

# ASSOCIATION OF COPY NUMBER VARIANTS WITH COAT COLOUR IN NGUNI CATTLE INVESTIGATED USING BOVINEHD SNP AND BIONANO OPTICAL MAPPING DATA

**Dlamini N.M<sup>1,2</sup>, Dzomba E.F<sup>2</sup>, Magawana M<sup>3</sup>, Ngcamu S<sup>3</sup> & Muchadeyi F.C<sup>1\*</sup>**

<sup>1</sup> Biotechnology Platform, Agricultural Research Council – Onderstepoort Veterinary Institute, Private Bag X5, Onderstepoort, 0110, South Africa

<sup>2</sup> Discipline of Genetics, School of Life Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville, 3209, South Africa; School of Agriculture, Earth and Environmental Sciences, Private Bag X01, Scottsville, 3209, South Africa

<sup>3</sup> KZN Department of Agriculture & Rural Development, Private Bag X9059, Pietermaritzburg 3200, South Africa

Corresponding Author: \*MuchadeyiF@arc.agric.za

## INTRODUCTION

- ❖ A Sanga-type breed with *B. taurus* and *B. indicus* ancestry, Nguni cattle have shown resistance to diseases and harsh climatic conditions in Africa.
- ❖ Nguni cattle have diverse colour patterns, and each animal has its unique pattern.
- ❖ Copy number variation (CNV) is described as segments of the DNA that are copy number variable when compared with a reference genome.
- ❖ It is associated with gene expression and may present a major genetic component of phenotypic diversity.

**In this study we characterized CNVs in South African Nguni cattle that were phenotyped for base coat colour, forehead stripe and colour-sidedness using Illumina's BovineHD genotype data and Bionano optical mapping.**

## MATERIALS & METHODS

### SNP Genotype Data

**SNP data generation:** Samples were collected from Bartlow Combine (n = 99) and Kokstad (n = 33) research stations.

### SNP quality assessment

Call rate of less than 90%; MAF (< 2%); Missing genotypes and genotyping failure of >10%; HWE ( $P < 0.001$ ).

**PCA-based clustering:** Rstudio software.

**Identification of Nguni Cattle CNVs:** PLINK v1.07 software (Purcell et al., 2007).

**Gene annotation:** Ensembl genome browser.



### Bionano Optical Mapping

**Sample collection:** Samples were collected from two farms, Agricultural Research Council research herd (n=3) and Mevamhlophe Farm (n=5).

