Selection Signatures in Canchim Beef Cattle

I. Urbinati[1]**, M. E. Buzanskas**[1]**, T. C. S. Chud**[1]**, F. B. Mokry**[2]**, L. C. A. Regitano**[3]**, R. H. Higa**[3] **and D. P. Munari**[1]**.**
[1] UNESP - Univ Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, SP, Brazil, [2] Univ Federal de São Carlos, SP, Brazil, [3] Brazilian Agricultural Research Corporation, Brazil.

ABSTRACT: Selection signature (SS) was assessed in this study by means of the integrated haplotype score (iHS) method, which determines the decay of homozygosity in the surroundings of a core single nucleotide polymorphism (SNP) marker. Canchim breed animals were genotyped using the Illumina BovineHD BeadChip; which has almost 800 thousand SNP markers. Genotype quality control (QC) was applied to exclude SNP with genotype calling score lower than 0.20; SNP with minor allele frequency lower than 0.01; and call rate for SNP and samples which were lower than 0.95 and 0.90, respectively. Only autosomal SNPs with known genome position were used. After the QC, 687,655 SNPs and 396 samples remained for SS analysis. Signals of SS were detected on chromosomes 5, 6, 8, and 14, indicating that these regions are conserved through recent generations.
Keywords: EHH, his, rehh package

Introduction

High-density genotyping is based on a high coverage of genetic markers (i.e. single nucleotide polymorphism – SNP) across the genome. Nowadays, the usage of SNP information has become very important for livestock, considering its applicability in animal breeding (e.g. genomic selection). In this context, selection signatures (SS) provide useful information about specific genomic regions that are conserved across generations. The SS regions can be detected by means of the extended haplotype homozygosity (EHH; Sabeti et al. (2002)) and integrated haplotype score (iHS; Voight et al. (2006)) methods.

Identification of SS in a population under selection allows to evaluate the relationship of SNPs/genes, in which a signal of SS is detected, with traits of interest. The knowledge of genomic regions with signals of SS could aid for future studies of genome-wide association and selection. Thus, the aim of this study was to analyze high-density genomic data on Canchim animals under selection in order to identify SS regions aiming at future applications in animal breeding.

Materials and Methods

A total of 400 animals (285 Canchim, 114 MA genetic group, and one Charolais), born between 1999 and 2005, and genotyped with the Illumina BovineHD panel (786,799 SNPs) were used in this study. Canchim beef cattle breed is a synthetic breed developed in Brazil by the Brazilian Agricultural Research Corporation (Embrapa). Canchim animals are composed by 62.5% - 37.5% Charolais - Zebu proportion; while MA genetic group are 65.6% - 34.4% Charolais - Zebu, approximately.

Genotype quality control (QC) was conducted using the snpStats package (Clayton (2012)), available in the R software (R Core Team (2013)). The QC excluded SNPs with calling score lower than 0.20, SNP with minor allele frequency (MAF) lower than 0.01; and call rate for SNP and samples (animals) lower than 0.95 and 0.90, respectively. Final data set comprised 687,655 autosomal SNPs and 396 animals.

The BEAGLE software (Browning and Browning (2007)) was used to infer the linkage phase and for haplotype building. The integrated haplotype score (iHS; Voight et al. (2006)) method was used to identify the SS regions. This method is available in the rehh package (Gautier and Vitalis (2012)), in R software (R Core Team (2013)). The iHS calculation utilizes integrals of the EHH statistics for both ancestral and derived marker alleles, through surrounding regions of a core SNP. Ancestral allele information was obtained from Utsonomyia et al. (2013).

Results and Discussion

Figure 1 depicts the suggestive regions of SS for each chromosome. Negative iHS values indicate long haplotypes carrying derived alleles, positive iHS values represent long haplotypes carrying ancestral alleles, and small or close to zero iHS values indicate similar rates of EHH decay for ancestral and derived marker alleles. Thus, extreme iHS values are indicative of SS (Voight et al. (2006)). Signals of SS were observed on chromosomes 5, 6, 8, and 14. Peaks of high iHS represent regions of high level of genomic conservation through generations under selection.

By surveying the literature, we found that Gasparin et al. (2007) and Machado et al. (2003) reported quantitative trait loci (QTL) on chromosome 5 associated with tick resistance and birth weight in beef cattle, respectively. Genome-wide association was observed for long-yearling scrotal circumference adjusted to 420 days of age on chromosomes 5 and 14 in the Canchim breed (Buzanskas (2013)). Choi et al. (2006) reported associations with carcass traits on chromosome 14. According to Wibowo et al. (2008), chromosome 14 has been widely explored to find genes/regions associated with traits of economic importance in bovine.

Conclusion

Signals of SS were detected on chromosome 5, 6, 8, and 14, indicating that these regions are conserved through recent generations. Future studies regarding genes and QTL in these SS regions will be carried out to understand the biological processes involved in the phenotypic expression of traits considered in the Canchim genetic evaluation program.

Acknowledgements

The database was provided by the Brazilian Agricultural Research Corporation (Embrapa; Empresa Brasileira de Pesquisa Agropecuária). I. Urbinati, M. E. Buzanskas and T. C. S. Chud would like to thank the São Paulo Research Foundation (FAPESP) (grantee n. 2013/09050-0, 2013/19335-2, and 2012/21891-8, respectively). F. B. Mokry was supported by a fellowship from the Coordination for Enhancement of Higher Education Personnel (CAPES/PNPD) (grantee n. 02663/09-0). D. P. Munari and L. C. A. Regitano were supported by a fellowship from the National Council of Technological and Scientific Development (CNPq).

Literature Cited

Browning, S.R. and Browning, B.L. (2007). Am. J. Hum. Genet., 81:1084-1097.

