

diseases, have lower weight gains, lower reproductive efficiency and lower milk production. Genome-Wide Association Studies (GWAS) use genotyping technologies for analyzing several single-nucleotide polymorphisms (SNPs) and relate them to clinical conditions and measurable traits. This kind of study usually identify SNPs in linkage disequilibrium with the real variant that contributes to the phenotype measured. A region around SNPs found to be associated with reactivity in a previous GWAS performed by our research group was delimited and investigated by a literature data-mining in the present study. The region was delimited by the rising of p-values around the associated marker found in chromosome 1. A list with all the genes located inside the region previously delimited, was constructed. The software Gene Relationships Across Implicated Loci (GRAIL) was used to perform the data-mining. This software uses our text-based definition of relatedness to identify a subset of genes, and it is based on two key methods: a text-based similarity measure that scores two genes for relatedness to each other based on text in PubMed abstracts and a novel statistical framework that assesses the significance of relatedness between genes in disease regions. The input file for GRAIL analyzes was made using the Gene name and the Gene ID of the homologous gene in Human genome.

Results

Initially, the input file consisted for 39 genes. After the analyses we were able reduce considerable the number of genes that can be associated with temperament in Guzerá cattle. The keywords related to the phenotype evaluated were temperament, behavioural and stress. These results show the genes in chromosome 1 that may be associated with behavioural traits and temperament in Guzerá cattle.

Conclusions

The bovine temperament is an important characteristic in the animal's well-being, and it affects the animals productivity. Given these facts, and the importance of this study for the breeding selection, the results shown in the

present work are very important for the temperament genetics from Guzerá and milk production in Brazil.

Supported by: Fapemig, CAPES, EMBRAPA.

14. SNP clustering on DBSCAN and sensitivity analysis

Vinicius C. Brum¹, Itamar L. Oliveira¹, Marcos Vinicius Silva², Wagner Arbex^{1,2}

¹Universidade Federal de Juiz de Fora, PGCC – UFJF

²Empresa Brasileira de Pesquisa Agropecuária – Embrapa

The aims of this work are the knowledge acquisition from genotype data of 56K single nucleotide polymorphism (SNP) chip in genome of three cattle breeds through the algorithm DBSCAN (Density-Based Spatial Clustering of Applications with Noises) and its sensitivity analysis with respect to changes on its initial conditions in this database. The database used has 2467 samples distributed in three cattle breeds. Two breeds are taurine cattle and have 56947 markers: Holstein (577) and Jersey (1024); and one breed is zebu cattle and have 54000 markers: Nellore (866). Through this work was possible to find markers that do not discriminate samples of these breeds, since they remain constant independently of their breeds. In addition, it was possible to analyze the algorithm's sensibility and show that it presents relative stability in this scenario, since the clustering results converged with no difficulty. Finally, this work demonstrates the possibility to build a model able to generalize and predict the unknown breed samples of cattle efficiently.

GENOME ASSEMBLY AND ANNOTATION

15. CGIIS: an automated tool to Close Gaps Inter and Intra Scaffolds

Raul Arias-Carrasco¹, Vinicius Maracaja-Coutinho¹

¹Universidad Mayor

Background

One of the main issues in whole genome sequencing is to fill the gaps on the final assembled genomic sequence, linking the different

RA

SP 6628 P. 216
2014
SP-PP-6628

First ISCB Latin American Student Council Symposium

Belo Horizonte, Brazil
October 27, 2014



Program Booklet