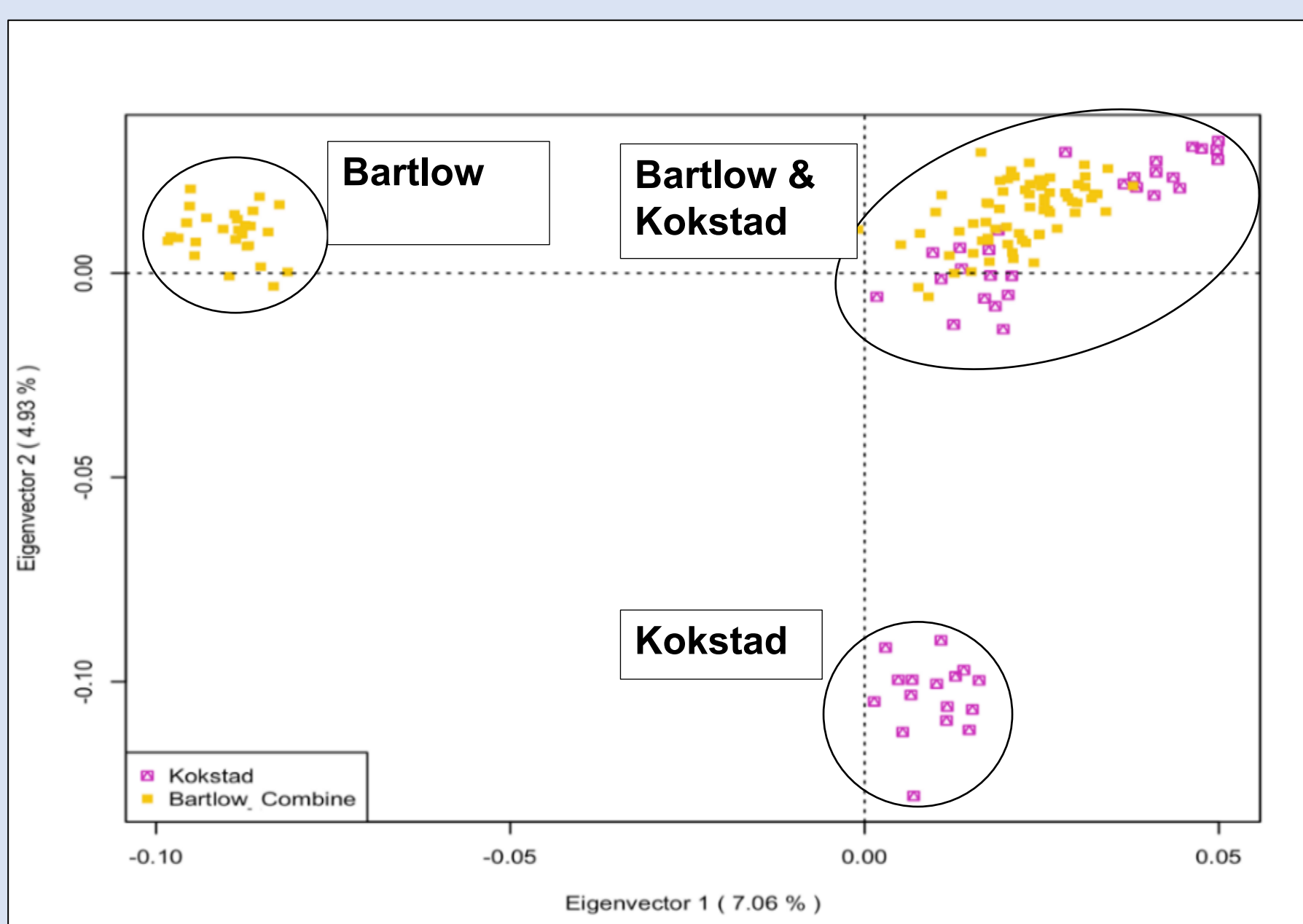
**gDNA isolation:** Bionano Prep SP Frozen Human Blood DNA Isolation Protocol.