Buzanskas, M.E. (2013). http://www.athena.biblioteca.unesp.br/exlibris/bd/bja/33004102030P4/2013/buzanskas_me_dr_jabo.pdf. Accessed on Jan. 2014.

Choi, I.S., Kong, H.S., Oh, J.D., et al. (2006). Asian Australas. J. Anim. Sci., 19:927–930.

Clayton, D. (2012). snpStats: R package. Version 1.8.1.

Gasparin, G., Miyata, M., Coutinho, L.L., et al. (2007). Anim. Genet., 38:453-459.

Gautier, M., and Vitalis, R. (2012). rehh: R package. Version 1.0.

Machado, M. B. B., Alencar, M. M., Pereira, A. P., et al. (2003). Genet. Mol. Biol., 26:259–265.

R core team. (2013). R: A language and environment for statistical computing. Version 3.0.1.

Sabeti, P.C., Reich, D.E., Higgins, J.M., et al. (2002). Nature, 419:832-837.

Utsunomiya, Y.T., Pérez O'Brien, A.M., Sonstegard, T.S., et al. (2013). PLoS ONE, 8:1-11.

Voight, B.F., Kudaravalli, S., Wen, X., et al. (2006). PLoS Biol., 4:0446-0458.

Wibowo, T.A., Gaskins, C.T., Newberry, R.C., et al. (2008). Int. J. Biol. Sci., 4:406-414.

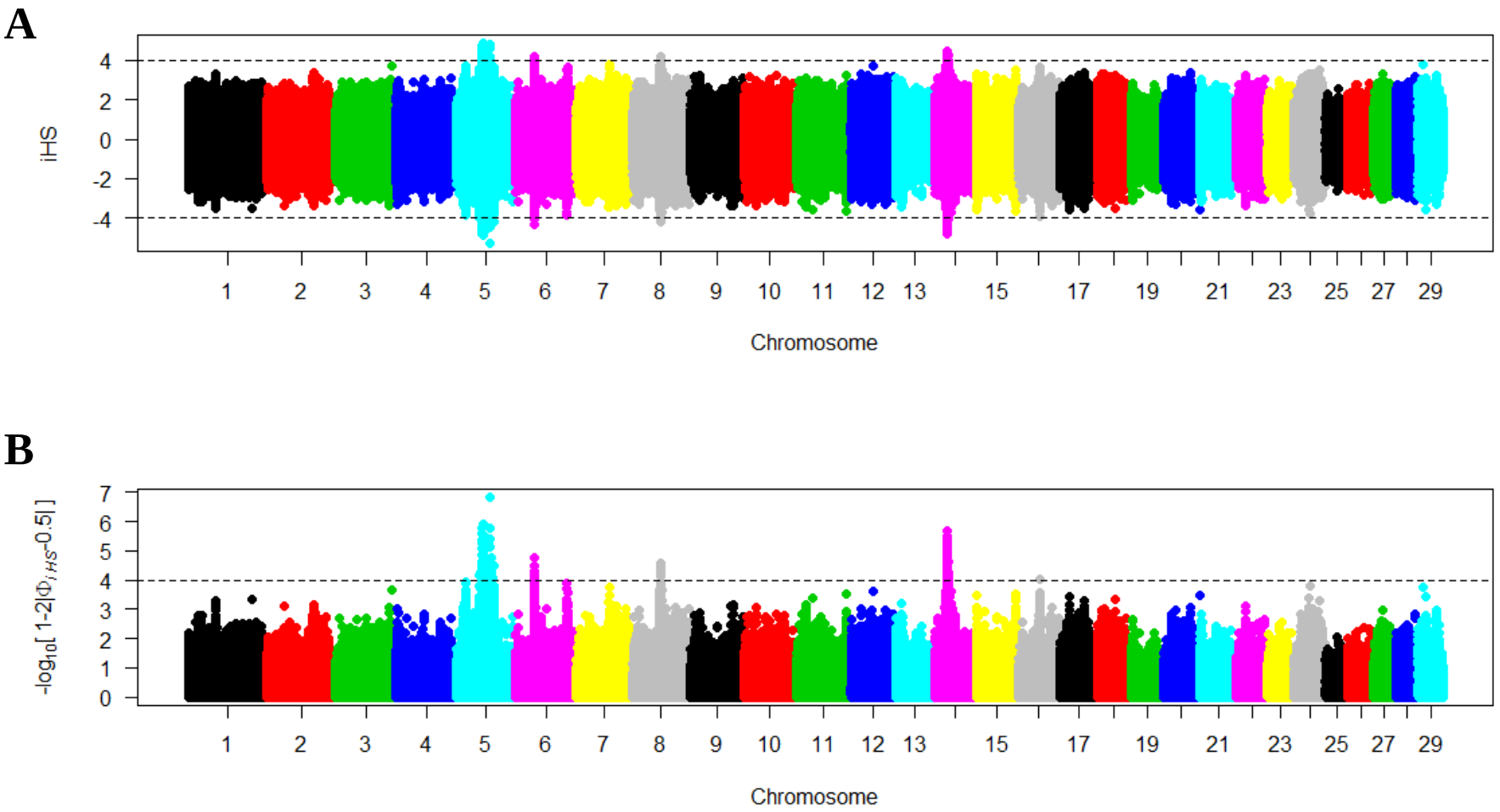

Figure 1. Integrated haplotype scores (iHS) for single nucleotide polymorphism (SNP) markers for each chromosome (A); and transformation of SNP iHS scores into $p_{iHS} = -\log_{10}[1 - 2\,|\Phi_{iHS} - 0.5|]$, where Φ(x) represent the Gaussian cumulative distribution function (B).