

Linkage disequilibrium analysis in Canchim beef cattle

Marcos E. Buzanskas*¹, Daniela A. Grossi², Flavio S. Schenkel², Ricardo Ventura², Luciana C. A. Regitano³, Maurício M. Alencar³, Danísio P. Munari¹

* Doctorate Student, Programa de Pós-Graduação em Genética e Melhoramento Animal, Universidade Estadual Paulista, Via de Acesso Prof. Paulo Donato Castellane s/n, CEP: 14884-900 - Jaboticabal, SP, Brazil; ¹Universidade Estadual Paulista – FCAV, Departamento de Ciências Exatas, Jaboticabal, SP; ²University of Guelph, Guelph, ON, Canada; ³ Embrapa Pecuária Sudeste, São Carlos, SP, Brazil

* marcosbuz@yahoo.com.br

Advances in technology and new methodologies applied to genomic evaluation are of great importance for improvements in the cattle industry. To begin a genomic evaluation, many factors should be taken into account; one of which is to evaluate the extent of linkage disequilibrium (LD) between genetic markers (e.g. single nucleotide polymorphism - SNP). LD is a non-random association of alleles at two or more loci in the same chromosome in which a deviation between the observed and the expected frequency exists. It results from migration, mutation, selection, and drift, or can also be created in livestock populations by crossing lines or breeds. The aim of this study was to investigate the extent of linkage disequilibrium in Canchim breed and MA genetic group using high density SNP panel. Canchim beef cattle are a synthetic breed developed at the facilities of Brazilian Agricultural Research Corporation (Embrapa). In this study we used animals from the Canchim breed (62.50% Charolais and 37.50% Zebu composition) and from the “MA” genetic group (65.60% Charolais and 34.40% Zebu composition). Genotype data from 400 individuals was collected for high density genotyping (BovineHD - Illumina® bead chip). SNP from Canchim (285 animals), MA (114 animals), and Charolais (1 animal), 195 males and 205 females born between 1999 and 2005 were analyzed. Genotype quality control excluded the following: sexual chromosomes, SNP with calling score lower than 0.15, call rate for animals and SNP lower than 0.85, SNP with deviations of the Hardy-Weinberg equilibrium ($P < 10^{-5}$), SNPs with excess of heterozygosity ($< 15\%$), and minor allele frequency lower than 5%. For LD analysis only SNPs within a maximum range of 0.5 Mb (mega base) were used. This limit was considered because there are too many comparisons to be made for the high density panel. The measures of linkage disequilibrium D and r^2 were calculated using $r^2 = D^2 / f(A)f(a)f(B)f(b)$, where $D = f(AB) - f(A)f(B)$, $f(AB)$, $f(A)$, $f(a)$, $f(B)$, and $f(b)$ are the observed frequencies of AB and alleles A , a , B , and b , respectively. Linkage phase between SNP was also verified by doing separate LD analyses of Canchim and MA animals. To determine the consistency of linkage phase it was used the sign of D and the square root of r^2 . Pearson correlations were obtained between signed $\sqrt{r^2}$ for the same SNPs and genomic distances and values above 0.80 were considered as consistent phase. Linkage phase was consistent between Canchim and MA, indicating that a multi-breed evaluation could be carried out, and findings in one of the genetic groups could be used in another. Thus, LD analysis can consider all animals of those two genetic groups as one population. Genomic distance from 0 to 0.0025 Mb (megabase), 0.0025 to 0.0050 Mb, and 0.0050 to 0.0075 Mb had average r^2 equal to 0.38, 0.34, and 0.31, respectively. Other genomic distances had average r^2 lower than 0.30. The threshold of 0.30 was determined to evaluate the extent of LD along the genome that could be useful for QTL (quantitative trait loci) mapping and genomic selection. In conclusion, useful LD was verified at short distances in this breed, showing that analyses such as genome wide association or selection signatures could be carried out in future studies. Furthermore, highly consistent linkage phase would facilitate genomic studies in this breed by expanding the source of information of animals.

Keywords: alleles, autosome, genomic selection, SNP, synthetic breed

Acknowledgments: CNPq grant (n. 142053/2010-4) and CAPES grant (n. 5285-11-9).