**Data collection:** Bionano Saphyr Chip (Bionano Genome).

**De novo Assembly and Structural Variant Calling:** Bionano Solve program.

**Gene annotation:** Ensembl genome browser.

## RESULTS AND DISCUSSION



**Figure 1.** PCA-based clustering of Bartlow Combine and Kokstad Nguni cattle populations.

**Table 1.** Summary statistics of number of CNVs (CNVs) and average length per CNV state per genetic cluster.

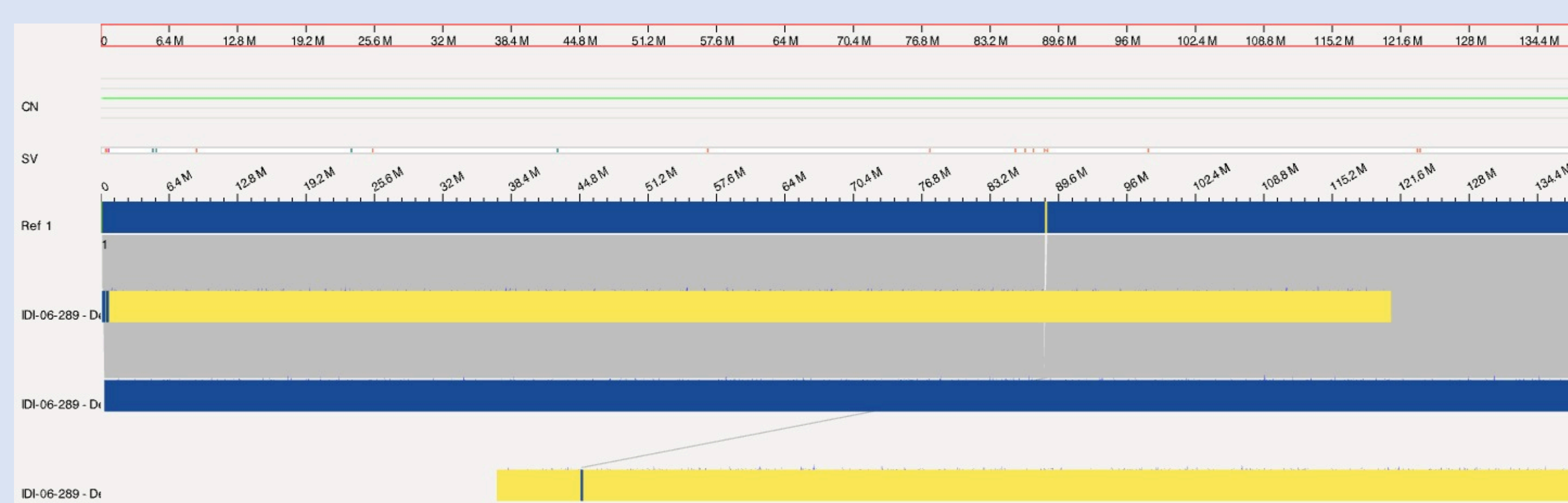
	Number of CNVs			Average length (Mb)		
State	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3
Gain	518	3 965	526	0.08±0.12	0.07±0.10	0.08±0.11
Loss	1 224	4 559	699	0.07±0.09	0.08±0.12	0.08±0.12
Mixed	-	-	-	-	-	-
Total	1 742	8 524	1 225	0.08±0.11	0.08±0.11	0.08±0.12

- ❖ Three clear genetic clusters were observed from the PCA (Figure 1).
- ❖ Across genetic clusters there were more loss CNVs than gain CNVs (Table 1, SNP data).
- ❖ A greater number of gains than loss events were recorded in all animals, based on the Bionano technology (Table 2).
- ❖ The highest number of gains were observed in sample Nguni\_33 (2013) and the highest number of loss events were found in animal Nguni\_54 (2093) (Table 2).

