance	23	27266576	27268501	LOC109577026
B×N	Stress response	11	10412204	10415095	HTRA2
B×N	Methane Production	20	41718533	41742666	RAI14
B×N	Taste	4	1,09E+08	1,09E+08	TASR38
B×N	Visual Perception	6	1,21E+08	1,21E+08	HMX1
B×N	Carbohydrate metabolism	18	2243625	2244812	LOC109573033

Discussion

In this study, two hundred and thirty-nine CNVs regions involving 2.29 Mbp (~0.092%) of the reference genome were identified in four animals from two crossbreds. CNVs ranged from 230 bp to 153,354 bp in size with an average length of 7,718 bp. Among these, 185 CNVs were deletions (losses) and 54 CNV duplications (gains). CNV analysis in numerous studies showed great variation in the number of CNV regions reported in cattle (Keel, *et al.,* 2016). A gene in B×N identified as LOC109577026 located on Chromosome 23 was shown to be involved with cellular response to heat stress (Sakatani *et al.*, 2012). Another gene identified that plays a role in adaptation, includes HtrA serine peptidase 2 (HTRA2) and also involved in cellular response to heat (Skorko-Glonek *et al.*, 1999) (Table 1). Reproduction is one of the key factors driving the economic efficacy through sustainable meat production in the beef industry. In this study, genes relating to animal development and fertility were detected. This include mitochondrial transcription termination factor 2 (MTERF2) located on Chromosome 5 in the B×N crossbreds and was previously identified as a candidate gene associated with reproductive traits in Nellore cattle (Oliveria Junior *et al.,* 2019). The armadillo repeat containing X-linked 1 gene located on the X Chromosome of the N×B crossbreds plays a role in embryogenesis during the development of embryos in humans (Kusuma *et al.,* 2010), while its function in cattle is still unclear. Additionally, genes relating to metabolism was detected in the both crossbreds. The carbohydrate sulfotransferase 6 (LOC109573033) located on Chromosome 18 in the B×N crosses is involved in carbohydrate metabolism and inflammatory response to pathogens (Tetas *et al.*, 2016)). While in the N×B crossbreds, the ENHO gene also known as adropin, located on Chromosome 8, plays a role in lipid metabolism and regulation of glucose homeostasis (Jasaszwili *et al.*, 2020). This is further supported by the GO terms cellular macromolecule biosynthetic process (GO: 0009059) and regulation of biological process (GO: 0050789). Overall, the functional genes identified were found to play a role in reproduction and metabolism in both crossbreds. It is therefore possible that these genes will play a role in adaptation to heat and other environmental stresses in both the Nguni-sired and Bonsmara-sired crossbreds.

Acknowledgements

References

Canavez, F.C., Luche, D.D., Stothard, P., Leite, K.R., Sousa-Canavez, J.M., *et al.* (2012) Journal of Heredity 103(3):342-348. https://doi.org/10.1093/jhered/esr153

Cho, S.J., Kang, M.J., Homer, R.J., Kang, H.R., Zhang, X., *et al.* (2006) Journal of Biological Chemistry, 281(12):8161-8168. https://doi.org/10.1074/jbc.m506770200

Goshu, H.A., Chu, M. and Yan, P. (2018) African Journal of Biotechnology 17(43):1313-1323. https://doi.org/10.5897/AJB2018.16464

Jasaszwili, M., Billert, M., Strowski, M.Z., Nowak, K.W. and Skrzypski, M., (2020) Molecules, 25(3):549 https://doi.org/10.3390/molecules25030549

Keel, B.N., Lindholm-Perry, A.K. and Snelling, W.M. (2016) Frontiers in genetics 7:207. https://doi.org/10.3389/fgene.2016.00207

Kusama, Y., Takayanagi, S., Tategu, M. and Yoshida, K., (2010) Experimental and Therapeutic Medicine *1*(2):395-399. https://doi.org/10.3892/etm_00000062

Liu, M., Fang, L., Liu, S., Pan, M.G., Seroussi, E., *et al.* (2019) BMC genomics 20(1):181. https://doi.org/10.1186/s12864-019-5552-1

Van Marle-Köster, E., Visser, C., Sealy, J., & Frantz, L., (2021) Sustainability, 13(8): 4388. http://dx.doi.org/10.3390/su13084388

Oliveira Júnior, G.A., Santos, D.J.A., Cesar, A.S.M. *et al.* (2019) J Animal Sci Biotechnol 10, 97. https://doi.org/10.1186/s40104-019-0403-0

Sakatani, M., Alvarez, N.V., Takahashi, M. and Hansen, P.J., (2012).. Journal of Dairy Science, 95(6):3080-3091. https://doi.org/10.3168/jds.2011-4986

Skorko-Glonek, J., Zurawa, D., Kuczwara, E., Wozniak, M. Wypych, Z. *et al.* (1999) Molecular and General Genetics 262(2):342-350. https://doi.org/10.1007/s004380051092

Stothard, P., Choi, J.W., Basu, U., Sumner-Thomson, J.M., Meng, Y., *et al.* (2011) BMC genomics 12(1):1-14. https://doi.org/10.1186/1471-2164-12-559

Tetas Pont, R., Downs, L., Pettitt, L., Busse, C. and Mellersh, C.S., 2016. Veterinary ophthalmology, 19(6):488-492. https://doi.org/10.1111/vop.12332

Zhan, B., Fadista, J., Thomsen, B., Hedegaard, J., Panitz, F. *et al.* (2011) BMC genomics 12(1):1-20. https://doi.org/10.1186/1471-2164-12-557