Genome wide Copy Number Variation (CNV) detection in Cinisara cattle breed

Journal:	*Italian Journal of Animal Science*
Manuscript ID	TJAS-2017-0055.R1
Manuscript Type:	Abstract Submission
Date Submitted by the Author:	07-Feb-2017
Complete List of Authors:	Di Gerlando, Rosalia; Universita degli Studi di Palermo, Scienze Agrarie e Forestali Sardina, Maria; Universita degli Studi di Palermo, Scienze Agrarie e Forestali Tolone, Marco; Universita degli Studi di Palermo, Sutera, Anna Maria; Universita degli Studi di Palermo, Scienze Agrarie e Forestali MASTRANGELO, Salvatore; Universita degli Studi di Palermo, Scienze Agrarie e Forestali Portolano, Baldassare; Universita degli Studi di Palermo, Scienze Agrarie e Forestali
Abstract:	Copy Number Variations (CNVs) are classes of polymorphic genomic regions including deletions, duplications and insertions of DNA fragments from at least 0.5 kb up to several Mb. CNV represents an important source of genetic variability that provides genomics structural information complementary to the single nucleotide polymorphism (SNP) data. Some CNVs have been shown to be important in both normal phenotypic variability and disease susceptibility in livestock. Several approaches to identify CNVs including FISH, aCGH, SNP array or NGS, were proposed and among these SNP genotyping is relatively low cost, high-throughput and high coverage method. The aim of this study was to identify the CNVs in 71

	animals of Cinisara breed using Illumina BovineSNP50 BeadChip v2. PennCNV software, which incorporates Log R ratio and B allele frequency at each SNP marker, was used to identify CNVs. Seven animals showed not shared CNVs, as well as autosomes 19, 21, 22. Chromosome 25 presented no CNVs at all. A final number of 322 CNVs were detected. The average number of CNVs was 4.5 per individual, with an average length and median size of 143.04 kb and 122.14 kb, respectively. All CNVs were grouped in CNV regions (CNVRs) and a total of 107 CNVRs, ranged from 50 to ~500 kb, were detected, which covered 4.90 Mb of polymorphic sequence and corresponded to 0.18% of the total genome length. In particular, we found 81 CNVRs with only gain (duplication), 22 with only loss (deletion), and four CNVRs with both. Furthermore, 8 CNVRs with >1%, 77 with >2.5%, and 22 with >5% frequency, were found. CNVRs having the highest frequency were located on Chr3:120501439-120647330 and Chr23:34673581-35007295, whereas the greatest number of genes was mapped in only one CNVR located on Chr 17:74123863-74393620. A total of 241 genes were included in the identified CNVRs. According to KEGG and DAVID database, most of the genes were involved in multiple signaling and signal transduction pathways in a wide variety of cellular and biochemical processes, such as immune response, adaptability, and olfactory receptors pathway. Further studies, using different algorithms and validating the CNVs discovered, will be conducted to corroborate these preliminary results on the CNVRs detected. These results will be used for the investigation of genomic changes and features of interest in the Cinisara breed, such as for association with functional or production traits and for biodiversity studies.

Genome wide Copy Number Variation (CNV) detection in Cinisara cattle breed

Rosalia Di Gerlando, Maria Teresa Sardina, Marco Tolone, Anna Maria Sutera, Salvatore Mastrangelo, Baldassare Portolano

Dipartimento Scienze Agrarie e Forestali, Università degli Studi di Palermo, Italy

Corresponding author: rosalia.digerlando@unipa.it

Copy Number Variations (CNVs) are classes of polymorphic genomic regions including deletions, duplications and insertions of DNA fragments from at least 0.5 kb up to several Mb. CNV represents an important source of genetic variability that provides genomics structural information complementary to the single nucleotide polymorphism (SNP) data. Some CNVs have been shown to be important in both normal phenotypic variability and disease susceptibility in livestock. Several approaches to identify CNVs including FISH, aCGH, SNP array or NGS, were proposed and among these SNP genotyping is relatively low cost, high-throughput and high coverage method. The aim of this study was to identify the CNVs in 71 animals of Cinisara breed using Illumina BovineSNP50 BeadChip v2. PennCNV software, which incorporates Log R ratio and B allele frequency at each SNP marker, was used to identify CNVs. Seven animals showed not shared CNVs, as well as autosomes 19, 21, 22. Chromosome 25 presented no CNVs at all. A final number of 322 CNVs were detected. The average number of CNVs was 4.5 per individual, with an average length and median size of 143.04 kb and 122.14 kb, respectively. All CNVs were grouped in CNV regions (CNVRs) and a total of 107 CNVRs, ranged from 50 to ~500 kb, were detected, which covered 4.90 Mb of polymorphic sequence and corresponded to 0.18% of the total genome length. In particular, we found 81 CNVRs with only gain (duplication), 22 with only loss (deletion), and four CNVRs with both. Furthermore, 8 CNVRs with >1%, 77 with >2.5%, and 22 with >5% frequency, were found. CNVRs having the highest frequency were located on Chr3:120501439-120647330 and Chr23:34673581-35007295, whereas the greatest number of genes was mapped in only one CNVR located on Chr 17:74123863-74393620. A total of 241 genes were included in the identified CNVRs. According to KEGG and DAVID database, most of the genes were involved in multiple signaling and signal transduction pathways in a wide variety of cellular and biochemical processes, such as immune response, adaptability, and olfactory receptors pathway. Further studies, using different algorithms and validating the CNVs discovered, will be conducted to corroborate these preliminary results on the CNVRs detected. These results will be used for the investigation of genomic changes and features of interest in the Cinisara breed, such as for association with functional or production traits and for biodiversity studies.