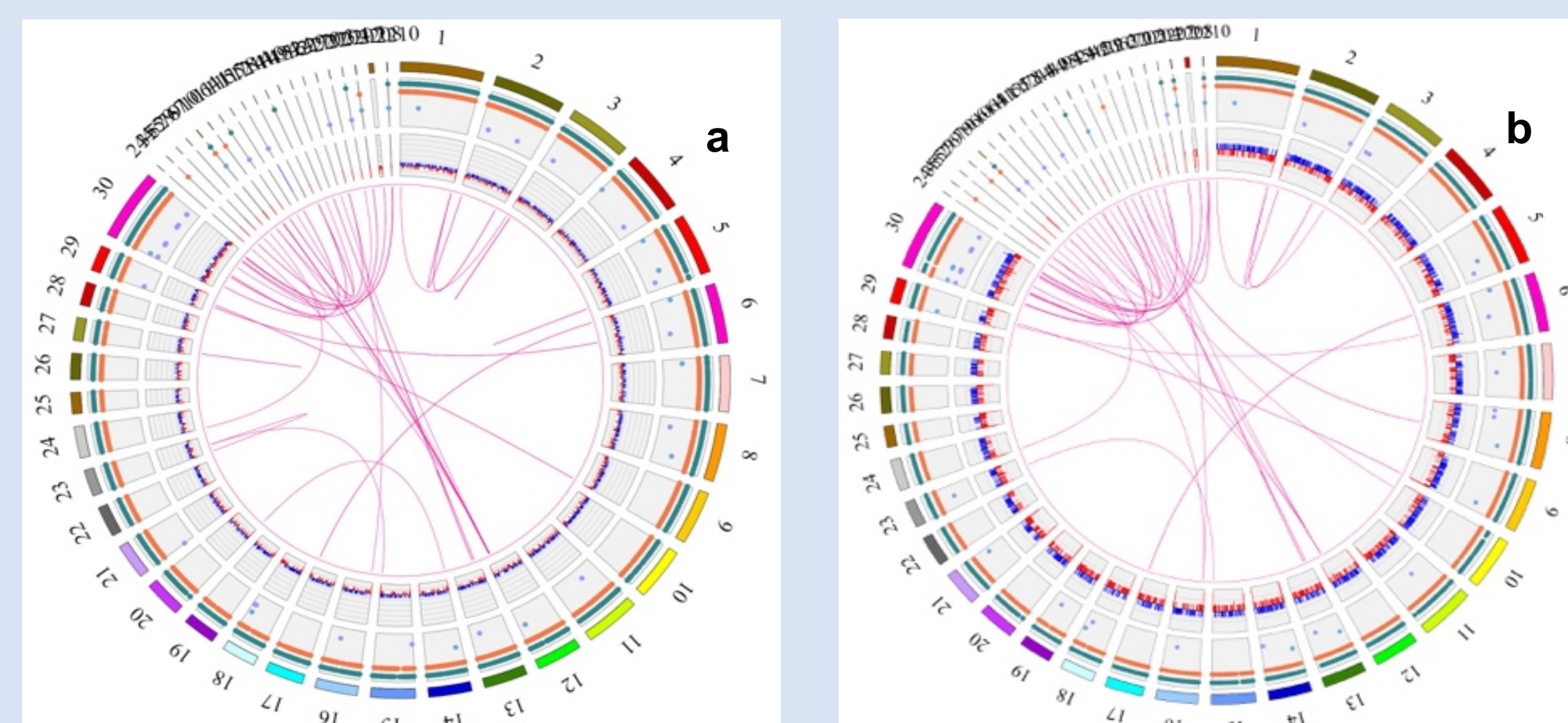
**Table 2.** Number of CNV calls per sample based on the Bionano technology.

Sample ID	Gain	Loss	#Total CNVs	Sum CNVs (Mb)	Coverage (%)
Nguni_33	2 013	1 783	3 796	660.70	26.54
Nguni_70	1 732	1 235	2 967	482.92	19.40
Nguni_16	1 230	805	2 034	296.05	11.89
Nguni_54	2 156	2 093	4 249	1 124.71	45.18
Nguni_109	1 789	1 728	3 517	808.25	32.47
IDI06-289	1 442	1 160	2 602	624.57	25.09
IDI20-06	1 987	1 980	3 967	1 167.46	46.90
IDI20-11	1 831	1 351	3 182	643.84	25.86
Total	14 180	12 135	26 314	726.06	29.17

- ❖ Chromosome rearrangements are depicted by the lines (coloured magenta) connecting one chromosome to another (Figure 2, Bionano data).
- ❖ After de novo assembly, Bionano maps were compared to the cattle reference genome (Figure 2).
- ❖ About 106 of the reported genes were associated with coat colour/pigmentation and included *PRKCA*, *CALML5*, *MAPK1*, *WNT5B*, *POMC*, *WNT3A*, *PLCB4* (SNP genotype data).
- ❖ A total of 35 569 candidate genes overlapped with the CNVRs including well-known coat colour genes of *KIT*, *KITLG*, *ASIP*, *TYR*, *TYRP1*, *WNT3*, *MAPK1* and *MC1R* (Bionano data).



**Figure 3.** Nguni sample IDI06-289 genome maps compared to the cattle reference genome.



**Figure 2.** Circos virtualization of detected structural variants in sample a) Nguni\_16 and b) IDI06-289 using Bionano data.



## CONCLUSION

- ❖ This study provides insights into the coat colour patterns observed in Nguni cattle and demonstrates the utility of Bionano optical mapping technology in coat colour genetics studies.
- ❖ It will be important to refine and validate these results by investigating more Nguni cattle from different herds.